A PARALLEL MATCHING ALGORITHM FOR STEREO VISION

Y. Nishimoto and Y. Shirai

Asada Research Laboratory
Kobe Steel, Ltd.
53-3, Maruyama, Gomo, Nada-ku, Kobe, Japan

Automatic Control Division
Electrotechnical Laboratory
1-1-4, Umezono, Sakura-mura Niihari-gun, Ibaraki, Japan

ABSTRACT

This paper proposes a parallel matching algorithm for feature-based stereo vision. Features are zero-crossing (ZC) points detected with various sizes of Laplacian-Gaussian filters. In order to obtain candidate intervals of disparity, the disparity histogram is computed all over the image. The image is, then, divided into small areas and the disparity histogram in each local area is computed within the candidate intervals. The local disparity histograms in all the channels are fed to the fusion evaluates and the most probable disparity is detected in each local area. Once the most probable disparity is detected, disparities for all the finest ZC points arc determined in the local area to obtain a high resolution disparity map. The, matching pairs are removed from a set of ZC points. A series of processes are iterated until no more disparities are determined.

Experiments with a sample scene reveals that the algorithm has advantages in efficiency and performance.

INTRODUCTION

Among various types of range finding methods, stereo vision is worthy of notice since it needs no active media. However, it is so difficult to match corresponding points in the two images that the stereo vision has not been fully established as a computer vision system.

There have been many matching algorithms proposed, which may be classified into feature based method and area-based method.

In feature-based method. Marr and Hildreth introduced the convolution operator V^2G, where V^2 is the Laplacian operator and G stands for the two-dimensional Gaussian distribution, and adopted the zero-crossing (ZC) of the V^2G-filtered image as the features to be matched (Marr and Hildreth, 1980). A problem with matching the ZC is that ZC points may appear randomly in the region o\' little intensity change. If we try to match ZCs including such random ones, the probability of false matching becomes large.

Another problem is concerned about how to match corresponding points. Marr and Poggio proposed the hierarchical matching algorithm, in which matching process is started in the coarsest channel of ZC and followed by finer ones (Marr and Poggio, 1979). As was successfully implemented by Grimson the algorithm is quite reasonable for a scene where the depth change smoothly (Grimson, 1981). However, it turns out to be inefficient in certain cases. The first is a case where a scene includes many objects at different positions as observed in usual room scenes. The inheritance of disparity information from a coarser channel to a finer one frequently fails, especially at the discontinuities of depth. The second is a case where the main feature consists of high spatial frequency component alone. A white wall with small scratches is one of the examples. In that case the matching process can not start in the coarsest channel.

In order to solve problems mentioned above, we have developed and implemented a parallel matching algorithm.

II OUTLINE 0F ALGORITHM

The general block diagram of the parallel matching algorithm is shown in Figure 1.

Features are the ZC points detected with various sizes of V^2G. In order to obtain candidate intervals of disparity, the disparity histogram is computed all over the image. The image is, then, divided into small areas and the disparity histogram in each local area is computed within the candidate intervals. The local disparity histograms in all the channels are fed to the fusion evaluator and the most probable disparity is detected in each local area. Once the most probable disparity is detected, disparities for all the finest ZC points are determined in the local area to obtain a high resolution disparity map. A senes of processes are iterated until no more disparities are determined.

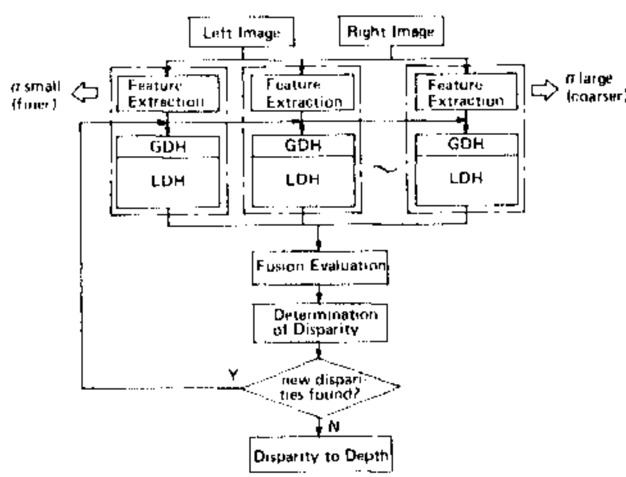


FIGURE 1 BLOCK DIAGRAM OF PARALLEL MATCHING ALGORITHM

(GDH : Global Disparity histogram
LDH : Local Disparity histogram)

III FEATURE EXTRACTION

Let the input image and $\nabla^2 G$-filtered image be denoted by $F^{(x)}(i, j)$ and $B_\sigma^{(x)}(i, j)$, where σ is the standard deviation of G and x means either right (R) or left (L). $B_\sigma^{(x)}(i, j)$ is given by the following equation.

$$B_\sigma^{(x)}(i,j) = \nabla^2 (G * F^{(x)}(i,j)) = (\nabla^2 G) * (F^{(x)}(i,j)) ,$$

where "$*$" denotes the convolution operation. The filtered image $B_\sigma^{(x)}(i, j)$ has no dc component and can be equally divided into positive and negative regions. The boundaries between the two regions turn out to be the ZC. Here, we define the ZC as unit vector $r_\sigma^{(x)}(i, j)$ along the boundary.

A ZC point corresponding to a small contrast is removed on the basis of the gradient value of the G-filtered image on that ZC, that is,

if $|\nabla (G * F^{(x)}(i,j))| < G_0$,
then the ZC is removed,

where G_0 is a predetermined threshold value.

IV DISPARITY HISTOGRAMMING

The calculation of disparity histogram is generally the most time-consuming process in the matching operation. Here, the global disparity histogram (GDH) is first computed to find an approximate disparity distribution. The GDH is defined as:

$$GDH_\sigma^{(R)}(d) = \frac{\sum_{(i,j) \in A} r_\sigma^{(R)}(i,j) \cdot r_\sigma^{(L)}(i+d,j)}{\sum_{(i,j) \in A} r_\sigma^{(R)}(i,j) \cdot r_\sigma^{(R)}(i,j)} ,$$

where d stands for a disparity and A is the whole image plane. Since points on the physical surface constitute clusters in space, the true matches tend to fall in some intervals of disparity, while the false matches randomly scatter. The GDH consequently gives the approximate disparity distribution of objects in the scene.

Now let the peak value of $GDH_\sigma^{(x)}(d)$ be H. Intervals S determined in the following equation is the disparity intervals for the local disparity histogram (LDH) to be calculated in.

$$S = \{ d \mid GDH_\sigma^{(R)}(d) \geq a \cdot H \} ,$$

where a is a constant value and $0<a<1$. S may generally consist of multiple intervals corresponding to objects in the scene. In this way we can limit the candidate intervals of disparity and greatly improve the efficiency of the LDH process.

The LDH represents the disparity distribution of true and false matches within window W_σ of $M_\sigma \times M_\sigma$ around ZC point P(i, j), where M_σ is determined as the average pitch of ZCs, that is, a function of σ. The LDH for a fixed window on the right image is defined as:

$$LDH_\sigma^{(R)}(d;i,j) = \frac{\sum_{(i',j') \in W_\sigma} r_\sigma^{(R)}(i',j') \cdot r_\sigma^{(L)}(i'+d,j')}{\sum_{(i',j') \in W_\sigma} r_\sigma^{(R)}(i',j') \cdot r_\sigma^{(R)}(i',j')} ,$$

where d f- S .

V FUSION EVALUATION

Now the fusion for each window area is evaluated using LDHs of all channels. The best fusion channel is first selected as the one such that the difference between the first and second largest peaks in LDH is the largest. Let the difference of the peaks in the best channel be $F^{(x)}(d; i, j; \sigma)$. The local fusion is established if the following condition is satisfied for a certain disparity d^*:

$$F^{(R)}(d^*; i, j; \sigma) \geq F_0 \quad \text{and} \quad F^{(L)}(d^*; i+d^*, j; \sigma) \geq F_0 ,$$

where F_0 is a predetermined threshold value. Disparity d^* can be regarded as the most probable disparity in W_σ.

VI DETERMINATION OF DISPARITY MAP

Once the most probable disparity d^* is obtained in W_σ, disparities for all ZC points in W_σ are determined in the following manner. Let the disparities between a ZC point in the right image and matching candidates in the left image be $d_1, d_2, \ldots, d_n$ in the finest channel, and the disparity which is the nearest to d^* among $d_1, d_2, \ldots d_n$ be d_{op}.

If $|d_{op} - d^*| \leq d_0$, then the disparity for the ZC is finally determined to be d_{op}, otherwise, the determination is postoponed,

where d_0 is the predetermined value. Once the pair $Q(P_m, P_n)$ is determined as a true match, pairs $Q(P_k, P_n)$ $(k \neq m)$ and pairs $Q(P_m, P_k)$ $(k \neq n)$ are regarded false matches and removed in the succeeding iterations, where P_m and P_n are a ZC point in the right and left image.

The processes from GDH calculation to this process are iterated until no more disparities are determined. The disparities are transformed to depths by means of predetermined geometrical relations and camera parameters in a straightforward manner.

VII EXPERIMENT

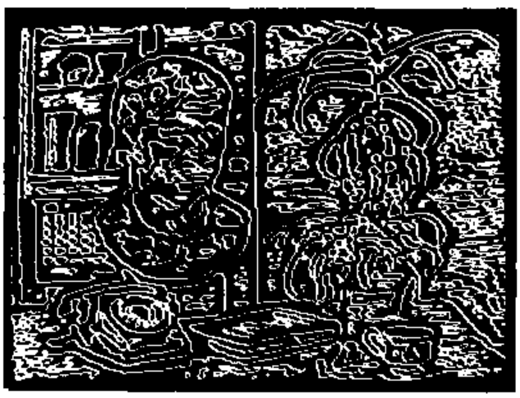
The parallel matching algorithm was tested with the scene shown in Figure 2, with the Prime 750 in the Information Computer System of Electrotechnical laboratory. The scene has many objects such as a plaster figure, a telephone, a book, plant, and a coffee cup in front of a shelf, white paper, and a calendar. The telephone has a glossy surface and the plant has quite a complex profile. In addition, the calendar has a periodical pattern which is, in principle, unsuitable for stereo vision.

The ZC points in the fine, middle, and coarse channel are shown in Figure 3, and the filtered ZC points in the fine channel is shown in Figure 4. The global disparity histogram from the right at the first iteration is shown in Figure 5. The disparities finally obtained are shown in Figure 6, where the larger disparities (nearer to the camera) are displayed by the brighter points. There are 9424 ZC points in all in the right image, of which about 6500 ZC points have correspondences and the left are those in the occluding or uncommon regions. Most of the disparities of 6500 ZC points were reasonably determined.

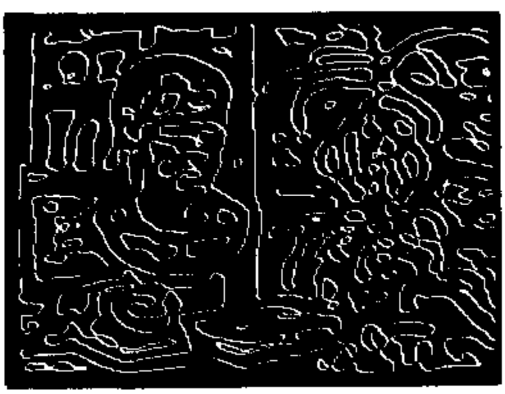
It is revealed that the ZC points whose correspondences are detected by a finer channel, mainly lie in the region near edges of objects, that is, the area of discontinuous depth, while the ZC points whose correspondences are detected by a coarser channel lie in the region of little change of depth, or in the periodical pattern on the calendar.

The cpu time is about 20 minutes except the calculation of feature extraction which could be quickly executed with special hardware.

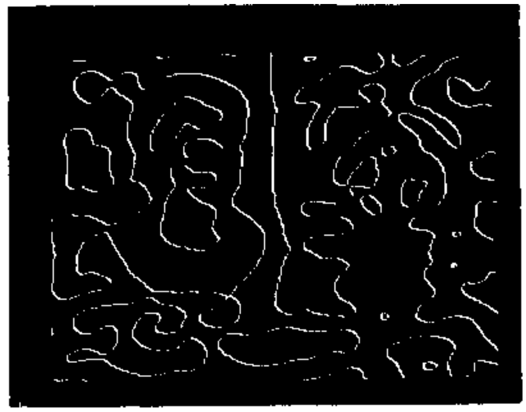
FIGURE 2 ORIGINAL IMAGE



(a) σ = 1.5 pixel (fine)



(b) σ = 3 pixel (middle)



(c) σ = 6 pixel (coarse)

FIGURE 3 ZC POINTS

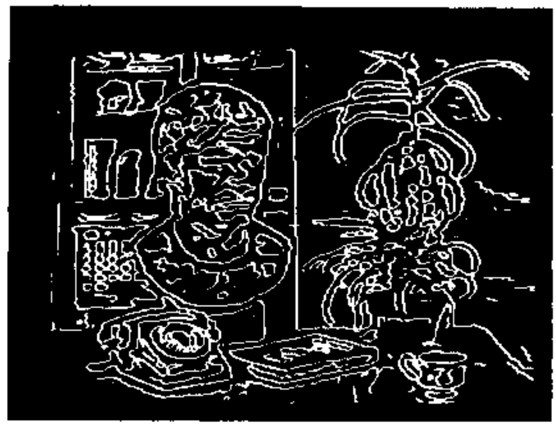


FIGURE 4 FILTERED ZC (FINE)

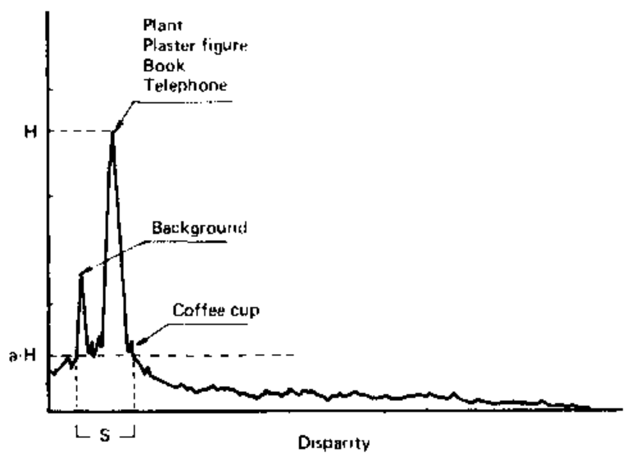


FIGURE 5 GLOBAL DISPARITY HISTOGRAM

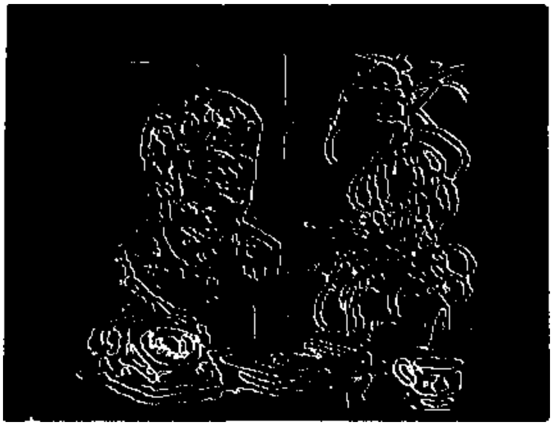


FIGURE 6 FINALLY OBTAINED DISPARITY

VIII CONCLUSION

A parallel matching algorithm for feature-based stereo vision is proposed.

It has the following features. Since the algorithm uses the features in the various resolutions evenly and complementally, it can deal with the correspondence problem according to the way how the depth changes on and between the physical surfaces of objects. For example, if an object has a surface on which the depth changes little, a coarser channel is automatically excited and the disparities over the wide area can be determined at a time. Introduction of GDH enables us to limit the search intervals of disparity and to reduce the amount of calculations. In addition, the algorithm is essentially suitable for hardware due to its parallel architecture.

However, there is room for improvement in the details of algorithm. Especially, the criterion of evaluation *of* fusion and the procedure of iteration need further refinement through experiments with various types of scenes.

REFERENCES

[1] Man, D. and Hildreth, E., "Theory of Edge Detection." Proc. R Soc. Lond. B. 207 (1 980) 1 87 217.

I2| Marr, D. and Poggio, T., "A Computational Theory *of* Human Stereo Vision." Proc. R. Soc. Eond. B 204 (1979) 301 328.

I3I Cirimson, WEE., "From Image to Surface." Mil Press, Cambridge (1981).