On the Noise and Scale Behaviour of Relational Descriptions

Claudia Fuchs, Felicitas Lang, Wolfgang Förstner
Institut für Photogrammetrie, Universität Bonn, Nußallee 15
e-mail: Claudia.Fuchs, Felicitas.Lang, and Wolfgang.Foerstner@ipb.uni-bonn.de
ISPRS Commission III Symposium '94, WG III/2

KEY WORDS: Relational description, quality evaluation, scale space, feature extraction, noise sensitivity

ABSTRACT:

The paper presents a concept for analysing the quality of relational descriptions of digital images. The investigations are based on the relational description automatically derived by a new coherent procedure for feature extraction providing a feature adjacency graph containing points, edges and segments and their relations. A new notion of scale (integration scale) is introduced, relating to a non linear function of the image, providing new stable descriptions. Based on the feature extraction we analysed the quality of the relational descriptions in dependency on the signal-to-noise ratio and on the control parameters of the feature extraction process, i. e. the significance level, the smoothing scale and the integration scale. First results on the quality of the features, focussing on their existence, distinct attributes and relations are given. The scope of this research is to predict the quality, especially probabilities of components of the relational description from a few measures depending on noise, scale and local properties of the image content. This is motivated by the applications we are dealing with, namely extracting man-made 2D or 3D structures by grouping procedures or image matching as both tasks are optimization problems where the probability of the unknown 2D or 3D model has to be maximized.

1 Motivation

Images may be represented at different levels of abstraction, relational descriptions being the most general ones. Therefore deriving relational image descriptions is a fundamental step in automatic scene interpretation. They not only contain features or primitives and their corresponding attributes, but also their mutual topological, geometrical and possibly semantic relations and their attributes. There are many applications relying on relational descriptions such as image matching (cf. Christmas *et al.* 1994), object or sensor orientation (cf. Vosselman 1992) or object reconstruction and recognition.

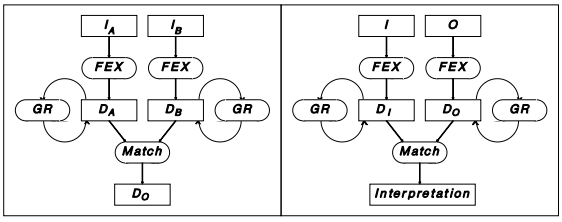

Figure 1: Relational description within image analysis tasks.

For two classical tasks the information flow is shown in Fig. 1. Given two images I_A and I_B, feature extraction (FEX) leads to the relational descriptions D_A and D_B. These may be immediately used for relational matching, e. g. in order to derive a 3D-description D_O of the scene. Perceptual grouping (GR) may be applied to the initial relational description (cf. Herault and Horaud 1992, Treutler 1992) in order to obtain a relational description which leads to better results in the matching stage. This may be due to the aggregation of features to higher level features or to the increase of consistency within the description. The same situation occurs when matching symbolic image description D_I to symbolic object description D_O in object location or interpretation problems (cf. Fig. 1.b). The relational description here again may be improved by grouping of features. In contrast to image matching where grouping is performed data driven, thus bottom-up, in order to use only very generic scene knowledge in object location or recognition problems, the grouping may be performed partly model driven, thus top-down. The motivation for the paper results from the view reconstruction and matching to be optimization problems where the probability $P(M \mid D, F)$ of the unknown model M out of several M_i of the object is maximized for given data D and the parameters used in the feature extraction process F (cf. Grimson and Huttenlocher 1990).
The models M_i may be anything specified by a set of parameters or a structure with attributed primitives and relations. We assume the model to be specified such that the true image structure can be predicted under the assumption that the feature extraction would be perfect.

From the joint probability $P(M_i, D, F)$ we obtain

$$P(M_i \mid D, F) = \frac{P(D \mid M_i, F) \cdot P(M_i, F)}{P(D, F)} \quad . \qquad (1)$$

This requires the knowledge of

- the probability a certain model M_i occurs together with the selected feature extraction process F,
- the conditional probability of a description D occurring for given M_i and F.

Maximizing (1) over i leads to an optimal selection/estimation $\hat{M}_i$ of M_i.

The feature extraction process F which is explicitely shown in (1) obviously influences the result. Splitting $P(M_i, F) = P(F \mid M_i) \cdot P(M_i)$ would suggest F to be dependent on the hypothesized model which e. g. could include a *model* dependent adaption of parameters within the feature extraction. We will not discuss this further and will assume F not to depend on M_i.
In case of two images to be matched, the goal would be to describe the mutual relation by some mapping function or to reconstruct the object based on two descriptions D_A and D_B. Omitting the dependency on F eq. (1) reads as $P(M_i \mid D_A, D_B) = P(D_A, D_B \mid M_i) \cdot P(M_i) \,/\, P(D_A, D_B)$. If we split the conditional probability $P(D_A, D_B \mid M_i) = P(D_A \mid D_B, M_i) \cdot P(D_B \mid M_i)$ we obtain $P(M_i \mid D_A, D_B) = P(D_A \mid D_B, M_i) \cdot P(D_B \mid M_i) \cdot P(M_i)/P(D_A, D_B)$ which, *in case* for given M_i the description D_A is not more or less likely if we know D_B, thus

$$P(D_A \mid D_B, M_i) = P(D_A \mid M_i) \tag{2}$$

reduces to

$$P(M_i \mid D_A, D_B) = \frac{P(D_A \mid M_i) \cdot P(D_B \mid M_i) \cdot P(M_i)}{P(D_A, D_B)} \quad . \tag{3}$$

Comparing (1) with (3) reveals that for both tasks, reconstruction and matching, we need the conditional probabilities $P(D \mid M_i)$. These explicitely describe the quality of the description D. The scope of the analysis is to predict

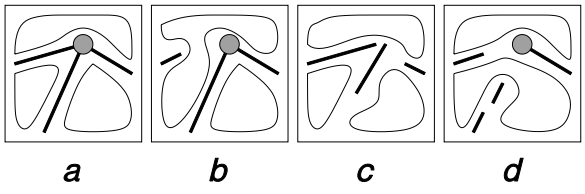

Figure 2: shows possible results of the feature extraction of a trihedral corner.

the probability of the relations between features from a few measures, which depend on the control parameters of the feature extraction process and on the local property of the image. E. g. a trihedral corner consists of 7 features with a great number of mutual relations and itself is described by two angles, two intensity contrasts and a noise variance making a full analysis intractable.
Fig. 2 shows several possibilities a corner may appear after feature extraction which, however, may be described by a small number of local distortions: the point may be detected or not, neighbouring segments may be merged, the edge may be partitioned into several parts. Other defects are spurious features which may occur due to noise effects.

This paper presents a first investigation of the quality of relational descriptions derived from digital images. We analyse the effect of noise and scale on the description of local image patterns. The analysis refers to our procedure for extracting features and their relations (Fuchs *et al.* 1993, Förstner 1994) and partly to the application we are dealing with, namely extracting man-made 2D or 3D structures from images.

2 Basic Tools

2.1 The Relational Description

The symbolic description of an image needs to represent the essential parts of an image related to some task. The quality of such a description has to refer to the ideal image, which requires to setup an appropriate image model. This image model itself is described using a symbolic representation. We assume the following model for the *ideal* image to hold:

1. The image is partitioned into segments S_j. Within the segments the intensity function or any other locally determinable function of the intensity is assumed to be piecewise smooth.
2. The boundaries of the segments are assumed to be piecewise smooth.

Thus, the ideal image $\mathcal{I}$ may be represented as the union of three sets of features, namely points $\mathcal{P} = \{\mathsf{P}_j\}$, linear features $\mathcal{L} = \{\mathsf{L}_j\}$ and segments $\mathcal{S} = \{\mathsf{S}_j\}$, with $\mathcal{I} = \mathcal{F} = \mathcal{P} \cup \mathcal{L} \cup \mathcal{S}$. Each feature may have attributes. In our context, we actually only use the following attributes:

$$\mathsf{P}_j \;:\; (x_j, y_j) \tag{4}$$
$$\mathsf{L}_j \;:\; \{(x_j, y_j), l(\mathsf{L}_j)\} \tag{5}$$
$$\mathsf{S}_j \;:\; \{(x_j, y_j) \mid (x_j, y_j) \epsilon \mathsf{S}_j\} \tag{6}$$

which we mentioned here for completeness.
There are many relations between these features which might be useful for image analysis. Here we consider the following relations:

$$\textsc{inc}\,(\mathsf{P}, \mathsf{L}) = \textsc{incidence}\,(\mathsf{P}, \mathsf{L}) \tag{7}$$
$$\textsc{sep}\,(\mathsf{S}_{j_1}, \mathsf{L}, \mathsf{S}_{j_2}) = \textsc{separates}\,(\mathsf{S}_{j_1}, \mathsf{L}, \mathsf{S}_{j_2}) \tag{8}$$
$$\textsc{nei}\,(\mathsf{P}, \mathsf{L}) = \textsc{neighbour}\,(\mathsf{P}, \mathsf{L}) \tag{9}$$
$$\textsc{nei}\,(\mathsf{P}, \mathsf{S}) = \textsc{neighbour}\,(\mathsf{P}, \mathsf{S}) \tag{10}$$

More complex local patterns may be described using these relations. The geometric attributes are reduced to those which are necessary to establish the four relations mentioned above. Therefore no further quality parameters, such as standard deviations, are analysed in the following. The estimated image also symbolically is described using the three feature types $\hat{\mathsf{P}}_i, \hat{\mathsf{L}}_i, \hat{\mathsf{S}}_i$ denoted with a hat, indicating 'estimated'. The feature extraction explicitely provides the attributes of the estimated features, especially the length $l(\hat{\mathsf{L}}_i)$, and the relations $\textsc{inc}(\hat{\mathsf{P}}, \hat{\mathsf{L}})$, $\textsc{sep}(\hat{\mathsf{S}}_{i_1}, \hat{\mathsf{L}}, \hat{\mathsf{S}}_{i_2})$ $\textsc{nei}(\hat{\mathsf{P}}, \hat{\mathsf{L}})$ and $\textsc{nei}(\hat{\mathsf{P}}, \hat{\mathsf{S}})$ using their exoskeleton.

We thus want to compare the two descriptions

$$D = \{\{\mathsf{P}_j\}, \{\mathsf{L}_j\}, \{\mathsf{S}_j\}; \{\textsc{inc}(\mathsf{P}, \mathsf{L})\}, \{\textsc{sep}(\mathsf{S}_{j_1}, \mathsf{L}, \mathsf{S}_{j_2})\}, \{\textsc{nei}(\mathsf{P}, \mathsf{L})\}, \{\textsc{nei}(\mathsf{P}, \mathsf{S})\}\} \tag{11}$$

of the given image and

$$\hat{D} = \{\{\hat{\mathsf{P}}_i\}, \{\hat{\mathsf{L}}_i\}, \{\hat{\mathsf{S}}_i\}; \{\textsc{inc}(\hat{\mathsf{P}}, \hat{\mathsf{L}})\}, \{\textsc{sep}(\hat{\mathsf{S}}_{i_1}, \hat{\mathsf{L}}, \hat{\mathsf{S}}_{i_2})\}, \{\textsc{nei}(\hat{\mathsf{P}}, \hat{\mathsf{L}})\}, \{\textsc{nei}(\hat{\mathsf{P}}, \hat{\mathsf{S}})\}\} \tag{12}$$

of the noisy image for different test patterns and parameters of the feature extraction.

The main scope of the analysis presented in this paper is to show how the quality, i. e. the probability of the mentioned relations are derivable from the probability of detection and the probability distribution of the attributes, especially the length $l(\mathsf{L}_j)$ of the linear features.

2.2 Feature Extraction

Feature extraction aims at replacing the image by a suitable representation here a symbolic one. Our feature extraction scheme aims at a simultaneous and coherent treatment of all features, i. e. points, edges and segments, and at reducing the number of control parameters to an absolute minimum in order to simplify the quality prediction of the feature extraction process.

The feature extraction is performed in three steps:

1. The image pixels are *classified* into three classes leading to segment, line and point regions. The classification is based on a measure for homogeneity

$$h = tr\overline{\Gamma_t g} \tag{13}$$

and a measure for isotropy

$$v = \sqrt{\lambda_2(\overline{\Gamma_t g})/\lambda_1(\overline{\Gamma_t g})} \quad . \tag{14}$$

Both use the average squared gradient

$$\overline{\Gamma_t g} = G_t * (\nabla_s g \cdot \nabla_s g^T) \tag{15}$$

where

$$\nabla_s g = (\frac{\partial}{\partial_x} G_s \ \frac{\partial}{\partial_y} G_s)^T * g = (g_x \ g_y)^T \tag{16}$$

where G_s and G_t indicate 2D-Gaussians with isotropic scales s and t.

Thus we use two scales s and t:

   (a) The *smoothing scale* s in (16) reduces the effect of noise in the image when calculating the gradient.

   (b) The *integration scale* t in (15) is used to take the full gradient information within the window G_t into account. Integrating the squared gradient reduces the bias at *lines*, moreover is able to detect the centre of symmetric lines, a property which we, however, do not further analyse in this paper.

With h we obtain homogeneous segments where $h < T_h$. The threshold is derived from the automatically estimated noise variance (Brügelmann and Förstner 1992) and a prespecified significance level α (cf. sect. 4.1). Within the nonhomogeneous regions we use $v > v_{\min}$ for classifying point and line regions. The threshold $v_{\min}$ can be chosen quite freely between 0.1 and 0.5 .

2. Within the point and line regions the *location* of the points $\hat{\mathsf{P}}_i$ and of the edges $\hat{\mathsf{L}}_i$ are obtained by non-optima suppression: edge pixels are defined by relative maxima of h across the edge. Points are defined by the relative minima of the trace of the estimated covariance matrix when locating them within a window

$$trC_{\hat{p},\hat{p}} = S \cdot tr(\overline{\Gamma_{t_p} g})^{-1} \tag{17}$$

with

$$S = g_x^2 * x^2 G_{t_p} + 2 g_x g_y * xy G_{t_p} + g_y^2 * y^2 G_{t_p} \tag{18}$$

being the estimated sum of the squared residuals of a least squares fit of all edge elements within the window G_{t_p}. Experiments suggest to use a larger integration scale t_p for locating the points than t_l for locating the edges.

3. The *relations* between the three feature types now can be derived from their exoskeleton. This skeleton lies between the (thin) points, the (thin) edges and the segments, possibly requiring to first get a free region of one pixel width around the points or edges using a dilation.

Summarizing the feature extraction is controlled by the following parameters: 1. the smoothing scale s, 2. the integration scale t_l for lines, 3. the integration scale t_p for points, 4. the significance level α, 5. the noise variance σ_n, and 6. the threshold $v_{\min}$ for the isotropy v. We will investigate the effect of the first four parameters on the feature extraction. The noise variance σ_n is automatically estimated from the data while the threshold $v_{\min}$ is fixed at 0.1 .

3 Principle of Analysis

The goal of the analysis is to predict the quality of the feature extraction, i. e. the probability that certain features are detected and have certain mutual relations. As the complexity of real structures is high we want to break down the complete probability into parts which are simple enough to be analysed and sufficient for measuring the quality of the complete description.
E. g. the possible distortions of a trihedral corner as first mentioned in sec. 1 (cf. Fig. 2) may be analysed locally by investigating the set of images shown in Fig. 3 .

Figure 3: From left to right: a blank image, an edge with contrast $\Delta g = 32$ [gr], corner images with contrast $\Delta g = 32$ [gr] and varying angle $\varphi = 30°, 60°, 90°, 120°, 150°$.

We use this example to show the line of our approach of analysing complex patterns. In order to arrive at realistic estimates for the quality values we run sufficient trials of noisy images, between 10 and 50, fixing all parameters. The noise is white and Gaussian, all experiments are performed with floating-point images.
The quality depends on the signal-to-noise ratio defined as

$$SNR = \frac{\Delta g}{\sigma_n} \tag{19}$$

where Δg is the contrast of the edge and σ_n is the noise standard deviation. We vary σ_n while keeping $\Delta g = 32$ [gr] fixed using $SNR = 0.7$ to $SNR = 8.0$ in steps of $\sqrt[4]{2}$.

We distinguish three types of measures:

1. Probabilities give direct access to the desired quality measures for matching.
2. The density of spurious features give indications about the proper choice of control parameters.
3. The quality of edge extraction with respect to grouping also gives indications on the proper choice of control parameters.

3.1 Probabilities

We determine three measures which will give us explicit access to probabilities of relations.

1. The probability
$$P(\mathsf{P}\,\text{recovered}) = P(\hat{\mathsf{P}} \mid \mathsf{P}) \tag{20}$$
of a given point P to be found.
This probability will depend on the angle of the corner. One can expect the system to have difficulties with the obtuse angle $\varphi = 150°$.
2. The probability
$$P(\hat{\mathsf{S}}\,\text{merged}) = P\left[\hat{\mathsf{S}}_r(\hat{\mathsf{L}}) \equiv \hat{\mathsf{S}}_l(\hat{\mathsf{L}})\right] \tag{21}$$
of two segments $\hat{\mathsf{S}}_r$ and $\hat{\mathsf{S}}_l$ on the right and left side of an extracted edge $\hat{\mathsf{L}}$ to be identical, in case of $\text{NEI}(\mathsf{S}_r, \mathsf{L}, \mathsf{S}_l)$ is needed to evaluate the decision to split the segment $\mathsf{S} = \mathsf{S}_r = \mathsf{S}_l$ in the prolongation of L (cf. Fig. 2.b and 2.d).
3. The coverage
$$c(\mathsf{L}) = \frac{\Sigma_i l(\hat{\mathsf{L}}_i \mid \mathsf{L})}{l(\mathsf{L})} \tag{22}$$
of an edge L by the extracted edges $\hat{\mathsf{L}}_i$ will serve us for predicting the likelihood of $\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}})$ (cf. sect. 5.1).

3.2 Spurious Features

Spurious features may disturb the image analysis. We use the following values to describe the property of the feature extraction in generating spurious features.

1. The density
$$d(\hat{\mathsf{P}}\,\text{spurious}) = \#(\hat{\mathsf{P}} \mid \neg\mathsf{P})/a\ [\#/128^2\,\text{pel}] \tag{23}$$
based on the number $\#\hat{\mathsf{P}}$ of extracted points $\hat{\mathsf{P}}$ which are not matched to a given point. As the number will increase with the area of the image $\mathcal{I}$ (strictly with the area of $\mathcal{L} \cup \mathcal{S}$) we refer this number to the image area. The unit of a is chosen to be $[128^2\,\text{pel}]$.
2. The density
$$d(\hat{\mathsf{L}}\,\text{spurious}) = \#(\hat{\mathsf{L}} \mid \neg\mathsf{L})/a\ [\#/128^2\,\text{pel}] \tag{24}$$
similarily measures the number of spurious edges $\hat{\mathsf{L}}$ with respect to the area a.
3. The average length
$$l(\hat{\mathsf{L}}\,\text{spurious}) = l(\hat{\mathsf{L}} \mid \neg\hat{\mathsf{L}})\ [pel] \tag{25}$$
will indicate whether short edges are allowed to be eliminated.

The three measures $d(\hat{\mathsf{P}}\,\text{spurious}), d(\hat{\mathsf{L}}\,\text{spurious})$ and $l(\hat{\mathsf{L}}\,\text{spurious})$ will be analysed on a blank image.

3.3 Quality of Edge Extraction

As we are interested primarily in gross disturbancies, we describe the quality of edge extraction by the following measures.

1. The density
$$d(\hat{\mathsf{P}}\,\text{on}\,\mathsf{L}) = \#(\hat{\mathsf{P}} \mid \mathsf{L})/l\ [\#/128\,\text{pel}] \tag{26}$$
is the number $\#\hat{\mathsf{P}}$ of extracted points $\hat{\mathsf{P}}$ erroneously sitting on an edge L with respect to a reference length l, here of [128 pel]. This gives the density of points sitting on a smooth (infinitely long) line.
2. The density
$$d(\mathsf{L}\,\text{partitioned}) = \#(\hat{\mathsf{L}} \mid \mathsf{L})/l\ [\#/128\,\text{pel}] \tag{27}$$
is the number $\#\hat{\mathsf{L}}$ of edges $\hat{\mathsf{L}}$ which are sitting on a given edge L with respect to a reference length l, here 128 pel. This gives the degree of partitioning of an (infinitely long) edge, or the edge segment density.
3. The average length
$$l(\hat{\mathsf{L}}\,\text{matched}) = l(\hat{\mathsf{L}} \mid \mathsf{L})\ [\text{pel}] \quad . \tag{28}$$
Observe that the $c(\mathsf{L}) = d(\mathsf{L}\,\text{partitioned}) \cdot l(\hat{\mathsf{L}}\,\text{matched})/l(\mathsf{L})$.

3.4 Incidence Relations $\text{INC}(\mathsf{P}, \mathsf{L})$, $\text{INC}(\mathsf{P}, \mathsf{S})$

For grouping the incidence relations $\text{INC}(\mathsf{P}, \mathsf{L})$ and $\text{INC}(\mathsf{P}, \mathsf{S})$ are of direct importance. We therefore determine $P(\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}}))$ and $P(\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{S}}))$ in dependence on the image structure, namely the angle φ of the corner.

We will see in sect. 5.1 that $P(\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}}))$ can be derived from $c(\mathsf{L})$ and $P(\hat{\mathsf{P}} \mid \mathsf{P})$.

3.5 The Transition Matrix

For analysing the relation between given and estimated features we build up a transition table $T = (t_{ij})$ indicating whether given and estimated features meet. For determining the relation $\text{MEET}(\hat{\mathsf{F}}_i, \mathsf{F}_j)$ we use the exoskeleton leading to regions $\mathcal{R}(\mathsf{F}_j)$ and $\mathcal{R}(\hat{\mathsf{F}}_i)$ around each feature and determine t_{ij} by
$$t_{ij} = \begin{cases} 0 & \text{if } \hat{\mathsf{F}}_i \cap \mathcal{R}(\mathsf{F}_j) = \emptyset \text{ and } \mathsf{F}_j \cap \mathcal{R}(\hat{\mathsf{F}}_i) = \emptyset \\ 1 & \end{cases} \tag{29}$$

Thus, $t_{ij} = 1$ if a given feature is close to an estimated. The distance threshold hereby is defined by the skeleton. $t_{ij} = 0$ definitely excludes any type of closeness.

The sums

$$p_j = \sum_i t_{ij} \quad \text{and} \tag{30}$$

$$m_i = \sum_j t_{ij} \tag{31}$$

have a very definite meaning:

1. p_j measures the *degree of partitioning*

   $p_j \geq 2$: a given feature F_j is partitioned into p_i features
   $p_j = 1$: a feature F_j is not partitioned
   $p_j = 0$: a feature F_j is lost.

2. m_i measures the *degree of merging*

   $m_i \geq 2$: m_i given features are merged into one estimated $\hat{\mathsf{F}}_i$
   $m_i = 1$: the estimated $\hat{\mathsf{F}}_i$ feature is not a merging of several given features
   $m_i = 0$: the estimated feature $\hat{\mathsf{F}}_i$ is spurious.

	P_a	m(P)	L_b	L_c	m(L)	S_d	S_e	m(S)	m(F)
$\hat{\mathsf{P}}_1$	0	**0**	1	0	**1**	0	0	**0**	**1**
$p(\hat{\mathsf{P}})$	**0**		**1**	**0**		**0**	**0**		
$\hat{\mathsf{L}}_2$	1	**1**	0	1	**1**	0	0	**0**	**2**
$\hat{\mathsf{L}}_3$	0	**0**	1	0	**1**	0	0	**0**	**1**
$\hat{\mathsf{L}}_4$	0	**0**	0	1	**1**	0	0	**0**	**1**
$\hat{\mathsf{L}}_5$	0	**0**	1	0	**1**	0	0	**0**	**1**
$\hat{\mathsf{L}}_6$	0	**0**	1	0	**1**	0	0	**0**	**1**
$\hat{\mathsf{L}}_7$	0	**0**	0	1	**1**	0	0	**0**	**1**
$\hat{\mathsf{L}}_8$	0	**0**	0	1	**1**	0	0	**0**	**1**
$p(\hat{\mathsf{L}})$	**1**		**3**	**4**		**0**	**0**		
$\hat{\mathsf{S}}_9$	0	**0**	0	0	**0**	1	1	**2**	**2**
$p(\hat{\mathsf{S}})$	**0**		**0**	**0**		**1**	**1**		
$p(\hat{\mathsf{F}})$	**1**		**4**	**4**		**1**	**1**		

Table 1: Transition table for the example shown in Fig. 4.

An example of such a transition matrix for F_j and $\hat{\mathsf{F}}_i$ shown in Fig. 4 is given on Table 1. The degrees p_j and m_i for partitioning and merging are given for each individual feature type and all features. The given point P_a was lost, the estimated point $\hat{\mathsf{P}}_1$ is spurious among the set of points. The edges L_b (right) and L_c (left) both are splitted. The two segments S_d and S_e have been merged into one estimated segment $\hat{\mathsf{S}}_9$. The off diagonal parts of the table indicate transitions from one feature type into another, e. g. the point P_a has been 'changed' into the edge $\hat{\mathsf{L}}_2$.

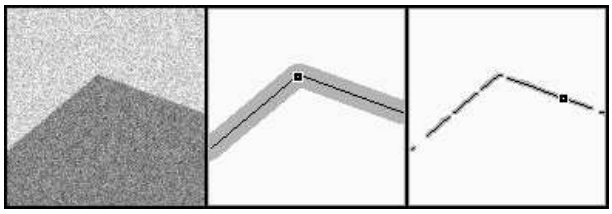

Figure 4: The 120° corner image with noise $\sigma_n^2 = 75\,[\mathrm{gr}^2]$, the 'ideal' features F and the 'extracted' features $\hat{\mathsf{F}}$ derived by feature extraction. The white parts correspond to the segments, the black pixel chains correspond to the linear features, points are indicated by black squares.

4 Results

4.1 Spurious Features

We first investigate the effect of the different control parameters on the generation of spurious features, namely the significance number α, the scales s and t_l in order to fix these values in further analysis.

Fixing the scales to $s = 1, t_l = 1.4$ and $t_p = 2$ and varying α we obtain the following results (cf. Table 2):

1) Increasing the significance level $S = 1 - \alpha$ decreases the number of spurious features as to be expected. The length of spurious edges is obviously constant. As the number of spurious features is already small for $\alpha = 0.05$, we will use this significance number throughout the following tests.

$\alpha =$	0.05	0.01	0.001	[unit]
$d(\hat{\mathsf{P}}$ spurious)	1.4	0.06	0	$[\#/128^2\mathrm{pel}]$
$d(\hat{\mathsf{L}}$ spurious)	1.4	0.04	0	$[\#/128^2\mathrm{pel}]$
$l(\hat{\mathsf{L}}$ spurious)	2.3	2.0	0	[pel]

Table 2: shows the dependency of spurious features on the significance number α.

2) The dependency of the length of the spurious edges on the smoothing scale s is shown in Fig. 5 . The integration scales are chosen to depend on s by $t_l = \sqrt{2}\cdot s$ and $t_p = 2\cdot s$. The length of the spurious edges obviously increases nearly proportional to s.

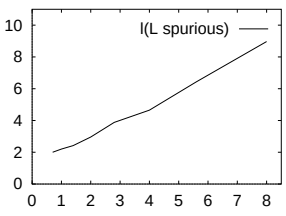

Figure 5: shows the average length of spurious edges $l(\hat{\mathsf{L}} \mid \neg\hat{\mathsf{L}})$ in dependency on s.

3) The dependency of spurious features on the integration

scale t_l is given in Table 3.

Obviously increasing t_l leads to a significant suppression of spurious features. As too large integration scales reduce the resolution of neighbouring edges, we will use the ratio $1 : \sqrt{2} : 2$ for $s : t_l : t_p$ in the following. As the smoothing scale s leads to a uniform reduction of noise, we fix it to $s = 1$ and will show in sect. 5 how to predict the performance for $s \neq 1$.

$t_l =$	1.0	1.4	2.0	2.8	4.0
$d(\hat{\mathsf{P}}$ spurious)	2.4	1.3	0.4	0	0
$d(\hat{\mathsf{L}}$ spurious)	9.7	1.4	0.2	0.1	0
$l(\hat{\mathsf{L}}$ spurious)	2.3	2.4	3.0	2.0	0

Table 3: shows the dependency of the spurious features on the integration scale t_l.

4.2 Edge Extraction

The quality of edge extraction depends on how easily the original edge can be recovered from the extracted edge segments. All the following results are based on tests on the edge image and on the five different corner images.

The average number $d(\mathsf{L}$ partitioned) (cf. (27)) of edge segments referring to a normalized reference length of 128 pel is shown in Fig. 6 (left) for all 6 images.

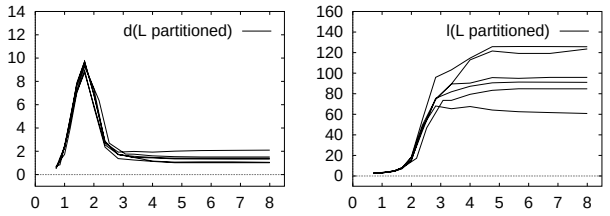

Figure 6: **Left:** density $d(\mathsf{L}$ partitioned) in dependency on the *SNR*. **Right:** average length $l(\hat{\mathsf{L}} \mid \mathsf{L})$ in dependency on the *SNR*.

The numbers are consistent for $SNR < 2.8$, thus for relatively noisy images. E. g. for $SNR = 1.8$ a line of 128 pixel length can be expected to be partitioned into approximately 9 segments. Their length can be expected to be about 8 pixels (cf. Fig. 6, right) leading to a coverage of approximately 60% (cf. Fig. 7, left). For values $SNR > 3.0$ the partitioning is not realistically estimable, as the used images are too small to contain edges of length $\gg 128$ pixels. Fig. 7 (left) however suggests that for $SNR > 4$, i. e. for $\sigma_n < \Delta g/4$, the full edge length can be expected to be recovered, independent of its length. This is supported by the probability $P(\hat{\mathsf{S}}$ merges) that two neighbouring segments which are separated by an edge are merged (cf. Fig. 8). For $SNR < 2$ this probability has been estimated to be close to 100%, while for $SNR > 3.5$ it is practically zero. Moreover, there is a slight tendency that increasing the integration scale t_l leads to a sharper distinction between good and bad: either edges are detected, then they are complete or they are not detected.

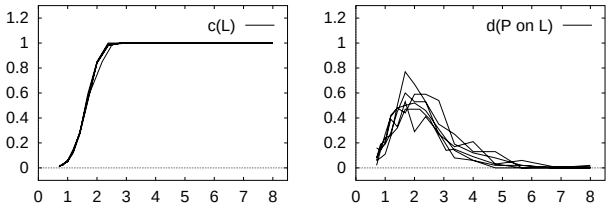

Figure 7: **Left:** coverage $c(\mathsf{L})$ in dependency on the *SNR* . **Right:** density $d(\hat{\mathsf{P}} on \mathsf{L})$ in dependency on the *SNR*.

However, even for $SNR > 4$ there is a small probability $P(\hat{\mathsf{P}}$ on $\mathsf{L})$ (cf. Fig. 7, right) that a spurious point $\hat{\mathsf{P}}$ partitions an edge. In this case, the point is linked to the two neighbouring edges with very high probability, as e. g. in Fig. 4 .

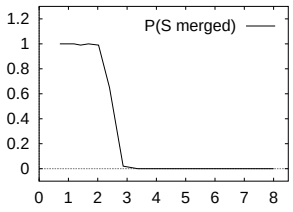

Figure 8: shows the probabilty $P(\hat{\mathsf{S}} merges)$ in dependency on the *SNR*.

4.3 Point Detection

The quality of detecting junction points depends on the angle of the adjacent edges. We therefore investigated the detectability of points and their relations to the adjacent edges and segments on corner images with varying angle φ ($30°, 60°, 90°, 120°$ and $150°$).

Fig. 9 shows the detectability $P(\hat{\mathsf{P}} \mid \mathsf{P})$ of points using the standard setup ($s = 1, t_l = \sqrt{2}, t_p = 2$) for the 5 corner images. In all cases except for $\varphi = 150°$ the detectability increases with *SNR*. For $\varphi = 30°$ and $120°$ the *SNR* needs to be larger for P to be detactable than for $\varphi = 60°$ and $90°$. For obtuse angles $\varphi = 150°$ the corner point is detected by chance.

In case the point actually is detected, we can estimate the probabilities that it is linked to its neighbouring edges or segments. The right column shows the average number of $\textsc{inc}(\hat{\mathsf{P}}, \hat{\mathsf{L}})$ which by division by 2 is identical to $P(\textsc{inc}(\hat{\mathsf{P}}, \hat{\mathsf{L}}))$. Obviously obtaining the relation $\textsc{inc}(\hat{\mathsf{P}}, \hat{\mathsf{L}})$ for $\varphi = 30°$ is not certain even for large *SNR*, while for $\varphi = 60°, 90°, 120°$ for large *SNR* one can expect to obtain both links.

The point-to-segment neighbourhood however can nearly never reliably be recovered. For $SNR > 3.5$ and acute angles $\varphi = 30°$ and $\varphi = 60°$ at least the link to the outer segment can be detected with high probability.

5 Predictions

We now want to show what type of predictions on attributes or relations of more complex patterns may be

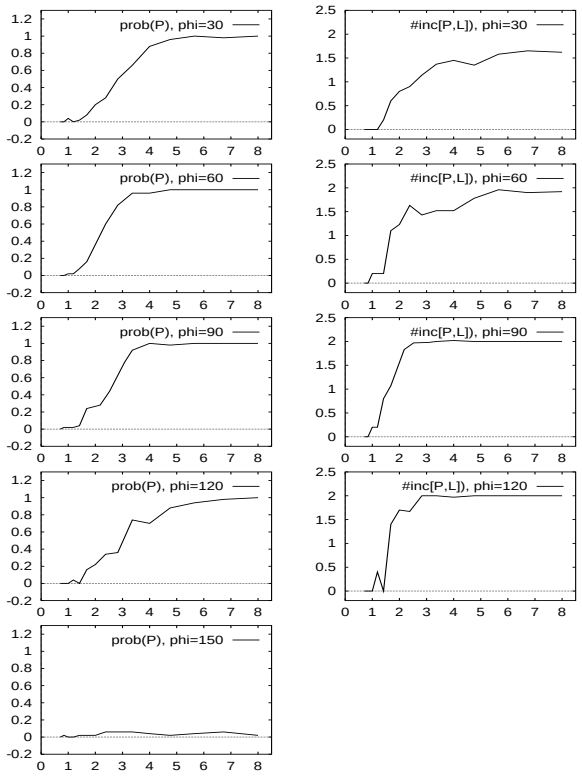

Figure 9: Dependency of a corner point on the angle φ. **Left column:** probabilty $P(\hat{\mathsf{P}} \mid \mathsf{P})$ in dependency on the *SNR*. **Right column:** average number of $\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}})$ in dependency on the *SNR*.

made based on the statistics of the occurence or values of simple attributes or relations.

5.1 Likelihood of $\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}})$ from $c(\mathsf{L})$

For identifying junctions it is important to know the probability a detected junction point $\hat{\mathsf{P}}$ is actually incident to its neighbouring edges $\hat{\mathsf{L}}_i$. E. g. there are four possibilities a corner may appear after feature extraction (cf. Fig. 10). Cases 2 and 3 will be equally likely.
Assuming $P(\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}}_{i_1})) = P(\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}}_{i_2}))$ we have with

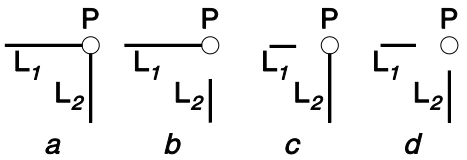

Figure 10: shows the four possibilities a corner can have relation INC to its neighbouring edges

$B[n,p]$ denoting the binomial distribution $\#\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}}) \sim B[2, P(\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}}))]$.
We now argue that the probability of $\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}})$ solely depends on the expected coverage $c(\mathsf{L})$ (cf. (22)). The probability of a spurious point $\hat{\mathsf{P}}$ falling into a partition of L not covered by some $\hat{\mathsf{L}}_i$ is assumed to be equal to the probability a point is not incident to an extracted edge $\hat{\mathsf{L}}_i$. Thus $P(\neg\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}})) = 1 - E(c(\mathsf{L}))$ or $P(\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}})) = E(c(\mathsf{L}))$. This yields the general expression for the probability of a corner $\hat{\mathsf{P}}$ being actually linked to $0, 1, 2$ of its 2 edges $\hat{\mathsf{L}}_i$

$$\#\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}}) \sim B[2, E(c(\mathsf{L}))] \quad . \tag{32}$$

The reasoning obviously is valid for angles between 60° and 120°. For acute angles the interference of the two edges reduces the likelihood of detecting both links, whereas for obtuse angles the number of detected points was not high enough to obtain reliable estimates.
E. g. if the coverage $c(\mathsf{L})$ of an edge is known to be 0.6 we find

$$P(\#\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}}) = 0) = 0.6^2 = 0.16 \tag{33}$$

$$P(\#\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}}) = 1) = 2 \cdot 0.6 \cdot 0.4 = 0.48 \tag{34}$$

$$P(\#\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}}) = 2) = 0.4^2 = 0.36 \tag{35}$$

thus

$$E(\#\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}})) = 2 \cdot 0.6 = 1.2 \quad . \tag{36}$$

In case of a trihedral junction the probabilities for the three edges to be incident with the junction point $\hat{\mathsf{P}}$ are different which leads to

$$E(\#\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}})) = \sum_{i=1}^{3} P(\text{INC}(\hat{\mathsf{P}}, \hat{\mathsf{L}}_i) \mid \mathsf{L}_j) \cdot P(\hat{\mathsf{P}} \mid \mathsf{P}) \tag{37}$$

where j relates to the three given edges on which the edges i fall.

Example: Let the trihedral corner with intensities 108, 128 and 160 [gr] be given (cf. Fig. 11). The contrasts of the three edges are 20, 32 and 52 [gr]. If we now select $SNR = 2.2$ for the left edge with $\Delta g = 32\,[\text{gr}]$ we can expect the middle edge with contrast 20 [gr] to be highly splitted and the neighbouring segments are merged as the $SNR = 1.37$ (cf. Fig. 7 and 8). The link of the estimated point to this edge also is unlikely as $c(\mathsf{L}) = 0.4$. On the other hand the right edge with contrast 52 [gr] will not be disturbed ($SNR = 2.2 \cdot 52/32 = 3.58$) and the left edge with medium contrast $\Delta g = 32\,[\text{gr}]$ will likely be splitted without allowing the neighbouring segments to merge.

The noisy corner image and the extracted features are shown in Fig. 11 (left and right resp.), thus confirming the prediction. □

5.2 Noise reduction by Smoothing

We have always used a fixed smoothing scale s. We now want to show how a change of s influences the feature extraction. Observe that in (16) $\nabla_s g = \nabla * G_s g$ thus convolution of g with G_s will reduce noise. Instead of g we could use $\overline{g} = G_s g$ which in case of $g = n$, thus just noise leads to $\overline{n} = G_s n$. For the variance $\sigma_{\overline{n}}^2$ we therefore obtain

$$\sigma_{\overline{n}}^2 = \int_x \int_y G_s^2(x,y)\,dx\,dy \cdot \sigma_n^2 \quad . \tag{38}$$

As the integral evaluates to $1/4\pi s^2$ we obtain for the noise in the filtered image

$$\sigma_{\overline{n}}^2(s) = \frac{\sigma_n^2}{4\pi^2 s^2} \quad . \tag{39}$$

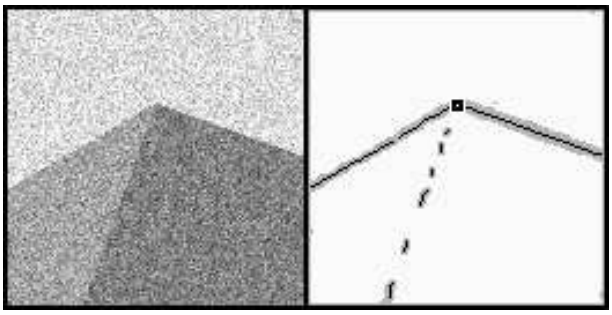

Figure 11: **Left:** the trihedral corner image with noise $\sigma_n^2 = 211\,[\mathrm{gr}^2]$. **Right:** result of the feature extraction. (For explanation cf. Fig. 4 .)

Thus the signal-to-noise ratio $SNR(s) = \Delta g/\sigma_{\overline{n}}(s)$ increases linearily with s.

Example: As the SNR in example Fig. 11 for the left and the middle edge were $SNR = 2.2$ and $SNR = 1.37$ resp., we change the smoothing scale s in a way that the SNR of the middle edge now is 2.2 . The left and the right edge then show $SNR = 2.2 \cdot 1.6 = 3.52$ and $SNR = 2.2 \cdot 2.6 = 5.72$ respectively. Thus, these both edges will certainly be extracted. However, the middle edge is at the boarder of being complete: the coverage $c(\mathsf{L}) >= 0.9$. The likelihood $P(\hat{\mathsf{S}}\,\mathrm{merges})$ that the neighbouring segments merge is approximately 20% . Two samples of the feature extraction are shown in Fig. 12: Obviously the edge is either complete or the two segments are merged. This supports the steep curve $P(\hat{\mathsf{S}}\,\mathrm{merges})$. We are not able to construct a case where the segments were separated and the edge being splitted within the edge region.

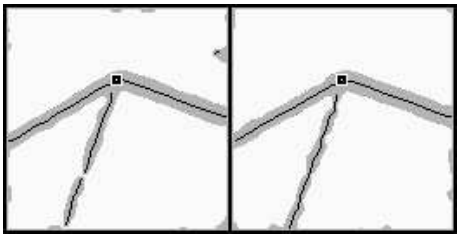

Figure 12: 2 results of the feature extraction of two instances of the noisy image shown in Fig. 11 with $s = 1.6$, $t_l = 2.24$ and $t_p = 4.0$. **Left:** the 2 segments are merged and the edge is splitted. **Right:** no splitting and no merging have been occured.

6 Conclusions

The paper provides basic tools for evaluating relational descriptions derived from digital images. Dependencies of the extracted features, their attributes and their relations on the control parameters of the feature extraction and on the local properties of the images have been analysed theoretically and by simulation using a set of images containing typical local image structures.
The contribution focusses on different aspects:
1. Investigations were done analysing the effect of the control parameters on the feature extraction. Besides an overall evaluation of our feature extraction module, the goal was to be able to choose these parameters in a top-down driven manner, of course depending on the application.
2. Analysing the occurence and the attributes of spurious features, especially in dependency on noise, scale and significance level directly might be used to improve the relational description given by the feature extraction leading to better results in the matching stage.
3. Analysing the quality of the feature extraction, i. e. probabilities of the occurence of features and their relations is necessary for any procedure optimizing the probability $P(M_i \mid D)$ as it is done within grouping and matching tasks. Difficulties in evaluating the quality of real structures which normally have high complexity are solvable by breaking down the structures into typical basic patterns. Examples were given demonstrating the predictability of the quality of the description of more complex patterns using the results of the analysis of only a few test patterns.
The choice of the test images was highly motivated by the applications we are dealing with, i. e. the reconstruction of man-made 2D or 3D structures, especially land use mapping and house detection. We choose the image of a trihedral corner due to its importance for image reconstruction and 3D-object recognition. Depending on the applications, of course, other structures are appropriate to be analysed.

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.

Christmas, B.; Kittler, J.; Petrou, M. (1994): Non-iterative contextual correspondence matching. In: Eklundh, J.-O. (Ed.), *Computer Vision-ECCV '94, Vol II*, pages 137–142. Lecture Notes in Computer Science, 801, Springer, 1994.

Förstner, W. (1994): A Framework for Low Level Feature Extraction. In: Eklundh, J.-O. (Ed.), *Computer Vision, ECCV '94, Vol. II*, pages 383–394. Lecture Notes in Computer Science, 801, Springer-Verlag, 1994.

Fuchs, C.; Löcherbach, Th.; Pan, H.-P.; Förstner, W. (1993): Land Use Mapping from Remotely Sensed Images. In: *Colloquium an Advances in Urban Spatial Information and Analysis, Wuhan, China*, 1993.

Grimson, W. E.; Huttenlocher, D. P. (1990): On the Verification of Hypothesized Matches in Model-Based Recognition. In: Faugeras, O. (Ed.), *Computer Vision - ECCV 90*, pages 489–498. Lecture Notes in Computer Science, 427, Springer-Verlag, 1990.

Herault, L.; Horaud, R. (1992): Figure-Ground Discrimination by Mean Field Annealing. In: Sandini, G. (Ed.), *Computer Vision ECCV '92*, pages 58–66. Lecture Notes in Computer Science, 588, Springer-Verlag, 1992.

Treutler, B. (1992): Verfahren zur Rekonstruktion von Knotenpunktsbereichen in Linienbildern. Master's thesis, Institut für Photogrammetrie, Universität Bonn, 1992.

Vosselman, G. (1992): *Relational Matching.* Lecture Notes in Computer Science 628. Springer-Verlag, 1992.