Input-output transformation in the visuo-oculomotor loop: modeling the ocular following response to center-surround stimulation in a probabilistic framework

L. Perrinet, J. Kremkow, F. Barthélemy, G. S. Masson

Institut de Neurosciences Cognitives de la Méditerranée
http://incm.cnrs-mrs.fr

CNRS / Université de la Méditérrannée
Marseille, France

Abstract

The machinery behind Ocular Following Response (OFR) is confronted to ambiguities which are efficiently resolved in the primate visual system. We study here a model of center-surround motion integration in a probabilistic framework and try to identify its different dynamical components by using contrast gain responses to the particular bipartite stimuli.

Motion integration may be modeled as an ideal observer in a probabilistic framework by using bayesian modeling [11] or statistical inference which proved to be successfully applied to the ocular following response [9]. Experiments on primates' oculo-motor recordings concentrated on bipartite stimuli optimized to study the dynamics of information integration for different levels of noise which provide evidence for an orientation selective suppressive effect of the surround on the contrast gain control of local stimuli [1]. We extend here our previous model to integrate different spatial cues: the information propagates to give a command response for ocular response that we could compare with the human behavioral response.

We present results which show that the hypothesis of independence of local measures succesfully accounts for the monotonic integration of the spatial motion signal but that another mechanism should be added to account for suppressive saturation. Adding this, we observed similar dynamics for the contrast gain control mechanisms observed in the behavioral data and in neuro-physiological through in-vivo cortical recording by optical imaging (see accompanying poster number 21 by Reynaud et al. [10]).

1 An "ideal observer" model for Ocular Following Response (OFR)

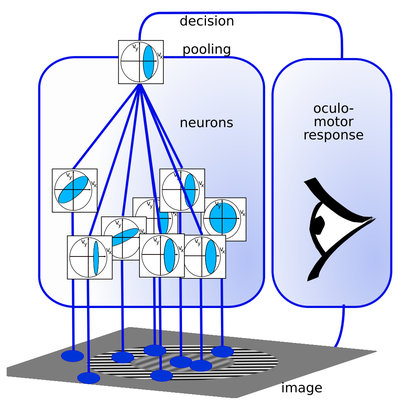


FIGURE 1: **Architecture of the model**. The model consists in the pooling of elementary infomation bits from neurons to provide a decision which reaches the oculomotor system to close the oculomotor loop by moving the eye's position. Every neuron integrates motion information from the image on its receptive field so as to provide a probabilistic representation of the local velocity. This information is then pooled across the different neurons (and hence across space) but also integrated in time, or potentially by other modalities (proprioception). A decision is formed —usually the Maximum A Posteriori (MAP) probability or the mean of the computed global velocity probability— which is transformed accordingly by the oculomotor system to produce an eye movement.

Perceiving motion is a difficult task and we use here the tools of *statistical inference* to model results observed in primates (humans and macaques) on the Ocular Following Response (OFR) [4, 5, 1]. This is particularly relevant in the visual system where this task relies on the integration of different information bits, such as visual local information in a receptive field, the information already detected by neurons which should be computed quickly and efficiently.

As described in [9], we model the observed change in eye position as proportional to a gain defined as the ideal response knowing the given information:

$$\gamma \propto E(\vec{v}|\mathrm{I}) = \int \vec{v} P(\vec{v}|\mathrm{I}) dP(\vec{v}|\mathrm{I}) \tag{1}$$

A difficulty is is to compute $P(\vec{v}|\mathrm{I})$ and a solution which seem to be implemented in the visual system consists in pooling the information from different neurons $\mathbf{n} \in \mathcal{P}$, where $\mathcal{P}$ is the total population of neurons, which have local receptive fields. It is then possible to evaluate the velocity $\vec{v}$ thanks to a stochastic model of the local translation in the image : $\mathrm{I}(\vec{x}, t) = \mathrm{I}(\vec{x} - \vec{v}dt, t - dt) + \nu$, where $\vec{x}$ is the position in the receptive field and ν is a Gaussian noise of variance $\sigma^2_{\mathbf{n}}$. This noise is inversely proportional to Michelson's Contrast (we note the full contrast image : $\mathrm{I}_{100} = C^{-1}.\mathrm{I}$).

Adding a Gaussian prior of variance $\sigma^2_{p,\mathbf{n}}$ favoring slow speed [11], it follows

$$\log P(\vec{v}|\mathrm{I}, \mathbf{n}) = Z - C^2.\mathrm{DI}_{100}(\vec{v})/\sigma^2/2 - (\log \|\vec{v}\|)^2/\sigma^2_{p,\mathbf{n}}/2 \tag{2}$$

where $\mathrm{DI}_{100} = C^{-2}.\mathrm{DI}$ is the contrast-normalized gradient constraint for the local image in the receptive field.

In the range of experiments that we describe here, the image is locally a grating:

$$\mathrm{I}_{100} = \sin(2\pi f(x - \vec{v}.t)) \tag{3}$$

for which DI_{100} is easy to compute: it is well approached by a quadratic function. Every node can thus be characterized by a mean $\vec{v}_{\mathbf{n}}$ and a covariance matrix $C_{\mathbf{n}}$ such that the density $P(\vec{v}|\mathrm{I}, \mathbf{n})$ is a Gaussian:

$$\mathcal{N}(\vec{v}_{\mathbf{n}}, C_{\mathbf{n}}) = \frac{1}{\sqrt{det(2\pi C_{\mathbf{n}})}}.exp(\frac{1}{2}(\vec{v} - \vec{v}_{\mathbf{n}})^T C^{-1}_{\mathbf{n}}(\vec{v} - \vec{v}_{\mathbf{n}}) \tag{4}$$

In the general case, we have $C_{\mathbf{n}}$ given by

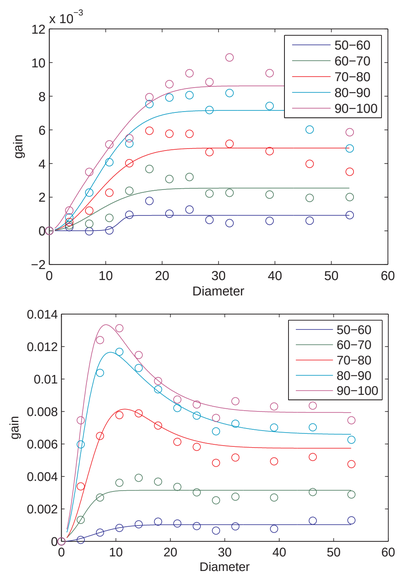
$$\begin{pmatrix} \sigma^2_{\mathbf{n}} & 0 \\ 0 & \sigma^2_2 \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \tag{5}$$

where θ is the orientation of the grating and $\sigma_2 >> \sigma_{\mathbf{n}}$ is the length of the distribution. In particular, the solution to the gain as a function of the contrast, that is the Contrast Response Function (CRF) is thanks to Eqs. 1,2 a Naka Rushton curve [6] of slope 2 [9]:

$$\gamma(C) = \frac{C^2}{C^2_{50} + C^2} \tag{6}$$

with $C_{50} \propto \frac{\sigma_p}{\sigma_{\mathbf{n}}}$. But knowing this integration for a single grating, how to combine these non-linear responses so as to given a global integration of the motion signal?

FIGURE 2: **Effects of size, temporal integration and latency of the ORF response to a center grating**. We present here the gain of the oculomotor response to a central grating (temporal frequency 10 Hz) alone as a function of its diameter. **(Top)** At low frequencies (0.12 cpd), the gain increase monotonotically with the diameter. The curves are well fitted by Eq. 14. **(Bottom)** However in more general conditions, the gain decreases after a given diameter (here 0.7 cpd). This corresponds to a surround suppression which is well captured by Eq. 15.



2 Analytical solution for spatio-temporal integration

A first analytical solution is given thanks to the hypothesis of independence of the measure of velocity on these different neurons [11] which permits to write:

$$\log P(\vec{v}|\mathrm{I}) = \sum_{\mathbf{n} \in \mathcal{P}} \log P(\vec{v}|\mathrm{I}, \mathbf{n}) \tag{7}$$

This ensures that from Eq. 2, $P(\vec{v}|\mathrm{I})$ as a product of Gaussian is also a Gaussian. In that case, the Maximum A Posteriori will again correspond to the mean (see Eq. 1). If the independence hypothesis holds, it follows that the resulting distribution $\mathcal{N}(\vec{v}_m, C)$ obeys

$$\begin{cases} C^{-1} = \sum C^{-1}_{\mathbf{n}} \\ C^{-1}.\vec{v}_m = \sum C^{-1}_{\vec{v}} \vec{v}_{\mathbf{n}} \end{cases} \tag{8}$$

with $C^{-1}_{\mathbf{n}}$ given by

$$\begin{pmatrix} \sigma^{-2}_{\mathbf{n}} & 0 \\ 0 & \sigma^{-2}_2 \end{pmatrix} \begin{pmatrix} \cos(-\theta) & \sin(-\theta) \\ -\sin(-\theta) & \cos(-\theta) \end{pmatrix} \tag{9}$$

Under this particular hypothesis, integration in time is monotonous and is proportional to the time of integration (that is from the onset of the computation). There is therefore under this hypothesis a direct equivalence between the integration time and the contrast of the image.

One should note that this model is similar to the feed-forward Linear/Non-Linear model [2], (see for instance the implementation of [3]) but could be also be implemented in a simpler manner using a spiking neurons network [7, 8]. However, due to the high level of redundancy of information in the visual hierarchy, it is highly likely that these computations are computed thanks to recurrent network with local, lateral or feed-back connections.

3 Effect of size and temporal integration

If we only consider a central grating of speed $\vec{v}_g$, it follows from Eq. 8 that the gain will be proportional to

$$\vec{v}_m = \frac{\sum_{\mathbf{n} \in \mathcal{P}_c} \frac{1}{\sigma^2_x}}{\sum_{\mathbf{n} \in \mathcal{P}_c} \frac{1}{\sigma^2_x} + \sum_{\mathbf{n} \in \mathcal{P}} \frac{1}{\sigma^2_{p,x}}}.\vec{v}_g \tag{10}$$

where $\mathcal{P}_c$ is the population of active central neurons. It appears first that the prior may be simply pooled as $\sigma_P = (\sum_{\vec{x} \in \mathcal{P}} \frac{1}{\sigma^2_{p,\vec{x}}})^{-1/2}$. Under these conditions, we see that the CRF will be again necessarily a Naka Rushton curve [6] of slope 2 with

$$C_{50} \propto \sigma_p \sqrt{\sum_{\mathbf{n} \in \mathcal{P}_c} \frac{1}{\sigma_x}} \tag{11}$$

To model the integration over $\mathcal{P}_c$, we may consider that the density (or weight) of neurons pooling responses for the OFR is a centered Gaussian of visual space with a width of ω we may thus write that

$$\sigma^2_P \sum_{\mathbf{n} \in \mathcal{P}_c} \frac{1}{\sigma^2_x} = \frac{C^2}{C^2_e} . \int_{O \le r \le d} \exp(-\frac{r^2}{2.\omega^2}).2.\pi.r.dr \tag{12}$$

$$= \frac{C^2}{C^2_e}.(1 - \exp(-\frac{d^2}{2.\omega^2})) \tag{13}$$

where d is the diameter and C_e a constant. It follows that from Eqs. 10,

$$\gamma(d) = \frac{C^2}{C^2_e} . \frac{1 - \exp(-\frac{d^2}{2.\omega^2})}{1 + \frac{C^2}{C^2_e}.(1 - \exp(-\frac{d^2}{2.\omega^2}))} \tag{14}$$

where C_e is a constant corresponding to the saturation of the maximum gain of the full field grating.

Even if this family of curves show a relatively good behavior for human responses (see Fig. 2,Top), they don't correspond to the observations for monkeys, especially for high grating frequencies (see Fig. 2, Bottom). In fact, these responses show a suppression after a specific contrast (the so-called *super saturation*) which could not be accounted for this integration model which could only yield monotonic CRFs. One has therefore to add another integration term which accounts for the surround inhibition. We may start by assuming that the surround suppression is initiated by the pooling of information on $\mathcal{P}_c$ toward the null velocity on a similar Gaussian distribution but with a larger size ω_i, similarly as the ratio of gaussian model of [**?**]. It thus comes

$$\gamma(d_c) = \frac{\frac{C^2}{C^2_e}(1 - \exp(-\frac{d^2_c}{2.\omega^2}))}{1 + \frac{C^2}{C^2_e}.(1 - \exp(-\frac{d^2_c}{2.\omega^2})) + \frac{C^2}{C^2_i}.(1 - \exp(-\frac{d^2_c}{2.\omega^2_i}))} \tag{15}$$

where C_i is a constant.

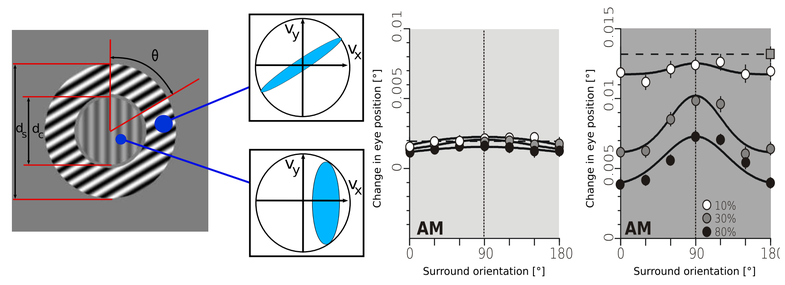


FIGURE 3: **Bipartite Stimuli and its solution**. **(Left)** To further evaluate the effect of the surround, the bipartite stimuli is defined by a central grating of diameter d_c, surrounded by a annulus of diameter d_s filled with a flicker stimulus parametrized by its orientation θ. The flicker corresponds in a first approximation to a null velocity stimulus but from the aperture problem, it presents an elongated probability profile along this orientation. The grating varried in contrast, the probability getting broader as noise increased. **(Right)** Results show that suppression was dependent on the contrast of the grating and weak for a low flicker contrast (Left). However, the suppression was stronger for a full contrast flicker and show a dependance of the suppression as a function of the orientation of the surround as predicted by Eq. 18.

4 Center-surround integration

To challenge the possible origin of the surround, we used the bipartite stimuli (see Fig. 3) for the OFR. it is constituted by a central excitation as before but surrounded by a perturbation of zero net velocity, the *flicker* stimulus. This stimulus is defined as the sum of a grating of speed $\vec{v}_f$ plus the same grating in the opposite direction :

$$\mathrm{I}^f_{100} = \sin(2\pi f(x - \vec{v}_f.t)) + \sin(2\pi f(x + \vec{v}.t)) \tag{16}$$

$$= 2.\sin(2\pi f x).\sin(2\pi f \vec{v}.t) \tag{17}$$

so that we see that $\mathcal{T}(\mathrm{I}^f_{100})(\vec{v}) \propto \|\mathrm{I}^f.\vec{v}\|^2$. The corresponding log-likelihood is a time-modulated quadratic corresponding to a line slanted in the direction perpendicular to $\vec{v}_f$ (see Fig.). It thus comes that the total likelihood $\mathcal{N}(s, \Sigma)$ is given by :

$$s = \sigma^2_X(\frac{1}{\sigma^2_X} + \frac{1}{\sigma^2_P} + \frac{cos(\theta)}{\sigma^2_S}) \tag{18}$$

where θ is the angle of the flicker with the central grating and σ^2_S the weighted integration of information on the flicker.

References

[1] F. V. Barthélemy, I. Vanzetta, and G. S. Masson. Behavioral receptive field for ocular following in humans: Dynamics of spatial summation and center-surround interactions. *J. Neurophysiol.*, (95):3712–26, 2006.

[2] M. Carandini, J. B. Demb, V. Mante, D. J. Tolhurst, Y. Dan, B. A. Olshausen, J. L. Gallant, and N. C. Rust. Do we know what the early visual system does? *J. Neuroscience*, 25(46):10577–97, 2005.

[3] M. Jazayeri and J. Movshon. Optimal representation of sensory information by neural populations. *Nature Neuroscience*, 2006.

[4] G. S. Masson and E. Castet. Parallel motion processing for the initiation of short-latency ocular following in humans. *J. Neuroscience*, 22(12): 5149–63, 2002.

[5] G. S. Masson and L. Stone. From following edges to pursuing objects. *Journal of Neurophysiology*, 88(5):2869–73, 2002.

[6] K. I. Naka and W. A. Rushton. S-potentials from luminosity units in the retina of fish (Cyprinidae). *J. Physiology*, 185(3):587–99, Aug. 1966.

[7] L. Perrinet. Feature detection using spikes : the greedy approach. *Journal of Physiology (Paris)*, 98(4-6):530–9, July-November 2004. doi: 10.1016/j.jphysparis.2005.09.012. URL http://incm.cnrs-mrs.fr/perrinet/publi/perrinet04tauc.pdf.

[8] L. Perrinet. Efficient Source Detection Using Integrate-and-Fire Neurons. In W. D. et al., editor, *ICANN 2005, LNCS 3696*, pages 167–72. URL http://incm.cnrs-mrs.fr/perrinet/publi/perrinet05icann.pdf.

[9] L. Perrinet, E. Barthélemy, Frédéric Castet, and G. S. Masson. Dynamics of motion representation in short-latency ocular following: A two-pathways bayesian model. In *ECVP*, 2005.

[10] A. Reynaud, F. Barthélemy, I. Vanzetta, G. S. Masson, and F. Chavane. Input-output transformation in the visuo-oculomotor loop: comparison of real-time optical imaging recording in v1 to the ocular following response to center-surround stimulation. In *FENS*, 2006.

[11] Y. Weiss, E. P. Simoncelli, and E. H. Adelson. Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6):598–604, Jun 2002.

Find this poster and more at http://incm.cnrs-mrs.fr/perrinet/dynn/LaurentPerrinet/Publications/Perrinet06fens