TR－A－0154

Direct Estimation of Multiple Disparities for Transparent Multiple Surfaces in Binocular Stereo

Masahiko Shizawa

1992.10.20

ATR視聴覚機構研究所

〒619-02 京都府相楽郡精華町光台2-2 ☎07749-5-1411

ATR Auditory and Visual Perception Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1411
Facsimile: +81-7749-5-1408

Direct Estimation of Multiple Disparities for Transparent Multiple Surfaces in Binocular Stereo

Masahiko Shizawa
Cognitive Process Department
ATR Auditory and Visual Perception Research Laboratories
Advanced Telecommunications Research Institute International
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-02, Japan

October 20, 1992

abstract

A closed-form single-shot stereo disparity estimation algorithm is proposed that can compute multiple disparities due to transparency directly from signal differences and variations on epipolar lines of a binocular image pair. The *transparent stereo constraint equations* have been derived by using a novel mathematical technique, the *principle of superposition*. A computationally tractable single-shot algorithm is derived by using the first-order approximation of the constraint equations with respect to disparities. The algorithm can compute multiple dispatities from only *two* images in contrast to the previous results for motion transparency that needed at least $n+1$ frames for n simultaneous motion estimates. The derived algorithm can be viewed as the SSD(sum of squared differences) for signal matching extended to deal with multiple disparities. However, the constraint is not dedicated solely to SSD method and several other implementations are possible. These possibilities are also discussed in this paper.

Keywords: Binocular stereo vision, Disparity, Transparency, Single-shot algorithm, Principle of superposition.

Acknowledgments

The author thanks Drs. Eiji Yodogawa and Katsunori Shimohara of ATR Auditory and Visual Perception Research Laboratories for their kind support.

Contents

1 Introduction

The use of binocular stereo vision for detecting disparities has drawn considerable attentions in computational vision [Julesz 1960, Marr & Poggio 1976, Marr & Poggio 1979, Burt & Julesz 1980a, Grimson 1981, Mayhew & Frisby 1981, Prazdny 1985, Pollard et al. 1985]. Stereo vision is also being investigated as an engineering method for passively obtaining the depth and structure of a scene [Baker & Binford 1981, Dhond & Aggarwal 1989, Lucas & Kanade 1981, Ohta & Kanade 1985, Okutomi & Kanade 1992]. Most stereo vision algorithms, however, have not incorporated transparency, i.e., multiple surface perception at the same retinal position. Recent psychophysics has revealed capabilities and limitations in the human perception of stereoscopic transparency [Akerstrom & Todd 1988, Weinshall 1989, Pollard & Frisby 1990, Weinshall 1991]. This transparency is also an important problem for engineering computer vision thst must handle realistic environments including complex occlusions, transparency, translucency, and several overlapping surfaces. These complex environments may be caused by fences, trees, bushes, glass, water surfaces, and windowpanes that are common in both natural and artificial scenes. Although a few papers have tried to handle transparency at the level of algorithms for the feature-based matching and area-based correlation [Pollard et al. 1985, Weinshall 1992], transparency had not been sufficiently recognized by the computer vision community in the past years.

Prazdny has pointed out that the discontinuities of a surface can be regarded as two surfaces in the vicinities of the discontinuities [Prazdny 1985]. He also proposed an algorithm based on the disparity gradient limit [Burt & Julesz 1980a]. Prazdny's algorithm does not have inhibitory neighborhood interactions and it can therefore handle multiple surfaces. A similar algorithm, the PMF algorithm has been proposed by Pollard et al. [Pollard et al. 1985], who claim that their algorithm is compatible with the human perception of multiple transparent surfaces [Pollard & Frisby 1990].

Feature-based matching, however, has trouble treating scenes containing pure transparency. In the case of physically pure transparency, appropriate features cannot be obtained by conventional feature detectors, such as edge and corner detectors, because they assume opacity of the scenes. Intensity-based matching is therefore more appropriate for dealing with pure transparency.

State-of-the-art computer vision does not have theories and tools that can handle complex environments including transparency. In fact, the transparency problem is more difficult than the segmentation problem that has been notorious in the history of computer vision. This paper therefore proposes a computational framework for direct single-shot estimation of multiple disparities from a binocular image pair. By using the assumption of additive superposition of image intensities, we derive direct linear constraints on the intensity variations along epipolar lines of the image pair. These constraints, derived under the assumption that n disparities exist at each image position, are approximated to first-order derivatives.

The proposed algorithm is *not* a special case of the previously reported multiple optical flow analysis [Shizawa & Mase 1990, Shizawa & Mase 1991a] because that analysis needs n-th order derivates for n-fold flow computation. Since the number of frames required for temporal differentiation is $n+1$, that analysis cannot be immediately applied to binocular stereo transparency. The algorithm proposed here can compute multiple disparities along an epipolar line from only *two* images. This algorithm can also be extended to multiple optical flow estimation from two frames.

Our previous paper [Shizawa 1992] proposed the constraint equation of binocular stereo transparency by applying the principle of superposition [Shizawa & Mase 1991b], but that paper neither derived an algorithm for solving the equation nor implemented it on computers. We instead discussed the relationships between the constraint of stereo transparency and the human psychophysics of ambiguous random-dot stereograms [Weinshall 1989], which are random-dot versions of the double-nail illusion [Krol & van de Grind 1980].

This paper uses the constraint equation proposed in the previous paper [Shizawa 1992] to derive a computational procedure for the single-shot algorithm by first-order approximations. A computer simulation is presented for one-dimensional(1D) signal matching.

The computational framework proposed here also concerns computational aspects of the physiological connections in the mammalian visual cortex and corpus callosum that take charge of binocular interactions between binocular retinal inputs [Hubel & Wiesel 1962, Hubel 1988].

The paper is organized as follows:

Section 2 describes an operator-based formal theory of stereo transparency based on the principle of superposition. (The use of the principle of superposition in deriving constraints for transparency was first proposed for the analysis of multiple motion [Shizawa & Mase 1991a, Shizawa & Mase 1991b].)

Section 3 derives the first-order approximation of the constraints by neglecting higher order terms. The constraints become linear in terms of image-intensity derivatives for the left and right images of the stereo pair. The coefficients of the constraint have the form of symmetrical polynomials of multiple disparities. The constraints are fitted to the image data by minimizing the square of the residual of the constraints with respect to their linear parameters within a local image patches. This derivation assumes that the local image structure is measured by the Gaussian-scaled derivatives similar to the image representation in the visual cortex. This minimization problem can be solved by a linear inverse operation, and the disparities can be obtained as n solutions of an n-degree univariate algebraic equation composed of the linear coefficients of the constraint. The overall algorithm has a quasi-linear single-shot nature.

Section 4 reports the results of preliminary simulations using computer-generated 1D signals. The algorithm successfully determines the multiple disparities at local windows.

Section 5 discusses other possible implementations of the theory proposed here. Constraints in the frequency domain are first described, and then the theory of stereo transparency is combined with the regularization theory.

Section 6 concludes this paper.

2 A Formal Theory of Stereo Transparency

In this chapter, the constraint of stereo matching is rewritten in terms of operator-based formalism. Then the principle of superposition (which we refer as PoS) and symmetrization are applied to derive the constraint equations of stereo transparency.

The intensity-based stereo matching constraint for an opaque surface can be written as

$$L(x,y) - R(x-D,y) = 0, \quad \text{or} \quad L(x+D,y) - R(x,y) = 0, \tag{1}$$

where $L(x,y)$ and $R(x,y)$ are intensity distributions at the left and right eyes, and D is a horizontal disparity. We use x for the coordinate along the epipolar line and y for the coordinate in the vertical direction. The constraint Eqs. (1) can be rewritten in terms of an *amplitude operator* $\mathbf{a}(D)$ and a *data distribution* $\mathbf{f}(x,y)$ as follows:

$$\mathbf{a}(D)\mathbf{f}(x,y) = \mathbf{0} \quad \text{where,} \quad \mathbf{a}(D) \equiv \begin{bmatrix} 1 & -\mathcal{D}(D) \\ -\mathcal{D}(-D) & 1 \end{bmatrix}, \quad \mathbf{f}(x,y) \equiv \begin{bmatrix} L(x,y) \\ R(x,y) \end{bmatrix}. \tag{2}$$

The term $\mathcal{D}(D)$ is a shift operator along the x-axis, and it transforms $f(x)$ into $f(x-D)$. Although the two equations (1) are equivalent each other, we formalize each of them to keep symmetrical operations in the operator formalism. Figure 1 illustrates the operation of the amplitude operator on binocular images.

The Taylor expansion of $f(x-D)$ around x is

$$f(x-D) = f(x) - D\partial_x f(x) + \frac{D^2}{2!}\partial_x^2 f(x) - \frac{D^3}{3!}\partial_x^3 f(x)\cdots = \sum_{i=0}^{\infty} \frac{(-D)^n}{n!}\partial_x^n f(x), \tag{3}$$

where $\partial_x^n f(x)$ denotes n-th order derivative of $f(x)$ with respect to x. The shift operator can therefore be written as the differential operator

$$\mathcal{D}(D) = \sum_{i=0}^{\infty} \frac{(-D)^i}{i!}\partial_x^i = \exp(-D\partial_x) = 1 - D\partial_x + \frac{D^2}{2!}\partial_x^2 - \frac{D^3}{3!}\partial_x^3 \cdots, \tag{4}$$

since Eq. (3) is equivalent to $f(x-D) = \mathcal{D}(D)f(x)$. Note that this operator is linear, i.e., $\mathcal{D}(D)(f_1(x) + f_2(x)) = \mathcal{D}(D)f_1(x) + \mathcal{D}(D)f_2(x)$ and $\mathcal{D}(D)0 = 0$. Furthermore, the shift operator forms a group with respect to its product, since

$$\mathcal{D}(D_\alpha)\mathcal{D}(D_\beta) \equiv \mathcal{D}(D_\beta)\mathcal{D}(D_\alpha) \equiv \mathcal{D}(D_\alpha + D_\beta). \tag{5}$$

The constraint (2) can be extended to the higher-order derivatives of the image intensity functions $L(x,y)$ and $R(x,y)$. We denote the (p,q)-th order partial derivatives of these by $L^{(p,q)}(x,y) = \partial_x^p \partial_y^q L(x,y)$ and $R^{(p,q)}(x,y) = \partial_x^p \partial_y^q R(x,y)$ and $\mathbf{f}^{(p,q)}(x,y) = [L^{(p,q)}(x,y), R^{(p,q)}(x,y)]^T$. Then the constraint can be extended to

$$\mathbf{a}(D)\mathbf{f}^{(p,q)}(x,y) = \mathbf{0}, \tag{6}$$

for $p, q = 0, 1, 2, \cdots$. In general, the whole set of (p, q)-th order derivatives at an image point can be considered to completely characterize the local image structure around that point.

Using the linearity of the shift operator, we can see that the amplitude operator $\mathbf{a}(D)$ is also a linear operator, i.e., $\mathbf{a}(D)\mathbf{0} = \mathbf{0}$ and $\mathbf{a}(D)\{\mathbf{f}_1(x, y) + \mathbf{f}_2(x, y)\} = \mathbf{a}(D)\mathbf{f}_1(x, y) + \mathbf{a}(D)\mathbf{f}_2(x, y)$.

The principle of superposition can be applied using above formalism, and we have a constraint of stereo transparency derived in an entirely formal way:

$$\mathbf{a}(D_n) \cdots \mathbf{a}(D_2)\mathbf{a}(D_1)\mathbf{f}^{(p,q)}(x, y) = \mathbf{0}, \tag{7}$$

where $\mathbf{f}^{(p,q)}(x, y) = \sum_{i=1}^{n} \mathbf{f}_i^{(p,q)}(x, y)$, and each $\mathbf{f}_i^{(p,q)}(x, y)$ is constrained by $\mathbf{a}(D_i)\mathbf{f}_i^{(p,q)}(x, y) = \mathbf{0}$. Unfortunately, however, the amplitude operator do not commute. In fact, we have

$$\mathbf{a}(D_1)\mathbf{a}(D_2) \equiv \begin{bmatrix} 1 + \mathcal{D}(-D_1)\mathcal{D}(D_2) & -\mathcal{D}(D_1) - \mathcal{D}(D_2) \\ -\mathcal{D}(-D_1) - \mathcal{D}(-D_2) & 1 + \mathcal{D}(D_1)\mathcal{D}(-D_2) \end{bmatrix} \tag{8}$$

and

$$\mathbf{a}(D_2)\mathbf{a}(D_1) \equiv \begin{bmatrix} 1 + \mathcal{D}(-D_2)\mathcal{D}(D_1) & -\mathcal{D}(D_2) - \mathcal{D}(D_1) \\ -\mathcal{D}(-D_2) - \mathcal{D}(-D_1) & 1 + \mathcal{D}(D_2)\mathcal{D}(-D_1) \end{bmatrix}. \tag{9}$$

Then

$$\mathbf{a}(D_2)\mathbf{a}(D_1) - \mathbf{a}(D_1)\mathbf{a}(D_2) = \begin{bmatrix} -\mathcal{D}(D_1 - D_2) + \mathcal{D}(D_2 - D_1) & 0 \\ 0 & \mathcal{D}(D_1 - D_2) - \mathcal{D}(D_2 - D_1) \end{bmatrix}, \tag{10}$$

and

$$\mathcal{D}(D_1 - D_2) - \mathcal{D}(D_2 - D_1) = 2\{-(D_1 - D_2)\partial_x - \frac{(D_1 - D_2)^3}{3!}\partial_x^3 - \frac{(D_1 - D_2)^5}{5!}\partial_x^5 - \cdots\}. \tag{11}$$

The order of the product of the amplitude operators therefore changes the meaning of the constraint, and the lefthand side of Eq. (7) cannot be zero for general image signals.

This problem can be remedied by the *symmetrization* of the amplitude operator of transparency with respect to the permutation group of the order of disparities $D_1, D_2, \cdots, D_n$. Then the constraint is valid in the sense of energy minimization.

We define the symmetrized amplitude operator for n-fold transparency by

$$\bar{\mathrm{a}}^{(n)}(D_{(1}, D_2, \cdots, D_{n)}) \equiv \frac{1}{n!} \sum_{\sigma \in P_n} \mathrm{a}(D_{\sigma(n)}) \cdots \mathrm{a}(D_{\sigma(2)})\mathrm{a}(D_{\sigma(1)}), \tag{12}$$

where P_n represents the permutation group of order n. When n equals 2 and 3, for example, we have the following expressions of the amplitude operators:

$$\bar{\mathrm{a}}^{(2)}(D_{(1}, D_{2)}) \equiv \frac{1}{2}\{\mathrm{a}(D_1)\mathrm{a}(D_2) + \mathrm{a}(D_2)\mathrm{a}(D_1)\} \tag{13}$$

and

$$\bar{\mathrm{a}}^{(3)}(D_{(1}, D_2, D_{3)}) \equiv \frac{1}{6} \{\mathrm{a}(D_1)\mathrm{a}(D_2)\mathrm{a}(D_3) + \mathrm{a}(D_2)\mathrm{a}(D_1)\mathrm{a}(D_3) + \mathrm{a}(D_3)\mathrm{a}(D_1)\mathrm{a}(D_2) + \mathrm{a}(D_3)\mathrm{a}(D_2)\mathrm{a}(D_1) + \mathrm{a}(D_2)\mathrm{a}(D_3)\mathrm{a}(D_1) + \mathrm{a}(D_1)\mathrm{a}(D_3)\mathrm{a}(D_2)\}. \quad (14)$$

3 Derivation of a Single-shot Algorithm

The constraints derived in Sec.2 are not easily used in computer implementations because they are highly nonlinear with respect to disparities $D_1, D_2, \cdots, D_n$. This section therefore derive a linear single-shot algorithm for binocular stereo transparency by using the first-order approximations of the amplitude operator $\mathbf{a}(D)$. We derive the constraints and algorithms only for the cases in which $n = 1$ or $n = 2$. More general descriptions, including the cases in which $n = 3$ or $n = 4$, are provided in Appendix A.

3.1 First-order Approximation of the Constraints

3.1.1 Algorithm for single disparity

We use the first-order approximation for the following shift operator:

$$\mathcal{D}(D) \approx 1 - D\partial_x. \tag{15}$$

Then the first-order approximation of the amplitude operator of the single disparity D becomes

$$\mathbf{a}(D) \approx \begin{bmatrix} 1 & -1 + D\partial_x \\ -1 - D\partial_x & 1 \end{bmatrix}. \tag{16}$$

Then the generalized constraint $\mathbf{a}(D)\mathbf{f}^{(p,q)}(x,y) = 0$ is approximated by the following equations:

$$\begin{aligned} r_L^{(1)}[p,q](x,y) &= \{L^{(p,q)}(x,y) - R^{(p,q)}(x,y)\} - DR^{(p+1,q)}(x,y) = 0 \\ r_R^{(1)}[p,q](x,y) &= \{R^{(p,q)}(x,y) - L^{(p,q)}(x,y)\} + DL^{(p+1,q)}(x,y) = 0. \end{aligned} \tag{17}$$

When p and q are both 0, this is equivalent to the equation used in the iterative signal matching for stereo vision [Lucas & Kanade 1981, Okutomi & Kanade 1992]. (In those papers, one of the two constraints is used to estimate disparity update values. In this paper, however, we use both of the constraints simultaneously for the sake of symmetry between the left and right eyes.) Now we have the following energy function for estimating disparity D around (x_0, y_0):

$$E(D(x_0,y_0)) = \int\int \{w[p,q](x-x_0, y-y_0)\}^2[\{r_L^{(1)}[p,q](x,y)\}^2 + \{r_R^{(1)}[p,q](x,y)\}^2]dxdy, \tag{18}$$

where $w[p,q](x,y)$ denotes positive weight functions that can be determined from the statistical properties of the image derivatives and the window function for integrating energy distribution. This function is integrated over the entire image. Different order constraints should be weighted differently, since each image intensity derivative has a different noise distribution. But that subject is not discussed in this paper. Solving $\frac{\partial E(D)}{\partial D} = 0$ for D, we have the following estimator for D at (x_0, y_0):

$$D(x_0, y_0) = \frac{b(x_0, y_0)}{a(x_0, y_0)}, \tag{19}$$

where

$$a(x_0, y_0) = \sum_{p,q} \int\int \{w[p,q](x - x_0, y - y_0)\}^2 \{(R^{(p+1,q)}(x,y))^2 + (L^{(p+1,q)}(x,y))^2\} dx dy \tag{20}$$

and

$$\begin{aligned} b(x_0, y_0) &= \sum_{p,q} \int\int \{w[p,q](x - x_0, y - y_0)\}^2 \\ &\times \{L^{(p,q)}(x,y) - R^{(p,q)}(x,y)\}\{R^{(p+1,q)}(x,y) + L^{(p+1,q)}(x,y)\} dx dy. \end{aligned} \tag{21}$$

3.1.2 An Algorithm for Two-fold Surfaces

It can be shown that the symmetrized amplitude operator for two-fold transparency is approximated into the following form:

$$\bar{\mathrm{a}}^{(2)}(D_{(1}, D_{2)}) \approx \begin{bmatrix} 2 - D_1 D_2 \partial_x^2 & -2 - (D_1 + D_2)\partial_x \\ -2 + (D_1 + D_2)\partial_x & 2 - D_1 D_2 \partial_x^2 \end{bmatrix}. \tag{22}$$

The residual functions therefore become

$$\begin{aligned} r_L^{(2)}[p,q](x,y) &= 2\{L^{(p,q)}(x,y) - R^{(p,q)}(x,y)\} - (D_1 + D_2) R^{(p+1,q)}(x,y) \\ &\quad - D_1 D_2 L^{(p+2,q)}(x,y) \end{aligned} \tag{23}$$

and

$$\begin{aligned} r_R^{(2)}[p,q](x,y) &= 2\{R^{(p,q)}(x,y) - L^{(p,q)}(x,y)\} + (D_1 + D_2) L^{(p+1,q)}(x,y) \\ &\quad - D_1 D_2 R^{(p+2,q)}(x,y). \end{aligned} \tag{24}$$

The energy function for estimating disparities D_1 and D_2 is

$$E^{(2)}(D_1(x_0, y_0), D_2(x_0, y_0)) = \int\int \{w_{p,q}(x - x_0, y - y_0)\}^2 [\{r_L^{p,q}(x,y)\}^2 + \{r_R^{p,q}(x,y)\}^2] dx dy. \tag{25}$$

This energy function is quadratic with respect to $s_1 = \frac{1}{2}(D_1 + D_2)$ and $s_2 = D_1 D_2$. By considering the condition for minimum energy $\frac{\partial E^{(2)}}{\partial s_1} = 0$ and $\frac{\partial E^{(2)}}{\partial s_2} = 0$, we have the following system of linear equations:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \tag{26}$$

where

$$
\begin{aligned}
a_{11} &= 4\sum_{p,q}\int\int\{w_{p,q}(x-x_0,y-y_0)\}^2[\{R^{(p+1,q)}(x,y)\}^2+\{L^{(p+1,q)}(x,y)\}^2]dxdy, \\
a_{12} &= 2\sum_{p,q}\int\int\{w_{p,q}(x-x_0,y-y_0)\}^2 \\
&\times\{L^{(p+2,q)}(x,y)R^{(p+1,q)}(x,y)-L^{(p+1,q)}(x,y)R^{(p+2,q)}(x,y)\}dxdy, \qquad (27)
\end{aligned}
$$

$$
a_{22}=\sum_{p,q}\int\int\{w_{p,q}(x-x_0,y-y_0)\}^2[\{L^{(p+2,q)}(x,y)\}^2+\{R^{(p+2,q)}(x,y)\}^2]dxdy, \qquad (28)
$$

and where

$$
\begin{aligned}
b_1 &= 4\sum_{p,q}\int\int\{w_{p,q}(x-x_0,y-y_0)\}^2 \\
&\times\{L^{(p+1,q)}(x,y)+R^{(p+1,q)}(x,y)\}\{L^{(p,q)}(x,y)-R^{(p,q)}(x,y)\}dxdy \qquad (29)
\end{aligned}
$$

and

$$
\begin{aligned}
b_2 &= 2\sum_{p,q}\int\int\{w_{p,q}(x-x_0,y-y_0)\}^2 \\
&\times\{L^{(p+2,q)}(x,y)-R^{(p+2,q)}(x,y)\}\{L^{(p,q)}(x,y)-R^{(p,q)}(x,y)\}dxdy. \qquad (30)
\end{aligned}
$$

We can solve this system of equations for parameters s_1 and s_2 as

$$
\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \qquad (31)
$$

Then the two disparities D_1 and D_2 ($D_1 \geq D_2$) can be obtained as solutions of the following quadratic equation of d:

$$
d^2 - 2s_1 d + s_2 = 0. \qquad (32)
$$

We can write them as

$$
D_1 = s_1 + \sqrt{s_1^2 - s_2} \qquad (33)
$$

and

$$
D_2 = s_1 - \sqrt{s_1^2 - s_2}. \qquad (34)
$$

For a single disparity, we have the condition,

$$
\Delta^{(2)} = s_1^2 - s_2 = 0, \qquad (35)
$$

where $\Delta^{(2)}$ denotes the discriminant of the algebraic equation (32).

The determination of the number of disparities is an important issue. We can use the condition (35) for the test of a single disparity in the two-fold disparity estimator. Furthermore, disparities are constrained to be real numbers. This is equivalent to the condition $\Delta^{(2)} \geq 0$. When $n > 2$, the constraints for multiple-root disparities and those for real solutions can be found in the basic mathematics of algebraic equations, such as the discriminant of algebraic equations, Descartes' law of signs, and Strum's theorem.

3.2 Computation of Local Image Structure by using Regularized Derivatives

We use the local Gaussian-scaled derivatives [Koenderink & van Doorn 1992, Florack et al. 1992] as local image representations for the stereo matching. It should be noted that the image representation itself is independent of the stereo transparency constraints. We use this representation because of its direct relationship to the information necessary for the constraint equations. Other image representations, such as Gabor elementary functions, are also possible. In this respect, biological implications for the physiology of image representation in the visual cortex and its binocular interactions [Hubel & Wiesel 1962, Hubel 1988] are also involved in this formalism.

The Gaussian-scaled regularized derivatives can be expressed as convolutions of images with derivatives of a Gaussian as follows.

$$\begin{aligned} L_\sigma^{(p,q)}(x,y) &= \int_{-\infty}^{+\infty} G_\sigma(x-u,y-v)L^{(p,q)}(u,v)dudv \\ &= \int_{-\infty}^{+\infty} G_\sigma^{(p,q)}(x-u,y-v)L(u,v)dudv, \end{aligned} \tag{36}$$

and

$$\begin{aligned} R_\sigma^{(p,q)}(x,y) &= \int_{-\infty}^{+\infty} G_\sigma(x-u,y-v)R^{(p,q)}(u,v)dudv \\ &= \int_{-\infty}^{+\infty} G_\sigma^{(p,q)}(x-u,y-v)R(u,v)dudv, \end{aligned} \tag{37}$$

where the two-dimensional Gaussian $G_\sigma(x,y)$ is defined as

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2}e^{-\frac{x^2}{2\sigma^2}} \cdot \frac{1}{2\pi\sigma^2}e^{-\frac{y^2}{2\sigma^2}} = \frac{1}{4\pi^2\sigma^4}e^{-\frac{(x^2+y^2)}{2\sigma^2}}. \tag{38}$$

We denote $\mathbf{f}_\sigma^{(p,q)}(x,y) = [L_\sigma^{(p,q)}(x,y), R_\sigma^{(p,q)}(x,y)]^T$.

Figure 2 shows a schematic diagram of the derived algorithm. (In the figure, to build the symmetry between the left and right eyes, the regularized derivatives are shown in mirror-reflected positions. Correspondingly, the connections in the right eye part are slightly different from the derived equations.)

4 Simulation Using One-dimensional Signals

Since the binocular stereo matching is reduced to 1D matching along epipolar lines, simulation of this matching is enough to test the validity of the algorithm. The y coordinate is therefore omitted from the following notations.

4.1 Experimental Condition and Results

Two original binocular pairs of signals were generated by using a random number generator and low pass filtering. Then the signals were shifted so that the two pairs had two different disparity maps for two surfaces (Fig. 3). Then the pairs of signals were additively superposed and subsampled to produce a pair of input signals for the left and right eyes. Figure 4 shows the signals generated for the following simulation.

To compute regularized derivatives, the input signals are convolved with the derivatives of Gaussian(Fig. 5). We used up to fifth-order derivatives as follows:

$$\begin{aligned}
G_\sigma^{(0,0)}(x,0) &= \frac{1}{2\pi\sigma^2}e^{-\frac{x^2}{2\sigma^2}} \\
G_\sigma^{(1,0)}(x,0) &= \frac{-x}{2\pi\sigma^4}e^{-\frac{x^2}{2\sigma^2}} \\
G_\sigma^{(2,0)}(x,0) &= \frac{x^2-\sigma^2}{2\pi\sigma^6}e^{-\frac{x^2}{2\sigma^2}} \\
G_\sigma^{(3,0)}(x,0) &= \frac{-x^3+3\sigma^2 x}{2\pi\sigma^8}e^{-\frac{x^2}{2\sigma^2}} \\
G_\sigma^{(4,0)}(x,0) &= \frac{x^4-6\sigma^2x^2+3\sigma^4}{2\pi\sigma^{10}}e^{-\frac{x^2}{2\sigma^2}} \\
G_\sigma^{(5,0)}(x,0) &= \frac{-x^5+10\sigma^2x^3-15\sigma^4}{2\pi\sigma^{12}}e^{-\frac{x^2}{2\sigma^2}} \qquad (39)
\end{aligned}$$

The Gaussian scale parameter σ was fixed to $4[pixel]$ in the simulation.

Figure 6 shows some of the results of the convolutions.

The weight function $w_{p,q}(x)$ for integration of the energy density is taken to be the square window function

$$w_{p,q}(x) = \begin{cases} 0, & (|x| > M) \\ 1, & (|x| \le M) \end{cases}, \qquad (40)$$

where M is the width of the window.

In this simulation, M was set at $121[pixel]$, and the estimation was done for $p = 0, 1, 2$ and 3 Figure 7 shows the final outputs of the two-fold multiple disparity estimator derived in Sec. 3. The true disparities and the disparity estimates are superimposed in the plots.

It should be emphasized that these results are obtained from the first-order approximation of the stereo transparency constraint. Taking this into account, the results is very promising. These estimates can be good initial guesses for the later stages of iterative refinement and complete reconstruction of the surfaces. Because the determination of the number of disparities is an important issue for applications, in future implementations, we will use the condition (35) to test for a single disparity in the two-fold disparity estimator.

5 Alternative Algorithms

5.1 Constraints in the Frequency Domain

The constraints can be transformed into the frequency domain representation by applying Fourier transformation. The residual functions become

$$\begin{aligned}\tilde{r}_L^{(2)}(\omega_x, \omega_y) &= 2\{\tilde{L}(\omega_x, \omega_y) - \tilde{R}(\omega_x, \omega_y)\} - (D_1 + D_2)(2\pi i)\omega_x \tilde{R}(\omega_x, \omega_y) \\ &\quad -D_1 D_2 (2\pi i)^2 \omega_x^2 \tilde{L}(\omega_x, \omega_y) \end{aligned} \tag{41}$$

and

$$\begin{aligned}\tilde{r}_R^{(2)}(\omega_x, \omega_y) &= 2\{\tilde{R}(\omega_x, \omega_y) - \tilde{L}(\omega_x, \omega_y)\} + (D_1 + D_2)(2\pi i)\omega_x \tilde{L}(\omega_x, \omega_y) \\ &\quad -D_1 D_2 (2\pi i)^2 \omega_x^2 \tilde{R}(\omega_x, \omega_y). \end{aligned} \tag{42}$$

These frequency domain constraints may be useful in building an algorithm using Gabor elementary functions for the image representation.

5.2 Regularization

The algorithm presented in Sec. 3 is based on the sum of square difference, so, the disparity in a local window is assumed to be constant. To solve our ill-posed problem, regularization theory [Poggio et al. 1985] can be applied to the framework proposed here. In this case, it is more convenient to impose smoothness on the linear parameters $s_i^{(n)}$ instead of the disparity parameters D_i. Then the energy functional becomes quadratic with respect to the linear parameters when approximated to first-order derivatives:

$$\mathcal{E}(\mathrm{s}^{(n)}) = \int\int [\{\sum_{p,q} \|G_\sigma \otimes \bar{\mathrm{a}}^{(n)}(\mathrm{s}^{(n)})\mathrm{f}^{(p,q)}(x,y)\|^2\} + \{S\mathrm{s}^{(n)}\}^2] dxdy, \tag{43}$$

where $\mathrm{s}^{(n)} = (s_1^{(n)}, s_2^{(n)}, \cdots, s_n^{(n)})$. The smoothing operator S should be chosen so that the original disparity parameters are regularized to an appropriate degree.

6 Conclusion

We have presented a single-shot algorithm for binocular stereo transparency. The derivation of this algorithm was based on the operator-based theory of transparency, the principle of superposition. The constraint in this paper was approximated by the first-order approximation of the amplitude operator. A quasi-linear single-shot algorithm was derived by minimizing the squared error of the approximated constraint. Simulation using a computer-generated 1D signal showed that good initial guesses can be computed by using the single-shot algorithm.

The algorithm presented here will be extended to an iterative method that can compute more precise disparity estimates in a way similar to the iterative image registration techniques [Lucas & Kanade 1981]. Appropriate control of the scale parameter will be a good strategy for large disparities. We may also be able to incorporate a statistical model for selecting the adaptive window [Okutomi & Kanade 1992]. Another possible implementations uses regularization theory. The occlusion analysis at depth boundaries[Little & Gillett 1990, Geiger et al. 1992] must be unified into the proposed framework. These environmental issues are under investigation.

It is possible to extend the algorithm into multiple optical flow estimation from only *two* frames. This extension will overcome some deficiencies of the three-frame algorithm for two motions [Bergen et al. 1990][Shizawa & Mase 1991b], such as weakness in the ability to handle acceleration. Further, since the presented algorithm can be viewed as a general signal matching algorithm, another application area may be the locations of multiple sound sources from binaural information. Our constraint-based theory of transparency can be regarded as a part of the computational theory of biological visual systems, since the computer-based system should use the physics of image formation that are common to animals and machines that must properly infer the real physical world in order to survive and to serve.

Appendix

A First Order Constraints for the Case of $n \leq 4$

A.1 Elementary Symmetric Polynomial Representations for Multiple Coexistent Disparities.

In this appendix, the argument of y is omitted. The coexistence of multiple disparities at each retinal position can be conveniently represented by the symmetric polynomials of D_i, which are defined as

$$
\begin{aligned}
s_1^{(1)} &= D_1^{(1)}, \\
s_1^{(2)} &= \frac{1}{2}(D_1^{(2)} + D_2^{(2)}), \\
s_2^{(2)} &= D_1^{(2)} D_2^{(2)}, \\
s_1^{(3)} &= \frac{1}{3}(D_1^{(3)} + D_2^{(3)} + D_3^{(3)}), \\
s_2^{(3)} &= \frac{1}{3}(D_1^{(3)} D_2^{(3)} + D_2^{(3)} D_3^{(3)} + D_3^{(3)} D_1^{(3)}), \\
s_3^{(3)} &= D_1^{(3)} D_2^{(3)} D_3^{(3)}, \\
s_1^{(4)} &= \frac{1}{4}(D_1^{(4)} + D_2^{(4)} + D_3^{(4)} + D_4^{(4)}), \\
s_2^{(4)} &= \frac{1}{6}(D_1^{(4)} D_2^{(4)} + D_2^{(4)} D_3^{(4)} + D_3^{(4)} D_4^{(4)} + D_4^{(4)} D_1^{(4)} + D_1^{(4)} D_3^{(4)} + D_2^{(4)} D_4^{(4)}), \\
s_3^{(4)} &= \frac{1}{4}(D_1^{(4)} D_2^{(4)} D_3^{(4)} + D_2^{(4)} D_3^{(4)} D_4^{(4)} + D_1^{(4)} D_3^{(4)} D_4^{(4)} + D_1^{(4)} D_2^{(4)} D_4^{(4)}), \\
s_4^{(4)} &= D_1^{(4)} D_2^{(4)} D_3^{(4)} D_4^{(4)}.
\end{aligned}
\tag{44}
$$

Multiple disparities $D_i^{(n)}$ for $n = 2, 3$, and 4 can be obtained as solutions of the following n-th degree univariate algebraic equations:

$$
\begin{aligned}
& d^2 - 2s_{2,1}d + s_{2,2} = 0, \\
& d^3 - 3s_{3,1}d^2 + 3s_{3,2}d - s_{3,3} = 0, \\
\text{and} \quad & d^4 - 4s_{4,1}d^3 + 6s_{4,2}d^2 - 4s_{4,3}d + s_{4,4} = 0.
\end{aligned}
\tag{45}
$$

A.2 Constraint Equations for the cases $n = 2, 3$, and 4

Using these representations, the residue functions $\mathrm{r}^{(n)}(x,y) = (r_L^{(n)}(x,y), r_R^{(n)}(x,y))$ $= \mathbf{a}^{(n)}(D_1, \cdots, D_n)\mathrm{f}(x,y)$ of the transparent stereo constraint equations can be written simply as couples of linear equations. The explicit forms for $n = 1, 2, 3$, and 4 are

$$
\begin{aligned}
r_L^{(1)}[p,q](x,y) &= \{L^{(p,q)} - R^{(p,q)}\} - R^{(p+1,q)} s_{1,1} \\
r_R^{(1)}[p,q](x,y) &= \{R^{(p,q)} - L^{(p,q)}\} + L^{(p+1,q)} s_{1,1}
\end{aligned}
\tag{46}
$$

$$
r_L^{(2)}[p,q](x,y) = 2\{L^{(p,q)} - R^{(p,q)}\} - 2R^{(p+1,q)} s_{2,1} - L^{(p+2,q)} s_{2,2}
$$

$$r_R^{(2)}[p,q](x,y) = 2\{R^{(p,q)} - L^{(p,q)}\} + 2L^{(p+1,q)}s_{2,1} - R^{(p+2,q)}s_{2,2} \quad (47)$$

$$\begin{aligned} r_L^{(3)}[p,q](x,y) &= 4\{L^{(p,q)} - R^{(p,q)}\} - 4R^{(p+1,q)}s_{3,1} - \{3L^{(p+2,q)} - R^{(p+2,q)}\}s_{3,2} \\ &\quad + R^{(p+3,q)}s_{3,3} \\ r_R^{(3)}[p,q](x,y) &= 4\{R^{(p,q)} - L^{(p,q)}\} + 4L^{(p+1,q)}s_{3,1} - \{3R^{(p+2,q)} - L^{(p+2,q)}\}s_{3,2} \\ &\quad - L^{(p+3,q)}s_{3,3} \end{aligned} \quad (48)$$

$$\begin{aligned} r_L^{(4)}[p,q](x,y) &= 8\{L^{(p,q)} - R^{(p,q)}\} - 8R^{(p+1,q)}s_{4,1} - 4\{2L^{(p+2,q)} - R^{(p+2,q)}\}s_{4,2} \\ &\quad + 4R^{(p+3,q)}s_{4,3} + L^{(p+4,q)}s_{4,4} \\ r_R^{(4)}[p,q](x,y) &= 8\{R^{(p,q)} - L^{(p,q)}\} + 8L^{(p+1,q)}s_{4,1} - 4\{2R^{(p+2,q)} - L^{(p+2,q)}\}s_{4,2} \\ &\quad - 4L^{(p+3,q)}s_{4,3} + R^{(p+4,q)}s_{4,4} \end{aligned} \quad (49)$$

A.3 Estimation Algorithms for the cases $n = 3$ and 4

The residual energy at (x_0, y_0) is defined as

$$E^{(n)}(x_0,y_0) = \sum_{p,q} \int\int \{w_{p,q}(x-x_0, y-y_0)\}^2 \{(r_L^{(n)}(x,y))^2 + (r_R^{(n)}(x,y))^2\} dx dy. \quad (50)$$

When $n = 3$, we have a simultaneous linear equation for $s_i^{(3)}$ as follows. We use the shorter notations $L^{(p,q)} = L^{(p,q)}(x,y)$, $R^{(p,q)} = R^{(p,q)}(x,y)$, and $w_{p,q} = w_{p,q}(x-x_0, y-y_0)$ in the following:

$$\begin{bmatrix} a_{11}^{(3)} & a_{12}^{(3)} & a_{13}^{(3)} \\ a_{12}^{(3)} & a_{22}^{(3)} & a_{23}^{(3)} \\ a_{13}^{(3)} & a_{23}^{(3)} & a_{33}^{(3)} \end{bmatrix} \begin{bmatrix} s_1^{(3)} \\ s_2^{(3)} \\ s_3^{(3)} \end{bmatrix} = \begin{bmatrix} b_1^{(3)} \\ b_2^{(3)} \\ b_3^{(3)} \end{bmatrix}, \quad (51)$$

where

$$\begin{aligned} a_{11}^{(3)} &= 16\sum_{p,q}\int\int (w_{p,q})^2[\{R^{(p+1,q)}\}^2 + \{L^{(p+1,q)}\}^2]dxdy \\ a_{12}^{(3)} &= 4\sum_{p,q}\int\int (w_{p,q})^2[\{L^{(p+1,q)}L^{(p+2,q)} - R^{(p+1,q)}R^{(p+2,q)}\} \\ &\quad +3\{R^{(p+1,q)}L^{(p+2,q)} - L^{(p+1,q)}R^{(p+2,q)}\}]dxdy \\ a_{13}^{(3)} &= -4\sum_{p,q}\int\int (w_{p,q})^2\{L^{(p+1,q)}L^{(p+3,q)} + R^{(p+1,q)}R^{p+3,q}\}dxdy \\ a_{22}^{(3)} &= 2\sum_{p,q}\int\int (w_{p,q})^2[5\{L^{(p+2,q)}\}^2 - 6L^{(p+2,q)}R^{(p+2,q)} + 5\{R^{(p+2,q)}\}^2]dxdy \\ a_{23}^{(3)} &= \sum_{p,q}\int\int (w_{p,q})^2[3\{R^{(p+3,q)}L^{(p+2,q)} + L^{(p+3,q)}R^{(p+2,q)}\} \\ &\quad -\{R^{(p+3,q)}R^{(p+2,q)} + L^{(p+3,q)}R^{(p+2,q)}\}]dxdy \\ a_{33}^{(3)} &= \sum_{p,q}\int\int (w_{p,q})^2[\{R^{(p+3,q)}\}^2 + \{L^{(p+3,q)}\}^2]dxdy \end{aligned} \quad (52)$$

and

$$
\begin{aligned}
b_1^{(3)} &= 16\sum_{p,q}\int\int(w_{p,q})^2\{L^{(p+1,q)}+R^{(p+1,q)}\}\{L^{(p,q)}-R^{(p,q)}\}dxdy \\
b_2^{(3)} &= 16\sum_{p,q}\int\int(w_{p,q})^2\{L^{(p+2,q)}-R^{(p+2,q)}\}\{L^{(p,q)}-R^{(p,q)}\}dxdy \\
b_3^{(3)} &= -4\sum_{p,q}\int\int(w_{p,q})^2\{L^{(p+3,q)}+R^{(p+3,q)}\}\{L^{(p,q)}-R^{(p,q)}\}dxdy \qquad (53)
\end{aligned}
$$

When $n=4$, the simultaneous linear equations becomes

$$
\begin{bmatrix} a_{11}^{(4)} & a_{12}^{(4)} & a_{13}^{(4)} & a_{14}^{(4)} \\ a_{12}^{(4)} & a_{22}^{(4)} & a_{23}^{(4)} & a_{24}^{(4)} \\ a_{13}^{(4)} & a_{23}^{(4)} & a_{33}^{(4)} & a_{34}^{(4)} \\ a_{14}^{(4)} & a_{24}^{(4)} & a_{34}^{(4)} & a_{44}^{(4)} \end{bmatrix}
\begin{bmatrix} s_1^{(4)} \\ s_2^{(4)} \\ s_3^{(4)} \\ s_4^{(4)} \end{bmatrix}
=
\begin{bmatrix} b_1^{(4)} \\ b_2^{(4)} \\ b_3^{(4)} \\ b_4^{(4)} \end{bmatrix}, \qquad (54)
$$

where

$$
\begin{aligned}
a_{11}^{(4)} &= 64\sum_{p,q}\int\int(w_{p,q})^2[\{R^{(p+1,q)}\}^2+\{L^{(p+1,q)}\}^2]dxdy \\
a_{12}^{(4)} &= 32\sum_{p,q}\int\int(w_{p,q})^2[2\{R^{(p+1,q)}L^{(p+2,q)}-L^{(p+1,q)}R^{(p+2,q)}\} \\
&\quad +\{L^{(p+1,q)}L^{(p+2,q)}-R^{(p+1,q)}R^{(p+2,q)}\}]dxdy \\
a_{13}^{(4)} &= -32\sum_{p,q}\int\int(w_{p,q})^2\{L^{(p+1,q)}L^{(p+3,q)}+R^{(p+1,q)}R^{(p+3,q)}\}dxdy \\
a_{14}^{(4)} &= 8\sum_{p,q}\int\int(w_{p,q})^2\{L^{(p+1,q)}R^{(p+4,q)}-R^{(p+1,q)}L^{(p+4,q)}\}dxdy \\
a_{22}^{(4)} &= 16\sum_{p,q}\int\int(w_{p,q})^2[5\{L^{(p+2,q)}\}^2-8L^{(p+2,q)}R^{(p+2,q)}+5\{R^{(p+2,q)}\}^2]dxdy \\
a_{23}^{(4)} &= 16\sum_{p,q}\int\int(w_{p,q})^2[2\{L^{(p+3,q)}R^{(p+2,q)}-R^{(p+3,q)}L^{(p+2,q)}\} \\
&\quad +\{R^{(p+2,q)}R^{(p+3,q)}-L^{(p+2,q)}L^{(p+3,q)}\}]dxdy \\
a_{24}^{(4)} &= -4\sum_{p,q}\int\int(w_{p,q})^2[2\{L^{(p+2,q)}L^{(p+4,q)}+R^{(p+2,q)}R^{(p+4,q)}\} \\
&\quad -\{L^{(p+2,q)}R^{(p+4,q)}+R^{(p+2,q)}L^{(p+4,q)}\}]dxdy \\
a_{33}^{(4)} &= 16\sum_{p,q}\int\int(w_{p,q})^2[\{R^{(p+3,q)}\}^2+\{L^{(p+3,q)}\}^2]dxdy \\
a_{34}^{(4)} &= 4\sum_{p,q}\int\int(w_{p,q})^2\{R^{(p+3,q)}L^{(p+4,q)}-L^{(p+3,q)}R^{(p+4,q)}\}dxdy \\
a_{44}^{(4)} &= \sum_{p,q}\int\int(w_{p,q})^2[\{R^{(p+4,q)}\}^2+\{L^{(p+4,q)}\}^2]dxdy \qquad (55)
\end{aligned}
$$

and

$$
\begin{aligned}
b_1^{(4)} &= 64\sum_{p,q}\int\int(w_{p,q})^2\{L^{(p+1,q)}+R^{(p+1,q)}\}\{L^{(p,q)}-R^{(p,q)}\}dxdy \\
b_2^{(4)} &= 96\sum_{p,q}\int\int(w_{p,q})^2\{L^{(p+2,q)}-R^{(p+2,q)}\}\{L^{(p,q)}-R^{(p,q)}\}dxdy
\end{aligned}
$$

$$
\begin{aligned}
b_3^{(4)} &= -32 \sum_{p,q} \int \int (w_{p,q})^2 \{L^{(p+3,q)} + R^{(p+3,q)}\}\{L^{(p,q)} - R^{(p,q)}\} dx dy \\
b_4^{(4)} &= -8 \sum_{p,q} \int \int (w_{p,q})^2 \{L^{(p+3,q)} - R^{(p+3,q)}\}\{L^{(p,q)} - R^{(p,q)}\} dx dy
\end{aligned} \tag{56}
$$

References

[Adelson & Bergen 1991] Adelson, E.H. and Bergen, J.R. 1991. The plenopic function and the elements of early vision. In *Computational Models of Visual Processing* (Landy M.S. and Movshon J.A. eds.), MIT Press, Cambridge, MA.

[Akerstrom & Todd 1988] Akerstrom, R.A. and Todd, J.T. 1988. The perception of stereoscopic transparency. *Percept. & Psychophys.* 44(5): 421-432.

[Baker & Binford 1981] Baker, H.H. and Binford, T.O. 1981. Depth from edge and intensity based stereo. Proc. 7th Int. Joint Conf. Artif. Intell. Los Altos. CA, pp.631-636.

[Bergen et al. 1990] Bergen, J.R., Burt, P., Hingorani, R., and Peleg, S. 1990. Computing two motions from three frames. In *Proc. IEEE 3rd Int. Conf. on Comput. Vision,* Osaka, Japan, pp.27-32.

[Blake & Zisserman 1987] Blake, A. and Zisserman, A. 1987. *Visual Reconstruction,* MIT Press, Cambridge, MA.

[Burt & Julesz 1980a] Burt, P.J. and Julesz, B. 1980. A disparity gradient limit for binocular fusion. *Science,* 208: 615-617.

[Burt & Julesz 1980b] Burt, P.J. and Julesz, B. 1980. Modifications of the classical notion of Panum's fusional area. *Percept.* 9: 671-682.

[Dhond & Aggarwal 1989] Dhond, U.R. and Aggarwal, J.K. 1989. Structure from stereo — a review. *IEEE Trans. Sys. Man, & Cybern.* 19(6): 1489-1510.

[Dhond & Aggarwal 1992] Dhond, U.R. and Aggarwal, J.K. 1992. Computing stereo correspondence in the presence of narrow occluding objects. In *Proc. IEEE Conf. on Comput. Vision and Pattern Recognition,* Champaign, IL, pp.758-760.

[Florack et al. 1992] Florack, L.M.J., Romeny, B.M.H., Koenderink, J.J., and Viergever, M.A. 1992. Scale and the differential structure of images. *Image & Vision Comput.* 10(6): 376-388.

[Geiger et al. 1992] Geiger, D., Ladendorf, B., and Yuille, A.L. 1992. Occlusions and binocular stereo. In Sandini G.(ed.):*Computer Vision — ECCV92, 2nd European Comference on Computer Vision,* Lecture Notes in Computer Science 588, Springer-Verlag, pp.425-433.

[Girod & Kuo 1989] Girod, B. and Kuo, D. 1989. Direct estimation of displacement histograms. In *Proc. OSA Meeting on Image Understanding and Machine Vision,* Cape Cod, MA, pp.73-76.

[Grimson 1981] Grimson, W.E.L. 1981. *From Images to Surfaces: A Computational Study of the Human Early Visual System.* Cambridge, MA. MIT Press.

[Hubel 1988] Hubel, D.H. 1988. *Eye, Brain, and Vision,* Freeman & Company, New York.

[Hubel & Wiesel 1962] Hubel, D.H. and Wiesel, T.N. 1962. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *J. Physiol.* 160: 106-154.

[Jones & Malik 1992] Jones, D.G. and Malik, J. 1992. A computational framework for determining stereo correspondence from a set of linear spatial filters. In Sandini G.(ed.):*Computer Vision — ECCV92, 2nd European Comference on Computer Vision,* Lecture Notes in Computer Science 588, Springer-Verlag, pp.395-410.

[Julesz 1960] Julesz, B. 1960. Binocular depth perception of computer generated patterns. *Bell Sys. Tech. J.* 38: 1001-1020.

[Kersten 1991] Kersten, D. 1991. Transparency and the cooperative computation of scene attributes. In *Computational Models of Visual Processing* (Landy M.S. and Movshon J.A. eds.), MIT Press, Cambridge, MA.

[Koenderink & van Doorn 1984] The structure of images. *Biol. Cybern.* 50: 363-370.

[Koenderink & van Doorn 1987] Representation of local geometry in the visual system. *Biol. Cybern.* 55: 367-375.

[Koenderink & van Doorn 1990] Koenderink, J.J. and van Doorn, A.J. 1990. Receptive field families. *Biol. Cybern.* 63: 291-298.

[Koenderink & van Doorn 1992] Koenderink, J.J. and van Doorn, A.J. 1992. Receptive field assembly pattern specificity. *J. Vis. Comm. & Image Representation* 3(1): 1-12.

[Krol & van de Grind 1980] Krol, J.D. and van de Grind, W.A. 1980. The double nail illusion. *Percept.* 9: 651-669.

[Little & Gillett 1990] Little, J.J., Gillett, W.E. 1990. Direct evidence for occlusion in stereo and motion. *Image and Vision Comput.* 8(4): 328-340.

[Lucas & Kanade 1981] Lucas, B.D. and Kanade, T. 1981. An iterative image registration technique with an application to stereo vision. *Proc. Image Understanding Workshop,* pp.121-130.

[Madarasmi 1992] Madarasmi, S., Kersten, D., and Pong, T.-C. 1992. A multi-layer approach to segmentation and interpolation with application to stereo vision. *Investigative Opthalmology and Visual Science Supplement(ARVO 1992)* 33: 1371.

[Marr 1982] Marr, D. 1982. *Vision —A Computational Investigation into the Human Representation and Processing of Visual Information—,* Freeman, San Francisco.

[Marr & Poggio 1976] Marr, D. and Poggio, T. 1976. A cooperative computation of stereo disparity. *Science.* 194:283-287.

[Marr & Poggio 1979] Marr, D. and Poggio, T. 1979. A theory of human stereopsis. *Proc. R. Soc. London Ser.B* 204: 301-328.

[Mayhew & Frisby 1981] Mayhew, J.E.W. and Frisby, J.P. 1981. Psychophysical and computational studies towards a theory of human stereopsis. *Artif. Intell.* 17:349-387.

[Noest & Koenderink 1990] Noest, A.J. and Koenderink, J.J. 1990. Visual coherence despite transparency or partial occlusion. *Perception,* 19(3): 384.

[Ohta & Kanade 1985] Ohta, Y. and Kanade, T. 1985. Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Trans. Pattern Anal. & Mach. Intell.* 7(2): 139-154.

[Okutomi & Kanade 1992] Okutomi, M. and Kanade, T. 1992. A locally adaptive window for signal matching. *Int. J. Comput. Vision* 7(2): 143-162.

[Penrose 1989] Penrose, R. 1989. *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics.* Oxford University Press, Oxford, UK.

[Poggio et al. 1985] Poggio, T., Torre, V., and Koch, C. 1985. Computational vision and regularization theory. *Nature,* 317: 314-319.

[Pollard et al. 1985] Pollard, S.B., Mayhew, J.E.W., and Frisby, J.P. 1985. PMF: A stereo correspondence algorithm using a disparity gradient limit. *Percept.* 14: 449-470.

[Pollard & Frisby 1990] Pollard, S.B. and Frisby, J.P. 1990. Transparency and the uniqueness constraint in human and computer stereo vision. *Nature.* 347: 553-556.

[Prazdny 1985] Prazdny, K. 1985. Detection of binocular disparities. *Biol. Cybern.* 52: 93-99.

[Shizawa 1992] Shizawa, M. 1992. On visual ambiguities due to transparency in motion and stereo. In Sandini G.(ed.):*Computer Vision — ECCV92, 2nd European Comference on Computer Vision,* Lecture Notes in Computer Science 588, Springer-Verlag, pp.411-419.

[Shizawa & Mase 1990] Shizawa, M. and Mase, K. 1990. Simultaneous multiple optical flow estimation. In *Proc. Int. Conf. on Pattern Recognition,* 1, Atlantic City, NJ, pp.274-278.

[Shizawa & Mase 1991a] Shizawa, M. and Mase, K. 1991a. A unified computational theory for motion transparency and motion boundaries based on eigenenergy analysis. In *Proc. IEEE Conf. on Comput. Vision and Pattern Recognition,* Maui, HI, pp.289-295.

[Shizawa & Mase 1991b] Shizawa, M. and Mase, K. 1991. Principle of superposition: a common computational framework for analysis of multiple motion. In *Proc. IEEE Workshop on Visual Motion,* Princeton, NJ, pp.164-172.

[Weinshall 1989] Weinshall, D. 1989. Perception of multiple transparent planes in stereo vision. *Nature.* 341: 737-739.

[Weinshall 1991] Weinshall, D. 1991. Seeing "ghost" planes in stereo vision. *Vision Res.* 31(10): 1731-1748.

[Weinshall 1992] Weinshall, D. 1992. A computational study of the matching of doubly ambiguous stereograms. Investigative Opthalmology and Visual Science Supplement(ARVO 1992) 33: 707.

[Yuille & Poggio 1984] Yuille, A.L. and Poggio, T. 1984. A generalized ordering constraint for stereo correspondence. M.I.T. Artif. Intell. Lab., Massachusetts Inst. Technol., Cambridge, MA, A.I. Memo. 777.

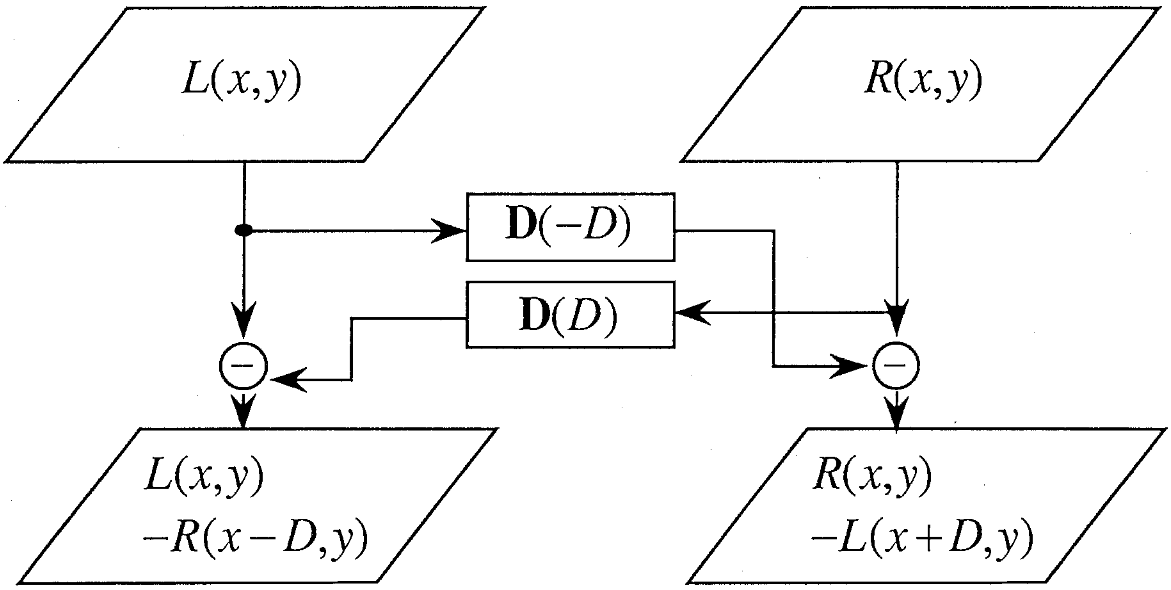


Fig. 1. Function of the amplitude operator.

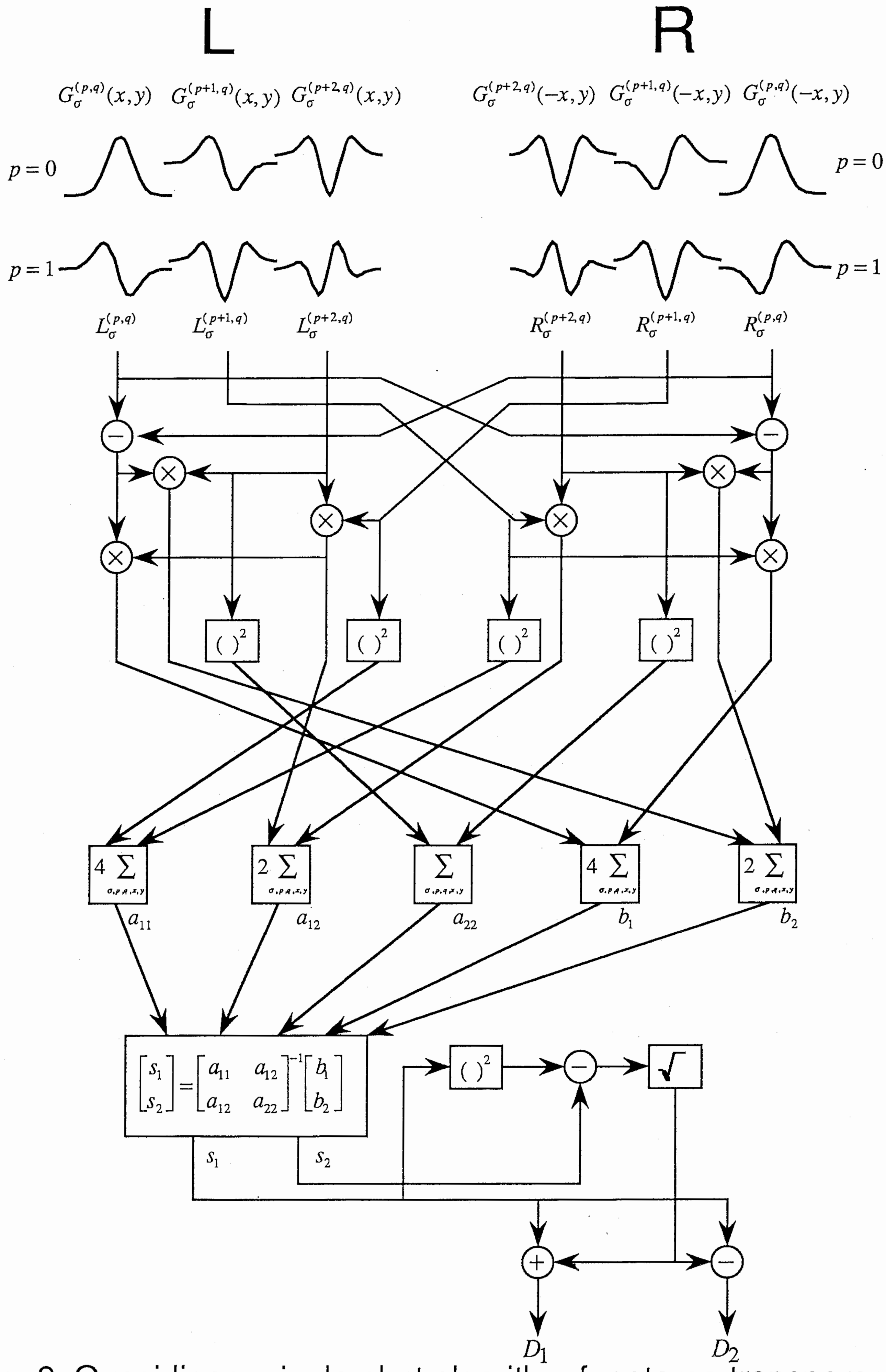
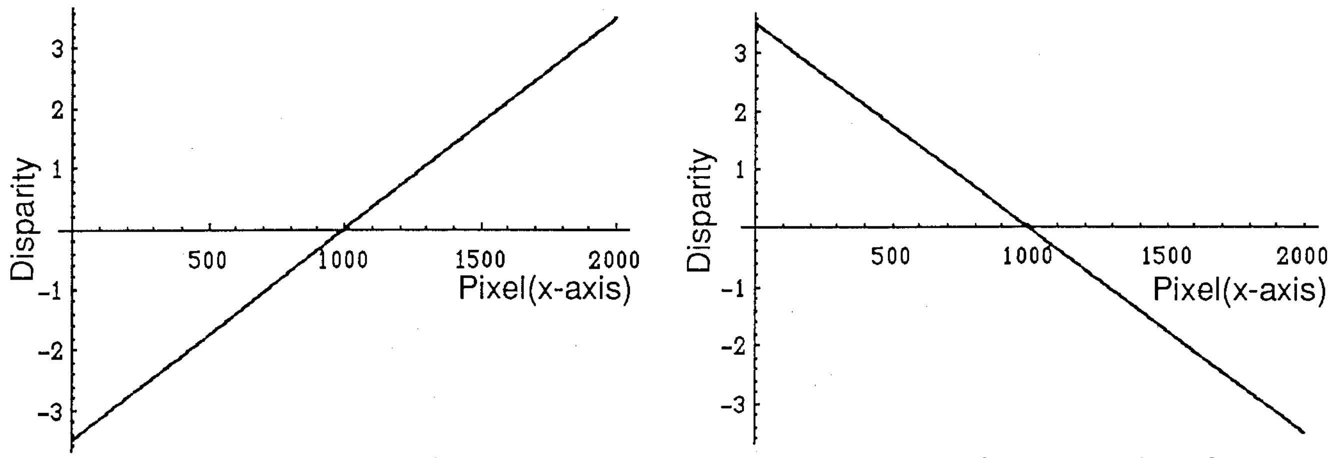
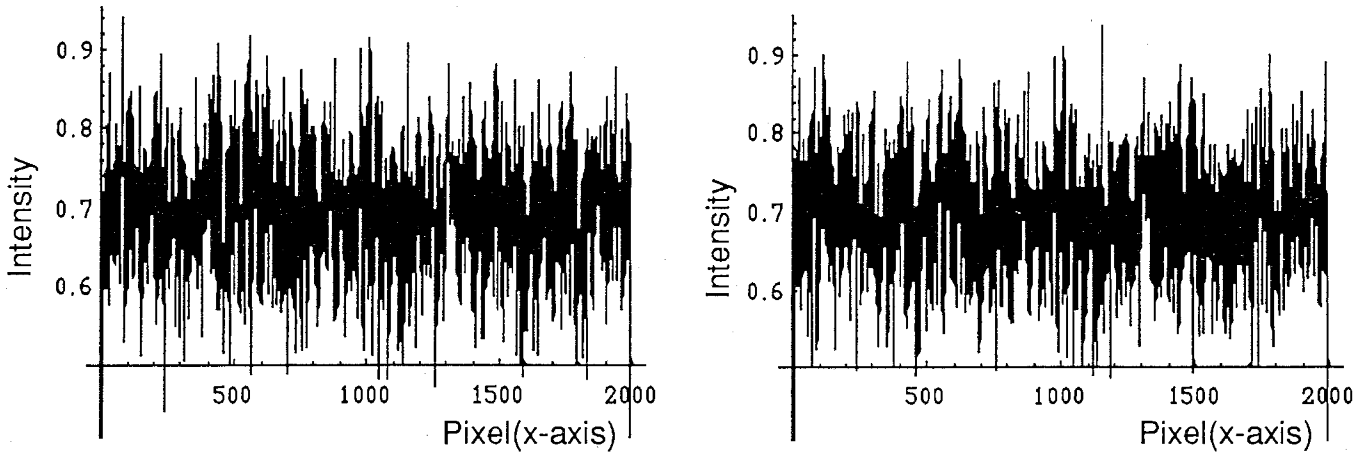


Fig. 2. Quasi-linear single-shot algorithm for stereo transparency.



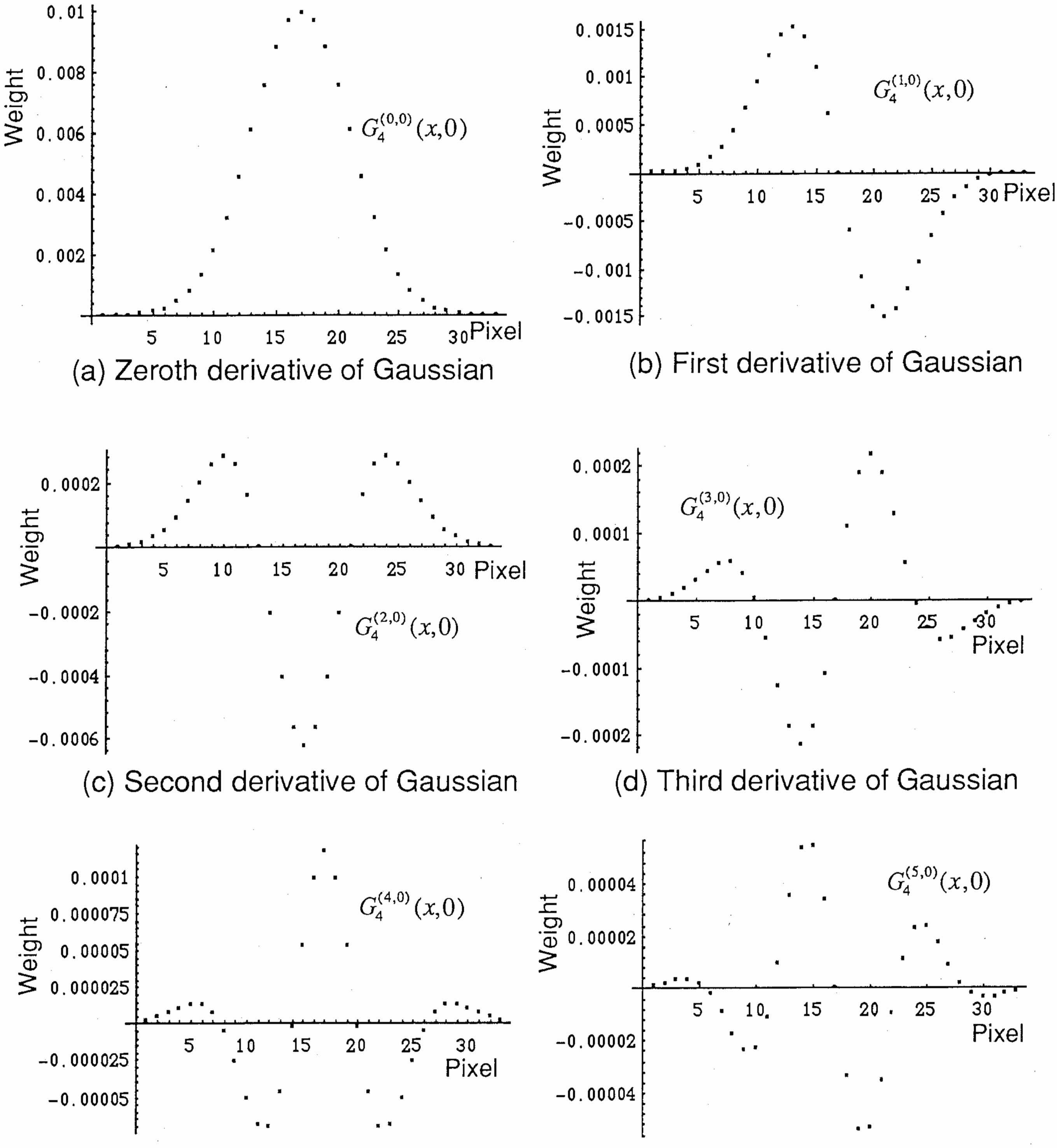
(a) True disparity map of surface #1 (b) True disparity map of surface #2

Fig. 3. True disparity maps of two surfaces for simulations.



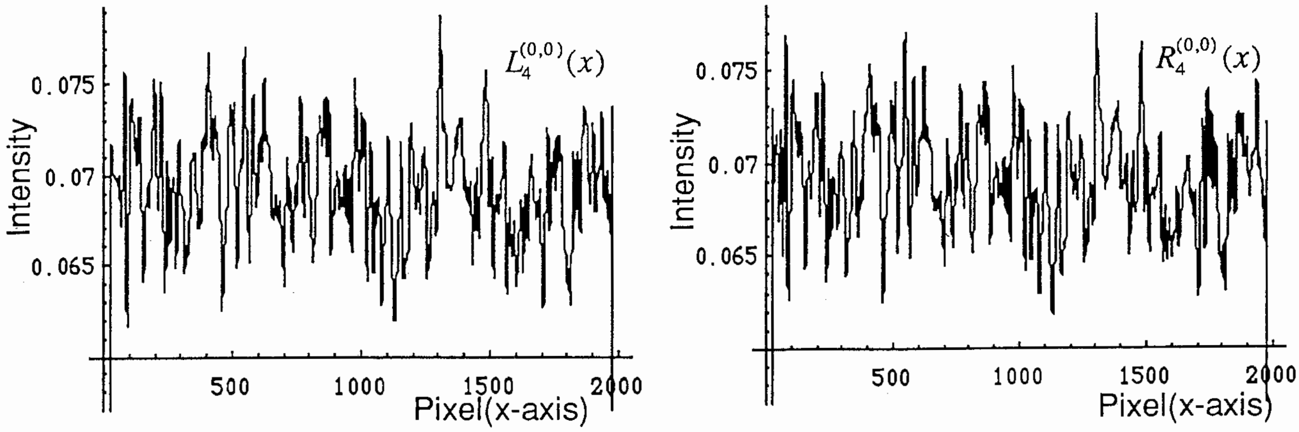
(a) Input signal of left eye (b) Input signal of right eye

Fig. 4. Binocular input signal involving stereo transparency.

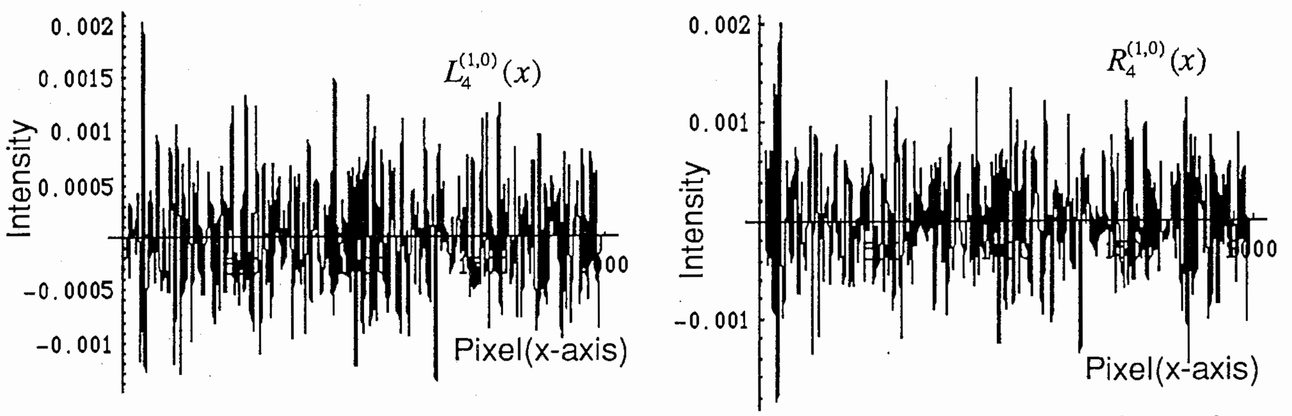


(a) Zeroth derivative of Gaussian
(b) First derivative of Gaussian
(c) Second derivative of Gaussian
(d) Third derivative of Gaussian
(e) Fourth derivative of Gaussian
(f) Fifth derivative of Gaussian

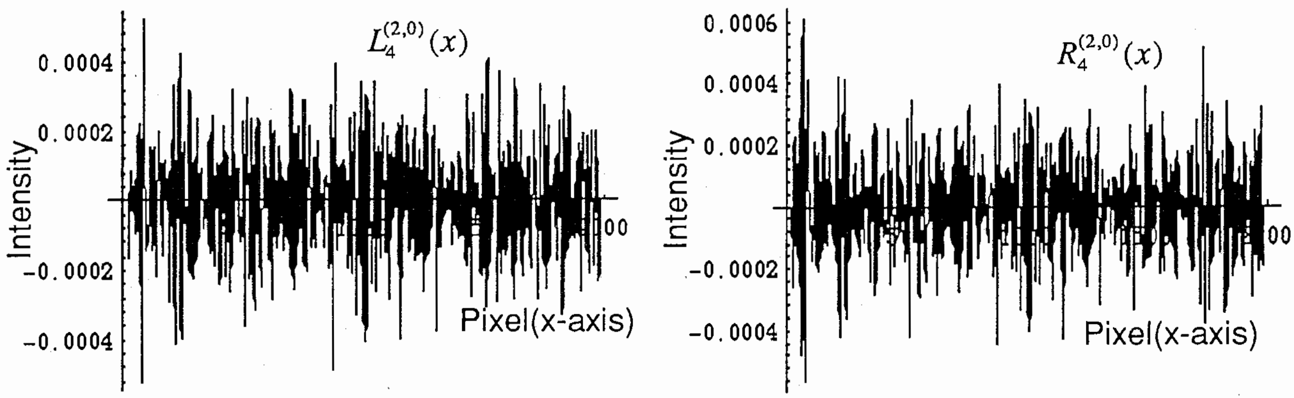
Fig. 5. Kernels for regularized derivatives (derivatives of Gaussian).



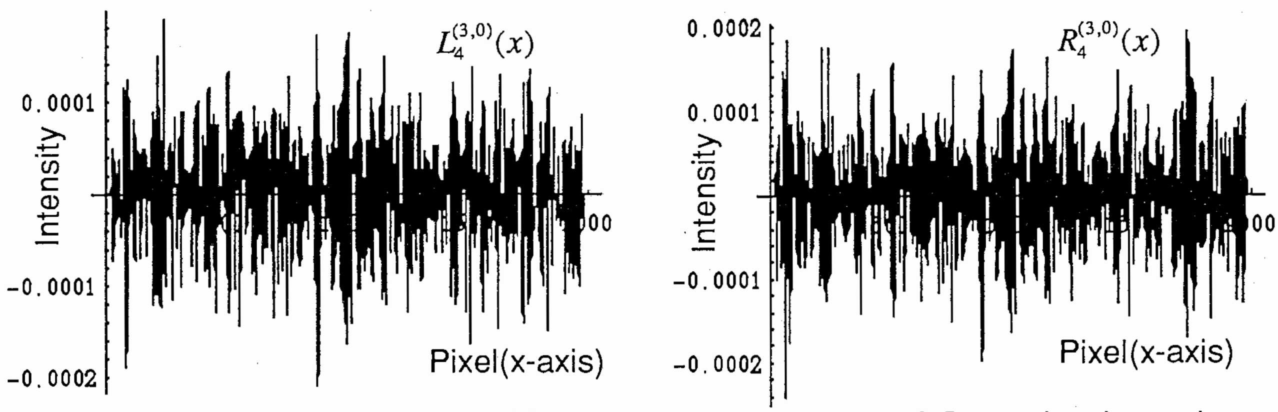
(a) Signals convolved with the Gaussian kernel



(b) Signals convolved with the first derivatives of Gaussian kernel



(c) Signals convolved with the second derivatives of Gaussian kernel



(d) Signals convolved with the third derivatives of Gaussian kernel

Fig. 6. Convolved signals with derivatives of Gaussian.

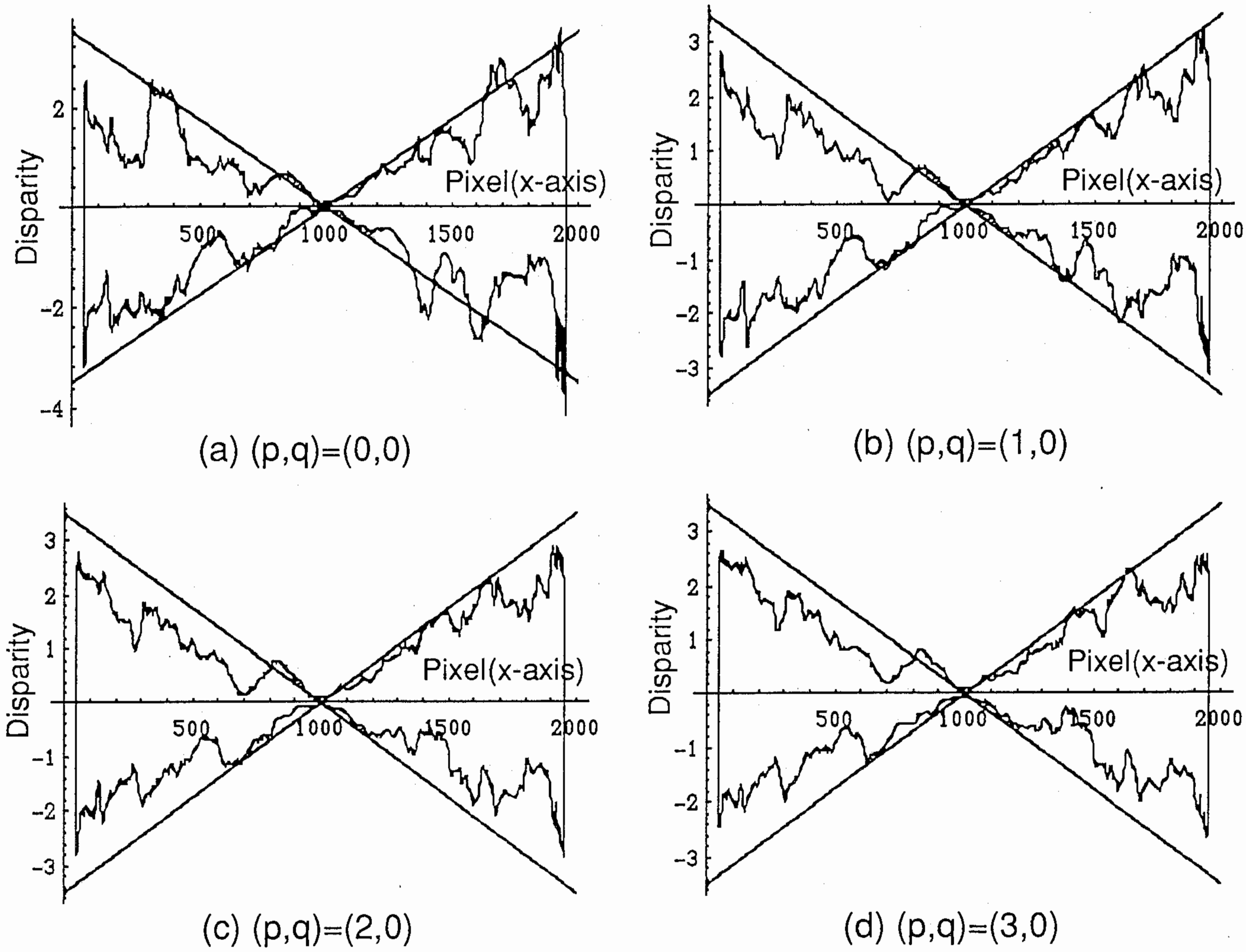


Fig. 7. Two-fold disparities estimated by the single-shot algorithm.