

Incorporation of texture information as an aid to the feature extraction
and classification with SPOT data

M. Michaelis

Institute of Photogrammetry and Engineering Surveys
University of Hannover, F.R.Germany

Abstract

Feature extraction and classification procedures, performed by digital image processing of remote sensing data from the french satellite SPOT are supported by the involvement of textural information. Classification results can significantly be improved by an a-priori image enhancement procedure which completely suppresses the so-called 'striping', caused by insufficiently calibrated detectors, without losing the essential image information. Methods are described allowing the extraction of a water net from the infrared band and the extraction of a road net from the panchromatic band of a SPOT scene. These methods especially take into account the orientation and energy of edge - and line like features.

1. Introduction

Feature extraction and classification of remote sensing data are important tools in the fields of map production and map revision. In order to support an extraction of special features and to enhance classification results textural image information like the orientation, frequency, phase and shape of edge- and line like structures can be involved in the evaluation process. The common characteristic of all methods using this texture information is, that each pixel in the original image is treated according to its relationship to the neighbourhood. The data of the french satellite SPOT with its high geometrical resolution offer a good base for investigations about the influence of texture phenomena on the quality of feature extraction and classification results, carried out at the institute of photogrammetry in the city of Hannover in the northern part of the Federal Republic of Germany.

2. Classification with SPOT data

In most cases original satellite images, like SPOT - data, contain noise, i.e. the 'striping' caused by insufficiently calibrated detectors of the sensor. Although striping can hardly be identified visually in the original bands - it produces differences of only 1-3 units in grey scale - a classification can respond very sensitively to this disturbance (misclassifications). Figure 3 illustrates the 'striping' - effect in the panchromatic band of a SPOT scene.

2.1 Image enhancement procedure

2.1.1 Orientation and phase estimate

By a so-called enhancement operation noise can be suppressed and features can be emphasized. A special property of the image enhancement is, that noise is filtered out, without losing the essential information. Moreover special features are sharpened without sharpening the noise.

In the beginning of the enhancement procedure a local orientation estimate of the original image is performed. This operation for every neighbourhood produces an output vector, the angle (argument) of which represents an estimation of the dominant orientation of an edge/line - like structure (Fig. 1). The length of the vector (magnitude) indicates the local directional energy, i.e. the strength

of the dominant structure. Thus each pixel is expressed as a complex number. The magnitude part is coded as a normal grey value, the argument is coded as a colour. The whole image is called a CONTEXT IMAGE reflecting the context of each pixel to its neighbourhood.

Before an ORIENT image is calculated, the structures must be detected. This is done by the use of filters, called 'kernels'. A cosine - filter is used to detect lines and a sine filter is used to detect edges. These functions form four quadratic filter pairs in four orientations. Each kernel pair picks up the amount of local energy (i.e. the edge/line content in the neighbourhood) in its own direction. When all contributions are summed, an isotropic energy estimate is formed and the dominant orientation is calculated [2].

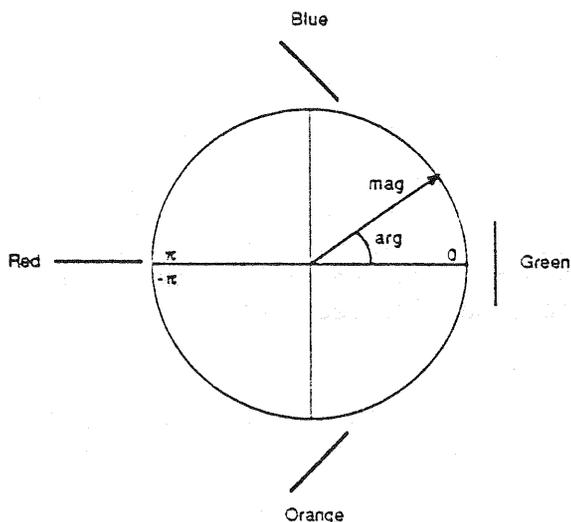


Figure 1: Orientation

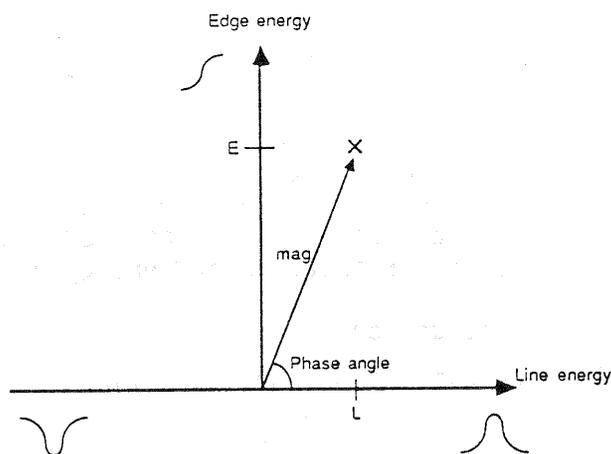


Figure 2: Phase

The orientation serves as an input for a so-called PHASE-operation, where the phase of lines and edges (e.g. the pixels position on the lines/edges) is calculated. The phase operation is performed with a kernel set, consisting of four complex valued kernels: a low-pass filter for smoothing, an isotropic (undirectional) high-pass filter, two cosine filters (line detecting) in two different orientations and four sine-filters (edge detecting) in four orientations. The cosine and sine filters are combined to form two directional filters, oriented along the dominant orientation in every neighbourhood, given by the argument value (colour) of the orientation image.

The polar presentation of the output result is demonstrated by figure 2. The argument value (angle) represents the phase estimate. A perfect edge will produce an output argument of 90° . A positive (white) line will give a zero argument and a negative (black) line an angle of 180° . The output magnitude value (vector length) is the total energy contribution from the line and edge estimates [2].

2.1.2 Pre-enhancement

After the performance of an orientation- and a phase-operation on the image, a so-called PRE-ENHANCEMENT is used to create two context images for the enhancement operation. The magnitude part of the first context image contains an estimate of the degree of inhomogeneity (occurrence of oriented structures), the argument gives the main orientation of this inhomogeneity. A small magnitude value indicates that the neighbourhood is homogeneous. The magnitude value of the second context image is an estimate of the local consistency of orientation (relative strength

of the main orientation) in the original image. A large value indicates that the degree of orientation consistency is high. The argument values are identical to the argument values in the first context image. The standard kernels used with the pre-enhancing are of averaging type.

2.1.3 Image enhancement

The last step in the enhancement procedure is the ENHANCEMENT itself. It is used to emphasize lines and edges, while smoothing homogeneous areas. The first context image of the pre-enhancement operation is used to control the degree of high-pass filtering relative to the degree of low-pass filtering. A low magnitude will reduce the contribution from the high-pass filter, hence the operation will smooth homogeneous areas. The second context image of the pre-enhancement operation is used to control whether the type of high-pass filtering should be directional or undirectional. The orientation of the directional high-pass filter is given by the argument values [2][5]. Figure 4 is the result of an image enhancement applied on the panchromatic band in figure 3. The 'striping' is completely suppressed, without a loss of fine textures like roads and field paths.

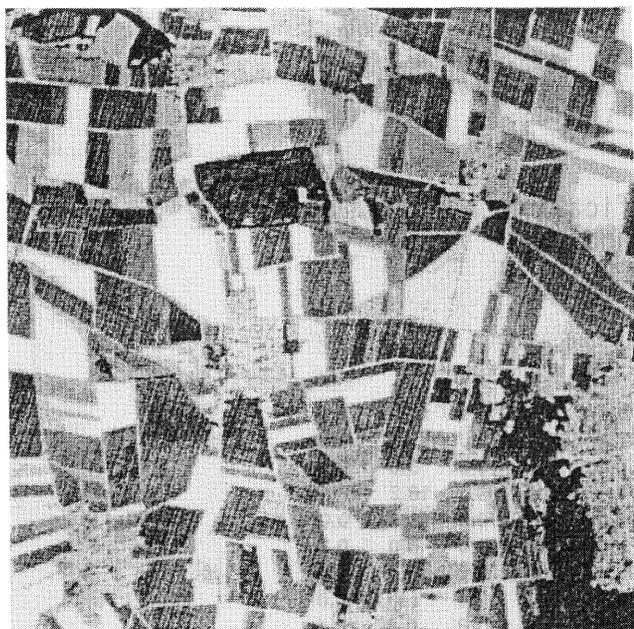


Figure 3: SPOT-pan before image enhancement

Figure 4: SPOT-pan after image enhancement

2.2 Classification results before image enhancement

The following classification example was performed with a SPOT data set of a rural area south of Hannover (fig. 3). Before the digital classification was carried out some field work has been done to gain the 'ground truth' needed for the training and the examination of the classification results. The data set consists of the infrared band and panchromatic band of SPOT, acquired on June 17, 1986. Due to the higher resolution of the panchromatic band the infrared band was resampled to a pixelsize of 10 m x 10 m. The remaining green and red multispectral channels were not used, because of their extremely high correlation to each other and to the panchromatic band [5].

The supervised classification method 'Maximum Likelihood' was applied to the radiometrically uncorrected channels of SPOT (table 1). The classification accuracy for the object class 'transport' is only about 65 % caused by small roads and field paths, most of them only between 3 m and 10 m wide. This leads to pixels, which are a mixture of two classes (e.g. roads and wheat). The large amount of pixels mixed by different adjacent kinds of crops leads to higher misclassifications

in the concerning classes. A relative high percentage of transport is classified as village, obviously caused by similar attributes of asphalted roads and buildings. A high percentage of wheat is classified as forest, caused by the sensor's 'striping' [5].

Table 1: Confusion matrix after ML-classification panchromatic/infrared band (%)

		P R E D I C T E D C L A S S										
		A	B	C	D	E	F	G	H	I	J	K
T	A	0.0	44.9	8.7	15.1	3.3	0.3	0.4	9.9	1.3	8.8	7.3
R	B	0.0	80.0	0.0	0.0	5.4	0.0	3.7	4.7	2.0	4.2	0.0
U	C	0.0	0.0	97.4	2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
E	D	0.0	0.0	0.0	72.3	0.0	0.0	0.0	0.0	0.0	19.7	8.0
	E	0.0	1.0	0.0	0.0	93.1	0.0	0.0	0.0	0.0	5.9	0.0
	F	0.0	0.0	0.0	0.0	0.0	96.3	0.0	0.0	0.0	0.0	3.7
C	G	0.0	7.7	0.0	0.0	0.0	0.0	92.3	0.0	0.0	0.0	0.0
L	H	0.0	1.3	0.0	0.0	0.0	0.0	0.0	97.2	0.7	0.7	0.1
A	I	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	97.5	0.2	0.6
S	J	0.0	16.4	0.8	6.0	5.0	0.6	0.0	1.2	0.0	64.3	5.7
S	K	0.0	0.0	0.0	3.6	0.0	0.3	0.0	0.0	0.0	9.2	86.9

Table 2: Confusion matrix after ML-classification enhanced panchromatic/infrared band (%)

		P R E D I C T E D C L A S S										
		A	B	C	D	E	F	G	H	I	J	K
T	A	0.0	49.0	8.7	15.2	3.0	0.2	0.3	4.5	0.9	10.5	7.7
R	B	0.0	85.7	0.0	0.0	8.6	0.0	1.4	0.0	0.0	4.3	0.0
U	C	0.0	0.0	98.3	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
E	D	0.0	0.0	0.0	73.9	0.0	0.0	0.0	0.0	0.0	18.2	7.9
	E	0.0	0.0	0.0	0.0	96.7	0.0	0.0	0.0	0.0	3.3	0.0
	F	0.0	0.0	0.0	0.0	0.0	99.3	0.0	0.0	0.0	0.0	0.7
C	G	0.0	4.2	0.0	0.0	0.0	0.0	95.8	0.0	0.0	0.0	0.0
L	H	0.0	0.4	0.0	0.0	0.0	0.0	0.0	98.2	0.4	1.0	0.0
A	I	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	97.3	0.4	0.6
S	J	0.0	19.4	0.8	6.2	1.7	0.3	0.0	0.0	0.0	65.7	5.9
S	K	0.0	0.0	0.0	3.9	0.0	0.1	0.0	0.0	0.0	9.2	86.8

- A = Null-Class
- B = Wheat / Barley
- C = Beet
- D = Root Crop
- E = Beans
- F = Rye
- G = Vegetables
- H = Deciduous
- I = Forest (mixed)
- J = Transport
- K = Village

2.3 Classification results after image enhancement

Table 2 shows the result of a classification with the enhanced panchromatic band and the original infrared band. The infrared channel was not to be enhanced, because 'striping' doesn't exist in this band. The suppression of noise in the panchromatic band leads to higher classification accuracies, especially for wheat (+6%), which previously was partly classified as forest.

3. Feature extraction with SPOT data

The extraction of special features in a satellite image is always useful for cartographic applications. Thus existing maps can be updated by an overlay with extracted features like roads or a water net. The following examples are an attempt to extract a water net from the infrared band and a road net from the panchromatic band of a SPOT - scene.

3.1 Extraction of a water net

The infrared band of a SPOT - scene of the city of Hannover was used, due to the high absorption of radiance on water in this channel. The water - features in this area are quite heterogeneous: Water channels, rivers and streams, lakes, pools, drains and a small port are existing (fig. 5).

3.1.1 Adaptive contrast stretch

In the first step of the feature extraction a so-called WALLIS-filter was used. This operation enhances grey scale images by an adaptive contrast stretch. The amount of stretch depends on statistical measures made in the neighbourhood. For every neighbourhood in the image a weighted mean value and the standard deviation are calculated. If the actual standard deviation in a neighbourhood is less than a desired one the centre pixel of that neighbourhood is set to an intensity value further from the mean value than before. If the actual standard deviation is greater than the desired one the center pixel is set to a value closer to the actual mean value. Figure 5 is the result of a WALLIS-operation applied to the infrared band. A kernel size of 7 x 7 was used. Lines and edges (high variance) are emphasized, which is a good pre-condition for a feature extraction [5].



Figure 5: SPOT-infrared
after WALLIS



Figure 6: SPOT-infrared
after FREICHEN

3.1.2 Conversion of edges to lines

The second step consists of an operation called FREICHEN, a combined edge/line detector. Due to the presence of lines (e.g. small channels, streams) and edges (e.g. lakes, pools) in the water net this operation was selected to perform an edge - line conversion. The kernel set used here consists of nine real valued 3 x 3 templates, which can be divided into three groups: edge detecting templates, line detecting templates and one averaging template. Figure 7 indicates, that edges have changed to lines. This is a good precondition for the next operation in the feature (line) extraction procedure [5].

3.1.3 Line detection

The line detecting operation for every neighbourhood produces an output vector, the angle (argument) of which represents an estimate of the line orientation. The magnitude value indicates the amount of estimated line energy, e.g. the strength of the line structure. The amount of estimated line energy is inversely weighted with the amount of edge energy within the neighbourhood. This weighting function can be controlled by an input parameter (THRESH), which specifies the number of times that the line energy must exceed the edge energy to produce an output:

```
IF      line energy
      ----- > THRESH  then
      edge energy

      OUT.r = line energy
ELSE
      OUT.r = 0.0
      OUT.phi = orientation estimate
```

Small values for the threshold - parameter will produce thicker output lines, high values cause thin lines to disappear [2]. Figure 7 is the result of the LINE operation.

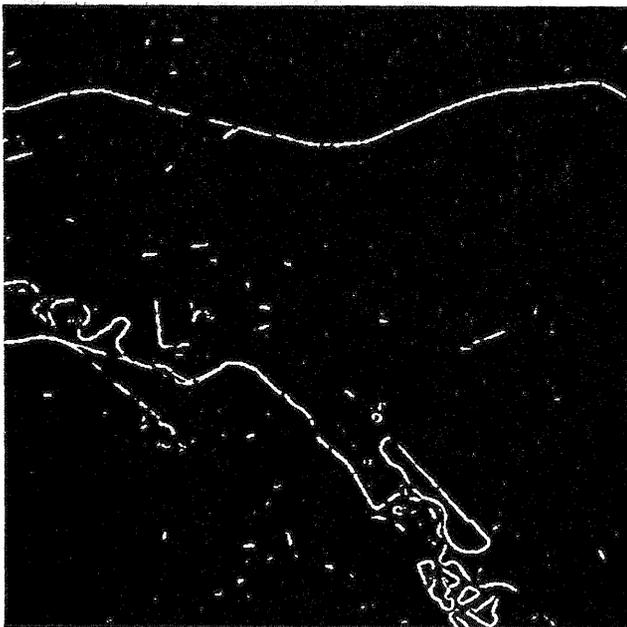


Figure 7: SPOT-infrared after line detection

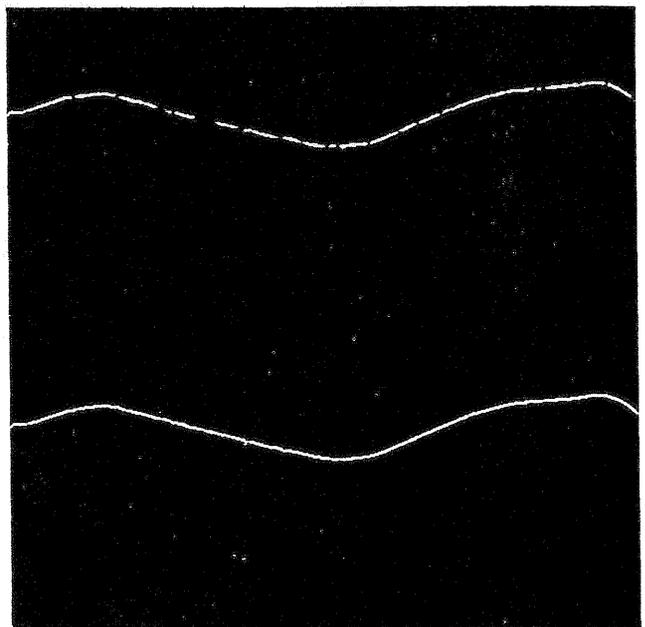


Figure 8: SPOT-infrared after line restoration

3.1.4 Restoration of lines

In many applications a part of the information, which ought to be emphasized, is weakened after a filter operation. In this case a CONSISTENCY operation can be useful, to restore weak or broken lines in an image. A kernel set, which consists of three kernels, one that detects lines and two averaging kernels (1 large, 1 small), are implied. The three kernel responses are called 'LINE', 'AVER' and 'AVERsm' in the formular below. 'CTXr' represents the magnitude value in the context image. 'OFF' adds an offset to the output magnitude values. 'K' adds a parameter controlled contribution from the context image to ensure that the result does not differ too much from the input [2].

$$\text{OUT.r} = \text{AMP} * \frac{\sqrt{\text{TA} * \text{TB}} + \text{K} * \text{CTXr}}{1.0 + \text{K}} + \text{OFF}$$

$$\text{TA} = \text{AVER} + \text{AVERsm}$$

$$\text{TB} = \text{LINE} + (\text{AVERsm} - \text{AVER})$$

$$\text{OUT.phi} = \text{line orientation}$$

The sum of the responses from the two average kernels (TA) give large values even for neighbourhoods containing just a small area of large pixel values (e.g. weak line). TB is the sum of the output from the line detecting kernel and the difference between the small and large averaging kernel responses. TB is large for thin lines, smaller for thicker or distorted lines and smaller still for small homogeneous areas. Hence, thin and weak lines will be emphasized while areas where the line consistency is too small are suppressed.

Figure 8 shows a part of the extracted large water channel in the north of the SPOT - scene before (top) and after (bottom) the consistency operation.

3.2 Extraction of a road net

The preconditions for an extraction of a road net are quite different from those of the water net extraction especially in built up areas. The contrast of roads in relation to other objects in the panchromatic band of SPOT is poor. The always present 'striping' often reaches the same spatial dimension as roads.

In the following example of a road net extraction, a geodetically rectified version of the original panchromatic band (Level 1 A) was used. The image was resampled to a pixelsize of 5m x 5m in order to ensure that small roads don't disappear during the filtering process. The whole extraction process was controlled by a visual comparison of the results with an extract of a 1:50 000 scale map (fig. 9).



Figure 9: Topographic map 1:50 000

3.2.1 High-pass filtering

Highways, streets, roads, paths and bridges compose the fine textural information or the high frequencies in the SPOT image. Thus a high-pass filter in the spatial domain could be applied. Generally a high-pass filter suppresses the lower frequencies, while the higher frequencies pass unaffected.

Figure 10 is the result of a high-pass-limited filtering (hplim) with a 15 x 15 kernel. Objects, like the large baroque garden in the upper image part, the gravel path situated in the right upper corner and the large highways come out very clearly. Streets in dense built-up areas (right part) are better identifiable than in low density housing areas (left part). The 'striping' influence is intensified, especially in homogeneous areas like the river, passing from north west to south east, and the channel with the port in the south west corner of the image (see map).

The difference of a hplim-filter to a high-pass filter (hp) is, that a hplim-filter completely suppresses frequencies outside the operating range, while a hp - filter suppresses them only to a certain level. Practically this means, that the curve of the hplim filter function is steeper.

Figure 11 shows the result after filtering the panchromatic band with a high-pass filter by means of a 9 x 9 kernel. Noise-frequencies are more suppressed now, indicated by the dark and more homogeneous portions of the object class water. Thus bridges crossing the river in the eastern part of the image can clearly be identified. High order streets like highways are placed in the foreground.

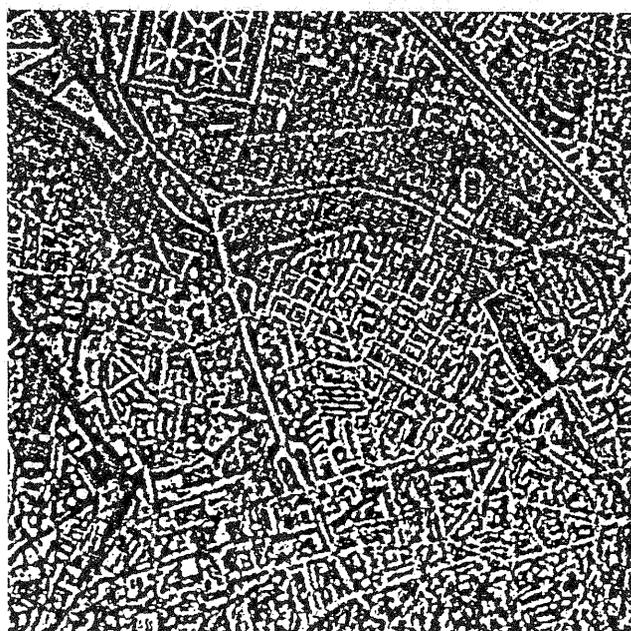
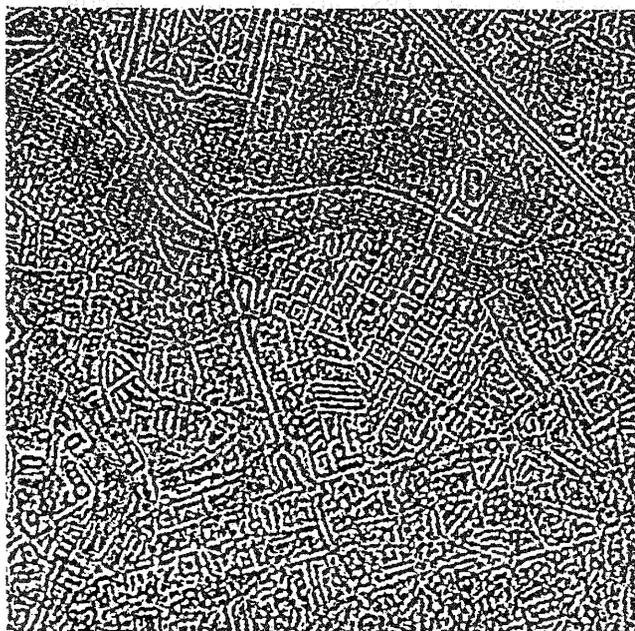


Figure 10: High-pass-limited filtered image

Figure 11: High-pass filtered image

3.2.2 Line detection

The hp- and hplim-filters, described before, emphasize special frequencies defined by simple filtering functions. They make no difference, whether an edge or a line is emphasized. The LINE operation has been discussed in section 3.1.3. Figure 12 is the result of a line detecting operation (15 x 15 kernel size) on the high-pass filtered image (fig.11). The first impression is, that edges, like the river and the channel disappeared because the 'striping' influence becomes stronger again. Edge-like features like the broad highways are shrunk, small edges are suppressed and lines are emphasized.

3.2.3 Restoration of lines

The LINE operation produces strong lines of higher consistency and weak or broken lines of lower consistency. The main portion of those objects with the lowest consistency is represented by the 'striping' and by isolated pixels, still left after the line operation.

A LINE CONSISTENCY operation as described in section 3.1.4 was applied to this image. Figure 13 indicates much more homogeneity. Isolated points mostly disappeared and the 'striping' is suppressed. Consequently objects, 'covered' by this noise before, like the large sports field in the north east part of the image and the bridges have been isolated and emphasized.

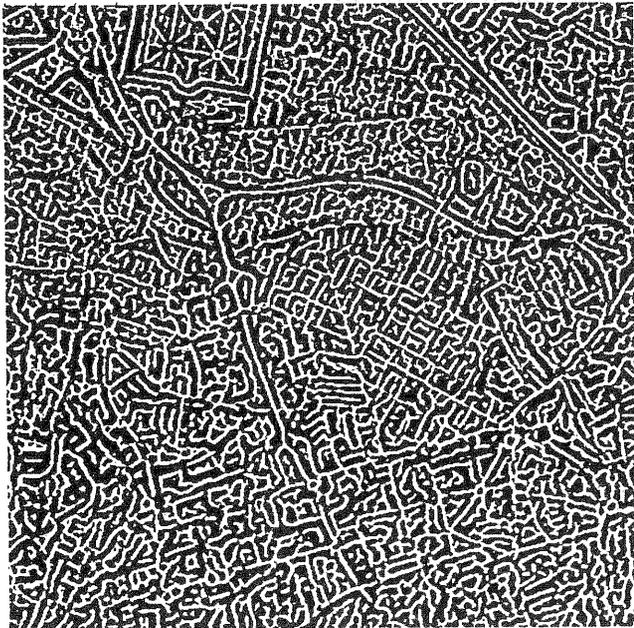


Figure 12: Detected lines

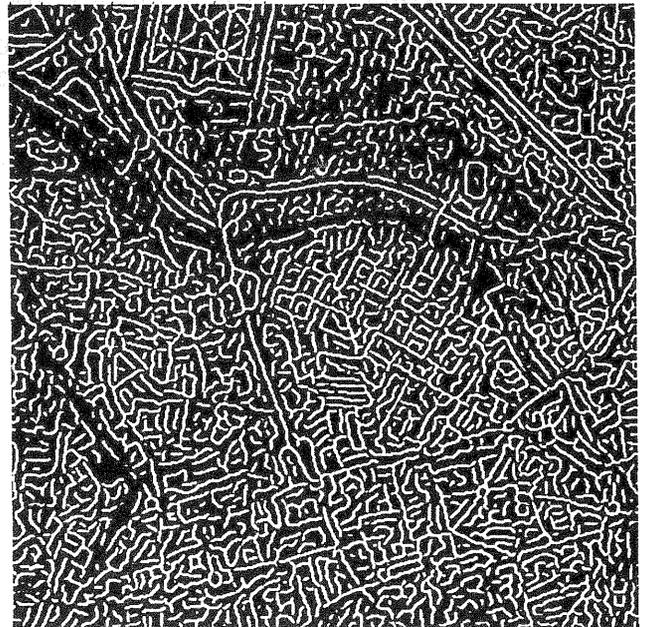


Figure 13: Lines after restoration

4. Conclusions

The incorporation of texture information in evaluation processes like the digital image classification and the extraction of features proves to be very helpful. Quality measures like the strength, orientation and phase of edge/line - like features in an image are suited, to localize the 'striping' in a SPOT image, a classification very sensitively responds to. By an image controlled enhancement procedure this noise can be suppressed without losing the essential image information, e.g. fine textures like roads and small field paths.

The energy, orientation and consistency of edge/line - like features are textural measures, which can be involved in the feature extraction e.g. of a water net or a

road net. This often requires a certain level of preprocessing with the original images. Thus features need to be emphasized before being extracted. The following line extraction operation detects lines by the use of kernel sets for different line directions. The orientations of the detected lines and their energy are coded in a complex valued context image. This image is used to control the restoration of lines weakened or broken by the extraction process.

The way, how to involve textural quality measures on an image like a SPOT scene always depends on it's geometrical and radiometrical quality and it depends on the user's conception. Thus the extraction of different features requires the use of different filtering functions with different kernel sizes. This means, that all methods described here can be standardized to a certain level only.

Once features have been extracted, they can be vectorized and transfered into the database of a geographic information system (GIS). In a GIS these data form a well suited base for the production and updating of topographic maps.

References:

- [1] CNES: SPOT, satellite-based remote sensing system. Paris, 1982.
- [2] Context Vision: Image Processing Manual, Vol 1. Linköping, 1987.
- [3] Gonzalez, R.C.: Digital Image Processing. Addison-Wesley
Wintz, P. Publishing Company, Massachusetts, 1977.
- [4] Konecny, G.: Die Methodik der Fernerkundung und ihre Anwendung zur
Erfassung thematischer Daten. ZfV 104, 1979.
- [5] Michaelis, M.: SPOT - Image Processing on the GOP 300 System.
Seminar on Photogrammetric Mapping from SPOT Imagery.
Hannover, Sept. 1987.
- [6] Rosenfeld, A.: Digital Picture Processing. Academic Press, USA, 1982.
Kak, A.C.