

Maximising Information Recovery from Rank-Order Codes

Sen B. ^a and Furber S. ^a

^aUniversity of Manchester, School of Computer Science,
Oxford Road, Manchester M13 9PL, U.K.

ABSTRACT

The central nervous system encodes information in sequences of asynchronously generated voltage spikes, but the precise details of this encoding are not well understood. Thorpe proposed rank-order codes as an explanation of the observed speed of information processing in the human visual system. The work described in this paper is inspired by the performance of SpikeNET, a biologically inspired neural architecture using rank-order codes for information processing, and is based on the retinal model developed by VanRullen and Thorpe. This model mimics retinal information processing by passing an input image through a bank of Difference of Gaussian (DoG) filters and then encoding the resulting coefficients in rank-order. To test the effectiveness of this encoding in capturing the information content of an image, the rank-order representation is decoded to reconstruct an image that can be compared with the original. The reconstruction uses a look-up table to infer the filter coefficients from their rank in the encoded image. Since the DoG filters are approximately orthogonal functions, they are treated as their own inverses in the reconstruction process.

We obtained a quantitative measure of the perceptually important information retained in the reconstructed image relative to the original using a slightly modified version of an objective metric proposed by Petrovic. It is observed that around 75% of the perceptually important information is retained in the reconstruction.

In the present work we reconstruct the input using a pseudo-inverse of the DoG filter-bank with the aim of improving the reconstruction and thereby extracting more information from the rank-order encoded stimulus. We observe that there is an increase of 10 – 15% in the information retrieved from a reconstructed stimulus as a result of inverting the filter-bank.

Keywords: neural encoding, rank-order codes, Difference of Gaussian (DoG), ganglion cells, image reconstruction, perceptual information recovery, model of retina, pseudo-inverse, Singular Value Decomposition, edge detection.

1. INTRODUCTION

The ganglion cells in the retina transform the information collected by the photoreceptors into sequences of identical spikes before passing it on to the brain down the visual pathway. This transformation function is mathematically represented by the Difference of Gaussian (DoG) function, mimicking the centre-surround structure of the retinal ganglion cell receptive fields on the retina.¹ How the brain reads the information from the multiple spike trains fired by a population of ganglion cells is a question yet to be answered. The rate code theory is the earliest recorded attempt to decode the neural codes.² It is based on experimental observations that the firing frequency of a spiking neuron varies in proportion to the stimulus strength. If we consider ourselves to be at the receiving end of a spiking neuron, we can estimate the locally applied stimulus strength by calculating the instantaneous frequency from two consecutive spikes fired by the neuron.

Experiments in recent times show that humans can recognise previously unseen scenes within 150msec after stimulus onset.³ If 50msec is considered as the time between stimulus onset and the start of neural response, this leaves about 100msec for visual processing. The information collected by the photoreceptors has to pass through about 10 synaptic stages to reach the cortex. Therefore, at each synaptic stage, a neuron gets about 10msec to propagate information by firing a spike. Again, most cortical neurons have a firing rate of below 100

Further author information: (Send correspondence to B. Sen)

B. Sen: E-mail: Basabdatta.Sen@postgrad.manchester.ac.uk, Telephone: +44 161 275 6294

S. Furber: E-mail: Steve.Furber@manchester.ac.uk, Telephone: +44 161 275 6129

spikes per sec. Thus, in a time window of 10msec, a neuron can fire at the most a single spike, or it may not even fire at all.^{4,5}

The rate code theory falters in giving an explanation of the speed of information processing in the human visual system. The retina must be using a much more efficient coding scheme to transmit information to the brain. To this effect, the rank-order codes⁷ proposed by Thorpe provide a plausible explanation. The idea behind the theory is the fact that the time taken by the integrate-and-fire neurons to reach threshold depends on the strength of the input. Thus the latency of a spike will be inversely proportional to the stimulus strength.

To test the performance of the codes, VanRullen and Thorpe built a model retina,⁸ using the DoG functions to filter the visual stimulus such that the magnitude of a certain coefficient of filtering corresponds to the activation level of a ganglion cell and thus to the latency of its firing a spike. Once encoded, the spikes are then decoded to reconstruct the visual stimulus and thus evaluate the performance of the rank-ordered code. Since the DoG is an approximately orthogonal function, it is treated as its own inverse in the decoding process. Also, since the true coefficients of filtering are lost in the rank-order encoding, a look-up-table of weights is used for the decoding. On visual assessment of the fidelity of the reconstructed picture, it is observed that by the time only 1% of the ganglion cells have fired their first spikes, the perceptually important parts of the input are fairly well reproduced in the output. Hence it is claimed that the rank-order codes are a viable means of information transfer in neurophysiologically inspired large-scale networks of asynchronously spiking neurons.⁹⁻¹¹

We started our work by simulating VanRullen's retinal model, but we aimed for a quantitative rather than a qualitative measure to test the information that is retained in the rank-order encoded image. The raw primal sketch is considered by Marr¹² to be a rich description of an image since it contains virtually all the perceptually important parts in an image. An algorithm proposed by Petrovic¹³ gives an objective measure for preservation of edges in a fused image with respect to two parent images. The algorithm uses the concept of primal sketch and assumes that all the perceptually important parts of a picture are contained in the edge contrast, orientation, and position. We used a slightly modified version of the algorithm to measure the perceptually important information that can be preserved in a rank-order encoded visual input with respect to the input itself. Using such a measure, we have shown¹⁴ that on an average, up to 75% of the perceptually important parts of the input stimuli can be retrieved. Further, more than 90% of this 75% information is retrieved by the time only up to 25% of the ordered spikes are decoded.

During the course of our previous work, we observed that even if we use the true coefficients of filtering during the decoding process, we do not get a perfect reconstruction. This indicates that we are not able to retrieve the maximum possible information from the rank-order codes. The loss of information is due to the fact that the DoG is not an exact orthogonal function and overlap errors are introduced during the encoding process.

In this work we aimed to achieve a perfect reconstruction of the visual stimulus using the true coefficients of filtering, anticipating a consequent increase in the information retrieval from the rank-ordered stimulus using the look-up-table. We used the pseudo-inverse of the filter-bank consisting of the DoG functions. We observed that on an average there is a 10% rise in the information retrieved from the rank-order encoded images by the time about 38% of the neurons have fired their first spikes. This therefore gives a better estimate of the maximum information available from the rank-order encoding of the image.

2. RANK-ORDER CODES: DEFINITION, SIMULATION, EVALUATION

The rank-order code hypothesis requires that an input stimulus applied to a population of neurons is encoded with an intensity-to-delay transformation function. It is based on the observation that the speed of information processing in the human visual system does not allow a neuron enough time to fire more than one spike. When a population of ganglion cells is stimulated with a visual input, the first spike corresponds to the most activated cell, the second spike to the second most activated cell, and so on. Thus, the ranking or the order of firing of the neurons in the population relates to the distribution of the input stimulus.

One of the advantages of using such a code is its invariance to both input intensity and contrast. The change in overall luminance and contrast of an input will not change the relative order of firing of the neurons, although there will be a change in the firing latency of each. Thus, there is an automatic normalization of the inputs.⁷

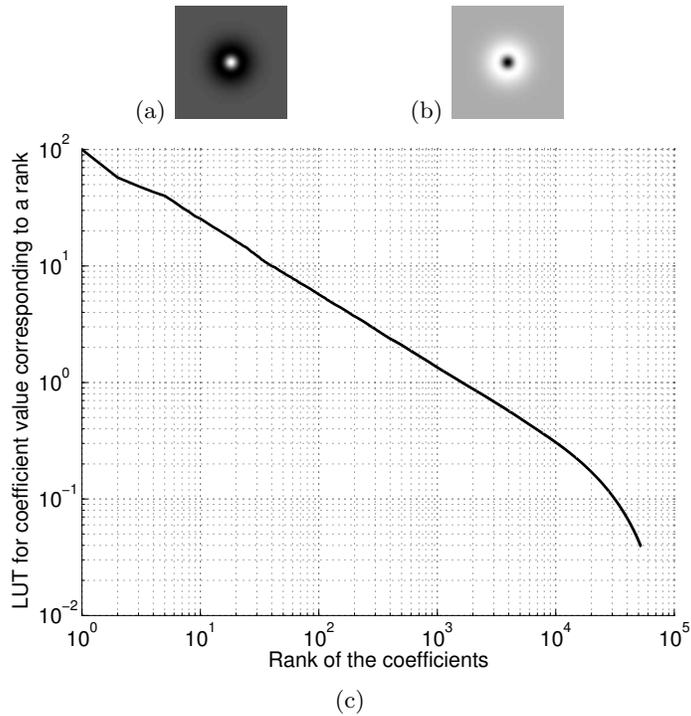


Figure 1. (a) On-Centre DoG (b) Off-Centre DoG (c) The LUT provide an average estimate of the activation level of each cell depending on its rank.

2.1. Simulating the Rank-Order Encoding of a Visual Stimulus

To test the theory, VanRullen and Thorpe encoded an image into rank-order by simulating a model retina. The centre-surround structure of the ganglion cell receptive field can be mathematically described by a Difference of Gaussian (DoG) function¹ as in Eq. (1). In this model, a simple DoG described by Field¹⁵ is used, where the width of the centre (σ_1) and the surround (σ_2) are in the ratio 1 : 3.

$$DoG(x, y) = \frac{1}{2\pi\sigma_2^2} \left(9e^{-\frac{(x^2+y^2)}{2\sigma_1^2}} - e^{-\frac{(x^2+y^2)}{2\sigma_2^2}} \right) \quad (1)$$

On and off centre cells in the retina, shown in Fig. 1 (a) and (b) respectively, are simulated by DoG functions of various sizes, the smallest being a 5×5 matrix, and the largest, a 767×767 matrix. The increase in size at each scale is such that if the size of the function at scale $s(i)$ is $n \times n$ with a standard deviation of d , then the size at scale $s(i+1)$ is $(2n+1) \times (2n+1)$ with a standard deviation of $2d$.

The linear filtering of an input image by the ganglion cell receptive fields is simulated by a taking the DoG as the basic filter to obtain a discrete wavelet transform of the image. Further, the transformation is in dyadic progression, i.e. as the size of the filter doubles, the sampling resolution of the transform is halved at each consecutive scale. The model uses DoG filters at eight scales to form a complete basis function of the wavelet transform. The filtering of the image I with the DoG at a certain scale s is given by the following equation:

$$C_s(k, l) = \sum_x \sum_y I(k-x)(l-y) \cdot DoG_s(x, y) \quad (2)$$

The set of data C derived from Eq. (2) are the true coefficients of filtering the input with the whole set of DoG functions. Each coefficient $C_s(i)$ corresponds to a certain scale s and a spatial location (k, l) and is considered as a neuron's membrane potential. The characteristic of an integrate-and-fire neuron model is that the higher the

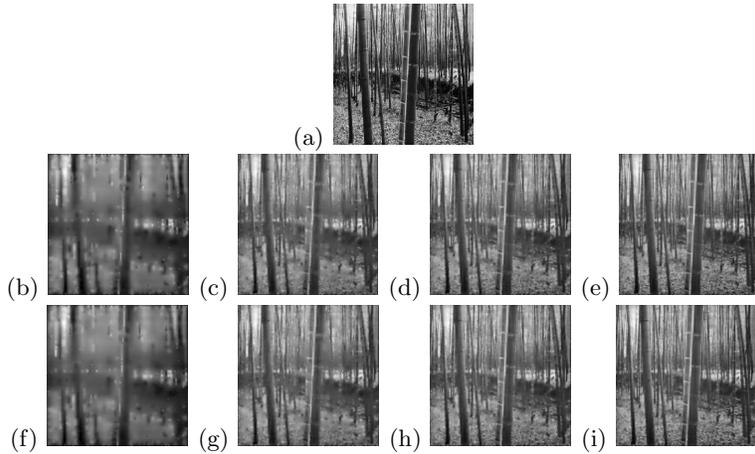


Figure 2. (a) input image (b)–(e) Reconstructed image using 1%, 5%, 10% and 20% respectively of the coefficients of filtering with the DoG. (f)–(i) Reconstructing rank-order encoded image using the first 1%, 5%, 10% and 20% of the values respectively from the LUT.

membrane potential, the sooner the threshold voltage is reached and the sooner a spike is fired by the neuron. The largest coefficient of filtering may be considered as corresponding to the neuron which is the most activated in the population and therefore is the first to fire a spike. Thus, each ganglion cell in the model retina is assigned a rank according to its activation level, which is represented by the coefficient of filtering at the corresponding spatial location and scale.

2.2. Decoding The Codes

To evaluate the quality of the rank-order codes, VanRullen and Thorpe used the rank-ordered spikes to reconstruct the input and thus estimate the quantity of information about the input carried by the codes. Since the DoGs are approximately orthogonal wavelet functions, they are treated as their own inverse in the reconstruction. The reconstruction is the reverse operation to that described in Eq. (2) and is written as

$$I_{rec}(k, l) = \sum_s \sum_k \sum_l C(x, y) \cdot DoG_s(x - k)(y - k) \quad (3)$$

But once rank-ordered, we only know the rank of a spike and have lost knowledge of the activation level that drove it above threshold. For this, VanRullen and Thorpe generated a look-up-table (LUT) of contrast values versus the rank of a neuron. The contrast value corresponding to each rank is estimated by averaging the true coefficients of filtering at that particular rank for 3000 images. Thus, a neuron is weighted according to its order of firing using the LUT, and the weight is fixed for a certain rank across all the input images.

In our simulation, we have generated a similar LUT using an array of around 70 images of resolution 256×256 and is shown in Fig. 1 (c).

2.3. A Qualitative Evaluation

To decode the rank-order encoding, values from the LUT are used in Eq. (3) instead of the true coefficients of filtering. The largest LUT value is associated with the first neuron to fire a spike, the second largest with the second firing neuron and so on. The image is reconstructed progressively as the spikes are fired. VanRullen and Thorpe showed that by the time 1% of the cells fire, the input is recognisable from the reconstruction. By the time 10% of the total population of ganglion cells have fired their first spike, the images are hardly differentiable. The reconstructed images using the coefficient set C as well as the LUT are shown in Fig. 2.

Thus, the rank-order code theory suggests that by the time only 10% of the total population of ganglion cells affected by a visual stimulus have fired just their first spikes, the brain has enough information to be able to view and perceive a complex scene.

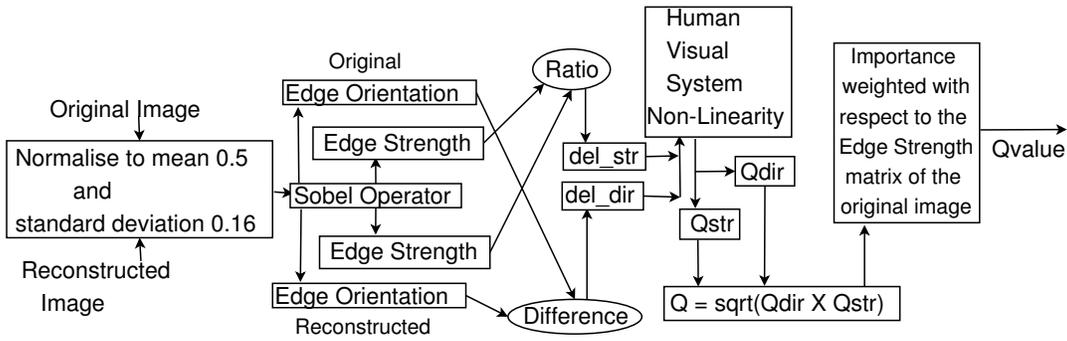


Figure 3. Flowchart of Petrovic’s algorithm as used in our work

3. QUANTITATIVE EVALUATION OF RANK-ORDER CODES

As described in Sect. 2, the conclusion about the information retained in the rank-ordered spikes is only a qualitative evaluation depending on the subjective judgment of perceptual differences in the original and the reconstructed image. Our aim is to obtain a quantitative measure of the perceptually important information retained in the reconstructed stimulus. This will create a benchmark for measuring the information encoded in rank-ordered spikes, and in turn will provide uniformity in evaluating the performance of the codes across a range of complex visual scenes.

Perceptual judgment of a picture depends on three factors, viz. spatial location, spatial frequency, and orientation. These factors are characterized by the edges in an image. An edge in an image corresponds to discontinuity in its intensity surface. It can be approximated by a piecewise linear curve composed of short, linear edge elements, known as edgels. Each edgel can be considered as a vector characterized by strength, direction and position. This approach is used in an algorithm developed by Petrovic¹³ to estimate the edge information preserved in an image formed by fusing two parent images. We used this algorithm, after slight modification, to measure the information content in the reconstructed stimulus with respect to the original. In this section, we give a brief description of how Petrovic’s algorithm is used here to measure the information recovered from rank-order encoded visual stimulus.

3.1. Appropriate Modification And Application Of Petrovic’s Algorithm

The flowchart of the algorithm as implemented in this work is shown in Fig. 3. Petrovic uses a Sobel Operator, which is a first order differentiator, to obtain the horizontal and vertical component of each edgel, whence the strength and direction are obtained.

If s_x and s_y are the horizontal and vertical components of an edgel, then the edgel strength is written as

$$E_{str} = |s_x| + |s_y|$$

and the orientation is given as

$$E_{dir} = \arctan\left(\frac{s_x}{s_y}\right)$$

The original image and the reconstructed image are normalized so that they have a common mean and standard deviation. Both images are processed through the Sobel filters to obtain the edge strength and orientation measure at every pixel position for each image.

Let $\delta_{str}(x, y)$ be the contrast ratio at the pixel position (x, y) . Then

$$\delta_{str}(x, y) = \frac{EI_{str}(x, y)}{ER_{str}(x, y)}$$

where EI and ER correspond to the edgel parameters of the original and the reconstructed images respectively. Similarly if $\delta_{dir}(x, y)$ is the difference in orientation at (x, y) , then

$$\delta_{dir}(x, y) = |EI_{dir}(x, y) - ER_{dir}(x, y)|$$

Both δ_{str} and δ_{dir} are normalized to lie in the range 0 to 1.

When comparing two images, the detection of the difference in edge luminosity and orientation at a certain position by the human eye depends on the detection threshold of the eye. The detection threshold is different for contrast and orientation, and is defined as the point at which the probability of detection of the difference by the human eye is 50%. With reference to this algorithm, this threshold is the point at which the edge strength or orientation has perceptually degraded to 50% of its original quality. The nonlinear behavior of the eye is defined as a psychometric function

$$Q = \frac{K}{1 + \exp^{-d(x+s)}}$$

where $x = [\delta_{str}, \delta_{dir}]$, d is the detection threshold, s is the steepness parameter, and K is a constant such that for optimal values of d and s , $Q = 1$. Petrovic uses results of subjective trials to set the value of the parameters $[d, s]$ to $[0.7, 11]$ and $[0.8, 24]$ for the contrast ratio and the orientation respectively.

The method described above gives a pixel to pixel measure of the perceptual information that has been preserved for the reconstructed image with respect to the original image, both for edge strength $Q_{strength}$ and orientation $Q_{orientation}$. From these two measures, we obtain a combined measure

$$Q = \sqrt{Q_{strength} \times Q_{orientation}}$$

Q is then importance weighted with the edge-strength values of the original picture. This is expressed as a normalized sum to give a single measure for the performance.

$$Q_{value} = \frac{\sum(W_{original} \times Q)}{\sum W_{original}}$$

3.2. Observations

The stimulus is reconstructed with each incoming spike as described in Sect. 2.2, and the information in the reconstruction measured with respect to the original using the algorithm described in Sect. 3.1. Thus, we obtain a plot of the information build up in the reconstructed image with each incoming spike. Figure 4 shows the information plot for two images. When applied over a wide variety of images, it is observed that, firstly, about 75% of the information can be retrieved from a rank-order encoded visual stimulus by the time only 20% of the cells in the population have fired their first spikes. Secondly, around 90% of this recoverable information comes through by the time only up to 15 – 20% of the cells have fired. Thirdly, the information is recovered much faster from an image with fewer edges. These observations seems to be biologically plausible as we notice the finer details in a picture some time after we first look at it.

4. TOWARDS PERFECT RECONSTRUCTION

The basis vector for image transformation in VanRullen’s retinal model is a set of DoG filters of different scales and types. The filters are nearly orthogonal although not exactly so. Also, since the filters are circularly symmetric, they are considered as their own inverse during the stimulus reconstruction, treating them as orthogonal. We observe that the information retrieved from a rank-order encoded visual stimulus is around 75% of the information content of the original image. Although the results are encouraging, it is observed that the quantification of information does not conform to the qualitative estimate of information content in a reconstructed stimulus with respect to the original. On visual inspection of the reconstructed images using true coefficients in Figures 4 and 5, we observe that they appear blurred with respect to the original as is expected from the flatter histogram (Fig. 5). Yet, we note that most of the edge information, i.e. the perceptually important parts in the original images are retained in these reconstructed images. Thus we would expect a Q_{value} , not equal to 1 as it is an

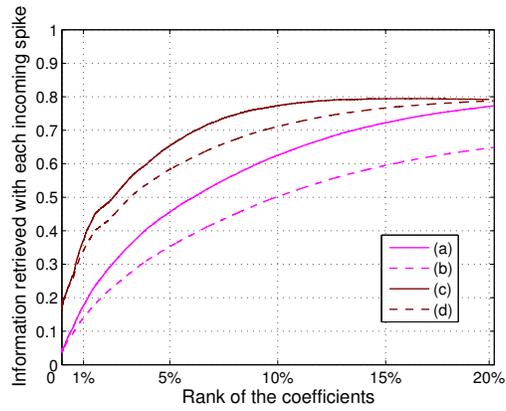


image 1



image 2



(e)



(f)



(g)



(h)

Figure 4. Plots (b) and (d) show the progressive information recovery from the rank-ordered encoded visual input for two images (image 1 and image 2 respectively). The corresponding information recovered when using the true coefficients of filtering are shown in plots (a) and (c). The reconstructed pictures at the point where 20% of the data have been used for information recovery in plots (a)–(d) are shown in images (e)–(h) respectively.

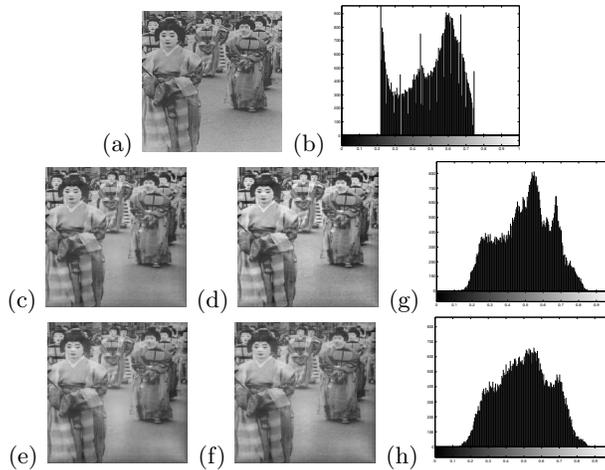


Figure 5. (a) An input image normalised to mean=0.5 and std.dev.=0.16 and (b) its histogram. (c)–(f) Reconstruction of the input using (c) 25% and (d) 50% of the true coefficients of filtering and (e) 25% and (f) 50% of data from the LUT. (g), (h) Histograms for the reconstructions shown in (d) and (f) respectively.

imperfect reconstruction (discussed in Sect. 4.1), yet having a value closer to 1. But as can be seen from the information plot in Fig. 4, the Q_{value} is considerably less than 1, signifying the recovery of less information than is actually anticipated by qualitative evaluation. This is because, as shown in Fig. 5, the contrast in the reconstructed images is decreased, thus causing a significant deviation in the pixel values with respect to the original image at regions of interest.

Our aim is to recover the maximum possible information from the rank-order codes and thus give a true estimate of their performance in terms of the encoded information content. In this section, we describe how we use the pseudo-inverse of a DoG filter-bank to get $Q_{value} = 1$ indicating a perfect stimulus reconstruction using the true coefficients of filtering. As expected, the Q_{value} for the decoded stimulus using the approximate coefficient values from the LUT also increases for every picture compared to the earlier method of reconstruction.

4.1. The DoG As Orthogonal Filters

VanRullen and Thorpe assume the DoG filters to be orthogonal wavelet transforms. This means that the responses of the filter when shifted along the image, or scaled, are uncorrelated, i.e. $\langle DoG_s(x, y), DoG_s(u, v) \rangle = 0$, where $(x, y) \neq (u, v)$. Now, if the filters are orthogonal, the reconstruction should be perfect. But this is not the case in Eq. (3) as $I_{rec} \neq I$. This is because the DoG filters are only approximately orthogonal, i.e. $\langle DoG_s(x, y), DoG_s(u, v) \rangle = \varepsilon$, where ε is small but not zero. Thus, the reconstruction is approximate and the reconstructed image is blurred which is characterized by a point-spread function (psf) of the response:¹⁶

$$I_{rec} = I * psf, \quad \text{where } psf = \sum_{1 \leq s \leq s_{max}} \frac{1}{\sigma_s^2} \langle DoG_s, DoG_s \rangle,$$

and s_{max} is the maximum scale. To obtain a perfect reconstruction using the true coefficients of filtering, we do away with this approximation and treat the DoG as non-orthogonal filters.

4.2. The DoG As Non-Orthogonal Filters

4.2.1. The DoG filter-bank

We consider the DoG filters as non-orthogonal filters and form a filter-bank $F^{M \times N}$ of the scaled and translated DoG filters. If the image size is $m \times n$, then $N = m \times n$, and $M = (8/3) \times m \times n$. Each DoG is a row vector of length N , and corresponds to a certain spatial location and scale. If it corresponds to a spatial location (k, l) , then it is in the K^{th} row, where $K = m(k - 1) + n(l - 1)$, and the element at (d_1, d_2) of the two dimensional representation of the DoG is at the L^{th} column of the K^{th} row, where $L = m(k + d_1 - 2) + n(l + d_2 - 1)$.

4.2.2. Reconstruction using exact data

We filter the input image using the filter-bank $F : C = F * I$, where C is the data set corresponding to the true coefficients mentioned above and I is the input image. We want to reconstruct an estimate I_R of the image I from the coefficients C , which we can write informally as $I_R = F^{-1} * C$. This is a case where we have M equations (the rows of F) in N variables (the elements of I) and $M > N$, so it is an over-determined system of equations. The solution may be obtained by taking the Moore-Penrose pseudo-inverse of F (appendix A.2), written as F^+ , by doing singular value decomposition on F (appendix A.1). This is the least squares solution for I_R and not an exact solution. But in the present case, the equations are consistent and I_R represents the exact solution. Therefore, if $I_R = C * F^+$, then $I_R = I$ and we have a perfect reconstruction of the input stimulus using the true coefficients.

We now order the coefficients in C according to their rank and perform a progressive reconstruction of I . The elements in C are arranged in descending order of their magnitudes, to represent a progressively decreasing activation level and hence progressively increasing latency of firing. The rows in F are permuted so as to exactly match the corresponding coefficient. Let $G = F(1)$, where $F(1)$ is the first row of F . We want a solution to $G * I = C$, which corresponds to an underdetermined set of equations with $M < N$. Since the equations are consistent, we apply the solution as described in Eq. (8) with $\Gamma = 0$ in Eq. (6) (Appendix A.3) to get $I_R(1) = C * G^+$. This will correspond to the image reconstruction when only the most stimulated cell has fired one spike. Using Petrovic's method, a quantitative measure $Q_{value}(1)$ of the information content in $I_R(1)$ with

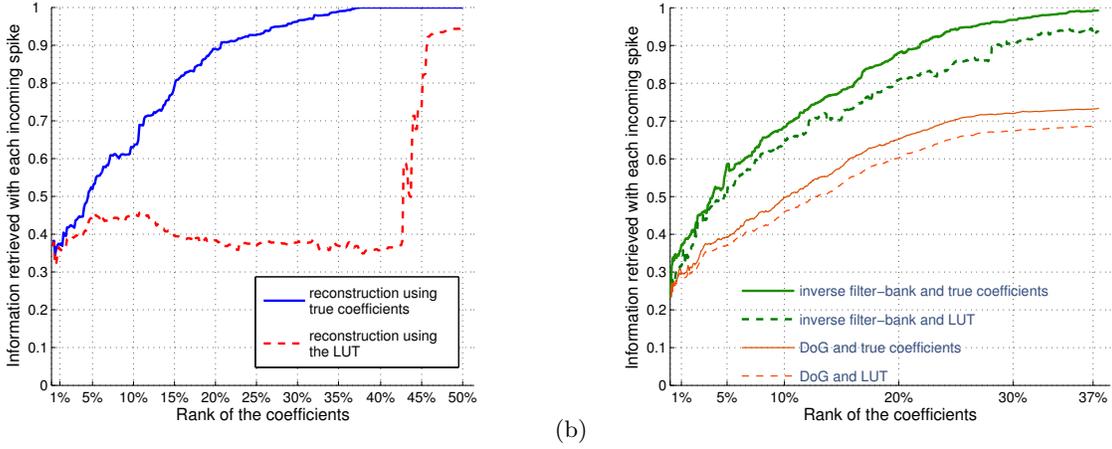


Figure 6. (a) Using the pseudo-inverse of the DoG filter-bank gives perfect reconstruction by the time upto 37.5% of the true coefficients of filtering are used. But for approximate coefficients from a LUT, the filter-bank is illconditioned as long as it is underdetermined, and its pseudo-inverse doesn't work properly. This anomaly is overcome by setting a cut-off threshold (0.1 in this work) for the singular values of the filter-bank before taking its pseudo-inverse. The cut-off threshold is used in the LUT results shown in (b). (b) A comparison of the progressive information recovery using the two methods of data reconstruction, viz. using the DoG as its own inverse vs. using the pseudo-inverse of the DoG filter-bank. Reconstruction using both the true coefficients of filtering and values from the LUT are shown.

respect to I is obtained. Next, we assign $G(2) = F(2)$. Now G has two non-zero rows, and $I_R(2) = C * G^+$, whence we obtain $Q_{value}(2)$. Continuing thus we find that $Q_{value}(t) = 1$ by the time $t = N$, as shown in Fig. 6(a). Now, $N/M = 3/8 = 0.375$. Thus, we obtain a full reconstruction of the input by the time we have 37.5% of the coefficients. On proceeding further, the number of rows in G becomes $(N + count)$ where $count = 1 \dots (M - N)$. Since $(N + count) > N$, the pseudo-inverse is calculated as given in Eq. (7) (Appendix A.3), with $\Gamma = 0$ in Eq. (6).

4.2.3. Reconstruction using approximate data

We now move on to the rank-order encoding of the visual stimulus, where we have lost the true coefficients of filtering and use the values from the LUT as described in Sect. 2.2. We do a progressive reconstruction as mentioned in Sect. 4.2.2 by progressively adding rows to G and doing $I_R(t) = G^+ * LUT(t)$. Interestingly, we observe that since the LUT is an average estimate of the actual data set C , the equations are inconsistent as long as they represent an underdetermined set of equations, i.e. $t < N$. This leads to erratic results giving values of $Q_{value} \approx 0$. The plot for the progressive reconstruction in such a case is shown in Fig. 6(a).

This behavior of the Q_{value} plot when done with data from the LUT can be explained thus: As long as $M < N$ in the matrix G , it is an ill-conditioned matrix (Appendix A.3). Thus, prior to calculating G^+ , we need to set $\Gamma \neq 0$ in Eq. (6) (Appendix A.3), unlike in Sect. 4.2.2, where we were working with exact data. By trial and error, we found the value of $\Gamma = 0.1$ as suitable for our present work. The plots for the progressive reconstruction using the true coefficients of filtering vs. that using data from the LUT are shown in Fig. 6(b). For comparison, information recovery plots using the DoG as their own inverse for image reconstruction are also shown on the same figure.

In our simulation we have used 32×32 image matrices as inputs due to time and memory constraints. Figure 7 shows the information recovered from three different rank-order encoded images, using both the methods of reconstruction viz. using the DoG as their own inverse vs. using the pseudo-inverse of the DoG filter-bank. We notice a distinct rise in the measured information in the reconstructed images using the latter method. The average information content by the time N out of M cells has fired is on average 90%, an increase of 10 – 15% over that in the reconstructed images obtained in Sect. 3.2. Scaling up the image sizes is expected to behave

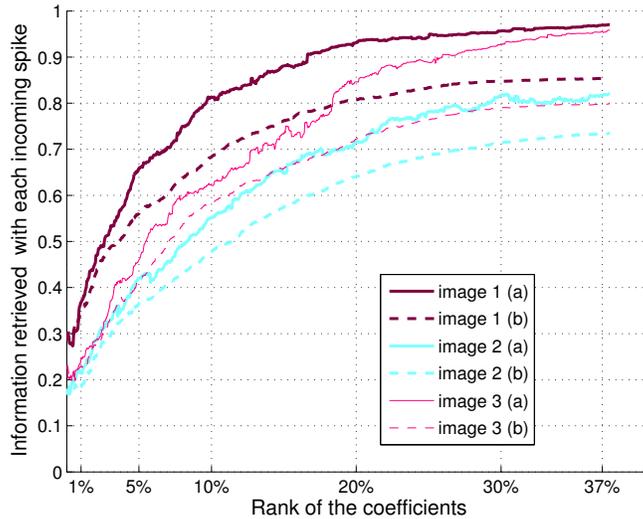


Figure 7. Information recovered from three different rank-order encoded images using the two methods of reconstruction viz. (a) pseudo-inverse of the DoG filter-bank and (b) the DoG as its own inverse.

essentially in the same manner although the exact value of information content may vary for different sizes of a certain image.

5. CONCLUSION

We present in this paper a simulation of the retinal model originally developed by Van Rullen and Thorpe to test the performance of rank-order codes. The stimulus is progressively reconstructed using each spike according to their order of firing, firstly, to evaluate the amount of information about the stimulus that is recovered from its rank-order encoded form, and secondly, to observe how many cells are required to fire their first spikes to convey the major part of the information in the input. We observe that indeed, by the time the first 10% of the ganglion cells have fired their first spikes, the quality of the image reconstruction is very good. For most of the images, the reconstruction seem to convey all the perceptually important information present in the input.

We then present how we quantify the above qualitative measures using Petrovic’s Edge Preservation Estimation algorithm. The results show that up to 75% of the information contained in an input is available in its reconstruction. Further our results also show that about 90% of the totally recoverable information is available when only up to 25% of the cells have fired their first spikes. Also, images having fewer edges, e.g. silhouettes, have a steep information recovery curve compared to those which contain a lot of fine detail, and the information recovery is complete by the time up to 20% of the cells have fired.

Next we discuss how Van Rullen and Thorpe’s method of stimulus reconstruction is sub-optimal to aid proper information recovery from the reconstructed image. We go on to describe how we obtain perfect reconstruction using the true coefficients of filtering and the pseudo-inverse of a DoG filter-bank, which is not possible using VanRullen and Thorpe’s method where the DoG is used as its own inverse. The difficulty of using this method with inexact data set from the LUT is discussed. We describe how application specific discretion is required while applying the Singular Value Decomposition technique for obtaining the pseudo-inverse of an ill-conditioned matrix, and how we have solved the problem successfully.

A comparison of the information build-up using both VanRullen’s method of reconstruction and our method presented in this paper is shown. The results show that there is a 10 – 15% increase in the information recovery from a reconstructed stimulus by the time 38% of the cells fire their first spikes.

The simulation is done using images of size 32×32 . The drawback of the method is that it needs a large amount of memory for storing and working with the filter-banks. Also, the computation time of the pseudo-inverse of a matrix is very slow and normally takes over 3 hrs. for a 32×32 image on a desktop computer. The method is scalable and produces consistent results for images of different sizes although the absolute values of the information recovered varies.

APPENDIX A.

A.1. Singular Value Decomposition

Singular Value Decomposition (SVD) is a robust and powerful technique for dealing with rectangular, singular or numerically close to singular matrices. It is defined as:¹⁷ Any $M \times N$ matrix A whose number of rows M is greater than or equal to its number of columns N , can be written as the product of an $M \times N$ column orthogonal matrix U , an $N \times N$ diagonal matrix W with positive or zero elements, and the transpose of an $N \times N$ orthogonal matrix V .

$$A^{M \times N} = U^{M \times N} \cdot W^{N \times N} \cdot (V^T)^{N \times N} \quad (4)$$

and

$$W_{N \times N} = \begin{pmatrix} w_1 & & & & \\ & w_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & w_N \end{pmatrix}$$

The diagonal elements of the matrix W are known as the singular values and this decomposition of the matrix A is known as the Singular Value Decomposition.

A.2. Pseudo-Inverse Of A Matrix Using SVD

If A is an $M \times N$ matrix, then there exists an $M \times N$ matrix A^+ , known as the pseudo-inverse of A , which is derived by applying an inverse operation on the SVD of A . Therefore, from Eq. (4) we have

$$A^+ = (U \cdot W \cdot V^T)^{-1} = (V^T)^{-1} \cdot \text{diag}(1/w_j) \cdot U^{-1} = V \cdot \text{diag}(1/w_j) \cdot U^T \text{ (Since } U \text{ and } V \text{ are orthogonal)} \quad (5)$$

A.3. Solving Linear Equations Using SVD

It is a common situation that some of the w_j 's in Eq. (4) are very small but non-zero. Thus, $(1/w_j)$'s are so large that they approach the machine's floating point precision and are therefore subject to round-off errors. Such matrices are called ill-conditioned matrices and some discretion is required before taking the inverse of the SVD.¹⁷ Below is a generalized method to solve linear equations using the SVD:¹⁸

1. For a matrix $A^{M \times N}$, where $M \geq N$:

Apply SVD to obtain $A = UWV^T$,

Find the largest absolute singular value $w_{max} = \max_{j=1 \dots n} |w_j|$

Set a threshold, (6)
 $t = \Gamma \cdot w_{max},$

where Γ is a constant and may be varied according to the requirement so long as the least square error is within acceptable limits.

Define
$$z_j = \begin{cases} (1/w_j) & \text{if } w_j > t \\ 0 & \text{otherwise} \end{cases}$$

so that
$$A^+ = VZU^T \quad (7)$$

2. If $M < N$:

Apply SVD to A^T to obtain $A^T = UWV^T$, so that $A = VWU^T$

Therefore
$$A^+ = UZV^T \quad (8)$$

ACKNOWLEDGMENTS

We are more than grateful to Tim Cootes, ISBE Manchester for his advise on the Pseudo-inverse and help with his notes. We extend our heartfelt thanks to Gavin Brown, Jon Shapiro, Mike Cumpstey and Joy Bose for their help and valuable advice from time to time.

REFERENCES

1. R. W. Rodieck, Quantitative analysis of cat retinal ganglion cell response to visual stimuli, *Vision Res.* **5**, pp. 583-601, 1965.
2. E. D. Adrian, *Basis Of Sensation*, Haffner Publishing Company, London, 1964.
3. S. Thorpe, D. Fize and C. Marlot, Speed of processing in human visual system, *Nature* **381**, pp. 520-522, 1996.
4. S. Thorpe and J. Gautrais, Rapid visual processing using spike asynchrony, in *Advances in Neural Information Processing System*, pp. 901-907, 1997.
5. M. Fabre-Thorpe, A. Delorme, C. Merlot and S. Thorpe, A limit to the speed of processing in ultra-rapid visual categorization of visual scenes, *J. Cog. Neurosc.* **13**, pp. 171-180, 2001.
6. J. Gautrais and S. Thorpe, Rate coding versus temporal order coding: a theoretical approach, *Biosystems* **48**, pp. 57-65, 1998.
7. S. Thorpe and J. Gautrais, Rank order coding in *Comp. Neurosc.:Trends in Research*, pp. 113-119, 1998.
8. R. VanRullen and S. Thorpe, Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex, *N. Comp.* **13**, pp. 1255-1283, 2001.
9. S. Thorpe, A. Delorme, R. Van Rullen and W. Paquier, Reverse engineering of visual system using networks of spiking neurons, in *IEEE International symposium on circuits and systems*, pp. 405-408, 2000.
10. A. Delorme, J. Gautrais, R. Van Rullen and S. Thorpe, SpikeNET: A simulator for modeling large networks of integrate-and-fire neurons, *Neurocomp.*, **26**, pp. 989-996, 1999.
11. S. Thorpe, R. Guyonneau, N. Guilbaud, J. Allegraud and R. Van Rullen, SpikeNet: Real time visual processing with one spike per neuron, *Neurocomp.*, **58-60**, pp. 857-864, 2004.
12. D. Marr, *Vision*, W. H. Freeman and Company, New York, 1982.
13. V. Petrovic and C. Xydeas, Objective evaluation of signal-level image fusion performance, *Opt. Engg.*, **44(8)**, 2005.
14. B. Sen and S. Furber, Information recovery from rank-order codes, in *Proceedings of International Workshop on Biologically Inspired Information Fusion*, pp. 8-13, 2006
15. D. J. Field, What is the goal of sensory coding? *N. Comp.* **6**, pp. 559-601, 1994.
16. L. Perrinet, M. Samuelides and S. Thorpe, Coding static natural images using spiking event times: do neurons cooperate? *IEEE Trans. On Neur. Net.*, 2004
17. W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, 1992.
18. Lecture Slides of Tim Cootes, ISBE, University of Manchester.
(http://www.isbe.man.ac.uk/~bim/MathsMethodsNotes/L3_linear_algebra3.pdf)