THE VERTEBRATE RETINA: A MODEL OF SPATIOTEMPORAL IMAGE FILTERING

Jeanny HERAULT and William BEAUDOT

Laboratoire de Traitement d'Images et de Reconnaissance de Formes

I.N.P.G. 46 Av. Félix Viallet 38031 Grenoble Cedex

RÉSUMÉ

Dans cet article, nous présentons un modèle du filtrage spatio-temporel qui a lieu dans le système visuel périphérique des vertébrés (système optique et rétine). Nous montrons qu'une implémentation efficace de ce traitement peut être réalisée sur une architecture numérique classique. Ce filtrage inspiré de la perception visuelle chez les vertébrés permet un traitement temporel analogique, la perception des contours et des couleurs ainsi que celle du mouvement. En outre, l'implémentation VLSI de cet algorithme est quasiment immédiate.

ABSTRACT

In this paper, we present a model of the spatiotemporal filtering that occurs in the peripheral visual system of vertebrates (optical and retinal systems). We show that an efficient implementation can be worked out on a conventional computer architecture. This biology-inspired perceptive filtering allows: a temporal analog processing, a colour and edge perception as well as a motion perception. Moreover, the VLSI implementation can be realised in a straightforward way.

1. Introduction

In this paper, we are interested in the very early stages of the biological vision in order to understand these mechanisms and to integrate them into the paradigm of smart visual sensor. Figure 1 shows a cross-section of the vertebrate eye. Before striking the retina, the continuous spatiotemporal image must pass through the crystalline lens which produces an non-homogeneous spatial filtering. Afterwards this resulting image strikes the retinal surface which is spatially sampled in a non-uniform way with four various photoreceptors in a chromatic respect. Finally, the luminous signal is converted into an electric signal which is altered by the neural filtering of the retina while keeping continuous in time and discrete in space.

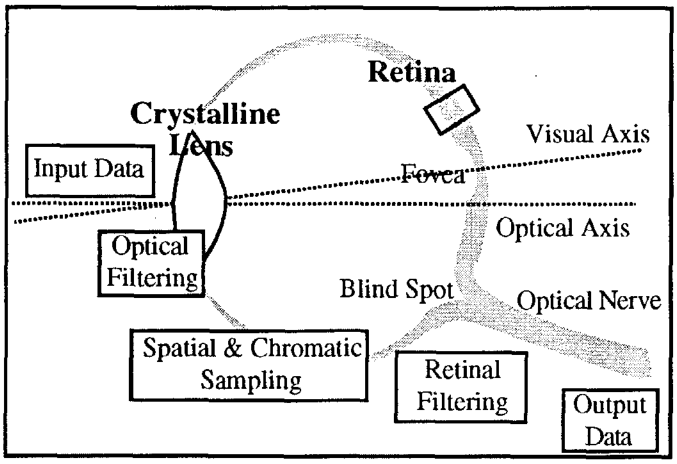

Fig 1. Overview of the Model

The neural structure of the retina can be modelled with some very simple electric components and thus, the underlying filtering can be easily analysed in a signal processing viewpoint with the Fourier and Z transforms (see section 2). Moreover, we propose a recursive implementation of the optical and neural processing that occurs in the vertebrate eye, which leads to a powerful simulation tool, that is simple and efficient [Beaudot and al. 1993]. In the third section, we show how the underlying filtering of the retinal model can be used to detect motion.

2. The Biology-Inspired Model

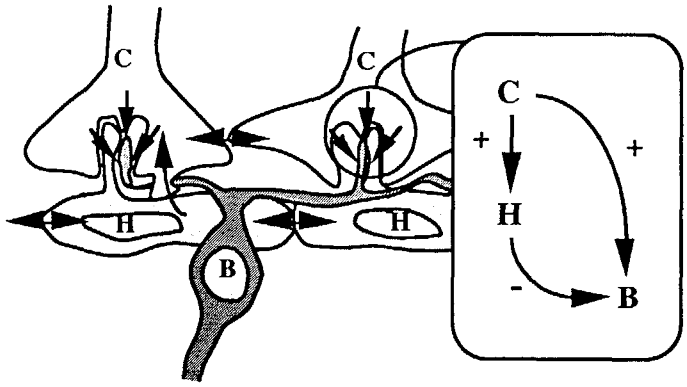

Fig 2. Synaptic transmission between the neurons of the first functional layer. C: cones, H: horizontal cells, B: bipolar cells. The inset shows the flow diagram of signal in the synaptic triad.

2.1 Towards a Mathematical Model

Figure 2 is an enlargement of the neural circuitry in the first functional layer of the retina (more exactly the Outer Plexiform Layer or OPL) which is well-known from a neurophysiological point of view [Dowling 1987]. Three types of neurons interact: the photoreceptor (C) transmits, after the transduction, the signal to the neighbouring photoreceptors (a first lateral diffusion) and to the two other types of neurons - horizontal-cells (H) and bipolar-cells (B) - in an antagonistic way. An horizontal-cell transmits its signal to its neighbours too (a second lateral diffusion) and to a bipolar-cell. Then the bipolar-cell receives two opposite signals which induce a spatiotemporal inhibition, if the temporal neural integration is taken into account.

In order to describe the functional properties of such a natural system, we need to design a model. We propose an electronic model (Figure 3) which is an extension of the Mead's model [Mead and Mahowald 1988]. In this model - analog in time and discrete in space - each lateral diffusion can be considered as a resistive network with a leaky integrator at each node to model the temporal properties of the neurons. Each lateral resistance models the electrical synapses (bidirectional transmission) between the same type of neurons and each vertical resistance models a chemical synapse or a unidirectional transmission. The differential amplifier is for the opposed effect of two synaptic transmissions onto the same neuron (here a bipolar-cell).

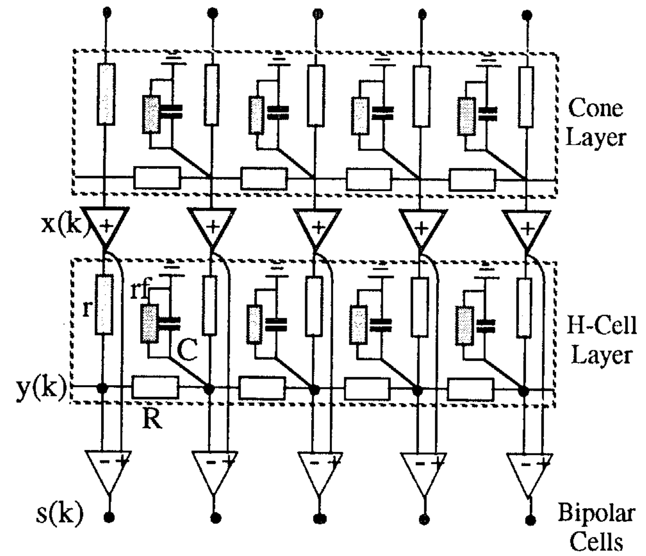

Fig 3. 1-D electronic model of the first functional layer.

We can apply the Kirchhoff's current law at each node of this circuit to obtain the differential equation of the dynamics of the model. Then, the relation between the input x(k,t) and the output y(k,t) of each resistive node is written:

$$y(k,t) = \frac{x(k,t) + \alpha \cdot [y(k-1,t) + y(k+1,t)] - \tau \cdot \dot{y}(k,t)}{1 + 2 \cdot \alpha}$$

This circuit is composed of two resistive grids whose transfer function $F(z_x,z_y,f_t) = Y/X$, using the Fourier transform for the analog time variable and the Z-transform for the discrete space variable, is given in the 2D case by:

$$\mathcal{F}(z_x, z_y, f_t) = \frac{1}{1 + \beta + 4\alpha - \alpha(z_x + z_x^{-1} + z_y + z_y^{-1}) + j2\pi f_t \tau}$$

The space constant α, the leaky constant β and the time constant τ are the parameters of each resistive grid and are related to the electrical components by: $\alpha = r/R$, $\beta = r/r_f$ and $\tau = r.C$. If we denote F_c and F_h the transfer function of each resistive grid, the total spatiotemporal transfer function $G(z_x,z_y,f_t) = S/X$ of the circuit (Fig. 3) is written:

$$G(z_x,z_y,f_t) = F_c(z_x,z_y,f_t) \, . \, [1 - F_h(z_x,z_y,f_t)]$$

The analysis of this transfer function can be very useful to bring out the functional properties of the model and hence the ones of the retina. The study of $F(z_x,z_y,f_t)$ can even lead to an efficient digital implementation of the resistive grid as we will see in the next subsection. Before going farther on, we must note the inseparability of time and space variables in this filter. This remark will have indeed a great importance for the motion processing in the retinal model (section 3).

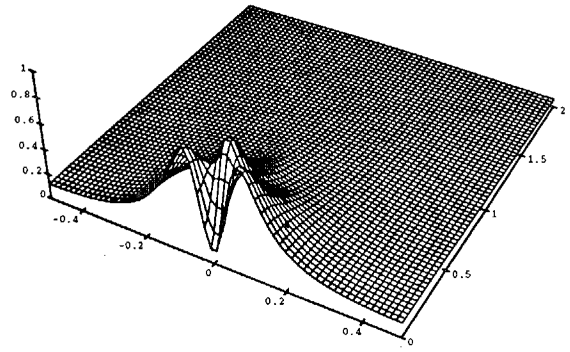

Fig 4. The transfer function of the model for 1D space dimension. left down: spatial frequency, up-right: temporal frequency.

In the signal processing viewpoint, this model acts as a spatio-temporal band-pass filter (figure 4) which induces some very useful properties [Srinivasan and al. 1982]:

- the edge and motion enhancement,
- the reduction of the spatiotemporal redundancy,

2.2 Towards a Digital Model

Let us consider the classical numerical approximation of the time derivative:

$$\dot{y}(k,t) \cong \frac{y(k,t) - y(k,t-\Delta t)}{\Delta t}$$

The digital approximation of the dynamics of a resistive grid is then written:

$$y(k,t) = \underbrace{\frac{\alpha \cdot [y(k-1,t) + y(k+1,t)]}{1 + 2 \cdot \alpha + \frac{\tau}{\Delta t}}}_{\text{Purely Spatial Filter}} + \underbrace{\frac{x(k,t) + \frac{\tau}{\Delta t} \cdot y(k,t-\Delta t)}{1 + 2 \cdot \alpha + \frac{\tau}{\Delta t}}}_{u(k,t)\ \text{Time Dependent Term}}$$

The term u(k,t) can be computed at time t and we showed that the 1-D transfer function Y(z)/U(z) [Beaudot 1992] can be put in the form:

$$Y(z_x)/U(z_x) = H(z_x) \, . \, H(z_x^{-1}) \qquad \text{where} \qquad \mathcal{H}(z) = \frac{1-\sigma}{1-\sigma \cdot z}$$

and σ a parameter related to α, β and τ. Thus, Y(z)/U(z) is the convolution of a causal first order recursive filter by an anticausal first order recursive filter. Then, in the signal domain, we obtain the digital filtering:

$$y(k) = \underbrace{u(k) * h^c(k)}_{y_1(k)} * h^a(k)$$

where $y^1(k) = \sigma \, . \, y^1(k-1) + (1 - \sigma) \, . \, u(k)$ for k = 1 to N

and $y(k) = \sigma \, . \, y(k+1) + (1 - \sigma) \, . \, y^1(k)$ for k= N down to 1

Each scanning needs only one addition and one multiplication per pixel. Thus the algorithm is simple and much more efficient than the use of direct and inverse FFTs. The extension of this decomposition into causal and anticausal recursive filters remains a good approximation in the 2D case. Figure 5 summarizes the 2D algorithm for the computation of the spatiotemporal processing of a resistive grid.

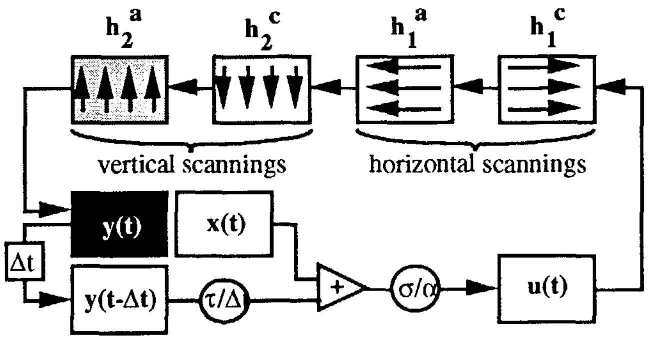

Fig 5. Summary of the 2D algorithm.

The simulation of the whole retinal processing consists simply to use this algorithm for each resistive grid. Figure 6 shows the results of this computation on a real images sequence which contains a strong spatiotemporal noise component.

Fig 6. Results of the OPL retinal processing on a noisy images sequence: the left one is the input and the right is the output.

The pedestrians are illuminated whereas only the edges of the static objects are enhanced. This processing would be very useful for the motion detection.

3. Application to the Motion Detection

3.1 Analysis of the Motion Sensitivity in the OPL

Let us consider a retinal input composed of a moving object with a translational velocity $\mathbf{V} = (v_x, v_y)$. We can easily show that there is a relation between its spatiotemporal spectrum $I(f_x,f_y,f_t)$ and its spatial spectrum $I(f_x,f_y)$:

$$I(f_x,f_y,f_t) = I(f_x,f_y) \,.\, \delta(f_t+v_x.f_x+v_y.f_y) \qquad (3.1)$$

where δ is the Dirac distribution. This expresses there exists a plan given by:

$$f_t+v_x.f_x+v_y.f_y = 0 \qquad (3.2)$$

which is very linked to the velocity direction of the moving object and where all its spectrum is restricted. The result of the OPL filtering on a spatiotemporal input is simply:

$$S(f_x,f_y,f_t) = G(f_x,f_y,f_t) \,.\, I(f_x,f_y,f_t) \qquad (3.3)$$

Let us substitute (3.1) for $I(f_x,f_y,f_t)$ in (3.3), then we can express the time frequency f_t in $G(f_x,f_y,f_t)$ with the velocity and the space frequencies according to (3.2):

$$S(f_x,f_y,f_t) = G(f_x,f_y,-v_x.f_x-v_y.f_y) \,.\, I(f_x,f_y) \,.\, \delta(f_t+v_x.f_x+v_y.f_y)$$

This last expression means that the spatiotemporal filtering of the first functional layer of the retina is equivalent to a purely spatial filtering $G(f_x,f_y,-v_x.f_x-v_y.f_y)$ when the input is a pattern satisfying (3.1). Figure 7 shows the space-equivalent transfer function of the OPL filter when the input is static: the retinal acts as an isotropic band-pass spatial filter.

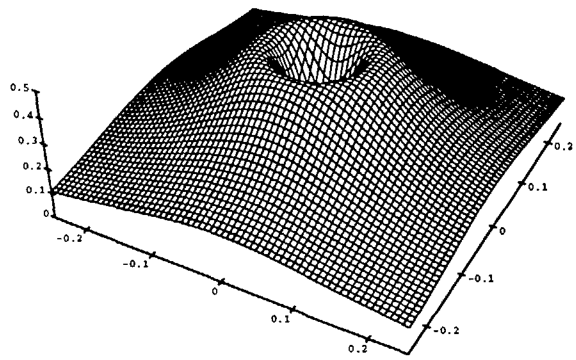

Fig 7. The 2-D transfer function of OPL for a static input.

Figure 8 shows the space-equivalent transfer function of the OPL filter for a moving object. The spatial filtering is no longer isotropic and highlights two directions in the spatial frequency domain: the medium spatial frequencies in the direction of the motion are enhanced whereas in the perpendicular direction of the motion they are more deadened.

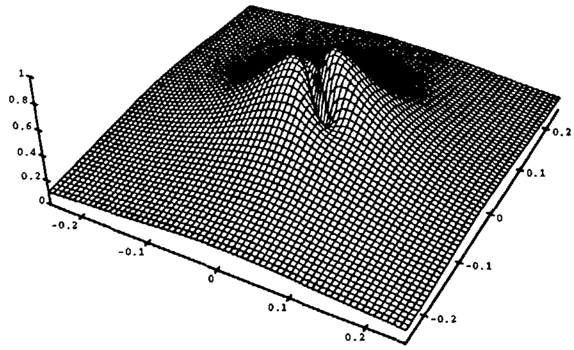

Fig 8. Module of the OPL transfer function for a moving input.

3.2 Analysis of the Motion Sensitivity in the IPL

The previous theoretical results and the figure 6 show that our retinal model can be very sensitive to the moving objects but also to the static ones. But we think also that the retina is the good place to make a motion detection. Then, what is the suitable processing after the OPL in order to realize a motion detection ?

We showed in [Hérault and Beaudot 1993] from the theory of the matched filtering that this filtering must be a temporal filter $\Psi(f_t)$ which is the convolution of a first order time high-pass filter with a time low-pass filter:

$$\Psi(f_t) = (-\, j2\pi f_t.\tau_1) \,/\, [\, (1+j2\pi f_t.\tau_1).(1+b+j2\pi f_t.\tau_2)\,]$$

This temporal filter has also a space-equivalent form when the input is a moving pattern. Figure 9 shows the resulting transfer function: the high spatial frequencies in the direction of motion are always enhanced and the very low frequencies in the direction of motion are always removed.

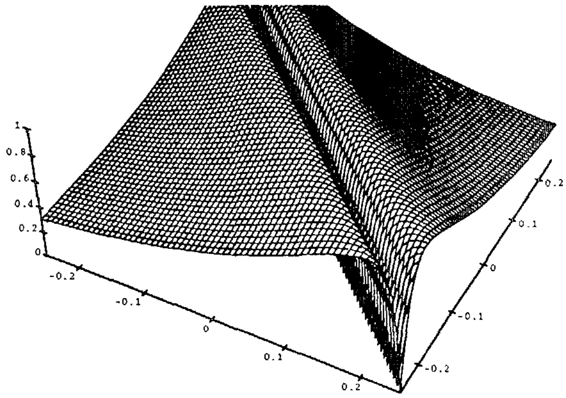

Fig 9. The space-equivalent transfer function of the IPL filtering for a moving input.

The existence of a temporal high-pass filtering in the second functional layer of the retina (more exactly the Inner Plexiform Layer or IPL) was already emphasized in [Richter and Ullman 1982], but the capacity of the retina to detect motion had never been clearly shown before. Thus, we identify Ψ with the IPL processing.

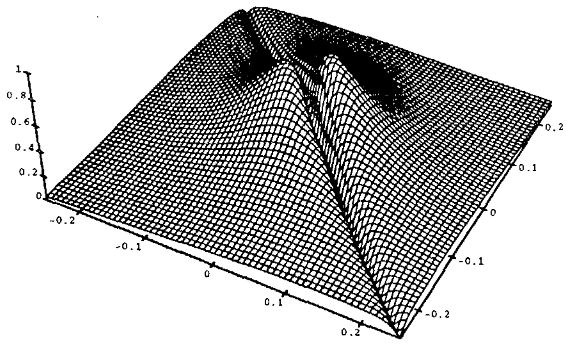

Fig 10. The space-equivalent transfer function of OPL o IPL for a moving input.

The convolution of the OPL filter $G(f_x,f_y,f_t)$ with the IPL filter $\Psi(f_t)$ leads to a new filter which has also a space-equivalent filter when the input is a moving pattern (figure 10). The retinal output is then a signal containing high spatial frequencies in the direction of motion and low spatial frequencies in the perpendicular direction.

In order to achieve a useful motion detection, it is needed to differentiate the bright moving objects on a dark background and the dark moving objects on a bright background. This is an intrinsic problem of the motion perception which is rarely taken into account. But it is known [Wässle and Boycott 1991] for a long time that there exists separate sensations ("channels") for the perception of darkness and lightness in the biological visual systems. In the retina, ON (for the lightness perception) and OFF-pathways (for the darkness perception) are always found as early as the bipolar level. Thus, we introduced a nonlinearity by separating the bipolar output into positive and negative signals and the IPL filtering was applied on each of them. Figure 11 shows the result of the motion detection on the negative part of the OPL output (pedestrians are darker than the background): static edges and spatiotemporal noise have been removed.

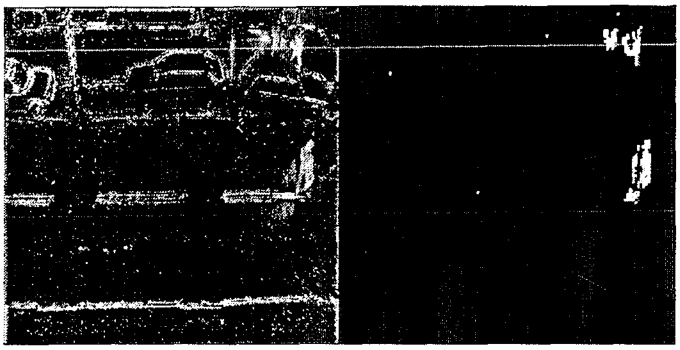

Fig 11. Results of the motion detection on a real images sequence. The left one is the result after the OPL processing and the right one is the result after the IPL processing on the negative part of the left one: the spatiotemporal noise has been removed.

4. Conclusion

In this paper, we proposed a simple and efficient biology-inspired system in order to perform some early visual processing (such as edge detection and motion detection), which can be worked out just as well on a classical computer architecture as on analog VLSI architecture. Moreover, the results shows that the signal processing tools are suitable to model, analyse, and replicate a biological neural system. Currently we are working on the motion selectivity and the motion estimation. The addition of some nonlinearities is also investigated so as to show that the retinal processing is locally adapted to the input data such as homomorphic filtering.

References

Beaudot W (1992) Traitement Spatio-Temporel d'Images par Modèle de la Rétine. Rapport de DEA, Laboratoire TIRF, Grenoble, France, June 1993

Beaudot W, Palagi P, Hérault J (1993) Realistic Simulation Tool for Early Visual Processing including Space, Time and Colour Data. International Workshop on Artificial Neural Networks 93, Barcelona, June 1993

Dowling JE (1987) The retina: An approachable part of the brain. MA: Havard University Press, Cambridge

Hérault J, Beaudot W (1993) Motion Processing in the Retina: About a Velocity Matched Filter. European Symposium on Artificial Neural Networks 93, Brussels, April 1993

Mead C, Mahowald M (1988) A silicon model for early visual processing. Neural Networks, 1:1, 91-97

Richter J, Ullman S (1982) A model for the temporal organization of X- and Y-type receptive fields in the primate retina. Biological Cybernetics 43, 127-145

Srinivasan MV, Laughlin SB, Dubs A (1982) Predictive coding: a fresh view of inhibition in the retina. Proceedings of The Royal Society of London, B 216, pp 427-459

Wässle H, Boycott BB (1991) Functional architecture of the mammalian retina. Physiological Reviews, Vol. 71, n°2, pp 447-480