A New Stereo Matching Approach Using Edges and Nonlinear Matching Process Objected for Urban Area

Mitsuteru Sakamoto[1], Wei Lu[1], Pingtao Wang[1] and Yukio Kosugi[2]

[1]Asia Air Survey Co., Ltd. R&D Department
8-10, Tamura-cho, Atsugi-shi, Kanagawa 243-0016, Japan
[2]Frontier Collaborative Research Center, Tokyo Institute of Technology
4259, Nagatsuta-cho, Midori-ku, Yokohama 226-8503, Japan

Abstract
Automatic recognition of buildings has been in demand for efficient digital mapping process or updating of existing information in geographical information applications. It requires effective stereo matching techniques that are applicable to urban area. However conventional techniques that make use of area-based, point-based or edge-based matching cannot generate satisfactory results for urban area because of occlusion or inability of mismatching recovery. This paper proposes a new stereo matching technique, which combines edge detection and nonlinear mapping, and is based on a process called Coincidence Enhancement Method (CEM) [4]. Edge segments, obtained from edge detection and stereo matching process, are used as constraints for edge enhancement in depth map. Experiments on edge detection, edge matching, CEM and edge enhanced CEM were performed with stereo aerial imageries of urban area. The results show that the proposed approaches of CEM with edge support are efficient for improving the matching results surrounding building's boundary.

I. INTRODUCTION

Up till now mapping processes of ground objects with aerial stereo image pairs still depend on professional operator and needs tremendous time and costs. While new systems known as Digital Photogrammetric Workstation (DPW) has been presented past several years, which has more effective functions for low cost mapping, it has not been realized to detect or map ground objects automatically. Computer vision is most widely studied technology and is becoming more and more promising thanks to easier availability of auxiliary information, such as height information from laser profiler finder or SAR. However, for urban area where occlusion happens frequently, the most fundamental technology such as stereo matching still face great difficulty, so is edge detection process, which output is one of the most crucial information for reconstruction of building model.

Edges that can be extracted with various kinds of filters are the most fundamental information for automatic recognition of manmade structures. If the structural boundary of a ground object can be extracted in the form of edge information, it will be not a much difficult job to reconstruct the 3 dimensional model of the original structure. There have been numerous researches regarding efficient and reliable edge detection. However, because of the limitation of image media, none of the acclaimed algorithms can produce complete and error free results.

Coincidence Enhancement (CE), being an extension of Hebb's rule, is a self-organizing process of neural network modeling. This process can be modeled by the principle of competition and consensus [4]. CE model can realize smooth projection between input signals and output pattern, which means that when the majority of the initial values are correct, the minority of erroneous data can be absorbed. This effect is very useful in reducing the wrong stereo matching result caused by local minimum. The key factor for CE model to function correctly is to ensure the reliability of the initial data, which is not an easy job in the case of computer vision, especially when dealing with real world images.

Recently, we have been developing a DPW system, and constantly making improvement for this system [1]. Even though this system has many superior functions, such as the ability of handling imageries of several hundred mega bytes at high speed even on personal computers of middle range cost, various digitizing function and so on, almost all of the processes still depend on manual operations. Our ultimate goal for the system is to introduce as many as possible automatic processing modules of high reliability to reduce the cost of production and increase the quality of digitized information.

The algorithms proposed in this paper aim at improving the precision and reliability stereo matching process by introducing CE process that is constrained by reliable edge information. The first section gives a general overview of the proposed algorithms. The second section describes in detail the improved algorithms for extracting reliable edge information. The third section describes the enhanced CE process with the reliable edge information functioning both as initial value and constraints during enhance process. Experiments have been conducted and the results show that the proposed algorithms are capable of improving the precision and reliability of stereo matching process.

1082-4006/00/0702-79$5.00

II. OVERVIEW OF THE PROPOSED ALGORITHMS

Objectives and tasks

Our objectives are to improve the reliability and precision of stereo matching for automatic building recognition. The reliability is improved by using highly reliable edge information that is extracted through strict constraints. The precision is improved by using CE process enhanced with edge constraints. The background and outline of the two strategies are as follows.

Edges are fundamental components of building. In this study, edge mainly servers for three purposes. One is for initialization of global search for stereo matching, which is equal to general registration in photogrammetry. The other is for constraining the local matching in CEM, whose brief description will be given below and the details will be given in section 4. The third one is for maintaining collinear condition during mapping between stereo image pairs.

Since edge information extracted by conventional approaches tends to be noisy and unreliable, our strategy is to only detect of highly reliable conjugate edges for general registration. Since in the process of mapping between stereo image pairs do not necessarily require conjugate edges, more edges that are not strictly matched can be used for maintaining uniform parallax.

CE is one kind of neural network model and gives very unique mapping process, which realizes nonlinear mapping from inputs to outputs. In this study, this technique is applied to stereo matching process, with the input and the output being the stereo image pairs. Since CEM produces matching results less sensitive to erroneous local minimum, it can reduce the possibility of mismatching caused by occlusions or noises. In order to reduce the effect of discontinuous shift vectors or parallax changes, we use the edge information of high reliability as constraints.

Process flow

Figure 1 shows the general process flow of our proposed algorithms.

In the first stage, edges are extracted using rectified stereo image pairs, which are across epipolar lines. The processes consist of several steps, such as edge enhancement, thinning, tracing of enhanced edge image, segmentation, filtering by geometrical constraint and template matching with the consideration of edge boundaries [2]. We apply two different types of edge operators for edge image enhancement and compare the result to choose the more effective one. The operators are SUSAN (Smallest Univalue Segment Assimilating Nucleus) operator [3] and Haar' s wavelet transformation respectively.

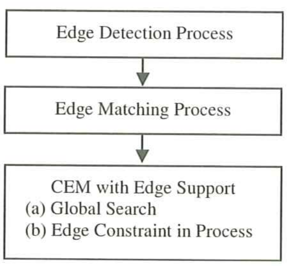


Figure 1. The general process flow

In the second stage, the edges extracted from stereo image pairs are matched with each other to form conjugate edge pairs. Geometric constraints from the orientation parameters, and image similarity calculated from edge dependent template are introduced to guarantee the high reliability of matching result. The matched edges are further trimmed to the same shapes according to their common section.

In the last stage, nonlinear matching process is performed. This process is based on CEM. CEM realizes nonlinear normalization at local area in the target system and acquiring of global agreements by iterative computation. In this approach we conceptually identify neuron's activity with matching or shifting vector. This network model has two phases in optimizing procedure of shifting vector. The first one is competition phase in which shifting vectors are individually searched for optimal displacement with certain constraints. The second one is consensus phase in which shifting vectors are operated to have local continuity of shifting vector or parallax. These two phases are repeated to achieve local normalization with respect to the edge matching results. We tried to extend CEM with the result of edge matching.

III. EDGE DETECTION FROM RECTIFICATION IMAGE

To simplify the task of edge detection and refinement, stereo image pairs are translated to rectified images using adequate perspective projection, as long as proper tie points are specified. This step includes photogrammetric operations such as estimation of orientation parameters by relative orientation. Rectification images provide a stereo model that has no y-parallax. As a result, original 2 dimensional matching becomes the equivalence of 1 dimensional matching (in x-direction, corresponding to epipolar line), and only extraction of edges across epipolar lines need to be considered.

The process flow of edge detection after rectification is shown in Figure 2 [2]. In this study, two types of edge detection operators are applied. One is the SUSAN operator proposed by Smith et al. and the other is wavelet transformation. SUSAN operator is a unique interest operator that pays attention to the variation of regional brightness distribution in a local window. The brief descriptions of each process are as follows.

(1) Pre-processing
As pre-processing, stereo image pairs are smoothed by 3 x 3 size median filter for noise reduction, and brightness or contrast adjustment is applied to have almost the same brightness.

(2) Edge enhancement
Two types of operators (SUSAN and wavelet) are used to enhance edge component to get the candidates of building's structural edge. Haar's wavelet is used here and operation is applied in x-direction. The result of level-1 process is selected for enhanced image. Binary edge images are also produced in this stage. Eq.1 shows the basic formula of wavelet transformation and Figure 3 shows Haar's mother wavelet.

$$(W_\varphi f)(b,a)=\int_{-\infty}^{\infty}\frac{1}{\sqrt{|a|}}\overline{\varphi\left(\frac{x-b}{a}\right)}f(x)dx \qquad (1)$$

where $f(x)$ is input signal, $\varphi(x)$ is mother wavelet function, a is scale parameter and b is translation parameter.

(3) Edge thinning
Applying mathematical morphology operation (dilation, erosion, hit and miss etc.), binary edge components are locally connected and thinned to trace edge segments.

(4) Edge tracing
Edge tracing is performed by 8-direction neighbor connection process. In this stage chained edge elements are extracted from raster edge images.

(5) Edge line segmentation
Since chained edge elements are not direct representation of object's shape, they must be approximated as line segments (line segmentation), which are the fundamental information in the following processes such as edge matching. The algorithm of line segmentation is as followings.

Consider edge element with m nodes $(n_1,.n_m)$.

(i) Approximate line L_i with node n_i and n_j. At initial $i = 1$, $j = m$.
(ii) Calculate distance d_k ($k = i+1,..j-1$) between each n_k and L_i.
(iii) If any d_k is greater than a pre-defined threshold dt, substitute $j -1$ for j and go to step (i).
(iv) Register L_i as line segmentation and substitute j for i.
(v) If i equals to j, process is finished, otherwise go to (i).

Line segments with length below a predefined threshold are regarded as of low reliability and discarded.

IV. EDGE MATCHING

The line segments obtained in the above process differ from each other in stereo image pairs in many ways. To make line segment usable for stereo matching process, the following algorithms are proposed to select and modify reliable line segment pairs.

We apply geometrical constraint and matching criterion of image texture for detecting conjugate edge segments. Candidates of conjugate edges are those of detected by processes of section 3. Figure 4 shows the flow of proposed process.

(1) Filtering with geometric constraint

In DPW system, the orientation parameters of an aerial image is known, and can be used as geometric constraints for candidates of conjugate points or line segments. In this step, it is assumed that all edge segments that have the same epipolar line in the other image are the candidates of conjugate pair. Combinations of conjugate candidates are filtered with the following geometric constraints.

(a) Edge candidates in the other image are supposed to exist in a certain search width, which means that there is an upper limit to parallax between conjugate edges.
(b) Conjugate edges have certain common section length across epipolar lines. In other words, the distance in real world between sharing section must exceed a specified threshold.
(c) Altitude calculated with edges must be within a pre-defined range. Anything out of this range will be rejected.
(d) Angle of elevation must be less than a specified threshold. This condition gives priority to detection of edges that form the top of buildings.

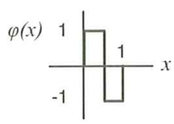


Figure 3. Haar's mother function

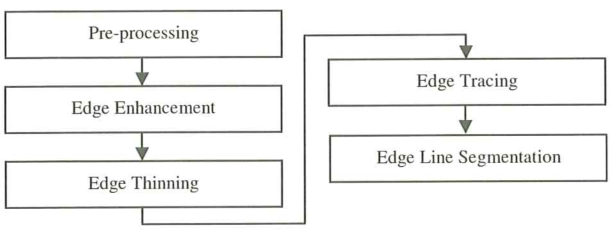


Figure 2. The process flow of edge detection

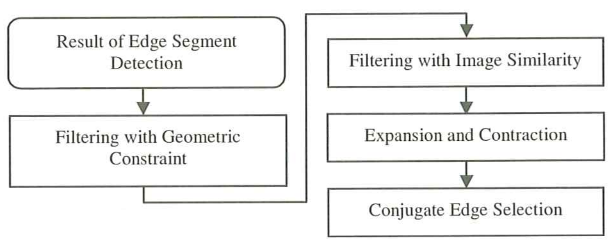


Figure 4. The process flow of edge matching

(2) Filtering with image similarity

In this phase, candidates of conjugate edges are further filtered with evaluation of image similarity, which is based on the correlation coefficient calculated with the templates shown in Figure 5. The templates are chosen to be both sides of edge segment between the overlapped range and the shape of template is parallelogram, which is parallel to edge segment and epipolar lines. The higher correlation coefficient of two templates is adopted for evaluation criterion and the combination with highest one is selected for candidate of conjugate edges.

(3) Expansion and contraction

Normally, the conjugate lines are not of the same length. The shorter edge of supposed conjugate edges is first extended to that of the longer one. If the image similarity described above is improved by this operation, expansion of edge is considered to be valid. This operation is executed for all candidates of conjugate edges. The next step is equalizing length of conjugate edges (contraction process). In this stage specific edges in other image have a possibility of referencing by multiple edges as a candidate of conjugate edges, so in such a case, the above processing is performed after copying those edges.

(4) Conjugate edge selection

As the last phase, conjugate edges are selected by the following criterion

(a) An edge pair is adopted unconditionally when in the assumed conjugate edge candidates refer to each other as the highest edge and not referenced by other edges.
(b) If an edge A in a image is referenced by multiple edges B_i in another image, and edge A has connectable edges in its neighborhood with a specified distance threshold, and the same condition applies to a B of B_i, the pair $\{A,B\}$ is selected.
(c) For cases similar to case (b), except that there are no connectable edges, the edge before being copied in step (3) is used, and the pair that gives the highest similarity is selected.

V. CEM ENHANCED WITH EDGE CONSTRAINT

Coincidence enhancement method

Coincidence Enhancement is an extension of Hebb's rule for self-organized process in neural network modeling. This process can be modeled by using the principle of competition and consensus. CE model can realize smooth mapping between input signals and output pattern. When take the left stereo image as input and the right as output, this concept is applicable to stereo matching.

The concept of stereo matching with CEM is illustrated in Figure 6. In the competition phase, each pixel or local area in image A tries to find an optimal poison in the search area. A shift vector is formed by connecting the original position in A and the optimal position in B. In consensus phase each shift vector is modified with surrounding shift vectors in consensus area. These operations are repeated at each local area. The final result becomes a nonlinear mapping between these two images.

For competition process it is very important to use the appropriate correlation function. Normally absolute value of difference between template area's brightness is used in terms of computational cost. To enhance object's features, pre-processing processes are necessary. The complexity Index (CI) calculated by Eq.2 and Eq.3 proposed by Kosugi et al. [5] is a good example. CI is capable of improving CEM objected for urban area. This has been ascertained by verification test.

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{2}$$

$$CI = \sum_{i \in S} \left(\left| \frac{\partial \nabla^2 f}{\partial x} \right| + \left| \frac{\partial \nabla^2 f}{\partial y} \right| \right) \tag{3}$$

The method of consensus operation depends on processed object. For example, it is efficient to use spline function for the matching objected for mountainous topography. Convergence property in matching by consensus operation mainly depends on the range of consensus area, shape of area, method of weighting. In this study we apply cross-shape consensus area and use median shifting vector as a consensus result.

The balance of competition and consensus operation is also an essential factor. If the influence of competition is too large,

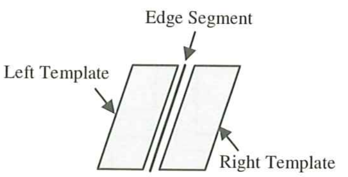


Figure 5. The layout of template for correlation calculation

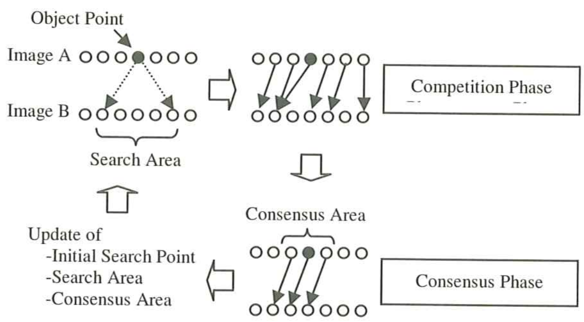


Figure 6. The principle of CEM

the number of mismatch area will increase. On the contrary, if the degree of consensus is too strong, matching results will be too smooth. Therefore these parameters must be adjusted properly according to the number of iteration.

Extension of CEM with edge constraints

CEM is a very effective approach for smoothly transforming object's matching. Since our target is for stereo matching in urban area, we make use of edge constraints to improve the performance of CEM.

Global search in initial stage of competition phase is very important, because if it fails to find approximate matching positions in large area, CEM cannot recover such errors due to its characteristics. We can avoid this problem with edge matching results described in Section 4. Up to now, to avoid of error in initial global search has been executed such as using more extensive area in competition phase, utilization of edge segment will offer alternative approach.

Conjugate edge segments gives shifting vectors and can be used as immovable area in both competition and consensus operation. This constraint brings prevention against deforming of some building's edges.

Furthermore, edge segments that are detected with the approaches described in Section 3 but not matched at Section 4 can also be used as efficient information in consensus phase. Because most of the edges detected in Section 3 are regarded to have nearly linear shift vector except for occluded region by collinear condition. The rules of edge constraint applied in CEM are as follows.

(i) Conjugate edge segments are immobile in competition and consensus operation.
(ii) Non-conjugate edge segments have uniform parallax shift.
(iii) Parallaxes of points put between edges don't exceed those edges.

VI. EXPERIMENTS AND RESULTS

Experimental environment

Experiments have been performed with stereo aerial imageries of urban area. Figure 7 shows a small portion of the stereo images. The image sizes are 700 x 700 respectively and the resolution is about 20 cm per pixel. There are several occluded areas and hidden regions in the shades.

Edge detection results

Figure 8 and Figure 9 show edge tracing results derived from images processed by SUSAN operator and Haar' s wavelet respectively. Similarly Figure 10 and Figure 11 show the results of segmentation operation of edge tracing. The final results of edge segment matching are illustrated in Figure 12 and 13. From the experiments, both approaches produce almost valid results, however we observe that SUSAN operator can detect more valid edge than wavelet does. This is because that at the stage of calculating binary edge images with wavelet, some of the important information on edges was lost by binarization process. On the other hand in SUSAN, edge components are detected by linking interest feature points, so that the points with weak interest feature can keep edge's information. We would try to apply this concept for using in wavelet operation as feature works.

Stereo matching with CEM

To verify the influence of balance between competition and

(a) (b)

Figure 7. A pair of stereo images. (a) Left image, (b) Right image

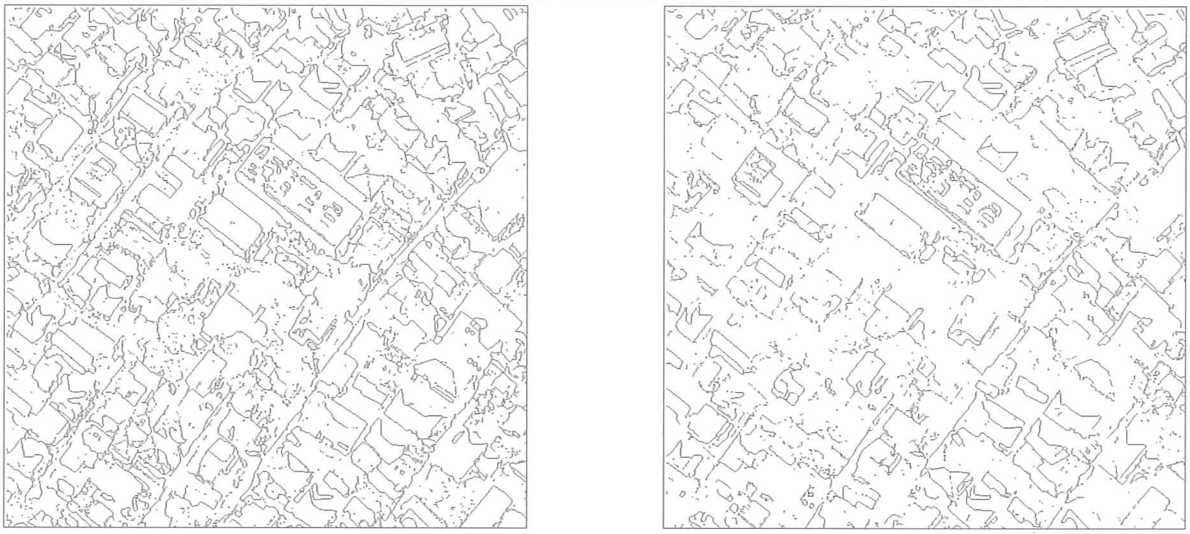


Figure 8. The edge tracing results by SUSAN operator

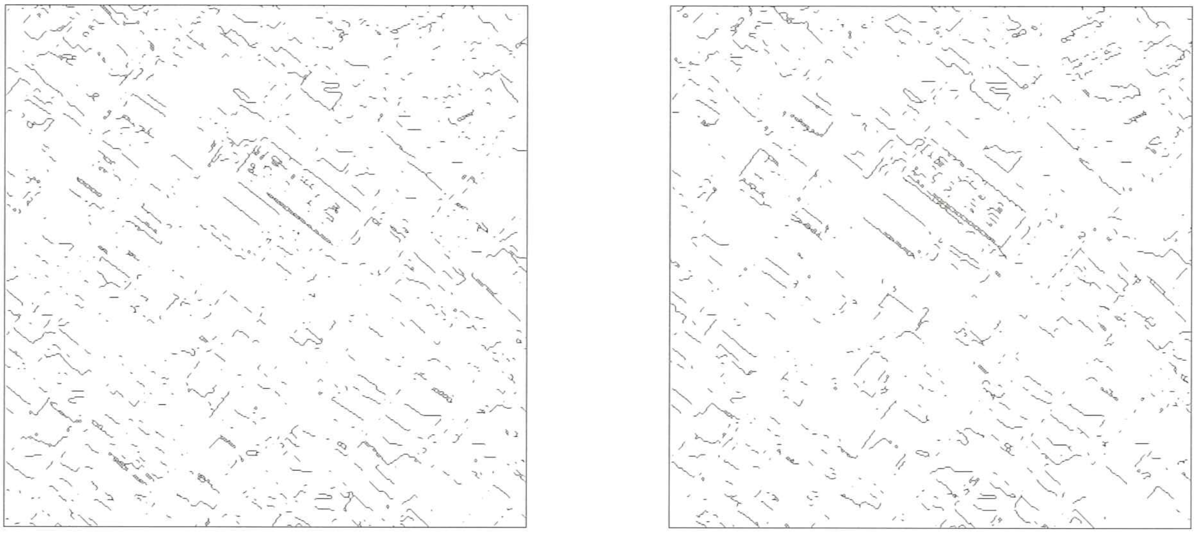


Figure 9. The edge tracing results by Haar' s wavelet

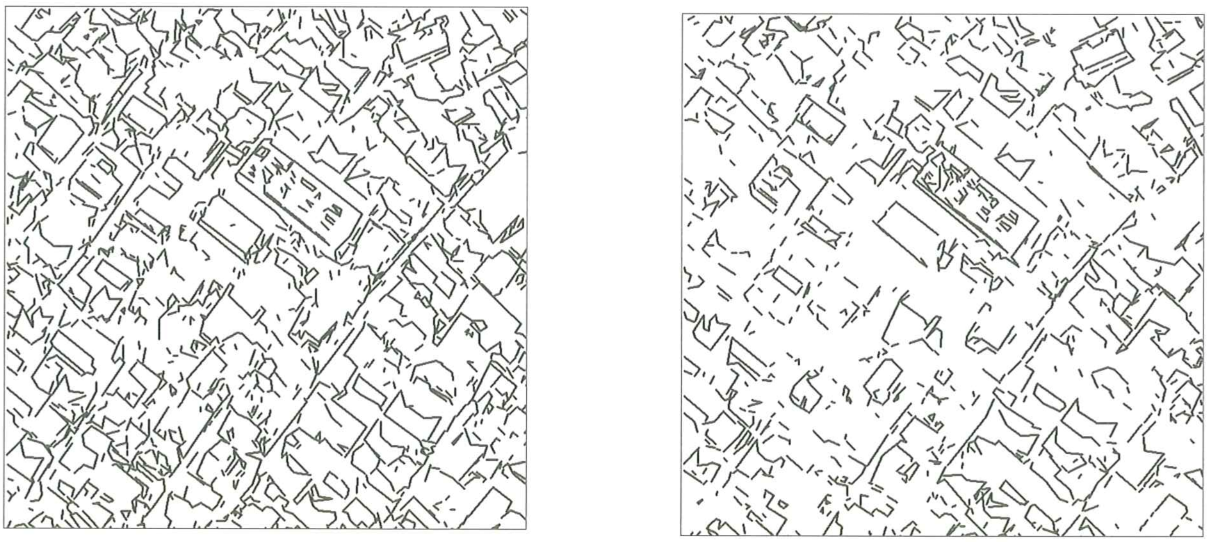


Figure 10. The result of line segmentation (SUSAN operator)

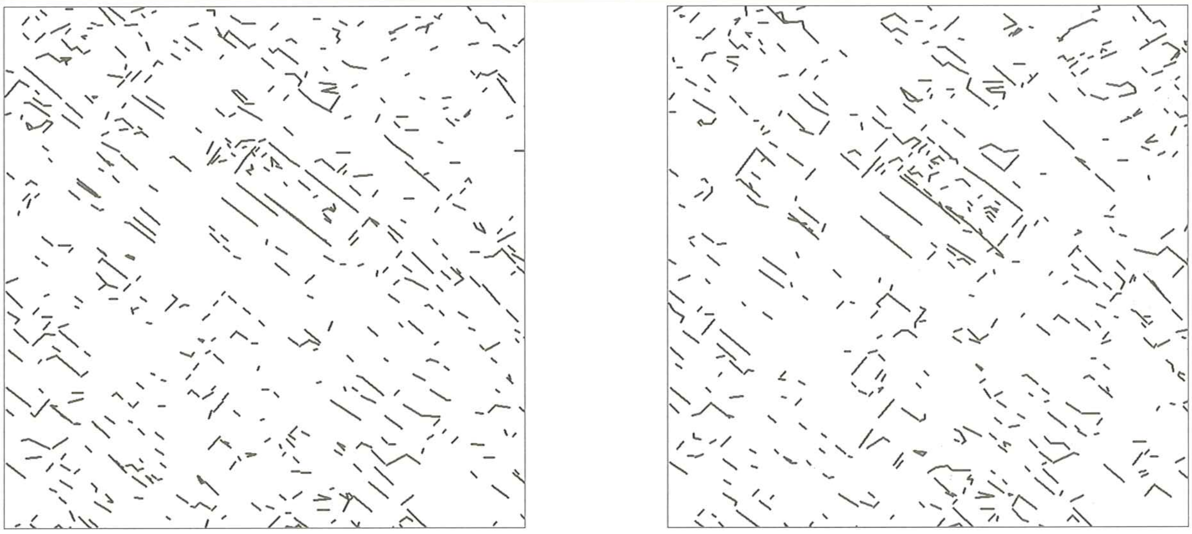


Figure 11. The result of line segmentation (Haar' s wavelet)

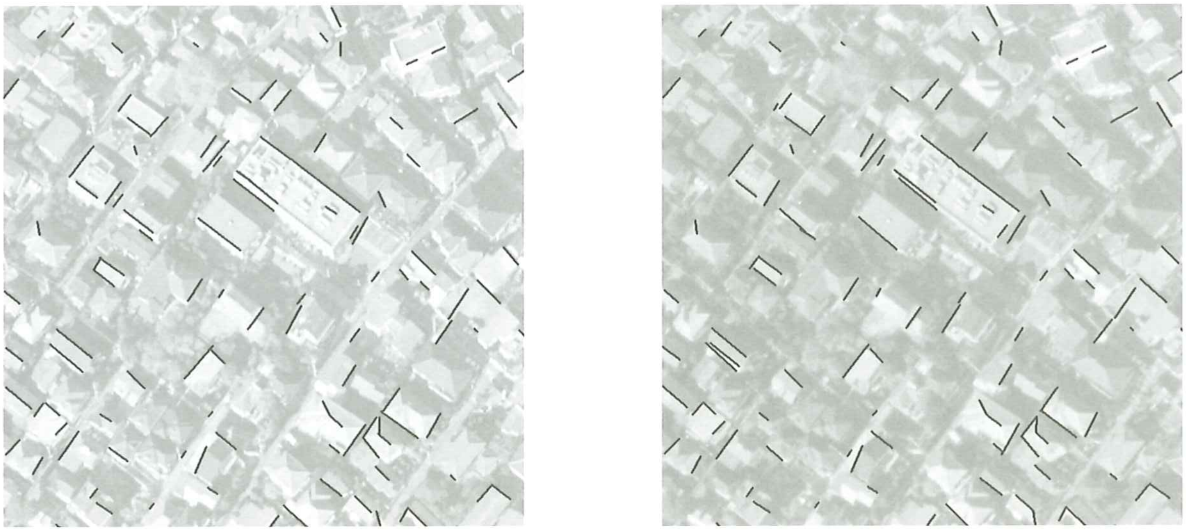


Figure 12. The result of edge matching (SUSAN operator)

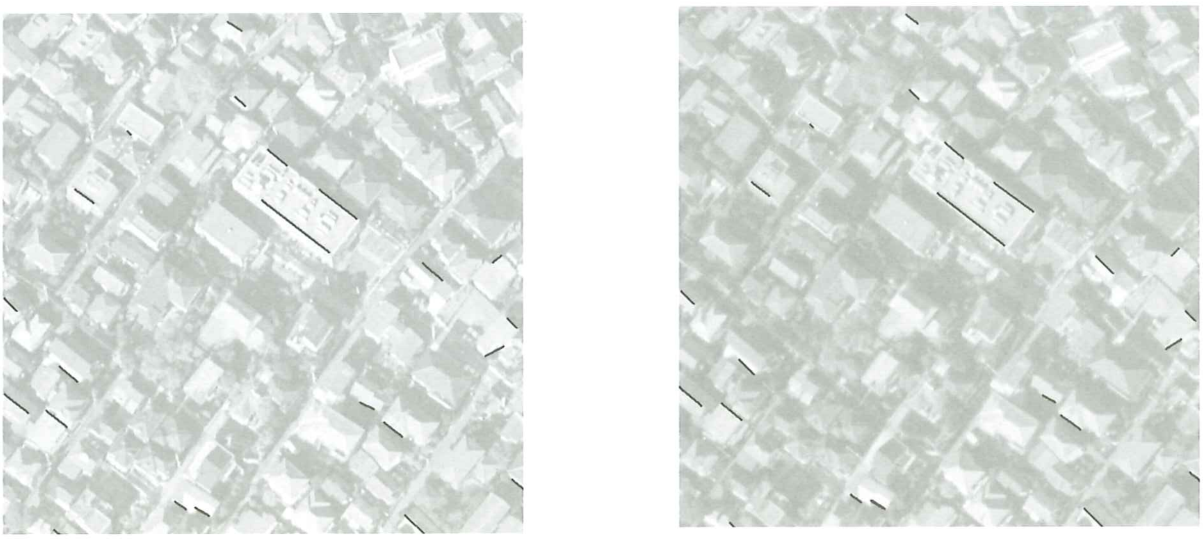
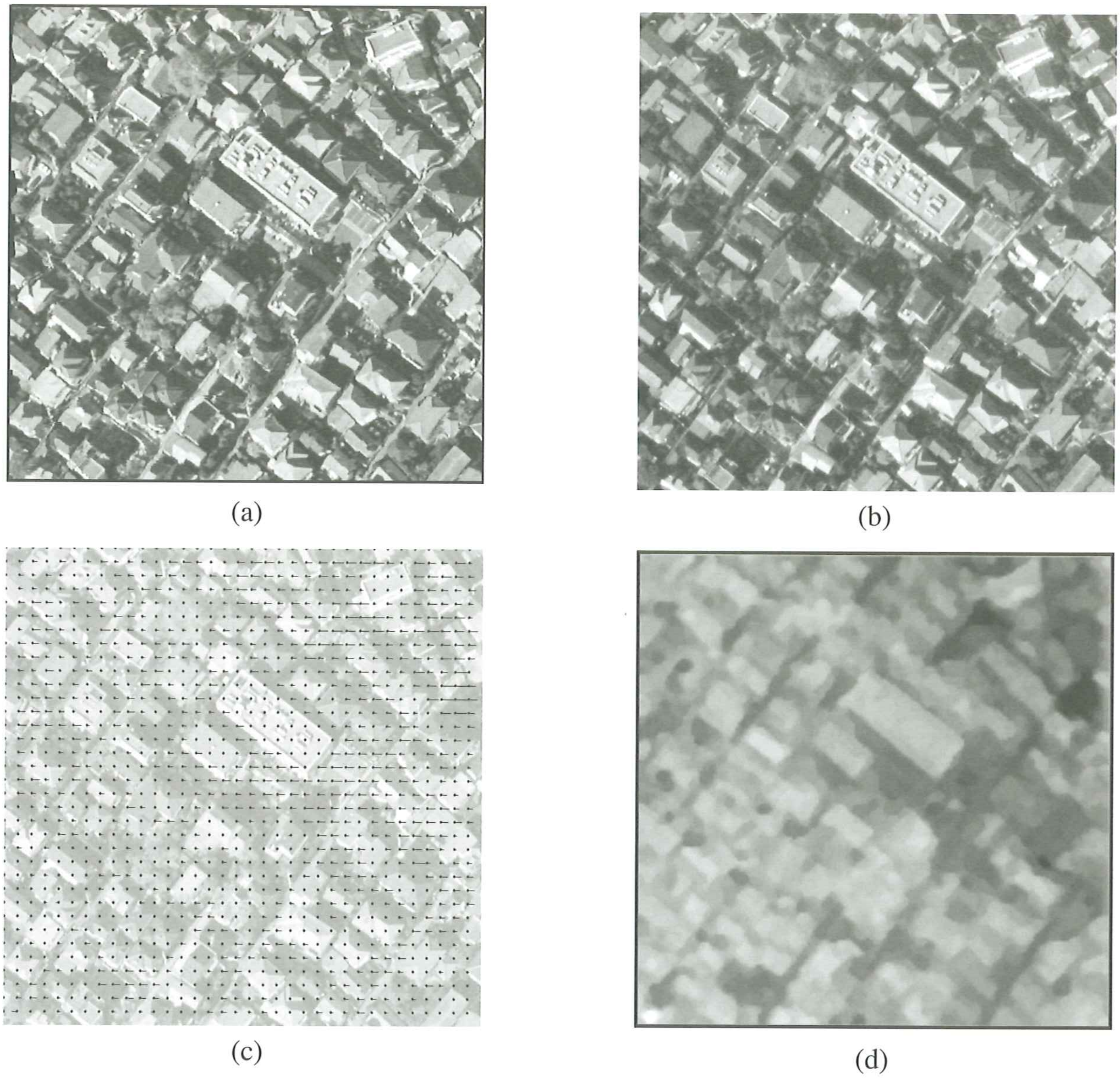


Figure 13. The result of edge matching (Haar' s wavelet)

consensus process, comparison experiments was conducted. Figure 14 shows the mapped image and shift vector from the left image to the right one by CEM with standard processing parameters. The depth map is also illustrated, which is calculated by shift vector and overlaid onto the left image. Figure 15 shows the result of a case when consensus operation is dominant compared to competition operation. It is noticed that enhanced consensus operation produces smoother shift vectors, yet overly strong application is not suitable for urban area.

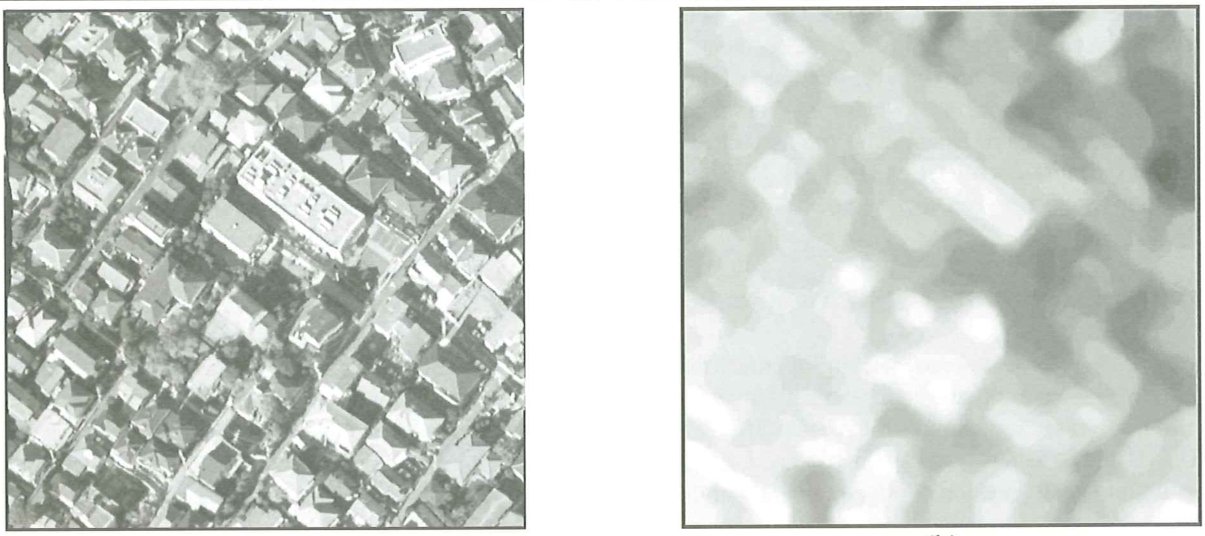
Figure 16 shows the images processed by CI operation with Eq.2 and Eq.3. The results processed by CEM with these images are illustrated in Figure 17. Figure 16 shows that CI operator enhances areas where brightness change is strong, especially for the edges of buildings. With regards of edge parts, CI operator has similar effect to the sum of absolute value of first-degree differential calculus in local area. By comparing Figure 17 with Figure 14, it is observed that the mapping result with CI operator gives almost acceptable depth map but leads more deformation. This result is caused by the areas of flat textures formed by CI operator, which consequently causes mismatching by local minimum. Therefore when using CI operator, techniques such as using only characteristic region for matching and interpolating other areas should be considered.

Figure 18 is the result of CEM with matched edge as constraints. Conjugate edges detected by SUSAN operator are used. In comparison with result in Figure 14, it is observed that shifts of parallax are improved at boundaries of buildings (See the comparison result of partially enlarged depth map images in Figure 19). However some areas have local disturbances because of consensus operation, which tries to decrease



(a) (b) (c) (d)

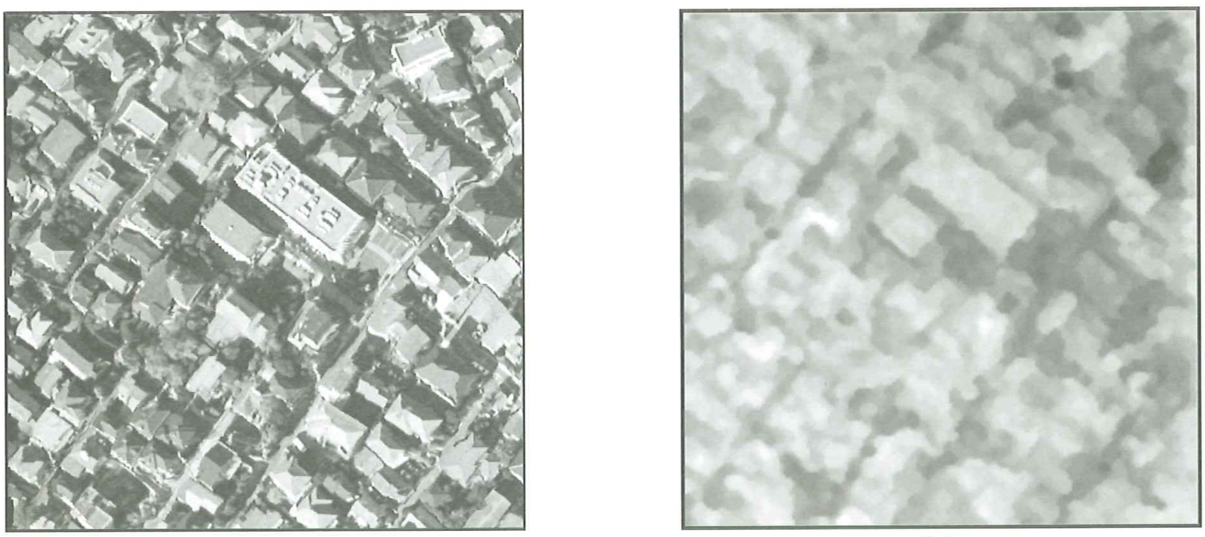
Figure 14. The CEM result with average parameters. (a) Mapped left image; (b) Original right image; (c) Shift vector; (d) Depth map projected to left image



(a) (b)

Figure 15. The CEM result with enhanced consensus operation. (a) Mapped left image; (b) Depth map projected to left image

Figure 16. The application result with CI operator



(a) (b)

Figure 17. The CEM result with CI operated images. (a) Mapped left image; (b) Depth map projected to left image

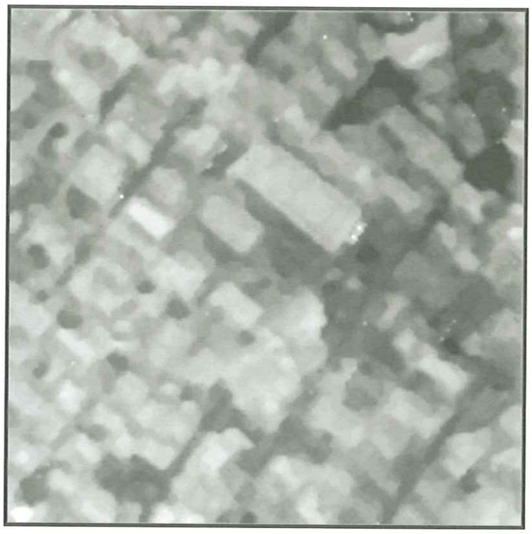
abrupt parallax shift. This is because that CEM process with edge constraint cannot deal only with edge segments. It also has some effects on other areas. In the next step of this study, perfect separation of process between edges and other areas must be investigated. For example, at the first stage, match only edge segments with CEM. In later stages, match adaptively the areas between edge segments.

VII. CONCLUSION AND FUTURE WORKS

We have introduced a new stereo matching approach that combines edge matching results and nonlinear matching with CEM. In edge detection and matching, two types of different operator, SUSAN and Haar' s wavelet are investigated. Experimental results show that detection of edge segment with SUSAN operator can provide more valid results. In the case of wavelet, another process has to be considered instead of binarization. It is verified that the proposed edge matching technique works satisfactorily.

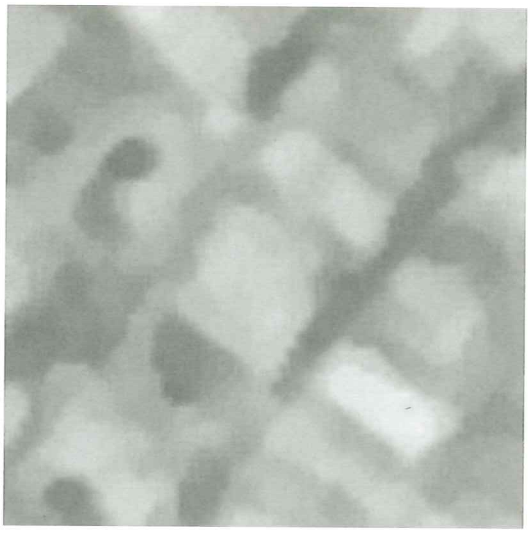
We presented an example of stereo matching result in urban area with CEM, where the conjugate edge segments are utilized as global search information and immovable points in CEM process. Experimental results show that the balance of parameters between competition phase and consensus phase is important. The efficiency of CI operator was tested and it is observed that CI operator can produce promising results for stereo matching of buildings. However more improvements are needed. CEM with edge segments constraints showed improving around building edges. In the present, CEM process of edge points and other areas are not definitely distinguished, therefore local disturbances occur at some areas. More inves-

(a)



(b)

Figure 18. The CEM result with edge support. (a) Mapp left image; (b) Depth map projected to left image



(a)

(b)

Figure 19. The comparison of projected depth map (partially enlarged images). (a) Depth map by CEM; (b) Depth map by CEM with edge support

tigation is necessary about this subject.

For further improvement, handling of edge points and other areas have to be separately processed in CEM with edge constraints. It is also important to investigate adjusting processing parameters adaptively for object's size in image.
In the next stage, we will try to study detecting buildings with edges and depth map derived from in this study.

REFERENCES

[1] Sakamoto, Mitsuteru, Osamu Uchida, Takeshi Doihara, Kazuo Oda, Wei Lu and Masayoshi Obata, 2000, Geo-Plotter a Softcopy Mapping System for Low Cost Digital Mapping Process, *IAPRS XXXIII-B4/2*, pp.889-892.

[2] Sakamoto, Mitsuteru, Takeshi Doihara and Yukio Kosugi, 2000, A Study on Improvements of Stereo Matching Techniques for Urban Area – Automatic Estimation of Camera Parameters and Utilization of Edge Segments, *Technical Report of IEICE*, 100(499):115-120.

[3] Smith, S. M. and J. M. Brady, 1997, SUSAN -A New Approach to Low Level Image Processing, *Int. Journal of Computer Vision*, 23(1):45-78.

[4] Kosugi, Y., M. Sase, H. Kuwatani et al., 1993, Neural Network Mapping for Nonlinear Stereotactic Normalization of Brain MR Images, *Journal of Computer Assisted Tomography*, 17(3):445-460.

[5] Kosugi, Yukio, Munenori Fukunishi, Mitsuteru Sakamoto, Wei Lu and Takeshi Doihara, 2001, Detection of Sheer Changes in Aerial Photo Images Using An Adaptive Nonlinear Mapping, *Proc. DMGIS 2001*, pp.145-146, Bangkok.