

COMPUTER SIMULATION OF ARTICULATORY ACTIVITY IN
SPEECH PRODUCTION

P. Mermelstein
Bell Telephone Laboratories, Incorporated
Murray Hill, New Jersey

A computer facility for modeling sound production resulting from exciting a moving articulatory system is described. The vocal tract is treated as an acoustic transmission line with variable cross-sectional area. The system allows on-line specification of the history of articulatory states (vocal-tract shapes) and excitation parameters, and generates the resultant acoustic signal for immediate evaluation by the experimenter. Novel features include graphical as well as numerical control of articulatory configurations, time-motion display of articulatory data and resonant-frequency versus time displays for the resulting acoustic signal. The facility is used for the study of the dynamics of articulatory movement and for speech synthesis.

Introduction

In modeling the process of human speech production we recognize five distinct stages. First, the message is organized on the linguistic level and structured grammatically. Second, the message is expressed phonologically in terms of a sequence of discrete units, labelled phonemes. Stress and intonation marks can be considered special phonemes and assigned special codes. Third, the string of phonemes is converted to a set of continuous signals controlling the musculature of the various articulators. This results in a highly complex integrated movement sequence in which generally all the articulators, the lips, the tongue, the mandible, etc., participate. Finally, the physical interaction of the vibrating vocal cords and the moving articulatory structure produces a continuous acoustic signal perceived as speech.

We are particularly interested in modeling the fourth or the articulatory movement stage of the speech event. We consider the speech event as composed of a sequence of phonetic segments specified in terms of one or more vocal-tract shapes and the corresponding excitation parameters. The acoustic signal is generated by simulating the excitation mechanism as well as the articulatory trajectory. In contrast, most previous speech production systems consider generation of the speech signal from the acoustic point of view and are controlled by specifying the variations

of vocal-tract resonant frequencies (formant frequencies) with time.^{1,2} Modeling the process at the articulatory level can be expected to be simpler because the articulators respond to muscular forces with predictable changes in their position and rates of movement. On the other hand, simple rules for resonant frequency changes are not necessarily realizable by the anatomy of the vocal tract. In fact, simple articulatory changes can have complex acoustic effects. This report is primarily concerned with generating the acoustic signal corresponding to given changes in the shape of the vocal tract. Accomplishment of that task is a prerequisite to studying the movements actually executed by the articulatory system when producing a particular utterance.

Recent studies show that no phoneme, not even the vowels, exhibits a unique set of resonant frequencies in all its occurrences in connected speech.³ Therefore, it is unreasonable to expect that a unique vocal-tract shape can be associated with all productions of that phoneme. Certain attributes or distinctive features of the excitation mechanism and of the articulatory state serve, however, to differentiate among the phonemes. Such features may include presence or absence of labial closure, of periodic or turbulent excitation, coupling of the nasal passage to the main vocal tract, etc. While incorporation of these features in the articulatory description is necessary to obtain an intelligible result, it is not sufficient. The human speech perception apparatus is easily confused by false cues that arise from unnatural articulations. Therefore, those parameters unspecified by the distinctive features of the phoneme must be specified with reference to the moving articulatory structure so that the result will correspond to the natural execution of the articulation. Contextual effects such as anticipation or inertia control these optional features, and one is interested in specifying in detail the nature of that control mechanism.

If the effects of context are expressed in rules for articulatory movement, only the essential features of the phonemes need be specified explicitly,

and the optional parameters will be determined automatically. The simulation system is intended to aid the postulation of such rules by serving as a convenient tool for the evaluation of their articulatory and acoustic effects. This report describes a facility for specifying a discrete sequence of excitation and articulatory states of the vocal tract, and the system simulates the articulatory movement between these states, thus producing a continuous acoustic signal. The segment corresponding to a particular phoneme is represented as a sequence of one or more such states, where the numerical description of the states may be functions of the adjacent phonemes. For example, sustained tongue tip closure in the absence of voicing is a feature of the consonant /t/, but the shape of the tongue hump for that production is controlled by the adjacent vowels. Aspiration, another context-dependent feature of the /t/, is achieved by optionally introducing states specifying the onset and termination of turbulent excitation. For the moment, all parameters of the excitation and vocal-tract shape must be specified explicitly, or they are assigned fixed default values.

The timing of articulatory movement can be controlled in open-loop or closed-loop modes. The closed-loop mode would require continuous position or velocity feedback; however, the rapidity of the articulatory movement does not allow for this. Hence we assume a ballistic-type movement based on the present and the target articulatory positions. The time interval for the movement is explicitly specified, and it is assumed that the forces generated by the musculature are adjusted accordingly. The time durations are in practice functions of the adjacent articulatory states and future models can be expected to compute them from the articulatory specifications. Where timing differences arise from causes other than the articulatory context, such as the duration difference in English vowels before voiced and unvoiced consonants, these cannot be handled at the articulatory level and would have to be introduced explicitly even where articulatory movement proceeds completely by rule.

The speech production apparatus is modeled as consisting of a periodic source of variable frequency and amplitude, the vocal cords, and/or a noise source, exciting a nonuniform acoustic transmission line continuously changing in shape. The shape, of course, controls the characteristic impedance at any point, and thereby the transfer function of the line. Articulatory states are represented in terms of sampled values of the vocal-tract

cross-sectional area at points along the tract.

Henke⁴ has postulated a set of rules governing tongue, lip and mandible movements based on analyzing frames from high-speed X-ray motion pictures of the moving vocal tract. However, he did not evaluate the resulting speech signal aurally. Our data are derived from acoustic measurement of the vocal-tract cross-sectional area⁵ and admit the postulation of similar rules. This report is primarily concerned with the implementation of a facility for immediate aural evaluation of utterances generated with the aid of postulated rules. Comparison of the computed trajectories of the vocal-tract resonant frequencies as functions of time with those measured from natural articulatory events yields a quantitative difference measure that can be used to complement the subjective evaluation of the audio signal. In addition, the specified movements are displayed for evaluation as sagittal projections convenient for comparison with X-ray pictures.

Static electric analogs of the vocal tract^{6,7,8} have been used for an extensive period as tools in the study of how deformations in the vocal tract control the characteristics of the speech signal. Dynamic electric analogs^{9,10} have been used for the synthesis of a restricted number of syllables, but they suffered from an inability to control the transmission-line elements over wide ranges at high speeds. Kadokawa and Nakata¹¹ have simulated the dynamic behavior of the vocal tract by computing a time-dependent transfer function and thus determining the variations with time of the vocal-tract resonant frequencies. Coker¹² uses a time-varying parametric representation of the vocal tract to compute the changing resonant frequencies and thus control a formant synthesizer. The primary obstacle to the development of a general synthesis facility has been the unavailability of a suitably flexible control system that allows quick and convenient specification or change of the parameters controlling the tract configuration. The computer-simulated system not only eliminates limitations to rapid configuration changes over a wide range of cross-sectional area values, but also permits experiments designed to search for convenient control strategies.

The presence of the human listener in the sound generation - evaluation - modification - sound generation feedback loop is necessary because at present we are unable to assign effective objective criteria to the significant attributes of

the acoustic result. Naturalness is a subjective criterion, and therefore the presence of the listener is essential. The experimenter also acts as a visual observer, evaluating shape and resonant-frequency variations with time. We endeavor to facilitate the experimenter's task so that in the light of his evaluation he can change the data variables quickly and conveniently. This is accomplished by allowing the experimenter to control the simulation using measures familiar to him, for example, lip opening or excitation frequency. He may build up the data for an utterance item by item, evaluating the result each time. Alternately, he may modify data previously entered and hear the modified audio signal in less than a minute.

The Simulation Method

The simulation principles are based on those used by Kelly and Lochbaum¹³ with considerable extensions and modifications. The vocal tract is represented as a non-uniform transmission line supporting longitudinal propagation only. Its characteristic impedance at all points is inversely proportional to the cross-sectional area in a plane perpendicular to the flow line. The tract is assumed hard-walled; the compliance of the soft tissues of the tract wall is ignored. The sampling frequency for the simulation, 24 kHz, is determined by the spacing of area samples, 1.5 cm, that is necessary to simulate the continuous tract up to a signal frequency of 6 kHz. Since the length of each section is considerably smaller than the wavelength even at this frequency, the effect of the distributed reflection of a section traversed by the wavefront in one sampling period can be represented by a lumped reflection coefficient centered on the section.

Figure 1 shows a block diagram of the vocal-tract model. Pressure samples propagate back and forth along the transmission line and suffer multiple internal reflections. Excitation may be periodic, simulating vocal-cord vibration, or turbulent corresponding to frication. For periodic excitation the samples of the excitation function are added to the samples of the reflected signal arriving at the glottis. Turbulent excitation may take place at the vocal cords or at any interior section. It is implemented by means of a white-noise generator and associated source impedance placed in series with the line. For computational simplicity, the tract is assumed lossless within its interior. Simple digital filters act at the tract boundaries, the vocal cords and the lips, to approximate the frequency-dependent losses so that

actually measured resonant-frequency bandwidths are properly matched. The glottal reflection and the labial radiation functions act essentially as high-pass filters; the labial reflection appears as a low-pass filter. The nasal passage is represented by an additional transmission line of fixed but nonuniform shape coupled to the main tract with the aid of a variable coupling parameter. It serves in the production of nasalized vowels and consonants. The system output is obtained by summing the outputs of digital filters approximating the free-space radiation characteristics at the lips and nostrils. Radiation through the lips is, of course, controlled by the variable lip area. If the opening is reduced, more of the energy arriving at the lips is reflected within the tract and less is radiated to the outside. Radiation through the tract wall is not explicitly included.

Articulatory movement is simulated by defining a sequence of articulatory states or targets and interpolating the position-dependent reflection coefficient values and oral-nasal coupling between them. Thus targets are not necessarily phonemic in nature but may be specified as frequently as desired in order to approximate natural articulatory movement as closely as possible. When the articulator positions are derived from a model of articulatory activity, the vocal-tract shapes computed therefrom are expected to act as the target states of the simulation program. Stress and intonation information are assumed embedded in the excitation frequency and amplitude parameters and will not be discussed here in detail.

Linear interpolation of the reflection-coefficient values implies that near sharp constrictions the area varies linearly with time, but where the area changes slightly or the reflections are small, the areas vary exponentially with time. The resonant-frequency curves resulting from this procedure resemble those observed in spectrograms more closely than the straight-line segments used in most formant synthesizers.

Implementation

The system is operational on a DDP-224 computer. All of the routines are written in FORTRAN, except for the most frequently executed loop of the wave propagation computations which was rewritten in assembly language. Among the noteworthy hardware features utilized is the use of input/output channels to feed the loudspeakers through the digital-to-analog converter from one memory module, while simultaneously reading the next record from tape or disk into a second module.

Cathode-ray tube display proceeds simultaneously with processor operations on an interrupt basis.

The execution time of the simulation is approximately 40 times real time. The sample propagation computations along the transmission line, although basically very simple, are the most time consuming. A special slave processor has been designed to carry out these operations under control of the general purpose computer. If implemented, it would cut the simulation time to four times real time. Immediate application to real-time speech synthesis is therefore limited and the system is considered primarily a research tool.

The excitation period serves as a time clock for the synthesis routine, and within each period parameters or parameter increments remain constant. Therefore an excitation period is specified even when the tract is excited by noise or not excited at all as in pauses within the utterance. This allows a relatively rapid, free-running simulation for hundreds of sampling periods before a check is made for new parameter values.

Once the control data have been specified the simulation runs without interaction, and writes the speech output on disk or tape. Thus the time duration of speech material generated in any run is limited only by the control table storage limitations and the patience of the experimenter. Because the pressure distribution along the vocal tract at any time within the utterance cannot be simply recreated, a change in any one target parameter requires the regeneration of the complete utterance.

Control and Use of the Simulation Facility

The control parameters for a particular synthesis task are specified in terms of two parameter lists. The first is a list of articulatory states defined in terms of sets of 13 area values. The second list is a time sequence of targets, nasal coupling values, target interval durations, excitation frequency values, excitation amplitude values, whether periodic or random excitation and if the latter, the place of excitation. Targets are specified by using character-string labels as pointers to entries in the table of articulatory states. Excitation parameters are conveniently separated from specification of the articulatory configurations and each is easily varied independently of the other. The same articulatory configuration may be repeatedly called by its label in the second list without having to specify it again in numerical terms. The program supplies

default values for most parameters, generally unchanged from the previous specification so that only changing parameters need be explicitly specified. Targets may be appended, inserted or deleted from the list conveniently and quickly. A typical sentence, the example discussed later, which consists of 19 phonemes required a sequence of 27 articulatory targets for acceptable synthesis.

The basic mode of data entry is through the on-line typewriter. Once a parameter list satisfactory to the experimenter has been entered, the program generates the output signal and records it digitally on tape or disk. Because the simulation runs slower than real time, the output signal is converted to analog form and played back only after its generation has been completed.

The above method of operation, though perfectly feasible, is not terribly convenient from the point of view of an experimenter not used to numerical specification of articulatory data. To visualize the articulatory configurations and movements between them, a time-motion display of the articulatory trajectory is projected at a rate proportional to the real-time movement rates specified. Any particular configuration may be selected for static display, graphically modified with the aid of a visible pointer, and returned to its slot in the target history of the utterance.

Another display that has found wide application to evaluate the course of articulatory events is one where the sequence of vocal-tract resonant frequencies are displayed on a time base. To ease the computational load, the resonances are computed from the stored shapes¹³ rather than the generated signal. This display resembles a conventional spectrogram in appearance. Based on his wide experience with spectrograms, an experienced speech scientist can rapidly pin-point situations where the acoustic effects of the specified articulatory trajectories deviate appreciably from normal patterns. The visual information he obtains thus complements his aural evaluation of the generated utterance.

Let us now consider in detail simulation of the production of the sentence, "This is a story about a man." This sentence was generated through trial and error procedures by modifying shapes derived from X-ray pictures and referring to the natural utterance for timing values. It demonstrates the quality of output available through careful work with the system. No explicit articulatory rules were used in this example. Instead the

experimenter's accumulated knowledge of articulatory dynamics and the acoustics of speech production were utilized. The automation of this process so that one will not have to resort to trial and error for every different sentence generated is the next task at hand.

In Figure 2 the spectrograms of the synthetic and natural utterances are compared. We transcribe that sentence in terms of phonemic symbols

/ðɪs ɪz ə 'stɔɪ ə 'baʊ ə 'mæn/. The apostrophes indicate stress marks which are considered in specifying the excitation frequency and target interval duration data. The excitation frequency function for the sentence is obtained by modifying the function determined from the intonation pattern. The excitation frequency is increased prior to and decreased relatively more rapidly after stress marks. For the voiced fricatives /ð/ (th) and /z/ we supply, in addition to periodic excitation, turbulent excitation at points 3 cm and 1.5 cm respectively behind the front end of the tract. In each case the vocal-tract shape is relatively constricted at that point. The unvoiced fricative /s/ is excited by noise only at the same point and using the same vocal-tract shape as /z/. The vowels /ɪ/, /e/, /o/ and /æ/ are periodically excited and have their articulatory configurations shaped by modifying the configuration appropriate for the isolated vowel in accordance with the articulatory influences of the neighboring consonants. In turn, the vowels are used to control the variable articulatory characteristics of the consonants. For example, the articulation of the initial /s/ is the same as that of the /ɪ/ except where the tongue tip is raised to produce the constriction. The glide /r/ is vowel-like in excitation and retroflexion of the tongue is not applied. The diphthong /aʊ/ is generated by linear reflection-coefficient interpolation between bounding target values that are somewhat modified versions of the isolated independent vowels /a/ and /ʊ/. The unvoiced stop /t/ has a transition part to an articulation exhibiting almost complete constriction at a point 3 cm from the front. This constricted articulatory state is sustained for a 21 msec interval before release. On release a short burst of noise is introduced to simulate affrication or turbulence; thereafter the amplitude of the periodic excitation is linearly increased from a zero value. The voiced stop /b/ has the articulation changing from that of the previous vowel /ə/ to that of the initial part of the following diphthong /aʊ/ while the lips are kept almost completely constricted. A very small but finite lip opening is helpful in simulating the acoustic effects of

energy radiated through the walls of the mouth. The nasal consonants /m/ and /n/ are produced with the oral tract completely constricted at the lips and 3 cm behind them, respectively, but with a high value of nasal coupling. The coupling is reduced but not eliminated for the intervening /æ/ vowel which is thus made nasalized.

One interesting result observed through trial and error experiments is that naturalness is enhanced by avoiding sustained articulations - those where two adjacement articulatory states are specified as identical. Except for the vowels in the stressed syllables and the fricatives, the articulations are continuously changing. This further invalidates the idea that speech can be produced by concatenating different stationary signals.

Conclusions

The simulation program produces intelligible speech if care is taken to modify articulatory shapes, time durations, and excitation frequency values to fit the contextual environment on a level at least as large as the sentence. Data appropriate for the production of an isolated word but not for the sentence in which it is embedded may make not only that word, but the complete sentence unintelligible. The quality of the material generated so far can be used as a goal for the quality to be achieved by implementing rules governing the movements of the individual articulators.

The program uses an excitation function that is stored and independent of the vocal-tract shape. The articulation, however, is known to affect both the wave-shape and the excitation frequency.¹⁵ The fixed wave-shape independent of frequency is but a rough approximation to the real situation. Detailed simulation of vocal-cord movement would, however, increase the execution time of the program by at least a factor of two. Because these factors are not central to the aspects of speech generation under consideration at the moment, they have not been included in the present program.

The vocal tract is assumed lossless in its interior, therefore, the resulting resonance bandwidths are fixed functions of frequency and independent of the articulation. The losses in general do depend on the surface area of the tract wall and are thus affected by the shape of the tract in the plane perpendicular to the flow directions. The contribution of this factor to lack of naturalness has not been precisely evaluated.

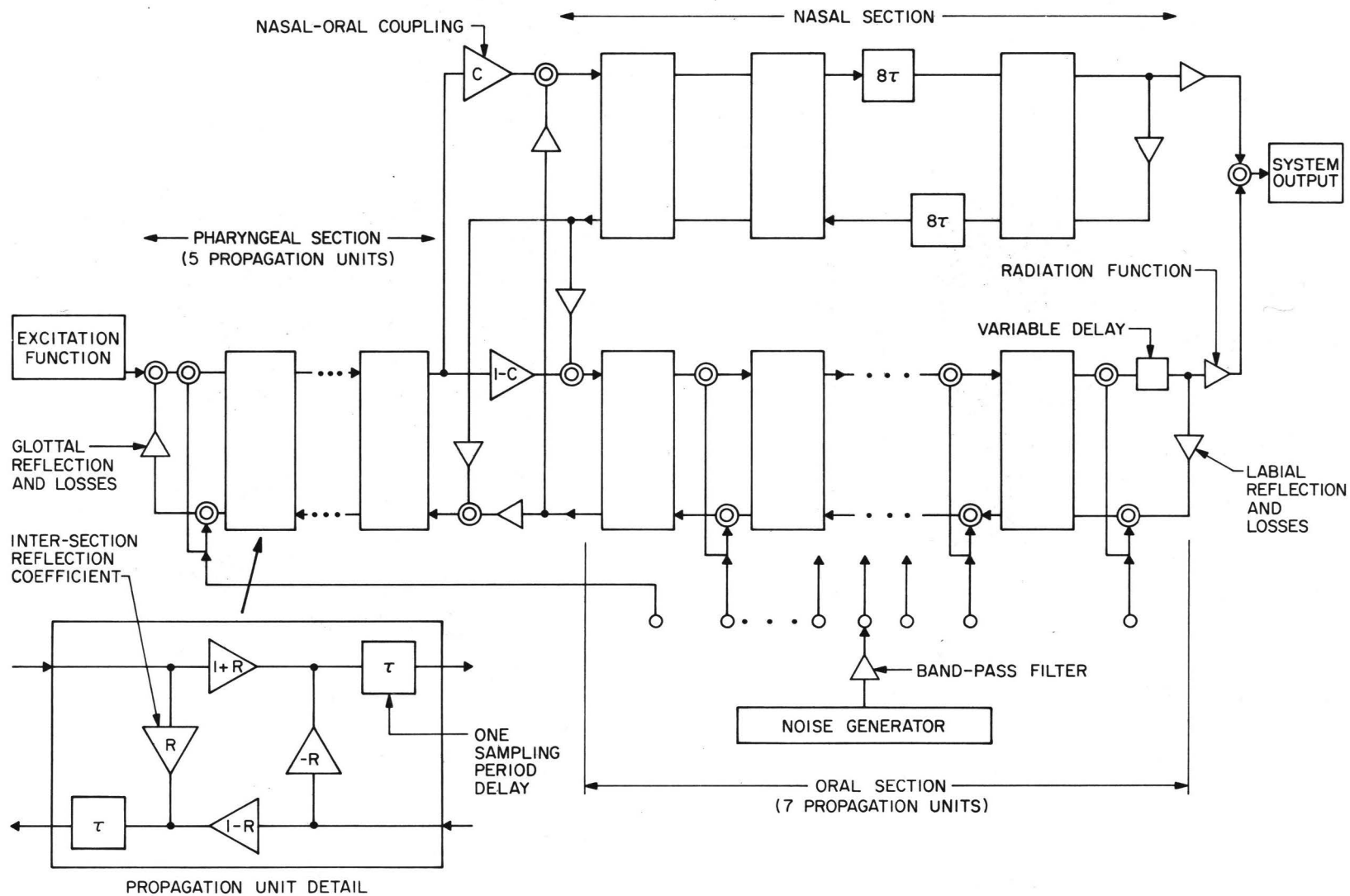
Articulations of /l/ and /r/, in particular, often exhibit sidebranches formed by the tongue and the hard palate that cannot be modeled precisely when one treats the vocal tract as one continuous tube. Our approximations are therefore expected to lead to differences in the acoustic result and must be recognized as limitations of the model.

A program that requires explicit specification of all articulatory and excitation parameters suited to the particular context cannot be said to exhibit extensive artificial intelligence. Instead, the simulation program must be viewed as a framework within which statements regarding articulatory movements can be expressed and evaluated. The framework reported is based on the consideration that utterances are organized on the articulatory level, and that organization can be studied most effectively at the same level.

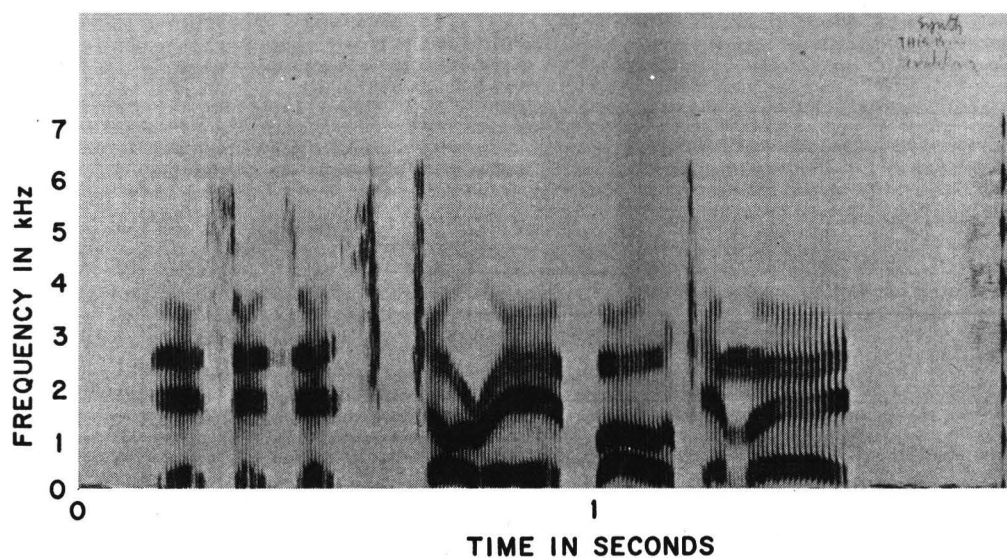
Our results up to this time concern organization at the lowest level. Here articulatory parameters are treated as independent, and interactions between them, except as reflected on the acoustics level, are ignored. For example, the nasal coupling parameter is used to control nasalization of vowels adjacent to a nasal consonant at the same time as the tongue shape is changing as required for the production of those vowels. Thus the adjacent vowels are perceived as nasalized. The production of the same effect on the acoustic level is much more complex. This example substantiates the basic motivation for speech generation through articulatory control, that this direction of attack can unravel the complex organization of the individual phonemes into the integrated utterance.

References

1. Holmes, J. N., Mattingly, I. G. and Shearme, J. N., "Speech Synthesis by Rule," *Language and Speech*, 7, 127 (1964).
2. Rabiner, L. R., "Speech Synthesis by Rule: An Acoustic Domain Approach," *Bell System Tech. Journal* 47, 17 (1968).
3. Lindblom, B., "Spectrographic Study of Vowel Reduction," *J. Acoust. Soc. Am.*, 35, 1773, (1963).
4. Henke, W., "Dynamic Articulatory Model for Speech Production Using Computer Simulation," Ph.D. Thesis, Mass. Inst. of Tech., Cambridge, Mass., 1966.
5. Schroeder, M. R., "Determination of the Geometry of the Human Vocal Tracts by Acoustic Measurements," *J. Acoust. Soc. Am.*, 41, 1002 (1967).
6. Dunn, H. K., "The Calculation of Vowel Resonances and an Electrical Vocal Tract," *J. Acoust. Soc. Am.*, 22, 740 (1950).
7. Stevens, K. N., Kasowski, S. and Fant, C. G. M., "An Electric Analog of the Vocal Tract," *J. Acoust. Soc. Am.*, 25, 734 (1953).
8. Fant, C. G. M., *Speech Communication Research*, Ing. Veitenshaps Akad., (Stockholm), 24, 331 (1953).
9. Rosen, G., "Dynamic Analog Speech Synthesizer," *J. Acoust. Soc. Am.*, 30, 20 (1953).
10. Hecker, M. H. L., "Studies of Nasal Consonants With an Articulatory Speech Synthesizer," *J. Acoust. Soc. Am.*, 34, 179 (1962).
11. Kadokawa, Y. and Nakata, K., "Analysis of Speech by Vocal-Tract Configuration," *J. Radio Res. Laboratories, Japan*, 11, 99 (1964).
12. Coker, C. H., "Speech Synthesis with a Parametric Articulatory Model," paper A-4, *Speech Symposium*, Kyoto, Japan, 1968.
13. Kelly, J. L., Jr. and Lochbaum, C. C., "Speech Synthesis," *Proceedings of the Fourth International Congress on Acoustics*, Copenhagen, 1962.
14. Mermelstein, P., "Determination of the Vocal-Tract Shape from Measured Formant Frequencies," *J. Acoust. Soc. Am.*, 41, 1283 (1967).
15. Flanagan, J. L. and Landgraf, L. L., "Self-Oscillating Source for Vocal-Tract Synthesizers," *Trans. IEEE AU-16*, 57-64 (1968).

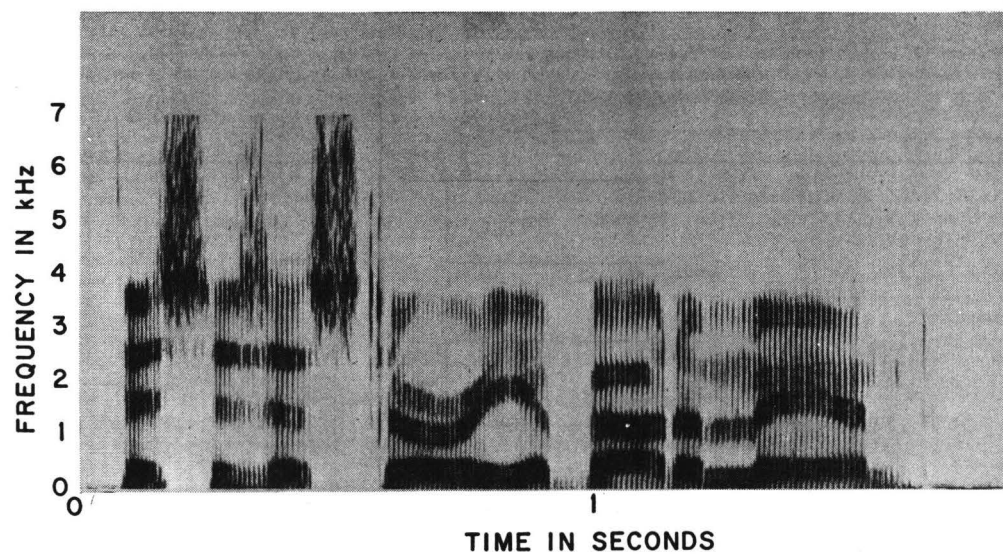


1. Block diagram of the simulator system.



2

MACHINE GENERATED SENTENCE



2

NATURAL SENTENCE

2. Spectrograms of machine generated and natural versions of "This is a story about a man."

A STATISTICAL STUDY OF VISUAL RESOLUTION THRESHOLDS

by

B. L. Hills, R. L. Beurle & M. V. Daniels,

Department of Electrical and Electronic Engineering,
University of Nottingham, England.

Summary

Measurements of threshold contrast required for various resolution tasks at various background light levels are reported. The predictions of a simple statistical model based on photon-noise limited detection are compared with the empirical observations. Correspondence is encouraging provided care is taken to account for the variation of all important parameters such as spatial and temporal integration, pupil area, etc.

Introduction

Since the early work of König in 1879, it has been known that visual acuity, i.e. the ability to discriminate fine detail in an object, progressively deteriorates as the level of illumination falls. Shlaer¹ has shown that this fall in visual acuity is dependent on the type of acuity pattern used. Thus at background luminances above 100 trolands, the gap in a 'Landolt C' pattern is easier to see than bars of a similar width in a parallel bar resolution pattern. At lower luminances, the reverse is true. The ultimate limit of acuity under optimum viewing conditions depends critically on the nature of the task. For vernier acuity, a displacement of 2 seconds of arc between the top and bottom halves of a thin vertical black line has been detected. For the Landolt 'C' and the parallel bar pattern, gaps and bar widths of the order of 25' are the best that have been resolved.

It is generally agreed that the fall in visual acuity at low levels of illumination is connected with the changeover from using the small cone summation areas of the fovea at high adaptation levels, to using the larger rod summation areas of the periphery at low levels. Also, these summation areas are known to increase in size with decrease in the background level, to which the eye is adapted. This will also reduce acuity.

This account is concerned with the recognition of simple patterns of varying contrast at relatively low light levels and is an extension of work on detection previously described by the authors². The empirical observations of detection threshold described in this reference have been compared with the predictions of a simple statistical model for a photon-noise limited detector. This work, which is to be published elsewhere, showed an encouraging correspondence, and the present paper describes the application of the same technique to resolution thresholds.

The Ideal Detector

The original idea of comparing the actual performance of the human visual system with an 'Ideal Detector' is due to Rose³, although similar theories were developed independently by de Vries and also Pirenne. An ideal detector (or ideal picture pickup device) is one whose performance is limited only by the statistical fluctuations in the number of incident photons picked up by the device. Presented with the task of detecting a pattern distinguished by a small luminance, ΔI , superimposed on a uniform background of luminance I , this detector can do no better than count the photons arriving within the area where the pattern is anticipated, and compare this with the mean density of arrival of photons in the background I . Background photons are considered indistinguishable from target photons, and all these incident photons are given equal weight in the output of the device (i.e., it is linear), so that every effectively absorbed quantum is taken into account in the final decision.

In adapting this concept, the ideal detector has been assumed to be subject to certain limitation analogous to those of the eye. For example, it is taken to have the ability to integrate temporally for a period which is assumed to vary as empirical observation suggests it does for the eye. Then again it is assumed that there is spatial integration over an area the extent of which depends on the background light level. Some assumptions must be made about the method of determining I . For example, it may be assumed that the background can be sampled on a sufficient number of occasions or over a sufficient area by the ideal detector for the background luminance I to be known precisely. Alternatively, it may be assumed that the measurement of I is subject to error in the same way as the measurement of $I + \Delta I$.

Now, it is generally accepted that the numbers of absorbed quanta fluctuate according to a Poisson distribution, the deviations from an average absorption of $\bar{n}$ quanta thus having a root mean square value of $\bar{n}^{0.5}$. Thus if $\bar{n}$ quanta are absorbed from the background, it is these fluctuations that interfere with the detection of a small change ΔI in illumination. If this small change in illumination, or 'signal', yields Δn absorbed quanta then, for a given reliability, the threshold of detection is given by

$$\Delta n = k \bar{n}^{0.5} \dots\dots\dots (1)$$

where k is a constant called the threshold signal-

to-noise ratio, its exact value depending on the required degree of certainty in detection. Equation (1) is the basic equation of the fluctuation theory of Rose. It is, however, an approximation, for it ignores the fluctuations in the signal. Thus, a more rigorous form of the equation is

$$\Delta n = k.(\bar{n} + \Delta n)^{0.5} \dots\dots\dots(2)$$

For all except the lowest background levels, this equation approximates to (1), since $\Delta n \gg \bar{n}$, but at very low background levels when this condition does not apply, equation (2) must be used.

In the experiments on detection, it was shown desirable to assume that in what has been called a primary recipient unit² the sensitivity varies radially over the effective summation area over which spatial integration takes place, a fact which has been recognised in the literature^{4,5}. The assumed variation is shown in Figure 1,^{2,5} this curve representing a radial variation of the form

$$S = \frac{S_o r_m^3}{(r_o + r_m)^3} \dots\dots\dots(3)$$

where r_m is a function of the background luminance, as shown in Figure 2. The adoption of this expression has been justified merely for convenience in calculation. It gives a good fit to the range of empirical results which it represents, but there is no reason to think that it has any biological significance. The summation time T and pupil area, α_p , were assumed to vary with background luminance in accordance with the curve shown in Figure 3 and the data given in Table I.

With this knowledge of the spatial and temporal summation characteristics of the retina, it now becomes possible to calculate the number of photons Δn in a single sample taken by the eye from a pattern illuminated by ΔI . The mean number of quanta $\bar{n}$ from the background I may similarly be calculated. Substitution of these in equation (1) now gives an equation for ΔI in terms of background luminance I . As a rough approximation this is of the form

$$\Delta I = K(I)^{0.5}$$

but the precise relationship is appreciably affected by the variation with light level of r_m , T and α_p . Since the technique of calculation is summarised in more detail in a companion paper⁸, it is not proposed to elaborate further on the details here.

It was found that for large objects at higher light levels the primary recipient detector was inadequate to explain the results, and it was essential also to assume the existence of elongated linear "edge" detectors². The evidence suggested that these had a transverse variation of sensitivity similar to the variation across the diameter of the units already referred to. This

could be the result of combining the output of a linear array of primary summation units. There was also evidence for a drop off in sensitivity towards the ends of these long summation areas and a law similar to equation (3) was assumed, replacing r_m with a length constant L_m which was also found to be a similar function of background level. Values assumed for all these parameters are given in Table I which also gives α_p the area of the natural pupil assumed in these calculations.

Since the detailed results of these experiments are to be published elsewhere, it will suffice here merely to say that for the patterns tested, discs, annuli and parallel bars, the correspondence between empirical thresholds and predicted thresholds was encouraging, discrepancies being typically less than 0.2 in logarithmic units to the base 10.

Resolution Measurements

The experiments reported here were designed to compare the threshold contrast for resolution at various light levels of the following pattern shapes:

- Parallel bars
- The "Foucault Fan" pattern
- The double disc pattern
- The 'Landolt C'
- A square equal to the gap in the 'Landolt C'
- A vernier acuity pattern

The shapes of these patterns are shown in Figure 4. Three different sizes of each pattern were used, and the object was to see whether the variations between resolution thresholds for different patterns could to any extent be accounted for by the application of the photon noise theory outlined above.

In these experiments the subjects were given adequate time to adapt to an evenly illuminated background, and the resolution pattern was superimposed on this, the colour temperatures being 3600°K and 3400°K respectively. The resolution thresholds were determined by the subject, who had control of the illumination of the pattern. He was asked to straddle the point at which the resolution feature of the pattern was just visible. It is thus possible to make a distinction between the "detection" threshold at which it is possible to see the pattern as a whole, and the recognition threshold at which it is possible to see the resolution feature (e.g. the bars in the parallel bar resolution pattern. With the exception of the square all the patterns have such a feature, e.g. the gap in the "Landolt C", the existence of two separate discs and the discontinuity in the vertical line. The square was included for comparison with the Landolt C where the task might be described as "seeing" the "missing square" where the gap is present.

The Results

The averaged increment thresholds for the three subjects are given in Table II. For

simplicity, the thresholds for the patterns will be compared two at a time.

One simplifying feature of the results is that the corresponding bar and fan patterns (as defined in Figure 4) have virtually identical increment thresholds. This is true for all three sizes and all background luminances. This is because the pointers indicating where 'z' is measured are centrally placed on the fan pattern. It has been found that if the pointers are used too near the top, the increment threshold for the fan is higher than for the bars. Thus, with the proviso that the pointers are used in the central region, the comparisons of performance that will be made with the parallel bar pattern would be the same if the comparisons had been made with the fan pattern.

In Figure 5, the first comparison is made between the increment thresholds for resolution of the Landolt 'C' and the parallel bar patterns. It is first noted that the increment thresholds for the bar patterns are considerably lower than the equivalent Landolt 'C' as defined in terms of z in Figure 4. There are major differences between the thresholds for these two patterns. For the large and medium sizes of both patterns, the increment threshold progressively decreases as the background level is lowered, but for the small patterns, the increment threshold is constant for background luminances of 10^{-4} mL and below.

It will be seen that vertical displacement of the curves for the bar pattern by 0.7 log units gives reasonably close fits to the curves for the Landolt C. This is true to some extent for all the patterns, excluding the squares, and might suggest that once some allowance has been made for the overall inefficiency factor of each pattern, they would all give much the same results. This is, however, believed to be an oversimplification, and indeed is only a rough approximation. Compared directly, there is little similarity between the results. For example, the large Landolt C and the small bar pattern require the same contrast threshold for resolution at a background level of -1.80 Log mL. Below this adaptation level, the threshold of the large Landolt C becomes progressively lower relative to that of the small bar pattern, while above this level, the reverse is true. Also for the highest background level, a contrast can be found for which all the bar patterns but none of the Landolt C's can be resolved, whereas at the lowest background luminances a contrast level which just allows all the bars patterns to be resolved, only leaves the small Landolt C unresolved. This again shows that the bar pattern becomes more visible in relation to this 'white' Landolt C if the background level is raised, at least for the range of background luminances used here. It will be suggested later that this is because the Landolt C is resolved using primary recipient units, while the bar pattern is resolved by line/edge detectors.

In Figure 6, the increment thresholds for the resolution of the 'double disc' and bar patterns are compared. The comparison shows some features similar to that between the Landolt C's and parallel bars. The double disc in general has a higher threshold, and the threshold increases relative to the bar pattern as the background level is raised. At the lowest background levels the thresholds for the double disc are much closer to the thresholds of the parallel bars than to those of the Landolt C. Thus, if we attribute the behaviour of the double disc to primary recipient units and that of the bar patterns to line/edge detectors, we must also explain the difference between Landolt C and double disc.

The increment thresholds for the resolution of the Landolt C and the detection of the squares are compared in Figure 7. For background luminances of 10^{-3} mL and above, the two sets of curves follow each other very closely. This suggests that in this region, resolution of the gap in these white Landolt C's is determined by the detectability of a square decrement of light equal in area to the gap, viewed against a similar background. This presupposes that increment and decrement thresholds for detection of a square are very similar. Below 10^{-3} mL, the increment thresholds for the squares fall progressively below those of the Landolt C. The curves for the squares are the steepest of all the patterns at the very low background levels.

In Figure 8, the increment thresholds for the resolution of the broken line are compared with those for the detection of the square. For background luminances above 10^{-4} mL, the curves are sufficiently close to suggest that the resolution of the broken line under these conditions is limited by the detectability of a square patch of light of side equal to the displacement between the top and bottom halves of the line. An interesting experiment might therefore be to compare the limit of vernier acuity as it is normally determined with the size of the smallest detectable black square, for a wide range of background luminances.

For background luminances above 10^{-3} mL, it will be seen from Table II that the increment thresholds for resolution of the medium double disc pattern and detection of the large square are very close. The same is true for the small double disc and medium sized square at background luminances above 10^{-2} mL.

During the experiment, the subjects were asked whether they could still detect the presence of a pattern when its increment in luminance had been lowered so that it could no longer be resolved. Their verbal reports indicate that at high background levels, they could at best only detect the patterns very faintly in this non-resolving condition. Particularly with the large and medium sizes of pattern, resolution came with detection. At the very low background luminances, most of the large patterns were still only faintly detectable, but the large Landolt C was bright enough to be readily detectable.

The medium sizes of pattern were all readily detectable, and the small sizes were very bright indeed compared with the background, and yet not resolved.

Extension of the Detection Model to Resolution

Thus far, the results have only been discussed in qualitative terms. In this section, a more quantitative approach is taken by extending the model of detection referred to earlier, to cover the calculation of resolution thresholds. Theoretical curves of increment threshold for resolution, ΔI_R , against background luminance, I , are derived for the parallel bars, Landolt C and 'discs' patterns, and the effect of different assumptions is examined.

For detection of an increment in luminance ΔI , on a background luminance I the response of a detector optimally positioned with respect to the stimulus so as to sample $I + \Delta I$, was compared with the response of a detector sampling the background, I , alone. It was argued that the difference between the means of these two responses was the 'signal', and that this signal must exceed the combined fluctuations in these two responses by a constant ratio in order for the signal to be detected.

For resolution, it is proposed that the responses to be compared should be from two identical detectors, X and Y , which are centred over different regions of the resolution pattern. For the parallel bar pattern, one is centred over a bar, and the other is centred over an adjacent space between the bars. For the Landolt C, one is centred over the gap, and the other over a segment of the ring. In the case of the double disc pattern, one is centred on one of the discs, the other over or near the gap between the two discs. Again, the difference between the mean responses of the two detectors is taken as the 'signal' and it is argued that this must exceed the combined fluctuations of the two responses by a constant ratio in order for the pattern to be resolved.

Predicted Resolution of Parallel Bar Pattern

For reasons discussed in the last section, the resolution of the bars in the parallel bar patterns will be assumed to be performed by line detectors. Figure 9(a) represents a line detector, X , centred on a bar and a line detector, Y , centred on an adjacent space between bars. The response of each detector to each point of the bar pattern is proportional to the product of the light intensity and the sensitivity of the detector at that point. Thus the effective areas, $\alpha'_E(X)$ and $\alpha'_E(Y)$ of the bar pattern in causing a response in the detectors X and Y respectively, may be found by integration. The 'signal' now becomes

$$\Delta n = C \Delta I \alpha_p T (\alpha'_E(X) - \alpha'_E(Y))$$

where C is a constant embodying geometrical factors and the number of quanta per unit of light. The integrated background which contributes to noise becomes

$$n + \Delta n = C \alpha_p T_E (2(I + I_D) r_m L_m + \Delta I (\alpha'_E(X) + \alpha'_E(Y)))$$

where I_D representing an "eigengrau" is unimportant except near absolute threshold. Thus from equation (2), by putting $Ck = K'$

$$\Delta I_R = \frac{K'}{(\alpha_p T)^{0.5} (\alpha'_E(X) - \alpha'_E(Y))} \times (2(I + I_D) r_m L_m + \Delta I_R \alpha'_E(X) + \Delta I_R \alpha'_E(Y))^{0.5} \dots (4)$$

At all but the lowest background luminances, the terms in I_D and ΔI_R on the right hand side of this equation are negligible.

The values of r_m , L_m , T and α_p were first all considered to be set by the background luminance alone. The theoretical curves thus derived are compared with the empirical data in Figure 10. The values of K' and I_D used were -1.22 and 7×10^{-7} mL respectively. It will be seen that the theoretical curves fit the empirical points well for the whole range of backgrounds. However, this good fit is thought to be misleading for reasons explained later.

In the section 'Results', it was noted that at low background luminances, the small bar pattern had to be very bright compared with the surrounding background, for it to be resolved. Therefore, in Figure 11, the theoretically fitted curves have been calculated assuming r_m , L_m , T and α_p are set by $(I + \Delta I/2)$, the mean luminance within the bar pattern. The fit for the medium and large patterns is hardly affected, but for the small pattern at low background luminances, the theoretical prediction is slightly optimistic.

In Table III, the absolute threshold values calculated from equation (4) assuming r_m , L_m , T and α_p are set by $\Delta I/2$ are compared with the measured values. Again, the fit is good for the large and medium patterns, and the prediction a little optimistic for the small bar pattern at low background luminances. This will be referred to later.

Resolution of the Landolt C Pattern

For reasons discussed in the last section, we shall take it that the position of the gap in the Landolt C is discovered by light integrated in primary recipient units with radial symmetry as in equation (3). Figure 9(b) indicates the positions of the centres of the two units whose responses are compared in order to resolve the pattern. Unit Y was centred over the gap of the C. In the first case, unit X was considered to be diagonally opposite Y , since this would give the maximum difference in response. The effective areas $a_E(X)$ and $a_E(Y)$ of the pattern

causing responses in the units X and Y respectively were calculated by numerical integration.

By adapting equation (4), the increment threshold for resolution, ΔI_R , is given by

$$\Delta I_R = \frac{K}{(\alpha_p T)^{0.5} (\alpha_E(X) - \alpha_E(Y))} \times \\ (2(I + I_D) \pi r_m^2 + \Delta I_R \alpha_E(X) + I_R \alpha_E(Y))^{0.5} \dots (5)$$

At all but the lowest background levels, the terms in I_D and ΔI_R on the right hand side of the equation are again negligible.

Figure 12 compares the empirical data with the theoretical curves assuming that r_m , L_m , T and α_p are functions of I alone or of $(I + \Delta I)$. The values of K and I_D used were -1.22 log units and 7×10^{-7} mL respectively. For both sets of assumptions, the theory predicts thresholds considerably lower than were in fact measured.

It was therefore arbitrarily decided to derive the theoretical curves assuming that the comparison unit, X, was located only one gap's width away from the unit, Y, on the Landolt C ring (see Figure 9(b)). In Figure 13, the resulting theoretical curves are compared with the experimental data. The same values of K and I_D were used. r_m , L_m , T and α_p were considered functions of I only. The fit is still not good, but is considerably better than that found when the unit X is diagonally opposite Y. The truth probably lies somewhere between, and it may be possible to improve this theory of the resolution threshold of the Landolt C by more refined positioning of the unit X. The unit should perhaps be positioned a fixed distance from Y rather than a distance related to the gap width.

Resolution of Double Disc Pattern

Resolution is again assumed to be performed by primary recipient units. Figure 9(c) represents the two cases considered. In both conditions, unit X was located at the centre of a disc, (see Figure 9). Unit Y was located (i) at the centre of the gap between the two discs, or (ii) at a distance of 1.64 times the disc radius from the centres of both discs. The effective areas $a_E(X)$ and $a_E(Y)$ of the discs in causing a response in the units X and Y respectively, were calculated by numerical integration. Theoretical increment thresholds for resolution were calculated for case (i) and case (ii) by substitution in equation (4).

Case (ii) provides the larger signal difference between the X and Y units and thus predicts the higher sensitivity. In Figure 14, the case (ii) predictions are compared with the empirical data for a value of K of -1.16 log units. The fit is reasonable at background luminances below 10^{-2} mL, but above this adaptation level, the experimentally determined increment thresholds are lower than those predicted. It is possible

that part of the discrepancy at high background luminances may be due to edge detection superseding area detection. However, if K had been taken as 1.22, as in Figures 10, 11 and 12, the agreement would have been close at high light levels, but the predicted thresholds would have been 0.6 log units too low at low light levels. A possible reason for this will be discussed later.

The theoretical values derived for case (i) with the unit Y positioned centrally between the two discs have not been plotted because the results showed no improvement.

Detection of Squares

Increment thresholds for the area detection of squares were calculated using equation (5), simplifying integration by assuming that a square is as detectable as an equal area disc for this type of detection. It can be shown numerically that the error in the calculated increment threshold in making this assumption is less than 0.5%. The theoretical increment thresholds thus derived are compared with those obtained experimentally in Figure 15. The values of K and I_D used were -0.98 log units and 7×10^{-7} respectively. The former was chosen to give the best approximation at low light levels and for background luminances below 10^{-2} mL the predicted values are in good agreement with the empirical measurements, but above this background level we find that the predicted thresholds are about 0.25 log units higher than those measured. To test whether or not this discrepancy is due to edge detection predominating at these light levels, theoretical edge detection curves were derived using equation (4), with the value of K' made equal to that of K used in the case of area detection. The theoretical area detection. The theoretical area and edge detection curves for the squares are shown in Figure 16. It will be seen that edge detection only supersedes area detection for the large square at the highest background luminance. Thus, the simple edge detection so far developed cannot explain the discrepancies at high background luminances. Summation of the four border effects is a possible factor.

A feature of the theoretical curve fitting to the data of this experiment is that the value of K which has been chosen for the squares (-0.98 log units) is considerably lower than the value of K' used for the resolution of the bar pattern (-1.22 log units). Indeed, if the value $K = -1.22$ log units had been taken, the fit would have been good at high levels of background luminance, at the expense of the fit at the lowest light level, which would then be optimistic by 0.24 log units.

General Comments

We have seen that depending on the choice of K , it is possible to fit theoretical curves either at the upper end or at the lower end of the background luminance range covered in these experiments.

If K is chosen for a good fit at the lower end, the theory is a little pessimistic at the upper end. If K is taken as -1.22 log units for the best compromise at the upper end, the theory is slightly optimistic at the lower end.

The discrepancy of up to 0.25 log units is in fact a very small one, considering that we are concerned with a range of five orders of background luminance, but it is, nevertheless, of interest to consider possible reasons for it. There is firstly the possibility that the inherent sensitivity or photon detecting efficiency changes with light levels. There is in fact a change from the use of rods at the lower end of the range used in these experiments, to the predominant use of cones at the upper end. This could readily alter the balance between the receptor sensitivity at the lower and upper ends of the range, by the small amount in question.

Secondly, there is some evidence that the inverse cubic law adopted for convenience in describing the distribution of sensitivity across a primary recipient unit may be unduly sharp in the centre. Moreover, the use of the same law to describe the transverse variation of sensitivity across each component of a boundary detector may also be unjustified, because almost any means by which the outputs from primary recipient units were combined to form linear detectors would introduce at least some rounding of the peak.

The data which originally justified the use of the inverse cubic law is not very sensitive to the sharpness of the central peak, but resolution threshold data, such as that for parallel bars, is very sensitive to the peak when the width of the bar is small compared with r_m . This could account for the discrepancy between the predicted and observed thresholds for the small parallel bar patterns at low light levels in Figure 11. The modification to the peak of the inverse cubic curve necessary to allow for this is indicated by the dotted curve in Figure 1. Some comparisons have also been made with published data on contrast sensitivity functions⁶ which support the idea of a slightly more rounded peak than the pure inverse cubic curve employed, and the question is being investigated further. At high light levels the rounding may be partly due to the influence of the optical system of the eye⁵.

Thirdly, there is the question of boundary detection taking over from the radially symmetrical primary recipient units. This has been considered earlier in relation to the square pattern, but there is some uncertainty regarding the relative sensitivity of the two processes and this can influence the balance between recognition thresholds at the two ends of the scale.

Discussion

The outstanding fact about the results reported here is that, within the limits of the

slight discrepancies which have just been discussed, it is possible to work on the assumption of such a uniform sensitivity of photon detection. Thus, once one has recognised the limitations of temporal and spatial summation in the retina, the efficiency of utilisation of the information in the incoming photon flux appears to be very nearly uniform over a very wide range of light levels. Considering the variety of visual tasks included in these experiments the uniformity of behaviour is surprising. Thus the theory gives some insight into the reasons for the characteristic differences in observed thresholds between the different forms of resolution test pattern.

The theory and its application to the experimental data is aimed at discovering how efficiently the eye uses the light in the retinal image. It has not primarily been concerned with the detailed nature of the mechanism by which the eye performs the operation. Thus, there is no attempt to identify the cause of the variation of r_m , I_m , etc. as a function of light level. Nor has there been any attempt to explain what neural mechanism would provide such uniform efficiency of extraction of information from the incoming light. This is an interesting field for conjecture, but there are many alternative possibilities to choose from, and a proper discussion of these must await a later occasion.

It is encouraging to find in the need to invoke line/edge or 'boundary' detectors some measure of correspondence with recent physiological findings⁷. For the present, whatever may be the biological significance of the summation areas and edge detectors, their practical significance as parameters determining the performance of the eye is a real one. They indicate clearly the nature of the limitations under which the eye operates when detecting patterns under conditions of poor visibility. Moreover, the concept of a photon noise limited threshold is supported by further experiments reported separately⁸ in which artificial noise, deliberately introduced into retinal images, appears to have an analogous effect on detection.

The work reported here has been on relatively simple patterns and, perhaps in consequence, fairly simple detecting mechanisms have come to light. The situation may well be complicated when one considers more complex patterns, especially if these are seen against an inhomogeneous background. It may well be that there exist more specialised detectors for particular patterns or features of patterns. Indeed, recent physiological evidence suggests there may be, but their existence cannot be deduced from these experiments. Then again, both types of detector discussed here are taken to represent the mean performance averaged over all the detectors which take part in the recognition of different patterns on different occasions. It is, however, clear from physiological evidence that there is appreciable variation between the response of different

ganglion cells and this may make the concept of an "average" detector shape of less value when more complex patterns are considered. The problem may resolve itself round the question of connectivity between less and more complex detectors. If the more complex detectors are formed by combining the outputs of simpler detectors then the former may be expected to be subject to the inherent limitations of the latter. If, on the other hand, more complex detectors can obtain their information directly from receptor cells, these limitations may not apply. The facts will only become apparent with further work.

References

1. Shlaer, S.: "The relation between visual acuity and illumination"; J. Gen. Physiol. 21, 165-188 (1937).
2. Beurle, R.L., Daniels, M.V. & Hill, B.L.: "Visual pattern detection"; I.E.E. conference on Pattern Recognition, Publication 42, published 1968 by the Institution of Electrical Engineers (Great Britain).
3. Rose, A.: "The relative sensitivities of the television pick-up tubes, photographic film and the human eye"; Proc. Inst. Radio Engineers, 30, 293-300 (1942).
Rose, A.: "The sensitivity performance of the human eye on an absolute scale"; J. Op. Soc. Am., 38, 196-208 (1948).
4. Blackwell, H.R.: "The effects of certain psychological variables upon target detectability. Eng. Res. Inst. Report 2455-12-F. University of Michigan (Ann Arbor, Michigan) (1958).
5. Fry, G.A.: Physiological irradiation across the retina, J. Op. Soc. Am., 55, 108-111 (1965).
6. Campbell, F.W. & Robson, J.G.: "Application of Fourier analysis to visibility of gratings"; J. Physiol. 197, 551-566 (1968).
7. Hubel, D.H. & Wiesel, T.N.: "Integrative action in the cat's lateral geniculate body"; J. Physiol. 155, 385-398 (1961).
Hubel, D.H. & Wiesel, T.N.: "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex"; J. Physiol. 160, 106-154 (1962).
Hubel, D.H. & Wiesel, T.N.: "Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat"; J. Neurophysiol. 28, 229-289 (1965).
8. "The effect of noise on visual pattern recognition". A companion paper by the same authors. International Conference on Artificial Intelligence, 1969.
9. Graham, C.H. & Kemp, E.H.: "Brightness discrimination as a function of the duration of the increment in intensity"; J. Gen. Physiol., 21, 635-650 (1938).
10. Graham, C.H. & Margaria, R.: "Area and intensity-time relationship in the peripheral retina"; Am. J. Physiol. 113, 299-305 (1935).
11. Barlow, H.B.: "Temporal and spatial summation in human vision at different background intensities"; J. Physiol. 141, 337-350 (1958).

<u>Log I (mL)</u>	<u>r_m</u>	<u>L_m</u>	<u>T_E(secs)</u>	<u>α_p(sq.mm)</u>
-∞	71'	280'	0.140	40.6
-5.05	40'	200'	0.140	38.8
-4.50	36'	200'	0.140	37.5
-3.92	32'	200'	0.130	35.8
-3.50	28'	200'	0.120	33.9
-3.00	22.5'	200'	0.107	32.0
-2.50	17.7'	168'	0.096	29.1
-1.80	10.6'	133'	0.081	25.0
-1.50	8.5'	119'	0.076	23.3
-0.80	5.3'	94'	0.065	19.2
+0.40	3.2'	63'	0.043	13.5

Table 1 Values of r_m, L_m, T_E and α_p used in the calculation of theoretical curves from Equations (4) and (5). Luminance measured in millilamberts and visual angle in minutes of arc.

	<u>Large</u>	<u>Medium</u>	<u>Small</u>
Predicted	-5.50	-4.81	-3.98
Measured	-5.50	-4.85	-3.52

Table III Comparison of predicted and measured absolute thresholds, in log (mL), for resolution of the parallel bar gratings.

<u>Pattern</u>	<u>Background Luminance</u>						
	<u>(1)</u>	<u>(2)</u>	<u>(3)</u>	<u>(4)</u>	<u>(5)</u>	<u>(6)</u>	<u>(7)</u>
Large Bars	-5.48	-5.07	-4.62	-4.16	-3.50	-2.79	-1.90
Large Fan	-5.38	-4.96	-4.50	-4.02	-3.39	-2.77	-1.85
Large C	-4.46	-4.36	-3.91	-3.53	-2.88	-2.27	-1.39
Large Discs	-5.19	-4.80	-4.24	-3.65	-3.03	-2.45	-1.52
Large Square	-4.95	-4.57	-3.97	-3.45	-2.82	-2.35	-1.40
Large Line	-4.56	-4.38	-3.89	-3.46	-2.96	-2.43	-1.64
Medium Bars	-4.85	-4.59	-4.18	-3.78	-3.16	-2.74	-1.81
Medium Fan	-4.71	-4.31	-4.06	-3.75	-3.12	-2.67	-1.81
Medium C	-3.79	-3.47	-3.34	-3.08	-2.73	-2.27	-1.37
Medium Discs	-4.54	-4.33	-3.84	-3.38	-2.90	-2.36	-1.44
Medium Square	-4.51	-4.11	-3.65	-3.14	-2.62	-2.17	-1.34
Medium Line	-4.05	-3.91	-3.63	-3.14	-2.68	-2.19	-1.41
Small Bars	-3.52	-3.44	-3.42	-3.27	-2.92	-2.52	-1.67
Small Fan	-3.52	-3.56	-3.39	-3.23	-2.89	-2.49	-1.68
Small C	-2.77	-2.65	-2.79	-2.63	-2.30	-1.96	-1.17
Small Discs	-3.41	-3.46	-3.20	-2.89	-2.62	-2.15	-1.32
Small Square	-3.96	-3.57	-2.95	-2.59	-2.19	-1.79	-1.04
Small Line	-3.22	-3.13	-2.90	-2.60	-2.33	-1.86	-1.04

Table II Increment thresholds, ΔI_R, in Log (mL), for resolution of the six patterns.

Background luminances (mL): (1) zero; (2) 7.4×10^{-6} ; (3) 1.2×10^{-4} ; (4) 9.4×10^{-4} ; (5) 1.5×10^{-2} ; (6) 1.5×10^{-1} ; (7) 2.5.

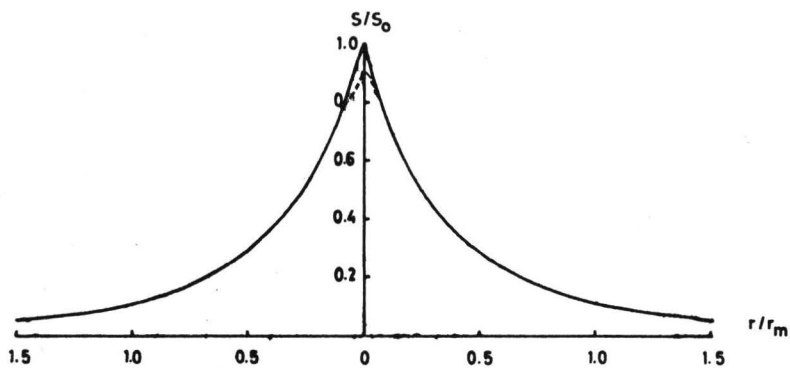


Fig.1 Sensitivity distribution across a detector whose sensitivity falls with distance from its centre according to the inverse cubic law adopted.

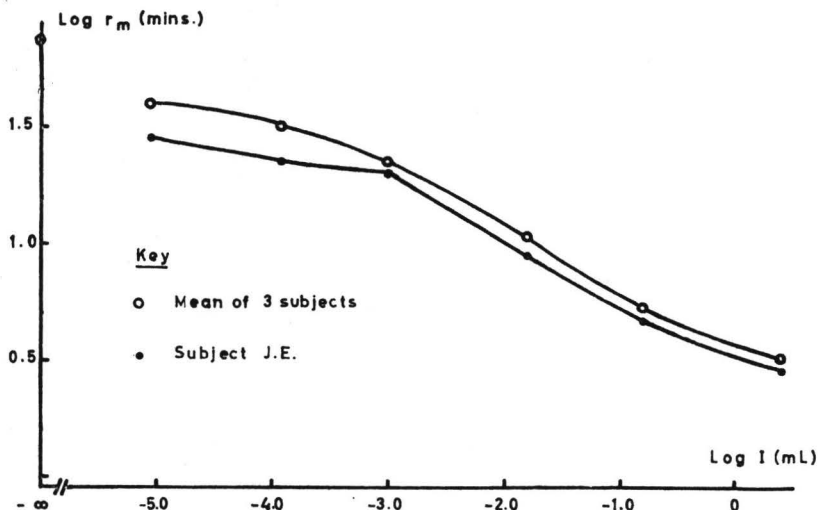


Fig.2 Variation of r_m with background luminance, I , derived from the data of Fig.2.

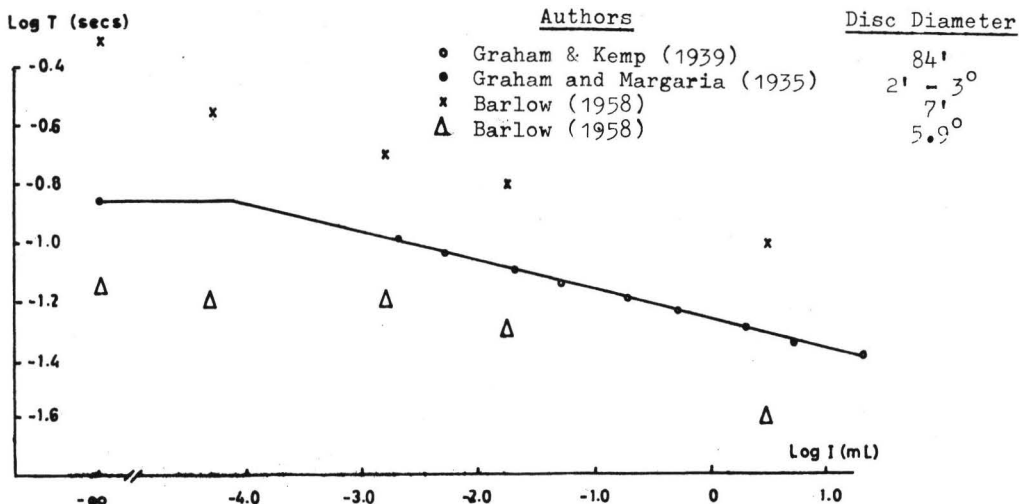


Fig.3 Variation of effective summation time, T_e , with background luminance, I , from three sets of data.

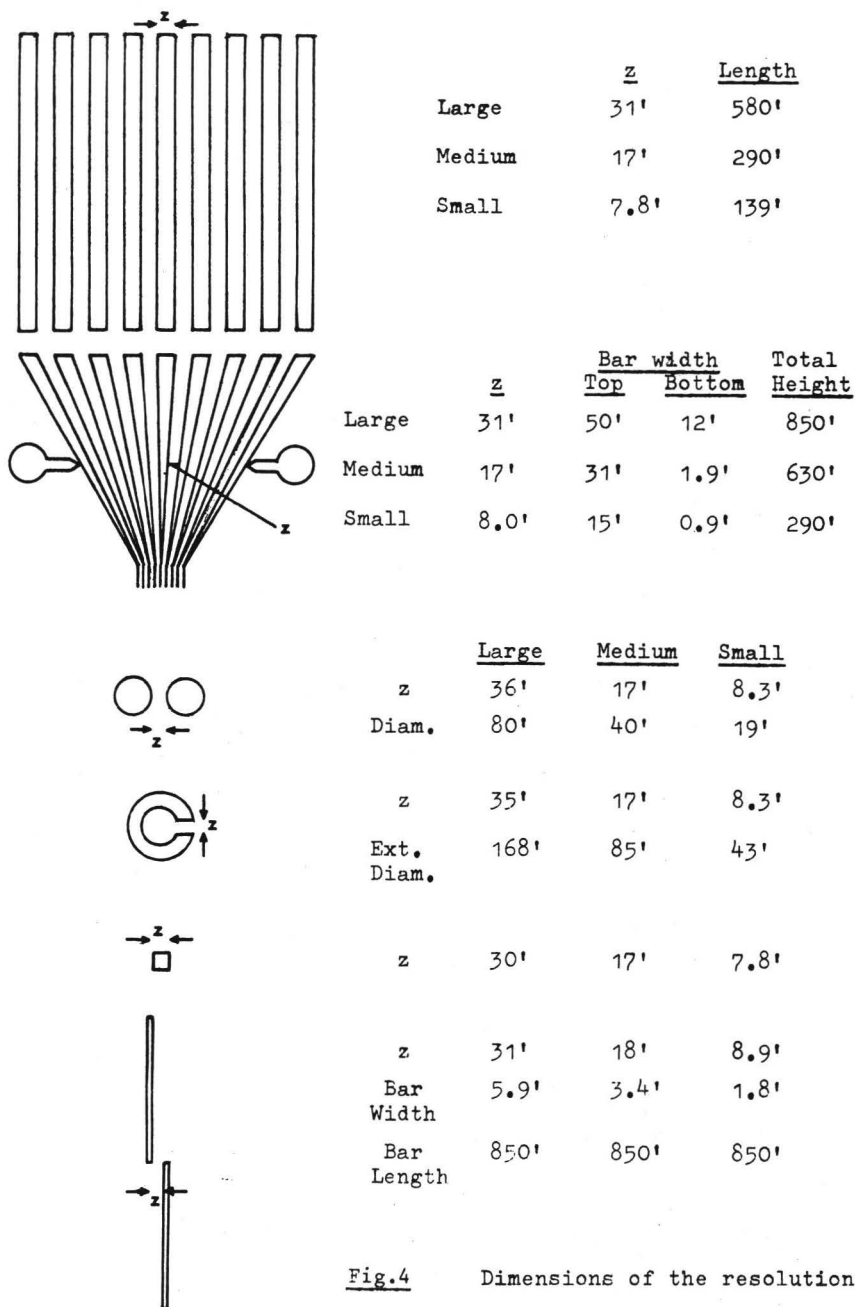


Fig.4 Dimensions of the resolution patterns.

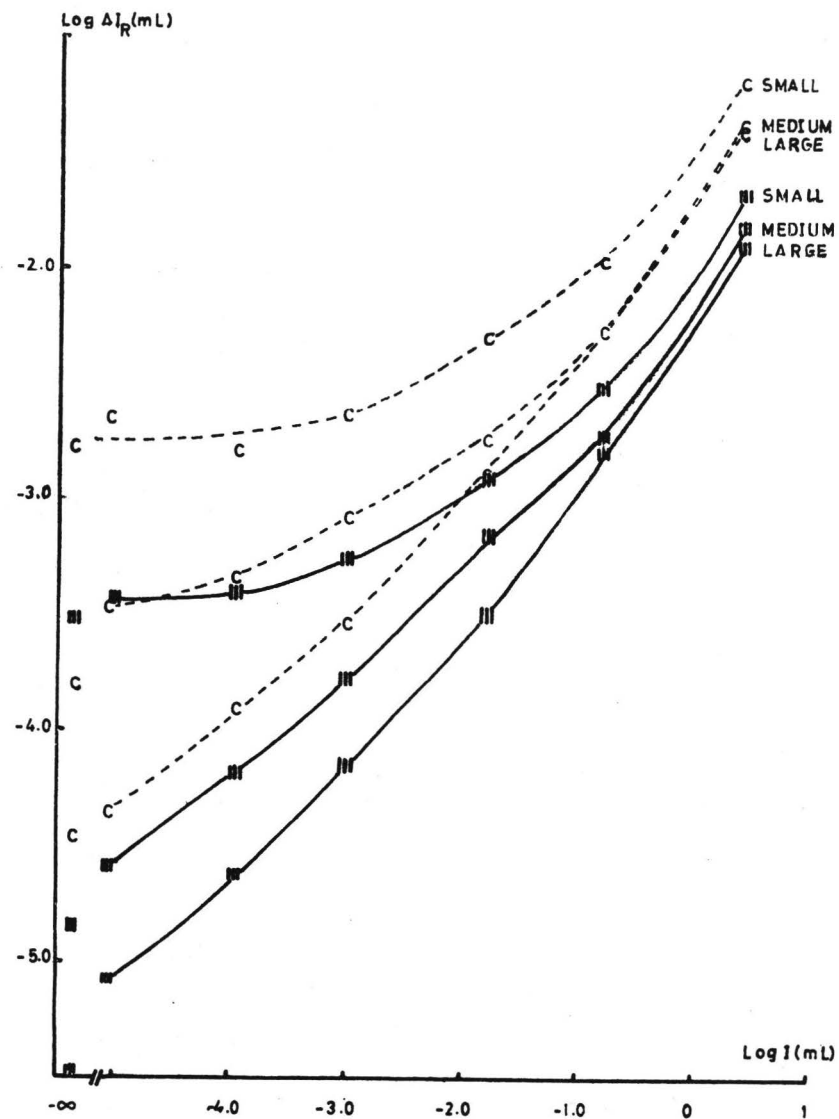


Fig.5 The variation of the increment threshold for resolution, ΔI_R , with background luminance, I . Comparison between the parallel bar gratings (III) and Landolt 'C's (C).

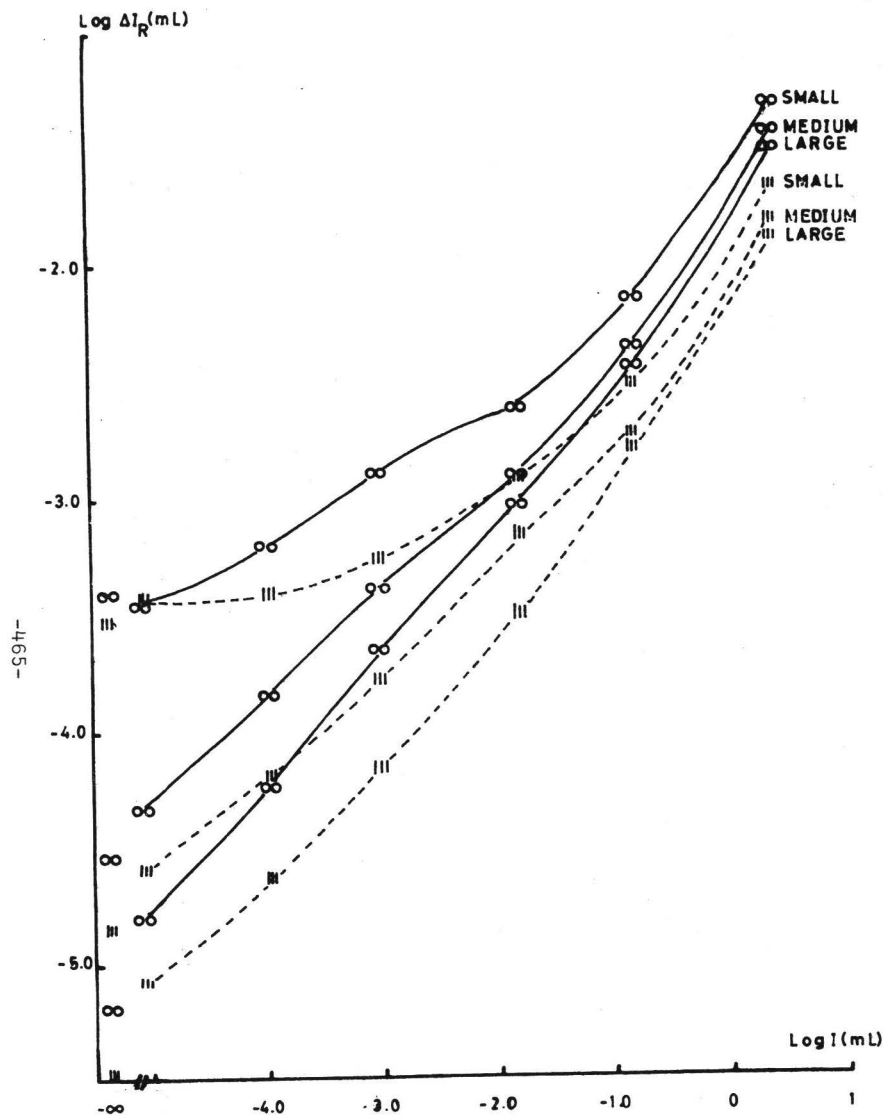


Fig.6 Variation of the increment threshold for resolution, ΔI_R , with background luminance, I . Comparison between the parallel bar gratings (|||) and discs pattern (∞).

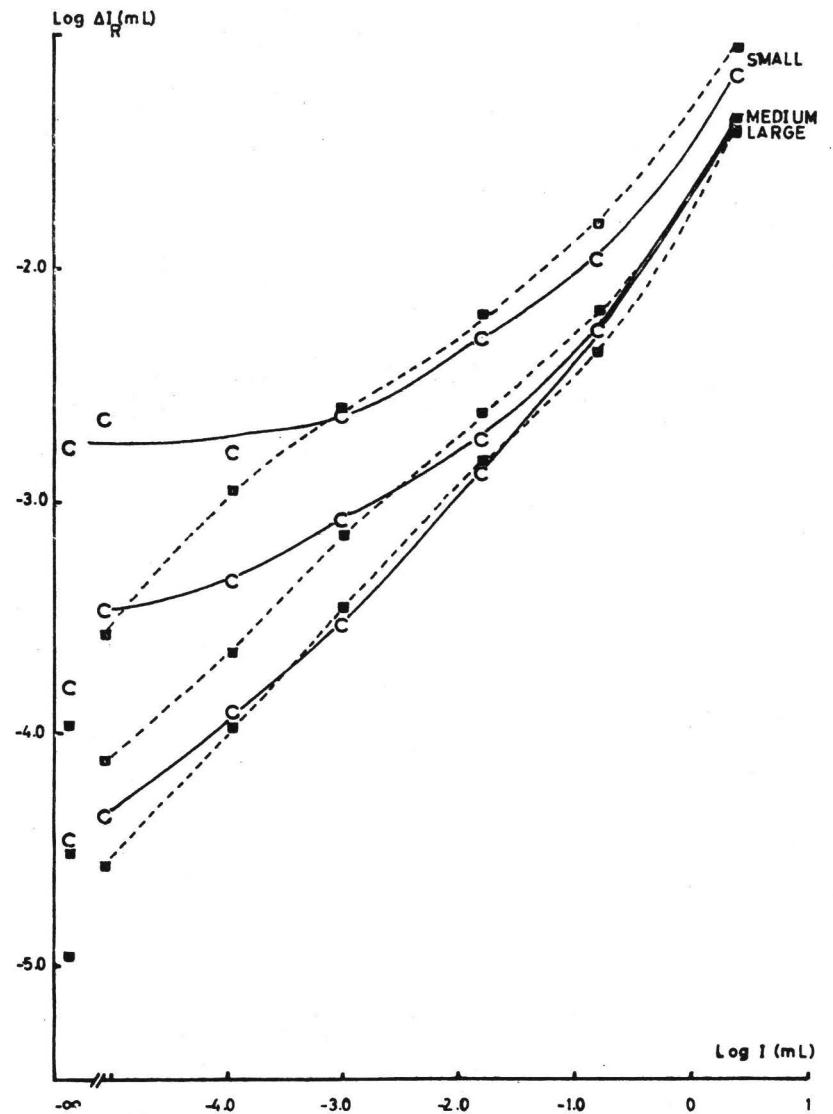


Fig.7

Variation of the increment threshold for resolution, ΔI_R , with background luminance, I . Comparison between the Landolt C's (C) and squares (■).

For the squares the thresholds are for detection.

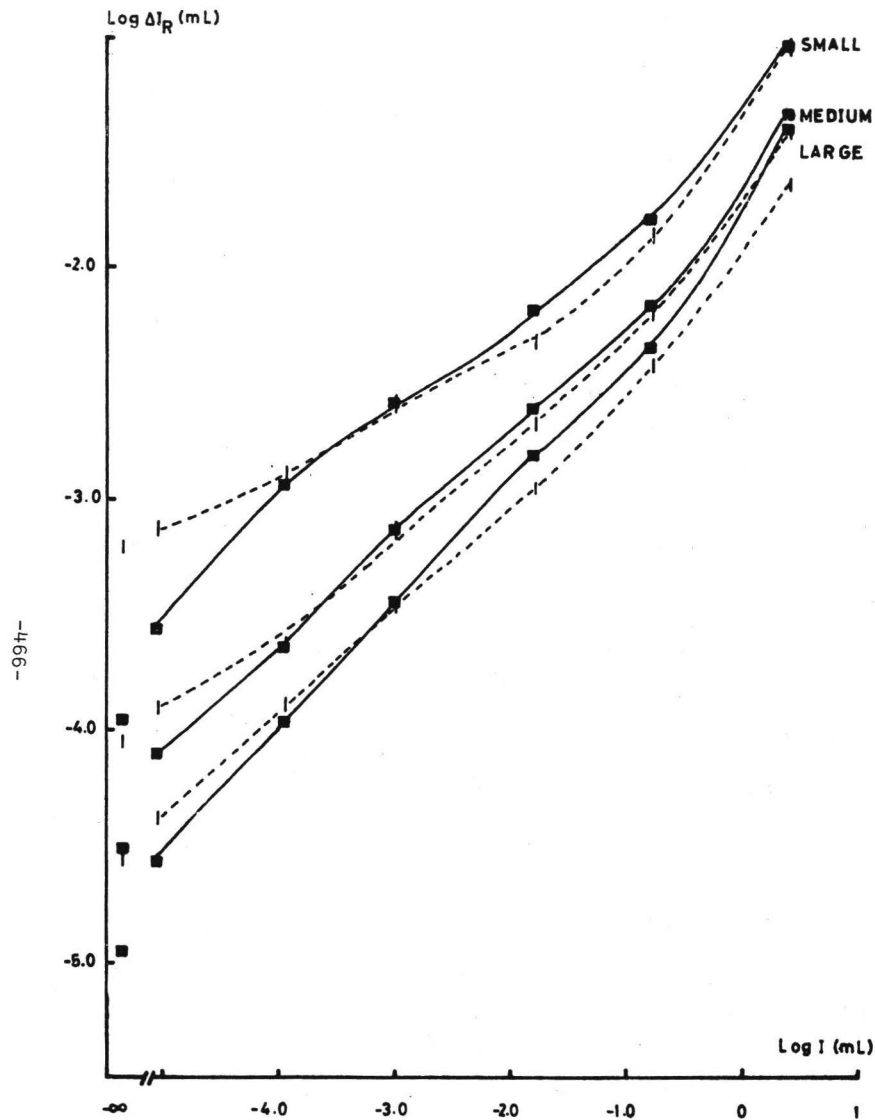
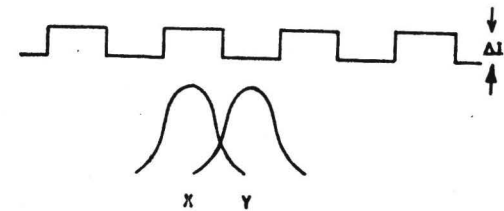


Fig.8 Variation of the increment threshold for resolution, ΔI_R , with background luminance, I . Comparison between the broken line patterns (I) and squares (■). For the squares, the thresholds are for detection.



(a) Parallel bars



(b) Landolt C



(c) Discs pattern

Fig.9 Positions of the comparison detectors X and Y considered in deriving theoretical increment thresholds for the resolution of the parallel bar gratings, Landolt C's and discs patterns.

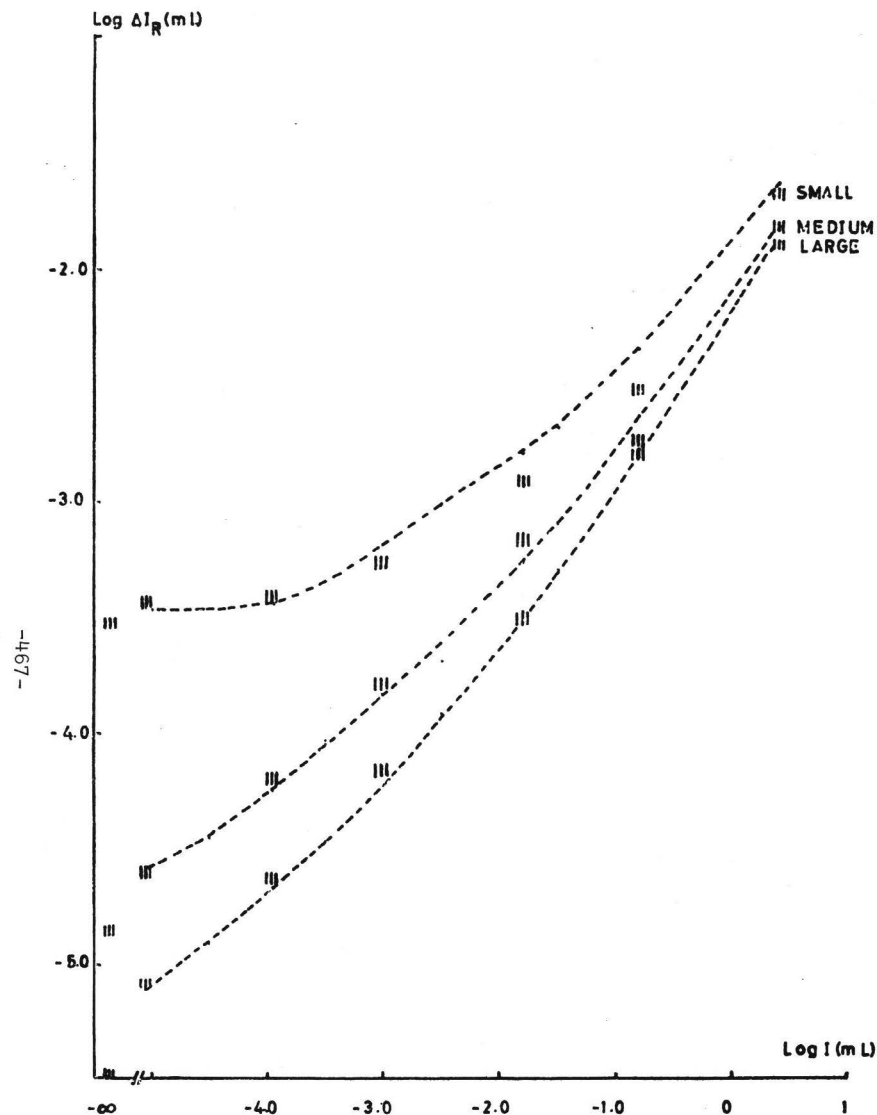


Fig.10

Experimental increment thresholds for the resolution of the parallel bar patterns fitted with theoretical curves derived from equation (3).
 $K' = -1.21$. r_m , L_m , T and α_p assumed to be set by background luminance, I , alone.

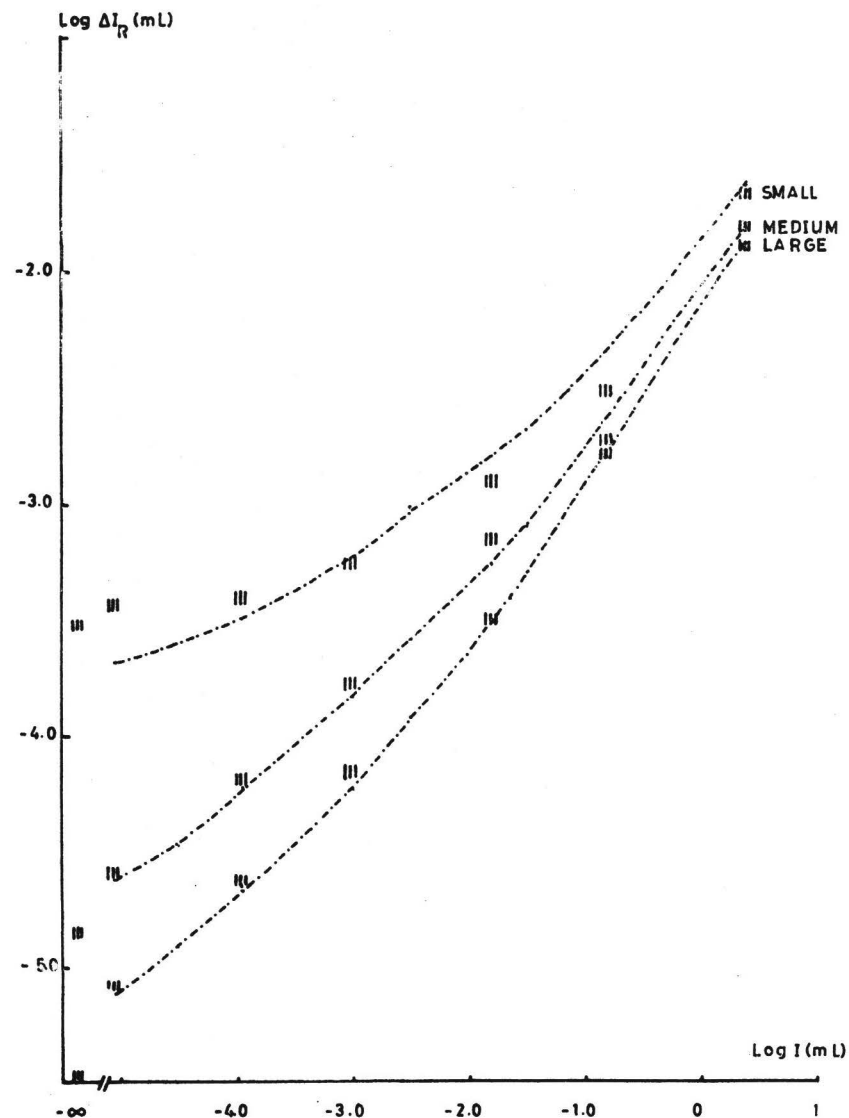


Fig.11

Experimental increment thresholds for the resolution of the parallel bar patterns fitted with theoretical curves derived from equation (3).
 $K' = -1.2$ log units. r_m , L_m , T_E and α_p assumed to be set by $I + \Delta I/2$.

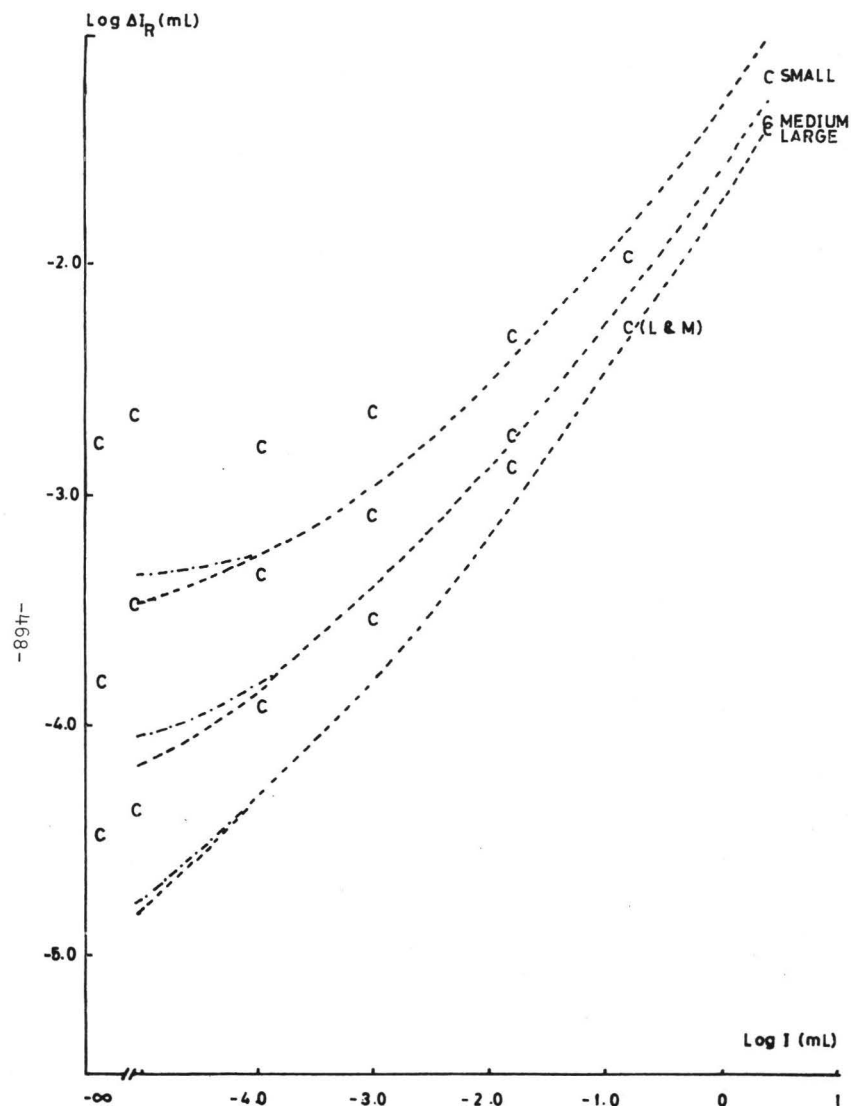


Fig.12

Experimental increment thresholds for the resolution of the Landolt C's fitted with theoretical curves derived from equation (4). Case (i) with detector X diagonally opposite Y. r_m , L_m , T_m and α_p assumed to be set by background, I , alone (----) or $I + \Delta I$ (- - - -). $K = -1.22 \log \text{ units}$

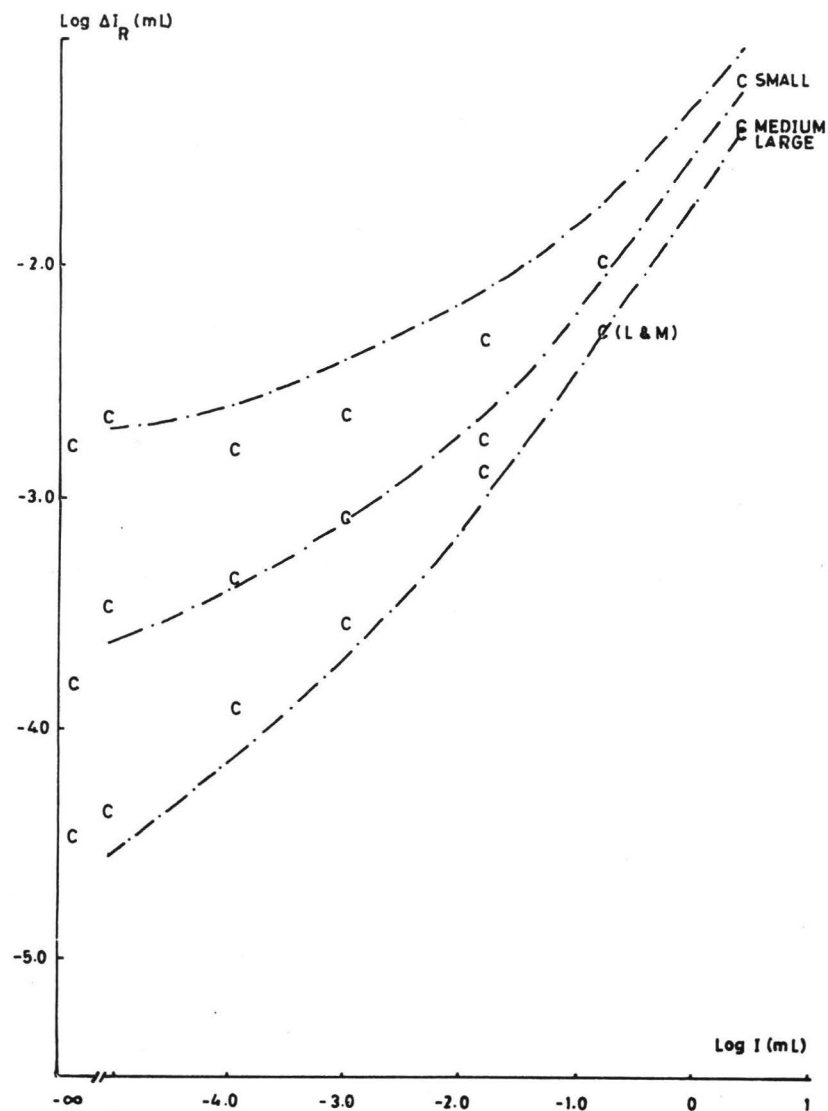


Fig.13

Experimental increment thresholds for the resolution of the Landolt C's fitted with theoretical curves derived from equation (4). Case (ii) with detector X a gap's width from Y. r_m , L_m , T_m and α_p assumed to be set by background, I , alone. $K = -1.22 \log \text{ units}$

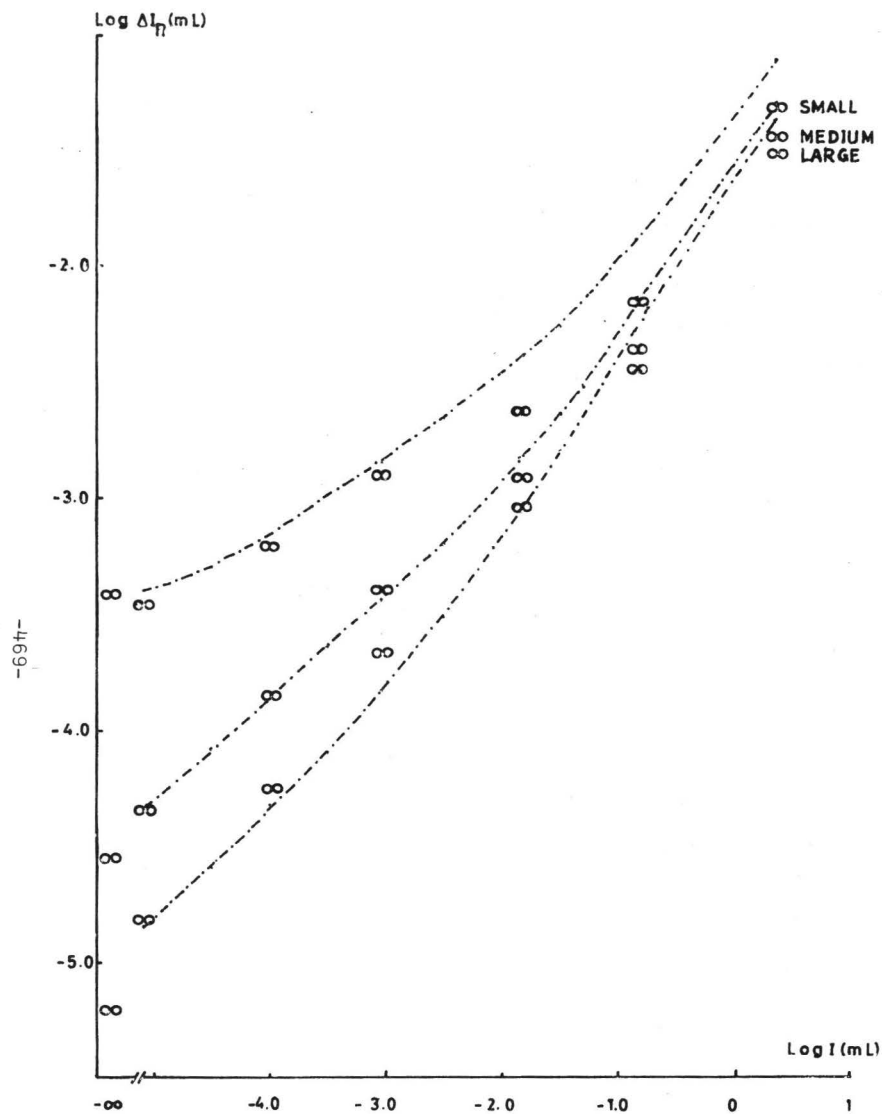


Fig.14

Experimental increment thresholds for the resolution of the discs patterns fitted with theoretical curves derived from Equation (4).
Case (ii), with detector Y 1.64 disc radii from the centres of both discs. r_m , L_m , T and α_p assumed to be set by $I + \Delta I$.
 $K = -1.16$ log units.

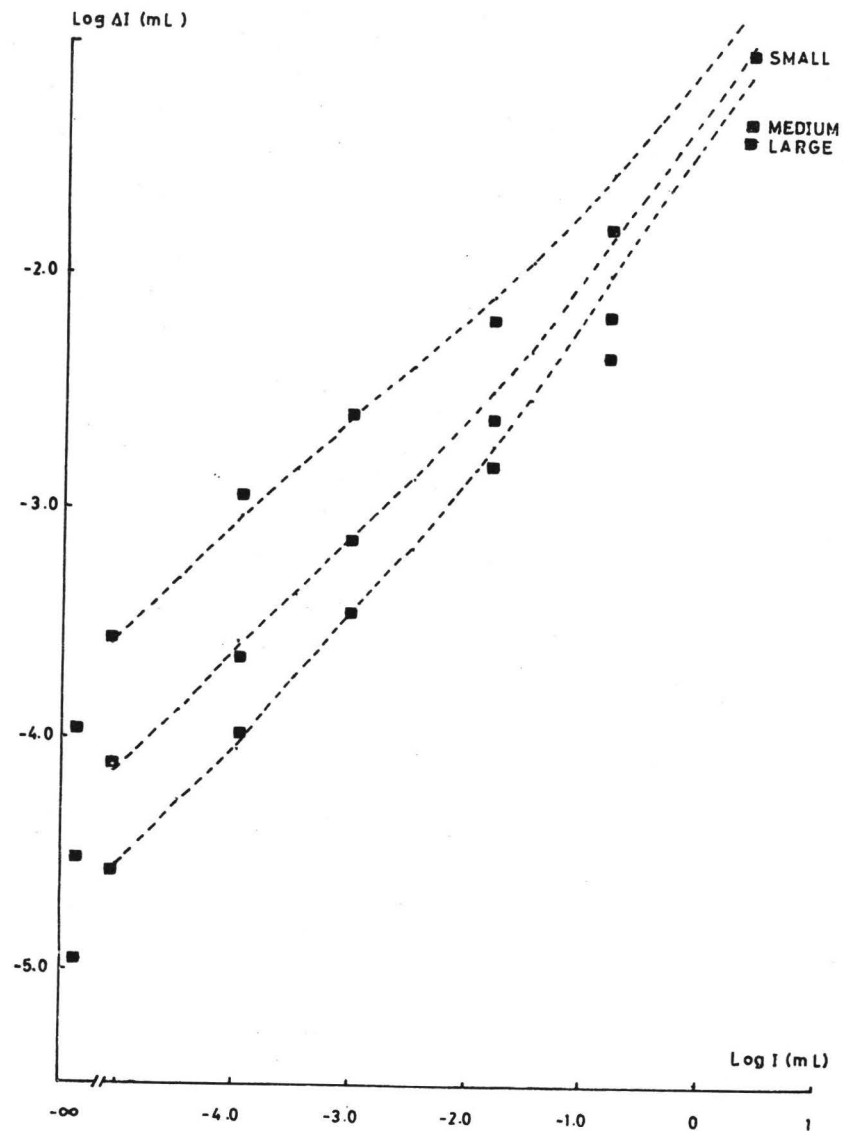


Fig.15

Experimental increment thresholds for the squares fitted with theoretical curves derived from the area detection equation (4).
 $K = -0.98$ log units.

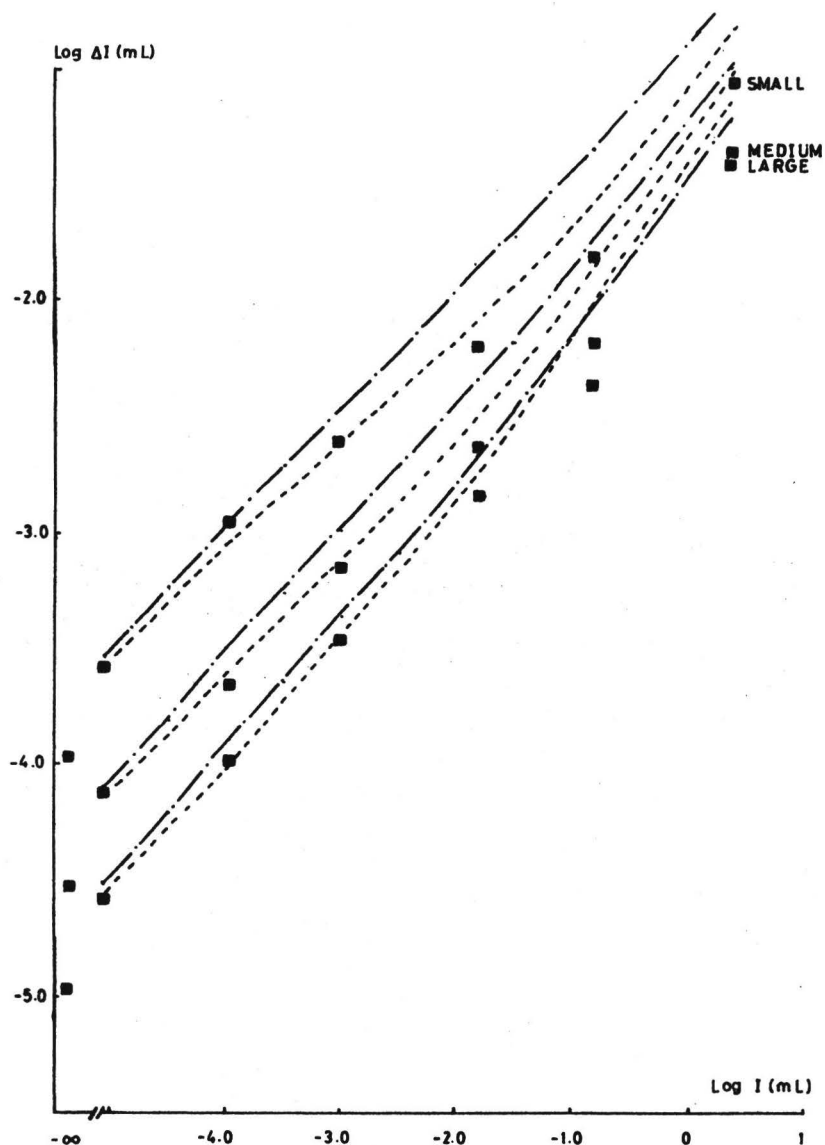


Fig.16 Experimental increment thresholds for the squares fitted with theoretical curves derived from the area detection equation (4) (---) and the line/edge detection equation (3) (-.-), $K \text{ \& } K' = -0.98 \text{ log units}$.