Machine Vision and Applications (1998) 11: 83–95

Machine Vision and Applications

Analog computation for phase-based disparity estimation: continuous and discrete models

Bruno Crespi[1], **Alex G. Cozzi**[2], **Luigi Raffo**[3], **Silvio Sabatini**[4]

[1] IRST, Istituto per la Ricerca Scientifica e Tecnologica, I-38050 Povo, Trento, Italy; e-mail: crespi@irst.itc.it
[2] Department of Neurophysiology, Ruhr University, Bochum, Germany
[3] DIEE, University of Cagliari, Piazza d'Armi, I-09123 Cagliari, Italy
[4] DIBE, University of Genoa, Via Opera Pia 11a, I-16145 Genova, Italy

Received: 27 October 1997 / Accepted: 18 June 1998

Abstract. The analog implementation of a phase-based technique for disparity estimation is discussed. This technique is based on the convolution of images with Gabor filters. The article shows that by replacing the Gaussian envelope with other envelopes, the convolution operation is equivalent to the solution of a system of differential equations, whose order is related to the smoothness of the kernel. A detailed comparison between the disparity estimates obtained using these kernels and those obtained using the standard filter is presented. The discretization of the model leads to lattice networks in which the number of connections per node required to perform convolution is limited to the first few nearest neighbors. The short connection length makes these filter suitable for analog VLSI implementation, for which the number of connection per node is a crucial factor. Experimental measures on a prototype CMOS 17-node chip validated the approach.

Key words: Phase-based technique – Differential equations – Disparity – Analog computation – VLSI

1 Introduction

The capabilities of the human (or animal) visual system to process sensory information are still beyond the power of today's computing machines even when perception is considered at the lowest levels. Although early vision processes can be defined in term of mathematically well-defined procedures, their implementation is still hampered by the huge amount of computing power needed to perform the mathematical procedures (filtering, correlations, matching, etc.). However, recent technological advances, such as the development of massive parallel computers and special-purpose VLSI circuits, indicate that it will soon be possible to develop real-time vision systems that combine different perceptual modalities.

In particular, analog VLSI technology has reached a point where it is realistic to imagine and project the construction of "neuromorphic" computers in which computation is realized in the systems dynamics. In contrast to digital system where computation is obtained by performing a well-defined sequence of logical operations, analog computation results as the output of the cooperative interaction of the (electronic) elements of a network. This approach is particularly attractive for those early vision applications in which flexibility of computation is less important than the speed to go from signals to action. In such devices, sensory information is directly mapped in the electrical variables, and computation is carried out massively in parallel with high efficiency and speed.

Effective collective computation (small area and low power consumption) on real-time signals is only possible when taking full advantage of the CMOS technology at transistor and/or circuit levels. Several examples are present in the literature, illustrating how powerful this approach can be. Among these contributions we can distinguish two schools of thought. The first one, problem-oriented, starts from a computational approach to visual tasks, formulated as variational problems. Analog VLSI circuits are considered as "analog computers" on which to map the related Euler-Lagrange equations [10, 12, 15]. The second one, stressing the similarities between the physics of analog CMOS and brain circuits, has accomplished the goal of reproducing the various blocks of first stages in the visual pathway: photoreceptors, retinal cells, and optical nerve. Applications of these blocks have been considered in relation to contrast adaptation, stereo correspondence, motion, etc. [6, 14, 22, 25].

The implementation of vision algorithms by means of analog circuits has to overcome several difficulties. One of the principal problem regards network connectivity. A high number of connections per node constitutes a serious obstacle for algorithm implementability. In analog computations, indeed, a physical link is necessary to exchange information between two nodes of a computational network. In this context, it is important to be able to devise algorithms that minimize the number of connections per node necessary to implement a given functionality [1, 14, 19, 21].

This work deals with the analog implementation of a technique for depth and motion estimation [2, 8, 9] that is based on the convolution of the incoming images with Ga-

Correspondence to: B. Crespi

bor filters. It is shown that the convolution operations can be realized by means of solutions of a linear system of differential equations (DE). The DE order is related to the number of interconnections per node required to perform the computation in the corresponding discrete network. The influence of the envelope change on the quality of the filter estimates is investigated in detail and compared with the standard convolution results.

The model presented in this work was used to design and fabricate a demonstrative analog chip that performs convolutions. The chip is constituted of a linear array of 17 basic cells, each cell is connected with its next and next nearest neighbors. This research activity was part of the Esprit Basic Research Project 8503 "Cormorant", whose objective was the design of algorithmic solutions for early vision processing that are suitable for analog hardware implementation.

2 The phase-difference-based technique

Stereoscopy is a technique used to extract depth from a pair of images taken from slightly different view points. Depth can be computed from the relative position of corresponding points in the two images. In a first approximation, the positions of corresponding points are related by a one-dimensional shift, the *disparity*, along the direction of the epipolar lines. In the simplest configuration, in which optical axes are parallel, epipolar lines are parallel to the line joining the optical centers of the cameras.

Several techniques have been developed to estimate the position disparity of corresponding points from pairs of epipolar lines. The phase-difference-based technique, proposed by [20], is based on the computation of the difference between the phases of the convolutions of the two stereo images with complex bandpass filters - the Gabor filters.

Since it is assumed that the two signals, denoted as $f_R(x)$ and $f_L(x)$, are *locally* related by a shift,

$$f_L(x + d(x_0)/2) \approx f_R(x - d(x_0)/2) \tag{1}$$

in the neighborhood of each point x_0 the local k Fourier components of $f_L(x)$ and $f_R(x)$ are related by a phase difference equal to $\Delta\psi(k) = \psi_L(k) - \psi_R(k) = k\,d$.

The local Fourier analysis is performed by convolving the images with a local envelope,

$$\begin{aligned} u_{L/R}(x, k_0) &= \int dy\; G(x-y)\, e^{i\,k_0\,(x-y)}\, f_{L/R}(y) \\ &= \rho_{L/R}(x)\, e^{i\,\psi_{L/R}(x)}\;, \end{aligned} \tag{2}$$

where $G(x-y)$ is a function that goes to zero as $|x-y|$ increases, and k_0 is the tuning frequency of the filter.

As a function of spatial position, the phase of the filter response, $\psi(x)$, has a quasi-*linear* behavior. In fact, the phase derivative $\psi_x(x)$ is generally close to the tuning frequency, $\psi_x(x) \approx k_0$. Linearity allows an accurate estimation of the shift from the phase difference by means of a second-order expansion in $d(x)$ [1]

$$\begin{aligned} \Delta\psi(x) &= \psi^L(x) - \psi^R(x) = \psi(x - d(x)/2) - \psi(x + d(x)/2) \\ &= \psi_x(x)\, d(x) + O(d^3(x))\;. \end{aligned} \tag{3}$$

[1] The second-order term is zero, $\psi_{xx}/2!\,(d^2(x) - d^2(x))/4 = 0$.

Using the average of the derivatives, disparity is given by

$$d(x) \approx 2\,\frac{[\Delta\psi(x)]_{2\,\pi}}{\psi_x^L(x) + \psi_x^R(x)}\;. \tag{4}$$

In its standard formulation, the local envelope is chosen to be a Gaussian:

$$G(x) = \frac{1}{\sqrt{2\,\pi}\,\sigma}\, e^{-\frac{(x-x_0)^2}{2\sigma^2}}\;;\quad \widehat{G}(k) = e^{-\frac{\sigma^2 k^2}{2}}\;. \tag{5}$$

The choice is dictated by the fact that the Gaussian envelope is best localized in both space and frequency: the product of the uncertainties in space and in frequency attains the minimum value, $\Delta x \cdot \Delta k = 1$. For other filters, the product of uncertainties is always greater, $\Delta x \cdot \Delta k > 1$. A relation between σ and k_0 is imposed by assuming that the bandwidth of the Gabor filter is one octave, that is $\lambda_0 = \frac{2\pi}{3}\sigma \approx 2.1\sigma$. This means the Gaussian envelope contains a complete wavelength λ_0 in the interval $[x_0 - \sigma,\, x_0 + \sigma]$.

Disparity computation is reliable when phases are linear. Around singular points where the amplitude vanishes, $\rho(x) = 0$, the phase develops stong nonlinearities. As a consequence, approximation (4) fails. Therefore, computation at point x is accepted only if

$$\left|\,(\psi'(x) - k_0)\,\right| < k_0\, T_S\;, \tag{6}$$

where $T_S \approx 0.4$ [8]. The left side of (6) measures the difference between the peak frequency, k_0, and the local frequency, $\psi'(x)$.

From the perspective of analog computation, an algorithm is implementable if it can be reduced to local operations. The phase-based approach to disparity estimation is local except for the convolution operation. However, if the Gaussian envelope is replaced by other kernels, convolution can be transformed into the solution of a set of differential equations whose order is related to the number of connections per node necessary to implement the filter in a discrete model.

3 Local Gabor-like filters

Because of the convolution theorem,

$$u(x) = \int dy\; G(x-y)\, e^{i\,k_0\,(x-y)}\, f(y) \tag{7}$$

can be expressed as the inverse Fourier transform of

$$\widehat{u}(k) = \widehat{G}(k - k_0)\,\widehat{f}(k)\;, \tag{8}$$

where $\widehat{u}(k)$, $\widehat{G}(k)$, and $\widehat{f}(k)$ indicate the Fourier transforms of function $u(x)$, $G(x)$, and $f(x)$, respectively.

If $\widehat{G}(k)$ is different from zero for every k, it is possible to write that

$$\widehat{G}^{-1}(k - k_0)\,\widehat{u}(k) = \widehat{f}(k) \tag{9}$$

or, by expanding the Gaussian kernel ($\mu = \sigma/\sqrt{2}$)

$$\left[1 + \mu^2(k-k_0)^2 + \frac{\mu^4(k-k_0)^4}{2!} + \ldots\right]\widehat{u}(k) = \widehat{f}(k)\;. \tag{10}$$

Since multiplication by $(k - k_0)$ in frequency space corresponds to the application of differential operator $[\,\frac{d}{dx} - i\,k_0\,]$

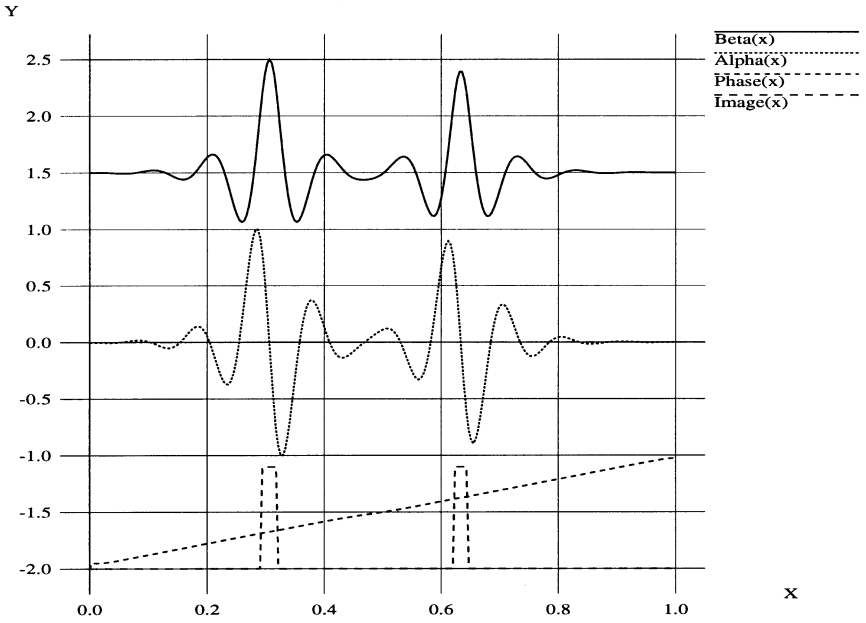


Fig. 1. Functions $\alpha(x)$ and $\beta(x)$, the phase, and the spike-like left 1D image

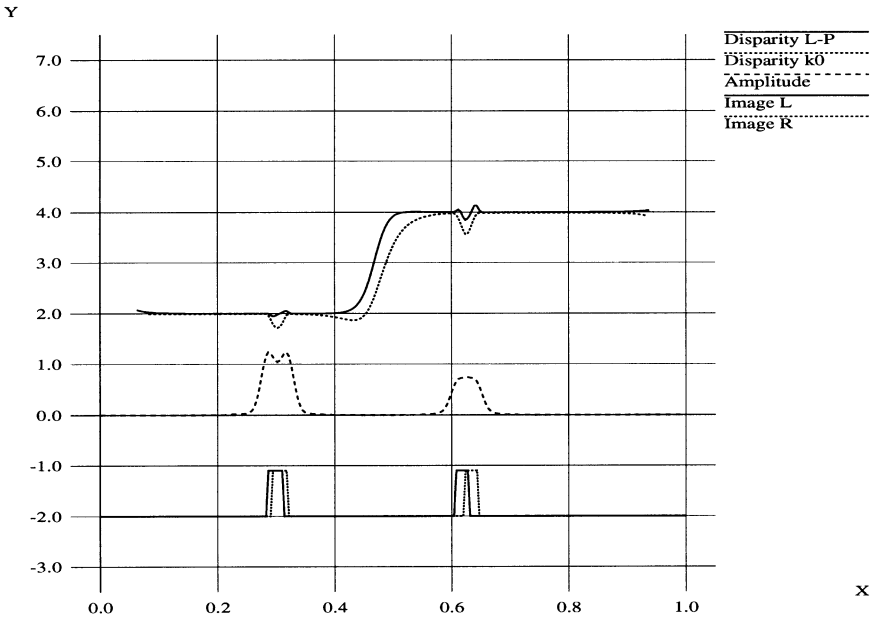


Fig. 2. The computed disparity (in pixels) using the local, $\psi'(x)$, and the tuning frequency, k_0, and the product of the amplitudes, $(\rho_L^2(x) \times \rho_R^2(x))$. In the *lower part*, the *left* and *right* 1D images are shown: the first spike is shifted by two pixels, the second by four pixels, as correctly estimated

in the x-space, the inverse Fourier transform of Eq. 9 generates a series of local envelopes. The envelope's smoothness increases with the order of the corresponding differential equation.

3.1 The cusp filter

The first-order expansion in $\mu = 1/\gamma$

$$\left[\gamma^2 + (k - k_0)^2\right] \widehat{u}(k) = \gamma^2 \widehat{f}(k) \tag{11}$$

generates the cusp envelope

$$C(x) = \frac{\gamma}{2}\, e^{-\gamma|x|} \ . \tag{12}$$

This implies that integral equation

$$u(x) = \int dy e^{-\gamma|x-y| + i\, k_0\,(x-y)}\, f(y) \tag{13}$$

can be turned into a linear differential equation in which the input signal $f(x)$ acts as a source term

$$(\frac{d}{dx} - i\, k_0)(\frac{d}{dx} - i\, k_0)\, u(x) + \gamma^2 u(x) = \gamma^2 f(x) \ , \tag{14}$$

i.e.,

$$-u_{xx}(x) + 2i\, k_0\, u_x(x) + (k_0^2 + \gamma^2)\, u(x) = \gamma^2\, f(x) \ . \tag{15}$$

Showing real and imaginary parts, $u(x) = \alpha(x) + i\,\beta(x)$, differential Eq. 15 can be written as the ODE system

$$\begin{aligned} -\alpha_{xx} - 2\, k_0\, \beta_x + (\gamma^2 + k_0^2)\, \alpha &= \gamma^2\, f(x) \\ -\beta_{xx} + 2\, k_0\, \alpha_x + (\gamma^2 + k_0^2)\, \beta &= 0 \ . \end{aligned} \tag{16}$$

The system can be rewritten as a single fourth-order differential equation. For function $\alpha(x)$, we find

$$-\alpha_{xxxx} + 2\,(\gamma^2 - k_0^2)\,\alpha_{xx} - E^2\alpha = \gamma^2\,(f_{xx} - E\,f)\ . \qquad (17)$$

In Fig. 1, the solution of the DE system for a one-dimensional (1D) image is shown. The two functions reproduce the real and imaginary parts of the "oscillating" cusp filter, as has to be expected for delta-like signals. The phase is close to a straight line. Figure 2 illustrates the disparity function obtained using both the local frequency, $\psi'(x)$, and the tuning frequency k_0. The computed disparity is correct: in fact, the first spike is shifted by two pixels, the second by four pixels, and the transition between the two disparity values is smooth. Ripples in the disparity function in correspondence to image structures are a consequence of the sharp peak of the cusp envelope.

Figure 6 shows the disparity map obtained from a stereo pair of computer-generated images using the cusp envelope. The computer-generated stereo pair is shown in Figs. 3 and 4. The true disparity map is displayed in Fig. 5 [2]

As usual, points where the algorithm is not reliable are drawn black. In general, points were disparity can be estimates are in the neighborhood of visual features, such as edges or textures. The size of the neighborhood depends on the width of the filter. The wavy behavior of the disparity estimates at the borders of the tiles corresponds to the ripples evidenced in Fig. 2. For this image, a comparison between errors and densities of the cusp and the Gaussian filters is shown in Fig. 7. The error measure is defined as the absolute difference between the true disparity (TD) and the estimated value (D). In the following, we always refer to the average value of the error defined as the normalized sum of the errors at the nonsingular points,

$$\text{error} = \frac{1}{N}\sum_{i.j} |TD_{i,j} - D_{i,j}|\ , \qquad (18)$$

where i and j run over the nonsingular points, and N is the number of nonsingular points. Density is the ratio between the number of nonsingular points and the total number of points.

Filters that are smoother at the origin and have a sharper spectrum decay can be generated from Eq. 10. However, this causes an increase of the order of the differential equations.

3.2 The quartic filter

If the k^4 terms are kept in the expansion of the Gaussian kernel. Eq. 10, the quartic envelope is found, see Fig. 8. Its analytical expression is (see Appendix A).

$$Q(x) = \frac{N}{\mu}\,\sin\left(\frac{\pi}{8} + \gamma_2|x|/\mu\right)\,e^{-\gamma_1|x|/\mu}\ , \qquad (19)$$

where $\gamma_1 = 2^{1/4}\cos\left(\frac{\pi}{8}\right)$, $\gamma_2 = 2^{1/4}\sin\left(\frac{\pi}{8}\right)$ and N is a normalization factor. The smoother behavior at the origin corresponds to the sharper decay of the frequency spectrum.

The solution of equation

[2] These images are orthographic projections generated with the ray-tracing program of the Computer Graphics group of Prof. Fellner at the Bonn University. Images and their description can be found at the following address: http://www-dbv.cs.uni-bonn/ ft/stereo.html.

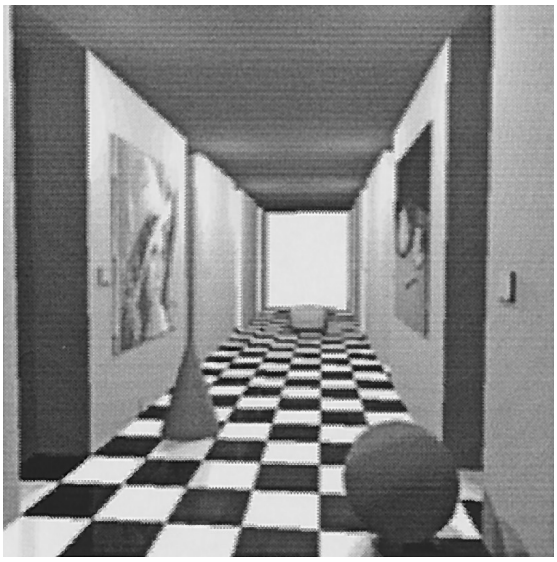


Fig. 3. The computer-generated "corridor image" (*left*), size 256 × 256 pixels

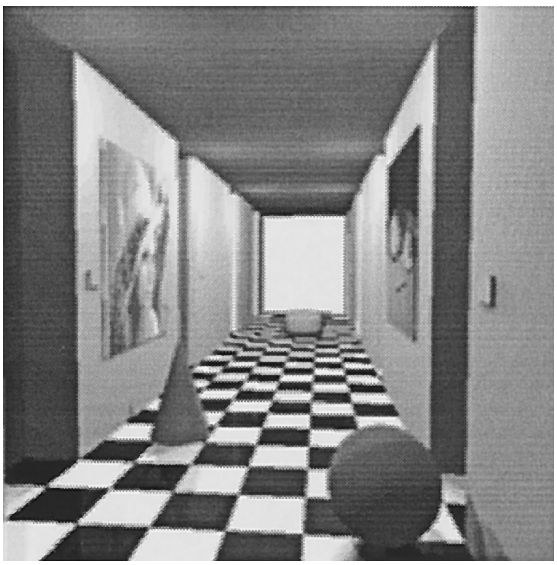


Fig. 4. The computer-generated "corridor image" (*right*), size 256 × 256 pixels

$$\mu^2\,(\frac{1}{i}\frac{d}{dx} - k_0)^2\,u(x) + \mu^4\frac{1}{2!}\,(\frac{1}{i}\frac{d}{dx} - k_0)^4\,u(x) + u(x) = f(x) \qquad (20)$$

i.e.,

$$\frac{1}{2}\mu^4\left(u_{xxxx} - 4i\,k_0\,u_{xxx} - 6\,k_0^2\,u_{xx} + 4i\,k_0^3\,u_x + k_0^4\,u\right) + \mu^2\left(-u_{xx} + 2i\,k_0\,u_x + k_0^2\,u\right) + \alpha = f(x) \qquad (21)$$

is the convolution of the source term, $f(x)$, with an "oscillating" quartic filter.

Making plain real and imaginary parts, Eq. 20 leads to the ODE system:

$$\frac{1}{2}\mu^4\left(\alpha_{xxxx} + 4\,k_0\,\beta_{xxx} - 6\,k_0^2\,\alpha_{xx} - 4\,k_0^3\,\beta_x + k_0^4\,\alpha\right) + \mu^2\left(-\alpha_{xx} - 2\,k_0\,\beta_x + k_0^2\,\alpha\right) + \alpha = f(x)\ , \qquad (22)$$

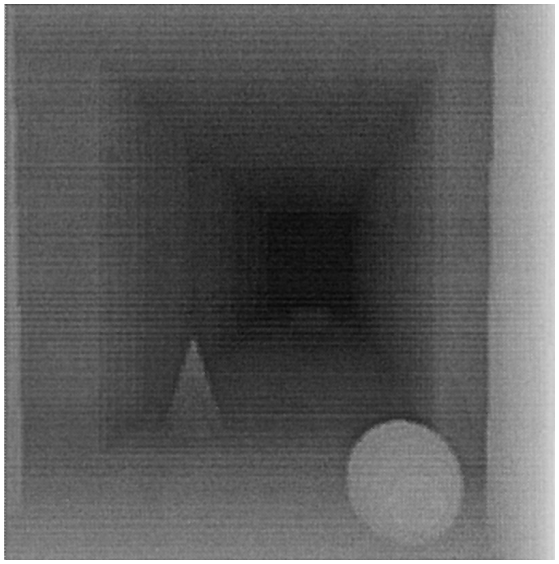


Fig. 5. True disparity map

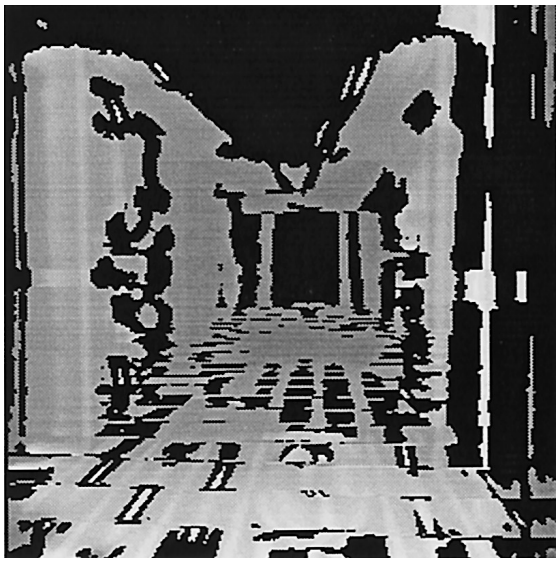


Fig. 6. Disparity map estimated using the cusp filter. Singular points are black

$$\frac{1}{2}\mu^4\left(\beta_{xxxx} - 4\,k_0\,\alpha_{xxx} - 6\,k_0^2\,\beta_{xx} + 4\,k_0^3\,\alpha_x + k_0^4\,\beta\right)$$
$$+\mu^2\left(-\beta_{xx} + 2\,k_0\,\alpha_x + k_0^2\,\beta\right) + \beta = 0\ .$$

Figure 9 shows the disparity function estimated using the quartic filter for the stereo signal of Fig. 1. The "ripples" found in the cusp case are not present.

4 Performance of the Gabor-like filters

From the computational perspective, the cusp and the quartic filters are interesting because they require low-order DE systems. In fact, in the corresponding discrete systems, the DE order is related to the number of connections per node necessary to realize the convolution, see Sect. 5. However, the change of envelope modifies the quality of disparity estimates [7]. In this section, a comparison among the cusp, the quartic, and the Gaussian filters is presented.

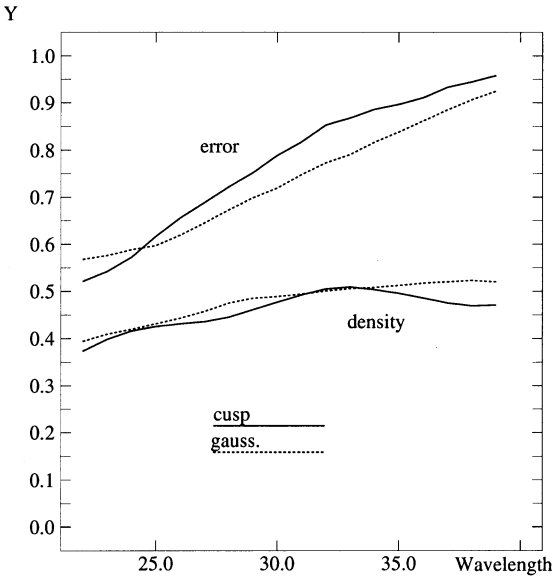


Fig. 7. Comparison between the cusp and the Gaussian filters for the image shown in Fig. 3 for several wavelengths. The y-axis represents the mean value of the error or the densitiy of singular points, as defined in (18)

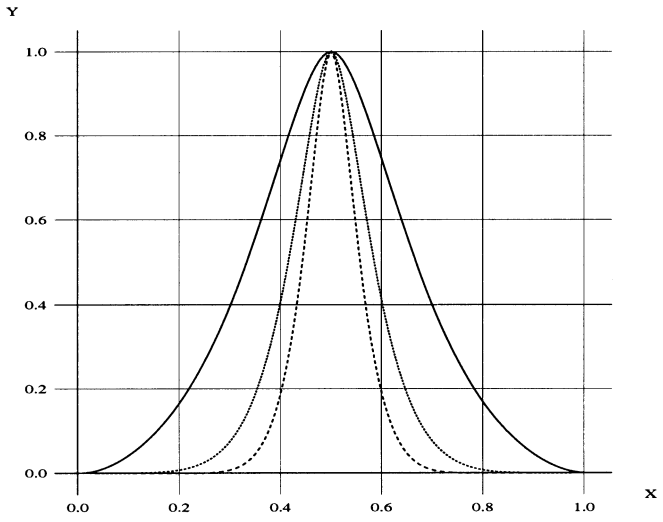


Fig. 8. The quartic envelope for 3 values of μ

4.1 Basic tests

For well-defined stereo signals characterized by simple disparity fields, the disparity estimates of the three filters are comparable. Figures 10 and 11 show the disparity estimates of the three filters for an image with spike-like features and for a random-dot image, respectively. In general, the cusp estimate is not as smooth as the quartic and the Gaussian estimates.

To obtain a significative comparison between the envelopes, the filters' performances were evaluated for a large number of random-dots 1D image pairs of size N pixels. Performance is measured in terms of error and density as defined in Sect. 3.1. It is important to notice that the random-dot 1D images used in the following simulations have zero mean, i.e., $\int f(x)\,dx = 0$, and that the shifts between left and right signals are integer.

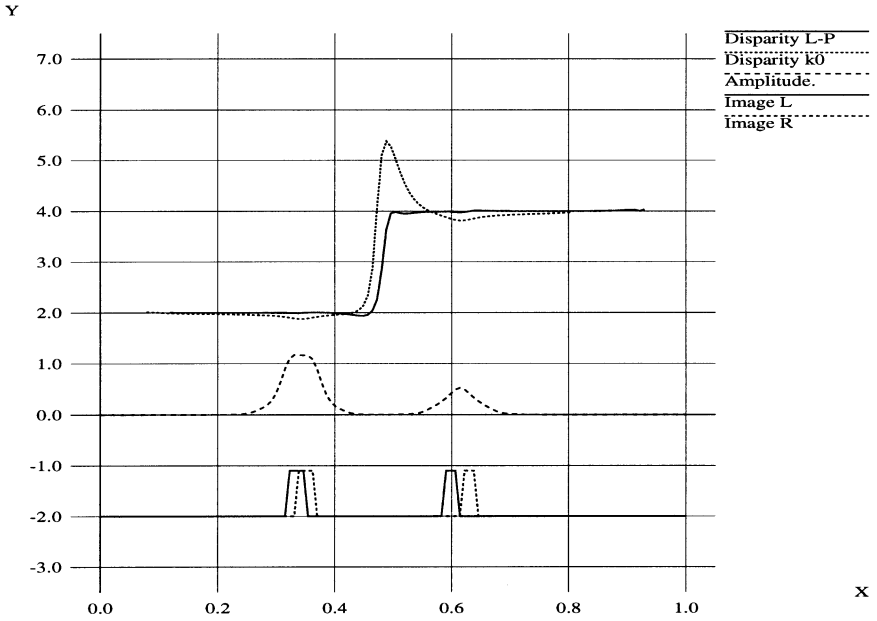


Fig. 9. The computed disparity (in pixels) using the local and the tuning frequency, and the product of the amplitudes, $(\rho_R^2(x) \times \rho_L^2(x))$ for the quartic model. In the *lower part*, the *left* and *right* 1D images are shown: the first spike is shifted by two pixels, the second by four pixels, as correctly estimated

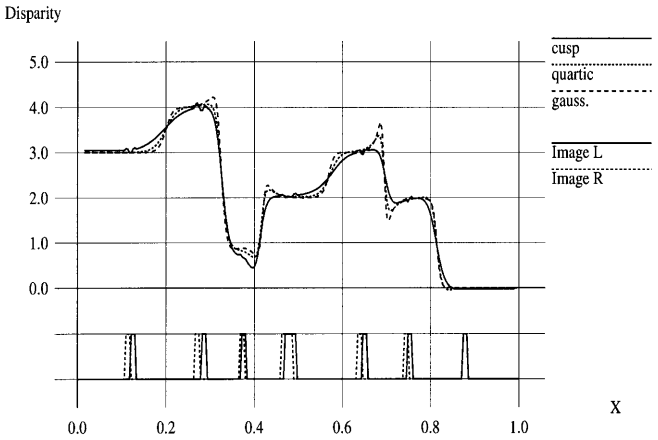


Fig. 10. Disparity estimates for spike-like features of the three filters, N = 256, λ_0 = 20 and β = 1

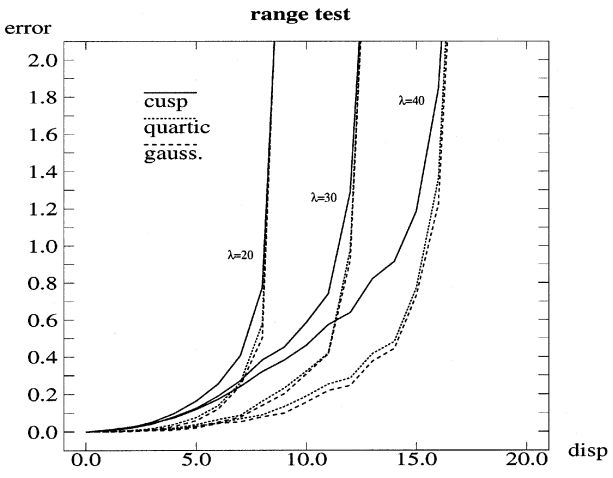


Fig. 12. Constant disparity field. Errors of disparity estimates for the three filters with $\lambda_0 = 20, 30, 40$ pixels and $\sigma = \lambda_0/2$ for 1D images of size N = 512 pixels

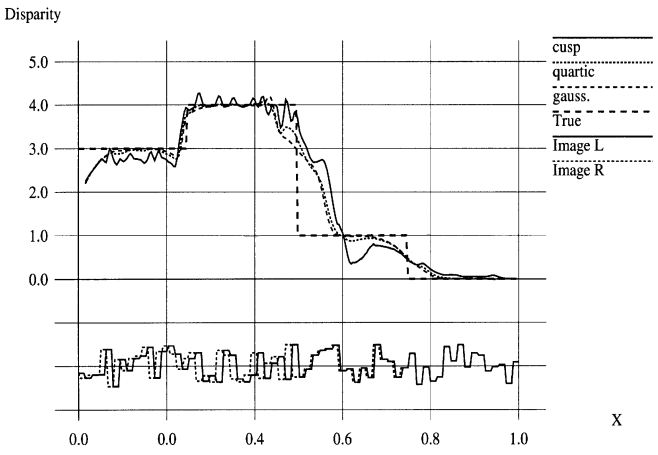


Fig. 11. Disparity estimates of the three filters for a random-dot image, N = 256, λ_0 = 20 and β = 1

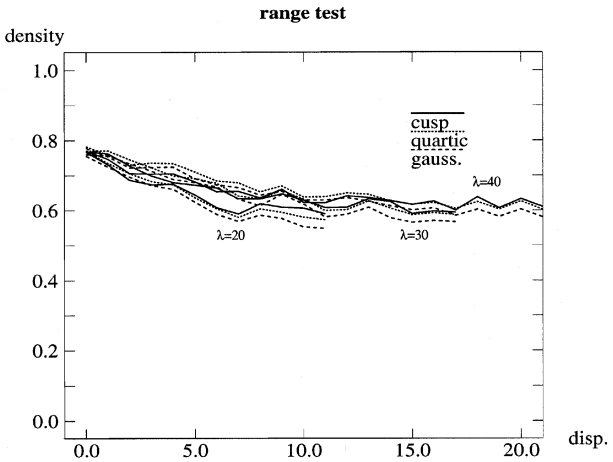


Fig. 13. Constant disparity field. Densities of disparity estimates for the three filters with $\lambda_0 = 20, 30, 40$ pixels and $\sigma = \lambda_0/2$ for 1D images of size N = 512 pixels

4.2 Stereo-signal complexity

At first, to test the estimation capability, constant disparity fields are used, i.e., the left image is obtained by a constant shift of the right one. This means that there are no occlusions. Figures 12 and 13 show the filters' performances for three different wavelengths. For fixed wavelength, the filters' behaviors follow the results found in [4], the error is almost flat for disparity in the range $[-\frac{\lambda_0}{3}, \frac{\lambda_0}{3}]$, then it grows rapidly. The sharp error increase is due to wraparound effects (aliasing). These are reduced using larger wavelengths. A simple increase of the filter's width, for fixed wavelength, does not change the error curves but improves only the densities. Using filter parameters that produce approximatively the same densities, the cusp error is slightly larger than the quartic and the Gaussian error, in particular for large wavelengths.

Analogous results are found for piecewise constant disparity fields. In this case, there are occlusions, therefore, the filters interpolation capabilities are also tested. Figure 14 shows the filters' performances for random-dot stereo pairs in which disparity assumes random values in the range $[0,\ D = 5]$ in 20 randomly generated intervals. An example of disparity estimates for random-dot images for four intervals is shown in Fig. 11. The error curves display minima for $\lambda_0 \approx 2.5\,D$, where D is the maximum disparity. This behavior is close to the one found in real stereo images, see [23]. Going from left to right, the error decreases because the wavelength increases (as explained above). However, when the filter becomes wider than disparity variations, the error starts to increase because the filter averages over close-by features of different disparities. The cusp performance is slightly inferior to the Gaussian performance, both in error and density.

An important issue is the influence of the stereo-signal complexity on the quality of disparity estimation. Here, we parametrize complexity as the number of intervals on which disparity is constant. Zero complexity means that the disparity is constant over the whole image, when the parameter increases there are several regions of different constant disparity and occlusions start to appear. Figure 15 shows the filters' performances as a function of complexity, i.e., as a function of the number of intervals.

4.3 DC sensitivity

Large differences among filters' performances arise when sensitivity to the zero-frequency signal content (DC sensitivity) is taken into consideration. DC sensitivity is caused by the nonzero value of integral $\int dy\, G(x-y)\, f(y)$. The cusp is much more sensitive to the zero-frequency components of the signals. Figures 16 and 17 illustrate the filters' performance as a function of the DC level added to the input signal. Note that 25% of DC is enough to dramatically decrease the cusp density. This effect can be reduced by increasing the number of oscillations within the envelope support, i.e., by increasing the filter width for a fixed wavelength, as shown in Figs. 18 and 19. However, filters can be simply modified in order to reduce DC effects by redefining convolution as

$$F^{DC}(x,k_0) = \int dy\; G(x-y)\left[e^{i\,k_0(x-y)} - \widehat{G}(k_0)\right] f(y)\;, \quad (23)$$

where $\widehat{G}(k)$ if the Fourier transform of envelope $G(x)$; for the cusp, we have

$$\widehat{G}(k_0) = \frac{\gamma^2}{\gamma^2 + k_0^2}\;,$$

and for the quartic,

$$\widehat{G}(k_0) = \frac{1}{1 + k_0^2/\mu^2 + 0.5\,k_0^4/\mu^4}\;.$$

Using this definition, the filters' performances become very similar even in presence of high zero-frequency content. Figures 20 and 21 show that performance for DC values of 50% and 100% is almost identical to performance with no DC added. Operation (23) is the difference between a convolution computed with tuning frequency k_0 and a convolution with $k_0 = 0$ weight by factor $\widehat{G}(k_0)$.

$$F^{DC}(x,k_0) = F(x,k_0) \;-\; \widehat{G}(k_0)\,F(x,k_0=0)\;. \quad (24)$$

5 Discrete models

Differential equations can be directly mapped on the discrete pixel space by means of IIR filter design techniques [16]. Given the shape of the convolution kernel, it is necessary to determine the values of coefficients a and b of the filter

$$\sum_{k=-O_n}^{O_n} a_k y(n-k) = \sum_{k=-I_n}^{I_n} b_{kl} f(n-k) \quad (n \in N) \quad (25)$$

that realizes the functionality required. Here, $f(n)$ and $y(n)$ are the input and the output of the processing node at location (n), and I and O determine the number of neighbors necessary to perform the computation.

5.1 The cusp filter

To determine the discrete model for the convolution with an oscillating decaying exponential, let us consider the discrete version of Eq. 15 i.e.,

$$\begin{aligned} &A_1\, e^{-i\,k_0}\, u(n+1) + A_0\, u(n) \\ &\quad + A_{-1}\, e^{+i\,k_0}\, u(n-1) = B\, f(n)\;, \end{aligned} \quad (26)$$

where A_1, A_0 and A_{-1} are real numbers and the exponential terms are explicited to simplify equations. Plugging in the discrete expression of the oscillating cusp

$$u(n) = \sum_{k=-\infty}^{+\infty} e^{-\gamma\,|n-k| + i\,k_0\,(n-k)}\, f_k\;, \quad (27)$$

three linear equations are derived: the first for $k \geq n+1$, the second for $k = n$, and the third for $k \leq n-1$:

$$\begin{aligned} &A_1 + A_0\, e^{-\gamma} + A_{-1}\, e^{-2\,\gamma} = 0\;, \\ &A_1\, e^{-\gamma} \;+\; A_0 \;+ A_{-1}\, e^{-\gamma} = B\;, \\ &A_1\, e^{-2\,\gamma} + A_0\, e^{-\gamma} + A_{-1} = 0\;. \end{aligned} \quad (28)$$

The system solution gives

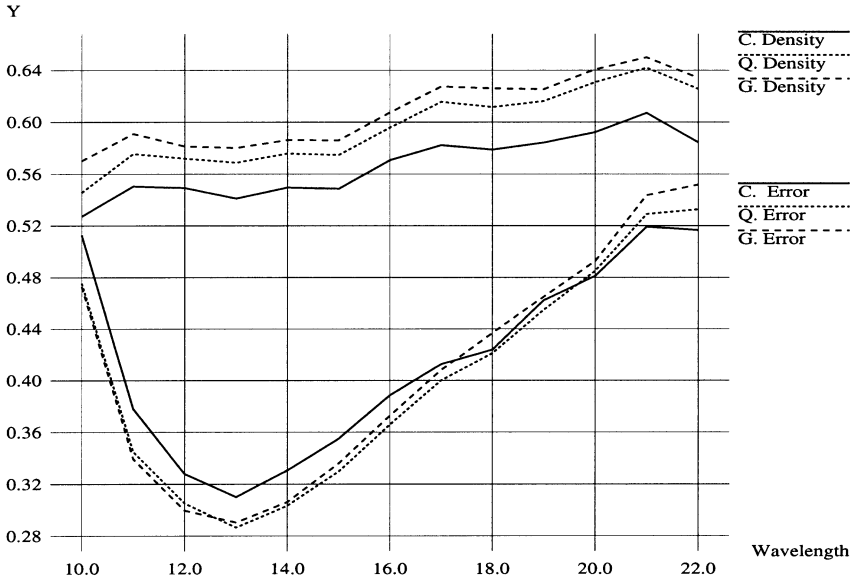


Fig. 14. Piecewise constant disparity field. Error and density for the three filters as a function of the wavelength, with $\sigma = \lambda_0/2$ and $N = 256$. In each stereo pair, disparity assumes random values in the range of [0, 5] pixels, in 20 randomly generated intervals. Each point represents the average error or the density on a large number (256) of random-dot 1D image pairs

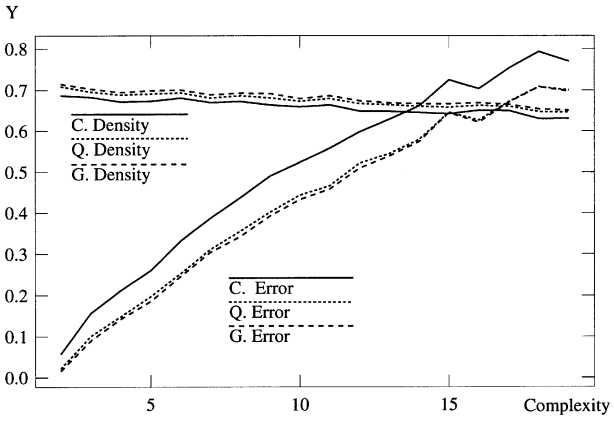


Fig. 15. Error and density as a function of the stereo-signal complexity (λ_0 = 20). Complexity is the numbers of intervals on which disparity is constant. The maximum disparity is D = 5 pixels

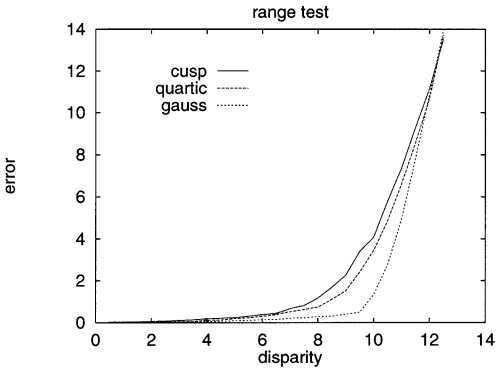


Fig. 16. Error of disparity estimates for an added DC level of 25%, N = 256, λ_0 = 32 and $\sigma = \lambda_0/2$

$$\begin{aligned} A_{-1} = A_1 &= -e^{-\gamma}, \\ A_0 &= 1 + e^{-2\gamma}, \\ B &= C(1 - e^{-2\gamma}), \end{aligned}$$

where C is an arbitrary constant.

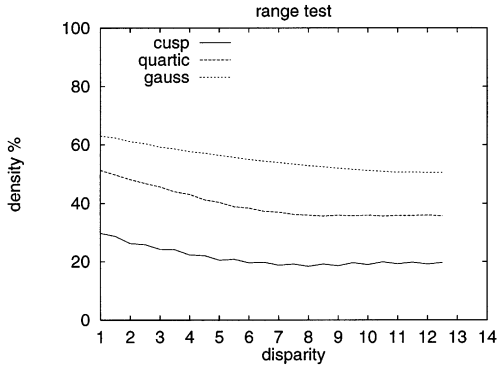


Fig. 17. Density of disparity estimates for an added DC level of 25%, N = 256, λ_0 = 32 and $\sigma = \lambda_0/2$

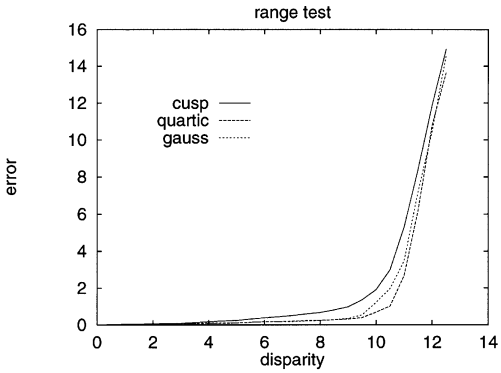


Fig. 18. Error of disparity estimates for an added DC level of 25%, N = 256, λ_0 = 32 and $\sigma = \lambda_0$

Separating real and imaginary parts, $u_n = \alpha(n) + i\,\beta(n)$, we find the discrete version of system (16)

$$\begin{aligned} &a\,\alpha(n+1) + c\,\alpha(n) + a\,\alpha(n-1) \\ &\quad -b[\beta(n+1) - \beta(n-1)] = B\,f_n\ , \\ &a\,\beta(n+1) + c\,\beta(n) + a\,\beta(n-1) \\ &\quad +b[\alpha(n+1) - \alpha(n-1)] = 0\ , \end{aligned} \tag{29}$$

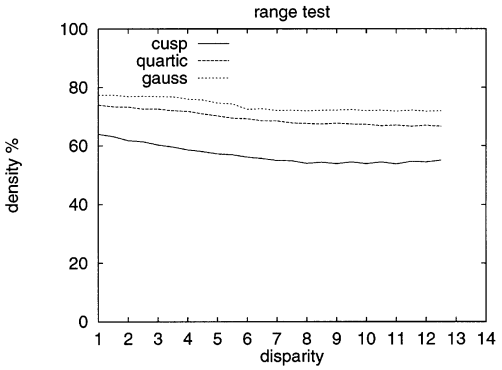


Fig. 19. Density of disparity estimates for an added DC level of 25%, $N = 256$, $\lambda_0 = 32$ and $\sigma = \lambda_0$

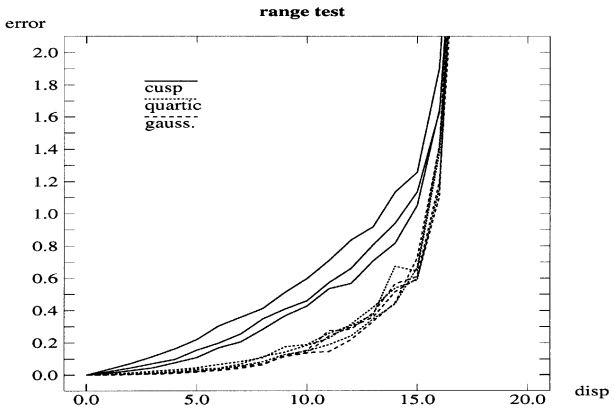


Fig. 20. Error for the DC cleaned filters for three levels of added DC, $d = 0\%$, $d = 50\%$ and $d = 100\%$; $N = 512$, $\lambda_0 = 40$ and $\sigma = \lambda_0/2$

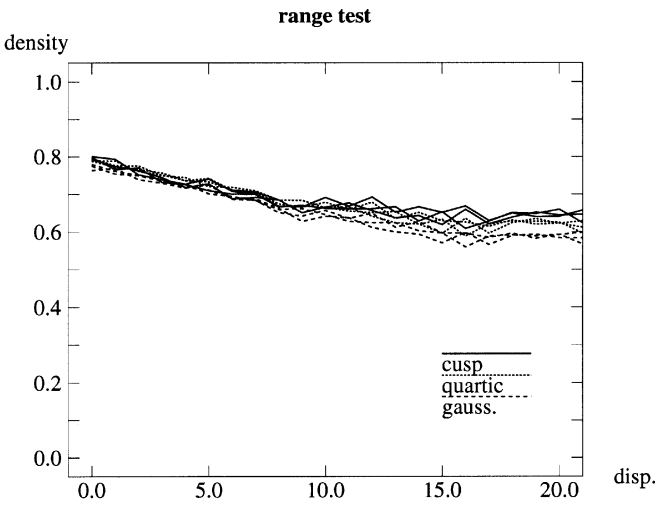


Fig. 21. Density for the DC cleaned filters for three levels of added DC, $d = 0\%$, $d = 50\%$ and $d = 100\%$, $N = 512$, $\lambda_0 = 40$ and $\sigma = \lambda_0/2$

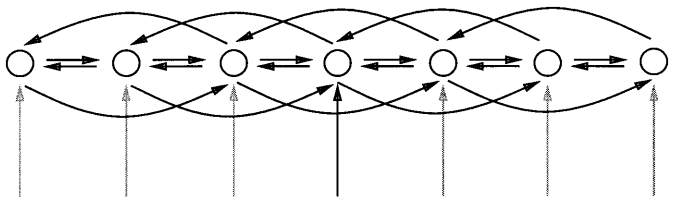


Fig. 22. Schematic representation of a "cusp" lattice network. Each node is connected with its nearest neighbors and with its next nearest neighbors (four connections per node)

where $a = -e^{-\gamma}\, cos(k_0)$, $b = -e^{-\gamma}\, sin(k_0)$, and $c = (1+a^2+b^2)$. Following steps similar to the ones used to derive Eq. 17, the above system can be rewritten as a single equation that involves the first four neighbors of each node, see Fig. 22. For $\alpha(n)$,

$$\begin{aligned} &a_2\,\alpha(n+2) + a_1\,\alpha(n+1) + a_0\,\alpha(n) \\ &+a_{-1}\,\alpha(n-1) + a_{-2}\,\alpha(n-2) \\ &= b_1\,f(n+1) + b_0\,f(n) + b_{-1}\,f(n-1)\ , \end{aligned} \tag{30}$$

with

$$\begin{array}{rcccl} a_{-2} = a_2 & = & a^2 + b^2 & = & \mathrm{e}^{-2\gamma}, \\ a_{-1} = a_1 & = & 2\,a\,c^2 & = & -2\mathrm{e}^{-\gamma}(1+\mathrm{e}^{-2\gamma})\cos k_0, \\ a_0 & = & 2\,a^2 + c^2 - 2\,b^2 & = & 1 + 4\mathrm{e}^{-2\gamma}\cos^2 k_0 + \mathrm{e}^{-4\gamma} \end{array}$$

and

$$\begin{array}{rcl} b_1 & = & -C\,(1-e^{-2\,\gamma})\,e^{-\gamma}\,cos(k_0-\phi), \\ b_0 & = & -C\,(1-e^{-4\,\gamma})\,cos(\phi), \\ b_{-1} & = & -C\,(1-e^{-2\,\gamma})\,e^{-\gamma}\,cos(k_0+\phi), \end{array}$$

where $\phi = 0$ for $\alpha(n)$, and C is an arbitrary constant. The same equation holds for $\beta(n)$, with $\phi = \pi/2$. The decay parameter γ and the tuning phase ϕ can be obtained from a proper choice of coefficients a, and b. In the appropriate limit, the discrete formalism leads to the result derived in the differential equation approach [5].

5.2 The quartic filter

Since the convolution equation with the quartic filter, Eq. 20, is a fourth-order differential equation, the connectivity of the discrete model has to be extended to the second nearest neighbors

$$\begin{aligned} &A_2\ e^{-2\,i\,k_0}u(n+2) + A_1\,e^{-i\,k_0}\,u(n+1) \\ &+A_0\,u(n) + A_{-1}\,e^{+i\,k_0}\,u(n-1) \\ &+A_{-2}\,e^{2\,i\,k_0}\,u(n-2) = \ B\,f(n)\ . \end{aligned} \tag{31}$$

The corresponding discrete equations are obtained by plugging in the discrete expression of the quartic filter,

$$\begin{aligned} u(n) = &\sum_{k=-\infty}^{+\infty} \sin\left(\frac{\pi}{8} + \gamma_2\,|n-k|/\mu\right) \\ &e^{-\gamma_1\,|n-k|/\mu + i\,k_0\,(n-k)}\,f(k)\ , \end{aligned} \tag{32}$$

with $\gamma_1 = 2^{1/4}\cos\left(\frac{\pi}{8}\right)$, $\gamma_2 = 2^{1/4}\sin\left(\frac{\pi}{8}\right)$. Five linear equations are derived (for $k \geq n+2$, for $k = n+1$, for $k = n$, for $k = n-1$ and for $k \leq n-2$). Exploiting symmetry $A_n = A_{-n}$, the number of equations is reduced from five to three. Coefficients $A_2 = A_{-2}$, $A_1 = A_{-1}$, and A_0 are the solution of the linear system:

$$\begin{pmatrix} s_4 + s_0 & s_3 + s_1 & s_0 \\ s_3 + s_1 & s_2 + s_0 & s_1 \\ 2\,s_2 & 2\,s_1 & s_0 \end{pmatrix} \begin{pmatrix} A_2 \\ A_1 \\ A_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ B \end{pmatrix}\ , \tag{33}$$

where

$$s_n = \sin\left(\frac{\pi}{8} + n\,\gamma_2/\mu\right) e^{-n\,\gamma_1/\mu}\ . \tag{34}$$

As done in the previous section, the complex discrete equation can be written as two coupled real equations (see Eq. 30). However, in this case, each node is connected to the first and to the second nearest neighbors.

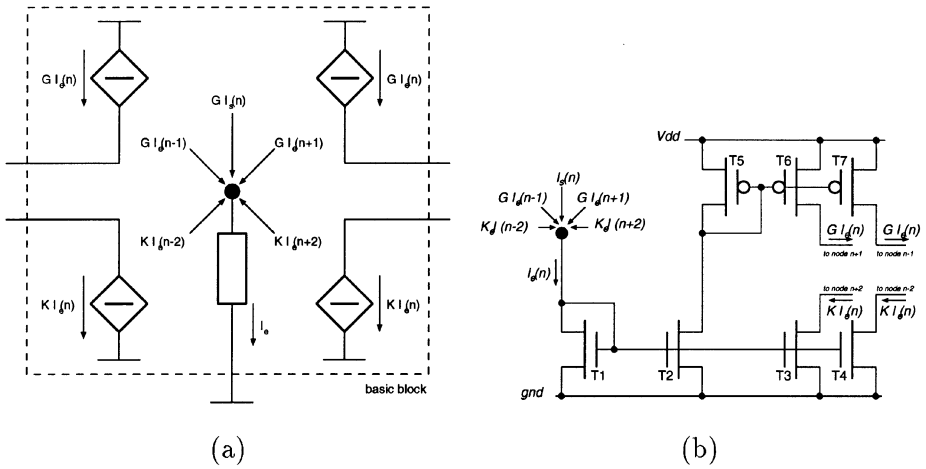


Fig. 23a. The node basic block and its internal circuital structure: the *black dot* represents the node which is stimulated by the photocurrent $I_s(n)$ and receives inhibitory (K) and excitatory (G) contributions from neighboring nodes. The sum of the current contributions from other nodes is performed by Kirchhoff current law at the node. The resulting excitation current, $I_e(n)$, acts, in turn, on neighboring nodes controlling the currents of four current generators. **b** Description of the node basic block at transistor level. $I_e(n)$ is copied (weighted according to G and K values) to provide excitation or inhibition to the neighboring nodes

In the present discussion, a lattice of infinite extent is considered. To simulate it, we assume a natural boundary condition: deleting the connections that would connect to nodes that fall outside the boundaries of the array. As stated in [21], this condition is equivalent to the Newman boundary conditions of zero normal derivative in partial differential equations. The stability of the computational scheme developed is discussed in Appendix B.

6 VLSI analog implementation

Analog recurrent filter operations with spatial kernels for smoothing and contrast enhancement can be implemented by using linear resistive networks with positive and negative resistors [1, 11, 14, 17], or by current generators driven by the voltage at other nodes (VCCS) [3]. Since our main interest is to investigate how complex functionalities can be achieved through cooperative computations, we choose to relate the lattice networks' equations to a circuit architecture of one-way-interacting elements implemented as current-controlled current sources (CCCS).

Let us thus consider a regular grid of nodes, in which each node is fed by a current generator whose value is proportional to the incident light at that point (feedforward contribution). By using a current mode technique, in which all signals are encoded by currents, the excitation e of the network is mapped on a current, the interaction can be implemented by CCCS that feed or sink currents according to the current values of neighboring nodes (see Fig. 23a).

In this way, the analog signal processing of our circuit architecture will be based on ratios of matched component values [24], thus eliminating the dependence on the performances of single devices. The sum of the current contributions from other nodes is performed by Kirchhoff current law at the node.

One advantage of this interpretation shows up at transistor level, where it is shown how to implement these interactions with "high" precision through current mirrors which provide, when the transistors are operating in the saturation region, a weighted copy of their input currents according to the W/L ratios of the transistors. To improve the matching between devices, the gain ratios of the mirrors are restricted to rational numbers, so that they can be implemented by using two sets of identical transistors connected in parallel (see Fig. 23b) [18]. A prototype 17-node VLSI circuit was fabricated by IRST on its CMOS 2.0 μm, N-well, single-poly, and double-metal technology (see Fig. 24). Figure 25 compares measured and expected impulse response for three different values of the phase: 0, $\pi/2$ and $\pi/4$. These results proves the feasibility of the approach to the generation of more complex functionalities.

7 Discussion and conclusions

In this paper, we have discussed the analog implementation of a phase-based technique for disparity estimaton through continuous and discrete computational models. It has been shown that the algorithm can be reduced to local operations. In fact, convolutions can be performed by means of the solution of a linear system of differential equations whose order depends on the smoothness of the local envelope used in the Gabor-like filters. The first and simplest choice, i.e., the cusp kernel that presents a discontinuity in the first derivative, leads to a fourth-degree differential equation. A successive approximation of the Gaussian kernel leads to the quartic filter that improves the smoothness of the disparity field estimates but requires higher degree differential equations.

The modifications of the system's performance due to the envelope's change were investigated in detail. The results of the simulation experiments indicate that the performance of the quartic filter is very close to the Gaussian's. The cusp filter is comparable if it is modified in order to eliminate the sensitivity to the zero-frequency content of the input signals.

The continuous models were translated into a self-consistent discrete lattice formalism, in which the degree of the differential equation system is related to the number

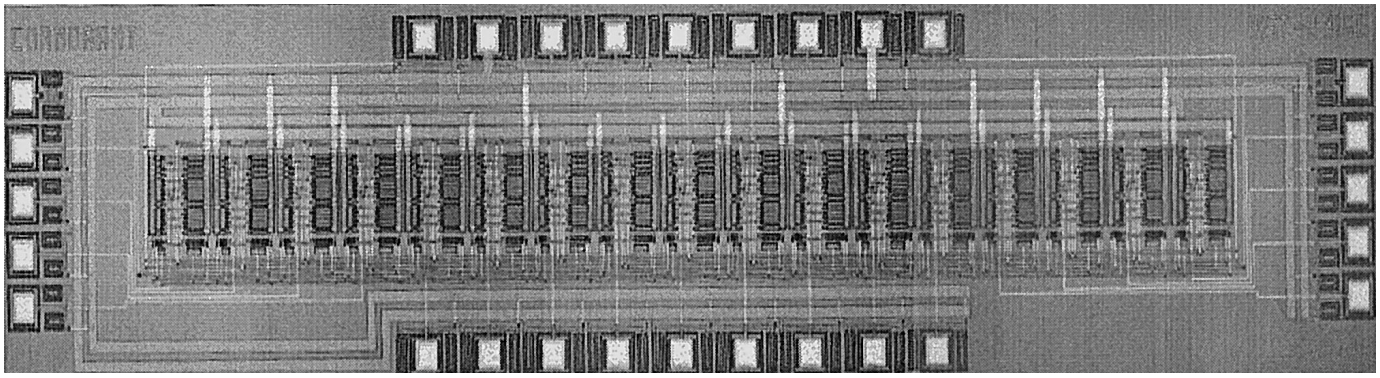


Fig. 24. Microphotograph of a 1D 2nd-order network chip

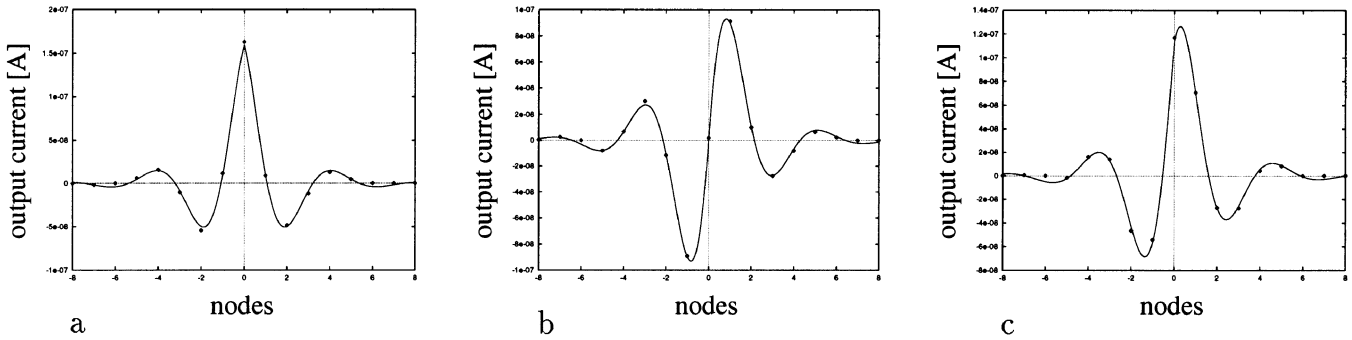


Fig. 25a–c. Comparisons between measured (*dots*) and theoretical (*continuous lines*) Gabor-like impulse responses of phases **a** 0, **b** $\pi/2$ and **c** $\pi/4$

of interconnections per node required to perform the computation. The implementation of a Gabor-like filter in which the Gaussian envelope is replaced by a decaying exponential requires that each lattice node is connected with the nearest four neighbors. The quartic filter requires to increase connectivity to eight neighbors per node. The limited number of connections is particularly helpful in the development of analog VLSI circuits. The possibility of implementing a computational procedure on an analog VLSI circuit is a very important issue. VLSI technology can lead to the development of microsystems capable of performing complex computations in microseconds.

The model for the cusp filter was used to design and fabricate a prototype analog chip that performs convolutions. The chip, fabricated by IRST is constituted by a linear array of 17 basic cells, each cell is connected with its 4 nearest neighbors. The development of a larger version of this circuit could lead to important applications because it is capable of performing convolutions with Gabor-like kernel in microseconds. Furthermore, it offers the possibility of comprising sensing and processing on the same silicon chip. Therefore, it could constitute an ideal low-cost, low-power microsystem for early-stage information processing in hybrid systems for depth and/or movement estimation [18].

Appendix A
Shape of the quartic kernel

The analytical expression of the quartic envelope can be easily derived by performing the inverse Fourier transform

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{e^{i\,k\,x}}{1+\mu^2\,k^2+\mu^4\,k^4/2!}\,dk \tag{35}$$

by means of the residue theorem ($i\,k = z/\mu$)

$$\frac{1}{2\pi}\int_{-i\,\infty}^{+i\,\infty}\frac{e^{\frac{x\,z}{\mu}}}{1-z^2+z^4/2!}\,dz = i\sum_n\frac{(z-z_n)\,e^{\frac{x\,z}{\mu}}}{1-z_n^2+z_n^4/2!}\,, \tag{36}$$

where z_n are the zeros of $\left[1+\mu^2\,(k-k_0)^2+\frac{\mu^4\,(k-k_0)^4}{2!}\right]$ that lie inside the integration contour. The contour is closed according to the sign of x. Since zeros are at $z_{1,2} = \pm\,(a + i\,b)$, and $z_{3,4} = \pm\,(\gamma_1 - i\,\gamma_2)$, where

$$\gamma_1 = 2^{1/4}\cos\left(\frac{\pi}{8}\right) = 2^{1/4}\sqrt{1+\frac{1}{\sqrt{2}}}\,,$$

and

$$\gamma_2 = 2^{1/4}\sin\left(\frac{\pi}{8}\right) = 2^{1/4}\sqrt{1-\frac{1}{\sqrt{2}}}\,,$$

we find

$$Q(x) = \frac{N}{\mu}\,\sin\left(\frac{\pi}{8}+\gamma_2\,|x|/\mu\right)\,e^{-\gamma_1\,|x|/\mu}\,, \tag{37}$$

where N is a normalization factor.

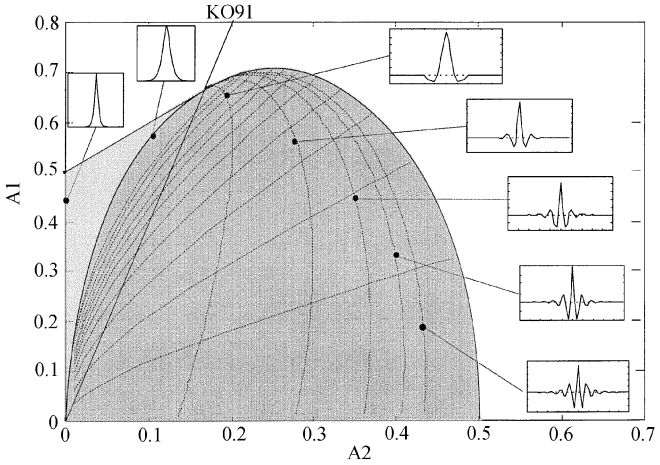


Fig. 26. Stability regions in the A_1, A_2 space. The resistive network is stable only if A_1 and A_2 values are in a grayed region. In the light-gray region, the operation performed is a double (i.e., cascade of) convolution with decaying exponentials. In the dark gray area, the operation is a convolution with a decaying oscillation given by (3). In this region the frequency decreases when A_1 increases and the dispersion of the envelope increases with A_2. *Pseudo-horizontal dotted lines* are the locus for function with the same spatial frequency. *Pseudo-vertical lines* are the locus for function with the same product frequency times dispersion. See text

Appendix B
Stability analysis

The output of a linear, space invariant, network described by (25) can be written as the convolution between the input excitation and a kernel given by the response to a spatial impulse. Hence, the convolution kernel can be found by applying a two-sided Z-transformation to both terms of (25) and anti-transforming the I/O transfer function. For a 1D network, the impulse response (i.e., the convolution kernel) is given by anti-transforming function

$$H(z) = \frac{Y(z)}{F(z)} = \frac{\sum_{k=-I}^{I} b_k z^{-k}}{\sum_{k=-O}^{O} a_k z^{-k}} \; . \tag{38}$$

It is noteworthy that, because of the Z-transform properties and the linearity of the network: (i) the overall input/output relation (i.e., Z-transform) for a system resulting from a linear combination of two independent subsystems (networks) with the same input can be expressed as the corresponding linear combination of the input/output relation (i.e., Z-transform) for each subsystem (network); (ii) the relation between the excitation at the nodes of a network driven by the output of a second network and the input stimulus of this second one (cascade of resistive networks) can be expressed as the product of the input/output relations (i.e., the Z-transfer functions) of the two networks.

The stability of a 2nd-order network can be analyzed according to [13] in the space of the parameters $A_1 = -a_1/a_0$ and $A_2 = a_2/a_0$. The network is stable if the corresponding A_1 and A_2 parameters are in the grayed regions in Fig. 26.

From the stability point of view A_1 and A_2 are sufficient for the analysis. Proportional variation of a_0, a_1, a_2 affects only the amplitude of the convolution kernel.

For each point inside the elliptical area ($A_1^2 < 4A_2 - 8A_2^2$), the convolution kernel is of 2nd-order type, as defined in (30). For each point outside the elliptical area to which correspond a stable network ($A_1^2 > 4A_2 - 8A_2^2$, $A_2 < 1/6, A_1 < A_2 + 0.5$, i.e., light-gray area in Fig. 26), the convolution kernel is the result of two successive convolutions with two decaying exponentials ($e^{-\gamma_1|n|}$ and $e^{-\gamma_2|n|}$). In particular on the frontier with the elliptical area, $\gamma_1 = \gamma_2$, we have the best approximation of a Gaussian (it is well known that the recurrent application of a convolution with a decaying exponential converges to a Gaussian kernel). The line marked with KO91 is the locus of the approximations of Gaussians found in [13]. It is worthy to note that, when A_2 is null, the network degenerates to a order-one network and the convolution kernel is a decaying exponential ($\gamma_1 = 0$ or $\gamma_2 = 0$).

References

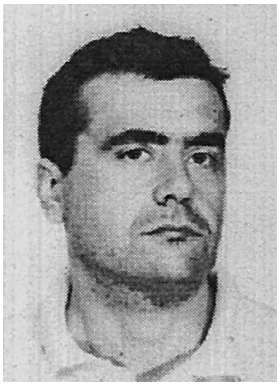
1. Arreguit X, Vittoz EA (1994) Perception systems implemented in analog VLSI for real-time applications. In: Proc. PerAc'94 Conference: From Perception to Action, Lausanne, Switzerland, 5–9 September 1994
2. Barron JL, Fleet D, Beauchemin S (1994) Performance of optical flow techniques. Int J Comput Vision 12(1): 43–77
3. Bisio GM, Bruccoleri M, Cusinato P, Raffo L, Sabatini SP (1994) An analog VLSI massively parallel module for low-level cortical processing in machine vision. In: Proc. of the IV Int. Conf. on Microelectronics for neural networks and fuzzy systems, IEEE MICRONEURO'94, 1994, Torino, Italy. IEEE Computer Society Press, pp 207–211
4. Cozzi A, Crespi B, Valentinotti F, Woergoetter F (1997) Performance of phase-based algorithms. Mach Vision Appl (Special issue on Performance Characteristics of Vision Algorithms), 5–6 (9):334–340
5. Cozzi A, Crespi B, Valentinotti F, Woergoetter F (1996) Continuous and discrete models for phase-based stereo vision. Technical Report, IRST 9610, Trento, Italy
6. Delbriick T (1993) Silicon retina with correlation-based velocity-tuned pixels. IEEE Trans Neural Networks 4: 529–541
7. Fleet DT, Jepson AI (1993) Stability of Phase Information. IEEE Trans Pattern Anal Mach Intell 15(12): 1253–1268
8. Fleet DJ, Jepson AD, Jenkin M (1991) Phase-based Disparity Measurement. Comput Vision Graphic Image Process: Image Understanding 53(2): 198–210
9. Fleet DJ, Jepson AD (1990) Computation of Component Image Velocity from Local Phase Information. Int J Comput Vision 5(1):77–104
10. Kobayashi H, Matsumoto T, Yagi T, Shimmi T (1993) Image processing regularization filters on layered architecture. Neural Networks 6: 327–350
11. Kobayashi H, White JL, Abidi AA (1991) An active resistor network for Gaussian filtering of images. IEEE J Solid State Circuits 26: 738–748
12. Koch C, Marroquin J, Yuille A (1986) Analog "neuronal" networks in early vision. Proc Natl Acad Sci 83: 4263–4267
13. Matsumoto T, Kobayashi H, Togawa Y (1992) Spatial versus temporal stability issues in image processing neurochips. IEEE Trans Neural Networks 3(4): 540–569
14. Mead CA (1989) Analog VLSI and Neural Systems. Addison-Wesley, Reading, Mass.
15. Poggio T (1985) Early vision: from computational structure to algorithms and parallel hardware. Comput Vision Graphics Image Process 31: 139–155
16. Rabiner LR, Gold B (1975) Theory and application of digital signal processing. Prentice Hall, Inc., Englewood Cliffs, New Jersey
17. Raffo L (1995) Resistive network implementing maps of G functions of any phase. Electron Lett 31(22): 1913–1914
18. Bisio GM, Raffo L, Sabatini SP: Analog VLSI primitives for perceptual tasks in machine vision. Neural Comput Appl (Special issue on Machine Vision) (in press)
19. Roska T, Vandewalle J (1993) Cellular Neural Networks. Wiley, Chichester

20. Sanger TD (1988) Stereo disparity computation using gabor filters. Biol Cybern 59: 405–418
21. Shi BE, Chua LO (1992) Resistive grid image filtering: Input/output analysis via the cnn framework. IEEE Trans Circuit Syst: I Fundamental Theory and Applications 39(7): 531–548
22. Shi BE, Roska T, Chua LO (1993) Design of linear cellular neural networks for motion sensitive filtering. IEEE Trans Circuit Syst: I Fundamental Theory and Applications 40: 320–331
23. Valentinotti F, Di Caro G, Crespi B (1996) Real-time parallel computation of disparity and optical flow using phase difference. Mach Vision Appl 9(3): 87–96
24. Vittoz EA (1994) Analog VLSI signal processing: Why, where, and how? J VLSI Signal Processing 8: 27–44
25. Wyatt JL, Keast C, Seidel M, Standley D, Horn B, Knight T, Sodini C, Lee H-S, Poggio T (1992) Analog VLSI systems for image acquisition and fast early vision processing. Int J Comput Vision 8: 217–230

Bruno Crespi graduated in physics at University of Trento, Italy. He received his Ph.D. in Physics from the University of Illinois at Urbana-Champaign. Since 1990 he has been working at the Istituto per la Ricerca Scientifica e Tecnologica in Trento, Italy. His research interests include artificial vision, pattern recognition, associative memories, and neural networks.

Alex Cozzi received the M.S. degree in computer science from the Università degli Studi, Milano, Italy, in 1994. He performed the thesis work at IRST (Istituto per la Ricerca Scientifica e Tecnologica), Italy. He is currently a research assistant and Ph.D. student in the group of “Computational Neuroscience” at the Dept. of Physiology, Ruhr-University Bochum, Germany. His research interests include artificial vision, robotics and object-oriented languages.

Silvio Sabatini was born in 1968. He received the M.S. degree and the Ph.D. in electronic engineering and computer science at the Department of Biophysical and Electronic Engineering (DIBE) of the University of Genoa, Italy, in 1992 and 1996, respectively. In 1996, he joined, as post-doctoral fellow, the “Physical Structure of Perception and Computation” (PSPC) Research Group at DIBE (http://www.dibe.unige.it/department/micro/pspc/home.html) working on: neuromorphic perceptual microsystems, biocybernetics of vision, cortical computational models, multidimensional signal representation.



Luigi Raffo received his degree in electronic engineering from the University of Genova, Italy and his PhD degree in electronic engineering and computer science in 1994 from the same university. In 1994, he joined the Microelectronic Group of the Department of Electrical and Electronic Engineering, University of Cagliari, as an assistant professor. He is currently associate professor of microelectronics. His research interest are in the field of VLSI architectures, systems and circuits. He is author of more than 40 international papers in his research fields.