Sampling-theory analysis of spatial vision

Wilson S. Geisler and David B. Hamilton

University of Texas, Austin, Texas 78712

Received February 19, 1985; accepted August 19, 1985.

Spatial-vision research has been largely concerned with measuring and understanding the consequences of receptive-field properties measured by single-unit recording. However, in order to understand spatial-information processing in the visual system, it is equally essential to know the densities and the distribution patterns of the receptive fields. If the receptive fields are not arrayed properly across the visual field, spatial information will be lost. It has been argued, on the basis of the Whittaker–Shannon sampling theorem, that the receptors of the fovea sample the retinal image at a high enough rate to preserve essentially all the available spatial information. In this paper we show how two-dimensional sampling theory can be used to determine which combinations of receptive-field shapes and sampling patterns would preserve all spatial information from the receptors. This analysis will prove useful for determining, in conjunction with electrophysiological and anatomical evidence, what spatial information is or is not being transmitted by a given stage of the visual pathway. It may also prove useful for developing and testing theories of spatial vision.

INTRODUCTION

The optics of the eye always blur the retinal image to some extent. Even under optimal conditions of small pupil diameter and accurate accommodation, foveal images are band limited to about 60 cycles/degree (c/deg). The photoreceptors must, of course, sample the retinal image at a number of discrete locations. In the fovea, the receptors have a diameter of around 30 sec of arc and therefore are taking around 120 samples of the retinal image per degree of visual angle. A number of investigators have argued, on the basis of the Whittaker–Shannon sampling theorem, that this is the most efficient receptor sampling scheme. According to the sampling theorem, if a signal is band limited to W c/deg, then the signal can be perfectly saved and reconstructed (if desired) by taking 2W, or more, regularly spaced samples per degree. The quantity 2W is called the Nyquist rate. Any fewer samples than the Nyquist rate will result in information loss. Thus it appears that the human receptor sampling scheme in the fovea is most efficient in the sense that it extracts all the information in the retinal image with the minimal number of samples (120/deg).

If the receptors (at least in the fovea) are preserving nearly all the information in the retinal image, what about the rest of the visual pathway? Each layer of neural units after the receptors is sampling the output of the layer before it. However, instead of sampling at small nonoverlapping points, the postreceptor units are sampling with relatively large overlapping receptive fields. How can we determine whether a level in the visual pathway is encoding all the information in the visual image? Some progress has been made in answering this question, but so far the analyses have been rather qualitative and nonrigorous. One goal of this paper is to show how two-dimensional sampling theory can be used to determine which combinations of receptive-field shapes and sampling patterns would preserve all spatial information in the retinal image and which ones would not.

The use of two-dimensional sampling theory to analyze the transmission of information in the visual pathway is of interest for several reasons. First, the theory may help in determining, in conjunction with electrophysiological and anatomical evidence, whether any given stage of the visual pathway is transmitting all the spatial information sent to it, and, if it is not, what information is being removed. Second, the theory may be useful in rejecting psychophysical or electrophysiological models of spatial vision by showing that a model implies loss of visual information known to be used by real visual systems. Finally, the theory may be of help in developing models of spatial vision. For example, it seems parsimonious to consider first models containing spatial sensors with receptive-field shapes that sample the retinal image with the minimum number of units. Sampling theory can be used to determine the class of models that have this property.

Only linear systems will be considered in this paper. Even so, there are a tremendous number of sampling schemes that will preserve all the spatial information. Thus it is useful to introduce some terminology for describing and categorizing different types of sampling schemes.

Sampling Unit

A sampling unit is a single linear spatial sensor that is completely described by giving its weighting function (receptive-field shape) and position. We allow both real-valued and complex-valued weighting functions. If a sampling unit has a real-valued weighting function it can be thought of as a single neural unit (e.g., a receptor or a simple cell). If a sampling unit has a complex-valued weighting function then it can be thought of as a pair of neural units, one for the real part and one for the imaginary part. For example, a pair of even- and odd-symmetric simple cells located at the same position can be regarded as a sampling unit with a complex weighting function.

Sampling Array

A sampling array is a collection of sampling units scattered across some region of the visual field that all have the same weighting function. The positions of the units form a sampling lattice. The cone photoreceptors in the fovea form a sampling array. Similarly, the collection of even-symmetric
simple cells with the same orientation and spatial-frequency bandwidth that cover the foveal region might be regarded as a sampling array. However, we allow some flexibility in defining sampling arrays. For example, a collection of even- and odd-symmetric pairs of simple cells with the same orientation and bandwidth might also be regarded as a single sampling array. In any event, at the cortical level a sampling array could be thought of as a spatial-frequency- and orientation-tuned channel over some region of the visual field.

**Sampling Scheme**

A sampling scheme is the set of all sampling arrays covering some region of the visual field at a given stage in the visual system. Thus the receptors in the fovea might be thought of as a sampling scheme consisting of one sampling array. On the other hand, the set of all simple cells whose receptive fields fall in or cover the fovea form a sampling scheme consisting of a large number of sampling arrays.

**Complete Scheme**

A sampling scheme is complete if it encodes all information sent to it without loss (e.g., without aliasing). If each sampling stage of the visual system is complete, all the information in the visual image will be faithfully transmitted.

**Nyquist Scheme**

A complete sampling scheme will be called a Nyquist scheme if the total number of sampling units per square degree equals the Nyquist rate. In two dimensions, the Nyquist rate is $4W^2$ samples/deg$^2$ for real-valued weighting functions and $2W^2$ samples/deg$^2$ for complex-valued weighting functions. In one dimension, the Nyquist rates are $2W$ and $W$, respectively. The Nyquist rate is generally the smallest sampling rate that a complete scheme can have. Thus Nyquist schemes are the most economical complete schemes (in terms of numbers of sampling units). Sakitt and Barlow proposed a cortical sampling scheme that they conjectured is economical in this sense.

**Locally Complete Scheme**

The weighting function (i.e., the receptive field) for a sampling array will pass only a certain range of spatial frequencies and orientations. If we make the sampling rate of the units high enough, then the sampling array will faithfully transmit all the information within the range of its weighting function. In this case we can think of the sampling array as being complete for this range of frequencies. If each sampling array in a sampling scheme is complete over the range of spatial frequencies and orientations encoded by that array, we say that the scheme is locally complete. For example, pyramid schemes of the sort often used in image processing (e.g., Ref. 7) are locally complete—the sampling arrays at each level of the pyramid completely represent the information within the frequency range that they cover. As we shall see, the so-called Gabor schemes for sampling spatial information are not locally complete.

**SAMPLEING ANALYSIS IN ONE DIMENSION**

Although we are primarily interested in the application of two-dimensional sampling theory, it is instructive first to develop the analysis in one dimension. We have tried to keep the analysis as simple and as intuitive as possible. Most of the basic facts of linear-system analysis used below are described in standard textbooks on the subject (e.g., see Ref. 11).

**Structure of the Problem**

The linear system that we wish to analyze is illustrated in Fig. 1. It is a two-stage system consisting of a receptor and a postreceptor stage; but the analysis can be easily extended to more stages. A visual stimulus, after being low-pass filtered by the optics of the eye, forms an intensity distribution $i(x)$ on the receptor layer. The retinal intensity distribution is then sampled by the receptors. We can describe this sampling process by a receptor weighting function $r(x)$ and a sampling function $s_0(x)$. If the receptor response is linear, then the response of the receptor array is given by the cross correlation of the receptor weighting function and the intensity distribution, all multiplied by the sampling function.

We assume that the receptors are spaced evenly. Thus the sampling function can be represented by a comb function. A comb function is a series of equally spaced impulse functions, each impulse representing a single sample point. The number of impulse functions per degree of visual angle is the sampling rate $NS$. In the fovea $NS \approx 120$ samples/deg.

The output of the receptors is then sampled by units with receptive fields having a weighting function $h_k(x)$. A weighting function could, for example, be implemented by assigning positive (excitatory) and negative (inhibitory) weighting factors to the individual receptor responses followed by a summation of the weighted responses. The form of the weighting function will, of course, be different for each type of sampling unit—the type of unit is represented by the subscript $k$. All the sampling units with the same weighting function form a sampling array. As mentioned above, there may be several different sampling arrays at any given step along the visual pathway.

The output in the space domain, $o_k(x)$, of a postreceptor sampling array is calculated by sampling the cross correlogram.
tion of the weighting function with the output of the receptors:

\[ o_b(x) = s_b(x) \times (h_b(-x) \times [s_b(x) \times [r(-x) + i(x)])], \quad (1) \]

where \( s_b(x) \) is the sampling function (another comb function) for the postreceptor array and \( \times \) is the operation of convolution. [Note that the cross correlation of two functions \( f(x) \) and \( g(x) \) is equal to \( f(x) \ast g(-x) \). In other words, \( g(-x) \) is the impulse-response function associated with the weighting function \( g(x) \).] The output from the whole sampling scheme is the output of all the sampling arrays; \( o_1, o_2, \ldots, o_m \).

The goals of the present analysis can now be stated more specifically:

1. For any weighting function \( h_b(x) \) how many samples/degree are needed in order to extract all the information within the frequency range covered by that weighting function? In other words, what is the minimum rate needed for the sampling array to be locally complete?
2. Which of the locally complete schemes are complete in the sense that they encode without loss all the spatial information in the visual image?
3. Which of these complete sampling schemes is a Nyquist scheme?

**Solution**

In order to answer these questions it is easiest to move into the frequency domain. Taking the Fourier transform of both sides of Eq. (1), we have

\[ o_b(u) = S_b(u) \ast (H_b(-u) \times [S_b(u) \ast [R(-u) \times I(u)]])], \quad (2) \]

where the capital letters represent the Fourier transforms of the corresponding functions in Eq. (1). We now evaluate Eq. (2) step by step.

**Receptor Sampling**

The first step is to multiply the transform of the retinal intensity distribution \( I(u) \) by the transform of the receptor weighting function \( R(-u) \). For convenience, we label the resulting function \( Z(u) \). We call \( z(x) \) [the inverse transform of \( Z(u) \)] the effective retinal intensity distribution. For small receptors, such as those in the fovea, \( z(x) \) is only a slight low-pass-filtered version of the retinal intensity distribution, \( i(x) \).

The second step is to convolve the transform of the effective intensity distribution \( Z(u) \) with the transform of the receptor sampling function \( S_b(u) \). Figure 2A shows the transform of a hypothetical effective intensity distribution, low-pass filtered at frequency \( W \) (e.g., 60 c/deg). Recall that the optics of the eye in the fovea low-pass filter the image at around 60 c/deg. Figure 2B shows the result of convolving \( Z(u) \) with \( S_b(u) \) when the sampling rate is set to the Nyquist limit \( 2W \). The curve in Fig. 2B is, of course, just the Fourier transform of the receptor output.

There are two important things to note about the curve in Fig. 2B. First, it consists of an infinite series of exact replicas of \( Z(u) \) equally spaced along the frequency axis. Second, the replicas just touch one other without crossing. This implies that no information has been lost in the process of sampling. The argument goes as follows: The original stimulus could be reconstructed if we had just the replica over the origin since it is exactly the Fourier transform of the input. But, we could isolate this replica by low-pass filtering the receptor output by a \( W \) c/deg. Low-pass filtering does not add information; therefore it must have already been there.

Figure 3A illustrates the calculations underlying Fig. 2B by pictorially convolving \( Z(u) \) and \( S_b(u) \). The Fourier transform of a comb function with a sampling rate \( NS \) is also a comb function but with a period of \( NS \). Thus, as shown in Fig. 3A, we can obtain the convolution by sliding a comb function \( [S_b(-u)] \) with period \( 2W \) along the frequency axis. At each position we multiply the comb function by \( Z(u) \), get the area under the product, and then plot the result at the current position of the comb function, which is marked with a cross.

As can be seen from Fig. 3A, if the sampling rate were less than \( 2W \), then the replicas would overlap one other. In the frequency ranges in which the replicas overlap, the exact shape of the replicas cannot be recovered; thus it is impossible to reconstruct the original image exactly. This information loss that is due to undersampling is called aliasing. Except for signal-to-noise effects resulting from quantal fluctuations or internal noise (which for present purposes we are ignoring), aliasing is the only source of information loss that is due to undersampling. Thus, with equally spaced samples, the amount of replica overlap is a good measure of the information lost in a sampling scheme.

Figure 3A also illustrates that a sampling rate greater than \( 2W \) would result in gaps among the replicas where no energy is present. The image could, of course, be completely reconstructed from the output obtained at this higher sampling rate, but this is an inefficient scheme because the extra sampling would be unnecessary.
Fig. 3.  A, The convolution of effective intensity distribution and receptor sampling function that was used to obtain Fig. 2B.  B, The convolution of postreceptor filter output with the postreceptor sampling function that was used to obtain Fig. 2D.  C, Illustration of undersampling with the postreceptor units.

Postreceptor Sampling: Real-Valued Weighting Functions

The third step in evaluating Eq. (2) is to multiply the Fourier transform of the receptor output by the Fourier transform of the weighting function \( H_k(-u) \) of the \( k \)th postreceptor sampling array. To begin with, we suppose that the sampling units have real-valued weighting functions. Figure 2C shows the result if the weighting function is that of a one-octave bandpass filter transmitting the highest range of frequencies in the retinal image.

The fourth and final step is to convolve the output of the bandpass filter with the postreceptor sampling function \( S_k(u) \). For certain bandpass filters it is possible to sample below the Nyquist rate implied by the highest frequency transmitted by the filter and still have a complete sampling array. Figure 2D shows the result of sampling the output in Fig. 2C at half the Nyquist rate. Again, note that the replicas just touch one other without overlapping; thus all the information within the 1-octave range has been preserved. Figure 3B illustrates how this result is obtained. It is possible to sample below the Nyquist limit because the aliased information falls outside the bandpass region covered by the 1-octave filter. However, this is not true for all bandpass weighting functions. Figure 3C shows that for a weighting function corresponding to a 1.3-octave bandpass filter, the sampling rate must be at the Nyquist limit of 2\( W \). The solid curves at the bottom show the result if the sampling is at the Nyquist rate. The dashed curves show the additional replicas that would be obtained if undersampling were attempted.

Using diagrams such as those in Fig. 3 it is not difficult to answer our first question, at least for low-pass or bandpass weighting functions. Namely, what sampling rate is required to extract all the information within the frequency range of the weighting function? Let \( u_h \) be the highest spatial frequency transmitted by the filter, and let \( u_l \) be the lowest frequency transmitted by the filter. The minimum sampling rate needed for a sampling array to be locally complete is given by

\[
NS = 2 \times b_u \times r/\text{trunc}(r),
\]

where \( b_u = u_h - u_l \) (the linear bandwidth), \( r = u_b/(u_h - u_l) \), and \( \text{trunc}(r) \) is the nearest whole number that is less than or equal to \( r \). Note that for a given linear bandwidth, the sampling rate needed to extract all the spatial information is smallest when \( r \) is a whole number:

\[
r = 1, 2, 3, 4, \ldots
\]

Thus we consider a weighting function to be efficient if \( r \) is a whole number and inefficient otherwise. Note that Eq. (3) correctly predicts the sampling rate needed for all three cases in Fig. 3.

The class of efficient weighting functions can also be expressed in the more familiar terms of octave bandwidth \( BW \). A receptive-field weighting function can sample spatial information efficiently if and only if

\[
BW = -\log(1 - 1/r)/\log(2),
\]

where \( r \) is a whole number. Thus for sampling efficiency, the bandwidths of the receptive-field weighting functions can only be one of the following: inf, 1, 0.58, 0.42, 0.32, \ldots octaves. (Note that an octave bandwidth of inf is a low-pass filter.) A result similar to Eqs. (3)-(5) can be obtained for multibandpass weighting functions, but we will not consider this case here since biological visual systems do not seem to use such receptive fields.

As defined above, a sampling scheme is the entire set of sampling arrays at a given level in the visual system. We can now answer our second question: Which sampling schemes encode without loss all information transmitted to them? A sampling scheme is complete if it satisfies two conditions: (a) Each sampling array must have at least the sampling rate specified by Eq. (3). (b) The weighting functions associated
with the sampling arrays must cover the entire frequency range transmitted to the sampling arrays. These criteria are not difficult to apply. If a proposed scheme does not satisfy these two conditions, one can then use graphical (e.g., as in Fig. 3) or analytical analysis to determine what information is being lost.

Finally, we consider the question of which complete sampling schemes are Nyquist schemes (i.e., require the minimum total samples). In order to be complete, a scheme must satisfy the two conditions above. In order to be a Nyquist scheme it must satisfy two further conditions: (c) The weighting functions associated with the sampling arrays should not have overlapping frequency ranges. Covering some of the same frequency range with two sampling arrays is a waste of samples. (d) The weighting functions must satisfy Eq. (4) [or equivalently Eq. (5)]. In other words, all the weighting functions must be efficient. If the total frequency range to be encoded by a sampling scheme is W, then a perfectly efficient sampling scheme will consist of a set of weighting functions satisfying Eq. (4) that break this frequency range up into nonoverlapping regions. Since the sampling rate needed in each frequency range is just twice the frequency range [Eq. (3)], the sampling rate of the whole sampling scheme is 2W, exactly the Nyquist limit. All other schemes require more samples.

Postreceptor Sampling: Complex-Valued Weighting Functions

If \( h(x) \) is a complex-valued weighting function, then it is of the form

\[
 h(x) = f(x) + ig(x),
\]

where \( f(x) \) and \( g(x) \) are real-valued functions. Clearly, a sampling unit with a complex weighting function could be implemented in a real visual system by a pair of receptive fields with weighting functions \( f(x) \) and \( g(x) \). The sampling-theory analysis for complex weighting functions is identical to the above analysis, except that the Fourier transform of a complex weighting function need not have a symmetric real part and an antisymmetric imaginary part. Therefore, when we multiply the transform of the weighting function together with the transform of the sampling array, we obtain a more complex set of transformations.

Following the analysis above, let \( F_k(u) \) be the Fourier transform of the real-valued weighting function of an arbitrary low-pass or band-pass filter. Now, define

\[
 H_k(u) = [F_k(u) + sgn(u)F_k(u)]/2,
\]

where \( sgn(u) = 1 \) for \( u \) positive and \( sgn(u) = -1 \) for \( u \) negative. \( H_k(u) \) is the transform of the complex weighting function that transmits exactly the same information as the real weighting function \( f_k(x) \). To see this, note that for positive spatial frequencies \( H_k(u) \) is identical to \( F_k(u) \) but for negative frequencies it is identically 0. Because of the symmetries of \( F_k(u) \) and the Fourier transform of the retinal output—\( Z(u) \) + \( S_u(u) \) (see Fig. 2B)—it is clear that the product \( [Z(u) + S_u(u)]H_k(-u) \) uniquely determines the product \( [Z(u) + S_u(u)]F_k(-u) \) and vice versa. In other words, the transfer functions \( H_k(u) \) and \( F_k(u) \) transmit exactly the same information. If, for example, \( f_k(x) \) is a Gaussian-damped sine wave (a Gabor function), then \( h_k(x) \) is (to close approximation) a complex weighting function whose real part is \( f_k(x) \) and whose imaginary part is another Gaussian-damped sine wave in which the sine wave has been phase shifted 90 deg. In this case, \( h_k(x) \) could be implemented in a visual system by a pair of real weighting functions consisting of an even- and an odd-symmetric Gabor function.

Since \( H_k(u) \) is identically zero for all negative frequencies, the minimum sampling rate for the sampling array to be locally complete is exactly the linear bandwidth of the weighting function

\[
 NS = b_w.
\]

Thus a sampling scheme using complex weighting functions, in the form of Eq. (6), will be complete if (a) each sampling array has a sampling rate at least as great as its bandwidth and if (b) the bandwidths of the sampling arrays collectively cover the entire frequency range transmitted to them. Finally, note that the locally complete Nyquist schemes must satisfy two further conditions: (c) The weighting functions associated with the sampling arrays cannot have overlapping frequency ranges, and (d) the sampling rate for each array must be at the above minimum. If the total frequency range to be encoded is 0 to \( W \) c/deg, then for the Nyquist schemes the total sampling rate is exactly \( W \) samples/deg. This is one half the Nyquist rate, but each sample requires two receptive fields, so the total sampling rate of neural units needed to implement the scheme is still the Nyquist rate of 2W.

## SAMPLING ANALYSIS IN TWO DIMENSIONS

The structure of the analysis in two dimensions is analogous to that of the one-dimensional case. In particular, the output of the sampling array and its Fourier transform are identical to the one-dimensional case [Eqs. (1) and (2)], except that the functions, convolutions, and Fourier transforms are two dimensional.

The sampling functions are now two-dimensional comb functions (sometimes called "bed-of-nails" functions) in which the density of samples may be different along the two axes. Let \( NS_x \) and \( NS_y \) be the sampling densities along the two axes \( x \) and \( y \). We define the sampling rate (in samples/deg\(^2\)) to be \( NS = NS_x \times NS_y \). In the present analysis we consider only rectangular sampling lattices, but completely analogous analyses can be carried out for other regular sampling lattices (e.g., triangular or hexagonal).

Figure 4, a representation of the two-dimensional Fourier domain, is useful for understanding sampling theory in two dimensions. The two diagonal axes labeled \( u \) and \( v \) represent spatial frequency in the \( x \) and \( y \) directions, respectively. The filled circles represent the Fourier transform of a comb function viewed from above the spatial-frequency plane. The small cross shows the position of the comb function (currently positioned at the origin).

## Receptor Sampling

Analogous to Fig. 2A, if the optics of the eye and the receptor size limit the maximum spatial frequency in the \( x \) and \( y \) directions to \( W \) c/deg, then the Fourier transform of the effective retinal intensity distribution will, at most, fill the entire square region in Fig. 4. In this case, the minimum
Fig. 4. Two-dimensional sampling schemes. The diagonal axes labeled \( u \) and \( v \) represent spatial frequency in the diagonal directions. Each square region represents the spatial-frequency range covered by one sampling array in a Nyquist scheme containing 20 arrays. The filled circles show the Fourier transform of the sampling function for the spatial-frequency range covered by the two large crosshatched squares. Also indicated are the densities of sampling units in each sampling array. The total density of sampling units is exactly at the Nyquist rate. The small pictures are examples of possible weighting functions.

The receptor sampling rate needed to encode the retinal image without information loss is obtained when setting \( NS_x = NS_y = 2W \) or \( NS = 4W^2 \). The reader can verify that this is the Nyquist rate in two dimensions by carrying out a pictorial convolution similar to those in Fig. 3 but with a plot like that in Fig. 4. Again we can obtain the convolution by sliding a comb function \( S(\cdot, -\cdot) \) all around the frequency plane. At each position we multiply the comb function by \( Z(\cdot, \cdot) \), get the area under the product, and then plot the result at the current position of the comb function. If \( NS_x = NS_y = 2W \), the replicas fill the entire spatial-frequency plane without overlap (cf. Fig. 2B).

**Postreceptor Sampling: Real-Valued Weighting Functions**

In order to analyze the output of a postreceptor sampling array, we first multiply the Fourier transform of the receptor output by the Fourier transform of the weighting function. When the (real-valued) weighting function corresponds to a low-pass or a bandpass filter, its transform will fit within some pair of rectangular regions placed symmetrically about the origin of the spatial-frequency plane in Fig. 4. (One of the regions contains the complex conjugate of the transform values over the other region.) For example, the shaded square region in Fig. 4 labeled 0 deg and the other shaded region placed symmetrically across the plane show the spatial-frequency region of an ideal bandpass weighting function that is oriented horizontally and has a bandwidth of 1 octave in both the \( x \) and the \( y \) directions. As before, the output of the postreceptor sampling array is given by the convolution (in the Fourier domain) of the sampling function with the product of the weighting function and the receptor output. Therefore the minimum sampling rate required for a sampling array to be locally complete is easily determined. Let \( b_u = u_h - u_l \) be the spatial-frequency range, in the upper half of the spatial-frequency plane, transmitted by the weighting function in the \( x \) direction, and let \( b_v = v_h - v_l \) be the range, in the upper half of the spatial-frequency plane, transmitted in the \( y \) direction. Also, let \( r_u = \max(0, |u_l|) / (u_h - u_l) \) and \( r_v = v_h / (v_h - v_l) \). Then the minimum sampling rate for extracting all the spatial-frequency information is

\[
NS = 2 \times b_u \times b_v \times r/\text{trunc}(r),
\]

where

\[
r = \begin{cases} r_u & \text{if } r_u = \text{trunc}(r_u) \\ r_v & \text{otherwise} \end{cases}
\]

Clearly, the sampling rate for any given spatial-frequency area \( (b_u \times b_v) \) is minimized when \( r \) is a whole number. For example, in Fig. 4, the crosshatched area labeled 0 deg represents a spatial-frequency area of \( W^2/4 \). The region is 1 octave in both the \( x \) and the \( y \) directions \( (r = 1) \); thus the number of samples needed to encode all the information in this spatial-frequency region is \( W^2/2 \). The bed-of-nails function in Fig. 4 shows one of the two possible arrangements of these samples. In particular, it represents a sampling rate in the \( x \) direction of \( W/2 \) \( (NS_x = W/2) \) and in the \( y \) direction of \( W \) \( (NS_y = W) \). The other equivalent sampling function has \( NS_x = W \) and \( NS_y = W/2 \). The reader can demonstrate the efficiency of this sampling scheme by pictorial convolution—the replicas of the crosshatched spatial-frequency region completely cover the entire spatial-frequency plane without overlap.

A two-dimensional sampling scheme is complete in the sense that it encodes all incoming information without loss if (a) each sampling array has at least the sampling rate specified by Eq. (7) and (b) the weighting functions associated with the sampling arrays cover the entire input frequency range. Furthermore, a complete scheme is a Nyquist scheme if (c) the weighting functions do not have overlapping frequency ranges and (d) the weighting functions satisfy Eq. (4) [or equivalently Eq. (3)].

If the total spatial-frequency range to be encoded by the sampling scheme is \( 2W \times W \) (e.g., Fig. 4), then a Nyquist scheme will consist of a set of weighting functions satisfying Eq. (4) that break the whole spatial-frequency range up into nonoverlapping regions. Since the sampling rate needed for each spatial-frequency region is twice the spatial-frequency area [Eq. (7)], the sampling rate of the whole scheme is just the Nyquist limit \( 4W^2 \).

**Postreceptor Sampling: Complex-Valued Weighting Functions**

As in the one-dimensional case, let \( F_k(u, v) \) be the Fourier transform of the real-valued weighting function of an arbitrary low-pass or bandpass filter. The Fourier transform of the complex weighting function that transmits exactly the same information as this real-valued weighting function is given by

\[
H_k(u, v) = [F_k(u, v) + \text{sgn}(v) \times F_k(u, v)]/2.
\]
EXAMPLES OF LOCALLY COMPLETE NYQUIST SCHEMES

Figure 4 illustrates a two-dimensional Nyquist sampling scheme that contains 20 sampling arrays. The highest spatial-frequency range is encoded by six sampling arrays, each composed of units whose real-valued weighting function is that of an oriented, 1-octave bandpass filter. (In the frequency domain these filters are pyramid shaped.) Each of these sampling arrays contains \( W/2 \) sampling units. Note that in Fig. 4 the orientation of the weighting function for each sampling array is indicated inside the frequency region covered by the array. The next two highest spatial-frequency ranges are encoded by similar sets of six sampling arrays whose weighting functions are identical to those covering the highest frequency range, except that their center frequencies have been moved down and their sampling rates have been lowered to \( W/8 \) and \( W/32 \), respectively. The lowest spatial-frequency range is encoded by two sampling arrays, each composed of units whose weighting function is that of an oriented low-pass filter. The two orientations are 0 and 90 deg, and the sampling rate of each array is \( W/32 \). The four photographs in Fig. 4 show the shape and the relative size of the weighting functions in each spatial-frequency range. We show weighting functions whose phase spectrum is constant at 90 deg (edge sensors), but other phase spectra, such as a 0-deg phase shift of every frequency component (bar sensors), would suffice.

Other Nyquist schemes can be obtained by encoding the low-pass region (the four small square regions in the center of Fig. 4) with a single sampling array of nonoriented low-pass receptive fields or, conversely, by breaking the region down into more 1-octave ranges. Similarly, one could encode all but, say, the highest spatial-frequency range in Fig. 4 with oriented or nonoriented low-pass weighting functions. Indeed, the encoding with bandpass weighting functions can be stopped at any level in Fig. 4. All these schemes require the minimum total sampling rate of exactly \( 4W \). However, note that these are the only locally complete Nyquist schemes using low-pass and bandpass weighting functions that are real valued and are 1 octave in both sampling directions. Of course, as implied by Eq. (7), there are other (locally complete) Nyquist schemes consisting of weighting functions that have bandwidths that are narrower than 1 octavé or that have bandwidths that are not equal in both the \( x \) and the \( y \) directions.

More locally complete Nyquist schemes are obtained by letting some of or all the weighting functions be complex valued. For example, every frequency region in Fig. 4 could be sampled by a pair of even- and odd-symmetric receptive fields. The odd-symmetric fields would look like those pictured in Fig. 4; the even-symmetric fields would appear similar but with the center of the receptive field in the center of an excitatory or inhibitory region (bar sensors).

The above set of sampling schemes appears to include all the possible locally complete Nyquist schemes with rectangular sampling lattices. A set of locally complete Nyquist schemes can be similarly derived for other regular sampling lattices.

OTHER NYQUIST SCHEMES

If the weighting functions for the different sampling arrays overlap significantly in the spatial-frequency domain, then a sampling scheme cannot be locally complete and Nyquist at the same time. Thus, strictly speaking, Nyquist schemes can be locally complete only for true low-pass and bandpass weighting functions such as those in Fig. 4. Deriving Nyquist schemes for weighting functions that overlap in the spatial-frequency domain (i.e., Nyquist schemes that are not locally complete) is more difficult.

If the requirement of being locally complete is dropped, the number of possible Nyquist schemes becomes extremely large. One class of Nyquist schemes that is not, in general, locally complete is the so-called Gabor schemes. In the one-dimensional case, a Gabor encoding employs complex-valued weighting functions of the form

\[
h_k(x) = w(x - nx_0)\exp[-ikux],
\]

where \( nx_0 \) and \( ku \) are, respectively, the position and the center frequency of the sampling unit. Each choice for the function \( w(x) \) results in a different sampling scheme. The form of \( w(x) \) is determined by how one wants to encode input signals. An entire sampling scheme is formed by letting \( k \) be \( 0, \pm 1, \pm 2, \ldots \), until \( |ku| \) reaches the highest spatial frequency to be encoded. We can think of each absolute value of \( k \) as representing a sampling array. The position of the sampling units within a sampling array is given by letting \( n \) be \( 0, \pm 1, \pm 2, \ldots \). Note that the phase between \( w(x) \) and the complex sine wave changes as the position of the unit changes (i.e., the symmetry of the weighting functions does not stay the same across the visual field).

In the Gabor schemes the spatial extent of the weighting functions and the number of sampling units is identical for all the sampling arrays [see Eq. (8)]. Thus the Gabor schemes are distinct from the locally complete Nyquist schemes described above.

We refer to all sampling schemes based on Eq. (8) as Gabor schemes. However, Gabor's original paper presented a special case of the above schemes designed to encode any signal in terms of the coefficients of Gaussian-damped sine waves (Gabor functions) of various frequencies and positions that when added together would reproduce the signal. Gaussian-damped sine waves are chosen because they are simultaneously more localized in the space (time).
and frequency domains than any other elementary function. However, it is important to note that Gabor’s scheme cannot be precisely obtained by sampling with weighting functions (receptive fields) that are themselves Gabor functions; that is, by letting \( w(x) \) in Eq. (8) be a Gaussian function. In fact, Gabor’s encoding requires a quite different form of weighting function, one that looks unlike any receptive-field profile ever reported or suggested in the vision literature (see Fig. 2 of Ref. 9).

Marcelja suggests that the spectral overlap between receptive fields in a Gabor scheme constructed with Gabor weighting functions is small enough that one could use Gabor weighting functions without much error. However, no calculations or estimates of the error are given.

The compactness of Gabor functions in the frequency domain also implies that Gabor’s original scheme should be nearly locally complete. This would not be true in an arbitrary Gabor scheme.

**DISCUSSION**

Although many sampling schemes proposed in the vision and image-processing literature are not Nyquist schemes, most are complete and locally complete. A cataloging of some of the sampling schemes that have been used in models of vision and in image processing is given in Table 1.

The cortical-sampling scheme proposed by Sakitt and Barlow is almost a Nyquist scheme. The center frequencies, bandwidths, and numbers of sampling units that they picked are close to those of a locally complete Nyquist scheme; thus there should be relatively little loss of information. Their weighting functions are cosine-phase Gabor functions; thus the scheme should be as close to locally complete as a Gabor scheme using Gabor functions. Sakitt and Barlow’s encoding scheme has sampling units with rather small orientation bandwidths (7.5 deg in the high-frequency units). Other locally complete Nyquist schemes with larger orientation bandwidths would be more consistent with psychophysical and electrophysiological data.

All the other sampling schemes in Table 1 (except the Gabor schemes) are not Nyquist schemes. Indeed, most of them contain at least \( \frac{1}{2} \) more sampling units than would be necessary with the right choice of weighting function and sampling positions. There appear to be two rational reasons for this apparent inefficiency, at least for the sampling schemes that have been developed for image-processing and computer-vision applications. First, if one wants to do image reconstruction by using some quick and simple (but inexact) interpolation functions, the image quality is often improved by increasing the sampling rate. Second, weighting functions are often picked for the ease and the speed with which the sampling-unit responses can be calculated on a serial computer.

Neither of these constraints may be relevant for biological vision. Presumably the cortex does not engage in image reconstruction, at least of the sort that is done in image processing. Also, biological visual processing is largely parallel. A system with many different weighting functions need not be more computationally expensive in processing time or storage (number of units) than having just a few types of weighting function.

**Table 1. Evaluation of Some Sampling Schemes Proposed in the Vision Literature**

<table>
<thead>
<tr>
<th>References</th>
<th>Sampling-Unit Type</th>
<th>Locally Complete</th>
<th>Nyquist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gabor⁴</td>
<td>C</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Kulikowski and Bishop¹⁴</td>
<td>C</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Sakitt and Barlow⁴</td>
<td>R</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Bastiaans¹⁰</td>
<td>C</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Watson⁶</td>
<td>C</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Robson⁵</td>
<td>C</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Burt and Adelson⁷</td>
<td>R</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Wilson and Gelb¹⁵</td>
<td>R</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

⁴ Contains our evaluations of proposed sampling schemes with respect to the terminology introduced in the text. \( R \), real-valued weighting function; \( C \), complex-valued weighting function; \( Y \), best described by the term; \( N \), not best described by this term; ? , insufficient information.

In the human fovea there are almost as many ganglion cells as there are photoreceptors. Most of these cells are reasonably linear, at least for small modulations around some background luminance, and have about the same shape of receptive field. The receptive fields are nonoriented and broadband enough that the ganglion-cell array should transmit most of the information received from the photoreceptor array.¹⁸

There are probably around 200 times as many neurons in area 17 of the primate visual cortex as there are ganglion cells, and around half of these are in layer IV where the axons from the lateral geniculate nucleus (LGN) project.¹⁶ Many of the units in layer IV are relatively linear simple cells. Thus there could conceivably be up to 50 complete representations of the retinal image in layer IV alone.

If the outputs from stages along the visual pathway up through layer IV are just approximate linear recordings of the retinal image, then two obvious questions arise. Why carry out the recording in the first place, and why are there so many more units in layer IV than earlier in the visual pathway?

There seem to be three possible reasons for linear recording in an information-transmission system. One reason is to increase the accuracy of information transmission. For example, transmission of signals through a noisy environment can be improved by recoding the signal onto a carrier frequency outside the bandwidth of the noise. However, it is not clear how the recoding that occurs in the visual pathway might increase signal-to-noise ratio in the transmission process.

Another possible reason for recoding is to attenuate selectively certain types of information that are either unnecessary or interfering. Ganglion-cell receptive fields may be designed to attenuate some of the relatively unimportant information about mean luminance in order to keep important pattern information within the limited dynamic range of the spike-generating neurons.

A third possibility is that recording is used to get the information into a form or a representation that is tailored for some processor at the receiving end of the transmission line. The encoding produced by the retinal and LGN weighting functions may make it easier to form the weighting functions of the cortical cells (e.g., consider the simple wiring schemes suggested by Hubel and Wiesel).¹⁹ The bandpass and orien-
tation-selective weighting functions of simple cells may, in turn, represent the image information in a form useful for computing the positions and orientations of edges, object boundaries, relative distance from stereopsis, etc.

This third reason for recoding may help explain why there are so many cells in layer IV. It is quite possible that there are many different visual processors or modules, each requiring a complete representation of the retinal image but each wanting the information represented in a different form. An edge-locating processor may want a complete representation encoded by odd-symmetric weighting functions, but another processor may want a complete representation encoded by odd-even pairs of weighting functions. Some processors may need only the output of a small subset of locally complete arrays.

The range of complete sampling schemes developed here underscores the difficulty of using psychophysical methods to determine what kind of sampling scheme(s) the human visual system is using. The large number of neurons in layer IV suggests that the cortical processing involved in different psychophysical tasks may be based on different sampling schemes (different complete sets of spatial-frequency channels). Thus evidence for all sorts of apparently conflicting schemes might be found.

It seems likely that an analysis of the sampling schemes in the visual cortex will require consideration of how their outputs are likely to be used. Simple information-summing mechanisms, such as those used in probability-summation models, are unlikely candidates as cortical processors. Much more likely are processors for grouping image data into edges, surfaces, and regions, computing the orientations and the distances of surfaces, and carrying out other sorts of basic image processing. Also with high likelihood are cortical processors, which extract perceptual invariance from superficially dissimilar images of a given physical object. Theories of how to compute these things (in parallel processing machines), given the constraints of imaging, neural computing, and the natural environment, may help reveal both the logic and the appropriate experimental tests of various cortical-sampling schemes.

ACKNOWLEDGMENTS

We thank Jack Yellot for helping us to get started on this project over a cocktail at the annual meeting of the Association for Research in Vision and Ophthalmology 1984. This research was supported by the National Institutes of Health under grant EY20668 to Wilson S. Geisler.

REFERENCES