Vision Res. Vol. 32, No. 3, pp. 433–446, 1992
Printed in Great Britain.

0042-6989/92 $5.00 + 0.00

The Field Topography of ERG Components in Man—I. The Photopic Luminance Response

ERICH E. SUTTER,* DUONG TRAN*

Received 10 September 1990; in revised form 15 February 1991

A technique of multi-input systems analysis is used to explore the field topography of ERG responses to local luminance modulation. Variations in amplitude and wave form are studied within the central 23°. Outside the fovea, the amplitude appears to follow a simple power law r^x as a function of eccentricity r where x is approximately $-2/3$. The largest inter-subject variability is found in the fovea. Some nasal–temporal asymmetry is observed in all subjects with higher response densities in the temporal field outside the blind spot. The topography of the luminance response shares all these properties with the density of retinal cones.

Electroretinogram Human Field Topography Perimetry Electrophysiology Luminance Photopic Cones

INTRODUCTION

Human visual function is highly inhomogeneous across the visual field. It is well known that visual acuity declines rapidly with distance from the center of fixation (Wertheim, 1980; Westheimer, 1982), while scotopic sensitivity and photopic flicker sensitivity (Tyler, 1987) increase with eccentricity. Local differences are also observed in the perception and discrimination of colors (Abramov & Gordon, 1977; Stabell & Stabell, 1984). With monocular viewing, significant differences in some of these properties are also found between the nasal and temporal hemi-retinas.

The local variability of many of these properties has been related to variations in retinal anatomy and histology. Scotopic vision relies on the presence of rods which are missing at the central fovea and abundant in the periphery. Color discrimination depends in part on the locally varying ratios of the three cone population and macular pigmentation (Stabell & Stabell, 1980; de Monasterio, McCrane, Newlander & Schein, 1985). The dependence of the diameter of cone outer segments has been implicated in local variations of temporal response characteristics and sensitivity (Tyler, 1985). The difference between central and peripheral acuity are related to the densities of the foveal and peripheral cone mosaics (Oesterberg, 1935; Curcio, Sloan, Kalina & Hendrickson, 1990) and the ganglion cell to receptor ratios which change from about 4:1 at the center to 1:1 at about 15° (Wassle, Grunert Rohrenbeck & Boycott, 1989). The topography of these and other anatomical properties of the retina have been compared between different primates including man (Stone & Johnston, 1981).

Anatomical inhomogeneities of retinal layers and supporting tissues are thought to be responsible, in part, for local variations in the susceptibility to retinal diseases which commonly manifest themselves first in fairly localized areas of highest susceptibility. In many cases, the distribution of the defects is a characteristic of the disease. The functional topography of the retina is, thus, of great importance to the clinician.

The only techniques presently available for the study of the functional topography of the human retina and for objective, non-invasive assessment of local retinal function in the clinic involve the recording of ERG signals. While techniques have been used to obtain ERG responses from small retinal areas, it has so far not been feasible to obtain ERG responses from a sufficient number of retinal locations to permit derivation of response topography maps. The main reason for this lies in the poor signal-to-noise ratio of ERG responses from small retinal areas. The derivation of signals from regions spanning as little as one degree of visual angle requires recording times which prohibit sequential recording from a large number of locations in the same session. In addition, the results of such sequential recordings would be difficult to compare due to signal variations during a recording session and from one session to the next.

As an alternative approach to field topography mapping. Sandberg and Ariel (1977) developed a hand-held stimulator ophthalmoscope that permits accurate

*Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115, U.S.A.

placement of a single small ERG stimuus for the derivation of a steady-state response from a few selected locations. This method of stimulation, clearly, does not permit recording from a large number of locations in the same session for the purpose of field mapping. Adequate signal-to-noise ratios are achieved by limiting the amount of information derived from the stimulated area to the amplitude and phase at one frequency.

A common problem encountered with focal recording techniques is related to light scattering in the eye. Even a relatively small fraction of scattered light falling on dark adapted retina can elicit a sizable response. Spurious responses may, thus, be observed when the focal stimulus is placed on an unresponsive area. The optic disc, being highly reflective, exhibits this problem particularly well. Its apparent responsiveness indicates how well the problem has been solved.

The obstacles to successful ERG field mapping can be overcome with techniques of simultaneous recording whereby small retinal areas are independently stimulated. Each area contributes to the overall response recorded from the cornea. If the temporal modulation of different areas is properly selected, their individual response contributions can be extracted from the record. The subsequent computation of the local responses is particularly simple if the stimulation of different locations is uncorrelated (orthogonal) over the entire recording period. The best known such technique uses orthogonal sinusoids, i.e. sinusoidal signals that are harmonically related to the recording period. The major problems with this approach have been discussed by Reagan and Milner (1981). Note that comparison of the local responses is difficult when they are all stimulated at different frequencies. Another simultaneous recording technique developed in this laboratory has made it possible to derive the responses from hundreds of locations in about the same time it would take to obtain one single local response. This technique is outlined in the methods section and is discussed in detail elsewhere (Sutter, 1991, 1992).

This first paper focuses on local variations in the luminance response. There is an abundance of lateral mechanisms throughout the retina. Such interactions already exist between receptors and are also mediated by horizontal and amacrine cells. They raise some questions concerning the usefulness of the concept of the local luminance response. Clearly, the response to a luminance change of any small retinal region must be affected to some extent by the luminance of the surrounding area. The smaller the region, the larger the expected surround effects. The validity of the concept of a local luminance response is primarily based on the observation that the pattern specific response is much smaller than the luminance response, even if the stimulated area is relatively small. In the central visual field, the concept appears to be useful for areas as small as a fraction of a degree. However, at this scale, one needs to be aware of the effects of contrast attenuation through blur and light scattering.

The stimulation and analysis techniques used in this study are unconventional and are described in detail elsewhere (Sutter, 1992). The outline presented in the methods section should provide the necessary background for the interpretation of the results. Additional details supplied in the appendices and in the referenced materials should make it possible to replicate our findings.

METHODS

Linear systems are described by their impulse response which can be measured directly using pulsed inputs. However, much better signal-to-noise ratios can be achieved in the same recording time if white noise is used as a test input. The impulse resonse $I(t)$ is then simply derived from the response as the cross-correlation function between input $x(t)$ and response $r(t)$:

$$I(t') = \int_0^\tau x(t - t')r(t)\mathrm{d}t. \qquad (1)$$

For correlational shifts larger than the duration τ of the impulse response, the recorded signal no longer shows any correlation with the input except for background noise. If a second input is now stimulated with the same input sequence delayed relative to the first input by a time interval $T > \tau$, then its impulse response will appear on the cross-correlation function starting at the correlational shift T. Thus, the impulse responses of both inputs can be extracted from the same record by a single cross-correlation. This procedure is illustrated in Fig. 1 for multiple inputs.

If the noise contamination is overall additive, then the simultaneous stimulation of multiple inputs with this technique does not increase the noise. The impulse responses of all the inputs are obtained with the same signal-to-noise ratio as if each one were individually stimulated for the same period of time. Fortunately, the noise in the ERG signal is indeed overall additive and the signal contributions from different inputs have zero correlation and do not contaminate each other with apparent noise.

Stimulation of all inputs with the same white sequence is the basic principle used in this study to obtain the responses of a large number of inputs. The stimulation of all inputs is exactly equivalent. It can be shown that among all the white noise inputs with a given amplitude range, those with a binary amplitude distributions are the most efficient. The test inputs used here belong to a special class of pseudo-random sequences called binary m-sequences. They distinguish themselves from other white binary sequences by special properties that make them particularly useful for the analysis of nonlinear systems (Sutter, 1987, 1992). The most important feature is that the interactions between a sequence of display frames due to the presence of nonlinearities are also found at specific locations on the cross-correlation function between the m-sequence and the response. If the m-sequence is properly selected, these terms will fall in between the first order responses considered above and

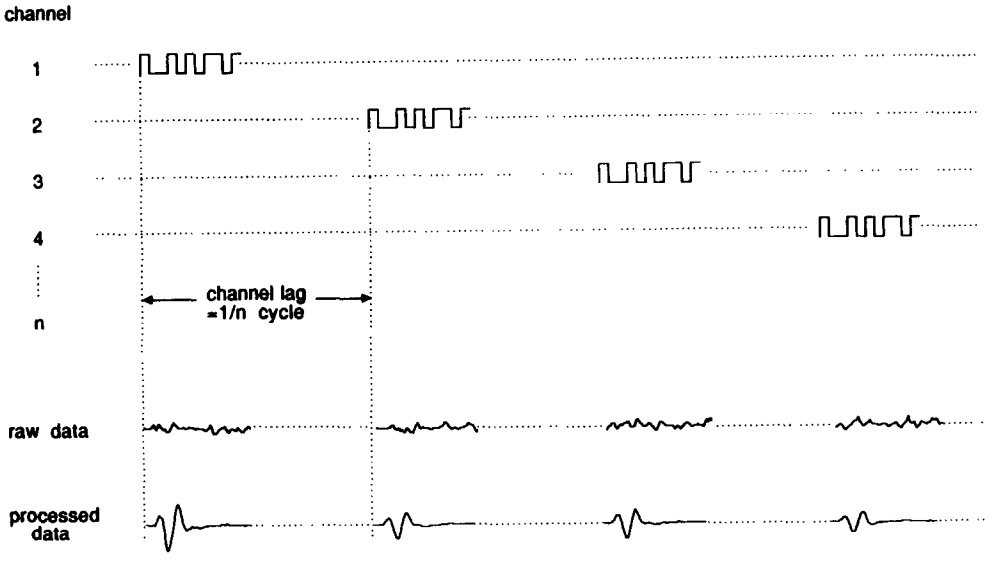

FIGURE 1. Schematic illustration of the multi-input systems analysis technique: each input channel is stimulated according to the same time series (binary m-sequence). A delay in the stimulation between channels (channel lag) renders the responses from different channels uncorrelated (orthogonal). They are extracted from the raw data by computation of cross-correlation between the m-sequence and the response cycle. The responses of the individual channels are found distributed along the cross-correlation cycle at intervals equal to the channel lag. For details see text.

both the first order and the dominant higher order terms are perfectly discriminable. In this study the dominant second order terms, i.e. the interaction between two consecutive frames and frames 30 msec apart (the two slices closest to the diagonal of the second order binary kernel) were also extracted and plotted. We found that these terms were not detectable in the noise under the experimental conditions used. They could only be seen after averaging over a large number of inputs. Since the first order term so clearly dominates, it was sufficient to restrict the considerations to the first order term. This does not mean, however, that the ERG response can generally be treated as linear. There is ample evidence for nonlinearities in the luminance ERG response (see e.g. Troelstra & Schweitzer, 1968; Sutter & Vaegan, 1990). When the experiments of the present study are performed at higher retinal illumination using dilated pupils, the dominant second order terms grow to appreciable size and their field topographies can be studied and compared. However, these aspects of the ERG are beyond the scope of this paper.

Binary m-sequences offer another important advantage. The computation of the cross-correlation (1) between the m-sequence input and the response can be reduced to a single Fast Walsh Transform (Sutter, 1991). This algorithm makes it possible to obtain the impulse responses for hundreds of inputs from records of up to a million data points in a fraction of a minute. No special hardware such as an array processor is required. This will be of great importance to future clinical applications of the technique.

The generation of binary m-sequences is so simple and fast that the multi-input stimuli can be generated on retrace between display frames (Sutter, 1991, 1992). Thus, no memory allocation is required for the stimulus sequence.

In order to achieve equivalence of the local stimulation, all the channel lags (see Fig. 1) were chosen equal, i.e. $1/k$th of a stimulation cycle, where k the power of 2 nearest to and larger than the number of stimulated inputs. The reason for the selecting a power of 2 is explained in Appendix A.

The stimulus matrix used in this study is a hexagonal array of fields scaled with eccentricity to provide approximately equal signal amplitude at all locations [Fig. 3(a)]. The scaling permits increased spatial resolution toward the center where the response density is higher. The hexagonal geometry was selected to achieve the closest possible packing of the localized fields as well as optimal isotropy of the array. Isotropy becomes particularly important when the array is scaled with eccentricity. An array of 241 hexagons was used to cover the central visual field. The field covered by the stimulus extended out to the 23° isometer, except in one high resolution experiment where the entire array was within the central 7°. The eccentricity scaling used in the two cases was somewhat different [compare Fig. 3(a) and Fig. 8(a)].

The stimulus hexagons were generated in black and white. The high luminance was 200 cd/m^2 and the low level was less than 3 cd/m^2, resulting in a mean screen luminance of approximately 100 cd/m^2. The brightness was modulated between the two levels in accordance with the m-sequence modulation technique outlined above.

The stimulus array was generated on a CRT color monitor (Hitachi HM3619A) by means of a customized Macintosh video card. The card had been modified and supplemented to permit updating of the brightness value or color of hundreds of fields on the screen during retrace of the raster scan.

Data collection

The length of the m-sequence used in all experiments was $2^{16} - 1$. The stimulation rate was equal to the

frame rate of the Macintosh video card, i.e. 67/sec. Thus, the net recording time for one stimulation cycle was 16 min. To avoid aliasing problems, the response signal was sampled synchronously with the video display at 8 samples per display frame, i.e. 536 samples/sec. The resulting sampling interval was, thus, 1.87 msec.

The record was obtained in 32 slightly overlapping segments of a tolerable length (32 sec). Data collection always started 1 sec after the beginning of stimulation to eliminate all effects due to stimulus onset.

A small but conspicuous red fixation spot marked the center of the display. The subject was instructed to maintain fixation during the 32 sec recording intervals. Clearly, shifts in fixation smaller than half the radius of the smallest hexagonal stimulus element have no significant effect on the local signals. The radius of the central hexagon was 40.5′ in the low resolution case and 14.3′ for the single high resolution experiment. All but one of the subjects were experienced observers and there was good reason to assume that they would be able to perform this simple task. Particularly subject GHP who participated in the high resolution recording, is highly experienced in psychophysical and electrophysiological experimentation. In addition, any break of fixation could be detected by monitoring of the ERG signal during recording, since even small saccadic eye movements generate considerable ERG artifacts. All record segments that showed such contamination were rejected and re-recorded. It is evident from the response topography plots themselves, particularly from the sharp central peak and the reproducibility of local features (see top row in Fig. 9), that fixation was not a problem in this study.

Burian–Allen electrodes were used for signal derivation. The pupil size was natural. The signal amplification was 200,000 and the filter setting was bandpass 10–100 Hz (Grass).

Computation of the local responses

Before processing, the recorded segments were spliced together to form one contiguous response cycle. To prevent discontinuities in the record, a linearly graded splice was used which covered the region of segment overlap (*ca* 1 sec).

The extraction of the local response contributions requires cross-correlation between the response cycle and the m-sequence as a cycle of 1 sec and -1 sec. It is accomplished by the ultra-fast m-transform algorithm (Sutter, 1991). The first order response components for all inputs are found distributed along the cross-correlation cycle at intervals equal to the channel lag of $1/k$th of the cycle length (Fig. 1, bottom). Due to the absence of significant nonlinearities, these components closely approximate the local impulse responses.

Artifact rejection

Artifacts from blinks and small eye movements can have a devastating effect in ERG recording and often require techniques of artifact rejection. This is a particularly important problem in field topography recordings where extremely small area local responses have to be extracted.

Most of the subjects were able to suppress blinks during the 30 sec recording segments. Occasionally, contaminated segments had to be rejected and re-recorded. However, it was not always possible to obtain adequate records without artifact rejection. The special multi-input systems analysis technique used in this study required the development of suitable means to deal with this problem. Conventional artifact rejection involving simple removal of contaminated record segments is not advisable in this case, since completeness of the record is necessary for accurate computation of the local responses. The removed segments would either have to be replaced by zeros or substituted from another recording. In the first case, the missing data results in cross-talk between the local responses, which has the appearance of noise contamination. In the second case, the required recording time is significantly increased. The artifact rejection technique used in this study involves replacing the segment containing the artifact by a response estimate computed from the entire record. A brief outline of the technique and an example of its effectiveness are presented in Appendix B.

Correction for scan delay

The presentation of the stimulus matrix on a raster scan CRT display results in slightly asynchronous stimulation of hexagons at different heights on the screen. At the frame rate of 67/sec the maximum asynchrony was approximately 13 msec. To accurately compensate for the scan delay of each stimulus hexagon, the screen was subdivided into 8 horizontal strips corresponding to the 8 data points collected during each frame. The hexagons in strip 5, for instance, is scanned five time bins after the hexagons in strip 0 at the very top of the screen. Its computed response requires a correction by 5 time bins. The response of each hexagon was, thus, advanced by a number of time bins equal to the number of the strip which contained the largest portion of its area. The timing accuracy of 1.86 msec, which was achieved with this correction, was needed for a comparison of response latencies and the estimation of response amplitudes by means of the method described below.

Note that the scanning of each hexagon also takes a finite amount of time. For the larger hexagons this time is close to 1 msec. However, since the high frequency cut-off used in this study was 100 Hz, the response to a 1 msec scan was indistinguishable from the response to ultra-brief flash stimulation of the same area.

Estimation of response amplitudes

In order to plot the topography of a specific response component, it is necessary to reduce each wave form to a single number which serves as a measure of its response amplitude. Peak-to-peak measurements are generally poor amplitude estimates since they are highly susceptible to noise contamination. A better amplitude estimate is found in the well known r.m.s. measure derived from the interval containing the major response

components. However, if the shape of the response is known, it can be used as a template for the derivation of a significantly more accurate amplitude estimate. A time interval (a,b) is selected which contains the relevant response components and the template t_i is normalized over this interval:

$$\mathbf{t} = \frac{\mathbf{s}}{\sqrt{\mathbf{s}\cdot\mathbf{s}}} \equiv \frac{\{s_i\}}{\sqrt{\frac{1}{(b-a)}\sum_{i=a}^{b}(s_i)^2}}. \quad (2)$$

The dot product between the normalized template t_i and the kth local response $(r_k)_i$ represents its amplitude estimate:

$$A_k = \mathbf{r}_k \cdot \mathbf{t} = \sum_{i=a}^{b} (r_k)_i t_i. \quad (3)$$

The relative high noise immunity of this amplitude estimate is related to its sensitivity to changes in wave shape. The same template can only be used within areas where little or no variations in wave forms are observed. Regions within which the same template could be used were established experimentally by averaging over different retinal areas. We found in most subjects that the changes depended mainly on eccentricity. Thus, in most cases, templates were derived as averages over concentric rings around fixation. Examples of averages used as templates are shown in Figs 4, 6 and 9. Because of considerable inter-subject variability (see Fig. 6) the templates were derived for each individual subject and only used within the set of traces from which they were derived.

A theoretical comparison of the amplitude measure based on templates with the well known r.m.s. measure is given in Appendix C. The performance of two measures was also evaluated experimentally by comparing the resulting response density plots. While the general features of the plots matched well, those derived by means of the r.m.s. measure were consistently more noisy.

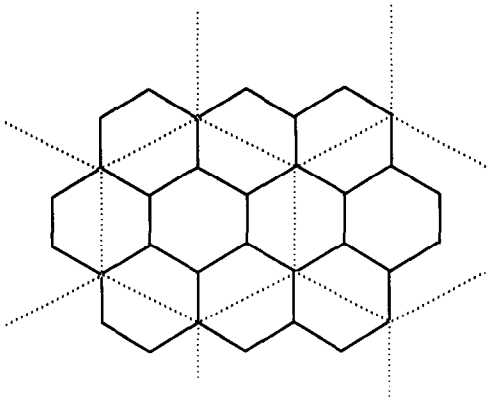

FIGURE 2. The interpolation technique used in the derivation of the three dimensional response density plots: the dotted lines delineate the stimulus hexagons. The interpolating hexagons (solid lines) centered on the stimulus hexagons are assigned the response density of the stimulus hexagon. The hexagons straddling a border between two stimulus hexagons are assigned the arithmetic mean of the two neighbors.

Derivation of amplitude densities

The local response amplitudes estimated by means of the above techniques are generated by stimulus elements of different sizes. They have no direct physiological meaning unless they are converted to response densities by normalization to unit retinal area. This is achieved by dividing each response amplitude by the area of the stimulus element that generated it.

The resulting local response density estimates can be represented as a three-dimensional surface over the stimulated area. The response topography plots as shown, for example in Fig. 8, are plotted with a finer hexagonal pattern than that used for stimulation.

The amplitudes for the finer pattern are estimated by a simple interpolation procedure as illustrated in Fig. 2. All the small hexagons which are centered on a stimulus hexagon (dotted lines) are assigned the amplitude density estimated from the corresponding response. The interpolating hexagons straddling a border between two stimulus hexagons are assigned the mean value of the two respective densities.

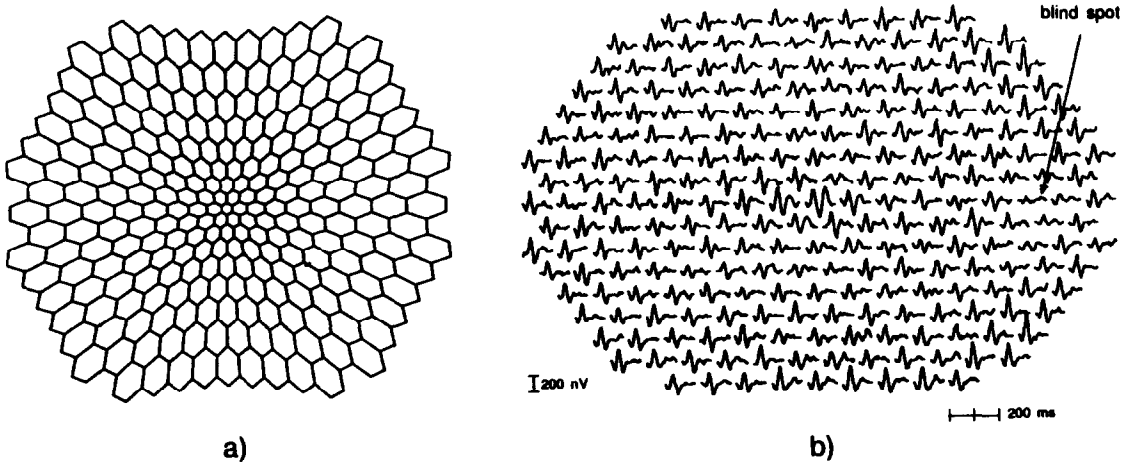

FIGURE 3. (a) The stimulus matrix of 241 hexagonal elements spanning the central 23° of the visual field. (b) The array of corresponding local luminance responses derived from a 16 min record of subject GHP.

RESULTS

The geometry of the stimulus matrix used in most experiments is shown in Fig. 3(a). It consists of 241 scaled hexagons and covers the central 23° of the visual field. The net recording time was 16 min or, when indicated, a multiple thereof. An example of an array of first order response traces corresponding to the stimulus matrix is shown in Fig. 3(b).

Note that since retinal illuminance was not controlled in this study, the absolute amplitudes as indicated by the vertical scale given in Fig. 3 (and also in Fig. 8) cannot be compared between subjects. All data presented here were recorded with natural pupils that could not be measured during recording. It is well known that, given the same stimulation, pupil size varies strongly between subjects and also in the same subject with age. Thus, retinal illuminance which is proportional to the square of pupil diameter was not controlled and might easily have varied by a factor of 3 or 4 between the subjects. This explains variations of the same magnitude between the local responses of the four subjects. However, these variations are of no major concern, since the focus of this study is on the response topography, i.e. on the relative response amplitudes at a large number of retinal locations.

It is also difficult to compare the sizes of these local responses with those obtained with conventional recording techniques. The impulse responses normally obtained through averaging over focal flash responses may appear larger because of the difference in the level of retinal adaptation. The mean retinal illuminance during stimulation is usually much lower. This also results in significantly larger rod contribution. In many regards, the stimulation used here resembles the so-called steady-state stimulation. However, the analysis does not just yield response amplitude and phase at one frequency, but results in the entire impulse response which is equivalent to the transfer function (assuming linearity). Again, a direct comparison of response amplitudes is not possible.

A well defined minimum in response amplitude is seen in Fig. 3(b) in the area of the blind spot. The residual response can be explained, at least in part, by the fact that the stimulus elements at around 10–15° eccentricity were relatively large, and no element fell completely within the receptor-free zone. Precision mapping of the blind spot onto the stimulus array was not possible, but simple hand plotting on two of the subjects suggested that most, if not all, of the residual response can be accounted for in this way. Even though light scattering within the eye may also contribute, it does not appear to be a major problem. This assertion finds additional support in the sharpness of the central peak and various other topographical features whose reproducibility is documented in the top row of Fig. 7.

Since not only ERG amplitudes, but also their wave forms are changing with location in the visual field, it was necessary to establish areas which required different templates for the estimation of response amplitudes. These areas were found by forming and comparing local averages. With most subjects, the major changes were with eccentricity. Figure 4 shows a series of template responses for subject GHP, together with the concentric rings from which they were derived.

The major differences in wave form are seen near the center of the 23° field. They consist predominantly in a latency increase with decreasing eccentricity which was observed in the later response components. This finding is in agreement with latency measurements reported by Brindley and Westheimer (1965) and Biersdorf (1982). Some subjects also showed slight

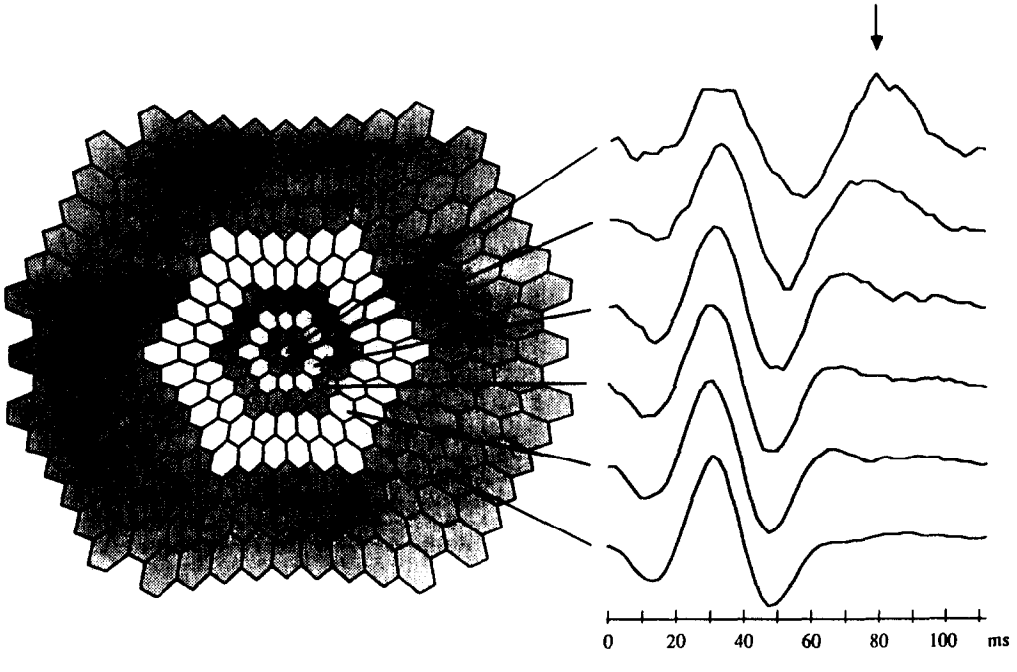

FIGURE 4. Averages of response wave forms derived from concentric rings of the data set shown in Fig. 3(b). Shape and peak latencies of the late components (marked with arrow) vary significantly with eccentricity.

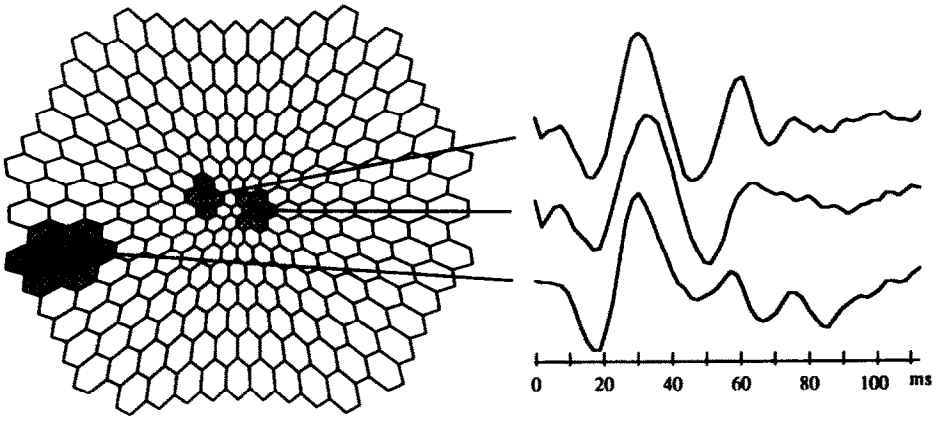

FIGURE 5. The largest local variability in response wave forms was seen in subject CSCH. Three examples of different wave form averages are shown here together with the patches from which they were derived.

nasal/temporal differences. The most pronounced local changes seen in one subject (CSCH) are shown in Fig. 5.

Inter-subject variations of local wave forms were also observed. Particularly small but significant differences in peak latencies were common. This is illustrated in Fig. 6, where signal averages derived from the same concentric rings are compared between subject CSCH and subject VG. These two subjects show the largest differences in response wave form observed in this study.

The latency differences seen in these two subjects may be attributed to differences in retinal illuminance due to uncontrolled pupil size. However, the oscillatory behavior in the later components of subject CSCH will require another explanation. It was not seen on other subjects with similar peak latencies.

Clearly, variabilities such as the ones seen in Fig. 6 do not permit the use of templates derived from one subject for amplitude estimation in another. Nor is it advisable to average together data from different subjects to derive low noise estimates for local template responses.

Because of the local variabilities and inter-subject differences, the templates used in the present study for the derivation of the response topography plots were derived for each subject by subdividing the stimulated field into areas with little variance in signal wave form, and averaging over these areas. Each template was then used only within the area from which it was derived. In most instances, the major local variations in response shape were with eccentricity. Thus, the templates were computed as normalized averages over concentric rings.

Three dimensional representations of retinal luminance response densities from the eyes of four different subjects are compared in Fig. 7. They span the range of inter-subject variability in response topography observed in recordings from eight healthy human eyes.

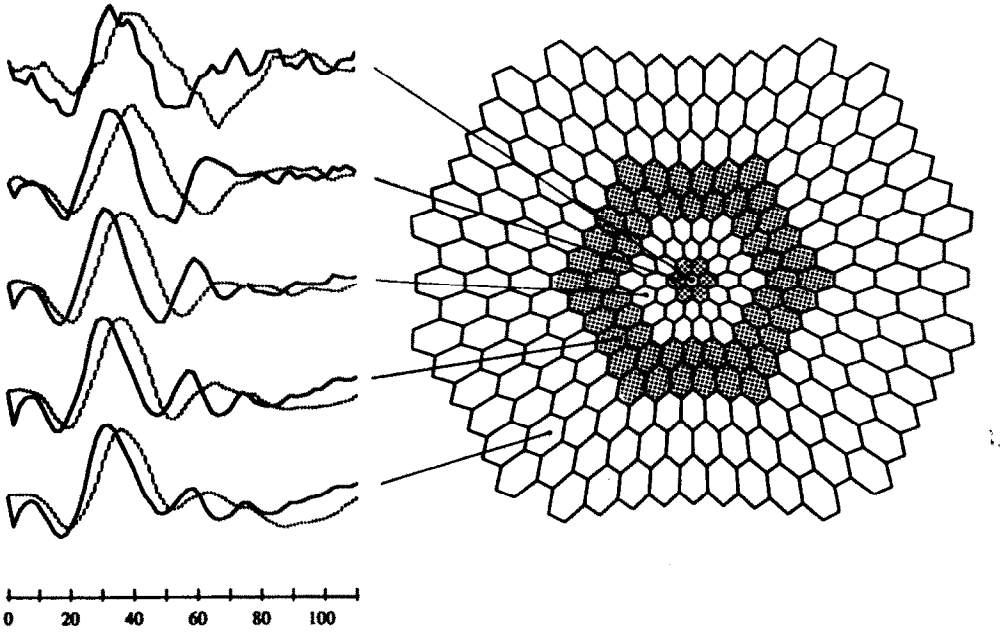

FIGURE 6. Illustration of the inter-subject variability in local response wave forms. Normalized ring averages are compared for subject CSCH (solid lines) and subject VG (dotted lines).

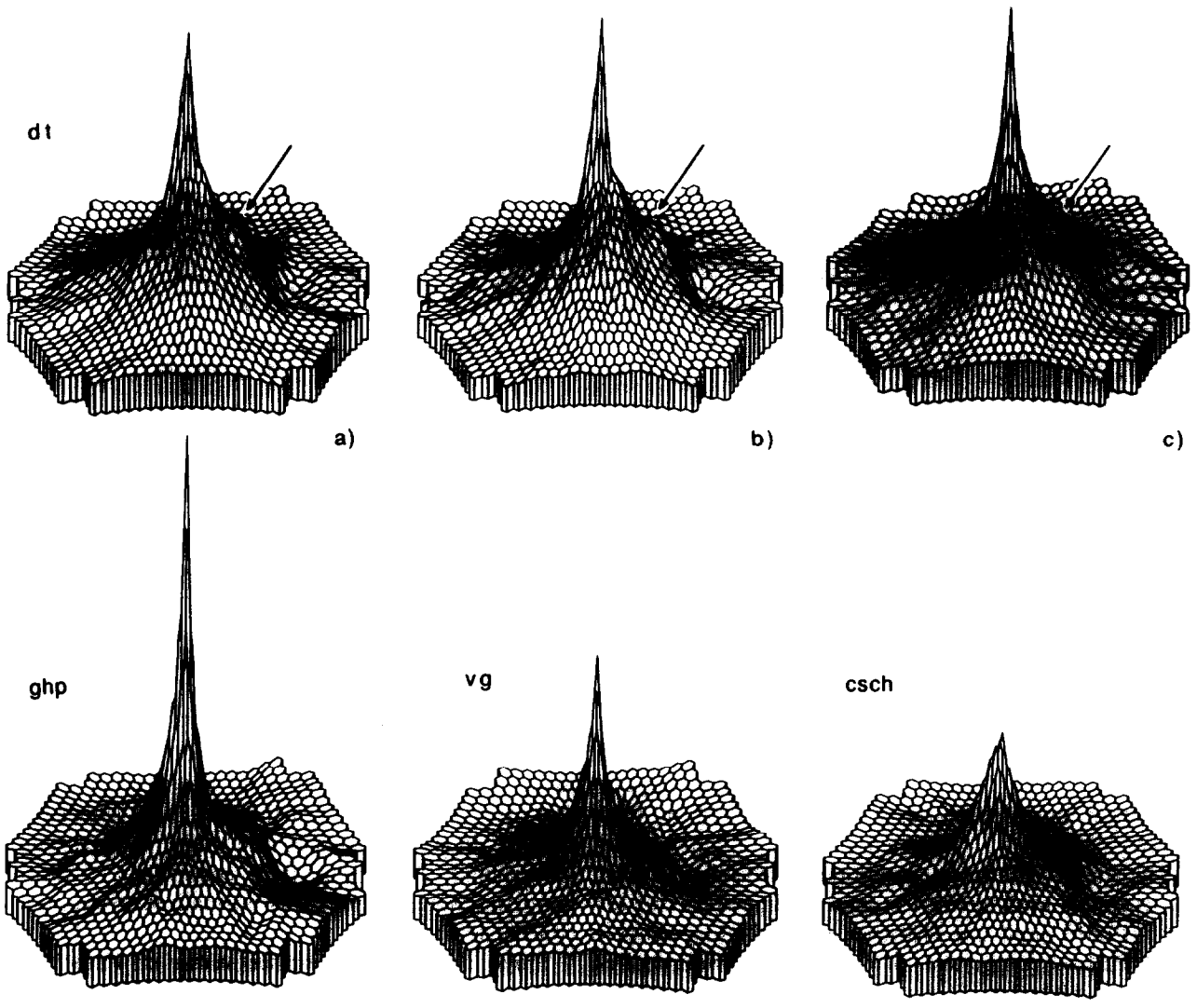

FIGURE 7. Luminance response topography of four different subjects (in all graphs the nasal and temporal fields are on the left and right, respectively, and the front represents the lower visual field). The top row illustrates the reproducibility of the plots on subject DT. Plot (a) is derived from the sum of three independent recordings obtained on different days. Plots (b) and (c) are obtained from two independent records. Note the common features, e.g. the slight bulge marked with arrow. The bottom row illustrates inter-subject differences. The plots were selected to span the full range of variability encountered in this study. When the plots are scaled to match outside the central 1.5°, the peak foveal response density varies by a factor of 2.7.

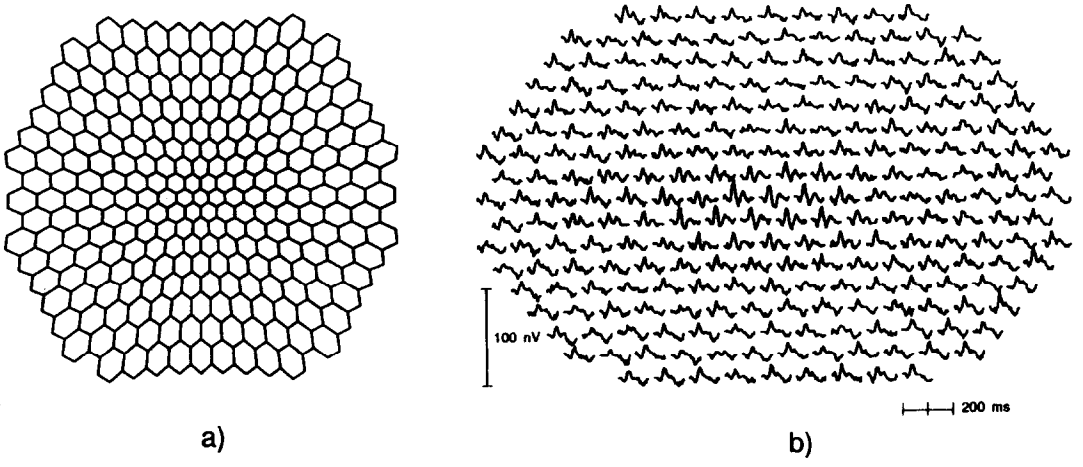

FIGURE 8. (a) The stimulus used in the high resolution experiment to test the central 7° of the visual field. The scaling of the hexagons with eccentricity is shallower than that used in the lower resolution experiments [Fig. 4(a)]. (b) The corresponding array of corresponding local luminance responses derived from a 80 min record of subject GHP.

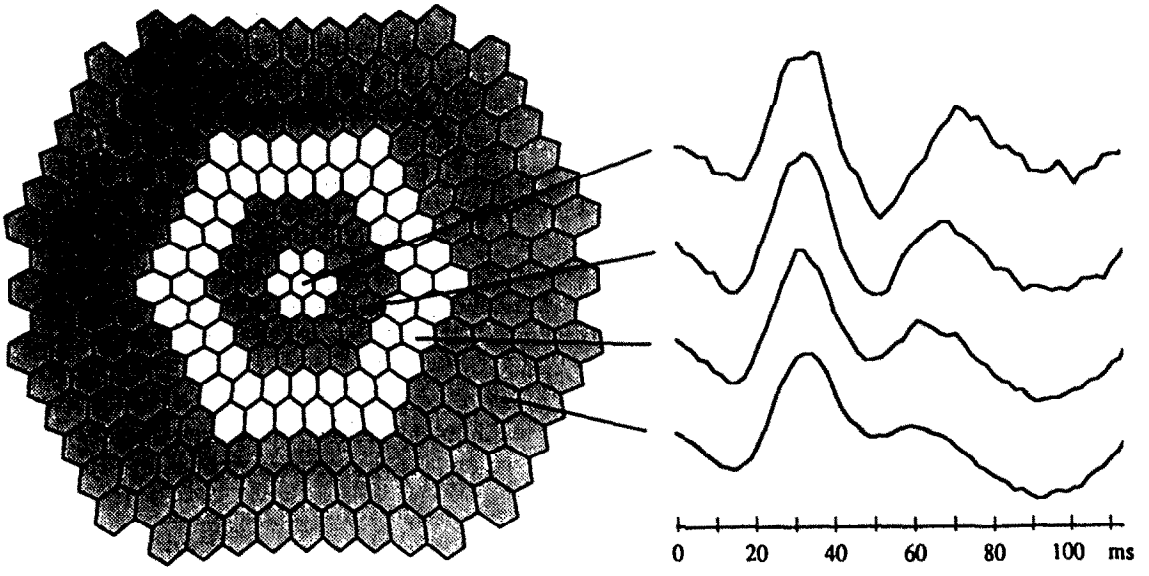

FIGURE 9. Averages of response wave forms derived from concentric rings of the data set shown in Fig. 8(b). Significant changes with eccentricity are only seen in the late components.

The amplitude measure was derived from the first 100 msec of the responses. With the filtering used, this interval covered all significant response components. The top row in Fig. 7 illustrates the reproducibility of plots derived from 16 minute records. Two independent plots obtained on different days from the same eye are shown in (b) and (c). The density plot (a) is derived from the averaged data of three such independent experiments. The major topographical features of this plot are already clearly visible from both of the independent recordings (b) and (c), particularly the comparatively high flanks in the extra-foveal region and a prominent bulge in the temporal field (indicated in Fig. 7 with arrow). Both features are not seen in any of the subjects of the bottom row.

The three examples of the bottom row illustrate the range of inter-subject variability observed in the response topography of the eight different eyes. Since the records were obtained with natural pupils, and pupil diameter could not be measured during recording, retinal illuminance was not calibrated. Thus, absolute response densities could not be compared. However, an approximate match in the topographies of most subjects could be achieved everywhere, except in the fovea, by multiplication of the plots with appropriate overall scaling factors. From the comparison of the matched plots, we conclude that the greatest inter-subject variability is in the fovea. A quantitative comparison between subjects of response densities as a function of eccentricity is given below.

The exceptionally high central response densities observed in the subject GHP prompted an attempt to achieve higher spatial resolution. It was necessary, however, to increase the net recording time to 80 min. Five complete 16 min cycles were recorded with the same m-sequence used for the lower resolution experiments and subsequently averaged together. In this experiment, the stimulus matrix shown in Fig. 8(a) covered the central 7°. The eccentricity scaling was somewhat

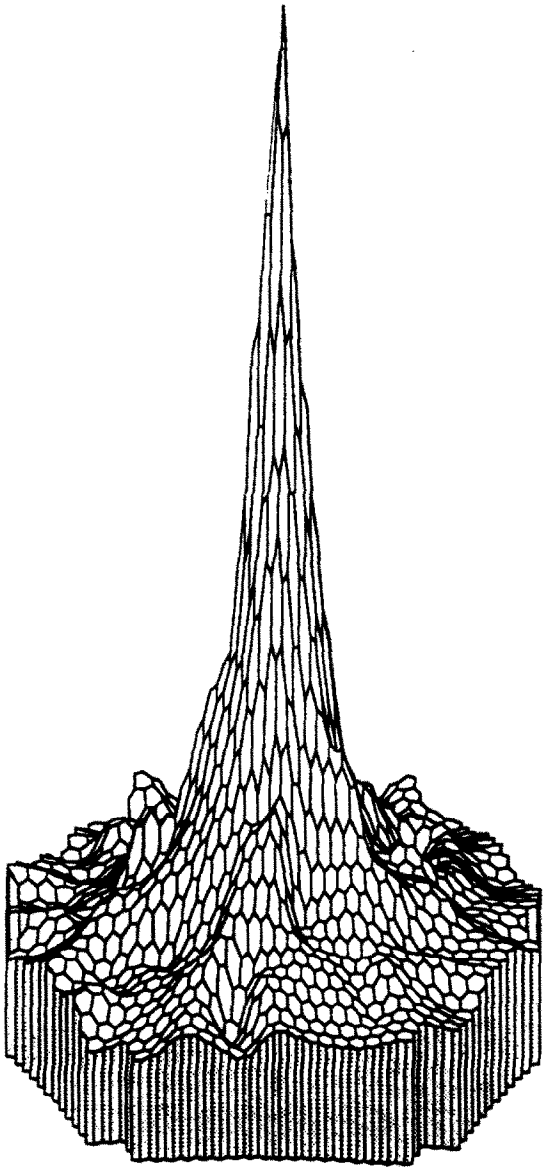

FIGURE 10. Response topography of the central 7° derived from the data set of Fig. 8(b).

shallower than that used in the lower resolution experiments [compare Fig. 5(a)]. The individual response traces are shown in Fig. 8(b).

As with the lower resolution data recorded from this subject, the greatest variability in response wave form is with eccentricity. Thus, the response templates for amplitude estimation were again derived as averages over concentric rings (Fig. 9). Significant changes are seen in shape and latencies of the late components, particularly within the central 2°, while the early phases of the *a*- and *b*-wave remain almost invariant.

In Fig. 10, the response density is plotted for the central 7°. The local density estimates were derived from the high resolution data of Fig. 8 by means of the templates of Fig. 9. A qualitative comparison of the high and low resolution plots from subject GAP shows good agreement in the essential topographical features. However, a quantitative comparison of the flanks of the distributions in the overlapping regions outside the fovea reveals a small discrepancy, which was attributed to contrast attenuation through the Burian–Allen electrode and the optics of the eye. A quantitative estimation of the effects of contrast attenuation in each experiment would be extremely difficult and requires instrumentation that was not available. However, based on the analysis by Sutter and Vaegan (1990), an appreciable effect of contrast attenuation has to be expected when the stimulus elements are only a fraction of a degree in diameter. These considerations also justify a minor

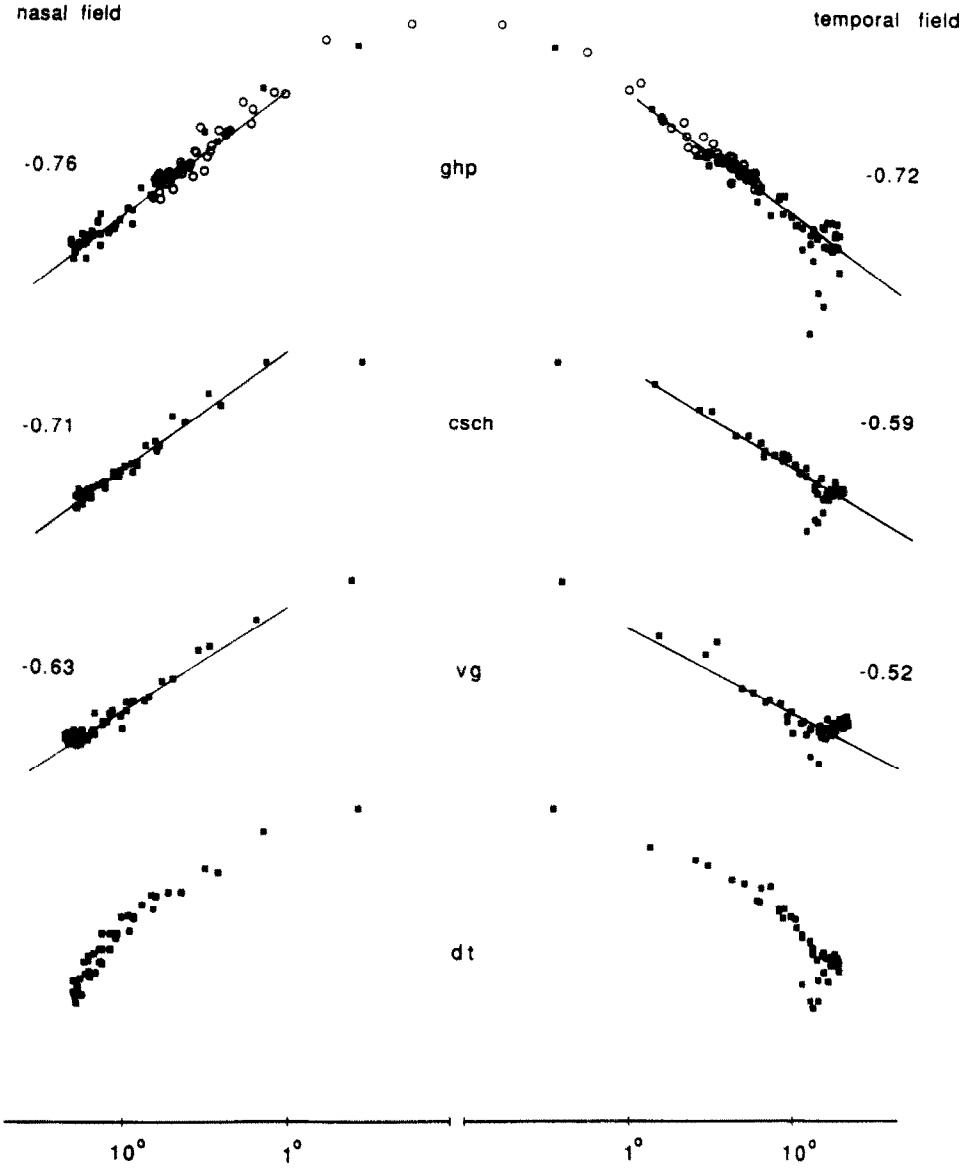

FIGURE 11. Response density as a function of eccentricity in double logarithmic coordinates. The data points on the left are derived from the nasal, and those on the right from the temporal field quadrant. Points of equal eccentricity are averaged together with the exception of those affected by the blind spot. The four different subjects of Fig. 8 are compared. All but one could be well fitted with a straight line with a slope of approx. −2/3. Points within the central 1° were excluded from the linear regression. The slopes shown next to each set were slightly steeper in the nasal field quadrant. Note that the plot for subject GHP includes data points from the high resolution recording (open symbols). A correction factor was applied to the high resolution data to compensate for contrast attenuation. The factor was established empirically by matching the regression lines for the two plots (see text).

correction used to achieve a match in the overlapping region of the two data sets was also used in Fig. 10.

The quantitative analysis of the response density as a function of eccentricity was done separately for the nasal and temporal field. Since only relative response amplitudes can be compared between subjects (see above), inter-subject differences are best tested by plotting the logarithm of the response density d as a function of eccentricity r. In regions where the densities differ only by a constant factor, the curves can be brought into coincidence through a vertical shift. In Fig. 11, logarithmic scale is used for the eccentricity r as well. For easy comparison, the data points from four different subjects are assembled in the same plot with a relative vertical displacement. Each curve is obtained from the amplitude density estimates of the nasal and temporal quadrants. The values for stimulus hexagons with the same eccentricity are averaged together with the exception of those for the stimulus hexagons invading the blind spot. Their response densities were plotted as separate points.

For the three subjects at the top of Fig. 11, the flanks outside the 1° isometer appear to be nearly parallel straight lines. To derive their slopes, they were fitted by means of linear regression. All points representing averages were weighted accordingly in the regression, and the points in the vicinity of the blind spot and inside the 1° isometer were excluded from the estimate. The slopes range from -0.63 to -0.76 in the nasal field and from -0.52 to -0.72 in the temporal field. In all subjects, the slope is steeper in the nasal field. This straight line fit in double logarithmic coordinates suggests a power law $d(r) = r^x$ where the power x is approx. $-2/3$.

The curve for subject DT presents a notable exception. Its flanks are rounded and, thus, not compatible with a simple power law. The reproducibility of this exceptional result is well documented in the top row of Fig. 7. There is no obvious explanation for this difference in the global response topography. However, it is worth noting that DT is the only subject of Asian origin and that racial differences in the cone distribution cannot be excluded.

In all the other subjects, the distribution of the sources of the local luminance response appears to be similar, except within the central 1°. The range of variabiity in the foveal response density among all eight eyes spans a factor of 2.7.

Note that for subject GHP, the low resolution data (solid symbols) and high resolution data (open symbols) were combined in the same plot. For this purpose, the two data sets had to be matched up in the region of overlap. It was found that the flanks of both sets in their double logarithmic representations were well fitted with straight lines. However, the slopes of the regression lines were slightly shallower in the case of the high resolution data. This can be explained as an effect of optical contrast attenuation. The effect is approximately proportional to the percentage of each stimulus element which is affected by blur. To compensate for it, each local response of the high resolution data set was multiplied by a factor proportional to the ratio circumference/area of the corresponding stimulus hexagon. The correction factor is, thus, k/C, where C is the circumference and k is a constant selected for best fit of the regression lines through the nasal and temporal flanks.

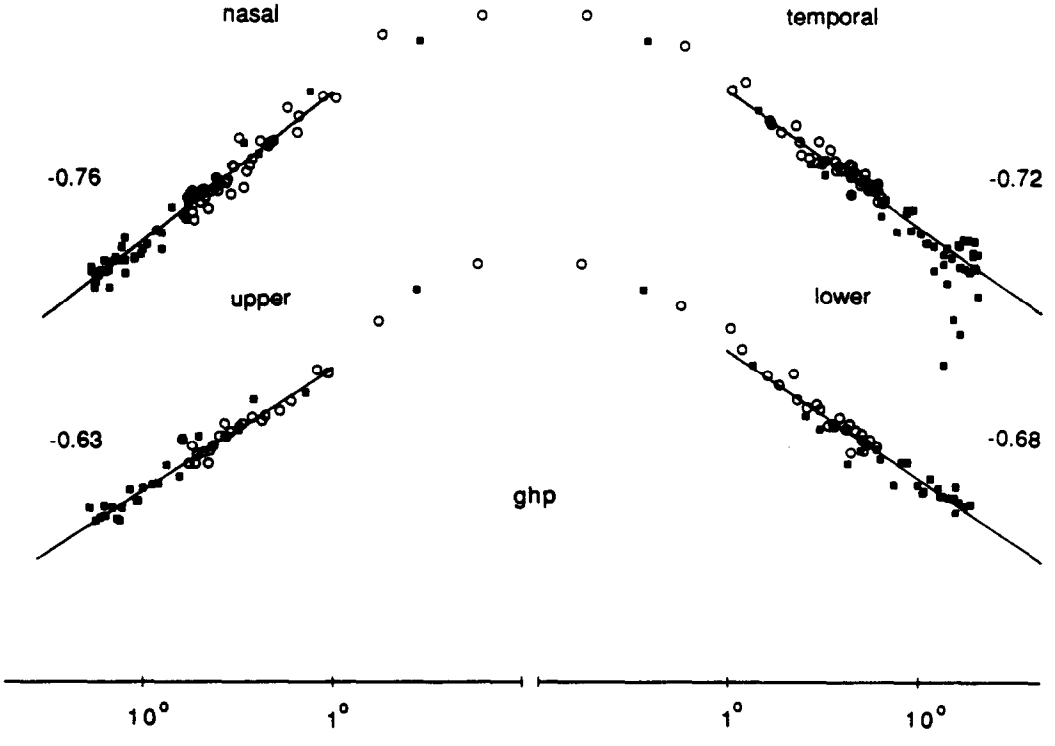

FIGURE 12. Comparison of response densities in the horizontal (top) and vertical (bottom) direction. The data points were derived from the nasal, temporal, upper and lower quadrants of the combined high and low resolution records from subject GHP. The response amplitudes from hexagons inside the central 1° were excluded in the derivation of the regression lines (see also Fig. 11).

It is clear from Fig. 11, that the luminance response topography has some degree of nasal/temporal asymmetry. For the first three subjects, the regression lines have steeper slopes in the nasal field. A notably faster nasal drop-off toward the periphery is also seen in subject DT. This asymmetry is mainly due to slightly higher response densities in the nasal field near the center.

The response density changes in the horizontal and vertical directions are compared in Fig. 12. The data points were derived from the combined results of the high and low resolution recordings from subject GHP. The points of the lower set were derived from the upper and lower quadrants of the hexagonal array and fitted with a straight line as described above. Again, the data points within the central 1° were excluded from the linear regression. This comparison shows a greater degree of symmetry in the vertical direction. The central peak of the distribution is significantly wider in the horizonal direction by extending farther into the nasal field. This central asymmetry and slight elongation in the horizontal direction is seen in the other subjects as well. It is reminiscent of a visual streak.

DISCUSSION

The above electrophysiological results have interesting parallels in anatomical properties of the outer retina. Cone densities have been studied in monkeys (Perry & Cowey, 1985), also in man by Oesterberg in 1930, and more recently by Curcio, Sloan, Packer, Hendrickson and Kalina (1990). These studies involved elaborate counting techniques in flat retinal wholemounts. The most important findings were:

1. Outside the central 1°, cone densities are approximately proportional to $r^{-2/3}$, where r is the eccentricity. In a plot of log (Cone Density) vs log (Eccentricity), the flanks outside the central 1° are well approximated on the nasal and temporal retina by straight lines with a slope of $-2/3$ (see Curcio, Sloan, Kalina & Hendrikson, 1987, Fig. 3).
2. The greatest inter-subject variability was observed within the central 1°.
3. At eccentricities larger than 20°, the cone densities are markedly higher on the nasal retina.

In all these essential features, there is surprisingly good agreement with response densities of the luminance ERG. This suggests that the strength of the relevant signal sources is predominantly determined by the receptor densities. Cone size and other histological properties of receptors which change with retinal location, appear to be less of a factor.

This study demonstrates that high resolution mapping of the field topography of the ERG to local stimulation is feasible. In 16 min of recording time, and less than one minute of data processing, it is possible to obtain response density plots with a central resolution of one degree or better. The technique promises clinical applcations. As a field mapping technique, it has much in common with psychophysical perimetry. There are, however, also important differences between the two methods. Static perimetry samples the field at a limited number of *points*, while the ERG topography tests many *small areas*. The duration of a perimetry test depends on the number of test locations, while the length of the ERG recording is generally only determined by the size of the stimulus elements. The number of areas tested in the present study (241) is significantly larger than the number of points that can be sampled in a perimetry test of the same length. In addition, ERG fields can be obtained from both eyes simultaneously. When better signal-to-noise ratios are made possible with brighter displays or pupil dilation, recording times can be significantly shortened. In this study, the high luminance level of the binary stimulus was limited to 200 cd/m^2 by the available CRT monitor.

Another important difference between the two field mapping techniques is found in the quantity derived from each tested location. In the case of psychophysical tests, a detection threshold is derived, while the result of the electrophysiological test is the local response characteristic described by a wave form or a series of wave forms (kernel slices). It is likely that much more information can be extracted from these wave forms, if the signal-to-noise ratio permits. Since the signal-to-noise ratio is proportional to the size of the stimulus elements, there is a trade-off between the spatial resolution and the amount of local information that can be derived from a record of a given length. By choosing large enough stimulus elements or by averaging over the responses of neighboring elements, it is possible to extract reliable estimates not only of response amplitude, but also of peak latencies and other features of the response traces. Ultimately, it should become possible to dissect the local response characteristics into components from specific retinal sources or layers. The use of different modes of local stimulation which selectively elicit responses from specific retinal sources can be helpful toward this end. The traditional stimulation with contrast reversing patterns will be explored in the next paper.

The response density measure used in this paper to derive topography plots is, of course, not the measure of choice for clinical applications. For the clinic, it will be necessary to establish criteria for the discrimination of abnormal local retinal responses from normal inter-subject variabilities. It has been clearly demonstrated in this study that the response variability in normal subjects depends on location. It is, thus, clear that the criteria for the detection of functional abnormality will have to be established locally.

REFERENCES

Abramov, I. & Gordon, J. (1977). Color vision in the peripheral retina. I. Spectral sensitivity. *Journal of the Optical Society of America, 67*, 195–202.

Biersdorf, W. R. (1981/1982). Temporal factors in the foveal ERG. *Current Eye Research, 1*, 717–722.

Brindley, G. S. & Westheimer, G. (1965). The spatial properties of the human electroretinogram. *Journal of Physiology, 179*, 518–536.

Connolly, M. & Van Essen, D. C. (1984). The representation of the visual field in parvocellular and magnocellular laminae of the lateral geniculate nucleus in the macaque monkey. *Journal of Comparative Neurology, 226*, 544–564.

Curcio, C. A., Sloan, K. R., Kalina, R. E. & Hendrickson, A. E. (1990). Human photoreceptor topography. *Journal of Comparative Neurology, 292*, 497–523.

Curcio, C. A., Sloan Jr., K. R., Packer, O., Hendrickson, A. E. & Kalina, R. E. (1987). Distribution of cones in human and monkey retina: Individual variability and radial asymmetry. *Science, New York, 236*, 579–582.

Golomb, S. W. (1982). *Shift register sequences*. Laguna Hills: Aegean Press, 44–45.

Hirose, T., Miyake, Y. & Hara, A. (1977). Simultaneous recording of electroretinogram and visual evoked responses. *Archives of Ophthalmology, 95*, 1205–1208.

de Monasterio, F. M., McCrane, E. P., Newlander, J. K. & Schein, S. J. (1985). Density profile of blue-sensitive cones along the horizontal meridian of macaque retina. *Investigative Ophthalmology and Visual Science, 26*, 289–302.

Oesterberg, G. A. (1935). Topography of the layer of rods and cones in the human retina. *Acta Ophthalmologica (Suppl.), 13(6)*, 1–102.

Perry, V. H. & Cowey, A. (1985). The ganglion cell and cone distributions in the monkey's retina: Implications for central magnification factors. *Vision Research, 25*, 1795–1810.

Reagan, D. & Milner, B. A. (1977). Objective perimetry by evoked potential recording: Limitations. *Electroencephalography and Clinical Neurophysiology, 44*, 393–397.

Sandberg, M. A. & Ariel, M. (1977). A hand-held two channel stimulator ophthalmoscope. *Archives of Ophthalmology, 95*, 1881–1882.

Schein, S. J. & de Monasterio, F. M. (1987). Mapping of retinal and geniculate neurons onto striate cortex of macaque. *Journal of Neuroscience, 7*, 996–1009.

Schiller, P. H. & Malpeli, J. G. (1977). Properties and tectal projections of monkey retinal ganglion cells. *Journal of Neurophysiology, 40*, 428–445.

Stabell, U. & Stabell, B. (1984). Color vision mechanisms in the extrafoveal retina. *Vision Research, 24*, 1969–1975.

Stabell, U. & Stabell, B. (1980). Variation of macular pigmentation and short wave cone sensitivity and eccentricity. *Journal of the Optical Society of America, 70*, 706–711.

Stone, J. & Johnston, E. (1981). The topography of primate retina: A study of the human, bushbaby, and new and old-world monkeys. *Journal of Comparative Neurology, 196*, 205–223.

Sutter, E. E. (1987). A practical nonstochastic approach to nonlinear time-domain analysis. *Proceedings, biomedical simulations resource, advanced methods of physiological system modeling* (Vol. I, pp. 303–315). Los Angeles: University of Southern California.

Sutter, E. E. (1991). The fast m-transform: A fast computation of cross-correlations with binary m-sequences. *SIAM Journal on Computing, 20*, 686–694.

Sutter, E. E. (1992). Pinter, R. B. & Nabet, B. (Eds), *Deterministic time-domain nonlinear analysis*. Cleveland, Ohio: CRC Press.

Sutter, E. E. & Vaegan (1990). Lateral interaction component and local luminance nonlinearities in the human pattern ERG. *Vision Research, 30*, 659–671.

Troelstra, A. & Schweitzer, N. (1968) Nonlinear analysis of the electroretinographic b wave in man. *Journal of Neurophysiology, 31*, 588–606.

Tyler, C. W. (1985). Analysis of human receptor density. *Investigative Ophthalmology and Visual Science, (Suppl.), 26*, 10.

Tyler, C. W. (1985). Analysis of visual modulation sensitivity. II. Peripheral retina and the role of photoreceptor dimensions. *Journal of the Optical Society of America, A2*, 393–398.

Tyler, C. W. (1987). Analysis of visual modulation sensitivity. III. Meridional variations in peripheral flicker sensitivity. *Journal of the Optical Society of America, A4*, 1612–1619.

Wassle, H., Grunert, U., Rohrenbeck, J. & Boycott, B. B. (1989). Cortical magnification factor and the ganglion cell density of the primate retina. *Nature, London, 341*, 643–646.

Wertheim, T. (1980). Peripheral visual acuity. Translated by I. Dunsky. *American Journal of Optometry, 57*, 915–924.

Westheimer, G. (1982). The spatial grain of the perifoveal visual field. *Vision Research, 22*, 157–162.

Acknowledgements—The authors wish to thank Dr Gunilla Haegerstrom-Portnoy for many helpful discussions and for her extensive participation in this study as an experiment subject. This project was supported by NIH grant EY06861 and the Smith-Kettlewell Eye Research Institute.

APPENDIX A

The task of stimulating many inputs by means of the same pseudo-random sequence with a delay (channel lag) between the different input channels is greatly facilitated by a special property of the binary m-sequences. This property is illustrated below using the example of a short m-sequence of length $2^3 - 1$.

The sequence is:

$$0\ 0\ 1\ 0\ 1\ 1\ 1\ldots \tag{A1}$$

Decimation of this sequence over four columns yields:

$$\begin{array}{cccc} 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ \cdot & \cdot & \cdot & \cdot \end{array} \tag{A2}$$

All the columns now contain the original m-sequence with a relative lag of 2 between columns from the last. For any binary m-sequence of length $2^k - 1$, it can be shown that if it is decimated over n columns and if the number of columns is a power of 2, then each column will contain the m-sequence with a relative lag between columns of $2^k/n$. Of course, the lag between the first and the last column has to be $2^k/n - 1$.

This property of m-sequences is extremely useful for fast real-time generation of m-sequence stimuli for multiple input systems.

APPENDIX B

The artifact rejection technique used in this study makes use of special features of the binary m-sequence technique. The cross-correlation between the m-sequence stimulus and the response signal which is used to extract the local impulse responses, represents a reversible transformation between raw and processed data. While in the raw signal the response is uniformly distributed, the artifacts are relatively sparse and localized in time. After processing, the reverse is the case: the response power is now concentrated in the sparse local impulse responses (compare Fig. 1), while the power of the artifacts has become distributed. These distinct properties of the two signal components and the ability to quickly move back and forth between the two domains can be used for their discrimination.

It is clear, that after data processing the signal power between the impulse responses is almost exclusively due to noise and artifact contamination. If these stretches are zeroed out and the back-transform is applied, one obtains the raw data set with most of the artifacts removed. Where the back-transform deviates significantly from the original signal, one simply replaces the contaminated stretch of raw data by the back transform. Everywhere else the original data are used. A subsequent forward transform of the corrected data leads to a significant improvement of the response estimates for a contaminated record (Fig. A1). This procedure can be iterated.

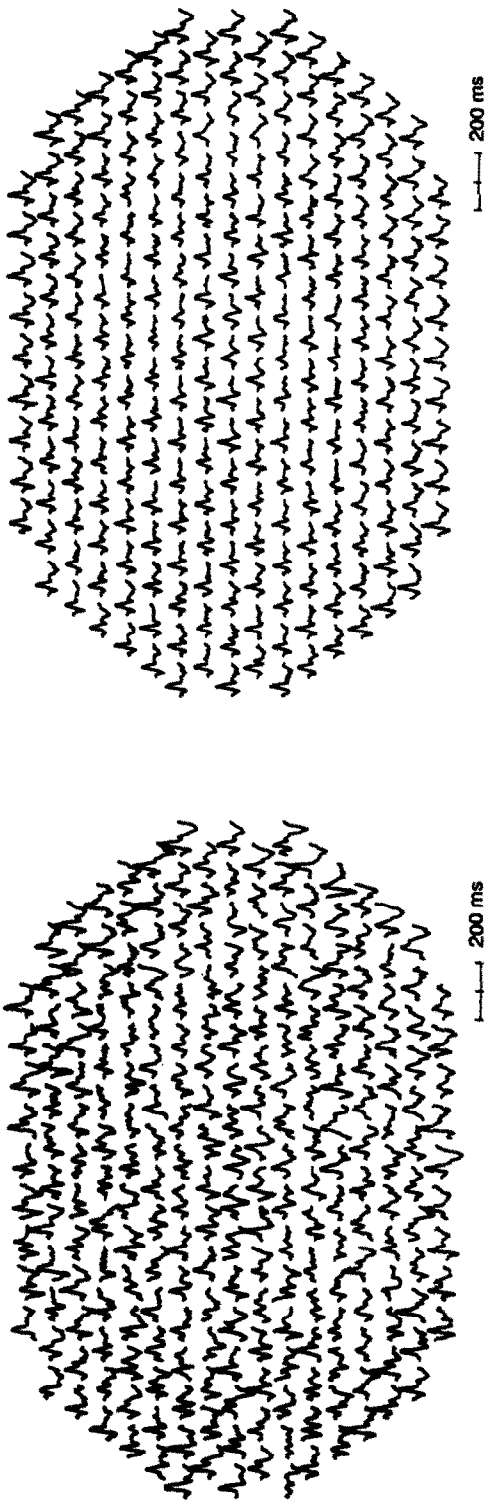

FIGURE A1. Example illustrating the effectiveness of the artifact elimination technique. (a) Response array extracted from a heavily contaminated 16 min record from subject VG. (b) The same data after artifact extraction.

This technique effectively replaces the contaminated segment by a response estimated from the entire record. Conceptually, this method is related to the well known procedure which replaces the extremes (outlayers) of an ensemble by the mean. It offers the great advantage that the response superimposed on the artifact is not discarded, even though no assumptions have to be made concerning the specific wave form of the artifacts. Please note, however, that this approach cannot be successful if the contamination resembles random noise rather than relatively sparse, isolated artifacts.

A more detailed description of the procedure is given by Sutter (1992).

APPENDIX C

A popular amplitude measure is the root mean square or r.m.s. in a given time interval (a, b):

$$A = \sqrt{\mathbf{r}\cdot\mathbf{r}} = \sum_{i=a}^{b} (r_i)^2 \tag{A3}$$

This measure requires no assumptions concerning the shape of the responses, except that their essential components are contained in the interval (a, b). With r.m.s. amplitude estimates, each individual response serves as its own template. They are less accurate since the noise appears in both factors of (2). In addition, there is always a positive contribution of the noise that needs to be estimated and subtracted. The r.m.s. of an estimated response r consisting of a signal s and additive noise n is given by

$$\sqrt{\mathbf{r}\cdot\mathbf{r}} = \frac{\mathbf{r}\cdot\mathbf{r}}{\sqrt{\mathbf{r}\cdot\mathbf{r}}} = \frac{\mathbf{s}\cdot\mathbf{s} + 2\mathbf{s}\cdot\mathbf{n} + \mathbf{n}\cdot\mathbf{n}}{\sqrt{\mathbf{r}\cdot\mathbf{r}}}. \tag{A4}$$

The last term in this expression is always positive and generates a systematic error. The contribution of this term can be estimated from an independent measurement of the noise and then subtracted. This leads to the following amplitude estimate:

$$A = \sqrt{\mathbf{rr}} - \frac{\mathbf{n}\cdot\mathbf{n}}{\sqrt{\mathbf{r}\cdot\mathbf{r}}} = \frac{\mathbf{s}\cdot\mathbf{s} + 2\mathbf{s}\cdot\mathbf{n}}{\sqrt{\mathbf{s}\cdot\mathbf{s}}}. \tag{A5}$$

The amplitude measure (A5) can be used if there is no appropriate template available. However, if the signal wave form s is known, the amplitude of a response can be estimated as the dot product with the normalized signal (template):

$$\mathbf{t} = \frac{\mathbf{s}}{\sqrt{\mathbf{s}\cdot\mathbf{s}}} \quad \text{response template} \tag{A6}$$

$$A = \mathbf{t}\cdot\mathbf{r} = \frac{\mathbf{s}\cdot\mathbf{s} + \mathbf{s}\cdot\mathbf{n}}{\sqrt{\mathbf{s}\cdot\mathbf{s}}} \tag{A7}$$

The resulting experimental error is only about half as large as with the corrected r.m.s. estimate (A5).