A Nonlinear Diffusion Model for Discontinuous Disparity and Half-Occlusions in Stereo

Jayant Shah

Mathematics Department, Northeastern University, Boston, Mass. 02115

Abstract[1]: *Most stereo algorithms do not take into account discontinuities in disparity and the fact that there are half-occlusions consisting of areas seen by one eye but not the other. At the same time, very few of them are formulated using the framework of energy functionals which are so succesfully used in other areas of computer vision such as image segmentation and surface representation. In this paper, a formulation is presented within such a framework taking into account the discontinuities and half-occlusions. The formulation follows directly from the assumption that when matching the left and right images, the order of points must be preserved. A model is derived consisting of two coupled energy functionals corresponding to the two eyes. They are coupled in the sense that the discontinuity locus determined by one eye also determines the occluded area in the image seen by the other eye. A nonlinear system of diffusion equations is derived by simultaneously applying gradient descent to these functionals. The diffusion equations are implemented by a straight-forward finite-difference scheme.*

1. Introduction

In binocular stereo vision, we have a pair of images as seen by two pinhole cameras ("eyes"). For the sake of simplicity, the image planes of the two cameras are assumed to be coplanar and situated behind the focal points.. The central problem in stereo vision is what is called the correspondence problem. One has to match the two images, pixel by pixel, so that one can compute the distances of objects from the observer.

Most of the recent stereo algorithms are feature-based algorithms [1,2,4,5]. Prominent features in the two images are extracted and matched so that a sparse but presumably accurate disparity map is obtained. The "area-based" methods involve matching patches of areas in the two images. The resulting disparity map is thus dense, but the disparity values represent averages over the patches. Very few of these algorithms use the framework of energy functionals which are so successfully employed in other areas of computer vision such as image segmentation and surface representation. The few algorithms which are based on energy functionals are usually cast in the framework of Markov random fields and use non-deterministic methods to find solutions.

What is more important is that these algorithms do not take into account discontinuities in disparity and the presence of unmatchable areas in each image due to half-occlusions (that is, the parts of the scene which are seen by one eye and not the other). Although the need for such a model has been expressed by many researchers, only recently have such models been proposed [3,7]. These models are formulated using the Bayesian framework. In contrast, this paper presents a similar, but simpler model within the framework of energy functionals. Nonlinear diffusion instead of annealing is used to find solutions.

The formulation follows naturally from the assumption that matching must be order-preserving. The initial form of the model that is derived is very similar to the model in [7]. However, the energy functional is too complex for implementation by the method of steepest gradient descent and it is necessary to approximate the model and look for approximate solutions. For this purpose, we consider two additional versions of the model as seen from the two eyes. The locus of discontinuity in disparity as seen from one eye is different from that seen from the other. In fact, as noted already in [3], a discontinuity in disparity perceived by one eye corresponds to an unmatchable patch in the image seen by the other eye. Hence the two discontinuity loci as seen from the two eyes are in a sense disjoint. We use this fact to our advantage for finding approximate solutions. Our strategy is to alternate between the two functionals when we apply the steepest descent. We fix the discontiuity locus as seen from the right eye and hence the occlusion for the left eye when we try to minimize the left functional and determine the left discontinuity locus. Similarly we fix the left discontinuity locus when we try to minimize the right functional and determine the right discontinuity locus. After some further approximations, we arrive at two coupled funtionals, each of which is similar in form to the image segmentation functional described in [8]. The functionals are coupled

[1] This research was partially supported under ARO Grant, No. DAAL03–91–G-0041.

in the sense that discontinuity found by each determines the occlusion area for the other.

In order to implement these functionals by the method of steepest descent, we follow Ambrosio and Tortorelli and convert the line processes representing the discontinuity loci into variables continuously defined over all of the corresponding image planes [9]. We obtain a nonlinear system of diffusion equations which is implemented by a straight-forward finite-difference scheme. Note that it is crucial that the disparity is a continuous variable and thus represented by floating point numbers in the numerical implementation. The reason is that unless the disparity is constant across an object, its right and left images will have different sizes and thus will occupy different numbers of pixels. If we use only integral values for disparity, then we are forced to match the images of the object on a pixel-by-pixel basis and some pixels will remain unmatched even though there is no occlusion.

Illustrative numerical examples are given as an empirical justification for the approach presented in this paper.

2. Formulation

Let R_l and R_r be the left and right image planes respectively. Let I_l and I_r be the images projected onto R_l and R_r. That is, we have the image intensity maps:

$$I_l : R_l \rightarrow \mathrm{R} \text{ and } I_r : R_r \rightarrow \mathrm{R} \tag{1}$$

Planes passing through the focal points of the two cameras intersect R_l and R_r in lines called the epipolar lines. The stereo correspondence problem consists in finding a map

$$f_l : R_l \rightarrow R_r \tag{2}$$

such that the match between I_l and $I_r \circ f_l$ is "best-possible". Note that f_l is required to map epipolar lines in R_l onto the corresponding epipolar lines in R_r. Alternatively, we may look for a best-possible map

$$f_r : R_r \rightarrow R_l \tag{3}$$

As is well-known, minimizing

$$\int\int_{R_l} |I_l - I_r \circ f_l|^2 \tag{4}$$

is an ill-posed problem and we must regularize the functional. The two constraints that are usually imposed on f_l (or f_r) are (i) smoothness and (ii) preservation of the order of points along the epipolar lines. In order to consider these constraints and possible discontinuities in the maps, consider first the 1–dimensional case obtained by restricting f_l to an epipolar line L_l. Let f_l map L_l onto the corresponding epipolar line L_r. Adopting the usual form of smoothness constraint as a penalty for large gradients, the first version of the regualized functional is

$$\int_{L_l} \left\{ |f_l'|^2 + \frac{4}{\sigma^2} |I_l - I_r \circ f_l|^2 \right\} \tag{5}$$

where f_l' is the derivative of f_l.

Preservation of the order of points along the epipolar lines means that if a point Q is to the right of a point P in L_l, then $f_l(Q)$ must be to the right of $f_l(P)$ also. In other words, f_l must be a smooth monotonically increasing function; that is, f_l' must be positive everywhere. This leads us to consider the modification

$$\int_{L_l} \left\{ |f_l'|^2 + \frac{1}{|f_l'|^2} + \frac{4}{\sigma^2} |I_l - I_r \circ f_l|^2 \right\} \tag{6}$$

If we start with an initial f_l such that $f_l' > 0$ and apply gradient descent to the above functional, f_l' will continue to remain positive. Notice that if we let $\sigma \rightarrow \infty$, the forcing term (the last term in the integrand) drops out and the limiting functional is minimized by setting $f_l = identity\ map$ so that $f_l' \equiv 1$.

We now turn to the question of half-occlusions and discontinuities. Consider the situation shown in Figure 1. Points B and D map onto the same point in L_l and the portion BCD of the object is occluded from the left eye. Similarly, points E and G map onto the same point in L_r and the portion EFG of the object is occluded from the right eye. Thus f_l must be discontinuous at B_l (= D_l) and be undefined over $E_l G_l$. Conversely, f_r must be discontinuous at E_r (= G_r) and be undefined over $B_r D_r$. The important point is that the segment over which f_l is undefined is exactly the set $L_l \backslash$(range of f_r) and the segment over which f_r is undefined is exactly the set $L_r \backslash$(range of f_l). The situation may be more clearly understood in terms of the graphs of f_l and f_r as shown in Figure 2.

FIGURE 1:

FIGURE 2:

It is convenient for computational purposes to extend f_l and f_r by filling in the dotted lines in Figure 2. (This corresponds to filling in the occluded portions of the object by replacing the portions BCD and EFG by Panum's limiting lines BD and EG.) Note that without other clues, we cannot determine the object geometry at points corresponding to the occluded areas and thus we are justified in assigning some convenient nominal values to the disparity in these areas. After such an extension, f_l and f_r have the same graph. While the discontinuities of f_l occur at points where $f_l' = \infty$, the occluded areas in L_l correspond to the segments where $f_l' = 0$. A similar characterization holds for f_r. Notice that the location of the discontinuity as seen from the left eye is quite different from the location of the discontinuity as seen from the right eye.

The monotonicity constraint excludes situations like the one shown in Figure 3 in which a small thin object, DE, occludes another object, AH, behind it. The situation is such that the segment CF of the object AH, which is *behind* DE is visible from both the eyes. In this case, f_l and f_r, shown in Figure 4, are no longer monotonic. For example, they reverse the order of points C, D and also that of points E, F. If we fill in the dotted lines corresponding to the occlusions, then the occlusions cannot be characterized by the vanishing of the derivative nor do f_l and f_r have the same graph. We can satisfy the monotonicity constraint in this case either by ignoring the segment CF (in which case the graph is similar to the graph shown in Figure 2) or by ignoring the object DE (in which case the graph is just the diagonal). The choice will depend on the relative sizes of the objects and the contrast in their image intensities.

FIGURE 3:

FIGURE 4:

Instead of choosing between f_l and f_r to formulate the model, it is natural to seek a formulation in terms of their common graph. This graph is a function, say d, defined over the SW-NE diagonal Z in Figure 2 such that $|d'| \leq 1$. Let x_l, x_r and x be the coordinates along L_l, L_r and Z with origin at the lower end of Z. Then Z maps onto L_l and L_r via the maps

$$x_l = \frac{1}{\sqrt{2}}(x - d) \text{ and } x_r = \frac{1}{\sqrt{2}}(x + d) \qquad (7)$$

The constraint of preservation of order now takes the form:

$$0 \leq \left(\frac{dx_l}{dx}\right)^2 / \left(\frac{dx_r}{dx}\right)^2 + \left(\frac{dx_r}{dx}\right)^2 / \left(\frac{dx_l}{dx}\right)^2 \leq \infty \quad (8)$$

at points in Z. In contrast to the formulation in [7], the disparity function d is forced to be continuous by the order-preservation constraint. The right and left discontinuities and occlusions now appear symmetrically as the set of points where $|d'| = 1$. This leads us to the functional

$$\begin{aligned} \hat{E}(d) = & \int\limits_{Z \setminus O} \left[\left(\frac{1-d'}{1+d'}\right)^2 + \left(\frac{1+d'}{1-d'}\right)^2 \right] \frac{dx}{\sqrt{2}} \\ & + \frac{4}{\sigma^2} \int\limits_{Z \setminus O} [I_l(x_l(x)) - I_r(x_r(x))]^2 \frac{dx}{\sqrt{2}} \\ & + 4\sqrt{2}\ \gamma length(O) \\ & + 4\nu(number\ of\ components\ of\ O) \\ & \text{where occlusion } O = \{x : |d'(x)| = 1\} \end{aligned} \qquad (9)$$

In the above functional, the constraint on the number of components of O is a form of figural continuity constaint and is included to ensure that the solutions will be sufficiently regular.

If we assume that d is approximately constant over each component of $Z \backslash O$, (that is $|d'| << 1$), then functional (9) may be simplified by using the approximation

$$\left(\frac{1+d'}{1-d'}\right)^2 + \left(\frac{1-d'}{1+d'}\right)^2 \approx 2 + 16(d')^2 \text{ if } |d'| << 1 \tag{10}$$

The resulting approximate functional is very similar in spirit to the functional given in [7] except that d now necessarily has to be continuous and cannot have discontinuities as it does in [7].

As usual with global formulations, the nonlinearity of functional (9) is too complex to permit algorithms for finding exact minimizers and we must look for approximate solutions. For this purpose, consider the pull-backs of the functional $\hat{E}(d)$ to L_l and L_r. For the pull-back to L_l, orient coordinates x_l and x_r from left to right and define the left disparity d_l by the equation

$$f_l = x_l + d_l \tag{11}$$

Note that since we have assumed the image planes to be coplanar, d_l is non-negative everywhere. Then if $|d'| << 1$,

$$\begin{aligned}\left(\frac{1+d'}{1-d'}\right)^2 + \left(\frac{1-d'}{1+d'}\right)^2 &= (f_l')^2 + \left(\frac{1}{f_l'}\right)^2 \\ &\approx 2 + 4(d_l')^2 \text{ if } |d'| << 1\end{aligned} \tag{12}$$

Define a functional $\hat{E}_l(d_l)$ over L_l as follows:

$$\begin{aligned}\hat{E}_l(d_l) = &\int\limits_{L_l \backslash (B_l \cup O_l)} \left\{ (d_l')^2 + \frac{1}{\sigma^2}[I_l - I_r \circ f_l]^2 \right\} dx_l \\ &+ \sum_{x_l \in B_l} \left(\nu + \gamma \left| f_l^+(x_l) - f_l^-(x_l) \right| \right) + \ \gamma length(O_l) \\ &+ \nu(number\ of\ components\ of\ O_l) \\ &\text{where occlusion } O_l = \{x_l : f_l'(x_l) = 0\}\end{aligned} \tag{13}$$

Similarly, define d_r and $\hat{E}_r(d_r)$ after orienting x_l and x_r from right to left. Then

$$\hat{E}(d) \approx C + 4\hat{E}_l(d_l) \approx C + 4\hat{E}_r(d_r) \tag{14}$$

where C is a constant. Hence we consider the functionals $\hat{E}_l(d_l)$ and $\hat{E}_r(d_r)$ in order to find approximate solutions.

Except for the presence of O_l and O_r, functionals $\hat{E}_l$ and $\hat{E}_r$ are very similar in form to the functionals used in segmentation problems which have been studied extensively. Since $B_l \cap O_l$ and $B_r \cap O_r$ are empty, a strategy to find approximate solutions would be as follows: Apply gradient descent alternately to $\hat{E}_l$ and $\hat{E}_r$. That is, alternately minimize $\hat{E}_l$ with respect to d_l and B_l keeping O_l fixed and minimize $\hat{E}_r$ with respect to d_r and B_r keeping O_r fixed. Determination of the discontinuity locus B_l when $\hat{E}_l$ is minimized also determines the occlusion O_r in L_r since $O_r = L_r \backslash$(range of f_l). Similarly the dicontinuity locus B_r determines the occlusion O_l in L_l.

We make three further simplifications before implementing this strategy. First, we follow the example of segmentation functionals and simplify the penalty term for the discontinuity locus by setting $\gamma = 0$. Next, we drop the constraint that f_l and f_r have the same graph. This is based on the heuristic that minimization of $\hat{E}_l$ and $\hat{E}_r$ without this constraint will still produce comparable optimal values for disparities. Finally we simplify the constraint that f_l and f_r must extend continuously across the occlusion sets O_l and O_r respectively such that they are constant over the occlusion set. We relax this requirement by merely extending d_l and d_r across the occlusion set and including the cost of $|d_l'|^2$ and $|d_r'|^2$ over the occlusion set in their respective energy functionals.

The 2–dimensional generalization is straight-forward. The final result is a pair of coupled functionals, E_l and E_r defined over R_l and R_r respectively as follows:

$$\begin{aligned}E_l(d_l) = &\int\!\!\int\limits_{R_l \backslash B_l} \|\nabla d_l\|^2 \\ &+ \frac{1}{\sigma^2} \int\!\!\int\limits_{R_l \backslash O_l} \left\{ [I_l - I_r \circ f_l]^2 \right\} + \nu |B_l| \\ &\text{where occlusion } O_l = R_l \backslash (\text{range of } f_r) \\ &\text{and } |B_l| = \text{ length of } B_l\end{aligned} \tag{15}$$

$$\begin{aligned}E_r(d_r) = &\int\!\!\int\limits_{R_r \backslash B_r} \|\nabla d_r\|^2 \\ &+ \frac{1}{\sigma^2} \int\!\!\int\limits_{R_r \backslash O_r} \left\{ [I_r - I_l \circ f_r]^2 \right\} + \nu |B_r| \\ &\text{where occlusion } O_r = R_r \backslash (\text{range of } f_l) \\ &\text{and } |B_r| = \text{ length of } B_r\end{aligned} \tag{16}$$

3. Implementation by Nonlinear Diffusion

In order to apply gradient descent, we have to compute the first variation of E_l and E_r. However, gradient descent with respect to B_l and B_r presents serious difficulties in implementation [9]. Therefore, we follow the method of Ambrosio and Tortorelli as described in [9] and replace B_l and B_r by variables which are defined *continuously* over R_l and R_r respectively:

$$\begin{aligned} \text{Replace } |B_l| \text{ by } &\frac{1}{2}\int\int_{R_l}\left\{\rho|\nabla w_l|^2+\frac{w_l^2}{\rho}\right\} \\ \text{and } |B_r| \text{ by } &\frac{1}{2}\int\int_{R_r}\left\{\rho|\nabla w_r|^2+\frac{w_r^2}{\rho}\right\} \end{aligned} \tag{17}$$

The values of w_l and w_r range between 0 and 1 and thus may be interpreted as the probability for the presence of a discontinuity in disparity at a point. Alternatively, we may think of w_l and w_r as blurred versions of B_l and B_r with the blurring radius $= \rho$. Since B_l and B_r are now spread out over all of R_l and R_r, it is necessary to modify also the integrals in the functionals since we no longer have the sets B_l, B_r, O_l and O_r. Since d_l is not smoothed wherever B_l is present, a simple choice is:

$$\begin{aligned} \text{replace } &\int\int_{R_l\setminus B_l}||\nabla d_l||^2 \text{ by } \int\int_{R_l}(1-w_l)^2||\nabla d_l||^2 \\ \text{and } &\int\int_{R_r\setminus B_r}||\nabla d_r||^2 \text{ by } \int\int_{R_r}(1-w_r)^2||\nabla d_r||^2 \end{aligned} \tag{18}$$

For dealing with the occlusions, notice that blurring of B_l and B_r implies that f_l and f_r are now continuous everywhere: f_l maps the blurred B_l onto a blurred version of O_l and f_r maps the blurred B_r onto a blurred version of O_r. Since f_l and f_r are required to be monotonic, their inverses are well-defined. Consequently, the pull-back $w_r \circ f_r^{-1}$ will be nearly equally 1 over the occluded part in R_l and nearly equal to 0 away from it. Therefore we

$$\begin{aligned} \text{replace } &\int\int_{R_l\setminus O_l}\{I_l - I_r\circ f_l\}^2 \\ \text{by } &\int\int_{R_l}\{I_l - I_r\circ f_l\}^2\left(1-w_r\circ f_r^{-1}\right)^2 \\ \text{and replace } &\int\int_{R_r\setminus O_r}\{I_r - I_l\circ f_r\}^2 \\ \text{by } &\int\int_{R_r}\{I_r - I_l\circ f_r\}^2\left(1-w_l\circ f_l^{-1}\right)^2 \end{aligned} \tag{19}$$

Appying simultaneous gradient descent to the resulting functionals is equivalent to simultaneous minimization of the following four functionals:

$$\begin{aligned} E_M(d_l) =& \int\int\{(1-w_l)^2||\nabla d_l||^2 \\ &+\frac{1}{\sigma^2}(I_l - I_r\circ f_l)^2\left(1-w_r\circ f_r^{-1}\right)^2\} \\ E_D(w_l) =& \int\int\{\rho||\nabla w_l||^2+\frac{w_l^2}{\rho}+2(1-w_l)^2||\nabla d_l||^2\} \\ &E_M(d_r), E_D(w_r) \text{ obtained by interchanging } l \text{ and } r \end{aligned} \tag{20}$$

We have to minimize $E_M(d_l)$ and $E_M(d_r)$ with respect to d_l and d_r respectively and minimize $E_D(w_l)$ and $E_D(w_r)$ with respect to w_l and w_r respectively. However, note that we could have equally well started with the system of functionals (20) right from the beginning, interpreting them as a general framework. The advantage of this viewpoint is that it provides a more flexible framework for designing stereo algorithms. Adopting this viewpoint, we further modify the functionals $E_D(w_L)$ and $E_D(w_r)$ to represent more faithfully the desired characteristics. In particular, we want the diffusion to proceed unimpeded in the occlusion zones.. However, since disparities have their x–derivatives theoretically equal to ±1 in the occlusion zones, the values of w will be fairly high, thus restricting diffusion. A simple remedy is to replace $||\nabla d_l||^2$ by $||\nabla f_l||^2$ in $E_D(w_l)$ and replace $||\nabla d_r||^2$ by $||\nabla f_r||^2$ in $E_D(w_r)$. An alternate choice is to modify $E_D(w_l)$ and $E_D(w_r)$ as follows: Define

with coordinate x_l oriented from left to right,

$$\left\|\widetilde{\nabla d_l}\right\|^2 = \left\{\max\left(0, \frac{\partial d_l}{\partial x_l}\right)\right\}^2 + \left(\frac{\partial d_l}{\partial y}\right)^2 \tag{21}$$

Replace $||\nabla d_l||^2$ in $E_D(w_l)$ by $\left\|\widetilde{\nabla d_l}\right\|^2$. Similarly define $\left\|\widetilde{\nabla d_r}\right\|^2$ with coordinate x_r oriented from right to left and replace $||\nabla d_r||^2$ in $E_D(w_r)$ by $\left\|\widetilde{\nabla d_r}\right\|^2$.

We now apply gradient descent to these new functionals by calculating their first variations. We get the following system of diffusion equations:

With x_l oriented from left to right, equations for d_l and w_l are

$$\begin{aligned} \frac{\partial d_l}{\partial t} =& \nabla\cdot\left((1-w_l)^2\nabla d_l\right) \\ &+\frac{1}{\sigma^2}[I_l - I_r\circ f_l]\left(\frac{\partial I_r}{\partial x_r}\circ f_l\right)\left(1-w_r\circ f_r^{-1}\right)^2 \\ \frac{\partial w_l}{\partial t} =& \rho\nabla^2 w_l - \frac{w_l}{\rho}+\frac{2}{\nu}(1-w_l)||\nabla f_l||^2 \end{aligned} \tag{22}$$

With x_r oriented from right to left, equations for d_r and w_r are

$$\begin{aligned}\frac{\partial d_r}{\partial t} =& \nabla \cdot \Big((1-w_r)^2 \nabla d_r\Big) \\ &+ \frac{1}{\sigma^2}[I_r - I_l \circ f_r]\left(\frac{\partial I_l}{\partial x_l}\circ f_r\right)\left(1 - w_l \circ f_l^{-1}\right)^2 \\ \frac{\partial w_r}{\partial t} =& \rho\nabla^2 w_r - \frac{w_r}{\rho} + \frac{2}{\nu}(1-w_r)\|\nabla f_r\|^2\end{aligned} \tag{23}$$

The boundary conditions are as follows. At the ends of each epipolar line, we require that d_l and d_r must be zero. There is no loss of generality by this assumption. If the disparities in fact are not zero at the end-points, then the situation is represented by making disparity discontinous at such end-points and having a corresponding occlusion area touching the end-points in the image of the other eye. The remaining boundary conditions are the homogenous Neumann boundary conditions; that is, the normal derivative of the variables is set equal to 0.

It is interesting to note that the diffusion equations for d_l and d_r behave like a "feature-based" algorithm. To see this, consider the 1–dimensional case and suppose that I_l and I_r are piecewise constant. Then the derivatives for I_l and I_r and hence the forcing terms in the diffusion equations are concentrated at the discontinuities (that is, the edges) of I_l and I_r. The diffusion process linearly interpolates the disparity values between its values at the sparse edge points. It is possible to change the smoothing term in the diffusion equations to have alternate forms of interpolation. For example, if we replace $||\nabla d_l||^2$ by $||\nabla d_l||^2/d_l^2$ and $||\nabla d_r||^2$ by $||\nabla d_r||^2/d_r^2$, we get interpolation by exponential curves. However adding d_l^2 and d_r^2 to the integrands of $E_l(d_l)$ and $E_r(d_r)$ respectively, presents a dilemma. On one hand, the presence of these terms will cause disparity to decay exponentially between the feature points even though it might in fact be constant. On the other hand, their presence lends stability to the algorithm. *It might be advisable to add these terms with a relatively small coefficient.* At present, we have omitted these terms from the algorithm so as to avoid having one more parameter and plan to investigate this issue in the future.

The diffusion equations are implemented by a simple finite-difference scheme. After each time step, d_l and d_r are adjusted so that f_l and f_r remain monotonic. This is a simple calculation. For example, after discretization,

With coordinate x_l oriented from left to right,
let $d_l^*(x_{l_0}) = \max\left(0, d_l(x_{l_0})\right)$
and for $i = 1, 2, \ldots$ let
$$d_l^*(x_{l_i}) = \max\left(x_{l_{i-1}} + d_l^*\left(x_{l_{i-1}}\right), x_{l_i} + d_l(x_{l_i})\right) - x_{l_i} \tag{24}$$

Let $f_l^* = x_l + d_l^*$. Then, f_l^* is monotonically nondecreasing. Compute d_r^* and f_r^* similarly after orienting the coordinate x_r from right to left. After each time step, replace d_l by d_l^* and d_r by d_r^*. Alternatively, we may leave d_l and d_r unadjusted after each time step (but still enforcing the restriction that they must be nonnegative), but use d_l^* and d_r^* to calculate only the occlusion terms $w_r \circ {f_l^*}^{-1}$ and the $w_l \circ {f_r^*}^{-1}$ as in [7]. In all the test cases, both schemes produced indistinguishable results. Note that $f_l(x_{l_i})$ and $f_r(x_{r_i})$ will not in general be integer-valued and the terms like $I_r \circ f_l$ are calculated by interpolation between the lattice points.

For illustration, we present an example[2] shown in Figure 5. Each image consists of 256×256 8–bit pixels. The initial value of the disparity was set equal to zero. Figure 6 portrays the numerically computed disparity values. The brighter areas indicate the higher disparity values. The maximum value of disparity is 76. It increases continuously from top to bottom along the stripes (roughly from 30 to 67) and is nonuniform within the ball also. Figure 7 depicts the values of w_l and w_r. The bright areas indicate the blurred version of the discontinuity locus of d_l and d_r while the dark areas correspond to occlusion. The uniform gray indicates approximately constant disparity. Since the disparity is set to zero at the left and right boundaries, we have large discontinuities there due to the presence of unmatched areas near the boundaries. Because the algorithm works by diffusion, the time it takes to converge depends on how far it has to diffuse the disparity values. In order to speed up the convergence and avoid local equilibriums, the following procedure was adopted. During the early stages, the images were smoothed (smoothing radius 20 pixels) and we set $\rho = 5$ pixels, $\sigma = 5$, $\nu = \infty$. As the disparity values stabilized, the smoothing radius was reduced to zero in stages. Finally, ν was decreased until the highest values of w_l and w_r were nearly 1.

FIGURE 5: IMAGES

[2] Grateful acknowledgement is made to Professor Jitendra Malik of UC, Berkeley for providing these pictures.

FIGURE 6: LEFT AND RIGHT DISPARITIES

FIGURE 7: w_l, w_r

As a second example, we present a random dot stereogram shown in Figure 8 in which each “dot” is a 4×4 pixel in size. Each image again consists of 256×256 pixels. The left image is obtained from a random-dot image by shifting the central 96×96 pixel square to the left by 8 pixels and shifting the 176×176 square surrounding it by 4 pixels to the left. The right image is obtained by shifting the same squares to the right by the same amounts. Thus the reconstructed object is a square “wedding cake”. Figure 9 portrays the disparity while Figure 10 shows w_l and w_r.

FIGURE 8: IMAGES

FIGURE 9: LEFT AND RIGHT DISPARITIES

FIGURE 10: w_l, w_r

4. References

1. Dhond, U.R. and Aggarwal, J.K.: ”Structure from Stereo - A Review”. IEEE Trans. Systems, Man, and Cybernetics, v.19, n.6, 1989.
2. Grimson, W.E.L.: ”A Computer Implementation of a Theory of Human Stereo Vision”. Philos. Trans. Roy. Soc. London, Series B292, v.292, 1982.
3. Geiger, D., Ladendorf, B. and Yuille, A.: ”Occlusions and Binocular Stereo”. EECV, 1991.
4. Grimson, W.E.L.: Computational Experiments with a Feature Based Stereo Algorithm”. IEEE Trans. PAMI, v.7, Jan. 1985.
5. Hoff, W. and Ahuja, N.: ”Surfaces from Stereo: Integrating Feature Matching, Disparity Estimation, and Contour Detection”. IEEE Trans. PAMI, v.11, n.2, 1989.
6. Marr, D. and Poggio, T.: ”A Computational Theory of Human Stereo Vision”. Proc. Royal Soc. London, v.B204, 1979.
7. Mumford, D. and Belhumeur, P.N.: ”A Bayesian Treatment of the Stereo Correspondence Problem Using Half-Occluded Regions”. Proc. IEEE CVPR, 1992.
8. Mumford, D. and Shah, J.: ”Boundary Detection by Minimizing Functionals, I”. Proc. IEEE CVPR, 1985.
9. Shah, J.: ”Segmentation by Nonlinear Diffusion”. Proc. IEEE CVPR, 1991.