M2.2

The Effect of Receptor Signal-to-Noise Levels on Optimal Filtering in a Sensory System

Mark D. Plumbley and Frank Fallside

Cambridge University Engineering Department,
Trumpington Street, Cambridge, CB2 1PZ, England.

ABSTRACT

In this paper, we consider image filtering (temporal and spatial) in a neural system for transmitting images through a limited capacity channel, in the case of a noisy image at the receptors. We use an extension of Shannon's formula for capacity of a Gaussian channel to determine the optimum filter to be used. For realistic image statistics, we show that the bandwidth of this filter is self-limiting, and it has a high frequency boost that disappears at low signal levels. This behaviour is mirrored in biological retinas.

1 INTRODUCTION

Many biological sensory systems are 'remote' from their higher processing centres in the brain. Signals from the human eye, for example, pass through the optic nerve via the Lateral Geniculate Nucleus to the brain[2]. The number of neural channels in the optic nerve is limited, and the energy used to drive these channels is a resource that an organism should not waste unnecessarily. Thus the sensory signals passing out of the remote sensor (e.g. the retina) down the limited-resource channel (the optic nerve) should be efficiently coded so as to obtain the maximum information capacity as possible (see Laughlin's review article[3] for a similar approach to the fly visual system). Related problems such as image coding are addressed from an engineering viewpoint in communication theory, but the principles are the same—maximise the amount of information transmitted for a given cost.

According to the well-know result from Shannon[6], the information capacity of a Gaussian channel with limited signal power in the channel is maximised if the signal power $S(f)$ and noise power $N(f)$ sum to a constant for $0 < f < B$, where B is the available channel bandwidth. Thus for white Gaussian noise ($N(f) = N_o$) a 'whitening' filter (such as a linear predictive coder) should be applied to a signal before transmission for the maximum channel capacity to be obtained.

Unfortunately, this is only strictly valid when the signal at the receptors is noise-free. If the input signal is itself noisy, a whitening filter may amplify noisy parts of the signal to the point where a significant amount of the available power is being used to transmit a low-quality, noisy signal. In this case, it would be more efficient to use more of the available signal power to transmit lower-noise parts of the signal spectrum, even though it may appear to be at a higher cost.

In a sensory system, any received signal is likely to be corrupted with a certain amount of noise. For example, a visual signal will be corrupted by shot noise due to the individual photons arriving at the receptors. The signals transmitted by spiking neurons include the shot-like noise induced by the spike train itself. To make optimal use of available resources, a sensory system should attempt to take account of this noise in the coding scheme used at the receptor. We might hope that the filtering performed in the eye, for example, is such as to maximise the information throughput under these conditions.

1.1 Outline

First we review Shannon's original result, with a noise-free signal at the receptor, leading to a whitening filter as the optimal for a power-limited channel with additive white Gaussian noise.

We then develop an optimal filter for the case of noise on the signal at the receptor as well as the channel, and show that this leads to a quadratic in the filter power gain.

2 NOISE-FREE RECEPTORS

Consider the noise-free receptor system shown in Fig. 1. $S_r(f)$

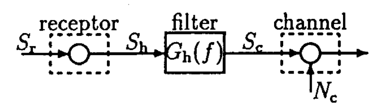

Figure 1: Filtering of noise-free signal before transmission

is the signal power at the receptor, $S_h(f) = S_r(f)$ is the signal power before the filter, $G_h(f)$ is the power gain of the filter, $S_c(f)$ is the signal power in the channel, and $N_c(f)$ is the noise power in the channel. We assume a channel bandwidth B.

Following Shannon [6], the capacity of a channel with signal spectrum $S(f)$, noise spectrum $N(f)$ and bandwidth B is given by

$$C_T = \int_0^B C(f)\,df$$

where

$$C(f) = \log\left(\frac{S(f) + N(f)}{N(f)}\right).$$

In our case, $S(f) = S_c(f) = G_h(f)S_r(f)$ and $N(f) = N_c(f)$. To transmit this information requires total power

$$P_T = \int_0^B S_c(f)\,df$$

thus to maximise C_T for a given power P_T we should maximise

$$J = \int_0^B C(f) - \lambda S_c(f)\,df$$

(where $\lambda \in \mathbf{R}$ is a Lagrange multiplier) by suitable choice of the filter power gain $G_h(f) \geq 0$. This leads to the condition [6]:

$$G_h(f)S_r(f) + N_c(f) = \gamma \qquad (1)$$

CH2977-7/91/0000-2321 $1.00 © 1991 IEEE

with $\gamma = 1/\lambda$.

In the case of white noise in the channel ($N_c(f) = N_o$), $G_h(f)S_r(f)$ should be constant. With this simplification, then, the filter should be a whitening filter, so that $G_h(f) \propto S_r(f)^{-1}$ within the bandwidth limit $0 \leq f \leq B$, with $G_h(f) = 0$ for $f > B$.

3 NOISE AT THE RECEPTORS

Consider a simple receptor system as shown in Fig. 2. The

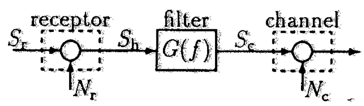

Figure 2: Receptor filtering before transmission

receptor picks up the signal of interest, with power spectrum $S_r(f)$, together with additive Gaussian noise $N_r(f)$ (which includes both external noise and noise internal to the receptor), to give the output $S_h(f) = S_r(f) + N_r(f)$ from the receptor. This is then passed through the filter with power spectral gain $G_h(f)$ to give the input to the neuron channel with power spectrum $S_c(f) = G_h(f)S_h(f)$. Finally, Gaussian noise $N_c(f)$ is added to the signal as it is transmitted through the channel. No bandwidth limit is assumed.

We assume that the only signal of interest arises from $S_r(f)$. Thus the 'signal' at the output of the system is

$$S(f) = G_h(f)S_r(f)$$

and the 'noise' is

$$N(f) = G_h(f)N_r(f) + N_c(f).$$

With these expressions for $S(f)$ and $N(f)$, as before, we maximise

$$J = \int_0^\infty C(f) - \lambda S_c(f)\,df$$

with

$$C(f) = \log\left(\frac{S(f) + N(f)}{N(f)}\right)$$

by suitable choice of the filter power spectral gain $G_h(f) \geq 0$. From the calculus of variations, the condition for this is

$$\frac{S_r(f)N_c(f)}{N(f)(S(f) + N(f))} - \lambda(S_r(f) + N_r(f)) = 0$$

or, dropping the '(f)'s and re-arranging,

$$(S_cN_r + N_cS_h)(S_c + N_c) - \gamma S_r N_c = 0 \tag{2}$$

where $S_h = S_r + N_r$, $S_c = G_hS_h$ and $\gamma = 1/\lambda$. If $N_r = 0$ (no input noise), we get

$$(S_c + N_c) = \gamma$$

as in (1). Otherwise, introducing signal-to-noise ratio terms $R_c = S_c/N_c$ and $R_r = S_r/N_r$, we have

$$(R_c + R_r + 1)(R_c + 1) - \frac{\gamma}{N_c}R_r = 0 \tag{3}$$

which leads to the solution

$$R_c = \frac{1}{2}\left(\sqrt{R_r^2 + \frac{4\gamma}{N_c}R_r} - (R_r + 2)\right) \tag{4}$$

from which G_h can be determined. The solution is valid (i.e. G_h non-zero) whenever

$$R_r > \frac{1}{(\gamma/N_c) - 1}. \tag{5}$$

Therefore, for white channel noise ($N_c(f) = N_{c_0}$), the optimal filter response is zero for all frequencies where the receptor signal-to-noise ratio R_r is less than a cut-off value $R_{r_{min}} = 1/(\gamma/N_{c_0} - 1)$. Fig. 3 shows a typical solution for $G_h(f)$, for white receptor and channel noise.

3.1 Bounding curves

For very low input noise, we would expect the solution given above to approach Shannon's result (1). It is also instructive to explore the other bounding curves, to give us a qualitative idea of the the general solution (3).

For the bounds, we shall consider approximations to the curve at extreme values of the receptor signal-to-noise ratio R_r, and the channel signal-to-noise ratio R_c. We would expect many real sensory signals to have high receptor signal-to-noise ratio (SNR) at low temporal and spatial frequencies, so we shall start with high receptor SNR.

1. $R_r \gg 1$ and $R_r \gg R_c$ (very high receptor SNR). This is the condition we might expect in a sensory system under normal operating conditions, at relatively low frequencies. In this case, the limiting noise is the channel noise N_c.

   Rearranging (3) we get

$$\frac{R_c^2}{R_r} + \left(1 + \frac{2}{R_r}\right)R_c + 1 + \frac{1}{R_r} = \frac{\gamma}{N_c} \tag{6}$$

$$S_c(1 + O(1/R_r) + O(R_c/R_r)) = \gamma - N_c(1 - O(1/R_r))$$

$$S_c = \gamma - N_c - N_cO(1/T_r) - (\gamma - N_c)(O(1/R_r) + O(T_c/T_r))$$

   giving an upper bound for R_c of

$$R_c + 1 \leq \gamma/N_c. \tag{7}$$

   In other words $S_c + N_c \approx \gamma$, confirming the analysis for the noise-free case.

   For white channel noise N_c, as for the noise-free case the channel signal power S_c should be constant, leading to the filter power gain

$$G_h \propto R_r^{-1} \tag{8}$$

   which can be seen on the first part of the curve in Fig. 3.

2. $R_c \gg R_r \gg 1$ (high receptor SNR, very high channel SNR). In this case, the receptor noise N_r starts to become significant. However, the major proportion of the signal S_c in the channel is still due to the filtered receptor signal G_hS_r, rather than noise.

   Again rearranging (3) we get:

$$1 + \frac{R_r + 2}{R_c} + \frac{R_r + 1}{R_c^2} = \frac{1}{R_c^2}\left(\frac{\gamma}{N_c}R_r\right)$$

$$R_c^2(1 + O(R_r/R_c) + O(1/R_c)) = \frac{\gamma}{N_c}R_r$$

$$R_c^2 = \frac{\gamma}{N_c}R_r(O(R_r/R_c) + O(1/R_c))$$

$$R_c^2 \leq \frac{\gamma}{N_c}R_r. \tag{9}$$

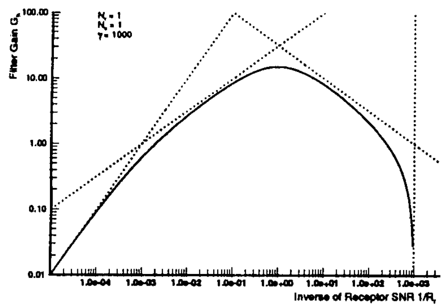

Figure 3: Typical boundaries for general solution

Substituting for R_c we get

$$(G_h N_r R_r (1 + 1/R_r))^2 \leq \gamma N_c R_r \tag{10}$$

$$\begin{aligned}(G_h N_r R_r)^2 &\leq \gamma N_c R_r (1 - O(1/R_r)) \\ &\leq \gamma N_c R_r\end{aligned}$$

or

$$G_h \leq N_r^{-1} (\gamma N_c / R_r)^{1/2}.$$

Thus in the case of white receptor and channel noise, we have approximately

$$G_h \propto R_r^{-1/2}. \tag{11}$$

This mode can be seen on the second part of Fig. 3. This is no longer a whitening filter: as the signal at the receptor reduces, the gain of the amplifier does increase a certain amount, but not as much as in the very low noise case.

3. $R_c \gg 1 \gg R_r$ (low receptor SNR, high channel SNR). This is similar to the previous case, but the 'signal' in the channel S_c is now mainly composed of noise at the receptor, rather than true signal S_r.

   Substituting for R_c in (9) we get

$$(G_h N_r (1 + O(R_r)))^2 \leq \gamma N_c R_r$$

so

$$(G_h N_r)^2 \leq \gamma N_c R_r \tag{12}$$

or

$$G_h \leq N_r^{-1} (\gamma N_c R_r)^{1/2}.$$

Thus for white noise N_c and N_r

$$G_h \propto R_r^{1/2} \tag{13}$$

The filter is no longer attempting to whiten the signal: to do so would be wasteful, since the 'signal' S_c in the channel is mainly due to the filtered receptor noise $G_h N_r$ rather than any useful signal. However, transmitting a small amount of very low quality signal such as this is still worthwhile, since the power cost is so low compared with the cost to transmit more of a higher quality signal in a frequency band which is already used significantly.

4. $R_r \ll 1$ and $R_c \ll 1$ (low receptor and channel SNR). This is really a limiting case, just before the optimal filter cuts off completely at the lower limit $R_{r_{min}} = 1/(\gamma/N_c - 1)$.

   Again from (3) we get:

$$(R_c + 1)^2 \approx \frac{\gamma}{N_c} R_r$$

$$R_c \approx \frac{1}{2} \left(\frac{\gamma}{N_c} R_r - 1 \right) \tag{14}$$

so, for the case of white noise,

$$G_h \propto \frac{\gamma R_r}{N_c} - 1 \tag{15}$$

Of course, from (5) we see that $R_c = 0$ for $R_r < 1/((\gamma/N_c) - 1)$ so cases 2, 3 and 4 above may not be approached unless $\gamma \gg N_c$, allowing a low threshold for R_r.

3.2 Typical Filter

Fig. 4 shows root power spectra of a typical optimal filter for

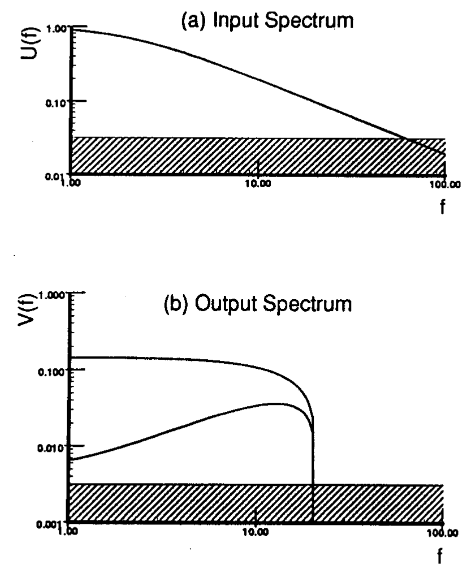

Figure 4: Optimum filter with low input noise.

signals with power spectral density of $S(f) = 1/(1 + (f/f_c)^2)$, with $f_c = 2$. The filter response is the lower curve of Fig. 4(b). In this case, the input SNR threshold is set relatively high, i.e. $R_{r_{min}} \approx 10$ with $\gamma \approx 1.4 N_c$, so there is a relatively swift transition from case 1 above to the cut-off.

In contrast, Fig. 5(a) compares this example with the equivalent filter for a factor of 100 reduction in SNR, and realignment of the operating point (to $\gamma \approx 25 N_c$). It can be seen that the high-frequency boost has all but disappeared, giving rise to the same qualitative effect as that observed in biological visual systems (Fig. 5(b)).

Laughlin has found that this effect in the fly visual system can be predicted with surprising agreement using a Linear Predictive Coding approach [3]. Essentially, this uses the Noise-Free

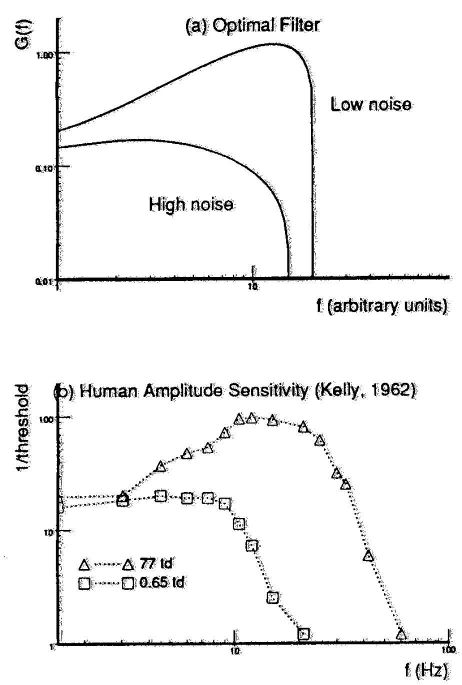

Figure 5: Optimal filter change with receptor noise level. Data in (b) adapted from [1]

receptor assumption of section 2, leading to a whitening filter as the optimum, but the receptor noise N_r is included in the receptor signal. It would be interesting to see how quantitative predictions from the optimal filter presented here differ from those.

4 DISCUSSION

Although the optimal filter in this paper is purely linear, the two noise sources lead to a quadratic in the filter coefficients. Of course, for a real biological system the situation will be more complex: non-white noise; signals transmitted using a pulse system instead of simple levels; the 'power' limitation not simply the mean square amplitude, to give some examples. However, we would expect the same principles to apply, and it would be interesting to try a quantitative analysis for a particular biological system, such as Lauglin's analysis for the fly visual system[3].

The analysis used for the filter in this paper is not restricted to filters, either spatial or temporal. It could be used instead to determine optimal gain parameters for any set of independent channels, where the total power is limited, but where it is not possible or desirable to move information from one channel to another. Simple linear filters cannot move signal frequency components to other frequencies: they can only scale the frequency components. An alternative is Principal Component Analysis (PCA), which can be used for data reduction by ignoring those components of lowest variance[7]. Earlier articles have already demonstrated a relationship between PCA and information theory when only noise on the inputs is considered[4, 5]. According to the analysis in this paper, if the noise on the input channels and the output channels is known or can be estimated, the outputs from the PCA should be scaled so that (3) holds.

For higher levels in the visual system, such as simple cells in the visual cortex[2] it is unlikely that linear analysis considering only second order statistics will be sufficient. However, the same principles may hold, so this may be an interesting avenue to pursue in future research.

5 CONCLUSION

We have considered the problem of optimal filtering of a noisy signal through a noisy channel, using a linear filter. A simple analysis leads to a quadratic for the filter gain G_h in terms of receptor signal-to-noise ratio (SNR) channel noise, and a Lagrange multiplier which fixes the operating point. For very high receptor SNR, this filter is close to Shannon's optimal 'whitening filter', or a Linear Predictive Coder

For realistic signal statistics, the bandwidth of this filter is self-limiting, with cut-off above a certain frequency as the receptor SNR drops below a critical value determined by the operating point and the channel noise level. In addition, as the overall SNR of the input signal is reduced, with a corresponding change in the operating point, the high-frequency boost evident for the high SNR case all but disappears.

Acknowledgements

The authors would like to thank the member of the Speech Vision and Robotics group at CUED for many stimulating discussions.

One of the authors (MDP) was supported by a grant from the UK Science and Engineering Research Council.

REFERENCES

[1] D. H. Kelly. Information capacity of a single retinal channel. *IRE Transactions on Information Theory*, IT-8:221–226, 1962.

[2] S. W. Kuffler, J. G. Nicholls, and A. R. Martin. *From Neuron to Brain: A Cellular Approach to the Function of the Nervous System*. Sinauer Associates Inc., Sunderland, MA, second edition, 1984.

[3] Simon B. Laughlin. Form and function in retinal processing. *Trends in NeuroSciences*, 10:478–483, 1987.

[4] Ralph Linsker. Self-organization in a perceptual network. *IEEE Computer*, 21(3):105–117, March 1988.

[5] Mark D. Plumbley and Frank Fallside. An information-theoretic approach to unsupervised connectionist models. In David Touretzky, Geoffrey Hinton, and Terrence Sejnowski, editors, *Proceedings of the 1988 Connectionist Models Summer School*, pages 239–245, San Mateo, CA., 1988. Morgan-Kaufmann.

[6] Claude E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37:10–21, 1949.

[7] Satosi Watanabe. *Pattern Recognition: Human and Mechanical*. John Wiley & Sons, New York, 1985.