ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 47-54

# From retinal circuits to motion processing: a neuromorphic approach to velocity estimation

Antonio Torralba & Jeanny Hérault

Laboratoire de Traitement d'Images et Reconnaissance de Formes,
INPG, 46 Avenue Félix Viallet, 38031 Grenoble Cedex, France

**Abstract.** A new energy-based algorithm for motion estimation is presented. It uses wide-band spatiotemporal filters (with non-separable variables) inspired from neuromorphic circuits which are simple, robust, noise resistant and easily implementable in analogue VLSI circuits for real-time on-chip estimation of optical flow-field. Results are presented for one and two dimensional motion, including an original solution to the aperture problem.

## 1. Introduction

Recently, several algorithms for motion estimation that can be implemented with VLSI circuits have been proposed [1]. They generally use correlation- or gradient-based methods. In this paper we present a new algorithm that belongs to the family of energy-based methods and that can also be implemented in analogue VLSI circuits. On the contrary of the narrow-band spatiotemporal Gabor-type wavelets normally used in energy-based methods [2, 3, 4] that require a heavy computational load, we propose here a new approach: it consists of wide-band velocity-tuned spatiotemporal filters inspired from neuromorphic circuits [5]. On one hand, this kind of circuits can be efficiently implemented and are robust with respect to mismatching tolerances. On the other hand, energy-based methods are known to be robust with respect to noise in the input signal. In this paper we present the one dimensional motion for seek of clarity, and we give the extension to the 2-D case, with an original solution to the aperture problem.

## 2. Model of translational motion and motion sensitive filters

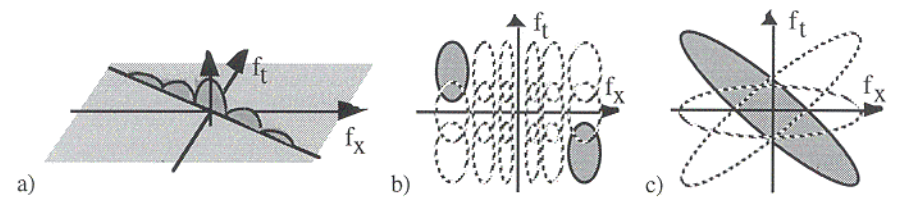



Fig. 1. a) Motion in the frequency domain, and its processing by b) a bank of numerous narrow-band Gabor-type filters (nine filters), c) our bank of wide-band filters (three filters).

The basic property of uniform motion is that the energy is contained, in the spatiotemporal-frequency domain, in a plane which orientation depends on motion velocity and direction. In the case of one-dimensional motion, the spatio-($x$) temporal-($t$) intensity function of a moving pattern with velocity v is given by $i(x,t)=i(x-vt)$, its Fourier transform being $I(f_x,f_t)=I(f_x).\delta(f_t+v.f_x)$, as illustrated in figure 1.a. The energy plane is determined by the equation $f_t+v.f_x=0$. Estimating the velocity will correspond to the problem of detecting the orientation of the energy plane in the spatiotemporal-frequency domain. In this approach, classical methods generally use a bank of narrow-band spatiotemporal filters to sample the spatiotemporal frequency domain. (figure 1.b). Each filter is tuned to some velocity and centred on some spatial frequency. The output of such a filter will be maximum for a moving pattern which

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 47-54

spectrum is contained inside of the pass band of the filter. For the one dimensional case, with the method proposed by Heeger [2] it is necessary to use at least nine filters, as shown in figure 1b. The technique we describe in this paper uses wide-band velocity-tuned spatiotemporal filters (figure 1c) inspired from neuromorphic circuits. It interestingly reduces the number of filters to be used to sample the spatiotemporal frequency domain, in the 1-D case, only three filters are necessary.

## 3. Neuromorphic motion sensor

Several neural networks with motion selectivity properties have been proposed in the literature [6, 7]. The one we propose here, figure 2.a, uses both excitatory and inhibitory lateral connections in order to provide motion selectivity (the local feedback connection takes into account the temporal behaviour, as it will be seen). The neural network will be analysed from a signal processing point of view by using an electronic circuit representation which is inspired from neuromorphic circuits [5]. The use of both inhibitory and excitatory connections will permit to obtain a simple circuit allowing a full control of all the filter parameters such as maximum spatial bandwidth, preferred direction and velocity selectivity.

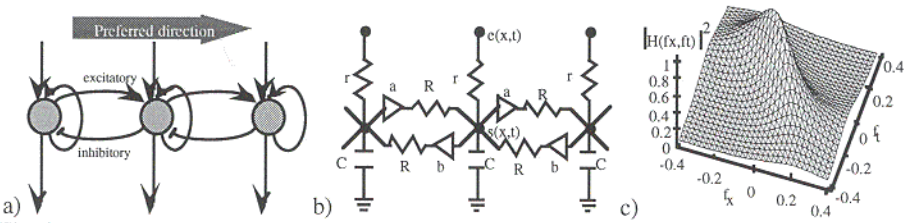



**Fig. 2 .** a) Neural network for motion estimation by temporal units with excitatory & inhibitory synapses. b) Equivalent neuromorphic circuit temporally causal and possibly causal/anti causal spatial behaviour. c) Example of spatio-temporal transfer function (compare with fig 1a,c).

Figure 2.b shows the neuromorphic circuit used to model the neural network. The input of the neurone is figured by the generator $e(x,t)$ and the synaptic resistance $r$. Its output state $s(x,t)$ is the voltage across a capacitor $C$ figuring the membrane capacitor, which will evolve continuously with time. The lateral "unidirectional" connections are composed of a gain (a or b, figuring synaptic coefficients that would be positive for the excitatory connections and negative for the inhibitory ones) and a resistance R figuring the corresponding synaptic resistance. This model is easily described by means of its Fourier analysis in spatial and temporal frequencies. The filter is discrete in space but continuous in time, and its differential difference equation is obtained by applying the Kirchoff's current law to each node:

$$e(x,t) = s(x,t)\cdot[1+2\alpha] - \alpha\left[a\cdot s(x-\Delta x,t) + b\cdot s(x+\Delta x,t)\right] + \tau\,\partial s(x,t)/\partial t$$

$e(x,t)$ is the input and $s(x,t)$ is the output. $\Delta x$ is the spacing between pixels. The other circuit parameters are: $\alpha=r/R$ and $\tau=r\,C$. By applying the Fourier transforms in spatial frequencies $f_x$, and temporal frequencies $f_t$, we can write:

$$H(f_x,f_t) = \frac{1}{1+\alpha\cdot\left[2-(a+b)\cos(2\pi f_x\Delta x)\right] + j\tau\left[2\pi f_t + \alpha/\tau\cdot(a-b)\sin(2\pi f_x\Delta x)\right]}$$

When $a+b>0$ this filter corresponds to a low-pass wide-band spatiotemporal filter. Figure 2.c shows the shape of the module of this function which complies with figure 1.c if $\sin(2\pi f_x\Delta x)\approx 2\pi f_x\Delta x$. It can be easily shown that the spatiotemporal stability of the filter is assured if the condition $|a+b|<2+1/\alpha$ is verified [8]. For better understanding, one can use the low spatial frequencies approximation: $\cos(x)\approx 1-x^2/2$

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 47-54

and $\sin(x)\approx x$. Considering $\Delta x=1$ (velocity will be given in pixels per second), the transfer function can be approximated by the expression:

$$H(f_x, f_t) \approx H_o \Big/ \left[1 + 2f_x^2/\Delta B_x^2 + j2\left(f_t + v_o f_x\right)/\Delta B_t\right] \qquad (1)$$

Where: $H_o=1/[1+\alpha(2-a-b)]$, which is the gain for low frequencies, and

$$\Delta B_x = \frac{1}{\pi\alpha^{1/2}}\sqrt{\frac{1+\alpha\left[2-(a+b)\right]}{a+b}},\ \Delta B_t = \frac{1}{\pi\tau}\left[1+\alpha\left[2-(a+b)\right]\right] \text{ and } v_o = \frac{\alpha(a-b)}{\tau}$$

For a pattern moving with velocity $v_o$, the energy is contained in the frequencies given by the equation $f_t+v_o f_x=0$, therefore the complex term in the denominator of (1) will be zero for the spatiotemporal frequencies that contain the input energy. Therefore, $\Delta B_x$ is the equivalent spatial bandwidth of the filter for a pattern moving with velocity $v_o$. For other input velocities the equivalent spatial bandwidth of the filter will be smaller. $\Delta B_t$ is a parameter that will determine the velocity selectivity of the filter. For numerical simulation we use $a+b=2$ which assures stability and allows to obtain very simple relations between the filter parameters and the components of the circuit. These relations are: $H_o=1$, $\alpha=1/(2(\pi\Delta B_x)^2)$, $\tau=\Delta B_x/\Delta B_t\cdot(2\alpha)^{1/2}$, $a=1+\tau v_o/(2\alpha)$ and $b=2-a$. As we can see, all the parameters ($\Delta B_x$, $\Delta B_t$ and $v_o$) can be chosen independently with the only restriction of $\Delta B_x\leq 0.5$ in order to be in the linear approximation of $\sin(2\pi f_x\Delta x)$. This would not be possible when considering only excitatory connections. In order to study the behaviour of the network in the time domain we will consider as input signal, an unit impulse with velocity $v$ and amplitude $A$: $A\cdot\delta(x-v\cdot t)$. The Fourier transform of such a signal is a Dirac plane $A\cdot\delta(f_t+v\cdot f_x)$. Therefore, the output will be the product $A\cdot H(f_x,f_t)\cdot\delta(f_t+v\cdot f_x)$. This can also be written as $A\cdot H(f_x,-v\cdot f_x)\cdot\delta(f_t+v\cdot f_x)$. This is equivalent to a purely spatial filter which is function of velocity $G(f_x)=H(f_x,-v\cdot f_x)$. Then, the output of the filter is the signal $A\cdot g(x-v\cdot t)$ where $g(x)$ is the inverse Fourier transform of $G(f_x)$. Using expression (1) and with $p=2\pi j f_x$ ($p$: Laplace variable) we can write:

$$G(p) \approx \frac{-2\pi^2\cdot\Delta B_x^2}{p^2 - 2\cdot\pi\cdot(v-v_o)\cdot\Delta B_x^2/\Delta B_t\cdot p - 2\cdot\pi^2\cdot\Delta B_x^2} = -\frac{P_1\cdot P_2}{(p-P_1)\cdot(p+P_2)}$$

$P_1$ and $P_2$ are positive numbers depending on filter parameters. The function can be decomposed in two parts, one causal ($P_1$) and one anti causal ($P_2$). Applying the inverse generalised Laplace transform we obtain the function:

$$g(x) \approx \frac{\pi\cdot\Delta B_x/\sqrt{2}}{\sqrt{1+2\cdot\dfrac{(v-v_o)^2}{\Delta V^2}}}\cdot\begin{cases} e^{P_1\cdot x} & \text{if } x<0 \\ e^{-P_2\cdot x} & \text{if } x\geq 0 \end{cases}$$

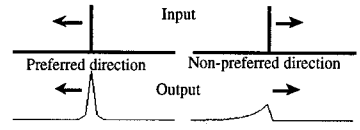
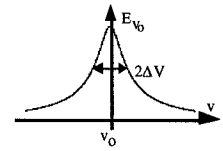



Figure at the right-hand side shows the spatial output of the filter for two directions of motion. The output amplitude is higher when motion is in the preferred direction. The energy of the output can be calculated as:

$$E_{v_o} = A^2\cdot\int_{-\infty}^{\infty}|G(f_x)|^2\,dx \approx \frac{A^2\cdot\pi\cdot\Delta B_x}{2\sqrt{2}\sqrt{1+2\dfrac{(v-v_o)^2}{\Delta V^2}}}$$




Where the velocity selectivity $\Delta V$ is calculated as $\Delta V=2\Delta B_t/\Delta B_x$. It shows the dependence of the energy on the velocity (figure): the output energy is maximum for motion at the preferred velocity ($v_o$). This expression will be also valid for input

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 47-54

patterns with white energy spectrum. This complies with the pre-processing function of the retina which is known to whiten the input signals [9] as we will discuss later.

## 4. Estimation of motion

As the energy is a function of velocity, it can be used for motion estimation. However, the output energy is also a function of input energy, therefore, we need to combine the outputs of several filters tuned to different velocities in order to have a measurement of motion, independently of input energy. If we use two filters tuned to opposite directions (velocities $v_o$ and $-v_o$), for a moving impulse or for a pattern with white spectrum, we can write using the energy expression given above:

$$1/E^2_{-V_o} - 1/E^2_{V_o} = K \cdot 8 \cdot v \cdot v_o / \Delta V^2$$

Where $K$ is a constant depending on filter parameters and input energy. This expression, linear for the input velocity, remains dependent on the input energy, and needs to be normalised. Using a third filter tuned to null velocity (spatially symmetric) with output energy $E_0$, we can obtain the expression:

$$1/E^2_{-V_o} + 1/E^2_{V_o} - 2/E^2_0 = K \cdot 4 \cdot v_o^2 / \Delta V^2$$

which doesn't depend on input velocity. Therefore, the velocity $v$ can be calculated as:

$$v = \frac{v_o}{2} \cdot \frac{E^2_{V_o} - E^2_{-V_o}}{E^2_{V_o} + E^2_{-V_o} - 2 \cdot E^2_{V_o} \cdot E^2_{-V_o} / E^2_0} \qquad (2)$$

This formula is valid for all velocities, the limit being given by the existing noise. Because this expression doesn't depend on filter parameters, it is not necessary to have filters very selective to velocity. On the contrary, filters with low selectivity will behave better with respect to noise. For all tested sequences we obtain good results with $\Delta V = v_o$. Figs. 1.c and 2.c show the shape of these filters in the frequency domain.

## 5. Application to "real" 1-D images

When dealing with real images, several problems arise. They come from i) the input, with noise and spatial aliasing (and temporal aliasing for time discretised sequences in numerical simulations), and ii) the motion estimation algorithm itself. The most important one is that input signals never satisfy a white spectrum condition.

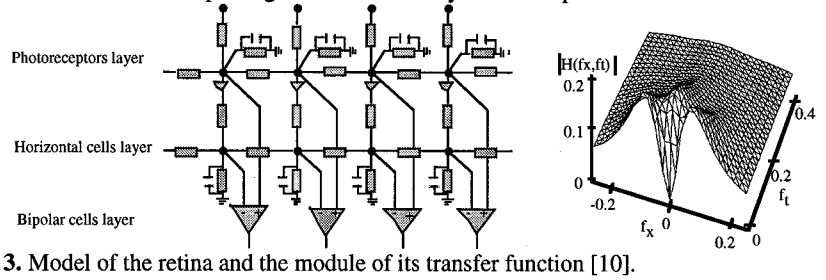



**Fig. 3.** Model of the retina and the module of its transfer function [10].

Atick [9] showed that natural images have roughly a $1/f$ spectrum that can be compensated by some pre-filtering of the input signal, just as the one realised by the retina in the visual system. Figure 3 shows the neuromorphic model of the retina and the shape of its transfer function [10]: in the bandwidth of the input signal, the retina behaves as a high pass filter which will compensate for the 1/f spectrum of the input signal. More, for high spatiotemporal frequencies it is low pass and reduces noise and aliasing. The output signal of our model of retina can now be applied to the three

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 47-54

velocity filters, in parallel as said in section 4: velocity is estimated from the combination of their output energies. In order to have local estimations of velocity we must integrate these energies only over a limited region, at the expense of accuracy. And in order to meet accuracy, we must integrate over large regions, at the expense of badly localised smooth velocity fields. In order to respect motion discontinuities we must use small integration regions, at the expense of biased velocity estimates. Integration is provided by a regularising network similar to that of the photoreceptor layer applied to the square of the outputs of the velocity-tuned filters.

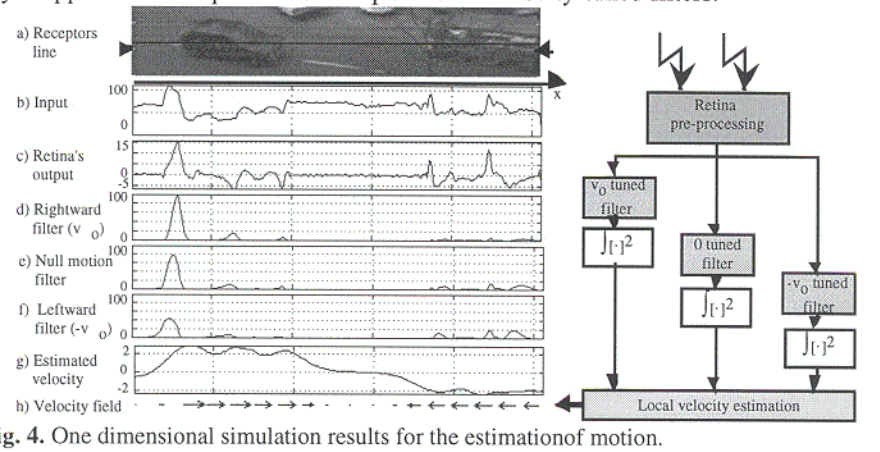



**Fig. 4.** One dimensional simulation results for the estimationof motion.

Fig. 4 shows the simulation results on a real sequence and the processing architecture. The parameters of the velocity-tuned filters are $v_O$=3, $\Delta V$=3 and $\Delta B_x$=0.15. The one dimensional input sequence is issued from a two dimensional one by taking only one line of receptors (a). The sequence consists in two vehicles moving in opposite directions. The input signal (b) is firstly pre-filtered by the retina. The output signal (c) has a zero mean value and has a wider bandwidth than the input signal. The output of the retina is applied to the three selective filters. Motion in the rightward direction (vehicle at left-hand side) produces a larger response in the filter tuned for rightward motion. The right-hand side vehicle moving leftward produces the largest output in the leftward motion tuned filter. The figure shows the squared output of the three filters before integration (d, e and f). Once local energy is computed we can use expression (2) to estimate the velocity at each pixel (g and h).

## 6. Formulas for the two-dimensional case

The hereabove presented results can be extended for two spatial dimensions, however, some new difficulty appears: the one of the aperture problem. In order to simplify expressions we will use vector notations: the position will be defined by the vector $\mathbf{x} = \{x, y\}$ and the spatial frequencies by the vector $\mathbf{f_S} = \{f_x, f_y\}$. The velocity vectors for a filter ($\mathbf{v_O}$) and for the image ($\mathbf{v}$) are respectively defined by:
$\mathbf{v_O} = \{v_{xO}, v_{yO}\} = \{v_O\, cos(\alpha), v_O\, sin(\alpha)\}$ and $\mathbf{v} = \{v_x, v_y\} = \{v\, cos(\beta), v\, sin(\beta)\}$
$v_O$ and $\alpha$ are the module and direction of the velocity-tuned filter, $v$ and $\beta$: same for the moving pattern.

### 6. 1. Transfer function

The neuromorphic circuit showed in figure 2.b can be extended to two dimensions by connecting each node to its four neighbours with both excitatory and inhibitory connections. The synaptic coefficients ($a$ and $b$ in the one dimensional model) are $a_x$

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 47-54

and $b_x$ for the horizontal connections and $a_y$ and $b_y$ for the vertical ones. The same analysis done previously can be extended to two dimensions. The low spatial frequencies approximation of the transfer function for a filter tuned to velocity $\mathbf{v_o}$ becomes:

$$H(f_s, f_t) \approx 1\Big/\left[\left(1+2|f_s|^2/\Delta B_s^2\right)+j\,2\left(f_t+v_o{}^T f_s\right)/\Delta B_t\right]$$

with $\gamma=r/R=1/(2(\pi\Delta B_s)^2)$, and $\tau=r\,C=\Delta B_s/\Delta B_t\cdot(2\gamma)^{1/2}$, $a_x=1+\tau v_o\cos(\alpha)/(2\gamma)$, $b_x=2-a_x$, $a_y=1+\tau v_o\sin(\alpha)/(2\gamma)$ and $b_y=2-a_y$. $\Delta B_s$ and $\Delta B_t$ are defined in the same way as $\Delta B_x$ and $\Delta B_t$ in the one dimensional case.

**6. 2. Velocity estimation for a moving impulse**

As in the one dimensional case, we will study the output of the filter when the input signal is a moving impulse and we will derive the way in which the velocity vector can be estimated. We model the input signal as $i(\mathbf{x},t) = \delta(\mathbf{x}-\mathbf{v}\cdot t)$ $=\delta(x-v\cos(\beta)\,t,\ y-v\cdot\sin(\beta)\cdot t)$, its Fourier transform is: $I(\mathbf{f_s},\ f_t) = \delta(f_t+\mathbf{v}^T\ \mathbf{f_s})$ $=\delta(f_t+v.\cos(\beta)\cdot f_x+v.\sin(\beta)\cdot f_y)$. The frequency spectrum of the filter's output is: $O(\mathbf{f_s}, f_t) = H(\mathbf{f_s}, f_t)\ I(\mathbf{f_s}, f_t) = H(\mathbf{f_s},\ -\mathbf{v}^T\ \mathbf{f_s})\ \ \delta(f_t+\mathbf{v}^T\ \mathbf{f_s})$. This means that it can be considered as a spatial pattern moving at velocity $\mathbf{v}$: $O(\mathbf{f_s},\ f_t) = G_v(\mathbf{f_s})\cdot\delta(f_t+\mathbf{v}^T\ \mathbf{f_s})$. Applying the inverse Fourier transform, we obtain the response of the filter: $o(\mathbf{x},t) = g_v(\mathbf{x}-\mathbf{v}\ t)$ the energy of which writes:

$$E_\alpha = \iint |g_v(\mathbf{x})|^2\,d\mathbf{x} = \iint \left|H\left(\mathbf{f_s}, -\mathbf{v}^T\mathbf{f_s}\right)\right|^2 d\mathbf{f_s} \approx 0.5\,\Delta B_s^2\,\pi\Big/\sqrt{1+\left(\Delta v_x^2+\Delta v_y^2\right)/\Delta v_0^2}$$

With: $\Delta v_x = v_o\,cos(\alpha)-v\,cos(\beta)$, $\Delta v_y = v_o\,sin(\alpha)-v\,sin(\beta)$, $\Delta v_o = 2\Delta B_t\,/\,\Delta B_s$

This expression is an approximation, it can be verified numerically that the error remains within 5%. Here again, the output energy will be maximum when the input impulse moves with the velocity ($\mathbf{v_0}$) to which the filter is tuned. Using again three wide-band filters tuned to velocities 0, $v_0\ \angle\alpha$ and$v_0\ \angle\alpha+\pi$ (opposite direction), we can estimate the component of the input velocity in the direction of the angle $\alpha$ by combining the output energies in the same manner as in the 1-D case:

$$\frac{v_o}{2}\cdot\frac{E_\alpha^2-E_{\alpha+\pi}^2}{E_\alpha^2+E_{\alpha+\pi}^2-2\cdot E_\alpha^2\cdot E_{\alpha+\pi}^2/E_{v=0}^2}\approx v\cdot cos(\alpha-\beta) \qquad (3)$$

Figure 5.a shows the meaning of this result for a moving dot. It is also valid when the input consists in a highly textured moving pattern with a white spectrum. In these situations, motion can be determined without any ambiguity even when the estimation is realised with local integration. With real images, as previously said, a prefiltering will be necessary in order to be near to the white spectrum condition.

**6. 3. Velocity estimation for a moving contour**

When the input signal is highly oriented (as in the case of a long contour), the condition of white spectrum will not be verified in any direction and the previous results will not hold. It is the well known "aperture problem" and then, with local image analysis, we will have access only to the normal component $\mathbf{v}$ of velocity. This situation doesn't exist in the one dimensional case. In order to study the filter's response in this situation we will consider as input signal an infinitely long moving contour: $i(\mathbf{x},t)=\delta(\mathbf{x}^T\cdot\mathbf{u_n}-v\cdot t)=\delta(x\cdot\cos(\beta)+y\cdot\sin(\beta)-v\cdot t)$ where $\mathbf{u_n}$ is an unitary vector normal to the contour, $v$ is the module of the normal velocity and $\beta$ is the angle between the contour and the vertical axis ($y$) of the image. In this case, it can be shown [8] that formula (3) turns to:

$$\frac{v_o}{2}\cdot\frac{E_\alpha^2-E_{\alpha+\pi}^2}{E_\alpha^2+E_{\alpha+\pi}^2-2\cdot E_\alpha^2\cdot E_{\alpha+\pi}^2/E_{v=0}^2}\approx\frac{v}{cos(\alpha-\beta)} \qquad (4)$$

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 47-54

The result is shown graphically in figure 5.b. This result and the previous one show that the estimation of velocity will be different with and without the aperture problem.

## 7. Two-Dimensional motion estimation and the aperture problem

In the general 2-D case, velocity estimation requires at least two sets of 3 filters tuned to two orthogonal velocity vectors (figure 5.a bold). Because the 0 velocity has no orientation, this leads to a bank of 5 filters (2 for velocities $+v_o$ and $-v_o$ in the horizontal direction, 2 for velocities $+v_o$ and $-v_o$ in the vertical direction, and one for velocity 0). We will show that this is not sufficient because of the aperture problem. 1/ We have said that for wide-band signals (dots or highly textured objects) the region of integration should be restricted in order to respect motion discontinuities. In this case, the result of section 6.2. will apply in order to estimate the velocity vector. 2/ However, in presence of the aperture problem, a large integration region is required, though producing poorly localised velocity fields. In this case, with small integration regions, the results of section 6.3 would apply.

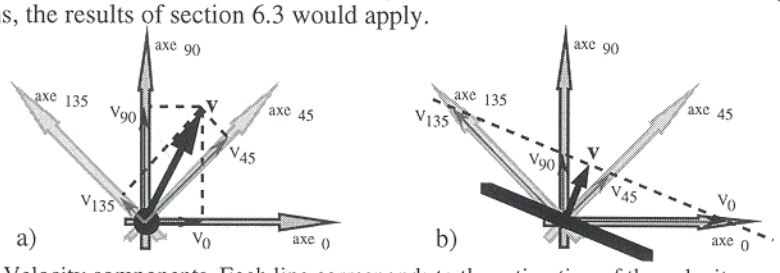



**Fig. 5.** Velocity components. Each line corresponds to the estimation of the velocity component in one direction. a) For a moving impulsion, and b) for a moving contour. (see text).

But, the problem we are faced now is the one of determining which case (aperture problem or not) is present at any location of the processed image. We propose an original method to solve it, by the use of two other sets of filters, shifted of 45° with respect to the previous ones. Fig. 5 shows the complete set of velocity tuned filters: in the directions of 0, 45, 90, 135°, and the estimated velocities in each case: we see that, only for the aperture problem, the affixes of the estimated vectors lie on a straight line which is parallel to the moving contour. This is a mean to determine which case is present at the current pixel location. Moreover, using four filters can help in exploiting the redundancy for noise reduction. The coherence between all the measurements can be used as a confidence value.

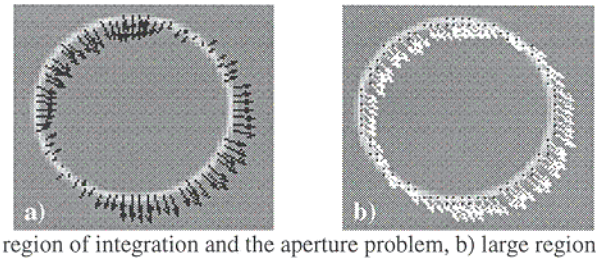



**Fig. 6.** a) Small region of integration and the aperture problem, b) large region of integration.

Figure 6 shows the results of simulation with a moving circumference. When using a small integration region the aperture problem is present and the algorithm produces normal-velocity estimates (fig. 6.a), however, when using an integration region of the size of the circle, the algorithm produces the real velocity estimation (fig. 6.b).

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 47-54

## 8. Results with a real sequence of images

Figure 7 shows the simulation results for a real sequence taken by a static camera. The scene contains four moving objects, one pedestrian on the upper-left side of the image, a taxi in the centre and two other vehicles at the low right-hand and left-hand sides of the image. The figure shows a) the input image, b) the output of the retina and c) the estimated velocity field. The estimation has been done with only five wide-band spatiotemporal filters. The integration region is large enough in order to prevent the aperture problem. This algorithm has been tested with other sequences providing results with the same accuracy as methods using Gabor-type filters [2] but, with a dramatic reduction of the complexity of the algorithm (5 filters instead of the 36 used in the method proposed by Heeger). It allows a fast computation: 1 image ($256^2$) per second for retina filtering and velocity estimation on a SUN UltraSparc work-station.

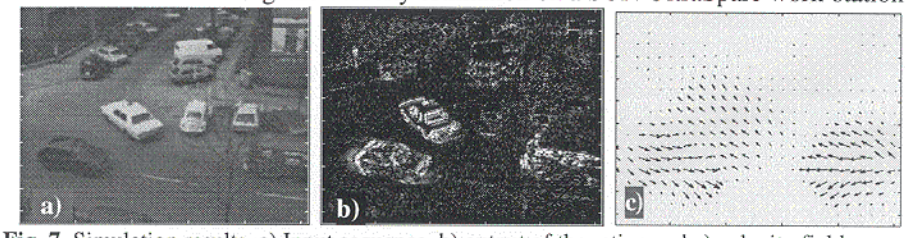



**Fig. 7.** Simulation results. a) Input sequence, b) output of the retina and c) velocity field.

## 9. Discussion

In this paper we have presented a new energy-based algorithm for motion estimation based on wide-band spatiotemporal filters that are inspired from neuromorphic circuits. The proposed architecture is simpler than other methods which use narrow-band Gabor-type wavelets, it is robust with respect to input and structural noise. The simplicity of the filters used here allows one to envisage their implementation in analogue VLSI circuits for real-time on-chip estimation of optical flow-field.

## 10. References

1. Sarpeshkar R., Kramer J., Indiveri G., Koch C. (1996) Analog VLSI architectures for motion processing: from fundamental limits to system applications. Proc. IEEE 84:969-987
2. Heeger D. J. (1987) Model for the extraction of image flow. Journal Optical Society of America, vol. 4, no. 8, 1455-71
3. Watson A. B., Ahumada A. J. (1985) Model of human visual-motion sensing. Journal Optical Society of America, vol. 2, no. 2, 322-42
4. Adelson E. H., Bergen J. R. (1985) Spatiotemporal energy models for the perception of motion. J. Opt. Soc. Am. A/vol. 2, no. 2, 284-299.
5. Douglas R., Mahowald M., Mead C. (1995) Neuromorphic analogue VLSI. Annu. Rev. Neurosci. 18:255-81.
6. Hirahara M., Nagano T. (1993) A neural network model for visual motion detection that can explain psychophysical and neurophysiological phenomena. Biol. Cyb., 68, 247-252.
7. Barlow HB., Levick WR. (1965) The mechanism of directionally selective units in rabbit's retina. Journal of Physiology, 178-549.
8. Torralba A. (1996) Traitements spatio-temporels non linéaires dans un modèle de vision d'inspiration biologique. Technical report, Laboratory TIRF, Grenoble, France.
9. Atick J., Redlich A. (1992) What does the retina know about natural scenes? Neural Comput. 4:196-210
10. Beaudot W. (1992) Le traitement neuronal de l'information dans la rétine des vertébrés. PhD thesis, Laboratory TIRF, Grenoble, France.