Analog processing for stereo vision

Bruno Crespi[a], Alex G. Cozzi[b], Luigi Raffo[c] and Silvio Sabatini[d]

[a] *IRST*, I-38050 Povo, Trento, Italy; e_mail:crespi@irst.itc.it

[b] Ruhr Univ. Bochum, Germany; cozzi@neurop2.ruhr-uni-bochum.de

[c] *DIEE*, Univ. of Cagliari- Piazza d'Armi- I-09123 Cagliari, Italy; e_mail: luigi@diee.unica.it

[d] *DIBE*, Univ. of Genoa- Via Opera Pia 11a- I-16145 Genova, Italy; e_mail: silvio@dibe.unige.it

Abstract

The implementation of early vision algorithms by means of analog circuits could lead to the fabrication of very fast and efficient sensors. However analog implementation has to overcome several difficulties. One of the major problems regards connectivity: because a physical link is necessary to exchange information between nodes of a computational network, it is important to devise algorithms that minimize the number of connections per node necessary to implement a given functionality [1, 7].

This work discusses the analog implementation of a technique for depth and motion estimation [5] that is based on the convolution of the input images with Gabor filter. 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.

Stereo vision, Phase-based technique, Analog Computation, VLSI.

1 Introduction: the phase-difference-based technique

The phase-difference-based technique for disparity estimation [10] is based on the computation of the difference between the phases of the convolutions of the two stereo signals, $f_{L/R}(x)$ with Gabor filters

$$u_{L/R}(x,k_0) = \int dy\, G(x-y)\, e^{i\,k_0\,(x-y)}\, f_{L/R}(y) = \rho_{L/R}(x)\, e^{i\,\psi_{L/R}(x)} \quad (1)$$

The local envelope, $G(x-y)$, is chosen to be a Gaussian

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

because this envelope is best localized in both space and frequency.

The phase of the filter response, $\psi(x)$, is a quasi *linear* function of spatial position. This means that phase derivative $\psi_x(x)$ is generally close to the tuning frequency, $\psi_x(x) \approx k_0$. The linear behavior allows an accurate estimation of the shift from the phase difference by means of a second order expansion in $d(x)$,

$$d(x) \approx 2\,\frac{[\Delta\psi(x)]_{2\,\pi}}{\psi_x^L(x) + \psi_x^R(x)}. \quad (3)$$

Approximation (3) is reliable only where phases are linear. Since in the vicinity of points where the amplitude vanishes, $\rho(x) = 0$, the phase develops strong nonlinearities. Computation at point x is accepted only if

$$\left|(\psi^{'}(x) - k_0)\right| < k_0\, T_S \quad (4)$$

where $T_S \approx 0.4$ [5]. 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.

2 Local Gabor-like filters

The convolution

$$u(x) = \int dy\ G(x-y)\, e^{i\,k_0\,(x-y)}\, f(y) \quad (5)$$

can be expressed as the inverse Fourier transform of

$$\widehat{u}(k) = \widehat{G}(k-k_0)\,\widehat{f}(k) \quad (6)$$

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

$$\widehat{G}^{-1}(k-k_0)\,\widehat{u}(k) = \widehat{f}(k) \qquad (7)$$

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

$$\left[1 + \mu^2\,(k-k_0)^2 + \frac{\mu^4\,(k-k_0)^4}{2!} + \ldots\right]\widehat{u}(k) = \widehat{f}(k). \qquad (8)$$

Since multiplication by $(k-k_0)$ in frequency space corresponds to the application of differential operator $[\frac{d}{dx} - i\,k_0]$ in the x-space, the inverse Fourier transform of eq. (7) generates a series of local envelops. The envelope's smoothness increases with the order of the corresponding differential equation.

2.1 The Cusp filter

The first order expansion in $\mu = 1/\gamma$ of eq.(8) generates the Cusp envelope

$$C(x) = \frac{\gamma}{2}\,e^{-\gamma|x|}\,. \qquad (9)$$

This implies that integral equation

$$u(x) = \int dy\, e^{-\gamma\,|x-y| + i\,k_0\,(x-y)}\,f(y) \qquad (10)$$

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

$$-\,u_{xx}(x) + 2i\,k_0\,u_x(x) + (k_0^2+\gamma^2)\,u(x) = \gamma^2\,f(x)\,. \qquad (11)$$

i.e.

$$\begin{aligned} -\alpha_{xx} - 2\,k_0\,\beta_x + (\gamma^2+k_0^2)\,\alpha &= \gamma^2\,f(x) \\ -\beta_{xx} + 2\,k_0\,\alpha_x + (\gamma^2+k_0^2)\,\beta &= 0\,. \end{aligned} \qquad (12)$$

where $u(x) = \alpha(x) + i\,\beta(x)$. Equivalently, the system can be rewritten as a single fourth order differential equation. For function $\alpha(x)$ we find

$$-\,\alpha_{xxxx} + 2\,(\gamma^2-k_0^2)\,\alpha_{xx} - E^2\,\alpha = \gamma^2\,(f_{xx} - E\,f)\,. \qquad (13)$$

Fig. 1 illustrates the disparity function obtained by means of the Cusp filter 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 2 pixels, the second by 4 pixels. Ripples in the disparity function in correspondence to image structures are a consequence of the sharp peak of the Cusp envelope.

Fig. 3 shows the disparity map obtained from a stereo pair of computer generated images using

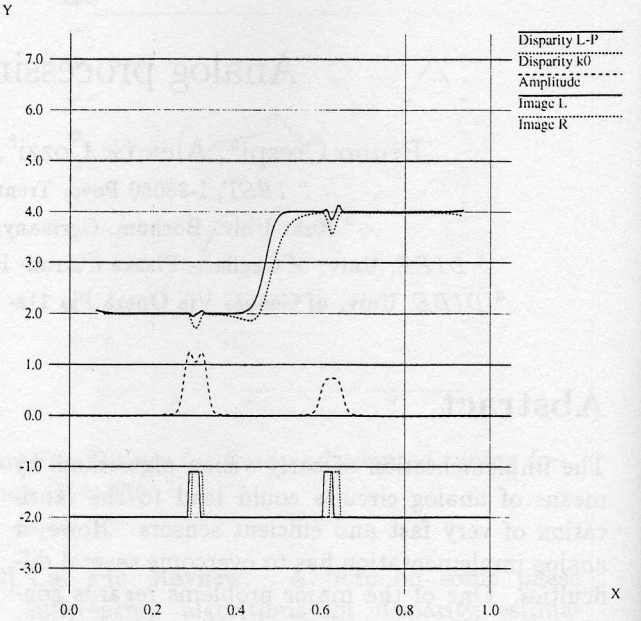


Figure 1: *Showing the computed disparity (in pixels) using the local, $\psi'(x)$, and the tuning frequency, k_0, and the product of the amplitudes, $(\rho_L^2(x) \times \rho_R^2(x))$. On the lower part, the left and right 1-D images are shown.*

the Cusp envelope. As usual, points where the algorithm is not reliable are drawn black. The wavy behavior of the disparity estimates at the borders of the tiles corresponds to the ripples evidenced in Fig. 1. For this image, a comparison between errors and densities of the Cusp and the Gaussian filters is shown in Fig. 4. The error measure is defined as the absolute difference between the true disparity and the value estimated by the algorithm. Density is the ratio between the number of points where disparity computation is accepted and the total number of points. Filters that are smoother at the origin and have a sharper spectrum decay can be generated form eq. (8). However, this determines an increase of the order of the differential equations.

2.2 The Quartic filter

If the k^4 terms are kept in the expansion of the Gaussian kernel. eq. (8), the Quartic envelope is found 5. Its analytical expression is

$$Q(x) = \frac{N}{\mu}\,\sin\left(\frac{\pi}{8} + \gamma_2\,|x|/\mu\right)\,e^{-\gamma_1\,|x|/\mu} \qquad (14)$$

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

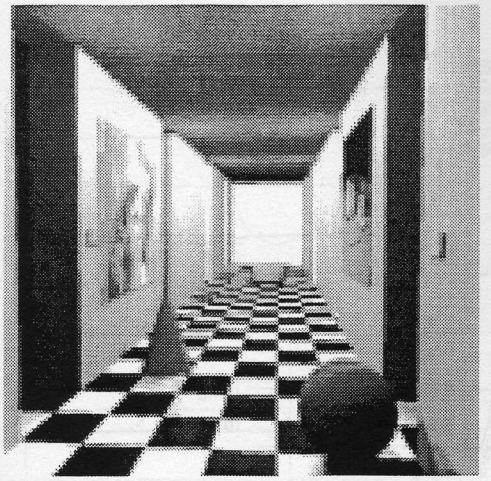


Figure 2: *Computer generated test image.*

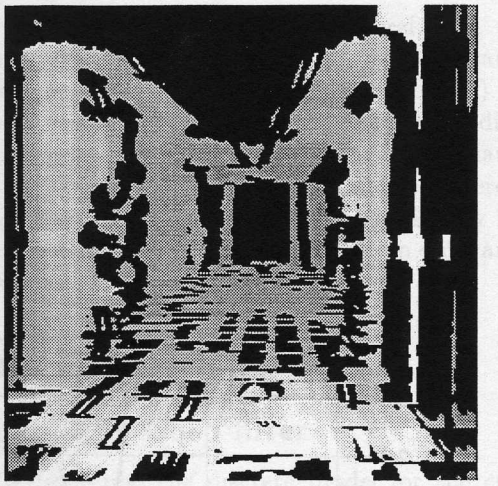


Figure 3: *Disparity map estimated using the Cusp filter. Singular points are black.*

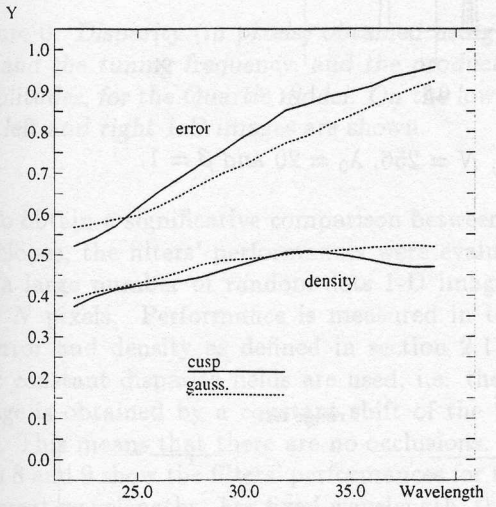


Figure 4: *Comparison between the Cusp and the Gaussian filters for the image shown in Fig. 2 as a function of the wavelength.*

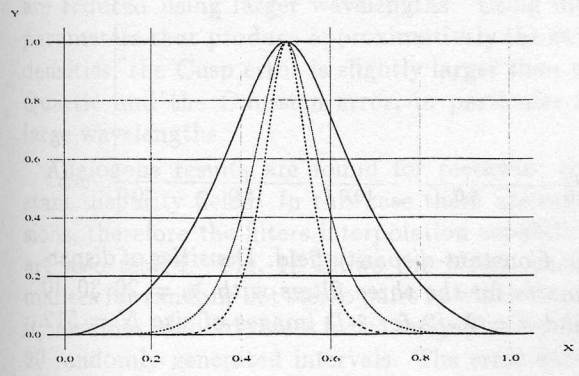


Figure 5: *The Quartic envelope for 3 values of μ.*

at the origin corresponds to the sharper decay of the frequency spectrum.

The solution of equation

$$\mu^2 \left(\frac{1}{i} \frac{d}{dx} - k_0 \right)^2 u(x) + \mu^4 \frac{1}{2!} \left(\frac{1}{i} \frac{d}{dx} - k_0 \right)^4 u(x) + u(x) = f(x) \tag{15}$$

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

Fig. 6 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.

3 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 the DE order is related to the number of connections per node necessary to realize the convolution, see Sect. 4. However, the change of envelope modifies the quality of disparity estimates [4]. In this section a comparison among the Cusp, the Quartic, and the Gaussian filters is presented.

For stereo signals characterized by simple disparity fields, the disparity estimates of the three filters are comparable. Fig. 7 shows 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.

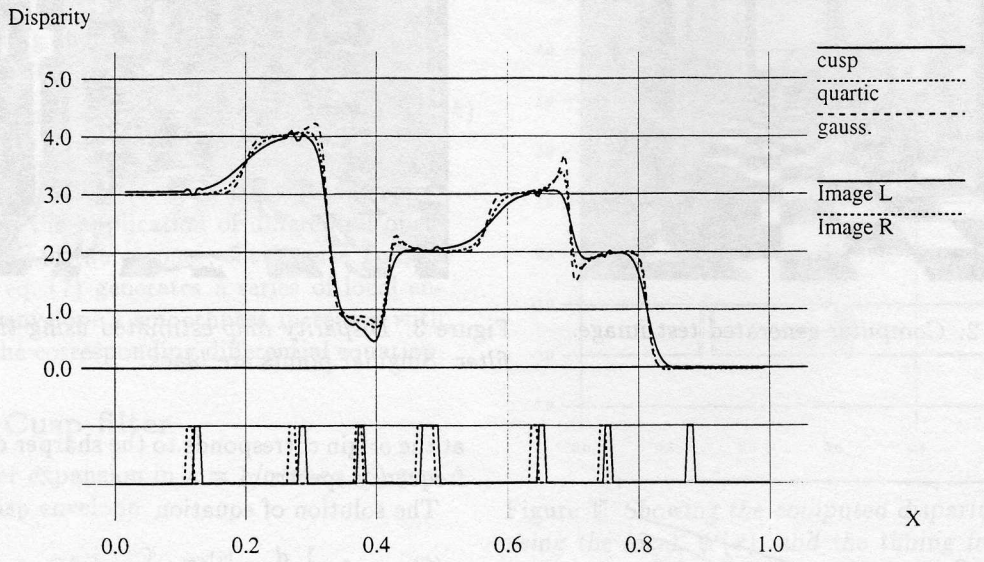


Figure 7: *Disparity estimates of the three filters,* $N = 256$, $\lambda_0 = 20$ *and* $\beta = 1$.

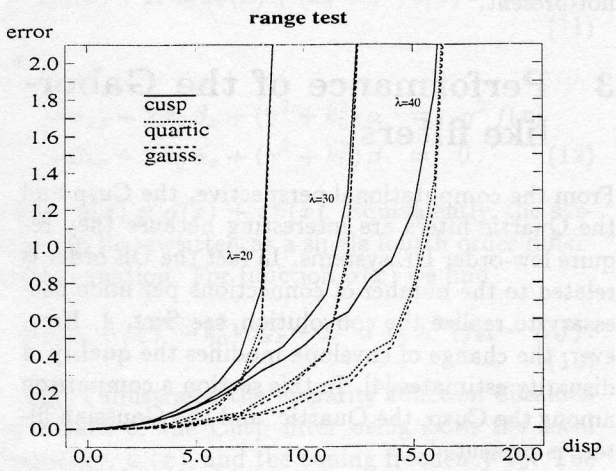


Figure 8: *Constant-disparity-field. Errors of disparity estimates for the three filters with* $\lambda_0 = 20, 30, 40$ *pixels and* $\sigma = \lambda_0/2$ *for 1-D images of size* $N = 512$ *pixels.*

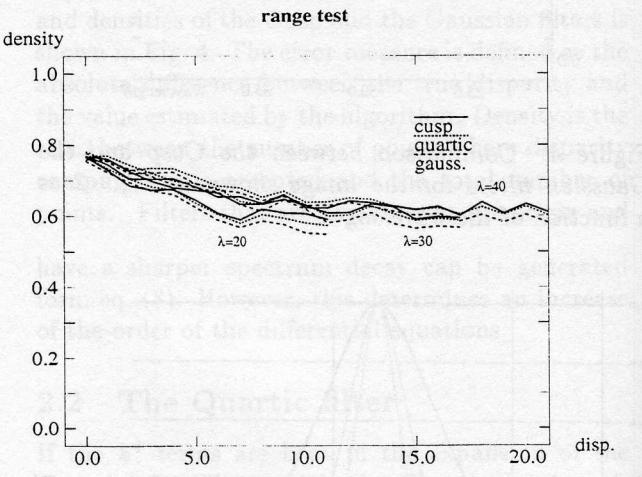


Figure 9: *Constant-disparity-field. Densities of disparity estimates for the three filters with* $\lambda_0 = 20, 30, 40$ *pixels and* $\sigma = \lambda_0/2$ *for 1-D images of size* $N = 512$ *pixels.*

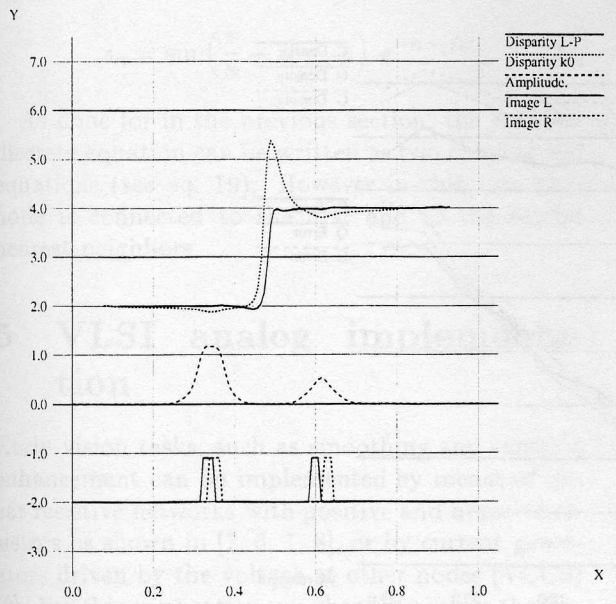


Figure 6: *Disparity (in pixels) obtained using the local and the tuning frequency, and the product of the amplitudes, for the Quartic model. On the lower part, the left and right 1-D images are shown.*

To obtain a significative comparison between the envelopes, the filters' performances were evaluated for a large number of random-dots 1-D images of size N pixels. Performance is measured in terms of error and density as defined in section 2.1. At first 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 8 and 9 show the filters' performances for three different wavelengths. For fixed wavelength, the filters behaviors follow the results found in ref. [3], 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. These are reduced using larger wavelengths. 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. Fig. 10 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. 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 ref. [11]. 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.

4 Discrete models

4.1 The Cusp filter

To determine the discrete model for the convolution with an oscillating decaying exponential let us consider the discrete version of eq.(11) i.e.

$$\begin{aligned} &A_1\, e^{-i\,k_0}\, u(n+1) + A_0\, u(n) \\ &+A_{-1}\, e^{+i\,k_0}\, u(n-1) = B\, f(n) \end{aligned} \tag{16}$$

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

$$u(n) = \sum_{k=-\infty}^{+\infty} e^{-\gamma\,|n-k| + i\,k_0\,(n-k)}\, f_k \tag{17}$$

three linear equations are derived. The system solution gives $A_{-1} = A_1 = -e^{-\gamma}$, $A_0 = 1 + e^{-2\gamma}$, and $B = C(1 - e^{-2\gamma})$, where C is an arbitrary constant.

Showing real and imaginary parts, $u_n = \alpha(n) + i\,\beta(n)$, we find the discrete version of system (12)

$$\begin{aligned} &a\,\alpha(n+1) + c\,\alpha(n) + a\,\alpha(n-1) - \\ &\quad b\,[\,\beta(n+1) - \beta(n-1)\,] = B\, f_n \\ &a\,\beta(n+1) + c\,\beta(n) + a\,\beta(n-1) + \\ &\quad b\,[\,\alpha(n+1) - \alpha(n-1)\,] = 0 \end{aligned} \tag{18}$$

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. (13), the above system can be rewritten as a single equation that involves the first 4 neighbors of each node, see Fig. 11. For $\alpha(n)$

$$\begin{aligned} &a_2\,\alpha(n+2) + a_1\,\alpha(n+1) + a_0\,\alpha(n) \\ &\quad +a_{-1}\,\alpha(n-1) + a_{-2}\,\alpha(n-2) = \\ &b_1\, f(n+1) + b_0\, f(n) + b_{-1}\, f(n-1) \end{aligned} \tag{19}$$

with

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

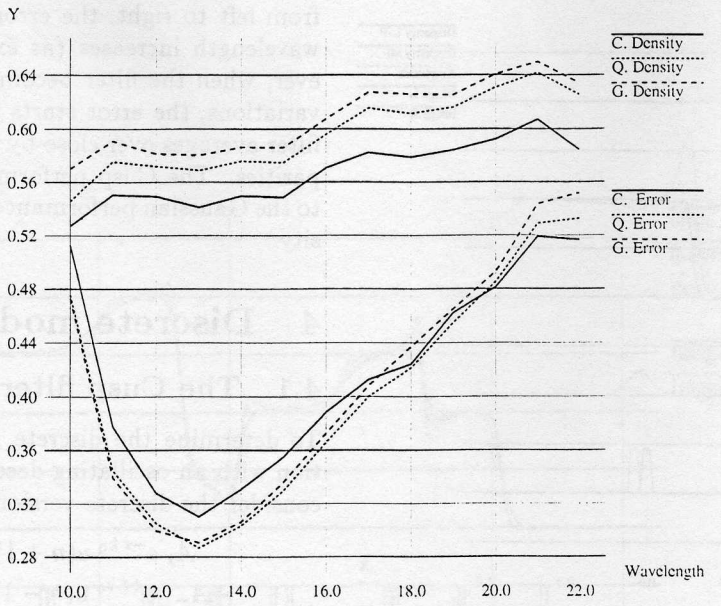


Figure 10: *Piecewise constant disparity field. Error and density for the three filters as a function of the wavelength, with* $\sigma = \lambda_0/2$ *and* $N = 256$ *. In each stereo pair, disparity assumes random values in the range of* $[0, 5]$ *pixels, in* 20 *randomly generated intervals.*

and

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

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

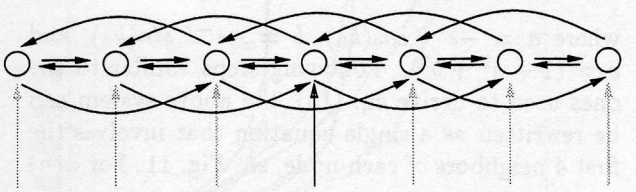


Figure 11: Schematic representation of a "Cusp" lattice network. Each node is connected with its nearest neighbors and with its next nearest neighbors (4 connections per node).

phase ϕ can be obtained from a proper choice of coefficients a, and b.

4.2 The Quartic filter

The discrete model for the convolution with an oscillating Quartic envelope

$$u(n) = \sum_{k=-\infty}^{+\infty} \sin\left(\frac{\pi}{8} + \gamma_2\,|n-k|/\mu\right) e^{-\gamma_1\,|n-k|/\mu + i\,k_0\,(n-k)}\,f(k) \qquad (20)$$

with $\gamma_1 = 2^{1/4}\cos\left(\frac{\pi}{8}\right)$, $\gamma_2 = 2^{1/4}\sin\left(\frac{\pi}{8}\right)$, can be derived in a similar way. Since convolution equation (16) is a fourth order differential equation, connectivity has to be extended to the second nearest neighbors

$$\begin{aligned} A_2\,e^{-2\,i\,k_0}\,u(n+2) + A_1\,e^{-i\,k_0}\,u(n+1) + \\ A_0\,u(n) + A_{-1}\,e^{+i\,k_0}\,u(n-1) + \\ A_{-2}\,e^{2\,i\,k_0}\,u(n-2) = B\,f(n) \end{aligned} \qquad (21)$$

From the discrete equation, 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 solution has to satisfy

$$\begin{pmatrix} s_4+s_0 & s_3+s_1 & s_0 \\ s_3+s_1 & s_2+s_0 & s_1 \\ 2\,s_2 & 2\,s_1 & s_0 \end{pmatrix} \begin{pmatrix} A_2 \\ A_1 \\ A_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ B \end{pmatrix} \qquad (22)$$

where

$$s_n = \sin\left(\frac{\pi}{8} + n\,\gamma_2/\mu\right)\, e^{-n\,\gamma_1/\mu} \qquad (23)$$

As done for in the previous section, the complex discrete equation can be written as two coupled real equations (see eq. 19). However in this case each node is connected to the first and to the second nearest neighbors.

5 VLSI analog implementation

Early vision tasks, such as smoothing and contrast enhancement can be implemented by means of linear resistive networks with positive and negative resistors as shown in [7, 6, 1, 8], or by current generators driven by the voltage at other nodes (VCCS) [2]. For this application, we choose to relate the lattice networks' equations to a circuit architecture of one-way-interacting elements implemented as current controlled current sources (CCCS). By using a current mode technique all signals are encoded by currents and the interaction can be implemented by current controlled current sources that feed or sink currents according to the current values of neighboring nodes. In this way, the analog signal processing of our circuit architecture will be based on ratios of matched component values [12], 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.

This technique allows to implement the interactions between the network nodes 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 [9].

Using the discrete model described in Sect. 4, a prototype 17-node VLSI circuit that perform the convolution of an input signal with the Cusp filter was fabricated by IRST [1] on its CMOS 2.0 μm, N-well, single poly, and double metal technology (see Fig. 12). Figure 13 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.

[1] Istituto per la Ricerca Scientifica e Tecnologica, Trento, Italy

6 Discussion and conclusions

This article discusses the analog VLSI implementation of a phase-based technique for disparity estimation. The initial step was to express the algorithm in terms of local operations. It is shown that 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 Cusp kernel – which is the simplest choice – leads to a fourth order differential equation. A successive approximation of the Gaussian kernel generates the Quartic filter that improves the quality of the disparity-field estimates but requires higher order differential equations.

The continuous models were translated into a self-consistent discrete lattice formalism in which the order of the differential equation system is related to the number 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 Cusp requires that each node is connected with the nearest four neighbors. The Quartic filter requires to increase connectivity to eight neighbors per node.

The Cusp filter model was used to design and implement a demonstrative analog chip that performs convolutions in real-time. The chip, fabricated by IRST, consists of a linear array of 17 nodes, in which each node is connected with its four nearest neighbors. The development of larger versions of this circuit could lead to important applications because it is capable of performing convolutions with Gabor-like kernels in microseconds. Besides, if the Quartic filter model would be implemented, performance very close to those of the "optimal" Gaussian filter could be reached. 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.

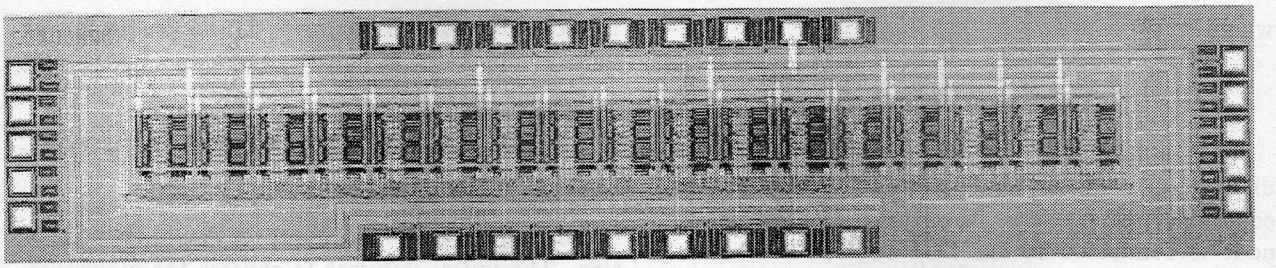
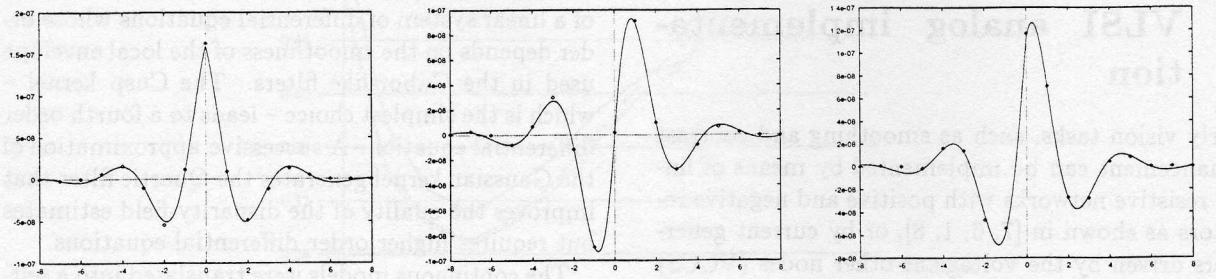


Figure 12: Microphotograph of a 1-D 2nd order network chip



(a) (b) (c)

Figure 13: Comparisons between measured (dots) and theoretical (continuous lines) Gabor-like impulse responses of phases 0 (a), $\pi/2$ (b) and $\pi/4$ (c).

References

[1] X. Arreguit and E.A. Vittoz. Perception systems implemented in analog VLSI for real-time applications. In *Proc. of the PerAc'94 Conference: From Perception to Action*, Lausanne, Sept.5-9, 1994, 1994.

[2] G.M. Bisio, M. Bruccoleri, P. Cusinato, L. Raffo, and S.P. Sabatini. An analog VLSI massively parallel module for low-level cortical processing in machine vision. In *Proc. IEEE-MICRONEURO94 - Torino (Italy)*, 1994.

[3] A. Cozzi, B. Crespi, F. Valentinotti, and F. Woergoetter. Performance of phase-based algorithms. *to appear on the Machine Vision and Application special issue on Performance Characteristics of Vision Algorithms.*

[4] D. J. Fleet and A. D. Jepson. Stability of Phase Information. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 15(12):1253–1268, 1993.

[5] D. J. Fleet, A. D. Jepson, and M. Jenkin. Phase–based Disparity Measurement. *CVGIP: Image Understanding*, 53(2):198–210, March 1991.

[6] H. Kobayashi, J.L. White, and A.A. Abidi. An active resistor network for Gaussian filtering of images. *IEEE J. on Solid State Circuits*, 26:738–748, May 1991.

[7] C.A. Mead. *Analog VLSI and Neural Systems.* Addison-Wesley, Reading, 1989.

[8] L. Raffo. Resistive network implementing maps of Gabor functions of any phase. *Electronics Letters*, 31(22):1913–1914, 1995.

[9] L. Raffo, S.P. Sabatini, G.M. Bo, and G.M. Bisio. Visual perception microsystems based on distributed analog VLSI processing. In *Proc. 2nd Int. Workshop on Mechatronical Computer Systems for Perception and Action, MCPA'97*, Pisa, Italy, February 1997.

[10] T. D. Sanger. Stereo disparity computation using Gabor filters. *Biol. Cybern.*, 1988.

[11] F. Valentinotti, G. Di Caro, and B. Crespi. Real-time parallel computation of disparity and optical flow using phase difference. *Machine Vision and Applications*, 9(3):87–96, 1996.

[12] E.A. Vittoz. Analog VLSI signal processing: Why, where, and how? *J. of VLSI Signal Processing*, 8:27–44, 1994.