**MVA '92** IAPR Workshop on Machine Vision Applications Dec. 7-9,1992, Tokyo

# Stereo Matching Based on Iterative Incompatibility Revision

Yasuo SEKI, Hisayuki TATSUMI
Kanagawa Institute of Technology

1030, Shimo-ogino, Atsugi-shi,
Kanagawa, 243-02, Japan

## ABSTRACT

The most important problem remaining in 3-D information recovery by stereopsis is the matching problem. A new segment-based stereo matching method, which achieves high stereo matching correctness for large vergence-angle stereo pairs, is presented. The algorithms used are (a) incompatibility dissolution by iterative revision, (b) limitation of allowable disparity range suitable for vergence-stereopsis, and (c) weighting the candidate evaluation measure using the number of the rows constituting a candidate. Among these, algorithm (a) is the most important, as it contributes most to matching correctness increase and can dissolve incompatibility, without the processing time increasing rapidly with the scene complexity as graph-theoretic methods. Thus, it is the main focus of this paper. We performed experiments with the algorithm using stereo image pairs from 39 various scenes. A quantitative evaluation revealed fairly good matching correctness results. Relative evaluation experiments showed that the algorithms mentioned above can improve matching correctness. And, any rapid increase in processing time does not occur for more complex scenes.

## INTRODUCTION

Stereopsis is an important method of passive 3-dimensional information recoverry. In stereopsis research, matching is the most significant problem remaining. Previously, we presented[1] a segment-based method which extracts matching candidates connected to a sweeping parameter, and then determines a matching pair from them. In that method, matching compatibility is achieved by sequentially determining the matching pair from candidates compatible to all the matching pairs previously obtained. That compatibility adjustment method is simple to perform, but uneven and not reasonable enough. Next, to evenly dissolve incompatibility between candidates, we presented[2] a stereo matching method that uses the clique method hierarchically based on an analog candidate evaluation measure. Experiments involving that method showed fairly good results with respect to matching correctness. That method can evenly adjust incompatibility, but has a problem with the processing time increasing drastically as the scenes become more complex.

Now, we present in this paper a segment-based stereo matching method, that uses an analog evaluation meaure, to solve the above problems and also to make it effective under larger vergence-angle configurations. The algorithms employed are : (a) iterative incompatibility revision, (b) disparity range limitation suitable to vergence-stereopsis, and (c) candidate-evaluation-measure weighting using the total number of rows common to the segments of both frames. Algorithm (a) is the most important one of the three, as it solves the process bursting of graph-theoretic methods. Iterative methods have been applied to several stereo matching methods.[3],[4]-[6] Among then, Barnard et al. Lloyd et al. and Nagata et al. are all based on probabilistic relaxation. Though they revise incompatibility iteratively, in the same way as the method in this paper, they update the incompatibility based on neighboring information, which constitutes the framework of relaxation. In contrast, this method adjusts incompatibility without referring to neighboring information. Furthermore, the probabilistic relaxation revision process is complex, while the revision process for this method is simple.

In Hwang et al., global matching is determined without revision and is performed in the regional units. Matching for every region is determined only once, by sequentially shifting the matching from the current region to an important neighboring region, and revision is done for a table of relations among regions, edges, and vertices. The revision is not based on any analog evaluation measure like the method presented in this paper. The results of quantitative experiments using the method introduced here, indicate fairly good matching correctness can be obtained for various scenes with an average large vergence-angle. The basic algorithm used in our new method is described in section 2, the three characteristic algorithms are described in section 3, and the performance and evaluation experiments are presented in section 4.

## BASIC ALGORITHM

In the stereo matching method introduced in this paper : An analog evaluation measure[2] is incorporated in a simple matching control structure[1]. Then, the method is extended to vergence-observation and color input images, and the three algorithms stated above are added. This paper intends to show that the three algorithms can produce fairly good matching correctness by improving on the matching correctness. The basic structure of this method is :

(s1) stereo image pair input,
(s2) vergence-to-parallel-axis image transformation,
(s3) longitudinal disparity correction,
(s4) image defferentiation,
(s5) edge extraction,
(s6) segment extraction,
(s7) matching candidate extraction and matching pair determination,
(s8) iterative incompatibility revision,
(s9) depth estimation,
(s10) intermediate-viewpoint-image generation.

In step s1, the input images are generally composed of R,G,B color images, and the stereo camera system is in vergence configuration (Fig.1). In the figure, the left- and right- camera viewpoints are $V_\ell$ and $V_r$, their optical axes are $V_\ell A_\ell$ and $V_r A_r$ and cross each other at an angle of $2\alpha$, their image plane coordinate systems are $o_\ell$-$x_\ell$-$y_\ell$, $o_r$-$x_r$-$y_r$ ( $x_\ell$ and $x_r$ are the horizontal axes ), the midpoint of line $V_\ell V_r$ is $O$, and the right-hand 3-D coordinate system fixed to the camera system is $O$-$X$-$Y$-$Z$. Optical axes $V_\ell A_\ell$ and $V_r A_r$ are assumed as an approximation to lie in the same plane, thus they lie in the $X$-$Z$ plane. An arbitrary point on a given object is point P. By step s2, the vergence-to-parallel-axis-stereopsis image transformation, the previous parallel-axis stereopsis algorithm can be applied without alteration. The vision field is preserved in this transformation. As the total number of pixels in a frame is made constant, the average vision stereoscopic angle of each pixel is preserved. The focal length of each virtual image plane after transformation, which is its distance from the corresponding viewpoint, is the arithmetic average of the left and right frame focal lengths.

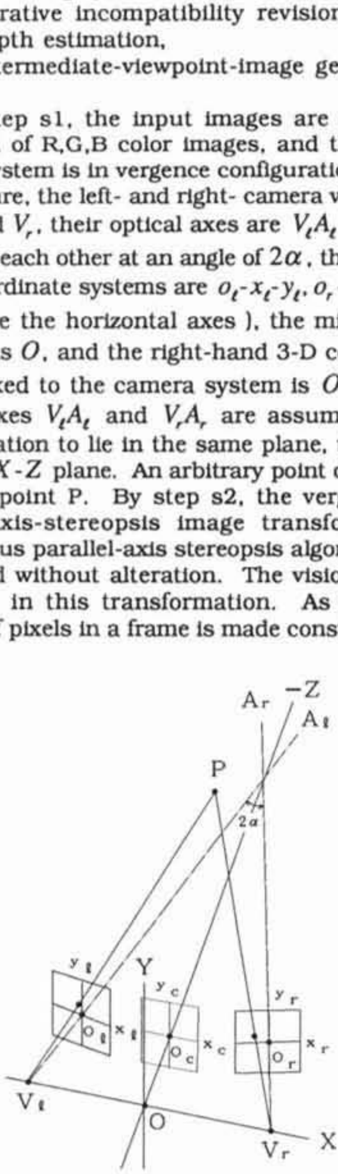


Fig.1 Stereo camera model for vergence configuration

In step s4, the image differentiation, the process of estimating gradient images from Gaussian-smoothed images has been extended to the color images. In step s5, edge extraction is performed in the same way as in ref.1. In step s6, segment extraction, we extract, on an edge line of one pixel in width, any maximal connected part which is bounded by branching points, tips, or $y$-coordinate optimal points. We call it a segment, an edge segment or a fundamental segment. We consider a composite segment, which is composed of two fundamental segments connected to each other, has no branching point in it, and is monotonous in its $y$-coordinate.

Step s7, matching candidate extraction and matching pair determination[1], is controlled by left-frame fundamental segments. A set of matching candidates is extracted from the segment pairs including a current left-frame fundamental segment. Then a matching pair is immediately determined from the matching candidate set. An allowable matching candidate composition in this paper is a pair of fundamental segments, or a pair made up of a fundamental segment and a composite segment[1]. In matching pair determination, the candidate with the smallest evaluation measure out of the candidates in the set corresponding to the current left-frame fundamental segment is determined to be the matching pair, without regard to previously determined matching pairs. As the candidate evaluation measure, we use an equally weighted sum of Euclidian distances for the segment shape and light intensity between the left and right frames. Step s8 is a newly added process for this paper, and will be explained in the next section. In step s9, the depth ( absolute $Z$-coordinates ) is estimated through disparity, and then a 3-D wireframe is generated in the computer[1]. In step s10, the 3-D wireframe is back-projected toward intermediate viewpoint $O$ to form a line image (the "intermediate-viewpoint image" ) on intermediate frame $o_c$-$x_c$-$y_c$ (Fig.1). By viewing the image, we can decide in principle without vagueness whether the segment-matching is correct or not.

## CHARACTERISTIC ALGORITHMS

### (I) Iterative Incompatibility Revision

In iterative incompatibility revision, matching pairs are modified based on their evaluation measures after they are tentatively obtained in step s7. The process of iterative incompatibility revision is :

(1) Matching candidates, including the current left-frame fundamental segment $\ell_s$, are listed from the smallest order of their evaluation

measure to at most the $n_{mc}$'th order (independent of $\ell_s$).

(2) For each iteration, the left-frame fundamental segments are scanned in a predetermined order ( for example, $y$-axis projection length of the segments ).

(3) For the current left-frame fundamental segment $\ell_s$, all the candidates listed in $\ell_s$ are examined in the order of smallest evaluation measure. Then, the first candidate found that is compatible with every tentative matching pair belonging to any other left-frame fundamental segment, is determined to be the updated tentative matching pair of $\ell_s$.

(4) Steps (2) and (3) are repeated unless the number of any case with $\ell_s$, in the current iteration, whose tentative matching pair is different from that in the previous iteration, is zero or small enough. Otherwise, the process stops.

A thinking experiment is shown for iterative incompatibility revision in Figs. 2 and 3. The relative magnitude of the evaluation measure for each matching candidate $C_i$ ( $i$ is serial number ) is shown in Fig. 2a. The incompatibility relation between candidates is shown by the graph in Fig. 2b. Each node on the graph corresponds to a candidate, and each edge of the graph shows that the two nodes connected to it are incompatible with each other in the depth uniqueness request.

The progress of the iterative incompatibility revision process, under the conditions corresponding to Fig. 2, is shown in Fig. 3. The symbol "−" in the registered candidate column shows that there is no candidate to be registered in the column. One candidate, for example c9, can belong to a number of different left-frame fundamental segments. Iterative number "0" corresponds to the matching pairs determined in step s7. The process converges at iteration number 2. The symbol "O"(circle) shows that the corresponding candidate is the tentative matching pair for the current left-frame fundamental segment and the current iteration number.

**(II) Allowable Disparity Range Limitation Suitable For Vergence Stereopsis**

The usual allowable disparity range limitation corresponds to the assumption that objects cannot be situated nearer (in $-Z$ coordinate) than a certain small distance from the stereo camera system in the foward direction and cannot be situated at all in the backward direction except for noise effect. In this paper, taking advantage of zoomed up vergence-stereopsis, we consider an allowable disparity range of width $\Delta d$, which is centered at disparity $d_c$ of the virtual parallel-axis frames between the vision lines of the left and right frames' centers (Fig. 1).

The usual method of determining the allowable disparity range and the method used in this paper for vergence-stereopsis are shown in Figs. 4(a) and (b), respectively. In the latter method, the allowable disparity range $d_{sp}$ is described by

$$d_c - \Delta d \le d_{sp} \le d_c + \Delta d$$

The allowable disparity half width $\Delta d$ is the frame width multiplied by, for example, 0.5. The method used in this paper has no useless sections in the allowable disparity range and is effective, as shown in Fig. 4.

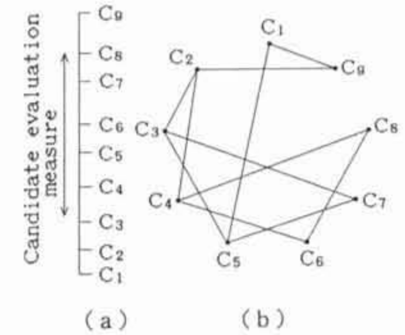
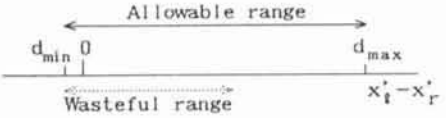


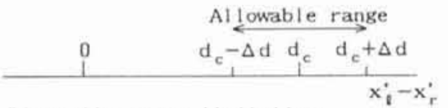
Fig. 2 Iterative inconsistency correction method

| Left-frame fundamental segment $l_S$ | Candidate pair | Tentatively determined pair<br>Repetition number 0 | 1 | 2 |
|---|---|---|---|---|
| 1 | − | | | |
| | C 9 | | | |
| | C 1 | ○ | ○ | ○ |
| 2 | C 7 | | ○ | ○ |
| | C 5 | | | |
| | C 3 | ○ | | |
| 3 | C 8 | | | |
| | C 6 | | ○ | ○ |
| | C 4 | ○ | | |
| 4 | − | | | |
| | C 9 | | | |
| | C 1 | ○ | ○ | ○ |

Fig. 3 Example of iterative inconsistency correction
( Progress of the algorithm )



(a) Previous disparity range limitation



(b) Disparity range limitation suitable for vergence-stereopsis

Fig. 4 Disparity range limitation suitable for vergence vision

### (III) Weighting Candidate Evaluation Measure by Common Height

Weighting the evaluation measure of a matching candidate by the height of the common range results in the preference for a candidate corresponding to a physical edge line with a larger effective length. Here, the common range is the range for the $y$-coordinates at which both frame segments belonging to a candidate coexist. To describe more precisely, at each row in the common range, the distances between features ( a shape feature, that is, an edge direction, and an edge intensity ) of both frames are averaged using weighting about the features. Then, the arithmetic average of the results about rows in the range is multiplied by the reciprocal of the common range width to produce the evaluation measure for the candidate.

## Performance and Evaluation Experiment

In this section, we describe performance and evaluation experiments for our stereo matching method. Relative evaluation experiments were carried out to determine the effect of the characteristic algorithms on matching correctness, and absolute (quantitative) evaluation experiments were performed to determine the matching correctness of the proposed method. All scenes used for the experiments are shown in Fig. 5. The input images for the experiments are $256 \times 240$ pixels compressed from $512 \times 480$ pixels using $2 \times 2$ pixel gray level averaging. In the experiments for the effect of the characteristic algorithms, the matching results were evaluated by comparing the intermediate viewpoint images from our method with images from our method with each characteristic algorithm removed or interchanged with an alternative algorithm. Results from the relative evaluation experiments for iterative incompatibility revision, the most important characteristic algorithm, are shown in Table 1. The symbols "+", "0" and "−" in the table indicate whether the matching correctness has increased, is unchanged, or has decreased, respectively.

The algorithm used for comparison is where matching candidate incompatibility adjustment is performed sequentially in step s7, as described in ref. 1, and any incompatibility adjustment as in step s8, is not performed afterwards. The table indicates that the iterative incompatibility revision algorithm has a very large effect on the matching correctness increase. The iterative incompatibility revision algorithm converges rapidly and converges usually in two or three iterations. In the convergence states, the tentative matching pairs did not change at all from those of the previous iteration, that is, incompatibility was fully dissolved. The matching results in the experiments were all obtained by forcedly stopping the processing after the third iteration. The stereo matching processing time was shown, by experiments, not to increase as sharply as the hierarchical clique method, even for complex scenes.

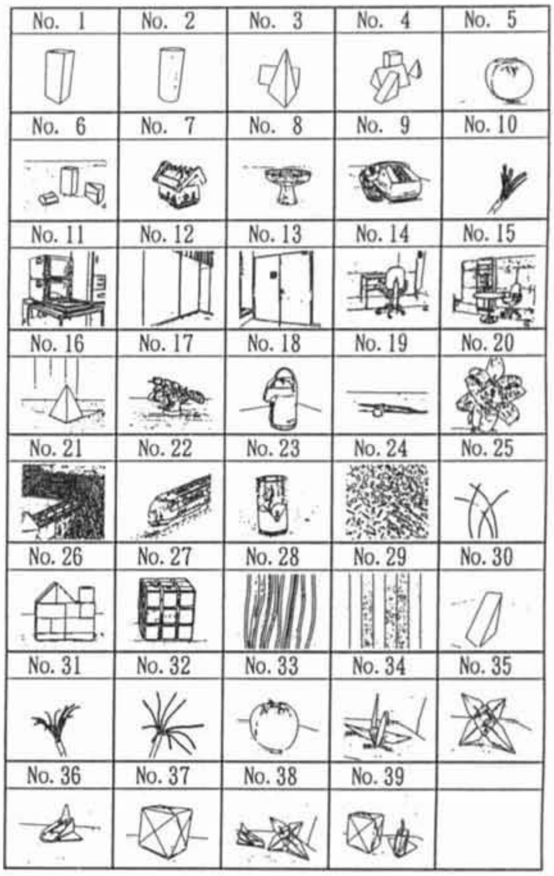


Fig. 5 Edge images showing sample scenes for matching experiments

Table 1 Iterative incompatibility revision

| Effect of algorithm | Number of scenes | % |
|---|---|---|
| + | 27 | 69 |
| 0 | 11 | 28 |
| − | 1 | 3 |

The stereo matching method, including the three characteristic algorithms, was used on 39 various scenes. The matching correctness judged by inspection was, roughly speaking, good. The matching correctness was evaluated quantitatively for those matching results. The fundamental viewpoints for the quantitative evaluation are described as follows :

(1) The quantitative evaluation targets are points on left-frame segments that are longer than a threshold and correspond to any structure lines or boundary lines between regions of different surface materials on the objects concerned.

(2) The possible evaluation categories for each target point defined for matching evaluated in the first place are :
   (a) "correct" : The target current point is matched to its correct right-frame point if the former is not occluded in the right frame, and is not allowed to be matched otherwise.
   (b) "false" : The target current point is matched to its false right-frame point if unoccluded, and is allowed if occluded.
   (c) "rejected" : The target current point is not allowed to be matched if unoccluded.

(3) If a target point is matched many times, the excess matching is counted in an incompatibility evaluation measure. If the excess matching is refused matching, the matching is naturally not counted in the measure.

Here, quantitative evaluation is carried out in the order of matching pairs belonging to left-frame fundamental segments with larger edge points. The quantitative evaluation results of this method for the 39 selected scenes are shown in Table 2. There, the quantities $r_c$, $r_f$, $r_r$, and $r_i$ are the rates of the correct, false, rejected and incompatible point number $n_c$, $n_f$, $n_r$, $n_i$ to the total interested edge point number $n_t$, respectively, where $r_c = 1 - r_f - r_r$ and $n_c = n_t - n_f - n_r$.

## Concluding Remarks

We presented a new stereo matching method that is characterized by three algorithms, iterative incompatibility revision, allowable disparity range limitation suitable for vergence-stereopsis, and weighting candidate evaluation measure by candidate height. Our experiments showed that each algorithm is sufficiently effective for increasing the matching correctness. The matching results for this method, using 39 various scenes, were shown to be fairly good using a quantitative evaluation.

Table 2 Matching correctness

| Scene no. | $n_t$ | $r_c$ | $r_f$ | $r_r$ | $r_i$ |
|---|---|---|---|---|---|
| 1 | 634 | 0.973 | 0. | 0.027 | 0. |
| 2 | 471 | 0.977 | 0. | 0.023 | 0. |
| 3 | 703 | 0.886 | 0. | 0.114 | 0. |
| 4 | 1001 | 0.680 | 0. | 0.151 | 0.169 |
| 5 | 705 | 0.781 | 0.054 | 0.165 | 0. |
| 6 | 927 | 0.755 | 0. | 0.191 | 0.054 |
| 7 | 1656 | 0.826 | 0.031 | 0.135 | 0.008 |
| 8 | 897 | 0.890 | 0.003 | 0.107 | 0. |
| 9 | 2051 | 0.634 | 0.191 | 0.159 | 0.016 |
| 10 | 1208 | 0.746 | 0.089 | 0.165 | 0. |
| 11 | 5442 | 0.853 | 0.020 | 0.126 | 0.001 |
| 12 | 2633 | 0.765 | 0.107 | 0.128 | 0. |
| 13 | 2839 | 0.706 | 0.077 | 0.217 | 0. |
| 14 | 2711 | 0.758 | 0.058 | 0.184 | 0. |
| 15 | 3702 | 0.773 | 0.061 | 0.166 | 0. |
| 16 | 689 | 0.609 | 0.060 | 0.331 | 0. |
| 17 | 1542 | 0.685 | 0.112 | 0.197 | 0.006 |
| 18 | 1382 | 0.712 | 0.019 | 0.269 | 0. |
| 19 | 756 | 0.762 | 0.020 | 0.218 | 0. |
| 20 | 2382 | 0.740 | 0.081 | 0.179 | 0. |
| 21 | 9672 | 0.672 | 0.101 | 0.217 | 0.010 |
| 22 | 2191 | 0.844 | 0.005 | 0.151 | 0. |
| 23 | 1760 | 0.817 | 0.014 | 0.169 | 0. |
| 24 | 5482 | 0.613 | 0.213 | 0.169 | 0.005 |
| 25 | 1076 | 0.669 | 0.048 | 0.283 | 0. |
| 26 | 1992 | 0.756 | 0.158 | 0.086 | 0. |
| 27 | 3394 | 0.504 | 0.146 | 0.350 | 0. |
| 28 | 4592 | 0.948 | 0.005 | 0.047 | 0. |
| 29 | 1708 | 1. | 0. | 0. | 0. |
| 30 | 523 | 1. | 0. | 0. | 0. |
| 31 | 1267 | 0.557 | 0.261 | 0.182 | 0. |
| 32 | 1904 | 0.836 | 0.012 | 0.152 | 0. |
| 33 | 714 | 0.695 | 0.032 | 0.273 | 0. |
| 34 | 1059 | 0.584 | 0.242 | 0.174 | 0. |
| 35 | 1359 | 0.716 | 0.135 | 0.149 | 0. |
| 36 | 934 | 0.737 | 0.100 | 0.163 | 0. |
| 37 | 1221 | 0.707 | 0.124 | 0.169 | 0. |
| 38 | 1518 | 0.757 | 0.081 | 0.162 | 0. |
| 39 | 1652 | 0.633 | 0.160 | 0.203 | 0.004 |
| Average | 2009 | 0.758 | 0.072 | 0.163 | 0.007 |

## References

(1) Seki Y. : "Stereo Matching Based on Edge Segment", Trans. IECE Japan (Section E), E69, 5, pp.666-674 (1986-05)

(2) Seki Y. : "Stereo Matching Using Segment Shape as Constraint", Trans. IECE Japan (Section D), J71-D, 11, pp.2357-2369 (1988-11).

(3) Barnard S.T. and Thompson W.B. : "Disparity Analysis of Images", IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-2, 4, pp.333-340 (1980-7).

(4) Lloyd S.A., Haddow E.R. and Boyce J.F. : "A Parallel Binocular Stereo Algorithm Utilizing Dynamic Programing and Relaxation Labelling", Computer Vision, Graphics, and Image Processing, 39, pp.202-225 (1987).

(5) Nagata H., Nakamura O., Minami T. : "Correspondence Search of Stereo Images Based on Relaxation with Epistemic Probability", Trans. IECE Japan (Section D), J72-DII, 9, pp.1397-1406 (Sept. 1989).

(6) Hwang J. J. and Hall E. L. :"Matching of Featured Objects Using Relational Tables from Stereo Images", Computer Graphics and Image Processing, 20, pp.22-42 (1982).

(7) Marr D. and Poggio T. : "A theory of human stereo vision", Artificial Intell. Lab., Massachusetts Inst. Technol., Cambridge, AI Memo 451, pp.1-89 (1977).