Avtometria 6 (1967) 113-120
Translated from Russian by S A Sullivan

MODEL OF THE DETECTION OF SIMPLE PATTERNS BY THE VISUAL SYSTEM

J. J. KULIKOWSKI

The primary image, which is described in the plane x_o, y_o by the function $i(x_o, y_o)$, is transformed by the process described by the following equation:

$$u(x,y) = \iint_{-\infty}^{\infty} i(x_o, y_o)a(x_o, x; y_o, y)dx_o dy_o \qquad (1)$$

where $u(x, y)$ is the distribution of optical signals in the plane of the secondary image; and $a(x_o, x; y_o, y)$ is the weighting function.

In the simplest case where the primary image is a point, (i.e. the primary image is described by the delta-function), the secondary image represents the weighting function. The transformation is called isotropic of it is independent of translation and rotation of the system of coordinates x, y [1]. The human visual system does not carry out an isotropic transformation; the resolving ability of the visual system is best in the fovea, and is better in horizontal and vertical directions than in the oblique.

The resulting characteristic of the transformation is a combination of the characteristics of the operation of the optics of the eye, and that of the nervous system, i.e. the transformation is composite.

Investigation of the transformation makes it possible, firstly, to trace the process of feature extraction from the plane image carried out by the nervous system, and, secondly, to draw conclusions about the detection ability of the human-operator. In this analysis we may hope to elucidate the more important question of image identification. [2].

Spatial-frequency characteristics.

In order to determine as generally as possible the manner in which fine details are distinguished, it is essential to introduce the spatial frequency characteristics of the system. These show at what density of details of the image (e.g. straight lines) it is possible to distinguish the details. In practice, as input signal it is convenient to introduce an image modulated along one of the spatial coordinates, e.g. an image whose brightness along the horizontal axis x is describes by the expression:

$$L = L(1+m \sin w_x x); \qquad (2)$$

$$w_x = 2\pi f_x, \qquad (3)$$

where L is the mean brightness;
m is the amplitude of the spatial modulation;
f_x is the spatial frequency (cycles/degree).

The determination of the spatial frequency characteristic consists of measuring the response of the system to the modulated image. Such a characteristic is called the modulation transfer function in one dimension, $A(w_x)$ or two dimensions, $A(w_x, w_y)$, and represents the Fourie transformation of the weighting functions $a(x_o)$ or $a(x_o, y_o)$ respectively, i.e.

$$A(w_x) = \int_{-\infty}^{\infty} a(x_o)\exp(-jw_x x_o)dx_o; \qquad (4)$$

$$a(x_o) = \int_{-\infty}^{\infty} a(x_o, y_o)dy_o. \qquad (5)$$

Using the following dependence it is comparatively easy to determine the frequency characteristic (i.e. the modulation transfer function).

$$A(w_x) = \frac{U(w_x)}{I(w_x)},$$

where $U(w_x)$ and $I(w_x)$ are the one dimensional Fourier transforms of the distributions corresponding to the second and first images, u and i. It should be noted that these characteristics have found application in physical optics and photography [3]. In the field of physiological optics using the above techniques which were introduced by Schade [4], it has been possible to separate the transfer function of the visual optics from the transfer function which describes the nervous system [5]. The determination of the modulation transfer function of the nervous system alone is the harder because the output is not susceptible to measurement. However it is possible to determine, for different spatial frequencies, the threshold values of the amplitude of modulation, i.e. the minimum values of the amplitude of modulation. The inverse of the threshold modulation amplitude, at which the image is visible,

$$S = \frac{1}{m_t} \qquad (6)$$

is called the contrast sensitivity. The contrast sensitivity as a function of frequency may be interpreted as the modulation transfer function of the system for threshold values. Fortunately, when investigating threshold values, it is possible to apply linear analysis, since the amplitudes of the signals are small. The measurement of threshold values is expedient not only because of the convenience of analysis, but because they can be used to determine the limits of the perception of fine details. (In the light of the 'human-operator problem'- Ed. Avtometria). A second method of determining the transfer function is based on comparative suprathreshold measurements [6], (in this case the system is noticeably non-linear).

The basic deficiency of both methods is that, at the time of measurement, the characteristics of the complete system are determined; this makes the analysis of the separate stages impossible. Therefore, when constructing models it is necessary to make use of electrophysiological data, and then to check whether the characteristics of the model developed correspond to the characteristics of the visual system.

Formulation of the problem.

The primary image is projected by the optical system of the eye onto the photoreceptors which are connected by means of various interstitial cells with the ganglion cells of the retina. The area of receptors, from which it is possible to excite a ganglion cell, is called the receptive field. The excitation can have positive or negative values.

From a series of investigations [7, 8] it is known that the retinae of highly organised mammals (cats, monkeys), have concentric receptive fields.(Fig. 1). This means that the values of the weighting function are equal at points equidistant from the centre of the receptive field, (Fig. 1a). The next stage of the transformation of the image is morphologically linked with the lateral geniculate body - which also has concentric fields (convergent zones of neurones unlike the receptive fields of the retina - Ed. Avtometria). The substantial difference in the forms of these fields can be revealed by the investigation of the receptive fields of the subsequent transforming stage - the visual cortex. [9]. Lines of equal weighting functions for some fields in the visual cortex have elliptical form. Thus it transpires that after a stage having concentric fields, there appear elongated fields.

The many-stage process of transformations in the visual analyser examined above, can be simplified to a two-stage functional model (Fig. 1c). A response of the first stage $u(x, y)$ appears when the sum of input signals $i(x_o, y_o)$ inside the receptive field R, exceeds a critical threshold value P_1:

$$u_1 = \max\left[0; \iint_{R_1} i(x_o, y_o)\, a_1(x_o - x_1, y_o - y_1)\, dx_o dy_o - P_1\right]. \qquad (7)$$

Similarly the second stage can be described by:

$$u_2 = \max\left[0;\ \iint_{R_2} u_1(x_1,y_1)\,a_2(x_1-x_2,y_1-y_2)\,dx_1\,dy_1 - P_2\right]. \qquad (8)$$

According to electrophysiological investigations, the field R_2 is elongated as a result of summation of regions of type R_1 along a straight line [9]. The influence of threshold values P_1, P_2, on the resultant threshold of discrimination can be determined experimentally by investigating the critical flicker frequency in psychophysical experiments.

Experimental Results.

The subject is placed 100 cms. from a cardboard surround subtending 10°, which is evenly illuminated (75 c/m²). In the centre of the surround there is an aperture, the dimensions of which can be altered, so that it subtends angles between the limits 3 mins. to 2 degs. Behind the aperture there is an oscilloscope screen on which the spatially modulated image is generated, using television techniques [4]. The average brightness of the screen is constant and equal to the brightness of the disk. The experimenter can change the frequency of spatial modulation, breadth and height of the aperture, the type of modulation, (sinusoidal, square and impulse), and other parameters. Each time the subject sets the amplitude of the modulation to threshold.

In the first series of experiments the threshold values of the modulation were determined for each spatial frequency. On the basis of the data acquired the contrast sensitivity function ($1/m_t$) was constructed as shown in logarithmic scale in Fig. 2. For frequencies of 5-6 c/deg. the function has a clearly expressed maximum- the maximum detectable frequency is about 60 c/deg. It is of interest to note that the data can be approximately described by the function [10]

$$S = A_+ \exp(-a_x f_x) - A_- \exp(-b_x f_x). \qquad (9)$$

The second term in the expression represents the action of inhibiting receptive field areas.

In the second experiment the areas over which periods could summate was determined. A modulated image of 5 c/deg (corresponding to maximum sensitivity) was viewed through a rectangular aperture of variable width; the width being varied by steps equal to one period. It was concluded that the contrast sensitivity increases in proportion to the square root of the width of the aperture up to a width corresponding to approximately 60 mins. (i.e. 5 cycles). Approximately the same dependence is obtained in the case of changing the height of the aperture between 6 mins. and 60 mins. In this case the contrast sensitivity increased in proportion to the square root of the line length. When the height and breadth of the aperture were increased simultaneously, so that the aperture remained square, the contrast sensitivity increased in proportion to the side of the square. It follows from this that summation occurs within a 60 mins. by 60 mins. area. It should be emphasised that the area of summation was different for different frequencies, (e.g. for 20 c/deg it is 33 mins by 33 mins).

In the third series of experiments, the subject set the threshold modulation amplitude of two superimposed images: one image modulated along the horizontal (giving vertical lines as before), the second modulated vertically (horizontal lines). As the two images have the same frequencies and amplitudes of modulation, the difference between maximum and minimum brightness is twice that in the case of a single image. However in spite of this, the threshold modulation decreases not by 50%, but only insignificantly; this may support the hypothesis that there exist separate 'channels' for horizontal and vertical lines. This means that in the model proposed in Fig. 1c, the threshold value of the second stage, P_2, is significantly greater than P_1. Similar conclusions also follow from results of experiments using different methods (masking) [11].

In the fourth series of experiments the connection between the frequency characteristics of images modulated along one dimension, and the ability to detect lines was established. These experiments are similar to those of the second series, in which there is only one period. If instead of one period of sinusoidal modulation, a line of width 6 mins. is used, the spatial spectrum is that of a square impulse.(Fig. 3). The threshold modulation value of such an impulse (i.e. a line) is determined by the point of intersection of the spectral characteristic and the threshold modulation function.

This intersection is near the minimum of the threshold modulation function (i.e. near the maximum of the contrast sensitivity function). Because the spectrum of lines narrower than 2 mins. tends to a flat line between 0 and 5 c/deg, the contrast sensitivity for such lines is proportional to their width [12]. The narrowest line capable of being distinguished was 1.2 secs. wide, which broadly agrees with results obtained by different methods of investigation.

From what has been said above, it follows that there is a physical difference between the ability to distinguish a single line and a series of lines, i.e. it depends on the different physical characteristics of the spatial spectrum. These results have made redundant hypotheses about the existence of special mechanisms detecting individual lines of widths of the order 0.5 to 1 secs, i.e. for lines narrower than the photoreceptor size.

The results of the 5th series of experiments determine the contrast sensitivity of lines of various lengths. Thus for lines of widths 0.6 mins and 6 mins, the lengths of which were increased from 6 mins initially, it was found that contrast sensitivity increased (the threshold values of modulation decreased correspondingly) proportional to the square root of the line length, up to the point where line length reached 60 mins. (Fig. 4). This is very similar to the phenomenon which occured in the second series of experiments, where the height of the aperture was varied, while a spatial modulation of 5 c/deg was viewed through it. Consequently one and the same mechanism detects the narrow lines and the 5 c/deg pattern. What is more, it was found that when a line 0.6 mins wide is reduced to a length of less than 6 mins, the contrast sensitivity function decreases more rapidly than the square root of length. For line lengths smaller than 2 mins, contrast sensitivity becomes proportional to the length.

It should be noted that the discrimination of two-dimensional objects of the same area, whose maximum dimension is not greater than 2 mins, becomes impossible. A rectangle can be distinguished from a square of the same height, when the length of the rectangle is of the order of 4mins. (The width of the rectangle should be less than 2 mins).

Let us note that this detection, which coincides with the ability to determine the angular orientation of the rectangle, is possible only when the contrast sensitivity ceases to be proportional to the line length, and becomes proportional to the square root of line length. This indicates that in the case of a line longer than 4 mins, there occurs summation of signals from component sources each with independent noise; this corresponds to the model shown in Fig. 1c. If a line is short enough, it is detected by a single concentric receptive field. In this case there is full proportionality of the contrast sensitivity to the area. As soon as the length of line exceeds a critical value, it is not located inside one concentric receptive field, and therefore the next field is excited; this leads to a summation of signals with independent noise. The 'second point' in the plane, which determines the orientation of the segment, is excited.

Most of the experiments above were carried out by the author in the physiology laboratory of the University of Cambridge.

Conclusions.

The suggested model of transformations in the visual system is based on data from electrophysiological investigations and does not contradict the investigation of spatial frequency characteristics of the whole system. In particular it is shown that:-

1) The second stage of transformation, consisting of elongation of the receptive field in one direction, is characterised by a higher threshold value than the first stage, which has concentric fields.

2) The summation of signals inside the concentric receptive field is proportional to the area and energy of these signals, but the summation of signals from several receptive fields is proportional to the square root of the line length.

3) The ability to discriminate straight lines has a direct link with the spatial frequency characteristics of human vision.

Bibliography.

1. L. Kovasznay. Image Processing. Proc. I.R.E. 1955, 43, 560-570.
2. Glazer and Zuckerman. Information and Vision. Ac. of Sci. U.S.S.R. 1961.
3. E. Linfoot. Fourier methods in optical image evaluation. Focal Press London 1964.
4. O. Schade. Optical and Photoelectric Analog of the Eye J.O.S.A. 1956, 46, 721-739.
5. Campbell and Green. Optical and Retinal factors affecting visual resolution. J.Physiol, 1965, 181, 576-593.
6. E.H.Lowry and S.S.de Palma. Sone-wave response of the visual system. J.O.S.A.1961, 51, 740-746.
7. S.Kuffler. Discharge Patterns and Functional Organisation of Mammalial Retina. Jour. Neurophysiol. 1953, 16, 37-68.
8. Wiesel. Receptive Fields of Ganglion Cells in the Cat's Retina. J. Physiol.(London) 1960, 153, 583-594.
9. D. Hubel. Transformation of Information in the Cat's visual system. Information Processing in the nervous system. XXII Intern. Congress Leiden 1962. Excerpta Medica.
10. J.J.Kulikowski, F.W. Campbell, J.G. Robson. Spatial and temporal frequency characteristics of human photopic vision, 2nd Internat. Biophysics Congress, Vienna, September, 1966.
11. F.W.Campbell, J.J.Kulikowski. Orientational selectivity of the human visual system. J. Physiol. 1966, 187, 437-445.
12. F.W.Campbell, D.G.Green, J.J.Kulikowski, J.G.Robson. The detectability of thin lines. (Unpublished).

Fig 1.

a) Sketch of receptive fields, concentric and elongated, and their weighting function.

b) Structural scheme of three stages of the visual system: retina, lateral geniculate body and visual cortex.

c) Structural scheme of the functional model of part of the visual system.

Fig 2.

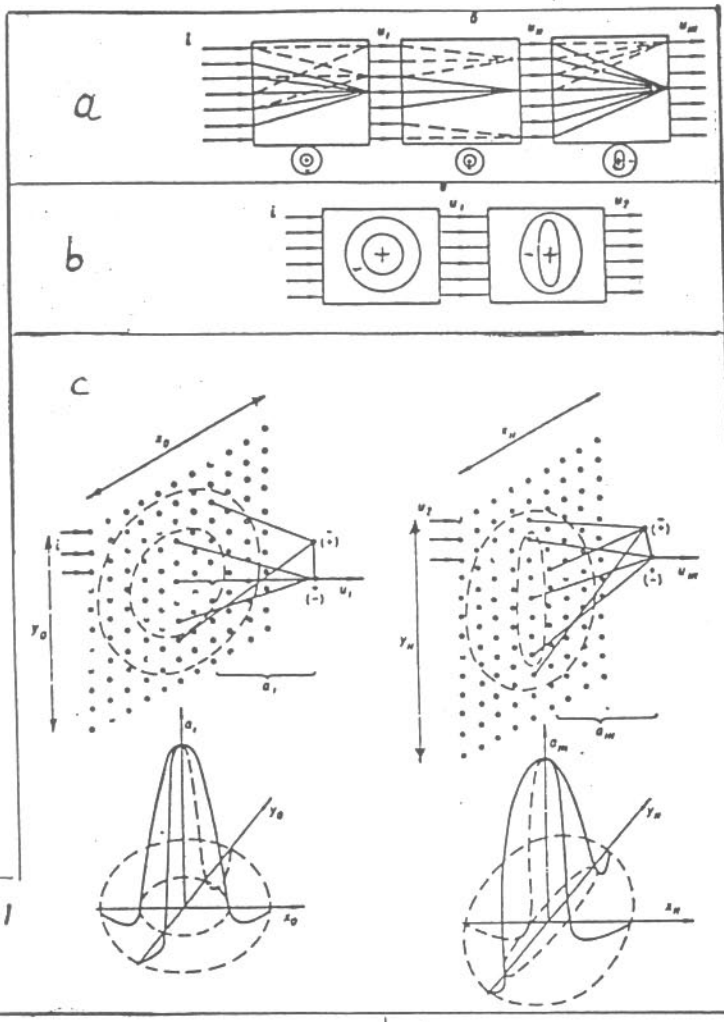
Contrast sensitivity (i.e. the inverse of amplitude of threshold modulation) as a function of spatial frequency.

Fig. 3.

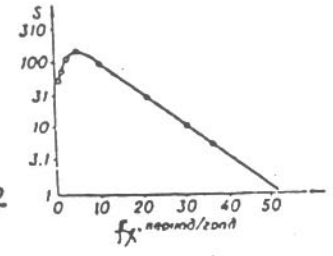
Spectrum of spatial frequency of the line width 6 mins, (dashed) and the threshold modulation function (solid line).

Fig 4.

Contrast sensitivity as a function of height of aperture (h), (the aperture is orientated perpendicular to the horizontal direction of spatial modulation,) for the case of a square wave modulation of 5 c/deg and for single impulses (i.e. lines) of width o.6 mins and 6 mins.



FIG·1



FIG·2

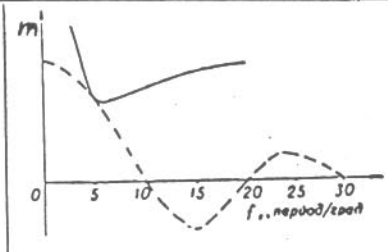
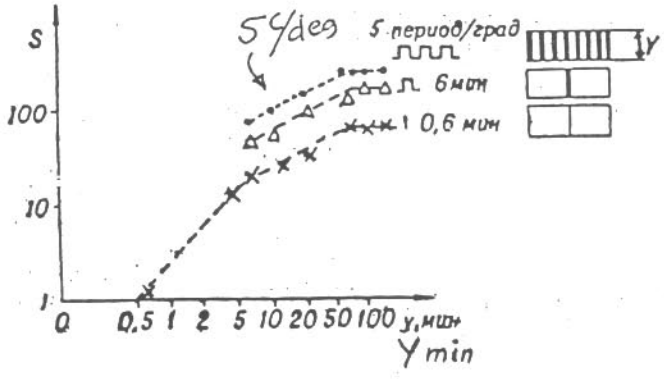


FIG 3



FIG·4