Tabu Search for Disparity Estimation

Robert Laganière* Amar Mitiche**
* Dept. of Computer Science
University of Ottawa, Ontario, CANADA, K1N 6N5
** INRS-Télécommunications
16, Place du commerce, Verdun, Québec CANADA, H3E 1H6

Abstract

In this paper, we propose to perform disparity estimation by minimizing an objective function that uses a fairly simple regularization function to control smoothing interaction of pixel neighbors where discontinuities are implicitly addressed. To accomplish the minimization of this objective function, we use the Tabu Search technique. Tabu search adapts to the particular structure of the problem it tries to solve, and thus performs an intelligent exploration of the state space.

1 Introduction

Stereopsis is the ability to recover the three-dimensional structure of a scene from images obtained from two sensors positioned according to a known configuration one with respect to the other [1]. A point in one of the images can be associated with a point in the other image because they are the images of the same point in the scene. Disparity vector $\delta(x,y)$ expresses the fact that the pixel located at (x,y) in one image is at $(x+\delta_x(x,y), y+\delta_y(x,y))$ in the other image. When disparity is known, 3-D structure can be obtained by simple triangulation. Computation of disparity, is a difficult problem although epipolar geometry establishes a number of constraints [2].

One possible approach to this problem is to express the problem of disparity estimation with a criterion function whose global minimum corresponds to the desired disparity field [3]. Typically, this objective function includes two terms: one makes the solution conform to the original image and the other is a regularization term for local coherence of the solution [4]. This regularization must, however, tolerate discontinuities in the disparity field which means that smoothing must be applied only over homogenous regions [5]. This kind of objective function can be written as:

$$F(\mathcal{D}) = \sum_{\delta_i \in \mathcal{D}} \psi(\delta_i, \mathbf{I}_r, \mathbf{I}_\ell) + \lambda \sum_{c \in C} \phi(\delta_j \,/\, j \in \mathbf{c}) \quad (1)$$

$\mathcal{D}$ is the disparity field, i.e. the set of all disparity vectors to be estimated. $\mathbf{I}_r$ and $\mathbf{I}_\ell$ are the right and left images; these constitute the observation. The function $\psi()$ measures the consistency of a candidate solution with the observation. $\phi()$ is a regularization function that defines how points of a particular clique $\mathbf{c}$ interact with each other; it measures how spatially coherent the proposed solution is. In this context, suitable estimation will be achieved if 1) $F(\ )$ is easily computable, 2) $\psi(\ )$ is such that it promotes plausible solutions with respect to the observations, 3) smoothing properties of $\phi(\ )$ preserve discontinuities and 4) minimization can be performed at a reasonable computational cost.

Under the assumption that the disparity field corresponds to the state of minimum energy of the objective function, it is essential to proceed to the global minimization of function (1). Simulated annealing in the context of MRF modeling has been widely used to perform this kind of task [6]-[9]. This technique can, theoretically, reach the global minimum of complex function but at a prohibitive computational cost. A technique called graduated non convexity (GNC) has also been used [10][11]. The method consists in successively applying descent algorithms on different convex approximations of the objective function that gradually converge to $F(\ )$. GNC can be seen as a deterministic annealing approach and its complexity strongly depends on the adopted mathematical formulation.

In this paper, we propose to perform disparity estimation by minimizing an objective function that uses a fairly simple regularization function to control smoothing interaction of pixel neighbors where discontinuities are implicitly addressed. To accomplish the minimization of this objective function, we use the Tabu Search technique [12][13]. This technique has shown its ability to obtain high quality solutions with modest computational effort in many operation research problems [13][14], dominating alternative methods, including Simulated Annealling. It has also been used in neural networks [15] and for image restoration [16]. The strength of the Tabu search technique comes from the fact that Tabu search adapts to the particular structure of the problem it tries to solve, and thus performs an intelligent exploration of the state space. This is in contrast with stochastic methods where a random exploration of the space is performed. Since Tabu search is one of the key aspects of the present work, we will first describe the basic concepts underlying this technique before presenting its application to the disparity estimation problem.

2 Tabu Search

Let $F(\mathbf{x})$, $\mathbf{x} = [x_1, ..., x_N]$, be an objective function that one wishes to minimize. And let the state space $\mathcal{X}$ be the set of all feasible solutions (i.e. $\mathbf{x} \in \mathcal{X}$). The goal is to move from one state to another in order to iteratively reach the global minimum of the function. The set of all allowed moves from a state $\mathbf{x}_i$ is designated by $\mathcal{D}_i$. The fact that state $\mathbf{x}_j$ is reached from state $\mathbf{x}_i$ by making the move $\mathbf{d} \in \mathcal{D}_i$ is written

$$\mathbf{x}_j = \mathbf{x}_i + \mathbf{d} \quad (2)$$

The Tabu procedure consists of choosing the *best* move $\mathbf{d}_i^o \in \mathcal{M}_i$, i.e., the one that will produce the highest reduction of the objective function, that is:

$$F(\mathbf{x}_i + \mathbf{d}_i^o) \leq F(\mathbf{x}_i + \mathbf{d}) \quad \forall \mathbf{d} \in \mathcal{D}_i \quad (3)$$

A gradient descent is thus performed: iteratively, the procedure will eventually bring the objective function to a local minimum. In such a situation, the *best* move will cause an increase in the value of the objective function. The Tabu search technique allows such a move when it is not possible to do better.

A key aspect of the Tabu Search is that, in order to avoid cycling (i.e. returning to an already visited state), each time a move to a new state is performed, all complementary moves (i.e., those that cancell this move) enter a *Tabu* status. This Tabu status remains valid for a certain set amount of time. In general, a move that is in a Tabu status will be forbidden. The fact that a move $\mathbf{d}$ is in a Tabu status ($\mathbf{d} \in \mathcal{T}$) for a certain amount of time constitues the short term memory. It forbids backward moves and thus allows the search to escape from a local minimum. Each time a local minimum is reached, it is compared to the best local minimum found so far, and if this new minimum is lower, then it becomes the new best current minimum.

However, in some circumstances, it may be advantageous to override the Tabu status of a given move. This may happen, for example, if by making a move currently Tabu, a lower value than the best current minimum would be reached or a large decrease of the objective function would be obtained. A long-term memory is therefore introduced, expressed by an *aspiration level condition* to allow a move despite its Tabu status if making this move potentially results in a reduction of the objective function. Aspiration level condition are formulated using boolean functions, $A(\mathbf{x}, \mathbf{d})$, that evaluate to true when the Tabu status should be overriden. The *best* move is then defined as follows:

$$\begin{aligned} F(\mathbf{x}_i + \mathbf{d}_i^o) &\geq F(\mathbf{x}_i + \mathbf{d}) \\ \forall \mathbf{d} &\in \mathcal{D}_i \ / \ A(\mathbf{x}_i, \mathbf{d}) \ OR \ \mathbf{d} \notin \mathcal{T} \end{aligned} \quad (4)$$

The stopping criterion of the search is generally based on a maximum number of iterations until no improvement of the current best minimum is obtained.

2.1 Algorithm

Below is a general description of Tabu Search; details of the algorithm can be found in [12][13]. The search can be initialized with a rough or random estimate:

1. Choose the best move, in the current set $\mathcal{D}$, that is not in a Tabu state unless an aspiration level condition allows its selection (equation (4)).
2. Apply the selected move and update the objective function value.
3. If the current value of the objective function is less than the best current minimum, then memorize the present state as the new current minimum.
4. Update the list of moves that have Tabu status.
   (a) Put all complementary moves to the selected move in a Tabu status.
   (b) Remove the Tabu status of the moves that have been Tabu for a sufficient number of iterations.
5. If the number of iterations since the last improvement of the best current minimum exceeds a preestablished threshold, then stop. Otherwise execute another iteration by returning to 1.

3 Disparity estimation

Tabu search will be used to minimize an objective funtion of disparity, $F(\mathcal{D})$. The disparity field consists of N disparity vectors occuring between the observations $\mathbf{I}_\ell$ and $\mathbf{I}_r$ (left and rigth images). Without loss of generality, we consider a particular configuration, where the two cameras are separated by a horizontal translation as illustrated in figure 1. With this configuration, epipolar lines are parallel and horizontal such that disparity is strictly horizontal, i.e., $\delta = (\delta_x, 0)$. This is not a restrictive case since image pairs can be tranformed to obtain horizontal epipolar by a simple process of image rectification [17][18].

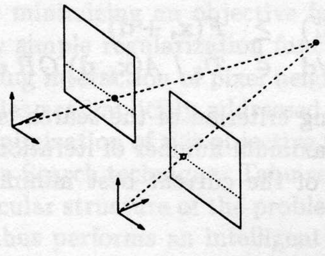

Figure 1. A parallel configuration of a stereoscopic system.

In the context of disparity estimation, we now define *consistency* and *coherence*.

3.1 Consistency

Consistency is based on correlation. More precisely, the consistency corresponding to the disparity value δ_i for point i is evaluated from the sum of square difference taken over a window $\mathcal{W}_i$:

$$\psi(\delta_i, \mathbf{I}_r, \mathbf{I}_\ell) = \sum_{(x,y)\in\mathcal{W}_i} (I_r(x,y) - I_\ell(x+\delta_i, y))^2 \quad (5)$$

where $I(x,y)$ is the pixel intensity at site (x,y).

While it should be possible to find a reliable match for most image points, it is however likely to happen that no interesting match can be found for some of them. This can occur for regions of occlusion, poor illumination conditions, important perspective distorsion, high level of noise, etc. Under these circumstances, the consistency term will give rise to high values for any disparity value. Because of the bad influence that can have these particular points, it would be preferable to discard the corresponding consistency terms. The reliability of a consistency term will, therefore, be evaluated from its minimum value over the allowed range of disparity:

$$\psi^i_{min} = \max_{\delta_i} \psi(\delta_i, \mathbf{I}_r, \mathbf{I}_\ell) \quad (6)$$

In order to normalize this reliability measure to range within the interval $[0,1]$ (1 is most reliable), the following function is applied:

$$f_i(\Psi_{min}) = \frac{1}{1 - \exp^{-\tau^2(\Psi^i_{min} - \theta)}} \quad (7)$$

The parameter θ is a threshold to mark the transition between reliable and unreliable terms occurs and τ controls the slope of this transition. This function is shown in figure 2. A similar function has been proposed in [19].

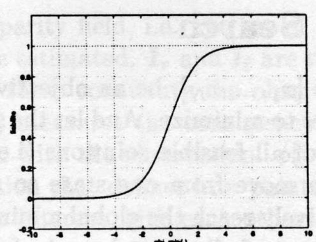

Figure 2. Graphical representation of f_i with $\tau = 1$ and $\theta = 0$.

Since consistency involves only the disparity at a given point, reliability can be estimated once, independently at each site, at the begining of the process. Disparity at site where the consistency term has been invalidated by the reliability measure will therefore be estimated only through local coherence.

3.2 Coherence

To define a reliable regularization term, we make the following observations. When all disparity vectors of a given clique belong to the same region (usually to some parts of the same object) then these vectors should be similar. The regularization term should therefore penalize solutions where dissimilar vectors occur. However, if these vectors correspond to different regions then dissimilarity should not be excessively penalized. This leads to the following function for $\phi()$ applied on cliques of dimension 2:

$$\phi(\delta_i, \delta_j) = -\exp^{-\beta^2(\delta_i - \delta_j)^2} \quad \forall(\delta_i, \delta_j) \in \mathcal{C} \quad (8)$$

This function, the graph of which is shown in Figure 3, has been initially proposed in [20] in the context of image restoration. For small differences, $\phi(\ )$ is similar to s^2 (the dotted curve in Figure 3) and this remains true until a certain threshold from which the penalty stabilizes.

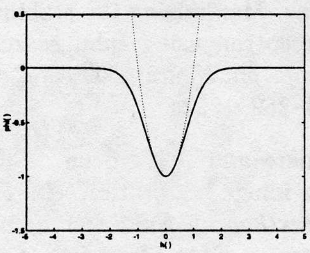

Figure 3. Graphical representation of $\phi(s)$ with $\beta = 1$, The dotted curve is $g(s) = s^2 - 1$.

3.3 Minimization procedure

Minimization of the resulting $F(\mathcal{D})$ is formulated as an integer programming problem: each δ_i can take integer values from 0 to δ_{max}.

Let us define the contribution C_i of a pixel i: this is the sum of all terms in the objective function where the term δ_i appears:

$$C_i = \sum_{(x,y)\in\mathcal{W}_i} (I_r(x,y) - I_\ell(x+\delta_i, y))^2 + 2\lambda \sum_{c\in\mathcal{C}\ /\ \delta_i\in c} - \exp^{-\beta^2(\delta_i-\delta_j)^2} \quad (9)$$

Note that the sum for $i = 1,..,N$ of the values of all contributions is not equal to the value of the objective function. Suppose that we are in a given state $\mathbf{x}$. We define an allowed move as one for which there is a change of a disparity value by 1 or 2. This gives $4N$ possible moves. We select the one having the smallest contribution C_i' that is not in a Tabu status. The corresponding reduction in the objective function is simply $C_i' - C_i$. When a move is completed, the complementary moves become Tabu. This means that a backward move to a previous disparity value cannot occur during its Tabu period unless 1) this move makes the objective function smaller than the best current minimum 2) the new contribution obtained by making this move is less than the smallest contribution obtained so far for the point under consideration.

Minimization should be done over the whole image. However, considering the fact that single disparity vector has a limited influence on the global objective function, spatially local minimization is appropriate. We have performed minimization over 3×3 windows which means that each minimization involves 9 variables. Moreover, if we carefully choose the configuration of these windows, then several minimizations can be carried out in parallel (see Figure 4). In order to propagate results of a spatially local minimization, these parallel minimizations are repeated with 16 different window configurations such that each pixel site is once the center of one window. With such a procedure, most sites are subjected to 9 minimization procedures.

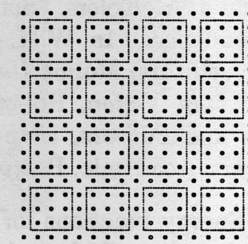

Figure 4. A parallel window configuration: Tabu search can be independently applied on each of these window.

4 Results

The proposed approach has been tested first on a pair of synthetic images (Figure 5). Reliability has been evaluated with $\theta = 2000$ and $\tau = 0.044$. Figure 5(c) shows how the values of f_i are distributed over the image. The parameters associated with the coherence term have been fixed to $\lambda = 800$ and $\beta = 0.05$. Minimization is performed with a maximum of 75 iterations per window and a duration of 15 for a Tabu status. δ_{max} has been fixed to 70.

In order to appreciate the quality of the estimated disparity field found by Tabu search, the results are presented in two different ways. First, the disparity field is used to compute the corresponding 3-D structure; this structure is shown in figure 5(d). Second, we have interpolated the right image by disparity compensation; this latter is shown in figure 5(e) together with the displaced image difference error in figure 5(f).

The same procedure has been applied on the image pair shown in figure 6(b). Resuts are presented in Figure 6(c)-(e).

5 Conclusion

We have presented a procedure to perform disparity estimation based on functional modelization of the problem. One main aspect of this work, is the use of the Tabu Search technique for functional minimization. We believe that this method exhibits an interesting potential to solve problems in computer vision where a global minimization is required.

References

[1] U.R. Dhond and J.K. Aggarwal. Structure from stereo - a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6):1489–1510, 1989.

[2] M. Jenkin and P.A. Kolers. Some problems with correspondence. In W.M. Martin and J.K. Aggarwal, editors, *Motion Understanding: Robot and Human Vision*, Robotics: Vision, Manipulation and Sensors, chapter 8, pages 269–295. Kluwer Academic, Boston, 1988.

[3] A.L. Yuille. Energy functions for early vision and analog networks. *Biological Cybernetics*, (61):115–123, 1989.

[4] K. Prazdny. Detection of binocular disparities. *Biological Cybernetics*, 52:93–99, 1985.

[5] D. Terzopoulos. Regularization of inverse visual problems involving discontinuities. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 8(4):413–424, July 1986.

[6] S. Geman and D. Geman. Stochatic relaxation, gibbs distribution and the bayesian restoration of images. *IEEE trans. on Pattern Analysis and Machine Intelligence*, (6):721–741, 1987.

[7] F. Heitz, P. Perez, and P. Bouthemy. Multiscale minimization of global energy functions in some visual recovery problems. *CVGIP: Image Understanding*, 59(1):125–134, January 1994.

[8] S.T. Barnard. Stochastic stereo matching over scale. *Int. Journal of Computer Vision*, 3:17–32, 1989.

[9] D. De Vleeschauwer. An intensity-based, coarse-to-fine approach to reliably measure binocular disparity. *CVGIP: Image Understanding*, 57(2):204–218, March 1993.

[10] A. Blake and A. Zisserman. *Visual reconstruction*. MIT Press, Cambridge, MA, 1987.

[11] A. Blake. Comparaison of the efficiency of deterministic and stochastic algorithms for visual reconstruction. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(1):2–12, January 1989.

[12] F. Glover. Tabu search- part i. *ORSA Journal on Computing*, 1(3):190–206, 1989.

[13] F. Glover. Tabu search- part ii. *ORSA Journal on Computing*, 2(1):4–31, 1990.

[14] A. Hertz D. de Werra. Tabu search techniques. *OR Spektrum*, (11):131–141, 1989.

[15] E. Mayoraz. Maximizing the stability of a majority perceptron using tabu search. In *Int. Joint Conf. on Neural Network*, volume II, pages 254–259, 1992.

[16] R. Laganière and A. Mitiche. Parallel tabu search for image restoration. In *Proc. of the IEEE Workshop on Nonlinear Signal and Image Processing*, Neos Marmaras, Greece, June 1995.

[17] N Ayache. *Construction et fusion de représentations visuelles. Applications à la robotique mobile.* PhD thesis, Université de Paris-Sud, Centre d'Orsay, 1988.

[18] D. Pelé. Calibration d'un système de prises de vues stéréoscopiques. géométrie épipolaire. rectification d'images. Technical Report PFI/IPT-MTS/01/90/DP, CCETT, Cesson-Sévigné, France, 1990.

[19] P. Anandan. Computing dense displacement fields with confidence measures in scene containing occlusion. In *Intelligent Robots and Computer Vision*, volume 521, pages 184–194. SPIE, 1984.

[20] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 14(3):367–382, 1992.

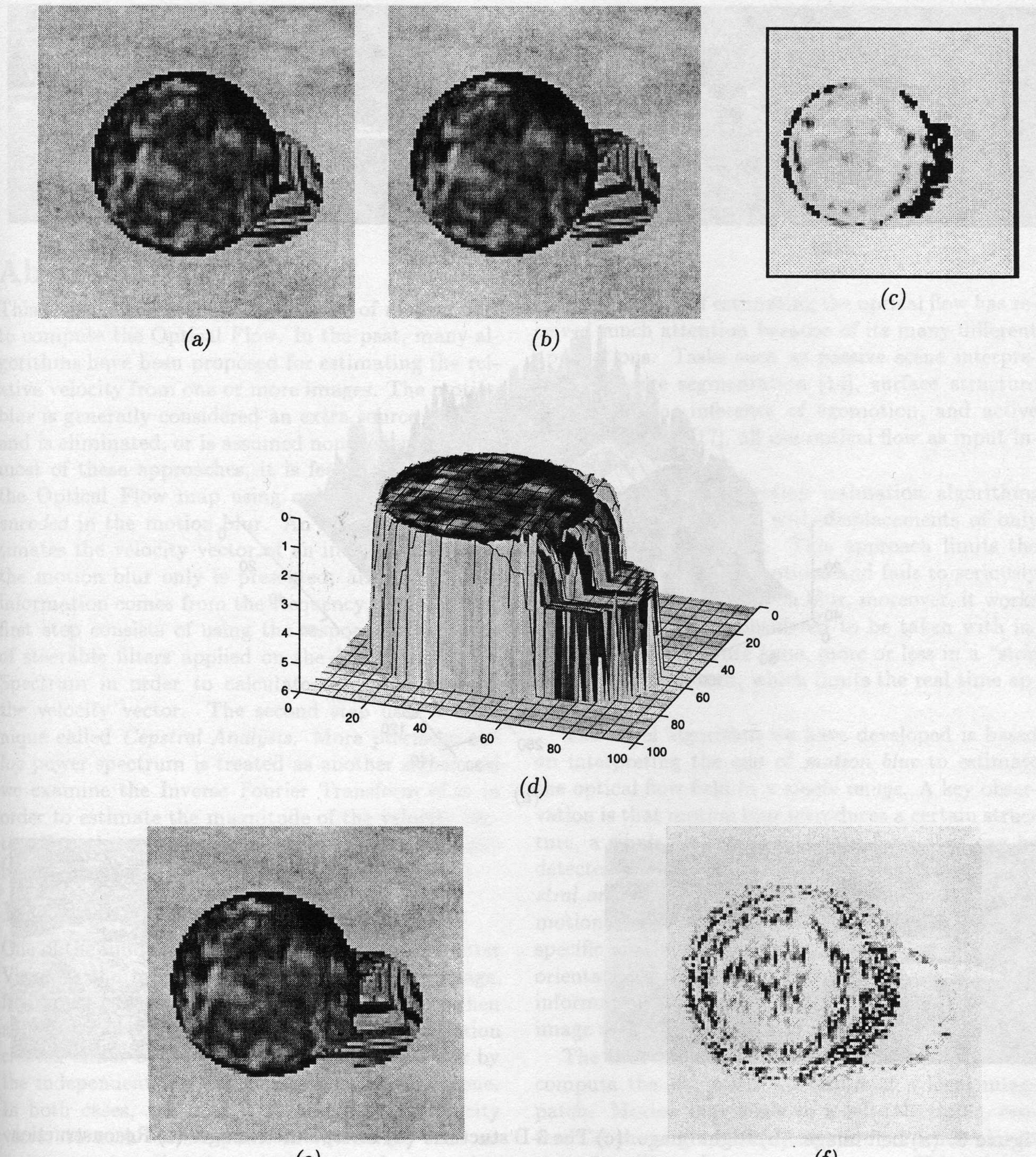

Figure 5. (a) Left image. (b) Right image. (c) Distribution of the reliability (black: $f_i = 0.0$, white $f_i = 1.0$) (d) The 3-D stucture. (e) Interpolated image. (f) Reconstruction error.

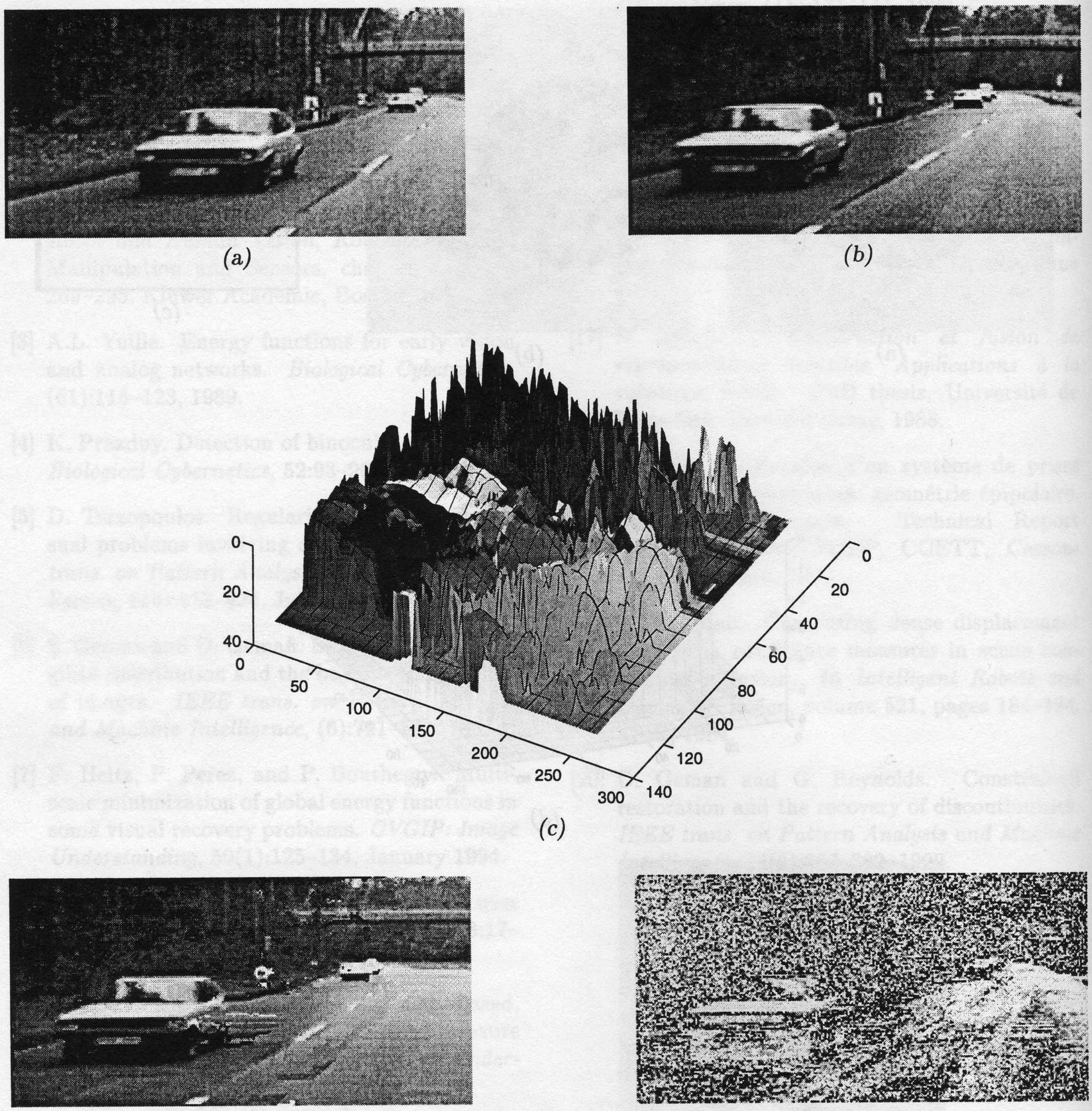

Figure 6. (a) Left image. (b) Right image. (c) The 3-D stucture. (d) Interpolated image. (e) Reconstruction error (white corrspond to square error= 0).