Cognitive Science Online, Vol.1, pp.1–7, 2003 http://cogsci-online.ucsd.edu

Nonlinear reverse-correlation with synthesized naturalistic noise

Hsin-Hao Yu
Department of Cognitive Science
University of California San Diego
La Jolla, CA 92093
`hhyu@cogsci.ucsd.edu`

Virginia R. de Sa
Department of Cognitive Science
University of California San Diego
La Jolla, CA 92093
`desa@cogsci.ucsd.edu`

Abstract

Reverse-correlation is the most widely used method for mapping receptive fields of early visual neurons. Wiener kernels of the neurons are calculated by cross-correlating the neuronal responses with a Gaussian white noise stimulus. However, Gaussian white noise is an inefficient stimulus for driving higher-level visual neurons. We show that if the stimulus is synthesized by a linear generative model such that its statistics approximate that of natural images, a simple solution for the kernels can be derived.

1 Introduction

Reverse-correlation (also known as white-noise analysis) is a system analysis technique for quantitatively characterizing the behavior of neurons. The mathematical basis of reverse correlation is based on the Volterra/Wiener expansion of functionals: If a neuron is modeled as the functional $y(t) = f(x(t))$, where $x(t)$ is the (one dimensional) stimulus to the neuron, any nonlinear f can be expanded by a series of functionals of increasing complexity, just like real-valued functions can be expanded by the Taylor expansion. The parameters in the terms of the expansion, called *kernels*, can be calculated by cross-correlating the neuronal responses to the stimulus, provided that the stimulus is Gaussian and white (Wiener, 1958; Lee & Schetzen, 1965; Marmarelis & Marmarelis, 1978).

Reverse correlation and its variants are widely used to study the receptive field (RF) structures of the sensory systems. In vision, the circular RF's of LGN neurons and the gabor-like RF's of simple cells in the primary visual cortex are revealed by calculating the first-order (linear) kernels. Neurons with more nonlinearity, such as complex cells, can also be studied by the second-order kernels (Szulborski & Palmer, 1990). However, reverse correlation is rarely applied to extrastriate visual areas, such as V2. One of the many factors that limit reverse correlation to the study of the early visual system is that Gaussian white noise is an inefficient stimulus for driving higher order neurons, since visual features that are known to activate these areas (Gallant et al., 1996; Hegdé & Van Essen, 2000) appear very rarely in Gaussian white noise.

The goal of this paper is to show that if we generate more "interesting" stimuli by training a linear generative model from natural images, solutions to the kernels can be obtained easily. We will proceed by first formulating the Volterra/Wiener series, describe the linear generative model of stimulus synthesis, derive the kernels, and then compare this scheme to other reverse-correlation methods using natural stimuli. The design of physiological experiments using this stimulus is in progress.

2 The Wiener series and reverse correlation

For simplicity, we will only consider systems of two inputs: $y(t) = f(x_1(t), x_2(t))$. Systems of more than two inputs (that is, driven by a stimulus of more than two pixels) follow the same mathematical form.

The Volterra series of f is given by:

$$\begin{aligned}
y(t) &= f(x_1(t), x_2(t)) \\
&= V_0 + V_1 + V_2 + \dots \\
V_0 &= k_1 + k_2 \\
V_1 &= \int k_1(\tau)x_1(t-\tau)d\tau + \int k_2(\tau)x_2(t-\tau)d\tau \\
V_2 &= \iint k_{11}(\tau_1,\tau_2)x_1(t-\tau_1)x_1(t-\tau_2)d\tau_1\tau_2 \\
&+ \iint k_{22}(\tau_1,\tau_2)x_2(t-\tau_1)x_2(t-\tau_2)d\tau_1\tau_2 \\
&+ \iint k_{12}(\tau_1,\tau_2)x_1(t-\tau_1)x_2(t-\tau_2)d\tau_1\tau_2
\end{aligned}$$

V_0 is the constant term. V_1 describes the linear behavior of the system. The kernels $k_1(\tau)$ and $k_2(\tau)$ are called the *first-order kernels*. V_2 describes the nonlinearity involving interactions between the two inputs. The kernels in V_2 are called the *second-order kernels*. There is a second-order kernel for each pair of inputs. $k_{11}(\tau_1,\tau_2)$ and $k_{22}(\tau_1,\tau_2)$ are called the *self kernels* and $k_{12}(\tau_1,\tau_2)$ is called the *cross kernel*.

In order to solve for the kernels, Wiener re-arranged the Volterra series such that the terms are orthogonal (uncorrelated) to each other, with respect to Gaussian white inputs (Wiener, 1958; Marmarelis & Naka, 1974; Marmarelis & Marmarelis, 1978).

$$\begin{aligned}
y(t) &= f(x_1(t), x_2(t)) \\
&= G_0 + G_1 + G_2 + \dots \\
G_0 &= h_1 + h_2 \\
G_1 &= \int h_1(\tau)x_1(t-\tau)d\tau + \int h_2(\tau)x_2(t-\tau)d\tau \\
G_2 &= \iint h_{11}(\tau_1,\tau_2)x_1(t-\tau_1)x_1(t-\tau_2)d\tau_1\tau_2 - P\int h_{11}(\tau,\tau)d\tau \\
&+ \iint h_{22}(\tau_1,\tau_2)x_2(t-\tau_1)x_2(t-\tau_2)d\tau_1\tau_2 - P\int h_{22}(\tau,\tau)d\tau \\
&+ \iint h_{12}(\tau_1,\tau_2)x_1(t-\tau_1)x_2(t-\tau_2)d\tau_1\tau_2
\end{aligned}$$

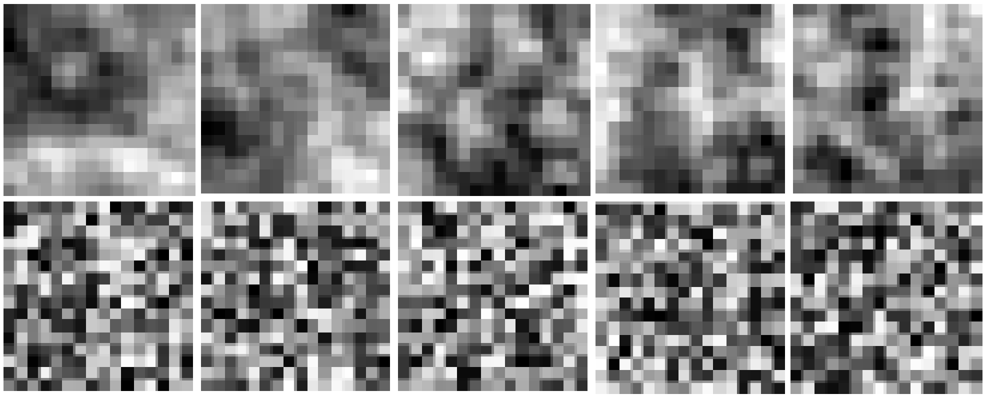

Figure 1: The stimuli (vector x, upper row) are synthesized by linearly transforming a white noise cause (vector s, lower row) via a linear generative model: $x = A\ s$. Matrix A is learned from samples of natural images.

where $x_1(t)$ and $x_2(t)$ are independent Gaussian white inputs, with equal power (or variance) P. The kernels are called the *Wiener kernels*.

Lee and Schetzen (Lee & Schetzen, 1965) showed that the Wiener kernels can be calculated by cross-correlating the neuronal response $y(t)$ with the inputs. For example, the first-order kernel $h_1(\tau)$ can be calculated from $\langle y(t)x_1(t-\tau)\rangle$, self-kernel $h_{11}(\tau_1, \tau_2)$ from $\langle y(t)x_1(t-\tau_1)x_1(t-\tau_2)\rangle$, and the cross-kernel $h_{12}(\tau_1, \tau_2)$ from $\langle y(t)t_1(y-\tau_1)x_2(t-\tau_2)\rangle^1$. See (Marmarelis & Naka, 1974; Marmarelis & Marmarelis, 1978) for details.

3 Synthesis of naturalistic noise and kernel calculation

3.1 The synthesis model

Instead of using Gaussian white noise for reverse correlation, we can linearly transform white noise such that the the statistics of the transformed images approximate those of natural images. This should produce a better stimulus for higher-order visual neurons since it contains more features found in nature.

More specifically, let the stimulus $x(t) = (x_1(t) \ldots x_n(t))^T$ be synthesized by:

$$x(t) = A\ s(t)$$

$$\begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} s_1(t) \\ \vdots \\ s_n(t) \end{bmatrix}$$

where $s(t) = (s_1(t) \ldots s_n(t))^T$ is white. The vector $s(t)$ is called the *cause* of the stimulus $x(t)$. The constant matrix A can be learned from patches of natural images by various algorithms, for example, Infomax Independent Component Analysis (Infomax ICA) (Bell & Sejnowski, 1995, 1996). In this case, the causes $s_1(t) \ldots s_n(t)$ are required to be Laplacian distributed.

---

[1] $\langle\ \rangle$ denotes expectation over t

Examples of the synthesized stimuli are illustrated in Figure 1. Visual features that occur very rarely in white noise, such as localized edges, corners, curves, and sometimes closed contours, are much more common after the A transformation.

Using linear generative models to synthesize stimulis for physiological experiments was also suggested in (Olshausen, 2001).

3.2 Kernel calculation

To calculate the kernels, one can follow Wiener and orthogonalize the Volterra series with respect to the distribution of the new stimulus, instead of Gaussian white noise. Here we provide a much simpler solution, using a trick that is similar to the treatment of non-white inputs in (Lee & Schetzen, 1965).

The derivation is illustrated in Figure 2. Instead of directly solving for the kernels of system f, we consider system f', which is formed by combining system f with the linear generative model: $f' = f \circ A$ (Figure 1b). The kernels of system f' can be calculated by the standard cross-correlation method, because its input $s(t)$ is white[2]. After f' is identified, we consider a new system f'', formed by combining f' with the inverse of the generative model: $f'' = f' \circ A^{-1}$ (Figure 1c). The kernels of system f'' can be easily obtained by plugging $s(t) = A^{-1}x(t)$ into the kernels of f', and expressing the kernels as functions of $x(t)$ instead of $s(t)$. But since $f'' = f' \circ A^{-1} = f \circ A \circ A^{-1} = f$, system f'' is equivalent to f. We therefore calculate kernels of f by transforming the kernels of f'.

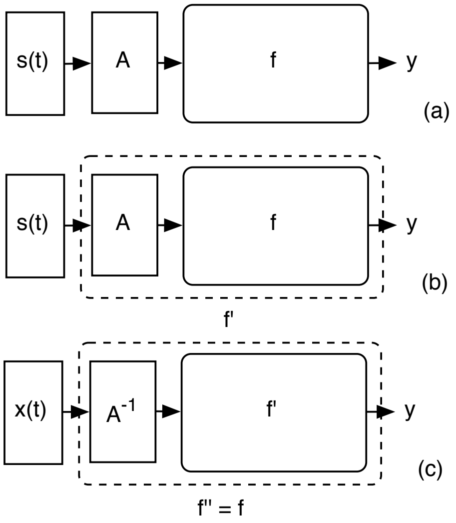

Figure 2: The derivation of formulas for kernels. (a) In order to calculate the kernels of system f, we form the system f' as in (b). Kernels of system f' can be obtained by the standard cross-correlation method because the input s is white. After the kernels of f' are identified, we construct system f'' as in (c). The kernels of system f'' can be obtained by transforming the kernels of f'. But since f'' is equivalent to f, this yields the kernels that we wanted in the first place.

---

[2] Note that $s(t)$ is Laplacian distributed, instead of Gaussian distributed. Kernels higher than the first order need to be calculated according to (Klein & Yasui, 1979; Klein, 1987).

Let $\phi_1(\tau) \dots \phi_n(\tau)$ be the first-order kernels of f', obtained by cross-correlating system response with white noise $s(t)$. The first-order kernels of the original system f, $h_1(\tau) \dots h_2(\tau)$, are simply

$$\begin{bmatrix} h_1(\tau) \\ \vdots \\ h_n(\tau) \end{bmatrix} = A^{-t} \begin{bmatrix} \phi_1(\tau) \\ \vdots \\ \phi_n(\tau) \end{bmatrix}$$

The second-order kernels of system f,

$$h_{ij}(\tau_1, \tau_2), \qquad i, j = 1 \dots n, \quad h_{ij}(\tau_1, \tau_2) = h_{ji}(\tau_1, \tau_2)$$

can be calculated from $\phi_{ij}(\tau_1, \tau_2)$, kernels of system f', by the following equation:

$$\begin{bmatrix} c_{11}h_{11}(\tau_1,\tau_2) & \dots & c_{1n}h_{1n}(\tau_1,\tau_2) \\ \vdots & & \vdots \\ c_{n1}h_{n1}(\tau_1,\tau_2) & \dots & c_{nn}h_{nn}(\tau_1,\tau_2) \end{bmatrix} = A^{-t} \begin{bmatrix} c_{11}\phi_{11}(\tau_1,\tau_2) & \dots & c_{1n}\phi_{1n}(\tau_1,\tau_2) \\ \vdots & & \vdots \\ c_{n1}\phi_{n1}(\tau_1,\tau_2) & \dots & c_{nn}\phi_{nn}(\tau_1,\tau_2) \end{bmatrix} A^{-1}$$

where $c_{ij} = 1$ if $i = j$, and $c_{ij} = \frac{1}{2}$ if $i \neq j$. Higher order kernels can also be derived.

3.3 Notes on implementation

First, since training ICA on natural images usually produces a matrix whose row vectors resemble gabor functions(Bell & Sejnowski, 1996), we can construct matrix A directly as rows of gabor patches. This is similar to the synthesis model in (Field, 1994), and has the advantage of not being biased by the particular set of images used for training. From this point of view, the synthesized stimulus is a random mixture of edges.

Second, the synthesis method described so far generates each frame independently. If ICA is trained on movies, we can synthesize image sequences with realistic motion (van Hateren & Ruderman, 1998; Olshausen, 2001). The frames in the sequences are correlated, but described by independent coefficients. The spatiotemporal kernels of neurons with respect to synthesized movies can also be derived by the same procedure.

4 Comparison to related work

To overcome the limitations of using Gaussian white noise for reverse correlation, researchers have recently started to use natural stimuli (Theunissen et al. (2000) in the auditory domain, and Ringach et al. (2002) in vision). They found RF features that were not revealed by white noise. The analysis strategy of these methods is to model receptive fields as linears filter with zero memory, and solve for the mean square error solution by regression (DiCarlo et al., 1998) or the recursive least square algorithm (Ringach et al., 2002). This involves estimating and inverting the spatial autocorrelation matrix of the stimulus.

The advantages of our approach using synthesized stimulus are:

- Dealing with natural images usually requires a large amount of memory and storage. In our method, unlimited number of frames can be generated on demand, once the synthesis matrix A is learned. Kernel calculation is also easier.
- In our method, all the statistics about the stimulus is contained in the matrix A, allowing us to formulate reverse correlation in terms of the Wiener series and derive formulas for higher order kernels, which can be important for studying the

non-linear behavior of neurons (Szulborski & Palmer, 1990). Higher order kernels for natural images are much more difficult to derive, due to their complicated (and largely unknown) statistical structure. The existing regression methods for natural image reverse correlation assume linearity and do not allow the calculation of higher order kernels.

- The synthesis model is motivated by the redundancy reduction theory of the early visual code (Barlow, 1961; Field, 1994; Olshausen & Field, 1996; Bell & Sejnowski, 1996), which states that the goal of early visual code is to transform the retinal representations of natural images to an independent, sparse code. If this theory is to be taken literally, the computation of the early visual system is essentially A^{-1}, and the synthesized stimulus $x(t)$ is represented as $s(t)$ by the first-order system (the primary visual cortex). Under this assumption, second-order neurons are receiving (Laplacian distributed) white noise stimuli. The kernels ϕ's can therefore be interpreted as the kernels of higher-order systems with respect to cortical codes, instead of retinal codes. This can be useful for interpreting the non-linear behavior of neurons(Hyvärinen & Hoyer, 2000; Hoyer & Hyvärinen, 2002)

5 Discussion

We have shown how to easily derive kernels for a specific form of naturalistic noise. As this stimulus has more of the features of natural stimulation, it should more strongly activate visual neurons and allow us to more efficiently explore receptive fields.

We are currently designing physiological experiments to test this procedure on simple and complex cells in the primary visual cortex of squirrels. Specifically,

- We will calculate first-order kernels using white noise, synthesized naturalistic noise, and natural images, and compare the quality of the receptive field maps.
- Examine if second-order kernels can be reliably calculated, and see if they help to predict the behavior of neurons.
- Analyze the relationship between h's (kernels with respect to retinal code) and ϕ's (kernels with respect to cortical code, under the whitening hypothesis), and examine if the coding hypothesis helps us understand the structure of the complex cells.

References

Barlow, H. B. (1961). The coding of sensory messages. In *Current problems in animal behavior.* Cambridge: Cambridge University Press.

Bell, A. J., & Sejnowski, T. J. (1995). An information maximisation approach to blind separation and blind deconvolution. *Neural Computation*, *7*(6), 1129–1159.

Bell, A. J., & Sejnowski, T. J. (1996). Edges are the "independent components" of natural scenes. *Advances in Neural Information Processing Systems*, *9*.

DiCarlo, J. J., Johnson, K. O., & Hsiao, S. S. (1998). Structure of receptive fields in area 3b of primary somatosensory cortex in the alert monkey. *Journal of Neuroscience*, *18*, 2626–2645.

Field, D. J. (1994). What is the goal of sensory coding. *Neural Computation*, *6*(4), 559–601.

Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. E., & Van Essen, D. C. (1996). Neural responses to polar, hyperbolic, and cartesian gratings in area V4 of the macaque monkey. *Journal of Neurophysiology*, *76*, 2718–2739.

Hegdé, J., & Van Essen, D. (2000). Selectivity for complex shapes in primate visual area v2. *The Journal of Neuroscience*, *20*(RGB 1 of 8).

Hoyer, P. O., & Hyvärinen, A. (2002). A multi-layer sparse coding network learns contour coding from natural images. *Vision Research*, *42*(12), 1593-1605.

Hyvärinen, A., & Hoyer, P. O. (2000). Emergence of phase and shift invariant features by decomposition of natural images into independent feature subspaces. *Neural Computation*, *12*(7), 1705-1720.

Klein, S. A. (1987). Relationship between kernels measured with different stimuli. In P. Z. Marmarelis (Ed.), *Advanced methods of physiological system modeling* (Vol. I, pp. 278–288). Plenum Press.

Klein, S. A., & Yasui, S. (1979). Nonlinear systems analysis with non-gaussian white stimuli: general basis functionals and kernels. *IEEE Transactions on Information Theory*, *IT-25*(4).

Lee, Y. W., & Schetzen, M. (1965). Measurement of the wiener kernels of a non-linear system by cross-correlation. *International Journal of Control*(2), 234–254.

Marmarelis, P. Z., & Marmarelis, V. Z. (1978). *Analysis of physiological systems: the white noise approach.* Plenum Press.

Marmarelis, P. Z., & Naka, K. (1974). Identification of multi-input biological systems. *IEEE Transactions on Biomedical Engineering*, *BME–21*(2).

Olshausen, B. A. (2001). Sparse coding of time-varying natural images. In *Society for neuroscience abstracts* (Vol. 27). Society for Neuroscience.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*, 607–609.

Ringach, D. L., Hawken, M. J., & Shapley, R. (2002). Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *Journal of Vision*, *2*(1), 12-24.

Szulborski, R. G., & Palmer, L. A. (1990). The two-dimensional spatial structure of nonlinear subunits in the receptive fields of complex cells. *Vision Research*, *30*(2), 249–254.

Theunissen, F. E., Sen, K., & Doupe, A. J. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *The Journal of Neuroscience*, *20*(6), 2315-2331.

van Hateren, J. H., & Ruderman, D. L. (1998). Independent component analysis of image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proc.R.Soc.Lond. B*(265), 2315–2320.

Wiener, N. (1958). *Nonlinear problems in random theory.* The M.I.T. Press.