Vision of details in space and time

Frans J.J. Blommaert

Vision of details in space and time

Front cover:

Contourplots of calculated internal images of the letters 'f' and 'b' of type font Courier 10, if the author views these letters from a distance of 0.5 m, at an eccentric position of 7 degrees and at an adaptation level of 1200 Td. For details see chapter 7.

Vision of details in space and time

Proefschrift

ter verkrijging van de graad van doctor aan de
Technische Universiteit Einhoven, op gezag van de
rector magnificus, prof. dr. F.N. Hooge, voor een
commissie aangewezen door het college van dekanen in
het openbaar te verdedigen op dinsdag 3 februari 1987
te 16.00 uur

door

Franciscus Johannes Josephus Blommaert

geboren te Deurne

Dit proefschrift is goedgekeurd door de promotoren:

Prof. Dr. Ir. J.A.J. Roufs

en

Prof. Dr. H. Bouma

aan Helma

aan mijn ouders

Contents

chapter 1

Introduction

1.1 History

In visual perception research, one is interested in the image the human visual system constructs from the outside world. Its imaging properties have been the subject of investigation for a long time, often stimulated by pure scientific interest and increasingly encouraged by technical demands of imaging devices.

This thesis reports on an investigation into spatial an temporal processing performed by the human visual system. The scope is restricted to relatively simple black and white stimuli, which are limited in space and/or time. An example of such a stimulus in the spatial domain is one out of the set of alphanumeric symbols, such as printed letters. In the temporal domain, signal-lights, short block-like flashes of a few milliseconds, are another example.

The problem may be tackled from different angles, for instance by trying to formulate physiological models that mimic certain perceptual imaging characteristics. In this investigation, however, the subject is approached using linear systems analysis and quantitative description of stimulus/response relations.

Since the time that Fourier techniques became popular, quite some effort has been spent on trying to characterise visual processing by suitable basic functions, like the responses to sinusoids or impulses. Such stimuli can be used to generalise the processing properties of the system under investigation. Since the visual system is rather complex, its general behaviour is commonly separated into different processes. For instance, colour processing and spatial or temporal processing are often regarded as more or less separated areas.

Even these areas are often subdivided according to, for instance, foveal versus eccentric vision, threshold versus suprathreshold perception or retinal versus more central processing. Such subdivisions are a natural consequence of both the complexity of visual processing as a whole, and the limited possibilities to analyse the input/output relations. In general, systems analysis is simplest if the system under investigation obeys the conditions of linearity and homogeneity (cf. Papoulis, 1962).

A number of earlier investigations suggest that the processing of the human visual system is rather non-linear in many aspects. This can for instance be inferred from the work of Stevens (1966) and Mansfield (1973), who determined brightness sensation as a function of stimulus luminance. It is also indicated by EEG responses on harmonically modulated light (Spekreijse, 1966) and by single cell firing rates as a function of stimulus luminance (e.g. Barlow and Levick, 1969). Although the relation between psychophysics and physiology is a matter of discussion, it seems obvious that non-linear processing of single cells is related to non-linear processing at the perceptual level. Linearity of processing then is only an adequate approximation if the state of equilibrium is maintained (i.e. the part of the retina concerned is kept

in a constant and well-defined adaptive state) and the amplitudes of the signals fall within the linearity range of the system. The signals accompanying stimuli consisting of <u>threshold</u> modulations of an otherwise constant adaptation level are assumed to fulfil these conditions (small-signal approach).

In the spatial domain, linear systems analysis in vision was for instance propagated by Schade (1956), van Nes, Koenderink, Nas and Bouman (1967) and Campbell and Robson (1968), who determined detection thresholds for sinusoidal gratings as a function of the spatial frequency of these gratings. The attenuation and phase shift of the sinusoids characterise the linear space-invariant system completely (Optical Transfer Function, OTF). It can be argued that in situations like the present, the frequency dependent attenuation (Modulation Transfer Function, MTF) is the only important characteristic. Using this approach, the spatial processing is characterised by a filter of which the properties can be obtained by determining the thresholds of sinusoids and assuming for instance constant maximum amplitude at threshold level. The reciprocals of the threshold amplitudes define the Contrast Sensitivity Function (CSF) which represents under these conditions the MTF but for a constant factor. From this, the response to arbitrary stimulus patterns can be calculated by linear combination of sinusoids. Campbell and Robson (1968) have shown that this approach is not sufficient for an accurate description of general spatial processing at threshold level.

In the temporal domain, a comparable approach was chosen by de Lange (1952). He measured thresholds for sinusoidally modulated light as a function of the temporal frequency and interpreted the reciprocals of these as the gain of the linear filter which characterises temporal processing. Roufs (1974a) showed that also in the temporal domain such a basic function does not suffice for an accurate description of general temporal processing at threshold level.

From these attempts to describe the action of the visual system, it was learned that, even under limited conditions, single spatial and temporal transfer functions cannot adequately describe these systems. Explanations for these failures were given by proposals on multiple channel processing in the spatial domain (Blakemore and Campbell, 1969) and a division into sustained- and transient-like processing in the temporal domain (Kulikowski and Tolhurst, 1973; Roufs, 1974a). These explanations are based on the idea that visual processing is not performed by one spatial or temporal filter but by different "channels" (cells or filters). According to this concept, signals of a subset of channels, working together on the basis of some summation rule, determine the total response. Those channels that are best tuned to the stimulus will dominate the response.

The spatial interpretation of a multiple channel system is elegantly formalised in the sunflower model of Koenderink and van Doorn (1978). The temporal concept has been treated in detail by Breitmeyer and Ganz (1976). Using a two-channel concept, Roufs (1974a) was able to describe thresholds for fast-changing temporal stimuli in an adequate way.

The research described sofar, is based on the Fourier approach of linear analysis, i.e.

one uses the property that sinusoids are eigenfunctions of a linear system. In such a line of thinking, it is logical to use harmonic functions as input stimuli in order to derive the filter characteristics of a process. When used as visual stimuli, however, these functions have the disadvantage that they are in principle infinitely extended either in space or in time, or both. The visual system is spatially inhomogeneous (e.g. van Doorn, Koenderink and Bouman, 1972) and suffers in the temporal domain of increasingly disturbing nonlinearities due to stochastic effects if the observation time increases (Roufs, 1974b). Therefore, the suitability of these stimuli may be questioned.

In order to deduce the processing characteristics of a linear system, it is also sufficient to specify the impulse response function in the time domain or the point spread function in the spatial domain. These functions are, by Fourier transform, uniquely related to the transfer functions in both temporal and spatial domains (see for instance Papoulis, 1962). Therefore, if impulse responses and point spread functions could be obtained avoiding the use of extended stimuli, they might prove to be precise tools for the specification of local temporal and spatial processing. This approach is the focus of this thesis.

In such an analysis of linear systems, impulse responses and point spread functions are usually obtained by using temporal and spatial pulses (approximations of Dirac-functions; Papoulis, 1962) as input signals. In order to determine these responses experimentally, a modified method of subthreshold summation was developed, which we called "perturbation technique" earlier (Roufs and Blommaert, 1975; Blommaert, 1977; Roufs and Blommaert, 1981; Blommaert and Roufs, 1981).

The theoretical essence of the technique was derived from mathematical physics, where it is common calculus to estimate small variations from a state of equilibrium by perturbing the steady state input signal by a small amount (Courant and Hilbert, 1962). This calculus can be extended to include perturbations of well-defined dynamic solutions of a process. In this way, variations of output signals due to the perturbating input, can fairly simply be estimated as low-order deviations from a well-defined solution.

In order to illustrate the technique, a hypothetical temporal response of the visual system to a short flash is shown in Fig. 1. In the example, the response is taken to be a monophasic deviation of the output signal from a steady state level (in order to understand the essentials, the exact nature of the response is not important). When modelling detection of such a flash, it is usually assumed that the flash will be seen if the amplitude of the response exceeds, at least once, a noisy threshold level "a".

For the sake of simplicity, it will be assumed for the moment that the system's noise is negligible. In that case, the flash will be detected if the peak value of the response equals or exceeds threshold level "a". Now, suppose a small perturbation is added to the input signal. This will result in a small linear increase of the output signal. The summation will only affect the threshold of the flash if the perturbing signal does not equal zero at the time at which the flash response reaches its peak value. It will be clear that at threshold the flash luminance can be decreased in close relation to
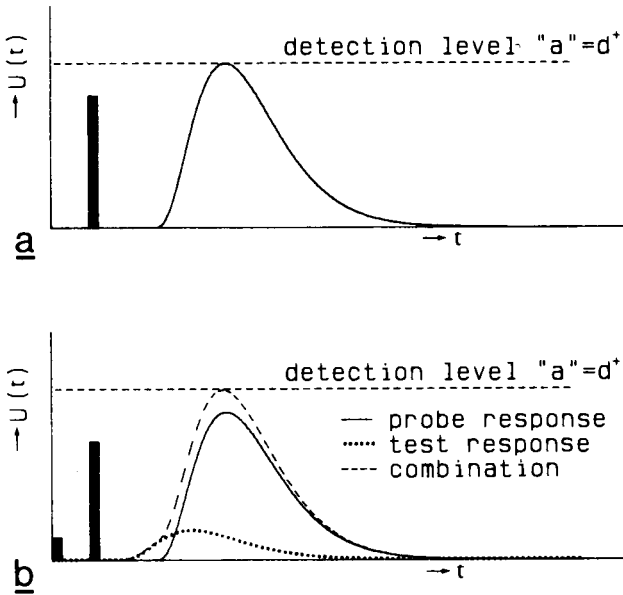
Figure 1: $\underline{a}$ A flash will be detected if the extreme value of its response $U_p(t)$ reaches threshold level "a". $\underline{b}$ In this case a perturbation flash is added to the input signal, yielding a hypothetical response $U_t(t)$. The total response will equal the algebraic sum of the individual ones $U_p(t) + U_t(t)$. Owing to the presence of the perturbation, threshold level "a" will be reached for a smaller intensity of the probe flash. The flashes are denoted by the black bars.

the response of the perturbating signal. This property underlies the principle of the perturbation technique as will be shown in the next chapter.

The perturbation technique is related to other subthreshold summation techniques (cf. Kulikowski and King-Smith, 1973; Hines, 1976; Ikeda, 1965) in the property that the effect of a subthreshold test stimulus is scanned by the response of a probe stimulus. In case of using the perturbation technique, however, special attention is given to achieve that the measured threshold variations are strictly related to the systems response to a certain input signal. Furthermore, the practical design of the experiments was chosen such that the effects of systematic sensitivity variations of the subjects were minimised.

As was pointed out before, an essential property should be fulfilled, namely that of linearity. From Bloch's and Ricco's law (cf. Roufs, 1972; Le Grand, 1967) it is suggested that, for short flashes and small points, ideal integration holds. Another experiment which supports the linearity hypothesis for small and local input signals was performed by Kulikowski and King-Smith (1973). They measured the effect of a subthreshold line on the detection threshold of a probe line situated nearby.

as a function of the test line luminance. Linearity was concluded from the linear relationship between the threshold of the probe line and the luminance of the test line. Such experiments, at threshold level, indicate that small-signal linearity is also a valid working hypothesis in case of signals which are extremely limited in space or time. In a number of perturbation experiments reported on in this thesis, the linearity assumption under these conditions will be tested.

In experiments where thresholds are determined, the noise properties of the visual system play an important role. This is reflected, for instance, by the fact that the detection process itself is a probabilistic one. Therefore, the effects of noise should be carefully evaluated. This can be found, for instance, in the chapters 3 and 5.

The strength of a basic function lies in its generalising properties, i.e. if such a basic function is known, responses of the system to arbitrary stimulus functions should be calculable. This property of basic functions will be used in order to predict visual responses on various details in space and time. For the temporal domain, predicted thresholds for a number of time-dependent fast-changing stimuli will be shown to be in quantitative agreement with experimentally determined ones if the conditions are chosen such that the signals of one of the two possible channels are dominant (chapter 3).

In the spatial domain, things are somewhat more complicated. However, for a limited set of slender details, thresholds can be successfully predicted on the basis of a single foveal point spread function, associated with the channel transfering the highest spatial frequencies (chapter 5). Its effectiveness will once more be tested in chapter 7, where letter confusions are calculated and confronted with experimental findings.

The content of this thesis can be divided roughly into three constituent parts:

- **Temporal processing.** The perturbation technique will be used here in order to determine impulse and step responses of the visual system (chapter 2). Next, predicted thresholds and latences on the basis of these experimentally determined impulse responses will be shown to be in fair quantitative agreement with experimental findings (chapter 3).

- **Spatial processing.** Experimental results on point and edge spread functions will be shown to be in quantitative agreement (chapter 4). On the basis of experimentally determined point spread functions, a minimal ensemble model is constructed which suffices for a quantitative prediction of thresholds for various spatial details (chapter 5). Chapter 6 is the report of a small investigation showing that the width of experimentally determined point spread functions depends on the temporal presentation.

- **Letter recognition.** Experimentally determined point spread functions will be used to estimate the effects of optical and neural processing on letter recognition if the letters are presented at a long viewing distance or in the parafovea (chapter 7). Letter recognition can also be investigated by contrast diminishing under otherwise normal viewing conditions. It will be shown how recognition thresholds depend on letter size and adaptation level (chapter 8).

The chapters 2, 4 and 6 are taken from articles already published. Since size reduction would have diminished their legibility substantially, it was decided to put the contents of these articles into the same print as used for the rest of this thesis.

1.2 Principles of the perturbation technique

a Theory

The model used as an approximation of spatial and temporal processing is visualised in Fig. 2. It consists of a linear element working either in the time or in the space domain. Its parameters generally will depend on the adaptive state of the eye. The linear operator is followed by a noisy peak detector. In the following, detection will be assumed to occur if the detection level "a" will be reached at least once (high threshold assumption).
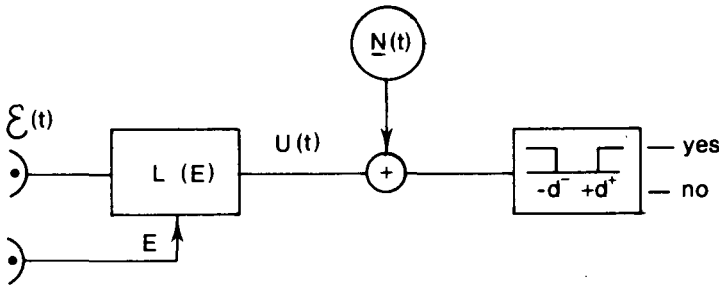
Figure 2: Hypothetical detection mechanism for stimuli varying either in time or in space. The stimulus signal is processed by a linear operator and the result is fed into a noisy peak detector. If the signal $U(t) + \underline{N}(t)$ at the detector's input exceeds level d^+ or $-d^-$, the stimulus will be detected.

For the sake of simplicity, it will be assumed for now that the effects of noise are negligible. Later on, its share in the detection process will be illustrated.
For the purpose of elucidating the perturbation technique, let us consider the visual response to a short flash, effectively a pulse. In that case, the linear element is considered to perform only temporal filtering.
For clarity, the impulse response is taken to be a monophasic fourth-order process as is depicted in Fig. 3 (in chapter 3, it will be shown that the impulse response looks like this in case the stimulus is a point source).
Using the model of Fig. 2, such a flash will be detected if the peak value of the response reaches the detection level $a = d^+$, at time t_{ex} at which the flash response reaches its extreme value (see Fig. 3a).
Now suppose, we add to the first flash a second flash of equal duration, delayed or advanced by a time interval τ and a factor of q lower in luminance. The response to
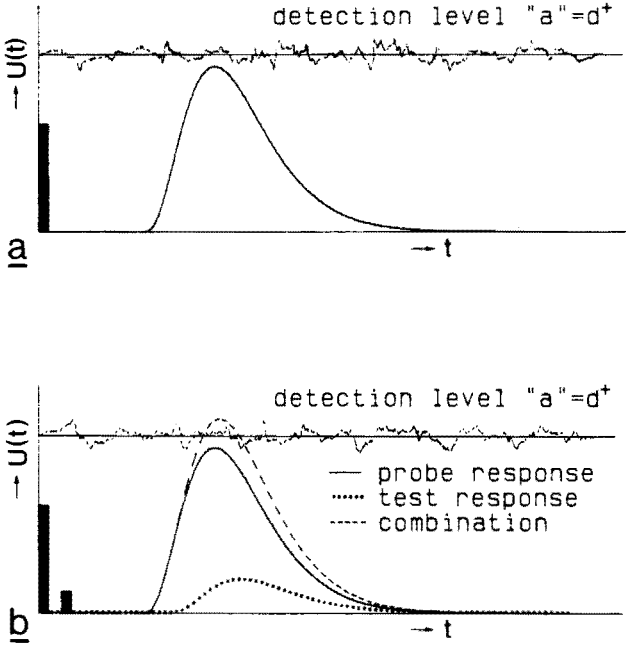
Figure 3: <u>a</u> A short flash will be detected if its response, represented by a fourth-order process, exceeds threshold level "a". In the illustrated case, no detection will occur. <u>b</u> The combination of probe and test flashes will be detected since the summed response (dashed curve) exceeds the threshold level, although neither of the individual ones does.

the combination of flashes then consists of the algebraic sum of the reponses to the two individual ones as is visualised in Fig. 3b.

On the basis of the model-assumptions, it can be derived for the threshold of a single flash that

$$\varepsilon_0 \vartheta U_\delta(t_{ex}) = a, \tag{1}$$

where:

ε_0 is the retinal illumination of the flash,
ϑ is the duration of the flash,
$U_\delta(t)$ is the impulse response function,
t_{ex} is time at which the impulse response attains its extreme value,
a is the detection level; $a = d^+ \cup -d^-$; $d^+, d^- > 0$.

For the threshold of the combination one obtains

$$\varepsilon_c \vartheta \{U_\delta(t_{ex}) + qU_\delta(t_{ex} - \tau)\} = a, \tag{2}$$

where:

ε_c is the retinal illumination of the probe flash,
q is the ratio of test and probe flash luminances,

τ is the time interval between both flashes, taken positive if the perturbation lags with respect to the probe stimulus.

From equations 1 and 2 it can be derived for the normalised impulse response $U_\delta^\star$ that

$$U_\delta^\star(t_{ex} - \tau) = \frac{U_\delta(t_{ex} - \tau)}{U_\delta(t_{ex})} = \frac{1}{q}\left(\frac{\varepsilon_0}{\varepsilon_c} - 1\right). \tag{3}$$

Eq. 3 shows that, in principle, from two threshold determinations a single value for the normalised impulse response $U_\delta^\star(t_{ex} - \tau)$ can be obtained. Thus by varying the time interval τ between probe and test flash, a discrete number of values for the impulse response can be found.

Instead of the normalised impulse response, also the absolute one could have been considered. As will be explained in chapter 2, taking the normalised form has the advantage that the effects of non-stationary sensitivity shifts on the shape of the experimentally determined impulse response can be minimised.

In illustrating the principle of the technique, it has been neglected that detection might occur at time instances which do not exactly equal t_{ez}. Furthermore, possible effects of noise should be taken into account. These are the subjects of the following paragraphs.

b The effect of a blunt probing peak

In explaining the perturbation technique, it is implicitly assumed that the response of the probe flash has a sharp dominant peak which makes it possible to scan the response of the test flash accurately. In practice, however, the dominant peak has a certain bluntness, which will cause a systematic error in the scanning of the desired response function. This error will not only depend on the degree of bluntness of the dominant peak but also on the shape of the response function under investigation. In general, it is likely that such an error will be larger if the scanned function has higher derivative values.

To illustrate the consequences of a blunt probing peak, a perturbation experiment was simulated using the impulse response of Fig. 3a. In this simulation, the q-factor (the ratio of test and probe flash luminances) was varied from 0.1 to 0.9. While allowing detection to occur at time instances differing from t_{ez}, the 'measured' responses were calculated according to eq. 3. The results for different q-values are shown in Fig. 4. From this figure it can be seen that, owing to the peak bluntness, the result of such an experiment will progressively deviate from the impulse response if the q-factor is larger, i.e. if the luminance of the test flash is higher. Furthermore, it can be seen that the deviation from the impulse response is largest at those time instances at which the derivative of the perturbing impulse response attains its largest value.

In the example of Fig. 4, a fourth-order process was chosen for illustration. The results, however, can be generalised to hold globally for arbitrarily shaped impulse responses: there will be a tendency to underestimate the slope at abscissa values where the derivative of the test function is largest.
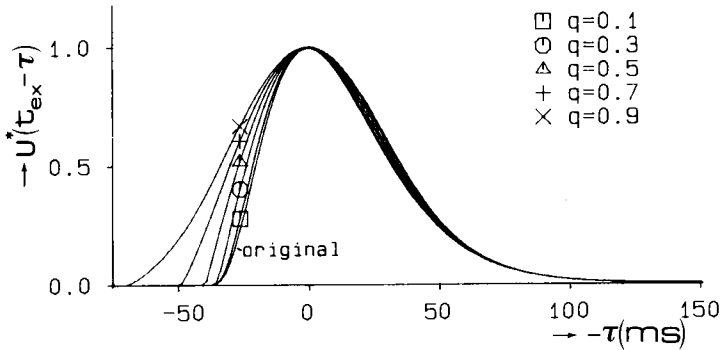
Figure 4: Demonstration of the effect of a blunt probing needle on the result of a perturbation experiment using the impulse response of Fig. 3a. If the q-factor (ratio of test and probe stimulus luminances) is small, the result will almost coincide with the test flash response. For larger q-values, however, the result will increasingly deviate.

The difference between a usual subthreshold summation experiment (see for instance Tolhurst, 1975) and the perturbation technique can be demonstrated with the aid of Fig. 4. Using the common form of a subthreshold summation experiment, one merely estimates the effect of a subthreshold test stimulus on the threshold of a probe. In that case, the practical value of the q-factor is of minor importance, as long as the desired effect can be demonstrated. Using the perturbation technique, however, one specificly seeks for the correct shape of some basic response function. In that case, one has to take care that systematic effects of peak bluntness are minimised.

To achieve this, small values of the q-factor should be chosen. As will be shown further on, this is also beneficial to reduce the effects of noise. Unfortunately, q cannot be chosen too small since in reality the threshold is a stochastic variable (see below) and the sample noise would make a precise measurement unpracticle (we experienced a value of q between 0.1 and 0.3 to be a good compromise).

c The influence of noise

If a stimulus, having a luminance nearby threshold, is repeatedly presented, one may notice that this stimulus will not be detected in all cases. This is caused by the presence of noise in the visual pathway and/or in the detection mechanism.

Hence, the threshold of a stimulus, defined as that luminance value at which the stimulus will be detected in 50% of the presentations, is not a deterministic quantity but essentially a probabilistic one. This is reflected in the so-called frequency-of-seeing curve or psychometric function. Such a curve can be measured by plotting the number of times a stimulus is detected against the luminance of the stimulus (see Fig. 5).

From such a curve, a 50% detection threshold can be extracted as that luminance
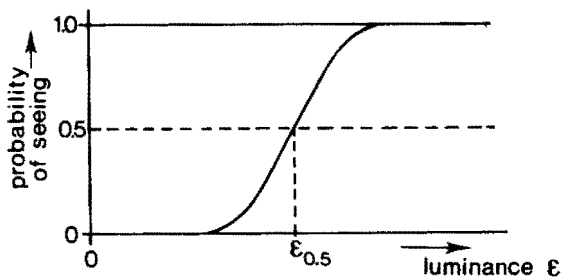
Figure 5: Schematised psychometric function. The probability of detecting a stimulus will increase with stimulus luminance according to a monotonic function, often assumed to be cumulative normal. The 50% detection threshold is defined by the luminance value $\varepsilon_{0.5}$ at which the psychometric function crosses the 50% detection level.

value at which the psychometric function crosses the 50%-level (e.g. Guilford, 1954).

If the determination of a frequency-of-seeing curve is repeated in time, a threshold value will be found which, in general, deviates from the value found before. The variation between these thresholds can only partially be explained by the stochastic process underlying the psychometric function (Roufs, 1974b). This means globally that the noise components which determine the shape of the psychometric function do not exactly equal those that are responsible for between-threshold variability.

This may be explained by non-systematic sensitivity shifts or the action of noise which contains rather low frequency components (see also chapter 3). If visual noise exhibits this property, a frequency-of-seeing curve will depend on the length of the time interval during which it is determined. This important property is seldomly recognized in the literature.

In order to be able to estimate the effects of noise, its structure and parameters should be known. For this purpose, various distribution functions can be postulated and fitted to experimentally determined psychometric functions.

The most important parameter, the slope of the psychometric function which is connected with the width of the distribution function of the stochastic variable, usually forms the basis of calculations of noise effects. The choice of the specific distribution seems to be relatively unimportant, although it is commonly assumed that the normal or log-normal distribution is the most plausible choice (Nachmias, 1981; Roufs, 1974b). Watson (1979) and Wilson and Bergen (1979) used Quick's (1974) non-linear distribution to describe the psychometric function. This distribution combines the advantage of simple mathematics with a fair similarity to the shape of the normal distribution.

Following the terminology of Quick, the psychometric function can be expressed as:

$$\Psi = 1 - \prod_{i=1}^{N}(1 - (1 - 2^{-|R_i|^\beta})), \tag{4}$$

where:

Ψ is the detection probability of the stimulus,

N is the number of samples constituting the response,

β is a parameter proportional to the slope of the psychometric function,

R_i is the response magnitude of the i'th sample.

It can easily be verified (cf. the appendix of chapter 3) that for the threshold condition of a stimulus with time function $f(t)$ and retinal illuminance ε:

$$\int_0^\infty \left| \varepsilon f(t) \otimes \frac{U_\delta(t)}{a} \right|^\beta dt = 1. \tag{5}$$

In this formula $\otimes$ denotes convolution.

Eq. 5 resembles the one given by Rashbass (1970) in his model on the visibility of transient changes of illuminance. A comparison with the present model will be discussed in chapter 2.

Using eq. 5, it can be calculated what the influence of noise would be on the result of an experiment using the perturbation technique. In Fig. 6, the results of such a simulation are shown if for the impulse response the fourth-order process of Fig. 3 is taken. Different values for β are indicated, and q was fixed at 0.1.
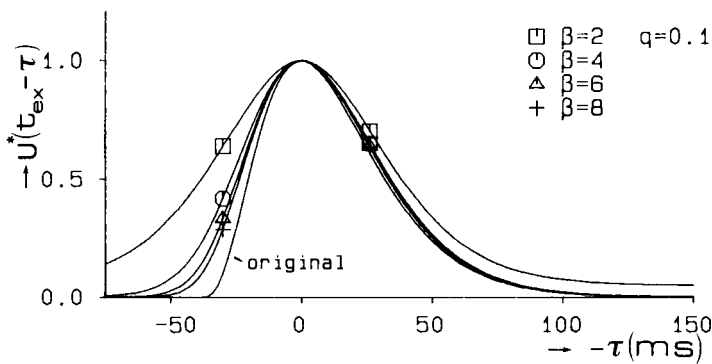
Figure 6: Demonstration of the influence of noise on the results of a perturbation experiment. For the calculations, the impulse response of Fig. 3a was taken. The q-factor was fixed at 0.1. The result can be seen to depend on the noise properties, reflected in the β-values shown in the legend. For higher β-values (steeper frequency-of-seeing curve) the result will be a fair approximation of the desired impulse response.

From the simulation it can be seen that for high β-values the effect of noise is negligible. For low β-values, however, the experimental result will increasingly deviate from the impulse response.

The correct value of β can thus be seen to be important, as was signalled for instance by Watson (1982).

As will be shown in chapter 4, we arrived at a β-value of 7.1, with a minimum of 6, by averaging across 13 subjects. In that case, the effect of noise on the results of perturbation experiments is well within experimental accuracy.

In the foregoing, principles of the perturbation technique were elucidated for the temporal domain. In the spatial domain, the influence of stochastic variables is diminished since the stimulus configuration can be chosen such that the subject only perceives the well-located response extremum (peak amplitude of the point spread function). In that case, the chance that other response areas might influence detection probability is ruled out. In the temporal domain, the subject is less capable of locating the extreme response amplitude.

References

Barlow, H.; Levick, W. (1969) Changes in the maintained discharge with adaptation level in the cat retina. J. Physiol. 202, 699-718.

Blakemore, C.; Campbell, F.W. (1969) On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. J. Physiol. 203, 237-260.

Blommaert, F.J.J. (1977) Spatial processing of small visual stimuli. IPO Ann. Progr. Rpt. 12, 81-86.

Blommaert, F.J.J.; Roufs, J.A.J. (1981) The foveal point spread function as a determinant for detail vision. Vis. Res. 21, 1223-1233.

Breitmeyer, B.G.; Ganz, L. (1976) Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression and information processing. Psychol. Rev. 83, 1-36.

Campbell, F.W.; Robson, J.G. (1968) Application of Fourier analysis to the visibility of gratings. J. Physiol. 197, 551-566.

Courant, R.; Hilbert, D. (1962) Methods of Mathematical Physics. Interscience Publishers, John Wiley and Sons, New York and London.

Doorn,A.J. van; Koenderink, J.J.; Bouman, M.A. (1972) The influence of retinal inhomogeneity on the perception of spatial patterns. Kybernetik, 10, 223-230.

Green, M. (1984) Masking by light and the sustained-transient dichotomy. Perc. & Psychoph. 35, 519-535.

Guilford, J.P. (1954) Psychometric Methods. McGraw-Hill New York.

Hines, M. (1976) Line spread function variation near the fovea. Vis. Res. 16, 567-572.

Ikeda, M. (1965) Temporal summation of positive and negative flashes in the visual system. J. Opt. Soc. Am. 55, 1527-1534.

Koenderink, J.J. Doorn, A.J. van (1978) Visual detection of spatial contrast; influence of location in the visual field, target extent and illuminance level. Biol. Cyb. 30, 157-167.

Kulikowski, J.J.; King-Smith, P.E. (1973) Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation. Vis. Res. 13, 1455-1478.

Kulikowski, J.J.; Tolhurst, D.J. (1973) Psychophysical evidence for sustained and transient detectors in human vision. J. Physiol. <u>232</u>, 149-162.

Lange, H. de (1952) Experiments on flicker and some calculations on an electrical analogue of the foveal system. Physica <u>18</u>, 935-950.

Le Grand, Y. (1967) Form and Space Vision. Indiana Univ. Press. Bloomington.

Mansfield, R.J.W. (1973) Brightness function: effect of area and duration. J. Opt. Soc. Am. <u>63</u>, 913-920.

Nachmias, J. (1981) On the psychometric function for contrast detection. Vis. Res. <u>21</u>, 215-223.

Nes, F.L. van; Koenderink, J.J.; Nas, H.; Bouman, M.A. (1967) Spatiotemporal modulation transfer in the human eye. J. Opt. Soc. Am. <u>57</u>, 1082-1088.

Papoulis, A. (1962) The Fourier Integral and its Applications. McGraw-Hill Book Company, New York, San Francisco, London, Toronto.

Quick, R.F. (1974) A vector magnitude model for contrast detection. Kybernetik <u>16</u>, 65-67.

Rashbass, C. (1970) The visibility of transient changes of illuminance. J. Physiol. Lond. <u>210</u>, 165-186.

Roufs, J.A.J. (1972) Dynamic properties of vision -I. Experimental relationships between flicker and flash thresholds. Vis. Res. <u>12</u>, 261-278.

Roufs, J.A.J. (1974a) Dynamic properties of vision -IV. Thresholds of decremental flashes, incremental flashes and doublets in relation to flicker fusion. Vis. Res. <u>14</u>, 831-851.

Roufs, J.A.J. (1974b) Dynamic properties of vision -VI. Stochastic threshold fluctuations and their effect on flash-to-flicker sensitivity ratio. Vis. Res. <u>14</u>, 871-888.

Roufs, J.A.J.; Blommaert, F.J.J. (1975) Pulse and step reponse of the visual system. IPO Ann. Progr. Rpt. <u>10</u>, 60-67.

Roufs, J.A.J.; Blommaert, F.J.J. (1981) Temporal impulse and step responses of the human eye obtained psychophysically by means of a drift-correcting perturbation technique. Vis Res. <u>21</u>, 1203-1221.

Schade, O.H. (1956) Optical and photo-electric analogue of the eye. J. Opt. Soc. Am. <u>46</u>, 721-739.

Spekreijse, H. (1966) Analysis of EEG responses in man. Thesis. Amsterdam.

Stevens, S.S. (1966) Duration, luminance and the brightness exponent. Perc. & Psychoph. <u>1</u>, 96-100.

Tolhurst, D.J. (1975) Sustained and transient channels in human vision. Vis. Res. <u>15</u>, 1151-1155.

Watson, A.B. (1979) Probability summation over time. Vis Res. <u>19</u>, 515-522.

Watson, A.B. (1982) Derivation of the impulse response: comments on the method of Roufs and Blommaert. Vis. Res. <u>22</u>, 1335-1337.

Wilson, H.R.; Bergen, J.R. (1979) A four mechanism model for threshold spatial vision. Vis. Res. <u>19</u>, 19-32.

14

chapter 2

Temporal impulse and step responses of the human eye obtained psychophysically by means of a drift-correcting perturbation technique[1]

Jacques A.J. Roufs
Frans J.J. Blommaert

Abstract

Internal impulse and step responses are derived from the thresholds of short probe flashes by means of a drift-correcting perturbation technique. The approach is based on only two postulated systems properties: quasi-linearity and peak detection. A special feature of the technique is its strong reduction of the concealing effect of sensitivity drift within and between sessions. Results were found to be repeatable, even after about one year. For a 1 deg foveal disk at 1200 Td stationary level, impulse responses of increments and decrements were found to be mirror-symmetrical. They were equal to the derivatives of the measured step responses. As a consequence the threshold of any fast-changing retinal illumination should be predictable. This will be tested in a subsequent paper. The transfer function of the system responding to a 1 deg stimulus shows a band-pass filter type of processing for transients, confirming quantitatively earlier findings. In contrast, a foveal point source on an extended background of 1200 Td, to which impulse and step responses also appear to be linearly related, gives rise to low-pass filter action of the system.

2.1 Introduction

This paper concerns the dynamic processing of visual stimuli near threshold level. In order to be able to predict thresholds of a fairly large class of time dependent stimuli, dynamic properties of the system have to be identified and the related parameters specified. The latter is usually done by deriving some basic response function from measurements. Several examples of this systems analysis kind of approach can be found in the literature (de Lange, 1952; Veringa, 1961; Kelly, 1961, 1969; Matin, 1968; Levinson, 1968; Sperling and Sondhi, 1968; Hallett, 1969a,b; Rashbass, 1970; Kelly and Savoie, 1978).

In an earlier series of papers, one of the present authors constructed such a model in two steps of refinement (Roufs, 1971; 1972a,b; 1973; 1974a,b,c). In these cases the basic response function was obtained by the thresholds of sinusoidal modulation and related to the thresholds of various type of one shot functions and perception

---

latency. The results could be explained by the assumption of two systems operating in parallel (Fig. 1). A low-pass filter, associated with the physiologically defined "sustained" cells (Cleland et al., 1971) and a strict band-pass type of filter associated with transient cells. In the case of sinusoidal modulation, the output of the former was suggested to cause the homogeneous brightness variations at low frequencies ("swell") and the output of the latter the typical percept seen at the high frequencies ("agitation"). The bandpass filter was found to process quasi-linearly and an impulse response was derived from the gaincurve by assuming minimum phase behaviour (but for a pure time delay). This impulse response was used to calculate the response of other transients by convolution, from which their thresholds could be calculated.
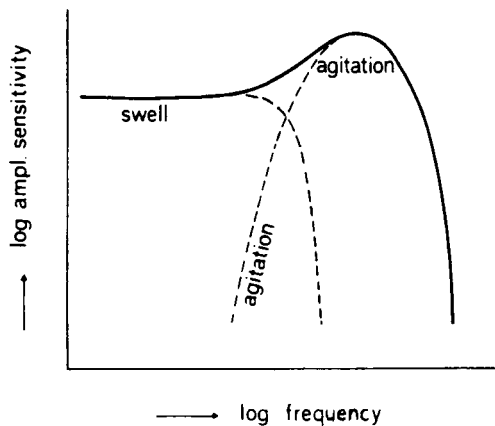
Figure 1: Schematic representation of the amplitude sensitivity curves of harmonically modulated light of a 1 deg foveal field as a composite of the curves of two constituent processes and the perceptual pheno- mena associated with the two output variables. Homogeneous bright- ness changes ("swell"; Roufs, 1972a, 1974a) are linked with a low-pass filter or sustained type of processing. The typical inhomogeneous per- cept accompanying flicker at middle and high frequencies ("agitation") is connected with a linear band-pass filter or transient type of process- ing.

Impulse responses or any response, however, can theoretically be derived from thres- hold measurements in a more direct way on the basis of quasi-linearity and peak detection by applying a special case of subthreshold summation, perturbation. This involves measurements of changes in threshold of a probe flash, the response of which is superimposed on the test flash response.

In practice, the derivation of responses from changes in a small quantity, as thres- holds usually are, is hampered seriously by imprecision due to non-stationary effects like drift in the thresholds themselves. Fortunately, as will be explained below, this disadvantage can be overcome by a special drift-correcting measuring technique, using

a sensitivity reference all the time. An important additional advantage of the use of a probe flash is that the perceptual attribute to be detected is the same for all kinds of test stimuli since the only stimulus the response to which exceeds the critical value to be detected is that of the probe flash.

Furthermore, in interpreting the results obtained with one short probe flash, no substantial corrections for probability summation have to be made, as are necessary for instance for prolonged stimuli like gated sinusoids (Roufs, 1973, 1974c).

In order to facilitate comparison, most experiments were performed with the same stimulus configuration and background levels used in previous experiments mentioned above. In this article it will be shown that:

- Transient responses of the visual system can be measured by means of a drift-correcting perturbation method within sufficient precision to make quantitative analysis applicable.

- The results can be understood on the basis of quasi-linearity and peak detection, which can be tested in combination.

- The system is found to react as a band-pass filter upon fast changes of luminance in the case of a 1 deg foveal stimulus. However, in case of a point source the system behaves as a low-pass filter.

In a subsequent article, it will be shown that threholds of several types of transients can be predicted accurately and simply from the impulse responses obtained in this way. (A short report of some essentials was published before, Roufs and Blommaert, 1975)

2.2 Methodological concepts

A model

Changes of retinal illumination caused by a stimulus on a steady background level E will be described by $\varepsilon_f f(t)$, ε_f being the amplitude factor and $f(t)$ the normalised time function. In this article we shall only consider small and fast changes of retinal illuminance (transients). For sufficiently large fields these evoke perceptual changes in the visual field ("agitation") which cannot be identified as brightness changes. For small fields, on the other hand, clear brightness increments or decrements may be observed. The model used is illustrated in Fig. 2.

Two deterministic systems properties are postulated. First, small changes are processed linearly:

$$L\{\varepsilon_f f(t)\} = \varepsilon_f U_f(t), \tag{1}$$

where L is a linear operator and $U_f(t)$ the response from the linear system to $f(t)$. Second, the stimulus $\varepsilon_f f(t)$ is seen if its response deviates at least by a magnitude "a" from the stationary reference level (peak detection). This might be a signal-to-noise criterion or an internal threshold level. Thus at threshold:

$$\varepsilon_f extr\{U_f(t)\} = a \tag{2}$$
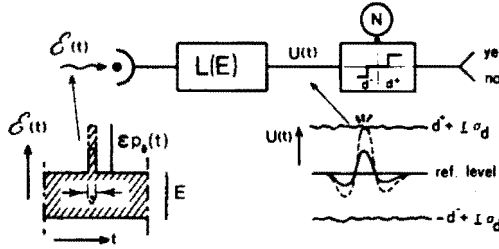
Figure 2: The working of the hypothetical mechanism for detecting fast luminance variations. At the lower left an example of such a variation of retinal illumination is shown. The signal, which is proportional to a (small) luminance variation is processed linearly by the first part L of the system. Response $U(t)$ leads to perception if the deviation from the stationary state exceeds a certain amplitude d^+ or $-d^-$. The standard variable of the stochastic process is $\underline{r}$.

$$a = d^+ \text{ or } -d^-; \ d^+, d^- > 0.$$

If the extremum happens to be positive $a = d^+$, otherwise $a = -d^-$ (symmetry can be concluded within the model from earlier experiments (Roufs, 1974a) but is not a necessary condition for the following). Eq. 2 states that if $U_f(t)/a$, the response to $f(t)$ in a-units is known, the threshold value of the amplitude factor ε_f can be calculated. The magnitude a is in fact thought to be a stochastic variable, giving rise to the psychometric function and involving some interesting invariances (Roufs, 1974c). In this article, however, the intrinsic stochastic properties are not essential and therefore, a will be treated for convenience as a deterministic quantity unless specified otherwise. In a subsequent paper the effect of stochastic variations of a will be dealt with in detail in connection with stimuli for which it is relevant (this model differs from the one proposed by Rashbass, 1970. However, it predicts the ellipse like figures as will be shown in full detail further on).

For all stimuli the values ε_f corresponding with a 50% detection probability will be taken as threshold. As an example, let us take a rectangular flash with an intensity increment ε and a duration ϑ, which is short compared to the time constants of system L. Denote this flash by $\varepsilon_p p(t)$. From its response, $\varepsilon_p U_p(t)$, we obtain the threshold value by applying eq. 2:

$$\varepsilon_p extr\{U_p(t)\} = a. \tag{3}$$

The system L is fully characterised by its unit impulse response $U_\delta(t)$. If the flash is short, the response $\varepsilon_p U_p(t)$ can be approximated by $\varepsilon_p \vartheta U_\delta(t)$. The threshold condition becomes in this case:

$$\varepsilon_p \vartheta extr\{U_\delta(t)\} = \varepsilon_p \vartheta U_\delta(t_{ex}) = a, \tag{4}$$

where t_{ex} is the time after stimulus onset at which $U_\delta(t)$ attains its extreme value.

Thus at threshold, stimulus factors are related to the extreme of the impulse response by:

$$\frac{1}{\varepsilon_p \vartheta} = \frac{U_\delta(t_{ex})}{a}.$$ (5)

The right hand side of the equation will be referred to later on as the norm factor of the unit impulse response. In order to predict thresholds of arbitrary fast changing stimuli by means of eq. 2, not $U_\delta(t_{ex})/a$ but $U_\delta(t)/a$ is generally needed, because the response to an arbitrary time function $\varepsilon_f f(t)$ is given by:

$$\varepsilon_f L\{f(t)\} = \varepsilon_f \frac{U_f(t)}{a} = \varepsilon_f \int_0^\infty f(t-\tau)\frac{U_\delta(\tau)}{a}d\tau.$$ (6)

Eq. 6 is a convolution of the input with the impulse response.

Perturbation approach

As said above, perturbation is a method that can be used to determine responses from measured thresholds, based on the assumed linearity and on peak detection. The essentials of the method are shown in Fig. 3.
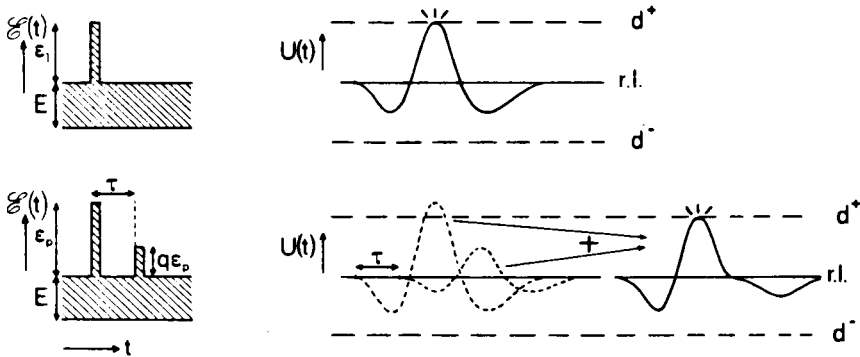
Figure 3: The principles of perturbation. The upper left is a short rectangular flash, effectively an impulse. Upper right represents the response of system L at threshold condition. The lower left is the combination of probe flash and smaller test flash. In the lower right part, the interaction of the two individual responses (dashed curves) and the resulting response (continuous curve) at threshold condition are shown. Notice that in this case the intensity of the probe flash in the combination must be larger than in the case of one isolated flash, reflecting the influence of the test flash response.

In order to probe the response to some stimulus unambiguously the response to the probe flash has to have one clear dominant phase which can trace the profile to be measured. (In Fig. 3 this would be the second phase.) If there is any doubt, this

can be tested within the same theoretical frame (see Roufs, 1974a, p840). Now take a combination of a short flash $\varepsilon_p p_\vartheta(t)$ and any other test transient $\varepsilon_f f(t)$, delayed some time τ, the response to which we want to determine. The threshold condition for the combination is:

$$extr\{\varepsilon_p \vartheta U_\delta(t) + \varepsilon_f U_f(t - \tau)\} = a. \tag{7}$$

Since we want to probe the response to $\varepsilon_f f(t)$ with the dominant phase of the probe flash response $\varepsilon_p \vartheta U_\delta(t)$, we shall have to make sure that for any τ no other combination of the phases of probe and test response meet the amplitude criterion a. In mathematical terms:

$$|extr\{\varepsilon_f U_f(t)\}| \ll |extr\{\varepsilon_p \vartheta U_\delta(t)\}|. \tag{8}$$

In fact, it is this inequality which characterises perturbation as a special case of subthreshold summation. The condition prescribed by eq. 8 is also determined by the stochastic nature of a and the necessity of keeping the joint probability of all other peaks negligible with respect to that of the dominant phase. This condition is especially relevant for prolonged stimuli. Thus, eq. 7 simplifies to:

$$\varepsilon_p \vartheta U_\delta(t_{ex}) + \varepsilon_f U_f(t_{ex} - \tau) = a. \tag{9}$$

It is convenient to use a preset ratio $\varepsilon_f/\varepsilon_p = q$ (see Fig. 3). Then from eq. 9:

$$\frac{\vartheta U_\delta(t_{ex})}{a} + \frac{q U_f(t_{ex} - \tau)}{a} = \frac{1}{\varepsilon_p(\tau)}. \tag{10}$$

By measuring ε_p at various values of τ, the wanted function $U_f(t_{ex} - \tau)/a$ can be found in principle by plotting $1/\varepsilon_p$ against $-\tau$. The varying second term in eq. 10 is superimposed on the constant first term. If the test stimulus is also a short rectangular flash of the same duration as the probe flash, eq. 10 is simply:

$$\frac{U_\delta(t_{ex})}{a} + q\frac{U_\delta(t_{ex} - \tau)}{a} = \frac{1}{\vartheta \varepsilon_p(\tau)}, \quad q \ll 1. \tag{11}$$

This is illustrated in Fig. 4.
According to eq. 11 there is a linear relationship between ε_p^{-1} and $U_\delta(t_{ex} - \tau)/a$. In practice, however, measuring the response according to eq. 11 has one serious disadvantage. Apart from the intrinsic spread due to the sampling procedure in determining the 50% thresholds, there is also a slow and relatively large sensitivity drift within and between sessions. (This is reflected for instance in the drift of repeatedly measured thresholds of a single flash, examples of which will be given later.) Both drift in the amplitude criterion a (or signal-to-noise ratio) and metabolically caused changes in $U_\delta(t_{ex})$ are likely candidates as a source of this drift. However, without loss of generality we shall attribute the drift to the former, stating:

$$a = a(t).$$

Since it takes time to do the measurements, special precautions have to be taken against the concealing effect of these variations. This can be done by means of a sensitivity reference as will be shown below.
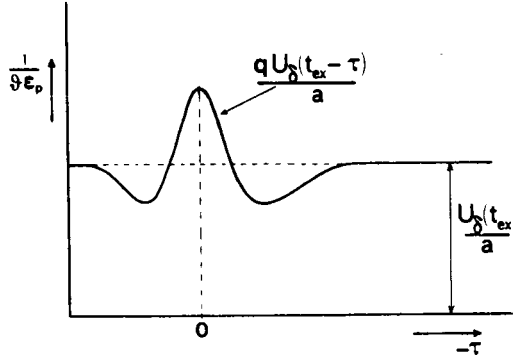
Figure 4: Illustration of how the impulse response might be extracted from measurements following eq. 10 if no sensitivity drift were to occur.

Two drift-correcting techniques

Two practical methods were investigated:

(a) The "slope" method

Differentiating eq. 11 with respect to q one obtains:

$$\frac{U_\delta(t_{ez} - \tau)}{a} = \frac{1}{\vartheta} \frac{\partial}{\partial q} \{\varepsilon_p^{-1}(\tau)\}. \tag{12}$$

The right hand side contains only experimental values. In practice, a series of slightly different q-values around $q = 0$ are used. Fig. 5 illustrates the principle for each value of τ; the values of $1/\varepsilon_p$ determined in the shortest possible time interval $[t_i, t_i + \Delta t_i]$ are plotted against q. In Fig. 5 we are thus probing along line $A - A'$.
Experimental data can be found in Fig. 9. If Δt_i is sufficiently small, drift can be neglected within this interval. From the slope of the line, $U_\delta(t_{ez} - \tau)/a(t_i)$ is calculated with eq. 12. the norm factor of eq. 5, using the intersection point $1/\varepsilon_p$ at $q = 0$, serves as a reference. It represents the sensitivity at time t_i. In Fig. 5 this is symbolised by arrow B. In determining the shape of the response, the effect of drift can now be reduced substantially by normalising eq. 12 with the norm factor

$$U_\delta^\star(t_{ez} - \tau) = \left[\frac{U_\delta(t_{ez} - \tau)}{a}\right]_{t_i} \left[\frac{a}{U_\delta(t_{ez})}\right]_{t_i} = \frac{1}{\varepsilon_p(t_i)} \frac{\partial}{\partial q}(\varepsilon_p). \tag{13}$$

After having finished the experiment for all chosen values of τ at different t_i's, the absolute value of the impulse response $U_\delta(t_{ez} - \tau)/a$ can be estimated by multiplying $U_\delta^\star$ by the norm factor averaged over all times t_i. In a manner of speaking we have shifted the effect of drift from the shape-determining procedure to the scale factor, where it hurts less and can be averaged out.
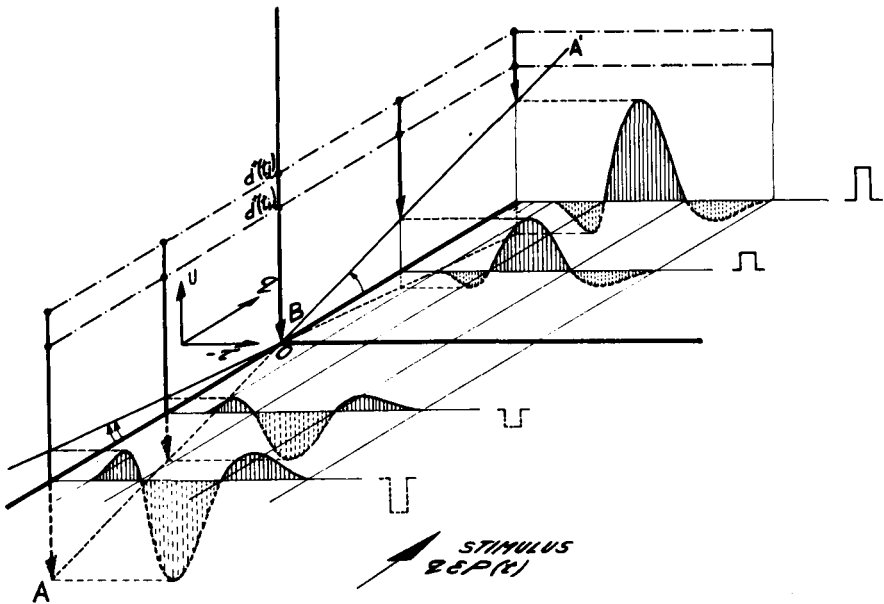
Figure 5: Illustration of the "slope method". The imaginary response of the system to various incremental or decremental short flashes, $q \varepsilon p_\vartheta(t)$, with a fixed probe-stimulus time interval. Negative q-values (see text) are associated with decremental flashes. The response amplitude at a fixed interval τ between probe and stimulus onset is probed at a time t_i with respect to the "amplitude ceiling" $d^+(t_i)$. The probing is symbolised by the arrows. The larger the amplitude of the response of the test stimulus at a given τ, the larger the slope found. If the measurements are repeated at a later time t_j, The "ceiling" has been drifted upwards but the slope is not affected provided there is no considerable drift in the time needed to do the measurements.

(b) The "Method of Pairs"

As a special case of the foregoing we usually used only two q-values, the accompanying probe flash thresholds being measured immediately after one another. The "fast pairs" have the advantages of a short execution time, symmetry in sampling strategy and simplicity of handling.

Their use will be elucidated only for the special case of an impulse response acting as a perturbation function. The derivation for the step response is analoguous. For simplicity, we will choose probe and test flash to have the same duration. The formulas for any other type of perturbation can be derived along the same lines. Suppose we use two q-values $q[q_1, q_2]$. From eq. 11 we obtain:

$$\frac{U_\delta(t_{ex})}{a(t_i)} + q_1 \frac{U_\delta(t_{ex} - \tau)}{a(t_i)} = \frac{1}{\vartheta \varepsilon_{p_1}}. \tag{14}$$

and

$$\frac{U_\delta(t_{ex})}{a(ti + \Delta t)} + q_2 \frac{U_\delta(t_{ex} - \tau)}{a(t_i + \Delta t)} = \frac{1}{\vartheta \varepsilon_{p_2}}. \tag{15}$$

If Δt is sufficiently small, then $a(t_i) \approx a(t_i + \Delta t)$. By elimination we obtain from equations 14 and 15:

$$\frac{U_\delta(t_{ex} - \tau)}{a(t_i)} = \frac{1}{\vartheta(q_1 - q_2)} \left(\frac{1}{\varepsilon_{p_1}} - \frac{1}{\varepsilon_{p_2}} \right). \tag{16}$$

The norm factor is:

$$\frac{U_\delta(t_{ex})}{a(t_i)} = \frac{1}{\vartheta(q_1 - q_2)} \left(\frac{q_1}{\varepsilon_{p_2}} - \frac{q_1}{\varepsilon_{p_1}} \right). \tag{17}$$

In order to obtain the response shape we again use the normalised expression, dividing 16 by 17:

$$U_\delta^\star(t_{ex} - \tau) = \frac{\varepsilon_{p_1} - \varepsilon_{p_2}}{q_2 \varepsilon_{p_2} - q_1 \varepsilon_{p_1}}. \tag{18}$$

In practice it is often convenient to simplify further. For istance by taking $q_1 = q$: $q_2 = -q$, meaning that the test flashes of the pairs of combinations are either an increment or decrement flash of equal amplitudes. Then eq. 16 simplifies to:

$$\frac{U_\delta(t_{ex} - \tau)}{a} = \frac{1}{2\vartheta q} \left(\frac{1}{\varepsilon_{p_+}} - \frac{1}{\varepsilon_{p_-}} \right), \tag{19}$$

where ε_{p_+} and ε_{p_-} are the thresholds of the positive and negative test flashes in combination with the probe.
Eq. 17 becomes:

$$\frac{U_\delta(t_{ex})}{a} = \frac{1}{2\vartheta} \left(\frac{1}{\varepsilon_{p_+}} + \frac{1}{\varepsilon_{p_-}} \right), \tag{20}$$

and eq. 18 is:

$$U_\delta^\star(t_{ex} - \tau) = \frac{1}{q} \left(\frac{\varepsilon_{p_-} - \varepsilon_{p_+}}{\varepsilon_{p_+} + \varepsilon_{p_-}} \right). \tag{21}$$

The unit impulse response, $U_\delta(t_{ex} - \tau)/a$, can be found again by multiplying the normalised unit by the average norm factor.
The response $\varepsilon_f U_f(t)/a$ of an arbitrary time function $\varepsilon_f f(t)$ can be derived with the aid of eq. 6, and its threshold ε_f can be predicted with eq. 2.
It is clear that this method only functions properly if the 50% thresholds of the pair elements are measured consecutively and sufficiently fast to make the effect of drift negligible. This implies a limited number of trials for each psychometric function. Precision can be improved by repetition and averaging the data obtained after applying equations 13 and 18. The effect of residual drift in the time interval needed to determine the thresholds of the pairs can be decreased by measuring the repetitions in counterbalanced order. The use of a reference implies that data of normalised responses, obtained at different sessions, can be averaged.
The perturbation method has two consequences for the measured responses:

(a) The position of the response on the time axis relative to stimulus onset is not known, because all points are measured relatively to t_{ex}.

(b) In case of the large stimuli, where there is no difference between the perceptual attributes of decremental and incremental flashes at threshold, there is a mirror ambiguity of the response with respect to the zero axis, since a in equations 5 and 11 might be d^+ or $-d^-$, depending on the sign of the extremum of $U_\delta(t)$. For small stimuli the response peaks of incremental flashes may meaningfully be called positive since they always give rise to brightness increments.

2.3 Apparatus and procedure

Apparatus

The stimulus was either a centrally fixated circular field of 1 deg, having a dark surround, or a foveal point source of 0.8 arc min on an 11 deg background. It was seen in Maxwellian view through an artificial pupil with a dia of 2 mm, provided with an entoptic guiding system to check the centre of the pupil of the eye (Roufs, 1963). The light was generated by a linearised RCA glow modulator, operated around a suitable working point (13 mA). The luminance of the background was set by attenuating "the working point luminance" by means of a neutral filter. The modulation of the background luminance was controlled electronically by function generators. The amplitude of the desired function could be quickly adjusted with a dB step-attenuator. The calibration of the dynamic stimuli was checked before every session by means of a photomultiplier tube, properly corrected with respect to spectral sensitivity. In the case of the point source superimposed on the background, the working point had to be taken very low in spite of a heavy neutral filter. Consequently, the light had to be monitored continuously in order to correct for temperature effects. The color was practically white. The background level was kept constant by keeping the working point current constant during the session. Its light output was checked before and after each session. The ratio q between test and probe flash, when applying perturbation (see Methodological concepts), was set in the way shown in Fig. 6.
The subject had one knob to release the stimulus, which was delayed for a convenient time interval. The beginning of the stimulus was marked by an acoustic signal. Three buttons enabled him to answer with "yes", "no" or "rejection" (when, for instance, the subject had blinked , moved his eyes, or in general was distracted from the stimulus in any way).

Procedure

In all cases the subject was dark-adapted for 30 min, and subsequently adapted for 5 min to the background luminance. The 50% detection threshold of the modulation was determined by means of a modified method of constant stimuli, as follows. For a certain modulation amplitude, 10 or 20 identical stimuli (depending on the experiment) were presented successively and the detected percentage was determined.
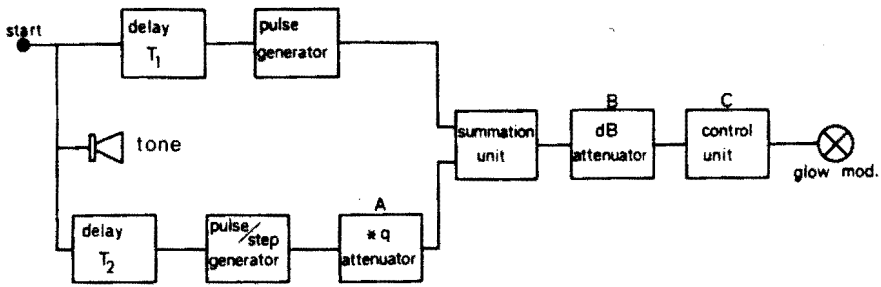
Figure 6: Block diagram of the circuitry applying the perturbation technique. The variable is $T_2 - T_1$ while the attenuation of A is such that the ratio of test and probe stimulus equals q. Attenuator B controls the amplitude of the combination. Unit C controls and linearises the glow modulator.

The dB attenuator was readjusted and the detected fraction was again determined. On average, 4 amplitudes taken in random order were needed to get sufficient data between 20% and 80% detection chance for approximating the psychometric function on a dB scale by a straight line (Roufs, 1974c).
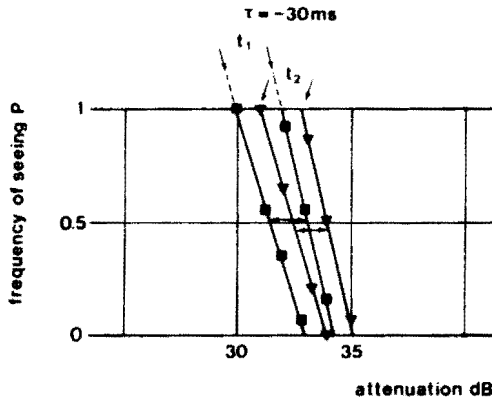
Figure 7: Examples of two pairs of psychometric functions associated with two identical stimulus pairs measured at different times t_i (impulse response, U_δ^*, $q = 0.15$, $\tau = 30$ ms, $E = 1200$ Td). The squares are measured at t_1, the triangles at t_2.

Immediately afterwards, in order to minimise the effect of non-stationary sensitivity changes, a different stimulus to be compared with the first was presented, following an identical routine. In most cases, the measurement of a fast pair was repeated an even number of times in counterbalanced order. If more than two different stimuli were to be compared, the whole set of stimuli was first completed in random order and then repeated in the reverse order and so on. The number of trials was about

800 per 30 min. Every 30 min or so, a pause of 15 min was interposed. The subject was not informed of either the type of stimulus or of its amplitude. Other relevant details or deviations from this procedure will be given in Results.

Two of the subjects were the authors. Subject FB had no previous experience in psychophysical work. He was 27 years old at the time of the experiments, had normal acuity, but was medium deuteranomalous. Subjects JP and JAJR were well trained and were respectively 29 and 46 yr old. Both had normal vision. JAJR having a slight correction.

2.4 Results

Symmetry and repeatability of measured impulse responses; a methodological reconnaissance

In practice, the usefulness of the perturbation technique strongly depends on the precision of the thresholds. According to the expressions in Methodological concepts, differences in thresholds basically determine the shape of the resulting responses. Due to non-systematic relatively slow sensitivity drifts, to order effects within sessions (probably due to fatigue) and to sudden non-systematic sensitivity shifts and learning effects, the spread between thresholds is considerably larger than within. Consequently, threshold variance is not simply inversely proportional to the number of trials constituting the samples of the psychometric functions (Roufs, 1974a,c).
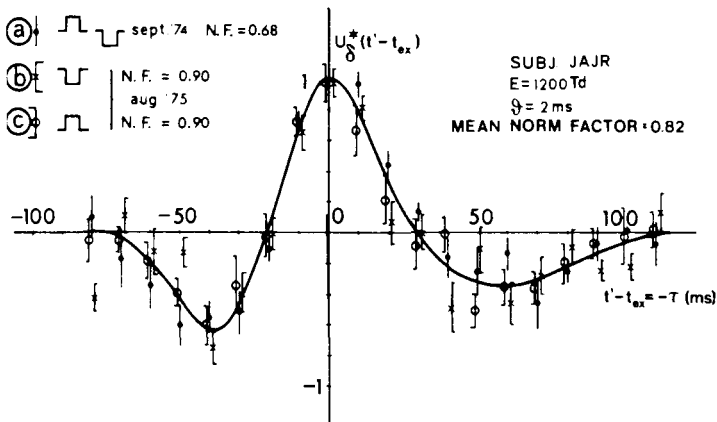
Figure 8: Three normalised impulse responses of subject JAJR, using rectangular pulses of 2 *ms* superimposed on a 1200 *Td* background. Curve fitted by eye. The vertical bars between the shrivels represent twice the Standard Deviations of the Mean (SDM).

To optimise precision, the sample strategy has to be chosen carefully. Important is

the rapidity of measurement. Looking for an adequate design, we tried out several strategies and these will be discussed further on.

Fig. 8 shows three normalised impulse responses of subject JAJR, using rectangular pulses of 2 ms, superimposed on a 1200 Td background. In experiment (a) the response was found by using the slope method (equations 12, 13).

q was changed in five small steps of 0.1. In this particular case the order of measurement was $q[0.1; -0.1; 0.2; -0.2; 0]$, implying partial counterbalancing. For every value of q a 50% threshold was determined (see apparatus and procedure) based on about seven fractions of 10 trials. The complete sequence of q's was repeated twice and the three thresholds at every q were averaged. Linear regression was applied to fit a straight line through the points as a function of q (see Fig. 9). The intersection point of this line with the ε_p^{-1} axis provides a good estimate of the sensitivity ε_p^{-1}. For a single 2 ms flash the norm factor $U_\delta(t_{ex})/a$ was found by using eq. 5 and averaging over all "slopes".
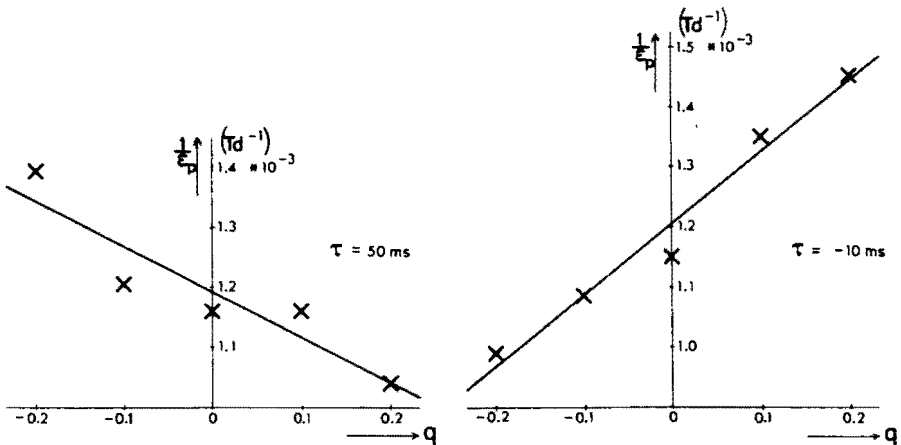
Figure 9: Two representative examples of sensitivity $1/\varepsilon_p$ as a function of the intensity ratio q. At the left the interval is $\tau = 50$ ms, right: $\tau = -10$ ms. The sign of the slope reflects the sign of U_δ^* according to eq. 13.

The complete experiment involved 9560 trials, distributed over seven sessions. The order of measurement on the τ-axis was random, as will be the case in all the following experiments. The mean slope of the 285 psychometric functions was -0.29 dB^{-1} (cf. Fig. 7), corresponding to a Crozier quotient $\sigma/\varepsilon_1 = 0.16$. Here ε_1 is the intensity increment of a single rectangular flash, with a detection probability of 50%. The symbol σ stands for the standard deviation of the pobability density function, which is the derivative of the psychometric function (see Roufs, 1974c).

In order to check the repeatability over longer periods of time, the experiment was done again about 1 yr later. At the same time we wanted to verify whether an incremental

and a decremental flash would give identical responses except for the signs, which is to be expected in the case of linear processing. Fig. 8(b) and (c) show the normalised impulse responses to decremental and incremental flashes respectively, of which the negative and positive norm factors are given in the legend. (As the probe flash was always incremental, the subject detected an incremental flash in all cases.) Note that the different periods of measurement result in different average norm factors. The points were calculated with eq. 18. In order to suppress the effect of sensitivity drift as much as we could and at the same time measure responses to decremental and incremental flashes in consecutive pairs, a very fast method was used in this case.
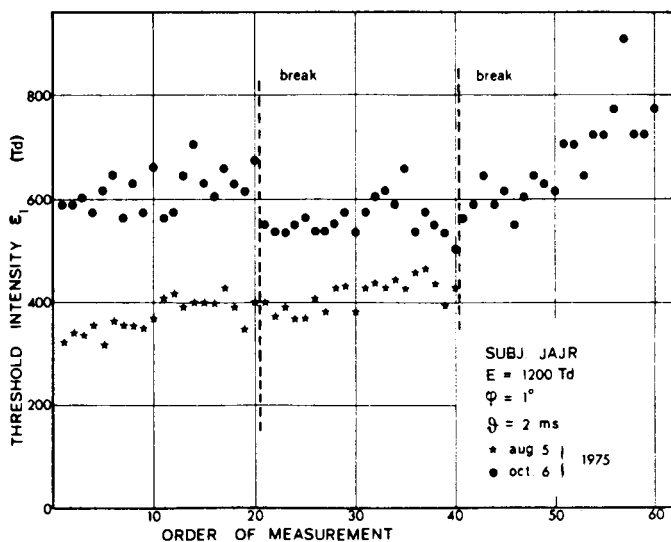
Figure 10: The 50% threshold of single rectangular flashes of 2 *ms* plotted against the order of measurement for two days. The result shows relatively large interthreshold spread superimposed on a systematic order effect. Note the partial recovery after the breaks and the excessive order effect when the subject becomes exhausted. Note also the relatively large difference between the two extreme days.

In experiments 8(b) and 8(c), we used only two q-values, $q[-0.3;0]$ and $q[0.3;0]$, respectively. Moreover, at every q-value, only one perceived fraction of 20 trials was used to determine the 50% threshold, provided this fraction was between 20% and 80%. This was done by using a constant slope, based on an a priori Crozier quotient of 0.19 for subject JAJR, arrived at from an average of many earlier measurements. Starting with $q = 0.3$ and $q = 0$, we took $q = -0.3$ and $q = 0$ immediately afterwards. To increase precision, this was repeated nine times and the pair differences were averaged. Experiments (b) and (c) each took 8000 trials during nine sessions.

Fig. 8 shows that: (a) the results are repeatable within experimental error, and (b) the shape of the responses of increments and decrements in normalised form is not

significantly different.

This supports the linearity hypothesis. It has to be kept in mind that, as said before, we do not know which of the responses (b) or (c) has a positive extreme value. Anyhow, the sign of the extremum of (b) is opposite to that of (c).

In order to predict thresholds of any type of fast modulation, a continuous response function is wanted and also a normalisation factor has to be known. To this end, a continuous curve was fitted by eye through all experimental points of Fig. 8. The overall averaged norm factor is the mean of the absolute value of the three separate norm factors. The typical shape of the response, the total time integral being about zero, supports the hypothesis of the band-pass filter processing for this stimulus condition (1 deg. dark surround).

The large number of thresholds of identical stimuli measured during one session provides statistical information of systematic threshold drift during one session.

Fig. 10 shows 50% thresholds ($q = 0$) obtained during experiments (b) and (c). A systematic increase of threshold is clear over a prolonged period of measurement. The relatively strong increase during an (exceptional) third session suggests fatigue as a probable cause. Fig. 10 also shows by way of example how large the differences between two days can be in extreme cases.
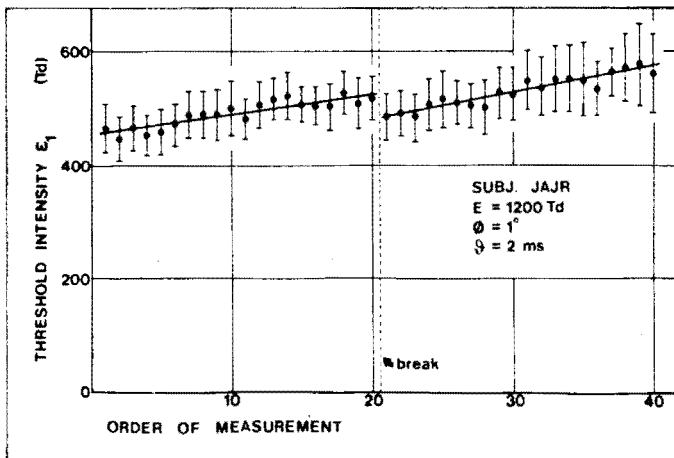
Figure 11: Mean thresholds of single rectangular 2 *ms* flashes, averaged over nine sessions, plotted against number of measuring order. The intervals between the two shrivels equal two SDM.

In Fig. 11 thresholds of nine sessions are averaged over each number of order. The small spread in the standard deviation of the means again reflects the overruling effect of interday variability in this case. The straight lines, obtained by linear regression have slopes of 2.2 Td/min and 2.9 Td/min respectively.

Relation between impulse and step response to a 1 deg field

An effective test of linearity can be found in the relation between step response and impulse response (the latter ought to be the derivative of the former).
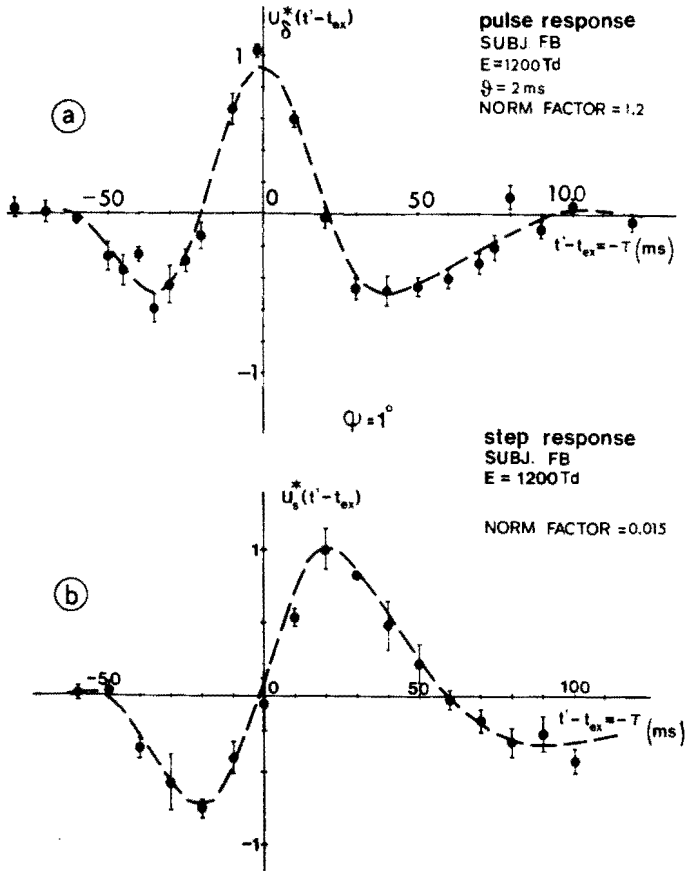
Figure 12: Figure (a) shows the normalised impulse response. The dots are the mean values obtained after reduction by dividing the response by its extreme value. This value is given in the legend as norm factor. Figure (b) is the normalised step response. The dashed curves are the results of a simultaneous fitting, the pulse response being the exact derivative of the step response. As a result of the fitting procedure, the extreme value of the impulse response curve does not quite equal 1.

In actual experiment, peak detection and linearity are in fact tested in combination. Fig. 12(a) shows an impulse response of subject FB obtained with a rectangular test flash of 2 ms at a 1200 Td background. The response was calculated with eq. 18 from two thresholds, q being 0.15 and -0.15. The calculated values were averaged

over five pairs, measured in counterbalance and involving 10 psychometric functions each having about four fractions of twenty trials. Per session two or three different points on the τ-axis were measured. The whole experiment took 21140 trials.

Fig. 12(b) shows the step response obtained for subject FB under identical conditions. Due to the much smaller threshold for a step, the q-values used here were 0.02 and -0.02. In this case all pairs on the τ-axis were measured in one session and averaged afterwards over five sessions. To complete the experiment, 14280 trials were needed. The dashed curves are the result of a simultaneous computer fitting of both pulse and step results, under the condition that the impulse response is the derivative of the step response.

The fair fit of the curves shows that within experimental error this condition is satisfied. The norm factors are calculated as before. For this subject too, order effects were found, comparable with those shown in Fig. 11. In addition, the averages of the single-flash thresholds within sessions reflect some effect of training (see Fig. 13).
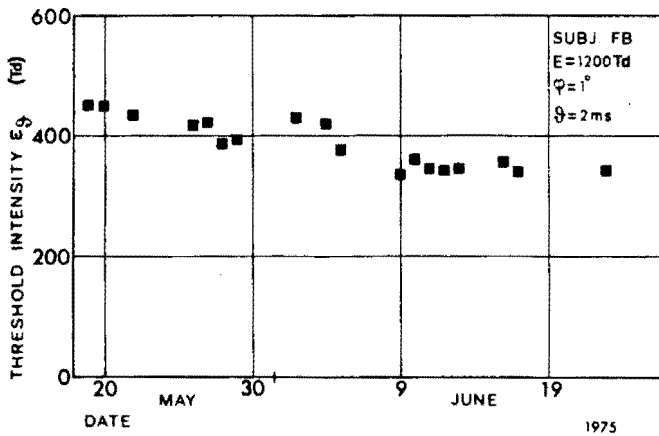
Figure 13: Daily averages of twelve to thirteen 50% thresholds of single rectangular 2 *ms* flashes derived from pairs according to eq. 17 and plotted against the date for subject FB showing a training effect.

In Fig. 14 the effect of the technique is demonstrated in graphs. For graph A only positive q-values are used, while for B, negative ones are used to calculate the response, as described in the legend. For graphs C and D consecutive pairs are used. The improvement is evident.

In order to investigate the suitability of the sample strategy used, Table 1 compares the predicted spread with the experimental findings. In row 1 the standard deviation of the mean of five values of the measured $U_\delta(t_j)$ averaged over all j's is shown. The data in the second row are calculated from the a priori SD's of the thresholds, applying Gaussian error propagation to formula 18. The a priori SD's of the 50%
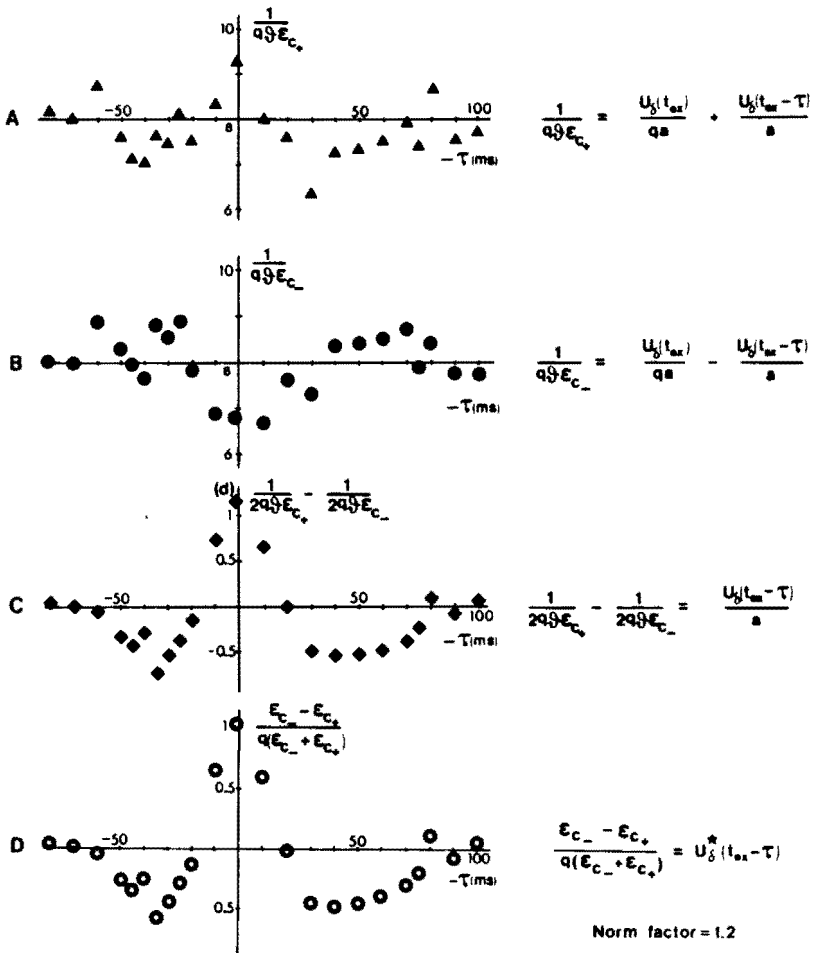
Figure 14: Four steps to demonstrate the effect of a reference on the accuracy of the measured pulse response. The graphs A and B are the average responses to incremental and decramantal flashes calculated with eq. 8, taking negative q's for the decremental flash. Graph C shows shows the impulse response calculated from consecutive pairs and using eq. 17. Finally, graph D shows the normalised response and the average norm factor. This response is used in Fig. 12(a).

Table 1

Standard deviation of	Pulse response $U_i^*(t_j)$	Step response $U_s^*(t_j)$
Results from the experimentally determined threshold pairs	0.07	0.10
Results from the threshold pairs, calculated a priori from the experimental slope of the psychometric function with eq. 19	0.06 (based on an average within-threshold value of $s(\varepsilon_1)/\varepsilon_1 = 0.03$, stemming from the slopes)	0.09 (based on an average within-threshold value of $s(\varepsilon_1)/\varepsilon_1 = 0.03$, stemming from the slopes)
Results from the experimental single threshold at identical flash intervals measured scattered within sessions[1]	0.18 (the value between thresholds within-sessions is $s(\varepsilon_1)/\varepsilon_1 = 0.08$)	0.21 (the value between thresholds within-sessions is $s(\varepsilon_1)/\varepsilon_1 = 0.07$)
Results calculated from measured threshold spread of singlets scattered all over the sessions	0.24 (the value between sessions is $s(\varepsilon_1)/\varepsilon_1 = 0.08$)	0.22 (the value between thresholds is $s(\varepsilon_1)/\varepsilon_1 = 0.02$)

[1]the number of trials for a singlet is equal to the number of trials used for a pair

thresholds are calculated with:

$$s(\varepsilon_p) = \frac{\sigma}{\sqrt{N \sum W_j}}. \tag{22}$$

Here N is the number of trials per sample, W_j are the Müller–Urban weighting functions and σ is the SD of the probability density function derived from the slope of the psychometric functions at the 50% point (Roufs, 1974c). To estimate these values we take the average slope of the psychometric functions and the average number of samples having fractions between 20% and 80% in the experiments. The actual experimental data and the predicted values are very close.

Since stastistical data of the norm factors obtained by the individual pairs are available, the spread in the results if no references were used can be calculated. The effect is demonstrated for singlets in row 3 if only measurements stemming from one session are used, in row 4 if data stem from different sessions.

Perceptual phenomena in connection with 1 deg fields

In all cases so far, the detection criterion was a typical change in the visual field termed "agitation" and not an increment in brightness.

Band-pass filter processing and de Lange characteristics

In Fig. 15(a) and (b) the gain curves of the subjects are shown as dashed lines. These are found by taking the modulus of the Fourier transforms of the impulse responses of Figs. 12(a) and (8) respectively.
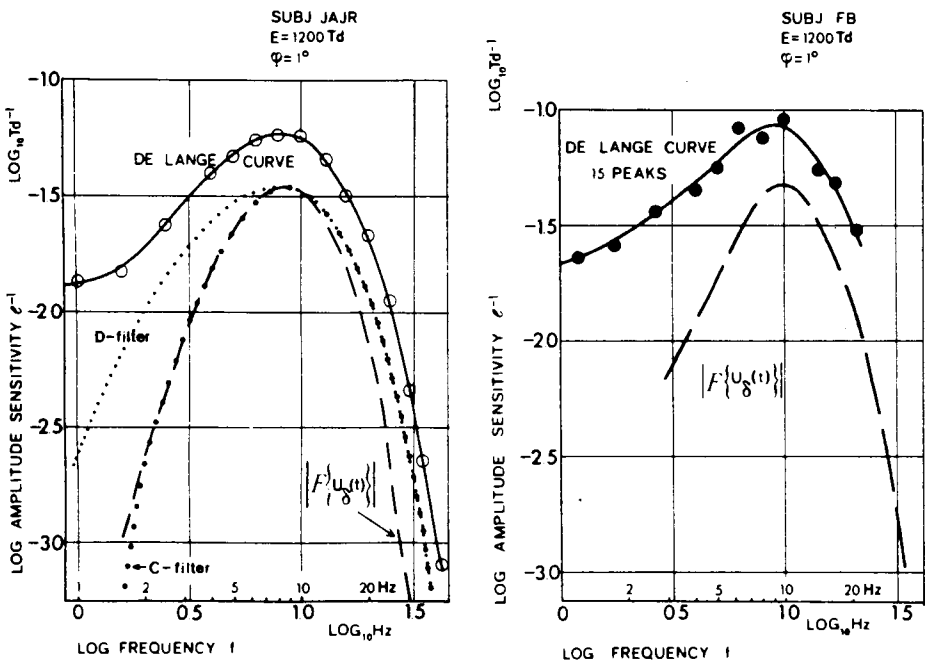
Figure 15: The absolute value of the Fourier transforms of the unit impulse response of two subjects are shown as a function of frequency (dashed lines). For comparison, measured amplitude sensitivities are plotted (continuous lines). For details see text.

The predicted gain curves have a band-pass filter character. This supports an earlier suggestion that for relatively large stimuli the perceptual attribute "agitation" which accompanies transients is linked with a separate variable (Roufs, 1974a). The fine and widely dotted lines in Fig. 13(b), belonging to the D and C filters taken from that article and Roufs (1974c) are calculations of the gain of this variable based on entirely different measurements. The C-type filter especially looks rather close. For comparison the de Lange flicker fusion curves are shown for the same subject and

for the same stimuli, although measured quite some time earlier. The experimental amplitude sensitivities are the reciprocals of the 50% thresholds obtained from psychometric functions of gated sinusoids. Subject JAJR terminated each trial himself and therefore the number of peaks was determined by his decision time, whereas subject FB was presented with exactly 15 fully fledged peaks in order to favour the same 'probability summation" effect for all frequencies. This limits the range at high values since the time that the gate is open becomes too short with respect to its slopes. The experimental amplitude sensitivity curve as a whole is situated about 0.25 log units upwards with respect to the one calculated from the impulse response, which is caused by "probability summation", and is consistent with the known Crozier quotient for single flashes (Roufs, 1974c). On top of that the flicker fusion curves at low frequencies tend to a horizontal asymptote, whereas the gain derived from the impulse response is almost zero at these frequencies. This supports the idea of the two channels, as explained in the introduction and illustrated in Fig. 1, the frequency content of an impulse response being mainly transferred in the high-frequency band-pass channel.

The effect of background luminance

Apart from a scale factor, the threshold curves of rectangular flashes over a long duration range are identical in shape at highly different background levels (Roufs, 1974a). The model suggests that this could be due to the similarity in shape of the impulse response at these levels, except for a change in time scale (Roufs, 1974a, Fig. 9). In principle this can now be verified. To that end an impulse response of the 1 deg stimulus was measured superimposed on a 2 Td adaptation level. This was the lowest level suited for reasonable measurement and consequently yielded the slowest obtainable response. Unfortunately, in the course of the measurements it became clear that we had taken the level a bit too low, since the subject found his measurements hampered by frequent fading of the image due to Troxler's effect, by irregular appearing clouds of "Eigenlicht" and other disturbing effects which usually accompany low level measurement. The measurements were experienced as very tiring. Probably as a result of all this, the spread was found to be considerably higher than at the 1200 Td level. Yet it seems worthwhile to compare the results. Fig. 16 shows the measured impulse response at the 2 Td level.

It was obtained by the pair method taking 15300 trials. From these trials 196 psychometric functions were obtained giving rise to 98 threshold pairs of which seven response data were averaged for any of the fourteen values of τ. In order to facilitate comparison, the normalised impulse response obtained for the same subject at the 1200 Td level is shown (solid line), the time unit being scaled up by a factor of 2.5.

Given the very large difference in luminance, there is indeed correspondence between these low level measurements and the data obtained in the previous section.
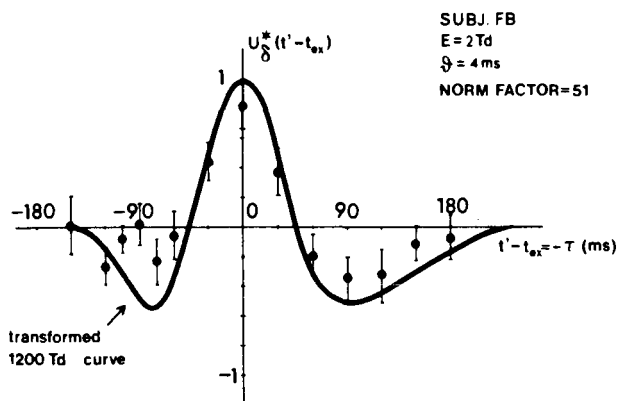
Figure 16: Dots represent the experimentally obtained impulse response on a 2 Td background level at various time intervals τ between probe and test stimulus onset. Bars indicate the SDM. The drawn line is the impulse response obtained at a 1200 Td level, its time unit being multiplied by a factor 2.5.

Impulse and step response of a point source

Impulse and step responses do not change much with background level, except for the time scale, if the stimulus dimension is large and the background is not too low. However, very small stimuli show differently shaped responses, which are more of the sustained type. The perceptual attributes to be detected also change. To demonstrate this, the results for a foveal point source are given in this section. The stimulus dia was 0.8 arc min, situated in the middle of four small and weak red fixation lights, positioned at the corners of a square having 25 arc min sides and superimposed on a 1200 Td 11 deg background, acting also as surround. The experimental techniques, the procedures and the theoretical assumptions were the same as in the foregoing sections.

In the case of the impulse response twelve values of τ were used. The results are averages of eight (50%) threshold pairs taking on the average 60 trials per psychometric function, amounting to 11520 trials in total. From the 192 psychometric functions the average Crozier quotient of $\sigma/\varepsilon = 0.14$ was found.

The measured values of the impulse and step responses are given in Fig. 17(a) and (b). The dashed curves are simultaneous computer fittings of which the impulse response is the exact derivative of the step response. The time scale is somewhat extended with respect to the responses in Fig. 12, which was to be expected since, for instance, the critical duration increases if the diameter decreases (Roufs and Meulenbrugge, 1967; Adler, 1970). The monophasic responses are of the sustained type, typical of low-pass filter action. The system can be described fairly well in
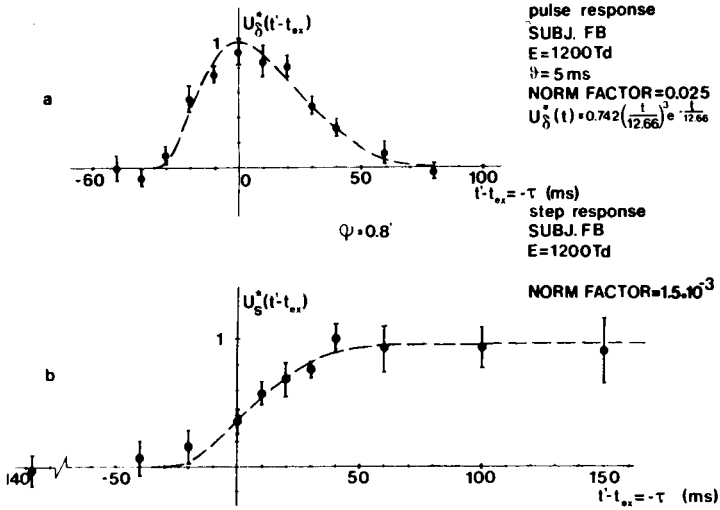
Figure 17: Pulse and step response of a point source on an extended background of 1200 Td. The vertical bars between the shrivels represent twice the SDM. The dashed curves are the results of a simultaneous fitting, the curve of Fig. (a) being the derivative of Fig. (b).

terms of a fourth order real pole. The normalised impulse response is:

$$U_\delta^\star(t) = 0.742\big(\frac{t}{12.66}\big)^3 e^{-\frac{t}{12.66}};$$

$$t \text{ in } ms; \ t_{ez} = 37ms.$$

The absolute value of the Fourier transform, being the gain curve of the system, is given in Fig. 18. The difference from the gain curve of a 1 deg field is obvious, as can be seen by comparing with Fig. 15a.

Perceptual phenomena with the point source

The perceptual attribute to be detected was either a brief or steplike brightness increment, looking homogeneous in space over the small stimulus. This attribute is comparable with the periodic brightness variation ("swell") in the case of low-frequency sinusoidal stimulation with 1 deg fields. In this case, the positive response peaks can without any doubt be attributed to brightness increments.

2.5 Discussion

Evaluation of procedures

The results of the first of the procedures used, the "slope' method described in Results, encouraged the approach. No significant deviations from linear regression were
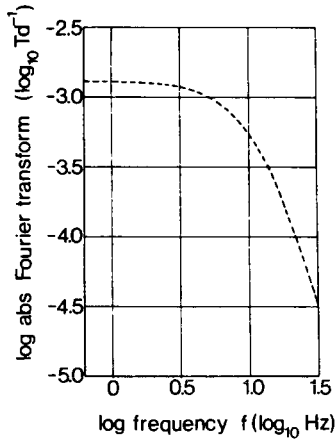
Figure 18: Gain curve of the transfer of a point source, obtained by taking the modulus of the Fourier transform of the impulse response of Fig. 17a.

found over the range of q-values used. The relative amplitudes of the phases of the determined impulse responses in fact suggest that q-values as large as 0.3 would have been allowed, leading to a larger effect and hence to improved experimental precision. The varying slopes of the regression lines, changing within the flash interval τ, and their relatively small standard deviation gave confidence in the method. However, a set of concatenated thresholds still takes so long that, in spite of time efficiency of the procedure, the chance of the results being obscured by sensitivity drifts was still thought not to be negligible. This led to the method of "pairs", allowing very fast measurement of concatenating pairs, although about the same experimental effort is needed to reduce the standard deviation of the averages to the desired level. From the results no clear evidence was found that this procedure is superior for the elimination of drift effects. However, the use of only two q-values proved to simplify the experimenters' task considerably. It is not essential to choose q-values that are symmetrical around $q = 0$, but it has the advantage of giving relatively large "pair differences" without introducing unsafe q-values.

The results of Fig. 14 and Table 1 show that the obscuring effects of drift are virtually eliminated by the use of such a reference. Although the resulting precision and the repeatability over a long period of time are quite satisfactory, the actual design of the experiments is rather laborious. Since about 700 trials were used for every point of the response curve, taking about 30 min in all. It needs some economising. The present set up gave refined information on the sources of variance, especially about the systematic drift, which we did not notice before (Roufs, 1972a). The information given in this article may lead to a more economic design.

A more fundamental type of potential error could have been the width of the dominant phase of the probe flash. The "probing needle" being too blunt. However, this is

obviously not the case. It can be seen by rough calculations that in the actual case the "bluntness error" is negligible in relation to experimental inaccuracy.

Validity of concepts

The detection model is supported by several experimental results: no significant deviation from the linear relation between ε_p^{-1} and q, according to eq. 11, was found. Neither was there any asymmetry when incremental and decremental flashes were used. The strongest support, however, is found in the correspondence between the impulse response and the derivative of the step response as shown in Figs. 12 and 17.

This implies that the threshold of any time function should be predictable by calculating its unit response by convolution and determining the reciprocal of its extreme value (equations 2 and 6). This appears to be the case in all cases we have investigated, as will be shown in a subsequent paper. In the case of stimuli that are considerably larger than point sources, precautions have to be taken to ensure that the function changes sufficiently fast, otherwise we get a mix-up of the two systems referred to in the introduction.

If the sign of the extreme value of the response is opposite to that of the dominant phase of the probe response, the ratio of d^+ and d^- has to be tested by measuring the threshold of a short incremental flash in relation to a decremental one and making use of eq. 5. (This is for instance the case for an incremental rectangular flash of a certain duration, as will be demonstrated in the related article.) In the case of the foveal 1 deg field this ratio was shown to be one (Roufs, 1974a).

In the literature several other models concerning the prediction of thresholds on the basis of systems analysis can be found. In comparing their performance we shall restrict ourselves to the most closely related work. Kelly and Savoie (1978) derived impulse and step responses from a postulated transfer function, the modulus of which was fitted to flicker fusion data, ignoring the difference in perceptual attributes at low and high frequencies. Their model is a linear filter followed by an asymmetric rectifier, which is effectively the same as our d^+ and d^- having different values under their conditions and from their point of view. They compared threshold ratios of an incremental flash of an 8 deg uniform field and of a temporal rectangular modulation of an 8 deg, 4 c/deg grating, both as a function of duration. From their calculated responses they concluded that the system is much more sensitive to negative than to positive response phases. However, neither Rashbass (1970) with a 17 deg field nor Herrick (1956) or Roufs (1974a) with a 1 deg field at comparable background levels found any significant asymmetry. As pointed out before (Roufs, 1974a, p. 836) asymmetry may be found and is even likely under certain conditions.

Rashbass (1970, 1976) proposed an interesting model to predict thresholds from luminance changes of which the essentials are also used by Broekhuijsen et al. (1976) and Watson and Nachmias (1977). In his view, a signal proportional to the luminance change passes successively a linear (band-pass) filter, a quadratic operator and an ideal integrator, operating over a time T which is large with respect to the duration

of the signal. The luminance change is thought to be detected if the output of the integrator surpasses a certain level. As pointed out before (Roufs, 1974a), this extra integrator is the essential difference between the models since a symmetric nonlinearity without integrator does not change the result. What would our experimental results mean in the Rashbass model? In Appendix A3 it is shown that our data of the 1 deg field are not quite as expected from his model, although the deviations are not dramatic. The point source results, however, deviate considerably, but this is to be expected if one accepts our step response, since it does not vanish for large times (which is required if the Rashbass model is used).

Hallett (1969a) determined impulse and step responses of the rod system. His model is almost identical with ours, but his detection criterion and consequently his technique is quite different. Detection is based on a relation between the signal amplitude threshold a at the output of the linear filter and the background level in order to obtain a detectable signal-to-noise ratio, assuming that all noise at the output is filtered quantum noise caused by the background. He increased the level of his 18 deg background impulsewise or stepwise in time and measured the discrimination threshold of a small (12 arc min) short flash in its centre as a function of the relative time shift of test flash and background. Hallet pointed out that non-stationary threshold variations prevented precise determination of the response (see also Hallett, 1969b). This might be overcome by using a reference as was done in the present work. Hallet found no satisfactory relation between impulse and step response. It is not clear whether this is due to the insufficient precision of his measurements or to the properties of the detection stage in his model.

Masking with a flashed background has been used in the fovea even earlier (e.g. Crawford, 1947; Boynton and Kandel, 1957; Sperling, 1965). Application of Hallett's model to the fovea is hampered by the consideration that, especially at the relevant levels, there is no square-root relation between intensity and threshold, indicating that photon noise is not a likely candidate as a noise source at the output of the linear filter. This is also unlikely on the basis of the shifts of the amplitude sensitivity curves of flicker experiments found as a function of the background intensity (Ruddock, 1969). On the contrary, the constancy of the ratio of the standard deviation of the density function associated with the psychometric curves and the 50% threshold itself over about 5 decades, strongly suggests that this noise is independent of background luminance (Roufs, 1974c).

Ikeda (1965) and Uetsuki and Ikeda (1970) used subthreshold summation based, although not explicitly, on the same assumptions. Their responses look similar to ours, apart from small details which might be caused by the fact that in their method the test response is not always small with respect to the probe response, so that different phase combinations may determine threshold.

Tolhurst (1975) also used subthreshold summation of gratings as a tool to separate the responses of the transient and sustained system. Norman and Gallistel (1977) looking for the impulse response as a basic function assumed the same systems properties. However, since they did not use the possibility of perturbation, they had

to postulate an extra set of properties about the shape of this function.

The nature of transfer

The band-pass character of the transfer of the output variable giving rise to "agitation" in the case of a 1 deg stimulus with a dark surround is confirmed. As a consequence, the impulse response has a multiphasic character. Comparing the modulus of the Fourier transform of the impulse response function with de Lange curves, shifted downwards 0.25 log units for reasons mentioned in Results, the gain at low frequencies is still significantly less than the de Lange curves would suggest. This gain curve can also be measured directly using the same technique with sinusoidal perturbation frequencies (Roufs and Pellegrino van Stuyvenberg, 1976). The results show also a definite band-pass character as will be shown in detail in a subsequent paper.

For small stimulus diameters this transient activity, described as the band-pass filter action, disappears as does the "agitation" percept. At low background levels on the other hand, it does not vanish so readily as the shape of De Lange curves would indicate (Roufs, 1972a). This difference between small and large stimuli suggests that band-pass filter action which accompanies "agitation" has to do with lateral interaction. Psychophysical evidence for sustained and transient activity in the human visual system was found for instance by Tolhurst (1975), Kulikowski and Tolhurst (1973), Johnson and Enoch (1978) and Philips and Singer (1974). Implications were analysed by Breitmeyer and Ganz (1976). There is of course overwhelming physiological evidence for sustained and transient activity in the retina and more central locations (Breitmeyer and Ganz, 1976). Also the type of sustained transfer we found for the point source (four RC networks in cascade) is known physiologically, for instance for the rod receptor in the bullfrog (Toyoda and Coles, 1975). Likewise, there is physiological evidence that lateral interaction in the case of larger stimuli results in more transient-like responses (Detwiler et al., 1978). However, we prefer not to take arguments of this kind into consideration because we find it premature and philosophically not quite correct to make detailed connections between physiological evidence and psychophysical data. In finding new support for the two output variables connected with threshold perception of time-dependent stimuli, the problem of their interrelation comes up. This has still to be investigated, especially in view of suprathreshold brightness variations accompanied by induced spatial inhomogeneities.

Individual differences

Fig. 19 allows intersubject comparison. The curves are essentially the same. Subjects FB and JP have a higher sensitivity (see norm factors) and smaller time constants reflected by the duration of the individual phases. Presumably this is due to age differences (see Apparatus and procedure).

In the literature it is not common to make any separate allowance for the effect of age on time constants and sensitivity. The response properties shown here result, as
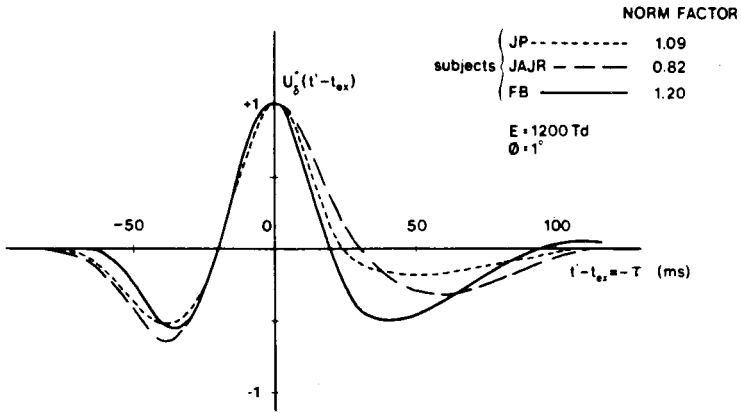
Figure 19: Comparison of normalised unit impulse responses of three subjects.

pointed out earlier (e.g. Roufs, 1974b), in an increase in reaction time and a decrease in critical flicker fusion frequency for older subjects. This is a common finding in literature (e.g. Weale, 1963).

Conclusions

Internal responses of one-shot functions can be determined by means of a drift-correcting perturbation technique with sufficient accuracy to allow quantitative analysis.

The results of impulse and step responses are consistent with quasi-linear processing of weak signals and peak detection. This implies that the threshold of any sufficiently fast transient should be predictable. (This will be confirmed for the tests in the related paper.)

The shape of the impulse response does not change essentially with the Plateau level, but the time scale does. On the other hand, if the stimulus dimensions become very small, the shape changes essentially.

For 1 deg foveal stimuli with dark surround, the system processes one-shot stimuli as a band-pass filter. The responses are of a transient type. For a 0.8 arc min point source on a large homogeneous background the processing is purely low-pass filtering, giving rise to sustained responses.

Between subjects only minor differences are found.

References

Adler, F.H. (1970) Adler's Physiology of the Eye: Clinical Application 5th edn. p. 589, Fig. 2101. Mosby, Saint Louis.

Boynton, R.M.; Kandell, G. (1957) On the responses in the human visual system as a function of adaptation level. J. Opt. Soc. Am. 47, 275-286.

Breitmeyer, B.G.; Ganz, L. (1976) Implications of sustained and transient channels
     for theories of visual pattern masking, saccadic suppression and information
     processing. Psychol. Rev. 83, 1-36.
Broekhuijsen, M.; Rashbass, C.; Veringa, F. (1976) The threshold of visual transients.
     Vis. Res. 16, 1285-1289.
Cleland, B.E.; Dubin, M.W.; Levick, W.R. (1971) Sustained and transient neurones
     in the cat retina and lateral geniculate nucleus. J. Physiol. 217, 473-496.
Crawford, B.H. (1947) Visual adaptation in relation to brief conditioning stimuli. Proc.
     Roy. Soc. B. 134, 283-302.
Detwiler, P.B.; Hodgking, A.L.; McNaughton, P.A. (1978) A surprising property of
     electrical spread in the network of rods in the turtle's retina. Nature 274, 562-
     565.
Hallett, P.E. (1969a) Impulse functions for human rod vision. J. Physiol. 202, 379-
     402.
Hallett, P.E. (1969b) The variations in visual threshold measurement.   J.Physiol.
     202, 403-419.
Herrick, R.M. (1956) Foveal luminance discrimination as a function of the duration
     of the decrement and increment in luminance. J. Comp. Physiol. Psychol. 59,
     473-443.
Ikeda, M. (1965) Temporal summation of positive and negative flashes in the visual
     system. J. Opt. Soc. Am. 55, 1527-1534.
Johnson, C.A.; Enoch, J.M. (1978) Human psychophysical analysis of receptive field-
     like properties: IV Further examination and specification of the psychophysical
     transient-like function. Doc. Ophtal. 41, 329-345.
Kelly, D.H. (1961) Diffusion model of linear flicker responses. J. Opt. Soc. Am. 59,
     1665-1670.
Kelly, D.H.; Savoie, R.E. (1978) Theory of flicker and transient responses III. An es-
     sential nonlinearity. J. Opt. Soc. Am. 68, 1481-1490.
Kulikowski, J.J.; Tolhurst, D.J. (1973) Psychophysical evidence for sustained and
     transient detectors in human vision. J. Physiol. 232, 149-162.
Lange, H. de (1952) Experiments on flicker and some calculations on an electrical
     analogue of the foveal system. Physica 18, 935-950.
Levinson, J.Z. (1968) Flicker fusion phenomenon. Science, New York. 160, 21-28.
Matin, L. (1968) Critical duration, the difference luminance threshold critical flicker
     frequency and visual adaptation, a theoretical treatment. J. Opt. Soc. Am.
     58, 404-415.
Norman, M.F.; Gallistel, C.R. (1978) What can one learn from a strength-duration
     experiment? J. Math. Psychol. 18, 1-24.
Philips, W.A.; Singer, W. (1974) Function and interaction of on- and off transients
     in vision I. Psychophysics. Exp. Brain Res. 19, 493-506.
Rashbass, C. (1970) The visibility of transient changes in luminance. J. Physiol.
     Lond. 210, 165-186.
Rahbass, C. (1976) Unification of two contrasting models of the visual incremental
     threshold. Vis. Res. 16, 1281-1283.

Roufs, J.A.J. (1963) Perception lag as a function of stimulus luminance. Vis. Res. 3, 81-89.

Roufs, J.A.J.; Meulenbrugge, H.J. (1967) The quantitative relation between flash threshold and and the flicker fusion boundary for centrally fixated fields. IPO Ann. Progr. Rpt. 2, 133-139.

Roufs, J.A.J. (1971) Threshold perception of flashes in relation to flicker. In: The Perception and Application of flashing Lights, pp 29-42. Hilger, London.

Roufs, J.A.J. (1972a) Dynamic properties of vision - I. Experimental relationships between flicker and flash thresholds. Vis. Res. 12, 261-278.

Roufs, J.A.J. (1972b) Dynamic properties of vision - II. Theoretical relationships between flicker and flash thresholds. Vis. Res. 12, 279-292.

Roufs, J.A.J. (1973) Dynamic properties of vision - III. Twin flashes, single flashes and flicker fusion. Vis. Res. 13, 309-323.

Roufs, J.A.J. (1974a) Dynamic properties of vision - IV. Thresholds of decremental flashes, incremental flashes and doublets in relation to flicker fusion. Vis. Res. 14, 831-851.

Roufs, J.A.J. (1974b) Dynamic properties of vision - V. Perception lag and reaction time in relation to flicker and flash thresholds. Vis. Res. 14, 853-869.

Roufs, J.A.J. (1974c) Stochastic threshold fluctuations and their effect on flash-to-flicker ratio. Vis. Res. 14, 871-888.

Roufs, J.A.J.; Blommaert, F.J.J. (1975) Pulse and step response of the visual system. IPO Ann. Progr. Rpt. 10, 60-67.

Roufs, J.A.J.; Pellegrino van Stuyvenberg, J.A. (1976) Gain curve of the eye to subliminal sinusoidal modulation of light. IPO Ann. Progr. Rpt. 11, 56-63.

Ruddock, K.H. (1969) Some aspects of human visual threshold performance in relation to temporal frequency characteristics. Opt. Commun. 1, 173-175.

Sperling, G. (1965) Temporal and spatial visual masking: I. Masking by impulse flashes. J. Opt. Soc. Am. 55, 541-559.

Sperling, G.; Sondhi, M.M. (1968) Model for visual luminance discrimination and flicker detection. J. Opt. Soc. Am. 58, 1133-1145.

Tolhurst, D.J. (1975) Sustained and transient channels in human vision. Vis. Res. 15, 1151-1155.

Toyoda, J.; Coles, A. (1975) Rod response to sinusoidally flickering light. Vis. Res. 15, 981-983.

Uetsuki, T.; Ikeda, M. (1970) Study of temporal visual response by summation index. J. Opt. Soc. Am. 60, 377-381.

Veringa, F. (1961) Enige natuurkundige aspecten van het zien van gemoduleerd licht. Thesis, Amsterdam.

Watson, A.B.; Nachmias, J. (1977) Patterns of temporal interaction in the detection of gratings. Vis. Res. 17, 893-902.

Weale, R. (1963) The Aging Eye. Lewis, London.

Appendix A1

In Rashbass's language the response $\phi(t)$ of his linear filter to a short rectangular flash at threshold is defined by

$$\int_0^T \phi^2(t)dt = 1. \tag{23}$$

The intensities of two combined flashes at threshold, having equal durations and an interval τ, can be expressed in the threshold of a single flash as a unit. The threshold intensities (A, B) of a pair are given by

$$\int_0^T [A\phi(t) + B\phi(t - \tau)]^2 dt = 1, \tag{24}$$

which leads to the equation of an ellipse:

$$A^2 + B^2 + 2ABL(\tau) = 1, \tag{25}$$

with:

$$L(\tau) = \int_0^T \phi(t)\phi(t - \tau)dt. \tag{26}$$

After differentiating with respect to B:

$$2A\frac{dA}{dB} + 2B + 2L(\tau)\left\{A + B\frac{dA}{dB}\right\} = 0. \tag{27}$$

Since in our experiment one flash is always small compared to the other, we look at the neighbourhood of $(A, B) = (1, 0)$. Eq. 27 becomes:

$$\left(\frac{dA}{dB}\right)_{1,0} = -L(\tau). \tag{28}$$

In our language the threshold condition for the combination is:

$$\frac{\varepsilon_A \vartheta U_\delta(t_{ex})}{a} + \frac{\varepsilon_B \vartheta U_\delta(t_{ex} - \tau)}{a} = 1, \; \varepsilon_A \gg \varepsilon_B. \tag{29}$$

Expressing the threshold intensities also in the threshold of one of the flashes alone, $\varepsilon_A/\varepsilon_0 = A; \varepsilon_B/\varepsilon_0 = B$, and using eq. 5 we obtain:

$$A + BU_\delta^*(t_{ex} - \tau) = 1. \tag{30}$$

κ, for any set of experimental pairs (A, B) in particular $(A, B) = (1, 0)$ we obtain:

$$\left(\frac{dA}{dB}\right)_{1,0} = -U_\delta^*(t_{ex} - \tau). \tag{31}$$
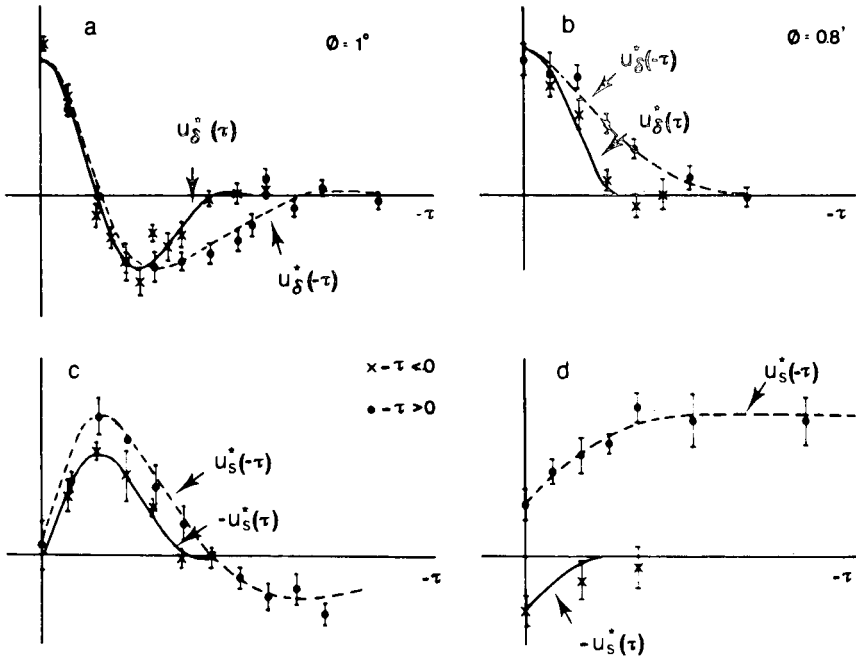
Figure 20: Plots (a) and (b) demonstrate the degree of symmetry around $\tau = 0$, at a 1200 Td background for a 1 deg and a 0.8 arc min stimulus. Plots (c) and (d) demonstrate the degree of asymmetry of the step response under the same conditions.

Appendix A2

Comparing eq. 31 with eq. 28 we see that the impulse response U_δ^* obtained from our material on the basis of our model is equivalent to the autocorrelation function $L(\tau)$ on the basis of Rashbass's model. An autocorrelation is essentially symmetric:

$$L(\tau) = L(-\tau). \tag{32}$$

However, our experimental data do not seem to fulfil this requirement too well as Figs. 20a and b show. On the other hand, the responses of the combination predicted by the present model lead to threshold pairs which mimic rather well the ellipses predicted by eq. 25 as seen in Fig. 21.
The shapes of these curves were basic to Rashbass's model. A combination of a short flash and a step expressed in their respective thresholds has in Rashbass's model the threshold criterion:

$$\int_0^T [C\phi_\delta(t) + D\phi_s(t - \tau)]^2 dt = 1. \tag{33}$$
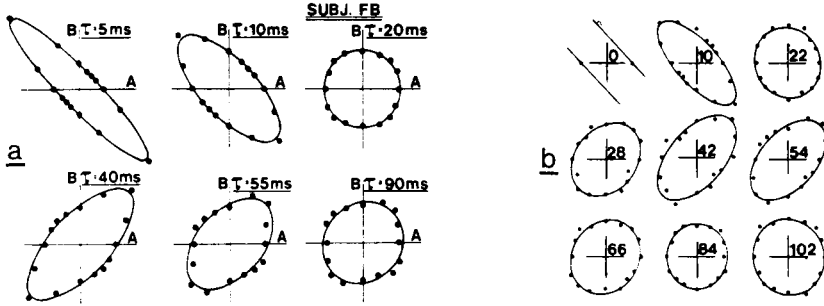
Figure 21: Plot (a) shows thresholds of flash combinations of incremental and decremental flashes predicted by eq. 29 on the basis of the experimentally determined unit impulse response. The ellipses are least square fittings. Plot (b), a copy of Rashbass's Fig. 4 (1970) is given for comparison. The experimental points were obtained according to eq. 25, through which he fitted the ellipses shown.

After some elaboration one obtains:

$$\left(\frac{dC}{dD}\right)_{1,0} = -\int_0^T \phi_\delta(t)\phi_s(t-\tau)dt$$

$$= \int_0^T \phi_s(t)\phi_\delta(t-\tau)dt, \tag{34}$$

$\phi_s(t)$ being the unit step response of the linear filter.
In our model the threshold criterion for the combination of an impulse and step is:

$$\frac{\varepsilon_C \vartheta U_\delta(t_{ex})}{a} + \frac{\varepsilon_D \vartheta U_s(t_{ex}-\tau)}{a} = 1. \tag{35}$$

Using again the thresholds ε_{C0} and ε_{D0} of impulse and step as intensity units, one obtains:

$$\frac{\varepsilon_C}{\varepsilon_{C0}} + \frac{\varepsilon_D}{\varepsilon_{D0}}U_s^\star(t_{ex}-\tau) = 1, \tag{36}$$

or:

$$C + DU_s^\star(t_{ex}-\tau) = 1, \tag{37}$$

and from this:

$$\left(\frac{dC}{dD}\right)_{1,0} = -U_s^\star(t_{ex}-\tau). \tag{38}$$

Appendix A3

The autocorrelation function of the step response, referred to as L_s by Rashbass is:

$$L_s = \int_0^T \phi_s(t)\phi_s(t-\tau)dt. \tag{39}$$

Combining equations 39 with eq. 34 and 38 one obtains:

$$\left(\frac{dC}{dD}\right)_{1,0} = -\frac{d}{d\tau}L_s = -U_s^\star(t_{ex} - \tau),\tag{40}$$

and since $L_s(\tau) \equiv L_s(-\tau)$:

$$\left(\frac{d}{d\tau}L_s\right)_\tau = -\left(\frac{d}{d\tau}L_s\right)_{-\tau}.\tag{41}$$

In our language:

$$U_s^\star(t_{ex} - \tau) = -U_s^\star(t_{ex} + \tau).\tag{42}$$

This means that our step responses should be anti-symmetric around t_{ex}, as defined for the pulse response, according to Rahbass's model. Fig. 20 shows that the experimental values deviate considerably from this behaviour, especially for the point source.

chapter 3

Prediction of thresholds and latency on the basis of experimentally determined impulse responses[1]

Frans J.J. Blommaert
Jacques A.J. Roufs

Abstract

As was shown before (Roufs and Blommaert, 1981), temporal impulse responses and step responses can be obtained psychophysically using a drift-correcting perturbation technique. In this paper, experimentally determined impulse responses are given for eight subjects using different experimental conditions, i.e. a 1 deg stimulus field at background luminances of 1200 Td and 100 Td, and a point source superimposed on an extended background of the same luminances, which is a possibility to separate transient and sustained processing. For a large class of stimuli, predictions of threshold curves and latency of different time functions are calculated on the basis of these measured impulse responses. Predictions are tested against experimental data. It will be shown that a simple model, only consisting of a linear filter followed by a noisy peak detector, suffices for a fair quantitative description of the available data.

3.1 Introduction

In Roufs and Blommaert (1981), further on referred to as 81/1, impulse- and step responses have been described which were determined from thresholds by means of a drift-correcting perturbation technique. These responses were obtained on the basis of only two systems properties viz. quasi-linearity and peak detection, and were shown to be mutually consistent. In this paper, experimentally determined impulse responses will be shown for 8 subjects, including 2 adaptation levels and 2 stimulus diameters. Differences between these responses will be discussed, as will be the effect of noise on the result of these perturbation experiments.

As pointed out in 81/1, the objective of these studies is to increase generalisation by providing means to predict thresholds of arbitrary time dependent stimuli. Using a model, as visualised in Fig. 1, this should at least be possible for the adaptation level on which the impulse response is determined, since the response of any time dependent change of retinal illumination can be calculated from this impulse response by convolution. However, this only applies if the stimulus-surround configuration and the probe are such that one of the two temporal systems, which behave as low-pass or band-pass (Roufs, 1974c; 81/1), is dominant. The resulting threshold can then be found by adjusting the input amplitude at the value for which the extremum of the response reaches the detection level.

---

[1] Submitted to Biological Cybernetics.
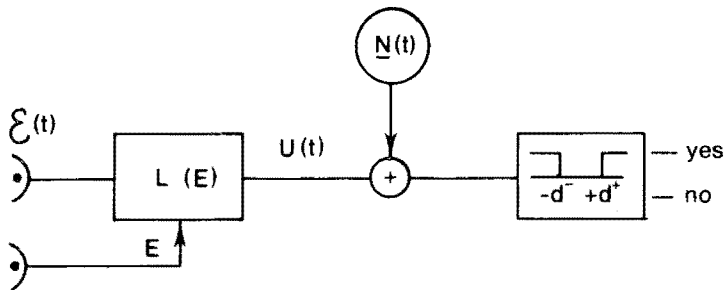
Figure 1: Model used to describe visual processing of time-dependent illuminance changes at threshold level. It consists of a linear operator followed by a noisy peak detector (appropriate for either the band-pass transient-like processing or the low-pass sustained-like transfer).

In the second part of this article, the prediction algorithm will be tested for a number of time functions using a 1 deg stimulus having a dark surround (favourable for the band-pass transient-like processing), and a point source at a large homogeneous background (favourable for the low-pass sustained-like processing).
For certain stimuli, corrections have to be made because of the stochastic nature of the detection process. This is the case if the response is such that prolonged or repeated exposure increases the detection chance noticeably. The resulting decrease of threshold is usually called to be due to "probability summation".
On the basis of the same model, also earlier latency measurements will be compared with predictions calculated from the measured impulse response.

3.2 Theoretical frame

A model

Using a model as depicted in Fig. 1, two deterministic system properties are postulated viz. linearity and peak detection. Small time changes of retinal illumination are processed linearly if:

$$L\{\varepsilon_f f(t)\} = \varepsilon_f U_f(t), \tag{1}$$

where:

L is the linear operator which characterises dynamic visual processing,

$f(t)$ is the normalised time function of the stimulus,

$U_f(t)$ is the response of the linear system to $f(t)$,

ε_f is the amplitude factor.

Stochastic aspects do play a role in detection processes, as was stressed for instance by Watson (1982). Its effects will be discussed later on in this paper and will be treated in detail elsewhere (Roufs, Piceni and Pellegrino van Stuyvenberg, forthcoming). If for a moment the stochastic aspects are neglected, the stimulus $\varepsilon_f f(t)$ is seen if some part of its response deviates at least by a magnitude "a"

from the stationary reference level (peak detection). Thus at threshold:

$$\varepsilon_f extr\{U_f(t)\} = a,\tag{2}$$

$$a = d^+ \text{ or } -d^-; \qquad d^+, d^- > 0.$$

If the extremum happens to be positive, $a = d^+$, otherwise $a = -d^-$.
An important property of linear systems is that the operator L is fully characterised
by its unit impulse response $U_\delta(t)$. From this impulse response, dynamic responses
to arbitrary changes in illuminance can be calculated according to:

$$\varepsilon_f L \left\{ \frac{f(t)}{a} \right\} = \varepsilon_f \frac{U_f(t)}{a} = \varepsilon_f \int_0^\infty f(t - \tau) \frac{U_\delta(\tau)}{a} d\tau.\tag{3}$$

In the following, eq. 3 will be used to calculate the dynamic time course of visual
responses on the basis of experimentally determined impulse responses.

Perturbation approach
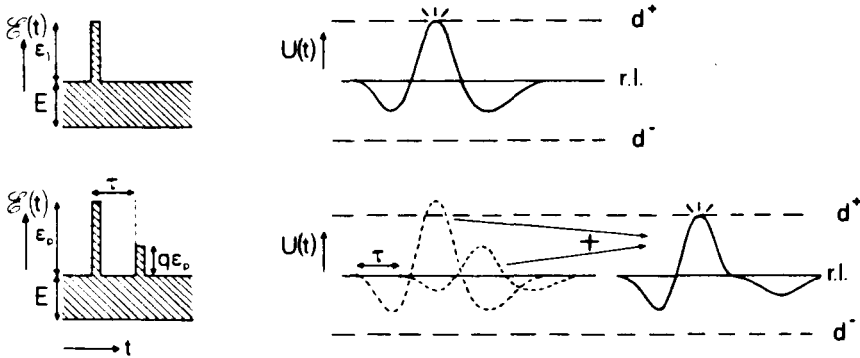
Figure 2: The principles of perturbation. The upper left denotes a short
rectangular flash. Upper right represents the (hypothetical) response
of system L at threshold. The lower left is the combination of probe
flash and smaller test flash, separated by a time interval τ. In the lower
right part, the two individual responses (dashed curves) and the sum
response (continuous curve) at threshold are shown. Notice that in this
case the intensity of the probe flash in the combination must be larger
than in case of one isolated flash, reflecting the influence of the test
flash response.

In order to determine impulse responses from measured thresholds, we chose stim-
ulus time functions as shown in Fig. 2 (for details on methodoloy see 81/1).
As a probe we used a rectangular flash of duration ϑ, short compared to the time
constants of the visual system. The response of the system to such a flash equals:

$$L\{\varepsilon_p f_p(t)\} = \varepsilon_p \vartheta U_\delta(t),\tag{4}$$

where:

   $U_\delta(t)$ is the impulse response of the system,
   $f_p(t)$ is the time function of the flash,
   ϑ is the duration of the flash,
   ε_p is the illuminance increment of the impulse.

Such an impulse response is schematised on the right hand side of Fig. 2a. The flash will be detected if the dominant peak of its response reaches "a":

$$\varepsilon_p \vartheta U_\delta(t_{ex}) = a, \tag{5}$$

where t_{ex} is the time at which $U_\delta(t)$ reaches its extreme value.

In Fig. 2a it is assumed that the impulse response has one clear dominant phase. This property is essential for the probe, used in the perturbation method, to be valid for the determination of responses. The impulse responses measured this far, as a probe response all exhibit this property.

In order to measure the complete impulse response, a combination of flashes was taken as shown in Fig. 2b. A relatively strong flash generates the probe response. The test stimulus, in this case a flash with the same duration, has to be essentially weaker. This second flash, being a factor q lower in amplitude than the probe, is delayed (or preceeds the large flash) a time interval τ . The total response, as schematised on the right hand side of Fig. 2b, then equals:

$$L\{\varepsilon_c f_p(t) + q\varepsilon_c f_p(t - \tau)\} = \varepsilon_c \vartheta U_\delta(t) + q\varepsilon_c \vartheta U_\delta(t - \tau). \tag{6}$$

Now we want to probe the response $U_\delta(t-\tau)$ with the dominant phase of the probe response $U_\delta(t)$. In that case we have to make sure that the threshold criterion is always met at the time t_{ex}. at which the dominant phase reaches the detection level a. Therefore we have to choose q such that:

$$\mid extr\{q\varepsilon_c\vartheta U_\delta(t - \tau)\} \mid \ll \mid extr\{\varepsilon_c\vartheta U_\delta(t)\} \mid, \tag{7}$$

$$\text{or: } q \ll 1.$$

If this condition is met, eq. 6 can be approximated by:

$$\varepsilon_c\vartheta U_\delta(t_{ex}) + q\varepsilon_c\vartheta U_\delta(t_{ex} - \tau) = a. \tag{8}$$

For a single value $U_\delta(t - \tau)$ for the impulse response at time $t_{ex} - \tau$, this leads to:

$$q\frac{U_\delta(t_{ex} - \tau)}{a} + \frac{U_\delta(t_{ex})}{a} = \frac{1}{\vartheta\varepsilon_c(\tau)}. \tag{9}$$

Since q, ϑ and τ are set values and ε_c is measured, the impulse response is known in principle but for a time delay t_{ex}. However, in practice there are non-stationary sensitivity variations which formally may be attributed to variations in "a". Therefore, shortly after the (fast) determination of the threshold ε_c for a certain τ, a sensitivity reference is determined by measuring the threshold of the

probe alone. For the threshold ε_p of the probe flash without the presence of the test flash, it follows from eq. 5 that:

$$\frac{U_\delta(t_{ez})}{a} = \frac{1}{\vartheta \varepsilon_p}. \tag{10}$$

Subtracting eq. 10 from eq. 9 and dividing by q:

$$\frac{U_\delta(t_{ez} - \tau)}{a} = \frac{1}{q\vartheta}\left\{\frac{1}{\varepsilon_c(\tau)} - \frac{1}{\varepsilon_p}\right\}. \tag{11}$$

For the normalised impulse response $U_\delta^\star(t_{ez} - \tau)$, we can derive:

$$U_\delta^\star(t_{ez} - \tau) = \left\{\frac{U_\delta(t_{ez} - \tau)}{a}\right\}\left\{\frac{a}{U_\delta(t_{ez})}\right\} = \frac{1}{q}\left\{\frac{\varepsilon_p}{\varepsilon_c(\tau)} - 1\right\}. \tag{12}$$

By varying the time interval τ between probe and test flash, a discrete number of points for the impulse response of the visual system can be found; a procedure which is followed in practice for the experimental determination of impulse responses.

3.3 Apparatus and procedure

In the impulse response experiments, the stimulus was either a centrally fixated circular field of 1 deg having a dark surround, or a foveal point source of about 1 arc min diameter on an extended adapting homogeneous background with a diameter of 10 deg. The stimulus was seen in Maxwellian view through an artificial pupil with a diameter of 2 mm, provided with an entoptic guiding system that enabled the subject to check whether the pupil of the eye was centered properly before the artificial pupil (Roufs, 1963). The lights were generated by linearised glow modulators, operated around a suitable working point. The luminance of the background was set by means of neutral density filters. The modulation of the background luminance was controlled electronically by function generators. The amplitude of the desired time function could be quickly adjusted using a dB step attenuator. The calibration of the dynamic stimuli was checked before every session by means of a photomultiplier tube, properly corrected with respect to spectral sensitivity.

The subject had one knob to release the stimulus, which was delayed for a convenient time interval (a few hundred milliseconds). In cases where a P800 mini computer guided the experiment, the stimuli were delayed for 300 ms. The beginning of the stimulus was marked by an acoustic signal. Three buttons enabled him to answer with "yes", "no" or "rejection" (in case of not being attentive or eye-blinks).

In all cases, the subject was dark adapted for 30 min, and subsequently adapted for 5 min to the background luminance. For the earlier experimentally determined impulse responses, the 50% detection thresholds of the modulation were determined by means of a modified method of constant stimuli.

For a certain modulation amplitude, 10 or 20 identical stimuli (depending on the experiment) were presented successively and the detected percentage was determined by the experimenter. The dB attenuator was readjusted, and the detected fraction for the new value of the stimulus amplitude was determined. On average, 4 amplitudes taken in random order were needed in order to get sufficient data between 20% and 80% detection chance for approximating the psychometric function on a dB scale by a straight line (Roufs, 1974c).

In order to minimise the effect of non-stationary sensitivity changes, the psychometric function of a second stimulus, to be compared with the first, was determined as fast as possible after or before the combination, and the measurements were repeated in counterbalanced order. Examples of time functions constituting such a "fast pair" are for example shown in Fig. 2. Together, the thresholds of such a pair constitute one value for the normalised impulse response according to eq. 12. For further details on the procedure, see 81/1.

In the majority of impulse response experiments, a somewhat different routine was followed, guided by a P800 mini computer and using the "constant slope" method as defined in 81/1. The strategy was based on the consideration that thresholds constituting a "fast pair" should be measured as quickly as possible after one another. This was done by determining the frequency-of-seeing at a single amplitude for the stimulus time function, provided the perceived fraction was between 20% and 80%. From this percentage, the 50% threshold was obtained by using a constant slope for the psychometric function based on the known Crozier quotient for this subject (the Crozier quotient σ/ε being defined as the standard deviation σ of the distribution function underlying the Ψ-function, divided by the threshold ε).

In practice, before every impulse response experiment, the Crozier quotient was determined by measuring about 40 psychometric functions for this subject of which the average was taken in order not only to increase the sample precision but also to eliminate bias. At the start of the experiment, the Crozier quotient, the sequence of different τ-values, and the starting value for the attenuation were entered into the computer which guided the experiment. From that point, the computer routine consisted of determining 50% thresholds on single perceived fractions for all chosen τ-values. Such 50% thresholds were usually determined from frequency-of-seeing scores based on 10 trials. At least 10 repetitions were generally needed to reduce the standard deviaton to a satisfactory level.

All relevant data measured in our laboratory for different purposes were gathered (see acknowledgement).

In the experiments where different time functions were used in order to check the model predictions, all thresholds were determined by the method of constant stimuli, using about 4 fractions of 10 or 20 trials for one psychometric function.

Eight subjects participated in the impulse response experiments. JAJR, FB, JP, LT, HR, HD, and KS were all male subjects of ages at the time of the experiments of 46, 27, 29, 32, 29, 28 and 35, respectively, whereas JW is a female subject of 25 years old. All subjects had normal acuity, although some of them used a slight

correction.

3.4 Results

Experimentally determined impulse responses

subject	method	stimulus diameter	background level	thresholds per point	q-values	Crozier quotient	β-value
JR(1)	psy-f	1 deg	1200 Td	15	+-0.1;+-0.2	0.16	7.2
FB	psy-f	1 deg	1200 Td	10	-0.15;+0.15	0.14	8.2
JR(2)	cs	1 deg	1200 Td	20	-0.3;+0.3	0.19	6.1
JP	psy-f	1 deg	1200 Td	16	+0.3	0.19	6.2
LT	cs	1 deg	1200 Td	24	+0.3	0.18	6.4
HD	cs	1 deg	1200 Td	24	+0.28	0.15	7.7
LT	cs	1 deg	100 Td	24	+0.3	0.18	6.4
HR	cs	1 deg	100 Td	24	+0.3	0.16	7.2
KS	cs	1 deg	100 Td	44	+0.1	0.16	7.2
JW	cs	1 deg	100 Td	28	+0.25	0.16	7.2
FB	psy-f	0.8'	1200 Td	16	+0.3	0.14	8.2
LT	cs	2'	100 Td	32	+0.27	0.18	6.4
HR	cs	2'	100 Td	48	+0.2	0.15	7.7

Table 1: Listing of stimulus conditions, methods and data on psychometric functions related to the impulse response experiments of which the results are shown in figures 3, 4 and 5. The codes given under 'method' stand for psy-f = psychometric function and cs = constant slope method, respectively. The values for β can be obtained from the Crozier quotient by $\beta = 1.15/(\sigma/\varepsilon)$, where σ/ε is the Crozier quotient (see Roufs, Piceni and Pellegrino van Stuyvenberg, forthcoming).

In Fig. 3, six experimentally determined impulse responses are shown for a 1 deg foveal stimulus with a dark surround, which favours the band-pass transient-like processing. The adaptation level was constant and equal to 1200 Td. Three of the responses have also been shown in 81/1 (for subjects JAJR an FB). The data points are calculated from "fast pairs" according to eq. 12. Averaging was carried out over a number of "fast pairs", the amount of which differed between the experiments, as shown in Table 1. The standard deviations of the means are indicated as the thin vertical lines between the points representing the mean
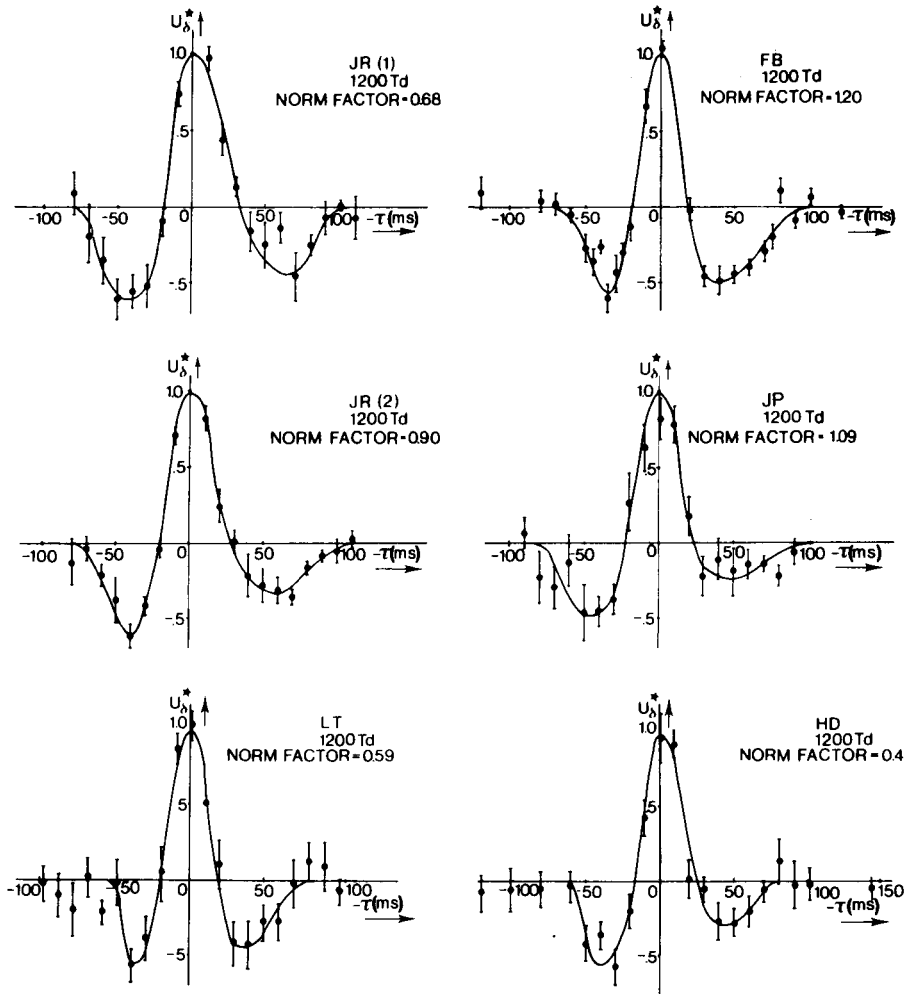
Figure 3: Six normalised impulse responses using a 1 deg stimulus field for five subjects at a background illuminance level of 1200 *Td*. The continuous curves are fitted by eye. The vertical bars between the horizontal shrivels represent two times the Standard Deviation of the Mean (SDM). The absolute values of the impulse responses can be found by multiplying the normalised ones with the norm factors given in the legends.

and the horizontal shrivels. In table 1, also the Crozier quotients and q-factors
(quotient of test and probe flash amplitudes) are listed.

The norm factors, from which the quantitative values for the impulse responses
can be calculated according to eq. 10, are given in the legends.

The continuous lines through the data are curves fitted by eye. These curves
were handled as the basic material for numerical computations as will be shown
furtheron.
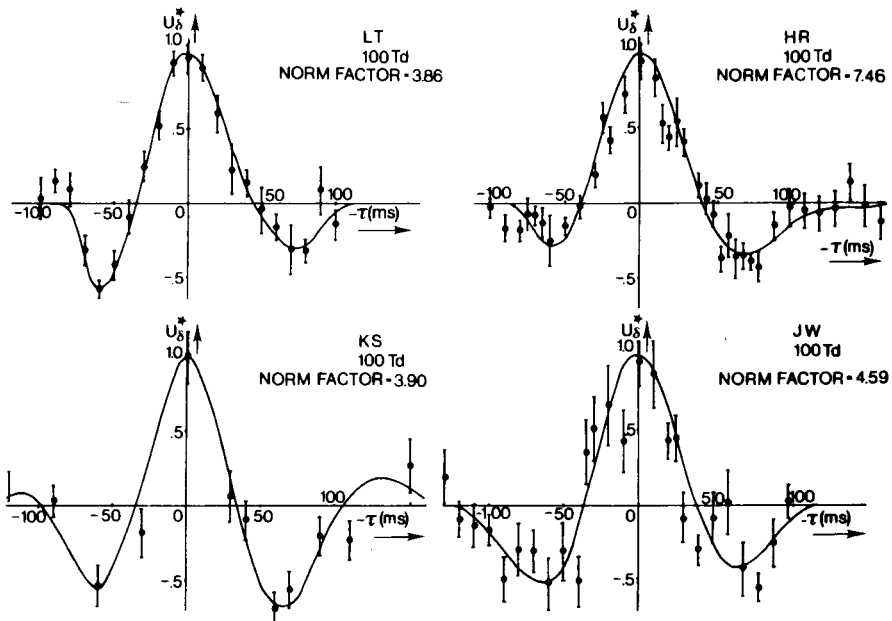
Figure 4: Four normalised impulse responses using a 1 deg stimulus
field for different subjects at a background illuminance level of 100 Td.
The continuous curves are fitted by eye. Twice the SDM's are indicated
as the thin vertical lines between the shrivels.

Fig. 4 gives the experimental results of four impulse responses for four different
subjects at a lower adaptation level of 100 Td. The stimulus again consisted of
a 1 deg disc with a dark surround. The circles are average values obtained from
"fast pairs", and the standard deviation of the mean is indicated. The continuous
lines are fitted by eye. The characteristic parameter settings of these experiments
are listed in Table 1.

In Fig. 5, experimentally determined impulse responses are shown for stimuli
approximating a point source which favoured low-pass sustained-like processing,
superimposed on adaptation levels of 1200 Td (subj. FB) and 100 Td (subjects
LT and HR), respectively. The surround consisted of a field of 10 deg diameter.
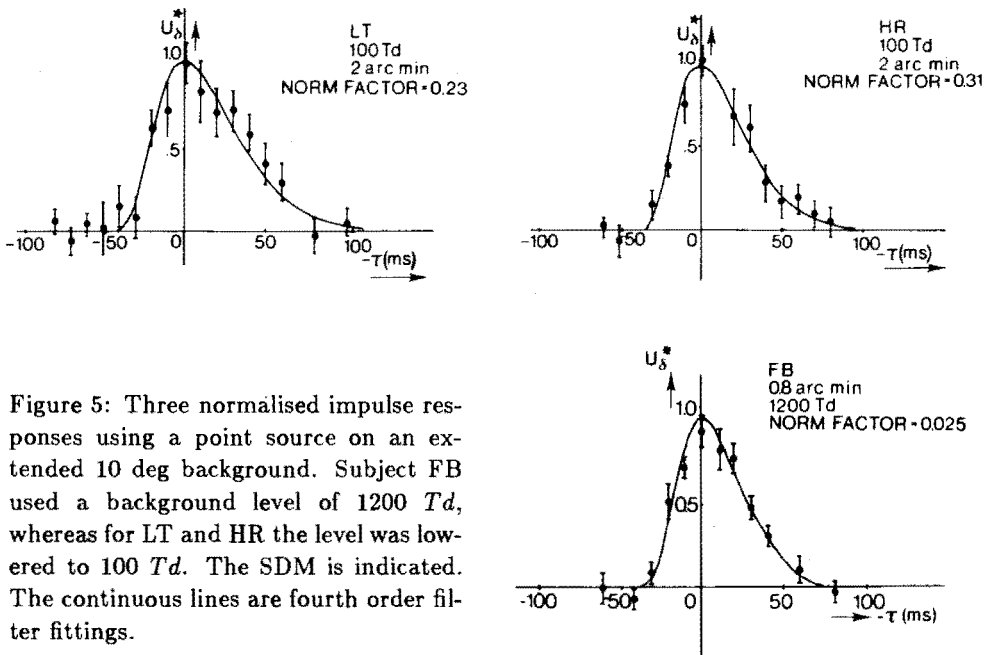
Figure 5: Three normalised impulse responses using a point source on an extended 10 deg background. Subject FB used a background level of 1200 Td, whereas for LT and HR the level was lowered to 100 Td. The SDM is indicated. The continuous lines are fourth order filter fittings.

For FB, the point diameter was equal to 0.8 arc min, whereas for LT and HR it was 2 arc min. The experimental data for the impulse responses are average values calculated from "fast pairs", the numbers of which are quantified in Table 1. The standard deviations of the means are indicated. The continuous curves are computer fittings of the following analytical expression:

$$U_\delta(t) = A(\frac{t}{T})^3 exp(-\frac{t}{T}).$$ (13)

In this expression, the time constant T and the amplitude factor A are the parameters that can be fitted in order to quantise the impulse response.

3.5 Intermediate discussion

Linearity of processing

The validity of experimentally determined impulse responses relies heavily on the linearity assumption. The experimental results can only be interpreted as valid impulse responses if this principle holds. This is also true for the prediction of thresholds of stimuli having different time functions. Only in case of linearity, one may hope for accurate predictions.

Therefore, in the preceding article (81/1), linearity of processing was tested in several ways. Firstly, the ratio q of probe and test stimulus was varied (including negative values for q). According to eq. 9, such a variation should yield a linear relationship between the proportionality factor q and the inverse of the threshold

of the combination of flashes. Deviations from linearity were not found in those
experiments.

A second, more thorough test, is based on a well known property of linear systems
namely that the impulse response should be the derivative of the step reponse.
Having this relationship in mind, step reponses were also measured using the
same perturbation approach as used for the determination of impulse responses.
Both for a 1 deg stimulus and for a point source this relationship was confirmed
(81/1; see figures 12 and 17).

Peak detection and the time structure of noise

If the visual system did not contain temporal noise, peak detection of the signals
at threshold level could be an exact description of the detection process. Since
dynamic visual responses are essentially noisy signals, peak detection will only
be an accurate approximation in case the effects of noise are negligible. If res-
ponse peaks are equal in magnitude, the effect of noise is greatest. This is, for
instance, reflected in threshold decreases (caused by probability summation) if
identical stimuli are repeated, as is the case for sinusoidal modulation in time
(Roufs, 1974c). If response peaks are not equal in magnitude, the contribution of
noise effects strongly depends on the noise properties.

In case of using the perturbation approach, the response signal contains a number
of peaks which are not equal in magnitude (at least if perturbation is properly
applied) but their presence still may obscure the interpretation of the determined
responses, as was signalled by Watson (1982). In fact, he doubted whether our
experimental results may be interpreted as impulse responses because of probabi-
lity summation effects. In our view, his reasoning is not valid, as will be explained
below. A more detailed analysis of this problem will be given elsewhere (Roufs et
al., forthcoming).

We have only poor knowledge about the time structure and the distribution of
noise in the visual system (Roufs, 1973, 1974b). The majority of information on
this subject stems from probability-of-seeing curves. These curves, however, are
subject to variation themselves, since they depend on stimulus parameters, false
alarm rates, detection criteria (cf. Nachmias, 1981), as well as the time it takes
to measure them.

Let us restrict ourselves to a fixed stimulus configuration, for instance a 1 deg
stimulus field modulated in time. Suppose, we want to measure the threshold for
a rectangular flash by determining the psychometric function for this flash. This
can be done by determining the detected fractions at a certain number of stimulus
intensities. An example of such an experiment is shown in Fig. 6a. In practice,
measurement of such a frequency-of-seeing curve will take a certain amount of
time T. Immediately afterwards, we again measure a frequency-of-seeing curve for
the same flash within the same time interval T. This result is plotted in Fig. 6b.
In this (realistic) example, it shows that, although the slopes of the psychometric
curves are about the same, the thresholds themselves are shifted on the intensity
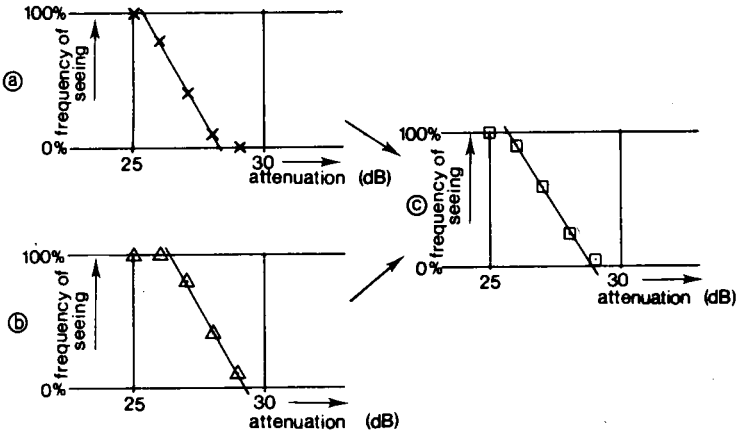
Figure 6: Two examples of psychometric functions ((a) and (b)) determined immediately after one another. The slopes of the curves are about the same, but the 50% threshold is shifted to a higher dB-value. In (c), the psychometric function is plotted if the data of (a) and (b) are averaged. Notice that the slope of the average curve is smaller than those of the individual ones.

axis as a result of relatively low noise-frequency components. If however, the rough data constituting the separate psychometric functions are averaged, a frequency-of-seeing curve emerges as shown in Fig. 6c. It is obvious that the slope of the curve is flattened by time-averaging the individual data points (slopes of -0.33 dB^{-1} and -0.29 dB^{-1} for (a) and (c), respectively).

From this example it is clear, that the frequency-of-seeing curve strongly depends on the time period T in which it is determined. Furthermore, if longer time periods T are taken for determining the psychometric function, the result will always tend to a flatter slope for this function.

If one interprets this issue in terms of stochastic processes, the distribution function underlying the psychometric function depends on the length of the sample of the stochastic process it is part of. Therefore, a psychometric function only contains frequency components of the internal noise that are higher than about $2\pi/T$ Hz, where T is the time interval in which the curve is determined. (In the foregoing article 81/1, it was stressed that special drift-correcting techniques should be employed in order to ascertain that the measured deterministic effects are large compared to the experimental spread, which can be considered as a by-product of low-frequency noise components which we called "relatively slow sensitivity shifts".)

This observation has an important implication for estimating the effect of probability summation, since this effect takes place over time periods (usually less than about 300 ms) that are small compared to T. Therefore, the low frequency components of noise only play a minor part in probability summation effects. On the

other hand, whatever information we have for calculating its effects is inferred from the psychometric functions. From the foregoing it will be clear that, in general, there will be an overestimation of the effect of probability summation if it is extrapolated from the properties of frequency-of-seeing curves, which overestimation is larger the longer it takes to measure these curves. The closest estimation of the effect of probability summation then will stem from psychometric curves measured in very short periods of time.

In the following, it will be shown what the effect of noise can be, as a function of the slope of the psychometric function, on the experimentally determined impulse responses.

It is assumed, as Watson (1982) did, that the detection chance of a stimulus can be formalised by using Quick's (1974) rule:

$$\Psi = 1 - \prod_{i=1}^{N}(1 - (1 - 2^{-|R_i|^\beta})) \tag{14}$$

where:

Ψ is the detection probability of the stimulus,

R_i is the response magnitude of sample i,

N is the number of samples constituting the response,

β is a parameter proportional to the width of the underlying probability density function.

On the basis of eq. 14, it can be derived (see appendix) that a temporal stimulus with time function $\varepsilon f(t)$ will be at threshold ε_{thr} if:

$$1 = \int_0^\infty \left| \varepsilon_{thr} f(t) \otimes \frac{U_\delta(t)}{a} \right|^\beta dt. \tag{15}$$

Using this formula, it can be calculated what the effect of noise will be on the results of the perturbation experiments as described in this paper. Doing this we follow Watson (1982) in his comment on our 81/1 article, who postulated a hypothetical impulse response function for the visual system which differs in shape from the measured ones reported on. The response was chosen in such a way that the modulus of its Fourier transform equals de Lange flicker sensitivity curves (which in our view envelopes two different processes). The impulse response mentioned above obeys the analytical expression:

$$U_\delta(t) = H(t)\left\{ \left(\frac{t}{4.94}\right)^8 e^{-\frac{t}{4.94}} - \frac{1}{12}\left(\frac{t}{6.58}\right)^9 e^{-\frac{t}{6.58}} \right\}, \tag{16}$$

$H(t)$ is the Heaviside (step) function.

Its shape is depicted in Fig. 7 (labelled original), and it shows a diphasic impulse response of which the first phase is largest. With the aid of eq. 15 it is calculated what the effect would be on a perturbation experiment, if this impulse response were valid, for different values of β. For large β-values, the effect of noise is negligible, and the "experimental result" equals the impulse response. For smaller
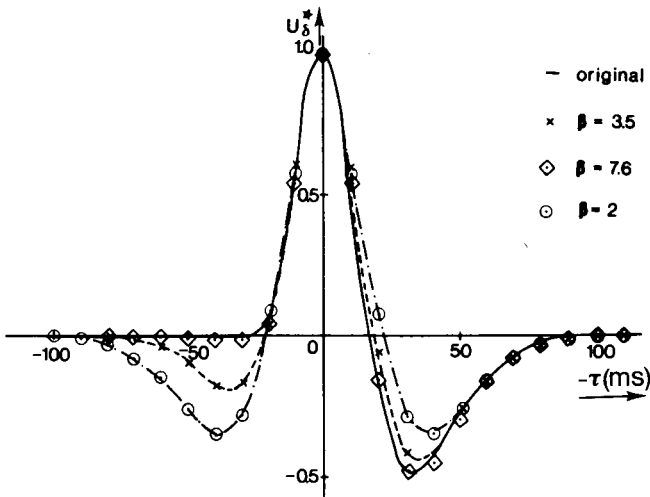
Figure 7: Illustration of the effect of noise on the perturbation approach for determining impulse responses. The impulse response advocated by Watson (1982) (continuous curve) is labelled "original". The noise parameter is β which is inversely proportional to the width of the distribution function underlying the psychometric function. It can be seen that the simulated result of the impulse response experiment approaches the theoretical impulse response for larger β-values.

values of β, the "experimental result" will increasingly deviate from the impulse response owing to an increasing effect of noise. On the basis of these observations, Watson questioned whether our experimental results could be interpreted as impulse responses of the system. He argued that β-values measured in practice usually range between 3 and 7, and that the values reported on in our paper (8.2 and 7.2 for different subjects) were atypically high.

We argue, however, that our β-values found from rapidly measured Ψ-functions more realistically reflect the noise parameter which is relevant for a short period during the perturbation in which the noise effects detection. Fast measurement of pairs of psychometric functions is the basis of this method. As is argued before, the fast drift-correcting technique is not only important to determine small threshold differences, but also to estimate the parameters of the stochastic process (underlying probability summation) adequately. From Table 1, where 13 β-values are listed, it can be seen that for our experimental conditions all β-values are above 6 with an average of 7.1.

The statistical analysis of the standard deviations of the measured response amplitudes for subject FB, as given in 81/1 (chapter 2; Table 1), showed that drift effects are indeed removed effectively and that its value is consistent with $\beta=7.6$. From Fig. 7, where the perturbation technique is simulated, it can be seen that for this β-value, the experimentally determined impulse response is in fair agreement

with the one corresponding to a peak detection process (labelled original). Starting from a three-phasic impulse response, the effect of noise is even smaller (numerical values will be given elsewhere; Roufs et al., forthcoming). In fact, in view of the foregoing, the β-value of 7.6 might even be a low estimate of the noise-exponent which governs "probability summation".

On the basis of these observations and other data to be treated further on, we have no reason to doubt the validity of the impulse responses obtained in the present conditions.

Impulse response variation

a Effects of background luminance

On the basis of earlier experimental results (Roufs, 1974a; 81/1) it might be expected that impulse responses do not change in shape if the background luminance is lowered, except for a scale factor of the time axis and a change in sensitivity. This expectation can now be evaluated by comparing the experimentally determined impulse responses as shown in figures 3 and 4 for the adaptation levels of 1200 Td and 100 Td, respectively. It can be seen that the three-phasic nature of the responses does not alter when the adaptation level decreases. This change merely results in a change in time scale of about a factor of 1.5 and a change in sensitivity as is reflected in the norm factors. Consequences of these effects will be evaluated in the second part of the article, where thresholds are evaluated for rectangular flashes over a large range of durations.

b Effects of stimulus dimension

As was shown in 81/1, impulse responses measured for a point source change in character compared to the ones measured for a 1 deg stimulus field. The results, as shown in Fig. 5, are monophasic responses, which can fairly accurately be described by a fourth-order process (81/1). Lowering the adaptation level in this case does not result in definitely slower impulse responses, as shown by the results of subjects HR and LT, which were determined at an extended background level of 100 Td.

As does the nature of the impulse response, so does the percept change when flashes are viewed at threshold level. For point sources, a pin-point-like brightness change is observed, usually attributed to a sustained type of processing. For a 1 deg stimulus field, the percept is typically a fast movement within the field ("agitation"), already in earlier papers coupled with a transient type of processing (81/1).

These findings are compatible with results of Kulikowski and Tolhurst (1973) and Watson and Nachmias (1977), who report on temporal response functions using sinusoidal gratings as stimuli. They showed that for high spatial frequencies these functions are also monophasic, accompanied by a pattern-like detection criterion,

whereas for low spatial frequencies the response functions exhibit a more oscillatory character coupled with a flicker percept.

c Intersubject variation

The results of figures 3, 4 and 5 show that the shape of the experimentally determined impulse responses does not vary much between subjects. Slight differences can be observed in sensitivity as is reflected in the norm factors. Furthermore, a small variation in the time constants can be observed. The differences, however, do not exceed the usual intersubject variation for psychophysical experiments.

3.6 Prediction of thresholds and latency

Methods

a Thresholds

The stimulus configuration, the apparatus and the adaptation procedures are identical to those described in "Apparatus and Procedure", unless stated otherwise. In all cases, except for the latency measurements, either a 1 deg foveal stimulus with a dark surround or a point source (0.8 arc min) on an extended (10 deg) background was used.

Thresholds were obtained using the method of constant stimuli and defined by the 50% detection probability. The thresholds belonging to one characteristic were obtained by varying the independent variable (such as duration or time interval) at random over its full range within one session, and repetitions in other sessions were made in a counterbalanced order over an even number of sessions.

In part of the experiments reported here, the authors (FB and JAJR) acted as subjects. Other subjects were HM, JP and RK, who all had normal vision and were 25, 28, and 18 years old respectively at the time of the experiments.

A number of experimental results used in this section were obtained in earlier years and have already been published (Roufs, 1974a, 1974b).

b Latency

The experimental configuration consisted of two adjacent, just overlapping circular fields of 1 deg, as shown in the inset of Fig. 15. The left hand field of the flashes was the reference with a fixed intensity, while the intensity of the right hand 400 ms flash was varied as a parameter. The subject was asked to set the test flash subjectively simultaneous with the reference flash. For further details see Roufs (1974b).

3.7 Rectangular increments over a large range of durations
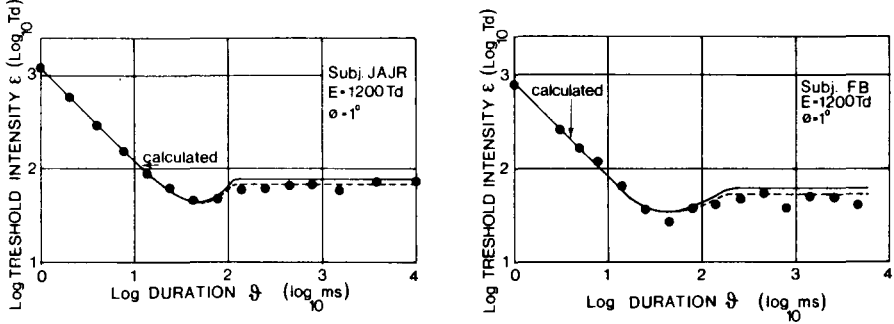
Figure 8: Thresholds of rectangular flashes as a function of the duration ϑ for two subjects. The data points are averages of two thresholds. The continuous curves are deterministic predictions on the basis of the impulse responses shown in Fig. 3. The dashed lines are corrections for probability summation at long durations.

In Fig. 8 threshold increments ε , measured for rectangular flashes, are plotted as a function of flash duration ϑ. The results of two subjects (FB and JAJR) are shown as the dots. These are averages of two 50% thresholds, which were measured in random order with respect to ϑ during the first session, the order of measurement of the second session being reversed. The close matching at short durations is probably a coincidence, since daily sensitivity variations can easily displace the theoretical and experimental curves with respect to each other by a perceptible amount.

The drawn curves of Fig. 8 are the predicted values calculated from the unit impulse response determined before. The calculation of the thresholds is based on eq. 3 as given in "Theoretical frame". According to the peak detection model, the stimulus is seen if the extreme value of the response surpasses "a". The threshold value of the amplitude factor of the retinal illumination ε_f, is therefore implicitly given by:

$$\varepsilon_f extr \left\{ \frac{U_f(t)}{a} \right\} = 1. \tag{17}$$

In the special case of a rectangular flash $\varepsilon p(t)$, having a duration ϑ and an increment of retinal illumination ε, eq. 3 transforms to:

$$\frac{\varepsilon U_p(t)}{a} = \varepsilon \int_0^\infty p_\vartheta(t - \tau) \frac{U_\delta(\tau)}{a} d\tau, \tag{18}$$

with

$$p_\vartheta(t) = 1 \text{ for } 0 < t \le \vartheta$$
$$= 0 \text{ for } t \le 0,\ t > \vartheta.$$

This convolution was calculated numerically for the unit impulse responses shown in Fig. 3 for subjects FB and JAJR. After applying eq. 17 to the responses obtained by eq. 18, the threshold predictions of the one degree stimulus with a background of 1200 Td and a dark surround were obtained. The dashed line is a correction for probability summation. It is obvious that prediction and experiment are very close. It has to be emphasized that the curves are not fitted in any way, since there are no free parameters left.
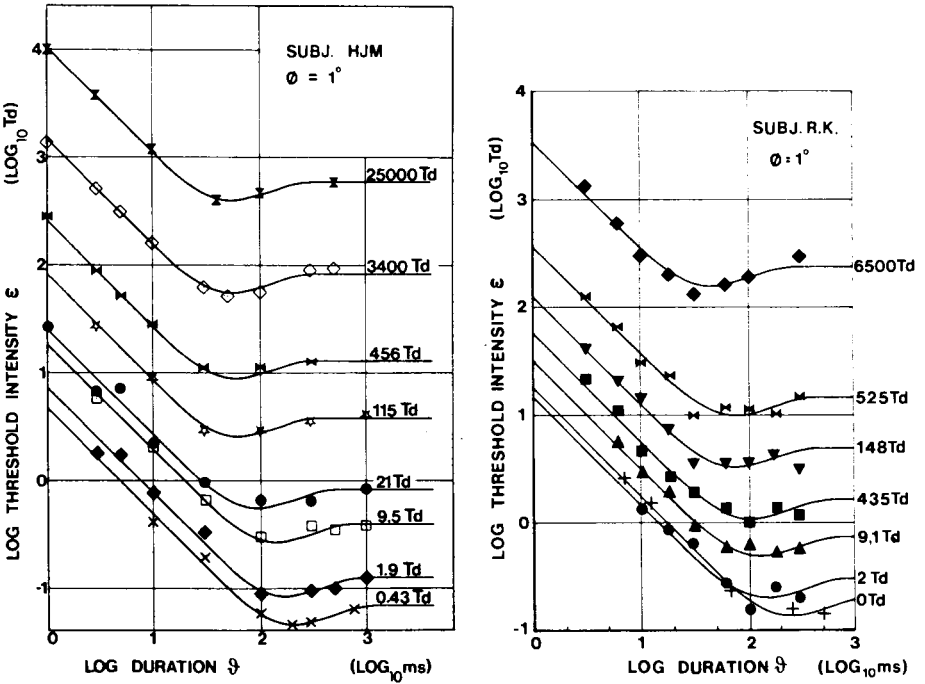
Figure 9: Thresholds for rectangular flashes as a function of the duration, at different background levels as indicated. The continuous curves all are templates of each other and are obtained from the reduced thresholds as shown in Fig. 10.

The typical shape of this type of threshold curve does not change with adaptation level as is demonstrated in Fig. 9. The curves can be brought together by a simple translation along the logarithmic axis, as was shown before by one of the authors (Roufs, 1974a). If the thresholds are divided by the asymptotic value for long durations ε_a, and the duration by a critical duration defined by the intersection point of the two asymptotes, one obtains a curve like the one in Fig. 10. Here, experimental data for different subjects measured at different adaptation levels,
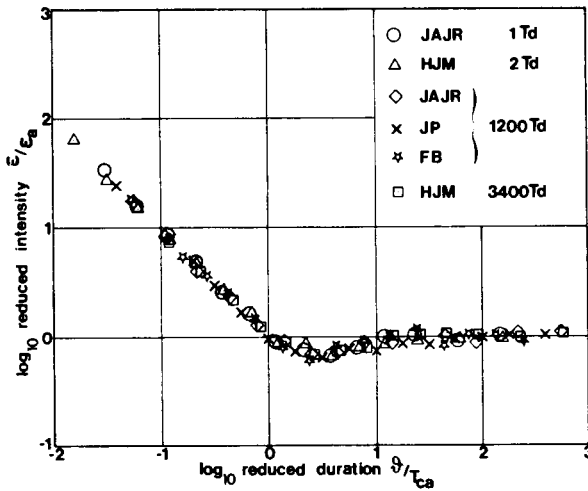
Figure 10: Reduced thresholds as a function of the duration of rectangular flashes, reduced with respect to the critical duration. As indicated in the legend, different observers took part while different background levels were used.

are reduced in that way. There seems to be not much difference between subjects either. In fact, the curve through the points of Fig. 9 is the same curve of Fig. 10 translated through the points given by one of the authors in an earlier paper (Roufs, 1972; Fig. 7).

At this point it is perhaps clarifying to relate the shape of the threshold curve to the response behind it. For durations short compared to the duration of the impulse response $U_\delta(t)$, eq. 18 becomes in good approximation:

$$\frac{\varepsilon U_p(t)}{a} = \varepsilon \vartheta \frac{U_\delta(t)}{a}. \tag{19}$$

Substituting this result into eq. 17, one obtains for the threshold ε_p:

$$\varepsilon_p \vartheta \, extr \left\{ \frac{U_\delta(t)}{a} \right\} = 1. \tag{20}$$

Since the maximum of the unit impulse reponse is constant, being equal to the norm factor attached to the reduced impulse response shown in Fig. 3, the equation says that the product of intensity increment and duration is constant, which is known as Bloch's law. This is reflected in the rectilinear asymptote at short durations, having a slope of minus one. Eq. 20 provides an easy way to check the prediction of thresholds at short durations, without the necessity of numerical integration. These thresholds, for instance those of Fig. 8, can be calculated for a certain duration by taking the reciprocal of the product of the norm factor (see Fig. 3) and the duration ϑ.
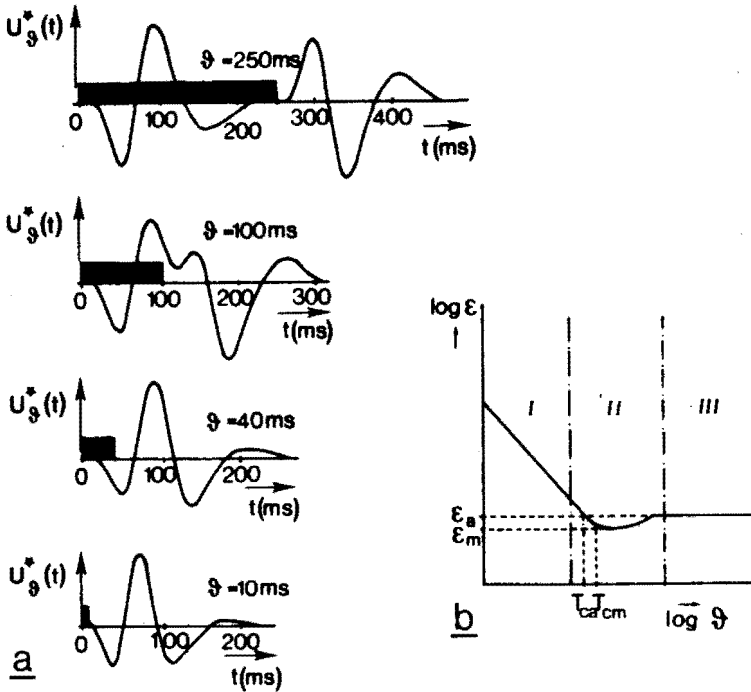
Figure 11: Calculated visual responses for different flash durations ϑ. For short durations, the responses have about the same shape as the measured impulse responses ($\vartheta = 10, 40\ ms$). For an intermediate duration of $\vartheta = 100\ ms$, two mingled step responses of the on- and off-going step are already recognizable. At a duration of $\vartheta = 250\ ms$, the step responses of the on- and off-going flanks are clearly separated. The duration of the flashes is indicated by the black bars. In Fig. 11b a schematised threshold-versus-duration curve is drawn, showing two definitions for the critical duration.

The effective integration time is closely related to the width of the dominant phase of the impulse response, a measure being found in the critical duration according to one of the definitions being in force. As pointed out before (Roufs, 1974a), it is unfit from a methodological point of view, to take the duration at which the curve parts from the asymptote for short durations. One practical definition would be to stick consequently to the intersection point of the two asymptotes as demonstrated in Fig. 11b. However, this time constant T_{ca} has the disadvantage that it is rather time consuming to measure since very long flash durations are involved. An alternative is to use the intersection point of the short duration asymptote and a line parallel to the abscissa and through the minimum threshold ϵ_m (Roufs, 1974a), see Fig. 11b.

The dip in the threshold curve at intermediate durations is a natural consequence of the additivity property. A rectangular blocklike flash is composed of a positive-

and a negative-going step function. At a certain duration, the second (dominant) phase of the response of the ongoing step adds to the first phase of the response of the negative-going step. This is illustrated in Fig. 11a. If the time difference between the onset and the offset of the stimulus is sufficiently large, only the dominant phases of the two responses (being equal in magnitude for this stimulus) determine the response.
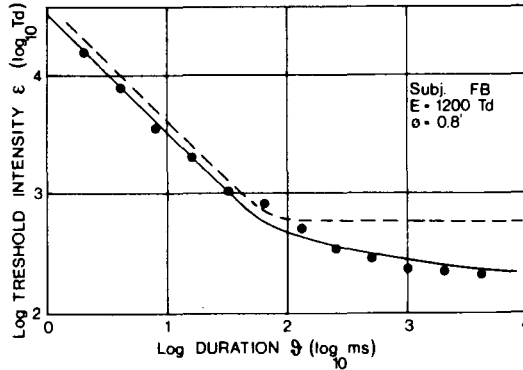
Figure 12: Measured thresholds for rectangular flashes as a function of duration for a point source superimposed on an extended background of 1200 Td. Note that the thresholds decrease monotonically as a function of the duration as opposed to the threshold course for a 1 deg stimulus field (cf. Fig. 8). The dashed curve is a deterministic prediction calculated from the impulse response for subject FB shown in Fig. 5. At longer flash durations a stochastic correction for probability summation is shown as the thick line. For clarity it was vertically fitted to the measured thresholds at short durations.

If the step response has no negative phase, no dip in the threshold curves can be expected. This is demonstrated for the next set of data, obtained with a point source at an extended background of 1200 Td. As shown in Fig. 5 and the preceding article 81/1, the impulse and step response have no oscillating character, and consequently the sum of the responses of positive- and delayed negative-going step can never surpass the maximum amplitude of a single step response. Experimental and predicted values are shown in Fig. 12.

No dip is predicted nor is it found. However, for durations larger than about the critical duration, threshold is lower than predicted. This is a result of the stochastic nature of the detection process. For large durations, the top of a block response becomes flat. Thus increasing the duration does increase the detection probability of the flash. The continuous line in Fig. 12 is a correction for this stochastic effect calculated on the basis of Quick's rule (eq. 14). The agreement between prediction and experimental values is also excellent for this configuration.

3.8 Comparison of prediction, derived from the measured impulse responses, with earlier experimental results

We can extend the number of prediction tests by making use of older data measured by the same subject (JAJR) under identical conditions and stimulus configuration. It gives an extra dimension to the test because it should tell us also something about repeatability.

Experiments with double flashes
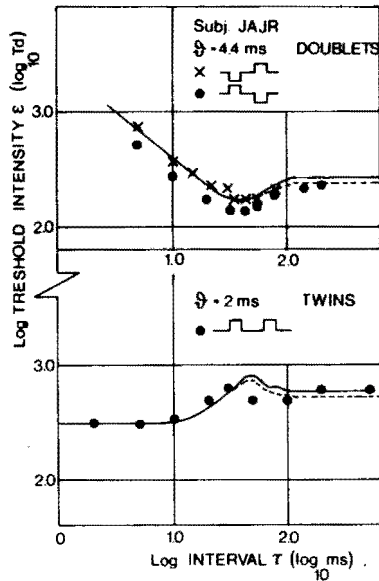
Figure 13: Thresholds for "doublets" and "twins" as a function of the time interval τ between two constituting flashes of which the time courses are schematically shown in the insets. The continuous curves are predictions based on the experimentally determined impulse responses for subject JAJR (averaged over the two curves given for this subject in Fig. 3). The dashed lines are corrections for probability summation at longer time intervals τ.

We are referring here mainly to experiments with twin flashes (Roufs, 1973), which consist of two identical increments with a variable time interval τ, and doublets (Roufs, 1974a), comprising two short flashes of opposite sign and separated by a variable time interval τ. The experimental results are plotted in Fig. 13 together with results calculated from the impulse response. It is obvious from the graphs that experiment and prediction are in good agreement. This is to be expected as far as the general shapes of the curves are concerned. However, the good repeatability

with respect to the vertical position is more surprising, having in mind that the
time gap between the concerning experiments is about 6 years. The norm factor for
instance, must be a good estimate for the sensitivity of this subject for transients,
despite the time difference between the measurements.
The small second bump in the curve for twin flashes is situated at those time
intervals τ where the detected peak of the response changes from sign. It is
probably not significant but caused by experimental error in determining the unit
impulse response.

Experiments on latency

For subject JAJR, data have been published on latency (Roufs, 1974b) which can
also be confronted with computations from the impulse response.
The variation of latency with stimulus intensity was measured as the difference
of perception lag of two step-functions in time. The double-flash method used
is based on subjective simultaneousness of the onset of the flashes. The onset
of the test flash with variable intensity was shifted with respect to the onset
of the reference flash having constant intensity, until the onsets appeared to be
simultaneous and no apparent movement was seen. For experimental details see
Roufs ibid.
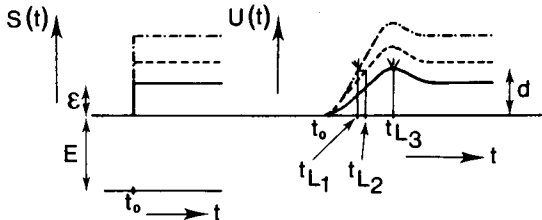
Figure 14: Visualisation of the model used for the prediction of latency.
The left-hand figure gives the variation of retinal illuminance as an
input signal on stepped stimulation. The right-hand column gives the
schematic signals at an appropriate point in the visual pathway. It
illustrates that the further the stimulus is above threshold, the more the
moment of threshold trespassing shifts in the direction of the starting
moment of the stimulus, so that according to the model the perception
lag is reduced.

The model used was identical to the one shown in Fig. 1, except that one postulate
was added. As shown in Fig. 15, we assume, as Roufs (1974b) did, that the time
the response needs to surpass the detection level "a" is the relevant magnitude.
This time-rising model needs the extra postulate, in comparison with the detection
model, that either at the intersection moment the delayed non-linear processing
is not yet observable (analoguous to De Voe's (1967a,b) observation in the wolf-
spider), or that the result of this type of response depends on the delay of the

action of movement detectors (see Foster, 1971), operating at a stage where the responses are still linear, the following non-linearity being irrelevant.

From the impulse response at the 1200 Td level, the step response was calculated and the rise time t_L for this step response, having different step intensities ε, was determined according to the implicit expression:

$$\varepsilon \frac{U_s(t_L)}{a} = 1. \tag{21}$$

The impulse response at the 1 Td level was not measured directly, but based on observed isomorphy of the reponses at different adaptation levels (cf. figures 3 and 4). From a set of curves like those in Fig. 9, but now for subject JAJR, the change of scale unit going from 1200 Td to 1 Td was estimated to be 4.0. From this value, and the known impulse response at 1200 Td, the impulse response at the 1 Td level can be calculated easily. As pointed out by Roufs (1974b), it is advantageous to express the intensity of the flash in its threshold value $\varepsilon_{.4}$ at the level for long durations. In this case, 400 ms was taken. Eq. 21 then becomes:

$$\frac{\varepsilon}{\varepsilon_{.4}} U_s^*(t_L) = 1, \tag{22}$$

where $U_s^*(t)$ is the step response normalised at its maximum.

The maximum latency is determined by the time of the reponse peak at threshold. In this case it was convenient to use the actual latency minus $\varepsilon_{.4}$ the actual value as the dependent variable.

Experimental results for two background levels are shown in Fig. 15, together with the calculated values (dashed curves). Since the time differences measured are relative with respect to the reference stimuli (indicated by stars), the scales belonging to the two background levels are translated along the latency axis in order to obtain the best fit.

3.9 General discussion

Predictive value of impulse responses

From figures 8, 12 and 13, it is clear that the quality of the prediction of thresholds on the basis of experimentally determined impulse responses is satisfactory in all cases investigated. It should be emphasized that, since no free parameters are involved, predicted thresholds are in quantitative agreement with measured ones. This means that both the shapes of the predicted threshold curves and the vertical position, i.e. the sensitivity, match the experimental results. Considering the experimentally determined impulse responses, the shapes of the threshold curves depend on the shape of the (normalised) impulse response, whereas the vertical positions depend on the norm factors. Both characteristics, therefore, seem to be accurate estimates for the purpose of predicting thresholds.

From Figs. 8 and 9, where thresholds for rectangular flashes at a 1 deg stimulus field are plotted, it is striking to observe a small but persistent dip that occurs in the
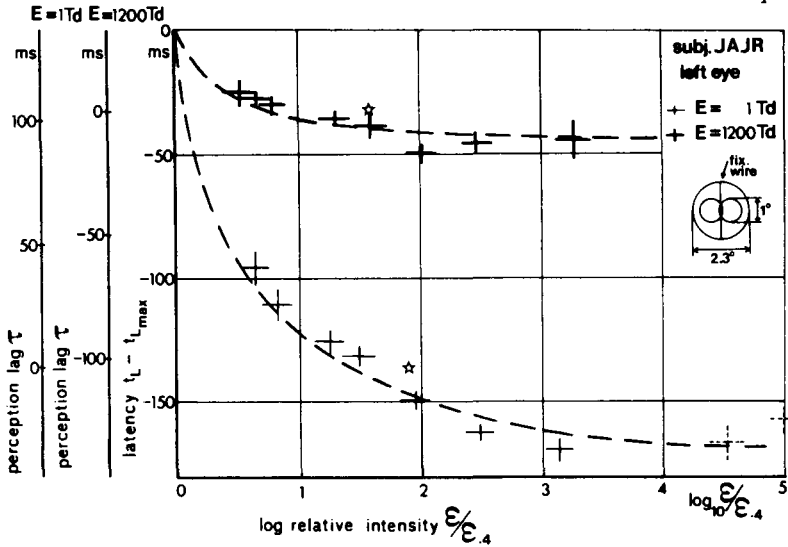
Figure 15: Perception lag of an incremental rectangular flash with a variable intensity ε of the increment compared to a constant reference flash. Both flashes lasted for 400 *ms*. The intensity of the increment is expressed appertaining to the low and high background levels at which the measurements were carried out. The time scales used with both levels are marked as such. The inset shows the stimulus configuration. The latency scale belongs to the theoretical dashed lines which are calculated from the averaged impulse response for subject JAJR (see Fig. 3).

curve at intermediate flash durations. The existence of such a dip seems to provide a suitable means for the purpose of evaluating the shape of the impulse responses as was explained in the concerning section. The step response should have a negative phase in order to predict this dip in the threshold curves. This datum puts a heavy constraint on the impulse response, since the impulse response is the primitive function of the step response. It can be derived that only three-phasic impulse responses are able to describe this particularity of the threshold curve for rectangular flashes. Even if probability summation plays a more dominant role than assumed in this paper (Watson, 1982), an impulse response of a diphasic nature (see for instance Fig. 7), is unable to explain such threshold variation; at least not if the model used possesses a realistic description of probability summation. Gorea and Tyler (1986) proposed a model based on a diphasic impulse response and assuming an extra non-linear element and an integrator. This model is able to produce a dip in the threshold-versus-duration curve, be it that for realistic values of their model parameters the dip is not deep enough in comparison with experimental values.

For longer flash durations, a stochastic correction on the deterministic prediction is proposed, based on the observation that, at these durations, the on- and off-

going step responses have dominant peaks of equal magnitude. In that case, the probability of detecting the flash is increased, leading to a lower threshold. As was stated before, such effects of probability summation may be formalised by using Quick's (1974) rule as an approximation to describe probability summation (for a discussion see Nachmias, 1981), provided β-values are taken that are large enough.

For a point source on the extended background level of 1200 Td, the effect of probability summation is much more pronounced (see Fig. 12). This is caused by the fact that step responses as measured for subject FB (81/1), and calculated from the impulse responses of Fig. 5, possess a flat plateau, the duration of which being about equal to the flash duration. This plateau increases the effect of probability summation and thus substantially lowers the threshold for increasing flash durations. As can be seen from Fig. 12, the prediction of the threshold curve on the basis of the monophasic impulse response and probability summation provides an accurate description of the data.

From Fig. 13, where thresholds for doublets and twin flashes are compared with prediction, it can be seen that the prediction of thresholds is also satisfactory if no integration of impulse responses is involved. For these stimuli, the total response according to the model simply consists of the algebraic difference or sum of two impulse responses separated by a time interval τ. The effect of probability summation on the deterministic threshold is again indicated. As was observed before, it is remarkable that the quantitative positions of the curves agree that well with the measured thresholds, having in mind that about six years elapsed between the doublet and twin experiments and the impulse response determination for subject JAJR. Seemingly, the norm factor, obtained from averaging over a vast amount of thresholds measured on different days, acts as an accurate measure for the average sensitivity of this subject even after a number of years.

In Fig. 15, experimental results on latency are plotted for subject JAJR, together with predictions based on the experimentally determined impulse responses. It is clear that the prediction is accurate for both the high and the low adaptation level. From this result it can be concluded that a time-rising model, based on experimentally determined impulse responses, is well able to describe this type of latency measurements.

Related temporal models

In a number of publications (Rashbass, 1970, 1976; Broekhuysen et al, 1976; Watson and Nachmias, 1977), a model was made explicit, consisting of a linear filter followed by a quadratic operator and an integrator. Using such a model, autocorrelation functions were determined that characterise the action of the linear operator. Thresholds calculated from this model, coincide for a fair amount with our predictions, be it that these calculations were not carried out quantitatively but solely with respect to the shape of threshold curves. We are mainly referring here to the ellipses as given for instance by Rashbass (1970) and Watson and Nachmias

(1977). As was shown in 81/1, our model is able to describe these ellipses as well if probability summation is taken into account. A disadvantage of the Rashbass model is that it uses an integrator with an integration time which is unknown and hard to quantify. Furthermore, as is shown in 81/1, some basic properties of autocorrelation functions do not agree with our experimentally determined impulse and step responses.

Krauskopf (1981) showed that his experimental data on the effects of adaptation to sawtooth light variation on incremental and decremental thresholds can not be explained on the basis of Rashbass's quadratic model, whereas they are consistent with the linear model used here.

As was pointed out by Broekhuysen et al. (1976) and Watson and Nachmias (1977), the two models can be unified by replacing the quadratic operator by an operator of the power β, in which β represents the effect of noise. As we pointed out before, β has to range between 6 and 8 according to the slope of psychometric functions obtained in short time intervals. In this way the non-linear operator together with the integrator will tend to a peak-detection mechanism having a slight correction for probability summation. This proposal of Watson (1979, 1982), therefore, is an elegant formalism for the prediction of thresholds.

de Lange characteristics

In an earlier series of papers, Roufs (1972a, b; 1973; 1974a, b, c) linked the description of thresholds for one-shot stimuli with thresholds for sinusoidally modulated light (de Lange characteristics). He concluded that at high temporal frequencies, threshold behaviour can be described very well with a band-pass filter of which the gain equals the the absolute value of the Fourier transform of the impulse response (Roufs and Pellegrino Van Stuyvenberg, 1976). This filter is linked with the percept "agitation" or "flicker". At low temporal frequencies a low-pas filter would be dominant, accompanied by a percept called "swell" (homogeneous brightness variation). Certainly, these characteristis are linked with sustained and transient processing, respectively, as often mentioned in the literature (Breitmeyer and Ganz, 1976; Kelly and Burbeck, 1984; Green, 1984). As is pointed out by these authors, sustained activity seems to dominate at high spatial frequencies, while at low spatial frequencies the transient system takes over. This property is reflected in the shape of the impulse responses determined for a 1 deg field (relatively low spatial frquency) compared to the impulse responses for a point source (high spatial frequency). The oscillatory impulse responses for the 1 deg field then should be the result of transient activity whereas the monophasic impulse responses for point sources should reflect sustained activity. In that sense, our data support Green (1984) when saying: "sustained and transient mechanisms use the synchrony of excitation and inhibition to trade off between good spatial and good temporal resolution. As a result, the human visual system is capable of operating over a greater range of conditions". This notion is also supported by variation in experimentally determined point spread functions if the stimulus presentation is

flashed, respectively, quasi-static (Blommaert, 1979).

Acknowledgement

References

Blommaert, F.J.J. (1979) Local visual responses in space and time. IPO Ann. Progr. Rpt. 14, 88-94.

Breitmeyer, B.G.; Ganz, L. (1976) Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression and information processing. Psychol. Rev. 83, 1-36.

Broekhuysen, M.; Rashbass, C.; Veringa, F. (1976) The thresholds of visual transients. Vis. Res. 16, 1285-1289.

De Voe, R.D. (1967a) Nonlinear transient responses from light-adapted wolfspider eyes to changes in background illumination. J. Gen. Physiol. 50, 1961-1991.

De Voe, R.D. (1967b) A nonlinear model for transient responses from light-adapted wolfspider eyes. J. Gen. Physiol. 50, 1993-2030.

Foster, D.H. (1971) A model for the human visual system and its response to certain classes of moving stimuli. Kybernetik 8, 69-84.

Gorea, A.; Tyler, C.W. (1986) New look at Bloch's law for contrast. J. Opt. Soc. Am. A 3, 52-61.

Green, M. (1984) Masking by light and the sustained-transient dichotomy. Perc. & Psychoph. 35, 519-535.

Kelly, D.H.; Burbeck, C.A. (1984) Critical problems in spatial vision. Crit. Rev. in Biomed. Eng. 10(2), 125-177.

Krauskopf, J (1981) Discrimination and detection of changes in luminance. Vis. Res. 20, 671-677.

Kulikowski, J.J.; Tolhurst, D.J. (1973) Psychophysical evidence for sustained and transient detectors in human vision. J. Physiol. 232, 149-162.

Nachmias, J. (1981) On the psychometric function for contrast detection. Vis. Res. 21, 215-223.

Quick, R.F. (1974) A vector magnitude model of contrast detection. Kybernetik 16, 65-67.

Rashbass, C. (1970) The visibility of transient changes of luminance. J. Physiol. Lond. 210, 165-186.

Rashbass, C. (1976) Unification of two contrasting models of the visual incremental threshold. Vis. Res. 16, 1281-1283.

Roufs, J.A.J. (1963) Perception lag as a function of stimulus luminance. Vis. Res. 3, 81-91.

Roufs, J.A.J. (1972a) Dynamic properties of vision-I. Experimental relationships between flicker and flash thresholds. Vis. Res. 12, 261-278.

Roufs, J.A.J. (1972b) Dynamic properties of vision-II. Theoretical relationships between flicker and flash thresholds. Vis. Res. 12, 279-292.

Roufs, J.A.J. (1973) Dynamic properties of vision-III. Twin flashes, single flashes and flicker fusion. Vis. Res. 13, 309-323.

Roufs, J.A.J. (1974a) Dynamic properties of vision-IV. Thresholds of decremental flashes, incremental flashes and doublets in relation to flicker fusion. Vis. Res. 14, 831-851.

Roufs, J.A.J. (1974b) Dynamic properties of vision-V. Perception lag and reaction time in relation to flicker and flash thresholds. Vis. Res. 14, 853-869.

Roufs, J.A.J. (1974c) Dynamic properties of vision-VI. Stochastic threshold fluctuations and their effect on flash-to-flicker sensitivity ratio. Vis. Res. 14, 871-888.

Roufs, J.A.J.; Pellegrino Van Stuyvenberg, J.A. (1976) Gain curve of the eye to subliminal sinusoidal modulation of light. IPO Ann. Progr. Rpt 11, 56-63.

Roufs, J.A.J.; Blommaert, F.J.J. (1981) Temporal impulse and step responses of the human eye obtained psychophysically by means of a drift-correcting perturbation technique. Vis Res. 21, 1203-1221.

Roufs, J.A.J.; Piceni, H.A.L.; Pellegrino van Stuyvenberg, J.A. (in preparation) The effect of noise on the analysis of the transient human visual system. To be submitted to Biol. Cyb.

Watson, A.B.; Nachmias, J. (1977) Patterns of temporal interaction in the detection of gratings. Vis. Res. 17, 893-902.

Watson, A.B. (1979) Probability summation over time. Vis. Res. 19, 515-522.

Watson, A.B. (1982) Derivation of the impulse response: comments on the method of Roufs and Blommaert. Vis. Res. 22, 1335-1337.

Appendix

Quick's (1974) rule for the detection probability of a stimulus, as used by Watson (1982), can be formalised by

$$\Psi = 1 - \prod_{i=1}^{N}\left(1 - \left(1 - 2^{-|R_i|^\beta}\right)\right). \tag{23}$$

Here,

Ψ is the detection chance of the stimulus ($0 \leq \Psi \leq 1$),

R_i is the response amplitude of the sample i occurring at $t_i < t < t_i + \Delta T$.

β is a parameter proportional to the slope of the psychometric function,

N is the number of samples constituting the temporal response function.

Note that R_i in eq. 23 is a dimensionless quantity.

In this equation it is implicitly assumed that positive and negative response values yield the same contribution with respect to probability summation.

The equation can be simply written as

$$\Psi = 1 - \prod_{i=1}^{N} 2^{-|R_i|^\beta}. \tag{24}$$

For a probability of seeing of $\Psi = 0.5$, it can be derived for the threshold condition that

$$0.5 = 1 - \prod_{i=1}^{N} 2^{-|R_i|^\beta}, \tag{25}$$

or:

$$2^{-1} = \prod_{i=1}^{N} 2^{-|R_i|^\beta}. \tag{26}$$

The product at the right hand side can be rewritten as a sum in the power of two:

$$\prod_{i=1}^{N} 2^{-|R_i|^\beta} = 2^{-\sum_{i=1}^{N} |R_i|^\beta}, \tag{27}$$

the threshold condition then becomes:

$$2^{-1} = 2^{-\sum_{i=1}^{N} |R_i|^\beta}, \tag{28}$$

and, after taking the logarithm:

$$1 = \sum_{i=1}^{N} |R_i|^\beta. \tag{29}$$

If the samples of the response function $R(t)$ are sufficiently densily taken, the numerical sum can be approximated by the integral:

$$1 = \frac{1}{\Delta T} \int_0^\infty |R(t)|^\beta dt, \tag{30}$$

or:

$$1 = \int_0^\infty \left| \frac{R(t)}{(\Delta T)^{1/\beta}} \right|^\beta dt. \tag{31}$$

The response function can be expressed as a convolution of the stimulus time function $S(t)$ and the impulse response $U_\delta(t)$. Therefore, and in order to get the dimensions correct, we can substitute for $R(t)$ without loss of generality:

$$R(t) = (\Delta T)^{1/\beta} S(t) \otimes U_\delta(t). \tag{32}$$

The threshold condition then becomes:

$$1 = \int_0^\infty |S(t) \otimes U_\delta(t)|^\beta dt. \tag{33}$$

If a new stimulus time-function $f(t)$ is defined by

$$\varepsilon \frac{f(t)}{a} = S(t), \tag{34}$$

then the threshold condition can be written as:

$$1 = \int_0^\infty \left| \varepsilon f(t) \otimes \frac{U_\delta(t)}{a} \right|^\beta dt, \tag{35}$$

which is condition 15 used for deriving the effect of noise on the impulse response experiments.

chapter 4

The foveal point spread function as a determinant for detail vision[1]

Frans J.J. Blommaert
Jacques A.J. Roufs

Abstract

A point spread function, chosen to link contrast sensitivity and stimulus dimension can be obtained from measured thresholds by assuming small-signal-linearity for the visual system. To that end a special case of summation of subthreshold signals (perturbation) is used, taking specific measures against the effect of sensitivity drift. The basic assumptions are tested simultaneously and confirmed. Other provisional assumptions like radial symmetry and homogeneity were evaluated along a horizontal and a vertical meridian through the fovea. In the fovea no deviation from radial symmetry was found. The effect of inhomogeneity within the central fovea, seems to be too small to cause a significant change in the point spread function. The validity for predicting thresholds of stimuli exposing larger areas is tested. Annuli with varying radii show no significant aberration if probability summation is taken into account. Predicted disk thresholds, however, show a large discrepancy with experiment for radii larger than 2 min arc. A possible extension of the model with multiple-detection units having tuned sizes is evaluated.

4.1 Introduction

In order to investigate characteristics of detail vision, stimuli like Landolt C's, lines or gratings are generally used. Sometimes, radially symmetrical stimuli like disks and annuli are used to investigate lateral processing over short distances (Fiorentini and Maffei, 1970; Westheimer, 1967). The spatial processing of the visual system is rather inhomogeneously over the retina. This is demonted by data on visual acuity (Le Grand, 1967) and thresholds for disks (Kishto, 1970; Wilson, 1970). A review is made by van Doorn et al.(1972). The effect is also found with thresholds of lines (Hines, 1976; Wilson and Bergen, 1979; Limb and Rubinstein, 1977). Wilson and Bergen, ibid, stressing the importance of local measurement of the inhomogeneous system, used line sources instead of gratings as stimuli. However, line sources of the usual type still stimulate retinal parts with different properties.
The use of point sources is an obvious further step, the more so, since it is not clear how much the discrepancy between measured thresholds and prediction based on a

---

[1] This chapter is the text from an article with the same title. It was published in 1981 in: Vis. Res. 21, 1223-1233.

line spread function (e.g. Wilson and Bergen, 1979) is disguised by effects of the mentioned extension of lines.

A point spread or weighting function can be used to characterise the effect of a near-threshold-point-source on the contrast sensitivity of neighbouring retinal points (e.g. van Meeteren, 1973; Kelly, 1975). However, the threshold changes which have to be measured to obtain such a function are very small. Consequently special measures have to be taken to eliminate the effects of unavoidable sensitivity drift in order to obtain the required precision.

In this report the results of a technique to obtain a point spread function from threshold measurements is investigated and the predictive power is tested. This point spread function should represent the simple processing of the visual system, thus including optical effects. We started a first approximation of a model based on four plausible assumptions, namely quasi-linearity, local homogeneity, local radial symmetry and peak detection. On the basis of these assumptions, the spread of the activity of a point source is measured by using a special case of subthreshold summation of signals. The response of the test stimulus is kept so small that it perturbates the signal of the probing stimulus. This will be elaborated on, further on. A major effort was put in testing the basic assumptions.

4.2 Theoretical formalism

We assume that detection of quasi-static luminance increments can be formalised by using a peak detection model, i.e. a stimulus is seen if and only if the extreme value of the model's responce $U(x,y)$ exceeds a certain level D. The threshold condition can be written as

$$extr\{U(x,y)\} = D. \tag{1}$$

Furthermore we postulate that within a small area of the fovea the processing is:

- homogeneous

- radially symmetric

- quasi-linear

The model is visualised in Fig. 1.

The response to an arbitrary small stimulus may, according to these assumptions, be written as a convolution integral with a local point spread function $U_\delta(x,y)$:

$$U(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_\delta(x - x', y - y')\varepsilon_f f(x',y')dx'dy'. \tag{2}$$

Here $f(x',y')$ is the distribution of retinal illumination of the stimulus , ε being the amplitude parameter and $U(x,y)$ is the response of the visual system. From this equation it can be seen that, if the unit point spread function $U_\delta(x,y)$ is known, the response to an arbitrary stimulus pattern can be calculated.
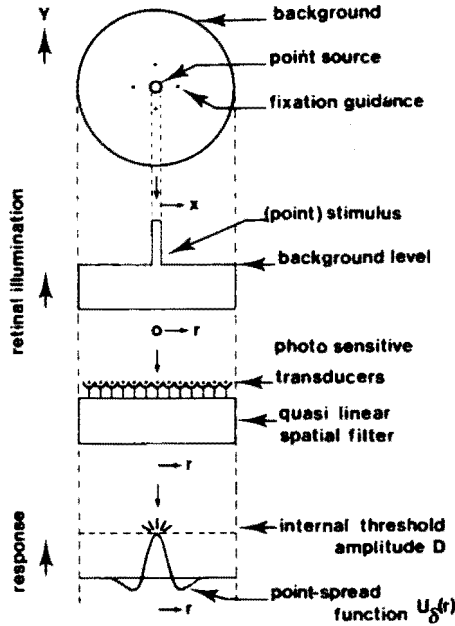
Figure 1: Visualised detection model.

As we only deal with radially symmetrical stimuli in this paper, it is convenient to use polar coordinates, the response to the arbitrary stimulus becomes:

$$U(r) = \int_0^{2\pi} \int_0^{\infty} U_\delta(|\vec{r} - \vec{r}'|)\varepsilon_f f(r')r'dr'd\varphi' \stackrel{\text{def}}{=} \varepsilon_f U'(r'). \tag{3}$$

In this equation, $U_\delta(r)$ is again the unit point spread function, which is radially symmetric according to the assumptions mentioned above; $U'(r)$ is the unit pattern response.

At threshold, the increment ε_f is given implicity by incorporating this result in eq. 1:

$$\varepsilon_f extr\{U'(r)\} = D. \tag{4}$$

Perturbation approach

The unit point spread function can be determined by subthreshold summation. The response of a subthreshold pointstimulus can be probed by the response of another point on the basis of the assumptions mentioned above. By varying the intensity of the probe, the sum of the peak value of its response and the response of the teststimulus at a certain distance can be brought at threshold level. Since radial symmetry is assumed, the test stimulus at distance r can be replaced by a thin annulus with radius r in order to increase the effect.

If the response of the teststimulus is kept sufficiently small with respect to that of the probe, one might say that the response of the teststimulus perturbates that of the probe. This is highly analogous to the situation in the temporal domain explained in detail in Roufs and Blommaert (1981). The spatial case is worked out in the Appendix.

4.3 Apparatus and procedure

A four-way pseudo Maxwellian view optical system was used as is shown in Fig. 2. The subject viewed monocularly an 11 deg uniform field with a retinal illuminance of 1200 Td. The stimuli used were superimposed on this field by using prisms. To facilitate fixation, four lights with a radius of 2 min arc were projected around the stimulus on a circle with a dia of 1 deg.

1	glow modulator
2	condensor
3	glass plate (mirror)
4	photodiode
5	neutral filter
6	stimulus
7	Maxwellian view lens
8	50% prism
9	ocular
10	2 mm artificial pupil
11	diverging lens
12	field lens
13	fixation marks
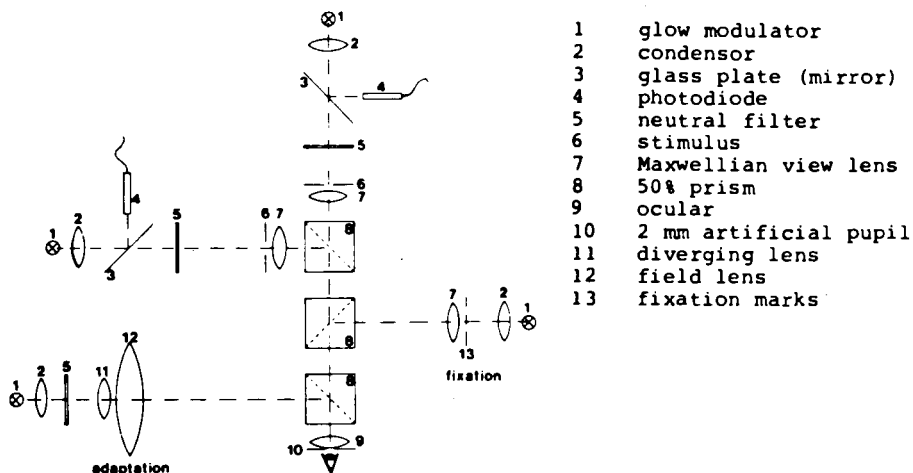
Figure 2: Scheme of the optical apparatus.

A 2 mm artficial pupil was used, which was provided with an entopic guiding system to check the centering of the pupil of the eye (Roufs, 1963). The lights were generated by linearised glow modulators. The time functions used as an approximation of quasi-static presentation of the stimuli consisted of pulses of about 500 $msec$, the beginning and the end of wich were smoothed in order to avoid transient phenomena. The luminance levels of the stimuli were kept in a constant ratio for reasons to be explained further on. The luminance levels were then simultaneously controlled by a dB step attenuator, after presetting the chosen ratio.

Via photodiodes, the luminance levels of the stimuli were recorded during the experiment, so that we were able to correct the results for minor luminance changes which occasionally occured as a result of substantial temperature changes in the control circuitry.

The apparatus was checked for possible straylight artefacts by optical inspection of suitable stimuli using a microscope. No degeneration of any stimulus could be observed so it was concluded that the optical spread of the apparatus was negligible compared with that of the optics of the human eye.

The subject had one knob to release the stimulus, which had to be detected. Three knobs enabled him to answer with "yes", "no" or "rejection". (Rejection" is used if the subject feels he is deprived from the stimulus by lack of attention, blinking etc.) All experimental data consisted of 50% luminance thresholds which were determined by the method of constant stimuli. At every intensity the detected fraction of 20 trails was determined. For one threshold, usually 3 to 4 different intensities were used, measured in random order for one psychometric function.

4.4 Eperiments

Experiment 1. Measuring a foveal point spread function

For this purpose we used a stimulus configuration as is shown in Fig. 3a. It consists of a small point-shaped stimulus (radius $r = 0.4$ min arc), surrounded by a concentric annulus of radius r_a and width Δr_a ($=0.4$ min arc). Both of these were superimposed on a constant background level of 1200 Td.
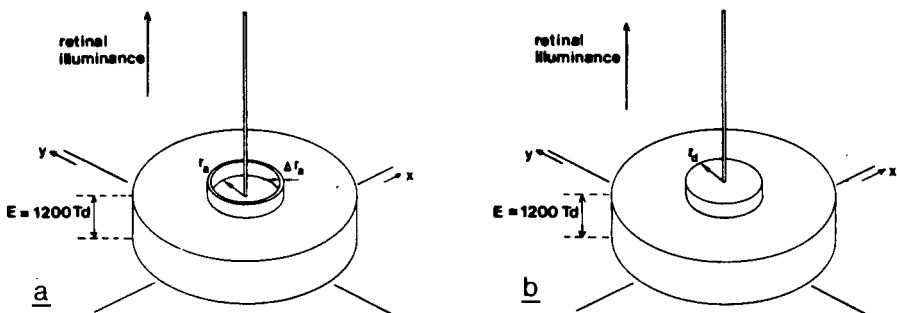
Figure 3: Schematic drawings of the stimulus configurations used in (a) point spread function experiment and (b) test on linearity of processing.

In this experimental set-up, the retinal illumination of the annulus was chosen in such a way that it was always subthreshold (about 0.4 times its threshold value).

Furthermore it was kept in a constant ratio q to the retinal illumination of the point source, the technical realisation of which was explained in the preceding section. In the experiment, we measured the threshold intensity of the point source in the presence of the perturbating annulus, compared with the threshold intensity of the point source without annulus. It can be derived (see appendix) from eq. 3 that a

discrete value for the point spread function (for one value of r, the radius of the annulus) can be calculated from the mentioned thresholds by:

$$U_\delta^*(r) = \frac{A_p}{qA_a} \left\{ \frac{\varepsilon_p}{\varepsilon_{p,a}} - 1 \right\}.$$ (5)

Norm factor:

$$\frac{U_\delta(0)}{D} = \frac{1}{\varepsilon_p A_p}.$$

Here,

$\qquad U_\delta^*(r)$ is the normalised unit point spread function,

$\qquad A_p$ is area of the point,

$\qquad A_a$ is area of the annulus,

$\qquad \varepsilon_p$ is threshold intensity of the point alone,

$\qquad \varepsilon_{p,a}$ is threshold intensity of the point surrounded by the annulus,

$\qquad q$ is constant retinal illuminance ratio of annulus and point source,

$\qquad D$ is threshold level for the response.

By using a number of annuli with different radii r, a discrete number of values for the point spread function can be found.

To minimise the effect of sensitivity shifts of the subjects' system on the resulting value of U_δ^* according to eq. 5 we took the following precautions:

- The paired thresholds ε_p and $\varepsilon_{p,a}$ were always measured fast after one another (fast pair). The quotient $\varepsilon_p/\varepsilon_{p,a}$ is then almost independent of sensitivity shifts of the subject. (for further details on technique and statistical analysis see Roufs and Blommaert, 1981).

- Repetitions of all experimental sessions were carried out in a counterbalanced order, within the fast pairs and within the complete experiment, to minimise the effect of systematic sensitivity changes like fatigue.

Fig. 4a shows the data of the normalised point spread function for one subject (FB). The absolute response, expressed in "D" units, can be found by multiplying the reduced values by the norm factor given in the legend.

This norm factor is found by averaging the threshold ε_p of the point alone over all 18 sessions necessary for the experiments 1 and 2. In this way we tried to acquire an optimal representation of the average sensitivity of the subject. Per session, one point of the curve containing about 1100 trials, was measured, calculated from 8 fast pairs according to eq. 5. The experimentally determined standard deviation of the mean is indicated. The order of measurement with respect to the r-axis was randomly chosen.

Experiment 2. Testing the linearity hypothesis

The stimulus configuration is shown in Fig. 3b and consists of a point-shaped stimulus superimposed on the centre of a subliminal disk on a 1200 Td background
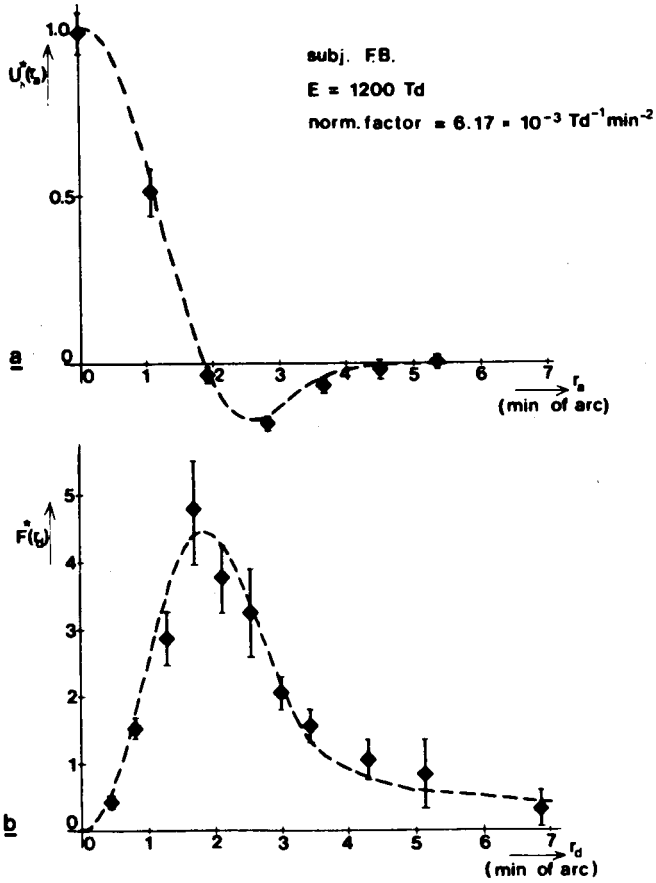
Figure 4: Experimental data of (a) normalised point spread function $U_\delta^*(r)$ and (b) the response to a disk at its centre, with radius r_d of the disk as a parameter. The dashed curves are the results of a simultaneous computer fit. They exactly obey linearity of processing.

level. The retinal illumination of the disk was again chosen in such a way that it was always subthreshold, while it was kept in a constant ratio to the retinal illumination of the point source. This configuration was chosen in order to imitate as closely as possible the experimental design of the first experiment. We argued that in trying to verify the linearity hypothesis, it would be better not to change the conditions with respect to the other basic assumptions.

Again we meaured the threshold of a point source, the same percept to be detected as before. Instead of the ring, a perturbating disk with a variable diameter was used. Again the thresholds of the point source with and without perturbation were measured in fast pairs.

The result of the experiment can be interpreted (since the experiment was carried out for a number of disks with different radii) as a radially symmetrical negative-going edge spread function. If this function is called $F(r)$, it can be calculated (see appendix) that:

$$F^*(r) = \frac{A_p}{q} \left\{ \frac{\varepsilon_p}{\varepsilon_{p,d}} - 1 \right\}. \tag{6}$$

Here,

ε_p is the threshold intensity of the point alone,

$\varepsilon_{p,d}$ is the threshold intensity of the point in the presence of the disk,

q is the illuminance ratio of point and disk.

In experiment 2 we determined the response of a (negative going) radially symmetrical edge. If linearity is true, there has to be a close relation with the response of the point-source of experiment 1. For a linear space invariant system the derivative of the response to a unit edge equals the response to a unit point. If the transformation to polar coordinates is taken into account, this can be formalised in the model by:

$$2\pi r U_\delta^*(r) = \frac{d}{dr} F^*(r). \tag{7}$$

The results for the radially symmetrical edge spread function, according to eq. 6, are shown in Fig. 4b. They were measured during 11 sessions, making an average of one point per session. All other conditions were the same as for the point spread function experiment.

The dashed curves are the result of a simultaneous fitting to all data of both experiments in such a way that the relation of the two curves exactly obeys eq. 7. The curves were obtained by first multiplying the data of Fig. 4a by $2\pi r$. Since we experienced good results if the fitting was done in the frequency domain, the Fourier components of these results were averaged with those of dF^*/dr of Fig. 4b (differentiating was carried out in the frequency domain). Finally, this average Fourier spectrum was transformed back to $U_\delta^*(r)$ and $F^*(r)$ respectively. As can be seen, both curves are highly consistent. This linearity test, which is commonly considered to be sensitive, appears to be positive

Experiment 3. Testing homogeneity and global radial symmetry.

For this purpose, the 50% thresholds of a point source ($r = 0.4$ min of arc) were measured as a function of eccentricity over a horizontal and vertical meridian through the fovea over a distance of about 3 deg. In the case of homogeneity, all thresholds in whatever direction and at whichever eccentricity have to be the same. In the case of radial symmetry with respect to a certain point, the variation of the threshold as a function of the distance to this point has to be independent of direction.
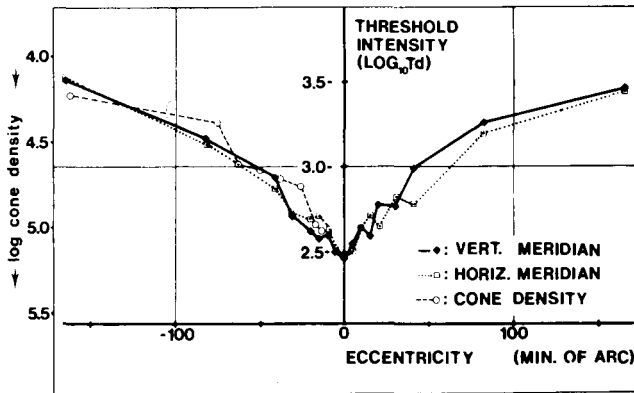
Figure 5: Incremental thresholds of a point source along a horizontal and a vertical meridian through the fovea. For comparison, cone density in N/mm^2 (from Österberg, 1935) is plotted as open circles.

Results with the fovea as a centre are shown in Fig. 5 (squares) on a semi-logarithmic scale. All points are the means of two thresholds. Of the four sessions, the first and second were used for measuring over the horizontal meridian, while the third and fourth sessions were reserved for the vertical meridian. The order of measurement with respect to eccentricity was randomly chosen in the first and third sessions, while the order in the second was the reserved one of the first and the order of the fourth was reversed from the third.

In order to estimate the effect of short range inhomogeneity, the points upto 40 arc min were averaged over all four directions. Within this range, the change of threshold with eccentricity can be described sufficiently accurate by a linear function as illustrated in Fig. 6a. The standard deviations of the mean are also indicated. The straight line through the data points is the result of a regression computation, giving all data points the same weight.

This finding was then incorporated in the model by assuming the retina to be linear-space-variant. The sensitivity was taken to vary with eccentricity according to the reciprocal of this threshold function.

Since sofar we have no data of point spread functions at different eccentricities available, we were not able to take the change of width of the point spread function into
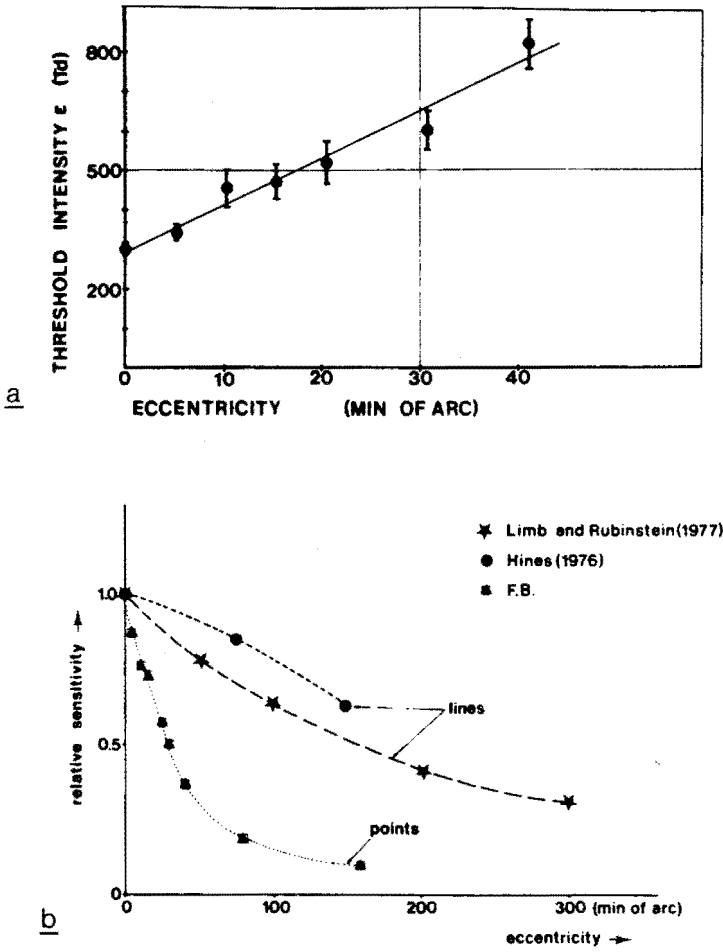
Figure 6: a Averaged thresholds of a point source over four directions through the fovea. The straight line is a regression fit, giving all points the same weight. b Sensitivity relative to the fovea for point and line stimuli. Sensitivity for points is derived from the thresholds as given in Fig. 5. Sensitivity for lines are taken from literature as shown in the legend.

account. However, since the eccentricities used here are small, we do not anticipate this to be a considerable effect, estimating on the data of Hines (1976).

Experiment 4. Testing the threshold prediction of extended stimuli.

To this end, we measured thresholds for annuli and disks with varying diameters and confronted these with the predicted thresholds from the model. An important difference with the foregoing experiments is that instead of a point larger areas have to be detected.
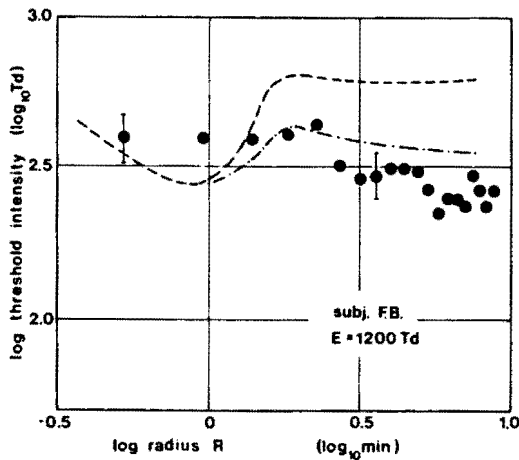
Figure 7: Thresholds for annuli as a function of the mean radius. The dashed curve is the prediction of the model. The dotted line gives a correction for probability summation of the deterministic prediction. The distance between two horizontal bars covers the average difference of two thresholds of which the means are indicated by the dots.

In Fig. 7 thresholds for annuli are plotted as a function of the mean radius. The experimental results are averages over two thresholds. The order of measurement with respect to the radius was randomly chosen in the first session and reversed in the second. The dashed line is the prediction according to equations 1 and 3.
The predicted values are in the right order of magnitude. For diameters larger than about 4 min of arc, the measured values are about 0.3 log units below the predicted ones. However, this is not sufficient to reject the model since we have not yet taken into account the effect cf stochastic fluctuations, which are responsible for the increase of detection probability if the stimulus area is enlarged.
Fig. 8 shows a number of experimental results for disk thresholds as a function of the radius. All data points are averaged over two threshold values. By using reversed order of measurement in different sessions, counterbalancing was again used as a tool to minimise the effect of systematic sensitivity drift on the final result. The dashed line is again the prediction from the model.
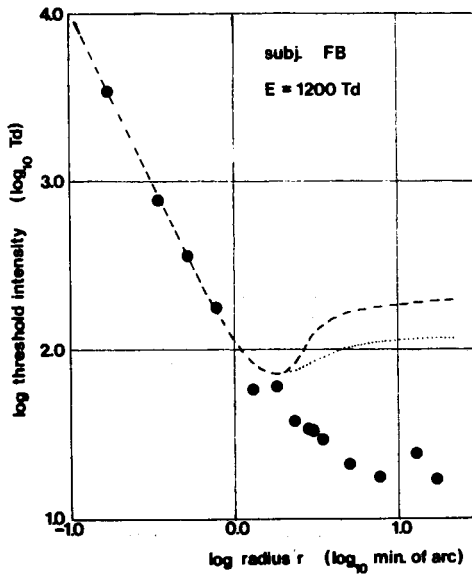
Figure 8: 50% thresholds for incremental foveal disks as a fuction of the radius. The dashed curve is the deterministic prediction from the model. The dotted line is a correction for probability summation.

Keeping in mind that there is no free parameter, the prediction for small diameters is very good. However, for diameters larger than about 4 min of arc, there is a distinct deviation between prediction and experiment. This will be discussed in greater detail further on.

4.5 Discussion

As one can clearly see from Fig. 4, the responses measured with the perturbation approach are fairly large in comparison with the spread.
The measured point spread function of our subject, has about the same shape as Fiorentini and Maffei's (1970) based on the effect of steady annuli with variable diameters on the incremental threshold of a small disk (dia 1.7 min arc). They also concluded for an excitatory centre surrounded by an inhibitory region. However, the spatial extension of our point spread function is somewhat less (width of centre of about 4 min arc against 7 min arc for Fiorentini and Maffei), which is probably due to our adaption level of 1200 Td against 100 Td in their experiments. A difference of about factor 2 due to these different adaptation levels, is to be expected on the basis of visual acuity data as a function of adaption level (see for instance Nakane and Ito, 1978).
It is of some interest to compare the present point spread function with line spread functions from literature assuming an infinite length, linearity and homogeneity. Using Hankel and Fourier algorithms, transformed data of Hines, 1976; Kulilowski and King-Smith, 1973; Limb and Rubinstein, 1977 are drawn in Fig. 9. Although one has to

take into account the effect of inhomogeneity, different line lengths and background levels used, the differences in the obtained point spread functions suggest that this kind of transformation is too simple to obtain a satisfactory point spread function.

authors	subject	adaptation level	line length
Hines (1976)	— MH.	30 $^{cd}/_{m^2}$	1.6°
Kulikowski and King-Smith (1973)	.....·· JJK.	5 $^{cd}/_{m^2}$	2.5°
Limb and Rubinstein (1977)	- - - - JQL.	17 $^{cd}/_{m^2}$	raised cosine half-width 1.5°
this paper	—·— F.B.	1200 Td	
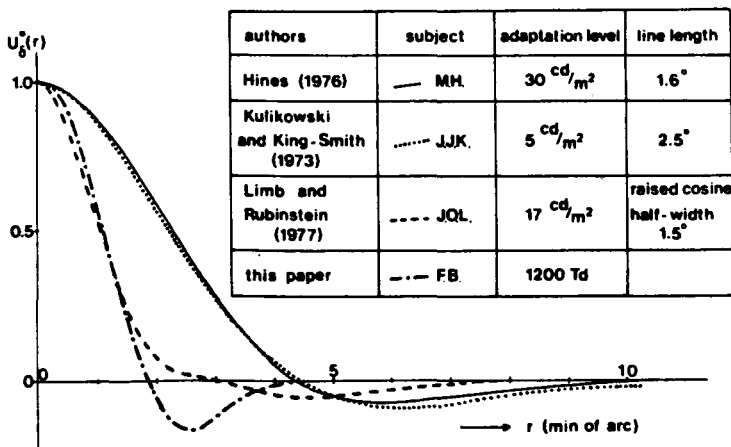
Figure 9: Normalised point spread functions derived from line spread functions measured by different authors and calculated on the basis of assumptions, mentioned in the text. For comparison the measured point spread function of Fig. 4. is shown.

According to Fiorentini et al. (1972), who found no summation if a disk is presented to one eye and an annulus to the other, the described interaction between signals of different retinal elements has to take place at a location before binocular convergence of signals occurs.

Furthermore, it is noteworthy that the width of the excitatory centre at this relatively high background level is not much larger than the optical spread given by Vos et al. (1976) for human eyes with pupil dia of 2 mm. This could mean that from the exitatory lateral processing in the fovea, the spread of the optical system is at least an important part. This probably is not the case at low levels of adaptation or in the periphery, as will be shown in a subsequent paper.

Fig. 4 shows that, within measuring precision, all experimental points lie on the dashed curves, which obey the basic linearity and peak detection postulates exactly. Linear lateral processing seems to be a good approximation in this case.

Looking at the three-dimensional plots of point-and disk-responses computed from the model, as shown in Fig. 10, it can be appreciated why $F^*(r)$ of Fig. 4b becomes zero for large values of r: For small disk radii the core of the point spread function determines the intergrating effect; for large radii the inhibiting part gradually diminishes the effect.

From Fig. 5, which shows the results for thresholds of a point source as a function of eccentricity, it can be seen that there is no significant difference between the results on the vertical and the horizontal meridian. Since we did not measure in oblique
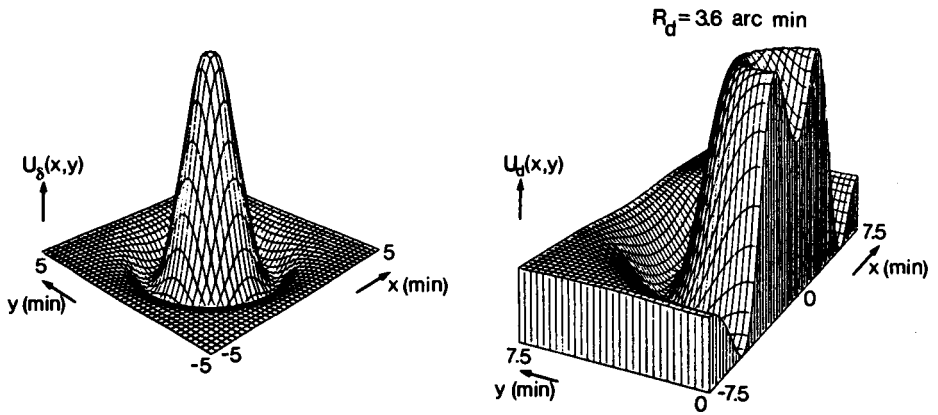
Figure 10: Left: three-dimensional plot of the point spread fuction of Fig. 4a. Right: intersection of 3D plot of calculated response of the model on a disk with radius $r_d = 3.6$ min arc.

directions, aberrations from radial symmetry of this eccentricity effect cannot be excluded. The magnitude of the threshold variations are compatible with findings of Kistho and of Wilson (1970), who measured with disks.

Note from Fig. 5 that the effect follows very closely the density of cones as a function of eccentricity as reported by Österberg (1935), although the connection is not obvious since other mechanisms take part. Compared with the effect of eccentricity on sensitivity as measured with lines (Hines, 1976; Wilson and Bergen, 1979 and Limb and Rubinstein, 1977), ours is more pronounced (Fig. 6b). One reason is probably that the spatial extension of lines flattens out the actual retinal inhomogeneity. Another is that our background luminance is higher.

One aspect of the measurement should not be left unmentioned, i.e. the possible effect of involuntary eye movements on the results of the experiments. As can be seen from Fig. 5, the fovea itself looks like some sort of singular point. Because of eye movements, the real singularity may very well be flattened out by the averaging effect that involuntary eye movements have. According to Ditchburn (1973), the magnitude of the eye movements in a situation where good fixation is maintained, can be described roughly by a normal distribution with a σ-value of 2 to 3 min arc. This would apply to our conditions (self release by subject after fixation; fixation periods of less than a second) and indicates a possible smoothing of the real inhomogeneity as a result from averaging within a gaussian window of 2-3 min arc.

Concerning the perturbation experiments, it should be kept in mind that small saccades are not likely to interfere much with the results of the experiment because the relative position of point- and perturbation-stimulus is not affected by eye movements.

The data of Fig. 5 were averaged over all four directions and replotted for a limited range in Fig. 6a. The increase of threshold as a fuction of eccentricity can be well approximated by a straight line over a distance of 40 min arc from the fovea. Taking the effect of inhomogeneity into account, the original model was transformed to a spatial variant one in such a way, that the norm factor of the point spread fuction decreases with the distance to the fovea as the reciprocal of the values prescribed by the straight line in Fig. 6a.

It turned out that in view of the distances involved the corrections to be applied to the point and edge spread funtion were small with respect to the experimental error.

In Fig. 7, the experimental thresholds for annuli are compared with the prediction of the model. Taking into account the effect of daily sensitivity variations, the prediction of thresholds of annuli with small radii is not bad. For radii larger than about 2 min arc there seems to be some discrepancy. However, this can be explained be thresholds decrease caused by the increase of probabiliy of detection of larger areas, the so called probability summation.

A rough estimation of the effect was carried out by treating the problem as a discrete one. It was assumed that foveal areas of 1 min had an independent chance of reaching a stochastic threshold value. On the basis of this, it was calculated how many independent areas annulus responses had in comparison to a point response (which, by definition, consistend of one area). By using stochastic calculations based on the slope of the measured psychometric functions (cf Roufs, 1974) the effect of increased probability on the 50% threshold was then calculated and the corrected values are shown as dotted lines in Fig. 7. It was found from calculations based on integration areas of 0.5 and 2 min^2, not shown here, that for the present stimulus dimensions the effect does not depend heavily on the knowledge of the exact integration area.

For the disks, Fig 8 gives the experimental thresholds, together with prediction. Now, for radii larger than 2 min arc, a substantial difference is observed between prediction and experiment, which increases as the diameter of the disk increases. A similar discrepancy between model and experiment is reported by investigators who work with line spread fuctions: if thresholds for bars are compared with theoretical thresholds from a single-unit model, the experimental thresholds are usually lower than the predicted ones (Kulikowsky and King-Smith, 1973; Hines, 1976). There just seems to be more integration by the visual system than is expected on the basis of point or line spread functions. With lines the results may be distorted by the effect of their extension over a relatively large part of the inhomogeneous retina. Using a point spread function, which would provide a clearer picture in that aspect the discrepancies found (for disks) are even larger.

One approach to deal with this partial disaccordance has been provided by among others: Sperling (1970), Thomas (1970), Koenderink and van Doorn (1978) and Wilson and Bergen (1979). These authors have propositions of the visual system in common, which consist of an ensemble of units (receptive fields or line spread functions) working at one spatial coordinate of the retina and that vary in extension and sensivity. Bagrash's (1973) psychophysical evidence for size-tuning seems to

support this idea. The most detailed model is worked out by Wilson and Bergen (1979) which has four line spread mechanisms working at each retinal eccentricity that differ in width and sensitvity and contains probability summation, which is formalised in an algotithm put forward by Quick (1974).

Exploring the potentialities of such a multiple unit model we assumed 4 isomorphic point spread mechanisms, the narrowest being equal to the measured one. The width of every next mechanism is increased by a factor 2 and the amplitude is fitted to the disk threshold data as shown in Fig. 11. Such an extension of the model is apparently able to account for the disk data. If the mechanisms are indepentent or weakly dependent like in the Wilson and Bergen model this does not alter the point and the edge spread fuctions nor does it change the prediction of the thresholds of annuli. (A test in relation with other stimulus shapes will be reported in a subsequent paper).
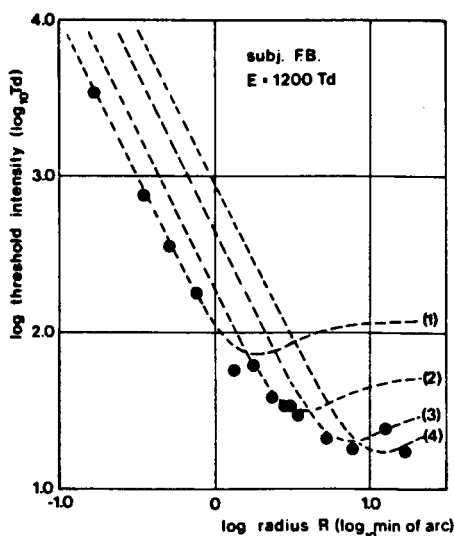
Figure 11: Illustration of fitting a "four-mechanism model" to disk thresholds as a function of the radius. The smallest mechanism (1) is described completely by the measured point spread function, probability summation included. Mechanism (2), (3) and (4) have the same shape but increase in width with a factor 2. The dashed curves are the predictions from the individual mechanisms.

In the view of the scatter in receptive field sizes found electrophysiologically in the retina (Fisher, 1973) and in the striate cortex (Hubel and Wiesel, 1974) a four mechanisms model seems rather artificially. On the other hand it might be a way to approximate a more naturel but also more difficult to handle model as the one proposed by Koenderink and van Doorn (1978), having a continuous receptive field size distribution.

Conclusions

Point and edge spread fuctions determined by using a point source as a probe are quantitatively mutually consistent.

If the point spread fuction is used to predict thresholds of larger stimuli the effect of increased detection probability extending the stimulus area does effect a non neglectable threshold decrease.

Quantitative correction based on independent area elements seems to be adequate for slim stimuli. However, "probability" summations cannot explain the threshold lowering for more "square" stimuli.

A multi-mechanism model, including probability summation can account for our results but is not necessarily the best one.

References

Bagrash, F.M. (1973) Size-selective adaptation: psychophysical evidence for size-tuning and the effects of stimulus contour and adapting flux. Vision Res. 13, 575-598.

Barbur, J.L.; Ruddock, K.H. (1980) Spatial characteristics of movement detection mechanisms. Human Vision-I Achromatic vision. Biol. Cybernet. 37, 77-92.

Blommaert, F.J.J. (1977) Spatial processing of a small visual stimuli. IPO Ann. Prog. Rep. 12, 81-86.

Ditchburn, R.W. (1973) Eye-Movements and Visual Perception. Clarendon Press, Oxford.

Doorn, A.J. van; Koenderink, J.J.; Bouman, M.A. (1972) The influence of the retinal inhomogeneity on the perception of spatial patterns. Kybernetik 10, 223-230.

Fiorentini, A.; Maffei, L. (1970) Transfer characteristics of excitation and inhibition in the human visual system. J. Neurophysiol. 33, 285-292.

Fiorentini, A.; Bayly, E.J.; Maffei, L. (1972) Peripheral and central contributions to psychophysical spatial interactions. Vision Res. 12, 253-259.

Fiorentini, A. (1977) Centre-surround concepts, luminance gradients and contrast. In Spatial Contrast, Report of a Workshop (Edited by Spekreijse H. and Tweel L.H. van der). 38-40. North Holland Amsterdam.

Fischer, B. (1973) Overlap of receptive field centers and represation of the visual field in the cat's optic tract. Vision Res. 13, 2113-2120.

Hines, M. (1976) Line spread fuction variation near the fovea. Vision Res. 16, 567-572.

Hubel, D.H.; Wiesel, T.N. (1974) Uniformity of monkey striate cortex: a parallel relationship between field size, scatter, and magnification factor. J. comp. Neurol. 158, 295-306.

Kelly, D.H. (1975) Spatial frequency selectivity in the retina. Vision Res. 15, 665-672.

Kistho, B.W. (1970) Variation of the visual threshold with retinal location. Vision Res. 10, 745-767.

Koenderink, J.J.; Doorn, A.J. van (1978) Visual Detection of spatial contrast: influence of location in the visual field, target extent and illuminance level. Biol. Cybernet. 30, 157-167.

Kulikowski, J.J.; King-Smith, P.F. (1973) Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation. Vision. Res. 13, 1455-1478.

Meeteren, A. van (1973) Visual aspects of image intensifaction. Thesis Utrecht.

Le Grand, Y. (1967) Form and Space Vision. Indiana Univ. Press. Bloomington.

Limb, J.O.; Rubinstein, C.B. (1977) A model of threshold vision incorporating inhomogeneity of the visual field. Vision Res. 17, 571-584.

Nakane, Y.; Ito, K. (1978) Study on standard visual acuity curves for better seeing in lighting design. J. Light Vis. Env. 4, 38-44.

Österberg, G. (1935) Topography of the Layer of Rods and Cones in the human Retina. Arnold Busch. Copenhagen.

Quick, R.F. (1974) A vector-magnitude model of contrast detection Kybernetik 16, 65-67.

Rodieck, R.W.; Stone, J. (1965) Analysis of receptive fields of cat retinal ganglion cells. J. Neurophysiol. 28, 833-849.

Roufs, J.A.J. (1963) Perception lag as function of stimulus luminance. Vision Res. 3, 81-91.

Roufs, J.A.J. (1974) Dynamic properties of vision VI Stochastic threshold and their effect on flash-to-flicker sensitivity ratio. Vision Res. 14, 871-888.

Roufs, J.A.J.; Blommaert, F.J.J. (1975) Pulse and step response of the visual system. IPO Ann. Prog. Rep. 10, 60-67.

Roufs, J.A.J.; Blommaert, F.J.J. (1981) Temporal impulse and step responses of human eye obtained psychophysically by means of a drift-correcting perturbation technique. Vision Res. 21, 1203-1223.

Sperling, G. (1970) Model of visual adaption and contrast detection. Percept. & Psychophys. 8, 143-157.

Thomas, J.P. (1970) Model of the fuction of receptive fields in human vision. Psychol. Rev. 77, 121-134.

Thomas, J.P. (1978) Spatial summation in the fovea: asymmetrical effects of longer and shorter dimensions. Vision Res. 18, 1023-1029.

Vos, J.J.; Walraven, J.; Meeteren, A. van (1976) Light profiles of the foveal image of a point source. Vision Res. 16, 215-219.

Westheimer, G. (1976) Spatial interaction in human cone vision. J. Physiol. 190, 139-154.

Wilson, H.R.; Bergen, J.R. (1979) A four mechanism model for threshold spatial vision. Vision Res. 19, 19-32.

Wilson, M.E. (1970) Invariant features of spatial summation with changing locus in the visual field. J. Physiol. 207, 611-612.

Appendix

For a small stimulus approximating a point source with retinal illuminance ε_p and radius r_0, the respons pattern of the visual system can, according to eq. 3, be written as

$$U_p(r) = \varepsilon_p \int_0^{2\pi} \int_0^{r_0} r' U_\delta(|\vec{r} - \vec{r}'|) dr' d\varphi'.$$

Due to the basic assumptions for determining threshold and assuming that the point is situated foveally, only the response for $r = 0$ is of interest, so

$$U_p(0) = \varepsilon_p \int_0^{2\pi} \int_0^{r_0} r' U_\delta(r') dr' d\varphi'.$$

If r_0 is taken sufficiently small, this can be approximated by

$$U_p(0) \approx \varepsilon_p A_p U_\delta(0),$$

where A_p is the area of the point.
The threshold condition of a point may now be formalised as

$$\varepsilon_p A_p U_\delta(0) = D. \tag{8}$$

If perturbation of a point response is applied with an arbitrary shape, eq. 8 changes into

$$\varepsilon_{p,pert} A_p U_\delta(0) + U_{pert}(0) = D. \tag{9}$$

Here $U_{pert}(0)$ is the response in the origin of the perturbating stimulus. $\varepsilon_{p,pert}$ is the illumination of the point source necessary for detection of the combination.
Of course, eq. 9 is valid only if the retinal illuminance of the perturbation is so small that detection is always governed by the extremum of the point response. Since D is in fact a stochastic magnitude, the intensity of the test stimulus should also be kept small in order to avoid a substantial contribution of its response to the detection chance.
In the first experiment we used an annulus with mean radius r_a and width Δr_a.
It can easily be verified from eq. 3 that for $\Delta r_a \ll r_0$ its response in the origin can be approximated by

$$U_\delta(0) = \varepsilon_a 2\pi r_a \Delta r_a U_\delta(r_a) = \varepsilon_a A_a U_\delta(r_a).$$

If $\varepsilon_a/\varepsilon_{p,a} = q$, where $\varepsilon_{p,a}$ is the threshold of the point-annulus configuration, we can write for the threshold condition of this combination

$$\varepsilon_{p,a}\{A_p U_\delta(0) + q A_a U_\delta(r_a)\} = D.$$

As pointed out above, q should be small.
Comparing the threshold with and without annulus $\varepsilon_{p,a}$ and ε_p we are able to derive for the normalised point spread function $U_\delta^\star(r_a)$

$$U_\delta^\star(r_a) = \frac{U_\delta(r_a)}{U_\delta(0)} = \frac{A_p}{q A_a} \left\{ \frac{\varepsilon_p}{\varepsilon_{p,a}} - 1 \right\}.$$

We can again find the absolute response by multiplying U_δ^* with its absolute value $U_\delta(0)$ expressed in "D" units.

In the second experiment we used a disk radius r_d as perturbation.

It can be seen from eq. 3 that for the response pattern of a uniform disk at $r = 0$ we can write (subscript "d" for disk)

$$U_d(0) = 2\pi\varepsilon_d \int_0^{r_d} rU_\delta(r)dr.$$

If $\varepsilon_d/\varepsilon_{p,d} = q$, this leads to the threshold condition

$$\varepsilon_{p,d}\{A_p U_\delta(0) + 2\pi q \int_0^{r_d} rU_\delta(r)dr\} = D.$$

From the threshold change we can derive for the response $F^*(r_d)$ on a negative-going radial symmetrical edge:

$$F^*(r_d) = 2\pi \int_0^{r_d} rU_\delta^*(r)dr = \frac{A_p}{q}\left\{\frac{\varepsilon_p}{\varepsilon_{p,d}} - 1\right\}. \tag{10}$$

By using disks with different radii r_d, we will find a discrete number of samples for $F^*(r_d)$ which has a unique relationship with the normalised point spread fuction $U_\delta(r)$ as is expressed in eq. 7.

chapter 5

Point spread function variation and visibility of details[1]

Frans J.J. Blommaert
Henny G.M. Heijnen
Jacques A.J. Roufs

Abstract

Point spread functions, meant to characterise local spatial transfer of the visual system, can be obtained psychophysically using a perturbation technique. Data of such point spread functions are shown for three experimental conditions: foveally at adaptation levels of 1200 Td and 10 Td, and in the parafovea at an eccentricity of 2 deg using a 1200 Td adaptation level. The results are consistent with earlier findings (Blommaert and Roufs, 1981). On the basis of such point spread functions a simplified multiple unit model was constructed, the parameters of which were fitted to thresholds of discs with varying diameter. Threshold predictions from this model for annuli, thin lines and broad lines were found to be in fair quantitative agreement with experimental results. It is argued that for a certain class of slender stimuli, including alphanumeric characters, thresholds can be described with a single channel model containing only the experimentally determined point spread function as a basic function.

5.1 Introduction

For quite some time, psychophysical investigators have tried to establish generalising rules in order to predict contrast thresholds of details (Blackwell, 1946, Lamar et al., 1947). In the last few decades, a systems analysis kind of approach has frequently been applied, based on the assumption that, at threshold level, the visual system operates linearly. Using this approach, one tries to characterise the visual system by a single operator like its Modulation Transfer Function, MTF (Campbell and Robson, 1968), its Line Spread Function, LSF (Hines, 1976), or its Point Spread Function, PSF (van Meeteren, 1973; Kelly, 1975; Barbur and Ruddock, 1980; Blommaert and Roufs, 1981).

On the basis of such functions, contrast thresholds of arbitrarily shaped stimuli can be calculated and tested against experimental results. As we pointed out in an earlier paper (Blommaert and Roufs, 1981), MTFs and LSFs have a serious disadvantage since they are measured using stimuli that are extended over a rather inhomogeneous retina. For the experimentally determined PSFs, this restriction does not hold since strictly local stimuli such as points and thin annuli are being used. On the other hand, the determination of PSFs is hampered by the fact that threshold *differences*, being

---

[1] Submitted to Spatial Vision.

second order effects, have to be measured. This difficulty can be overcome by using special drift-correcting techniques as described by Roufs and Blommaert (1981).

After the pioneering work of Campbell and Robson (1968), it is commonly assumed that the visual system does not operate as a single channel, but that at a single retinal position channels with different spatial properties operate in parallel.

This approach has been worked out in the frequency domain by many investigators. However,in view of the inhomogeneous nature of the retina and uncertainties about phase preservation, the frequency approach does not seem to be quite appropriate for the purpose of calculating thresholds of details.

In the spatial domain there exist some examples of multiple channel models like the sunflower model of Koenderink and van Doorn (1978) and the four-mechanism model of Wilson and Bergen (1979). These models seem to offer a better connection with the approach based on experimentally determined local systems-functions. Considering that the Wilson and Bergen model is based on essentially one-dimensional spread functions of lines traversing an inhomogeneous retina, a model based on point spread functions seems more attractive.

Unfortunately, the method developed sofar allows only the measurement of the point spread function of the narrowest channels. Therefore, we replaced the continuous ensemble of channels by four clearly separated channels, the same number as Wilson and Bergen used, and of which the narrowest is represented by the experimentally determined point spread function. On the basis of a few assumed properties, the parameters of the other three channels can be estimated, and prediction of thresholds of other stimuli are admissible for tests.

Since in our approach experimentally determined point spread functions play an important role as a characteristic function, experimental data are shown for measured PSFs under different experimental conditions, e.g. at a different adaptation level and eccentricity. At the photopic level of 1200 Td, the model-parameters were fitted to the data in order to predict thresholds of differently shaped stimuli like annuli, thin lines and rectangles.

The model does fit the data well, but it is argued that such a detailed model is not really necessary to predict thresholds for the important class of stimuli which are constructed from thin lines and annuli (thus including alphanumeric symbols). In order to predict thresholds for such stimuli, it suffices to use a subset of the ensemble model containing only point spread functions measured at different retinal positions. It is argued that detection of these stimuli is mainly determined by the smallest units, which also cause the experimentally determined point spread functions.

5.2 Methodological concepts

In the analysis of linear systems, a point spread function is commonly defined as the response to an infinitesimally small unit mass. In the case of optical transfer, the unit mass can be interpreted as an ideal point source whose energy equals 1 (see for instance Jones, 1958). For the eye acting as a linear system for small signals, this mass can further be specified as an ideal point source of which the product of area

and retinal illumination equals 1. The response is to be understood as a signal which generates the internal image and of which the strength is measured psychophysically.
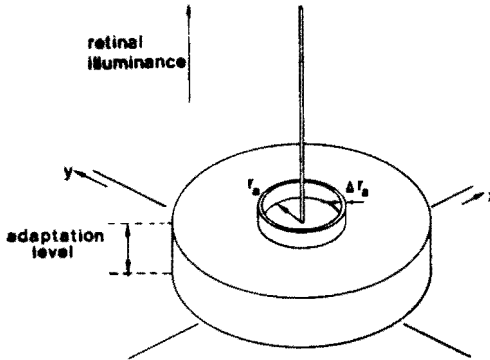
Figure 1: Schematic drawing of the stimulus configuration used in the point spread function experiments.

For the purpose of measuring point spread functions, the stimulus configuration of Fig. 1 was chosen. It consists of a point source situated in the centre of an annulus whose radius could be parametrically varied. Both point and annulus were superimposed on a large homogeneous background field of 10 deg.

The basic assumptions are local radial symmetry and linear spatial processing of the small signals at threshold level. Furthermore we assume that detection may be approximated by a peak detector. On the basis of these assumptions we are able to show that a point spread function can be obtained by probing the response of a subthreshold annulus with the response of a point source situated in the centre of the annulus (for a more detailed description of theoretical issues see Blommaert and Roufs, 1981).

Keeping in mind that multiple channels operate in parallel, it is implicitly assumed here that the responses are dominated by the narrowest units, or in case of a continuous ensemble, by a collection of the smallest units.

Let $\varepsilon_p A_p U_\delta(r)$ be the visual response of the point source having an area A_p and an increment of retinal illumination ε_p, $U_\delta(r)$ being the unit point spread function. The response $\varepsilon_a U_a$ of a thin annulus with radius r_a and width Δr_a, at the position of the probe in the centre of the annulus, will then be:

$$\varepsilon_a U_a = \varepsilon_a 2\pi r_a \Delta r_a U_\delta(r_a) = \varepsilon_a A_a U_\delta(r_a), \tag{1}$$

where A_a = area of the annulus.

Now we keep $\varepsilon_a/\varepsilon_p = \rho$ at a constant value and ρ is small enough to ensure that the ratio of the response of the annulus and that of the probe is sufficiently small and at

the same time the annulus response is also sufficiently below threshold. In formula:

$$extr\{\varepsilon_a U_a\} \ll extr\{\varepsilon_p U_p\}, \tag{2}$$

or:

$$\varepsilon_a A_a U_\delta(0) = q\varepsilon_p A_p U_\delta(0), \tag{3}$$

$$q \ll 1, \text{ say about } 0.3,$$

or:

$$\rho\varepsilon_p A_p U_\delta(0) = q\varepsilon_p A_p U_\delta(0), \tag{4}$$

$$\rho = q\frac{A_p}{A_a}. \tag{5}$$

The amplitude of the response of the combination in the origin at threshold is given by:

$$extr\{\varepsilon_p A_p U_\delta(0) + \varepsilon_a A_a U_\delta(r_a)\} = D, \tag{6}$$

where D is the amplitude for 50% detection probability. If the response of the annulus is kept small with respect to that of the point probe:

$$\varepsilon_a A_a \frac{U_\delta(r_a)}{D} \ll \varepsilon_p A_p \frac{U_\delta(0)}{D}, \tag{7}$$

then from eq. 6:

$$A_p\frac{U_\delta(0)}{D} + \rho A_a\frac{U_\delta(r_a)}{D} = \frac{1}{\varepsilon_p(r_a)}. \tag{8}$$

By varying r_a, a discrete number of values for the point spread function, expressed in D-units, can be found from the variable part $1/\varepsilon_p(r_a)$. Subtracting the reciprocal of the threshold intensity $\varepsilon_p(0)$ of the point source alone leads to the simple formula:

$$\frac{U_\delta(r_a)}{D} = \frac{1}{\rho A_a}\left(\frac{1}{\varepsilon_p(r_a)} - \frac{1}{\varepsilon_p(0)}\right). \tag{9}$$

In principle, the point spread function (in D-units) can be obtained from eq. 9. However, in practice it takes a considerable amount of time to perform these measurements. Sensitivity changes during this time may cause a large spread. Therefore, the threshold of the probe alone is measured as a sensitivity reference immediately before or after the determination of the threshold of the combination. Dividing $U_\delta(r_a)$ by its extreme value $U_\delta(0)$ we find for the normalised point spread function:

$$U_\delta^\star(r_a) = \frac{A_p}{\rho A_a}\left(\frac{\varepsilon_p(0)}{\varepsilon_p(r_a)} - 1\right). \tag{10}$$

In order to minimise the effect of sensitivity shifts of the subject's visual system on the resulting values of $U_\delta^\star(r_a)$ according to eq. 10, we took the following precautions:

- The paired thresholds $\varepsilon_p(r_a)$ and $\varepsilon_p(0)$ were always measured quickly one after the another (fast pair). The quotient $\varepsilon_p(0)/\varepsilon_p(r_a)$ is then almost independent of the sensitivity shifts of the subject.

- Repetitions of all experimental sessions were carried out in a counterbalanced order, within the fast pairs and within the complete experiment to minimise the effect of systematic sensitivity changes like fatigue.

For further details on methodology see Roufs and Blommaert (1981). Within the model, the response to an arbitrary stimulus pattern $\varepsilon_s S(x,y)$ can be calculated by convolution:

$$\frac{\varepsilon_s U(x,y)}{D} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\varepsilon_s S(x',y')}{D} U_\delta(|\vec{r}-\vec{r}'|)dx'dy',$$

$$|\vec{r}-\vec{r}'| = \sqrt{(x-x')^2+(y-y')^2}.\tag{11}$$

At a later stage, space variancy of the point spread function will be added.

5.3 Apparatus and procedure

1	glow modulator
2	condenser
3	glass plate
4	photodiode
5	prism
6	shutter
7	mirror
8	neutral filters
9	stimulus masks
10	lenses
11	fixation mask
12	artificial pupil
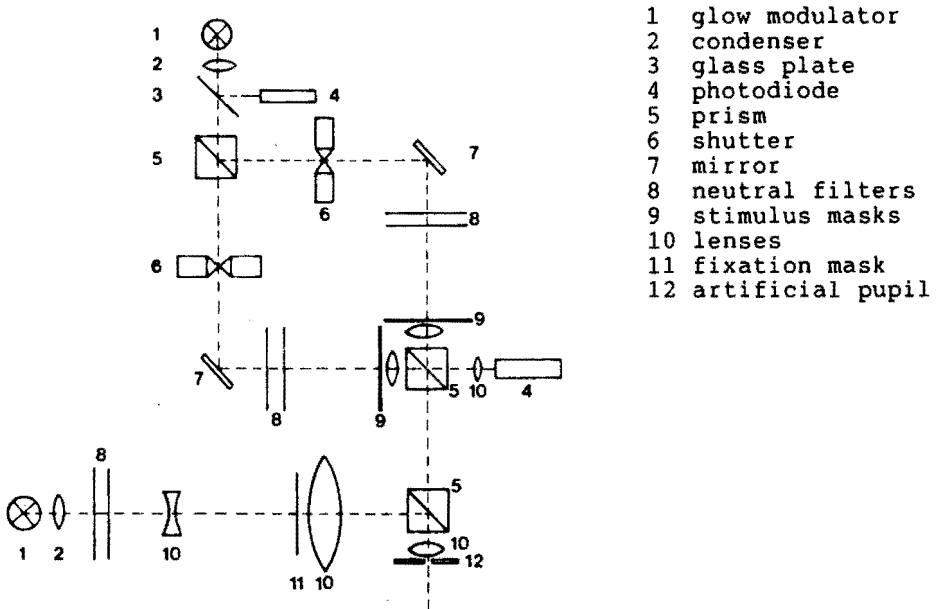
Figure 2: Scheme of the optical apparatus

A three-way pseudo Maxwellian view optical system was used as is shown in Fig. 2. The subject viewed monocularly a uniform adaptation field which subtended an angle of 10 degrees. The stimuli used were superimposed on this field by using prisms. In the case of point spread function experiments, both probe and test stimulus were generated by using only one glow modulator, the light from which was split into two equal beams by means of 50% prisms. In this way, the ratio of luminance levels of test and probe stimulus was kept constant in spite of substantial slow luminance changes

in the control circuitry due to temperature effects. Once this ratio was adjusted by means of neutral density filters, both luminace levels could be varied simultaneously by a one-dB step attenuator.

With a photodiode, properly corrected with respect to spectral sensitivity, the luminance of the glow modulator was measured during every stimulus presentation, thus enabling us to correct afterwards the actual luminance levels.

A fixation aid enabled the subject to fixate his eye at a specified point. During all experiments, an artificial pupil with a diameter of 2 mm was used. It was provided with an entoptic guiding system to check the centering of the pupil of the eye. (Roufs, 1963).

Since in the experiment rather small stimuli were used, diffraction had to occur, leading to a broadening of the light beam at the place of the artificial pupil. Therefore, for every stimulus used, correction factors were measured by means of a highly sensitive photomultiplier, to compensate for the loss of light due to diffraction.

The light was generated by linearised glow modulators. The time functions used as an approximation of quasi-static stimulus presentation consisted of pulses of about 100 ms, whose on- and off-going steps were made approximately Gaussian (each of about 400 ms duration) in order to avoid transient phenomena. The subject had one knob to release the stimulus which had to be detected. Three knobs enabled him to answer with "yes", "no" or "rejection" ("rejection" was used if the subject felt he had been deprived from the stimulus due to lack of attention, blinking etc.).

All experimental data consisted of 50% detection thresholds determined by the method of constant stimuli. At every stimulus intensity the detected fraction of 20 trials was determined. For one threshold usually 4 different intensities were used, measured in random order for one psychometric function. Precision was improved by averaging over measurements repeated in randomised blocks. An exception was made for the point spread function experiments where the thresholds were determined for only one perceived fraction of 20 trials, provided the score was between 20% and 80%. This was done using a constant slope based on an earlier measured Crozier quotient, defined as the standard deviation σ of the distribution underlying the psychometric function, divided by the 50% threshold ε. For subject HH, the Crozier quotient taken was $\sigma/\varepsilon = 0.23$, being the average of earlier experimental findings. This was done in order to measure the elements of a fast pair as quickly as possible after one another in an attempt to minimise the obscuring effects of sensitivity shifts.

For the purpose of measuring point spread functions, we used a small point-shaped stimulus (radius $r = 0.7$ arc min) surrounded by a concentric annulus of radius r_a and width $\Delta r_a = 0.23$ arc min. Both of these were superimposed on a constant background level as shown in Fig. 1. In the experimental set-up, the retinal illumination of the annulus was chosen to be always below its threshold (about 0.3 times its threshold value).

In the experiment we measured fast pairs consisting of two thresholds viz. the threshold of the point source in the presence of the annulus and the threshold of the point source without annulus.

5.4 Results

Point spread function determination and variation with subject, background level and eccentricity
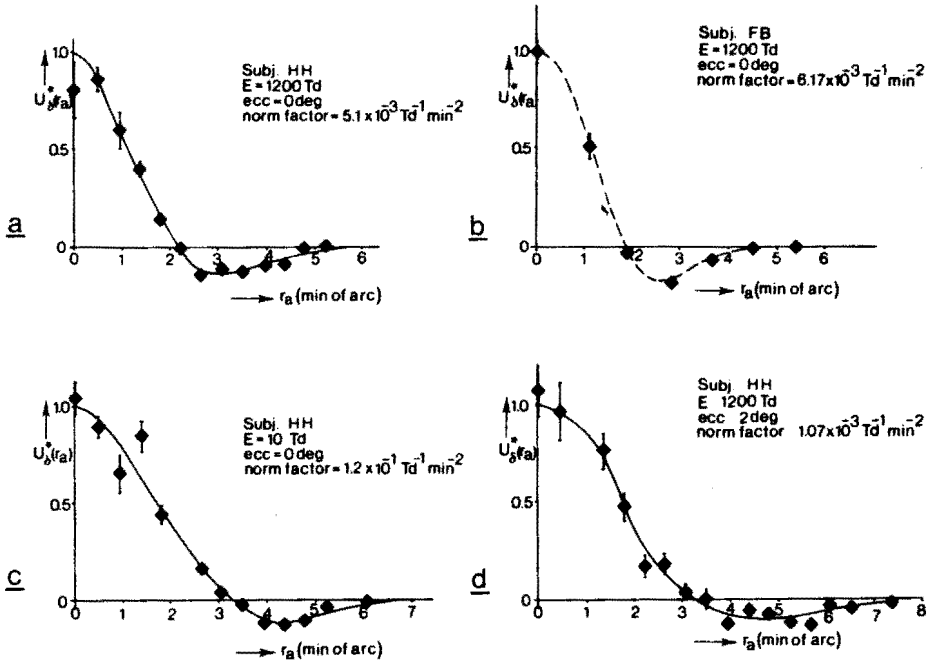
Figure 3: Experimental data of normalised point spread functions $U_\delta^*(r_a)$ measured for two subjects and a few experimental conditions as indicated in the legends. The drawn curves are polynomial fits, whereas the dashed line of 3b is the result of a simultaneous computer fit which exactly obeys linearity of processing (for details see Blommaert and Roufs, 1981).

In Fig. 3, data on four normalised point spread functions are shown. Fig. 3a shows a PSF for subject HH measured foveally at an adaptation level of 1200 Td. The one in Fig. 3b (taken from Blommaert and Roufs, 1981) was measured for subject FB under identical conditions. Fig. 3c shows a PSF for subject HH, measured foveally again, but here at the mesopic level of 10 Td. Finally, Fig. 3d gives results of a point spread function measurement at our standard photopic level of 1200 Td but now at an eccentricity of 2 deg. The vertical lines are twice the standard deviations of the mean.

The absolute values of the point spread functions can be found by multiplying the reduced values by the norm factors given in the legends. These norm factors were obtained by averaging the thresholds $\varepsilon_p(0)$ over all sessions of each experiment. The

data points are averaged over 8 threshold quotients according to eq. 10. The experimentally determined standard deviations of the mean are indicated. The order of measurement with respect to the r-axis was randomly chosen.

5.5 Intermediate discussion

Evaluation of point spread functions

First of all, the basic hypothesis have to be considered, namely radial symmetry, linearity of processing and peak detection. As to radial symmetry, no signs have yet been found that this assumption is violated at threshold level in the limited area of the retina considered. That is, no perceptual clues have been noticed, nor threshold variations, that might contradict the radial symmetry assumption of spatial interaction for the local stimuli used.

As to linearity of processing, a test on linearity was reported in an earlier paper (Blommaert and Roufs, 1981) which confirmed linearity. Further support for the hypothesis comes from data on line and edge spread functions (Kulikowski and King-Smith, 1973; Roufs and Polstra, 1982). Furthermore, Roufs and Polstra (1982) were able to determine the response on a subliminal bar pattern, the result of which was in quantitative agreement with the calculated response on the basis of a measured line spread function. It seems that responses at threshold are small enough for linearity of processing to be an accurate approximation.

With respect to peak detection, in all experiments where linearity was tested, peak detection was implicitly tested too. Since every test turned out to be positive, there is no reason to abandon this detection formalism.

The results of Fig. 3 show that point spread functions can be obtained rather accurately using the perturbation method. From Figs. 3a and 3b it can be concluded that different subjects yield comparable results: a scale factor of about 1.2 has to be expected as a between subject variation.

Comparing the point spread function of Fig. 3a with the one of Fig. 3c, it is striking that, although the overall shape is broadened, the PSF at 10 Td has the same shape as the one at 1200 Td. This experimental fact has an important implication. The optical point spread is namely independent of the background level since an artificial pupil is used. This implies that the resulting point spread function must be dominated by neural effects.

Isomorphy also holds for the PSF of Fig. 3d, measured at an eccentricity of 2 deg. To facilitate comparison of shapes, the results of Fig. 3 are replotted in Fig. 4 in such a way that every radius axis has been scaled to fit the PSF of Fig. 3a. This means that the axis of Fig. 3b has been stretched out by a factor of 1.2 and the axis of 3c and 3d have been shrunk by a factor of 1.5. From Fig. 4 it can be concluded that to a fair approximation all determined point spread functions are isomorphic.

The variation in width may be compared with visual acuity measured in the parafovea (see for instance Sloan, 1968). Since the size of a point spread function is expected to be inversely proportional to visual acuity, the broadening of experimentally determined
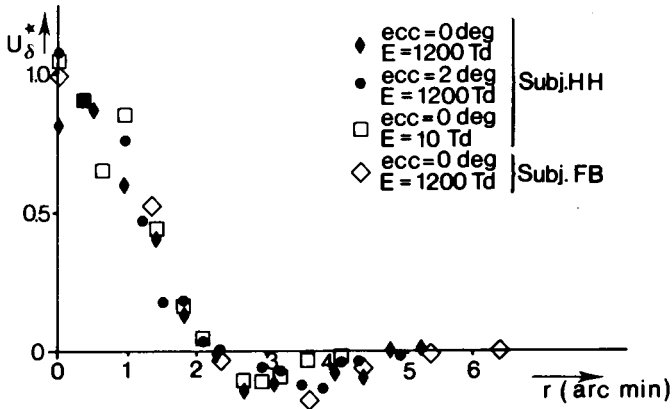
Figure 4: Experimental data of normalised point spread functions, reduced with respect to the zero crossing of the PSF of Fig. 3a. For details see text.

PSFs, at a mesopic level as well as at the eccentricity of 2 deg, is in quantitative agreement with expectations on the basis of visual acuity data.

Although we are aware that it is dangerous to take physiological properties to explain psychophysical data, a comparison of the properties of PSFs and those of receptive fields may be relevant. Physiological data on receptive field sizes of ganglion cells as a function of eccentricity (Wiesel, 1960; Fischer, 1973) show the same tendency in sizes, that is, receptive field sizes increase roughly in proportion to distance from the fovea.

With respect to the effect of an adapting background on the width of the receptive fields, two a priori possibilities should be considered. First, one could argue that the width of the receptive field changes with background. To our knowledge, however, there is no clear physiological evidence of comparable changes in the width of receptive fields of ganglion cells while changing the adaptation level. This would mean that the change in point spread function measured foveally at 1200 Td and 10 Td could not be explained by a change in the size of the same ganglion cell receptive fields. A possible explanation could then be that receptive fields of different sizes are present in the fovea; a collection of units that vary in sensitivity or excitability if the adaptation level is changed. Then, at high adaptation levels narrower units dominate, while at the lower levels wider units do so. Since visual acuity is a continuous monotonic function of the adaptation level, it is suggested that the scatter in receptive field sizes is not discrete but more or less continuous. This is compatable with physiological data on scatter in receptive field sizes (Fischer, 1973). Measuring a point spread function at 1200 Td would then mean that the threshold is determined by a subset of the ensemble of units, which on average is 1.5 times narrower than the average subset at 10 Td.

The point spread function determined at the eccentricity of 2 deg suggests that the

narrowest receptive fields of the fovea no longer exist at this retinal position. This variation of receptive field sizes is indeed also known from the literature (Hubel and Wiesel, 1974) It should be kept in mind, however, that properties of receptive fields need not be the only relevant factors for perception.

"Point spread function" type results have also been obtained by among others Fiorentini and Maffei (1970) and Detlef (1973). Both the latter papers reported experimental results on the threshold variation of a point-shaped stimulus due to the presence of subthreshold discs and/or annuli. Qualitatively they show the same results as in this paper but for a scale factor, which is not surprising in view of the foregoing. Other related work is reported by Spillmann (1971), who used Hermann grids in an attempt to determine the size of so-called "perceptive fields", receptive-field-like functions which may account for the Hermann grid illusion. The perceptive field sizes they report are much larger (zero crossing at 7 to 15 min of arc) than the PSFs reported here, suggesting that the smaller ensemble units do not account for the Hermann grid illusion.

Variation in size is also reported by several investigators who measured line spread functions of the human visual system. Hines (1976) for instance determined LSFs at different retinal positions, which show the same tendency as found here, i.e. LSFs broaden with increasing distance from the fovea. Roufs and Polstra (1982) reported on experimentally determined line spread functions at different adaptation levels. They also found that line spread functions remained isomorphic, yet broadened if the adaptation level was lowered.

Another interesting question is to what extent optical processes take part in the measured point spread functions. Since our experiments were carried out using a 2 *mm* artificial pupil, the part of optical deformation can be estimated using the optical point spread function given by Vos et al. (1976), which is caused by diffraction and stray effects in the optical media. It can easily be verified that the measured PSF equals the convolution of the optical PSF and its neural counterpart. In formula:

$$PSF(r) = PSF_{neural}(r) \otimes PSF_{optical}(r), \qquad (12)$$

where $\otimes$ denotes convolution.

Since from the above equation, the measured and optical PSF are known as a function of the radius r, the neural component of the point spread function can be calculated. This was carried out by dividing the Hankel transform of the measured PSF by the Hankel transform of the optical PSF. The resulting frequency spectrum was then transformed back again to the space domain, thus yielding the neural point spread function. A result of such a computation is shown in Fig. 5 for the point spread function of Fig. 3a. (In doing so, it is implicitly assumed that the optical point spread of subject HH's left eye equals that of an average observer.)

The share of optical effects like scatter and diffraction can be seen to be rather limited. This is certainly the case for the point spread functions of Figs. 3c and 3d due to the fact that these are broader than the one of 3a.

At this point, one may wonder what exactly is represented by a point spread function. In our view, there is no doubt that spatial processing as depicted by the centre-
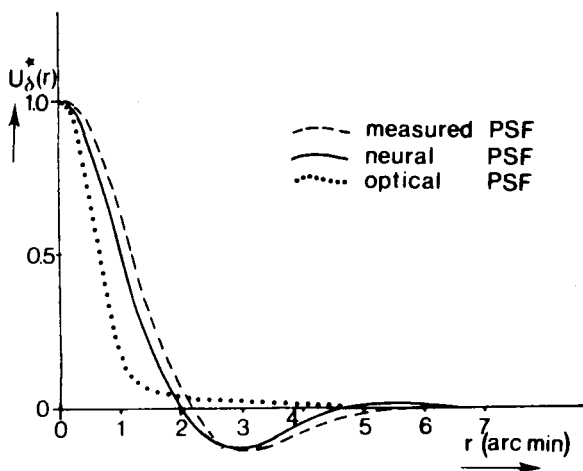
Figure 5: Neural point spread function as calculated from the experimentally determined one of Fig. 3a and the known optical spread for 2 *mm* pupils as given by Vos et al. (1976).

surround structure of point spread functions essentially occurs in the human visual system. Furthermore, we think that, on the basis of psychophysical and physiological as well as morphological investigations, visual mapping can be described by the action of an ensemble of units which differ in size and sensitivity (see for instance Koenderink and van Doorn, 1978). Depending on the type of stimulation, a different subset of the ensemble becomes more manifest then others.

In this notion, a point spread function is a sort of average response of an ensemble subset. For a certain stimulus, those units respond most that "fit in" best with the stimulus structure. For instance, if the response to a point source is measured, the smallest units will determine the response. If on the other hand the response to a disc of larger dimension is measured, the response will be governed by units that are broader. In the following, predictions of such a model will be confronted with experimental findings.

5.6 Estimating the parameters of a simplified ensemble model

With the aim of specifying a model on detail detection and/or detail vision, several starting-points may be chosen. One possibility might be that it is constructed to fit in with detailed physiological findings. Another point of view would be that psychophysical findings on detail vision should be well fitted with as few parameters as possible. Our aim was somewhere in between, in that the model should have the most important features of physiological findings on receptive field properties, yet be economical to handle.

As a first order approximation we assume that scatter in receptive field sizes may be represented by the action of only four units differing in width, the narrowest one being equal to the experimentally determined point spread function of Fig. 3a. The other

three units were chosen to be progressively wider, each time by a factor of two (the four units were chosen to link up with the Wilson and Bergen model (1979) on line spread functions). Spatial variance is incorporated in such a way that all units vary linearly in width as a function of retinal eccentricity. The scale factor comes from the experimentally determined PSF at 2 deg eccentricity (Fig. 3d), i.e. width increases linearly from 1 to 1.5 over a range of 2 deg.
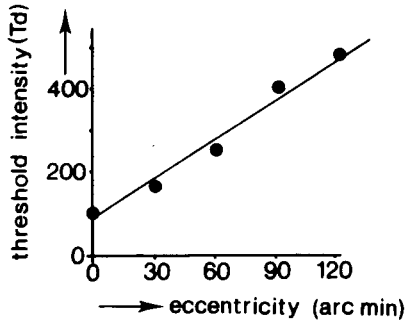
Figure 6: Thresholds variation as a function of retinal eccentricity for a point source with a diameter $\phi = 1.4$ arc min. The regression line is shown as well.

Spatial variance is also expressed in variation of sensitivity as a function of eccentricity. This property is reflected in a variation in response magnitude of the units, the rate of which was estimated for subject HH in an experiment where point thresholds were measured as a function of retinal position.

The averages of two threshold determinations for a 1.4 min of arc stimulus are plotted in Fig. 6. It shows that, within a limited spatial range, threshold elevation as a function of eccentricity can be approximated fairly well by a straight line. The sensitivity factors of the four units were therefore taken to be inversely proportional with distance to the fovea, according to the straight line of Fig. 6. This is consistent with earlier findings of Blommaert and Roufs (1981).

For the prediction of thresholds we have to adopt a formalism for the detection of stimuli, especially if larger areas are involved. Since stochastic variables are involved in the detection process, stimuli which excite larger response areas may have a considerable threshold decrease caused by an increased chance of detection. This effect will be demonstrated by the following experiment. The stimulus consisted of a number of point sources with diameter $\phi = 1.2$ arc min, situated on a circle with a diameter of 52 arc min. The fixation point was situated in the centre of the circle. The number of points was varied from 1 to 16. Since all point sources were identical and situated at the same distance from the fovea, threshold variation as a function of number had to be the result of chance effects if small deviations from isotrophy are neglected. Spatial summation effects were also neglected since the distance between the point sources (≈ 10 arc min) was chosen such that their responses could only interact
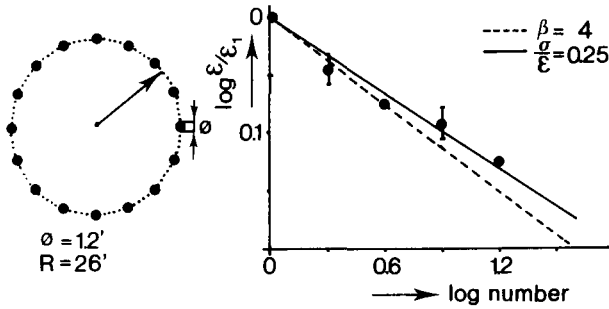
Figure 7: Left: stimulus configuration used for the determination of threshold variation as a function of the number of point sources constituting the stimulus. Right: experimental results. The lines are different theoretical predictions. For details see text.

through very wide units, which are highly insensitive to points. The experimental results as shown in Fig. 7 indicate that the logarithm of the threshold decreases about linearly with the logarithm of the number of point sources used as the stimulus. The straight lines through the data points are two separate predictions. The continuous line is calculated assuming detection of the stimulus if at least one of the points is detected (Roufs, 1974). The stochastic parameter, the Crozier quotient $\sigma/\varepsilon = 0.25$, is found from the slopes of the psychometric function for subject HH. The threshold lowering according to this calculation may be seen to be in fair agreement with the experimental results. The dashed line follows a prediction according to the non-linear summation given by Quick's (1974) rule, choosing the exponent to be 4, as Wilson and Bergen (1979) did:

$$U = \{ \sum_{visual field} U_i^4(\vec{r})\}^{\frac{1}{4}}, \tag{13}$$

U_i being the reponse to one dot.

On the basis of a Weibull (1951) distribution, it can be shown that the exponent β, as predicted by the slope of the psychometric function or the Crozier quotient $\sigma/\varepsilon = 0.25$ of the isolated dots, equals $\beta = \frac{1.15}{\sigma/\varepsilon} = 4.6$. The dashed line is a reasonable first-order approximation, and $\beta = 4$ will be used further on in order to link up with literature.

At this point, there are still three free parameters left which have to be estimated, namely the ratios of the sensitivity factors of the three wider units compared with the points spread function sensitivity. We decided to estimate these by means of an experiment in which thresholds for discs with variable diameters were determined. The results are shown in Fig. 8. The experimental inaccuracy is indicated by the thin vertical lines, which denote the range between the two thresholds constituting the data points.

The four dashed lines are the theoretical threshold-area curves of the four assumed individual units. The narrowest one, labelled PSF, is seen to fit the data in the Ricco area. The other curves have been shifted vertically by eye in order to fit the data.
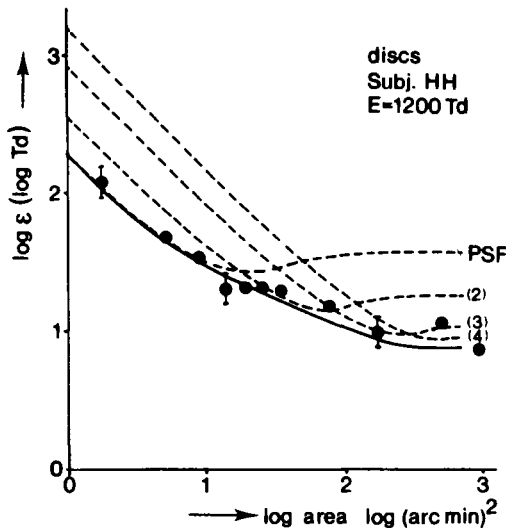
Figure 8: Experimentally determined thresholds for discs as a function of area. The dashed curves are computations for the four individual units, labelled PSF, 2, 3 and 4. They were fitted by eye to the experimental data within appropriate area-intervals. The continuous curve is the overall fit containing non-linear summation of the ensemble units.

It can be seen that every unit covers the threshold prediction in a certain range of disc diameters. The continuous line follows if the Wilson and Bergen interpretation of Quick's rule is applied to the response of the whole ensemble.

5.7 Model evaluation and discussion

For the purpose of testing the detection model, thresholds for annuli, thin lines and broad lines were measured, the results of which are shown in Fig. 9. In Fig. 9a thresholds of thin annuli are plotted against area (the radius r was varied while the width was kept constant at $\Delta r_a =0.23$ arc min). It can be seen that the threshold is almost independent of area. Fig. 9b gives the determined thresholds of thin lines (width 1.75 arc min), and in Fig. 9c thresholds for broader line (12.25 arc min) are shown. The dashed lines in these figures are the predicted threshold versus area curves of the four individual units and the bold lines are the resulting predictions of the ensemble model.

Bearing in mind that after the fitting of units 2, 3 and 4, there are no free parameters left in the model, threshold predictions are seen to be in satisfactory agreement with experimental results. Variations as shown here have to be expected owing to sensitivity variations between sessions and between days.

From the model predictions some interesting features can be extracted. In the first place, it can be observed from Figs. 9a and 9b, that for annuli and thin lines the
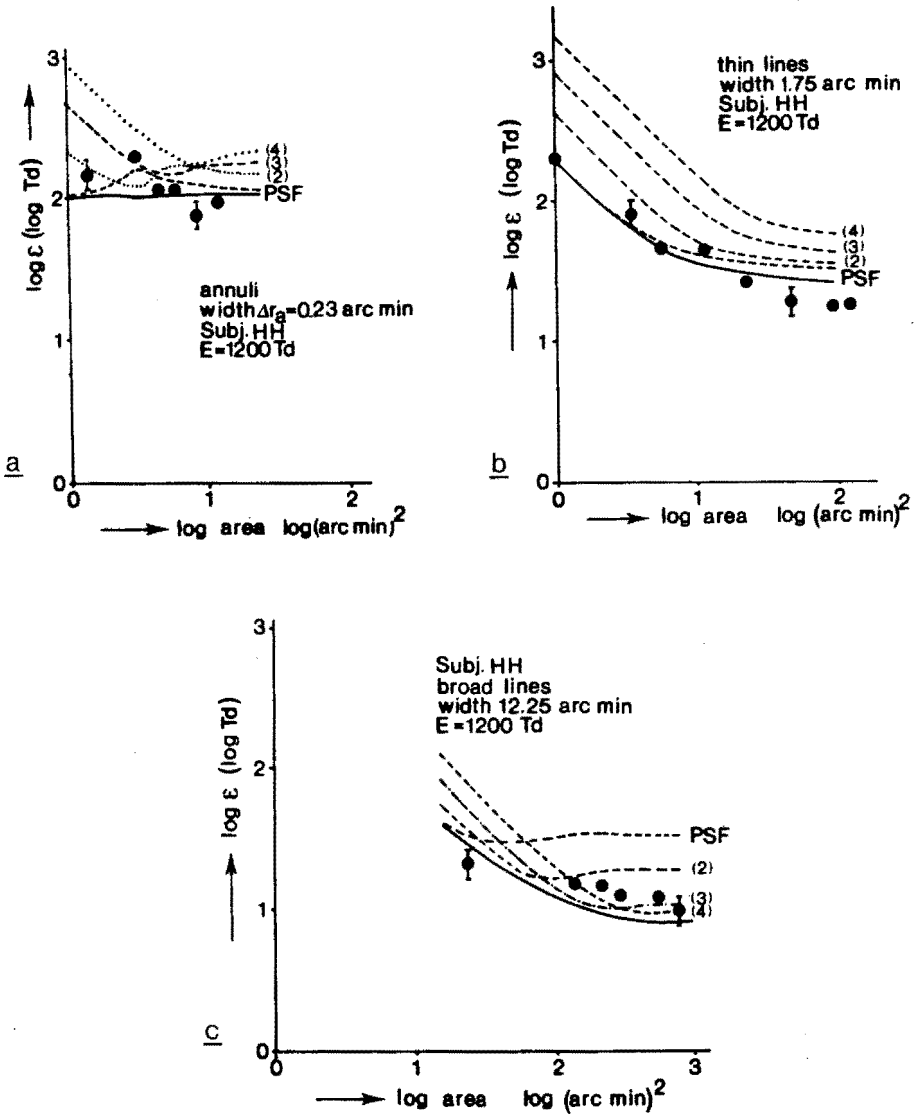
Figure 9: Experimentally determined thresholds for annuli (9a), thin lines (9b) and broad lines (9c) as a function of stimulus area. The data points are averages of two thresholds. The average distance between these thresholds is indicated by the length of the thin vertical lines. Characteristic stimulus dimensions are given in the legends. The broken curves are predictions for the four individual ensemble units. The heavy lines are quantitative overall predictions of the ensemble model.

model prediction is almost entirely determined by the experimentally determined point spread function. For the more chubby stimuli like discs and broad lines, prediction of thresholds is gradually taken over by the broader units if the area of the stimulus increases.

This observation is an interesting one since it means that if a stimulus is extended in only one dimension, or put otherwise: if a stimulus is built up of thin lines, threshold prediction is governed mainly by the experimentally determined point spread function.
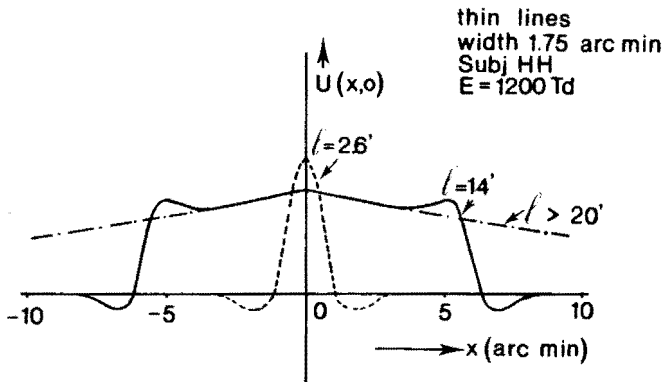
Figure 10: Cross-sections of response profiles for thin lines in the direction of the lines, calculated for different line lengths from the point spread function. Line width is 1.75 arc min. Retinal inhomogeneity is incorporated.

Support for this notion comes from perceptual phenomena observed when viewing these stimuli at threshold level. Annuli, when viewed at threshold level, just look like clear rings if their radii are larger than about 3 arc min; if the radius is smaller than 1.5 arc min they look like points. This type of stimulus behaviour follows exactly the shape of calculated responses on the basis of just the PSF. For thin lines it is even more convincing as is demonstrated in Fig. 10.

Here, calculations are shown for three thin lines with lengths of respectively $l = 2.6'$, $l = 14'$, $l > 20'$. Profiles are shown of cross sections through calculated responses based on the PSF. For short lengths (up to about 3 arc min) lines will be seen as single points. For the line length of about 14 arc min, the typical percept is that only the middle and both endings of the line segment are seen. For the longer lines, only the middle is detected, which stresses the effect of retinal inhomogeneity. The calculated profiles exactly depict these perceptual phenomena.

On the basis of these observations it might be concluded that the function of small units of an ensemble is the visual mapping of details, whereas broader units are responsible for the mapping of coarser features. It should be noted that the class of thin stimuli contains a very important one, namely alphanumeric symbols (for instance type letters at a normal viewing distance).

This can be appreciated if one interprets the stimuli of Fig. 9a as a type letter "o" (or number "zero") with varying viewing distance. Likewise the stimuli of Fig. 9b may be interpreted as the letter "l" (without shrivels) at different viewing distances (it can also be interpreted as a "slash" symbol or the number "one").
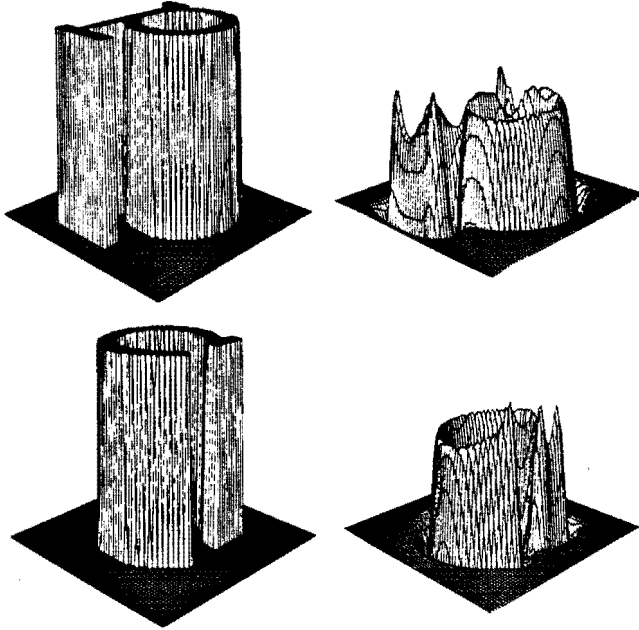
Figure 11: Three-dimensional plots of, left: luminance profiles of the lower case letters "p" and "c" of type font Courier 10. At the right: calculated visual responses on the basis of a measured point spread function if the letters are viewed foveally, at an adaptation level of 1200 Td, from a distance of 0.35 m.

In this interpretation the experimentally determined point spread function, which can be regarded as the basic function, becomes a powerful tool for determining visual processing of alphanumeric symbols. Fig. 11 shows by way of example the response of a few letters calculated by convoluting their luminance profiles with the foveal point spread function.

A stimulus set comparable to ours was used by Lamar et al. as far back as 1947. They measured thresholds for discs and line segments of variable length and width in order to formulate empirical rules that describe the behaviour of detail thresholds as a function of area. The relationships they find between area and threshold are consistent with the threshold courses depicted by the experimental results of Figs. 8 and 9.

The number of channels we used is based on the Wilson and Bergen (1979) four mechanisms, which originally were designed to predict thresholds on the basis of line spread functions. Although line spread functions are of considerable value for

determining one-dimensional coarse features of spatial processing, their ability to deal with two-dimensional detail patterns is limited. Furthermore, derivation of point spread functions from line spread functions or MTFs is not obvious since the retina is inhomogeneous.

As was signalled by Kelly and Burbeck (1984), the experimentally determined PSF is narrow in comparison with calculated ones from line spread functions and modulation transfer functions. This might be explained by increased action of, on average, wider ensemble units in the experimental determination of line spread functions and modulation transfer functions.

In conclusion, we think that the concept of an ensemble model is a reasonable approximation, be it that a lot of effort has yet to be done to specify such a model in full detail, especially when time is incorporated.

8 Conclusions

- Point spread functions determined under different experimental conditions (two subjects, adaptation levels and retinal positions), are found to be isomorphic.

- The observed variations in width factors are in agreement with visual acuity data.

- Optical spread is found to be only a minor component of the determined point spread functions.

- A model based on point spread functions, and assuming linearity, is able to describe visual processing of "slim" stimuli like alphanumeric symbols.

- In order to predict thresholds for more "chubby" stimuli like squares and discs, it suffices to extend the model to the simplified multiple unit one.

References

Barbur, J.L.; Ruddock, K.H. (1980) Spatial characteristics of movement detection mechanism. Human vision -I Achromatic Vision. Biol.Cyb. 37, 77-92.

Blackwell, H.R. (1946) Contrast thresholds of the human eye. J. Opt. Soc. Am. 36, 624-643.

Blommaert, F.J.J.; Roufs, J.A.J. (1981) The foveal point spread function as a determinant for detail vision. Vis. Res, 21, 1223-1233.

Campbell, F.W.; Robson, J.G. (1968) Application of Fourier analysis to the visibility of gratings. J. Physiol. 197, 551-566.

Detlef, P. (1973) Analyse Flachenhafter biologischer Empfangssysteme mit Hilfe von Detektionsexperimenten. Ph. D Thesis, Karlsruhe, West Germany.

Fiorentini, A.; Maffei, L. (1970) Transfer characteristics of excitation and inhibition in the human visual system. J. Neurophysiol. 33, 285-292.

Fischer, B. (1973) Overlap of receptive field centers and representation of the visual field in the cat's optic tract. Vis. Res. 13, 2113-2120.

Hines, M. (1976) Line spread function variation near the fovea. Vis. Res. 16, 2113-2120.

Hubel, D.H.; Wiesel, T.N. (1974) Uniformity of monkey striate cortex: a parallel relationship between field size, scatter and magnification factor. J. Comp. Neur. 158, 295-306.

Jones, R.C. (1958) On the point and line spread functions of photographic images. J. Opt. Soc. Am. 48, 934-937.

Kelly, D.H. (1975) Spatial frequency selectivity in the retina. Vis. Res. 15, 665-672.

Kelly, D.H.; Burbeck, C.A. (1984) Critical problems in spatial vision. Crit. Rev. Biom. Eng. 10, 125-177.

Koenderink, J.J.; van Doorn, A.J. (1978) Visual detection of spatial contrast; influence of location in the visual field, target extent and illumination level. Biol. Cyb. 30, 157-167.

Kulikowski, J.J.; King-Smith, P.E. (1973) Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation. Vis. Res. 13, 1455-1478.

Lamar, E.S.; Hecht, S.; Schlaer, S.; Hendley, C.D. (1947) Size, shape and contrast in detection of targets by daylight vision I. J. Opt. Soc. Am. 37, 531-545.

Meeteren, A. van (1973) Visual aspects of image intensification. Thesis. Utrecht.

Quick, R.F. (1974) A vector-magnitude model of contrast detection. Kyb. 16, 65-67.

Roufs, J.A.J. (1963) Perception lag as a function of stimulus luminance. Vis. Res. 3, 81-91.

Roufs, J.A.J. (1974) Dynamic properties of vision- VI. Stochastic threshold fluctuations and their effect on flash-to-flicker sensitivity ratio. Vis. Res. 14, 871-888.

Roufs, J.A.J.; Blommaert, F.J.J. (1981) Temporal impulse and step responses of the human eye obtained psychophysically by means of a drift-correcting perturbation technique. Vis. Res. 21, 1203-1221.

Roufs, J.A.J.; Polstra, J. (1982) Line and edge-spread funtions of the visual system elicited by a TV display in situ. IPO Ann. Progr. Rpt. 17, 91-100.

Sloan, L.L. (1968) The photopic acuity-luminance function with special reference to parafoveal vision. Vis. Res. 8, 901-911.

Spilmann, L. (1971) Foveal perceptive fields in the human visual system measured with simultaneous contrast in grids and bars. Pflug. Arch. 326, 281-299.

Vos, J.J.; Walraven, J.; Meeteren, A. van (1976) Light profiles of the foveal image of a point source. Vis. Res. 16, 215-219.

Weibull, W. (1951) A statistical distribution function of wide applicability. J. Appl. Mech. 18, 293-297.

Wiesel, T.N. (1960) Receptive fields of ganglion cells in the cat's retina. J. Physiol. 153, 583-594.

Wilson, H.R.; Bergen, J.R. (1979) A four mechanism model for threshold spatial vision. Vis. Res. 19, 19-32.

chapter 6

Local visual responses in space and time[1]

Frans J.J. Blommaert

6.1 Introduction

Perturbation of the response of a probe stimulus by suitable space or time functions has been used as a technique for measuring impulse responses (Roufs and Blommaert, 1981; Blommaert and Roufs, forthcoming) and point spread functions (Blommaert and Roufs, 1981; Blommaert, Heijnen and Roufs, forthcoming) of the human visual system. The perturbation approach applied here in the space-time domain does not essentially differ from those used in the space and time domains separately. The ultimate goal of this approach is to find some basic functions from which visual responses on arbitrary stimuli can be quantitatively predicted at threshold level. In the space-time domain this could be a Green's function which, by definition, is the response of a linear system to a unit impulse in space and time. In this paper, experimental results are presented that can be interpreted as one-dimensional intersections of a two-dimensional Green's function for foveal vision.

6.2 Theoretical frame

Our starting point consists of a number of assumptions about the nature of spatio-temporal processing which have been tested and confirmed in the individual space and time domains:
- quasi-linearity of processing both in time and space,
- peak detection as a formalism to describe threshold determination,
- homogeneity of the fovea within the extent of lateral processing,
- radial symmetry.

If these assumptions hold good, a local Green's function $U_\delta(r,t)$ of the fovea can be defined as the response to a unit impulse in space and time. Note that this function is radially symmetrical by definition. (For details on Green's functions see Courant and Hilbert, 1961). The response of the visual system to an arbitrary stimulus may, on this basis, be written as the convolution of this stimulus with the Green's function:

$$U(\vec{r},t) = \int_0^{2\pi} \int_0^\infty \int_0^\infty U_\delta(|\vec{r}-\vec{r}'|, t-t')\varepsilon(\vec{r}',t')r'd\varphi'dr'dt'. \tag{1}$$

Here $\varepsilon(\vec{r},t)$ is the distribution of retinal illumination of the stimulus.
In Fig. 1 the stimulus time configurations and time functions are shown that make it possible, in principle, to measure a Green's function of the eye.

---

[1] This chapter is the text from an article with the same title published in 1979 in: IPO Ann. Progr. Rpt. <u>14</u>, 88-94.
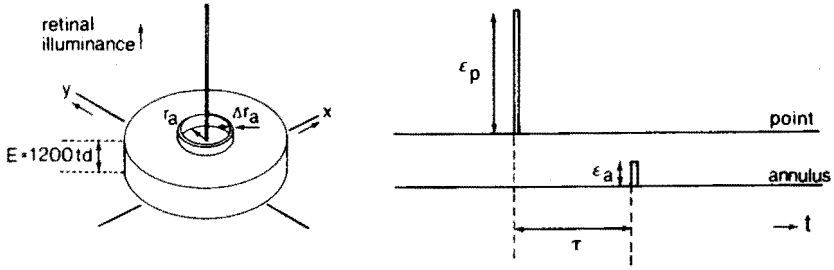
Figure 1: Stimulus configuration and time functions used in the perturbation experiments. Duration of flashes is 10 ms.

The 1200 Td background level is assumed only to play a role as a parameter-setting in the description of visual processing at threshold level. It will be assumed below that the point-radius and the annulus-width are small compared to spatial resolution, and the time-functions short compared to the temporal resolution of the visual system.
In this case, and assuming that the point is situated in the origin, the response of the visual system to the stimulus combination point + annulus may be approximated by:

$$U_c(r,t) = \varepsilon_p A_p U_\delta(r,t) + \varepsilon_a \frac{A_a}{2\pi} \int_0^{2\pi} U_\delta(|\vec{r} - \vec{R}|, t - \tau) d\varphi', \qquad (2)$$

$$|\vec{r} - \vec{R}|^2 = r^2 + R^2 - 2rR\cos\varphi'.$$

where $U_c(r,t)$ is the response to the combination. $\varepsilon_p, \varepsilon_a$ = retinal illumination of point and annulus, respectively. A_p, A_a = area of point and annulus , respectively.
In the actual experiments, the luminance ratio $q = \varepsilon_a/\varepsilon_p$ was kept constant and chosen so that the annulus was always subthreshold. Thus we established that detection was governed by the extreme value of the point response. Furthermore, we can define without loss of generalisation the point response as situated in the origin. Considering this, the threshold condition simplifies to:

$$U_c(0, t_{ex}) = \varepsilon_p A_p U_\delta(0, t_{ex}) + \varepsilon_a A_a U_\delta(R, t_{ex} - \tau) = D. \qquad (3)$$

Here, t_{ex} is the time at which $U_\delta(0,t)$ reaches the extreme value D. In order to minimise the effect of sensitivity changes (see Roufs and Blommaert, 1981), the threshold of the combination was always compared to ε_0, the threshold of the point alone, for which we can derive:

$$\varepsilon_0 A_p U_\delta(0, t_{ex}) = D. \qquad (4)$$

The threshold pair $[\varepsilon_p, \varepsilon_0]$ will be denoted in the following by 'fast pair'.
From equations 3 and 4 one value for the normalised Green's function can be found to be:

$$U_\delta^*(R, t_{ex} - \tau) = \frac{U_\delta(R, t_{ex} - \tau)}{U_\delta(0, t_{ex})} = \frac{A_p}{qA_a}\left\{\frac{\varepsilon_0}{\varepsilon_p} - 1\right\}. \qquad (5)$$

By varying the radius R of the perturbating annulus and the stimulus-onset-asynchrony τ, it is possible in principle to get a total mapping of the Green's function in the space-time domain. However, this procedure would be very time-consuming, for which reason we started by measuring temporal responses for only four different values of the radius R of the annulus.

6.3 Methods

The subject viewed an 11 deg homogeneous background field with a retinal illuminance of 1200 Td monocularly. Both point-stimulus and perturbating annulus were projected on top of this field by using prisms. To facilitate fixation, four weak fixation lights with a radius of 2 min of arc were projected around the stimulus on a circle with a diameter of 1 deg. A 2 mm artificial pupil was used, which was provided with an entoptic guiding system which enabled the subject to check the centering of the pupil of the eye (Roufs, 1963). The lights were generated by linearised glow modulators (for details see Roufs, 1972).
All experimental data consisted of 50% luminance thresholds measured by the method of constant stimuli. Every threshold determination took 3 to 4 luminance levels, each of which was presented 20 times to the subject. The subject had one knob which he used to release the stimulus which was delayed for a convenient preset time interval. Three other knobs enabled him to answer with "yes", "no" or "reject". The subject had normal acuity (1.5) but was moderately deuteranomalous.

6.4 Results

In Fig. 2, four temporal responses are shown, measured according to eq. 5, with annuli having different radii as indicated in the legends. The experimental points of Fig. 2a and 2b are arithmatic means of eight pair-thresholds, while in the figures 2c and 2d, four pair-thresholds are used at every τ-value. Standard deviations of the means are indicated as vertical bars through the data points. During a session, one fast pair was measured at each τ-value, in random order with respect to τ. During the next session, the order of measurement was reversed. Also the order of threshold determination within the fast pairs was reversed in consecutive sessions. By fully counterbalancing in this way, we tried to minimise the effect on the mean experimental results of systematic sensitivity changes in the subject (for further details on technique see Roufs and Blommaert, 1981).

6.5 Discussion

Fig. 2a shows the limiting case in the space domain where the perturbating shape is identical with that of the point source. If the basic assumptions concerning linearity and peak detection hold, this result can be interpreted as a local impulse response of the visual system. For this special case, these assumptions were tested by also measuring a step response for a point source, which then has to be the primitive
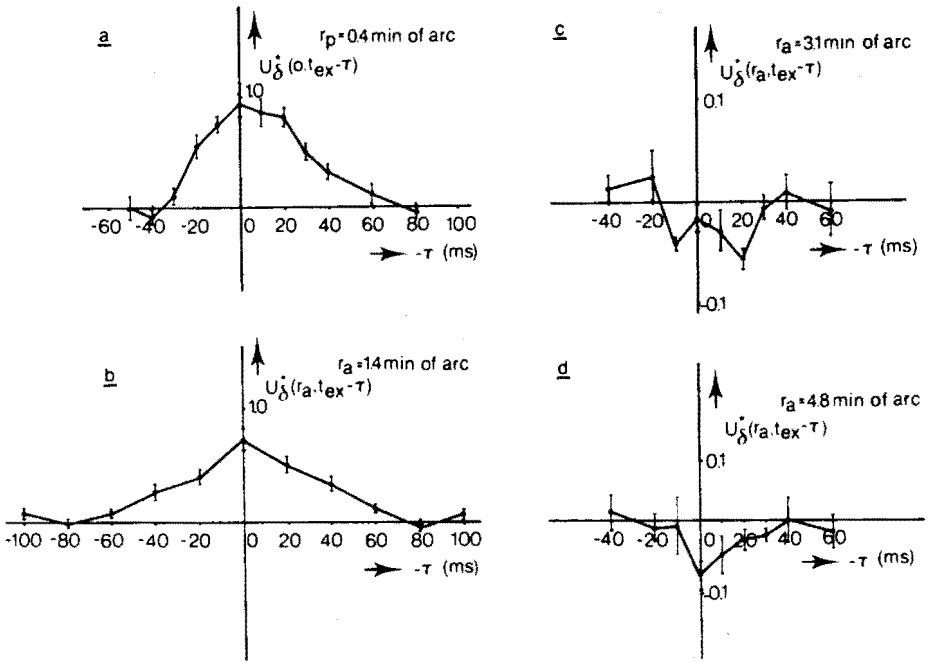
Figure 2: Temporal response functions, measured by perturbation, of annuli at their centers. Four different radii are used, as indicated in the legends. 2a is the limiting case where the perturbation shape is identical to the point source itself.

function of the impulse response. The experimental results confirmed linearity and peak detection in combination (for details see Roufs and Blommaert, 1981). The response of Fig. 2b has about the same temporal shape and extension as the impulse response of Fig. 2a, while the amplitude is a factor of 0.7 smaller.

Although the measured effects of the responses as shown in Figs. 2c and 2d hardly go beyond experimental error, it can be concluded that the resulting responses essentially have a negative sign. Because of the poor signal-to-noise ratio, shape and extension of the responses cannot be judged very accurately. A first-order approximation, however, could be that the temporal shape of all four responses measured is the same. In this case, the problem of finding a local Green's function is greatly simplified because separation of variables in space and time is then possible. Under this condition the Green's function may be written as the product of the temporal impulse response function and the spatial point spread function.

In order to work out this idea, we compared the approximate response values for $\tau = 0$ with an experimentally determined quasi-static point spread function (Blommaert and Roufs, 1981) as is shown in Fig. 3. It can be seen that the shape of the transient curve looks somewhat more extended than the quasi-static one.

This finding is confirmed by among others Wilson and Bergen (1979), who measured
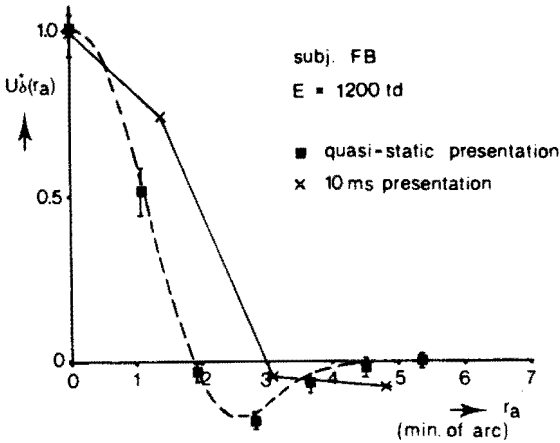
Figure 3: Point spread functions for quasi-static and transient presentation, respectively. Dashed curve is a simultaneous fitting to point- and edge spread function in the static case.

line spread functions for 'transient' and 'sustained' time functions. They also report that, in case of transient presentation of stimuli, the measured line spread function is broader than for sustained presentation.

Other circumstantial support for more extended lateral processing when stimuli are presented briefly comes from an experiment where thresholds for discs are measured as a function of the radius. In Fig. 4, such thresholds for transient presentation (10 ms flashes) are compared with thresholds for quasi-static presentation (500 ms rectangular pulses with smoothed beginning and end). All data are means of two thresholds that were measured in counterbalanced order. It can be seen that spatial integration of the visual system takes place over a more extended area in the case of transient presentation, which can be explained by the presence and activity of broader point spread functions.

In conclusion, we think that in search for a local Green's function, separation of variables in a first order approximation is a promising approach. The validity of this approximation, however, has to be carefully tested, together with the provisional assumptions as formulated in the theoretical framework.

Note

The stimulus configuration used here (shown in Fig. 1) is rather similar to the one described by van der Wildt and Vrolijk (1981)[1] and Vrolijk and van der Wildt (1985)[2] in their study on "jumping flashes". An essential difference, however, is that the luminances of the point and annulus (in their case an ellipse instead of an annulus) were chosen such that

---

[1]Vis. Res. 21, 1765-1771.
[2]Vis. Res. 25, 1413-1421.
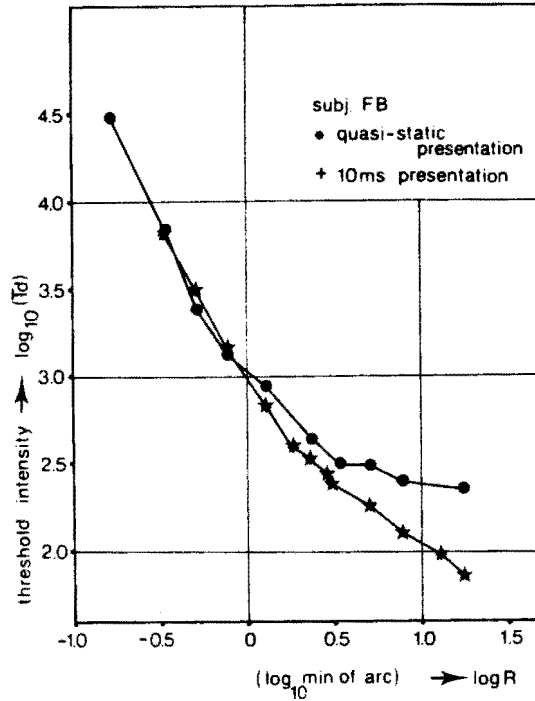
Figure 4: Comparison of disc thresholds as a function of radius, for quasi-static and transient presentations respectively. Data for quasi-static presentation are vertically fitted in Ricco area $(log R < 0)$. In the case of transient presentation there is substantially more integration, reflected by the lower threshold for increasing disc size.

both stimuli had about the same detection probability. As pointed out before, we took care that the annulus was always well below threshold, which is an essential condition for the perturbation technique. From the experimental results found for "jumping flashes", it might be expected that an inhibitive effect would be found from the subthreshold annulus on the threshold of the point source. Its maximum effect would occur at a r-value differing from zero and depending on the annulus diameter, since a constant velocity was found for propagating inhibition. Owing to the small annulus diameters used here and to the poor signal-to-noise ratio, no delay can be observed in our Figs. 2c and 2d, although the inhibititive effect itself is present in the latter results.

Summary

The experimental results presented here have to be seen in the broader context of finding a model of the visual system which is able to describe local processing in space and time at threshold level. With this aim in mind, time functions were mea-

sured by means of a perturbation technique which can be interpreted as time-response functions in a coordinate (x, y) of the visual system if a unit impulse is presented in another coordinate (x', y'). This interpretation is only allowed if the following assumptions are satisfied: quasi-linearity, peak detection homogeneity and radial symmetry of processing. The extent of validity of these provisional assumptions, that were already confirmed for the space and time domains separately, has still to be investigated.

References

Blommaert, F.J.J. (1977) Spatial processing of small visual stimuli. IPO Ann. Progr. Rpt. 12, 81-86.

Blommaert, F.J.J.; Roufs, J.A.J. (1981) The foveal point spread function as a determinant for detail vision. Vis. Res. 21,1223-1233.

Blommaert, F.J.J.; Roufs, J.A.J. (forthcoming) Prediction of thresholds and latency on the basis of experimentally determined impulse responses. In preparation.

Courant, R.; Hilbert, D. (1961) Methods of mathematical Physics I. Third printing. Interscience Publishers Inc. New York, London.

Roufs, J.A.J. (1963) Perception lag as a function of stimulus luminance. Vis. Res. 3, 81-91.

Roufs, J.A.J. (1972) Dynamic properties of vision - I. Experimental relationships between flicker and flash thresholds. Vis. Res. 12, 261-278.

Roufs, J.A.J.; Blommaert, F.J.J. (1975) Pulse and step response of the visual system. IPO Ann. Progr. Rpt. 10, 60-67.

Roufs, J.A.J.; Blommaert, F.J.J. (1981) Temporal impulse and step responses of the human visual system obtained psychophysically by means of a drift-correcting perturbation technique. Vis. Res. 21, 1203-1223.

Wilson, H.R.; Bergen, J.R. (1979) A four-mechanism model for threshold spatial vision. Vis. Res. 19, 19-32.

chapter 7

On estimating optical and neural imaging factors in letter confusions[1]

Frans J.J. Blommaert

Abstract

For the purpose of quantifying models on letter recognition, similarities are often specified in terms of stimulus properties. In this paper, an approach is advocated based on similarities between internal letter responses[2], i.e. it is argued that optical and retinal factors play a more prominent role in letter confusions than is usually assumed. To illustrate this, letter responses were calculated on the basis of earlier experimentally determined point spread functions (Blommaert, Heijnen and Roufs, forthcoming). Next, data on confusion matrices from Bouma (1971) were taken to evaluate different measures which might be useful for quantifying letter response similarities. In the analysis of experimental data, Luce's (1959, 1963) choice model has been used. It was found that if response similarities were expressed in terms of differences between response contours, a fair first order approximation of Bouma's experimental results could be formulated (overall correlation coefficient of 0.91). Other measures like correlation between spatial frequency spectra of letter responses were found to be less successful. The method used provides a means to relate quantitatively stimulus features and optical and neural factors to letter confusions.

7.1 Introduction

Since text is an important source of visual information, it is of interest to study letter recognition, as is reflected in the amount of literature on this subject.

A popular method to learn about the nature of visual letter recognition is by measuring letter confusions when the letter stimuli are degraded to a certain extent with the aim of creating a situation where subjects make mistakes. The measured letter confusions are often presented as a so-called confusion matrix, of which each cell is filled with the amount of a particular confusion between letters (see for instance Bouma, 1971; Townsend, 1971; Geyer and Dewald, 1973). Such stimulus degradation can be achieved in different ways: the letters may be presented at an eccentric position or at a large viewing distance (Bouma, 1971; Bouwhuis and Bouma, 1979), they may be presented briefly or masked by other stimuli (Townsend, 1971; for a review see Breitmeyer and Ganz, 1976), or they may be presented at low contrasts (van

---

[1] Submitted to Spatial Vision.

[2] The expression ' internal letter response' will frequently be used in this paper. It indicates the internal image which the visual system constructs from a letter stimulus. In the following, the internal images of letters are approximated by calculating the effects of optical and neural processing on letter configurations.

Nes, Timmers and Blommaert, 1980; van Nes and Jacobs, 1981; Blommaert, 1980; Blommaert and Timmers, forthcoming).

The results of such experiments have led to models which describe particular confusions more or less accurately and provide explanations regarding the nature of these confusions (Luce, 1963; Atkinson and Kinchla, 1965; Townsend, 1971; Keren and Baggen, 1981). Using these models, letter similarities are often derived from the letter configurations themselves.

As an alternative, we argue that letter similarities may be derived from *letter images* which result from the early stages of visual processing on the letter stimuli. This notion is not new (cf. Garner and Haun, 1978) and is justified by the argument that at the recognition stage only preprocessed letter images are available and not the letter configurations themselves.

Psychophysical visual research has made considerable progress during the last few decades. Some models have been proposed, based on threshold experiments with simple stimuli like lines and points, which can describe and explain visual processing upto the level where detection occurs (Koenderink and van Doorn, 1978; Wilson and Bergen, 1978; Blommaert and Roufs, 1981, Blommaert, Heijnen and Roufs, forthcoming). On the basis of these models, it seems possible to describe more accurately the result of such processing and estimate its effects on the recognition and confusion of alphabet letters. An exploration of the possibilities of this kind of approach is the aim of this article.

Calculations in the spatial frequency domain have been advocated by a number of psychophysicists (Campbell and Robson, 1968; Ginsburg, 1980). Letter properties are, according to this approach, represented by frequency components, which are of a global nature and essentially non-local (Cavenagh, 1984). If the characteristics are sought in the space domain, letter features will behold their local character and their strengths will depend on the stimulus conditions and the nature of peripheral processing. Feature extraction in the space and frequency domain will be compared. From the derived internal letter images, letter similarities can be calculated, which of course do not contain the effects of bias factors. These latter factors may cause differences among letters in terms of their tendency to be named, based on for instance familiarity or personal preferences. Using a modified version of Luce's (1959, 1963) choice model, it will be shown that a fair amount of letter confusions can be explained by optical and neural factors. To demonstrate this, earlier experimental results from Bouma (1971) will be used.

7.2 Background and scope of the analysis

Outline

The global scheme, used as an outline in the present analysis, is depicted in Fig. 1. It explicitly shows that recognition of letters is thought to be split up into two distinct and essentially serial components. The first stage consists of a transformation of the stimulus into some internal image. This transformation will be the summed

TEXT

```
┌─────────────────────────────────────┐
│ - optical deformation               │        VISUAL
│ - transduction from light to        │        ANALYSIS
│                  nerve signals      │
│ - retinal processing                │
└─────────────────────────────────────┘
```

INTERNAL IMAGE

```
┌─────────────────────────────────────┐
│ - feature extraction                │
│                                     │
│ - activation of higher visual       │        RECOGNITION
│                         units       │        STAGE
│ - selection between alternative     │
│                         items       │
└─────────────────────────────────────┘
```
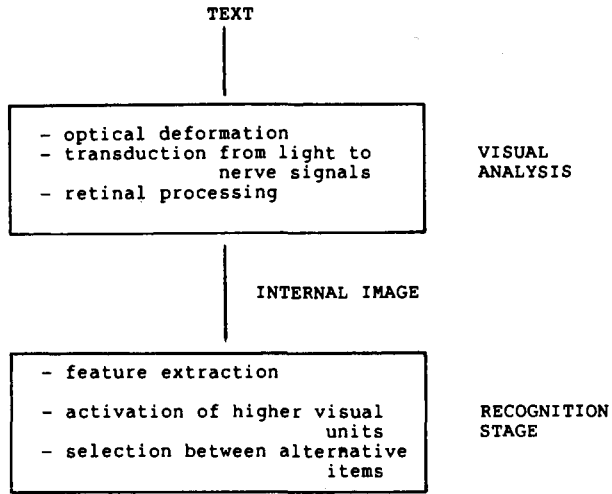
Figure 1: Outline used in the present analysis on letter confusions. The peripheral stages of visual processing and the cognitive mechanism are thought to be successive mechanisms.

effect of early stages of visual processing and will therefore contain for instance deformation by the eye optics and retinal processing. The second stage contains the actual recognition process and comprises for instance activation of higher visual units, memory functioning and selection between alternative items.

In this article, emphasis will be put on the first process, and its global effects on letter recognition will be calculated. Of course, in doing so, some assumptions will have to be made about the nature of the recognition process. In order to avoid misinterpretations as much as possible, these assumptions will be fairly global. Anyway, various models have already been proposed to account for a description of the recognition process. Therefore, in the following we will restrict ourselves mainly to evaluating differences between and correlations among spatial letter responses and letter spectra, respectively.

Confusion matrices

A confusion matrix depends on a number of parameters. First of all, letter font is a variable. It will also make a difference whether the letters are of undercast or uppercast type font. Furthermore, the specific experimental conditions are important. It is clear that masking, illumination level, temporal presentation of the stimuli, contrast and retinal position will all be parameters which may control the outcome of a recognition experiment considerably (see also Blommaert and Timmers, forthcoming). Global effects might be inferred from the way in which psychophysical characteristics depend on these parameters. They obviously affect the first stage of the recognition scheme of Fig. 1. The second stage, however, is left unaltered. Therefore, differences

in confusion matrices obtained under different experimental conditions might be attributed largely to changes in the initial stage of the recognition scheme; i.e. only the internal imaging of letters changes by a change of the experimental conditions.

The outcome of a recognition experiment furthermore depends on cognitive and personal factors such as perceptual and response bias. Therefore, the description of confusion matrices in perceptual response features will be (essentially) incomplete. Deviations will not only be of a nonsystematic nature (noise) but also systematic as a result of these cognitive factors.

	a	s	z	x	e	o	c	n	m	u	r	v	w	d	h	k	b	t	i	l	f	g	p	j	y	q
a	62	3			7	1		6	5	1				5	2	3		3				2				
s	33	10	3		14	3		8	8	7	2	2		2								3	3			2
z	19	6	12	2	14		2	4	1		11	2	1	1				2	4		4	8	4	2		1
x	13	3	6	32	9	2		6	1	2	4	4	2	1		5			2		1	4	2			1
e	19	3			34	5	2	7	10	7	2	1	2	2								2	2			2
o	5				11	57	7	4	4	3		1	1	2		1						1	3			
c	8	2			29	19	19	2	1		4	2		2			1				1	6	2			2
n	6				1		1	56	21	4					2	6	2						1			
m	9	2			1			4	79	1			1		1							1	1			
u							2	11	9	56		2	5	3	3	1	5	1								2
r	1	1	1		5					1	55							12	4		13	4	3			
v				1	1					1	2	72	17										1	5		
w				1	1					3	1	22	70											2		
d	4					1			1	2				83			2	1				1				5
h								2	2						80	2	14									
k	2			4	2			2						2	17	51	14	2	2	2						
b							1	2	1	2					27	3	62						2			
t	2	2			2		2			1	3			2	2	5	4	59	11		5					
i							1							4	2	2		8	69	10	2		2			
l							1							4	4		1	3	50	26	2			9		
f						2	2						5	2				3	2		81	2		1		
g	3				6	2		2	1	4	4	2		3			2					25	9	3	3	31
p									1	1	1		2			8						2	84	1		
j											5			1	3	5	3	1						82		
y																										
q	2				2	3			1	2				5	1							6	3	1		74

Figure 2: Experimental results on letter confusions taken from Bouma (1971). The rows give stimulus letters, the columns indicate response letters. The letters have been arranged in groups of three to five, such that within the groups confusions are relatively common (after Bouma). Confusions were enhanced by presenting the isolated letters at a distance of 0.5 m and a retinal eccentricity of 7 deg. Data were averaged over 10 subjects to whom each letter was presented 20 times.

In Fig. 2, an example of an experimentally determined confusion matrix is shown. It is taken from Bouma (1971), and will be used further on in order to evaluate calculated letter confusions. The letters (of the undercast type font Courier 10) were presented at an eccentric position of 7 deg and a viewing distance of 0.5 m.

Point spread functions

In order to quantify early stages of visual processing, point spread functions will be used, which quantify the spread of the visual system if it is stimulated with an infinitesimally small point source. Such point spread functions were determined experimentally by means of a perturbation method. Assuming linearity, radial symmetry and peak detection, measured threshold differences were converted to discrete values for a point spread function.

Although at every retinal position receptive fields or point spread functions with different size and sensitivity are operating, as a result of the slim nature of characters only the smallest point spread function, which indeed is found by the method mentioned above, determines the letter response (Blommaert and Roufs,1981; Blommaert, Heijnen and Roufs, forthcoming).

An example of these experimental results is given in Fig. 3. It contains reduced experimental values as a function of the radius r of four PSF's obtained under different experimental conditions. In order to facilitate comparison, the response amplitudes are normalised at their maximum value and the radius is reduced using a scale factor for the other conditions than 1200 Td in the fovea (subj. HH). The data can be interpreted as the result of optical and neural processing if a point is presented foveally at a background level of 1200 Td.

It clearly shows that an inner excitative zone is surrounded by an inhibitive one, comparable with receptive field like structures measured physiologically (Kuffler, 1953; Wiesel, 1960). The height and width of point spread functions generally will depend on adaptation level and retinal position.

7.3 Theoretical frame

The concept of a point spread function is of important value only if visual processing is linear in good approximation. A test on linearity turned out to be positive (Blommaert and Roufs, 1981). Encouraged by this result, we assume that visual processing of slender details can be characterised by a point spread function, which by definition equals the visual response to a unit point source. In that case, the response of the human visual system to arbitrarily shaped slender details (so including letters) can be calculated by convolution:

$$U(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(x',y')U_\delta(|\vec{r} - \vec{r}'|)dx'dy', \tag{1}$$

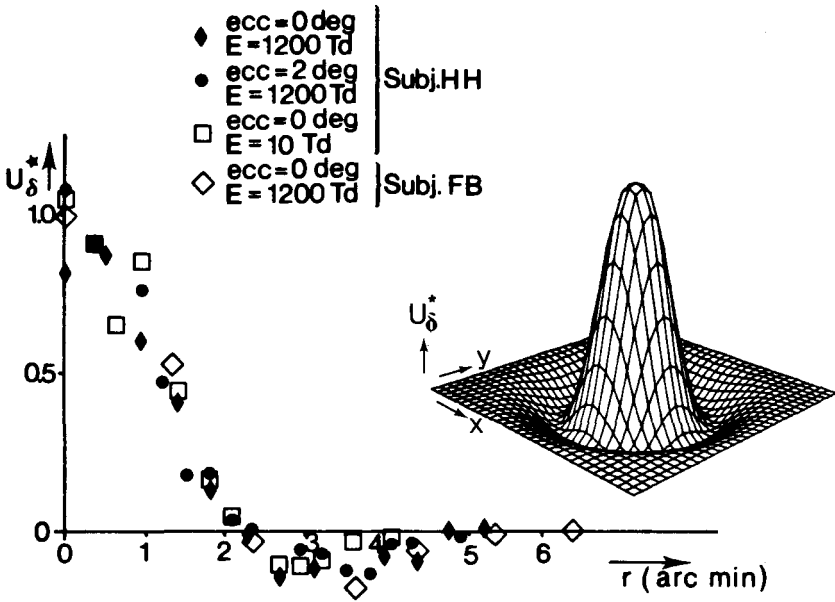$$|\vec{r} - \vec{r}'| = \sqrt{(x-x')^2 + (y-y')^2},$$

where:

Figure 3: Data on experimentally determined point spread functions taken from Blommaert, Heijnen and Roufs (forthcoming). Results are normalised at the maximum and reduced with respect to r by a scale factor depending on subject, adaptation level and retinal position with respect to the point spread function for subject HH at the 1200 Td adaptation level. The result can be interpreted as the spread due to optical factors and neural processing if the fovea is stimulated with a point source. For easy comparison with later figures, a 3D-plot, fitted to the experimental results of subject HH (0 deg; 1200 Td), is also given.

$U_\delta(r)$ is the (radially symmetrical) point spread function,
$S(x, y)$ is the luminance distribution of the stimulus,
$U(x, y)$ is the visual image of the stimulus.

If summation is carried out with sufficient precision, eq. 1 can be approximated by:

$$U(k, l) = \Delta x \Delta y \sum_i \sum_j S(i, j) U_\delta(k - i, l - j). \tag{2}$$

In this expression, all continuous functions of eq. 1 have been replaced by their discrete equivalences. Eq. 2 was in fact used for the numerical calculation of visual responses on letter stimuli.

The Fourier transform $F_l(u, v)$ of a letter response can be expressed as:

$$F_l(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-j(ux + vy)} U(x, y) dx dy. \tag{3}$$

Since $U(x, y)$ equals the convolution of stimulus and point spread function (see eq. 1), eq. 3 can be written as:

$$F_l(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-j(ux+vy)} S(x, y) \otimes U_\delta(x, y) dx dy, \qquad (4)$$

where the symbol $\otimes$ stands for convolution.

A handy property of Fourier transforms is that convolution of two functions in the space domain equals multiplication of their spectra in the frequency domain (see for instance Bracewell, 1975). Eq. 4 then transforms into:

$$\begin{aligned} F_l(u, v) &= (\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-j(ux+vy)} S(x, y) dx dy) \times (\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-j(ux+vy)} U_\delta(x, y) dx dy), \\ &= S(u, v) \times U_\delta(u, v), \qquad (5) \end{aligned}$$

where $S(u, v)$ is the Fourier spectrum of the letter stimulus and $U_\delta(u, v)$ that of the point spread function.

Since point spread functions have been assumed to be radially symmetrical (Blommaert and Roufs, 1981), the Fourier transform of a point spread function equals its Hankel transform:

$$\begin{aligned} U_\delta(u, v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-j(ux+vy)} U_\delta(x, y) dx dy, \\ &= 2\pi \int_{0}^{\infty} r J_0(\rho r) U_\delta(r) dr. \qquad (6) \end{aligned}$$

In this equation, $r = \sqrt{x^2 + y^2}$ and $\rho = \sqrt{u^2 + v^2}$ while $J_0(\rho r)$ are Bessel functions of the first kind.

7.4 Methods

All calculations were carried out on a VAX 11/780 computer, supplied with NAGLIB mathematical subroutines. A separate software packet (DISSPLA) was used in order to generate computer plots.

Matrix convolution

Visual responses on letters were calculated with the aid of discrete matrix convolution according to eq. 2. To this end, letters of the type font Courier 10 were fed into the computer. After visual inspection it was decided that matrices of 16×32 pixels were large enough in order to describe the letters in a sufficiently natural way. In these matrices, short letters, ascenders and descenders have heights of about 13, 17 and 17 pixels, respectively, while the line width equals 2 picture elements. Examples of some discrete letters are shown in Fig. 4 for four sizes.

Discrete data of point spread functions, as used for the description of peripheral visual processing, were interpolated from the measured data points (Blommaert et al., forthcoming; subj. HH; E = 1200 Td; ecc. = 0), using spline interpolation subroutines. The matrices used to approximate the two experimental conditions used by Bouma (1971) contained 31×31 and 59×59 pixels respectively.

Figure 4: Examples of some discrete letters implemented in the computer to simulate the Courier 10 type font, shown for four sizes. The letters were framed in matrices of 16×32 pixels. If the height of the letters decreases, the frayed effect of discretising can be seen to diminish.

Fourier transforms

Fourier spectra of letters were obtained using discrete Fourier transform techniques. If appropriate, the Fourier spectra were multiplied by the Hankel transforms of point spread functions according to eq. 6. The Hankel transforms were calculated also in a discrete fashion using standard subroutines for the approximation of Bessel functions of the first kind.

7.5 Calculated internal images of letter stimuli

a Space domain

In order to evaluate calculations on letter response similarities, we searched for experimental results on letter confusions for lower case letters and obtained under stimulus conditions where the letters subtended only small angles on the retina. The model used to calculate the letter responses by applying only one dominating point spread function is expected to work best for such results.

Therefore, the stimulus conditions were chosen identical with the ones reported on by Bouma (1971). In two different experimental paradigms, he obtained correct recognition scores of about 55% using IBM Courier 10 type letters. In the first condition the isolated letters were presented foveally at a distance of 3.35 m (further on to be named "foveal-distant vision"). In the second condition, he presented them at an eccentricity of 7 deg and at a viewing distance of 0.5 m (named "eccentric-near vision" further on). In both cases, the presentation time was long (200 ms and 3 s, respectively), and the adaptation level photopic (1800 Lux ($E \approx 5000\ Td$) and 400 Lux ($E \approx 1000\ Td$), respectively).

Point spread functions were measured before under slightly different conditions

(Blommaert and Roufs, 1981; Blommaert, Heijnen and Roufs, 1981, forthcoming). In order to simulate the experimental paradigm for foveal-distant vision, a point spread function was used which was determined foveally at a photopic adaptation level of 1200 Td (see Fig. 3). In doing so, we assumed that the width of a point spread function does not shrink substantially in the range between 1200 Td and 5000 Td. This assumption seems a reasonable approximation, judged by experimental data on visual acuity (Le Grand, 1953).

The experimental condition of eccentric-near vision is not so easy to specify since no point spread functions were determined at 7 deg eccentricity. Therefore the point spread function has to be extrapolated from other experimental data. To begin with we assume here, as we did before (Blommaert, Heijnen and Roufs, forthcoming) that point spread functions at different retinal locations are isomorphic and their widths depend linearly on eccentricity. This is in accordance with proposed models on contrast detection (cf. Koenderink and van Doorn, 1978; Wilson and Bergen, 1979). It is also in agreement with visual acuity, which varies inversely proportional with distance from the fovea (cf. Le Grand, 1953; Sloan 1968). The point spread function can then be specified by extrapolating a width factor from experimental data. On the basis of our point spread function results, the estimated width increase would amount to a factor of 3.6. Within experimental precision this factor is in agreement with visual acuity (Le Grand, 1967; Sloan, 1968), although a factor of 7 also would have been possible. In order to link up with the mentioned psychophysical models we chose the width factor to be 3.6, the more so as in our concept a factor of 7 would mean that the experimental conditions for foveal-distant and eccentric-near vision would be the same. Therefore, an advantage of taking this factor to be 3.6 is that the sensitivity to variations of viewing distance and eccentricity might be estimated.

In performing the calculations, it was furthermore assumed that the shape of a point spread function does not depend on the sign of the stimulus, i.e. the shape is assumed to be the same regardless of the point stimulus being darker or brighter than the surround. This assumption is probably valid in first order approximation, according to visual acuity data (van Norren, 1981).

For the two different experimental conditions, calculated letter images are shown in Fig. 5, using the letters 'f' and 'c' as stimuli.

The letter images for the eccentric-near vision condition (Fig. 5a) are calculated using a point spread function matrix of 31×31 elements (width of pixel elements 1.08 arc min). For the foveal-distant vision condition (Fig. 5b), the point spread function was discretised in a 59×59 matrix (width of pixel elements 0.16 arc min). It can be seen that the letter images of Fig. 5a still contain more detail information than the ones of Fig. 5b. This is caused by the fact that the point spread function used for simulating the eccentric-near stimulus condition was estimated to be about a factor of 2 less broad. Although the point spread function for foveal-distant vision is broader with respect to the stimulus dimensions, the actual width (in arc minutes) is smaller.

For the foveal-distant vision condition (Fig. 5b) the difference between the responses for the letters 'f' and 'c' is hardly visible. Furthermore, they look very much like the
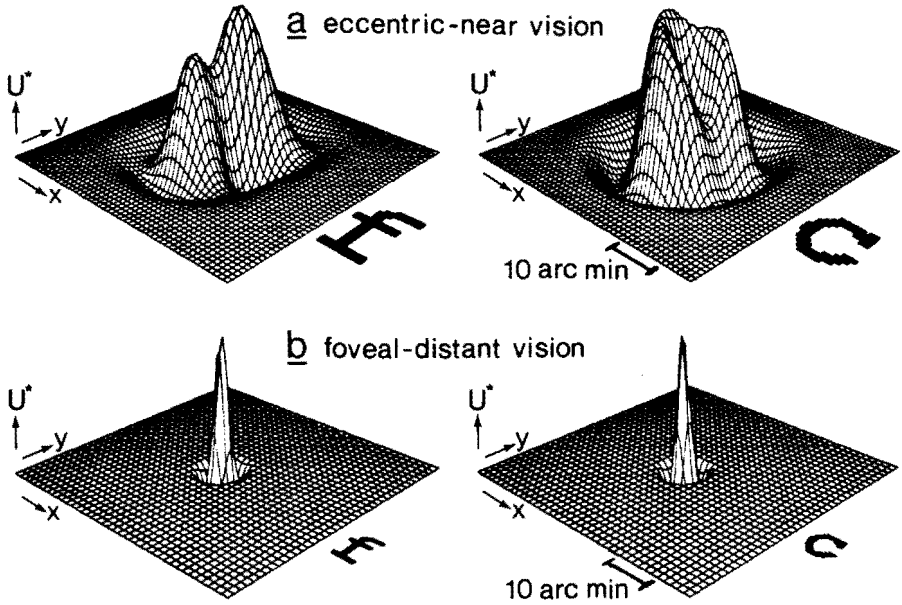
Figure 5: Calculated internal letter response profiles for the letters 'f' and 'c' for the two simulated experimental conditions of eccentric-near vision (5a) and foveal-distant vision (5b). Note how the letter features are progressively revealed if the width of the point spread function is increased with a factor of two with respect to letter dimension, as is the case for the foveal-distant condition. The projected letter shapes are not on scale.

point spread function itself. In the limiting case that the letters are infinitesimally small, the responses will even be equal to the point spread function. Looking at the calculated letter response profiles of Figs. 5 and 6, it seems a bit puzzling that the responses for both stimulus conditions may yield the same average correct score in recognition experiments. It should, however, be kept in mind that the letters are processed at different retinal locations where sensitivity and noise properties may be quite different, implying that the relation between the internal image of a stimulus and its detection or recognition is not obvious, and also that the width factor for eccentric-near vision has been kept at 3.6.

For comparison, contour plots of the same letter responses are shown in Fig. 6. It shows nicely how the different letter features are deformed as a result of peripheral visual processing.

letter f   letter C

a

eccentric-near
vision
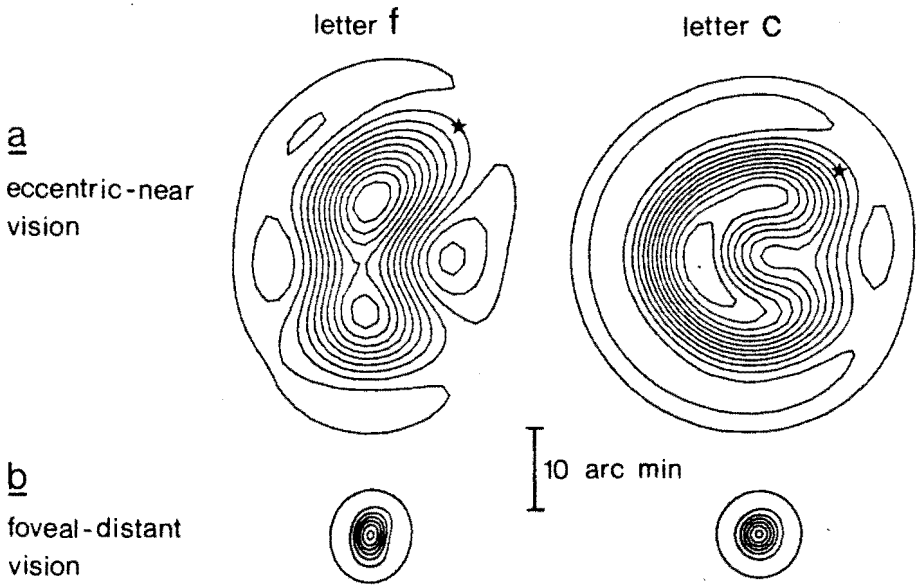
b

foveal-distant
vision

10 arc min

Figure 6: Same responses as shown in Fig. 5 now as contour plots.
The lines are intersections of the letter responses with equally distant
horizontal planes. In Fig. 6a, zero-crossings are marked with $\star$. Note
that the original letter features are much more visible in this figure than
in Fig. 5.

b The frequency domain

Two-dimensional frequency spectra of letters were calculated using Fast Fourier trans-
form algorithms (see for instance Bracewell, 1975). In order to simulate the stimulus
conditions of eccentric-near and foveal-distant vision, respectively, these spectra were
multiplied by Hankel transforms of the point spread functions according to eq. 6,
in order to make a fair comparison with the spatial domain analysis possible. The
widths of the used point spread functions were chosen the same as for the calcula-
tions in the space domain. In this way, complex frequency spectra were obtained, of
which the absolute values were taken.

Some results are shown in Fig. 7 for the letters 'f' and 'c'. The shown spectra are
the absolute values of the Fourier transforms of Figs. 5a and 5b.

As can be seen from Fig. 7, the letter spectra for the eccentric-near vision condition
(Fig. 7a) contain more high frequency components than the ones for the foveal-
distant vision condition (Fig. 7b). This reflects finer detail information. Note the
difference in spectra for the letters 'f' and 'c': the spectrum of the response of the
letter 'c' is almost radially symmetrical owing to the circular shape of this letter.
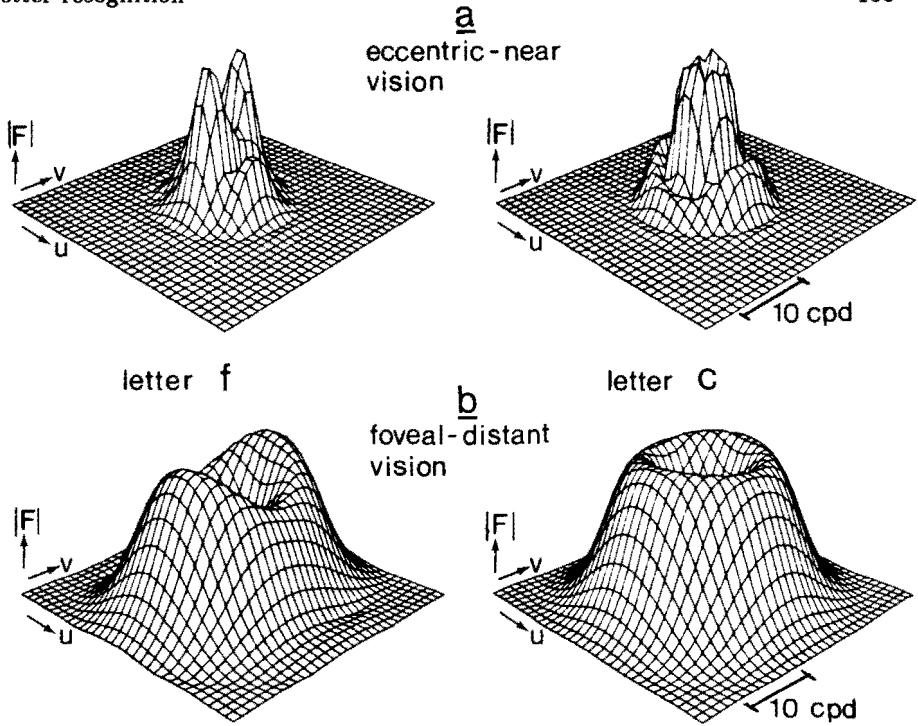
Figure 7: Three-dimensional representation of the absolute values of the frequency spectra of the responses from Fig. 5. For the eccentric-near vision condition, the spectra show more structure than for the foveal-distant condition, reflecting the effect of more or less fine detail information in the spatial response.

7.6 Working hypothesis for the evaluation of calculated letter confusions

It can be seen from the calculated letter responses of Figs. 5 and 6 that the major part of information, which characterises a particular letter, is lost if it is viewed under the chosen stimulus conditions. This loss of information will result in an enhanced chance that the stimulus letter will be confused with another letter.

For the recognition of a letter, a subject must have some internal memory image of the letter features which distinguish a particular letter from the others. The way in which such a letter image is represented in the subject's memory is not known, although a lot of interesting theories have been developed (Cavenagh, 1984). The storage might be related to space coordinates or to frequency coordinates. It might even be a storage of specific letter cues or a storage of inner and outer contours.

In the recognition process, the visual letter responses will be compared somehow with these memory images of letters, and the chance of naming a particular letter will depend on the similarity of the letter response with the memory interpretation of itself

and that of other letters. Therefore, the number of times in which a particular letter confusion occurs will be a measure for the resemblance of the letter response with the memory interpretation of the particular letter to be confused with. If response bias is also incorporated we can write:

$$p(j \mid i) = a_i b_j p(m_j \mid U_i), \tag{7}$$

where:

- $p(j \mid i)$ is the probability that letter j is named if letter i is the stimulus.

- $p(m_j \mid U_i)$ is the probability that the memory representation m_j of letter j is activated if U_i is the response to letter i,

- b_j is a response bias factor for letter j (the same choice as Luce (1963),

- a_i is a normalisation factor for stimulus letter i, necessary to ensure that in a decreased universe the ratio of the probabilities is kept equal, there sum being one.

Knowledge of the specific memory interpretation of letters is not available. Therefore, a straightforward calculation of letter confusions is impossible without making assumptions on the nature of these memory items. In the following we will try to get around this difficulty by comparing the obtained letter responses with that of other letters. In doing so, the similarity between letter response and memory representation may be reflected by:

$$p(m_j \mid U_i) = p(U_j \mid U_i) p(m_j \mid U_j), \tag{8}$$

where:

- $p(U_j \mid U_i)$ is the probability that the letter response U_i is confused with the letter response U_j.

Substituting eq. 8 in eq. 7 one obtains:

$$p(j \mid i) = a_i b_j p(U_j \mid U_i) p(m_j \mid U_j). \tag{9}$$

The structure of eq. 9 is alike the one defining Luce's (1959, 1963) choice model. In fact, if $p(m_j \mid U_j) = 1$, eq. 9 would be identical to the formulation of the choice model. In the present interpretation the terms have a somewhat different meaning as will be discussed further on.

For the probability that letter i is named if letter j is the stimulus, it can likewise be written:

$$p(i \mid j) = a_j b_i p(U_i \mid U_j) p(m_i \mid U_i). \tag{10}$$

For the correct scores on i and j, it follows:

$$p(i \mid i) = a_i b_i p(m_i \mid U_i), \tag{11}$$
$$p(j \mid j) = a_j b_j p(m_j \mid U_j). \tag{12}$$

It is our goal to derive a quantity from the above equations which is proportional to the chance $p(U_i \mid U_j)$ that the two letter responses U_i and U_j are confused. Under the assumption that $p(U_i \mid U_j) = p(U_j \mid U_i)$ it can be derived from equations 9, 10, 11 and 12 that the similarity measure $S(i,j)$ between the responses U_i and U_j can be estimated according to:

$$S(i,j) = \hat{p}(U_i \mid U_j) = \sqrt{\frac{p(i \mid j)p(j \mid i)}{p(i \mid i)p(j \mid j)}}. \tag{13}$$

This expression is equivalent to the nominator $\hat{\eta}_{ij}$ for letter similarities sometimes derived for estimating the parameters of the choice model (cf. Townsend, 1971). In the following, eq. 13 will be used for the purpose of comparing experimental data on letter confusions with calculated letter similarities.

An advantage of using eq. 13 is that it is not important how the letter responses are represented, since only response similarities have to be determined. The calculations may be carried out in the space domain as well as in the frequency domain, and make use of any comparison mechanism. This property will be exploited in the next chapter in order to compare calculations on letter response similarities, which are based on different algorithms. The basic notion is that there might be found a monotonic relation between experimental and theoretical similarity quantities. It should be kept in mind, however, that the calculations are global and will only hold in first order approximation.

In adopting the choice model, the effect of bias parameters is effectively evaded. However, in order to be able to calculate confusion matrices using eq. 9, the products $b_i p(m_i \mid U_i)$ have to be known. They can be estimated on the basis of equations 9 to 12 from experimentally determined asymmetries in confusion scores using an approximation similar to the one given by Townsend (1971):

$$b_i p(m_i \mid U_i) = \frac{1}{N} \sum_{j=1}^{N} \sqrt{\frac{p(i \mid i)p(i \mid j)}{p(j \mid i)p(j \mid j)}}, \tag{14}$$

under the condition that $p(i \mid j)$ and $p(j \mid i) \neq 0$.

As pointed out by Gilmore et al. (1979) and Keren and Baggen (1981), more accurate approximations can be found. We judge eq. 14, however, to be of sufficient precision for the present analysis.

Note:

In one of Bouma's confusion matrices, that of foveal-distant vision, 'illegible' answers were included. The percentages were found to depend on the specific stimulus letter. Due to the nature of the model, the effect of these answers is eliminated in equations (13) and (14).

7.7 Results

Calculated letter response similarities evaluated against experimental letter confusion data

a Correlation of responses in the space domain

For the two experimental conditions used by Bouma (1971), correlation coefficients were calculated between all spatial letter responses using Pearson's r:

$$r = \frac{\sum_i \sum_j (U_1(x_i, y_j) - \overline{U_1})(U_2(x_i, y_j) - \overline{U_2})}{\sqrt{\sum_i \sum_j (U_1(x_i, y_j) - \overline{U_1})^2 \sum_i \sum_j (U_2(x_i, y_j) - \overline{U_2})^2}}. \tag{15}$$

where:

$U_1(x_i, y_j)$ is the response of letter (1) in the coordinates (x_i, y_j).

$U_2(x_i, y_j)$ is the response of letter (2) in the coordinates (x_i, y_j).

$\overline{U_1}$ is the average response on letter (1).

$\overline{U_2}$ is the average response on letter (2).

Such a correlation measure will be high if at corresponding coordinates the letter responses have corresponding amplitudes (if letter (1) equals letter (2) than $r = 1$). If there is a decreasing correspondence between the response amplitudes of different letters, the correlation r will decreasingly deviate from unity.
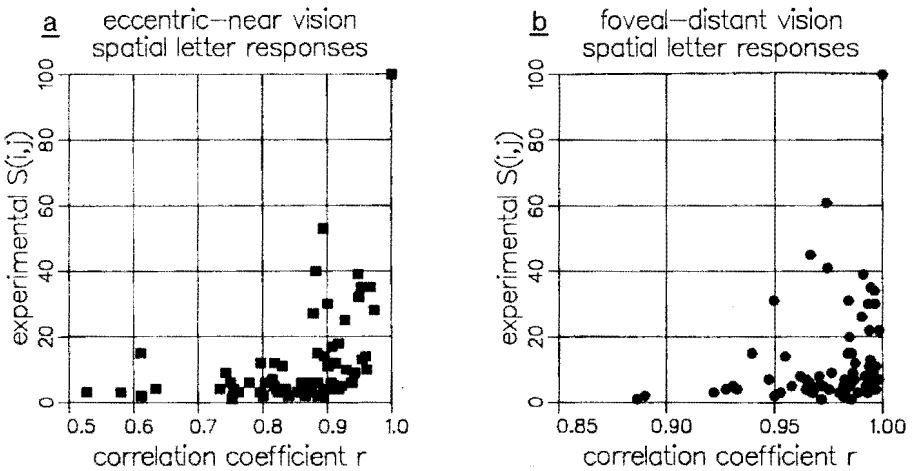
Figure 8: Calculated correlation coefficients between spatial letter responses, plotted against experimentally determined letter similarities $S(i, j)$ as estimated from Fig. 2 using eq. 13. Data for $S(i, j) = 0$ have been omitted for clarity. The correlation coefficients are calculated between spatial letter reponses using Pearson's r.

It is obvious that such a correlation measure will depend on the positions of the letter responses with respect to each other. Therefore, the correlation factors were

maximised by translating the letter-response matrices with respect to each other within a square of 11×11 matrix elements. In a manner of speaking, responses of different letters were shifted in the (x, y)-plane with respect to each other in order to obtain an optimal fit.

It was found that, in general, a pronounced maximum correlation existed if the positive- and negative response areas of the one letter response matrix coincided as much as possible with those of the other letter-response matrix. Due to the nature of the correlation coefficient and that of the letter responses, the calculated correlation coefficients between letter responses were always found to be high.

This can be seen in Fig. 8, where the theoretical results have been plotted against experimentally derived similarity scores $S(i, j)$ according to eq. 13 for eccentric-near and foveal-distant vision, respectively. For clarity, the data for experimental similarity $S(i, j) = 0$ have been omitted in the figure, as they will be in all figures of this paragraph.

As can be seen from Fig. 8, a definite though noisy relation exists between the theoretical values and experimental ones: if the experimentally derived similarity scores are high, the letter response correlations will be high also, and on the other hand, if the correlation coefficients are low, experimental similarity scores are also low.

If the results for the two conditions are compared, it is striking that for the foveal-distant vision condition all correlation coefficients are found to be extremely high $(r > 0.81)$. This result may be appreciated if the response profiles for the letters 'c' and 'f' of Figs. 5b and 6b are kept in mind. Apart from that, the scatter diagram for eccentric-near vision shows relatively less noise.

b Correlations between frequency components

Pearson's r correlation coefficients have also been calculated between the discrete frequency spectra of letter responses. To this end, the absolute values of the complex Fourier spectra were taken. This implicates that in the calculations, the phase of the different components is neglected. An advantage of doing so is that the position of the letters with respect to each other becomes irrelevant. A disadvantage of the procedure is that some specific letter properties are lost in the calculations, such as mirroring with respect to the horizontal and vertical axis. Certain letter response correlations will, due to this effect, be extremely high like correlations between the letters 'p' and 'b'.

In Fig. 9, calculated correlation coefficients are plotted against experimental similarity scores $S(i, j)$ according to eq. 13.

As can be seen from Fig. 9, the relation found earlier between the correlation coefficient and the experimental similarity measure for the space domain as depicted by Fig. 8, also roughly holds for the frequency domain: although there exists a definite relationship between similarity rates and correlations for both space and frequency domain, this relation is rather noisy. The results suggest that other measures might be more appropriate for the purpose of quantifying the similarity between letter responses.

a   eccentric—near vision
    letter response spectra

b   foveal—distant vision
    letter response spectra

c   eccentric—near vision
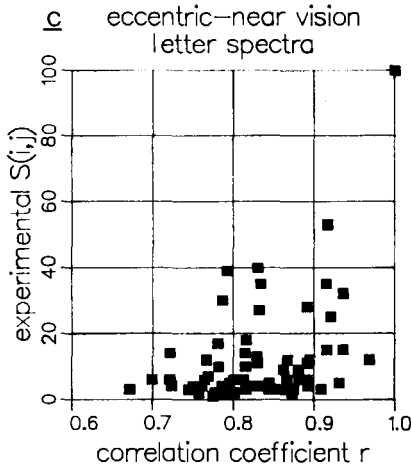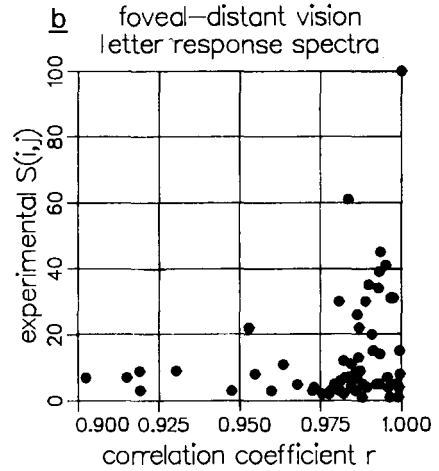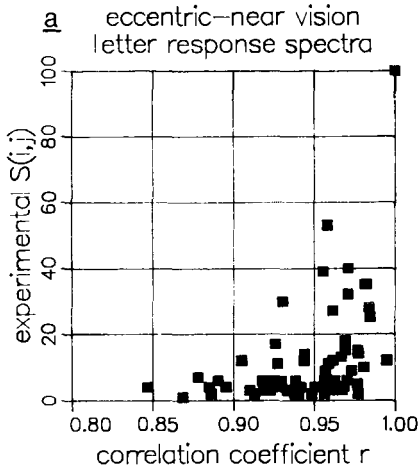   letter spectra

Figure 9: 9a and 9b.   Calculated cor-
relation coefficients between frequency
spectra of letter responses, plotted
against experimentally determined sim-
ilarity rates $S(i,j)$ for eccentric-near
and foveal-distant vision, respectively.
9c Correlation coefficients between fre-
quency spectra of the original letter pro-
files plotted against experimental $S(i,j)$
for eccentric-near vision.   Data for
$S(i,j) = 0$ have been omitted.

For comparison, correlation coefficients were calculated between the letter spectra
themselves, i.e. without multiplying the letter spectra with the Hankel transform
of the point spread function according to eq. 5. This approach, though not quite
proper since the effects of peripheral visual processing are neglected, is sometimes
used in the literature to evaluate the quality of letter types (Maddox, 1980). The
computational results are shown in Fig. 9c plotted against the similarity scores for
eccentric-near vision. As can be seen from this figure, less correspondence with the
experimental data is visible in comparison with the results of Fig. 9a.

c Differences between letter responses

From a psychophysical point of view, it is more logical to search for a measure
which relates *thresholds* of letter details with confusion scores. In such a concept,

difference functions of letter responses might be a valuable starting point. One might ask which letter differences will be detected easily and which will not be detected at all. To this end, we subtracted spatial letter responses from each other, thus obtaining two-dimensional letter difference functions. After that, the difference functions were summed according to Quick's (1974) rule for contrast detection (see also Blommaert, Heijnen and Roufs, forthcoming).
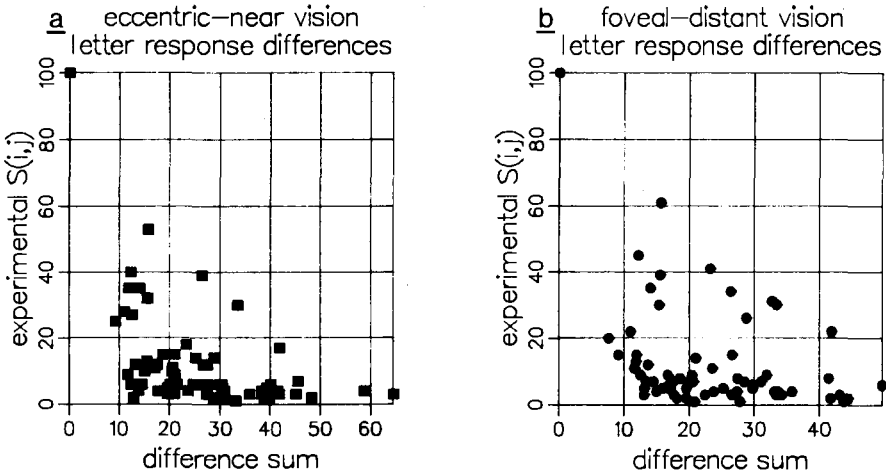

Figure 10: Difference sums of spatial letter responses plotted against experimentally determined $S(i,j)$ estimated according to eq. 13. They were obtained by subtracting spatial letter responses from each other and summation of the results according to Quick's (1974) law. For clarity, data for $S(i,j) = 0$ have been omitted. The horizontal axes have been scaled to fit the range of calculated difference sums.

These sums were treated as measures for the similarity between responses. A large sum is then thought to correspond to easy detectability of the letter difference and a small sum to hard detectability, thus yielding high similarity. The results of these computations are shown in Fig. 10 for the two experimental conditions. Since the choice of the origins again plays a role in determining the resulting sum of differences, the letter responses were translated with respect to each other and the minimum sum was taken as a measure for similarity.

Results of these computations should yield a sort of psychometric function: on the one axis the 'strength of the stimulus' (letter response difference sum) and on the other axis the detection probability ('not confusion'). Although the plotted relation between response differences and experimental similarity rates can be seen to exhibit the well-known S-shape roughly, the results are too noisy to gain confidence in the algorithm.

d Contour differences between letter responses

When looking at Courier 10 type letters under the approximate experimental condi-
tions of eccentric-near and foveal-distant vision, it is striking that almost all of the
inner structure of the letters seems to have disappeared: they all look like dark blots
with a more or less outstanding contour. Therefore, an algorithm was worked out
based on contour differences between letter responses. To this end letter contours
were defined as the zero-crossings (crossing of the profile with the zero response
plane) of the letter response profiles (see Fig. 6), yielding shapes as shown in Fig.
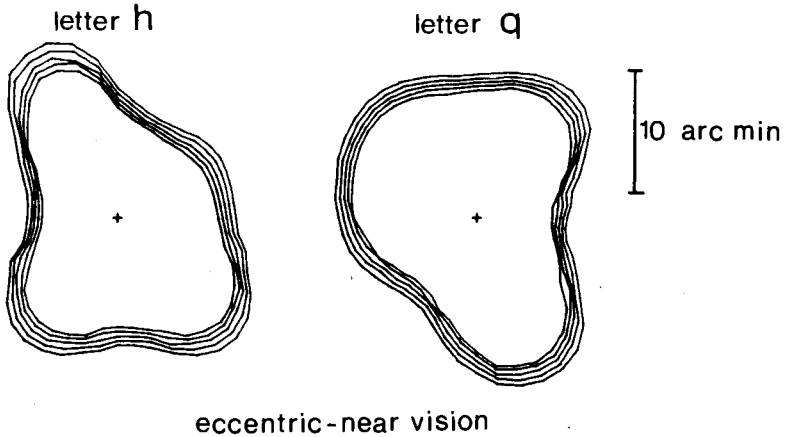11 for some examples.

letter h   letter q

10 arc min

eccentric-near vision

Figure 11: Contour functions for the letters 'h' and 'q', that were used
for calculations on contour difference sums. For the actual computa-
tions, the outer contours were taken, representing the zero-crossings
of the calculated responses. The crosses are the centers of gravity of
the responses, acting as the origin for the coordinate system. Contours
were (discretely) calculated in polar coordinates using 48 angles around
the origin.

In Fig. 11, contours were taken from intersections of planes at 0.00, 0.06, 0.11, 0.17,
0.22 and 0.28 of the maximum response height. Fortunately, there is a considerable
congruency between the different slices so that the exact choice of the height of the
plane is not critical.
In order to make size variation and letter rotation easy to calculate, the contours were
computed in polar coordinates; while the centre of gravity of the responses was defined
as the origin. Size variation is then easily incorporated as a parameter variation in
the r-direction whereas rotation becomes a translation of the angle variable.
With the aim of finding a measure for contour differences between letter responses,
the contour functions were fitted to each other in size by a method of least squares
and the resulting contour difference was expressed in units of area that was left
between the fitted contours. Such contour differences were calculated between all
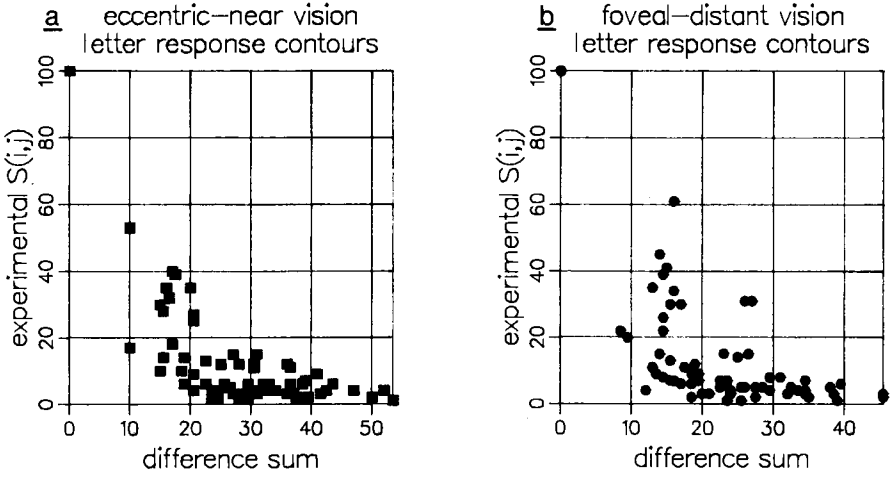
Figure 12: Contour difference sums plotted against experimentally determined letter similarities $S(i,j)$ estimated from the confusion matrix of Fig. 2 according to eq. 13. Data for $S(i,j) = 0$ have been omitted. The horizontal axes have been scaled to fit the range between extreme high and low difference sums.

letter responses and are shown as the scatter diagrams of Fig. 12. Rotations were permitted over a small angle of 15 deg and the minimum contour difference was taken to represent similarity. From the results it can be seen that this measure looks rather promising for the purpose of quantifying similarity between letter responses, especially for the eccentric-near vision condition. Although there is still some scatter left, a rather smooth relation can be seen to emerge between contour difference sums and experimantal similarity rates in Fig. 12a. For the foveal-distant vision condition, the result is again more noisy. Possible causes will be discussed further on.

In the foregoing, results for $S(i,j) = 0$ have continually been omitted. Furthermore, it is of interest which particular letter similarities do occur (in view of the goal of this paper, a detailed analysis into the structure of confusions has been avoided). In order to give an example, theoretical contour similarities have been plotted against experimental $S(i,j)$ in Fig. 13 for the letters 'a' and 'l' in the eccentric-near vision condition.

On the basis of a few additional assumptions, theoretical confusion matrices can be derived from the similarity measures given in this paragraph. Results of such a computation will be given in the next section.
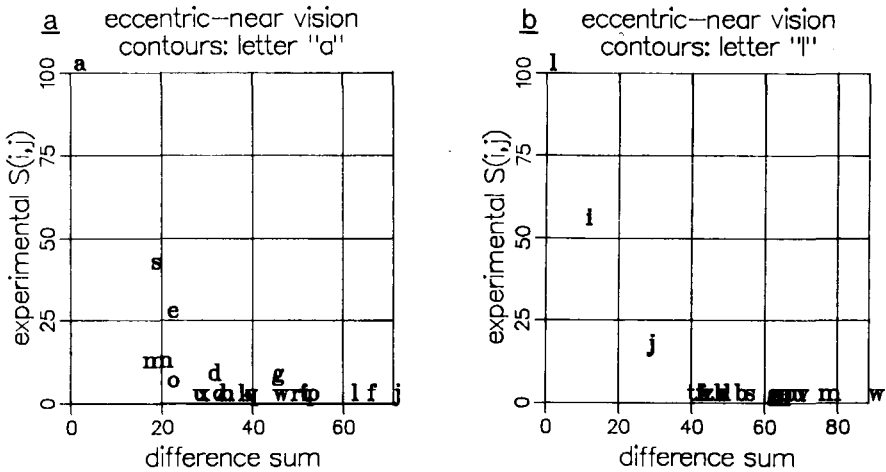
Figure 13: Calculated response similarities for two stimulus letters ('a' and 'l') plotted against experimental $S(i,j)$ for eccentric-near vision. The computations are based on contour difference sums. All letter data are displaced a few units in order to improve visibility as can be seen from the the positions of the letters 'a' and 'l' in Fig. 13a and 13b, respectively. They are shifted from the point (0,100).

7.8 On calculating confusion matrices from theoretical similarity measures: an example

In order to quantify a confusion model, a theoretical measure like the contour difference sum might well be used. This will be demonstrated in the following.

First, a conversion from similarity "strength" (difference sum) to similarity "probability" is needed. This was done by using the psychometric function for contrast detection. Using this function, two parameters had to be estimated, namely the threshold level Δ and the slope of the psychometric function. In order to simulate some variation between subjects, the slope was chosen a factor of two smaller than for normal contrast detection of subject FB, thus yielding a slope of 0.34 log-units per 100 percent in the 50% point (see Blommaert and Timmers, forthcoming).

The threshold parameter Δ was fitted to obtain an average correct score of 55%, as equal to Bouma's experimental result for eccentric-near vision. The products of correction and bias parameters were estimated from the confusion matrix of Fig. 2, according to eq. 14.

The confusion matrix calculated according to eq. 9 is shown in Fig. 14, and may directly be compared to the one of Fig. 2. In order to facilitate comparison, the scores of the calculated and experimental confusion matrix are plotted against each other in the scatter diagram of Fig. 15.

It shows a fair correlation of 0.71 for correct scores and 0.74 for confusions, while

	a	s	z	x	e	o	c	n	m	u	r	v	w	d	h	k	b	t	i	l	f	g	p	j	y	q
a	51	5		1	8	4		7	21	2																
s	15	11	6	7	14	20	2	11	8	2	1	2		1	1											
z	1	9	41	12	3	2		1		1	13	1		1	6	2	2	1	1							4
x	3	9	18	37	4		1	14	8	3	1	2		1												
e	6	5		1	32	22	6	4	14	5	4															
o	10	10	1		28	23	3	3	13	3	4															
c	1	7	1	4	39	8	24	3	1	2	2	9														
n	17	6		4	3	1		32	33					1			1									
m	29	3		1	6	2		11	44	1	1															
u	3	6	1	4	17	2		1	4	44	12	4			1											1
r			3								92										2	1				
v		2			2	1	1				2	83	6													1
w					1				1	2		28	68													
d	3	6	4		2	4					1			78				1								
h	1	2	7	2				2							56	6	23									
k	1		17	1				2							17	53	2	5	2							
b	2	2	4		5	2		3			1				30	1	49									
t		1									1							96	1							
i																		1	72	26						
l																			69	29					2	
f								11										1	1		87					
g		2	1		1	1					3	10									1	30				50
p		1	2		1	1					24	4											68			
j																		5	2					93		
y																										
q		3	7		2	1					2	1	8								7					67

Figure 14: Confusion matrix calculated on the basis of detection probability of letter similarities according to the contour difference sums of Fig. 12a. The simulated experimental condition is that of eccentric-near vision. The bias factors were estimated from asymmetries in the confusion matrix of Fig. 2 according to the choice model. The results may be directly compared with Fig. 2.
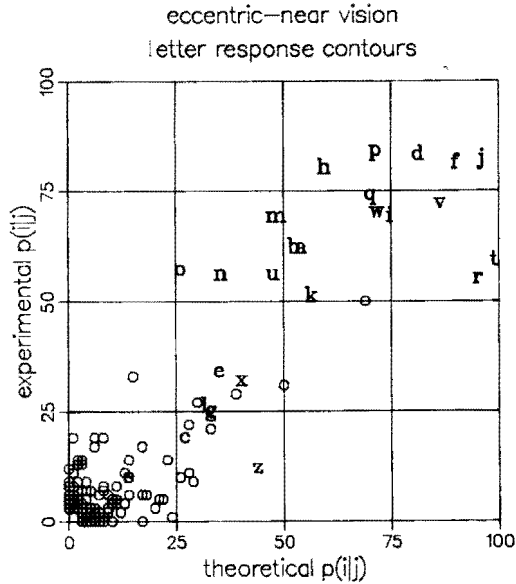
eccentric—near vision
letter response contours

Figure 15: Scatter diagram of correct and incorrect scores respectively between Bouma's experimental data on eccentric-near vision (Fig. 2) and the theoretical matrix of Fig. 14 obtained from contour difference sums. The correct scores are indicated by the letters and incorrect scores by ○. Data in the neighbourhood of the origin are omitted. The calculated correlation coefficients are $r = 0.71$ for correct scores and $r = 0.74$ for incorrect scores. The overall correlation equals $r = 0.91$.

the overall correlation coefficient amounted to 0.91. This is high compared to the ranges that Bouma gave as correspondence between confusion matrices obtained for his individual subjects and the averages of the other nine (correlations from 0.54 to 0.89 for correct scores and from 0.23 from 0.82 for summed incorrect scores). From this point of view, the calculated confusion matrix is closer to the average one than most of Bouma's subjects.

In the original confusion matrix, five cases occur where another than the stimulus letter has the highest naming rate (for the letters s, z, c, l, g; see Fig. 2). Four of these cases are correctly predicted by the theoretical confusion matrix.

One may wonder to what extent the prediction of the confusion matrix of Fig. 14 is determined by the similarity factors and what might be the contribution of the bias factors. Therefore, a theoretical similarity matrix for eccentric-near vision was calculated from contour difference sums applying the same values for the threshold and slope parameters of the psychometric function as used for the prediction of the overall confusion matrix of Fig. 14. It was found that a correlation coefficient of $r = 0.69$ existed between the so-obtained similarity matrix and the one derived from Bouma's results according to eq. 13. This value indicates that the prediction of

the confusion matrix of Fig. 14 is determined for a fair amount by the similarity parameters derived from contour difference sums.

7.9 Discussion

Few attempts have been made to estimate the share of early stages of visual processing on the internal imaging of text, and isolated letters in particular. Ginsburg (1980) is one of the few investigators who calculated its effects on letters by filtering in the frequency domain according to transfer functions based on experimentally determined MTF's, assuming parallel spatial channels and their specific filter properties. The resulting letter responses have in common with our calculated responses that the deformation caused by optical and neural factors can be shown.

It is not really important whether letter responses are represented in the frequency domain or in the space domain, since both transformations uniquely relate the internal space with the outside world. The issue, however, becomes a relevant one if a particular description is chosen in order to mimic the actual visual processing involved. In that case, it seems logical to choose for a description in the space domain, being a more likely candidate for an analogous formalism. This is physiologically and morphologically argued by the fact that "receptive fields are restricted to small local areas and do not cover the entire visual field as would be required" (Cavenagh, 1984). Furthermore it was found psychophysically by means of a selective adaptation paradigm (Perizonius et al., 1985) that spatial frequency channels behave as strictly local units. In that case, Fourier analysis looses its strength, which in the first place lays in a global description of structures over larger areas. Ginsburg's (1980) argument that calculations in the frequency domain are "simpler and provide more insight" is discutable, since this is merely a matter of taste and familiarity with a certain kind of analysis.

If the computed letter responses, as for instance shown in Fig. 5, are compared with the real perceptual images if the letters are viewed under the stimulus conditions of eccentric-near and foveal-distant vision, the resemblance is rather convincing, especially with regard to the contours. This resemblance between theory and percept stimulated further research in the chosen direction.

As a working hypothesis we assumed that a confusion score can be written as the product of letter response similarity, a correction factor and a bias factor (eq. 9). The correction factor is then thought to compensate for the fact that the letter responses are not compared with the real memory item as they should have been, but instead with the letter responses themselves. The bias factors β_i used in the classical interpretation of the choice model then would equal the product of our bias and correction factors. In other words, the classical β_i's would still contain perceptual components.

If the similarities are derived at stimulus level, i.e. if letter similarities are calculated from the letter profiles themselves, the same working hypothesis can be used in principle. In that case, however, the correction factors should not only contain the differences between letter responses and memory responses, but also the result of

optical and neural processing. The "bias" factors, in that case, are probably heavily dependent on perceptual factors. In this view it is not surprising that Keren and Baggen (1981) found a high correlation between bias and similarity factors in the data of Gilmore et al. (1979). In fact, there should be according to our eq. 14.

By using the letter responses as input for the model instead of the letter profiles, one hopes to acquire a more appropriate description of letter confusions.

Similarity and dissimilarity features do exist both in the stimulus letter and in the letter responses. The features will, however, not be identical. By the action of peripheral visual processing, some of them may have disappeared and some of them may be deformed. This is one of the reasons we think that models based on feature lists (Keren and Baggen, 1981; Geyer and Dewald, 1973), always need weighting factors for every feature, which according to the foregoing are not constant but depend on the way in which stimuli are deformed in the visual pathway. By applying suitable visual transformations on the stimuli, it may be hoped for that the features get their weights in a more natural way.

In order to calculate letter response similarities, we tried out several formalisms. The specific choices shown here, might be discussed and could have been extended in several ways. However, the chosen algorithms are not meant as the final solving of the problem but merely as a few illustrative examples of how the problem might be tackled.

If one chooses the procedure followed in this paper for evaluating theoretical similarity measures (eq. 13), one has to accept a rather noisy result. Some factors can be allocated:

- In order to formalise the characteristics of early visual processing, the point spread function of a single subject was taken. In determining the confusion matrices, however, Bouma used ten subjects of which the visual acuity ranged from 1.4 to 2.0. Furthermore, Bouma's subjects viewed the stimuli with both eyes and natural pupils, as opposed to the experimental design for the point spread function where only one eye and an artificial pupil with a diameter of 2 mm was used. As mentioned earlier, also the adaptation levels were not quite the same.

- In order to be able to calculate letter responses, a discrete version of these letters had to be implemented. Although the differences with the original printed letters are not estimated as drastic, some effect of this procedure may be present in the results.

- The estimate $S(i,j)$ used for extrapolating letter similarity measures from experimentally determined confusion matrices is probably biased. This is caused by a pecularity if the choice model is used in practice. As pointed out by Keren and Baggen (1981) approximation of the similarity measure $S(i,j)$ according to eq. 13 leads to an underestimation of the parameters due to the presence of empty cells in the original data.

- According to signal-to-noise criteria, at least on average 5 presentations per cell of the matrix are necessary to obtain confident estimates for confusion proba- bility from experimantal confusion scores (cf. Houtsma, 1983). For Bouma's experimental confusion matrices used in the present analysis, this criterion was fulfilled, be it that the results were averaged over 10 subjects.

It is mainly because of these "noise" factors that the results of the similarity ana- lysis should be watched with a global eye. Such results cannot be expected to be more precise than the uncertainty in the basic quantities. In order to obtain higher order approximations, confusion matrices and point spread functions should better be determined for the same subjects under identical experimental conditions.

As mentioned before, there exists some uncertainty concerning the width that should be chosen for the point spread function which is meant to match the eccentric-near vision condition. The uncertainty in the width parameter amounts to a factor of 2, due to experimental inaccuracy in visual acuity results and also due to variation of acuity between subjects. Therefore, it is of interest to compare the derived similarity measures for the foveal-distant and the eccentric-near conditions (as the differences in these results are due to a factor of 2 in width variation of the point spread function). The correspondenc between the different similarity measures, as shown in the section Results, differ from one measure to another. For contour difference sums, a high correlation of $r = 0.97$ was found between the two conditions, whereas for correlations between spatial responses Pearson's r amounted to 0.88. It can be concluded from these results that the computed similarity measures are not particularly sensitive to width variation of the point spread function.

From the scatter diagrams shown in Results, it is striking that the amount of scatter found for the foveal-distant condition is generally larger than for the eccentric-near condition. A possible cause might be that the point spread function used to simulate the foveal-distant condition is chosen a little too broad. The difference in adapta- tion levels between the recognition and point spread function experiments (5000 Td against 1200 Td) might have attributed to this phenomenon, judged by precise data on visual acuity (Ikeda et al., 1980). Since in this condition the x-height of the letters on the retina is only 2 arc min, the effect of eye movements might also have been a possible noise source.

In performing the calculations, the basic idea was that there might exist a similarity measure, derived from calculated letter responses, which exhibits a monotonic relation with experimentally estimated similarities $S(i,j)$. Some of the measures used can be seen to be more suited than others. Although the correlations between frequency spectra (Fig. 9) show a certain relationship with experimental $S(i,j)$'s, the result is not convincing. Probably, correlation between spatial frequency spectra does not stress the right cues in the internal imaging of letters. In that respect, the correlation between spatial letter responses (Fig. 8) shows a slightly better figure. In our opinion, however, the contour difference sum (Fig. 12) yields the most appropriate measure obtained sofar for quantifying letter similarities. Apparently, the visual system is unable, for the investigated conditions, to discriminate between the inner parts of

letter responses as much as would be depicted by the response difference sum (Fig. 10).

The quantities that were used as illustrations in this paper were not the only ones investigated. Some other measures used will be shortly commented on:

- Since response difference sums (Fig. 10) apparently stress the inner response parts too much, Weber-like functions like $\Delta U(x,y)/U(x,y)$ were tried ($\Delta U(x,y)$ standing for the letter response difference function and $U(x,y)$ for the letter response function of the stimulus letter or of the "memory" letter or of an average of both). Using such a quantity a calculational difficulty exists, since $U(x,y)$ will become zero for a number of (x,y)-values. Therefore, fractions like $\Delta U/(C + U)$, C being a constant, were tried. The results were slightly better than in the case only ΔU-functions are taken (Fig. 10), but worse than for the contours of Fig. 12. This may, however, be caused by the fact that responses were not fitted in any way as was the case for the contours.

- In case of stimulus conditions like eccentric-near and foveal-distant vision the retinal image of the letters becomes very small compared to local resolution of the eye. In view of this, one wonders whether the visual system is able to perform an precise analysis as used for the calculation of contour differences. Therefore, a more global measure was introduced in which contour areas within angles of 45 degrees were averaged, yielding rather global contours containing only eight directions around the origin (x- and y-axis and oblique directions). The results were judged to be equal or slightly better than the ones shown in Fig. 12, implying that the quantity of contour difference sums might be optimised.

From the calculational results given in this paper, it shows that letter response contours play an important role in confusion scores. This conforms Bouma's analysis based on the derivation of upper cue values for stimulus properties (cue value being defined as that fraction of correct and incorrect scores to which perception of the property has contributed). He concludes: "outer parts provide stronger cues than inner parts; in particular height-width quotients, letter extensions, vertical and oblique outer parts and also outer gaps seem to induce corresponding letter responses (in our terminology: letter scores) with a bias towards frequent letters". Apart from the fact that our analysis was based on internal letter response functions instead of the letter profiles themselves, all relevant properties named by Bouma are represented somehow in our contour function concept. Therefore, both analytical results are qualitatively in agreement.

It should be emphasized that the conclusion that letter contours play a dominant role in recognition might not be extended to other stimulus conditions. In fact, it is obvious that recognition is not only based on contours but also on the inner structure of letter responses. In the chosen stimulus paradigms, however, letter responses have apparently been degraded such that information on the inner parts has become largely unavailable for the recognition mechanism. Judging from calculated responses,

this would probably not be the case if letters were presented foveally at a normal reading distance, while the contrast of the letters was diminished in order to increase confusability.

Ackowledgement

References

Atkinson, R.C.; Kinchla, R.A. (1965) A learning model for forced-choice detection experiments. Brit. J. Math. and Stat. Psych. 18, 183-206.

Blommaert, F.J.J. (1980) Letter recognition at low contrasts. IPO Ann. Progr. Rpt. 15, 76-82.

Blommaert, F.J.J.; Heijnen, H.G.M.; Roufs, J.A.J. (1981) Point spread functions, analysis and use. IPO Ann. Progr. Rpt. 16, 61-66.

Blommaert, F.J.J.; Roufs, J.A.J. (1981) The foveal point spread function as a determinant for detail vision. Vis. Res. 21, 1223-1233.

Blommaert, F.J.J.; Heijnen, H.G.M.; Roufs, J.A.J. (forthcoming) Point spread function variation and visibility of details. In preparation

Blommaert, F.J.J.; Timmers, H. (forthcoming) Letter recognition at low contrast levels: effects of letter size. In preparation.

Bracewell, R. (1975) The Fourier Transform and its Applications. McGraw-Hill, New York.

Breitmeyer, B.G.; Ganz, L. (1976) Implications of sustained and transient channels for theories on visual pattern masking, saccadic suppression and information processing. Psychol. Rev. 83, 1-36.

Bouma, H. (1971) Visual recognition of isolated lower-case letters. Vis. Res. 11, 459-474.

Bouwhuis, D.G.; Bouma, H. (1979) Visual word recognition of three-letter words as derived from the recognition of the constituent letters. Perc. & Psychoph. 25, 12-22.

Campbell, F.G.; Robson, J.G. (1968) Application of Fourier analysis to the visibility of gratings. J. Physiol. 197, 551-566.

Cavenagh, P. (1984) Image transforms in the visual system. In: Figural Synthesis. Ed. by Dodwell and Caelli. Erlbaum. Hillsdale, New Jersey.

Garner, W.R.; Haun, F. (1978) Letter identification as a function of perceptual limitation and type of attribute. J. Exp. Psychol.: Human Perc. and Perf. 4, 199-209.

Geyer, L.H.; Dewald, C.G. (1973) Feature list and confusion matrices. Perc. & Psychoph. 14, 471-482.

Gilmore, G.C.; Hersh, A.; Caramazza, A.; Griffin, J. (1979) Multidimensional letter similarity derived from recognition errors. Perc. & Psychoph. 25, 425-431.

Ginsburg, A.P. (1980) Specifying relevant spatial information for image evaluation and display design: an explanation of how we see certain objects. Proc. SID. 21, 219-227.

Houtsma, A.J.M. (1983) Estimation of mutual information from limited experimental data. J. Acoust. Soc. Am. 74, 1626-1629.

Ikeda, K.; Noda, K.; Yamaguchi, S. (1981) A relation between adaptation luminance and visual acuity for the Landolt ring under the uniform background. J. Light & Vis. Env. 4, 22-31.

Keren, G.; Baggen, S. (1981) Recognition models of alphanumeric characters. Perc. & Psychoph. 29, 234-246.

Koenderink, J.J.; van Doorn, A.J. (1978a) Visual detection of spatial contrast; influence of location in the visual field, target extent and illumination level. Biol. Cyb. 30, 157-167.

Kuffler, S. (1953) Discharge patterns and functional organization of the mammalian retina. J. Neurophysiol. 16, 37-68. J. Opt. Soc. Am. 68, 845-849.

Le Grand, Y. (1953) Optique Physiologique III: L'Espace Visuel. "Revue Optique". Paris

Le Grand, Y. (1967) Form and Space Vision. Indiana Univ. Press. Bloomington.

Luce, R.D. (1959) Individual Choice Behaviour. New York: Wiley.

Luce, R.D. (1963) Detection and recognition. In: Handbook of Mathematical Psychology. Ed. by Luce, Bush and Galanter. J. Wiley and Sons. New York and London.

Maddox, M.E. (1980) Two-dimensional spatial frequency content and confusion among dot-matrix characters. Proc. SID. 21, 31-39.

Nakane, Y.; Ito, K. (1978) Study on standard visual acuity for better seeing in lighting design. J. Light & Vis. Env. 2, 38-44.

Nes, F.L. van; Jacobs, J.C. (1981) The effect of contrast on letter and word recognition IPO Ann Progr. Rpt. 16, 74-83.

Norren, D. van (1981) Visual acuity in a condition of traffic sign viewing: the effects of luminance changes. Am. J. Optom. & Phys. Opt. 58, 699-705.

Perizonius, E.; Schill, W.; Geiger, H.; Rohler, R. (1985) Evidence on the local character of spatial frequency channels in the human visual system. Vis. Res. 25, 1233-1240.

Quick, R.F. (1974) A vector-magnitude model of contrast detection. Kybernetik. 16, 65-67.

Sloan, L.L. (1968) The photopic acuity-luminance function with special reference to parafoveal vision. Vis. Res. 8, 901-911.

Townsend, J. (1971) Theoretical analysis of an alphabetic confusion matrix. Perc. & Psychoph. 9, 40-50.

Wiesel, T.N. (1960) Receptive fields of ganglion cells in the cat's retina. J. Physiol. 153, 583-594.

Wilson, H.R.; Bergen; J.R. (1979) A four mechanism model for threshold spatial vision. Vis. Res. 19, 19-32.

chapter 8

Letter recognition at low contrast levels: effects of letter size[1]

Frans J.J. Blommaert
Han Timmers

Abstract

With the aim of obtaining a better understanding of the role that early stages of visual processing play in letter recognition, contrast variation was used in an attempt to measure recognition thresholds for lower case alphabet letters. Thus frequency-of-recognition curves were measured for alphabets with varying x-heights (x-height being defined here as the vertical visual angle that the letter 'x' subtends to the eye). Since variation of the adaptational state of the eye changes the characteristics of primary visual processing in a quantifiable way, recognition thresholds were measured at a high (150 cd/m^2) and a low (0.9 cd/m^2) adaptation level. It was found that thresholds decrease if letter size increases in a way comparable with data on visual acuity in the literature. At the lower adaptation level recognition thresholds became overall higher, which is also in accordance with visual acuity data. Furthermore, it was found that the slopes of the frequency-of-recognition curves for alphabets as a function of log contrast decrease with letter size. It will be argued that this is mainly caused by an increasing dispersion of internal representations of individual letters on the internal psychological scale as letter size decreases. An algorithm will be presented which is designed to quantify this dispersion in terms of dispersion of recognition thresholds for suitably chosen alphabet subsets.

8.1 Introduction

In order to study legibility of letters and words, stimulus conditions are often chosen such that recognition scores are the dependent variables. In that case, the scores should be limited between 0% and 100%. A number of methods can be used to obtain this stimulus condition. Stimuli may be presented briefly or they may be masked by another stimulus (for a review see Breitmeyer and Ganz, 1976); they may be presented at an eccentric position (Bouma, 1971) or at such a long viewing distance that the stimulus details subtend viewing angles comparable with the spatial resolution of the eye (Bouma, 1971, 1973; Bouwhuis and Bouma, 1979). Another - rather uncommon - way to reduce legibility of letters and words is to diminish the contrast of the symbols with respect to background (Spencer, 1976; Ginsburg, 1978; Timmers, 1978; van Nes and Jacobs, 1981). A disadvantage of diminishing legibility in such a way is that the situation of reduced contrast is relatively rare in normal texts, except in cases where VDU's are improperly used or contrast of printed text is diminished by gloss.

---

[1] Submitted to Perception.

On the other hand, contrast variation is the most common tool in investigating early stages of visual processing from a psychophysical point of view. Measurements of contrast thresholds and brightness matches are used to find general rules that govern these processes.

These investigations have recently led to a number of models that predict thresholds for variously shaped stimuli rather well (Koenderink and van Doorn, 1978, 1982; Wilson and Bergen, 1979; Roufs and Bouma, 1980; Blommaert and Roufs, 1981; Blommaert, Heijnen and Roufs, 1981). Therefore, an advantage of contrast reduction as a tool for studying letter and word recognition could be that its results can be linked up with the steadily increasing knowledge on properties of the peripheral visual pathway.

This paper describes experiments in which controlled contrast reduction was used to determine 50% recognition thresholds for letters as a function of letter size. These thresholds may be treated as a measure of letter recognition as a function of letter size, and can be compared with characteristics of peripheral vision, like acuity.

Recognition thresholds for alphabets with varying letter size were determined at two adaptation levels, viz. 150 cd/m^2 and 0.9 cd/m^2, since variation of background luminance changes the processing of the peripheral system in a well-known way (Blackwell, 1946; Nakane and Ito, 1978). In the course of the investigation it became clear that the measured thresholds have to be evaluated carefully, since the stimulus set (the alphabet) tends to become less homogeneous with respect to legibility as letter size decreases. That is to say, the legibility of different letters does not change in the same ratio when the x-height is varied. By postulating that letter recognition at low contrasts is mainly determined by detail detection, this effect can be quantified by extrapolating thresholds for lowest and highest scoring alphabet subsets.

8.2 Apparatus and procedure

Apparatus

Using a two-channel tachistoscope, contrast was controlled as depicted in Fig. 1. During the stimulus presentation, the stimulus field (which contained the letter) and the adaptation field (which was blank except for suitably chosen fixation marks) had different luminances L_s and L_a. Care was taken that the summed luminances $L_s + L_a$ were always equal to the adaptation level L_o.

Before and after stimulus presentation the luminance of the adapting field was automatically fixed at L_o, while the luminance of the stimulus field was fixed at zero. In this way, a stimulus presentation was realised without a visible change in the surround or background luminance.

In the following, we take as a measure of contrast a commonly used psychophysical definition:

$$C = \frac{L_w - L_b}{L_w},\qquad(1)$$

where L_w is the luminance of the white paper and L_b that of the black symbol.
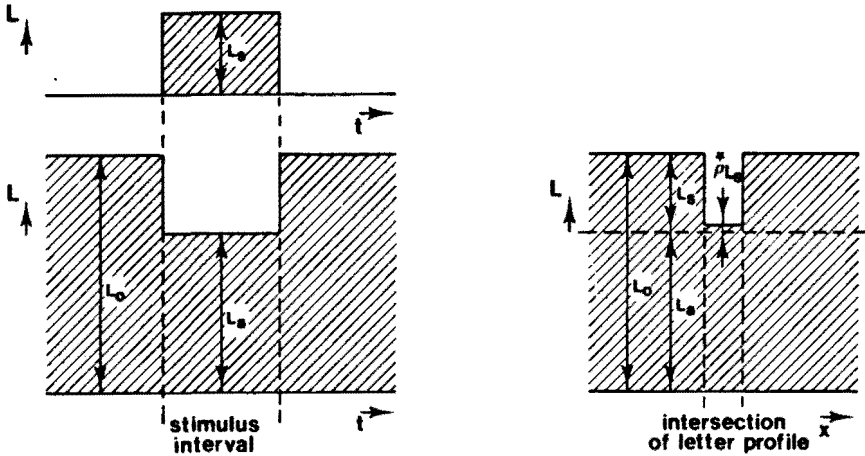
Figure 1: Top left: luminance time function of stimulus field. Bottom left: luminance time function of adapting field. Right: spatial intersection of luminance profile during stimulus presentation.

It is easily verified that the contrast of the symbols is continuously controllable between maximum ($C \approx 0.9$) and minimum ($C = 0$) by changing the ratio of L_s and L_a.

Using eq. 1 we find:

$$C = \frac{L_0 - (L_a + \rho^* L_s)}{L_0} = \left(1 - \rho^*\right)\frac{L_s}{L_0}. \qquad (2)$$

where ρ^* is the reflection coefficient of the black letter relative to that of the white paper.

Before the experiments were started, this relative reflection coefficient was determined by measuring the luminances of symbol and surround using a Pritchard photometer (1970-PR). As a normal printed letter is small, it is usually fairly difficult to measure its luminance. Since we worked with "press on" letters, however, this could easily be done as we had large symbols available. If the respective luminances are measured for various surround luminances, a constant ratio has to be found between L_b and L_w which equals ρ^*.

From these measurements we estimated the reflection coefficient to be $\rho^* = 0.08$. This means that the highest obtainable contrast, according to eq. 2, is 0.92.

The letter type we used in the experiment was "Eurostile bold extended", which is shown in Fig. 2. This letter type has the advantage that it is available in a large variety of sizes, which served our purpose as indicated in the introduction.

Vision was binocular, with natural pupils. The distance of the subject's eyes to the stimulus was 0.57 m except for a set of x-heights 4.6, 5.5, 6.8, and 8.3 min of arc, where the distance was doubled, and another set of 3.6, 4.5, and 5.5 min of arc, where the distance was tripled. By using an overlap, a possible effect of changing the

a   b   c   d   e
f   g   h   i   j
k   l   m   n   o
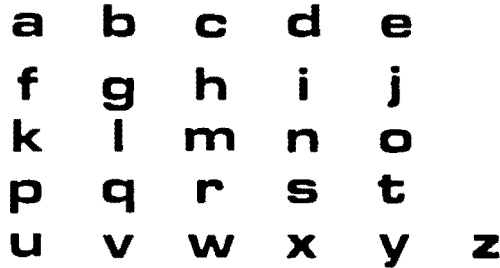p   q   r   s   t
u   v   w   x   y   z

Figure 2: The letter type 'Eurostile bold extended' as used in the experiments ('press on letters').

optical transfer function of the eye by changing its accommodation level and pupil diameter could be estimated.

In order that the stimuli should be presented foveally, subjects were instructed to fixate at the space between the halves of a thin line, showing a gap at the intended point of fixation where the stimuli were presented. This excludes serious masking effects.

Procedure

The recognition[1] threshold of a letter can be determined by hiding it in a suitably chosen stimulus set and measuring a frequency-of-recognition curve for this letter.

This stimulus set should have the property of being rather homogeneous with respect to the internal representation of its items, and in order to minimise the effect of chance scores it should not be too small. As mentioned before, our aim is to measure an average recognition threshold of all alphabet letters with a certain x-height in order to acquire a legibility measure of an alphabet with a certain letter size. This may be achieved by measuring recognition thresholds of individual letters and averaging over them.

Apart from the fact that this procedure would be very time consuming, it is not easy to compose adequate stimulus sets for each letter. We therefore decided to measure frequency-of-recognition curves for total alphabets, thus using the alphabet as the stimulus set. This choice has certain advantages:

- The stimulus set is large enough ($n = 26$) to minimise the effect of chance scores.

- What is measured is exactly what we aimed at, viz. an average recognition threshold of the total alphabet at a certain x-height.

- It is to be expected that, at a certain x-height, the alphabet will be fairly homogeneous with respect to the internal representation of the constituent

---

[1]Throughout this paper we will use the term "recognition" although one might prefer the term "identification" instead, since a closed and well-known stimulus set is used.

letters in the internal space. It is known from the literature (Ginsburg. 1978) that recognition thresholds for alphabets can indeed be determined in such a way.

Before determining the correct recognition score of an alphabet with a certain x-height, a chosen contrast was set. Luminances of both stimulus field and adapting field could be instantaneously varied and measured as output voltages of two photodiodes, one placed in the stimulus field and the other in the adapting field. Before the start of every session the photocell voltages were calibrated against the Pritchard photometer, so that the experimenter was always informed of the different luminance levels and could change the contrast immediately.

The letters were presented for 64 ms in a quasi-random order. At suitably chosen contrasts the correct recognition score of one total alphabet was determined. For every x-height. equidistant contrasts were chosen on a logarithmic scale (mostly 0.1 log units, sometimes 0.2 log units). The stimulus presentation was self-released. For every x-height, the contrasts used were quasi-randomly chosen. Before measuring at a certain x-height. 10 letters of this size were presented at an intermediate contrast to give the subject some indication of the x-height to be used. After that a complete alphabet was presented, to determine the correct score. After measuring the frequency-of-recognition curve at a certain x-height. another letter size was randomly chosen and the same procedure was followed.

Two adaptation levels were used 150 cd/m^2 and 0.9 cd/m^2. Two subjects (the authors) participated in the experiments. Both have normal acuity (1.5 as measured with a Landolt C-chart at $L = 90\ cd/m^2$) at a distance of 5 m and HT was medium protanomalous and FB medium deuteranomalous.
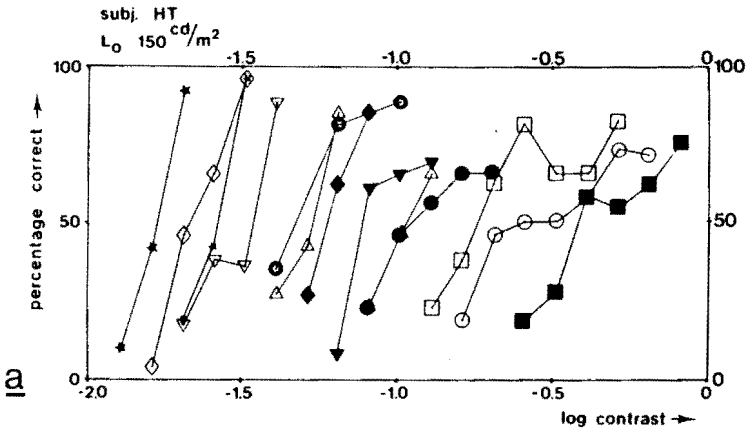
8.3 Experimental results

### Frequency-of-recognition curves and 50% recognition thresholds

Percentages of correct recognition scores on single alphabets are plotted against contrast on a semi-logarithmic scale for both subjects in Figs. 3a and 3b at the adaptation level of 150 cd/m^2.

From the results it is clear that, as letter size increases, decreasingly less contrast is needed to obtain the same percentage correct. Thus, recognition thresholds decrease monotonically with increasing x-height of the alphabet. Since every data point consists of only one alphabet, the frequency-of-recognition curves are rather noisy. From these results, 50% recognition thresholds were derived by fitting regression lines through the data points. They are shown as crosses in Fig. 8. Corresponding thresholds at the lower adaptation level are also shown in this figure. They are derived from figures 3c and 3d, where the recognition scores are shown for the low adaptation level of 0.9 cd/m^2.

A striking feature to be observed in Fig. 3 is the tendency of the slopes of the frequency-of-recognition curves to decrease as letter size decreases. Since the curves are plotted on a semi-logarithmic scale this would mean, in terms of a Thurstonian
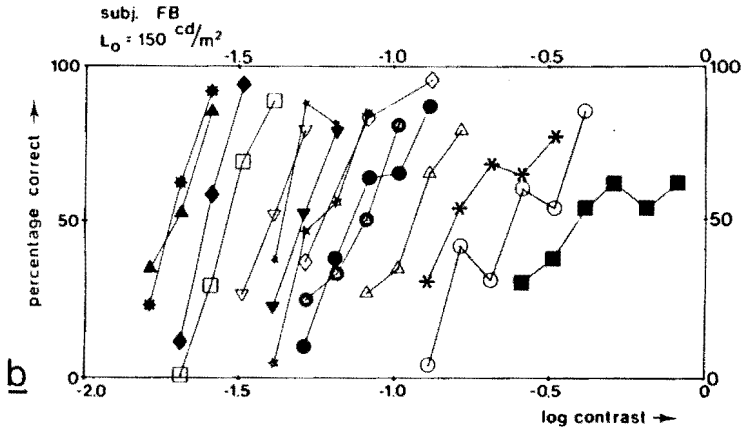
Figure 3: Percentage of correct scores as a function of log contrast. Each data point stands for the correct score on a single alphabet. The parameter is letter size, indicated by the x-height and expressed in min of arc. The straight lines connect correct scores for alphabets with the same size. Fig. a and b give results for the 150 cd/m^2 condition.
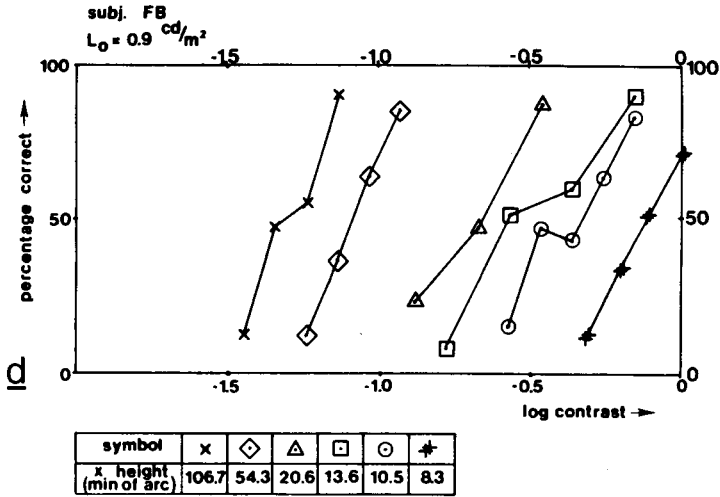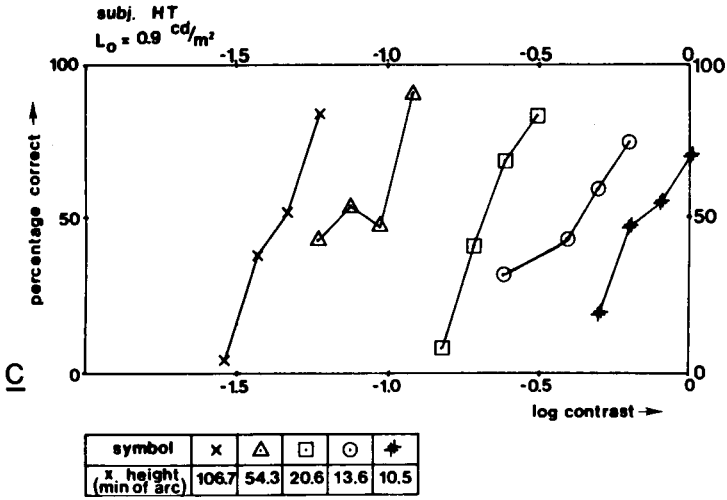
Figure 3: Continuation of page 158. Fig. c and d show the results for $L = 0.9 \ cd/m^2$.

model, that the width of the distribution function mapping the image of the alphabet on the psychological scale, divided by the mean, increases if letter size decreases (Thurstone, 1959; Bock and Jones, 1968).

As a rough interpretation, the frequency-of-recognition curve of an alphabet can be seen as a sort of average of individual recognition curves of the 26 letters constituting the alphabet. There may be several reasons why this curve becomes shallower:

- The slopes of all or some subset of individual frequency-of-recognition curves becomes shallower with decreasing letter size.

- The frequency-of-recognition curves for individual letters disperse on the log contrast axis for small sizes.

- For some reason the chance scores increase, which would lead to an enlarged number of false responses at relatively low contrasts for that particular size.

These possibilities may all cause the composite frequency-of-recognition curve to become skewer. This would imply that the composite curve cannot be approximated by a cumulative Gaussian distribution, the derived recognition threshold being a questionable measure of the average recognition threshold of the 26 letters constituting the alphabet.

In order to disentangle the experimental facts, we follow an indirect approach. This approach emerged from the perceptual phenomena observed when performing the experiments. Both subjects noted that, for large x-heights, recognition of letters was probably closely related to detection of letter details. They observed that, at these letter sizes, they either saw nothing at all, or if they did detect anything, they were fairly definite about their answer being correct. Therefore, the slope of the psychometric function for contrast detection (of subject FB) was compared with the slope of the frequency-of-recognition curves at large letter sizes. They turned out to be very close (both about 0.17 log units per 100%).

By postulating that letter recognition at low contrasts is mainly determined by detail detection, 50% recognition thresholds of well or poorly recognizable letters can be found by extrapolation. This will be dealt with in the next section.

Evaluation of results

If contrast detection of details governs the recognition of low-contrast letters, an interesting point is that for contrast detection the quantity σ/m is an invariant. Here, σ is the width of the distribution function mapping the image of the stimulus on the psychological scale and m is its threshold value. It can be proved that this means that the slope of the psychometric function as a function of log contrast is constant (Roufs, 1974). This constancy is important since it is difficult to tell which details of a particular letter are important for its recognition and thus what detail size is relevant.

To demonstrate this, slopes of $p(logC)$ curves for contrast detection in the 50% point are plotted against the diameter of discs in Fig. 4 (for details see Blommaert and Roufs, 1981).
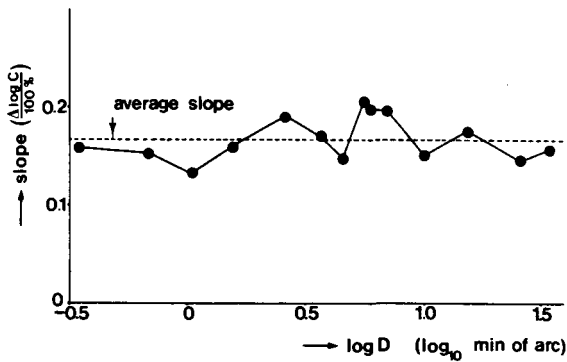

Figure 4: Slopes of psychometric functions for contrast detection of discs as a function of diameter. The data are averages of four threshold functions. The dashed line is the mean of all these slopes.

Apart from the inevitable sample spread, the slope seems to be constant for all disc sizes used. In order to carry out quantitative calculations, we have to choose an analytical expression for the total shape of the psychometric function. The cumulative log normal distribution has the practical advantage for us that the calculations become relatively simple. Although this choice can certainly be discussed, we will use this shape throughout this paper. The differences in the available distributions are anyhow too subtle to influence the rough treatment of data given in this paper.

Now in order to make an analysis possible we assume that:

Frequency-of-recognition curves of all letters and all letter sizes have the same shape as the psychometric function for contrast detection.

If the assumption is valid, the experimental data can be treated differently. For determining the recognition score for an individual letter there is only one free parameter, viz. the position of the frequency-of-recognition curve on the contrast axis. Thus, if at a certain contrast the recognition score of a letter is known, its recognition threshold can be calculated provided the score is not 0% or 100%.

However, in practice we do not only have a score at one single contrast but a number of scores at different equally spaced contrast values (see Fig. 5a). We therefore can now look for a parameter that fixes the frequency-of-recognition curve more precisely. Such a parameter is for example the average correct score in a certain contrast interval, using all correct scores collected for this letter. If the average recognition score for a certain letter were known, the 50% recognition point of this letter could be calculated unambiguously, provided the used contrasts are taken into account.

In practice, we calculated tables which related every possible average correct score to the position of the frequency-of-recognition curve on the contrast axis. These tables
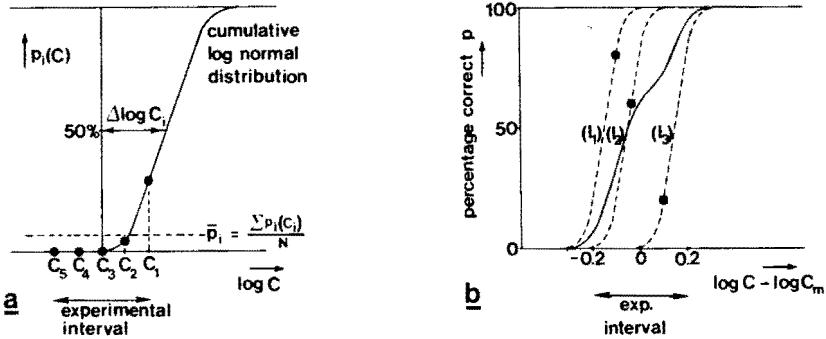
Figure 5: 5a. Illustration of the algorithm used to extrapolate the 50% recognition thresholds of a letter l_i from its average correct score p_i within the experimental interval. If the average correct score is known, the position of the psychometric function is uniquely depicted. Hence the 50% recognition point can be determined. 5b. Hypothetical example of using the extrapolation algorithm. In an experimental interval of 0.4 log units, average correct scores for the letters l_1, l_2 and l_3 are respectively 80%, 60% and 20% (denoted by •). It can then be calculated that the shifts of the individual 50% recognition points relative to the middle of the experimental interval are -0.15 for l_1, -0.05 for l_2 and +0.15 for l_3, respectively. The dashed lines are the three individual threshold recognition curves; they are all templates of the Ψ-function for contrast detection. The continuous curve is the average recognition curve for the alphabet subset (l_1, l_2, l_3). Note that the slope is flattened by the position difference of the curves for individual letters.

enabled us to estimate the 50% recognition point of alphabets, or alphabet subsets. As an example, it can be calculated how a composite frequency-of-recognition curve emerges from three individual ones as is illustrated in Fig. 5b. The frequency-of-recognition curves for the letters l_1, l_2 and l_3 are by definition all templates of the psychometric curve for contrast detection. In this example the average correct scores are respectively 80% for l_1, 60% for l_2 and 20% for l_3. Our tables then tell us that their 50% points with respect to the middle of the experimental interval are -0.15, -0.05 and 0.15 $logC$ respectively. The solid line then represents the composite frequency-of-recognition curve for the alphabet subset (l_1, l_2, l_3) calculated as the average of the three individual ones (note that this curve is more shallow than the individual ones).

The assumption we put forward in this section is rather hazardous, therefore we wanted to test its validity. This is effectively done by using the same method to describe the shapes of the frequency-of-recognition curves for the extreme alphabet sizes, i.e. the largest and the smallest x-heights used in the experiment.

The results are shown in Fig. 6. The dashed curves are derived by calculating the

positions of the frequency-of-recognition curves for individual letters based on their average correct scores. The scores were averaged afterwards over the 26 letters constituting the alphabet. As can be seen from Fig. 6, the predictions are fairly good for the largest letter sizes, as expected. By good fortune, the prediction is equally good for HT as for FB, bearing in mind that the derivations are based on the slope of the psychometric function for FB.

For the smallest letter sizes there is some overestimation of the calculated slopes compared to the actual experimental findings.
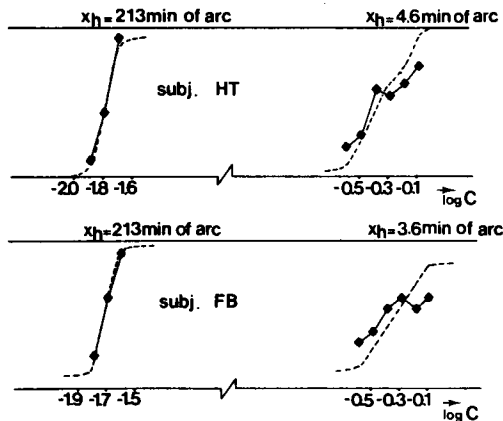

Figure 6: Illustration of the use of the algorithm to describe the shape of frequency-of-recognition curves for total alphabets. The data (diamonds) for smallest and largest x-heights are taken from Fig. 4. The dashed lines represent the predicted frequency-of-recognition curves. The prediction shows a fair fit for the largest x-heights, whereas for the smallest letter sizes it tends to be somewhat too steep.

A plausible explanation for this second order discrepancy between calculated and experimental results may be that for small letter sizes contour similarities (letter cues: Bouma, 1971) play an important role in the recognition scores. These cues enhance the possibility to choose an almost illegible letter at low contrast correctly since the class of possible letters to identify is diminished by certain contour similarities, as for instance the round envelope of the letters e, o, c or the elongation of the subset j, i, l. This effect leads to a smaller slope than calculated on the basis of just detail detection and letter identification, since the effect enhances the probability for correct scores at low contrasts due to the choice of letters from a contour-selected-subset.

For an accurate description of the shape of the frequency-of-recognition curves for small x-heights, we think it is necessary to introduce a second-order term to account for the effect of contour similarities. However,it is not our aim to give a full description of the frequency-of-recognition curves but just to determine upper and lower boundaries for the 50% thresholds of recognition of alphabets with varying x-heights.

For this purpose, the proposed formalism seems a reasonable approximation.

Fig. 7 gives some examples in which this scheme is tested in practice. Frequency-of-recognition curves for the five best and five worst scoring letters are calculated in the described way and plotted together with the correct scores. For the worst recognizable letters there seems to be a small bias, presumably caused by similar letter contours and hence leading to 50% threshold calculations which are a bit too low.

This paradigm seems able to explain the upper and lower boundaries for the 50% recognition thresholds of alphabets at various x-heights as a function of contrast.

As a result, Fig. 8 shows the extrapolations on a double logarithmic scale. This figure also presents the results for the adaptation level of 0.9 cd/m^2. The same analysis was applied here, using the same slope for the psychometric function. In doing so, it is assumed that this slope does not change with adaptation level.

This assumption is certainly valid for flashes, where a constant slope is found if the adaptation level is varied over as much as 5 log units (Roufs, 1974). It also holds for quasi-static with varying size, judged by σ/m constancy (Blackwell, 1963). When the 50% boundaries were determined for letter recognition at the 0.9 cd/m^2 level, the results fitted the experimental data as well as the ones shown in Fig. 8 for 150 cd/m^2. The crosses are the result of linear regression fits to the rough data from which 50% thresholds of the average alphabets were derived.

8.4 Discussion

The early stages of visual processing may be modelled in the space domain (see for instance Koenderink and van Doorn, 1978) as well as in the frequency domain (Campbell and Robson, 1968). The approaches are not essentially different since spatial frequency is inversely proportional to length. For the following reasons we prefer to work with spatial coordinates:

- Frequency analysis is not quite a realistic description of visual processing since receptive fields are restricted to small local areas and do not cover the entire visual field as would be required (Cavenagh, 1985).

- Detail processing, as in the case of isolated letter stimuli as we used, is confined to a strictly local area of the fovea; in which case a local analysis in spatial coordinates is more suitable.

Global results, however, should be similar irrespective of the kind of analysis applied. Ginsburg (1980) measured detection and recognition thresholds for Snellen letters as a function of letter size. His experimental results on letter recognition are in fair agreement with our results (cf. Fig. 8) at the high adaptation level, although he does not mention any letter dispersion, being the threshold range of optotypes at small sizes.

His interpretation is in terms of frequency analysis. From the experimental results he concluded that 2.36 cycles/letter higher spatial frequency was needed to identify
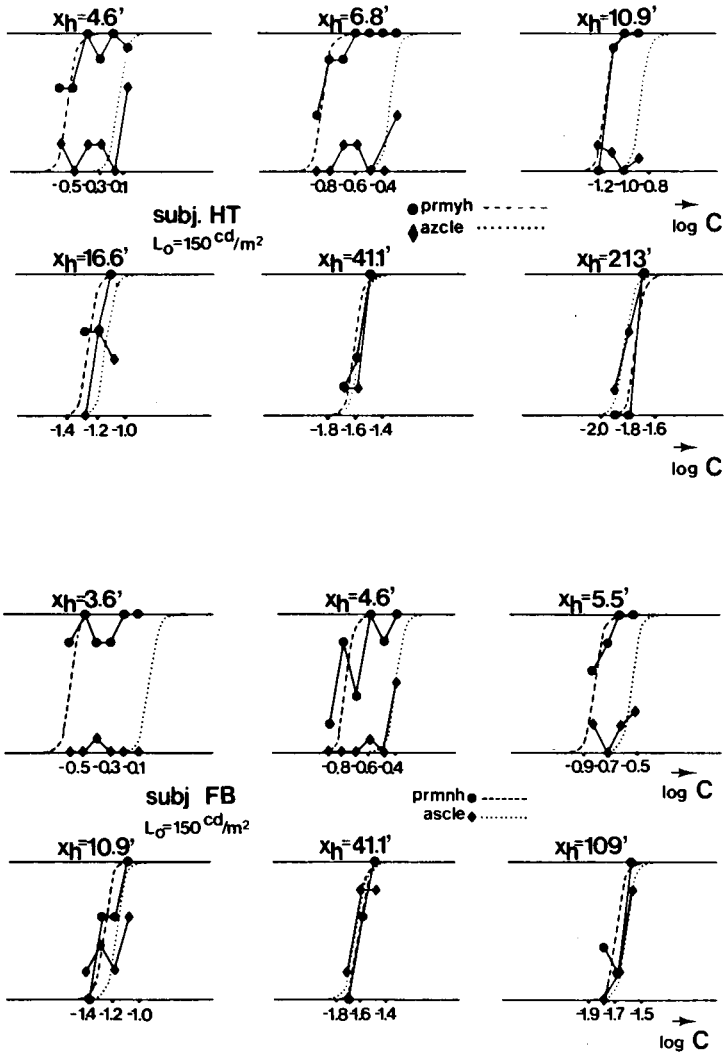
Figure 7: Some examples of the way in which 50% recognition thresholds were determined for best and worst alphabet subsets. The dashed lines represent the Ψ-curves for the best subset; the dotted lines for the worst subset. The black symbols indicate the actual correct scores for the respective subsets. Note that these subsets have almost the same composition for both subjects.

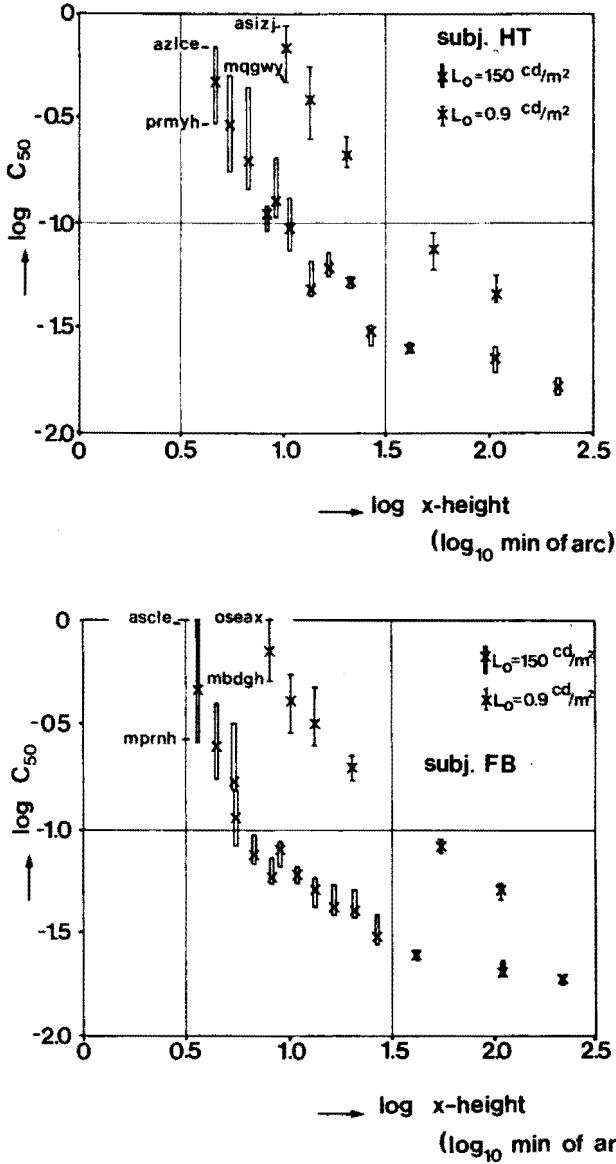Figure 8: 50% Recognition thresholds as a function of x-height on a double logarithmic scale, measured at two adaptation levels. The bar and stripe endings indicate the extrapolated thresholds for worst and best subsets. These subsets are seen to differ with adaptation level. The symbols indicate 50% recognition thresholds for total alphabets calculated by fitting regression lines to the rough data of Fig. 4.

than to detect a letter. This is about equivalent with an experimental result that shows that a letter is recognized if details of a letter are detected that are on average 2.36 times smaller than the global letter dimension. It should furthermore be noticed that the bandwidth of 2.36 cycles/letter is a stimulus property. Other letter types would yield a different bandwidth if the letter details are in a different proportion to the global letter dimension.

It may be asked how the specific time course of a stimulus presentation influences the experimental results. The stimulus duration we used in the recognition task was 64 ms, an onset time which is sufficiently long to avoid large effects of transient phenomena. However, these factors cannot be ruled out altogether since it is well known that the visual system is very sensitive to abrupt changes in luminance like the on- and offset of a stimulus (Roufs and Blommaert, 1981). If these effects are present in the experimental results, they are reflected in an increased sensitivity for larger letters (see for instance Wilson and Bergen, 1979).

In order to eliminate these effects altogether, one should arrange a stimulus presentation in which the onset and offset of the stimulus are gradual functions of time. Since the observations made in this paper are fairly global, it seems not necessary to take into account second order effects caused by temporal factors.

We would like to emphasize that the formalism we put forward in this paper for describing frequency-of-recognition curves for individual letters does not imply that we think that letter recognition at low contrasts is determined by *letter detection*. Rather, we tend to think that letter recognition at low contrasts is mainly determined by *detail detection*. Or, in other words, a letter may be recognized correctly if relevant parts of it (details) are detected. These details will then activate a number of letters from which the most probable one is chosen and reported.

A crucial assumption in developing the formalism for extrapolating 50% recognition thresholds is that the slope of the psychometric function for contrast detection on a semi log scale does not depend on stimulus size. From Fig. 4 it can be concluded that, within experimental error, this condition is satisfied. A far more extensive study on this matter was performed by Blackwell (1965). He measured psychometric functions under a large diversion of stimulus conditions, including a variety of disc sizes and adaptation levels, and fitted cumulative normal distributions to them. He concluded that within reasonable limits the ratio σ/m was constant for all stimulus conditions investigated (σ being the standard deviation of the distribution and m the 50% threshold; see Roufs, 1974). This implies that, on a semi logarithmic scale, slopes of psychometric functions were found to be constant.

From Fig. 6, where frequency-of-recognition curves are calculated for complete alphabets, it can be observed that the calculated curves for subject HT fit the data as well as those for subject FB, although both calculations are based on the slope of the psychometric function for contrast detection of FB. This leads us to the question of how the slopes of psychometric functions vary between subjects. Both Blackwell (1963) and Roufs (1974) present data on this matter. From these it can be seen that slopes may vary a limited amount between subjects. Thus agreement of slopes

between subjects is not remarkable after all. The variety of slopes is greater if experiments are performed under different experimental conditions. Both Roufs (1974) and our data show psychometric functions that are steeper than the ones reported on by Blackwell (1963) and Sachs et al. (1971). This may be due to the circumstance that both Blackwell and Sachs et al. averaged over psychometric functions obtained in different sessions. In this case, relatively slow sensitivity shifts may tend to decrease the slope of the composite psychometric functions.

From the test of the extrapolation formalism by trying to describe frequency-of-recognition curves of total alphabets as shown in 6, it is clear that the overall description is excellent for the largest letter sizes but gives rise to some deviation for the smallest sizes. As was suggested in the section concerned, this deviation is probably due to response bias caused by letter similarities. Experimental support for this notion comes from confusion matrices for lower case alphabets measured under similar stimulus conditions (for instance long viewing distance: Bouma, 1971). They show that letter confusions are more frequent if letters have increasingly common details and/or contours.

In view of the shortcomings of our extrapolation formalism it could be worthwhile measuring recognition thresholds for individual letters or suitably chosen subsets of letters. For this purpose, an approach adopted by Lupker (1979) looks promising; he used a stimulus set containing for uppercase letters (T, X, L, V) and eight meaningless configurations sharing more or less detail and/or contours with these letters. Such a stimulus set could very well act as a suitable means to determine recognition thresholds of individual letters or in general to investigate the role of detail visibility in letter recognition more thoroughly.

From Fig. 8 it can be concluded that contrast thresholds for recognition of alphabets decrease monotonically if letter size decreases. In terms of detail vision, such a dependence can be explained from data on visual acuity measured with Landolt rings (see for instance Nakane and Ito, 1979).

Also the finding that at the low adaptation level letter recognition is 3 to 4 times worse than at the high adaptation level is in quantitative agreement with these acuity findings. The parallel outcome of the letter recognition data presented and visual acuity data in the literature stresses the practical value of visual acuity as a yardstick for judging reading situation designs. Visual acuity decreases very rapidly as a function of retinal eccentricity (Le Grand, 1967). Thus one should expect the same tendency to hold for correct letter recognition. This is confirmed by data from Bouma (1971) and van Nes and Jacobs (1981), who measured recognition scores as a function of the position of the letters on the horizontal meridian through the fovea.

A striking feature of Fig. 8 is that the thresholds of different letters indeed show quite a lot of scatter if x-height decreases. A global indication of this alphabet divergence is also depicted by the correct scores on the main diagonal of confusion matrices measured under similar experimental conditions. The scatter of correct scores for individual letters in a confusion matrix can be interpreted as a measure of the divergence of letters with that particular font and size and measured under those

particular conditions. This is reflected in literature data in that they do not agree very much (see for instance Townsend, 1971; Bouma, 1971; Geyer, 1977).

From these investigators, Townsend was the only one who used upper case letters. From his results one might conclude that that the amount of scatter is somewhat less for upper case than for lower case letters. This may be due to the fact that upper case letters all have the same height, whereas for lower case letters, ascenders and descenders often have a correct score which is higher than those for short letters (Bouma, 1971), which leads to an enhanced alphabet inhomogeneity.

In order to measure visual acuity, still often Snellen letters are used. Considering the amount of scatter found for small letter sizes, this seems a rather inaccurate way to measure visual acuity.

From our data, it is difficult to forecast what might be the effect of contrast diminishing on word recognition or the reading process in general. At low contrasts, word recognition will evidently be impaired. Van Nes and Jacobs (1981), who measured recognition scores for letters and words as a function of contrast and retinal eccentricity, conclude that contrast diminishing leads to a shrinking of the horizontal visual reading field and thus to a slowing down of the normal reading process.

From our results it can be concluded that it is not advisable to use low adaptation levels for reading situations, which of course was suggested long ago by visual acuity data. Therefore too, it might be expected that black letters on a white background would yield a more efficient information transfer than white letters on a dark background. This would especially be true for visual display units, where reading is already often hampered by a bad type font and unsharpness of letter edges. This expectation is confirmed by data from Bauer and Cavonius (1980).

Another situation in which contrast of letters is diminished is where there is an effect of gloss. This may occur when VDU's are improperly installed or when text printed on glossy paper is read with unfavourable lightning conditions.

We conclude from the results of our experiments and analysis that contrast detection of letter details plays a dominant role in the recognition of letters at low contrasts. It therefore seems fruitful to attempt to relate contrast detection models to the letter recognition process. The first steps in that direction have been taken by Blommaert, Heijnen and Roufs (1981) and Blommaert (forthcoming).

Acknowledgement

References

Bauer, D.; Cavonius, C.R. (1980) Improving the legibility of visual display terminals through contrast reversal. In Ergonomic Aspects of Visual Display Terminals.

Ed by E. Grandjean and E. Vigliani. London: Taylor and Francis (1980) pp. 137-142. Proc. of Int. Workshop. Milan 1980.

Blackwell, H.R. (1946) Contrast thresholds of the human eye. J. Opt. Soc. Am. 36, 624-643.

Blackwell, H.R. (1963) Neural theories of simple visual discriminations. J. Opt. Soc. Am. 53, 129-160.

Blommaert, F.J.J.; Roufs, J.A.J. (1981) The foveal point spread function as a determinant for detail vision. Vis Res. 21, 1223-1233.

Blommaert, F.J.J.; Heynen, H.G.M.; Roufs, J.A.J. (1981) Characterization of contrast transfer and prediction of contrast thresholds. Foundation of Biophysics (Netherlands) Ann. Report 80-81, p.200-208.

Blommaert, F.J.J. (forthcoming) On estimating optical and neural imaging factors in letter confusions. In preparation.

Bock, R.D.; Jones, L.V. (1968) The Measurement and Prediction of Judgement and Choice. Holden-Day San Francisco, Cambridge, London, Amsterdam.

Bouwhuis, D.G.; Bouma, H. (1970) Visual word recognition of three-letter words as derived from the recognition of the constituent letters. Vis. Res. 11, 459-474.

Bouma, H. (1970) Visual recognition of isolated lower-case letters. Vis. Res. 11, 459-474.

Bouma, H. (1973) Visual interference in the parafoveal recognition of initial and final letters of words. Vis. Res. 13, 767-782.

Breitmeyer, B.G.; Ganz, L. 1976) Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression and information processing. Psych. Rev. 83, 1-36.

Campbell, F.W.; Robson, J.G. (1968) Application of Fourier analysis to the visibility of gratings. J. Physiol. Lond. 197, 551-566.

Cavanagh, P. (1985) Image Transforms in the Visual System. In: Figural Synthesis. By Dodwell and Caelli Eds. Erlbaum.

Geyer, L.H. (1977) Recognition and confusion of the lower case alphabet. Perc. Psychoph. 22, 487-490.

Ginsburg, A. P. (1978) Relationship between the detection and Recognition of Snellen letters and contrast sensitivity for the normal and abnormal visual system. J. Opt. Soc. Am. 68, 1455.

Ginsburg, A.P. (1980) Specifying relevant spatial information for image evaluation and display-design: an explanation of how we see certain objects. Proc. S.I.D. 21/3, 219-227.

Koenderink, J.J.; Doorn, A.J. van (1978) Visual detection of spatial contrast: influence of location in the visual field, target extent and illumination level. Biol. Cyb. 30, 157-167.

Koenderink, J.J.; Doorn, A.J. van (1982) Invariant features of contrast detection: an explanation in terms of self-similar detector arrays. J. Opt. Soc. Am. 72, 83-87.

Lupker, S.J. (1979) On the nature of perceptual information during letter perception. Perc. Psychoph. 25, 303-312.

Nakane, Y.; Ito, K. (19..) Study on standard visual acuity curves for seeing in lighting design. J. Light. & Visual Env. 28, 1978, 38-44.

Nes, F.L. van; Jacobs, J.C. (1981) The effect of contrast on letter and word recognition. IPO Ann. Progr. Rpt. 16, 74-83.

Sachs, M.B.; Nachmias, J.; Robson, J.G. (1971) Spatial frequency channels in human vision. J. Opt. Soc. Am. 61, 1176-1186.

Spencer, H.; Reynolds, L.; Coe, B. (1977) The effects of image/background contrast and polarity on the legibility of printed materials. Readability of Print Research Unit, Royal College of Art.

Roufs, J.A.J. (1974) Dynamic properties of vision-VI. Stochastic threshold fluctuations and their effect on flash-to-flicker sensitivity ratio. Vis. Res. 14, 871-888.

Roufs, J.A.J.; Blommaert, F.J.J. (1981) Temporal impulse and step responses of the visual system obtained psychophysically by means of a drift-correcting perturbation technique. Vis. Res. 21, 1203-1221.

Roufs, J.A.J.; Bouma, H. (1980) Towards linking perception research and image quality. J. Soc. Inf. Displ. Eng. 21, 1980, no. 3, pp. 247-270.

Thurstone, L.L. (1927) Psychophysical Analysis. Am. J. Psych. 38, 368-389.

Timmers, H. (1978) An effect of contrast on the legibility of printed text. IPO Ann. Progr. Rpt. 13, 64-67.

Townsend, J.T. (1971) Theoretical analysis of an alphabetic confusion matrix. Perc. Psychph. 9, 40-50.

Wilson, H.R.; Bergen, J.R. (1979) A four mechanism model for threshold spatial vision. Vis. Res. 19, 19-32.

Synopsis

In this thesis, the central question is in what way the human visual system processes details in space and in time and how this can be modelled. Details should be interpreted here as black and white stimuli that are limited in one of these dimensions. An example of a stimulus which is limited in space is a printed letter at a normal viewing distance. In the time domain, a signal-light, lasting a few milliseconds is another example.

In order to make generalisations possible, systems theory was used. Unfortunately, the visual system is non-linear in many aspects. Since, in practice, the state of the visual system is determined by the background level and the signals at visibility level remain relatively small, linearising around these levels was applied. This is the reason why linear systems theory can be used in order to analyse the relations between input (detail) and output signal (visual response on detail). The goal is to predict visual responses on arbitrary details and their detection thresholds, on the basis of the mentioned quantitative description.

An important part of the investigation consisted of the determination of basic functions which may be used to quantify visual processing of details. To this end, impulse responses were chosen for the time domain and point spread functions for the space domain. These functions have in common that they represent responses on Dirac-functions: they are the response on an 'infinitely' short flash in time and an 'infinitely' thin point in space, respectively. In order to measure these basic functions, an earlier developed perturbation technique was used, in which the visual response on such a Dirac-pulse is scanned by a suitable probe. The experimental methods used (of which in chapter 1 the global features are treated and in chapter 2 the details), are of a psychophysical nature. The experimental measures consisted mainly of 50% detection thresholds.

It was found that impulse responses and point spread functions can be measured sufficiently accurately to make a quantitative analysis possible (chapter 2, 3, 4, 5). Furthermore, the measurements are shown to be reproducable and consistent for different subjects. Impulse responses determined at different background levels do not change in shape if the spatial configuration remains the same (isomorphy; chapter 2, 3). Isomorphy is also found if point spread functions are determined at a different adaptation level or at another retinal position (chapter 5). Furthermore, it was found that the neural part of point spread functions is dominant if this is estimated from the (known) optical spread (chapter 5).

The usefulness of the chosen basic functions depends on the validity of the model assumptions. In this case, small-signal linearity of the visual process and peak detection of the signals at threshold level were the main assumptions. Realising their importance, these hypotheses were tested in several ways, such as by determining step responses in the time domain and an edge spread function in the space domain. Results of the experiments should obey a unique relation with impulse responses and point spread functions, as dictated by the assumptions. This relation was confirmed within experimental error (chapter 2, 4).

Another part of the investigation consisted of testing the generalising power of impulse responses and point spread functions. For this purpose, thresholds were calculated for various stimuli which differ from Dirac-functions, and these were confronted with experimental findings. For the time domain, where the calculations were based on experimentally determined impulse responses, the prediction was good in all cases investigated (chapter 3). The way in which the visual system processes temporal stimuli depends on the spatial configuration of the stimulus. From literature it is known, that for low spatial frequencies 'transient'-like processing occurs. This transfer, defined in electrophysiology, relatively favours fast illuminance changes. A 'sustained'-like transfer, dominating at high spatial frequencies, is active at slow variations and static stimuli. Furthermore it is known that the 'transient'-like system operates as a band-pass filter and the 'sustained' one as a low-pass transfer. The monophasic impulse response for point sources' linked with the perceptual impression of homogeneous brightness change ('swell'), is in agreement with this property. For more extended stimuli, the impulse responses are triphasic, and connected with a percept which may be described as a spatially inhomogeneous disturbance of brightness ('agitation). By making suitable stimulus choices, both channels were analysed separately, and also in both cases an adequate prediction of thresholds could be obtained (chapter 3).

For the space domain, where point spread functions were used as a starting point, the predictions are found to be accurate for a class of "slender" stimuli among which printed letters. For stimuli having a more "chubby" character like squares and bars, a systematic deviation is found between prediction and experiment. This deviation can be overcome by assuming that at a single retinal position different channels or point spread functions operate in parallel. In this concept, the experimentally determined point spread functions represent the smallest channels (chapter 5).

The nature of the measured point spread functions (in the spatial domain) not only depends on adaptation level and retinal position but also on the duration of the stimulus. It is found that a point spread function becomes broader if the stimulus presentation is shorter. This is in agreement with the common opinion that the visual system has a high resolving power in the space domain or in the time domain separately, but not at the same time in both domains.

In chapter 7 of this thesis, the usefulness of point spread functions is tested by calculating letter confusions from them if the letters are presented at a large viewing distance or, at a normal viewing distance but at a retinal position situated a distance of 7 deg from the fovea. These conditions, which hamper recognition, are often used to measure the quality of letter types and to determine the recognition probability and associated confusion matrix. It is found that a major part of the confusions can be attributed to optical and neural imaging properties which are represented in the experimentally determined point spread functions.

Letter confusions can also be provoked by diminishing the contrast of letters under otherwise normal viewing conditions. This is a second method which may be used to determine the quality of letter types.Experimentally determined recognition thresholds, appear to depend, apart from adaptation level and letter size, on global

psychophysical characteristics like acuity. Since detail vision depends on adaptation level, it has to be expected that relations exist between recognition probability, letter size, letter contrast and adaptation level. These relations were analysed experimentally, and explained within a global model. Furthermore, it is shown that the alphabet becomes inhomogeneous with respect to recognition of its constituent items if letter size decreases. This means that this type of approach may be useful in designing letter types, in order to decrease the confusion chances between letters.

It seems justified to conclude that, the technique treated in this thesis offers a fair description for detail vision at threshold level. Therefore, an extension to supra-threshold visibility percepts like brightness looks promising, if certain conditions are fulfilled.

Samenvatting

In dit proefschrift staat de vraag centraal op welke manier het menselijke visuele systeem details in plaats en tijd verwerkt en hoe deze verwerking modelmatig beschreven kan worden. Met details worden hier stimuli met een beperkte uitgebreidheid in deze dimensies bedoeld. Met betrekking tot de plaats kan men daarbij bijv. denken aan typeletters op een normale leesafstand. Een voorbeeld van een in de tijd beperkte stimulus is een signaal-licht van enkele milliseconden.

Om generalisatie mogelijk te maken is gebruik gemaakt van systeemtheorie. Ongelukkigerwijze is het oog in het algemeen niet lineair. Omdat in de praktijk de toestand van het visuele systeem bij de hier behandelde details bepaald wordt door hun achtergrond en de signalen aan de grens van de zichtbaarheid relatief klein blijven, heeft linearisering rond deze nivo's als uitgangspunt gediend. Daarom kan gebruik worden gemaakt van de lineaire systeemanalyse waarbij de relaties tussen ingangssignaal (detail) en uitgangssignaal (responsie op detail) relatief eenvoudig zijn. Het uiteindelijke doel is dan om op grond van die kwantitatieve beschrijving de responsie van het visuele systeem op andere details en daaruit detektie te voorspellen.

Een belangrijk gedeelte van het onderzoek betrof het meten van basisfunkties waarmee de visuele verwerking van details gekarakteriseerd kan worden. Gekozen is voor het meten van impulsresponsies in het tijddomein en puntspreidfunkties in het plaats-domein. Beide funkties hebben met elkaar gemeen dat ze responsies voorstellen op stimulering met een Dirac-funktie: het zijn de responsie op respektievelijk een "oneindig" korte flits in de tijd en een "oneindig" smalle punt in de plaats. Voor het meten van deze basisfunkties werd reeds eerder een zogenaamde storingstechniek ontwikkeld, waarbij de visuele responsie op zo'n Dirac-puls als het ware gescanned wordt met een geschikt gekozen sonde. De meetmethoden die gebruikt zijn (waarvan in hoofddstuk 1 algemene zaken aan de orde komen en in hoodstuk 2 de details), zijn psychofysisch van aard. De metingen bestaan voornamelijk uit het bepalen van 50% zichtbaarheidsdrempels.

Het blijkt dat impulsresponsies en puntspreidfunkties met voldoende nauwkeurigheid gemeten kunnen worden om een kwantitatieve analyse mogelijk te maken (hoofdstuk 2, 3, 4, 5). Tevens blijken de experimentele resultaten herhaalbaar te zijn en consistent voor verschillende proefpersonen. Impulsresponsies gemeten bij verschillende adaptatienivo's veranderen niet van vorm (isomorfie) als de spatiale konfiguratie gelijk blijft (hoofdstuk 2, 3). Isomorfie werd ook gevonden indien puntspreidfunkties op een ander adaptatienivo of een andere retinale positie bepaald werden (hoofdstuk 5). Tevens werd gevonden dat het neurale aandeel in puntspreidfunkties overheerst indien dit afgeschat wordt aan de (bekende) optische spreiding (hoofdstuk 5).

De bruikbaarheid van de gekozen basisfunkties hangt sterk af van de korrektheid van de uitgangshypothesen. De gekozen aanpak gaat uit van quasi-lineariteit van de visuele verwerking en piekdetektie van signalen op drempelnivo. Gezien het belang van deze hypothesen zijn ze op verschillende manieren getest, o.a door het bepalen van stapresponsies in het temporele domein en een randspreidfunktie in het spatiale domein. Resultaten van deze metingen dienen een eeneenduidige, door de uitgangshypothesen

voorgeschreven, relatie te hebben met gemeten impulsresponsies en puntspreidfunkties. Binnen de experimentele nauwkeurigheid werd hieraan voldaan (hoofdstuk 2, 4).

Een ander belangrijk gedeelte van het onderzoek bestond uit het testen van de generaliseerbare kracht van de gemeten basisfunkties. Met dit doel zijn voorspellingen berekend voor drempels van andere, van Dirac-funkties verschillende stimuli, en zijn deze vergeleken met experimenteel bepaalde drempels. Voor het tijddomein, waar berekeningen gebaseerd zijn op gemeten impulsresponsies, is de voorspelling in alle gevallen binnen de meetnauwkeurigheid in overeenstemming met het experiment (hoofdstuk 3). De manier waarop het visuele systeem temporele stimuli verwerkt is afhankelijk van de spatiale konfiguratie van de stimulus. Uit de literatuur blijkt dat bij lage spatiale frekwenties 'transient'-achtige verwerking plaatsvindt. Deze in de electrofysiologie gedefinieerde overdracht is relatief gunstig voor snelle veranderingen. Bij hoge spatiale frekwenties blijkt een soort 'tonische' overdracht te domineren en deze is hoofdzakelijk werkzaam bij langzame variaties en statische stimuli. Gebleken is ook dat het 'transient'-achtige systeem zich als een bandfilter gedraagt en het tonische als een laagdoorlatend filter. Voor een puntbron gedraagt de impulsresponsie zich, daarmee in overeenstemming, monofasisch, waarbij de perceptieve indruk een homogene helderheidsverandering is ("deining"). Voor meer uitgebreide stimuli is de gemeten impulsresponsie trifasisch, terwijl het percept omschreven kan worden als een spatiaal inhomogene verstoring van de helderheid ("onrust"). Door een geschikte keuze te maken voor de stimuli, zijn beide kanalen afzonderlijk onderzocht waarin ook in beide gevallen een adequate beschrijving van drempels gerealiseerd is (hoofdstuk 2, 3).

Voor het plaatsdomein, waar puntspreidfunkties als uitgangspunt voor de berekening dienen, blijken de voorspellingen te voldoen voor een verzameling van "slanke" stimuli zoals typeletters. Voor stimuli met een "molliger" karakter zoals vierkantjes en balkjes, is een systematische afwijking gevonden tussen voorspelling en experiment. Deze kan worden verklaard door aan te nemen dat op een retinale positie meerdere kanalen of puntspreidfunkties van verschillende afmeting naast elkaar werkzaam zijn. In dit concept vertegenwoordigen gemeten puntspreidfunkties de smalste kanalen (hoofdstuk 5).

De aard van gemeten puntspreidfunkties (in de plaats) hangt niet alleen af van adaptatienivo en retinale positie maar ook van de stimulusduur. Het blijkt dat een puntspreidfunktie breder wordt naarmate de stimulusduur korter is. Dit is in overeenstemming met de algemene idee dat het visuele systeem een hoog scheidend vermogen heeft in de plaats of in de tijd afzonderlijk, maar niet tegelijkertijd in plaats en tijd (hoofdstuk 6).

In Hoofdstuk 7 van het proefschrift wordt de bruikbaarheid van puntspreidfunkties getest bij het berekenen van letterverwarringen die optreden wanneer typeletters op grote afstand aangeboden worden, of op een korte afstand doch een retinale positie die 7 graden buiten het blikpunt ligt. Deze, de herkenning bemoeilijkende kondities worden wel gebruikt om de kwaliteit van lettertypen te meten door de identifikatie-

waarschijnlijkheid en de daarbij behorende verwarringsmatrix te bepalen. Het blijkt dat een belangrijk gedeelte van deze verwarringen verklaard kan worden uit optische en neurale afbeeldingseigenschappen zoals die in gemeten puntspreidfunkties vertegenwoordigd worden.

Letterverwarringen kunnen ook uitgelokt worden door het kontrast van letters te verlagen, waarbij kijkafstand en retinale positie verder normaal gehouden worden (hoofdstuk 8). Dit is een tweede methode om de kwaliteit van lettertypen te bepalen. Experimenteel bepaalde herkenningsdrempels van letters blijken behalve van de grootte van de letters en het adaptatienivo af te hangen van globale psychofysische karakteristieken zoals scherpte. Daar de mate van detailzien afhangt van van het adaptatienivo moet men verwachten dat er relaties bestaan tussen identifikatie-waarschijnlijkheid, lettergrootte, letterkontrast en adaptatienivo. Deze zijn experimenteel bepaald en door een globaal model verklaard. Tevens wordt aangetoond dat voor kleine letterafmetingen, herkenningsdrempels niet gelijk zijn voor alle letters van het alfabet. Dit betekent dat dit type benadering een methode inhoudt om bij letterontwerp rekening te houden met een zo klein mogelijke verwisselingskans.

Als konklusie lijkt gerechtvaardigt dat de in dit proefschrift behandelde techniek een goede beschrijving oplevert voor het waarnemen van details op de grens van het zichtbare. Daarom lijkt een uitbreiding naar bovendrempelige percepten zoals helderheid veelbelovend, mits aan zekere randvoorwaarden is voldaan.

Curriculum vitae

2 juni 1948	geboren te Deurne.
sept. 1960 - juli 1965	Peellandcollege te Deurne, HBS-b.
sept. 1965 - nov. 1975	TH-Eindhoven, Technische Natuurkunde. Afstudeeronderwerp: Visuele dynamica.
dec. 1975 - jan. 1978	Wetenschappelijk assistent TH-Einhoven, gedetacheerd op het Instituut voor Perceptie Onderzoek (IPO) voor het verrichten van onderzoek op het gebied van visuele kontrastoverdracht.
april 1978 - april 1982	Wetenschappelijk medewerker in dienst van de Stichting Biofysica van de Nederlandse Organisatie voor Zuiver Wetenschappelijk Onderzoek (ZWO), gedetacheerd op het IPO voor het verrichten van onderzoek op het gebied van visuele kontrastoverdracht.
aug. 1985 - sept. 1986	Wetenschappelijk assistent TH-Eindhoven, gedetacheerd op het IPO voor het verrichten van onderzoek aan visuele kontrastoverdracht en het gebruik van leesloupes door slechtzienden.
sinds sept. 1986	Wetenschappelijk medewerker TU-Eindhoven. Onderwerp: onderzoek aan perceptieve aspekten van LC-TV's (contractresearch voor Philips).

Stellingen

I

Ten onrechte gebruikt Watson (1982) het feit dat difasische impuls-
responsies in de electrofysiologie gebruikelijker zijn dan trifasische,
als argument om psychofysische trifasische meetresultaten te ver-
werpen als impulsresponsies.

Watson, A.B. (1982)
Derivation of the impulse response: comments on the method of Roufs
and Blommaert. Vis Res. 22, 1335-1337.

II

Door het feit dat frekwentie-selektieve adaptatie slechts lokaal werkt,
boet frekwentie-analyse van spatiale visuele overdracht aan kracht in.

Perizonius, E.; Schill, W.; Geiger, H.; Röhler, R. (1985)
Evidence for the local character of spatial frequency channels in the
human visual system. Vis. Res. 25, 1233-1240.

III

Het door Ginsburg (1980) aangedragen argument dat frekwentie-
analyse van spatiale visuele overdracht eenvoudiger is en meer inzicht
biedt dan een beschrijving in het plaats-domein, geldt hoogstens voor
onderzoekers die ruime ervaring hebben met Fourier-analyse.

Ginsburg, P.A. (1980)
Specifying relevant spatial information for image evaluation and display
design: an explanation of how we see certain objects. Proc. SID 21,
219-227.

IV

De helling in het 50%-punt van een experimenteel bepaalde psy-
chometrische kromme wordt minder steil naarmate de meting langer
duurt.

Hoofdstuk 3 van dit proefschrift.

V

Bij het lezen met behulp van een handloupe is er een kategorie mensen die tekst aan de linkerkantlijn lezen met het rechteroog en tekst aan de rechterkantlijn met het linkeroog.
Blommaert, F.J.J.; Neve, J.J. (forthcoming)
On the field width and field displacement of loupes in reading. Submitted to J. Opt. Soc. Am. -A.

VI

De uitdrukkingen positief kontrast en negatief kontrast, zoals die vaak gebruikt worden bij tekst op beeldschermen, zijn niet alleen ambigu doch ook ongelukkig gekozen gezien het subjektieve karakter van de adjektieven en gezien de wetenschappelijke onzekerheid over welk kontrast te prefereren is.

VII

Afgezien van de onmiskenbare voordelen die tekstverwerkers bieden bestaat het gevaar dat hun tekortkomingen de lay-out van officiële documenten gaan bepalen.

VIII

De krantekop "Rellen eisen gewonden" lijkt niet alleen de konstatering van een feit te zijn maar schijnt tevens voor sommige deelnemers de definitie van een spelregel in te houden.

IX

Aanhangers van een no-nonsense beleid dienen te bedenken dat men, door het voortdurend met beide benen op de grond staan, niet ver komt.

Frans J.J. Blommaert

3 februari 1987