A Framework for Low Level Feature Extraction

Wolfgang Förstner

Institut für Photogrammetrie, Universität Bonn, Nußallee 15, D-53115 Bonn
e-mail: wf@ipb.uni-bonn.de

Abstract. The paper presents a framework for extracting low level features. Its main goal is to explicitely exploit the information content of the image as far as possible. This leads to new techniques for deriving image parameters, to either the elimination or the elucidation of "buttons", like thresholds, and to interpretable quality measures for the results, which may be used in subsequent steps. Feature extraction is based on local statistics of the image function. Methods are available for blind estimation of a signal dependent noise variance, for feature preserving restoration, for feature detection and classification, and for the location of general edges and points. Their favorable scale space properties are discussed.
Keywords: low level features, keypoints, edge detection, segments, local image statistics, noise estimation, restoration, adaptive thresholds, scale space, quality evaluation.

1 Introduction

Feature extraction is the first crucial step of all image analysis procedures which aim at symbolic processing of the image content. Basic features of nearly all symbolic, i. e. non-iconic image descriptions, are points, edges and regions. The research in feature extraction is rich and dozens of procedures have been proposed for the extraction of these feature types. However, no coherent theory seems to be available suited to extract features of all types simultaneously. The lack of a theoretical basis for feature extraction was the stimulus to search for the framework documented in this paper. It had to fulfill the following requirements:

1. Since feature extraction is meant to support image interpretation, modelling needs to start in object space, from which via the sensing model an image model can be derived. This excludes all models starting at the grid structure of the digital image.
2. Feature extraction has to treat the basic features simultaneously in order to avoid the necessity of developing conflict resolution strategies.
3. For self-diagnosis, only models including stochastic components for describing the image content are suitable. This, at the same time, allows to reduce the number of "buttons" controlling the result and to retain those with a clear interpretation.
4. The features should show "nice behaviour" over scale space (cf. [24]) and should have small bias supporting coarse-to-fine strategies. Therefore, only nonlinear filters seem to be suited for feature extraction (cf. [1]).

The proposed framework intentionally restricts to low level features. No attempt is made to include specific scene knowledge. The following four steps for low-level feature extraction are discussed: 1.) estimation of noise characteristics, 2.) information preserving restoration 3.) feature detection and 4.) feature location. Unification of the steps is obtained by analysing the local autocovariance function or - equivalently - the local power spectrum. This technique has received a great deal of attention for more than 15 years, due to its versability in representing local image properties both using geometric and statistical tools (cf. [10], [17], [16], [2], [14], [23], [18]).

The novelty of the proposed approach lies in the *integration* and - as a by product - in the *simplification* of existing techniques for feature extraction and the provision of statistically founded *evaluation measures* for making the quality of the individual steps transparent and *objective*.

2 Image Model

Describing image analysis, with feature extraction being a part of it, requires the setting up of models of the scene to be recovered, of the sensing process used for observation, of the image as information memory, and of the tools used for inverting the imaging process, yielding an estimated or inferred description of the scene (cf. [12], [20]). The image model is derived in three steps.

The Ideal Continuous Image: We first assume the camera to be modelled as a pinhole camera, the lighting to be diffuse and the light sensitive image area to be of unlimited resolution.

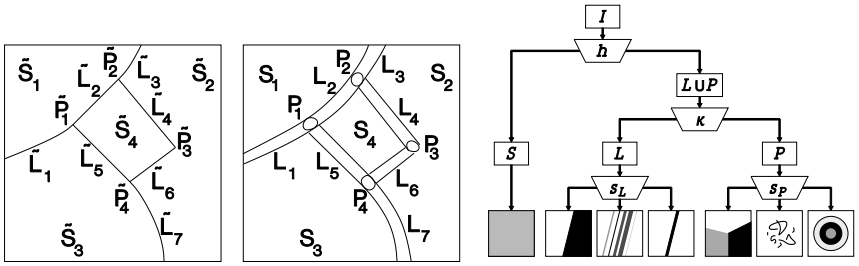
The image area $\mathcal{I}$ therefore consists of the ideal or true *homogeneous segments* $\tilde{\mathcal{S}}_i$, where the intensity function $\tilde{f}(x,y)$ or some locally computable function of $\tilde{f}$ is assumed to be *piecewise smooth*. The segments are assumed to show *piecewise smooth boundary lines* $\tilde{\mathcal{L}}_j$. *Points* $\tilde{\mathcal{P}}_k$ are either boundary points of high curvature or junctions (cf. Fig. 1a). A classification of all image points (x,y) is thus possible:

$$\mathcal{I} = \tilde{\mathcal{S}} + \tilde{\mathcal{L}} + \tilde{\mathcal{P}} = \bigcup_{i=1}^{\tilde{n}_s} \tilde{\mathcal{S}}_i + \bigcup_{j=1}^{\tilde{n}_l} \tilde{\mathcal{L}}_j + \bigcup_{k=1}^{\tilde{n}_p} \tilde{\mathcal{P}}_k \tag{1}$$

The Real Continuous Image: Assuming a real objective, more general lighting condition and a light sensitive image area of limited resolution, in general leads to, though continuous, blurred images; the blur generally being non-homogenous and anisotropic. Due to the blurring defining a possibly local and anisotropic scale $\Sigma_1(x,y)$ we obtain *segment-regions* $\mathcal{S}_i$, often referred to as *blobs*, *line-regions* $\mathcal{L}_j$ and *point-regions* $\mathcal{P}_k$ with a partitioning similar to eq. (1) (cf. Fig. 1b). As far as they are observable, the true points $\tilde{\mathcal{P}}_k$ and lines $\tilde{\mathcal{L}}_j$ are assumed to lie in the point- and line-regions.

The Observed Image: The observed image is a sampled and noisy version of the real image, now referring to a row-columm grid (r,c): $g(r,c) = f(r,c) + n(r,c)$. The noise is at least caused by the Poisson-process of the photon flux,

Fig. 1. shows the structure of the ideal image (a), with points, lines (edges), and segments as basic features and of the real image (b) containing point-, line- and segment-type regions. (c) shows the classification tree for image features.



by the electronic noise of the camera and - in case g is rounded to integers - by the rounding errors. For modelling purposes an approximation of $n(r,c)$ by a Gaussian distribution seems to be sufficient. The variance in general depends on the signal and in case of no over exposition can be assumed to be linearily dependent on f, thus $\sigma_n^2(r,c) = a + bf(r,c)$.

The task of feature extraction is to recover the position of the points and lines and the mutual relations between all features in order to obtain a relational description in the sense of a feature adjacency graph (cf. [9]) used for further processing steps.

3 Feature Extraction

3.1 Local Image Characteristics

We use measures for locally characterizing the image: the average squared gradient and the regularity of the intensity function with respect to junctions and circular symmetric features (cf. [8]).

The Average Squared Gradient: With the gradient $\nabla g = (g_x, g_y)^T$ we obtain the squared gradient $\boldsymbol{\Gamma} g$ as dyadic product

$$\boldsymbol{\Gamma} g = \nabla g \nabla g^T = \begin{pmatrix} g_x^2 & g_x g_y \\ g_y g_x & g_y^2 \end{pmatrix} . \tag{2}$$

The rotationally symmetric Gaussian function with centre $\mathbf{o}$ and standard deviation σ is denoted by $G_\sigma(x,y) = G_\sigma(x) * G_\sigma(y)$. This yields the *average squared gradient* image

$$\overline{\boldsymbol{\Gamma}_\sigma g}(x,y) = G_\sigma * \boldsymbol{\Gamma} g = \int\int \boldsymbol{\Gamma} g(u,v) G_\sigma(x-u, y-v) du dv . \tag{3}$$

The three essential elements of $\overline{\boldsymbol{\Gamma}_\sigma g}$ can be derived by three convolutions.

Thus we have to distinguish *two* scales:

- the *natural scale* σ_1, which is used to describe the blurring process (cf. above) or to generate the classical, 'natural' pyramids $G_{\sigma_1} * g$.
- the *artificial scale* $\sigma \doteq \sigma_2$, which is used to integrate the nonlinear function $\boldsymbol{\Gamma} g$ of g [15] [19] leading to the notion of an 'artificial' image pyramid $G_{\sigma_2} * \boldsymbol{\Gamma} g$.

Remark: The power spectrum $P_g(\boldsymbol{u})$ can be characterized by the effective bandwidth $\boldsymbol{B}_g \doteq \int \boldsymbol{u}\boldsymbol{u}^T P_g(\boldsymbol{u}) d\boldsymbol{u} / \int P_g(\boldsymbol{u}) d\boldsymbol{u}$ which, due to the moment theorem [22] is closely related to the average squared gradient by $\boldsymbol{B}_g = \overline{\boldsymbol{\Gamma} g}/(4\pi^2\sigma_g^2)$ with the constant variance σ_g^2. We will use this relation for *estimating the natural local scale.* •

Obviously the average squared gradient grasps essential parts of the *statistics* of the image function and - assuming local stationarity - may be estimated easily from (2) and (3). Diagonalization $\overline{\boldsymbol{\Gamma} g} = \boldsymbol{T}\boldsymbol{\Lambda}_g\boldsymbol{T}^T = \lambda_1(g)\boldsymbol{t}_1\boldsymbol{t}_1^T + \lambda_2(g)\boldsymbol{t}_2\boldsymbol{t}_2^T$ leads to an intuitive description of $\overline{\boldsymbol{\Gamma} g}$ with three parameters, each having very specific interpretations going beyond the ones discussed in [8], [5], and [13] and being the basis for the proposed framework:

1. The trace $h \doteq tr\overline{\boldsymbol{\Gamma} g} = \lambda_1(g) + \lambda_2(g) = \overline{\| \nabla g \|^2} \doteq \sigma_{g'}^2 = \sigma_{g_x}^2 + \sigma_{g_y}^2$ yields the total energy of the image function at (x, y), the edge busyness. We will use $tr\overline{\boldsymbol{\Gamma}_\sigma g}$ for *measuring the homogeneity* of segment-type features. It is approximately χ_m^2-distributed, where m is the number of pixels involved in the integration, allowing to fix a threshold solely based on some prespecified significance level and the noise variance. Assuming the image model $g = f + n$ with white noise, $\sigma_{g'}^2$ may be split into a signal and a noise component $\sigma_{g'}^2 = \sigma_{f'}^2 + \sigma_{n'}^2$. This will be the basis for *estimating the noise variance* and for *determining regularization factors* during restoration.
2. The ratio $v = \lambda_2/\lambda_1$ of the eigenvalues yields the degree of orientation or of an isotropy. If $\lambda_2 = 0$, we have anisotropic texture or straight general edges with arbitrary cross-sections. Using the local approximation $z = au + bv^2$ of a curved edge, one can easily show that the square κ^2 of the curvature of the isophote is given [7]: $\kappa^2 = v/\sigma^2$, ($\mid 2b \mid < \mid a \mid$). We will therefore use κ^2 for measuring the *smoothness of line-type features.*
3. The largest eigenvalue is an estimate for the *local gradient* of the texture or the edge. Due to the squaring, the phase information is lost [16]. The variance of this orientation is proportional to $\lambda_2/(\lambda_1 - \lambda_2)$, yielding an additional interpretation and showing that the variance of orientation is large for $\lambda_1 \approx \lambda_2$.

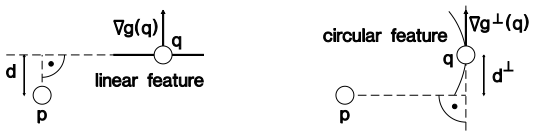
Regularity Measures: Junctions and circular symmetric features can be distinguished by analysing the local gradient field.

We use the weighted sums of the squared distances $d(\boldsymbol{p}, \boldsymbol{q})$ and $d^\perp(\boldsymbol{p}, \boldsymbol{q})$ of the reference point $\boldsymbol{p}$ to the line passing through $\boldsymbol{q}$ in the direction parallel and orthogonal to $\nabla g(\boldsymbol{q})$ resp. (cf. Fig. 2 a, b), where the weights are the squared gradient magnitude and the window is given by the Gaussian $G_\sigma(\boldsymbol{p})$.

The regularity measure for junctions is given by

$$S(\boldsymbol{p}, \sigma) = \int\int d^2(\boldsymbol{p}, \boldsymbol{q}) \parallel \nabla g(\boldsymbol{q}) \parallel^2 G_\sigma(\boldsymbol{p} - \boldsymbol{q}) d\boldsymbol{q} = tr\{\boldsymbol{\Gamma}^p(\boldsymbol{p}) * [\boldsymbol{p}\boldsymbol{p}^T G_\sigma(\boldsymbol{p})]\} \quad (4)$$

Fig. 2. relation between $\mathbf{p}$ and a feature at $\mathbf{q}$, the degree of fit of $g(x, y)$ at $\mathbf{p}$ wrt. a junction (a), circular symmetric feature (b).



for corners or junctions. $S^{\perp}(\boldsymbol{p}, \sigma)$ is obtained by replacing d by $d^{\perp}$.

Both types are unified by BIGÜN [2] showing them to be special spiral type features. Whereas BIGÜN uses the regularity measures for classifying pixels, we will also use them for locating features. S and $S^{\perp}$ correspond to the keypoint map $\hat{K}$ in [23]). They, however, are simpler and – what is essential – again allow statistical testing, as they are independent and approximately χ^2_m-distributed, with m being the number of pixels involved (cf. [5]).

An Important Link: The measures $\overline{\boldsymbol{\Gamma} g}$, S and $S^{\perp}$ have been independently motivated. However, they are not only loosely coupled by having the three elements g_x^2 , $g_x g_y$ and g_y^2 of the squared gradient $\boldsymbol{\Gamma} g$ in common. Much more, they are linked by a very important relation, namely by the covariance matrix $\boldsymbol{C}_{\hat{p}\hat{p}}$ of the estimated feature location $\hat{\boldsymbol{p}}$ (cf. [8]).

Assuming a *fixed* position $\boldsymbol{p}_0$ of the window minimizing $S_{p_0}(\boldsymbol{p}, \sigma) = \int\int (\boldsymbol{p} - \boldsymbol{q})^T \nabla g(\boldsymbol{q}) \nabla g^T(\boldsymbol{q}) (\boldsymbol{p} - \boldsymbol{q}) G_\sigma(\boldsymbol{p}_0 - \boldsymbol{q}) d\boldsymbol{q}$ with respect to $\boldsymbol{p}$ we obtain the estimate $\hat{\boldsymbol{p}}(\sigma) = \left(\int\int \boldsymbol{\Gamma} g(\boldsymbol{p}) G_\sigma(\boldsymbol{p}_0 - \boldsymbol{p}) d\boldsymbol{p}\right)^{-1} \int\int \boldsymbol{\Gamma} g(\boldsymbol{p}) G_\sigma(\boldsymbol{p}_0 - \boldsymbol{p}) \boldsymbol{p} d\boldsymbol{p}$. With (3) this can be written as $\hat{\boldsymbol{p}}(\sigma) = (\overline{\boldsymbol{\Gamma}_\sigma g})^{-1} \overline{\boldsymbol{\Gamma}_\sigma g \cdot \boldsymbol{p}}$. With the estimated variance factor $\hat{\sigma}_0^2 = S(\hat{\boldsymbol{p}}, \sigma)/(m-2)$, assuming m pixels being used, the covariance matrix of the estimated position is given by

$$\boldsymbol{C}_{\hat{p}\hat{p}}(\sigma) = \frac{S(\hat{\boldsymbol{p}}, \sigma)}{m-2} \cdot (\overline{\boldsymbol{\Gamma}_\sigma g})^{-1} . \tag{5}$$

The regularity measure $S(\hat{\boldsymbol{p}}, \sigma) = S(\boldsymbol{p}, \sigma) - \hat{\boldsymbol{\beta}}^T(\sigma)\, \overline{\boldsymbol{\Gamma}_\sigma g}\, \hat{\boldsymbol{\beta}}(\sigma)$ at $\hat{\boldsymbol{p}}$ is reduced by the bias $\hat{\boldsymbol{\beta}}(\sigma) = \hat{\boldsymbol{p}}(\sigma) - \boldsymbol{p}$ weighted with $\overline{\boldsymbol{\Gamma}_\sigma g}$. Thus the average squared gradient $\overline{\boldsymbol{\Gamma}_\sigma g}$ and the regularity measure S are needed for determining the covariance of the location of a junction point. A similar reasoning holds for circular symmetric features, e. g. isolated points, replacing S by $S^{\perp}$ and $\boldsymbol{\Gamma} = \nabla g \nabla^T g$ by $\boldsymbol{\Gamma}^{\perp} = \nabla^{\perp} g \nabla^{\perp T} g$. The weight of $\hat{\boldsymbol{p}}$, being proportional to the inverse of (5), equals $\overline{\boldsymbol{\Gamma}_\sigma g}$, motivating the weight $tr\boldsymbol{\Gamma} g = \|\nabla g\|^2$ in (4).

We are now prepared to discuss the four steps of feature extraction.

3.2 Step 1: Estimation of Noise Characteristics

The noise characteristics are decisive for thresholding. Thresholding always is performing a hypothesis test of some kind. Therefore, thresholds should depend

on the distribution of the test statistic and the significance level. Whereas the distribution of the test statistic can be derived from the noise characteristics, the significance level may be fixed for all tests and/or be related to some cost function making the choice of thresholds a transparent operation.

We have developed a blind estimation scheme for the case of a linear increase of the noise variance with the signal $\sigma_n^2 = a + bf$ (cf. [3]), as real images show a significant linear portion in the noise variance. It is based on the ability to separate signal and noise in the distribution of $g_x^2 + g_y^2$ (cf. also [25]). The procedure also gives quality measures for these estimates, revealing them to be quite accurate, (in general $< 10\%$) which is consistent with the simulation results in [21] and fully sufficient.

3.3 Step 2: Information Preserving Restoration

Restoration in general aims at recovering the original signal $\tilde{f}$ from g, undoing the effects of blur and noise. This requires the estimation of the noise characteristics (cf. step 1), of the possibly local scale and of the image smoothness. We first discuss how to automatically estimate the image smoothness, assuming the local scale to be known, estimated (cf. below) or negligible.

We want to apply an adaptive Wiener filter which exploits the local statistics of the image function referring to the image model. The homogeneity of the image function g is measured by the local variance of the gradient. As the gradient varies in direction, we use the diagonalization of $\overline{\boldsymbol{\Gamma} g}$ which can be written as $\overline{\boldsymbol{\Gamma} g} = \overline{\boldsymbol{\Gamma} f} + \sigma_n^2 \boldsymbol{I} = (\sigma_{f_1}^2 + \sigma_{n_1}^2)\boldsymbol{t}_1\boldsymbol{t}_1^T + (\sigma_{f_2}^2 + \sigma_{n_2}^2)\boldsymbol{t}_2\boldsymbol{t}_2^T$. The eigenvalues of $\overline{\boldsymbol{\Gamma} g}$ can be split into the variances $\sigma_{f_1}^2$ and $\sigma_{f_2}^2$ of the slopes in the principle directions 1 and 2 and the corresponding (equal) variances $\sigma_{n_1}^2$ and $\sigma_{n_2}^2$ of the first derivative of the noise, resp. allowing to estimate $\sigma_{f_i}^2, i = 1, 2$ from $\overline{\boldsymbol{\Gamma} g}$, in case σ_n^2 is known. We can now restore the image by optimizing the energy function which explicitely reflects the image model, especially $E(\nabla \hat{f}) = 0$:

$$E = \sum_{r,c} \left(\frac{g(r,c) - \hat{f}(r,c)}{\sigma_n(r,c)} \right)^2 + \sum_{r,c} \left[\left(\frac{\hat{f}_1(r,c)}{\sigma_{f_1}(r,c)} \right)^2 + \left(\frac{\hat{f}_2(r,c)}{\sigma_{f_2}(r,c)} \right)^2 \right] \qquad (6)$$

$\hat{f}_i(r,c), i = 1, 2$ are the derivatives of the restored image in the local principle directions 1 and 2. As *all* standard deviations can be estimated from the data *no tuning of parameters is necessary*. The procedure can be extended to second derivatives, thus a more general image model [26], and to more general surfaces [27].

The *selection* of a proper possibly tensor valued *scale* $\boldsymbol{\Sigma}_1$ can also be based on the local image statistics [6] [19]. Approximating the local power spectrum by a Gaussian, one can show $\boldsymbol{\Sigma}_1 = \sigma_g^2 \overline{\boldsymbol{\Gamma} g}^{-1}$ (cf. the remark above). Using a small integration scale for $\overline{\boldsymbol{\Gamma} g}$ and a large integration scale for σ_g^2, reduces bias and gathers the image contrast resp.

3.4 Step 3: Feature Detection

The classification of all image pixels into the three classes can be interpreted as feature detection in the sense that only the existence and approximate location of the features is of primary concern. The segmentation of the image area (cf. Fig. 1c) leads to a *ternary image* whose connected components give initial limitation of the search space and – in case of point- and line-regions – steer the location procedure. Moreover, a classification of the point- and line-regions can be performed.

Given the discrete version $\boldsymbol{L}$ of the average squared gradient $\overline{\boldsymbol{\Gamma}_\sigma g}$ and the regularity functions S and $S^\perp$. Since the covariance of an estimated point is $D(\hat{\boldsymbol{p}}) \propto S \cdot \boldsymbol{L}^{-1}$ or $D(\hat{\boldsymbol{p}}) \propto S^\perp \cdot \boldsymbol{L}^{\perp -1}$ the 2×2-matrix $\boldsymbol{L}$ and $\boldsymbol{L}^\perp$ also have to be checked for size and form using the eigenvalues λ_i common to $\boldsymbol{L}$ and $\boldsymbol{L}^\perp$.

Detecting Regions: The procedure starts with the classification of pixels into homogeneous and nonhomogeneous ones by investigating $L \doteq tr\boldsymbol{L} = L_{11} + L_{22} = \lambda_1 + \lambda_2$. If $L < T_h$, the pixel will be considered as homogeneous.

Approximating the χ^2_m-distribution of h by a normal distribution which holds for windows larger than 5×5 the local threshold T_h is given by $T_h(\boldsymbol{x}) = const. \cdot \sigma_n^2(\boldsymbol{x})\,(1 + z_{1-\alpha}/\sqrt{m})$ where $\sigma_n^2(\boldsymbol{x})$ is the local noise variance, *const.* a factor depending on the convolution kernels for determining $\boldsymbol{L}$, e. g. 3/8 when using the Sobel-operator, m the number of pixels involved during integrations with G_σ and $z_{1-\alpha}$ the α-quantile of the normal distribution. The threshold only holds for unfiltered images. Otherwise the reduction in noise variance has to be taken into account, which for a Gaussian with σ_1 leads to a reduction factor of $1/(4\pi\sigma_1^2)$.

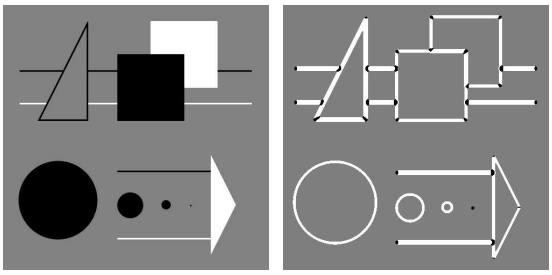
Classifying Point and Edge Regions: The second step is to investigate the form of $\boldsymbol{L}$ or $\boldsymbol{L}^\perp$. It is described by the ratio $v = \lambda_2/\lambda_1$. If $v = 1$, thus $\lambda_2 = \lambda_1$ then the form of $\boldsymbol{x}^T\boldsymbol{L}\boldsymbol{x}$ is circular indicating isotropy of the gradient caused by a corner, an isolated point or by isotropic texture. It easily can be derived from the form factor $q = 4det\boldsymbol{L}/tr^2\boldsymbol{L} = 4\lambda_1\lambda_2/(\lambda_1 + \lambda_2)^2$ via $v = \lambda_2/\lambda_1 = (1 - \sqrt{1-q})/(1 + \sqrt{1-q})$. As it is related to the curvature κ^2 of the isophote by $\kappa^2 = v/\sigma^2$ (cf. above), it is useful to distinguish points and smooth edges. The threshold T_κ should reflect the minimum curvature κ_{min} required for an image curve to contain a corner. T_κ can be made independent on the used scale σ. If $\|\kappa\| > T_\kappa = \kappa_{min}$, the pixel will be considered to be a candidate for a point pixel, otherwise for an edge pixel.

3.5 Step 4: Feature Location

The precise location of the point- and line-type features requires generic models which in principle allow to estimate the real-valued position of points and edges of the ideal image using the appropriate local scale of the feature. We use the two models for points, namely junctions and circular symmetric features, optimizing, i. e. minimizing, the regularity measures S (4) and $S^\perp$ from section 3.1.2. We may also use S for the location of edges, as they can be seen to be special junctions with two edges of the same orientation meeting.

Fig. 3. shows the ternary images (b, d; labels from left to right) indicating point and edge regions for a line entering an area (a). The local minima of S (bright) within the point, edge and segment regions are given in (c) and (e). Observe the unbiasedness of the optima. The integration scale is $\sigma = 4$ in (b, c). The centre of the *line* not the two edges is indicated in (c). In (d) and (e) with $\sigma = 0.7$ the two edges are detected, which are 3 pixels apart

Fig. 4. shows the test image from ROSENTHALER et al. with the ternary image. All junctions and corners have been detected. The integration scale is $\sigma = 2$.

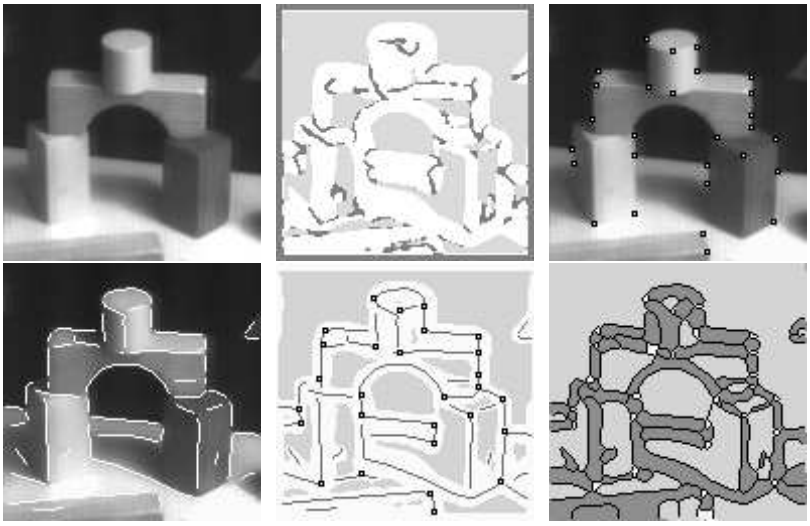


Locating Point-Type Features: There are two ways to estimate the location of point features:

- One-Step Procedure: First pixels are classified as junctions or circular features (cf. below). Pixels where S or $S^{\perp}$ shows a local minimum in a 8-neighborhood are considered to be point-type pixels. This yields approximate integer positions (r_0, c_0). A subpixel position (r, c) may be obtained by a parabolic fit to S or $S^{\perp}$, resp.
- Two-Step Procedure [8]: We first search for a relative maximum of the weight $1/tr(\boldsymbol{L}^{-1})$ in a 8-neighbourhood. *Independent on the classification* this yields the optimal window position, i. e. where the highest accuracy for the location is to be expected. After classification the subpixel position then is estimated yielding $S(\hat{\boldsymbol{p}}, \sigma)$. For circular features similar relations hold, yielding $\hat{\boldsymbol{p}}^{\perp}$ and $S^{\perp}(\hat{\boldsymbol{p}}^{\perp}, \sigma)$.

Classifying Junctions and Circular Symmetric Features: The classification of points (cf. Fig. 1c) can be based on the ratio $s_P = S/S^{\perp} \sim F_{m,m}$ [5], which again is Fisher distributed in case of white noise and orthogonal kernels for determining the derivatives $g_r(r, c)$ and $g_c(r, c)$. Two one-sided tests may be

Fig. 5. shows the recovery of the image structure with the proposed procedure (labels from left to right). The pixels of the original image (a) are classified into point, line, and segment regions (b), corresponding to the structure of the image model. The extracted and located points and edges are shown in (c) and (d) Including the blobs in (e), we obtain a partial reconstruction of the true image. The mutual relation between all feature types can be derived from the exoskeleton (f).



applied: 1. If $s_P < F_{m,m,\alpha}$ then $S \ll S^{\perp}$; the point is hypothesized to be a junction. 2. If $s_P > F_{m,n,1-\alpha}$ then $S \gg S^{\perp}$; the point is hypothesized to be an isolated point or, more generally, a circular symmetric feature. Otherwise no decision is made. It is meaningful to hypothesize a junction in this case. The neighborhood of $\boldsymbol{p}$ may still show regularities, e. g. being a spiral type feature [2], or may be pure texture.

Locating Edges: Similarily the location of edges may be performed in two ways:

- One-Step Procedure: Pixels where S shows a local minimum in the direction of the maximal gradient are considered as edge pixel. The direction of the gradient may be taken from the largest eigenvector of $\boldsymbol{L}$. The procedure is then similar to the one of CANNY [4].
- Two-Step Procedure: Again we first maximize $1/tr(\boldsymbol{L}^{-1})$ in the direction of the maximal gradient. This may be performed to subpixel accuracy already yielding acceptable results (cf. Fig. 5d). The integer position can then be used as the centre of a window within the optimal edge position is estimated finally.

Classifying Edge- and Line-type Features: The test whether a linear feature is a line or an edge (cf. Fig. 1c) can be based on a test comparing the

significance of the vectors $(\overline{g}_x, \overline{g}_y)$ and $(\overline{g}_{xx}, \overline{g}_{xy}, \overline{g}_{yy})$ of the smoothed first and second derivatives with $\overline{z} = G_\sigma * z$. The test statistic

$$s_L = \frac{\left[(\overline{g}_x/\sigma_{\overline{g}_x})^2 + (\overline{g}_y/\sigma_{\overline{g}_y})^2\right]/2}{\left[(\overline{g}_{xx}/\sigma_{\overline{g}_{xx}})^2 + (\overline{g}_{xy}/\sigma_{\overline{g}_{xy}})^2 + (\overline{g}_{yy}/\sigma_{\overline{g}_{yy}})^2\right]/3} \sim F_{2,3} \tag{7}$$

is used for a two-sided test: if $s_L < F_{2,3,\alpha}$ the feature is supposed to be a line, if $s_L > F_{2,3,1-\alpha}$ the feature is supposed to be an edge, otherwise it is assumed to be oriented texture. The kernels for the derivatives are supposed to be mutually orthogonal. The standard deviations are to be derived by error propagation.

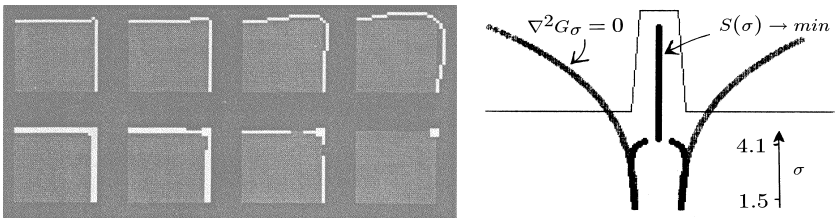
In both cases, the two-step-procedure gives satisfying results, as the locally optimal window for estimation is used. The one step procedure, however, is theoretically more transparent allowing to analyze the scale behaviour more easily.

The principle and an example for the techniques are shown in figures 3 to 5.

3.6 Scale Space Properties

The favorable scale space properties of this approach have been found be HEIKKILÄ [15]. Compared to zero crossings of $\nabla^2 G_{\sigma_1} g$ no bias occurs at corners. In addition less spurious effects occur when σ becomes large.

Fig. 6.



(a) shows effect of the scale for zero crossings of $\nabla^2 G_\sigma g$ at a corner compared with the minima of the regularity measure S. Points are indicated thick: No bias occurs (taken from HEIKKILÄ 1989). (b) shows the scale behaviour of the zero crossings of $\nabla^2 G_\sigma g$ and of the regularity measure S for a 1D-bar edge. The scale dependent bias (cf. BERZINS 1984) of the zero crossings is clearly visible. The bias of the feature location with S is below 10 % of the width of the bar edge, here with a width of $2a = 6$.

The interference of multiple edges can be studied analytically. We only give the result of a cross-section through a line, i. e. a bar-edge of width 2a. Assuming the squared gradient $g_x^2(x) = \delta(x+a)+\delta(x-a)$ convolution with $x^2 G_\sigma(x)$, the one dimensional version of 4, yields $S(x,\sigma) = (x+a)^2 G_\sigma(x+a) + (x-a)^2 G_\sigma(x-a)$.

At larger scales the procedure selects dark or light lines instead of edges (cf. Fig. 3). The bifurcation takes place at that scale where the second derivative $S(x, \sigma)$ vanishes at $x = 0$. This is at $\sigma = 2a/(5-\sqrt{15}) \approx 1.5a$ (cf. Fig. 6). The maximum bias is approximately at scale $\sigma = 1.2a$ and has a size of approximately 10% of the width of the bar. At scales below $\sigma = a/2$ the bias is negligible.

4 Summary

We have sketched a framework for an integrated approach to low-level feature extraction. It covers all steps from the raw digital image to an initial symbolic image description. The main scope was to provide statistically sound measures for steering the individual steps and for evaluating the results. All analysis steps are explicitely motivated by the chosen image model. The feature extraction allows an efficient implementation and shows favorable scale space properties.

The techniques reveal fruitful links to existing ones suggesting to exploit these relations for further extension of the concept. This especially holds for transferring the techniques to range images or truely 3D/volumetric images, which can be achieved by using the 3D-version of the smoothness measures [11]. The extension to multispectral/multichannel is straight forward by weighting the channel information according to the individual noise characteristics [3]. The extension towards texture analysis may use the quadratic filter approaches [18] to an advantage. As the goal was not to explicitely include high-level or scene knowledge, the integration with interpretation modules having top-down queries to the low-level processing needs to be investigated.

Acknowledgements:
I thank Udo Tempelmann for the programming, Claudia Fuchs for providing the examples, Maggie Kugel for TEXing the manuscript and Uwe Weidner for the layout of the paper. This work has been supported by the Deutsche Forschungsgemeinschaft.

References

1. V. Berzins. Accuracy of Laplacian Edge Detectors. *CVGIP*, 27:185–210, 1984.
2. J. Bigün. A Structure Feature for Some Image Processing Applications Based on Spiral Functions. *CVGIP*, 51:166–194, 1990.
3. R. Brügelmann and W. Förstner. Noise Estimation for Color Edge Extraction. In W. Förstner and S. Winter, editors, *Robust Computer Vision*, 90–107. Wichmann, Karlsruhe, 1992.
4. J. Canny. A Computational Approach to Edge Detection. *IEEE T-PAMI*, 8(6):679–698, 1986.
5. W. Förstner. *Statistische Verfahren für die automatische Bildanalyse und ihre Bewertung bei der Objekterkennung und -vermessung*, Volume 370 of Series *C*. Deutsche Geodätische Kommission, München, 1991.
6. W. Förstner. Determination of Local Scale in an Image. Technical report, Institut für Photogrammetrie, Bonn, 1993.

7. W. Förstner. Feature Extraction in Digital Photogrammetry. *Photogrammetric Record*, 14(82):595–611, 1993.
8. W. Förstner and E. Gülch. A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Circular Features. In *Proceedings of the Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken*, 281–305, 1987.
9. C. Fuchs, T. Löcherbach, H.-P. Pan, and W. Förstner. Land Use Mapping from Remotely Sensed Images. In *Colloquim on Advances in Urban Spatial Information and Analysis, Wuhan*, 1993.
10. G. H. Granlund. In Search for a General Picture Processing Operator. *CVIP*, 8:155–178, 1978.
11. G. H. Granlund. Image Sequence Analysis. In S. J. Pöppl and H. Handels, editors, *Mustererkennung*, 1–18. DAGM, 1993.
12. U. Grenander. Advances in Pattern Theory. *The Annals of Statistic*, 17:1–30, 1989.
13. R. M. Haralick and L. G. Shapiro. *Robot and Computer Vision*. Addison-Wesley, 1992.
14. D. Heeger. Optical Flow from Spatiotemporal Filters. In *Proceedings of 1st ICCV*, 181–190, 1987.
15. J. Heikkilä. Multiscale Representation with Förstner Operator. *The Photogrammetric Journal of Finland*, 11(2):40–59, 1989.
16. M. Kass and A. Witkin. Analyzing Oriented Patterns. *CVGIP*, 37:362–385, 1987.
17. H. Knutson and G. H. Granlund. Texture Analysis Using Two-Dimensional Quadrature Filters. In *Workshop Computer Architecture for Pattern Analysis and Image Data Base Management, Pasadena*, 1983.
18. T. S. Lee, D. Mumford, and A. Yuille. Texture Segmentation by Minimizing Vector-Valued Energy Functionals: The Coupled Membrane Model. In *Computer Vision – ECCV '92, Proceedings*, 165–173. Springer, 1992.
19. T. Lindeberg and J. Gårding. Direct Computation of Shape Cues by Multi-Scale Retinoptic Processing. Technical Report 117, Computational Vision and Active Perception Laboratory, Stockholm University, 1993.
20. D. E. McClure. Image Models in Pattern Theory. In A. Rosenfeld, editor, *Image Modelling*, 259–275. Academic Press Inc. , Orlando Florida, 1980/81.
21. P. Meer, J. Jolion, and A. Rosenfeld. A Fast Parallel Algorithm for Blind Estimation of Noise Variance. *IEEE T-PAMI*, 12(2):216–223, 1990.
22. A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. Electrical Engineering. McGraw-Hill, 2 edition, 1984.
23. L. Rosenthaler, F. Heitger, O. Kübler, and R. von der Heydt. Detection of General Edges and Keypoints. In *Computer Vision – ECCV '92, Proceedings*, 78–86. Springer, 1992.
24. V. Torre and T. A. Poggio. On Edge Detection. *IEEE T-PAMI*, 8(2):147–163, 1986.
25. H. Vorhees and T. Poggio. Detecting Blobs as Textons in Natural Images. In *Image Understanding Workshop, LA, Proceedings*, 1987.
26. U. Weidner. Informationserhaltende Filterung und ihre Bewertung. In B. Radig, editor, *Mustererkennung, Proceedings*, 193–201. DAGM, Springer, 1991.
27. U. Weidner. Parameterfree Information-Preserving Surface Restoration. In *Computer Vision – ECCV '94, Proceedings*, 1994.

This article was processed using the LaTeX macro package with LLNCS style