We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

4,400
Open access books available

117,000
International authors and editors

130M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Towards a Biologically Plausible Stereo Approach

José R.A. Torreão and Silvia M.C. Victer
Universidade Federal Fluminense
Brazil

1. Introduction

A computational model for stereoscopic disparity estimation has been recently introduced which reformulates, in neurophysiologically plausible terms, the traditional image matching approach (Torreão, 2007). The left and right stereo images are assumed viewed through the receptive fields of cortical simple cells, modeled as Gabor functions (Marcelja, 1980). The Green's functions of a matching equation (whose uniform solutions are also Gabor functions) are used to filter the receptive-field modulated inputs, introducing different relative shifts between them. A measure of the local degree of matching of such shifted inputs can then be used for the estimation of stereoscopic disparities, in a way which is reminiscent of the energy model for the responses of cortical complex cells (Adelson & Bergen, 1985; Ohzawa et al., 1997).

Although based on well-established neurophysiological concepts, the Green's function approach is still far from being biologically plausible. For instance, it assumes as cortical inputs the original irradiance images, disregarding the transformations performed by the earlier stages of the visual pathway − namely, the retina and the lateral geniculate nucleus. As a further step towards a fully biological stereo, we here introduce an improved version of the algorithm in (Torreão, 2007), incorporating a new concept of signal coding by input-dependent receptive fields (Torreão et al., 2009).

The classical receptive field description, which assumes a fixed, stimulus-independent spatial organization, has recently been challenged by neurophysiological indications that the receptive-field structure does change with neuronal input (Allman et al., 1985; Bair, 2005; David et al., 2004). In (Torreão et al., 2009), a signal coding scheme has been introduced, where the parameters of the coding functions are obtained from the Fourier transform of the coded signal, and, in (Torreão & Victer, 2010), such scheme has been taken up as a model for stimulus-dependent center-surround (CS) receptive fields, such as those found in the retina and in the lateral geniculate nucleus. Assuming that the role of the CS structures is that of decorrelating natural images, as suggested in (Attick & Redlich, 1992) and (Dan et al., 1996), center-surround receptive fields which code whitened versions of the input signals have been obtained. We show that, by incorporating a similar center-surround coding module into the Green's function stereo algorithm, we are able to obtain better quality disparity estimates, through an approach which is closer to the neurophysiological situation.

This chapter is organized as follows. In Section 2, we review the Green's function disparity estimation approach. Next, we present the stimulus-dependent model for cortical and center-surround receptive fields (Section 3), and in Section 4 we couple the latter with the Green's function stereo. The chapter ends with our concluding remarks.

2. Green's function stereoscopy

The estimation of binocular disparities is a key issue for stereo processing, both in natural and in artificial vision. Binocular disparities arise from positional differences of scene features projected in the two retinae, or in the two cameras of a binocular system. The computational approach for disparity estimation which seems more firmly grounded on neurophysiological findings is the one based on the energy model for cortical cells (Ohzawa et al., 1997). According to this model, a quadrature pair of *simple* cells responsive to the stimulation of either eye (*binocular* cells) filter the stereo images through their receptive fields, modeled by Gabor functions (Marcelja, 1980). The squared magnitude of the output of the quadrature pair is then assumed to constitute the stereo response of the *complex* cortical cell feeding on the pair. By considering a population of such model complex cells, it has been possible to obtain disparity estimates both from synthetic and from real-world stereograms (Chen & Qian, 2004; Qian, 1994).

In (Torreão, 2007), an alternative approach for disparity estimation, preserving most of the neurophysiologically plausible features of the energy model, has been introduced, starting from an image matching equation. In stereoscopy, assuming that the epipolar geometry is known, the matching equation takes a one-dimensional form, such as

$$I_l(x + U, y) = I_r(x, y) \tag{1}$$

where $U \equiv U(x, y)$ denotes the disparity field, and where I_l and I_r stand for the left and the right input images. Typically, an equation such as (1) is employed for the estimation of $U(x, y)$, given the stereo pair, but in (Torreão, 2007) it was taken as a constraint over the matching images, such that, given I_r, its perfect match is to be found, for simple forms of the disparity field.

2.1 Uniform disparity field

For instance, assuming uniform disparity, $U(x) = u$, where u is a constant, we can take a second-order Taylor series expansion of Eq. (1), to obtain

$$\frac{u^2}{2} I_l''(x) + u I_l'(x) + I_l(x) = I_r(x) \tag{2}$$

where the primes denote differentiation with respect to x (for simplicity, we henceforth omit the dependences on y). Eq. (2) can then be solved via the Green's function approach (Torreão et al., 2007), to yield

$$I_l(x) \approx I_r(x - u) = \int_{-\infty}^{\infty} G_u(x, x_0) I_r(x_0) dx_0 \tag{3}$$

where

$$G_u(x, x_0) \equiv G_u(x - x_0) = \begin{cases} \frac{2}{u} \sin\left(\frac{x - x_0}{u}\right) e^{-\left(\frac{x - x_0}{u}\right)}, & \text{if } x > x_0 \\ 0, & \text{if } x < x_0 \end{cases} \tag{4}$$

is the Green's function of Eq. (2), which amounts to the solution to that equation when its right-hand side is the impulse function, $\delta(x - x_0)$ (we are here assuming an unbounded image domain: $x \in [-\infty, \infty]$).

From Eq. (3) we then find that, to second order in u, convolving the right image, $I_r(x)$, with the Green's function $G_u(x)$ effects a shift of that image, to the right, by the fixed amount u. Similarly, the convolution with $G_u(-x)$ (keeping u as a positive parameter) would effect a similar shift to the left. More generally, we can consider a complex Green's kernel, by introducing the quadrature pair to $G_u(x - x_0)$, given as

$$H_u(x - x_0) = \begin{cases} \frac{2}{u} \cos\left(\frac{x - x_0}{u}\right) e^{-\left(\frac{x - x_0}{u}\right)}, & \text{if } x > x_0 \\ 0, & \text{if } x < x_0 \end{cases} \tag{5}$$

which yields a homogeneous solution to Eq. (2) − that is to say, a solution to that equation when its right-hand side is identically zero.

Thus, being

$$K_u(x - x_0) = G_u(x - x_0) + iH_u(x - x_0) \tag{6}$$

the complex Green's kernel, we would obtain, for the rightwards shifted version of $I_r(x)$,

$$I_r(x - u) = \int_{-\infty}^{\infty} K_u(x - x_0) I_r(x_0) dx_0 \tag{7}$$

up to second order in u.

2.2 Linear disparity field

The disparity estimation approach of (Torreão, 2007) is based on a similar Green's kernel as that of Eq. (6), but for a differential equation which approximates a linear matching constraint, that is to say, one for which the disparity field takes the form $U(x) = u + vx$, for u and v constants. More specifically, a rightwards shift is there performed by the complex kernel

$$K(x, x_0) = \begin{cases} 2k e^{[ik(x - x_0)]} e^{-\left[\frac{(x+a)^2 - (x_0+a)^2}{2\sigma^2}\right]}, & \text{if } x > x_0 \\ 0, & \text{if } x < x_0 \end{cases} \tag{8}$$

where σ and a are positive constants, and $k = a/\sigma^2$. The kernel $K(x, x_0)$ is the complex Green's function to the equation

$$\frac{1}{2}\left(\frac{\sigma^2}{a}\right)^2 I_l''(x) + \left(\frac{\sigma^2}{a} + \frac{\sigma^2}{a^2}x\right) I_l'(x) + \left(1 + \frac{x^2 + 2ax + \sigma^2}{2a^2}\right) I_l(x) = I_r(x) \tag{9}$$

whose homogeneous solution is the Gabor function $e^{ikx} e^{-\frac{(x+a)^2}{2\sigma^2}}$.

When $|x|$ and σ are both much smaller than a, Eq. (9) can be approximated as

$$\frac{1}{2}\left(\frac{\sigma^2}{a}\right)^2 I_l''(x) + \left(\frac{\sigma^2}{a} + \frac{\sigma^2}{a^2}x\right) I_l'(x) + I_l(x) = I_r(x) \tag{10}$$

which, up to first order in σ^2/a, corresponds to an image matching equation for the linear displacement field

$$U(x) = \frac{\sigma^2}{a} + \frac{\sigma^2}{a^2}x \qquad (11)$$

Similarly as in the uniform case, a leftwards image shift would be effected, in this linear disparity model, by the kernel $K^{(-)}(x, x_0) = K(-x, -x_0)$.

2.3 Green's function disparity estimation

Disparity estimation, in the Green's function approach, proceeds thus: the input images, I_l and I_r, are each multiplied by a Gabor function, yielding complex signals. These are then filtered, respectively, by the $K(x, x_0)$ and the $K^{(-)}(x, x_0)$ kernels, which effect spatial shifts and phase changes in the Gabor-modulated inputs (see below). The optimal spatial shift at each image location can thus be obtained by evaluating the match of the filtered images when different values of the kernel parameters are employed. This yields an estimate of the disparity map encoded by the stereo pair.

The relation of the Green's function approach to the neurophysiological models of stereoscopy stems from the following property of the Green's kernels: when filtering a Gabor-function modulated signal, they yield similarly modulated outputs, but for spatially shifted versions of the signal. For instance, let us consider the result of filtering the complex signal

$$I_1(x) = e^{i\kappa x}e^{-\frac{x^2}{2\sigma^2}}I_r(x) \qquad (12)$$

by the kernel $K(x, x_0)$. This can be shown to yield (Torreão, 2007)

$$I_2(x) = e^{i(\kappa x + \psi)}e^{-\frac{x^2}{2\sigma^2}}I_r(x - u) \qquad (13)$$

where $I_r(x - u)$ is given by Eq. (7), for $u = \sigma^2/a$, and where $\psi = \kappa u$. Thus, filtering signal $I_1(x)$ by the kernel $K(x, x_0)$ essentially preserves its Gabor modulating factor (with the introduction of a phase), but spatially shifts the modulated image.

Assuming that, locally − $i.e.$, under the Gaussian window of the Gabor modulating function −, the disparity between the right and left input images is well approximated by u, such that $I_l(x) \approx I_r(x - u)$, we would be able to rewrite Eq. (13) as

$$I_2(x) = e^{i(\kappa x + \psi)}e^{-\frac{x^2}{2\sigma^2}}I_l(x) \qquad (14)$$

Together with Eq. (12), this would then mean that $I_1(x)$ and $I_2(x)$ correspond, respectively, to a right and a left stereo images, as seen through the receptive fields of a quadrature pair of simple cortical cells, according to the so-called *phase-shift* model of stereo responses (Fleet et al., 1991).

Thus, operating on the equivalent to the *right-eye* cortical input (*viz.*, the right-eye *retinal* image as seen through the simple cell right-eye receptive field), the Green's kernel $K(x, x_0)$ produces the equivalent to the *left-eye* cortical input (*viz.*, the left-eye *retinal* image as seen through the simple cell left-eye receptive field).

If we change right for left and left for right in the foregoing development, it also becomes valid for the Green's kernel $K^{(-)}(x, x_0)$. Namely, being

$$I_1^{(-)}(x) = e^{i\kappa x} e^{-\frac{x^2}{2\sigma^2}} I_l(x) \tag{15}$$

the Gabor-modulated signal to be filtered by $K^{(-)}(x, x_0)$, we obtain

$$I_2^{(-)}(x) = e^{i(\kappa x - \psi)} e^{-\frac{x^2}{2\sigma^2}} I_l(x + u) \tag{16}$$

where, once again, $u = \sigma^2/a$. And Eq. (16) can also be rewritten, similarly to Eq. (14), as

$$I_2^{(-)}(x) = e^{i(\kappa x - \psi)} e^{-\frac{x^2}{2\sigma^2}} I_r(x) \tag{17}$$

Thus, operating on the equivalent to the *left-eye* cortical input (*viz.*, the left-eye *retinal* image as seen through the simple cell left-eye receptive field), the Green's kernel $K^{(-)}(x, x_0)$ produces the equivalent to the *right-eye* cortical input (*viz.*, the right-eye *retinal* image as seen through the simple cell right-eye receptive field).

Therefore, if we now compute the local match between the signals $I_2(x)$ and $I_2^{(-)}(x)$ − for instance, through the measure

$$R(x) = |I_2(x) - I_2^{(-)}(x)|^2 \tag{18}$$

− we will be effectively assessing the match between a left- and a right- cortical images, under the assumption that the local disparity is $2u$ (recall Eqs. (13) and (16)). Different u values can be obtained by changing the parameter a of the Green's kernels, while keeping σ and κ fixed, as proposed in (Torreão, 2007). This allows the estimation of the local disparity, for instance, as

$$d(x) = \frac{2\sigma^2}{\bar{a}(x)} \tag{19}$$

where $\bar{a}$ is that a value for which the measure $R(x)$ is minimized.

Incidentally, the measure $R(x)$ affords the comparison of the Green's function approach with the energy model for disparity estimation (Qian, 1994; Qian & Miakelian, 2000). Using Eqs. (13) and (16), we can rewrite it as

$$R(x) = e^{\frac{-x^2}{\sigma^2}} |I_r(x - u) - e^{i\Delta\psi} I_l(x + u)|^2 \tag{20}$$

where $\Delta\psi = -2\kappa u$. When considered for $x = 0$, Eq. (20), apart from the minus sign in front of the phase factor (which can be easily accomodated by assuming a phase difference of π between the Gabor modulating functions in Eqs. (12) and (15)), is similar to what is predicted for the complex cell response, by the *phase-and-position-shift* energy model. In the present case, the position shift is given by the disparity measure, $d = 2u$, and the phase shift, by $\Delta\psi$.

3. Stimulus-dependent receptive fields

The receptive field (RF) of a visual neuron defines the portion of the visible world where light stimuli evoke the neuron's response, and describes the nature of such response,

distinguishing, for instance, excitatory and inhibitory subfields (Hubel & Wiesel, 1962). In the standard description, the spatial organization of the RF remains invariant, and the neuronal response is obtained by filtering the input through a fixed receptive field function. Lately, this classical view has been challenged by neurophysiological experiments which indicate that the receptive field organization changes with the stimuli (Allman et al., 1985; Bair, 2005; David et al., 2004). Motivated by such findings, we have proposed a model for stimulus-dependent receptive field functions − initially for cortical simple cells (Torreão et al., 2009), and later for center-surround (CS) structures, such as found in the retina and in the lateral geniculate nucleus (Torreão & Victer, 2010). In what follows, we present a brief description of the model as originally proposed for simple cells − which is mathematically easier to handle −, later extending it to the CS structures.

3.1 Cortical receptive fields
Let us first consider a one-dimensional model. Being $I(x)$ any square-integrable signal, it can be expressed as

$$I(x) = \int_{-\infty}^{\infty} e^{i[\omega x + \varphi(\omega)]} e^{-\frac{x^2}{2\sigma^2(\omega)}} * e^{i\omega x} d\omega \tag{21}$$

where the aterisk denotes a spatial convolution, and where $\sigma(\omega)$ and $\varphi(\omega)$ are related, respectively, to the magnitude and the phase of the signal's Fourier transform, $\tilde{I}(\omega)$, as

$$\tilde{I}(\omega) = (2\pi)^{3/2}\sigma(\omega)e^{i\varphi(\omega)} = \int_{-\infty}^{\infty} e^{-i\omega x} I(x) dx \tag{22}$$

Eq. (21) can be verified by rewriting the integral on its right-hand side in terms of the variable ω', and taking the Fourier transform (FT) of both sides. Using the linearity of the FT, and the property of the transform of a convolution, we obtain

$$\tilde{I}(\omega) = (2\pi)^{3/2} \int_{-\infty}^{\infty} \sigma(\omega')e^{i\varphi(\omega')} e^{-\frac{\sigma^2}{2}(\omega-\omega')^2} \delta(\omega-\omega')d\omega' \tag{23}$$

and, by making use of the sampling property of the delta,

$$\tilde{I}(\omega) = (2\pi)^{3/2}\sigma(\omega)e^{i\varphi(\omega)} \tag{24}$$

which is exactly the definition in Eq. (22).

Also, if we make the convolution operation explicit in Eq. (21), it can be formally rewritten as

$$I(x) = \int_{-\infty}^{\infty} < e^{i[\omega x + \varphi(\omega)]} e^{-\frac{x^2}{2\sigma^2(\omega)}}, e^{i\omega x} > e^{i\omega x} d\omega \tag{25}$$

where the angle brackets denote an inner product,

$$< f(x), g(x) > = \int_{-\infty}^{\infty} f(x)g^*(x) dx \tag{26}$$

with $g^*(x)$ standing for the complex conjugate of $g(x)$. Comparing Eq. (25) to a signal expansion on the Fourier basis set,

$$I(x) = \int_{-\infty}^{\infty} < I(x), e^{i\omega x} > e^{i\omega x} d\omega \tag{27}$$

we conclude that, in so far as the frequency ω is concerned, the Gabor function

$$e^{i[\omega x + \varphi(\omega)]} e^{-\frac{x^2}{2\sigma^2(\omega)}} \tag{28}$$

is equivalent to the signal $I(x)$. Thus, we have found a set of signal-dependent functions, localized in space and in frequency, which yield an exact representation of the signal, under the form

$$I(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i[\omega x + \varphi(\omega)]} e^{-\frac{(x-\xi)^2}{2\sigma^2(\omega)}} d\xi d\omega \tag{29}$$

which amounts to a Gabor expansion with unit coefficients (the above result can be easily verified, again by making the spatial convolution explicit in Eq. (21)).

In (Torreão et al., 2009), the above development has been extended to two dimensions, and proposed as a model for image representation by cortical simple cells, whose receptive fields are well described by Gabor functions (Marcelja, 1980). In the 2D case, Eq. (21) becomes

$$I(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_c(x,y;\omega_x,\omega_y) * e^{i(\omega_x x + \omega_y y)} d\omega_x d\omega_y \tag{30}$$

where

$$\psi_c(x,y;\omega_x,\omega_y) = e^{i[\omega_x x + \omega_y y + \varphi(\omega_x,\omega_y)]} e^{-\frac{(x^2+y^2)}{2\sigma_c^2(\omega_x,\omega_y)}} \tag{31}$$

is the model receptive field, with $\varphi(\omega_x,\omega_y)$ being the phase of the image's Fourier transform, and with $\sigma_c(\omega_x,\omega_y)$ being related to its magnitude, as

$$\sigma_c(\omega_x,\omega_y) = \frac{1}{(2\pi)^{3/2}} \sqrt{|\tilde{I}(\omega_x,\omega_y)|} \tag{32}$$

The validity of Eq. (30) can be ascertained similarly as in the one-dimensional case. Moreover, as shown in (Torreão et al., 2009), the same expansion also holds with good approximation over finite windows, with different σ_c and φ values computed locally at each window. Under such approximation, it makes sense to take the coding functions $\psi_c(x,y;\omega_x,\omega_y)$ as models for signal-dependent, Gabor-like receptive fields.

3.2 Center-surround receptive fields

A similar approach can be followed for neurons with center-surround organization, as presented in (Torreão & Victer, 2010). The role of the center-surround receptive fields — as found in the retina and in the lateral geniculate nucleus (LGN) — has been described as that of relaying decorrelated versions of the input images to the higher areas of the visual pathway (Attick & Redlich, 1992; Dan et al., 1996). The retina- and LGN-cells would thus have developed receptive field structures ideally suited to whiten natural images, whose spectra are known to decay, approximately, as the inverse of the frequency magnitude — *i.e.*, $\sim (\omega_x^2 + \omega_y^2)^{-1/2}$ (Ruderman & Bialek, 1994). In accordance with such interpretation, we have introduced circularly symmetrical coding functions which yield a similar representation as Eq. (30) for a whitened image, and which have been shown to account for the neurophysiological properties of center-surround cells.

Specifically, we have

$$I_{white}(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(r;\omega_x,\omega_y) * e^{i(\omega_x x + \omega_y y)} d\omega_x d\omega_y \tag{33}$$

where $I_{white}(x,y)$ is a whitened image, and where $\psi(r;\omega_x,\omega_y)$ is the CS receptive field function, with $r = \sqrt{x^2 + y^2}$. Following the usual approach (Attick & Redlich, 1992; Dan et al., 1996), we have modeled the whitened image as the result of convolving the input image with a zero-phase whitening filter,

$$I_{white}(x,y) = W(x,y) * I(x,y) \tag{34}$$

where $W(x,y)$ is such as to equalize the spectrum of natural images, at the same time suppressing high-frequency noise. The whitening filter spectrum has thus been chosen under the form

$$\tilde{W}(\omega_x,\omega_y) = \frac{\rho}{1 + \kappa\rho^2} \tag{35}$$

where κ is a free parameter, and $\rho = \sqrt{\omega_x^2 + \omega_y^2}$.

On the other hand, the signal-dependent receptive field has been taken under the form

$$\psi(r;\omega_x,\omega_y) = -\frac{e^{i\varphi(\omega_x,\omega_y)}}{\pi r} \left\{ 1 - \cos[\sigma(\omega_x,\omega_y)\pi r] - \sin[\sigma(\omega_x,\omega_y)\pi r] \right\} \tag{36}$$

where φ is the phase of the Fourier transform of the input signal, as already defined, while $\sigma(\omega_x,\omega_y)$ is related to the magnitude of that transform, as

$$\sigma(\omega_x,\omega_y) = \frac{\rho}{\pi} \sqrt{1 - \left[1 + \frac{\rho\tilde{W}(\omega_x,\omega_y)|\tilde{I}(\omega_x,\omega_y)|}{4\pi} \right]^{-2}} \tag{37}$$

Eq. (37) can be verified by introducing the above $\psi(r;\omega_x,\omega_y)$ into Eq. (33), and taking the Fourier transform of both sides of that equation.

We remark that the most commonly used model of center-surround receptive fields, the difference of Gaussians (Enroth-Cugell et al., 1983), has not been considered in the above treatment, since it would have required two parameters for the definition of the coding functions, while our approach provides a single equation for this purpose.

Fig. 1 shows plots of the coding functions obtained from a 3×3 fragment of a natural image, for different frequencies. The figure displays the magnitude of $\psi(r;\omega_x,\omega_y)$ divided by σ, such that all functions reach the same maximum of 1, at $r = 0$. Each coding function displays a single dominant surround, whose size depends on the spectral content of the coded image at that particular frequency (when the phase factor in Eq. (36) is considered, we obtain both center-on and center-off organizations). For $\rho = 0$, σ vanishes, and the coding function becomes identically zero, meaning that the proposed model does not code uniform inputs. At low frequencies, the surround is well defined (Fig. 1a), becoming less so as the frequency increases (Fig. 1b), and all but disappearing at the higher frequencies (Fig. 1c). All such
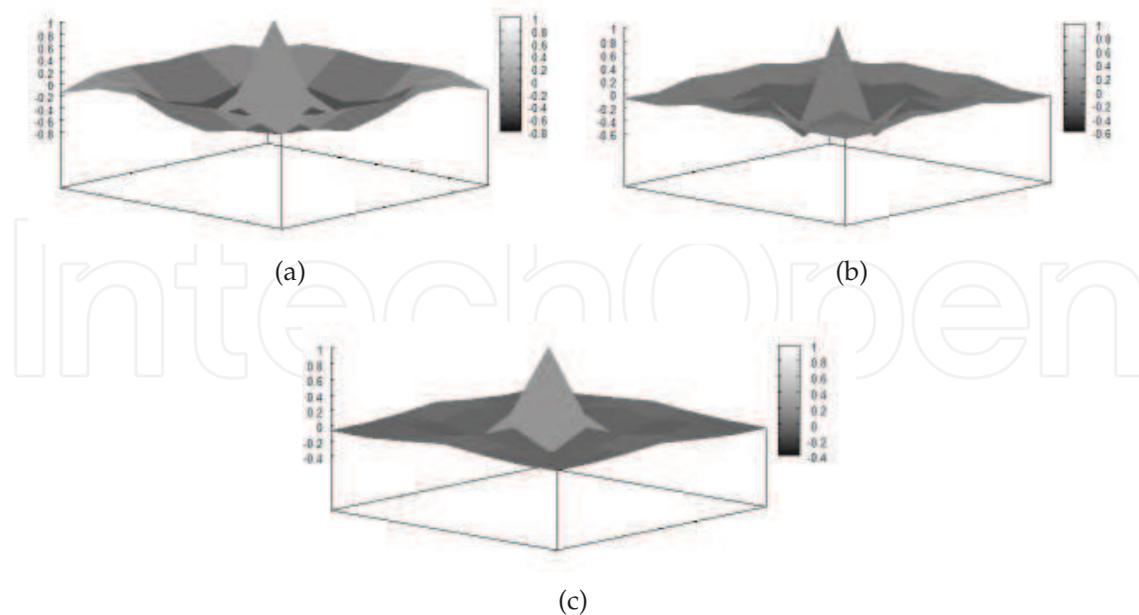
Fig. 1. Plots of the magnitude of the coding functions of Eq. (36), obtained from a 3×3 fragment of a natural image. The represented frequencies, (ω_x, ω_y), in (a), (b) and (c), are (0,1), (2,0), and (3,1), respectively.

properties are consistent with the behavior of retinal ganglion cells, or of cells of the lateral geniculate nucleus.

Fig. 2 shows examples of image coding by the signal-dependent CS receptive fields. The whitened representation is obtained, for each input, by computing Eq. (33) over finite windows. As shown by the log-log spectra in the figure (the vertical axis plots the rotational average of the log magnitude of the signal's FT, and the horizontal axis is $\log \rho$), the approach tends to equalize the middle portion of the original spectra, yielding representations similar to edge maps which code both edge strength and edge polarity. We have observed that the effect of the κ parameter in Eq. (35) is not pronounced, but, consistent with its role as a noise measure, larger κ values usually tend to enhance the low frequencies.

In the following section, we will use the whitened representation of stereo image pairs as input to the Green's function algorithm of Section 2, showing that this allows improved disparity estimation through an approach which is closer to the neurophysiological situation.

4. Green's function stereo with whitened inputs

We have incorporated a whitening routine into the stereo matching algorithm of Section 2, such that the Green's function procedure is now performed over signals which emulate the neurocortical input from the lower visual areas. Thus, what we have identified as *retinal* images, in Section 2.3, become the whitened representations of the stereo pair, obtained, through Eq. (33), by means of center-surround, signal-dependent receptive field functions. Disparity estimation proceeds as described, with the proviso that, similarly as in (Torreão, 2007), instead of choosing precisely that disparity which minimizes the matching measure $R(x)$ (Eq. (18)), we take, as our estimate, the sum of all disparities considered, each weighted by the inverse of the corresponding $R(x)$ value. This yields a dense disparity map, avoiding the need of interpolation. The preprocessing alignment of the stereo pair, through
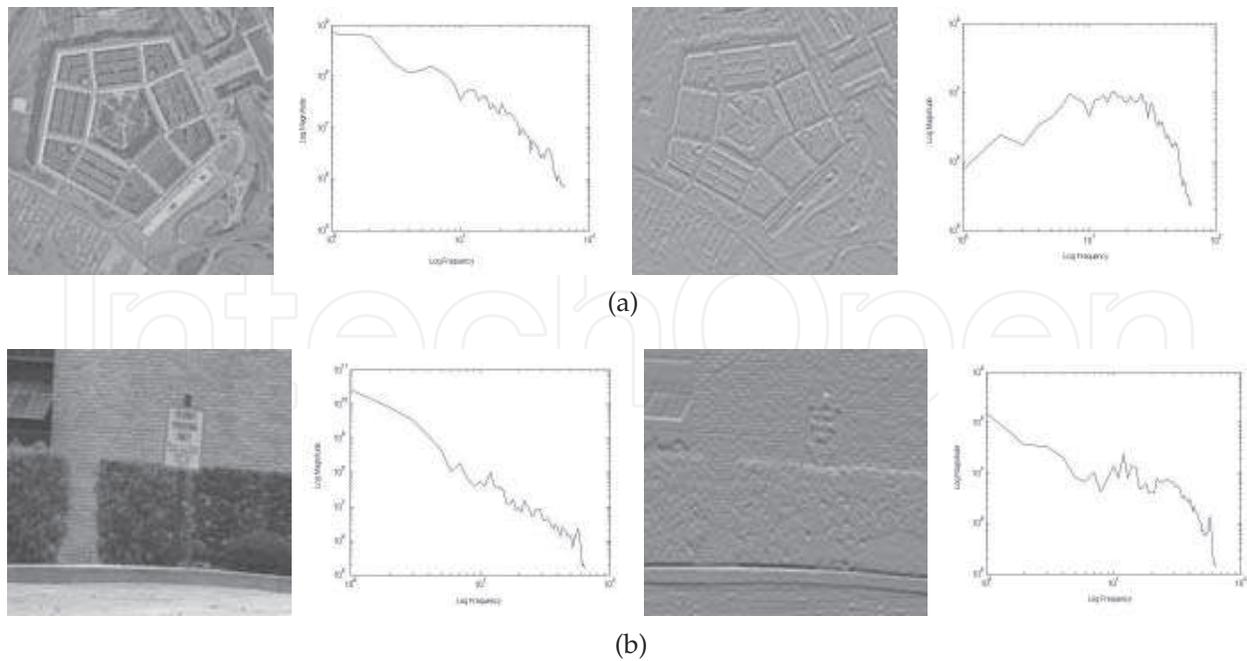
(a)

(b)

Fig. 2. a) Left: input image and its log-log spectrum. Right: whitened image and its log-log spectrum. Similarly for b). We have used 3×3 windows, with $\kappa = 0.05$.

uniform-disparity matching, has also been followed (Torreão, 2007), as a means for handling large overall disparities.

Figs. 3 to 6 depict some results of the whitened Green's function stereo, along with those yielded by the original approach. It is apparent that the former generally affords better spatial resolution, as well as sharper disparity definition. Being based on whitened inputs, the new algorithm proves less sensitive to image features not depth related, as can be seen in the background region of the meter stereo pair (Fig. 6), more uniformly captured by the whitened approach than by the original one, which is biased by the complex edge structure in the region.

5. Conclusion

In this chapter, we have reviewed the Green's function stereoscopy (Torreão, 2007) − a neurophyisiologically-inspired stereo matching approach −, along with a recently introduced model for signal-dependent receptive fields (Torreão et al., 2009; Torreão & Victer, 2010). By coupling the Green's function algorithm with a whitening representation of the stereo inputs, based on center-surround, signal-dependent receptive field functions, we have been able to obtain better disparity estimates, through an approach which is closer to the neurophysiological situation. We are presently working on the incorporation, into our stereo algorithm, of the Gabor-like, signal-dependent receptive model of Section 3.1. This will allow a more realistic parallel with the cortical mechanisms of biological stereo vision.
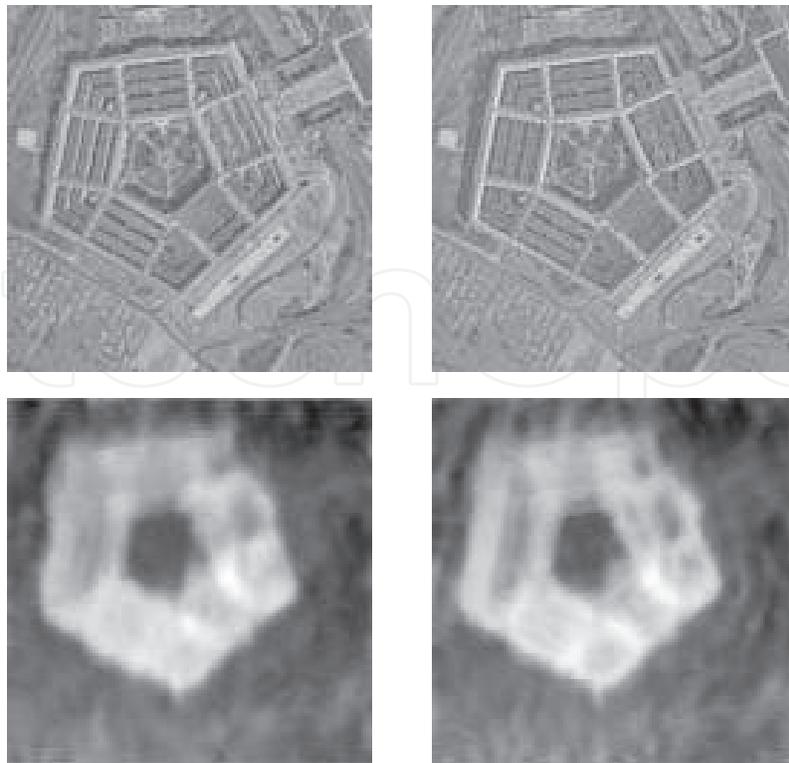
Fig. 3. Top: Pentagon stereo pair. Bottom: Disparity maps obtained through the whitened Green's function approach (left) and through the original approach of (Torreão, 2007) (right).
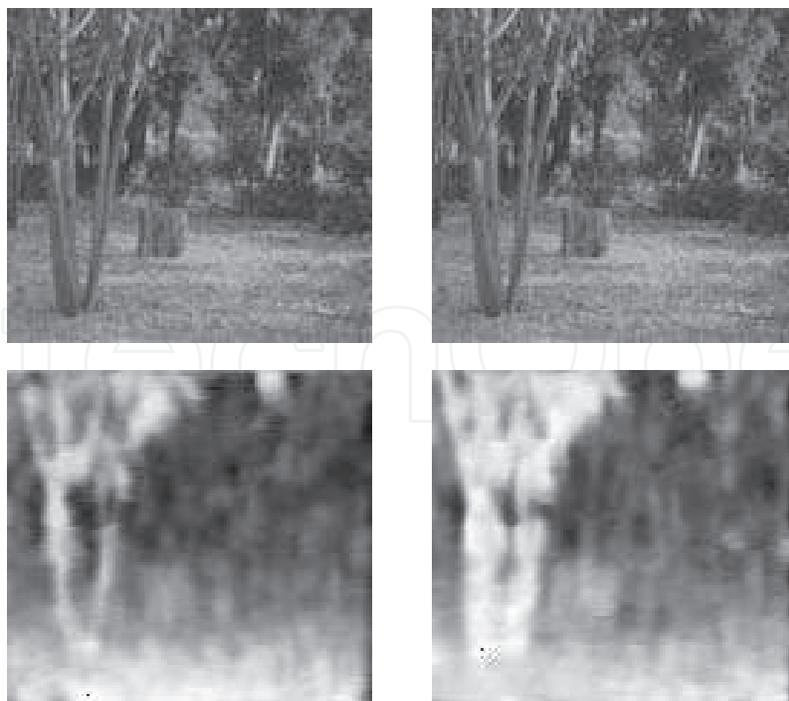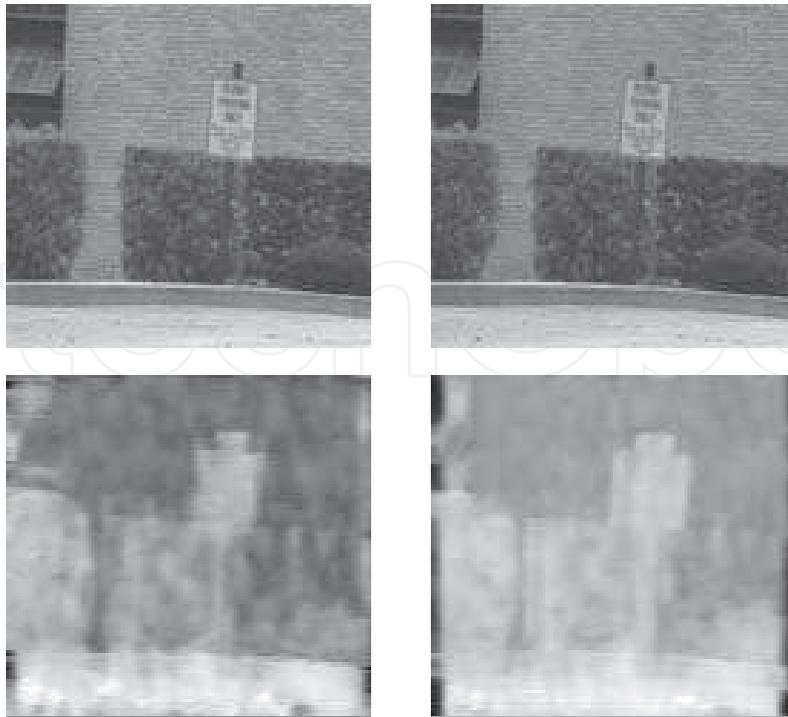

Fig. 4. Top: Tree stereo pair. Bottom: Disparity maps obtained through the whitened Green's function approach (left) and through the original approach of (Torreão, 2007) (right).

Fig. 5. Top: Shrub stereo pair. Bottom: Disparity maps obtained through the whitened Green's function approach (left) and through the original approach of (Torreão, 2007) (right).

Fig. 6. Top: Meter stereo pair. Bottom: Disparity maps obtained through the whitened Green's function approach (left) and through the original approach of (Torreão, 2007) (right).

6. References

Adelson, E.H. & Bergen, J.R. (1985). Spatiotemporal energy models for the perception of motion. J. Opt. Soc. Am. A 2, pp. 284–299.

Allman, J.; Miezin, F. & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparison in visual neurons. Annu. Rev. Neuroscience 8, pp. 407–430.

Atick, J.J. & Redlich, A.N. (1992). What does the retina know about natural scenes? Neural Comp. 4, pp. 196–210.

Bair,W. (2005). Visual receptive field organization. Current Opinion in Neurobiology 15(4), pp. 459–464.

Chen, Y. & Qian, N. (2004). A coarse-to-fine disparity energy model with both phase-shift and position-shift receptive field mechanisms. Neural Comput. 16, pp. 1545–1577.

Dan, Y., Atick, J.J. & Reid, R.C. (1996). Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. J. Neurosci. 16(10), pp. 3351–3362.

David, S.V.; Vinje, W.E. & Gallant, J.L. (2004). Natural stimulus statistics alter the receptive field structure of V1 neurons, The Journal of Neuroscience 24(31), pp. 6991–7006.

Enroth-Cugell, C.; Robson, J.G.; Schweitzer-Tong, D.E., & Watson, A.B. (1983). Spatiotemporal interactions in cat retinal ganglion cells showing linear spatial summation, J. Physiol. 341, pp. 279–301.

Fleet, D.J.; Jepson, A.D. & Jenkin, M.R.M. (1991). Phase-based disparity measurement. Comp. Vis. Graphics Image Proc. 53(2), pp. 198–210.

Hubel, D.H. & Wiesel, T. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. 160, pp. 106–154.

Marcelja, S. (1980). Mathematical description of the responses of simple cortical cells. J. Opt. Soc. Am. A 70, pp. 1297–1300.

Ohzawa, I.; DeAngelis, J.C. & Freeman, R.D. (1997). Encoding of binocular disparity by complex cells in the cat's visual cortex. J. Neurophysiol. 77(6), pp. 2879–2909.

Qian, N. (1994). Computing stereo disparity and motion with known binocular cell properties. Neural Comput. 6, pp. 390–404.

Qian, N. & Mikaelian, S. (2000). Relationship between phase and energy methods for disparity computation. Neural Comput. 12, pp. 303–316.

Ruderman, D.L. & Bialek, W. (1994). Statistics of natural images: Scaling in the woods, Phys. Rev. Lett. 73(6), pp. 814–817.

Torreão, J.R.A. (2007). Disparity estimation through Green's functions of matching equations. Biol. Cybernetics 97, pp. 307–316.

Torreão, J.R.A.; Fernandes, J.L.; Amaral M.S. & Beltrão, L. (2007). Green's functions of matching equations: A unifying approach for low-level vision problems In: *Vision Systems: Segmentation and Pattern Recognition*, G. Obinatta and A. Dutta, pp. 381-396, I-Tech Education and Publishing, ISBN 978-3-902613-05-9, Vienna, Austria.

Torreão, J.R.A.; Fernandes, J.L. & Victer, S.M.C. (2009). A model for neuronal signal representation by stimulus-dependent receptive fields, *Proceedings of the 19th International Conference on Artificial Neural Networks - ICANN 2009*, Limassol, Cyprus, Springer, vol. I, pp. 356–362.

Torreão, J.R.A. & Victer, S.M.C. (2010). A model for center-surround stimulus dependent receptive fields, *Proceedings of the 20th International Conference on Artificial Neural Networks - ICANN 2010*, Thessaloniki, Greece, Springer, vol. I, pp. 305–310.

Advances in Stereo Vision

Edited by Prof. Jose R.A. Torreao

Stereopsis is a vision process whose geometrical foundation has been known for a long time, ever since the experiments by Wheatstone, in the 19th century. Nevertheless, its inner workings in biological organisms, as well as its emulation by computer systems, have proven elusive, and stereo vision remains a very active and challenging area of research nowadays. In this volume we have attempted to present a limited but relevant sample of the work being carried out in stereo vision, covering significant aspects both from the applied and from the theoretical standpoints.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jose R.A. Torreao and Silvia Victer (2011). Towards a Biologically Plausible Stereo Approach, Advances in Stereo Vision, Prof. Jose R.A. Torreao (Ed.), ISBN: 978-953-307-837-3, InTech, Available from: http://www.intechopen.com/books/advances-in-stereo-vision/towards-a-biologically-plausible-stereo-approach

INTECH
open science | open minds