View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by INRIA a CCSD electronic archive server

Another look at the retina as an image dithered scalar quantizer

Khaled Masmoudi, Marc Antonini, Pierre Kornprobst

To cite this version:

Khaled Masmoudi, Marc Antonini, Pierre Kornprobst. Another look at the retina as an image dithered scalar quantizer. International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2010), Apr 2010, Desenzano del Garda, Italy. paper 21. hal-00481337

HAL Id: hal-00481337

https://hal.archives-ouvertes.fr/hal-00481337

Submitted on 6 May 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ANOTHER LOOK AT RETINA AS AN IMAGE DITHERED SCALAR QUANTIZER

Khaled Masmoudi, Marc Antonini

Univ. Nice Sophia Antipolis-CNRS-I3S
Sophia Antipolis, France
kmasmoud, am@i3s.unice.fr

Pierre Kornprobst

INRIA-NeuroMathComp
Sophia Antipolis, France
pierre.kornprobst@sophia.inria.fr

ABSTRACT

We explore, in this paper, the behavior of the mammalians retina considered as an analog-to-digital converter for the incoming light stimuli. This work extends our previous effort towards combining results in neurosciences with image processing techniques [1]. We base our study on a biologically realistic model that reproduces the neural code as generated by the retina. The neural code, that we consider here, consists of non-deterministic temporal sequences of uniformly shaped electrical impulses, also termed as *spikes*. We describe, starting from this spike-based code, a dynamic quantization scheme that relies on the so-called *rate coding* hypothesis. We, then, propose a possible decoding procedure. This yields an original quantizing/de-quantizing system which evolves dynamically from coarse to fine, and from uniform to non-uniform. Furthermore, we emit a possible interpretation for the non-determinism observed in the spike timings. In order to do this, we implement a three-staged processing system mapping the anatomical architecture of the retina. We, then, model the retinal noise by a dither signal which permits us to define the retina behavior as a non-subtractive dithered quantizer. The quantizing/de-quantizing system, that we propose, offers several interesting features as time scalability as well as reconstruction error whitening and de-correlation from the input stimuli.

1. INTRODUCTION

The human visual system conveys information as a set of electrical impulses called spikes. Spikes [2] appear very early in the chain of treatment of the human visual system. At the retina level, after a chain of internal treatments, ganglion cells convert an analogous signal into a series of spikes called spike trains forming the neural code. Spikes have the same shape and amplitude and which yields a binary-like neural code.

In order to experiment the behavior of the retina as a quantizer, we implement a three-staged system based on a biologically realistic retinal model of introduced in [3]. The considered simulator is one of the most complete ones generating a spike-based output which, furthermore, successfully reproduce actual neurophysiologic recordings. The model maps the anatomical structure of the retina. This structure is strongly related to the retina functional architecture. Indeed, the retina is a succession of layers. The output of each one is the input of the following. The progression of light stimuli, from the outermost light receptors layer, to innermost ganglionic layer, involves several processing mechanisms.

The innermost ganglion cells layer of the retina emit spikes to convey information over the optic nerve [2]. By opposition, retinal cells of the outer stages do note fire spikes. As the input stimuli gets through these stages, the input signal is filtered, but still has the form of a graded continuous electrical signal. These cells are, from innermost to outermost, amacrine cells (in the inner plexiform layer), the Bipolar cells (in the outer plexiform layer), the horizontal cells and finally light receptors. Only the ganglion cells are responsible for signal discretization. In the following we focus on the cells of the three deepest retina layers, involved in the generation of the visual neural code, namely bipolar, amacrine and ganglioon cells. These cells form the main stages responsible for the shaping of the spiking retina code. The paper is organized as follows: In Section 2, we present the model of the three-staged system. Then, in Section 3,we specify the bioinspired quantization/de-quantization algorithm that we implemented. Finally, in Section 4, we explore the overall system behavior and emit the hypothesis of non subtractive dither to interpret the retinal noise.

2. A BIOLOGICALLY REALISTIC RETINA MODEL

We describe, in this section, the input/output map of the mammalians retina. In order to do this, we base our work on the biologically realistic retina model introduced in [3]. We restrain our study to the temporal behavior of the retina, thus the spatial filtering blocks are ignored in the following description. Furthermore, only the three deepest retina layers in the model are considered, as they are the main stages responsible for the shaping of the spiking retina code. Although this model does not take into account some features of the biological retina, such as lateral connections, still it renders the main biological properties of the actual retina. The three-staged simplified model, as implemented in this work, is described through Sections 2.1 to 2.3.

2.1. Bipolar cells layer: The gain control stage

Biological systems need, often, to adjust their operational range to match the input stimuli magnitude range [4]. Interestingly, fast magnitude adaptation mechanisms are largely observed in the bipolar cells. Here, the bipolar cells rescale an input current $I(t)$ to generate an output potential $V_B(t)$.
Let us consider a time duration ΔT such that $I(t)$ is constant across $[0, \Delta T[$. In the following, we define the stimulus signal:

$$I(t) = \begin{cases} I_j, \text{ if } t \in [0, \Delta T[\\ 0 \text{ otherwise }, \end{cases} \quad (1)$$

so that, we can study our system behavior in piecewise fashion. The gain control procedure, as introduced in [3], is defined by:

$$\frac{dV_B(t)}{dt} + g_B(t)V_B(t) = I(t), \quad (2)$$

where g_B represents a variable leakage term. The expression of g_B, for a potential V_B encompasses spatial filtering. As the spatial aspect of the retina behavior is ignored in the current study, the spatial filter is set to a Dirac impulse. Referring to [1], the gain control expression

is developed to get the following:

$$\frac{dV_B(t)}{dt} - g_B^0\left(e^{-\frac{t}{\tau_B}} - 1\right)V_B(t) + \frac{\lambda_B}{\tau_B}\left(\int_0^t V_B^2(t-s)e^{\frac{-s}{\tau_B}}ds\right)V_B(t) = I(t), \quad (3)$$

where g_B^0, τ_B, and λ_B are constant scalar parameters. The output bipolar voltage V_B, as we defined it, is the input of the subsequent inner plexiform layer stage (IPL).

2.2. Inner plexiform layer: The non-linear rectification stage

We consider the signal, of voltage V_B, as generated by the bipolar cells of the retina. This current is subject to a non-linear rectification by the amacrine cells in the IPL. The output of the IPL is a corrected current I_G. A biologically realistic model of this rectification [3] is given by:

$$I_G(t) = N\left(\varepsilon\, T_{w_A,\tau_A}(t) * V_B(t)\right), \quad (4)$$

where T_{w_A,τ_A} is a linear transient filter (see [3] for a formal definition), ε, w_A, and τ_A are constant scalar parameters, and N is a function expressed as follows:

$$N(v) = \begin{cases} \dfrac{I_A^2}{I_A - \lambda_A(v - V_A)}, \text{ if } v < V_A \\ I_A + \lambda_A(v - V_A), \text{ if } v \geqslant V_A, \end{cases}$$

where I_A, V_A, and λ_A are constant scalar parameters. Developing (4), we get the following expression for I_G (see details in [1]):

$$I_G(t) = N\left(\varepsilon\left(V_B(t) - \frac{w_A}{\tau_A}\int_0^t V_B(t-s)e^{\frac{-s}{\tau_A}}ds\right)\right),$$

I_G is the current input of the last retina stage, namely the ganglionic layer, which produces the neural code of the retina.

2.3. Ganglion cells layer: The spike generation stage

The ganglionic layer is the deepest one tiling the retina. The ganglion cells are the neurons that generate the spiking output of the retina. A formalization for spike generator neurons in the retina is proposed in [3]. The model chosen is the widely used noisy leaky integrate and fire (nLIF) [5]. $I_G(t)$ is the input stimulus of this spike generator layer, and $V(t)$ is its output voltage. We study $V_G(T)$ behavior in the time bin $[0, \Delta T[$, which amounts studying the spike emission timings $(T_i)_{i\geqslant 0}$. $(T_i)_{i\geqslant 0}$ are defined by the following:

$$\begin{cases} V(T_i) = \delta, \forall\, i \geqslant 0,\ T_i \in [0, \Delta T[, \\ V(t) = V_R, \forall\, i \geqslant 0,\ \forall\, t \in [T_i, T_i + T_{ref}[. \end{cases} \quad (5)$$

where δ is the integration threshold of the neuron, and T_{ref} its refractory time. In the following, the refractory time will be neglected as $T_{ref} \ll \Delta T$. Whenever the voltage V reaches δ, the neuron fires a spike, then reinitializes its voltage to V_R, the reset potential. Once the spiking mechanism is specified (cf. (5)), the model defines the behavior of $V_G(t)$ in the time bin $[T_i, T_{i+1}[$, as $V_G(t)$ obeys to the following differential equation:

$$c_G\frac{dV(t)}{dt} + g_G V(t) = I_G(t) + \eta(t), \forall\, t \in [T_i, T_{i+1}[, \quad (6)$$

where g_G is a constant conductance, c_G is a constant capacitance, and η is a random noise which will be further discussed in Section 4.2. Then, solving (6), we get:

$$V_G(t) = \left(\frac{1}{c_G}\int_{T_i}^t (I_G + \eta)(s)e^{\frac{g_G(s-T_i)}{c_G}}ds + V_R\right)e^{-\frac{g_G(t-T_i)}{c_G}}, \quad (7)$$

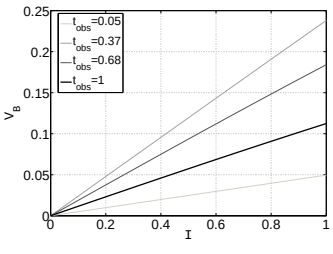


Fig. 1. $V_B(I)$: A one-to-one map associating each input current I to a bipolar output potential V_B. Maps are shown for different observation durations t_{obs}, ranging from $t_{obs} = \Delta T$ (thick line) to $t_{obs} = \frac{\Delta T}{20}$ (thin line).

In this section, we specified the model transform that leads to the generation of spikes, here restricted to the time transform. In Section 3, we introduce a bio-plausible coding scheme, and specify the corresponding decoding procedure.

3. A QUANTIZATION/DE-QUANTIZATION ALGORITHM BASED ON RATE CODING

In Section 2, we presented a biologically realistic three-staged model for spike generation in the retina. Our aim, in this Section, is to specify the algorithm that we implemented in order to experiment our bioinspired quantizer. In order to do this, we first study each stage transfer function separately, then we propose a possible decoding process to recover the initial input.

3.1. Coding pathway

An interesting feature, that we emphasize in this model, is its dynamics as it involves time t. Our approach is to study each stage, for a given observation time $t = t_{obs}$, then explore how this behavior evolves as t_{obs} varies.

3.1.1. The gain control stage

For each given stimulus maximum value I_j (cf. (1)), we solve the differential equation in (3) using the Runge-Kutta solver [6]. Examples of resulting solutions $V_B(t, I_j)$, for different values of I_j, are shown in [1]. Thus, we estimate $V_B(t_{obs}, I)$ for all possible values of I_j. We, then, infer the one-to-one map $V_B(t_{obs}, I)$ by observing the value of the potential V_B, at a given observation time t_{obs}, across the solutions $(V_B(t, I_j))_{j\in\mathbb{N}}$. This leads to the mappings shown in Figure 1. These results prove that, in the restrained domain of our model assumptions correctness, the gain control in the bipolar cells layer is linear. The linear slope $G_{t_{obs}}$ of the gain is, obviously, dependent on the observation time t_{obs}.

3.1.2. The non-linear rectification stage

After the stimulus is rescaled in the gain control stage, it gets non-linearly rectified in the second IPL stage. Computing the transform in (4), we obtain the mappings shown in the Figure 2, each one corresponding to an observation time t_{obs}. It appears that, for short observation times, input is quasi-linearly rescaled, while for longer observation times, non linearity is accentuated. This implies that, the instantaneous behavior of the IPL stage is a linear gain control, while as observation goes on, emphasize is made on the high amplitude IPL inputs.

3.1.3. The spike generation stage: the rate coding approach

The current I_G, that is generated by the IPL stage, passes through the ganglionic stage yielding a spike-based code. Here we consider the so-called *rate coding* hypothesis to interpret the coding mechanism of the retina. This is the most commonly used theory. The rate coding assumes that, in a given predefined time bin ΔT, the

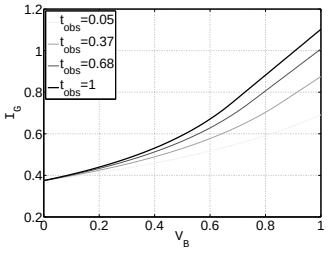


Fig. 2. $I_G(V_B)$: Non linear IPL rectification mapping each V_B value into an output current I_G. Maps are shown for different exposition durations t_{obs}, ranging from $t_{obs} = \Delta T$ (thick line) to $t_{obs} = \frac{\Delta T}{20}$ (thin line).

count of spikes convey the major part of the stimulus information [7]. Through the two preceding stages, input current I is rescaled by a static gain control slope $G_{t_{obs}}$ and corrected by a static non linear function. Thus I_G is supposed constant over the time interval $[0, t_{obs}] \subset [0, \Delta T[$. This assumption is bio-plausible for a sufficiently restrained observation time t_{obs}, then, we get:

$$V_G(t) = \left(\frac{I_G + \eta}{c_G} \int_{T_i}^{T_i+t} e^{\frac{g_G(s-T_i)}{c_G}} ds + V_R \right) e^{-\frac{g_G(t-T_i)}{c_G}},$$

where $t \in [T_i, T_{i+1}[$. $V_G(t)$ is a periodic function of time, and finding the firing timing T_{i+1}, knowing T_i, is equivalent to the deduction of the period P of $V_G(t)$. This yields the following formula for the computation of the count of emitted spikes N (see details in [1]):

$$\begin{aligned} N &= \left\lfloor \frac{\Delta T}{P} \right\rfloor \\ N &= \left\lfloor \frac{g_G \Delta T}{c_G \, log\left(1 + \frac{g_G(\delta - V_R)}{I_G + \eta - g_G \delta}\right)} \right\rfloor . \end{aligned} \quad (8)$$

We compute the function in (8) for different values of I_G. The results as shown in [1] demonstrate that the ganglion cell is a quasi-uniform scalar quantizer after a very short transitory stage around zero.

Based on a biologically realistic model of the retina, we have defined now a rate coding scheme for temporal signals. We propose a possible decoding algorithm in Section 3.2.

3.2. Decoding pathway

Our aim, in this Section, is to recover $\tilde{I}$, the estimation of the input I, knowing its rate code N, and the model parameters. Though the coding scheme in Section 3.1 is strongly related to actual biological retina behavior, we do not claim that the proposed decoding algorithm is the one that is actually employed in the visual cortex.

The decoding algorithm goes exactly the opposite way of the coding one, from the reverse ganglionic layer to the reverse gain control. First, we recover $\tilde{I}_G$, the estimation of I_G (cf. (4)). For this, we apply the following reverse mapping:

$$\tilde{I}_G = \frac{g_G(\delta - V_R)}{e^{-\frac{g_G \Delta T}{c_G N}} - 1} + g_G \delta. \quad (9)$$

Second, we recover $\tilde{V}_B$, the estimation of V_B, knowing $\tilde{I}_G$. For this, we infer the reverse IPL stage mapping through a look up table. The voltage $\tilde{V}_B$, corresponding to values of $\tilde{I}_G$ that do not match the table elements, are computed by spline interpolation.

Finally, we recover the input signal $\tilde{I}$, by the reverse bipolar gain control. As the gain control in the first coding stage is linear, the reverse gain control is a simple division.

Obviously, the recovered signal $\tilde{I}$ does not match exactly the original I. This is due to the floor operator in the spike generation mechanism (cf. (5)). The behavior of the coder/decoder system is, thus, analogous to a quantizer/ de-quantizer. We investigate the characteristic behavior of the bioinspired quantizer, that we just defined, in Section 4.

4. THE OVERALL SYSTEM BEHAVIOR REGARDLESS TO RETINAL NOISE

4.1. Case of a noiseless ganglion cell quqntizer

Let us cascade the three layers of our system. We aim at defining the characteristic behavior of the bioinspired quantizer as defined in Sections 3.1 and 3.2, and explore the evolution of it across time. It appears that, as the observation time t_{obs} increases, our system goes from coarse to fine, and from uniform to non-uniform.

The refining is intuitive and confirmed by actual neurophysiologic experiments. Indeed the visual cortex perceives global aspects of the stimulus first, then as time goes acquire more information about sharp features.

Then the model quantizer is non-uniform. High magnitude signals are mapped accurately, by a small quantization step, while small magnitude signals are coarsely rendered. This is due to the non-linear rectification in the IPL stage. Indeed, this rectification compresses the dynamic range of small magnitude signals around zero and span higher ones in a linear fashion, this before the generation of spikes in the ganglion cells. This tendency to non-uniformity is accentuated as the gain control gets higher across time.

Figure 3 shows an example map of a reconstructed input $\tilde{I}$ as a function of an input I, using the bioinspired quantizer, and this at two different observation timings. Yet, telecommunication systems are

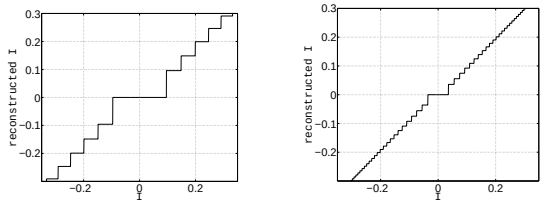


Fig. 3. Input signal/reconstructed signal characteristic: Behavior evolution of the cascaded three stages of the bioinspired model. On the left $t_{obs} = 0.075\,\Delta T$, on the right $t_{obs} = 0.27\,\Delta T$.

already implemented for dynamic signal range compression, namely compandor circuits. Companding is a technique that is widely used in telecommunication [8] making the quantization steps unequal, as the IPL stage does in our case. It is also interesting to denote that companding is preceded, for audio recordings, by a variable-gain amplifier, which is locally linear, in the same manner as the bipolar cells gain control loop described above.

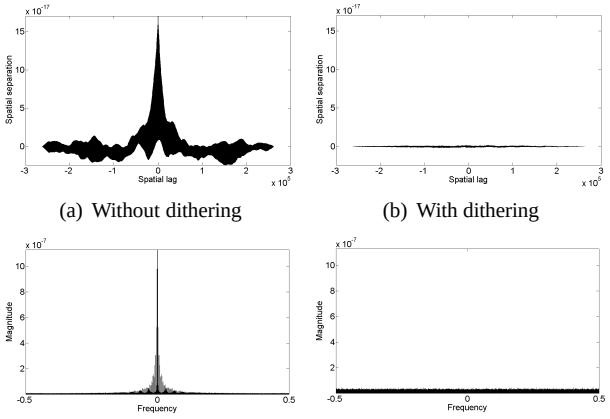
4.2. Case of a noisy ganglion cell quantizer: Is the ganglion cell a non-subtractive dithered quantizer?

An issue that neuroscientists encountered, in the understanding of the neural code, is the trial-to-trial variability of the retinal neural code. Indeed, given a single visual stimulus, spikes timings in the retina output are not exactly reproducible across trials. Here, we make the proposal that the retinal noise could be a random dither noise signal [9].

The hypotheses made to explain the phenomena underlying the spikes timings irregularity yielded two different points of view. The first is that the precise timings of individual spikes convey a large amount of information [10], and the second assumes that such a variation is a

random instantiation of a desired firing rate [7]. This implies that the spikes timings variability convey either information or noise [11].
In the following, we admit that the quantizing ganglion cell is subject to a noise, and we give a possible interpretation of its role in the stimuli coding/decoding process. Up to our knowledge, little have been done to explicit the probability distribution of such a noise. In the literature, it is generally and empirically assumed that the retinal noise η is Gaussian [12]. Thus, we can suppose that η has a triangular probability distribution function (pdf) with no loss of biological plausibility. Furthermore, we suppose that the dynamic range of η is twice wider than the quantization step of the ganglion cell. Under the restriction of these hypotheses correctness, we mapped the retinal noise η into a dither signal. As we do not subtract the dither signal in the de-quantization process, we talk about non-subtractive dithered system (NSD) [9]. Although, not intuitive, adding such a random dither signal to the input stimulus allow the quantizer to have interesting features. Mainly, the quantization error $\epsilon = (I - \tilde{I})$ and the input stimuli I are de-correlated. This feature is clearly demonstrated when computing the cross correlation between ϵ and I as shown in the Figures 4(a) and 4(b).

Besides, quantization error is whitened so that error is uniformly distributed over the stimulus spectrum. Figures 4(c) 4(d) show a comparison between the spectra of the ganglion cell quantizer with and without NSD.



(a) Without dithering (b) With dithering

(c) Without dither (d) With dither

Fig. 4. 4(a) 4(b) Cross correlation of the quantization error and the input stimuli. The abscissa represents the spatial lag and the ordinate the cross correlation magnitude. 4(c) 4(d)Noise whitening using a dithered quantizing ganglion cell: A comparison of reconstruction error spectra between non-dithered and dithered quantizing ganglion cell. The test image is Lena. The observation time is $t_{obs} = 55ms$.

The whitening and de-correlation features engenders a greater reconstruction error in terms of mean squared error (see [9] for further details). Though, the visual quality of the reconstruction $\tilde{I}$ is better when using a dithered system. Figure 5 shows the great impact of an NSD ganglion cell quantizer when compared to a non-dithered one, for the same observation time t_obs.

5. DISCUSSION

We presented a bioinspired quantizer/de-quantizer mapping the retina behavior. The model of the retina that we adopted, though restrained to its temporal aspect, reproduces many mechanisms involved in the actual biological system. Our quantizer behavior evolves dynamically, and thus, it permits scalability as it goes from coarse to fine

(a) Without dither (b) With dither

Fig. 5. Comparison of the reconstruction visual quality between non-dithered and dithered quantizing ganglion cell. The test image is Lena. The observation time is $t_{obs} = 55ms$.

across time. Interestingly, the quantizer evolves also from uniform to non-uniform, but in contradiction with traditional Lloyd-Max quantizers, renders high magnitudes precisely while it maps low magnitudes coarsely. Besides, we emitted a biologically plausible hypothesis that supposes the retinal noise distribution to have specific characteristics, yielding the definition of a non-subtractive dithered system. We do not claim that the retinal noise is a dither signal, but still such a hypothesis is seducing by the noise whitening and de-correlation features it allows. Our future work aims at adding several mechanisms of the retinal processing that are not taken into account in the current model. Namely, spatial filtering and lateral inhibitions are two important features that will be integrated in the upcoming model. Our goal is to infer, starting from a sufficiently realistic model, a decoding algorithm that could decipher actual neural recordings.

6. REFERENCES

[1] K. Masmoudi, M. Antonini, and P. Kornprobst, "Another look at retina as an image scalar quantizer," in *IEEE International Symposium on Circuits and Systems*, 2010.

[2] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code*, The MIT Press, Cambridge, MA, USA, 1997.

[3] A. Wohrer and P. Kornprobst, "A biological retina model and simulator, with contrast gain control," *Journal of Computational Neuroscience*, vol. 26, no. 2, pp. 219–249, April 2009.

[4] F. Rieke, "Temporal contrast adaptation in salamander bipolar cells," *Journal of Neuroscience*, vol. 21, no. 23, pp. 9445–9454, December 2001.

[5] W.Gerstner and W.Kistler, *Spiking Neuron Models : Single Neurons, Populations, Plasticity*, Cambridge University Press, 2002.

[6] J.R. Dormand and P.J. Prince, "A family of embedded runge-kutta formulae," *Journal of Computational and Applied Mathematics*, pp. 19–26, 1980.

[7] E.D. Adrian, "The impulses produced by sensory nerve endings," *Journal of Physiology*, vol. 61, no. 1, pp. 49–72, March 1926.

[8] A.B. Clark et al, "Electrical picture-transmitting system," US Patent assigned to AT& T, 1928.

[9] R.A. Wannamaker, S.P. Lipshitz, J. Vanderkooy, and J.N. Wright, "A theory of nonsubtractive dither," *IEEE Transactions on Signal Processing*, vol. 48, no. 2, pp. 499–516, 2000.

[10] D. Perkell and T. Bullock, "Neural coding," *Neurosciences Research Program Bulletin*, vol. 6, pp. 221–343, 1968.

[11] M.N. Shadlen and W.T. Newsome, "Noise, neural codes and cortical organization," *Findings and Current Opinion Cognitive Neuroscience*, vol. 4, pp. 569–79, 1998.

[12] A. Wohrer, *Model and large-scale simulator of a biological retina, with contrast gain control*, Ph.D. thesis, University of Nice-Sophia Antipolis, 2008.