On Visual Ambiguities Due to Transparency in Motion and Stereo *

Masahiko Shizawa

ATR Communication Systems Research Laboratories, Advanced Telecommunications Research Institute International, Sanpeidani, Inuidani, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

Abstract. Transparency produces visual ambiguities in interpreting motion and stereo. Recent discovery of a general framework, *principle of superposition,* for building constraint equations of transparency makes it possible to analyze the mathematical properties of transparency perception. This paper theoretically examines multiple ambiguous interpretations in transparent optical flow and transparent stereo.

1 Introduction

Transparency perception arises when we see scenes with complex occlusions such as picket fences or bushes, with shadows such as those cast by trees, and with physically transparent objects such as water or glass. Conventional techniques for segmentation problems using relaxation type techniques such as coupled MRF(Markov Random Field) with a line process which explicitly models discontinuities[5][13], statistical decision on velocity distributions using statistical voting[1][2][3] or outlier rejection paradigm of robust statistics[14] and weak continuity[15], cannot properly handle these complex situations, since transparency is beyond the assumptions of these techniques. More recently, an iterative estimation technique for two-fold motion from three frames has been proposed[16].

The *principle of superposition(PoS),* a simple and elegant mathematical technique, has been introduced to build motion transparency constraints from conventional single motion constraints[25]. PoS resolves the difficulties in analyzing motion transparency and multiple motions at the level of basic constraints, i.e., of *computational theory* in contrast to conventional *algorithm* level segmentation techniques[21]. Using PoS, we can analyze the nature of transparent motion such as the minimum number of sets of measurements, signal components or correspondences needed to determine motion parameters in finite multiplicity and to determine them uniquely. Another advantage is its computational simplicity in optimization algorithms such as convexity of the energy functionals.

In this paper, the constraints of the two-fold transparent optical flow is examined and ambiguities in determining multiple velocities are discussed. It is shown that conventional statistical voting type techniques and a previously described constraint-based approach[23][24] behave differently for some particular moving patterns. This behavioral difference will provide a scientific test for the biological plausibility of motion perception models regarding transparency.

Then, I show that transparency in binocular stereo vision can be interpreted similarly to transparent motion using PoS. The constraint equations for transparent stereo matching are derived by PoS. Finally, recent results in studies on human perception of multiple transparent surfaces in stereo vision[19] are explained by this computational theory.

* Part of this work was done while the author was at NTT Human Interface Laboratories, Yokosuka, Japan.

Lecture Notes in Computer Science, Vol. 588
G. Sandini (Ed.)
Computer Vision – ECCV '92

2 Principle of Superposition

2.1 The Operator Formalism and Constraints of Transparency

Most of the constraint equations in vision can be written as,

$$a(\mathbf{p})f(\mathbf{x}) = 0. \tag{1}$$

where $f(\mathbf{x})$ is a *data distribution* on *data space* $\mathcal{G}$. $f(\mathbf{x})$ may be the image intensity data itself or outputs of a previous visual process. $\mathbf{p}$ is a point on a parameter space $\mathcal{H}$ which represents a set of parameters to be estimated and $a(\mathbf{p})$ is a linear operator parametrized by $\mathbf{p}$. The linearity of the operator is defined by $a(\mathbf{p})\{f_1(\mathbf{x}) + f_2(\mathbf{x})\} = a(\mathbf{p})f_1(\mathbf{x}) + a(\mathbf{p})f_2(\mathbf{x})$ and $a(\mathbf{p})0 = 0$. We call the operator $a(\mathbf{p})$ the *amplitude operator*. The amplitude operator and the data distribution may take vector values.

Assume n data distributions $f_i(\mathbf{x})(i = 1, 2, \cdots, n)$ on $\mathcal{G}$, and suppose they are constrained by the operators $a(\mathbf{p}_i)(\mathbf{p}_i \in \mathcal{H}_i, i = 1, 2, \cdots, n)$ as $a(\mathbf{p}_i)f_i(\mathbf{x}) = 0$. The data distribution $f(\mathbf{x})$ having transparency is assumed to be an additve superposition of $f_i(\mathbf{x})$ as $f(\mathbf{x}) = \sum_{i=1}^{n} f_i(\mathbf{x})$. According to PoS, the transparency constraint for $f(\mathbf{x})$ can be represented simply by

$$a(\mathbf{p}_1)a(\mathbf{p}_2)\cdots a(\mathbf{p}_n)f(\mathbf{x}) = 0.^2 \tag{2}$$

It should be noted that if the constraint of n-fold transparency holds, then the constraint of m-fold transparency holds for any $m > n$. However, parameter estimation problems based on the constraint of m-fold transparency are ill-posed because extra parameters can take arbitrary values, i.e. are indefinite. Therefore, appropriate multiplicity n may be determined by a certain measure of well-posedness or stability of the optimization as in [24].

2.2 Superposition Under Occlusion and Transparency

An important property of the transparency constraint equation is its *insensitivity to occlusion*. If some region of data $f_i(\mathbf{x})$ is occluded by another pattern, we can *assume* that $f_i(\mathbf{x})$ is zero in the occluded region. The transparency constraint equation still holds because of its linearity. Therefore, in principle, occlusion does not violate the assumption of additive superposition.

In the case of transparency, there are typically two types of superposition: additive and multiplicative. Multiplicative superposition is highly non-linear and therefore substantially violates the additivity assumption. However, taking the logarithm of the data distribution transforms the problem into a case of additive superposition.

3 Visual Ambiguities in Motion Transparency

3.1 The Constraint Equations of Transparent Optical Flow

In the case of optical flow, the amplitude operator in spatial and frequency domains are defined by[24]

$$a(u, v) \equiv u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + \frac{\partial}{\partial t}, \quad \tilde{a}(u, v) \equiv 2\pi i(u\omega_x + v\omega_y + \omega_t), \tag{3}$$

[2] For this constraint to be satisfied strictly, the operator $a(\mathbf{p}_i)$ must commute, i.e., $a(\mathbf{p}_i)a(\mathbf{p}_j) = a(\mathbf{p}_j)a(\mathbf{p}_i)$ for $i \neq j$.

where (u,v) is a flow vector. Then, the fundamental constraints of optical flow can be written as $a(u,v)f(x,y,t)=0$ and $\tilde{a}(u,v)F(\omega_x,\omega_y,\omega_t)=0$ where $f(x,y,t)$ and $F(\omega_x,\omega_y,\omega_t)$ denote a space-time image and its Fourier transform[9][10][11][12]. Using PoS, the constraints for the two-fold transparent optical flow are simply $a(u_1,v_1)a(u_2,v_2)f(x,y,t)=0$ and $\tilde{a}(u_1,v_1)\tilde{a}(u_2,v_2)F(\omega_x,\omega_y,\omega_t)=0$ where (u_1,v_1) and (u_2,v_2) are two flow vectors which coexist at the same image location.

These two constraints of two-fold motion transparency can be expanded into

$$d_{xx}u_1u_2+d_{yy}v_1v_2+d_{xy}(u_1v_2+v_1u_2)+d_{xt}(u_1+u_2)+d_{yt}(v_1+v_2)+d_{tt}=0, \qquad (4)$$

where components of $\mathbf{d}=(d_{xx},d_{yy},d_{xy},d_{xt},d_{yt},d_{tt})$ are for example $d_{yt}=\frac{\partial^2}{\partial y\partial t}f(x,y,t)$ for the spatial domain representation and $d_{yt}=(2\pi i)^2\omega_y\omega_tF(\omega_x,\omega_y,\omega_t)$ for the frequency domain representation. Therefore, we can simultaneously discuss brightness measuments and frequency components.

3.2 The Constraint Curve of Two-fold Transparent Optical Flow

Equation (4) is quadratic in four unknowns u_1,v_1,u_2 and v_2. Therefore, if we have four independent measurements or signal components $\mathbf{d}^{(k)}(k=1,2,3,4)$, a system of four quadratic constraint equations denoted by E_k will produce solutions of a finite ambiguity. The solution can be obtained as intersections of two cubic curves in velocity space as shown below. This is the two-fold transparent motion version of the well-known fact that the intersection of two lines which represent the single optical flow constraint equations in the velocity space (u,v) uniquely determines a flow vector.

From E_1 and E_2, we can derive rational expressions $u_2=G_u(\mathbf{d}^{(1)},\mathbf{d}^{(2)};u_1,v_1)$ and $v_2=G_v(\mathbf{d}^{(1)},\mathbf{d}^{(2)};u_1,v_1)$ which transform the flow vector (u_1,v_1) into (u_2,v_2) and vice versa. The concrete forms of these rational expressions can be written as

$$G_u(\mathbf{d}^{(i)},\mathbf{d}^{(j)};u,v)=\frac{q_y^{(i)}q_t^{(j)}-q_t^{(i)}q_y^{(j)}}{q_x^{(i)}q_y^{(j)}-q_y^{(i)}q_x^{(j)}},\quad G_v(\mathbf{d}^{(i)},\mathbf{d}^{(j)};u,v)=\frac{q_t^{(i)}q_x^{(j)}-q_x^{(i)}q_t^{(j)}}{q_x^{(i)}q_y^{(j)}-q_y^{(i)}q_x^{(j)}}, \qquad (5)$$

where $q_x^{(i)}\equiv(d_{xx}^{(i)}u+d_{xy}^{(i)}v+d_{xt}^{(i)})$, $q_y^{(i)}\equiv(d_{xy}^{(i)}u+d_{yy}^{(i)}v+d_{yt}^{(i)})$ and $q_t^{(i)}\equiv(d_{xt}^{(i)}u+d_{yt}^{(i)}v+d_{tt}^{(i)})$.

If we have three measurements/components $\mathbf{d}^{(i)}$, $\mathbf{d}^{(j)}$ and $\mathbf{d}^{(k)}$, then the equation $G_u(\mathbf{d}^{(i)},\mathbf{d}^{(j)};u,v)=G_u(\mathbf{d}^{(i)},\mathbf{d}^{(k)};u,v)$ gives the constraint for the velocity (u,v) in the case of two-fold transparent optical flow. This equation can be factored into the form of $q_x^{(i)}G_{uv}(\mathbf{d}^{(i)},\mathbf{d}^{(j)},\mathbf{d}^{(k)};u,v)=0$ where

$$\begin{aligned}G_{uv}(\mathbf{d}^{(i)},\mathbf{d}^{(j)},\mathbf{d}^{(k)};u,v)=&\;q_x^{(i)}q_y^{(j)}q_t^{(k)}+q_y^{(i)}q_t^{(j)}q_x^{(k)}+q_t^{(i)}q_x^{(j)}q_y^{(k)}\\&-q_x^{(i)}q_t^{(j)}q_y^{(k)}-q_y^{(i)}q_x^{(j)}q_t^{(k)}-q_t^{(i)}q_y^{(j)}q_x^{(k)}.\end{aligned} \qquad (6)$$

If $q_x^{(i)}=0$ then we can substitute the i by another index i' which is not equivalent to i, j and k. Then $q_x^{(i')}=0$ cannot hold if we have transparency, because two equations $q_x^{(i)}=0$ and $q_x^{(i')}=0$ imply single optical flow. Thus, we can substitute i by i' without loss of generality. Therefore, the cubic equation $G_{uv}(\mathbf{d}^{(i)},\mathbf{d}^{(j)},\mathbf{d}^{(k)};u,v)=0$ with respect to u and v gives the constraint curve on the velocity space (u,v) under the assumption of two-fold transparency. Intersecting points of two curves in uv-space

$$C_1:\;G_{uv}(\mathbf{d}^{(1)},\mathbf{d}^{(2)},\mathbf{d}^{(3)};u,v)=0,\quad C_2:\;G_{uv}(\mathbf{d}^{(1)},\mathbf{d}^{(2)},\mathbf{d}^{(4)};u,v)=0, \qquad (7)$$

provide the candidate flow estimates for (u_1,v_1) and (u_2,v_2). By using (5), we can make pairs of solutions for $\{(u_1,v_1),(u_2,v_2)\}$ from these intersections.

The Three-fold Ambiguity of Four-component Motion. In the space-time frequency domain, there exists a three-fold ambiguity in interpreting the transparent motion of four frequency components, since there are three possible ways to fit two planes so that they pass through all four points (frequency components) and the origin. Figure 1 provides the predicted visual ambiguity due to this fact. If we have two image patterns A and B each of which has frequency components along just two space directions ($\{G_1, G_2\}$ for A and $\{G_3, G_4\}$ for B), and they move with different velocities $\mathbf{v}_A$ and $\mathbf{v}_B$, their superposed motion pattern has three-fold multiple interpretations.

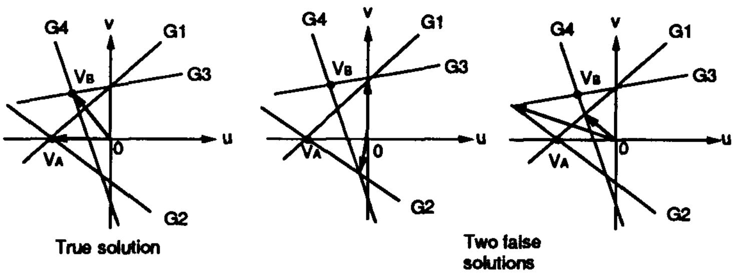

Fig. 1. The three-fold ambiguity of transparent motion

3.3 Unique Solution from Five Measurements or Components

If a system of five constraint equations E_k of five independent measurements or five frequency components $\mathbf{d}^{(k)}(k = 1, 2, 3, 4, 5)$ are available, we can determine two velocities uniquely. The system of equations can be solved with respect to a vector of five 'unknown parameters',

$$\mathbf{c} = (c_{xx}, c_{yy}, c_{xy}, c_{xt}, c_{yt}) = \left(u_1 u_2, v_1 v_2, \frac{1}{2}(u_1 v_2 + v_1 u_2), \frac{1}{2}(u_1 + u_2), \frac{1}{2}(v_1 + v_2) \right), \quad (8)$$

as a linear system. Component flow parameters u_1, u_2, v_1 and v_2 can be obtained by solving two quadratic equations, $u^2 - 2c_{xt}u + c_{xx} = 0$ and $v^2 - 2c_{yt}v + c_{yy} = 0$. We denote their solutions as $u_\pm = c_{xt} \pm \sqrt{c_{xt}^2 - c_{xx}}$ and $v_\pm = c_{yt} \pm \sqrt{c_{yt}^2 - c_{yy}}$. There are constraints $c_{xt}^2 \geq c_{xx}$ and $c_{yt}^2 \geq c_{yy}$ for the existence of real solutions. We now have two possible solutions for (u_1, v_1) and (u_2, v_2) as $\{(u_1, v_1), (u_2, v_2)\} = \{(u_+, v_+), (u_-, v_-)\}$ and $\{(u_1, v_1), (u_2, v_2)\} = \{(u_+, v_-), (u_-, v_+)\}$. However, we can determine a true solution by checking their consistency with the remaining relation $c_{xy} = \frac{1}{2}(u_1 v_2 + v_1 u_2)$ of (8). Therefore, we have a unique interpretation for the general case.

Behavioral Difference Against Conventional Schemes. The significance of transparent motion analysis described above is its capability of estimating multiple motion simultaneously from the *minimum* amount of image information, i.e. minimum measurements or signal components as shown above. In this subsection, I show that conventional techniques by statistical voting of constraint lines on velocity space (e.g. [1][2][3][4]) cannot correctly estimate multiple flow vectors from this minimum information.

Figure 2(a) shows an example of moving patterns which produces this behavioral difference between the proposed approach and conventional statistical voting. The two moving patterns A and B are superposed. Pattern A has two frequency components which may be produced by two plaids G_1 and G_2; its velocity is $\mathbf{v}_A$. The other pattern B, which has velocity $\mathbf{v}_B$, contains three frequency components produced by three plaids G_3, G_4 and G_5. If the superposed pattern is given to our algorithm based on the transparent optical flow constraint, the two flow vectors $\mathbf{v}_A$ and $\mathbf{v}_B$ can be determined uniquely as shown in the previous subsection. Figure 2(b) shows plots of conventional optical flow constraint lines on the velocity space (u, v). There are generally seven intersection points only one of which is an intersection of three constraint lines but other six points are of two constraint lines.[3] The intersection of three lines is the velocity $\mathbf{v}_B$ and can be detected by a certain peak detection or clustering techniques on the velocity space. However, the other velocity $\mathbf{v}_A$ cannot be descriminated from among the six two-line intersections!

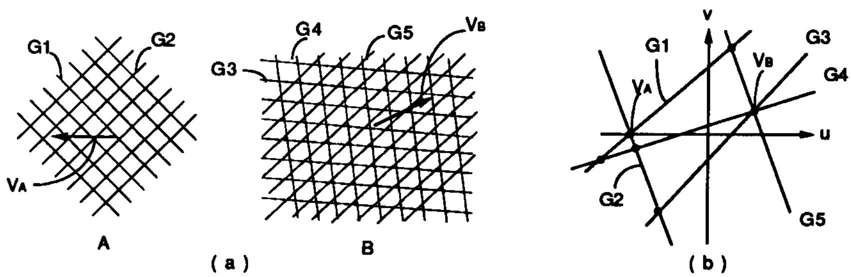

Fig. 2. Moving pattern from which statistical voting schemes cannot estimate the correct two flow vectors

4 Visual Ambiguities in Stereo Transparency

In this section, the transparency in stereo is examined by PoS. Weinshall[19][20] has demonstrated that the uniqueness constraint of matching and order preservation constraint are not satisfied in the multiple transparent surface perception in human stereo vision. Conventional stereo matching algorithms cannot correctly explain perception of transparency, i.e., multiple surfaces[20]. My intention is not to provide a stereo algorithm for the transparent surface reconstruction, but to provide stereo matching constraints which admit and explain the transparency perception.

4.1 The Constraint of Stereo Matching

The constraints on stereo matching can also be written by the operator formalism. We denote the left image patterns by $L(x)$ and the right image patterns by $R(x)$ where x denotes a coordinate along an epipolar line. Then, the constraint for single surface stereo

[3] Figure 2(b) actually contains only five two-line intersections. However, in general, it will contain six.

with disparity D can be written as

$$\mathbf{a}(D)\mathbf{f}(x) = \mathbf{0} \quad \text{where,} \quad \mathbf{a}(D) \equiv \begin{bmatrix} 1 & -\mathcal{D}(D) \\ -\mathcal{D}(-D) & 1 \end{bmatrix}, \quad \mathbf{f}(x) \equiv \begin{bmatrix} L(x) \\ R(x) \end{bmatrix}. \tag{9}$$

$\mathcal{D}(D)$ is a shift operator which transforms $L(x)$ into $L(x-D)$ and $R(x)$ into $R(x-D)$.[4] It is easy to see that the vector amplitude operator $\mathbf{a}(D)$ is linear, i.e. both $\mathbf{a}(D)\mathbf{0} = \mathbf{0}$ and $\mathbf{a}(D)\{\mathbf{f}_1(x) + \mathbf{f}_2(x)\} = \mathbf{a}(D)\mathbf{f}_1(x) + \mathbf{a}(D)\mathbf{f}_2(x)$ hold.

Figure 3 is a schematic diagram showing the function of the vector amplitude operator $\mathbf{a}(D)$. The operator $\mathbf{a}(D)$ eliminates signal components of disparity D from the pair of stereo images, $L(x)$ and $R(x)$, by substitutions.

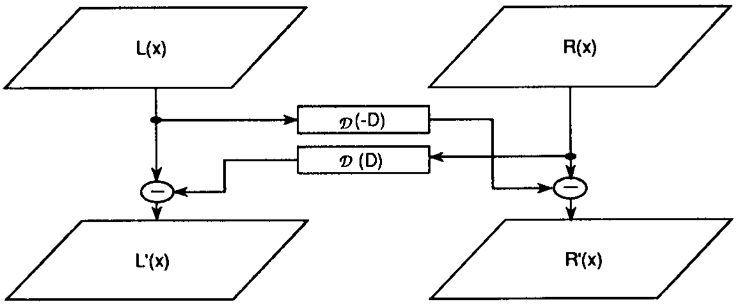

Fig. 3. Function of the amplitude operator of stereo matching

4.2 The Constraint of Transparent Stereo

According to PoS, the constraint of the n-fold transparency in stereo can be hypothesized as

$$\mathbf{a}(D_n) \cdots \mathbf{a}(D_2)\mathbf{a}(D_1)\mathbf{f}(x) = \mathbf{0}, \tag{10}$$

where $\mathbf{f}(x) = \sum_{i=1}^{n} \mathbf{f}_i(x)$, and each $\mathbf{f}_i(x)$ is constrained by $\mathbf{a}(D_i)\mathbf{f}_i(x) = \mathbf{0}$. It is easily proved using the commutability of the shift operator $\mathcal{D}(D)$ that amplitude operators $\mathbf{a}(D_i)$ and $\mathbf{a}(D_j)$ commute, i.e. $\mathbf{a}(D_i)\mathbf{a}(D_j) = \mathbf{a}(D_j)\mathbf{a}(D_i)$ for $i \neq j$ under the condition of constant D_i and D_j. Further, the additivity assumption on superposition is reasonable for random dot stereograms of small dot density.

4.3 Perception of Multiple Transparent Planes

In this section, the human perception of transparent multiple planes in stereo vision reported in [19] is explained by the hypothesis provided in the previous section.

We utilize a random dot image $P(x)$. If $L(x) = P(x-d)$ and $R(x) = P(x)$, then the constraint of single surface stereo holds for disparity $D = d$, since

$$\mathbf{a}(d)\mathbf{f}(x) = \begin{bmatrix} P(x-d) - \mathcal{D}(d)P(x) \\ P(x) - \mathcal{D}(-d)P(x-d) \end{bmatrix} = \begin{bmatrix} P(x-d) - P(x-d) \\ P(x) - P(x-d+d) \end{bmatrix} = \mathbf{0}. \tag{11}$$

[4] We can write this shift operator explicitly in a differential form as $\mathcal{D}(D) = \exp(-D\frac{\partial}{\partial x}) = 1 - D\frac{\partial}{\partial x} + \frac{D^2}{2!}\frac{\partial^2}{\partial x^2} - \frac{D^3}{3!}\frac{\partial^3}{\partial x^3} \cdots$. However, only the shifting property of the operators is essential in the following discussions.

In [19], a repeatedly superimposed random dot stereogram shown in Fig.4 is used to produce the transparent plane perception. This situation can be represented by defining $L(x)$ and $R(x)$ as

$$L(x) \equiv P(x) + P(x - d_L), \quad R(x) \equiv P(x) + P(x + d_R), \tag{12}$$

where d_L and d_R are shift displacements for the pattern repetitions in left and right image planes. According to [19], when $d_L \neq d_R$, we perceive four transparent planes which correspond to disparities $D = 0, d_L, d_R$ and $d_L + d_R$. The interesting phenomenon occurs in the case of $d_L = d_R = d_C$. The stereogram produces a *single* plane perception despite the fact that the correlation of two image patterns $L(x)$ and $R(x)$ has three strong peaks at the disparities $D = 0, d_C$ and $2d_C$.

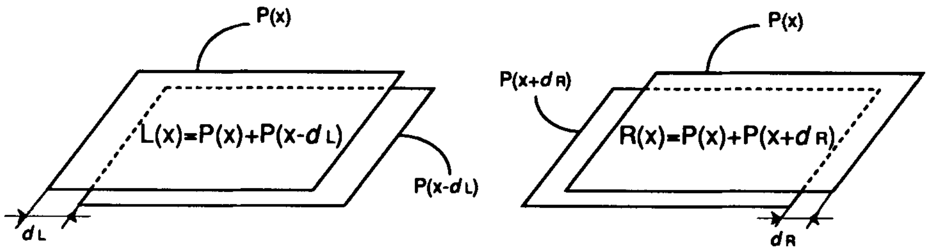

Fig. 4. The stereogram used in the analysis

From the viewpoint of the constraints of transparent stereo, these phenomena can be explained as shown below.

First, it should be pointed out that the data distribution $\mathbf{f}(x)$ can be represented as a weighted linear sum of four possible unique matching components.

$$\mathbf{f}(x) = \alpha\mathbf{f}_1(x) + \alpha\mathbf{f}_2(x) + (1-\alpha)\mathbf{f}_3(x) + (1-\alpha)\mathbf{f}_4(x), \tag{13}$$

where

$$\mathbf{f}_1(x) = \begin{bmatrix} P(x) \\ P(x) \end{bmatrix}, \ \mathbf{f}_2(x) = \begin{bmatrix} P(x-d_L) \\ P(x+d_R) \end{bmatrix}, \ \mathbf{f}_3(x) = \begin{bmatrix} P(x-d_L) \\ P(x) \end{bmatrix}, \ \mathbf{f}_4(x) = \begin{bmatrix} P(x) \\ P(x+d_R) \end{bmatrix}, \tag{14}$$

and

$$\mathbf{a}(0)\mathbf{f}_1(x) = 0, \quad \mathbf{a}(d_L + d_R)\mathbf{f}_2(x) = 0, \quad \mathbf{a}(d_L)\mathbf{f}_3(x) = 0, \quad \mathbf{a}(d_R)\mathbf{f}_4(x) = 0. \tag{15}$$

Note that the weights have only one freedom as parameterized by α.

When assuming $d_L \neq d_R$, the following observation can be obtained regarding the constraints of the transparent stereo.

1. The constraint of single surface stereo $\mathbf{a}(D_1)\mathbf{f}(x) = \mathbf{0}$ cannot hold for *any* values of disparities D_1.
2. The constraint of two-fold transparent stereo $\mathbf{a}(D_2)\mathbf{a}(D_1)\mathbf{f}(x) = \mathbf{0}$ can hold only for two sets of disparities $\{D_1, D_2\} = \{0, d_L + d_R\}$ and $\{D_1, D_2\} = \{d_R, d_L\}$ which correspond to $\alpha = 1$ and $\alpha = 0$, respectively.

3. The constraint of three-fold transparency $\mathbf{a}(D_3)\mathbf{a}(D_2)\mathbf{a}(D_1)\mathbf{f}(x) = \mathbf{0}$ can hold only when $\alpha = 1$ or $\alpha = 0$ as same as the two-fold transparency. However, one of the three disparities can take arbitrary value.
4. The constraint of four-fold transparency $\mathbf{a}(D_4)\mathbf{a}(D_3)\mathbf{a}(D_2)\mathbf{a}(D_1)\mathbf{f}(x) = \mathbf{0}$ can hold for arbitrary α. The possible set of disparities is unique as $\{D_1, D_2, D_3, D_4\} = \{0, d_L, d_R, d_L + d_R\}$ except the cases of $\{D_1, D_2\} = \{0, d_L + d_R\}$ and $\{D_1, D_2\} = \{d_L, d_R\}$ which correspond to $\alpha = 1$ and $\alpha = 0$, respectively.
5. The constraints of more than four-fold transparency can hold, but some of the disparity parameters can take arbitrary values.

We can conclude that the stereo constraint of n-fold transparency is valid only for $n = 2$ and $n = 4$ by using the criterion of Occam's razor, i.e., the disparities should not take continuous arbitrary values. Then, in both cases for $n = 2$ and $n = 4$, the theory predicts coexistence of four disparities 0, d_L, d_R and $d_L + d_R$.

When $d_L = d_R = d_C$, the constraint of single surface stereo $\mathbf{a}(D_1)\mathbf{f}(x) = 0$ can hold only for $D_1 = d_C$, since

$$\begin{aligned}\mathbf{a}(d_C)\mathbf{f}(x) &= \begin{bmatrix} \{P(x) + P(x - d_C)\} - \mathcal{D}(d_C)\{P(x) + P(x + d_C)\} \\ \{P(x) + P(x + d_C)\} - \mathcal{D}(-d_C)\{P(x) + P(x - d_C)\} \end{bmatrix} \\ &= \begin{bmatrix} \{P(x) + P(x - d_C)\} - \{P(x - d_C) + P(x + d_C - d_C)\} \\ \{P(x) + P(x + d_C)\} - \{P(x + d_C) + P(x - d_C + d_C)\} \end{bmatrix} = \mathbf{0}. \quad (16)\end{aligned}$$

Therefore, the case of $d_L = d_R$ *must* produce the single surface perception, if we claim the criterion of Occam's razor on disparities.

5 Conclusion

I have analyzed visual ambiguities in transparent optical flow and transparent stereo using the principle of superposition formulated by parametrized linear operators. Ambiguities in velocity estimates for particular transparent motion patterns were examined by mathematical analyses of the transparent optical flow constraint equations. I also pointed out that conventional statistical voting schemes on velocity space cannot estimate multiple velocity vectors correctly for a particular transparent motion pattern. Further, the principle of superposition was applied to transparent stereo and human perception of multiple ambiguous transparent planes was explained by the operator formalism of the transparent stereo matching constraint and the criterion of Occam's razor on the number of disparities.

Future work may include development of a stereo algorithm based on the constraints of transparent stereo. The research reported in this paper will not only lead to modification and extension of the computational theories of motion and stereo vision, but will also help with modeling human motion and stereo vision by incorporating transparency.

Acknowledgments. The author would like thank Dr. Kenji Mase of NTT Human Interface Laboratories and Dr. Shin'ya Nishida of ATR Auditory and Visual Perception Laboratories for helpful discussions.

He also thanks Drs. Jun Ohya, Ken-ichiro Ishii, Yukio Kobayashi and Takaya Endo of NTT Human Interface Laboratories as well as Drs. Fumio Kishino, Nobuyoshi Terashima and Kohei Habara of ATR Communication Systems Research Laboratories for their kind support.

References

1. C.L.Fennema and W.Thompson: "Velocity Determination in Scenes Containing Several Moving Objects," *CGIP*, Vol.9, pp.301-315(1979).
2. J.J.Little, H.Bülthoff and T.Poggio: "Parallel Optical Flow Using Local Voting," *Proc. 2nd ICCV*, Tampa, FL, pp.454-459(1988).
3. R.Jasinschi, A.Rosenfeld and K.Sumi: "The Perception of Visual Motion Coherence and Transparency: a Statistical Model," *Tech. Rep. Univ. of Maryland*, CAR-TR-512(1990).
4. D.J.Fleet and A.D.Jepson: "Computation of Component Image Velocity from Local Phase Information," *IJCV*, Vol.5, No.1, pp.77-104(1990).
5. D.W.Murray and B.F.Buxton: "Scene Segmentation from Visual Motion Using Global Optimization," *IEEE Trans. PAMI*, Vol.9, No.2, pp.220-228(1987).
6. B.G.Schunck: "Image Flow Segmentation and Estimation by Constraint Line Clustering," *IEEE Trans. PAMI*, Vol.11, No.10, pp.1010-1027(1989).
7. E.H.Adelson and J.A.Movshon: "Phenomenal Coherence of Moving Visual Patterns," *Nature*, **300**, pp.523-525(1982).
8. G.R.Stoner, T.D.Albright and V.S.Ramachandran: "Transparency and coherence in human motion perception," *Nature*, **344**, pp.153-155(1990).
9. B.K.P.Horn and B.G.Schunck: "Determining Optical Flow," *Artificial Intelligence*, Vol.17, pp.185-203(1981).
10. E.H.Adelson and J.R.Bergen: "Spatiotemporal Energy Models for the Perception of Motion," *J.Opt.Soc.Am.A*, Vol.2, pp.284-299(1985).
11. J.G.Daugman: "Pattern and Motion Vision without Laplacian Zero Crossings," *J.Opt.Soc.Am.A*, Vol.5, pp.1142-1148(1987).
12. D.J.Heeger: "Optical Flow Using Spatiotemporal Filters," *IJCV*, **1**, pp.279-302(1988).
13. S.Geman and D.Geman: "Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images," *IEEE Trans. PAMI*, **6**, pp.721-741(1984).
14. P.J.Besl, J.B.Birch and L.T.Watson: "Robust Window Operators," *Proc. 2nd ICCV*, Tampa, FL, pp.591-600(1988).
15. A.Blake and A.Zisserman: **Visual Reconstruction**, MIT Press, Cambridge, MA(1987).
16. J.R.Bergen, P.Burt, R.Hingorani and S.Peleg: "Transparent-Motion Analysis," *Proc. 1st ECCV*, Antibes, France, pp.566-569(1990).
17. A.L.Yuille and N.M.Grzywacz: "A Computational Theory for the Perception of Coherent Visual Motion," *Nature*, **333**, pp.71-74(1988).
18. N.M.Grzywacz and A.L.Yuille: "A Model for the Estimate of Local Image Velocity by Cells in the Visual Cortex," *Proc. Royal Society of London*, Vol.B239, pp.129-161(1990).
19. D.Weinshall: "Perception of multiple transparent planes in stereo vision," *Nature*, **341**, pp.737-739(1989).
20. D.Weinshall: "The computation of multiple matching in stereo," *14th European Conference on Visual Perception(ECVP)*, page A31(1990).
21. D.Marr: **Vision**, Freeman(1982).
22. R.Penrose: **The Emperor's New Mind: Concerning Computers, Minds, and The Laws of Physics**, Oxford University Press, Oxford(1989).
23. M.Shizawa and K.Mase: "Simultaneous Multiple Optical Flow Estimation," *Proc. 10th ICPR*, Atlantic City, NJ, Vol.I, pp.274-278(June,1990).
24. —: "A Unified Computational Theory for Motion Transparency and Motion Boundaries Based on Eigenenergy Analysis," *Proc. IEEE CVPR'91*, Maui, HI, pp.289-295(June,1991).
25. —: "Principle of Superposition: A Common Computational Framework for Analysis of Multiple Motion," *Proc. IEEE Workshop on Visual Motion*, Princeton, NJ, pp.164-172(October,1991).