Image Preprocessing for Feature Extraction in Digital Intensity, Color and Range Images

Wolfgang Förstner

Institute for Photogrammetry, Bonn University
Nussallee 15, D-53115 Bonn, e-mail: wf@ipb.uni-bonn.de
http//:www.ipb.uni-bonn.de

Abstract. The paper discusses preprocessing for feature extraction in digital intensity, color and range images. Starting from a noise model, we develop estimates for a signal dependent noise variance function and a method to transform the image, to achieve an image with signal independent noise. Establishing significance tests and the fusion of different channels for extracting linear features is shown to be simplified .

1 Motivation

Signal analysis appears to be one of the most interesting problems common to Geodesy and Photogrammetry. Objects of interest primarily in Geodesy are the gravity field or the topography of the ocean. Objects of primary interest in Photogrammetry are photographic or digital images and all elements of topographic maps. Digital elevation models (DEM) and paths of sensor platforms of air or space vehicles are among the common interests to both fields.

Seen from the type of result Geodesists and Photogrammetrists, however, appear to show increasing differences, if one looks at both research and practical applications: Geodesists mainly are interested in the form, shape or position of geometric objects, especially point, scalar, vector or tensor fields. There was a long period of common interest in point determination, orientation, calibration and DEM derivation, all triggered by the strong tools from estimation and adjustment theory. Since about 10 years photogrammetric research has moved – away from Geodesy – to image interpretation, aiming at recovering not only the geometry of the objects but also their meaning (cf. the review by Mayer 1999).

This development did not really come over night, as the derivation of geometric structures from discretized continuous signals, especially structure lines in DEM appears to be a first step towards extracting symbolic information from sensor data. But this type of transition from an iconic, often raster type of description, to a symbolic, often vector type of description appears to be only slowly penetrate the research in Geodesy.

Interestingly enough, the signal processing part in image processing has only partly been realized by Geodesists, though the type of questions are very similar:

predicting a possibly piecewise smooth continuum from sampled data. This task is very similar in both fields, thus also the type of approaches are similar, e. g. using Wiener prediction for restoring images. Of course there are differences in the underlying models. The main difference is the open world Photogrammetry has to cope with, i. e. the impossibility to impose very hard constraints on the observations, as this is the case e. g. in Physical Geodesy.

This is the reason why the author chose a topic for this paper in the overlapping area of Geodesy and Photogrammetry: the processing of two dimensional data prior to the derivation of geometric structures of these signals. Such two dimensional signals are either digital or digitized, black and white, color and multi spectral images or directly measured or derived DEM's. In both cases we assume the data to be given in a regular raster [1].

Starting with modeling the noise behavior of the given data, we develop methods for estimating the noise characteristics. As extracting geometric features starts with detecting signals in noise, which can be interpreted as hypothesis testing, we simplify processing by a noise variance equalization of the given signal. A locally adaptive Wiener filter can be used to smooth the signal depending on the local information content, aiming at smoothing homogeneous areas while preserving edges, lines, corners and isolated bright or dark points. In all cases we discuss digital images and DEM's, or so called range images, in parallel. The fusion of several channels in color images appears to be a challenge as there is no unique way to integrate the information.

The goal of this preprocessing is to exploit the statistical model of the signal as far as possible and thus reduce the number of control parameters for the algorithms. Actually we only need a significance level for distinguishing between signal and noise, as all other properties are estimated from the given data. The procedure for intensity images is partly described in an early version by Förstner 1994 and extended and thogoughly investigated by Fuchs 1998. First investigations into the performance of the feature extraction procedure are presented by Fuchs *et al.* 1994.

The paper collects research at the author's institute of about one decade. The references contain some review papers which point to further reading and alternative approaches.

Notation: We distinguish continuous signals with coordinates (x, y) and discrete signals with coordinates (r, c), r and c standing for rows and columns. Partial derivative operators are $\partial_x \doteq \partial/\partial x$ etc., discrete versions are $\partial_r \doteq \partial/\partial r$ etc. resp. Where necessary, stochastical variables are underscored, e. g. $\underline{g}$. The nor-

[1] This suggests to use Fourier techniques. But this is by no way reasonable: The objects shown in images do not show any periodicity nor is the statistical behavior homogeneous. Wavelet techniques seem to overcome some of the limitations, especially the assumption of homogeneity. But, at least in their Kronecker-version of basis-functions, they lack of providing rotation invariant image properties, and of including higher level knowledge. Only Gabor-wavelets have been widely used for more than two decades for describing texture in digital images for quite some time.

mal density is G_s with standard deviation s as parameter in both coordinate directions.

2 The Image Model

A digital image can be treated as a random process.

In a first approximation it therefore can be characterized by its mean and its variance-covariance structure, thus its first and second moments. In general both are inhomogeneous, thus location dependent, and anisotropic, thus orientation dependent. Moreover, we have to expect the noise statistics to be dependent on the signal.

The distribution of the random process can be expected to be quite complicated in case we start to model the generating process in detail, e. g. using the geometrical-physical model of the sensor. For simplicity we assume the stochastic process to be Gaussian. This allows to easily derive thresholds in the subsequent analysis steps. It is motivated by the central limit theorem and the experience, that deviations form the Gaussian distribution usually can be circumvented by a slight modification of the model.

For modeling reasons we distinguish three types of multi valued images:

1. the *true image* $\underline{\tilde{\boldsymbol{f}}}(x, y)$ is continuous. It is used for describing the image which we would have obtained with an ideal sensor with infinite resolution, no bias and no noise. The true image is modeled as a stochastic process. However, in our context we only impose very limited constraints on the true image: The image area $\mathcal{I}$ is assumed to be partitioned into regions $\mathcal{R}_i$ covering the complete image area. Thus:

$$\mathcal{I} = \bigcup_i \mathcal{R}_i, \quad \mathcal{R}_i \cap \mathcal{R}_j = \emptyset \quad \forall\, i, j \tag{1}$$

   Within each region the image function is assumed to be homogeneous, in this context smooth, with low gradient or low curvature in all channels. We will specify these properties later, when needed.
2. the *ideal image* $\underline{\boldsymbol{f}}(x, y)$ also is continuous but a blurred version of the true image. In a first instance it is used to explain the limitations of any feature extraction concerning the resolution. It also is used to cover the limited resolution of the sensor internal filtering processes. Technically it is necessary to fulfill the sampling theorem. The blurring also could cover small geometric errors, e. g. due to lens distortion.
3. the *real image* $\underline{\boldsymbol{g}}(r, c)$ is a sampled and noisy version of the ideal image. The sampling is indicated by the indices (r, c), representing *r*ows and *c*olumns of the digital image.

   The image $\underline{\boldsymbol{g}}$ therefore can be written as:

$$\underline{\boldsymbol{g}}(r, c) = \underline{\boldsymbol{f}}(r, c) + \underline{\boldsymbol{n}}(r, c) \tag{2}$$

As noise usually is small we can always assume it to be additive by assuming its characteristic to be signal dependent, i. e. dependent on $\boldsymbol{f}$.

In addition to the variance properties the correlations between neighboring pixels may be taken into account. In original images taken with a good camera or scanner the noise correlations between neighboring pixels are comparably small, i. e. usually less than 50 %. We therefore neglect spatial correlations and only discuss the variance of the noise.

We now discuss the different noise models in detail.

2.1 Intensity Images

Images of CCD-Cameras The noise in single channel intensity images $g(r, c)$ with CCD-Cameras contains three basic components:

1. The noise characteristics is dominated by the Poisson distribution of the photon flux (cf. Dainty and Shaw 1974). This holds, in case the intensity $\underline{g}$ is proportional to the number $\underline{N}$ of the photons. The Poisson distribution is characterized by $\mathsf{E}(\underline{N}) = \mathsf{D}(\underline{N}) = \mu_N = \sigma_N^2$. Thus this noise variance component increases linearly with the intensity.
2. The rounding errors, showing a variance of $1/12$.
3. Electronic noise, being independent on the intensity.

Therefore an adequate model for the noise is

$$\sigma_n^2(g) = a + b\, g \tag{3}$$

In the ideal case both parameters a and b are positive in order to guarantee a positive noise variance. The parameter a should be larger than $1/12$, as it covers the rounding errors. In case the image acquisition device performs a linear transformation on the intensities, the linear behavior of the noise variance still can be observed, however, possibly only in the interval $[g_{\min}, g_{\max}]$. Thus the following two conditions should hold

$$\sigma_n^2(g_{\min}) \geq 1/12, \quad b \geq 0 \tag{4}$$

In case the intensity is not proportional to the number N of photons, but e. g. proportional to $\log N$, the model eq. (3) needs to be modified.

Observe, the noise characteristics are dependent on the intensity which varies throughout the image, thus is totally inhomogeneous. However, for the same intensity it is invariant to position. This is one of the main reasons why it is not reasonable to use Fourier techniques.

In case the sensitivity of the individual pixels depends on the mean intensity in the neighborhood for increasing the radiometric resolution, the simple model eq. (3) does not hold anymore. Without knowing the type of adaptivity of the sensitivity to the intensity in the neighborhood there is no simple rule how to model the noise variance.

Digitized Photographs Digitized photographs, especially digitized aerial images show quite complex noise characteristics. In a first instance one would expect a similar increase of the noise variance with the signal as in CCD-images. Due to film characteristics, film development and scanner characteristics, especially also the scanner software, the noise characteristics will not follow a simple function of the intensity.

We now assume the noise characteristics only depends on the intensity, but in an arbitrary manner. Thus we have:

$$\sigma_n^2(g) = s(g), \quad s(g) \geq 1/12 \tag{5}$$

2.2 Color Images

Color images $\underline{g} = (\underline{g}_k)$ are multichannel images, where each of the K channels represents a specific spectral band. In principle the number of bands is not restricted to three, as in RGB-images, but may be any number, as in multi spectral images.

The sensitivity of the sensors or the filters used for separating the spectral bands may overlap. This results in correlations of the signals $\underline{g}_{k'}$ and $\underline{g}_{k''}$ in different channels k' and k''.

However, the *noise* of the different channels can certainly be assumed to be independent, unless the multi channel sensor does not transform originally independent channels, e. g. by a color transformation.

In the subsequent image analysis steps we do not refer to the colors in the way the visual system perceives them. We interpret the different channels as the result of independent physical sensing processes and therefore intensionally do *not* perform any color transformation.

The result of this argumentation is simple: Every channel has a statistically independent noise characteristic, either following eq. (3) leading to:

$$\sigma_{n_k}^2(g_k) = a_k + b_k\, g_k \quad k = 1, ..., K \tag{6}$$

with $2K$ parameters a_k and b_k specifying the noise variances or, following eq. (5), to

$$\sigma_{n_k}^2(g_k) = s_k(g_k) \quad k = 1, ..., K \tag{7}$$

with K different functions s_k.

2.3 Range Images

Range images in principle are single channel images as intensity images are. Depending on the type of sensor the noise can be assumed to depend on the distance, e. g. in case of time of flight sensors or image stereo sensors. The noise characteristics also may depend on the direction of the sensing ray and on the slope of the surface. The type of dependency not really has been investigated.

On the other hand sensor internal error sources or the resampling process, transferring irregular data into a regular grid, may dominate and lead to a noise

variance being constant over the full range image. We therefore assume the noise to be constant in range images. Making it dependent on position, orientation or other attributes can be taken into account if necessary.

As edge extraction in range images is based on the two channel image:

$$\underline{\boldsymbol{g}}(r,c) = \nabla d(r,c) = \begin{pmatrix} d_r \\ d_c \end{pmatrix} = \begin{pmatrix} \dfrac{\partial d}{\partial r} \\ \dfrac{\partial d}{\partial c} \end{pmatrix} = \begin{pmatrix} \partial_r * d \\ \partial_c * d \end{pmatrix} \tag{8}$$

we want to discuss the noise characteristics of this two channel image.

The covariance matrix of the two values $\underline{g}_k(r,c), k = 1,2$ is diagonal:

$$\boldsymbol{\Sigma}_{gg} = \mathsf{D}(\nabla d) = \sigma_d^2 \, t\boldsymbol{I} \tag{9}$$

where the factor t depends on the filter kernel used for determining the gradient, and $\boldsymbol{I}$ is the unit matrix. This holds if the differentiation kernels $\partial_x(x,y)$ and $\partial_y(x,y)$ are orthogonal, thus

$$\int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} \partial_x(x,y)\partial_y(x,y)\, dx\, dy = 0 \tag{10}$$

We often use Gaussian kernels

$$G_s(x,y) = \frac{1}{2\pi s^2} \mathrm{e}^{-\dfrac{(x^2+y^2)}{2s^2}} \tag{11}$$

or its derivatives, e. g.

$$G_{x;s}(x,y) = \frac{\partial}{\partial x} G_s(x,y) = -\frac{x}{s^2} G_s(x,y) \tag{12}$$

Then (10) obviously holds.

E. g. the Sobel kernels

$$\partial_{r,S} \doteq \left(\frac{\partial}{\partial r}\right)_S = \frac{1}{8} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \tag{13}$$

$$\partial_{c,S} \doteq \left(\frac{\partial}{\partial c}\right)_S = \frac{1}{8} \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \tag{14}$$

are discrete approximations of $G_{x;s}(x,y)$ and $G_{y;s}(x,y)$ with $s \approx 0.7$ which can be proved by comparing the energies of the two filter kernels:

$$\sum_{r,c} \partial_{r,S}^2 = \frac{3}{16}, \qquad \int_{x,y} G_{x;s}^2(x,y) dx\, dy = \frac{1}{8\pi s^4} \tag{15}$$

which are identical for $s = 0.6787\ldots$.

We now obtain the factor t_S in (9)

$$t_S = \left(\frac{1}{8}\right)^2 ((-1)^2+(-2)^2+(-1)^2+0^2+0^2+0^2+1^2+2^2+1^2) = \frac{12}{64} = \frac{3}{16} \quad (16)$$

which is due to the linear dependency of the gradient of the original distances:

$$d_r = \frac{1}{8}(- d(r-1,c-1) - 2d(r-1,c) - d(r-1,c+1) \quad (17)$$

$$+ d(r+1,c-1) + 2d(r+1,c) + d(r+1,c+1)) \quad (18)$$

$$d_c = \frac{1}{8}(- d(r-1,c-1) + d(r-1,c+1) - 2d(r,c-1) \quad (19)$$

$$+ 2d(r,c+1) - d(r+1,c-1) + d(r+1,c+1)) \quad (20)$$

and identical with the energy in (15a).

As the sum of products of the coefficients for d_r and d_c sum to 0, the two partial derivatives are uncorrelated, showing the Sobel to be a consistent approximation of the first Gaussian derivatives.

3 Noise Variance Estimation

The noise variance needs to be estimated from images. There are three possible methods to obtain such estimates:

1. *Repeated images*: Taking multiple images of the same scene without changing any parameters yields repeated images. This allows to estimate the noise variance for each individual pixel independently. This certainly is the optimal method in case no model for the noise characteristics is available and can be used as a reference.
   The method is the only one which can handle the case where there is no model for the noise characteristics.
   We used it for finding out the noise model for the scanner component of digitized aerial images (cf. fig. 1ff, taken from Waegli 1998).
   The disadvantage of this method is the need to have repeated images, which, e. g. in image sequences is difficult to achieve.
2. *Images of homogeneous regions:* Images of homogeneous regions, thus regions with piecewise constant or linear signal, allows to estimate the noise variance from one image alone.
   The disadvantage is the requirement for the segmentation of the images into homogeneous regions. Moreover, it is very difficult to guarantee the constancy or linearity of the true intensity image within the homogeneous regions. Small deviations from deficiencies in the illumination already jeopardize this method.
   The method is only applicable in case the noise only depends on the signal.
3. *Images with little texture:* Images with a small percentage of textured regions allow to derive the noise variance from the local gradients or curvature. For the larger part of the image they can be assumed to have approximately zero

mean. Thus presuming a small percentage of textured regions assumes the expectation of the gradient or the curvature in the homogeneous regions to be negligible compared to the noise.

Also this method is only applicable in case the noise characteristics is only depending on the signal.

We want to describe this method in more detail. We first discuss the method for intensity images. The generalization to range images is straight forward.

3.1 Estimation of the Noise Variance in Intensity Images

Case I: Constant Noise Variance The idea is to analyze the histogram of the gradient magnitude of the image in the area where there are no edges and no texture. The procedure given here is similar to that proposed in (Förstner 1991).

We now need to specify the model for the ideal image f. We assume that a significant portion $\mathcal{H}$ of the image area $\mathcal{I}$ is homogeneous, thus shows locally constant intensity, thus $\mu_f = const.$. Adopting notions from statistical testing $H_0 = (r,c) \in \mathcal{H}$ is the null-hypothesis, i. e. the hypothesis a pixel belongs to a homogeneous region. Thus

$$\mathsf{E}(\nabla \underline{f} | H_0) = \mathbf{0} \tag{21}$$

The other area $\mathcal{I} - \mathcal{H}$ covers edges and textured areas with significantly larger gradients.

Then, as to be shown, the histogram of the homogeneoity measure $h = |\nabla g|$ shows exponential behavior in its left part representing the noise in the image and arbitrary behavior in the right part representing the edges:

We assume the intensities to be Gaussian distributed with fixed mean and random noise. Assuming the simple gradient kernels

$$\left(\frac{\partial}{\partial r}\right)_0 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \qquad \left(\frac{\partial}{\partial c}\right)_0 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \tag{22}$$

neglecting the scaling factores $1/2$, we obtain the gradient

$$\nabla g = \begin{pmatrix} g_r \\ g_c \end{pmatrix} = \begin{pmatrix} g_{r+1,c} - g_{r-1,c} \\ g_{r,c+1} - g_{r,c-1} \end{pmatrix} \tag{23}$$

which is Gaussian distributed with covariance matrix

$$\mathsf{D}\left(\begin{matrix} \underline{g}_r \\ \underline{g}_c \end{matrix} \middle| H_0 \right) = \sigma_{n'}^2 \, \boldsymbol{I} \tag{24}$$

Here we use the convention

$$\sigma_{n'}^2 = \sigma_{n_r}^2 = \sigma_{n_c}^2 \tag{25}$$

which in general is given by (cf. eq. (15b), Fuchs 1998)

$$\sigma_{n'}^2 = \int_{x,y} G_{x;s}^2(x,y) dx\, dy = \frac{1}{8\pi s^4} \sigma_n^2 \, \sigma_n^2, \quad \text{or} \quad \sigma_{n'}^2 = \sum_{r,c} \partial_r^2(r,c) \, \sigma_n^2 \tag{26}$$

In our case eq. (22) leads to

$$\sigma_{n'}^2 = 2\sigma_n^2 \tag{27}$$

The squared gradient magnitude measures the homogeneity h

$$h_\nabla(r,c) = |\nabla g(r,c)|^2 = g_r^2(r,c) + g_c^2(r,c) \tag{28}$$

It is the sum of two squares of Gaussian variables.

In case the mean $\mu_g = \mu_f$ of $\underline{g}$ is constant in a small region, thus the model eq. (21) holds, the squared gradient magnitude is χ_2^2 or exponentially distributed with density function (neglecting the index ∇ for simplicity)

$$p(h|H_0) = \frac{1}{\mu_h}\,\mathrm{e}^{-\frac{h}{\mu_h}} \tag{29}$$

and mean

$$\mathsf{E}(\underline{h}|H_0) = \mu_h = 4\sigma_n^2 \tag{30}$$

Therefore we are able to estimate the parameter μ_h from the empirical density function in the following way:

1. Set the iteration index $\nu = 0$. Specify an approximate value $\sigma_n^{(0)}$ for the noise standard deviation. Use $\mu_h^{(0)} = 4\sigma_n^{2(0)}$ as approximate value for the gradient magnitude.
2. Determine all $h(r,c)$
3. Take the mean $m^{(\nu)}$ of all values $h(r,c) < \mu_h^{(\nu)}$. Its expected value is given by

$$\mu_m^{(\nu)} = \frac{\int_{h=0}^{\mu_h^{(\nu)}} h p(h|H_0)dh}{\int_{h=0}^{\mu_h^{(\nu)}} p(h|H_0)dh} = \frac{\mathrm{e}-2}{\mathrm{e}-1}\,\mu_h^{(\nu)} \tag{31}$$

   in case the edges or textured areas do not significantly contribute to this mean. Thus a refined estimate $\mu_h^{(\nu+1)}$ for μ_h is given by:

$$\mu_h^{(\nu+1)} = \frac{\mathrm{e}-1}{\mathrm{e}-2}\,m^{(\nu)} \approx 2.392\,m^{(\nu)} \tag{32}$$

4. Set $\nu = 1$ and repeat step 3.

Usually, only two iterations are necessary to achieve convergence. A modification would be, to take the median of the values $h(r,c)$ as a robust estimate and compensate for the bias caused 1) by taking the median instead of the mean and 2) by the edge pixels (cf. Brügelmann and Förstner 1992).

This procedure can be applied to every channel in a multi channel image, especially in color images or in gradient images of range images.

Fig. 1. *shows three image sections of 300 × 300 pixels from three different aerial images.*

Fig. 2. *shows histograms image sections of 300 × 300 pixels. Observe that the frequencies of the intensities vary heavily. Only for intensities with high frequencies one can expect to obtain reliable estimates for the noise variance (from Waegli 1998).*

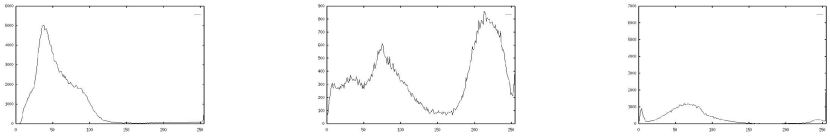


Fig. 3. *shows the noise standard deviation with 32 grey level intervals as a function of the intensity. The estimated variances increase with intensity but not linearly and showing larger variation in areas with low intensity frequency (from Waegli 1998).*

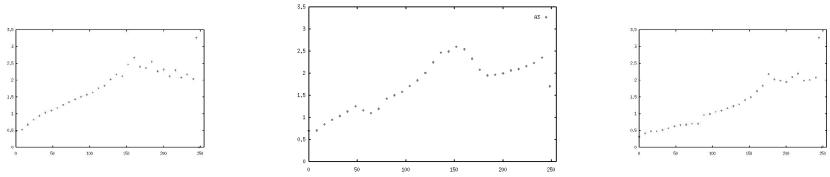
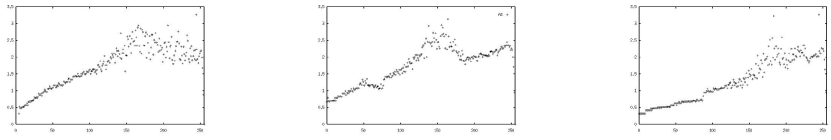


Fig. 4. *shows the noise standard deviation with no partitioning as a function of the intensity. The effect of the intensity frequency onto the reliability of the estimated noise variances is enlarged, but taking groups of 8 intensities is fully acceptable (from Waegli 1998). The small step in the right figure is likely due to the scanner software.*



Case II: General Noise Variance In case the noise variance is not constant over the whole image area and can be assumed only to depend on the intensity, we need to parameterize the noise variance function $\sigma_n^2 = s(g)$ in some way.

The easiest possibility is to assume it to be continuous. Then we can partition the range $[0..G]$ of all intensities g into intervals $I_\gamma, \gamma = 1..\Gamma$ and assume the noise variance to be constant in each interval.

Thus we repeat the procedure of subsection 3.1 for each intensity interval under the condition $g \in I_\gamma$.

The choice of the intervals obviously requires some discussion, as it may significantly influence the solution. Taking a set of constant intervals may lead to intervals where no intensities belong to, even in case one would restrict to the real range $[g_{\min}, g_{\max}]$. Therefore the intervals should be chosen such that

1. they contain enough intensity values..
   The number should be larger than 100 in order to yield precise enough estimates for the noise variances, which in this case has a relative (internal) accuracy better than 10 %. The number of intervals should be chosen in dependency of the expected roughness of $s(g)$. For aerial images we have made good experiences with intervals between 1 and 8 grey values on image patches of 300 × 300 pixels (cf. Waegli 1998 and figs. 3 and 4).
2. they contain an equal number of intensities. This may easily be achieved by using the histogram of the intensities.

Case III: Linear Noise Variance The case of linear noise variance can be handled in a special way. The parameters a and b of the linear noise variance function $\sigma_n^2(g) = a + bg$ can be determined without partitioning the intensity range, but by directly performing a robust weighted linear regression on the gradient magnitudes. The weights, can be derived from the χ_2^2-characteristics of the gradient magnitudes, whereas the robustness can be achieved by excluding all pixels with too large gradient magnitudes, making the threshold dependent on the estimated noise variance function (cf. Brügelmann and Förstner 1992).

3.2 Noise Estimation in Range Images

In range images we assume the curvature to be small. Thus we assume the Hessian of the image function to have zero expectation:

$$\mathsf{E}(\boldsymbol{H}(\underline{f})|H_0) = \mathsf{E}\left(\left(\begin{array}{cc}\underline{f}_{rr} & \underline{f}_{rc}\\ \underline{f}_{cr} & \underline{f}_{cc}\end{array}\right)\middle| H_0\right) = \mathbf{0} \tag{33}$$

Then using the kernels

$$c_1 = \begin{pmatrix} 1 & 0 & -1\\ 0 & 0 & 0\\ -1 & 0 & 1\end{pmatrix} \quad c_2 = \begin{pmatrix} 0 & 1 & 0\\ -1 & 0 & -1\\ 0 & 1 & 0\end{pmatrix} \tag{34}$$

which both measure the local torsion lead to the homogeneity measure

$$h_\tau = (c_1 * g)^2 + (c_2 * g)^2 \tag{35}$$

which again is χ^2_2-distributed in case the mean curvature is locally zero. The expectation of h_τ is

$$\mathsf{E}(h_\tau) = 8\sigma_n^2 \tag{36}$$

allowing to use the same procedure for estimating σ_n^2 from the histogram of h_τ.

4 Variance Equalization

The estimated noise variance function may be used when thresholding on functions of the intensity image. This was the line of thought in the feature extraction up to now (cf. Förstner 1994, Fuchs 1998 and sect. 6).

The disadvantage is the need to refer the threshold to some center pixel of an operating window. E. g. when thresholding the gradient magnitude determined with the Sobel operator, one needs to use a representative intensity in the 3×3-window for determining the signal dependent noise variance. Actually all 9 intensity values have different noise variance, making a rigorous error propagation tedious, without being sure that the rigorous calculations are much better than the approximate calculations.

In order to avoid this situation one can transform the image such that the intensities in the resulting image have equal noise variance.

4.1 Principle

The idea is to find a pixel wise transformation

$$\dot{g} = T(g) \tag{37}$$

such that $\sigma_{\dot{g}} = \sigma_0 = const.$ is independent on g. From

$$\sigma_{\dot{g}}^2 = \sigma_0^2 = \left(\frac{dT}{dg}\right)^2 \sigma_g^2 = \left(\frac{dT}{dg}\right)^2 s(g) \tag{38}$$

we find

$$dT = \sigma_0 \frac{dg}{\sqrt{s(g)}} \tag{39}$$

and obtain

$$\dot{g} = T(g) = \sigma_0 \int_{l=0}^{g} \frac{dl}{\sqrt{s(l)}} + C \tag{40}$$

where the two parameters σ_0 and C can be chosen freely, e. g. that $T(g_{\min}) = g_{\min}$ and $T(g_{\max}) = g_{\max}$ resulting in an image $\dot{g} = T(g)$ with the same intensity range.

This *variance equalization* simplifies subsequent analysis steps as the noise characteristics is homogeneous throughout the image.

4.2 Linear Variance Function

In case $\sigma_n^2(g) = s(g) = a + b\ g$ we can evaluate the integral and obtain:

$$\dot{g} = t(g) = \frac{2\sigma_0}{b}\sqrt{a + b\ g} + C \tag{41}$$

again with the possibility to freely choose the two parameters σ_0 and C.

4.3 General Variance Function

In the case of general $s(g)$ and piecewise linear approximation the integral can also be evaluated algebraically.

Alternatively, one could sample $s(g)$ and determine

$$\dot{g} = t(g) = \sigma_0 \sum_{l=0}^{g} \frac{1}{\sqrt{s(l)}} + C \tag{42}$$

5 Information Preserving Filtering

After having characterized the image noise and possibly transformed the image we now want to increase the signal to noise ratio by filtering the image. Usually signal and noise cannot completely be separated. Therefore any filtering meant to suppress noise at the same time suppresses signal. Depending on the signal and the noise model different filters may be optimal in increasing the signal to noise ratio.

We want to develop a filter which locally leads to a best restoration, i. e. prediction of the underlying signal. We start with the most simple case, where both signals are Gaussian with known homogeneous and isotropic statistics, requiring a Wiener filter as optimal filter. We then modify this filter to locally adapt to the signal content, which we estimate from the given data. The presentation follows (Förstner 1991).

5.1 The Wiener Filter

The Wiener filter starts from the model eq. (2)

$$\underline{\boldsymbol{g}} = \underline{\boldsymbol{f}} + \underline{\boldsymbol{n}} \tag{43}$$

with

$$\mathsf{E}(\underline{\boldsymbol{f}}) = \mathsf{E}(\underline{\boldsymbol{n}}) = \mathbf{0} \tag{44}$$

$$\mathsf{D}(\underline{\boldsymbol{f}}) = \boldsymbol{\Sigma}_{ff} \qquad \mathsf{D}(\underline{\boldsymbol{n}}) = \boldsymbol{\Sigma}_{nn} \qquad \mathsf{Cov}(\underline{\boldsymbol{f}}, \underline{\boldsymbol{n}}) = \boldsymbol{\Sigma}_{fn} = \mathbf{0} \tag{45}$$

Under these conditions the filter leading to best, i. e. most precise results is the Wiener filter (Wiener 1948, Moritz 1980, p. 80):

$$\widehat{\underline{\boldsymbol{g}}} = \boldsymbol{\Sigma}_{fg}\boldsymbol{\Sigma}_{gg}^{-1}\underline{\boldsymbol{g}} = \boldsymbol{\Sigma}_{ff}(\boldsymbol{\Sigma}_{ff} + \boldsymbol{\Sigma}_{nn})^{-1}\underline{\boldsymbol{g}} = \boldsymbol{W}\underline{\boldsymbol{g}} \tag{46}$$

A slightly different approach has been developed in and (Weidner 1994a and Weidner 1994b). It integrates a variance component estimation for determining the local statistical behavior of the signal and the noise (cf. Förstner 1985). The approximate solution proposed in (Förstner 1991) and given and generalized for range images below, has proven to be similar in performance but computationally much more efficient.

5.2 Approximation of the Autocovariance Function

The statistics of the image function theoretically can be described by its auto covariance function. This, however, would not capture the locally varying structure. This structure also shows severe orientation dependencies at least at edges but also in textured areas, e. g. within agricultural areas.

Therefore we want to describe the local structure of the image function f by an inhomogeneous and anisotropic auto covariance function $k(\boldsymbol{x})$, depending on the difference $\boldsymbol{x} = (r_2 - r_1, c_2 - c_1)^\mathsf{T}$.

The auto covariance function can be assumed to decay smoothly with increasing distance from $\boldsymbol{x} = \boldsymbol{0}$. In a second order approximation it can be characterized by the curvature at $\boldsymbol{x} = \boldsymbol{0}$ (cf. Moritz 1980, p. 175). As an example, assume the auto covariance function has the form:

$$k(\boldsymbol{x}) = \sigma_f^2 \mathrm{e}^{-\frac{1}{2}\boldsymbol{x}^\mathsf{T}\boldsymbol{S}\boldsymbol{x}} \tag{47}$$

with a symmetric positive definite matrix $\boldsymbol{S}$, then its Hessian containing the second derivatives at $\boldsymbol{x} = \boldsymbol{0}$ is given by

$$\boldsymbol{H}_k = \begin{pmatrix} k_{rr} & k_{rc} \\ k_{rc} & k_{cc} \end{pmatrix} = -\sigma_f^2 \boldsymbol{S} \tag{48}$$

Now, due to the moment theorem (cf. Papoulis 1984), the Hessian of the auto covariance function is directly related to the variances and covariances of the gradient of the true signal f by

$$\boldsymbol{H}_k = -\mathsf{D}(\nabla f) = -\begin{pmatrix} \sigma_{f_r}^2 & \sigma_{f_r f_c} \\ \sigma_{f_r f_c} & \sigma_{f_c}^2 \end{pmatrix} = -G_t * \begin{pmatrix} f_r^2 & f_r f_c \\ f_c f_r & f_c^2 \end{pmatrix} = -\overline{\nabla f \, \nabla^\mathsf{T} f} \tag{49}$$

This relation allows to capture the essential part of the auto covariance function from a quite local computation, namely the local dispersion of the gradient of the image function.

As we do have no access to the true image f, but the noisy image g we need to estimate $\mathsf{D}(\nabla f)$. This either can be achieved by iteratively estimating f or by using the relation between $\mathsf{D}(\nabla f)$ and $\mathsf{D}(\nabla g)$:

$$\mathsf{D}(\nabla g) = \mathsf{D}(\nabla f) + \sigma_{n'}^2 \boldsymbol{I} \tag{50}$$

The covariance matrix $\mathsf{D}(\nabla g)$ and $\sigma_{n'}^2$ can be estimated from the given image using (49) and the procedure given in sect. 3.1. With the eigenvalue decomposition $\widehat{\mathsf{D}}(\nabla g) = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^\mathsf{T}$ we obtain

$$\widehat{\mathsf{D}}(\nabla f) = \boldsymbol{C}\widehat{\boldsymbol{\Lambda}}\boldsymbol{C}^\mathsf{T} \tag{51}$$

with

$$\widehat{\lambda}_i = \max(\lambda_i - \sigma_{n'}^2, 0) \tag{52}$$

5.3 An Adaptive Wiener Filter for Intensity Images

In principle the vectors $\underline{\boldsymbol{g}}$ etc. in eqs. (43) – (46) contain the complete image. This is due to the correlation structure of $\underline{\boldsymbol{f}}$, which links all pixels, leading to a full covariance matrix $\boldsymbol{\Sigma}_{ff}$.

The filter we propose is realized as a, possibly iterated, convolution with a small but signal dependent kernel.

In order to motivate this setup, assume the vector $\underline{\boldsymbol{g}}$ contains only the 9 values of a 3×3 window. In a first instance, the covariance between intensity values is assumed to only depend on the distance, thus is homogeneous and isotropic, and can be represented by an auto covariance function

$$\mathsf{Cov}(g(r_1, c_1), g(r_2, c_2)) = k(d_{12}) \tag{53}$$

where $d_{12} = \sqrt{(r_2 - r_1)^2 + (c_2 - c_1)^2}$ is the distance between the pixels. Assuming the correlation to fall off rapidly enough, the Wiener filter with isotropic covariance function can be written as convolution with a small completely symmetric kernel

$$\boldsymbol{w}^{(i)} = \begin{pmatrix} w_1 & w_2 & w_1 \\ w_2 & w_0 & w_2 \\ w_1 & w_2 & w_1 \end{pmatrix} \tag{54}$$

In case the covariance function is anisotropic the kernel will only be point symmetric:

$$\boldsymbol{w}^{(a)} = \begin{pmatrix} w_1 & w_2 & w_3 \\ w_4 & w_0 & w_4 \\ w_3 & w_2 & w_1 \end{pmatrix} \tag{55}$$

As $\mathsf{E}(\boldsymbol{f}) = \boldsymbol{0}$ the weights do not sum to 1 in general. In order to be able to handle signals with $\mathsf{E}(\boldsymbol{f}) \neq \boldsymbol{0}$ the elements need to sum to 1, thus $w_0 + 4w_1 + 4w_2 = 1$, in order that a constant signal is reproduced by the filtering.

The idea now is to locally determine the covariance function and approximately determine the 5 parameters w_0 to w_4 in $\boldsymbol{w}^{(a)}$

We propose a simple expression for the convolution kernel $\boldsymbol{w}$. It should fulfill the following conditions:

1. It should allow strong smoothing in homogeneous areas of the image or, equivalently, in case the signal has the same or a lower variance than the noise. Therefore it needs to depend on the variance of the image noise.
2. It should preserve the sharpness of edges and corners. Therefore it needs to depend on the local auto covariance function.

Thus the idea is to make the coefficients w_i dependent on the curvature $\boldsymbol{S}$ of the auto covariance function and the noise variance σ_n^2.

We propose a simple weight function

$$w(\boldsymbol{x}) = \frac{C}{1 + \frac{1}{2}\frac{\boldsymbol{x}^\mathsf{T}\overline{\nabla f\, \nabla^\mathsf{T} f}\boldsymbol{x}}{2\,\sigma_{n'}^2}} \tag{56}$$

the constant C being used for normalization.

It has the following properties:

1. For a constant noisy signal $\underline{g} = f + \underline{n}$ the estimated covariance matrix $\widehat{\mathsf{D}}(\nabla f) = \overline{\nabla\widehat{f}\,\nabla^\mathsf{T}\widehat{f}} \approx \mathbf{0}$, thus all weights are equal. The weighting kernel is the box filter R_3

$$\boldsymbol{w} = \frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \tag{57}$$

2. in case of a steep straight edge in r-direction we will have

$$\mathsf{D}(\nabla f) = \begin{pmatrix} \sigma_{f_r}^2 & 0 \\ 0 & 0 \end{pmatrix} \tag{58}$$

   thus for $\sigma_{f_r}^2 >> \sigma_{n'}^2$ we obtain the weight matrix

$$\boldsymbol{w} = \frac{1}{3}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{59}$$

3. Finally, in case of a sharp symmetric corner we will have

$$\mathsf{D}(\nabla f) = \frac{\sigma_{f'}^2}{2}\boldsymbol{I} \tag{60}$$

   thus for $\sigma_{f'}^2 >> \sigma_{n'}^2$ we obtain

$$\boldsymbol{w} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{61}$$

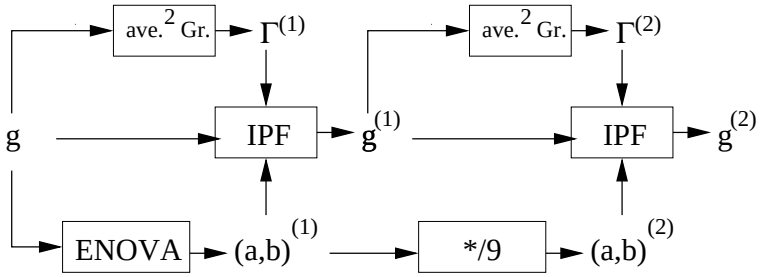
   which does not change the signal. The same situation holds for an isolated bright or dark spot.

Thus the smoothing properties are those intended.

Three points are worth mentioning when implementing this filter:

1. As mentioned above, the true signal is not available. A possibility to obtain the characteristics of the true signal f would be to use the estimated covariance matrix of the gradients of f in eq. (52).
2. In order to avoid the eigenvalue decomposition at each pixel, one could alternatively use an estimate $\widehat{f}$ to determine the covariance matrix $\mathsf{D}(\nabla f)$. In a first implementation one could use the signal g as an estimate for f.

Fig. 5. *shows an iterative version of the information preserving filter. The matrix* $\boldsymbol{\Gamma}$ *denotes the dispersion* $\mathsf{D}(\nabla g \nabla^{\mathsf{T}} g)$ *or* $\mathsf{D}(\boldsymbol{H}^2(d))$ *of the gradients or the Hessian resp..* ENOVA *denotes the estimation of the noise variance, here assumed to yield the two parameters* a *and* b *of a linear noise variance model.* IPF *is the simple version information preserving filtering with locally adaptive weights. Observe, we reduce the variances by a factor of 9, which is the maximum reduction which can be achieved by filtering with a* 3×3*-kernel.*



A better approximation of the true signal could be obtained by starting with a smoothed version $\overline{f} = G_s * f$ for determining the the covariance matrix $\mathsf{D}(\nabla \overline{f})$. Then adaptive weights are more realistic. This especially holds, in case the noise is not very small. Otherwise, the covariance matrix would not show the full structure of the signal.

3. In case the noise is very large, we would like to smooth more than with a box filter (57). Then we could apply the filter iteratively. This is indicated in Fig. 5.

In our implementation we use the filter kernels (Roberts gradient):

$$\left(\frac{\partial}{\partial r}\right)_R = \frac{1}{2}\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \qquad \left(\frac{\partial}{\partial c}\right)_R = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \tag{62}$$

for determining the gradients and the filter kernel

$$R_4 = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \tag{63}$$

instead of G_t for the integration according to eq. (49).

5.4 An Adaptive Wiener Filter for Range Images

In range images the trend $\mathsf{E}(\underline{d})$ cannot be assumed to be locally constant, or, equivalently, $\mathsf{E}(\nabla \underline{d}) \neq \mathbf{0}$. The most simple assumption would be to assume a linear trend, or, equivalently $\mathsf{E}(\boldsymbol{H}(\underline{d})) = \mathbf{0}$ (cf. above).

Starting from the range image d, we then can use the gradient image ∇d as two channel image: $\boldsymbol{f} = \nabla d$. Using the relation $\boldsymbol{H}(d) = \nabla\nabla d$ the mean quadratic

gradient $G_t * (\nabla \boldsymbol{f}\, \nabla^\mathsf{T} \boldsymbol{f})$ then is given by:

$$\overline{\boldsymbol{H}^2(d)} = G_t * (\nabla \boldsymbol{f}\, \nabla^\mathsf{T} \boldsymbol{f}) = G_t * \boldsymbol{H}^2(d) \tag{64}$$

Thus we can use the weight function

$$w(\boldsymbol{x}) = \frac{1}{1 + \frac{1}{2} \frac{\boldsymbol{x}^\mathsf{T} \overline{\boldsymbol{H}^2(d)} \boldsymbol{x}}{2\,\sigma_{n''}^2}} \tag{65}$$

6 Fusing Channels: Extraction of Linear Features

We now want to discuss one of the many algorithms for extracting geometric features from digital images, namely the extraction of edge pixels, which then may be grouped to linear image features (Fuchs 1998).

Edge pixels are meant to be borders of homogeneous regions. Thus they show two properties:

1. The homogeneity is significantly larger than in homogeneous regions.
2. The homogeneity is locally maximum across the edge.

Therefore we need to discuss detection and localization of edge pixels.

The problem we want to solve is how to integrate the information of different channels of a multi-channel image.

6.1 Detecting Edge Pixels

The detection of edge pixels can be performed as a hypothesis test, namely calling all pixels edge pixels in case the homogeneity $h_1 = |\nabla g|^2$ is significant with respect to the noise in the image. Thus edge pixels actually are pixels which are non-homogeneous, thus indicate places where there may be signal, e. g. and edge.

Detecting Edge Pixels in Intensity Images We first assume constant noise variance.

In order to capture the full information in the neighborhood of a pixel we test the averaged homogeneity $\overline{h}_1 = G_t * h_1$ for significance. This at the same time allows to detect thin lines, not only edges, as the smoothing caused by the convolution with G_t then covers both sides of a thin line.

We use the test statistic

$$\underline{z}_1 = \frac{\overline{\underline{h}_1}}{\sigma_{n'}^2} = \frac{\overline{|\nabla \underline{g}|^2}}{\sigma_{n'}^2} = \frac{tr\left(G_t * (\nabla \underline{g} \nabla^\mathsf{T} \underline{g})\right)}{\sigma_{n'}^2} \tag{66}$$

which in case the pixels lies in a homogeneous region is χ_2^2-distributed. Observe, we need to use the same differentiation kernel for determining the partial derivatives ∇g of the signal as for determining $\sigma_{n'}^2$ (cf. eq. (26)).

Fig. 6. *shows first a section of a digitized aerial image. The large compound image below shows left the result of the information preserving filter (IPF)and below the difference to the original, and right the result of a smoothing with a Gaussian with* $\sigma = 1.4$ *and below the difference to the original. Observe: the IPF smoothes the vegetation areas and the roof tilings quite strongly without smearing out the edges and the line being a hair of the analog film. The Gaussian filter does not smooth the vegetation areas and the tilings too much but at the same time smoothes the edges. The difference images demonstrate that the edges are better preserved using IPF. The notion* information preserving *obviously needs an application dependent discussion, as a user might be interested in the roof tilings.*

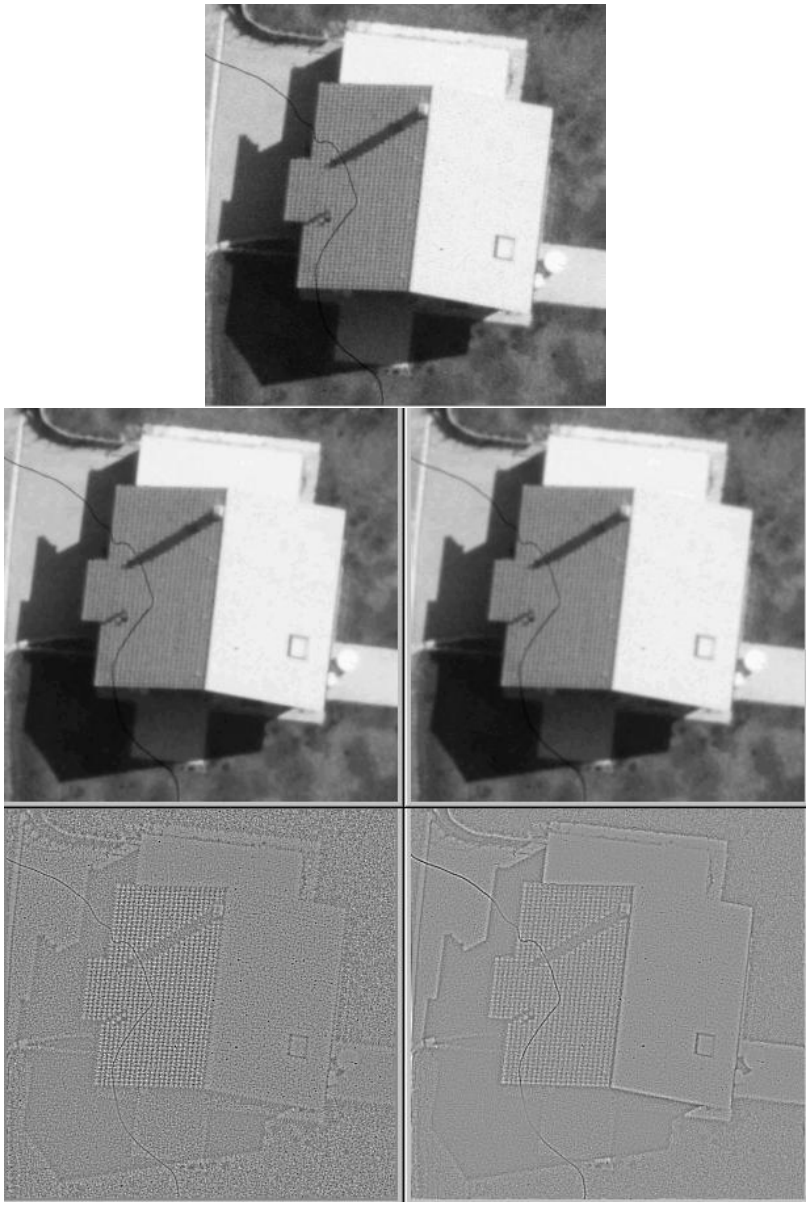
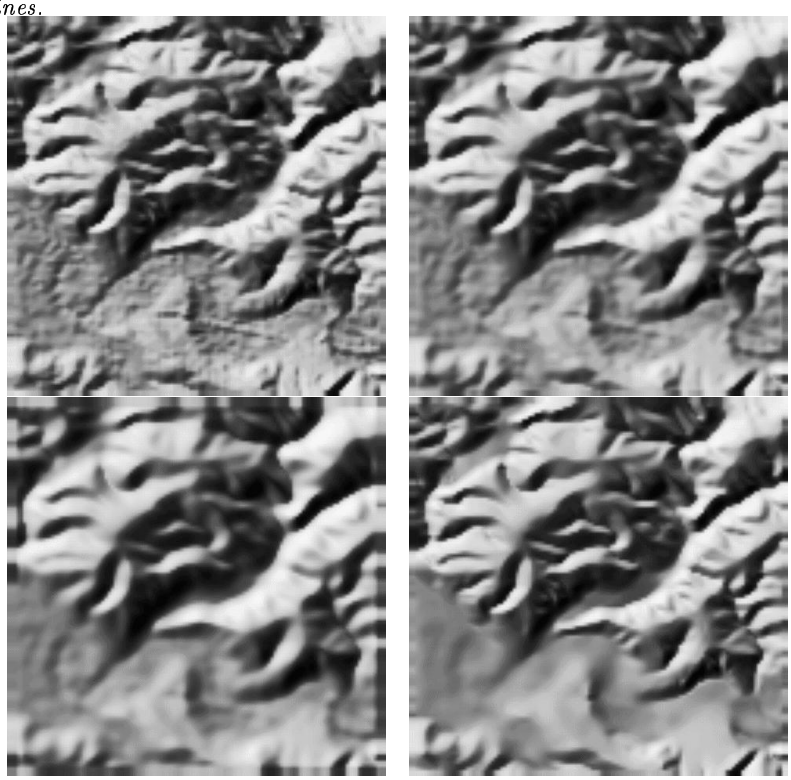


Fig. 7. *shows top left the original of a DHM, shaded. Top right shows the effect of a linear filter with $G_{0.7}$ below left the effect of a linear filter with $G_{1.2}$ and below left shows the effect of the information preserving filter, in the rigorous version from Weidner 1994a). Observe the smoothing effect in the homogeneous areas and the presservation of structure lines.*



Thus pixels with

$$\frac{\overline{|\nabla g|^2}}{\sigma^2_{n'}} > \chi^2_{2,\alpha} \tag{67}$$

are significantly non-homogeneous, thus likely to be edge pixels.

In case the noise variance is signal dependent the determination of $\sigma^2_{n'}$ would need to take the different variances of the neighboring pixels into account. An approximation would be to assume all pixels involved in the determination of the gradient to have the noise variance of the center pixel. However, first applying a variance equalization, leads to an efficient procedure as $\sigma^2_{n'}$ is constant for the complete image $\dot{g}$.

Detecting Edge Pixels in Color Images In color images we need to integrate the information in the different channels. Applying the tests individually

in general would lead to conflicts, as there may be an edge in one channel where there is no edge in the other image. The significance of the individual results is difficult to evaluate.

We therefore proceed differently and determine a homogeneity measure on the multi valued image.

Again we first assume the different channels to have constant noise variance. Then we can measure the homogeneity by

$$\underline{z} = \sum_{k=1}^{K} \frac{\overline{\underline{h}_{k;1}}}{\sigma^2_{n'_k}} = \sum_{k=1}^{K} \frac{\overline{|\nabla \underline{g}_k|^2}}{\sigma^2_{n'_k}} \tag{68}$$

which now is χ^2_{2K}-distributed in case the multi-channel image is homogeneous. This is due to the fact that $\underline{z}$ is the sum of squares of $2K$ normally distributed variates. Thus pixels with

$$\sum_{k=1}^{K} \frac{\overline{|\nabla g_k|^2}}{\sigma^2_{n'_k}} > \chi^2_{2K,\alpha} \tag{69}$$

are significantly non-homogeneous, thus likely to be edge pixels.

Assuming general noise behavior again reveals the noise equalization to simplify matters. Here we obtain an even more simple expression for the test statistic:

$$\underline{z} = \frac{1}{\sigma^2_{\dot{n}'}} \sum_{k=1}^{K} \overline{\dot{\underline{h}}_k} = \frac{1}{\sigma^2_{\dot{n}'}} \sum_{k=1}^{K} \overline{|\nabla \dot{\underline{g}}_k|^2} \tag{70}$$

which shows that we just add the homogeneities of the normalized channels and refer to the common noise variance $\sigma^2_{\dot{n}'}$ of the gradients.

Detecting Edges in Range Images Edges in range images are pixels which do not lie on flat surfaces, thus are expected to be pixels where the curvature is significant compared to the noise.

In range images we start with the gradient image $\boldsymbol{g} = \nabla d$ as two-channel image. We want to use the same argument for fusing the channels here.

Assuming constant noise variance we instead of the homogeneity h_1 derived from the first derivatives, obtain for the homogeneity h_2 derived from the second derivatives, using $h_1(z) = z_x^2 + z_y^2$

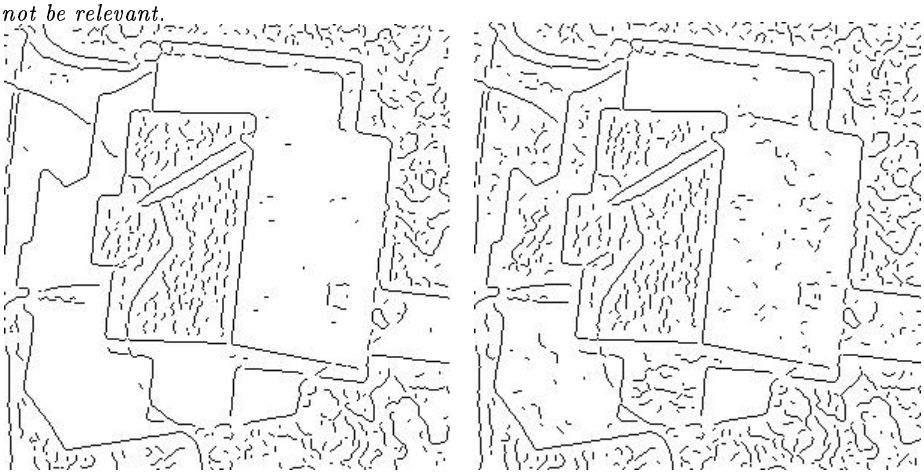
$$h_2 = h_1(d_x) + h_1(d_y) = (d_{xx}^2 + d_{xy}^2) + (d_{yx}^2 + d_{yy}^2) \tag{71}$$

$$= d_{xx}^2 + 2d_{xy}^2 + d_{yy}^2 \tag{72}$$

$$= tr\boldsymbol{H}^2(d) = \lambda_1^2(\boldsymbol{H}) + \lambda_2^2(\boldsymbol{H}) = \kappa_1^2 + \kappa_2^2 \tag{73}$$

in analogy to (64). Observe, h_2 is the quadratic variation. It only is zero in case the pixel's surrounding is flat, as only if the two principle curvature a both zero the homogeneity measure h_2 is zero.

Fig. 8. *shows the edges from the original image (left) and from the information preserving filtered image. Observe the low contrast edges, detected in the image filterd with IPF ate the expense of getting additional, statistically significant edges, which might not be relevant.*



Thus the generalization of the homogeneity measure to multi-channel images together with the use of the gradient as two channel-image for describing the form in a range image leads to a very meaningful result.

The differentiation kernels for determining the second derivatives could be approximations of the corresponding Gaussian's

$$\frac{\partial^2}{\partial x^2} G_s(x,y) = \frac{x^2 - s^2}{s^4} G_s(x,y) \tag{74}$$

$$\frac{\partial^2}{\partial x \partial y} G_s(x,y) = \frac{xy}{s^4} G_s(x,y) \tag{75}$$

$$\frac{\partial^2}{\partial y^2} G_s(x,y) = \frac{y^2 - s^2}{s^4} G_s(x,y) \tag{76}$$

which are orthogonal. Thus normalization with the noise variances leads to the test statistics

$$\underline{z} = \frac{d_{xx}^2}{\sigma_{n_{xx}}^2} + \frac{d_{xy}^2}{\sigma_{n_{xy}}^2} + \frac{d_{yy}^2}{\sigma_{n_{yy}}^2} \tag{77}$$

The noise variances in the denominators can be explicitly derived for the case of the Gaussian kernels eq. (74):

$$\sigma_{n_{xx}}^2 = \sigma_{n_{yy}}^2 = \int\int_{-\infty}^{\infty} \left(\frac{\partial^2}{\partial x^2} G_s(x,y) \right)^2 dx\, dy\, \sigma_n^2 = \frac{3}{16\pi s^6} \sigma_n^2 \tag{78}$$

$$\sigma_{n_{xy}}^2 = \int\int_{-\infty}^{\infty} \left(\frac{\partial^2}{\partial x \partial y} G_s(x,y) \right)^2 dx\, dy\, \sigma_n^2 = \frac{1}{16\pi s^6} \sigma_n^2 \tag{79}$$

which has shown to be a good approximation for discrete kernels approximating the Gaussian.

The test statistic is χ^2_3-distributed in case the region is homogeneous, which can be used to perform a statistical test for edge detection in range images.

6.2 Localizing Edge Pixels

The statistical test on pixels which are significantly non-homogeneous leads to edge areas. Within these edge areas the edge or the boundary between homogeneous regions is to be expected.

A classical approach to detect edges is to take those pixels where the slope lines show an inflection point. In a one dimensional signal these positions are given by $g'' = 0$, and $g'''g' < 0$, where the second condition is needed to avoid non valid inflection points. The generalization to two dimensional signals motivates the zero crossings of the Laplacian $g_{rr} + g_{cc}$ as edges, with a similar constraint to avoid false edges.

Obviously this technique cannot be generalized to multichannel images. As – in one dimension – edges can be defined by the maxima of g'^2 we easily can generalize this into two dimensions, by looking for local maxima of $|\nabla g|^2 = tr(\nabla g \nabla^\mathsf{T} g)$ in the direction of the gradient. We actually use the locally averaged squared gradient $G_t * |\nabla g|^2 = tr[G_t * (\nabla g \nabla^\mathsf{T} g)]$. This has the advantage of higher stability, and at the same time allows to detect bright or dark *lines*. The orientation of the gradient can be determined by the eigenvectors of $G_t * (\nabla g \nabla^\mathsf{T} g)$.

Generalization now is easy, as we only need to take the possibly weighted sum of these averaged squared gradients of the individual channels:

$$\sum_{k=1}^{K} \frac{G_t * (\nabla g_k \nabla^\mathsf{T} g_k)}{\sigma^2_{\dot{n}'_k}} = \frac{1}{\sigma^2_{\dot{n}'}} \sum_{k=1}^{K} G_t * (\nabla \dot{g}_k \nabla^\mathsf{T} \dot{g}_k) \tag{80}$$

where in the second expression the normalization with $\sigma^2_{\dot{n}'}$ is not necessary for location.

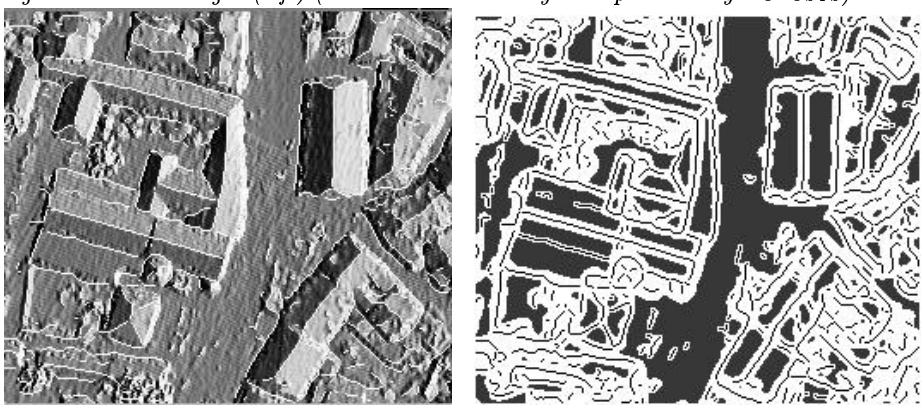
In range images the squared gradient $\nabla g \nabla^\mathsf{T} g$ just needs to be replaced by the squared Hessian $\boldsymbol{H}^2(d)$, which again demonstrates the simplicity of the approach.

7 Outlook

The paper presented tools for preprocessing intensity, color and range images for feature extraction. The idea was to exploit the full knowledge about the image as far as possible. The examples demonstrated the feasibility of the approach.

Obviously the individual steps may be conceptually integrated to a much larger extent. The iterative version of the information preserving filter obviously is an approximation, which needs to be analyses and possibly overcome by a more rigorous solution. The edge detection should be linked with the information

Fig. 9. *shows the first derivative image of a range image having a density of appr. 0.6 m (acknowledging* TOPOSYS*) with the edges overlaid (left) and homogeneous regions together with the edges (left) (The data have kindly been provided by* TOPOSYS*)*



preserving filter in order to exploit the *inhomogeneous* accuracy of the intensity values after the adaptive filter. The detection of edges in range images may take advantage of the range images in homogeneous, i. e. flat regions. Finally methods should be developed which allow to integrate preknowledge about the form of edges in order to increase the accuracy and the resolution of feature extraction.

References

BRÜGELMANN, R.; FÖRSTNER, W. (1992): Noise Estimation for Color Edge Extraction. In: FÖRSTNER, W.; RUWIEDEL, S. (Eds.), *Robust Computer Vision*, pages 90–107. Wichmann, Karlsruhe, 1992.

DAINTY, J. C.; SHAW, R. (1974): *Image Science.* Academic Press, 1974.

FÖRSTNER, W. (1985): Determination of the Additive Noise Variance in Observed Autoregressive Processes Using Variance Component Estimation Techinque. *Statistics & Decisions*, Supplement Issue 2:263–274, 1985.

FÖRSTNER, W. (1991): *Statistische Verfahren für die automatische Bildanalyse und ihre Bewertung bei der Objekterkennung und -vermessung*, Band 370 der Reihe C. Deutsche Geodätische Kommission, München, 1991.

FÖRSTNER, W. (1994): A Framework for Low Level Feature Extraction. In: EKLUNDH, J. O. (Ed.), *Computer Vision - ECCV '94, Vol. II*, Band 802 der Reihe LNCS, pages 383–394. Springer, 1994.

FUCHS, C.; LANG, F.; FÖRSTNER, W. (1994): On the Noise and Scale Behaviour of Relational Descriptions. In: EBNER, HEIPKE, EDER (Ed.), *Int. Arch. f. Photogr. and Remote Sensing, Vol. 30, 3/2*, Band XXX, pages 257–267, 1994.

FUCHS, C. (1998): *Extraktion polymorpher Bildstrukturen und ihre topologische und geometrische Gruppierung.* DGK, Bayer. Akademie der Wissenschaften, Nünchen, Reihe C, Heft 502, 1998.

MAYER, H. (1999): Automatic Object Extraction from Aerial Imagery – A Survey Focussing on Buildings. *Computer Vision and Image Understanding*, 74(2):138–149, 1999.

MORITZ, H. (1980): *Advanced Physical Geodesy.* Herbert Wichmann Verlag, Karlsruhe, 1980.

PAPOULIS, A. (1984): *Probability, Random Variables, and Stochastic Processes.* Electrical Engineering. McGraw-Hill, 2. edition, 1984.

WAEGLI, B. (1998): Investigations into the Noise Characteristics of Digitized Aerial Images. In: *Int. Arch. for Photogr. and Remote Sensing, Vol. 32-2*, pages 341–348, 1998.

WEIDNER, U. (1994): Information Preserving Surface Restoration and Feature Extraction for Digital Elevation Models. In: *ISPRS Comm. III Symposium on Spatial Information from Digital Photogrammetry and Computer Vision, Proceedings*, pages 908–915. SPIE, 1994.

WEIDNER, U. (1994): Parameterfree Information-Preserving Surface Restauration. In: EKLUNDH, J. O. (Ed.), *Computer Vision - ECCV '94*, Band 802 der Reihe LNCS, pages xxx–xxx. Springer, 1994.

WIENER, N. (1948): *Cybernetics.* MIT Press, 1948.