A STEREO MATCHING METHOD USING THE GREEN'S FUNCTION

Susumu Hattori, Shunji Murai

Institute of Industrial Science,
University of Tokyo
7-22-1, Roppongi, Minatoku,
Tokyo, 106,Japan

Hitoshi Ohtani

Topcon Co. Ltd.
75-1, Hasunumacho, Itabashiku,
Tokyo, 174,Japan

ABSTRACT

This paper proposes a new idea of a stereo matching technique based on the minimal potential energy criterion. The method is concerned with the Marr's human stereopsis model. According to his model, the real 3-D world is recognized by interpolating identified disparities. We think, however, matching and interpolation should be done simultaneously. Our method realizes this by use of the Green's function, and is indicative of human and machine matching strategic similarity. It is shown that the potential energy minimum criterion is reduced to minimizing a quadratic form, and minimization process can be localized. The method consists of two parts; 1) Picking up disparity candidates using a support function. 2) Selection of true disparities by searching for a surface passing through disparities and being the flattest.

INTRODUCTION

According to the well-known Marr's human stereopsis model [1], in the human eye-brain system, retinal images are narrow-band-pass-filtered with the Laplacian of Gaussian (LOG) filters with different scales, and extracted left and right edges are matched using the coarse-to-fine principle (We term a zero-crossing 'edge' in this paper). The 3-D continuous world is recognized by interpolating identified disparities.

In his model matching edges is apparently only the resolution problem. He suggested that in matching two conditions must be satisfied; uniqueness and continuity. But since his model consists of local and sequential processes, these conditions are hard of being stably and consistently satisfied.

To satisfy the conditions, some kind of global optimization is necessary. We propose in this paper a new matching method based on the potential energy minimum criterion, which match and interpolate disparities at the same time. From our instinct it seems more natural and compatible to the actual human visual system than the Marr's sequential model.

Our method consists of two parts. The former half is to pick up disparity candidates for each edge by use of a support function [2], and the latter half is to select true disparities by search for the flattest (potential energy minimum) surface which passes through one disparity candidate at each edge. This paper first treats the latter half for the convenience of discussion.

INTERPOLATION OF A SURFACE USING THE GREEN'S FUNCTION

Grimson [3] suggested with reference to the Marr's model that disparities which have been identified should be interpolated so as to minimize

$$U = \frac{1}{2}\iint (u_{xx}^2 + 2u_{xy}^2 + u_{yy}^2)\,dxdy, \tag{1}$$

which means the surface should be as smooth as possible over the object space. Eq.(1) corresponds to the functional of a bending plate with the Poisson's ratio being zero. We think, however, the following functional being minimized is more suitable;

$$U = \frac{1}{2}\iint (u_x^2 + u_y^2)\,dxdy. \tag{2}$$

For Eq.(2) allows steep changes in disparities at edges, and yet this can not be realized with Eq.(1). Eq.(2) produces the flattest surface, and physically is concerned with the membrane equation. A typical differential equation of a membrane and its functional are

$$\nabla^2 u = -f(x,y),\; E = \frac{1}{2}\iint (u_x^2 + u_y^2 - fu)\,dxdy \tag{3}$$

with f(x,y) being a distribution of an external force and u(x,y) being a displacement of the membrane.

We use Eq.(3) for interpolating disparities, in which a disparity corresponds to a displacement of the membrane. We assume that a disparity is caused by a virtual force working at the edge. With a simple circular boundary **B** on which disparities assume zero (Fig.1), the solution of Eq.(3) is expressed by

$$u(x,y) = -\iint_A f(\xi,\eta)\, G(x,y,\xi,\eta)\, d\xi\, d\eta, \qquad (4)$$

where $G(x,y,\xi,\eta)$ is a Green's function [4];

$$G(x,y,\xi,\eta) = \frac{1}{4\pi} \ln \left[\frac{(x-\xi)^2 + (y-\eta)^2}{R^2 + \frac{1}{R^2}(x^2+y^2)(\xi^2+\eta^2) - 2(x\xi+y\eta)} \right] \qquad (5)$$

The Green's function can be recognized as an influential function for a unit impulse of a force.

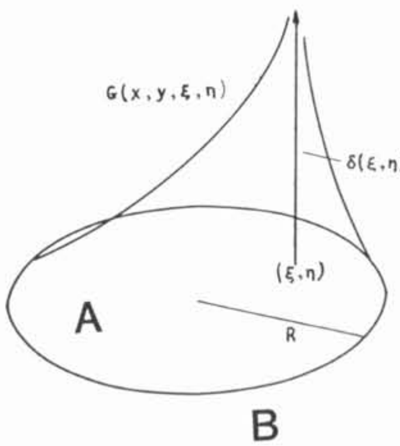

Fig.1 Green's function

By substituting Eq.(5) into Eq.(2), we obtain

$$U = \iint_A \iint_A G(x_1,y_1,x_2,y_2)\, f(x_1,y_1)\, f(x_2,y_2)\, dx_1\, dy_1\, dx_2\, dy_2 \qquad (6)$$

Then we convert Eq.(6) into a digital form suitable for computer processing. We number edges column-scan-wise as $i=1,2,\ldots,N$ at which disparity candidates exist. Let image coordinates of the ith edge be denoted by x_i, y_i, its disparity by u_i, an unknown force loaded on the ith edge by $F_i = F(x_i, y_i)$ (a distribution f is replaced by a force F for a digital image), and the value of Eq.(5) at edge i when an impulse is loaded on edge j by $h_{ij} = G(x_i,y_i,x_j,y_j)$. In addition we introduce following vector and matrix notations;

$$\mathbf{u} = (u_1, u_2, \cdots, u_n)^T,$$

$$\mathbf{F} = (F_1, F_2, \cdots, F_n)^T, \quad H = (h_{ij})$$

Then Eqs.(4) and (6) are converted to digital forms;

$$\mathbf{u} = H\mathbf{F} \quad (4') \qquad U = \mathbf{F}^T H \mathbf{F} = \mathbf{u}^T H^{-1} \mathbf{u} \quad (6')$$

Thus selecting the disparities is reduced to minimizing the quadratic form. Note the following;

(1)The space of admissible solutions is not convex. Therefore we can not use conventional descending approaches. In this paper optimization was done by trying every possible combination of disparity candidates.

(2)Actually an influential range of the Green's function is not wide. If a unit force is loaded on a center of the domain **A** in Fig.1, with the radius being 128 pixels, the function decreases to 1/10 of the peak value in about 100 pixels. This means that the global solution of Eq.(6') can be approximated by a local solution in every finite domain. This implies the reason why a man can recognize the object space stereooptically by a continuous local matches along with eye movement.

(4)Diagonal elements in the matrix H become infinite, because of the singularity of the Green's function. To vanish the singularity, we modify the definition of the function from an influential surface which is generated by an impulse force to one generated by a finite force loaded on a small area. For a square pressure working on a pixel p(x,y), the modified Green's function is expressed by

$$G'(x,y,\xi,\eta) = \frac{1}{\varepsilon^2} \int_{-\varepsilon/2}^{\varepsilon/2} \int_{-\varepsilon/2}^{\varepsilon/2} G(x,y,\xi+\alpha,\eta+\beta)\, d\alpha\, d\beta \qquad (7)$$

where the width of a pressure, ε, was set to 1/10 of the pixel width. The values were evaluated by numerical integration.

One attention must be paid here that in every candidate set one true disparity always must be involved. Otherwise as a result of matching a false disparity is selected where no true disparity is involved.

SUPPORT FUNCTION

Prazdny [2] proposed a concept of a support function to match edges. His idea is similar to relaxation[6] and gives candidates determining locally flattest surfaces. Any disparity candidate, i_k, of an edge i receives support from any surrounding edge j (j_1, j_2..., are its disparity candidates). Prazdny evaluated the amount of support by the function

$$S_P(r_j, \Delta d_j) = \frac{1}{\sqrt{2\pi}\, r_j} \exp\left[-\frac{c}{2}\left(\frac{\Delta d_j}{r_j}\right)^2\right], \qquad (8)$$

where c is a parameter, r_j denotes the distance between i and j and Δd_j denotes the minimal possible disparity difference between i_k and $j_{k'}$. The total support is given by a summation of contribution of surrounding edges. According to our experiments Prazdny's function worked relatively well, but support values vary with edge density, which make quantitative judge less significant. Hence we used the following modified support[6].

$$S_H = \sum_j S_P(r_j, \Delta d_j) \Big/ \left(\sum_j \frac{1}{\sqrt{2\pi}\, r_j}\right). \qquad (9)$$

S_H assumes unity if correspondence of edges is perfect. We found with this

function correct candidates were successfully picked up, if they exist. But at some edges where no true disparity exists, it still picks up false similar candidates. To evade this, the following additive procedures were proved useful.
1)Bi-directional searches (i.e., independent searches from the left image to the right, and from the right to the left) are done and only commonly picked-up disparities are left.
2)A search range smaller than Marr suggested[1] is set. In our algorithm images are reduced by sampling after LOG-filtering every 1,2 and 4 pixels to save memories and computation time. And the constant search range $K=\pm 3$ pixels is used.

EXPERIMENTS AND DISCUSSION

Fig.2 shows a pair of stereo images of 512x512 pixels in size. Distortions due to orientation were already rectified. They contain typical difficulties in matching, i.e., occlusion by buildings and a wall, noises by halasion on roofs and periodical patterns in parking lots. Fig.3(a),(b) and (c) shows y-directional edges extracted from images obtained by LOG-filtering with peak frequencies, ω_p= 1/8(rad/pixel), ω_p= 1/4 and ω_p = 1/2 respectively.
We notice the following: No features exist in a shady space between two high buildings, so that we have no means to know its height. When we try to fuse, however, we feel it as high as the ground. This seems to be an effect of 'a priori knowledge'.

Firstly using Eq.(9) disparity candidates were picked up by setting $c=0.24$, $S_H=0.7$, $K=\pm 3$. We checked visually that picked-up candidates always contained true disparities.

Secondly Fig.4(a) was obtained as a result of selecting disparities so as to minimize Eq.(6'). Brighter tone denotes a smaller disparity (higher altitude). The radius R in Eq.(5) was set to 128 pixels for the reduced 64x64 pixel images. This value was used commonly in the following experiments.

Thirdly after matching Fig.3(b), we obtained Fig.4(b) as a disparity map. Initial approximations of disparities were given from Fig.4(a).

In actual calculation Fig.4(b) was subdivided to 9 parts as shown in Fig.5 by solid lines. Each area is independently processed including a surrounding area enclosed by dotted lines. For example, the oblique line portion is for processing Area A5.

Identified disparities were inspected visually and proved to be all true but a portion marked by in Fig.4(b), which is wrongly matched due to noises by a shadow. Further in Fig.4(b) there are some edge pairs which escaped from picking-up, because of the smaller search range.

Fig.4(c) is a result of a similar process as in Fig.4(b). The central part of buildings is dented because lacking of matched edges on the roof due to halation. And buildings swell in the y-direction because lateral constraint is missing. To resolve these, x-directional edges should be utilized, though the Marr's model neglects them.

CONCLUSION

A matching method based on the potential minimum criterion is described. Matching edges is replaced by minimization of a quadratic form by use of the Green's function. With this method high matching precision was obtained. For further development high speed calculation and utilization of x-directional edges are necessary.

REFERENCES

1. D.Marr and T. Poggio, A theory of human stereo vision, A.I. Memo, No.451, MIT,1977.
2. K.Prazdny, detection of binocular disparities, Biol. Cybern. 52, 1985, 93-99
3. W.E.L. Grimson, A computational theory of visual surface interpolation, Phil. Trans. R. Lond. B 298, 1982,395-427.

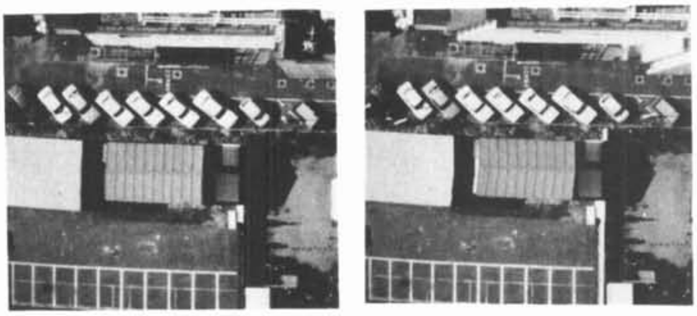

Fig.2 Stereo images (512x512 pixels)

(a)

(b)

(c)

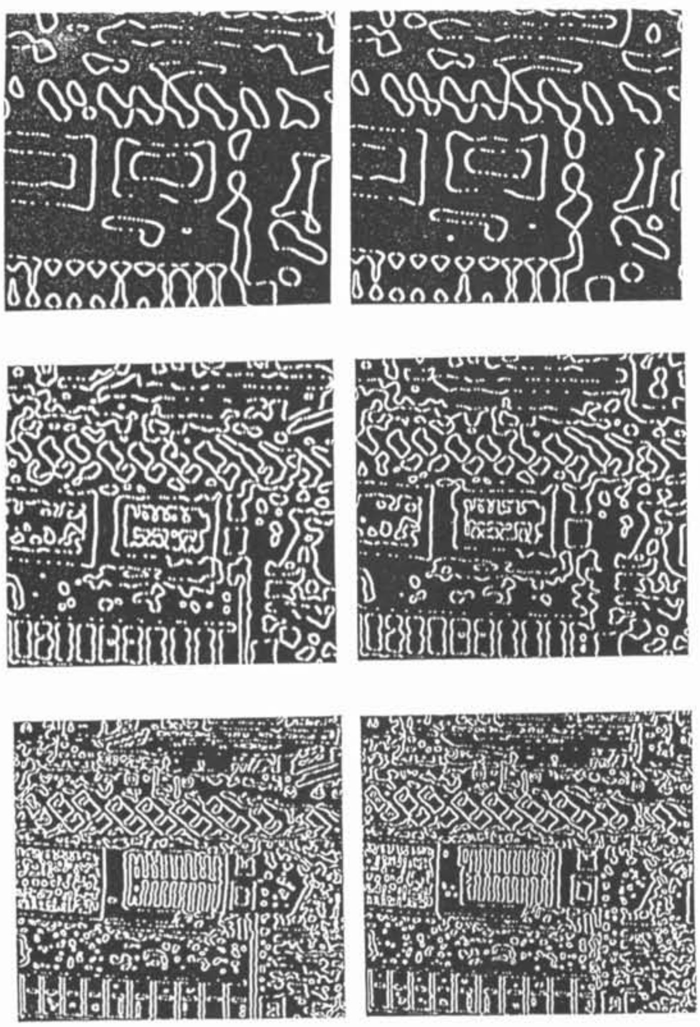

Fig.3 Y-directional edge map pairs

(a)

(b)

(c)

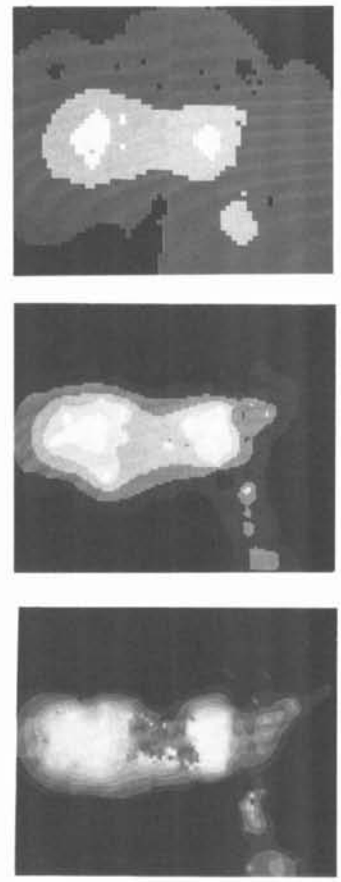

Fig.4 Disparity maps

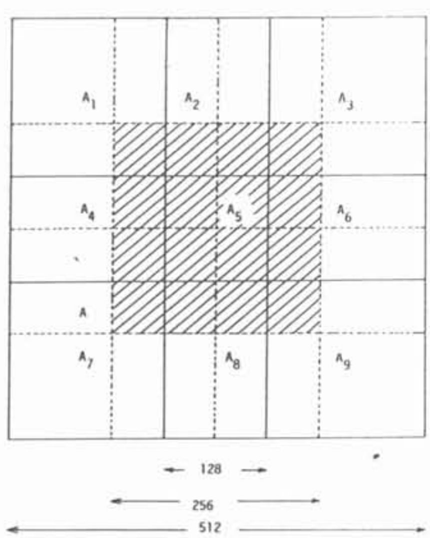

Fig.5 Image subdivision for independent processing

4. Michael D. Greenberg, Application of Green's functions in Science and Engineering, pp.71-92, Prentice-Hall, New Jersey, 1971.

5. S.T. Barnard and W.B. Thompson, Disparity Analysis of Images, Trans. on Pattern Analysis and Machine Intelligence, Vol.PAMI-2, No.4, July, 1980, pp.333-340

6. S. Hattori, H. Ohtani and S. Murai, Matching of edges by support functions (in preparation), 9th Asian Conference on Remote Sensing, Bangkok,1988