Summary and Conclusions

1. The purpose of this study was to describe how the responses of neurons in inferior temporal (IT) cortex represent visual stimuli. In the preceding paper we described the responses of IT neurons to a large set of two-dimensional black and white patterns. The responses to different stimuli showed temporal modulation of the spike trains. This paper develops a method for quantifying temporal modulation and shows that the stimulus determines the distribution over time, as well as the number, of spikes in a response.

2. The responses were quantified using an orthogonal set of temporal waveforms called principal components. The principal components related to each neuron were extracted from all the responses of that neuron to all of the stimuli, regardless of which stimulus elicited which response. Each response was then projected onto the set of principal components to obtain a set of coefficients that quantified its temporal modulation. This decomposition produces coefficients that are uncorrelated with each other. Thus each coefficient could be tested individually, with univariate statistics, to determine whether its relation to the stimulus was nonrandom.

3. The waveforms of the principal components are unconstrained and depend only on the responses from which they are derived; hence, they can assume any shape. Nonetheless, the 21 neurons we analyzed all had principal components that belonged to only one of two sets. The two sets could be characterized by their first principal component, which was either phasic or tonic. This suggests that these neurons may use as few as two different mechanisms in generating responses.

4. The first principal component was highly correlated with spike count, and both were driven by the stimulus. Higher principal components were uncorrelated with spike count, yet some of them were also driven by the stimulus. Thus the principal components form a richer description of the stimulus-dependent aspects of a neuronal response than does spike count.

5. Bootstrap tests showed that several principal components (usually 3 or 4) were determined by the stimulus. Since higher principal components were not correlated with the spike count, the stimulus must have determined the distribution of spikes in the response as well as their number. However, it is possible that the number and distribution of spikes are both determined by the same characteristics of the stimulus. In this case, the temporal modulation would be redundant, and a simple univariate measure would be sufficient to characterize the stimulus-response relationship. The next paper in this series shows that the distribution of spikes is independent of their number and conveys additional information about the stimulus (12).

Introduction

The goal of this study is to develop a new methodology for studying the way neuronal spike trains represent visual patterns. Such an approach must be capable of analyzing data from experiments in which more than one
stimulus parameter is varied (18). A general
to determine what stimulus fea-
tures a visual neuron is sensitive to, and how
these features are represented by the neuron,
requires a complete description of both the
stimuli and the responses. A complete quan-
titative description of the stimuli was achieved
in this study by restricting the stimulus set to
black and white Walsh patterns.

In the preceding paper we stimulated single
inferior temporal (IT) neurons with the mem-
bers of a stimulus set based on complete or-
thogonal two-dimensional Walsh functions.
The responses of IT neurons to these stimuli,
which can be regarded as fundamental features
of any black and white picture, appeared to
show modulation of spike distribution as well
as modulation of spike count (18). In this pa-
per we quantify the temporal modulation,
taking into account the distribution as well as
the number of spikes in the responses. Our
goal in this paper is to test whether temporal
modulation is determined by stimulus fea-
tures.

To quantify temporal modulation and to
identify intrinsic response characteristics that
could potentially represent stimulus features,
we used the Karhunen-Loeve transform. This
transform provides a description of the re-
sponses in terms of orthogonal waveforms
called principal components. The coefficients
of these principal components are uncorrel-
ated with each other and thus can be analyzed
separately. Principal component analysis has
been used for many years in electrophysiology
for the study of evoked potentials (7) but has
not been applied to the study of spike trains.
A separate set of principal components was
extracted for each neuron from all of its re-
sponses, regardless of which stimulus had been
presented. A bootstrap statistical test of stim-
ulus-response relation developed in the pre-
ceding paper (18) was then applied to the
principal components of each neuron to de-
termine which of them were stimulus-depen-
dent.

The waveform of the principal components
for each neuron are free to assume any shape,
i.e., they are not constrained to be analytical
functions such as sines and cosines. Instead,
the shapes of the principal components are
determined by the second-order statistics (co-
variances) of that neuron's data. Our results
showed that the sets of principal components
from the 21 cells analyzed fell into just two
groups.

When the principal components were
tested, it was found that the coefficients of
three or four of them depended on the stim-
ulus presented. Since only the first principal
component was correlated with spike count,
our findings established that the distribution
of spikes, as well as their number, depended
on stimulus features. Thus the results prove
that stimulus features determine the temporal
modulation of the response. Some of this work
has appeared in abstract form (15, 16, 17).

METHODS

The data used in this paper were recorded from
single neurons of inferior temporal (IT) cortex of
awake rhesus monkeys. The methods for training,
single-unit recording, and data collection were
described in detail in the preceding paper (18). The
stimulus set consisted of either 64 or 128 black and
white stimulus patterns based on two-dimensional
Walsh functions. The details of the computation
of the Walsh functions, and how the stimuli were
presented, were also given previously.

In order to completely quantify the responses,
we applied an orthogonal transform, the Karhunen-
Loeve transform, to the responses. The Karhunen-
Loeve transform was chosen instead of more fa-
miliar analytic transforms, such as the Fourier or
Walsh transforms, because it has three unique
properties. The basis vectors used by the Karhunen-
Loeve transform, called principal components, are
ordered to represent decreasing proportions of the
total variance across the data. The Karhunen-Loeve
transform is also optimal, in the mean-square-error
sense, for representing the data with a truncated
sum of components. The most important property
needed for our analysis is that the coefficients of
the decomposition are uncorrelated, which can not
be guaranteed for any other orthogonal transform
(1, 6).

The Karhunen-Loeve transform was applied to
spike density representations of the individual re-
sponses. For each neuron, spike density functions
of the individual responses to a given stimulus and
their means were formed as previously described
(18). The standard error of the mean response to
each stimulus was also calculated at each point in
time. The individual spike densities (sampled at 1-
ms intervals) were low-pass filtered with a nonre-
cursive digital filter (−3 dB at 17 Hz, length, 41)
and were then resampled at 6-ms intervals to
obtain a response vector with 64 elements, thereby
spans 6 × 64 = 384 ms. This resampling allows
any response to be represented by a point in a sixty-
four-dimensional vector space. This vector space.
dimensionality was chosen because it gives rise to a covariance matrix of dimension 64 × 64, which is the largest from which our computer (PDP-11/44) can conveniently extract principal components.

**Influence of background activity**

In many analyses of single-unit data the activity during an unstimulated period is subtracted from the response during stimulus presentation, as an adjustment for the background level of the neuron. However, the spike train is generated by a stochastic process, and the statistics of that underlying process may well be different in the stimulated and unstimulated states. Even if the underlying statistics are stationary (i.e., the same in the stimulated and unstimulated condition), the quantity to be subtracted must still be estimated. Because they are estimated, both the stimulated measure and the background measure will contain noise, which will be reflected in the variances of these estimates. Since variances always combine additively, subtracting a measure of the unstimulated activity from the stimulated measure will increase the variance of the measure of the response. Furthermore, it is not necessary to account for the activity during the unstimulated period if the analysis is confined to stimulus-driven intervals.

We were interested only in comparing the responses to different Walsh stimuli. Hence, our analysis was confined to the stimulus-related portions of the spike trains. We did not include any measure of neuronal activity during unstimulated intervals.

**Extraction of the principal components**

The principal components of the responses are calculated using the variances and covariances of the data set, and thus are characteristic of the population from which they are derived. To extract the principal components from the data, the average of all the responses is first subtracted from each individual response. This subtraction is not the same as subtracting the background rate from the pre-stimulus interval. Once calculated from the responses to all the stimuli, the average response can be considered a constant (i.e., noise free), and since the same average is subtracted from each response, the variance of the response measure is not increased. The net effect of this subtraction is a translation of all the responses to a new origin in the response space. This leads naturally to the quantification of the responses by a representation of their deviations from that average response (1).

The variances and covariances for each neuron’s deviation from the average were calculated from its sixty-four-dimensional vector and averaged to form a covariance matrix. These calculations are represented in matrix notation as

\[ \Sigma_0 = \text{E}[x_i - \bar{x} (x_i - \bar{x})] \]  

(1)

where \( \text{E}[\cdot] \) is the expectation operator (i.e., it computes the average), \( \bar{x} \) is the average of all the response vectors in the data set, and \( x_i \) is the \( i \)-th response. (Note that a vector, \( x_i \), is a column vector, whereas its transpose, \( x_i' \), is a row vector. Hence their product, \( x_i x_i' \), is a square matrix.) If the response vector has dimension \( N \), the covariance matrix, \( \Sigma_D \), has dimension \( N \times N \). The diagonal elements of \( \Sigma_0 \) are just \( \bar{x}^2 \), the variances of the spike density function at each point in time. The off-diagonal elements, \( \Sigma_{ij} \), are the covariances between points of the spike density function at different times. The principal components are the eigenvectors of this matrix. Eigenvectors have the special property that they change only by a scale factor (called the eigenvalue) but not direction, when multiplied by their matrix (21).

The eigenvectors and the related eigenvalues were extracted from the covariance matrix by the QR method (algorithm 254 of the collected algorithms of the Association for Computing Machinery) (5) and used to form an \( N \times N \) matrix with the eigenvectors as orthogonal columns

\[ \Phi = [\phi_1 \cdot \cdot \cdot \phi_{n-1}] \]  

(2)

The eigenvectors of a matrix are only determined to within a scale factor; i.e., their orientations in space, but not their lengths, are specified. The scale is chosen (by convention) so that each eigenvector has unit length.

**Calculation of coefficients**

The set of response transforms was computed for each neuron with the basis vectors derived from those responses. The means and standard errors of the first 10 principal components were plotted. The proportion of the variance of the data accounted for by each principal component was graphed, along with the 95% confidence interval of the eigenvalue estimate, in a score plot (11). Tests of the statistical properties of the decomposition were done using the Mahalanobis-D statistic (11) and the bootstrap simulation method (3, 18).

The KL transform decomposition is defined as the projection of the deviation of the response from the mean response (\( \bar{z} \)) onto the set of principal components, \( \Phi \). This can be represented as a matrix transformation of the response vector

\[ y = \Phi (x - \bar{z}) \]  

(3)

where \( x_i \) is the \( i \)-th response vector, \( \bar{z} \) is the average response, \( y \) is the transform vector, and \( \Phi \) is the matrix of eigenvectors from Eq. 2. The components of the vector \( y \) are the coefficients of the KL transformation for response \( i \).

The significance of the KL transformation in producing uncorrelated coefficients can be appreciated from the matrix description of the data. In general, the covariances in the covariance matrix
of the response data (off-diagonal terms) are not zero (i.e., the points in a response are serially correlated in time). After transforming the data to some other domain, e.g., the frequency domain by means of a Fourier transform, the covariance matrix of the transformed data will not be diagonal. In other words, the coefficients of the transform will, in general, be correlated with each other. Of all linear transforms only the KL transform produces a transform domain covariance matrix that is guaranteed to be diagonal, i.e., all covariances between coefficients of the decomposition are zero.

RESULTS

The results shown in this paper are based on the analysis of the responses from a subset (21 of 48) of differentially responsive neurons reported in the preceding paper that responded to members of the Walsh based stimulus set (18). The neurons chosen for analysis were those with the greatest total number of successfully completed trials per stimulus. These neurons were recorded from five hemispheres in four monkeys.

Temporal modulation of neuronal responses

Figure 1A shows the low-pass filtered spike densities of the responses from one neuron to the 64 positive contrast stimuli (cf. Fig. 7, Richmond et al., 18); Fig. 1B shows the responses to the 64 contrast reversed stimuli for another neuron (cf. Fig. 8, Richmond et al., 18). In both panels of Fig. 1, the mean spike density is shown by the solid line. The dotted lines show the standard errors of the mean computed at each point in time. Each diagram shows the response beginning 40 ms after stimulus onset and lasting 384 ms. The spike density functions represent the overall pattern of activity in the responses rather than the exact position of any individual spike.

Principal components

Figure 2A shows the principal components for the neuron of Fig. 1A, and Fig. 2B shows the principal components for the neuron of Fig. 1B. In the preceding paper neuronal responses were represented by their average spike densities, the overall probability that a spike would occur at some time after the stimulus onset. This average spike density is shown in the upper left of each panel in Fig. 2 (labeled AVG). The horizontal bar in this panel alone indicates zero probability of spike occurrence. The first eight principal components are also shown ($\phi_0 - \phi_7$). The principal components represent the deviation from the average probability of spike occurrence. The principal components form an orthonormal set (i.e., they are independent of each other and each has unit length). The horizontal bars in these panels indicate zero deviation from the average spike density function. Each trace lasts 384 ms.

Relationship of principal components to spike count

Scatter diagrams were plotted to show the relationship between the spike count and the first principal component for the mean response to each stimulus for each neuron. The correlation of the first principal component with the spike count was 0.82 or higher for each of the 21 neurons. An example of this result is shown in Fig. 3A, where the correlation is 0.89. Therefore, the spike count and first principal component are representing closely related characteristics of the spike train. Since the principal components are by definition uncorrelated, the first and second principal component is highly correlated with the spike count, the second and higher principal components should be largely uncorrelated with the spike count. The correlation of the second principal component with the spike count was 0.33 or less for all 21 neurons. An example from the neuron used in Fig. 3A is shown in Fig. 3B, where the correlation is only 0.19.

Efficiency of principal component representation

The amount of variance across all of the data that is represented by each principal component is shown in Fig. 4 (solid fulling lines, left ordinate). The dotted lines show the 95% confidence intervals for the variance estimates. The rising solid lines (right ordinate) show the cumulative variance, which is the same as the total power, that can be accounted for by including progressively more principal components.

For all the neurons analyzed, the first principal component represented between 18 and 48% of the variance across the data set, whereas the second principal component represented between 5 and 12% of the variance. The results in Fig. 4 illustrate the extremes, Fig. 4A showing the case with the largest, and Fig. 4B the case with the least amount of vari-
FIG. 1. The spike density functions that represent the mean responses to each of 64 Walsh stimuli. A and B show the responses from 2 different IT neurons. Each set is arranged in an $8 \times 8$ grid. The responses are in positions that correspond to the stimuli shown in Fig. 2 of the preceding paper (18). The solid lines show the mean responses and the dotted lines show the standard error for the 2 neurons described in Figs. 3 and 4 of Ref. 18. The horizontal base line is at zero probability of spike occurrence. To emphasize the waveforms of the responses, each spike density for a given neuron was drawn with the same scale. The number of responses per stimulus was 4 or 5 in both $A$ and $B$. 
FIG. 2. Principal components for the 2 neurons shown in Fig. 1. The average response is located in the upper left corner of each panel. The other curves represent the first 8 principal components for each neuron. The principal component set in A is characterized by a tonic first principal component. The set in B is characterized by a phasic first principal component. The base line is at zero probability of spike occurrence for the AVG curve and at zero deviation from the average for the others. The principal components are normalized to have vector length equal to one. All the principal components were drawn with the same vertical scale.
Fig. 3. Correlation of principal components with spike count. A: the correlation of the spike count with the first principal component, \(\phi_0\), for the mean response to each of the 64 Walsh stimuli is good (\(r = 0.89\)). B: the correlation of the spike count with the second principal component, \(\phi_1\), is poor (\(r = 0.19\)).

The difference in the amount of variance represented by the first principal component occurred largely because of a redistribution of variance among the first 10 principal components. For the neuron shown in Figs. 1A, 2A, and 4A, 90% of the variance is represented by the first 9 components, whereas for the neuron shown in Figs. 1B, 2B, and 4B, 90% of the variance is represented by the first 11 principal components. The amount of variance encompassed by the first principal component was not related to whether the neuron had been tested with 64 or 128 visual stimuli.

SIMILARITY OF THE PRINCIPAL COMPONENTS ACROSS IT NEURONS. The general shapes of the first three principal components appeared to fall into only two classes. Each panel of Fig. 5 shows sets of principal components from several neurons overlaid on one another, revealing the similarity of the principal components within each class. We differentiated the two groups, sustained and phasic, on the basis of the first principal component, since it is the one whose shape is unconstrained in any way at the time of derivation. The set in Fig. 5A (showing examples from three neurons) has a sustained (tonic) first principal component, \(\phi_0\). The second principal component, \(\phi_1\), is biphasic, and \(\phi_2\) appears triphasic. Nine of the 21 neurons tested belong to this set. The second family of principal components is shown in Fig. 5B (with examples from four neurons). The first component, \(\phi_0\), is monophasic, \(\phi_1\) is a biphasic function, and \(\phi_2\) has a triphasic weighting. Twelve of the 21 neurons in our population fall into the second set.

To test the hypothesis that these principal components form two distinct populations, we
A

Normalized Variance

Cumulative Variance

N = 313
N₉₀ = 9

Component

B

Normalized Variance

Cumulative Variance

N = 515
N₉₀ = 11

Component

FIG. 4. Scree plots. Normalized variance that is accounted for by each principal component (descending solid line). The horizontal axis is a logarithmic scale of the ordinal principal component number (i.e., φₙ is number one). The left ordinate shows the percentage of variance accounted for across the whole data set. The dotted lines define the 95% confidence interval for the variance. The rising solid line shows the cumulative variance (right ordinate). These examples show the two extremes of principal component variance distribution. N is the number of responses from the neuron. N₉₀ is the number of components needed to account for 90% of the power or the total variance of the responses.
extracted new principal components from a set composed of the first principal components of the 21 neurons. The new first principal component divides the neurons into two non-overlapping groups of nine and twelve. The Mann-Whitney $U$ test showed this to be significant at the $P < 0.001$ level (19).

SPECTRAL DECOMPOSITION OF IT NEURONAL RESPONSES. The individual responses of each neuron were represented as weighted sums of that neuron's principal components. The coefficients of the representation quantify the temporal modulation of the responses. These weighting coefficients ($y$, calculated with Eq. 3 above) form a spectral decomposition of the responses. The coefficients for the individual responses to each stimulus were averaged to obtain a mean and variance for each principal component. Figure 6 shows examples of responses to three stimuli with their line spectra. Fig. 6, A, B, and C show the mean spike densities as solid lines with the standard errors as dotted lines. In Fig. 6, D, E, and F, the cor-
responding line spectra of those responses are shown. These examples are from the neuron illustrated in Fig. 1A. The principal components used for the decompositions shown in Fig. 6 are illustrated in Fig. 2A. Each line spectrum consists of 10 vertical hollow bars on a horizontal base line. Each hollow bar is proportional in length to the value of the mean of the coefficients of one principal component, and the coefficients of the first 10 principal components are shown, from left to right. The thin vertical line in each hollow bar shows the standard error of the mean for that coefficient.

If the coefficients of the first principal component, $\phi_1$, are similar in two responses and those of the second principal component, $\phi_2$, are different, the spike counts will be similar in the two responses but the distributions of the spikes will be different. In the example shown in Fig. 6, the spike count was $9.0 \pm 6.7$ (SD, $n = 4$) for Fig. 6B and $9.0 \pm 11.0$ (SD, $n = 4$) for Fig. 6C. The coefficients of the first component were similar, whereas the coefficients of the second principal components were similar in magnitude but of opposite sign. Hence the difference in the coefficients of the second principal component quantifies the difference between the single-peaked spike density in Fig. 6B and the double-peaked spike density in Fig. 6C.

**Response space**

The principal components can be regarded as defining a multidimensional response space. The neuron’s response to a particular pattern corresponds to a point, or a vector of transform coefficients, in that space. Spaces with more than two dimensions cannot be drawn. Thus, to aid visualization of the distribution of the responses, their locations in the space are projected onto planes to form scatter diagrams. In Fig. 7, $A$ and $B$, the amounts of the first principal component needed to represent the responses of one neuron for all patterns are plotted along the abscissa. The amounts needed for the second principal component are shown on the ordinate of Fig. 7A, and the amounts of the third principal component are
FIG. 7. Scatter diagrams of neuronal response space. Each panel is a two-dimensional projection of a multidimensional neuronal response space for 1 neuron. A shows the projection of the responses onto the $\phi_0$ vs. $\phi_1$ plane. The horizontal
shown on the ordinate of Fig. 7B. The numbers in the boxes show the locations of the average response to that stimulus in this plane. The size of a box shows the standard errors of the mean. The values are scaled to have equal variances for equal distances along both axes. Use of axes based on equal variances emphasizes the differentiability of the patterns in the scatter plots at the expense of obscuring the relative amount of power represented by each principal component.

The horizontal and vertical lines in the lower left corner show how far apart two patterns in the response space must be to be significantly different along that dimension ($P < 0.05$). This 95% confidence interval is calculated by using the Mahalanobis-D statistic, a distance measure that is approximately $\chi^2$ distributed and which gives the reliability of separations along each axis in a multivariate space ($11$).

Each panel shows the projections in a different direction through the multidimensional response space. Figure 7 shows that some stimuli which are indistinguishable along one dimension in the response space can be distinguished along another dimension. For example, in Fig. 7A, Walsh stimulus 07 can be distinguished from stimulus 32 along the abscissa (i.e., $\phi_0$) but not along the ordinate (i.e., $\phi_1$), whereas stimulus 55 can be distinguished from stimulus 51 along the ordinate but not along the abscissa. Using the third component, shown in Fig. 7B, still other stimuli become distinguishable which were not differentiable in Fig. 7A (e.g., stimuli 41 and 59).

**Significance of the principal components**

Although the representations of the various patterns are scattered through the response space as in Fig. 7, few of them reach the 95% confidence interval of the Mahalanobis-D statistic for being different from each other along any individual response axis. Given the large number of stimuli, about 5% of the separations would be expected to reach the 95% confidence interval by chance alone. We therefore examined whether the responses to the stimuli were in fact randomly distributed in the response space.

Parametric statistical tests are based on knowledge of the underlying distribution from which the samples are drawn. In general, such tests rely on knowledge that the normal distribution, or some other analytically described distribution, is an adequate representation of the data. When the data come from populations whose distribution can not be assumed, distribution-free tests are used. One class of distribution-free tests, the computational tests, manipulate the raw data by random resampling to estimate relevant population parameters.

Using the bootstrap statistical method, a computational method described in the preceding paper (18), we tested the null hypothesis that the distributions of the coefficients came about by a random relation between the responses and the stimuli. This can be done using the coefficients of the Karhunen-Loeve transform individually because the coefficients are uncorrelated. This procedure generated an estimate of the probability density function of each coefficient under the hypothesis that the responses were independent of the stimuli. This hypothetical density function was then compared with the density function calculated from the actual data set.

Figure 8 shows density functions for the spike count and the first three principal components for one neuron. In each panel, the solid line shows the distribution function generated by the bootstrap simulation, and the dashed line shows the distribution function of the actual coefficients found in the experiment. The difference between the two was tested with the nonparametric Kolmogorov-Smirnov test. In all four panels the difference was significant ($P < 0.05$). For three of the neurons, only one principal component was determined by the stimuli (significance was taken as $P < 0.05$). For 18 neurons two or more principal com-

![Fig. 7 (continued)](image)

*axis* represents the magnitude of the coefficient of the first principal component and the vertical axis represents the magnitude of the coefficient of the second principal component. The numbers are positioned at the mean response to stimuli with that number. The box around the number represents the standard errors (the scales have been normalized to give equal variance for equal distances on both axes). $B$ shows the $\phi_0$ vs. $\phi_1$ plane. There are patterns that can be separated along $\phi_0$ that were not separable along either $\phi_0$ or $\phi_1$ (see text). The horizontal and vertical fiducial lines in the lower left corner indicate the 95% confidence interval (based on the Mahalanobis-D statistic) for differentiating between 2 of the mean responses.
components were determined by the stimuli. Table 1 shows the distribution of the number of neurons that had each number of significant principal components. The portions of the responses that show at least some nonrandom stimulus-driven time modulation typically add up to \(\sim 50-80\%\) of the power of the responses.

**TABLE 1. Distribution of the number of significant principal components**

<table>
<thead>
<tr>
<th>No. Significant Principal Components</th>
<th>No. Neurons</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

For each neuron, the coefficients of the principal components were tested individually to see whether they were randomly related to the stimuli. The number given is the number of principal components that were significantly nonrandom by the Kolmogorov-Smirnov test \((P < 0.05)\).

**DISCUSSION**

These experiments had as their major goal the identification of those parameters of the response waveform of IT neurons that were driven by the stimulus. We reached this goal in two steps. The first step was the development of a method to represent and quantify the temporal modulation of the responses of visual system neurons with principal components. The second step showed, with the bootstrap statistical method, that temporal modulation of the spike train, as represented by individual principal components, was related to the particular stimulus presented.

The neuronal spike trains were converted into estimates of the probability of spike occurrence (i.e., spike densities) for individual responses. This method preserves temporal modulation. The Karhunen-Loève transform (i.e., Eq. 3) decomposed these spike density functions into their principal components. The coefficients of the principal components constitute quantitative descriptions of the
temporal modulation of the neuronal responses.

Sensitivity to stimulus parameters was tested in two ways. First, the data were displayed in a multidimensional response space (cf. the scatter diagrams of Fig. 7) to determine whether the separation between responses to different stimuli exceeded 95% confidence intervals. These confidence intervals were computed under the assumption that the data were normally distributed (11). Second, to test for stimulus-dependence in a distribution-free manner, a bootstrap technique was also used (3, 18). This analysis showed that three or four principal components for each neuron were driven differentially by the stimulus ($P < 0.05$) for the majority of tested neurons.

Only the first principal component was correlated with the spike count (Fig. 3). Since the coefficients of the principal components are uncorrelated with each other (11), and since more than one coefficient is determined by the stimulus, the stimulus-response relation must be multidimensional. Hence, no single measure, such as spike count, can represent all of the stimulus-dependent modulation of a response. This is the major finding of the experiments reported in this paper.

Principal component waveforms

Another result of this experiment is that the sets of principal components from all the neurons could be separated into two distinct groups. These two groups are characterized by the shape of the first principal component, $\phi_0$, which was either phasic or tonic. This does not mean that a given neuron produces only phasic or tonic responses. Neurons with a phasic $\phi_0$ could have sustained responses to some stimuli. Similarly, neurons with a tonic $\phi_0$ could have transient responses to some stimuli. Therefore, characterization of neurons in IT as either phasic or tonic is valid only for the ensemble of a neuron's responses to many different stimuli, and this determination can not always be made on the basis of the neuron's responses to a few individual stimuli.

When a system responds to various inputs, the responses often have stereotyped or common elements. Identification of those common elements can give insight into the characteristics of the system under study. Principal components are the eigenvectors of a covariance matrix, and as such they represent intrinsic modes of the statistical properties of a neuron. The response to any stimulus will contain some of these modes. Thus the principal components are a reflection of the intrinsic mechanism underlying the generation of responses.

There is no a priori reason that the principal components need take any particular shape. The principal components are constrained only by the characteristics of the data. In this study the only external constraints affecting these characteristics were the two steps of low-pass filtering done by the Gaussian convolution ($-3$ dB at 14 Hz) and before the resampling ($-3$ dB at 17 Hz). Since the filtering was the same for all neurons, this could not account for the difference in $\phi_0$ across neurons. The similarity of principal component shapes within these two groups suggests that IT neurons may use as few as two mechanisms to encode stimulus characteristics in spike train modulation.

**Source of temporal modulation**

Our analysis shows that temporal modulation of neuronal activity by different visual stimuli is significant. The temporal modulation must arise from summation of excitatory and inhibitory influences occurring with varying time courses. For the visual system, models of receptive field structure in retinal ganglion cells (4, 10) and striate cortex (9) utilized spatial distributions of excitatory and inhibitory mechanisms to show how the responses to spots, edges, bars, and sinusoidally modulated gratings could arise. For example, the properties of simple cells have been modeled with spatially separated excitatory and inhibitory regions, whereas those of complex cells have been modeled with large regions of overlap of excitatory and inhibitory regions. Recently this overlap in complex cells has been shown to have inhomogeneous spatial mixing of inhibition and excitation (20). Such separated regions could have different time constants and could provide a source for temporally modulated signals.

There are few data that directly address the nature of receptive field structure in IT cortex. Neither the very earliest experiments by Gross et al. (8), who used bars and rectangles as well as more complex stimuli, nor the more recent experiments by Pollen et al. (13), who used stimuli based on Gabor func-
tions, i.e., sine-wave gratings modulated by a Gaussian envelope, have found evidence for receptive field substructure in IT neurons. The stimulus-dependent temporal modulation found in our studies, however, may indicate the existence of a subtle receptive field substructure.

These speculations suggest how temporal modulation might arise from earlier levels. Another possible source of the temporal modulation is feedback. Connections with neurons in the same, or later, levels of processing could result in a network that would account for the temporal modulation we found.

**Stimulus-dependent temporal modulation**

The dimensionality of the stimulus-driven part of the response is determined by those principal components that have nonrandom coefficients (usually two to four). This analysis suggests, but does not prove, that more than one component of the response waveform conveys independent information about the stimulus. An alternative explanation is that these principal components are all driven by the same feature of the stimulus. To exclude this alternative explanation we must show that the coefficients of the principal components are not only uncorrelated but also statistically independent. Only when data have a normal distribution, does uncorrelated necessarily imply independent (2). As seen in Fig. 8, our data do not seem to have a normal distribution. Hence, it is possible that the principal components, although uncorrelated, convey redundant information. The next paper in this series uses information theory to show that the principal components do indeed contain independent information about stimulus parameters (12).

REFERENCES