Rate Coding Versus Temporal Order Coding: What the Retinal Ganglion Cells Tell the Visual Cortex

2 authors, including:

Simon Jonathan Thorpe
Université Toulouse III - Paul Sabatier
229 PUBLICATIONS   **20,876** CITATIONS

Rate Coding Versus Temporal Order Coding: What the Retinal Ganglion Cells Tell the Visual Cortex

Rufin Van Rullen
Simon J. Thorpe
Centre de Recherche Cerveau et Cognition, Faculté de Médecine Rangueil, 31062 Toulouse Cedex, France

It is often supposed that the messages sent to the visual cortex by the retinal ganglion cells are encoded by the mean firing rates observed on spike trains generated with a Poisson process. Using an information transmission approach, we evaluate the performances of two such codes, one based on the spike count and the other on the mean interspike interval, and compare the results with a rank order code, where the first ganglion cells to emit a spike are given a maximal weight. Our results show that the rate codes are far from optimal for fast information transmission and that the temporal structure of the spike train can be efficiently used to maximize the information transfer rate under conditions where each cell needs to fire only one spike.

1 Introduction

How do neurons transmit information? This question is a central problem in the field of neuroscience (Perkel & Bullock, 1968). Signals can be conveyed by analog and electrical mechanisms locally, but over distances information has to be encoded in the spatiotemporal pattern of trains of action potentials generated by a population of neurons. The exact features of these spike trains that carry information between neurons need to be defined. The most commonly used code is one based on the firing rates of individual cells, but this is by no means the only option. In recent years a strong debate has opposed partisans of codes embedded in the neurons' mean firing rates and researchers in favor of temporal codes, where the precise temporal structure of the spike train is taken into account (Softky, 1995; Shadlen & Newsome, 1995, 1998; Gautrais & Thorpe, 1998). Here we address this question of neural coding in the context of information transmission between the retina and the visual cortex.

The retina is a particularly interesting place to study neural information processing (Meister & Berry, 1999). First, it is relatively easy to stimulate and record retinal cells. Furthermore, the general architecture and functional organization of the retina are remarkably well known (Rodieck, 1998). There is probably no other place in the visual system where one can define more

rigorously what information needs to be represented, how many neurons are available to do it, and how long the transmission should last.

A widely used simplification states that the information transmitted from the retina to the brain codes the intensity of the visual stimulus at every location in the visual field. Although this strong statement can certainly be discussed, it is clear that the aim of retinal coding is to transmit enough information about the image on the retina to allow objects and events to be identified.

One can also consider that the different types of ganglion cells "tile" the entire retina, so that there is little or no redundancy in the number of neurons, which should therefore be kept to an absolute minimum. In monkeys, in particular, the number of ganglion cells is roughly 1 million. However, the limited redundancy in the number of neurons encoding a given stimulus feature at a particular location does not mean that there is no overlap between ganglion cells' receptive fields. In fact, in the cat retina, between 7 and 20 ganglion cells have receptive field centers that share a given common position in the visual field (Peichl & Wässle, 1979; Fischer, 1973). Correlations among the responses of neighboring ganglion cells also demonstrate that they do not operate as independent channels (Arnett & Spraker, 1981; Mastronarde, 1989; Meister, Lagnado, & Baylor, 1995). Nevertheless, it is not clear in the literature how much of that correlation in the output firing pattern can be explained by shared common inputs (from photoreceptors, bipolar, horizontal, or amacrine cells) to the ganglion cells (Brivanlou, Warland, & Meisler, 1998; DeVries, 1999; Vardi & Smith, 1996).

Finally, data on the speed of visual processing provide severe limitations to the time available for information transmission through the visual system. First, recorded neuronal latencies can be extremely short: responses start at around 20 ms in the retina (Sestokas, Lehmkuhle, & Kratz, 1987; Buser & Imbert, 1992) and approximately 10 ms later in the lateral geniculate nucleus (LGN) (Sestokas et al., 1987); the earliest responses in V1 already exhibit selectivity to stimulus orientation around 40 ms poststimulus (Celebrini, Thorpe, Trotter, & Imbert, 1993; Nowak, Munk, Girard, & Bullier, 1995). In inferotemporal cortex (IT), face-selective responses begin between 80 and 100 ms after stimulus presentation (Perrett, Rolls, & Caan, 1982) and show selectivity to face orientation even at the very start of the response (Oram & Perrett, 1992). Taken together, these data indicate that visual processing should rely on very short transmission times, on the order of 10 to 20 ms between two consecutive processing stages and less than 50 ms between the retina and the cortex. The same conclusions can be derived from psychophysical observations on monkeys and humans, in a task where subjects must decide whether a briefly flashed photograph of a natural scene contains a target category such as an animal, food, or a means of transport. Monkeys can respond as early as 160 ms after stimulus presentation, and human subjects around 220 ms (Thorpe, Fize, & Marlot, 1996; Fabre-Thorpe, Richard, & Thorpe, 1998; VanRullen & Thorpe, 1999).

Given the large number of synaptic stages involved, it appears here again that information processing and transfer should not last more than about 10 ms at each processing stage, and probably less than 50 ms between the retina and the brain (the delay of transduction in the photoreceptors has to be taken into account). Therefore, further computation should rely on very few spikes per ganglion cell.

As Meister and Berry (1999) argue, computations that use a very restricted number of spikes are difficult to conciliate with the common view stating that retinal encoding uses the firing frequencies of individual ganglion cells. Classically, ganglion cells are thought to encode their inputs in their output firing frequency (Warland, Reinagel, & Meister, 1997), and the process of retinal spike train generation is supposed to be stochastic, that is, subject to a Poisson or pseudo-Poisson noise. As an alternative, Meister and Berry (1999) review a number of arguments for taking into account the temporal information that can be derived from the very first spikes in the retinal spike trains. For example, this information could be represented by synchronous firing among neurons.

Here we introduce another temporal coding scheme, based on the order of firing over a population of ganglion cells. One can consider the ganglion cells as analog-to-delay converters; the most strongly activated ones will tend to fire first, whereas more weakly activated cells will fire later or not at all. Under such conditions, the relative timing in which the ganglion cells fire the first spike of their spike train can be used as a code (Thorpe, 1990). A more specific version of this hypothesis uses only the order of firing across a population of cells (Thorpe & Gautrais, 1997). This coding scheme has already been proposed to account for the speed of processing in the visual system (Thorpe & Gautrais, 1997, 1998). We have also shown (VanRullen, Gautrais, Delorme, & Thorpe, 1998) that it can be successfully applied to a computationally difficult task, such as detecting faces in natural images.

We will compare the performances of this code with two classical implementations of rate coding: one that relies on the spike count and the other on the mean interspike interval to estimate the ganglion cells' firing frequencies over a Poisson spike train.

The approach that we use tests the efficiency of the different coding schemes for reconstructing the input image on the basis of the spike trains generated by the ganglion cells (Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1997). More precisely, we suppose (for simplicity) that a natural input image is briefly presented to the retina, preceded and followed by a dark uniform field. Previous experiments (e.g., Thorpe et al., 1996) suggest that under these conditions, information should be available to the visual cortex as early as 50 ms poststimulus. We take the position of an imaginary observer "listening" to the pattern of spikes coming up the optic nerve and trying to derive information about the input image. Of course, we do not consider the role of the visual system in general as being to reconstruct the
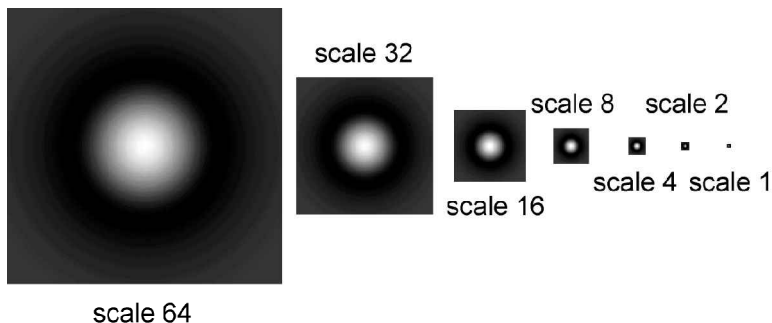
scale 32

scale 8   scale 2

scale 16

scale 4 scale 1

scale 64

Figure 1: Difference of gaussians: ON-center cells' receptive fields. The scale 128 filter is not represented, although it was used in our simulations.

image in the brain. Rather, this reconstruction should be seen as a form of benchmark—a test of the potential of a particular code.

First, we introduce a simple model of the architecture of the retina and its functional organization, independent of the way information will be represented. Then we describe our rank order coding scheme, and show how it can be applied to retinal coding. Finally, we compare both a "noise-free" and a "noisy" version of this code with two rate-based coding models (one in which the information is embedded in the spike count, the other in the mean interspike interval), by estimating the quality of the input image reconstruction that they provide as a function of time.

2 Retinal Model

2.1 Wavelet-like Transform. We designed a model retina. Our model ganglion cells compute the local contrast intensities at different spatial scales and for two different polarities: ON- and OFF-center cells. We can consider this decomposition as a wavelet-like transform, using differences of gaussians (DoG) as the basic filters (Rodieck, 1965). The spatiotemporal properties of our model ganglion cells match those of X-type cells: they use linear spatial summation between the center and surround regions of the receptive field, and there is no temporal component in the input-output function (Buser & Imbert, 1992). The ganglion cells' receptive fields are shown in Figure 1.

We used the simple DoG described by Field (1994), where the surround has three times the width of the center. An OFF-center filter was simply an inverted version of the ON-center receptive field. The narrowest filters (at scale 1) were 5×5 pixels in size, and the widest 767×767 pixels (at scale 128). Furthermore, these filters were normalized so that when the input pattern is identical to the filter itself, the result of the convolution at this given scale

should be 1. The result of the application of these filters at any position and scale is the output of the wavelet-like transform, which produces a set of analog values, corresponding to the activation levels of our model ganglion cells. According to wavelet theory (Mallat, 1989), the wavelet-like reconstruction will simply be obtained by applying on the reconstructed image, for each scale and position, the corresponding ganglion cell's receptive field, multiplied by the corresponding activation value.

More precisely, the contrast at a particular position (x, y) and scale (s) is defined as:

$$Contrast_{\mathrm{Im}}(x, y, s) = \sum_i \sum_j (\mathrm{Im}(i + x, j + y) \cdot DoG_s(i, j))$$

where DoG_s denotes the DoG filter at scale s and (i, j) spans the width and height of the DoG_s filter. Using these contrast values, the reconstruction $\mathrm{Im}_{\mathrm{Rec}}$ of the image Im is obtained by

$$\mathrm{Im}_{\mathrm{Rec}}(i, j) = \sum_x \sum_y \sum_s Contrast_{\mathrm{Im}}(x, y, s) \cdot DoG_s(x - i, y - j),$$

where s spans the range of spatial scales and (x, y) spans the image width and height.

2.2 Subsampling. For computational reasons, as well as for biological plausibility, the spatial resolution of the transform varies together with the scale of the filters, so that when the scale is doubled, the resolution is divided by 2. More precisely, the narrowest convolutions (at scale 1) are computed for every pixel in the original image, whereas the filters at scale 2 are applied once every 2 pixels horizontally and vertically. Therefore, the number of neurons per image is no more than 8/3 times the number of pixels in the original image.

Let n be the number of pixels in the input image; the number of ganglion cells is then

$$2 \cdot (n + n/4 + n/16 + n/64 + \cdots + n/16, 384).$$

This organization scheme is detailed in Figure 2.

All natural images that will be used in the following simulations are 364×244 pixels in size, and the number of ganglion cells will then be approximately 236,000.

Of course, we do not claim that this precise architecture is biologically realistic. The real organization of ganglion cells in the mammalian retina has numerous differences from our model. First, the actual recorded receptive field sizes do not span as many octaves as our model receptive fields do. The ratio between the biggest and the smallest receptive field sizes across
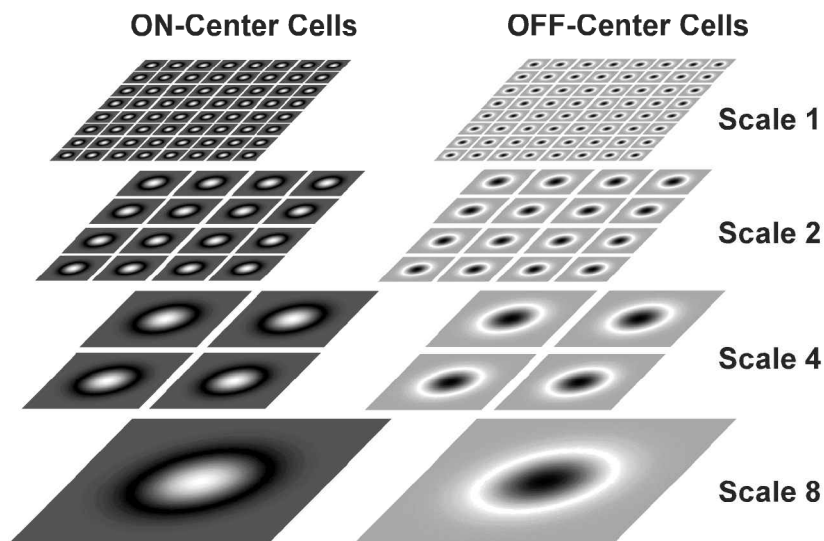
Figure 2: Retinal organization. The image is encoded through a bank of filter maps with two different polarities (ON- and OFF-center cells) and eight different scales (only four are shown here). The sampling resolution is inversely proportional to the scale.

the entire retina is less than 100, and only around 2 at a given eccentricity (Croner & Kaplan, 1995). However, one could argue that even if there were cells with very large receptive fields, they would be very rare and therefore difficult to record. Another notable difference is that in biological visual systems, the ganglion cells at each spatial scale are not equally distributed over the retina.

On the other hand, the model we used allows the information in an image to be fully encoded with a wavelet-like transform, and therefore we are able to address the real issue that concerns us here: What would be the most efficient way of transmitting this information to the brain using spiking neurons? Given the architecture, how do these neurons convert an analog intensity value, representing the local contrast in their receptive field, into a succession of firing events, the spike train? What is the optimal way of doing so in terms of the maximization of information transmission?

In the following sections, we consider a variety of coding schemes that could be used to transmit information about the image to the brain. First, we describe a method that uses an analog-to-delay mechanism coupled with a coding scheme in which the order of firing in the retinal ganglion cells is used to code the information in the image (Thorpe, 1990; Thorpe & Gautrais, 1998). Later we compare the efficiency of this code with more conventional

rate-based codes using either counts of the total number of spikes produced by each cell or measures of the interspike interval.

3 Rank Order Coding

The result of the convolution computed by our model ganglion cells on the original image is an analog value, which can be thought of as the neuron's membrane potential. It has to be converted into a spike train that will be transmitted to the cortex via the optic nerve.

3.1 Analog-to-Rank Conversion. Since we want to show that the relative order in which the ganglion cells will generate their first spike can be used as a code, we can simply assign a rank to each neuron as a function of its activation. The most activated will fire first, and so on. This is supported by the characteristics of integrate-and-fire neurons: the higher the membrane potential is, the sooner the threshold will be reached, and the sooner a spike will be emitted.

We do not need to model the absolute or relative timing of the spike precisely, since the only relevant variable that will be used for decoding is the neuron's rank. However, it is possible to assign a latency to each neuron, and we will implement such a function in section 4.1 in order to compare our model with rate coding schemes.

3.2 Rank Order Decoding Using Image Statistics: Contrast = f(Rank). Now that the mechanism responsible for order encoding is described, we need to evaluate the quality of that code. A good and simple way of estimating the information about the visual stimulus that is carried by the spikes along the optic nerve is to use these spikes to reconstruct the input stimulus.

With the kind of wavelet-like transform that is computed by our model ganglion cells, the reconstruction of the image is simply obtained by once again applying the DoG filters on the result of the previous convolution. Each ganglion cell's receptive field needs to be added to the reconstruction image, at the position corresponding to its center, with a multiplicative value equal to the result of the previous convolution—this neuron's activation level. The problem with our rank order code is that this result has been "forgotten" through the analog-to-order conversion. If the output of the ganglion cells were simple analog values, there would not be any such problem, but the only information available in our case is the relative order in which the ganglion cells fired. If neuron i fired first, what does it mean in terms of the contrast intensities in the original image? If neuron j fired right after neuron k, how relevant is the information transmitted by neuron j compared to that conveyed by neuron k? A simple way of answering these questions is to associate each possible order with the average contrast value that drove the corresponding neuron above threshold. This is roughly equivalent to a sort of reverse correlation analysis.

We presented our model ganglion cells with more than 3000 natural images (364×244 pixels), sorted the obtained contrast values for all scales and polarities in decreasing order for each image, and then computed the average contrast value obtained at the location of the first neuron to generate a spike. This location, of course, varied with the different images. The same was done for all second spikes in the images, and so on, until the last neuron that fired. This procedure is described by the following equation. The average maximal contrast *MaxC* is defined by:

$$MaxC = \frac{1}{card\{I\}} \sum_{Im \in \{I\}} \max_{(x,y,s)} \left\{ \left| \sum_i \sum_j Im(x+i, y+j) \cdot DoG_s(i,j) \right| \right\}$$

where $\{I\}$ denotes the ensemble of natural images, of cardinal ($card\{I\}$), Im spans the ensemble $\{I\}$, s spans the range of spatial scales, (x, y) spans the image width and height, and (i, j) spans the width and height of the DoG_s filter. Note that because we use a subsample when the scale is greater than 1, the range of possible coordinates (x, y) depends on the scale s. For simplicity, this does not appear in the above equation or in the following ones.

The average contrast at rank r is obtained with the same procedure, replacing the function max() (which returns the value of rank 1) by the function rank$_r$() (returning the value of rank r).

We obtained a list of results that we can consider as a particular kind of look-up table (LUT) that allows looking up the most likely contrast value for a spike with a given rank. The LUT is plotted in Figure 3. Note that the absolute contrast values obtained here depend on the normalization that we applied on the DoG filters, as well as the intensity levels in the input images, which were in the range 0 to 255 in our simulations. Therefore, these values have been normalized: 100% denotes the average maximal contrast, normalized over the whole set of 3000 images. Furthermore, no preprocessing was applied on the set of input images; different images might thus span a very different range of intensity and contrast levels. Consequently, the variance of contrast values obtained over the image set was relatively high, as shown in Figure 3. Note that a high variance is not optimal for rank order decoding, because it means that the contrast value attributed to a ganglion cell firing with a certain rank can be strongly over- or underestimated.

3.3 Qualitative Results: Some Examples of Image Reconstruction. Now that we can estimate which contrast value corresponds to a spike arriving with a given rank, it is possible to reconstruct the input image.

The reconstruction is empty at the beginning of the process; that is, it is initialized with a gray level of 128. Each time a spike is received from a ganglion cell, the DoG filter of the corresponding scale and polarity is added to the reconstruction, at the corresponding position, and with the multiplicative value corresponding to its rank (the equations describing
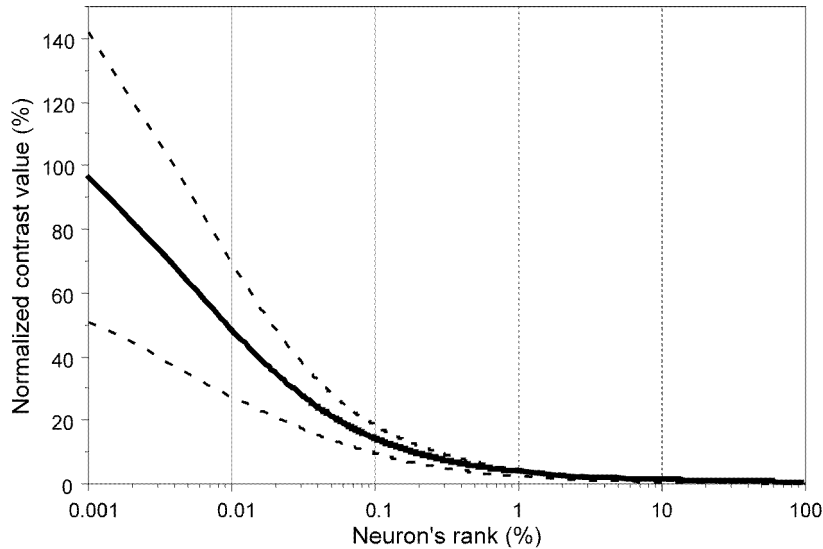
Figure 3: Mean contrast values as a function of the ganglion cells' firing rank (as a percentage of the total number of neurons). *X*-axis in log scale. *Y*-axis as a percentage of the maximum contrast value, averaged and normalized over more than 3000 images. Standard deviation is plotted using dashed lines.

the reconstruction process with an order code are detailed in section 4.2). Therefore, we can stop the spike propagation mechanism at any time and consider how much information has been received when a given percentage of the ganglion cells has fired.

Some examples of reconstructed images are presented in Figure 4, together with the proportion of neurons that have fired. This percentage cannot be greater than 50%. Since a contrast at a given position and scale in the input image is either positive or negative, a pair of ON- and OFF-center ganglion cells coding for the same scale and position cannot both fire. From Figure 4 it appears that even when less than 1% of the retinal ganglion cells have fired one spike, the identity of objects in the image is often clear.

4 Rate Coding Models

The code described above seems efficient for rapid information transmission between the retina and the brain. Now we would like to compare it with other classical codes based on neuronal firing rates. Thus, we adapted our model to make it suitable for that kind of code.

The architecture of the retina is the same as before, but the neurons are allowed to generate more than one spike, and the whole spike train can
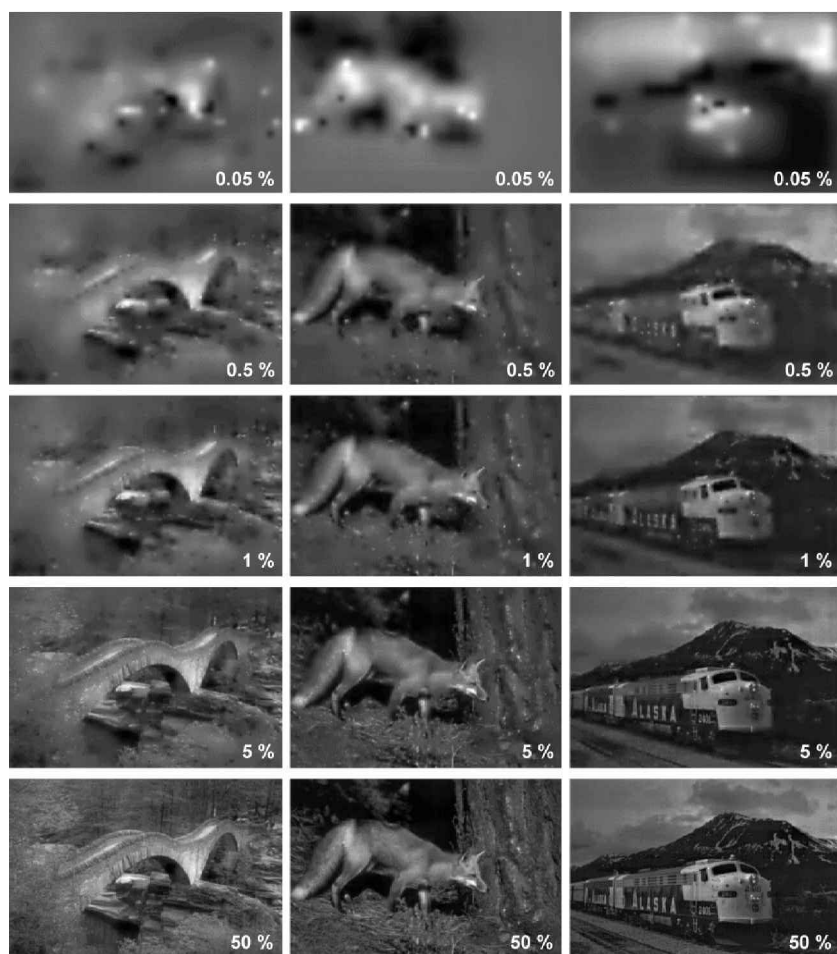
Figure 4: Examples of image reconstruction at different steps. The percentage of neurons that have generated a single spike is indicated for each image.

then be used for rate decoding. The mechanisms of spike train generation are discussed below.

4.1 Poisson Spike Train Generation. In the rank coding scheme, spikes were propagated with a rank that depends on the activation level of the neurons that generated them. There was no need to specify the precise timing of each spike, since only the relative order was relevant. Now if we assign a delay to the first spikes generated by the ganglion cells, used by the order coding (see section 3.1), as well as for the rest of the spikes in the spike

train, used for rate decoding, we will be able to compare the performances obtained with order coding and rate coding under comparable conditions. Note that in the case of our rank order coding scheme, any negative strongly monotonic (decreasing) function would have the same effect. The stronger the input is, the shorter the latency is. The only thing that would vary would be the amount of time needed to encode a given signal.

The problem here is to choose a function that would transform the contrast intensity in a neuron's receptive field into a firing latency or a firing frequency. The properties of real neurons can help us decide what function should be chosen.

Croner and Kaplan (1995) proposed that the response of ganglion cells in the primate retina to a given contrast can be described by a Michaelis-Menten function. This kind of function has also been applied to model the response of various sensory neurons (Lipetz, 1971), including horizontal cells in the fish retina (Naka & Rushton, 1966).

The Michaelis-Menten function can be written as

$$R(C) = a \cdot C/(b + C),$$

where R is the response (firing rate, Hz) to a given contrast C, a represents the maximum response, and b is the contrast at which the response is $a/2$. Therefore, if we consider the mean interspike interval I to be inversely proportional to the neuron's firing rate, we have

$$I(C) = 1/R(C)$$
$$= 1/a + (b/a) \cdot (1/C).$$

Here b/a can be thought of as the inverse of the neuron's contrast gain G, and $1/a$ as the neuron's refractory period Ref. Thus, with these new notations, the mean interspike interval and the ganglion cell's firing rate can be rewritten as

$$I(C) = Ref + 1/(G \cdot C)$$
$$R(C) = G \cdot C/(1 + Ref \cdot G \cdot C).$$

The neuron's firing rate R appears to be the combination of a "free" firing rate $G \cdot CW$ proportional to the contrast in the neuron's receptive field, and a refractory period Ref. The concept of "free" firing rate was introduced by Berry and Meister (1998). It represents the rate that would be produced if there was no refractory period. However, this refractory period has been found to play a major role in the ganglion cell's spike generation mechanism (Lankheet, Molenaar, & van de Grind, 1989), and is also thought to be responsible for the reproducibility and precision of the timing of firing events in retinal spike trains (Berry, Warland, & Meister, 1997; Berry & Meister, 1998), under conditions where interspike intervals are relatively short.

The values of the parameters *Ref* and *G* must now be chosen. As an example, Croner and Kaplan (1995) give for a "typical" ganglion cells the approximate values 200 for *G* (in Hz . contrast^{-1}) and 10 ms for *Ref*. The refractory period *Ref* acts as a constraint on the observed firing rate *R*, so that *R* cannot be greater than $1/Ref$. Therefore, lower values of *Ref* will benefit a count code, because they will allow the neuron to fire at a higher maximum rate. Most ganglion cells in the retina have a firing rate inferior to 100 Hz; thus, a refractory period $Ref = 0.005$ s, restraining the range of possible firing rates to 0 to 200 Hz seems a reasonable value for both the order and count codes. The parameter *G*, the contrast gain of the ganglion cell, will determine how the "free" firing rate varies with the input contrast. Let $MaxC = 1$ (i.e., 100%) denote the maximal contrast that can fall into a ganglion cell's receptive field. As an example, if *MaxF*, the maximal "free" firing rate, is chosen equal to 2000 Hz, then $G = MaxF/MaxC = 2000$ (using the previous values) implies that all contrast values above 50% will give a free firing rate $G \cdot C$ greater than 1000 Hz (i.e. $MaxF/2$), that is, an observed firing rate *R* greater than 166 Hz when $Ref = 5$ ms (in our case, according to the LUT obtained previously, this should not be more than 0.01% of the neurons). This maximum value of the free firing rate (2000 Hz) is in the same range as the one found by Berry and Meister (1998). In our simulations, we will therefore use a contrast gain $G = 2000$ and a refractory period $Ref = 5$ ms. The Michaelis-Menten function transforming the input contrast into a firing rate with the above parameter values is plotted in Figure 5, together with the estimated firing rate distribution derived from the distribution of contrast intensities obtained in section 3.2. The great majority of neurons appear to have low theoretical firing rates. However, this can be an advantage in two ways. First, the system acts as if it was minimizing the neurons' firing rates and thus the energetic cost for a given amount of transmitted information (Baddeley et al., 1997). Second, most of the neurons are computing in the semilinear part of their regime, very similar to what would be expected for an optimal wavelet reconstruction.

Finally, the spike train is obtained by applying a Poisson law on the theoretical firing rate *R* described above. This Poisson process is a probabilistic mechanism known to induce noise on the precise timing of the firing events, but to keep the observed firing rate identical when considered over relatively large time windows.

Note that a consequence of the application of a Poisson law is that only the ganglion cells with strictly positive firing rates will be allowed to fire. Thus, there will not be any spontaneous activity in the model. Moreover, when an ON-center ganglion cell is activated, the OFF-center cell coding for the same scale and position cannot fire. These are limitations of the stochastic process of spike train generation that we use, which are clearly different from what would be expected in real systems.
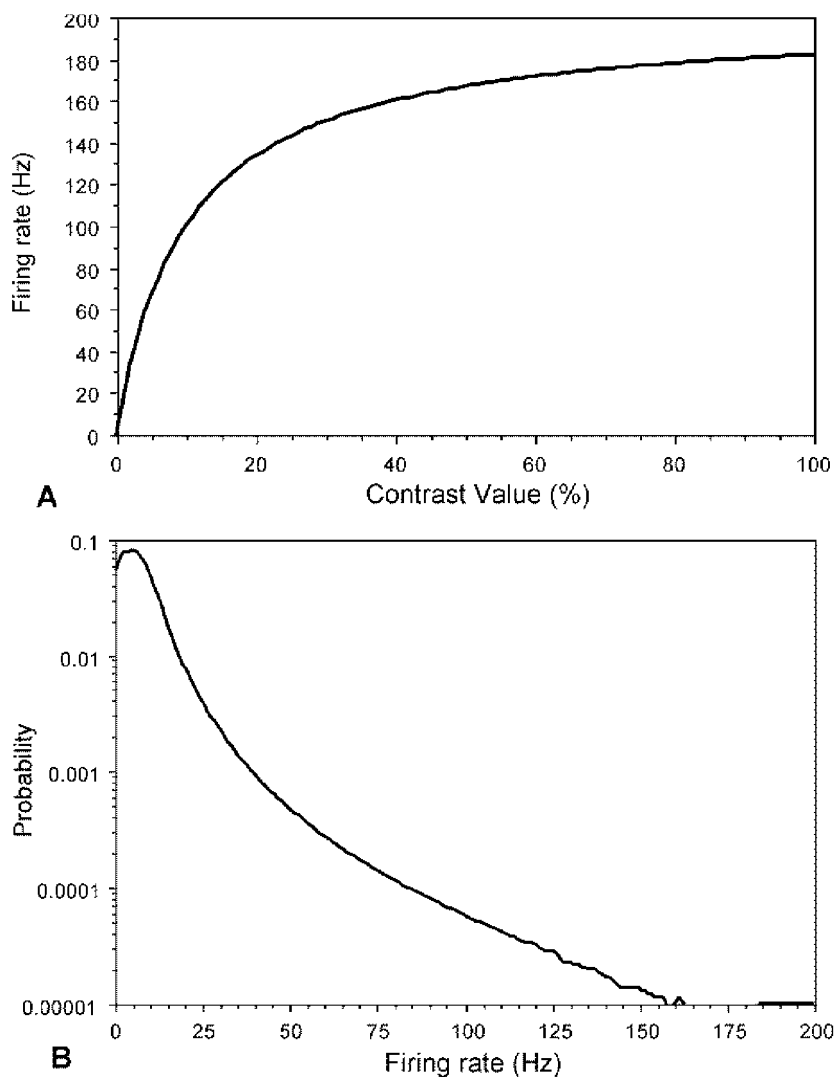
Figure 5: (A) Theoretical firing rate as a function of the input contrast (in percentage, notation as in Figure 3). Michaelis-Menten function, parameters $G = 2000$, $Ref = 0.005$ s (see text for details). (B) Distribution of firing rates as derived from the distribution of contrast levels in Figure 3, calculated over more than 3000 natural images (y-axis in log scale).

4.2 Different Rate Codes. We will test the performances of two different codes that try to use this spike train in order to estimate mean firing rate. The first is a count code that assigns to each ganglion cell a value proportional to the number of spikes it has generated. Another code will estimate the mean interspike interval (ISI) of the spike train and assign to a given ganglion cell a firing rate inversely proportional to its mean ISI. This kind of code requires at least two spikes in a spike train to induce nonzero estimated rates. We will call this second code an ISI code. Clearly, ISI coding is potentially more accurate than the other tested codes, because the interval between two spikes is an analog value that can be quickly transmitted and arbitrarily precise.

To keep comparisons with the order coding model as simple as possible, we will let the delay of the first spike emitted by a ganglion cell be the mean ISI I of the spike train, that is, the inverse of the neuron's firing rate R.

For each code, the decoded value can be interpreted as an estimation of the contrast in a given neuron's receptive field. Thus, we can define, for a particular neuron n (described by its position and scale (x_n, y_n, s_n)), the estimated contrast value decoded at time t by each of the considered codes.

For the order code, we have:

$$\begin{aligned} Contrast_{\text{Order}}(n, t) &= 0 && \text{if } t < \text{Latency }(n) \\ &= \text{sign}(n) \cdot \text{LUT}(\text{rank}(n)) && \text{if } t \geq \text{Latency }(n) \end{aligned}$$

where sign (n) is the sign of neuron n, -1 if n is an OFF-center cell, $+1$ if n is an ON-center cell; rank (n) is the rank of discharge of neuron n; and LUT (r) is the average contrast value corresponding to rank r (see section 3.2).

The estimated contrast value of neuron n at time t with a count code is defined as:

$$Contrast_{\text{Count}}(n, t) = \text{sign}(n) \cdot \text{count}(n, t),$$

where count (n, t) is the number of spikes emitted by neuron n at time t.

Finally, the estimated contrast value derived by an ISI code will be:

$$\begin{aligned} Contrast_{ISI}(n, t) &= 0 && \text{if count}(n, t) < 2 \\ &= \frac{\text{sign}(n) \cdot (\text{count}(n, t) - 1)}{t_{\text{count}(n,t)} - t_1} && \text{if count}(n, t) \geq 2 \end{aligned}$$

where t_k denotes the time at which neuron n fires its kth spike

In the previous equations, the obtained contrast values can be negative. This is only a simplification of the fact that neurons in our model can be of two different types: ON- and OFF- center.

The reconstructed image at time t, with a particular code, can now be defined as:

$$\text{Im}_{\text{Rec}}(i, j, t) = \sum_n Contrast_{\text{Code}}(n, t) \cdot DoG_{s_n}(x_n - i, y_n - j)$$

where s spans the range of spatial scales, and n spans the whole set of neurons. This equation is in fact derived from the one introduced in section 2.1. However, here the contrast value applied to a filter at a given position and scale is not the exact value extracted by convolution, but rather an estimation of that value according to a particular code.

Now we can compare the performances of an order code, a count code, and an ISI code under the same conditions. For rank order decoding and image reconstruction, we will simply use the process described above: the reconstructed image is empty at the beginning of the process; when the current time-step corresponds to a neuron's latency, the filter of the corresponding scale and polarity is added to the reconstruction, at the corresponding position and with a multiplicative factor corresponding to the neuron's rank. This multiplicative factor is derived from the LUT obtained previously (see section 3.2).

The spike train decoding and the image reconstruction for the count code can be simply obtained: the reconstructed image is empty at the beginning of the process, and every time a spike is received, whatever its position in the spike train, the filter of the corresponding scale and polarity is added to the reconstruction, at the corresponding position. There is no need for a multiplicative factor, since the neuron's activity is supposed to be encoded in the number of spikes emitted (the neuron's observed firing rate R) rather than any other variable.

For the ISI code, the estimated firing rate of a ganglion cell is recomputed every time the cell generates a spike, starting with the second spike. The change in the estimated firing rate can update the image reconstruction by adding or subtracting the corresponding filter at the corresponding position with a multiplicative value equal to this change. Here, as for the count code, the analog variable used for reconstruction is the ganglion cell firing rate. One should keep in mind that this rate does not depend linearly on the local contrast but is transformed through a nonlinear Michaelis-Menten function.

Note that for the three different codes, the reconstructed images that are used for result analysis are normalized, so that they all span the whole range of gray levels (here between 0 and 255). Thus, there is no need to define precisely the absolute value of the maximum contrast applied to the ganglion cell with the shortest latency (for an order code), the highest count (for a count code), or the smallest interspike interval (for an ISI code). Only the relative values attributed to different ganglion cells will have an actual consequence on the reconstruction.

5 Comparison

5.1 Performance Estimation: Mutual Information and Mean Square Error. To evaluate the performances obtained with the different models with the different spike train types, we have to define a measure of the quality of the image reconstructions obtained with the different codes.

A widely used measure in image processing is the simple mean square error (MSE), defined by

$$\text{MSE} = \frac{1}{\text{width} \cdot \text{height}} \sum_i \sum_j (\text{Im}_1(i, j) - \text{Im}_2(i, j))^2,$$

where Im_1 and Im_2 are the two grayscale images to compare and (i, j) spans the image width and height. However, a number of problems are associated with this measure. It takes into account only the difference in intensity levels between two images. For instance, an image obtained by dividing by a factor of 2 all intensity levels in the original image will often yield a much higher MSE than a "flat" image obtained by assigning to every pixel the mean intensity level of the original image. Yet the former contains all the information present in the original image, whereas the latter contains virtually no information at all.

To overcome these limitations, one can also use the mutual information measure (Rieke et al., 1997). We assume that the original and the reconstructed images are two random variables X and Y (respectively) of distributions P(X) and P(Y). The mutual information I is the average information that observations of Y provide about X (and conversely):

$$I = \int \int P[X, Y] \log_2 \left(\frac{P[X, Y]}{P[X]P[Y]} \right) [dX][dY].$$

The mutual information (MI) between two images can be thought of as the average information (in bits) that a pixel of one of the images provides about the corresponding pixel in the other. For example, the self MI (MI between an image and itself) of an 8-bit gray level image is 8 if and only if the image actually contains 256 different gray levels and they are equally distributed. The MI is a more accurate measure of the relative information between two images than the MSE, except that it does not account for differences in gray level. At one extreme, the reconstructed image could be the exact negative of the original, and the MI would be maximal at the same time.

Therefore, a good coding scheme should lead to a good reconstruction in terms of both MI and MSE.

5.2 Results. Fifty images were randomly chosen among the 3000 natural images used previously (see section 3.2), and presented to our model retina. The contrasts at different scales and polarities were computed using the DoG filters.

The first step was to estimate the theoretical limit performance of the model retina that we used. For that purpose, we reconstructed the 50 images on the basis of the exact analog contrast values obtained above. For each position in the visual field and each scale (one should keep in mind that the visual field is subsampled when the scale increases), the corresponding filter

(ON-center if positive contrast, OFF-center if negative) was multiplied by the absolute analog contrast value obtained for that neuron and applied on the reconstruction. This algorithm is directly derived from the wavelet theory (see the equations in section 2.1). On the basis of these reconstructions, we were able to calculate the MSE and MI theoretical limits of the model, with the same image set that will be used for estimating the performance of the different codes. Then the contrast values were used to compute the firing rate R and the first spike latency I of all ganglion cells in the retina. Time was divided into 1 ms time steps, and for each ganglion cell, a Poisson spike train was generated. A cell could emit one or more spikes at any given time step, with a probability depending on its theoretical firing rate with a Poisson law.

For the order code, every cell generated its first and only spike during the time step corresponding to its latency I. To make the comparison with (noisy) rate codes as fair as possible, a second version of this order-based reconstruction procedure was applied using latencies subject to a gaussian stochastic distribution, centered on the theoretical latency I, with a standard deviation $\sigma = I/5$. This standard deviation can be thought of as a 20% time jitter applied on a neuron's latency. The value of 20% means that the minimum jitter (corresponding to the most activated cells) will in any case be greater than 1 ms. Under these conditions, according to the average neuron's firing frequency distribution (see Figure 5), the average time jitter observed over an infinite time period would be approximately 60 ms. Note that the stochastic distribution applied can lead to negative latencies with nonzero probability for any neuron considered. The (very rare) observed negative latencies were replaced by zero latencies. It is worth underlining the fact that the new procedure introduced here is not a different coding scheme from the order code. Rather, it is a test of how the rank order coding scheme can resist a decrease in timing precision.

For each image, the same exact spike trains were used for decoding and reconstruction purposes by the different rate codes described. At different time steps (1, 2, 4, 8, . . . , 1024 ms for the rate codes as well as the noisy order code, $Ref + 1$, $Ref + 2$, $Ref + 4$, . . . , $Ref + 1024$ ms for the order code), the reconstructed images derived by all different codes were normalized and saved for further analysis. The average MI and MSE of the reconstructions could then be plotted for each code as a function of time. Note that this renormalization process, an adjustment of gray levels between 0 and 255, can have an influence on only the estimated MSE, not on the MI, which is independent of the actual gray-level values. The MSE can thus be found to be relatively high with input images that span very few gray levels, that is, low-contrast images (one should keep in mind that there was no preprocessing or normalization on the set of input images). Results are presented in Figures 6 and 7.

5.3 Analysis. In terms of MI as well as MSE, the order code (both the "reliable" and "noisy" versions of that code) is better than either of the firing-rate-based codes (count code and ISI code). Both rate-based codes
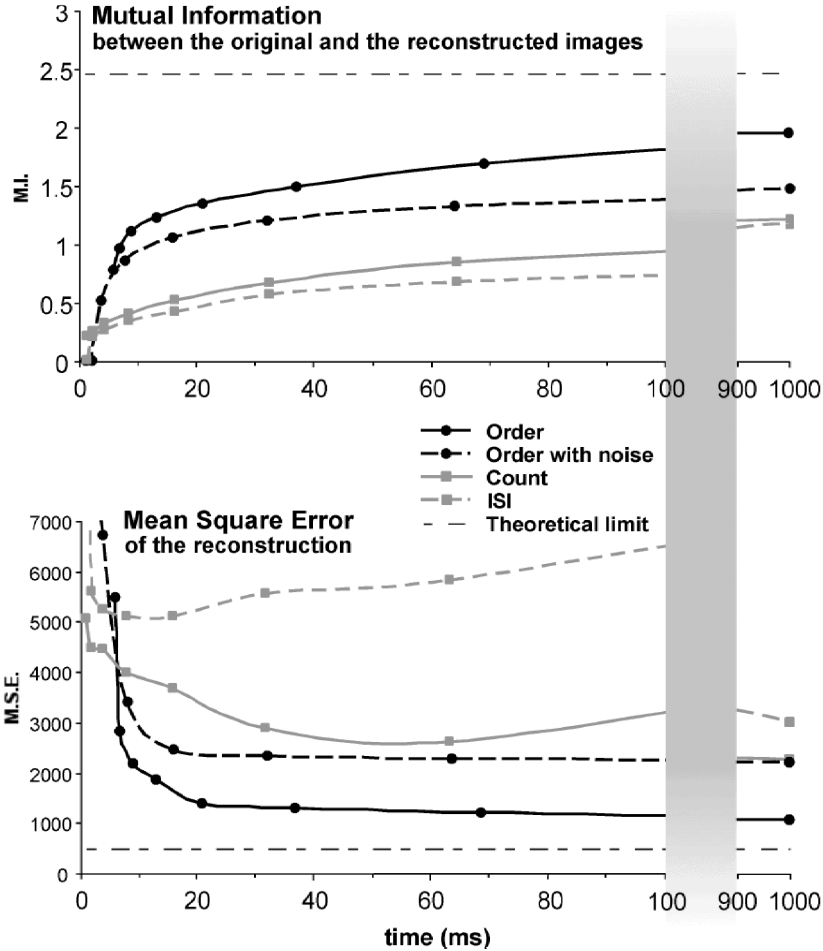
Figure 6: Results. Mutual information and mean square error of the images reconstructed by the different codes between 0 and 100 ms after stimulus presentation, as a function of time. The theoretical limit performance of the model is plotted as a dotted line. Performance between 900 and 1000 ms is also indicated on the right.

appear to be limited by the probabilistic process in the Poisson spike train generation, as well as the relatively low average firing rate R. Nevertheless, it is worth noticing that the Poisson nature of the spike firing in the two rate codes means that some information can indeed be transmitted during the first 5 ms after stimulus presentation, whereas with the ("reliable") order
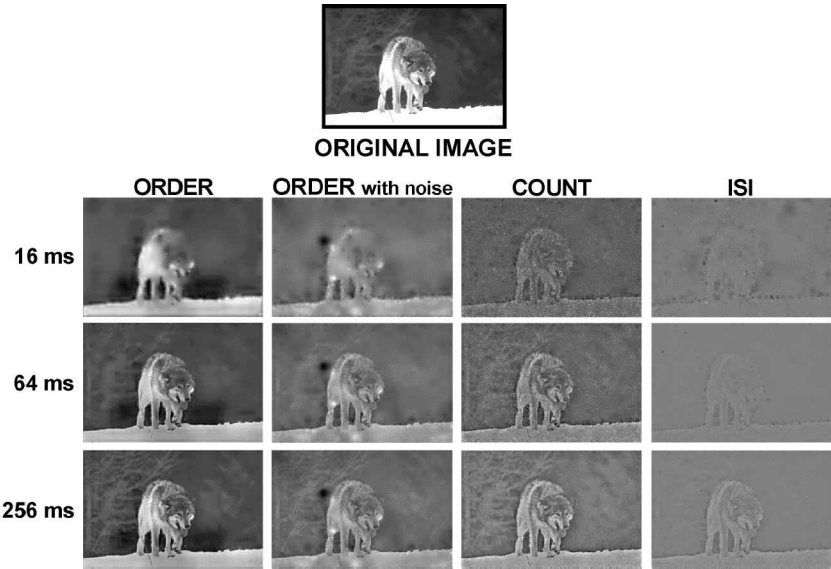
Figure 7: Image reconstructions obtained with the different codes at various time steps after stimulus presentation.

code, the first 5 ms are totally blind. Recent studies support the idea that this advantage explains how rate codes can read out neuronal responses over short time windows (Panzeri, Treves, Schultz, & Rolls, 1999). However, this early information transmitted by the rate codes is clearly not sufficient here to lead to a better reconstruction than the one obtained with the order code. The saturation of the observed firing rate R at 200 Hz is clearly a limiting factor. Indeed, count codes rely on the number of spikes received from each ganglion cell and therefore give better results with high firing rates. Also, the observed firing rate is not linearly dependent on the input contrast, which is not optimal for a wavelet-like transform.

These limitations appear more clearly on the examples of reconstructed images in Figure 7. The information transmitted by the two rate codes during the first milliseconds is subject to a noise that cannot be compensated for by the redundancy in the number of spikes. This noise implies spurious high firing rates from ganglion cells not necessarily carrying the most relevant information. As a consequence of the contrast saturation at these points, less importance will be attributed to other regions.

The count code, in the first few milliseconds of computation, when most of the neurons have not fired at all and the great majority of the rest have fired only once, acts as a binary code. Even after 100 ms, if the highest observed firing rate is around 200 Hz, which corresponds to a spike count of 20 spikes,

the count code can differentiate only between 20 different contrast levels. In the same time, with an order code the system can already signal 100 different contrast levels (because the time step of the simulation is 1 ms).

With the ISI code, even after more than 200 ms, a ganglion cell that has generated only two spikes, but very close in time (on the order of 1 ms), will be attributed much more importance than one that has already generated 10 spikes but separated by 20 ms on average. This illustrates why this code can be reliable only after every neuron has fired a considerable number of spikes (the more the better in the presence of noise, and at least two if the spike train is a noise-free process).

In the case of a rank order coding scheme, one spike per neuron is sufficient to generate the reconstruction. Nevertheless, it is not a binary code. The importance given to a spike depends on when it has been emitted relative to the other cells in the population. Since a ganglion cell will fire sooner if it is highly activated, the first information to appear on the reconstruction will be the most relevant. Indeed, more than 90% of the information that can possibly be transmitted with this code is sent during the first 100 ms. Furthermore, with the parameter values that have been chosen (see section 4.2), the first 1% of the ganglion cells have fired after only 15 ms of computation. It can be seen on Figure 4 that most of the time, this information is sufficient to identify the object in the image. However, a reconstruction with a rank order coding scheme cannot exactly reproduce the original intensity levels, as can be observed on Figure 7, because it applies to every image the average statistical distribution of intensities that we derived from thousands of natural images (see section 3.2). The specific distribution of intensity levels corresponding to a particular image is "forgotten" by the system during the analog-to-rank conversion. But in many cases, this is actually an advantage: it can be seen as an automatic renormalization of inputs. In particular, this scheme is completely invariant to changes in global contrast or mean luminance level. An image at low contrast will produce a similar result, whereas with a Poisson rate code, the effects of noise would be even greater.

Finally, the resistance of rank order coding to a considerable time jitter applied on the neurons' latencies needs to be stressed. There are numerous arguments for taking into account millisecond-precise firing events. One could argue, on the other hand, that this precision will hold only for relatively high input contrasts. Here we have shown that with a time jitter of at least 1 millisecond for the most activated neurons and of the order of tens of milliseconds for the great majority of neurons, the rank order coding scheme can still outperform rate codes based on Poisson stochastic spike trains. This is even more surprising because the statistical rank distribution of natural images (the average contrast value corresponding to a neuron firing with a given rank) has been computed (see section 3.2) under conditions where the firing order was fully reliable. If the ganglion cells' firing process actually was unreliable, this would certainly lead to a different statistic over the ensemble of natural images, reflecting the fact that the first neurons to fire

are not always the most activated. There is no doubt that using this new statistical distribution instead of the one that we applied here would improve the order reconstruction, under conditions where the neurons' latencies are not fully reliable.

In summary, codes based on the firing probability of the neurons are not optimal for information transmission between the retina and the brain. It is important to make use of the precise temporal information carried by the spike trains, when it is available. Furthermore, assigning a greater impact to the first spikes generated in the retina is a good way of maximizing the information transfer rate.

Under these conditions, the very first spikes occurring after stimulus presentation can carry enough information to allow considerable further cortical processing.

6 Discussion

It has already been argued (Gautrais & Thorpe, 1998) that codes based on the neurons' firing rates are unlikely to be efficient enough for fast information processing. During the first 10 ms of computation (i.e., restricting the number of spikes per neuron to zero or one), n neurons can transmit only $\log_2(n + 1)$ bits of information with a count code, whereas $\log_2(n!)$ bits of information can be transmitted with an order code. While ISI codes could transmit an arbitrarily large amount of information with two spikes per neuron, no information at all is transmitted by the first spike. It is also worth noticing that these calculations constitute an upper bound on information transmission under noise-free conditions. Using a Poisson process for spike train generation makes things even worse. As Gautrais and Thorpe (1998) argued, to overcome these limitations, rate codes need to be either time-consuming, or they require a high redundancy in the number of neurons encoding a simple analog value (see also Shadlen & Newsome, 1998). In the retina, the limited number of ganglion cells does not allow for such a redundancy. Here we have shown, using an information transmission approach, that rate codes require too much processing time to be efficient under these conditions.

Our analysis of the problems arising as one tries to decode the neurons' firing rates has also shed light on other interesting questions. In particular, the ISI code has not proved as efficient as expected. In our implementation of the ISI code, the estimated firing rate was updated only when a new spike was received. But if a neuron has received two spikes in a few milliseconds and no information at all for the next 20 ms, should the estimated interspike interval be updated, or should the estimated firing rate remain high? In the former case, the loss of computational efficiency is clear, because all values need to be updated at each time step. In the latter, the interpretation of the firing rate is far from optimal, as shown in our simulations. Perhaps one could devise a better decoding mechanism, taking into account some

combination of the information in the count and the ISI code. However, this would require additional decoding machinery in the postsynaptic cell.

Another argument with which rate codes cannot easily cope is the existence of fast postsynaptic plasticity (Abbott, Varela, Sen, & Nelson, 1997; Markram & Tsodyks, 1996). A count code is fairly useless in the presence of a fast synaptic depression, which will attribute some meaning to only the first few spikes received. However, this kind of synaptic advantage to the first spikes generated—those that carry the most relevant information—is exactly what one would need in order to decode a rank order code.

On the other hand, the idea of a rank order coding scheme is also associated with a number of limitations, but these limitations are more theoretical than practical. First, the order code as described here requires a general "reset" between the processing of two successive images, so that the first spikes emitted in response to a new image are not considered as the last spikes triggered by the previous one. Visual saccades or micro-saccades (Rodieck, 1998; Martinez-Conde, Macknik, & Hubel, 2000) could constitute the functional basis for such a reset mechanism. A much simpler solution, however, is to restrict our investigation to conditions where the input image is flashed briefly, preceded and followed by a uniform dark field. The response of the retina is a short wave of asynchronous spikes, and the time between two successive waves can be made long enough for recovery (i.e., reset) to occur. Indeed, the very fast scene processing demonstrated by Thorpe et al. (1996), Fabre-Thorpe et al. (1998), and VanRullen and Thorpe (1999) was achieved under precisely these conditions. Another point worth discussing is the assumed "globality" of the order code. For reconstructing the input image, we have used the order of firing across the whole neuron population. One could argue that this is neither realistic nor fair, when the rate coding mechanisms described are essentially local. However, as we underscored before, the idea of a global image reconstruction certainly does not correspond to what real cortical neurons do. A cortical neuron receiving retinal information via the LGN will need to know only the order of firing of afferents inside its own receptive field. Such local order decoding schemes have been shown to account for simple-cell oriented (Thorpe & Gautrais, 1997) and more complex receptive field properties (VanRullen et al., 1998). Therefore, if global information is required for the whole image reconstruction, more local information is sufficient for cortical processing. In the same way, the rather wide range of spatial scales used in our model was necessary to reconstruct the whole image in an efficient way. However, a limited number of spatial scales will be sufficient to describe fully the visual information falling into one cortical neuron's receptive field (for example, only three spatial scales will describe the stimulus in a 15×15 pixel cortical neuron's receptive field). It is reasonable to believe that what we have demonstrated here for the whole retinal image will still hold for a single cortical neuron receptive field, which can be considered as a much smaller image. In conclusion, most of the theoretical problems associated with the idea of a rank

order code can be overcome with very few assumptions about the simulated visual processing and the underlying cortical organization. The main question remaining, which we have left open until now, concerns the assumed temporal precision of the neurons' latencies.

There are numerous arguments for taking into account the precise temporal structure of the spike trains generated by the ganglion cells. The timing of firing events has been found to be highly reproducible, with a timing jitter as low as 1 ms (Berry et al., 1997; Reich, Victor, Knight, Ozaki, & Kaplan, 1997). This temporal structure has been shown to carry several times more information than the mean firing rate alone (Berry et al., 1997; Naka & Sakai, 1991; Reich, Victor, & Knight, 1998). More important, the latency of the first spike generated by ganglion cells has been found to vary systematically with the input contrast, a property that could not be explained as a trivial consequence of the change in firing rate amplitude (Sestokas, Lehmkuhle, & Kratz, 1991). We have shown here that it is easy to take advantage of this temporal information and that it requires few assumptions about the neurons' properties. Furthermore, the integrate-and-fire ganglion cells that we used for that purpose are well known to reproduce the temporal properties of retinal spike trains described above (Reich et al., 1997, 1998).

However, most studies refer to the mean firing rate as the only relevant variable of a ganglion cell's response (Warland, Reinagel, & Meister, 1997). This idea is also widely used to model neural computation throughout the visual system (Heller, Hertz, Kjaer, & Richmond, 1995; Gerstner & van Hemmen, 1992), despite the fact that numerous studies have reported a high temporal precision of individual spikes (millisecond or submillisecond range) in various visual areas such as V1 (Richmond & Optican, 1990; Victor & Purpura, 1996), V2, V3 (Victor & Purpura, 1996), IT (Nakamura, 1998), or middle temporal cortex (Bair & Koch, 1996). Importantly, the latency of the first spike of a spike train has been found to be the most reliable (Mainen & Sejnowski, 1995). The question of the efficiency of rate codes is the subject of a strong and still open debate (Softky, 1995; Shadlen & Newsome, 1995). In this context, we believe that our approach constitutes a strong argument in favor of fast codes that use the temporal structure of the spike train instead of a simple mean firing rate, estimated from the spike count or the mean ISI.

The idea that the mean firing rate might not be the optimal way to describe neural activity is not new (Perkel & Bullock, 1968; Thorpe & Imbert, 1989). But many of the alternative temporal coding schemes that are currently attracting so much interest (Bair, 1999) involve spike synchronization as a way of increasing information transfer (Meister et al., 1995; Engel, Konig, Kreiter, Schillen, & Singer, 1992; Singer, 1999; Eckhorn, 1994). Here we have explored another type of temporal coding scheme that relies on spike asynchrony rather than synchrony and have shown that it could be applied to fast information transfer between the retina and the visual cortex. In particular, we show that many scenes can be recognized when only 1 to 2% of the neurons have fired one spike. This is important, given the speed

with which the visual system is known to operate. A pure rate code would not be adequate.

The rank order code, as well as the two rate codes examined here, do not constitute the complete set of possible coding schemes, and further work is needed to evaluate other candidate codes. For example, although this is hardly plausible from a biological point of view, the rate codes implemented here could be tested under noise-free conditions. Another option would be to choose another spike distribution (e.g., a gamma function) for the spike train instead of the Poisson function that some studies consider as a nonoptimal description of the statistical properties of neuronal spike trains (Baddeley et al., 1997; Reich et al., 1998). To overcome the limitations induced by the nonlinear contrast-to-rate transformation (Michaelis-Menten function), an efficient rate code could also estimate the neuron's "free" firing rate (see section 4.1) instead of its theoretical firing rate, subject to a saturation at high contrast levels (Berry & Meister, 1998).

Other temporal codes could also be investigated. First, the rank-order coding scheme could be applied to the first spikes of the Poisson spike trains that were used for rate decoding in our simulations. Nevertheless, previous considerations on the precision of the timing of individual spikes, and on the relevance of the first spike latency, show that it is perfectly reasonable to use millisecond-precise firing events. Another temporal code could extract the neurons' exact firing latencies instead of their rank of discharge. This would allow the fast transmission of analog values. Gautrais and Thorpe (1998) showed that this was indeed a very powerful code. However, it implies that the brain should have an idea of the time of stimulus presentation (zero-time), which is not realistic. Moreover, such a code would be dependent on the contrast and luminance levels in the input image, whereas one of the advantages of the order coding scheme presented here is its invariance to such changes. To be as complete as possible, codes taking into account synchronous firing events among neurons should also be tested. Recordings of neural responses in the LGN have demonstrated that correlated activity can occur with a millisecond precision (Alonso, Usrey, & Reid, 1996), and could be used to reinforce the thalamic input to visual cortex or serve as an additional information channel (Dan, Alonso, Usrey, & Reid, 1998).

Finally, it should be possible to evaluate "hybrid" codes, regrouping features from two or more of the codes described above. For example, we could design a code in which the high temporal precision of the first spike is used to encode as much information as possible on the basis of rank order coding, and afterward the neuron switches to a more sustained Poisson-like discharge for the rest of the spike train. This latter analog information could then refine the one derived by the order code.

It is important to stress that the model retina that we used in this article is unrealistic for a number of reasons. First, the model neuron's properties do not include an adaptation mechanism, which is known to influence the response of real ganglion cells (Smirnakis, Berry, Warland, Bialek, & Meister,

1997). Second, there were no lateral interactions between our ganglion cells, whereas these interactions have been found to play a major role in retinal processing (Nirenberg & Latham, 1998; Meister et al., 1995; Brivanlou et al., 1998). A recent study (Berry, Brivanlou, Jordan, & Meister, 1999) has also demonstrated that ganglion cells are sensitive to moving stimuli in such a way that they can help in predicting the next position of the visual stimulus. To model the ganglion cells' population and topography more accurately would also require implementing a fovea and a resolution decrease with eccentricity (Croner & Kaplan, 1995; Curcio & Allen, 1990; Crook, Lange-Malecki, Lee, & Valberg, 1988).

But here the idea was to keep the model as simple as possible, in order to show that our hypothesis could lead to efficient results when compared to other rate-based codes. That is why we did not try to mimic the mammalian retina in detail. At first sight, it seems obvious that any improvement in the architecture of the model would benefit the rank order coding scheme as much as the rate codes, though this statement could be directly investigated by making our model retina more similar to the mammalian retina. Furthermore, there is little doubt that such a scheme, being biologically more plausible, could save neurons, spikes, and computational time.

An interesting alternative to the purely theoretical model we described here would be to apply to real retinal spike trains the different decoding mechanisms compared in this study. A recent experiment by Stanley, Li, and Dan (1999) has demonstrated the possibility of reconstructing spatiotemporal visual inputs from an ensemble of spike trains simultaneously recorded in the LGN. In this study however, the variable used to describe the neuron's responses was their time-varying firing rate. Our results suggest that other variables reflecting the precise temporal structure of the spike trains could constitute a powerful way of describing neural responses, and this theory could benefit from direct experimental investigation.

7 Conclusions

We have shown that codes based on the ganglion cells' mean firing rate, as derived from the spike count or the mean interspike interval in a Poisson spike train, cannot account for the efficiency of information transmission between the retina and the brain. We have introduced instead a coding scheme based on the relative order in which the ganglion cells emit their first spike in response to a given visual stimulation and shown that this rank order coding, although computationally very simple, can lead to a very good stimulus reconstruction, even over relatively short time periods and even when firing latencies are subject to a 20% time jitter. Provided that we take into account the temporal structure of the spike train and assign a greater impact to the spikes with the shortest latencies, it appears that the very first spikes generated in the retina can carry sufficient information for further cortical processing. In addition, these results demonstrate that the

idea of temporal coding refers not only to using spike synchrony as a way of multiplexing information with multiple channels; spike asynchrony is a powerful alternative that should not be neglected.

Moreover, these results do not apply only to retinal coding. They explain how, given a certain architecture, the neurons can maximize their information transfer rate. Thus, this scheme could be used for stimulus encoding in the input layer of any kind of artificial neural network, but could also account for information transfer between any two layers of such a network. Indeed, we used this code in simulations with SpikeNET (Delorme, Gautrais, VanRullen, & Thorpe, 1999), and it has proved powerful, especially when the next layers of processing also use a rank order coding scheme.

Finally, the proximity between the neural coding scheme presented here and wavelet-based compression, as well as MPEG or JPEG technologies, should be pointed out. Millions of years of natural selection have no doubt produced an image compression scheme in the optic nerve that is highly optimized; engineers could well find that the main features of modern image compression techniques were already discovered millions of years ago.

Acknowledgments

References

Abbott, L. F., Varela, J. A., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science, 275*, 220–224.

Alonso, J. M., Usrey, W. M., & Reid, R. C. (1996). Precisely correlated firing in cells of the lateral geniculate nucleus. *Nature, 383*, 815–819.

Arnett, D., & Spraker, T. E. (1981). Cross-correlation analysis of the maintained discharge of rabbit retinal ganglion cells. *J. Physiol. (Lond.), 317*, 29–47.

Baddeley, R., Abbott, L. F., Booth, M. C., Sengpiel, F., Freeman, T., Wakeman, E. A., & Rolls, E. T. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc. R. Soc. Lond. B. Biol. Sci., 264*, 1775–1783.

Bair, W. (1999). Spike timing in the mammalian visual system. *Curr. Opin. Neurobiol., 9*, 447–453.

Bair, W., & Koch, C. (1996) Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Comput., 8*, 1185–1202.

Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature, 398*, 334–338.

Berry, M. J., & Meister, M. (1998). Refractoriness and neural precision. *J. Neurosci., 18*, 2200–2211.

Berry, M. J., Warland, D. K., & Meister, M. (1997). The structure and precision of retinal spike trains. *Proc. Natl. Acad. Sci. U.S.A., 94*, 5411–5416.

Brivanlou, I. H., Warland, D. K., & Meister, M. (1998). Mechanisms of concerted firing among retinal ganglion cells. *Neuron, 20*, 527–539.

Buser, P., & Imbert, M. (1992). *Vision*. (R. H. Kay, Trans.). Cambridge, MA: MIT Press.

Celebrini, S., Thorpe, S. J., Trotter, Y., & Imbert, M. (1993). Dynamics of orientation coding in area V1 of the awake monkey. *Visual Neurosci., 10*, 811–825.

Croner, L. J., & Kaplan, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. *Vision Res., 35*, 7–24.

Crook, J. M., Lange-Malecki, B., Lee, B. B., & Valberg, A. (1988). Visual resolution of macaque retinal ganglion cells. *J. Physiol., 396*, 205–224.

Curcio, C. A., & Allen, K. A. (1990). Topography of ganglion cells in human retina. *J. Comp. Neurol., 300*, 5–25.

Dan, Y., Alonso, J. M., Usrey, W. M., & Reid, R. C. (1998). Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nat. Neurosci., 1*, 501–507.

Delorme, A., Gautrais, J., VanRullen, R., & Thorpe, S. J. (1999). SpikeNET: A simulator for modeling large networks of integrate and fire neurons. *Neurocomputing, 24*.

DeVries, S. H. (1999). Correlated firing in rabbit retinal ganglion cells. *J. Neurophysiol., 81*, 908–920.

Eckhorn, R. (1994). Oscillatory and non-oscillatory synchronizations in the visual cortex and their possible roles in associations of visual features. *Prog. Brain Res., 102*, 405–426.

Engel, A. K., Konig, P., Kreiter, A. K., Schillen, T. B., & Singer, W. (1992). Temporal coding in the visual cortex: New vistas on integration in the nervous system. *Trends Neurosci., 15*, 218–226.

Fabre-Thorpe, M., Richard, G., & Thorpe, S. J. (1998). Rapid categorization of natural images by rhesus monkeys. *Neuroreport, 9*, 303–308.

Field, D. J. (1994). What is the goal of sensory coding? *Neural Comput., 6*, 559–601.

Fischer, B. (1973). Overlap of receptive field centers and representation of the visual field in the cat's optic tract. *Vision Res., 13*, 2113–2120.

Gautrais, J., & Thorpe, S. J. (1998). Rate coding vs temporal order coding: A theoretical approach. *Biosystems, 48*, 57–65.

Gerstner, W., & van Hemmen, J. L. (1992). Universality in neural networks: The importance of the "mean firing rate." *Biological Cybernetics, 67*, 195–205.

Heller, J., Hertz, J. A., Kjaer, T. W., & Richmond, B. J. (1995). Information flow and temporal coding in primate pattern vision. *J. Comput. Neurosci., 2*, 175–193.

Lankheet, M. J., Molenaar, J., & van de Grind, W. A. (1989). The spike generating mechanism of cat retinal ganglion cells. *Vision Res., 29*, 505–517.

Lipetz, L. E. (1971). The relation of physiological and psychological aspects of sensory intensity. In W. R. Lowenstein (Ed.), Handbook of sensory physiology (Vol. 1, pp. 191–225). New York: Springer-Verlag.

Mainen, Z. F., & Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science, 268*, 1503–1506.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *Inst. Electrical Electronics Engrs. Trans. on Pattern Analysis and Machine Intelligence, 11*, 674–693.

Markram, H., & Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature, 382*, 807–810.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2000). Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. *Nat. Neurosci., 3*, 251–258.

Mastronarde, D. N. (1989). Correlated firing of retinal ganglion cells. *Trends Neurosci., 12*, 75–80.

Meister, M., & Berry, M. J. (1999). The neural code of the retina. *Neuron, 22*, 435–450.

Meister, M., Lagnado, L., & Baylor, D. A. (1995). Concerted signaling by retinal ganglion cells. *Science, 270*, 1207–1210.

Naka, K. I., & Rushton, W. A. H. (1966). S-potentials from luminosity units in the retina of fish (cyprinidae). *Journal of Physiology, London, 185*, 587–599.

Naka, K., & Sakai, H. M. (1991). The messages in optic nerve fibers and their interpretation. *Brain Res. Brain Res. Rev., 16*, 135–149.

Nakamura, K. (1998). Neural processing in the subsecond time range in the temporal cortex. *Neural Comput., 10*, 567–595.

Nirenberg, S., & Latham, P. E. (1998). Population coding in the retina. *Curr. Opin. Neurobiol., 8*, 488–493.

Nowak, L. G., Munk, M. H. J., Girard, P., & Bullier, J. (1995). Visual latencies in areas V1 and V2 of the macaque monkey. *Visual Neurosci., 12*, 371–384.

Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *J. Neurophysiol., 68*, 70–84.

Panzeri, S., Treves, A., Schultz, S., & Rolls, E. T. (1999). On decoding the responses of a population of neurons from short time windows. *Neural Comput., 11*, 1553–1577.

Peichl, L., & Wässle, H. (1979). Size, scatter and coverage of ganglion cell receptive field centres in the cat retina. *J. Physiol. (Lond.), 291*, 117–141.

Perkel, D. H., & Bullock, T. H. (1968). Neural coding. *Neurosciences Research Program Bulletin, 6*, 221–348.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research, 47*, 329–342.

Reich, D. S., Victor, J. D., & Knight, B. W. (1998). The power ratio and the interval map: Spiking models and extracellular recordings. *J. Neurosci., 18*, 10090–10104.

Reich, D. S., Victor, J. D., Knight, B. W., Ozaki, T., & Kaplan, E. (1997). Response variability and timing precision of neuronal spike trains in vivo. *J. Neurophysiol., 77*, 2836–2841.

Richmond, B. J., & Optican, L. M. (1990). Temporal encoding of two-dimensional patterns by single units in primate primary visual cortex. II. Information transmission. *J. Neurophysiol., 64*, 370–380.

Rieke, F., Warland, D., de Ruyter van Steveninck, R. R., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

Rodieck, R. W. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Res., 5*, 583–601.

Rodieck, R. W. (1998). *The first steps in seeing*. Sunderland, MA: Sinauer Associates.

Sestokas, A. K., Lehmkuhle, S., & Kratz, K. E. (1987). Visual latency of ganglion X- and Y-cells: A comparison with geniculate X- and Y-cells. *Vision Res., 27*, 1399–1408.

Sestokas, A. K., Lehmkuhle, S., & Kratz, K. E. (1991). Relationship between response latency and amplitude for ganglion and geniculate X- and Y-cells in the cat. *Int. J. Neurosci., 60*, 59–64.

Shadlen, M. N., & Newsome, W. T. (1995). Is there a signal in the noise? *Curr. Opin. Neurobiol., 5*, 248–250.

Shadlen, M. N., & Newsome, W. T. (1998). The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *J. Neurosci., 18*, 3870–3896.

Singer, W. (1999). Time as coding space? *Curr. Opin. Neurobiol., 9*, 189–194.

Smirnakis, S. M., Berry, M. J., Warland, D. K., Bialek, W., & Meister, M. (1997). Adaptation of retinal processing to image contrast and spatial scale. *Nature, 386*, 69–73.

Softky, W. R. (1995). Simple codes versus efficient codes. *Curr. Opin. Neurobiol., 5*, 239–247.

Stanley, G. B., Li, F. F., & Dan, Y. (1999). Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *J. Neurosci., 19*, 8036–8042.

Thorpe, S. J. (1990). Spike arrival times: A highly efficient coding scheme for neural networks. In R. Eckmiller, G. Hartman, & G. Hauske (Eds.), *Parallel processing in neural systems* (pp. 91–94). Amsterdam: North-Holland, Elsevier.

Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*, 520–522.

Thorpe, S. J. & Gautrais, J. (1997). Rapid visual processing using spike asynchrony. In M. C. Mozer, M. Jordan, & T. Petsche (Eds.), *Advances in neural information processing systems, 9* (pp. 901–907). Cambridge, MA: MIT Press.

Thorpe, S. J., & Gautrais, J. (1998). Rank order coding: A new coding scheme for rapid processing in neural networks. In J. Bower (Ed.), *Computational neuroscience: Trends in research* (pp. 113–118). New York: Plenum Press.

Thorpe, S. J., & Imbert, M. (1989). Biological constraints on connectionist models. In R. Pfeifer, Z. Schreter, F. Fogelman-Soulié, & L. Steels (Eds.), *Connectionism in perspective* (pp. 63–92). Amsterdam: Elsevier.

VanRullen, R., Gautrais J., Delorme A., & Thorpe, S. J. (1998). Face processing using one spike per neuron. *Biosystems, 48*, 229–239.

VanRullen, R., & Thorpe, S. (1999). *Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects*. Unpublished paper, Centre de Recherche Cerveau et Cognition.

Vardi, N., & Smith, R. G. (1996). The AII amacrine network: Coupling can increase correlated activity. *Vision Res., 36*, 3743–3757.

Victor, J. D., & Purpura, K. P. (1996). Nature and precision of temporal coding in visual cortex: A metric-space analysis. *J. Neurophysiol., 76*, 1310–1326.

Warland, D. K., Reinagel, P., & Meister, M. (1997). Decoding visual information from a population of retinal ganglion cells. *J. Neurophysiol., 78*, 2336–2350.