Information Recovery from Rank-Order Encoded Images

Basabdatta Sen[1] and Steve Furber[1]

[1]School of Computer Science
University of Manchester
Oxford Road
Manchester M13 9PL UK

Abstract. The work described in this paper is inspired by SpikeNET, a system developed to test the feasibility of using rank-order codes in modelling large-scale networks of asynchronously spiking neurons. The rank-order code theory proposed by Thorpe concerns the encoding of information by a population of spiking neurons in the primate visual system. The theory proposes using the order of firing across a network of asynchronously firing spiking neurons as a neural code for information transmission. In this paper we aim to measure the perceptual similarity between the image input to a model retina, based on that originally designed and developed by VanRullen and Thorpe, and an image reconstructed from the rank-order encoding of the input image. We use an objective metric originally proposed by Petrovic to estimate perceptual edge preservation in image fusion which, after minor modifications, is very much suited to our purpose. The results show that typically 75% of the edge information of the input stimulus is retained in the reconstructed image, and we show how the available information increases with successive spikes in the rank-order code.

1 Introduction

How does a population of retinal ganglion cells encode visual information into sequences of action potentials? The firing-rate code theory proposed by Adrian [1] says that a population of neurons encode information entirely in the frequencies of firing of the individual neurons. If we imagine ourselves to be at the receiving end of these spikes, then we need at least two spikes from a neuron to determine its firing frequency. More recently, experiments have shown that at each synaptic stage of the Human Visual System (HVS), a neuron has about 10msec to propagate information by firing a spike. Most cortical neurons have a firing rate of below 100 spikes per second. Thus, in a time window of 10msec, a neuron can fire at the most a single spike, or it may not fire at all [2–5]. These findings question the plausibility of the rate-code theory as applied to the HVS [6].

The rank-order code theory proposed by Thorpe [7] overcomes the timing constraints mentioned above. The hypothesis is that the input stimulus applied to a population of neurons is encoded with an intensity-to-delay transformation function. When applied to a spiking neural model of the retina, it is observed that the perceptually important parts of the input stimulus are well reproduced by the time only the first 1% of the neurons have fired their first spikes [8].

At this point, we propose to quantify the above results by measuring the perceptual similarity between the input image and an image reconstructed from the rank-order encoding of the input. An algorithm proposed by Petrovic [9] gives an objective measure for the preservation of edges in a fused image with respect to two parent images. In this paper we describe how we have used a slightly modified version of the algorithm to measure the perceptually important edges that are preserved in a rank-order encoded visual input with respect to the input itself.

2 Rank-Order Codes: Definition and Performance

The rank-order code theory proposes that the latency of a spike fired by a neuron will be inversely proportional to the applied stimulus strength. The exact latency at which a neuron fires is not critical here. Rather, it is the rank-order of the first spike generated by each neuron in a population that is important. Thus, the change in overall luminance and contrast of an input stimulus will not change the relative order of firing of the neurons, although there will be a change in the firing latency of each, resulting in an automatic normalization of the inputs [7].

2.1 Rank-Order Encoding of an Image

To test the performance of rank-order codes, Van Rullen and Thorpe built a model retina [8]. The centre-surround structure of the retinal ganglion cell receptive fields are represented by Difference of Gaussian (DoG) functions, with the width of the surround three times that of the centre [10], as shown in Fig. 1(a). On-centre and off-centre DoG filters at eight scales are used to simulate the different sizes of ganglion cells in the retina, and are shown in Fig. 1(b)(c). The input image is filtered using this set of sixteen

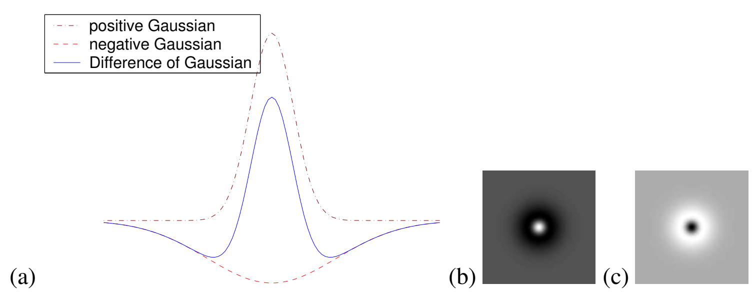

Fig. 1. (a) The one-dimensional DoG function. (b) On-Centre Off Surround and (c) Off-Centre On-Surround DoG functions.

DoG filters. The sampling resolution of filtering decreases with increasing scale of the DoG filter. The result of filtering is a set of sixteen matrices containing the convolution coefficients. A coefficient value models the activation level of a neuron which drives it towards the firing threshold. The largest coefficient value corresponds to the neuron which is the first in the population to spike. The coefficients are arranged in descending order to rank the neurons according to their latency of firing their first spike. What we then have is the input image encoded as rank-ordered coefficients [8].

2.2 Decoding the Rank-Ordered Data: Stimulus Reconstruction

When a population of neurons fire, we only know the order in which the neurons fire, not the stimulus values that drove them above threshold. True rank-order encoding means that we must throw away the true coefficients and adopt a generic activation level value corresponding to the rank of firing of a neuron. In other words, a neuron is weighted according to its order of firing, and the weight is fixed for a certain rank across all input images. This is done by using a look-up table (LUT) for the average coefficient value corresponding to a certain rank. The coefficient value corresponding to each rank is generated by averaging the true coefficients of filtering at that particular rank for a set of images. VanRullen and Thorpe [8] generated the table using three thousand images. In our simulations, we generated a similar LUT using an array of thirty images of resolution 256×256. The average coefficient values are plotted against rank on a log-log scale in Fig. 2, which shows that they closely follow a power law. The plot is fitted with an equation of the form $y = C \times x^{-a}$ where $a \simeq 0.63$ and $C \simeq 100$. To test the

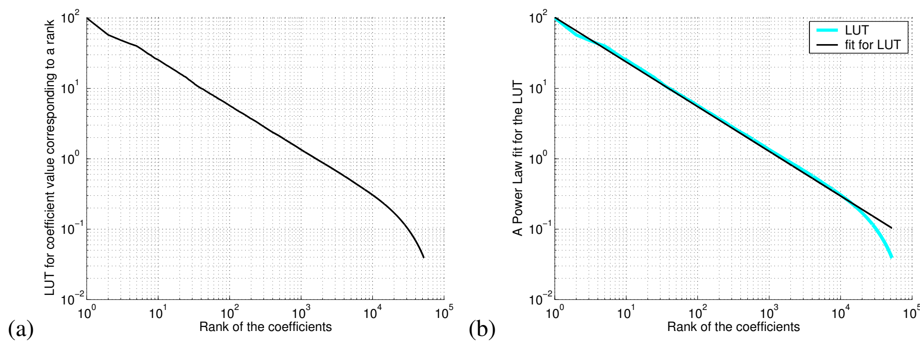

Fig. 2. (a) A normalised LUT plot on the logarithmic scale. (b) Power Law fit for the LUT.

performance of rank-order codes, VanRullen and Thorpe reconstructed the original input stimulus using rank-ordered coefficients read from the LUT. They observed that by the time the first 1% of the spikes have arrived, the subject of the reconstructed picture is fairly recognisable. The reconstructed pictures from our emulation of VanRullen's model are shown in Fig. 3.

3 An Objective Metric for the Performance of Rank-Order Codes

We see in Fig. 3 that there is a distinct perceptual difference between the reconstructed and original pictures. We wish to find a measure for the proportion of the information

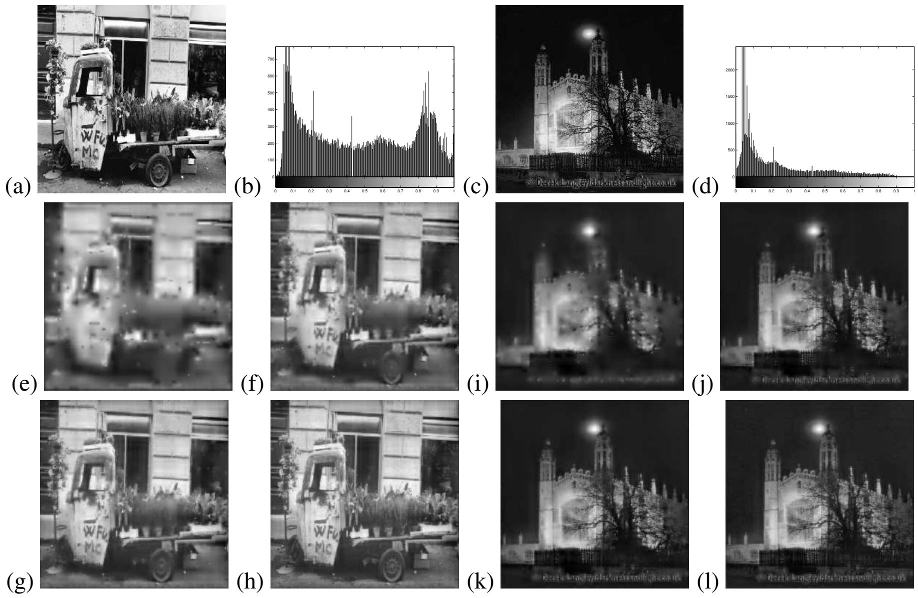

Fig. 3. (a),(c) Original images and (b),(d) their respective histograms. (e)–(h) Reconstruction of the image in (a) using 1%, 5%, 10% and 50% of the LUT values. (i)–(l) Similar data for the image in (c).

in the input image that is transmitted by the rank-order encoding, and to estimate the way that this information builds up through the arrival of successive spikes from a neural population. Petrovic et al. [9] have developed an algorithm for measuring the perceptual loss suffered by an image during the fusion of two parent images, which we have adapted to our purpose. The flowchart of the algorithm as used in our application is shown in Fig. 4.

As seen in Fig. 5(d), the histogram of the reconstructed picture has a greater spread than that of the original Fig. 5(b). To overcome this discrepancy, both the original and the reconstructed pictures are normalised to a common mean (0.5) and standard deviation (0.16) before providing them as inputs to the algorithm. The two normalised images are then passed through a Sobel first-order differentiator to detect the magnitude and direction of all of the edges in each. From this data, we derive the contrast ratio (del_{str}) and difference in orientation (del_{dir}) of the two pictures at each pixel position. This linear data is then scaled to conform to the non-linear behaviour of the HVS as defined by the psychometric function $Q = K/(1+exp^{-d(x+s)})$ where $x = \{del_{str}, del_{dir}\}$, and K is a constant such that for optimal values of d and s decided by results of subjective evaluation [9], $Q = 1$. The method described above gives a measure of the perceptual information that has been preserved in the reconstructed image with respect to the original image, both for edge strength (Q_{str}) and orientation (Q_{dir}), and is shown in Fig. 6(g). The mean $Q = \sqrt{Q_{str} \times Q_{dir}}$ is then importance-weighted with the edge-strength values of the original picture (W_{orig}). This is expressed as a normalized sum to give a single

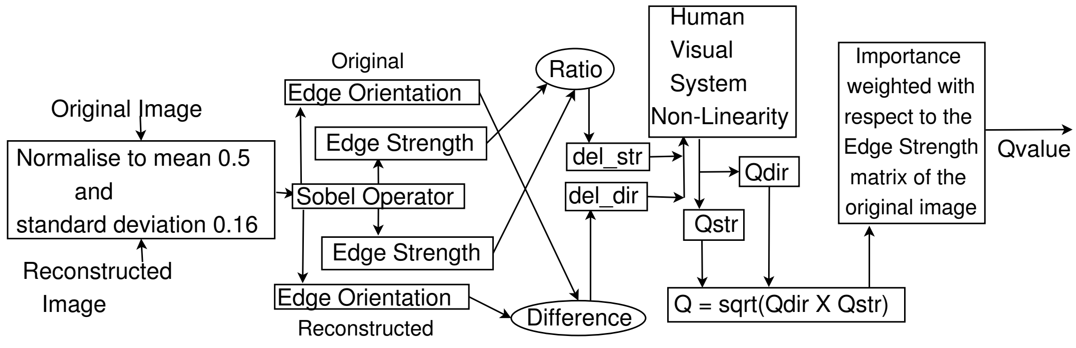

Fig. 4. The Edge Preservation Estimation Algorithm

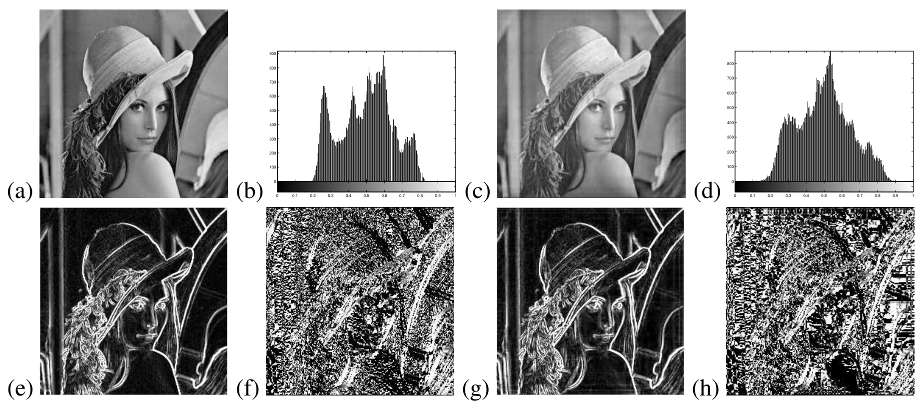

Fig. 5. (a) Original and (c) Reconstructed images along with (b), (d) their respective histograms. (e), (g) Edge Magnitude and (f), (h) Edge Orientation of the original and reconstructed images respectively.

measure $Q_{value} = \sum(W_{orig} \times Q) / \sum W_{orig}$ for the performance. We apply the algorithm illustrated in Fig. 4 to the images shown in Fig. 6(a)–(e) and their reconstructions. Plots of the information build-up (Q_{value}) with the arrival of each spike are shown in Fig. 6(f). It is observed that, on average, the reconstruction using the LUT retrieves upto 75% of the information contained in the original image. Again, of the total information retrieved by the algorithm, more than 90% is obtained by the time 15% – 20% of the neurons have fired their first spikes .

4 Conclusion

In this work we first discussed our replication of Thorpe's work of simulating the visual processing in the retina using rank-order codes. We then wanted a means to measure

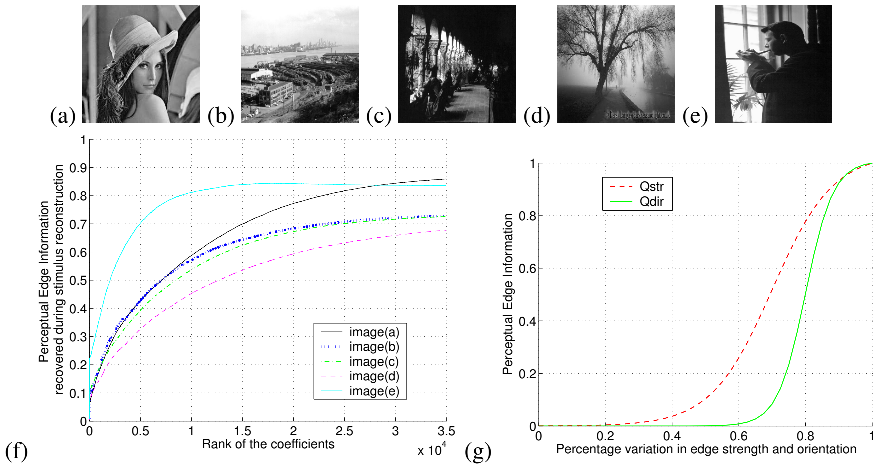

Fig. 6. (a)–(e): Five images for which the rate of information retrieval from their rank-order encoding is plotted in (f); (g) Perceptual Edge Preservation measure.

the performance of the rank-order codes in terms of the preservation of the perceptually important information of the original image. For this, we used the objective edge preservation estimation algorithm devised by Petrovic et al. Our results indicate that more than 90% of the information that the model is capable of coding can be retrieved very early on in the process, by the time only 20% of the neurons have fired their first spikes.

References

1. Adrian, E.D.: Basis of sensation. Hafner Publishing Company, London (1928)
2. Thorpe, S., Imbert, M.: Biological constraints on connectionist models. In: Connectionism in perspective, 63-92, Elsevier, Amsterdam (1989)
3. Thorpe, S., Fize, D., Marlot, C.: Speed of processing in human visual system. Nature **381**, 520–522 (1996)
4. Fabre-Thorpe, M., Delorme, A., Merlot, C., Thorpe, S.: A limit to the speed of processing in ultra-rapid visual categorisation of novel scenes. J. Cog. Neurosc. **13**, 171–180 (2001)
5. Thorpe, S., Gautrais, J.: Rapid visual processing using spike asynchrony. In: Advances in neural information processing systems, 901–907 (1997)
6. Gautrais, J., Thorpe, S.: Rate coding versus temporal order coding: a theoretical approach. Biosystems **48**, 57–65 (1998)
7. Thorpe, S., Gautrais, J.: Rank order coding. In: Computational Neuroscience: Trends in research, 113–119 (1998)
8. Van Rullen, R., Thorpe, S.: Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. Neural Comp. **13**, 1255–1283 (2001)
9. Petrovic, V., Xydeas, C.: Objective evaluation of signal-level image fusion performance. Optical Engineering **44(8)** (2005)
10. Field, D.J.: What is the goal of sensory coding? Neural Comp. **6**, 559–601 (1994)