INTEREST OPERATOR AND FAST IMPLEMENTATION

Lü Yan
Institute of Digital Photogrammetry (IDP)
Wuhan Technical University of Surveying and Mapping (WTUSM)
Wuhan, The P. R. China
Commission III

Abstract

Good feature extraction is a reliable preprocessing step for good image matching. Three kinds of interest operators for feature extraction - the well-known Moravec's operator, the Förstner's operator, and a new operator - are compared each with in this paper. The experiments show that the new operator is superior to the former two in respects of speed and accuracy. The fast implementation of feature extraction on a standard serial computer with language C is briefly explained.

1. Introduction

It is well known, that image matching is a key step for computer vision and digital photogrammetry. Many efforts have been made by computer scientists and photogrammetrists to improve the precision and reliability of image matching. One way to do this is to use feature-based matching instead of area-based matching, or to integrate these two matching methods. Actually, area-based matching point by point, e.g. Least Squares Matching, cannot reliably be used in featureless areas of nearly homogeneous brightness or in areas with severe scale changes. Good feature extraction is necessary for feature-based matching and image understanding. Therefore selecting good features and providing reliable and accurate approximate values for succeeding fine correlation attracts ever increasing interest. In the field of computer vision and pattern recognition many different operators have been developed for feature extraction. This paper only discusses on the widely used Moravec operator /Barnard 1982/, and the relatively new Förstner operator, which is based on a deepgoing theoretical research, and is implemented in the software package InduSURF/Schewe 1987/. Then we propose a new operator, that includes two versions, one of which is especially improving Förstner's operator. The newly developed operator has some advantages over the two operators mentioned before. One of the advantages is its simplicity, which allows it to run rather fast. Another feature is that the threshold it requires, can be determined easily. Furthermore, with the new operator one does not need considering the weighted centre of gravity within the window, which must be calculated in the case of Moravec's and Förstner's operators to obtain an optimal point /Förstner 1987/.

We first outline the basic points of these operators, compare them, and give some examples. Followings briefly describe the fast implementation of feature extraction on standard serial computers with the C language.

2. Different Interest Operators

2.1 Moravec's Operator

Moravec's interest operator /Moravec 1979/ searches for points that have high variance between adjacent pixels by measuring the distinctness of a local piece of the image from its surroundings. The main points for the calculation can be summed up as follows :

a. Interest value (IV) for each point : minimum of four variances (horizontal, vertical, and two main diagonals) at one window, e.g. 5 by 5 (Fig.1). Variances are calculated by summing up the squares of the differences of gray values of adjacent pixels along the four directions respectively.

b. Interest points (IP) are those with IV > T (threshold). The threshold T is chosen empirically to produce some fraction of the total image points.

c. Suppression of local non-maxima, i.e. checking IV of each interest point against its neighbors to try to avoid bunching points in highly-textured area. The size of the suppression window is a free parameter, which is dependent on the expected density of interest points, e.g. the window size from 5x5 /Charles 1983/ to 11x11 /Collins et al. 1987/ have been used.

In brief, Moravec's operator selects points with the largest-minimun variance of gray level differences in four directions. It is conceivable, that the process of search and calculation point by point is timeconsuming, therefore it is only applied to one image in the pyramid of reduced-resolution versions, which is also for filtering high-frequency noise /Charles 1983/. As for the empirical determination of the threshold T, there are no certain rules or recommended experience values. Research by Collins on the utilization of transputer arrays for real time matching showed, that the use of low threshold values in Moravec's operator gives poor consistency, and difficulties in matching caused high computational cost. Using a high threshold to overcome this problem leads to a non-uniform utilization of the transputer array again at high computational costs. Their prelimilary conclusion is that Moravec's operator is inadequate for satellite imagery, and one should study modifications of this operator, as well as alternative interest operators /Collins et al. 1987/.

It should be pointed out, that Hannah /1980/ modified this operator by consideration of both ratios of the variance in the four directions, and ordinary image intensity variance over large areas. This modified operator seems to locate a better selection of both strong and subtle features / Barnard 1982/, but it still requires high computational costs.

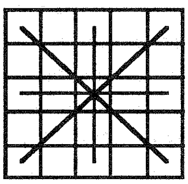


Fig.1

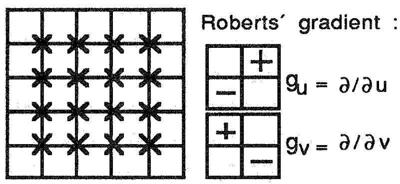


Fig.2

2.2 Förstner's Operator

This operator searches for distinct points, corners, and centres of circular image features. The extraction consists of two steps : the window selection and the feature location. In many papers /Paderes et al. 1984, Förstner 1986, Förstner et al. 1987/ the authors already explained the origin, theoretical ground and application of the operator. For simplicity we do not repeat it here. The main steps for the calculation can be summed up as follows :

a. Calculations of Robert gradients /Rosenfeld and Kak 1982/ and covariance matrix $\underline{Q}$ window by window (Fig.2)

$$\underline{Q} = \underline{N}^{-1} = \begin{bmatrix} \Sigma g_u^2 & \Sigma(g_u g_v) \\ \Sigma(g_u g_v) & \Sigma g_v^2 \end{bmatrix}^{-1} \qquad (1)$$

b. Interest values q and w (weight) for each point :

$$q = \frac{4\,\mathrm{Det}\,\underline{N}}{(\mathrm{tr}\,\underline{N})^2} \qquad w = \frac{1}{\mathrm{tr}\,\underline{Q}} = \frac{\mathrm{Det}\,\underline{N}}{\mathrm{tr}\,\underline{N}} \qquad (2)$$

c. Interest points are those points, whose q and w are greater than the given threshold /Förstner et al. 1987/ :

$$w = \begin{cases} w & \text{if } q > q_{lim} \text{ and } w > w_{lim} \\ 0 & \text{else} \end{cases} \tag{3}$$

$$q_{lim} = 0.5 \sim 0.75 \qquad w_{lim} = \begin{cases} f \cdot w_{mean} & f = 0.5 \sim 1.5 \\ c \cdot w_{med} & c = 5 \end{cases}$$

w_{mean} - the average of the weights of all window positions in the image.

w_{mean} - the median of the weights taken over the whole image.

f, c - experience values.

d. Suppression of local non-maxima in a spiral manner within a progressively enlarged window in order to avoid clustering relative maxima (Fig.3). The size of the suppression window is also a free parameter. The authors /Förstner et al. 1987/ think, that for feature based matching windows of 5x5 or 7x7 pixels lead to satisfying results in any cases.

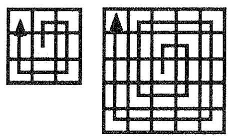


Fig.3

From the formulas (1) (2) (3) and the Fig.2 it is not difficult to find out that the process for search and calculation point by point is also time-consuming, and the determination of the proper threshold, especially for w, is both timeconsuming and a little troublesome. According to Luhmann et al. /1986/ Förstner's operator requires more time than Moravec's. Furthermore, as the author concedes, with steps (a~d) it can only gain optimal windows, not points / Förstner 1986, Förstner et al. 1987/. Therefore there must be further search for optimal points within the optimal windows. In fact, the larger the window, the greater the bias of the centre point at the window from the optimal point is, because of the influence between individual pixels of the window. By only applying 3 by 3 windows instead of 5 by 5 or 7 by 7, the problem (only this problem) disappears (see the variant of Förstner's operator of Fig. 8-7 and Fig. 9-7).

Incidentally, Hannah's operator /Hannah 1974/ searches for points, whose autocorrelation function of the gray level is steep in all directions. As the covariance matrix directly measures the curvature of the 2D-autocovariance within the window, Förstner's operator, except for the normalization, is essentially identical with Hannah's operator, but it is more simple to be calculated /Förstner 1986/.

2.3 A New Operator

The new operator is divided into two steps. In the first step with a **ground operator**, which has rather simple form, one can selects those points, which are to be used in the second step. As shown later, the first step leads to a great reduction of data.

2.3.1 The **Ground Operator**

At each point, the four gradients to the neighboring pixels are calculated (Fig.4). A point is kept only for the second step, if at least two of the absolutes of the gradients are larger than a threshold dg. Empirically, the lower bound of dg for the ground operator is between 5~10 with 3 by 3 local average filtering, and between 8~15 without any filtering.

Fig.8-2 and Fig.9-2 illustrate one of the functions of the ground operator : it leads to a great information reduction, e.g. the total number of pixels considered, is reduced from 4900 (70x70) to 300 for sample0, from 4900 to 602 for sample1, and for Fig.10, which is taken from the correlation test organized by Working Group III/4 of ISPRS, from 57600 (240x240) to 7274, 9799, 10020 respectively for photo 3, 5, 19.

As for the determinnation of the threshold dg it can be manully set to achieve visually acceptable results, before one have found a method, in which the optimal threshold dg for different images could be automatically defined.

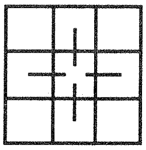


Fig.4

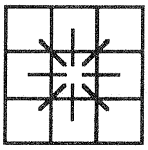


Fig.5

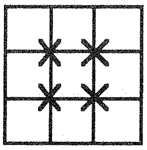


Fig.6

2.3.2 Selection of Interest Points

For the second step two versions are presented as below.

Version I :

I-1. Interest values for points selected by the ground operator (Fig.5) :

$$IV = \sum_{i=1}^{8} \text{abs}\,(dg_i) \tag{4}$$

where dg_i is the difference of gray levels between two adjacent pixels.

I-2. Suppression of local non-maxima in the manner, in which the compared window is progressively enlarged. The function is the same as those of Moravec and Förstner operator. Once an IV within the compared window (the light hatched parts of Fig.7, which indicates suppression windows are 3x3 and 5x5) is greater than the centre's value of the suppression window, the comparison stops, and the centre's IV is set to zero. The size of the suppression window can be selected according to the density of interest points expected for the results.

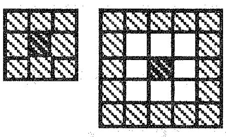


Fig.7

The second version is also based on the above-mentioned ground operator. This version can be considered as an improvment of Förstner's operator in respects to speed and accuracy.

Version II:

II-1. Interest value: In accordance with the ground operator q and w values are calculated with formula (1) and (2) (as for Förstner's operator), only using 8 adjacent points (Fig.6) rather than 5 by 5 or 7 by 7 windows.

q_{lim} = 0.32 ~ 0.5 which could be selected as constant, corresponding to the ratio 3.2 ~ 2.4 of the semiaxes of the error ellipse.

II-2. Suppression of local non-maxima by the same method refer to version I.

It should be reminded, that the threshold of w is no longer necessary, because of using the ground operator. The reason is, that the q value must be small, if w is small. As a result, only one fixed threshold of q is enough, e.g. 0.5. Significantly, due to applying the ground operator one cannot only reduce numerous calculations, but also determine the threshold dg easily (substitute for the troublesome threshold of w in Förstner's operator). Furthermore, as the interest values q and w are calculated for points, not for windows, the bias of the centre of the window from the optimal point disappears too (Fig.8-4, Fig.9-4). The other advantages of Förstner's operator, such as distinct points, and corners, which allow for accurate correlation, as well as that the selection can be evaluated by the standard deviation of the estimated position, are still preserved.

In the case of using Förstner's operator i, in which all points are calculated pixel by pixel, it is possible that the q value is rather large, even though w value is very small. Fig.11 gives an extreme example. The upper three windows are taken from window 3), 14) and 15) of Fig.5 in Förstner's paper / Förstner 1986/. They can be applied for explanation of two facts : 1) for Förstner's operator the threshold for w is indispensable; 2) the merits of using the ground operator. For each template a variant is made, whose gradients have been reduced significantly, but are still proportional to the original ones : e.g. the original values of 0, 80, 160, 240 have been reduced to 0, 1, 2, 3 for the variants (see graylevel matrices of Fig.11). According to formulae (1), (2) one can calculate interest values w and q for these six windows. Obviously, the q values of all six templates are large enough to be selected as optimal windows. But there exists a world of difference between corresponding w values of the original and its variant. As mentioned above, the point with the small w can also have a large q. From images with small w we can directly perceive through our senses, that they are inadmissible as optimal windows. This shows, that it is essential for using Förstner's operator over the whole image to take a threshold for w, else in areas void of the features many "feature points" could occur too. In other words, for Förstner's operator the threshold for w is indispensable. However, if the ground operator is used for the same six windows, three windows with small w must certainly be eliminated. Clearly, the computation of the ground operator is much simpler.

It should be pointed out, that if images contain many linear structures perpendicular to the epipolar lines (such as ISP photo 19 in Fig.10) and camera model is known or epipolar constraints are available, Version II just as Föstner's operator is not appropriate, because both operators cannot selected points on edges. Points on edges, perpendicular to the epipolar lines, however, could achieve very high matching accuracy. In this case Version I could be considered as alternative. As for the case with the epipolar constraint the author will prepare another paper /Lü 1988, also see Zhang 1988/.

Additionlly, although the determination of threshold dg of the new operator is simpler than of threshold w of Föstner's operator, it needs further investigation, how to automatically and optimally define it for different images.

3. Results and Programming with the C Language

Using the three operators, we have experimented on all twelve pairs of photos of the correlation test organized by Working Group III/4 of ISPRS, and others, such as Simens's star, children's portraits, and so on. The results show, that the new operator is superior to the other two operators in respects of speed and accuracy. As space is limited, it is impossible to publish all the results we calculated. Here we have listed only a small part of them, as shown in Tab.1 and Fig.8 ~ Fig.10. The general impression of all calculations is, that the results of Förstner's operator were better than Moravec's, and the ones of the new operator better than Förstner's. The computation time consumed by the new operator for any image was less than by the other two operators.

By the way, sample0 and sample1 were designed as 70x70 pixels, which is same size as the test photo in the benchmarktest /Förstner et al. 1987/, in order to compare speed (Tab.1). Besides the programming technique the used programming language may be of importance to speed up the computations. By applying the C language instead of FORTRAN one can use the advantages of pointer arrays and work directly at the address level to avoid numerous data exchanges, which especially occured in programming Förstner's operator in sequential manner. Certainly C also has many other characteristics, the pointer is one among other things, but one of the most important. It could be said, for digital photogrammetry, which includes a lot of array operations, language C is highly recommendable.

4. Conclusions

This paper presented a new interest operator. The results achieved with different images demonstrate, that the new operator is superior to Moravec's and Förstner's operators in respect to speed and accuracy. Version II of the new operator could be considered as an improvement of

Förstner's operator, because it cannot only run much faster than Förstner's, due to applying a very simple ground operator, and exploiting the threshold dg instead of w used by Förstner's, (the former can be determined more easily than the latter), but also preserve the advantages of Förstner's operator, such as distinct points and corners, which allow for accurate correlation, as well as the evaluation of the selection by the standard deviation of the estimated position. Furthermore, the new operator does not need to search optimal points within the optimal windows any more, which must be done in case of using Förstner's or Moravec's.

Although the range of the threshold dg of the new operator is quite small, its automatic and optimal determination needs further investigation.

5. Acknowledgements

The author had been a visiting scholar of the Institut of Geodäsie and Photogrammetrie (IGP) of the Eidgenössischen Technischen Hochschule (ETH), in Zürich, Switzerland. At the time the author wrote this paper. The author wishes to thank Prof. Dr. A. W. Grün for his fervent encouragement and support, and Dr. Novak for his comment and elaborately correcting the manuscript. She would also like to express her appreciation to all other colleagues of IGP for their sincere cooperation.

References :

/Barnard 1982/ S.T.Barnard and M.A.Fischler, "Computational Stereo", SRI Internatioal Technical Note 261, 1982

/Chales 1983/ E.T.Chales, "An Analysis of Interest Operators for FIDO"

/Collins et al. 1987/ K.A.Collins, J.B.G.Roberts, A.Anthony, J-P Muller, "Transputer arrays for Real-time Feature Point Stereo Matching," Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, switzerland, June, 1987

/Förstner 1986/ W.Förstner, "A Feature based Correspondence Algorithm for Image Matching," Int. Arch. of Photogrammetry, Vol. 26-III, Rovaniemi, Finland, 1986

/Förstner et al. 1987/ W.Förstner and E. Gülch, "A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features," Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, switzerland, June, 1987

/Hannah 1974/ M.J.Hannah, "Cmputer Matching of Areas in Stereo Imagery," Ph.D. dissertation, AIM 239, Computer Science Department, Stanford University, Stanford, California, 1974

/Luhmann et al. 1986/ T.Luhmann and G.Altrogge, "Interest Operator for Image Matching," Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, switzerland, June, 1987

/Lü 1988/ Lü Yan, "A Simple Algorithm for Edge Detection," IGP Bericht, ETH, Zürich (to be published), 1988

/Moravec 1979/ H.Moravec, "Visual Mapping by a robot rover," Proc. 6th Int. Joint Conf. Artificial Intell., vol. 1, Tokyo, Japan, August, 1979

/Paderes et al. 1984/ F.E.paderes, E.M.Mikhail, W.Förstner, "Rectification of Single and multiple Frames of Satellite Scanner Imagery Using Points and Edges as Control," NASA Symposium on Mathmatical Pattern Recognition and Image Analysis, 1984

/Rosenfeld et al. 1982/ A.Rosenfeld and A.C.Kak, "Digital Picture Processing," 1982

/Schewe 1987/ H.Schewe, "Automatic Photogrammetric Car-body measurement," Proc. of 41th Photogrammetric Week, Stuttgart, 1987

/Zhang 1988/ Z.Zhang, "A New Approach of Epipolar-line Image Matching for Improving the Performance of SODAMS system," Present paper to the 16th ISPRS Congr., Commission III, Kyoto, Japan, 1988

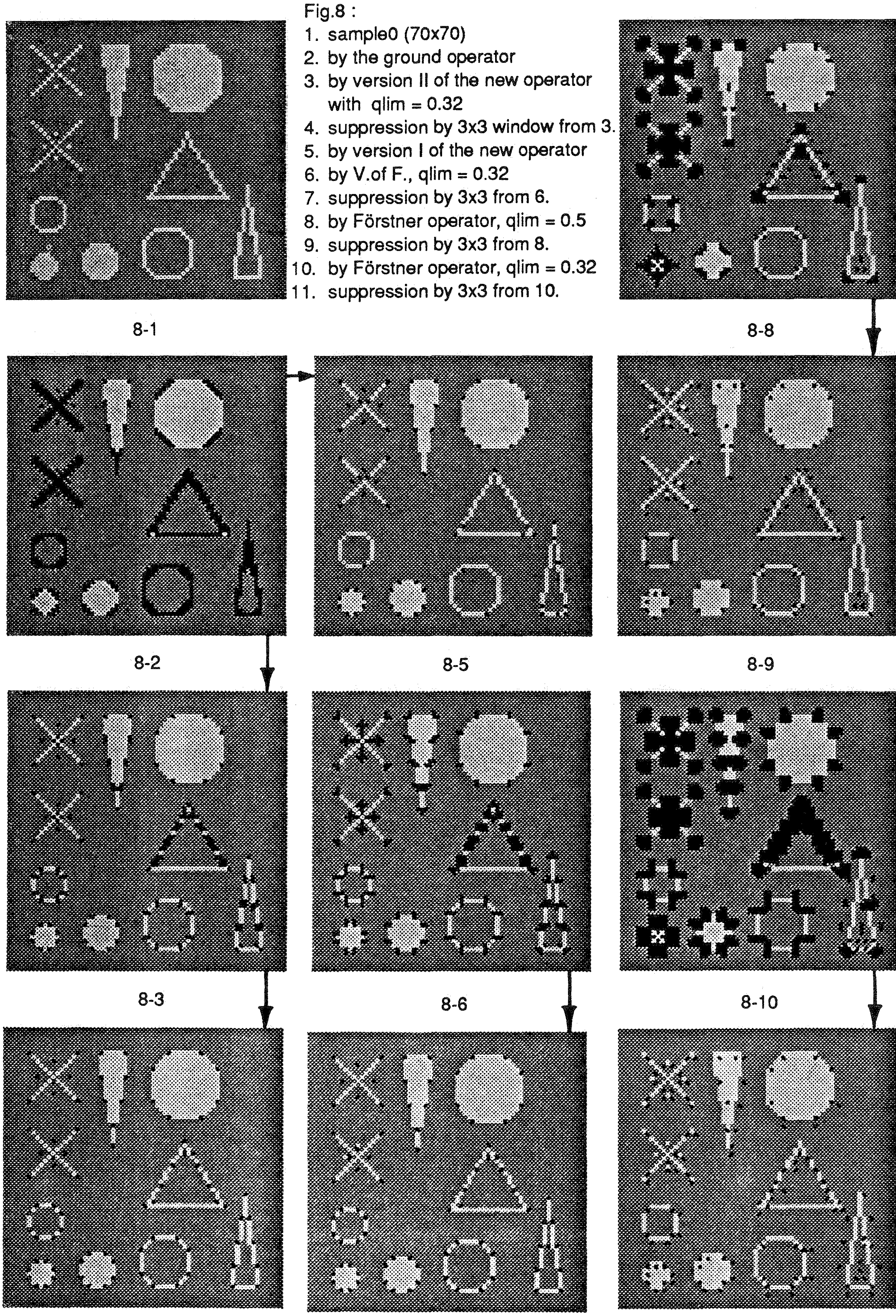


Fig.8 :

1. sample0 (70x70)
2. by the ground operator
3. by version II of the new operator with qlim = 0.32
4. suppression by 3x3 window from 3.
5. by version I of the new operator
6. by V.of F., qlim = 0.32
7. suppression by 3x3 from 6.
8. by Förstner operator, qlim = 0.5
9. suppression by 3x3 from 8.
10. by Förstner operator, qlim = 0.32
11. suppression by 3x3 from 10.

Fig.8

Fig.9 :
1. sample1 (70x70)
2. by the ground operator
3. by version II of the new operator with qlim = 0.32
4. suppression by 3x3 window from 3.
5. by version I of the new operator
6. by V.of F., qlim = 0.32
7. suppression by 3x3 from 6.
8. by Förstner operator, qlim = 0.5
9. suppression by 3x3 from 8.
10. by Förstner operator, qlim = 0.32
11. suppression by 3x3 from 10.

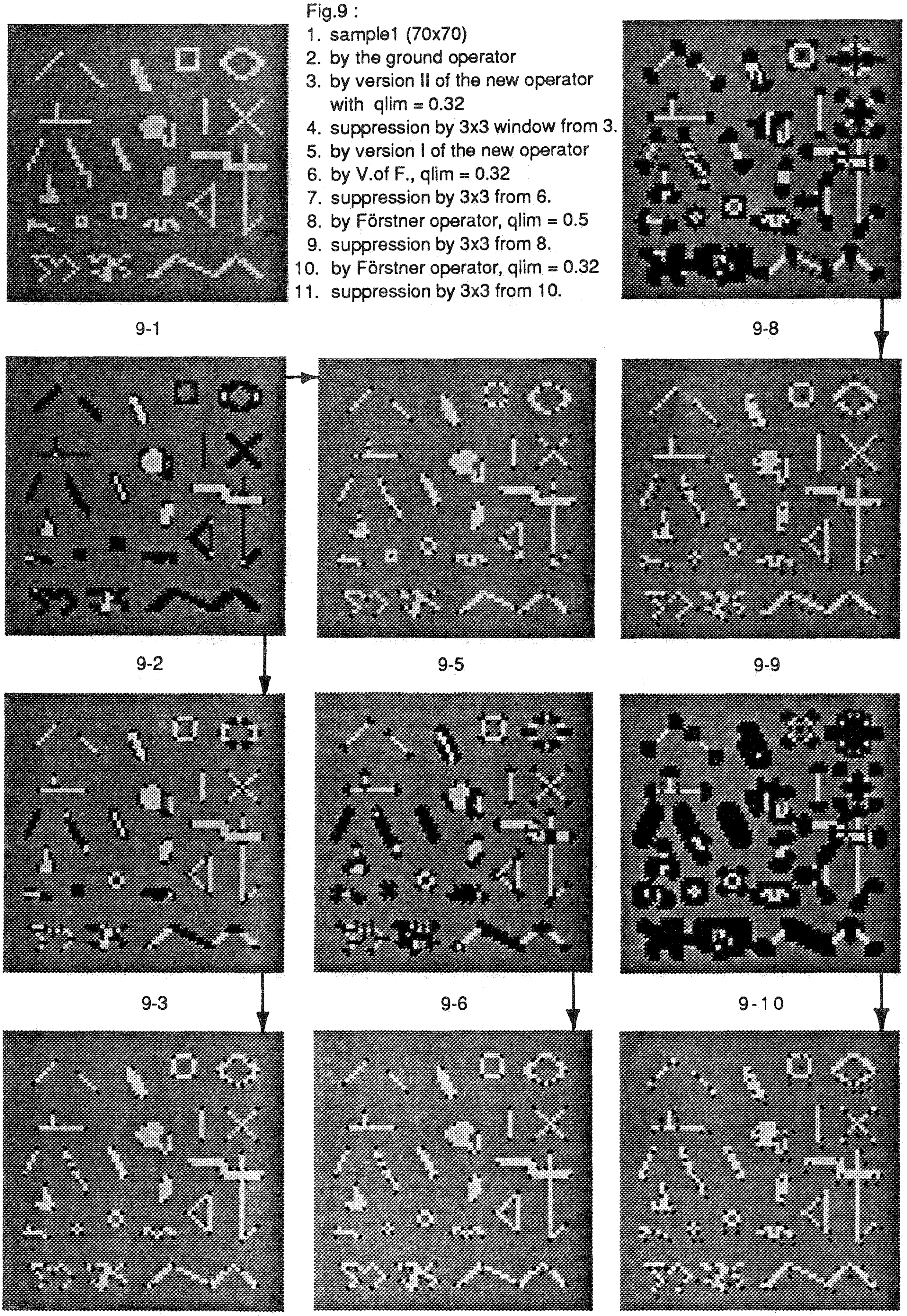


Fig.9

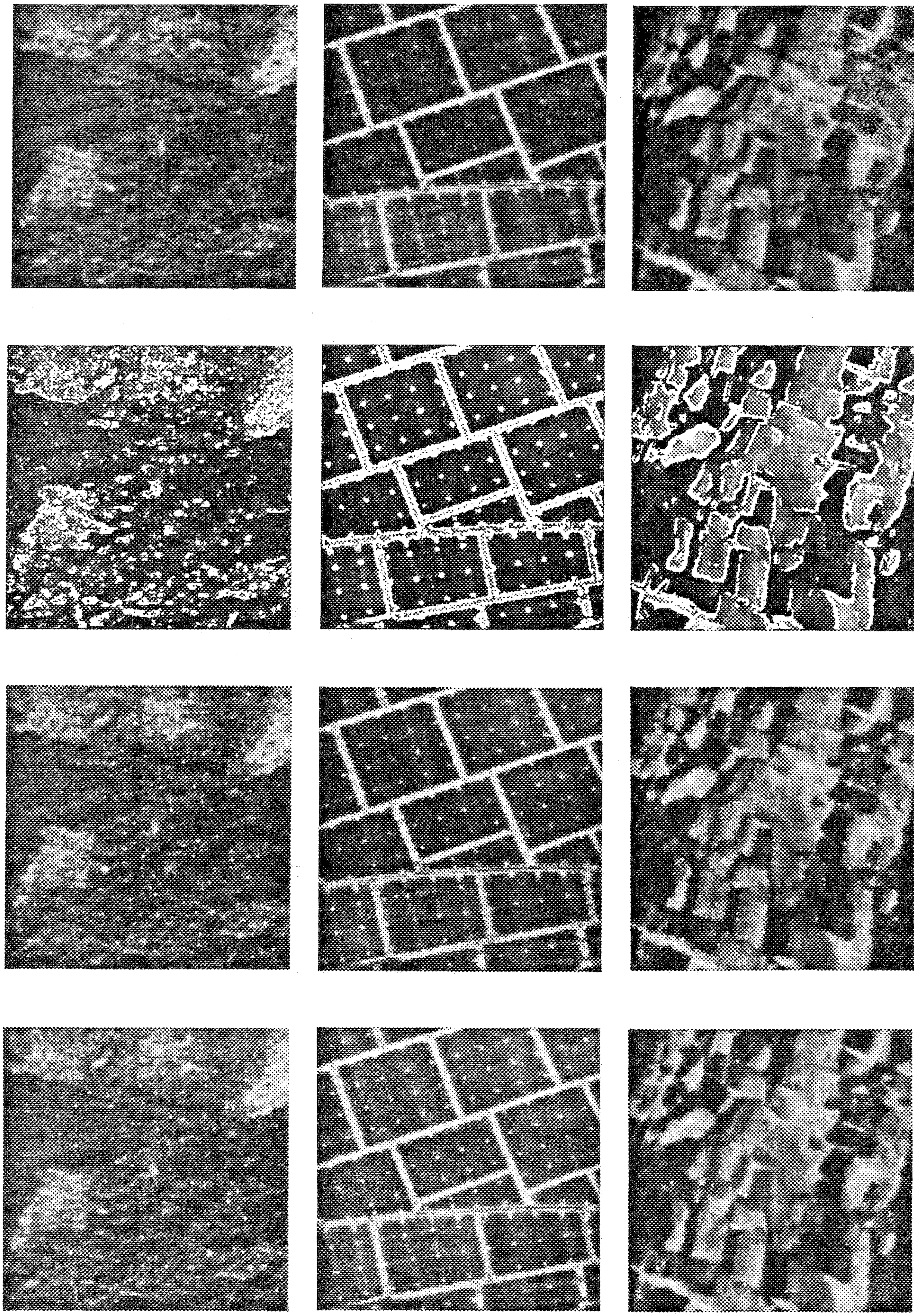


Fig.10 from top to bottom : 1) original photo; 2) by the ground operator; 3) by version I of the new operator; 4) by version II of the new operator.

from left to right : 1) ISP photo 3; 2) ISP photo 5; 3) ISP photo 19.

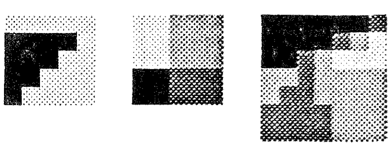
Tab.1 Benchmarktest for feature extraction on computer SUN 3/280 and SUN 3/110

file name	photo size (pixel)	suppression window size	Operator	threshold **	number of points	CPU time (seconds) on *** SUN 3/280	SUN 3/110
sample0	70x70	5x5	Förstner	400/0.50	60	0.58	1.01
				400/0.32	66	0.58	1.06
sample0	70x70	5x5	V. of F. *	100/0.50	44	0.61	1.05
				100/0.32	64	0.62	1.08
sample0	70x70	5x5	Ly-1	20	49	0.16	0.34
sample0	70x70	5x5	Ly-2	20/0.50	46	0.26	0.43
				20/0.32	66	0.24	0.42
sample1	70x70	5x5	Förstner	400/0.50	90	0.62	1.04
				400/0.32	84	0.62	1.08
sample1	70x70	5x5	V. of F.*	100/0.50	81	0.62	1.04
				100/0.32	82	0.60	1.05
sample1	70x70	5x5	Ly-1	20	75	0.14	0.32
sample1	70x70	5x5	Ly-2	20/0.50	79	0.29	0.46
				20/0.32	79	0.28	0.50
ISP photo 3	240x240	9x9	Förstner	850/0.50	352	8.04	13.50
				1000/0.50	301	7.88	13.60
ISP photo 3	240x240	9x9	V. of F.*	275/0.50	361	7.78	13.08
				325/0.50	298	7.96	13.20
ISP photo 3	240x240	9x9	Ly-1	13	407	2.42	4.50
				14	377	2.37	4.18
ISP photo 3	240x240	9x9	Ly-2	13	309	3.19	5.59
				14	267	3.06	5.29

* V.of F. is also Förstner's operator, but with 3x3 window

** w/q for Förstner operator & V.of F.; dg for Ly-1; dg/q for Ly-2

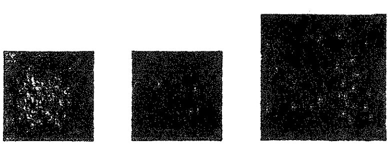
*** SUN 3/110 without the accelerator is same as SUN 3/75
According to Förstner/Förstner 1987/, for 70x70 pixel image it ran 2.1 seconds on SUN 3/75 with Förstner's operator.



$$\begin{bmatrix} 240 & 240 & 240 & 240 & 240 \\ 0 & 0 & 0 & 0 & 240 \\ 0 & 0 & 0 & 240 & 240 \\ 0 & 0 & 240 & 240 & 240 \\ 0 & 240 & 240 & 240 & 240 \end{bmatrix} \quad w = 119631 \quad \frac{q}{4} = 0.16$$

$$\begin{bmatrix} 240 & 240 & 160 & 160 & 160 \\ 240 & 240 & 160 & 160 & 160 \\ 240 & 240 & 160 & 160 & 160 \\ 0 & 0 & 80 & 80 & 80 \\ 0 & 0 & 80 & 80 & 80 \end{bmatrix} \quad w = 32000 \quad \frac{q}{4} = 0.14$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 80 \\ 0 & 0 & 0 & 0 & 80 & 160 & 240 \\ 0 & 0 & 80 & 160 & 240 & 240 & 240 \\ 160 & 160 & 80 & 160 & 160 & 160 & 160 \\ 160 & 80 & 80 & 160 & 160 & 160 & 160 \\ 80 & 80 & 80 & 80 & 160 & 160 & 160 \\ 80 & 80 & 80 & 80 & 160 & 160 & 160 \end{bmatrix} \quad w = 57200 \quad \frac{q}{4} = 0.14$$



$$\begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 3 & 3 & 3 \\ 0 & 3 & 3 & 3 & 3 \end{bmatrix} \quad w = 19 \quad \frac{q}{4} = 0.16$$

$$\begin{bmatrix} 3 & 3 & 2 & 2 & 2 \\ 3 & 3 & 2 & 2 & 2 \\ 3 & 3 & 2 & 2 & 2 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad w = 5 \quad \frac{q}{4} = 0.14$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 & 3 & 3 & 3 \\ 2 & 2 & 1 & 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 2 & 2 & 2 \end{bmatrix} \quad w = 9 \quad \frac{q}{4} = 0.14$$

Fig.11 Three windows selected by Förstner's operator /Förstner 1987/and their graylevel matrices as well as interest values w, q