Master des Sciences de la Matière

École Normale Supérieure de Lyon
Université Claude Bernard Lyon 1

STAGE 2007-2008
DIPOPPA Mario
M2
Option Physique

Spatio-Temporal Correlations of Natural Images during Fixational Eye Movements

The spatio-temporal correlation of natural images during fixational eye instability is studied. The statistics of eye movements and static natural images are first analyzed separately and then combined to obtain an analytical formula for the correlation function. Previous work on this subject is presented and important shortcomings are pointed out. A more rigorous and general formula is derived and tested numerically.

keywords :
Theoretical Neuroscience,Visual System, Natural Images, Non-equilibrium Statistical Physics, Stochastic Differential Equations

Max Planck Institute for Biological Cybernetics
Spenmanstraße 41,
72076 Tübingen (Ger)
Independent Research Group:
Computational Vision and Neuroscience

Supervisor
Dr. Matthias BETHGE

Date
31/07/08

Contents

Chapter 1

Introduction

1.1 Purpose of the project

The subject of this project is the *spatio-temporal correlation* of the *visual input* during *fixational eye movements*. This is part of a larger project which investigates the function of the primate retina, a fundamental part of the visual system. In previous works the *spatial* correlation of the visual input has already been studied [1, 2]. The starting point of our work are two recent studies concerning this topic [3, 4]. The original contribution of this project is a more correct understanding of the *temporal* correlation in the visual input. In particular we provide a more general and rigorous model for the spatio-temporal correlation. Our interest in a characterization of the spatio-temporal statistics of the visual input originates from the following reasons:

- If we want to understand how a system achieves a certain task, it is worth to know the general features of the input and the output. Then, if we have a model for the system, we can enter the input into it and test whether the model responds with the expected output.
- A successful idea in biology is to suppose that systems have evolved in a way to respond as efficiently as possible to the properties of the input. Barlow applied this concept to the visual system assuming that the system removes the highly redudant spatial information existing in *natural images* [5]. Atick, based on this assumption, predicted physiological properties of retinal ganglion cells of *spatial* light-sensitivity [6] that fitted very well to the experimental data [7]. Therefore, thanks to our input model, it is possible to test again this concept by trying to predict the *spatio-temporal* light sensitivity.
- An important controversy is the role of fixational eye movements. It is known that even if we are fixating at a certain point, the eyes make small uncontrolled movements [8]. As this kind of movement would blur a digital camera image, how is it possible that we can still see in an accurate way? An interesting result is that when we try to compensate this movement artificially vision tends to fade [9]. Why the primate eye has evolved in that way is not known. Several hypotheses have been suggested about the role of fixational eye movements:
  - It is necessary to refresh the retina in order to maintain the stimulus.
  - It decorrelates the spatial information [3] having a role of a filter similar to that described by Atick [6]. We will show that the previous theoretical work on this hypothesis in [3] made oversimplifying assumptions in the calculations of the correlation function.
  - It could improve the ability of edge detection [10].

- It could be used to encode spatial image structures into temporal ones.
- In the retina there is a point of high acuity, the fovea, (see Figure 1.1). The fixational instability moves this point around. In this way a large area is covered, in which the details can be analyzed with high spatial resolution.

1.2 Organization

During the internship I was supervised by Dr. Matthias Bethge, and I received help and suggestions from the others members of the group. The main topics studied by the *Computational Vision and Neuroscience Group* are the visual system and the behavior of neurons. Theoretical Neuroscience studies these subjects from a theoretical point of view and tries to apply knowledge from Computer Science, Bioinformatics, Physics, Biology and Psychology.

During my project I mainly imported methods from non-equilibrium statistical physics and scale-invariant systems to develop analytical models. Then I used numerical computation to verify these results. As part of the project, I studied the theory behind the retinal structure and the visual system in general. Even if this did not completely concern the topic of my project, this work was very useful to acquire some background of theoretical neuroscience since I want to continue in this direction during my PhD.

I want to thank Dr. Matthias Bethge to have supervised me during the internship. I want also to thank Philipp Berens, Alexander Ecker, Sebastian Gerwinn, Reshad Hosseini, Philipp Lies, Jakob Macke, Fabian Sinz, and Jonas Kosten, for all the help I received from them.

1.3 The Retina

The retina is a fundamental part of the visual system. All visual information is processed by this organ before being transmitted to the brain. Photons coming from the natural environment are refracted by the cornea and the lens in the eye, which operates similarly to the lens of a camera (see Figure 1.1). Then, photons traverse the retina that is transparent to light and excite the photoreceptors in the deeper layer (see Figure 1.2). Photoreceptors "count" photons within a certain range of energy: there are photoreceptors specialized to chromatic light (cones) and others to non-chromatic light (rods). These neuronal cells transduce light into an electrical signal. The signal is then transmitted to bipolar and horizontal cells. Horizontal cells play a lateral processing role, connecting several neuronal cells of the same layer, while the bipolar cells connect directly the photoreceptors with the ganglion cells. Amacrine cells play also a role of lateral processing. These cells connect several other cells in the layer where ganglion and bipolar cells are in contact. Ganglion cells are the final stage of this circuit projecting the signal to the Lateral Geniculate Nucleus (LGN). From the LGN the visual information is further transmitted to the primary visual cortex.

One of the reasons why the retina is studied intensively by theoretical neuroscientists is that it is supposed to be a forward system. In contrast to most other neural systems in the mammalian brain, the retina can therefore be studied as an isolated system which is not affected by the neural activity in the rest of the brain. Another reason is that the input and the output of the system are easy to study. The input, is well-defined and is namely the light. The output passes through the optic nerve, a unique channel that behaves like

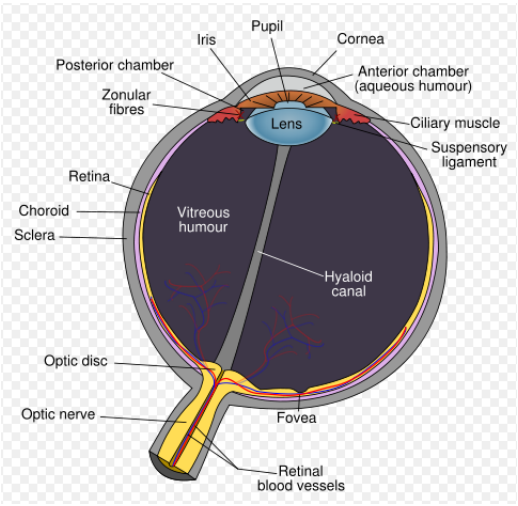

Figure 1.1: Scheme of a human eye cross section. The retina is the pink membrane in the inner part of the eye. All the information are collected by the ganglion cells and transmitted to the optic nerve. The fovea is a dip in the retina and is the point of highest acuity.

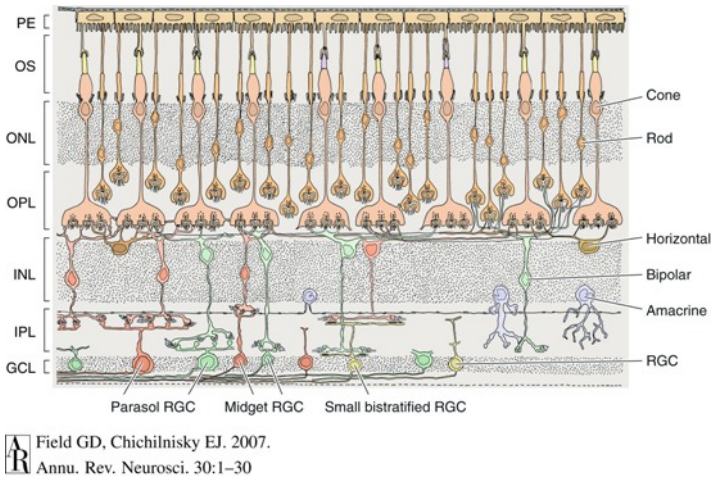

Figure 1.2: Scheme of the primate retina cross-section, from [12]. The neural circuit of the retina is composed by five different type of cells: photoreceptors (cones and rods), horizontal cells, bipolar cells, amacrine cells and retinal ganglion cells (RGC).

a bottleneck in the visual system. Therefore it is easy to read out the electrical signal output from the LGN. This makes the retina a unique model system for studying the neural basis of visual processing.

Chapter 2

Spatio-temporal correlation of natural images

2.1 Problem Statement

We are looking for an analytical expression of the spatio-temporal correlation of the retina's visual input. The retina transforms light rays coming from a three-dimensional world in a two-dimensional signal that we call *visual input*[1]. During fixation, there are two processes that make the visual input time-dependent: the fixational instability and the movement of the objects in the natural scene. Since we want to understand the consequences of fixational eye movements, only *static natural images* during these movements were investigated in this project.

The retinal layers are made of discrete arrays of neuronal cells (photoreceptors at the input and ganglion cells at the output) and the system is obviously finite in size but, in order to make analysis simple, we suppose the system to be continuous and infinite in size. We will discuss afterwards why this assumption is reasonable. We assume, moreover, that images are converted into a non-chromatic scale and so the visual input is a scalar function $I(\vec{x}, t)$ measuring light intensity.

The only two components needed to obtain the *visual input* are the *static natural image* $S(\vec{x})$ and the *fixational movement* $\vec{\xi}(t)$. Intuitively, the following simple relation relates these functions.

$$I(\vec{x}, t) = S(\vec{x} + \vec{\xi}(t)) \tag{2.1}$$

We would like to calculate the input correlation function of $I(\vec{x}, t)$. We assume the following properties for our model:

- $I(\vec{x}, t)$ has zero-mean and the correlation function is

$$C_{II}(\vec{x}, \vec{x}', t, t') = \left\langle \left\langle I(\vec{x}, t) I(\vec{x}', t') \right\rangle_{\Xi} \right\rangle_{\vec{\xi}} \tag{2.2}$$

  where $\langle \cdots \rangle_{\Xi}$ denotes the average over natural images and $\langle \cdots \rangle_{\xi}$ over fixational movements.

  We assume that the way how the light intensity is spatially distributed in natural images and the way how the eye moves during time are independent processes.

[1] All long the report we will use this $2D$-notation: the spatial coordinates $\vec{x} \equiv (x_1, x_2)$, the radius $x \equiv \sqrt{x_1^2 + x_1^2}$ and the gradient $\vec{\nabla} = (\partial_{x_1}, \partial_{x_2})$

Intuitively we can suppose that the first one is an ergodic and translational invariant process. Therefore we can replace the average over infinitely many different natural images by the average over the spatial domain for an infinite-size image[2]:

$$C_{II}(\vec{x}, t, t') = \left\langle \left\langle I(\vec{x}' + \vec{x}, t) I(\vec{x}', t') \right\rangle_{\vec{x}'} \right\rangle_{\vec{\xi}} \tag{2.3}$$

- For the same reasons as above, the static natural image correlation could be obtained by average over the spatial domain:

$$C_{SS}(\vec{x}) = \left\langle S(\vec{x}' + \vec{x}) S(\vec{x}') \right\rangle_{\vec{x}'} \tag{2.4}$$

We study the process in the Fourier domain in order to avoid convolution computations. It is possible that

$$\left\langle S(\vec{x}' + \vec{x}) S(\vec{x}') \right\rangle_{\vec{x}'} = FT_{\vec{x}}^{-1} \left\{ \left| \hat{S}(\vec{u}) \right|^2 \right\}, \tag{2.5}$$

where FT is the Fourier transform operator.

By symmetry reasons we assume that $\left|\hat{S}(\vec{u})\right|^2 = \hat{C}_{SS}(\vec{u})$ is isotropic and depends only on the radius u. The two-dimensional Fourier transform of a function depending only on the radius is a 0-order Hankel transform [19].

$$FT_{\vec{x}}^{-1} \left\{ \hat{f}(u) \right\} = H_x^{-1} \left\{ \hat{f}(u) \right\} = \int_0^\infty du \hat{f}(u) u J_0(ux), \tag{2.6}$$

where J_0 is the 0-order Bessel function[3].

In this way , by using the rotational symmetry, we can effectively reduce our analysis from two dimensions to one dimension.

- We model the fixational eye movements as a Gaussian process.

Definition: a stochastic process $\vec{\xi}(t)$ is a Gaussian process if $\left\{ \vec{\xi}(t_1), \cdots, \vec{\xi}(t_n) \right\}$ is normally distributed for any sequence $\{t_1, \cdots, t_n\}$ [16].

Additionally we assume that the components ξ_1 and ξ_2 are independent (2.7), zero-mean (2.8) and isotropic (2.9):

$$\left\langle \xi_1(t) \xi_2(t') \right\rangle = 0, \tag{2.7}$$

$$\left\langle \vec{\xi}(t) \right\rangle = \vec{0}, \tag{2.8}$$

$$\left\langle \xi_1(t) \xi_1(t') \right\rangle = \left\langle \xi_2(t) \xi_2(t') \right\rangle = \left\langle \xi(t) \xi(t') \right\rangle \tag{2.9}$$

[2]Since in the numerical analysis we work with finite-size images, we average over the spatial domain AND over several natural images.

[3]for simplicity, all along the report we'll use the following notation: $\hat{f}(\vec{u})$ is the Fourier Transform in the frequency domain $\vec{u}$ of $f(\vec{x})$ and $\hat{f}(u)$ is the 0-order Hankel Transform of $f(x)$

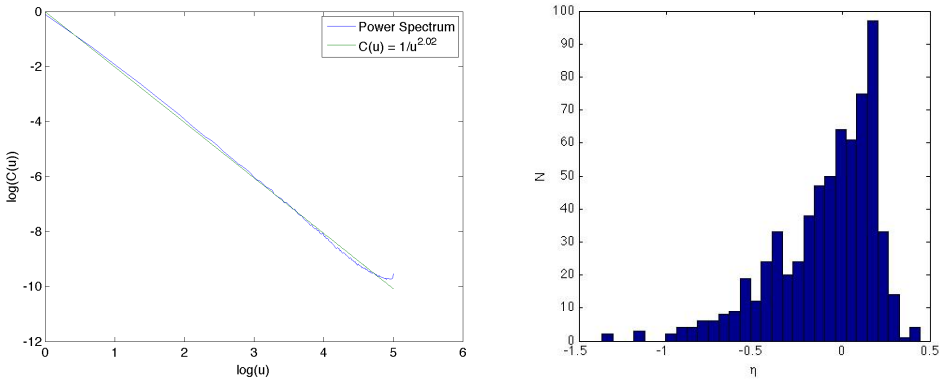

Figure 2.1: Left: power spectrum of the static natural images ensemble fitted with a power law distribution. Right: Histogram of η for every single image. We see that η is continuously distributed without any cluster. Therefore averaging over all images seems to be justified.

2.2 Static Image Statistics

2.2.1 Power Spectrum

We now focus on the statistics of *static natural images*. We used a database of pictures taken by van Hateren with a digital camera (*Kodak DCS420*) around the Netherlands [11]. From this database we selected a set of 1000 images, discarding images showing man-made objects (see Figure 2.2).

We first converted the image pixels into log intensities as a rough model for the process carried out by the photoreceptors [20]:

$$S(\vec{x}) \rightarrow \log(S(\vec{x})) - \langle \log(S(\vec{x})) \rangle_{\vec{x}} \tag{2.10}$$

Numerically, we computed the average power spectrum by first calculating the power spectra of each individual image and then averaging. We found a power law distribution (see Figure 2.1):

$$\hat{C}_{SS}(u) \propto \frac{1}{u^{2-\eta}} \tag{2.11}$$

with $\eta = -0.02$.

The same result, but with slightly different values for η was found in others work [2, 12].

Several hypotheses have been put forward to explain why the power spectrum has this distribution. In particular, it has been suggested that this is due to the presence of edges and scale invariance of the natural image statistics [2].

The exact value of η can vary from data set to data set. Therefore, we treat η as a free parameter in our analysis of the effect of fixational eye movements on the correlation function. For the numerical calculations we generate pink noise (see Fig. 2.2) according to equation instead of using natural images. In this way, we can precisely control the parameter η. Pink noise images are generated by convolving white noise $R(\vec{x})$ with an appropriate linear filter:

$$S(\vec{x}) = FT_{\vec{x}}^{-1} \left\{ \frac{1}{u^{1-\eta/2}} FT_{\vec{u}} \left[R(\vec{x}) \right] \right\} \tag{2.12}$$

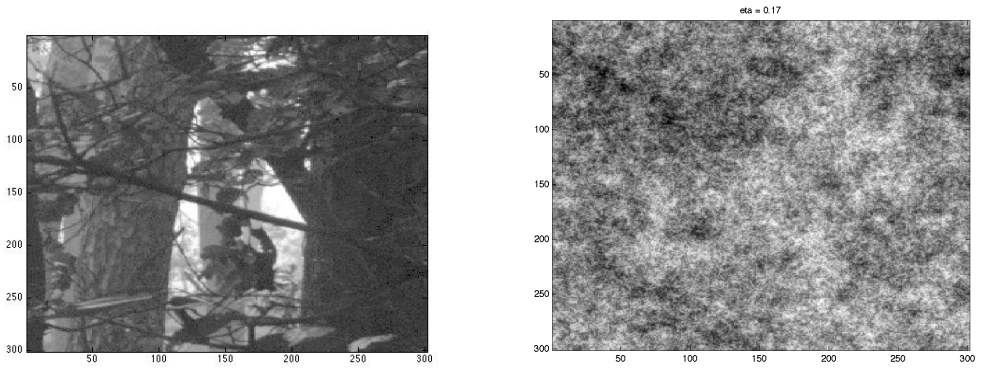

Figure 2.2: Left: sample of a wood scenario from the set of static natural images. Right: image generated by pink noise with $\eta = 0.17$.

2.2.2 Correlation Function

From (2.2.1) the correlation function is (see [17]):

$$C(x) = H_x^{-1}\left\{\frac{1}{u^{2-\eta}}\right\} = 2^{-1+\eta}\frac{\Gamma(\eta/2)}{\Gamma(1-\eta/2)}x^{-\eta} \tag{2.13}$$

for $0 < \eta < 2$.

However, numerically we found a negative constant in addition to that, probably due to the fact that the image has finite size.

Ruderman found an approximation formula computed by the Fourier Transform of the power spectrum (2.2.1) with a cutoff in the low frequencies $u < \lambda$ [2]:

$$C(x) = -C_1(\eta) + C_2(\eta)x^{-\eta} \tag{2.14}$$

The coefficients are (up to a multiplicative factor):

$$C_1(\eta) = \lambda^\eta/\eta \tag{2.15}$$

and

$$C_2(\eta) = 2^\eta \cos(\pi\eta/2)\frac{\Gamma^2(1/2-\eta/2)}{\pi\Gamma(1-\eta)}\Gamma(\eta) \tag{2.16}$$

This formula is only valid for $\eta \neq 0$. According to Ruderman [2] the limit for $\eta \to 0$ of (2.14) exists and is

$$C(x) = -\log(\lambda x) \tag{2.17}$$

In this case there is a discontinuity for $\eta = 0$ (see Figure 2.3). But if the limit exist that should not happen. In the following, we will show, that (2.17) is not the correct limit of (2.14). To this end, the different terms of the coefficient (2.16) are computed around $\eta = 0$:

- Since $\Gamma'(1) = -\gamma$ (see [17]), where γ is the Euler constant, we can make the following approximation for $\eta \approx 0$:

$$\Gamma(\eta) = \frac{\Gamma(1+\eta)}{\eta} \cong \frac{1-\gamma\eta}{\eta} \tag{2.18}$$

- Using other properties of the Gamma function (see [17]) we have:

$$\frac{\Gamma(1/2-\eta/2)}{\sqrt{\pi}\Gamma(1-\eta)} = \frac{2^{\eta}}{\Gamma(1-\eta/2)} = 2^{\eta}\frac{\sin(\pi\eta/2)}{\pi}\Gamma(\eta/2) = 2^{\eta}\frac{\sin(\pi\eta/2)}{(\pi\eta/2)}\Gamma(1+\eta/2) \tag{2.19}$$

If then we do a first order approximation around $\eta = 0$, we obtain:

$$2^{\eta}\frac{\sin(\pi\eta/2)}{(\pi\eta/2)}\Gamma(1+\eta/2) \cong 2^{\eta}\left(1-\frac{\gamma}{2}\eta\right) \tag{2.20}$$

- We further compute (see [17])

$$\frac{\Gamma'(1/2)}{\Gamma(1/2)} = -2-\gamma-\sum_{n=1}^{\infty}\left[\frac{1}{n}-\frac{1}{n+1/2}\right] \tag{2.21}$$

and

$$\sum_{n=1}^{\infty}\left[\frac{1}{n}-\frac{1}{n+1/2}\right] = 2\left[\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\frac{1}{5}+\cdots\right] = 2(1-\log(2)). \tag{2.22}$$

Together with $\Gamma(1/2) = \sqrt{\pi}$ [17] it follows for $\eta \approx 0$:

$$\frac{\Gamma(1/2-\eta/2)}{\sqrt{\pi}} \cong 1+\left(log(2)+\frac{\gamma}{2}\right)\eta. \tag{2.23}$$

- Finally, we have:

$$2^{\eta}\cos(\pi\eta/2) \cong 1+2\eta. \tag{2.24}$$

Combining the results (2.18), (2.20), (2.23) and (2.24) yields

$$C_2(\eta) \cong \frac{1}{\eta}\left[1+(3log(2)-\gamma)\,\eta\right] \tag{2.25}$$

Now we can write down the expansion of Rudermann formula (2.14), around $\eta \approx 0$:

$$C(x) \cong \frac{1}{\eta}\left\{\left[1+(3\log(2)-\gamma)\,\eta\right]\left[1-\eta\log(x)\right]-\left[1+\lambda\eta\right]\right\} = \tag{2.26}$$

$$= -\log\left(\frac{e^{\gamma}}{8}\right)-\log(\lambda)-\log(x)+O(\eta) \tag{2.27}$$

Therefore the correct limit of for $\eta \to 0$ reads:

$$C(x) = -\log\left(\lambda x\frac{e^{\gamma}}{8}\right) \tag{2.28}$$

As one can see in Figure 2.3, there is no discontinuity anymore when using our new result for the limit (Eq.(2.28)).

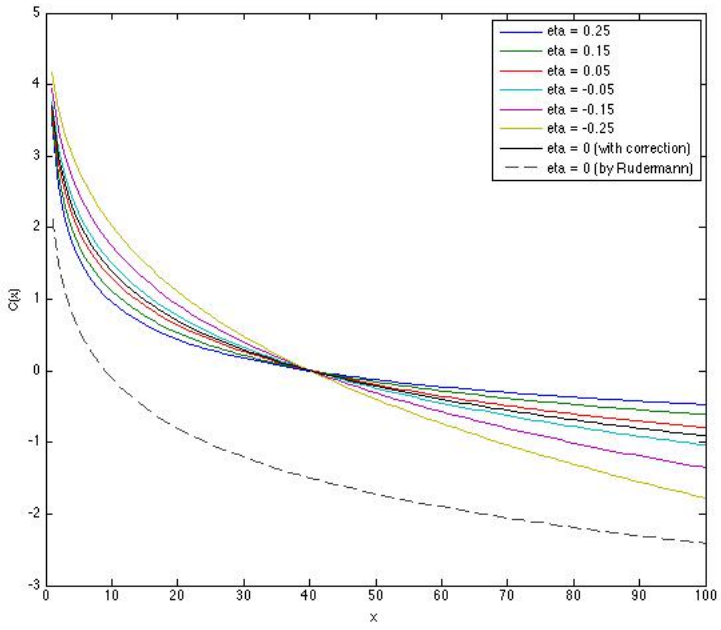

Figure 2.3: We can see that the curve defined by the Ruderman formula for $\eta = 0$ (black dashed line) is far off form the other cases (colored lines) when $\eta \neq 0$ but small. Hence, it cannot correctly describe the limit of $\eta \neq 0$ (colored lines) for $\eta \to 0$. In contrast, the curve defined by our new formula (black solid line) perfectly falls into place with the other curves (colored lines).

2.3 Modeling the eye movement process

Fixational eye movements are composed of different kind of processes [8]:

- *Tremors* are wave-like movement with frequencies $\sim 90\ Hz$ and amplitude $\sim 1\ \mu m$.
- *Drifts* are slow movements of the eye covering several photoreceptors with amplitude $\sim 10\ \mu m$ and duration $\sim 0.3\ s$.
- *Microsaccades* are fast "jumps" of the eye restoring the position after a drift movement near the inital position of the gaze. The amplitude is $\sim 10\ \mu m$ and duration is $\sim 0.02\ s$.

We can compare these measures with the width of the cones in the fovea $\sim 2.5\ \mu m$ and the width of the rods near the fovea $\sim 1\ \mu m$ [20].

It has been argued that tremors do not influence processing of the visual information in the retina. Therefore, we consider only drifts and microsaccades in the analytical models for the movement process $\vec{\xi}(t)$. We will determine the statistical properties of these processes, studying in particular $\langle \Delta\xi^2 \rangle$, where $\Delta\xi \equiv \xi(t) - \xi(t')$, because it will be used in the final model of the correlation function.

Similar to [10] we will use stochastic processes for modeling fixational eye movements. Importantly, however, we do not assume stationarity since we fix the initial position $\vec{\xi(0)} = 0$, and define $I(\vec{x}, t = 0) = S(\vec{x})$. This choice implements the fact that it makes sense to define the time of saccading to a new fixation point as the starting point for the fixational eye movements.

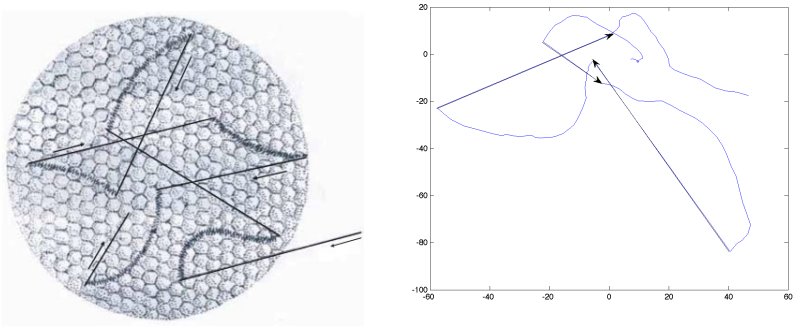

Figure 2.4: Left: Scheme of fixational eye movements, from [8]. *Tremors* are represented by the small amplitude and high frequency movements, *drifts* by the broad and slow curved lines (superimposed to the tremors movements) and *microsaccades* are represented by the straight lines. Right: Simulated eye movement using one of the proposed biological based processes (process D, we will describe later that)

Here, we will restrict our analysis to Gaussian processes in order to facilitate an analytic treatment. Even if these processes do not describe exactly the real fixational eye movements they provide us with a qualitative understanding of the features of the spatio-temporal correlation function.

2.3.1 Process A

The simplest Gaussian process that we can use is a Wiener Process, defined by the following relation:

$$\frac{d\xi(t)}{dt} = \beta R(t) \tag{2.29}$$

where β represents a velocity parameter and $R(t)$ stands for Gaussian white noise, whose correlation function is defined to be delta correlated: (2.30):

$$\left\langle R(t)R(t')\right\rangle = \delta(t-t') \tag{2.30}$$

This process describes the situation in which the eye makes completely random movements.

From [18] we have:

$$\left\langle \xi(t)\xi(t')\right\rangle = \beta^2 \min\left\{t,t'\right\} \tag{2.31}$$

Now we can write:

$$\left\langle \Delta\xi^2\right\rangle = \left\langle \xi(t)^2\right\rangle + \left\langle \xi(t')^2\right\rangle - 2\left\langle \xi(t)\xi(t')\right\rangle \tag{2.32}$$

Defining $\Delta t = |t-t'|$ we obtain:

$$\left\langle \Delta\xi^2\right\rangle = \beta^2\Delta t \tag{2.33}$$

2.3.2 Process B

While fixating the eye makes correcting movements onto the fixation point by means of microsaccades. The following process introduces knowledge about the fixation point:

$$\frac{d\xi(t)}{dt} = -\alpha\xi(t) + \beta R(t) \tag{2.34}$$

The first term in the right side of the equation constitutes a drift term which drives the eye position back to the fixation point.

This process has already implemented by [10]. But, we stress again, that we are fixing the initial condition and therefore, we will obtain a non-stationary process.

The solution of the equation is

$$\xi(t) = \beta e^{-\alpha t} \int_0^t ds e^{\alpha s} R(s). \tag{2.35}$$

We can now compute the correlation of the process. Since $R(t)$ is the only random variable we can put the average inside the integral:

$$\left\langle \xi(t)\xi(t') \right\rangle = \beta^2 e^{-\alpha(t+t')} \int_0^t ds \int_0^{t'} ds' e^{\alpha(s+s')} \left\langle R(s)R(s') \right\rangle \tag{2.36}$$

Suppose $t > t'$, using (2.30) we can integrate over s' obtaining:

$$\left\langle \xi(t)\xi(t') \right\rangle = \beta^2 e^{-\alpha(t+t')} \int_0^{t'} ds' e^{2\alpha s'} = \frac{\beta^2}{2\alpha} e^{-\alpha t} \left(e^{\alpha t'} - e^{-\alpha t'} \right) \tag{2.37}$$

Considering the case $t < t'$ as well, we have in general:

$$\left\langle \xi(t)\xi(t') \right\rangle = \frac{\beta^2}{2\alpha} \left[e^{-\alpha|t-t'|} - e^{-\alpha(t+t')} \right] \tag{2.38}$$

If we define $t_0 \equiv \min\{t, t'\}$ it follows:

$$\left\langle \Delta\xi^2 \right\rangle = \frac{\beta^2}{\alpha} \left[1 - e^{-\alpha\Delta t} - \frac{1}{2} e^{-2\alpha t_0} \left(1 - e^{-\alpha\Delta t} \right)^2 \right] \tag{2.39}$$

2.3.3 Process C

As another case, we consider a model for the drift movements. We can assume that during these movements the eye does not know the position but the velocity follows a brownian motion with friction:

$$\frac{dv(t)}{dt} = -\alpha v(t) + \beta R(t) \tag{2.40}$$

$$v(t) = \frac{d\xi(t)}{dt} \tag{2.41}$$

and we add the initial condition: $v(0) = 0$. The variable α has the role of a friction parameter of the term reducing the eye velocity.

This is again a Gaussian process since this process could be obtained by integrating (2.35). We know that a linear operator transforms a Gaussian process in another Gaussian process [16]. For the same reason we can compute the correlation of this process integrating the correlation of process B (2.38):

$$\left\langle \xi(t)\xi(t') \right\rangle = \int_0^t ds \int_0^{t'} ds' \left\langle v(s)v(s') \right\rangle = \frac{\beta^2}{2\alpha} \int_0^t ds \int_0^{t'} ds' \left[e^{-\alpha|s-s'|} - e^{-\alpha(s+s')} \right] \tag{2.42}$$

Consider the first term in the integral (2.42). We suppose $t > t'$ and then we split the integral in two parts. In the first one the variable s is integrated from 0 to t' and in the second one from t' to t.

$$\int_0^t ds \int_0^{t'} ds' e^{-\alpha|s-s'|} = \int_0^{t'} ds \int_0^{t'} ds' e^{-\alpha|s-s'|} + \int_{t'}^t ds \int_0^{t'} ds' e^{-\alpha(s-s')} \tag{2.43}$$

In this way we have that he first integral of (2.43) is symmetric with respect to the axes $s = s'$. Therefore it can be integrated just over the area below that axes:

$$\int_0^{t'} ds \int_0^{t'} ds' e^{-\alpha|s-s'|} = 2\int_0^{t'} ds \int_0^{s} ds' e^{-\alpha(s-s')} = \frac{1}{\alpha^2}\left[2\alpha t' - 2\left(1 - e^{-\alpha t'}\right)\right] \tag{2.44}$$

Then we solve the second integral of (2.43):

$$\int_{t'}^{t} ds \int_0^{t'} ds' e^{-\alpha(s-s')} = \frac{1}{\alpha^2}\left[1 - e^{-\alpha t'} - e^{-\alpha(t-t')} + e^{-\alpha t}\right] \tag{2.45}$$

Consider now the second term in the integral of (2.42). We can solve it directly:

$$-\int_0^{t} ds e^{-\alpha s'} \int_0^{t'} ds' e^{-\alpha s'} = \frac{1}{\alpha^2}\left[-1 - e^{-\alpha(t+t')} + e^{-\alpha t} + e^{-\alpha t'}\right] \tag{2.46}$$

Using together (2.44), (2.45) and (2.46) and generalizing also for the case $t < t'$ we finally obtain:

$$\left\langle \xi(t)\xi(t') \right\rangle = \frac{\beta^2}{2\alpha^3}\left[2\alpha \min\left\{t, t'\right\} + 2e^{-\alpha \min\{t,t'\}} - 2 + 2e^{-\alpha \max\{t,t'\}} - e^{-\alpha|t-t'|} - e^{-\alpha(t+t')}\right] \tag{2.47}$$

From this equation we can compute (2.32) for this process:

$$\left\langle \Delta\xi^2 \right\rangle = \frac{\beta^2}{\alpha^3}\left[\alpha\Delta t - \left(1 - e^{-\alpha\Delta t}\right) - \frac{1}{2}e^{-2\alpha t_0}\left(1 - e^{-\alpha\Delta t}\right)^2\right] \tag{2.48}$$

2.3.4 Process D

Finally, consider a more sophisticated process which is generated by the equations of process C (2.40) and which is periodically restored near the center position. This is not a Gaussian process anymore but we can see how it mimics the fixational eye movement very well (see Figure 2.4).

2.4 Casile-Rucci approximation

Casile and Rucci proposed an approximation of the correlation function of the visual input, under the hypothesis that the fixational movement does not move too far away from the fixation point $\xi \approx 0$ [3]. We will show that this approximation is flawed, even within this assumption.

Casile and Rucci expanded the input up to the first order of Taylor series:

$$I(\vec{x}, t) = S(\vec{x} + \vec{\xi}(t)) \cong S(\vec{x}) + \vec{\nabla}S(\vec{x}) \cdot \vec{\xi}(t) \tag{2.49}$$

They insert equation (2.49) in the general correlation function (2.3):

$$C_{II}(\vec{x}, t, t') \cong \left\langle \left\langle \left[S(\vec{x}' + \vec{x}) + \vec{\nabla}S(\vec{x}' + \vec{x}) \cdot \vec{\xi}(t)\right]\left[S(\vec{x}') + \vec{\nabla}S(\vec{x}') \cdot \vec{\xi}(t')\right] \right\rangle_{\vec{x}'} \right\rangle_{\vec{\xi}} \tag{2.50}$$

By symmetry we can neglect the terms with $\langle \xi_i(t) \rangle$. Then we can compute the terms with $\left\langle \xi_i(t)\xi_j(t') \right\rangle$:

$$\left\langle \partial_{x_1} S(\vec{x}' + \vec{x}) \partial_{x_1} S(\vec{x}') \right\rangle_{\vec{x}'} \left\langle \xi_1(t)\xi_1(t') \right\rangle_{\vec{\xi}} + \left\langle \partial_{x_2} S(\vec{x}' + \vec{x}) \partial_{x_2} S(\vec{x}') \right\rangle_{\vec{x}'} \left\langle \xi_2(t)\xi_2(t') \right\rangle_{\vec{\xi}} \tag{2.51}$$

Using the isotropy property (2.9) we can easily rearrange this expression in order to obtain the following approximation of the correlation function:

$$C_{II}(\vec{x},t,t') \cong \left\langle S(\vec{x}'+\vec{x})S(\vec{x}')\right\rangle_{\vec{x}'} + \left\langle \vec{\nabla}S(\vec{x}'+\vec{x})\cdot\vec{\nabla}S(\vec{x}')\right\rangle_{x'} \left\langle \xi(t)\vec{\xi}(t')\right\rangle_{\vec{\xi}} \tag{2.52}$$

It is possible to show that:

$$\left\langle \vec{\nabla}S(\vec{x}'+\vec{x})\cdot\vec{\nabla}S(\vec{x}')\right\rangle_{\vec{x}'} = FT_{\vec{x}}\left\{u^2\left|\hat{S}(\vec{u})\right|^2\right\} \tag{2.53}$$

Therefore, inserting (2.53) in (2.52), using the 0-order Hankel function (2.6) and recalling that the first term of this last equation represents just the static image correlation (2.4), we obtain:

$$C_{II}(\vec{x},t,t') \cong C_{SS}(\vec{x}) + H_x^{-1}\left\{u^2\hat{C}_{SS}(u)\right\}\left\langle \xi(t)\xi(t')\right\rangle \tag{2.54}$$

Rucci and Casile concluded that u^2 compensates the power law $1/u^2$ of static natural images and so the signal is partially decorrelated, and could be one of the reasons why fixational eye movements exist.

It is easy to show that, for $t \neq 0$, $\langle\xi(t)\xi(t'=t)\rangle = 0$ does not always hold, as for example in the process A $\left\langle\xi^2(t)\right\rangle \propto t$. Therefore, if the Rucci-Casile formula was true, we would have $C_{II}(\vec{x},t,t'=t) \neq C_{SS}(\vec{x})$. This is in contradiction to the fact that by translational invariance, we have:

$$C_{II}(\vec{x},t,t'=t) = \left\langle S(\vec{x}'+\vec{x}+\vec{\xi}(t))S(\vec{x}'+\vec{\xi}(t))\right\rangle_{\vec{x}'} = \left\langle S(\vec{x}'+\vec{x})S(\vec{x}')\right\rangle_{\vec{x}'} = C_{SS}(\vec{x}) \tag{2.55}$$

In general, the use of truncated Taylor expansion as an approximation is very dangerous if one aims at computing correlation function and we will derive an exact formula later. For now, we locate the flaw in the Rucci-Casile calculation by showing that the second term in the Taylor expansion must not be ignored.

$$I(\vec{x},t) \cong S(\vec{x}) + \vec{\nabla}S(x)\cdot\vec{\xi}(t) + \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\partial_{x_i}\partial_{x_j}S(\vec{x})\xi_i(t)\xi_j(t) \tag{2.56}$$

Inserting (2.56) in (2.3) we find a second-order term that was not take into account by Rucci and Casile:

$$S(\vec{x}'+\vec{x})\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\partial_{x_i}\partial_{x_j}S(\vec{x'})\xi_i(t')\xi_j(t') + S(\vec{x}')\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\partial_{x_i}\partial_{x_j}S(\vec{x'}+\vec{x})\xi_i(t)\xi_j(t) \tag{2.57}$$

Then using again the isotropic property (2.9) this term becomes:

$$\left[\frac{1}{2}\nabla^2 S(\vec{x'})S(\vec{x}'+\vec{x})\right]\xi(t')\xi(t') + \left[\frac{1}{2}\nabla^2 S(\vec{x'}+\vec{x})S(\vec{x}')\right]\xi(t)\xi(t) \tag{2.58}$$

It is possible to show that:

$$\left\langle \nabla^2 S(\vec{x}'+\vec{x})S(\vec{x}')\right\rangle_{\vec{x}'} = -FT_{\vec{x}}^{-1}\left\{u^2\left|\hat{S}(\vec{u})\right|^2\right\} \tag{2.59}$$

If we neglect all third-order and higher-order terms of the process, the correct second-order approximation is obtained by adding (2.58) to (2.54):

$$C_{II}(x,t,t') \cong C_{SS}(x) - H_x\left\{u^2\hat{C}_{SS}(u)^2\right\}\left[\frac{1}{2}\left\langle\xi^2(t)\right\rangle + \frac{1}{2}\left\langle\xi^2(t')\right\rangle - \left\langle\xi(t)\xi(t')\right\rangle\right] \tag{2.60}$$

Note, that the term neglected by Casile and Rucci is even larger than the other term with the consequence that we obtain the opposite sign for the change in the correlation function $C_{II}(x,t,t') - C_{SS}(x)$. This is easily recognized by the fact that $\left\langle [\xi(t) - \xi(t')]^2 \right\rangle \geq 0$ implies $\left\langle \xi^2(t) \right\rangle + \left\langle \xi^2(t') \right\rangle \geq 2 \left\langle \xi(t)\xi(t') \right\rangle$. Furthermore, we now find $C_{II}(x,t,t'=t) = C_{SS}(x)$ as it has to be expected for the correct formula. Nevertheless, one has to be very careful with this approximation as it still may deviate strongly from the truth.

Chapter 3

The final model

We want to build up a more general model. The average position ξ is not always small. For example in the process A, using (2.31), we have:

$$\xi \equiv \sqrt{\langle \xi^2(t) \rangle} = \beta\sqrt{t} \tag{3.1}$$

so that ξ is growing boundlessly with time.

We can study the visual input in the space-frequency domain. Using the relation (2.1) we obtain:

$$I(\vec{u},t) = \int dx_1 \int dx_2 S(\vec{x} + \vec{\xi}(t)) e^{-i\vec{x}\cdot\vec{u}} = e^{i\vec{\xi}(t)\cdot\vec{u}} S(\vec{u}) \tag{3.2}$$

Translating also the correlation function into the Fourier domain, the convolutions can be replaced by multiplications.

$$\left\langle \left\langle I(\vec{x}' + \vec{x}, t) I(\vec{x}', t') \right\rangle_{\vec{x}'} \right\rangle_{\vec{\xi}} = FT_{\vec{x}}^{-1} \left\{ \left\langle I(\vec{u}, t) I(-\vec{u}, t') \right\rangle_{\vec{\xi}} \right\} \tag{3.3}$$

Inserting (3.2) in (3.3) yields:

$$FT_{\vec{x}}^{-1} \left\{ \left\langle S(\vec{u}) e^{i\vec{\xi}(t)\cdot\vec{u}} S(-\vec{u}) e^{-i\vec{\xi}(t')\cdot\vec{u}} \right\rangle_{\vec{\xi}} \right\} = FT_{\vec{x}}^{-1} \left\{ |S(\vec{u})|^2 \left\langle e^{i\Delta\vec{\xi}\cdot\vec{u}} \right\rangle_{\vec{\xi}} \right\} \tag{3.4}$$

Note that this means that we have separated the correlation of the static natural image from the gaze movement process.

$$C_{II}(\vec{x}, t, t') = FT_{\vec{x}}^{-1} \left\{ \hat{C}_{SS}(u) \left\langle e^{i\Delta\vec{\xi}\cdot\vec{u}} \right\rangle_{\vec{\xi}} \right\} \tag{3.5}$$

This is a very important step as we can now study the two different problems separately: the contribution of spatial correlations between pixel intensities and the temporal correlation of image the movement. The solution given by (3.5) is fully general, without approximation, and valid also if the movement is not a Gaussian process. We can already see that when $t = t'$ and hence $\Delta\vec{\xi} = 0$ the value of the exponential function is 1 and the correlation function (3.5) reduces to $C_{SS}(\vec{x})$ as to be expected. Now we show that when the movement is a Gaussian, zero-mean, and isotropic process, and the movements along the two axes are independent, the correlation takes a very simple form.

First, because of the independence of ξ_1 and ξ_2 (2.7), we have:

$$\left\langle e^{i(\Delta\xi_1 u_1 + \Delta\xi_2 u_2)} \right\rangle_{\vec{\xi}} = \left\langle e^{i\Delta\xi_1 u_1} \right\rangle_{\xi_1} \left\langle e^{i\Delta\xi_2 u_2} \right\rangle_{\xi_2} \tag{3.6}$$

If we assume that the fixational eye movements are modeled by a Gaussian process this yields:

$$\left\langle e^{i\Delta\xi_1 u_1} \right\rangle = \int d\Delta\xi e^{i\Delta\xi u_1} P\left(\Delta\xi|t,t'\right) \tag{3.7}$$

Furthermore, it follows that the distribution of $\Delta\xi$ is then Gaussian too: the vector $(\Delta\xi(t,t'),\xi(t'))$ can be obtained by a linear combination of $(\xi(t),\xi(t'))$, which is normally distributed. Therefore, and because ξ is a Gaussian process, $(\Delta\xi(t,t'),\xi(t'))$ will be normally distributed with Gaussian density: $P(\Delta\xi(t,t'),\xi(t'))$ [16]. After marginalization over the joint density $P(\Delta\xi(t,t')) = \int d\xi(t')P(\Delta\xi(t,t'),\xi(t'))$ we obtain:

$$P\left(\Delta\xi|t,t'\right) = \frac{1}{\sqrt{2\pi\left\langle\Delta\xi^2\right\rangle}} e^{-\frac{\Delta\xi^2}{2\left\langle\Delta\xi^2\right\rangle}} \tag{3.8}$$

From (3.7) and (3.8) we have:

$$\int d\Delta\xi e^{i\Delta\xi u_1} \frac{1}{\sqrt{2\pi\left\langle\Delta\xi^2\right\rangle}} e^{-\frac{\Delta\xi^2}{2\left\langle\Delta\xi^2\right\rangle}} = e^{-\frac{\left\langle\Delta\xi^2\right\rangle u_1^2}{2}} \tag{3.9}$$

Inserting (3.9) in (3.6) and then in (3.5) yields:

$$C_{II}(\vec{x},t,t') = FT_{\vec{x}}^{-1}\left\{\hat{C}_{SS}(u)e^{-\frac{\left\langle\Delta\xi^2\right\rangle u_1^2}{2}} e^{-\frac{\left\langle\Delta\xi^2\right\rangle u_2^2}{2}}\right\} \tag{3.10}$$

Thus, we finally obtain[1]

$$C_{II}(x,t,t') = H_x^{-1}\left\{\hat{C}_{SS}(u)e^{-\frac{\left\langle\Delta\xi^2\right\rangle u^2}{2}}\right\} \tag{3.11}$$

We tested our model for processes A, B, and C numerically obtaining results that fit very well to the analytical formula (3.11) (see Figures 3.1, 3.2 and 3.3):

- For $\left\langle\Delta\xi^2\right\rangle$ in the analytical formula, the equations (2.39), (2.33) and (2.48) have been used for the processes A, B and C, respectively. For $\hat{C}_{SS}(u)$ we assumed a power law distribution (2.2.1).
- In the numerical calculations we used a discretized version of (2.34), (2.29) and (2.40) to simulate the processes A, B, and C, respectively. Furthermore we used pink noise intead of natural images (see Figure 2.2) as the correlation function does not depend on higher order correlation. We then let these images move rigidly according to this dynamics.

The physical interpretation of this formula is the following. The exponential in (3.11) blurs the static correlation in high frequencies and only the low frequencies survive $u \ll 1/\sqrt{\langle\Delta\xi^2\rangle}$. Therefore in the spatial domain, the correlation will be reduced for small distances. The physical reason is that, when the eye moves, the small details will be blurred, while the correlation of large objects will not be affected.

[1] In the numerical computation, since we are dealing with system of finite size L we should divide the term in the exponential by: $\frac{L^2}{(2\pi)^2}$

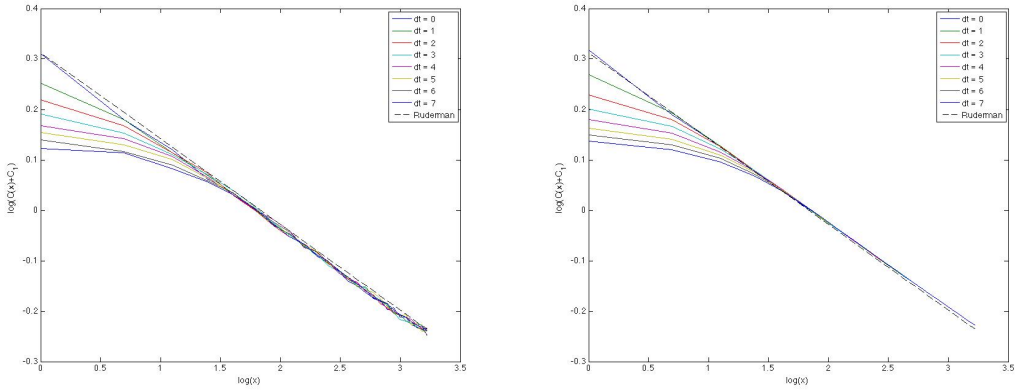

Figure 3.1: Spatio-temporal correlation for the process A, with static images generated by pink noise with $\eta = 0.17$ and computed at initial time $t_0 = 0$. Left: numerical data. Right: analytical result

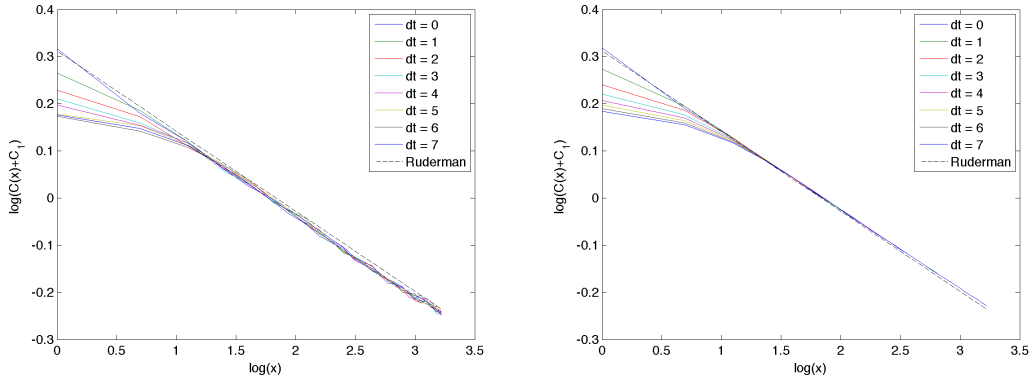

Figure 3.2: Spatio-temporal correlation for the process B, with static images generated by pink noise with $\eta = 0.17$ and computed at initial time $t_0 = 0$. Left: numerical data. Right: analytical result

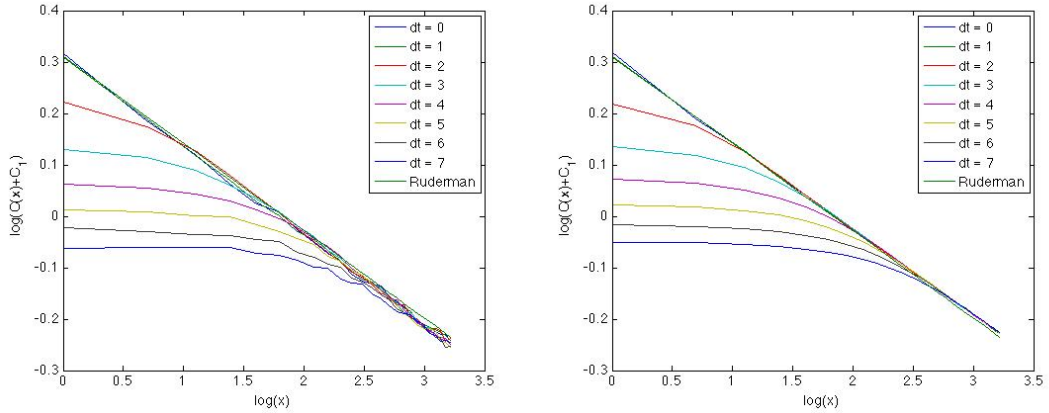

Figure 3.3: Spatio-temporal correlation for the process C, with static images generated by pink noise with $\eta = 0.17$ and computed at initial time $t_0 = 0$. Left: numerical data. Right: analytical result

3.1 Conclusion

It is a long standing hypothesis that sensory neurons are adapted to the statistics of the typical input signals. In particular, the redundancy reduction hypothesis put forth by Barlow [5] states that the goal of sensory coding is to transform the input signals into statistically independent components. Theoretical studies have successfully derived several *spatial* response properties of retinal ganglion cells from this principle including color opponency [13], center-surround bandpass filtering [6] and contrast gain control [14]. A few studies have also applied the redundancy reduction principle to movies or sequences of natural images [15] but ignored the global jitter imposed by fixational eye movements. The role of fixational eye movements in the context of the redundancy reduction hypothesis has been studied only recently by Rucci and Casile [3]. Their study points to an important contribution of the fixational eye movements to the temporal structure of the visual input collected in the retina. Their theoretical analysis, however, is based on invalid approximations of the correlation function.

In this project, we have reanalyzed the effect of fixational eye movements on the correlation function of the visual input. In particular, we studied three different Gaussian processes to model these eye movements. An exact expression has been found which shows that the effect on the spatio-temporal correlation function can be described by a Gaussian low pass filter whose cut-off frequency decreases with increasing temporal difference $t - t'$.

If each neuron would respond to the temporal differences $a(x,t) := I(x,t) - I(x,t')$ rather than to the pixel intensities itself, then the fixational eye movements would have exactly the opposite effect on the spatio-temporal correlation function of the $a(x,t)$. They would act as a Gaussian high pass filter whose cut-off frequency decreases with increasing temporal difference $t - t'$. High-pass filtering is desirable in terms of redundancy reduction since the power spectrum of natural images decreases with increasing spatial frequencies. However, the power spectrum decays like a power law which cannot be turned into a flat spectrum by a Gaussian high-pass filter.

In oder to achieve perfect redundancy reduction, it is therefore still necessary to fine-tune the filter shapes of the neurons. This is an important point that contradicts earlier claims by Rucci and Casile [3] that fixational eye movements would render a fine-tuning of the receptive field properties unnecessary. Beyond the scope of this report, we will continue to work out the receptive field properties that are essential for achieving decorrelated neural responses. We intend to publish our results in a scientific journal as soon as the study is completed.

Bibliography

[1] D. J. Field. *Relations between the statistics of natural images and the response properties of cortical cells*. J. Opt. Soc. Am. A, **4**, 2379-2394 (1987).

[2] D. L. Ruderman. *Origins of Scaling in Natural Images*. Vision Res., **37**, 3385-3398 (1997)

[3] M. Rucci, A. Casile. *Fixational instability and natural image statistics: Implications for early visual representations*. Network **16**, 121-138 (2005)

[4] D. W. Dong, J. J. Atick. *Statistics of natural time-varying images.* Network: Computation in Neural Systems, **6**, 345-358 (1995)

[5] H. B. Barlow. *Unsupervised learning*. Neural Comp., **1**, 295-351 (1989)

[6] J. J. Atick, A. N. Redlich. *What Does Retina Know about Natural Scenes?* Neural Computation **4**, 196-210 (1992)

[7] F. L. Van Ness, B. A. Bouman. *Spatial Modulation transfer in the human eye*. J. Opt. Soc. Am., **57**, 401-406 (1967)

[8] S. Martinez-Conde, S. L. Macknik. *The role of fixational eye movements in visual perception.* Nature Reviews, Neuroscience **5**, 229-240 (2004)

[9] L. A. Riggs, F. Ratliff. *The effects of counteracting the normal movements of the eye*. J. Opt. Soc. Am. **42**, 872-873 (1952).

[10] M-O Hongler, Y. L. de Menes, A. Beyeler, J. Jacot. *The Resonant Retina: Exploiting Vibration Noise to Optimally Detect Edges in an Image*. IEEE Trans. on Pattern and Mach Intell. **25**, 1051-1062 (2003).

[11] J.H. van Hateren, A. van der Schaaf. *Independent component filters of natural images compared with simple cells in primary visual cortex*. Proc. R. Soc. Lond. B, **265**, 359-366 (1998)

[12] G. D. Field, E. J. Chichilnisky. *Information Processing in the Primate Retina: Circuitry and Coding*. Annu. Rev. Neurosci., **30**, 1-30 (2007)

[13] G. Buchsbaum, A. Gottschalk. *Trichromacy, Opponent Colours Coding and Optimum Colour Information Transmission in the Retina*. Proc. of the Royal Society of London B, **220**, 89-113 (1983).

[14] F. Sinz, M. Bethge: *Redundancy Reduction in Natural Images - Quantifying the Effect of Orientation Selectivity and Contrast Gain Control*. -, - (accepted) (2008)

[15] M. P. Eckert, G. Buchsbaum. *Efficient Coding of Natural Time Varying Images in the Early Visual System*. Philos. Trans.: Bio. Sci., **339**, 385-395 (1993).

[16] B. Øksendal *Stochastic Differentials Equations. An Introduction with Applications*. Springer, (2005)

[17] L. Råde, B. Westergren. *Springers Mathematische Formeln*. Springer (1996)

[18] A. Papoulis, S. U. Pillai. *Probability, Random Variables and Stochastic Process*. McGraw Hill (2002)

[19] A. Papoulis. *System and Transforms with Application in Optics*. McGraw Hill (1968)

[20] B. A. Wandell. *Foundations of Vision*. Sinauer Associates (1995)