An automatic assessment scheme for steel quality inspection

Klaus Wiltschi[1], Axel Pinz[1] and Tony Lindeberg[2]

[1]Institute for Computer Graphics and Vision, TU-Graz
Münzgrabenstraße 11, A-8010 Graz, Austria.
Email: {wiltschi, pinz}@icg.tu-graz.ac.at, http://www.icg.tu-graz.ac.at

[2]Computational Vision and Active Perception Laboratory,
Department of Numerical Analysis and Computing Science,
KTH, S-100 44, Stockholm, Sweden.
Email: tony@nada.kth.se, http://www.nada.kth/~tony

Abstract

This paper presents an automatic system for steel quality assessment, by measuring textural properties of carbide distributions. In current steel inspection, specially etched and polished steel specimen surfaces are classified manually under a light microscope, by comparisons with a standard chart. This procedure is basically two-dimensional, reflecting the size of the carbide agglomerations and their directional distribution.

To capture these textural properties in terms of image features, we first apply a rich set of image processing operations, including mathematical morphology, multi-channel Gabor filtering, and the computation of texture measures with automatic scale selection in linear scale-space. Then, a feature selector is applied to a 40-dimensional feature space, and a classification scheme is defined, which on a sample set of more than 400 images has classification performance values comparable to those of human metallographers. Finally, a fully automatic inspection system is designed, which actively selects the most salient carbide structure on the specimen surface for subsequent classification.

The feasibility of the overall approach for future use in the production process is demonstrated by a prototype system. It is also shown how the presented classification scheme allows for the definition of a new reference chart in terms of quantitative measures.

Keywords: multi-scale analysis, automatic scale selection, multi-channel texture analysis, active inspection system, carbide classification.

i

Contents

1 Introduction

In the production of high speed steel, the rolling affects the micro-structure of the steel, which in turn influences the mechanical properties (Fischmeister, Paul & Karagöz 1988, Beiss & Wähling 1989). Specifically, the distribution of carbide is essential, since cracks propagate within the carbide agglomerations. In current quality control, the properties of the steel are assessed manually by comparison with a standard chart (figure 2), containing representative patterns for each steel class. These classifications as well as the standard chart itself, however, are not based on quantitative measures but on the qualitative visual impression of the metallographers.

The standard technique for classifying these carbide distributions is two-dimensional, where the first dimension, called *degree*, basically corresponds to scale (to the size of the largest carbide agglomeration) and the second dimension, denoted *type*, basically reflects the directional distribution (how strongly the net structure of carbide has been stretched by rolling). Thus, this classification task deals with the problem of how to characterize scaled textures, i.e. structures with scale and shape continuously distributed over a certain range. The goal of this article is to present an integrated prototype system, which performs this classification in a fully automatic manner, and thus carries out inspection of high speed steel specimens according to their carbide distribution.

To capture the textural properties mentioned above in terms of image features, we first develop a rich repertoire of image processing tools, based on morphological operations (Serra 1982, Serra 1988), Gabor filtering (Bovik, Clark & Geisler 1990, Jain & Bhattacharjee 1992) and scale-space methods (Witkin 1983, Koenderink 1984, Florack, ter Haar Romeny, Koenderink & Viergever 1992, Lindeberg 1994, Florack 1997). In particular, to reliably estimate the size of the dominant structures in the image, we explore recently developed techniques for automatic selection of significant scales (Lindeberg 1994, Lindeberg 1996*b*, Lindeberg 1996*a*) and use this scale information in the computation of textural descriptors (Bigün, Granlund & Wiklund 1991, Gårding & Lindeberg 1996, Lindeberg & Gårding 1993). For each of these classes of image processing operations, a set of image features is presented, which capture size and shape information and altogether result in a 40-dimensional feature space. To select discriminative features from this feature space, and to reduce its dimensionality, we apply feature selection methods (Devijver & Kittler 1982, Kittler 1986, Fukunaga 1990) before the subsequent parametric classification is carried out. Finally, an active inspection system is presented, which integrates the presented algorithms and substantially improves the computational speed, compared to a more traditional static analysis.

The paper is organized as follows: We start our presentation with a description of the image material used (section 2. The first method, described in section 3.1, is a data-specific scheme to extract a so-called characteristic structure of the carbide agglomerations. This structure is used for capturing features corresponding to the manually assessed properties and also serves as a reference for the subsequent scale selection module. Based on two multi-scale techniques, multi-channel Gabor filtering (section 3.2) and automatic scale selection in linear scale-space (section 3.3), features are defined for estimating the scale and the shape of textures. In sections 3.4 and 3.5, we describe the feature selection and classification approaches. Finally, we present

the results of scale and shape estimation (section 4.1) and of using the proposed classification scheme (section 4.2), for constructing a fully automatic steel inspection scheme (section 5) as well as for defining a new objectively defined reference chart (section 6).

2   Material

To assess the carbide distribution of high speed steel, specimens of a few cm^3 are taken. The specimen surface is polished and etched to show the carbides. By means of a light microscope and a standard CCD-camera, images are taken at a magnification of 1:100, where the rolling direction of the specimens is always aligned horizontally, which results in a horizontal preferential direction. Figure 1 shows such an image with a medium-sized strongly stretched carbide distribution, where the carbide particles are represented by white.[1]
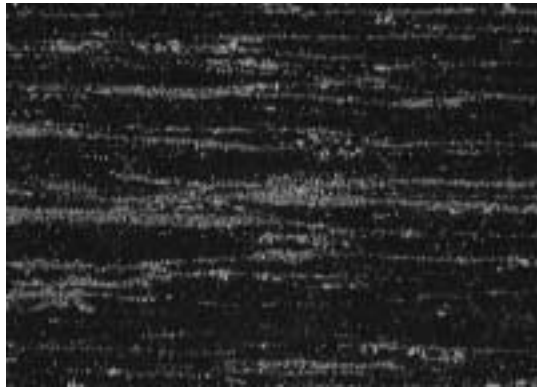

Figure 1: A microscopic image of a medium-sized, strongly stretched carbide distribution at a magnification of 1:100 (white pixels = carbides; 700×500 pixels; 1 pixel $\approx 1,27\mu m$).

2.1   Standard chart for carbide distributions

The *standard chart* depicted in figure 2 shows the large variety of the appearance of carbide distributions. It distinguishes the carbide distributions by their *type* (i.e. from line-shaped to net-shaped) and by their *degree* (i.e. the size of their agglomerations). The 4 different distribution types (increasing from top to bottom) and the 7 different distribution degrees (increasing from left to right) are arranged in the rows and columns of the standard chart. The forms of carbide distributions occurring in the production are continuously distributed over the whole range of types and degrees.

It should be noted, however, that the reference images have been chosen based on the qualitative expert knowledge to properly represent all forms of the carbide distributions. This implies that there is no quantitative definition of the classes. As described in section 6, our new automatic classification scheme, opens up the possibility of defining such a reference chart by quantitative measures.

---

[1]The images used throughout this work have a size of 700×500 pixel and each pixel corresponds to a surface area of 1.27 μm× 1.27 μm.
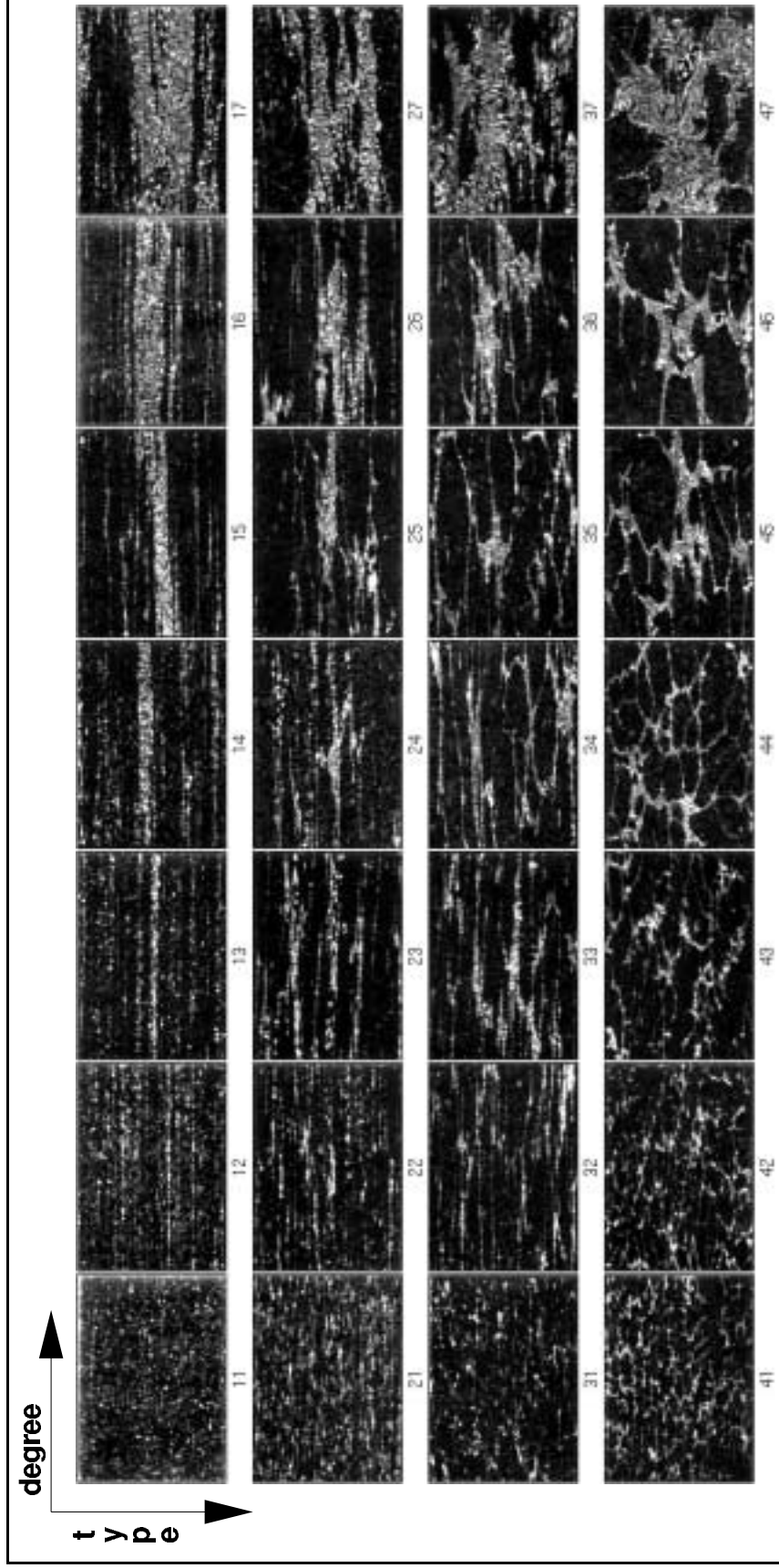
Figure 2: The microscopic standard chart for the characterization of carbide distributions of high speed steels. Steel quality increases from lower right to upper left. The first number of the class (the row index) denotes the type and the second number (the column index) denotes the degree of the carbide distribution. The degree corresponds to scale (size of the carbide agglomeration). The type reflects the directional distribution (stretching of the net-shaped structure due to rolling).

3

2.2 Carbide distributions - a scaled texture

The degree of the carbide distribution mainly reflects the size of the largest carbide distribution. Therefore, the standard chart can be interpreted as the representation of 4 different textures at 7 discrete scales, where the type constitutes the different texture types and the degree defines the scale of the texture. Figure 3 justifies this interpretation by comparing a resized subwindow of reference image 22 with reference image 23. It is obvious that the resized subimage shows a rather similar carbide distribution except the spatial resolution. This property is one of the motivations for using multi-scale techniques to characterize the carbide distributions.

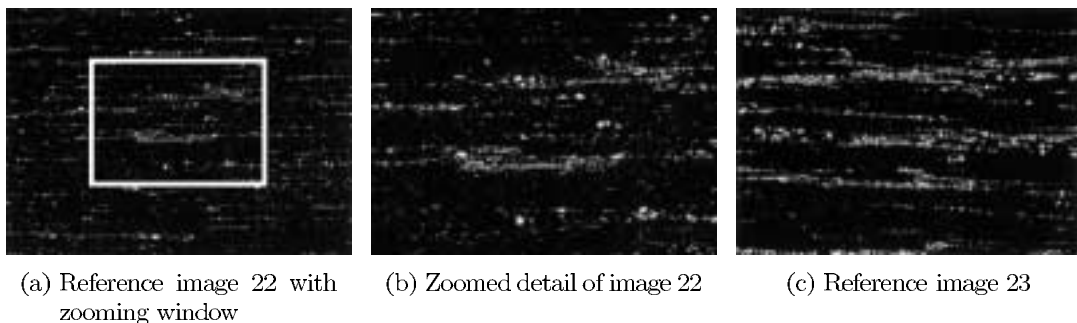| (a) Reference image 22 with zooming window | (b) Zoomed detail of image 22 | (c) Reference image 23 |

Figure 3: Scaled texture nature of the reference images. As can be seen, the zoomed detail (b) of reference image 22 (a) shows a rather similar carbide distribution as the direct neighboring reference image 23 (c).

2.3 Datasets

In addition to the reference images, a set of carbide distribution images has been captured from high speed steel specimens at a magnification of 1:100. These 438 images have been independently classified by 4 metallographers according to the 28 reference classes of the standard chart. The number of identical classifications gives a measure of the reliability of the manual classifications. Based on these manual classifications, we define the most frequent classification as the *true class* used for the classification experiments. For example, for an image manually classified as 23, 23, 24 and 33 by the 4 metallographers the true class is defined to be 23. In 9 cases, no true class can be found due to 4 different manual classifications. These images are omitted, yielding a dataset of 429 images to capture the expert knowledge.

	correct classifications worst / mean / best [%]	classifications within one-class-deviation worst / mean / best [%]
degree	77 / 81 / 86	96 / 97 / 99
type	85 / 88 / 92	100 / 100 / 100

Table 1: Worst, mean and best performance of the manual classifications of the metallographers, estimated by the deviation from the true class.

The reliability of the manual classifications is illustrated in table 1, where the performance of each metallographer is estimated by the deviation from the true class.

4

Of course, this assessment is an overestimate of the reliability of the manual classifications, since the judgment of the true class is solely based on the 4 manual classifications. For comparison, the performance values achieved by tolerating a one-class deviation are listed as well.

3 Methods

The scale of the carbide distributions ranges from fine irregular structures, i.e. textures with small texture primitives, to large structures, i.e. image-sized texture primitives. Obviously, any analysis of these images has to explicitly or implicitly determine the size of the operators to use. In the following sections, we present techniques to handle the large scale and shape range of the carbide distribution images, based on multi-channel texture analysis and automatic scale selection in linear scale-space. We start by presenting a data-specific scheme to extract a so-called characteristic structure of the images, which is used both for defining image features for classification and in a verification module of the automatic scale selection step.

3.1 A characteristic structure of carbide distributions

In the manual classification of carbide distributions, the size and the shape of the largest carbide agglomeration are the main factors that influence the classification. As a first processing step, we shall capture this information in the form of characteristic structure, by generating a characteristic binary carbide structure, basically by global thresholding, a special filtering operation and a size-adapted morphological scheme. The method is based on (Wiltschi, Pinz & Hackl 1995, Wiltschi, Pinz & Hackl 1996), and the results are illustrated using the reference images of the standard chart. In general, the parameters of the operations have been determined empirically, using the images of the standard chart and of the carbide image dataset.

3.1.1 Segmentation

To extract the carbide agglomerations, a distinction between the carbide and background areas is necessary. Segmentation (Gonzalez & Woods 1992, chapter 7) based on the detection of discontinuities does not work due to the punctuate nature of the image data, whereby already slight smoothing previous to segmentation suppresses significant structures in the fine carbide distribution images. Due to the contrast of the carbide particles, however, a global thresholding can be applied to properly segment the carbides. The unimodal histograms of the carbide distribution images (figure 4.b) prevent the application of the classical optimal threshold detection scheme (Gonzalez & Woods 1992, p. 447), based on the approximation of two peaks in the histogram. Thus, we use a threshold detection based on the first derivative of the histogram, by detecting a specific slope value. This slope value is determined by empirically chosen parameters, which yield a reliable segmentation of all available images, and is adapted to the brightness of the image. In figure 4.b, the detected threshold is marked and figure 4.c shows the resulting binary image for reference image 14.

5

Next, isolated points outside the carbide agglomerations are erased, which is imperative for two reasons. First of all, these points are ignored during the manual inspection, and secondly, they turn out to disturb the robustness of the subsequent morphological closing. Therefore, the number of white pixels in a 7x7 neighborhood is computed for all pixels in the binary image. Based on the empirical mean value and the maximum value of these numbers from the whole image, a threshold is determined and all pixels with fewer white neighbors than this threshold are erased (figure 4.d). Similar filtering operations can be found in (Haralick & Shapiro 1992).

The subsequent morphological closing (Serra 1982, Serra 1988) yields connected regions for the dense carbide areas (Chermant, Coster & Gougedon 1989, Russ 1991, Fischmeister et al. 1988). To connect the white pixels within the agglomerations isotropically, which is especially important for net-shaped structures, we use a circular structuring element of 9 pixel diameter. The diameter has been determined empirically to optimize the connectivity of the resulting regions. Figure 4.e shows the result of this morphological closing. As can be seen, the dense carbide particle areas form connected regions in the resulting closed binary image.
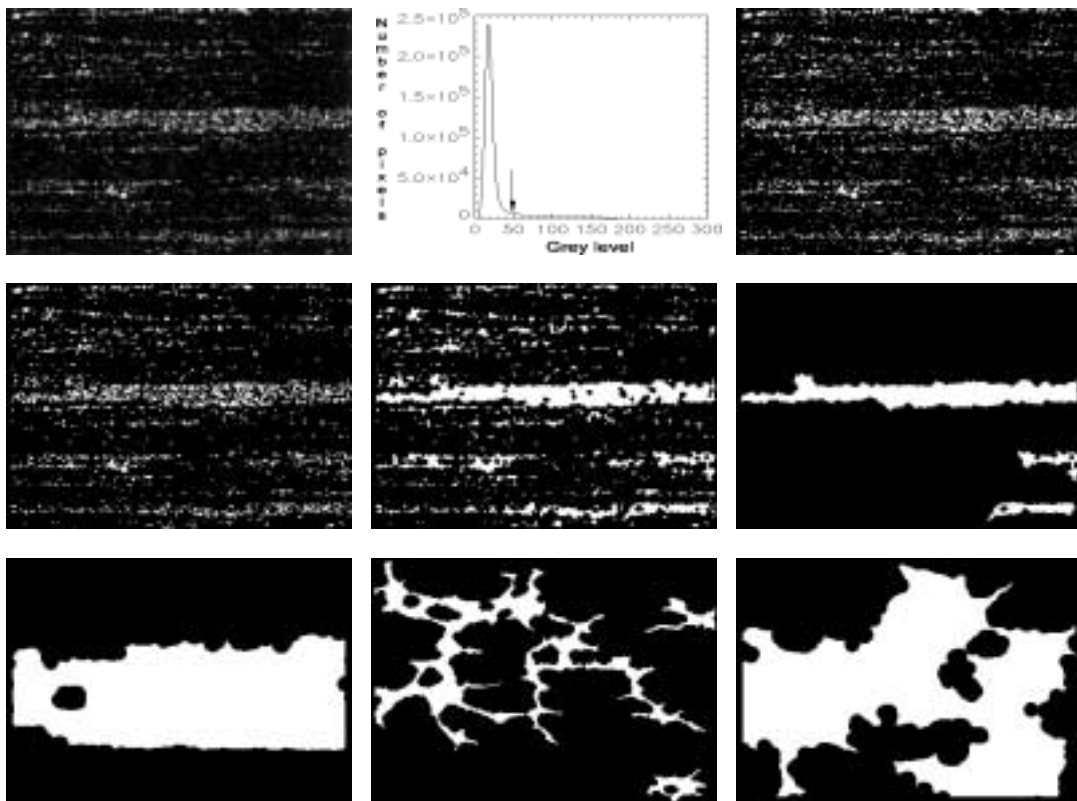

Figure 4: Generation of the characteristic binary structure illustrated for (a) reference image 14; (b) smoothed histogram with marked threshold, (c) resulting binary image, (d) filtered binary image, (e) morphologically closed image and (f) characteristic structure. For comparison (g)-(i) show the characteristic structures for reference images 17, 44 and 47.

6

3.1.2 Definition of the characteristic structure

After the abovementioned segmentation and morphological steps, a region labeling algorithm is applied to separately locate each carbide cluster. Using this information, we select a set of relevant clusters

$$CS = \{(C_i | A(C_i) \geq d_0 A_{max}) \; \wedge \; (A(C_i) \geq d_1)\} \, , \tag{1}$$

where $A(C_i)$ denotes the area of cluster C_i, i.e. the number of pixels, A_{max} is the area of the largest cluster in the present image and in discussion with the experts we have empirically chosen $d_0 = 0.1, d_1 = 10$ pixel to capture the significant information. Subsequently, each cluster $C_i \in CS$ is morphologically closed with a circular structuring element with a cluster-size dependent diameter. This size-adapted closing yields a set of 'smoothed' clusters (figure 4.f), where insignificant details, like small holes, are removed. The new set of clusters defines the *characteristic structure* of a carbide distribution. For comparison, figures 4.g-i show the characteristic structures of three other reference images.

3.1.3 Feature extraction

The properties assessed by the metallographers to determine the degree of the carbide distribution can be approximated by features defined from the characteristic structure in the spatial domain. Some of the features that will be described subsequently are based on the methods used in quantitative micro-structural analysis (Exner & Hougardy 1990), such as linear projection values, whereas others are new approaches to model the characteristic information, such as evaluations of skeletons. The extraction of significant features for distinguishing between different types of carbide distributions, however, turns out to more complicated, and we will therefore define additional features in the frequency domain. In this presentation, we just describe the more significant of the altogether 32 features, we have used to arrive at the final assessment scheme (see (Wiltschi 1998) for more details).

Shape features. The *number of clusters* in the characteristic structure is a feature which clearly discriminates fine from medium-sized as well as coarse carbide distributions. An important feature for the determination of the carbide distribution degree is the *area of the largest cluster A_{max}*, which is given by the number of pixels of the largest cluster.

The *Hotelling transform* is used to determine the principal axes (eigenvectors) and the corresponding eigenvalues of the scatter matrix for the largest cluster of the characteristic structure. The direction of maximum spread and its perpendicular direction are indicated by these axes and the corresponding eigenvalues give the variance in these directions. The *smallest eigenvalue of the scatter matrix λ_2* estimates the width of the carbide-lines and consequently increases with increasing degree. Additionally, this feature tends to increase with increasing type, especially for coarse carbide distributions. The *largest eigenvalue of the scatter matrix λ_1* is also used as a feature to capture the size information in the principal direction. The ratio between the smallest and the largest eigenvalues defines the *eccentricity of the largest cluster $\mathcal{E} = \lambda_2/\lambda_1$*, which allows us to distinguish elongated from compact structures.
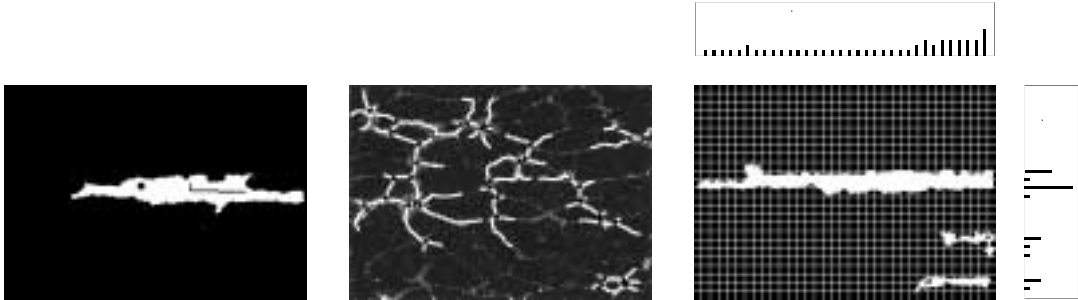
7

Figure 5: Features based on the characteristic structure; (a) area of the largest cluster A_{max} of reference image 25; the principal axes and the corresponding eigenvalues are shown by the dark lines; (b) detection of ramifications in the carbide distribution structure; skeleton (white) and the detected nodes (black) for reference image 44; (c) horizontal and vertical linear projection of the characteristic structure of reference image 14; the distributions of the number of structure segments per projection line are shown.

Figure 5.a illustrates these features extracted from the largest cluster of the characteristic structure using reference image 25. In order to get a more reliable information for fine carbide distributions, the *mean values* over all clusters of the characteristic structure are calculated as well for the following features: smallest eigenvalue of the scatter matrix λ_2, largest eigenvalue of the scatter matrix λ_1 and eccentricity of the largest cluster $\mathcal{E}$.

Skeleton of characteristic structure. To assess the net-shaped structure of a carbide distribution, the skeleton of the characteristic structure is determined by the *classical skeleton algorithm* from (Pavlidis 1982, section 9.2). Afterwards, we locate the nodes of this skeleton by convolution with the 16 masks of 3×3 pixels representing all possible forms of nodes in this skeleton. The number of these nodes yields an approximation of the *number of ramifications* in the characteristic structure and serves to evaluate the type. Moreover, we use the *path length of the skeleton* as a feature, which is given by the number of pixels of the skeleton. Figure 5.b illustrates this scheme for reference image 44. Due to the sensitivity of the skeleton algorithm to details in the boundary of an object, not all nodes in the skeleton (figure 5.b) correspond to a ramification of the carbide distribution.

Linear projections. As shown in figure 5.c, parallel, horizontal and vertical lines with a constant distance are used to determine various features along these lines. This linear projection scheme is used in quantitative micro-structural analysis (Exner & Hougardy 1990) as well. In order to capture the extensions of the carbide agglomerations, the length of agglomeration segments and the number of segments per line in the characteristic structure are determined and the number of intersections per line in the skeleton are calculated to assess the shape of the carbide distribution. The distributions of these values are evaluated statistically, where the *maximum* and *mean* values of the *number of vertical segments, length of vertical segments, length of horizontal segments* and *number of vertical intersections* turn out as features capturing significant information for the characterization of the carbide distributions. Addition-

8

ally, the *number of non-empty horizontal lines in the structure* and *in the skeleton* are used as features for the distinction of band-shaped and net-shaped structures.

(a) $|\mathcal{F}\{I(x,y)\}|$ (zoomed 1:2)   (b) Radially summing up $|\mathcal{F}|$ (c) Approximation with f_g
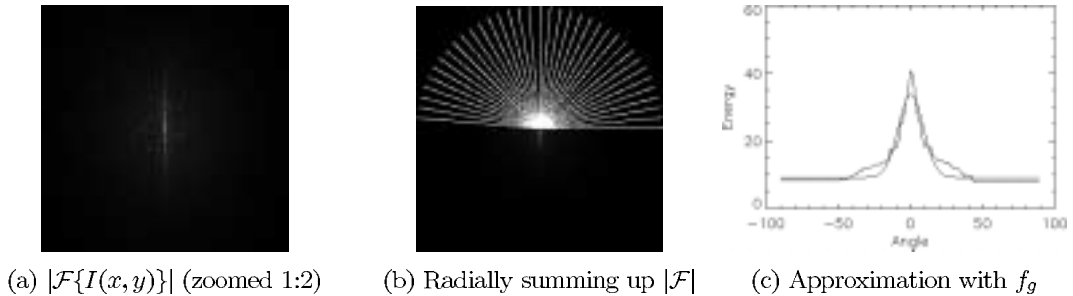
Figure 6: Evaluation of the spectrum illustrated for reference image 14 (a); the spectrum $|\mathcal{F}\{I(x,y)\}|$ (b) is summed up radially to get an directional energy distribution which is approximated by a Gaussian function (c).

Frequency-based orientation estimate. The amount of stretching of the net-shaped structure of carbides during the production process (rolling) determines the type of the carbide distribution and can be described by the degree of orientation of the structure. In order to assess the directional information of the carbide distribution images, we sum up the Fourier spectrum $|\mathcal{F}\{I(x,y)\}|$ radially as illustrated in figure 6, which yields an energy distribution over the angle of orientation from $0°$ to $180°$. Such techniques are also described in (Gonzalez & Woods 1992, p. 511ff). The flanks of this distribution are erased, so that they cannot disturb the subsequent evaluation of the degree of orientation of the preferential direction, which is always horizontal due to the specimen preparation process. Afterwards, this directional energy distribution is approximated by a Gaussian function f_g parameterized by

$$f_g = p_1 \ \exp\left[-\frac{1}{2}\frac{(x-p_2)^2}{p_3^2}\right] + p_4, \tag{2}$$

where p_1 reflects the amplitude, p_2 the preferred direction, p_3 the width and p_4 the energy offset of the energy distribution. Here, the parameter p_3 provides a significant feature for the *degree of orientation* of the carbide distribution structure, which increases with increasing type.

3.2 Multi-channel texture analysis

The multi-channel Gabor filtering technique, introduced by (Bovik et al. 1990, Jain & Farrokhnia 1991, Bovik 1991, Jain & Bhattacharjee 1992) provides a frequency and orientation selective image representation. In particular, the energy in the different frequency and orientation channels yields an estimate of the amount of structure at a specific scale and of a certain orientation.

The impulse response of a two-dimensional *real even-symmetric Gabor filter* is given by

$$h_{res}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \ \exp\left[-\frac{1}{2}\left(\frac{x'^2}{\sigma_x^2}+\frac{y'^2}{\sigma_y^2}\right)\right] \ \cos(2\pi Fx'), \tag{3}$$

9

(a) Spatial domain representation

(b) Fourier domain representation

(c) Multi-channel half-peak support
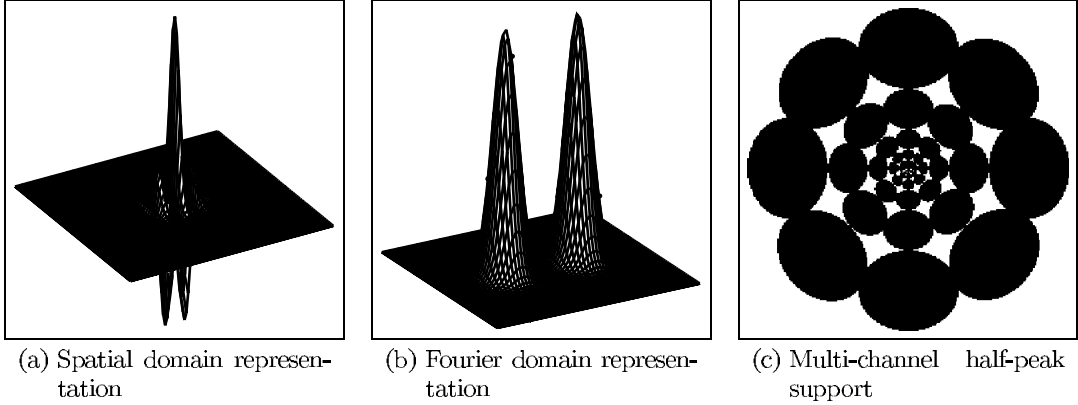
Figure 7: Real even-symmetric Gabor filter (a) in the spatial and (b) in the Fourier domain ($F = 8\sqrt{2}$ cpi, $\phi = 0°$, $B_f = 1$ octave and $B_\phi = 45°$); (c) half-peak support of multi-channel Gabor representation with $F_i = 1\sqrt{2}, 2\sqrt{2}, \cdots, 64\sqrt{2}$ cpi, $\phi = 0°, 45°, \cdots, 135°$, $B_f = 1$ octave and $B_\phi = 45°$.

where the radial frequency $F = \sqrt{U^2 + V^2}$ measured in 'cycles per image' (cpi) is the frequency of a sinusoidal planar wave with orientation $\phi = \arctan(V/U)$. The standard deviations of the Gaussian enveloped along the principal axes are denoted by σ_x and σ_y, and $(x', y')^T$ denotes the coordinate system rotated by the angle ϕ. The representation in the frequency domain is

$$H_{res}(u, v) = \exp\left[-\frac{1}{2}\left(\frac{(u' - F)^2}{\sigma_u^2} + \frac{v'^2}{\sigma_v^2}\right)\right] + \exp\left[-\frac{1}{2}\left(\frac{(u' + F)^2}{\sigma_u^2} + \frac{v'^2}{\sigma_v^2}\right)\right] \quad (4)$$

where $\sigma_u = 1/(2\pi\sigma_x)$ and $\sigma_v = 1/(2\pi\sigma_y)$. (Figure 7.a-b shows such a filter in both the spatial and the frequency domains.) The half-peak magnitude *frequency bandwidth* B_f and the *orientation bandwidth* B_ϕ of the Gabor filter defined by equation (3) and (4) is given by

$$B_f = \log_2\left(\frac{\pi F\sigma_x + \sqrt{\ln 2/2}}{\pi F\sigma_x - \sqrt{\ln 2/2}}\right) \quad \text{and} \quad B_\phi = 2\arctan\left(\frac{\sqrt{\ln 2/2}}{\pi F\sigma_y}\right), \quad (5)$$

where B_f is measured in octaves and B_ϕ in degrees.

An almost complete multi-scale representation for an image with size $2^N \times 2^N$ pixels can be constructed by filtering with a *Gabor filter set* using $k = N - 1$ radial frequencies

$$F_i = F_0 \, 2^{iB_f}; \quad F_0 = \sqrt{2}; \quad i = 0, \cdots, k - 1, \quad (6)$$

and l orientations (tuned to the orientation resolution needed)

$$\phi_j = jB_\phi; \quad j = 0, \cdots, l - 1, \quad (7)$$

with frequency bandwidth $B_f = 1$ octave and orientation bandwidth $B_\phi = 180°/l$. In the following, this representation is referred to as a *multi-channel Gabor representation*. The half-peak support of such a Gabor filter set is illustrated in figure 7.c.

10

3.2.1 Scale and shape estimation based on energy distributions

Let $r(x,y)$ be the discrete single *Gabor filter response* of an input image $i(x,y)$. The energy E of this response $r(x,y)$ yields a measure of the amount of structure in the image $i(x,y)$, which corresponds to the specified frequency channel with respect to scale (frequency) and orientation. It is given by

$$E = \sum_{u,v} |R(u,v)|^2 = \sum_{x,y} [r(x,y)]^2. \qquad (8)$$

Thus, for a multi-channel Gabor representation of an image, the distribution of the energy $E(F_i, \phi_j)$ over all frequency channels F_i and orientation channels ϕ_j can be calculated. The energy distribution with respect to the frequency channels contains information about the scale of the structures in the image, whereas the energy distribution with respect to the orientation channels explicitly shows the directional information of the image (figure 8).
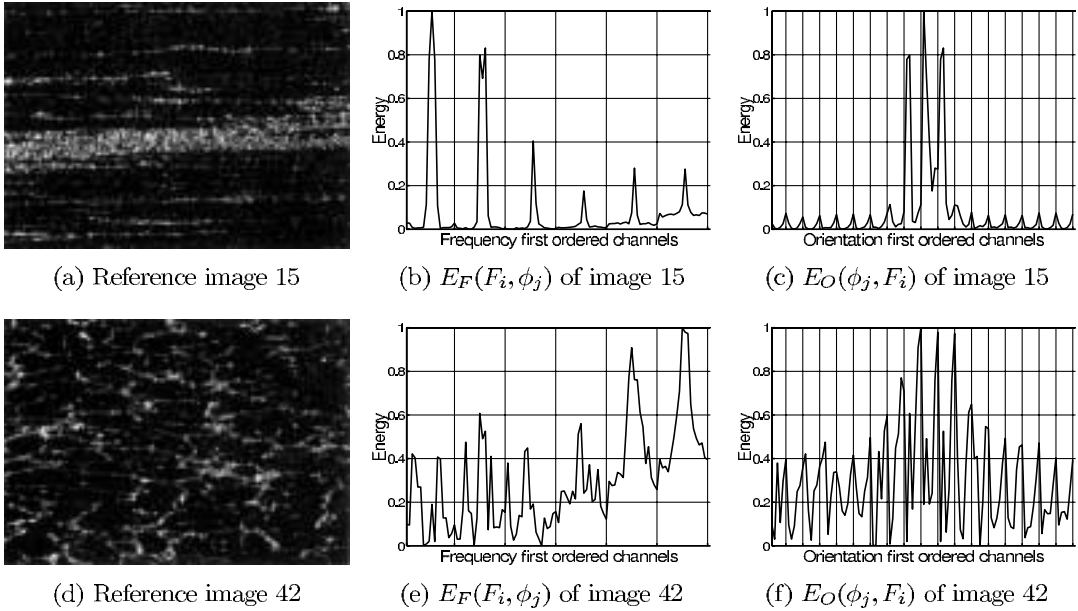

(a) Reference image 15   (b) $E_F(F_i, \phi_j)$ of image 15   (c) $E_O(\phi_j, F_i)$ of image 15

(d) Reference image 42   (e) $E_F(F_i, \phi_j)$ of image 42   (f) $E_O(\phi_j, F_i)$ of image 42

Figure 8: Energy distributions of a multi-channel Gabor representation using 6 frequency channels $F_i = F_0, 2F_0, \cdots, 32F_0$ ($F_0 = \sqrt{2}$ cpi) and 18 orientation channels $\phi_j = 0°, 10°, \cdots, 170°$ with $B_f = 1$ and $B_\phi = 10°$ calculated for (a) reference image 15 and (d) image 42; (b) and (e) show the corresponding frequency energy distribution $E_F(F_i, \phi_j)$ and (c) and (f) the corresponding orientation energy distribution $E_O(\phi_j, F_i)$.

The energy distribution $E(F_i, \phi_j)$ is first ordered according to the frequency channels F_i, where the energy values of the different orientation channels ϕ_j are arranged according to ascending orientation within each frequency channel. This ordering is referred to as *frequency energy distribution* $E_F(F_i, \phi_j)$. In the second arrangement called *orientation energy distribution*, $E_O(\phi_j, F_i)$ the energy values are ordered according to the orientation channels ϕ_j and within each of these channels the energy values are arranged in ascending frequency order. Obviously, the frequency energy distribution shows large energy values in frequency channels corresponding to the

11

scales of the structures in the images (figure 8.b and e). Furthermore, the orientation energy distribution of a strongly oriented pattern like figure 8.a shows only high energy values for the orientation channel corresponding to the orientation of the structure (figure 8.c), whereas an almost isotropic pattern like figure 8.d yields a wide distribution of the energy over all orientation channels (figure 8.f).

Based on these properties of the energy distribution, we derive the following features to assess the scale corresponding to the degree and the shape reflecting the type of the carbide distribution images. In order to reduce the influence of the high energy values in the highest frequency channels for all orientations, we estimate the scale with regard to dominant orientation, i.e. the *degree estimate of maximum Gabor energy* E_{fm} is given by

$$p_{fm}(i) = \max_{j=0,\cdots,l-1} E_F(F_i, \phi_j); \quad i = 0, \cdots, k-1 \tag{9}$$

$$E_{fm} = \frac{1}{k-1} \sum_{i=0}^{k-1} \frac{p_{fm}(i)}{\Sigma(p_{fm})} \; i; \quad \Sigma(p_{fm}) = \sum_{i=0}^{k-1} p_{fm}(i) \tag{10}$$

The directional information decisive for the type of the carbide distributions is the relation between horizontal and non-horizontal structures. Structures with orientation $45°$ and $135°$ have the same impact on the type of the carbide distribution. Thus, the orientation energy distribution is assessed between the orientations $\phi = 0°$ to $90°$ by the *reduced orientation energy distribution*

$$E'_O(\phi_j, F_i) = \begin{cases} E_O(\phi_j, F_i) & \text{if} \quad j = 0, \\ 1/2 \left[E_O(\phi_j, F_i) + E_O(\phi_{l-j}, F_i) \right] & \text{if} \quad j = 1, \cdots, l/2. \end{cases} \tag{11}$$

Analogously to the degree feature, the estimation of the type is performed scale-adapted, i.e. with respect to the dominant frequency channels to only assess the directional information for significant scales. Using the dominant frequency channel F_d determined by the value of E_{fm}, we define the *type estimate of maximum Gabor energy* E_{om} by

$$p_{om}(j) = \sum_{i=s-1}^{s+1} E'_O(\phi_j, F_i); \quad j = 0, \cdots, \frac{l}{2}, \quad s = \begin{cases} 1 & \text{if} \quad d = 0, \\ d & \text{if} \quad 0 < d < k, \\ k-1 & \text{if} \quad d = k, \end{cases} \tag{12}$$

$$p'_{om}(j) = p_{om}(j) - \min_{r=0,\cdots,l/2} p_{om}(r), \quad E_{om} = \frac{2}{l} \sum_{j=0}^{l/2} \frac{p'_{om}(j)}{\Sigma(p'_{om})} \; j; \quad \Sigma(p'_{om}) = \sum_{j=0}^{l/2} p'_{om}(j) \tag{13}$$

3.3 Automatic scale selection in linear scale-space

To handle the inherent multi-scale nature of image data, the notion of scale-space theory has been developed (Witkin 1983, Koenderink 1984, Lindeberg 1994, Florack 1997). For any N-dimensional image $f : \mathbb{R}^N \to \mathbb{R}$, its *scale-space representation* $L : \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$ is defined by

$$L(\cdot; t) = g(\cdot; t) * f(\cdot) \tag{14}$$

where $g : \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$ denotes the N-dimensional Gaussian kernel:

$$g(\cdot; t) = \frac{1}{(2\pi t)^{N/2}} e^{\frac{-(x_1^2 + \cdots + x_N^2)}{2t}} \qquad (15)$$

and the variance $t \in R^+$ of the Gaussian kernel is referred to as the *scale parameter*. Based on this representation, *scale space derivatives* are defined by

$$L_{x^\alpha}(\cdot; t) = \partial_{x_1^{\alpha_1} \cdots x_N^{\alpha_N}} L(\cdot; t) = (\partial_{x_1^{\alpha_1} \cdots x_N^{\alpha_N}} g(\cdot; t)) * f(\cdot) \qquad (16)$$

with corresponding *normalized derivatives* given by

$$\partial_{\xi_i^n} = t^{n/2} \partial_{x_i^n}. \qquad (17)$$

Several feature detectors can be formulated as (linear or non-linear) combinations of partial derivatives. Specifically, scale levels for feature detection can be selected by detecting local extrema over scales of such differential geometric descriptors (Lindeberg 1993, Lindeberg 1994, Lindeberg 1996b, Lindeberg 1996a). Here, we shall use this framework for constructing feature detectors for evaluating the degree of the carbide distributions, which basically corresponds to the size of the carbide agglomerations. Furthermore, we shall compute directional information from scale-tuned second moment descriptors, using the scale information from the scale selection module. In the following, a methodology is described for reliable extraction of characteristic features for estimating the degree and type of carbide distributions (based on (Wiltschi, Lindeberg & Pinz 1997a, Wiltschi, Lindeberg & Pinz 1997b)).

3.3.1 Determining the significant scale

Since the agglomerations mainly form blob- or ridge-like structures, (see figure 2), we use a *ridge detector* to capture these image features. Building upon earlier methods for ridge detection (Haralick 1983, Eberly, Gardner, Morse, Pizer & Scharlach 1994, Koenderink & van Doorn 1994) it is defined as follows (Lindeberg 1996a): Introduce a local (p, q)-coordinate system at each image point, defined by the mixed second-order derivative being zero (i.e. $L_{pq} = 0$). Then, we can detect (possibly elongated, bright) blob features from points which are simultaneously maximal with respect to space and scales in

$$-L_{pp,norm} = -t \, L_{pp}, \qquad (18)$$

where L_{pp} is the principal curvature having the largest absolute value.

To rank features detected by means of equation (18) on significance, we weight the normalized response $L_{pp,norm}$ at each scale-space extremum by $\sqrt{t}$, which gives the *significance measure*

$$S(E) = -\sqrt{t} \, L_{pp,norm}(x, y). \qquad (19)$$

An intuitive motivation for this weighting is that the width of the ridge feature can be expected to be proportional to $\sqrt{t}$.

Certain image structures give rise to multiple responses. To suppress overlaps, a scale-space maximum A is rejected if there exists another maximum B and

$$(\text{center}(A) \in \text{support region}(B)) \wedge$$
$$(\text{center}(B) \in \text{support region}(A)) \wedge \tag{20}$$
$$\left(t_A/t_B \in \left[\tfrac{1}{\alpha}, \alpha\right]; \ (\alpha > 1)\right) \ \wedge \ (S(A) < S(B))$$

with $\alpha = 4$ corresponding to a ratio of 2 between the blob radii. Figure 9.a shows the result of applying these operations to reference image 33, where the detected ridges are marked by their elliptical support region $\mathcal{S}_{ell}$ (equations (21)-(24)).

(a) Elliptical support (10 most significant extrema)

(b) Characteristic structure

(c) Euclidean distance map within largest ellipse

(d) Local extrema in distance map (c)

(e) Non-overlapping largest circular background regions for (c)

(f) Verified structures (10 most significant scale-space extrema)
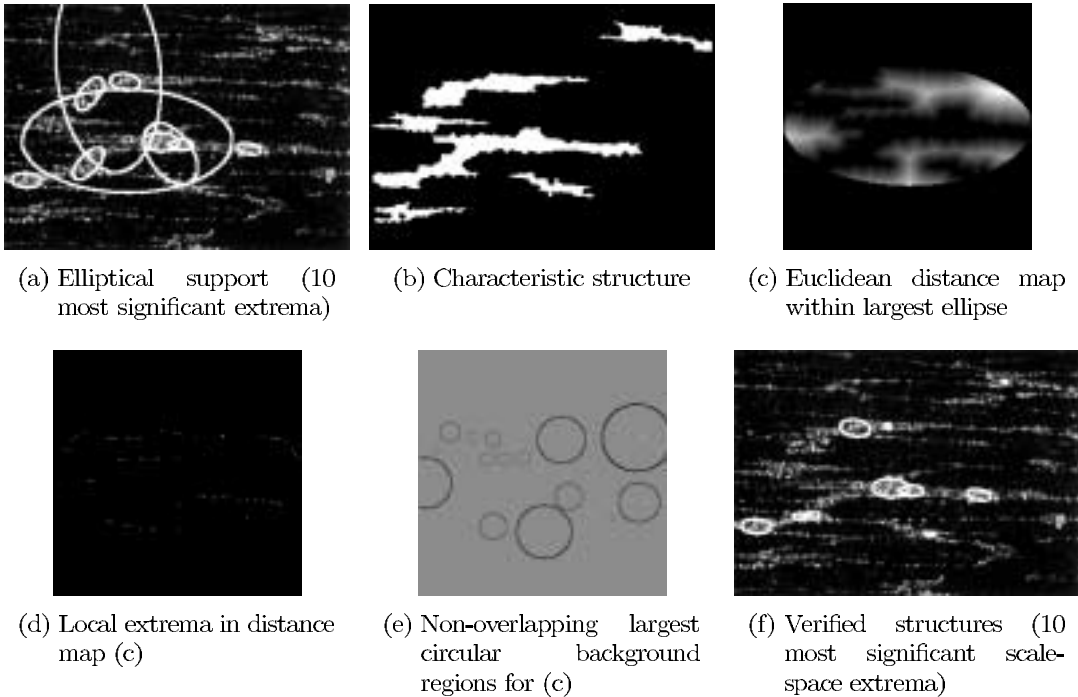
Figure 9: Verification of the most significant scale-space extrema using the characteristic structure.

To suppress spurious responses from the feature detection module (false alarms), we use the following verification mechanism, which is based on the morphologically constructed characteristic structure of the carbide distributions (section 3.1.2). An *elliptical half-peak support region* $\mathcal{S}_{ell}$ is associated with each scale-space extremum E_a detected at scale t_a based on the two principal curvatures $L_{pp,norm}/L_{qq,norm}$ as well as the orientation of the ridge. Using an idealized two-dimensional elliptical blob model

$$B_{ell}(x,y; \ t_1, t_2) = \frac{1}{\sqrt{2\pi t_1}} \exp\left(-\frac{x^2}{2t_1}\right) \frac{1}{\sqrt{2\pi t_2}} \exp\left(-\frac{y^2}{2t_2}\right), \tag{21}$$

the ratio of the principal curvature values in dependence on the scale t is given by

$$\frac{L_{pp,norm}}{L_{qq,norm}} = \frac{t + t_2}{t + t_1} \tag{22}$$

14

for $t_1 < t_2$. The local extremum over scale for the idealized blob B_{ell} occurs at (Lindeberg 1996a)

$$t_{ideal} = \frac{1}{4}(t_1 - t_2) \pm \sqrt{\frac{1}{16}(t_1 - t_2)^2 + t_1 t_2} \qquad (23)$$

Then, the ratio t_2/t_1 corresponding to the major and the minor axis of the ellipse can be estimated as

$$\frac{t_2}{t_1} = \frac{1}{2}\left(3\,\frac{L_{pp,norm}}{L_{qq,norm}} - 1\right), \qquad (24)$$

where we set $t_1 := t_a$ and use an upper bound of $t_2/t_1 \leq 4$ to prevent overestimation of the support region.

Within each elliptical support region $\mathcal{S}_{ell}$, the binary characteristic structure (figure 9.b) is used to determine the 'largest non-overlapping circular background regions'. This concept is illustrated by figure 9.c-e for the largest ellipse in figure 9.a. A Euclidean distance map (distance r from the characteristic structure) is built (figure 9.c), local maxima in r_i are detected (figure 9.d) and are used to draw circular background regions. In cases of overlaps, the smaller circles are eliminated, yielding the 'non-overlapping' map of n remaining circles (figure 9.e). Then, the areas $A_i = r_i^2 \pi$ of these background regions are used to suppress feature responses with interfering substructures. This suppression is performed by multiplying the significance measure $S(E)$ of equation (19) by

$$W = \left(\frac{C}{C + \frac{A_{opp}}{A_{\mathcal{S}_{ell}}}}\right)^m, \qquad (25)$$

where the area $A_{\mathcal{S}_{ell}}$ of the elliptical support region $\mathcal{S}_{ell}$ is related to a measure

$$A_{opp} = \sum_{i=1}^{n} A_i \qquad (26)$$

and $C = 0.5$ and $m = 2$ are constants. The parameter values occurring here were chosen to maximize the classification performance of the resulting size description features R_{det} and $A_{det}(\mathcal{S}_{ell})$ according to the degree of the carbide distribution. Figure 9.f shows the most significant structures after this verification procedure.

After the transformation (25) of the significance values $S(E)$, the k most significant scale-space extrema are selected. The largest of these extrema is the *most significant blob* and its *radius* R_{det} and the *area* of its elliptical support region $A_{det}(\mathcal{S}_{ell})$ are used as size description features for the classification of the carbide distribution according to the degree, i.e.

$$R_{det} = \sqrt{2\,t_{det}\,\log 2}; \quad \text{with} \quad t_{det} = \max_{i=1...k}(t_i), \qquad (27)$$

$$A_{det}(\mathcal{S}_{ell}) = \pi\,\kappa\,R_{det}^2 \quad \text{with} \quad \kappa = \sqrt{\frac{1}{2}\left(3\,\frac{L_{pp,norm}}{L_{qq,norm}} - 1\right)}, \qquad (28)$$

where $\kappa \leq 2$ according to the upper bound used in the support region estimation and $k = 10$ proved to be a reliable choice to detect the manually determined largest agglomerations.

15

3.3.2 Assessing the texture shape

To represent directional distributions, the second moment matrix is a useful texture descriptor (Bigün et al. 1991, Gårding & Lindeberg 1996, Lindeberg & Gårding 1993). Given a symmetric normalized window function w, the *windowed second moment matrix* can be defined by

$$\mu_L(\mathrm{q}) = \int_{\mathrm{x} \in \mathbb{R}^2} (\nabla L(\mathrm{x}))(\nabla L(\mathrm{x}))^T w(\mathrm{q} - \mathrm{x})\, d\mathrm{x}, \tag{29}$$

where $L : \mathbb{R}^2 \to \mathbb{R}$ denotes the image brightness and $\nabla L = (L_x, L_y)^T$ its gradient. Denoting the windowing operation by E_q, equation (29) can be written as

$$
\begin{aligned}
\mu_L(q) &= \begin{pmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{pmatrix} = E_q \begin{pmatrix} L_x^2 & L_x L_y \\ L_x L_y & L_y^2 \end{pmatrix} \\
&= E_q((\nabla L)(\nabla L)^T),
\end{aligned} \tag{30}
$$

and from the components of μ_L, the following descriptors can be defined

$$P = E_q(L_x^2 + L_y^2), \quad C = E_q(L_x^2 - L_y^2), \quad S = 2E_q(L_x L_y). \tag{31}$$

P is a measure for the strength of the operator response, C and S contain directional information, which can be summarized in two anisotropy measures

$$Q = \sqrt{C^2 + S^2}, \qquad \tilde{Q} = Q/P. \tag{32}$$

The *normalized anisotropy* $\tilde{Q} \in [0; 1]$ is zero, if and only if $E_q(L_x^2) = E_q(L_y^2)$ and $E_q(L_x L_y) = 0$ and $\tilde{Q} = 1$ if and only if $E_q(L_x L_y) = E_q(L_x^2)E_q(L_y^2)$. A rotationally symmetric gray-level pattern has $\tilde{Q} = 0$ and a translationally symmetric pattern has $\tilde{Q} = 1$.

When computing this descriptor in practice, the gradient vectors are calculated at *local scale* t_l and we use a Gaussian window function g with *integration scale* t_i (Lindeberg 1994, chap. 14) to define the *multi-scale windowed second moment matrix* μ_L as

$$\mu_L(\mathrm{q}; t_l, t_i) = \int_{\mathrm{x} \in \mathbb{R}^2} (\nabla L(\mathrm{x}))(\nabla L(\mathrm{x}))^T g(\mathrm{q} - \mathrm{x}; t_i)\, d\mathrm{x}. \tag{33}$$

The directional distribution (*type*) of the carbide distribution can be modeled by evaluating $\mu_L(\mathrm{q}; t_l, t_i)$ based on the scale information of the scale selection scheme described in section 3.3.1. The normalized anisotropy $\tilde{Q}$ is calculated for

$$t_l = \gamma_l * t_{det}, \tag{34}$$

where $\gamma_l = 0.5$ is chosen to maximize the classification performance of the resulting shape description feature $\tilde{Q}$ according to the *type*.

Assuming a globally valid significant scale for each image, the integration scale is set to the size of the image. Global values of P, C and S are computed by summing up the corresponding gradient expressions over the whole image and determining global values of Q and $\tilde{Q}$ afterwards. This yields a shape description feature $\tilde{Q}_{det}$ for classifying carbide distributions according to their *type* given by

$$\tilde{Q}_{det} = \tilde{Q}(\mu_L(x, y; t_l, \text{image size})). \tag{35}$$

16

3.4 Feature selection

To combine the size and shape information derived from the multi-scale techniques (sections 3.2 and 3.3) with the features modeling the manual assessment of the carbide distribution (section 3.1), we form a normalized feature vector of altogether 40 features. To evaluate the quality of each feature and to find the best feature subset, we applied suboptimal feature selection algorithms (Devijver & Kittler 1982, Kittler 1986, Fukunaga 1990).

One reason for the need for feature selection is the so-called *curse of dimensionality* or *peaking phenomenon*, which means the effect that the classification error for a finite number of training samples assumes a minimum for a certain optimal number of features and increases if further features are added (Raudys & Jain 1991, Devijver & Kittler 1982). Another advantage of feature selection is that the reduction of the dimensionality of a classification process improves the *generalization ability* of the designed classifier. In finite-sample-size situations, the parameter estimation errors increase rapidly with dimensionality and the classifier becomes finely tuned to the data used for learning. Therefore, feature selection improves the classification performance on unknown samples by trading off error probability for estimation errors. Additionally, the computational efficiency is improved by calculating only significant features.

As one would already expect from the results in (Pudil, Novovičová & Kittler 1994, Zongker & Jain 1996), the floating sequential algorithms yield better results on the present data than the simple sequential methods and the plus l take away r algorithm. The branch and bound algorithm is not applicable to the problem, because the criterion used for class separability does not guarantee the requested monotony and a brute-force calculation of optimal subsets is prohibited by the computational effort for this large number of features. Hence, we used sequential floating algorithms in our experiments. As a criterion of class separability, we measured the performance of the applied classifier. To get a good estimation of the classification performance, we used a rotation scheme with three disjunct sample subsets of all images. The samples of each class are evenly distributed over the three subsets. The feature selection is performed three times using two-thirds of the dataset and a leave-one-out estimate. Furthermore, a test performance is calculated on the remaining third of the dataset. The final classification performance values of both steps — the selection and the test classification — are the average values over the three sample subsets.

3.5 Classification approaches

Carbide distributions occur continuously distributed over the degree axis and the type axis of the standard chart (figure 2). Features reflecting the corresponding size and/or shape information of the carbide distribution naturally show a unimodal class distribution. Furthermore, non-parametric classification methods need more samples to yield good estimates of the probability distributions of the present problem than parametric ones. Thus, parametric classification algorithms are more appropriate to model the continuum of the appearances of carbide distributions. Additionally, the expert knowledge is directly introduced into the automatic classification process using a training set for parametric classifiers, where the true class is given by the

17

classifications of the metallographers (see section 2.3). Unfortunately, the application of the Mahalanobis distance classifier is prevented by the fact that class covariance matrices become singular for n_j samples in a class ω_j using a dataset with D features, if $n_j < D$ (Duin 1995, Hoffbeck & Landgrebe 1996). Obviously, this is the case for the present dataset. Thus, our experiments are mainly based on the minimum distance classifier.

In order to independently investigate the scale and the shape estimation of the carbide distributions, the classification is performed separately for the degree of the carbide distributions (*degree classification*) and according to the type (*type classification*). This is achieved by combining all reference classes from one column (row) of the standard chart to form one new class yielding 7 *degree classes* (4 *type classes*).

4   Results

In the following, we illustrate that the degree and type features described in chapter 3.2 and 3.3 model the scaled texture nature of the carbide distribution by capturing the size and shape information. Furthermore, the performance of the proposed classification scheme is described (section 3.5) using the best found feature subsets.

4.1   Scale and shape estimation

In our experiments we use a multi-channel representation with the radial frequencies $F_i = \sqrt{2}, 2\sqrt{2}, \cdots, 32\sqrt{2}$ cpi ($i = 0, \cdots, 5$), the orientations $\phi_j = jB_\phi (j = 0, \cdots, 17)$, the frequency bandwidth $B_f = 1$ octave and the orientation bandwidth $B_\phi = 10°$. These small orientation channels have to be used to capture the fine differences concerning the directional information of the carbide distributions. The highest frequency channels $F_6 = 64\sqrt{2}$ and $F_7 = 128\sqrt{2}$ cpi proved to contain no significant information of the carbide distribution. This filter configuration has been found by determining the best classification performance of the resulting degree and type features for $B_\phi = 5°, 10°$ and $20°$. Figure 10 shows all reference images with type 2 filtered with the Gabor channel corresponding to the dominant frequency given by E_{fm}. Obviously, the radial frequency values of the selected Gabor filter do approximate the scale of the dominant structure in the images.
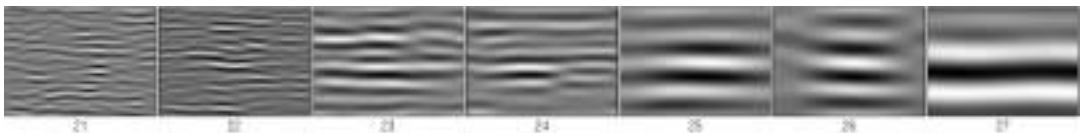

Figure 10: Filter response of Gabor filter corresponding to E_{fm} for all reference images with type 2.

The size information assessed by the degree feature $A_{det}(S_{ell})$ derived from automatic scale selection (section 3.3) is illustrated in figure 11. The elliptical support region corresponding to the area of the most significant blob $A_{det}(S_{ell})$ is shown for all reference images of type 4. One can clearly recognize that the scale of the largest structure in the carbide images is captured by this descriptor. The shape feature $\tilde{Q}_{det}$ is calculated using this scale information to define the pre-smoothing of the image.
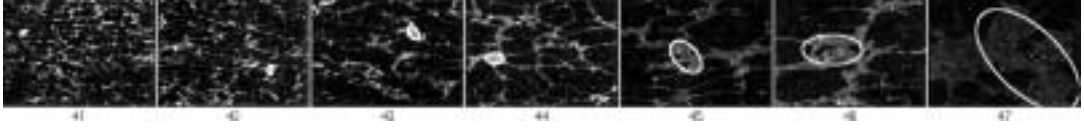
18

Figure 11: Elliptical support region corresponding to $A_{det}(S_{ell})$ for all reference images with type 4.

The selection of a local scale $t_l = 0.5t_{det}$ for the calculation of the fine normalized global anisotropy $\tilde{Q}_{det}(\gamma_l = 0.5)$ maximizes the classification performance according to the type of the carbide distribution.

The discriminative information of the scale features E_{fm}, $A_{det}(S_{ell})$ and the shape features $E_{om}, \tilde{Q}_{det}(\gamma_l = 0.5)$ is further illustrated in figure 12, where the mean and the standard deviation of the features from all samples of one class are shown for the 7 degree and the 4 type classes, respectively. The scale feature $A_{det}(S_{ell})$ has continuously increasing mean values with increasing degree and E_{fm} shows mean values with inverse tendency, except for degree 7. The shape features E_{om} and $\tilde{Q}_{det}(\gamma_l = 0.5)$ show decreasing mean values with increasing type, except for type 1. In general, the deviation of the scale features is smaller than the deviation of the shape features. In the case of E_{om}, this is mainly due to the fact that the finest carbide distributions have a wide distribution of the energy over all orientation channels irrespective of the type. The fact that almost no anisotropic (band-, or line-like) structures occur in the very fine-scale carbide structures (figure 2) results in a low anisotropy value, as for the net-shaped structures. This yields a large variance of the values of $\tilde{Q}_{det}(\gamma_l = 0.5)$ for type 1.
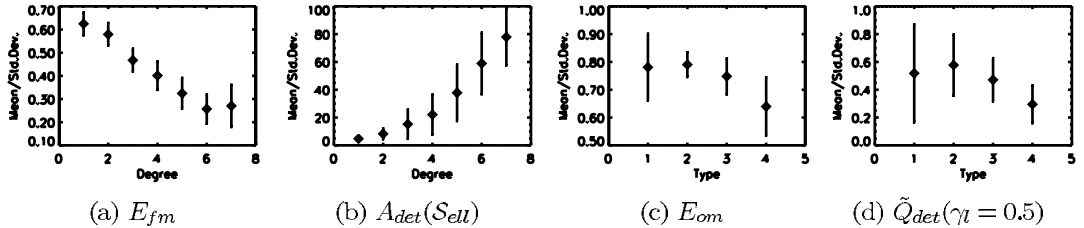

(a) E_{fm} (b) $A_{det}(S_{ell})$ (c) E_{om} (d) $\tilde{Q}_{det}(\gamma_l = 0.5)$

Figure 12: Discriminative power of the features (a) E_{fm}, (b) $A_{det}(S_{ell})$, (c) E_{om} and (d) $\tilde{Q}_{det}(\gamma_l = 0.5)$; for all 7 degree and 4 type classes, respectively, the mean and the standard deviation over all samples of the class is shown.

4.2 Classification performance

The results of the experiments using sequential floating selection for the degree and type classification are shown in figure 13, where the x-axis represents the size of the feature subsets and the y-axis shows the average performance values of the best found subsets over the three dataset partitions. The well-known effect that fewer features yields a better performance for finite sample sets (Raudys & Jain 1991) can be seen from the diagrams. The performance values achieved with the forward and the backward procedure are quite similar with slightly better results using the forward

selection. In general, type classification requires more features than degree classification to achieve optimal performance. The performance on the test data is about 5% lower than the performance during the selection for medium-sized subsets, whereas the performance values are approximately equal for small and for large subsets. On the one hand, this illustrates the leave-one-out estimate yields realistic values and on the other hand it shows that the selection leads to *overtraining* for medium-sized subsets, i.e. the selected features are fitted to the finite data at hand.

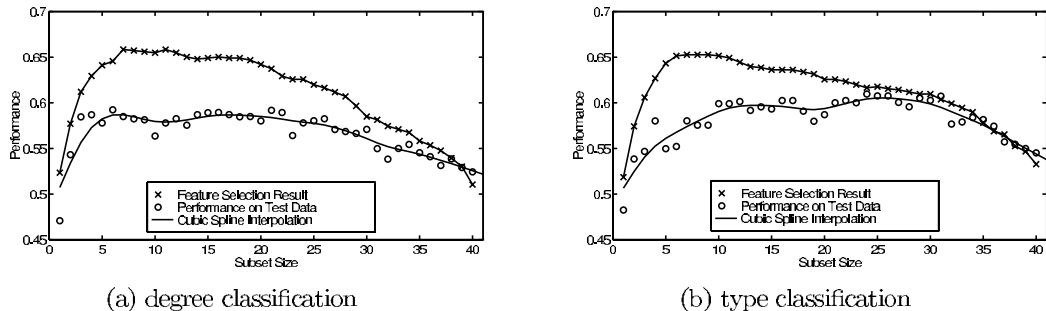(a) degree classification   (b) type classification

Figure 13: Results of sequential forward floating selection with a leave-one-out performance estimation of a minimum distance classifier for the degree (a) and type (b) classification; the performance values are averaged over three dataset partitions yielding a rotation estimate for the test data.

Due to the deviation between the classifications of the four metallographers, which the true class is based on (section 2.3), we also calculate the performance allowing a one-class-deviation. Table 2 summarizes the achieved best average performance values during selection and on the test data. Furthermore, the size values of the feature subsets yielding the highest test performance are given. As described in section 3.4, the Mahalanobis distance classifier is not applicable for the degree classification due to the limited sample size. But the number of samples for the 4 type classes is large enough for this classifier. The results of the sequential forward and backward floating selection for the type classification using a Mahalanobis distance classifier illustrate the potential improvement (approx. 10%) for the present dataset using this more adaptive classifier. We achieved an average performance during selection of almost 80% and a test performance of more than 70%.

classification	performance selection [%] (exact)	test [%] (exact)	subset size	test [%] tolerant
degree (minimum dist.)	65	60	4	95
type (minimum dist.)	65	61	10	92
type (Mahalanobis dist.)	79	72	8	96

Table 2: Average performance values of minimum distance classifier with best feature subsets found by sequential floating selection for degree and type classification - values for exact and tolerant classification (allowing a one-class-deviation) and the corresponding feature subsets size.

Due to the rotation estimate of the classification performance using 3 disjunct test sample sets, 3 feature selection experiments have been performed for each floating

algorithm as well. Using one selection algorithm for one classification approach yields different results for different datasets. In order to finally define a 'best' subset for the degree and type classification, we investigate the found feature sets of all 6 sequential forward and backward selection experiments concerning the occurrence frequency of the features. For this purpose the feature subset with size 4 for the degree and 10 for the type classification are used. We select the most frequent features, which occurred at least twice in these subsets, for the final 'best' feature subsets listed in table 3. In case of highly correlated features, only one of them is chosen. The resulting 'best' feature sets of size 5 and 10 for the degree and type classification yield the same performance values as for the best found subsets in the feature selection experiments. Obviously, the feature selection improves the classification runtime, because only at most 25% of the features must be calculated. Furthermore, the classification performance for these feature sets is about 5% higher than using all features.

feature name	degree	type
number of clusters	$\checkmark$	$\checkmark$
area of the largest cluster A_{max}	$\checkmark$	
smallest eigenvalue of the scatter matrix λ_2		$\checkmark$
eccentricity of the largest cluster $\mathcal{E}$		$\checkmark$
mean eccentricity of clusters		$\checkmark$
mean length of vertical segments	$\checkmark$	
number of non–empty horizontal lines in structure		$\checkmark$
path length of the skeleton		$\checkmark$
mean number of vertical intersections		$\checkmark$
degree of orientation (Gaussian)		$\checkmark$
radius of the most significant blob R_{det}	$\checkmark$	
fine normalized global anisotropy $\hat{Q}_{det}(\gamma_l = 0.5)$		$\checkmark$
degree estimate of maximum Gabor energy E_{fm}	$\checkmark$	
type estimate of maximum Gabor energy E_{om}		$\checkmark$

Table 3: List of all 14 features contained in the final 'best' feature subsets for the degree and type classification, where the presence in the subsets is marked by '$\checkmark$'.

5 Fully-automatic steel inspection

During the manual inspection of the carbide distribution of high speed steel, the whole surface of the steel specimen is taken into account. The metallographer determines the *most salient carbide distribution* at a lower magnification and classifies this focus-of-attention according to the standard chart at a magnification of 1:100. This technique is modeled by a multi-resolution approach by actively controlling the magnification of the microscope during the determination of the *focus-of-attention*. The approach follows the *purposive active vision* paradigm (Bajcsy 1988, Aloimonos 1990) by controlling imaging parameters such as the view direction, the focus and the illumination and by classifying only the representative most salient carbide distribution.

The automatic detection of the most salient carbide distribution starts scanning the whole specimen surface at a magnification of 1:20 (figure 14.a) with an overlap of 10%, which corresponds to half the image size at a magnification of 1:100, to prevent

21

(a) Carbide distribution (1:20)   (b) Detected representative area (1:50)   (c) Detected most salient carbide distribution (1:100)
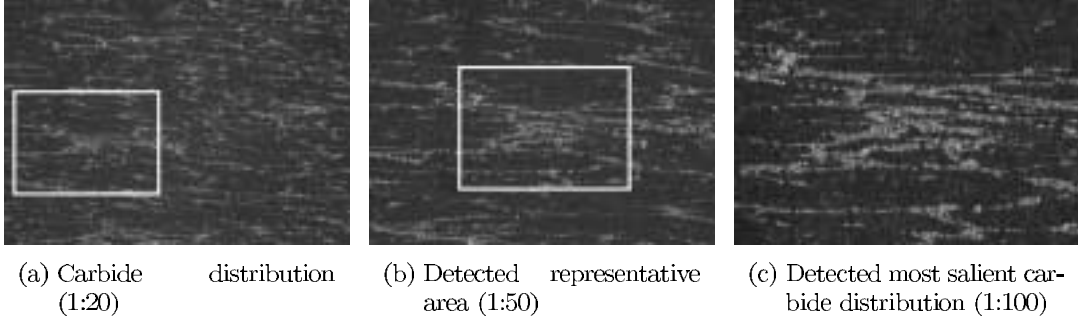
Figure 14: Illustration of the focus-of-attention scheme to actively detect and zoom to the most salient carbide distribution for classification on the steel specimen surface. The white frames mark the detected representative areas, which are visible at the next higher magnification

the system from missing the most salient area. After moving the specimen stage to this position and raising the magnification to 1:50, the illumination and the focus are adjusted to guarantee well defined imaging conditions. Then, the most salient area is tracked by the same detection to ensure a more precise localization of this area (figure 14.b), which is finally classified at 1:100 (figure 14.c). The most salient carbide distribution is detected by determining the largest agglomeration using the morphological operations described in section 3.1.1. This focus-of-attention algorithm leads to a selective evaluation of a specific area of the specimen surface, which is 16 times faster than taking images side by side (see also (Wiltschi & Pinz 1996)).

In order to automatically evaluate the carbide distribution of a high speed steel specimen using this active vision scheme, we have set up a prototype *active inspection system*. The system consists of a light microscope with a xyz-controllable specimen stage, a switchable objective revolver, a controllable light source, a graphics workstation, a CCD-camera, and a control device connected to the workstation (figure 15.a). Figure 15.b depicts a flow chart of the complete system, showing the main algorithmic modules of the integrated inspection technique.

6 Quantitatively defined reference chart

To overcome the subjectivity of the qualitative definition of the currently used standard chart, we propose a new reference chart based on the most significant features found by the feature selection experiments (table 3). To construct a new quantitatively defined reference chart of representative carbide distribution images in consideration of the expert knowledge, the following scheme is applied to the present carbide image dataset: Using the 'best' feature subset for the classification according to the 28 reference classes of the standard chart, the most representative image within each of the 28 classes is determined. This *most representative image* of class ω_j is represented by a feature vector $\vec{x}_r^j$, which has the minimal Euclidean distance to the class mean vector μ_j, i.e

$$\delta_E^2(\vec{x}_r^j, \mu_j) = \min_{i=0,\cdots,n_j} \delta_E^2(\vec{x}_i^j, \mu_j), \tag{36}$$

22

Active Detection of Most Salient Carbide Distribution
• Segmentation of largest carbide agglomeration
• Active Zoom to detected area

Segmentation of Characteristic Structure
• Automatic global thresholding
• Morphological operations

Spatial Feature Extraction
• Shape parameters
• Skeleton extraction
• Linear projections

Multi–channel Texture analysis
• Gabor filtering
• Energy–based feature extraction

Automatic Scale Selection in Linear Scale–Space
• Scale estimation by detection of scale–space maxima
• Scale–tuned shape estimation

Supervised Statistical Parametric Classification
=> Carbide Classification for large steel surface area

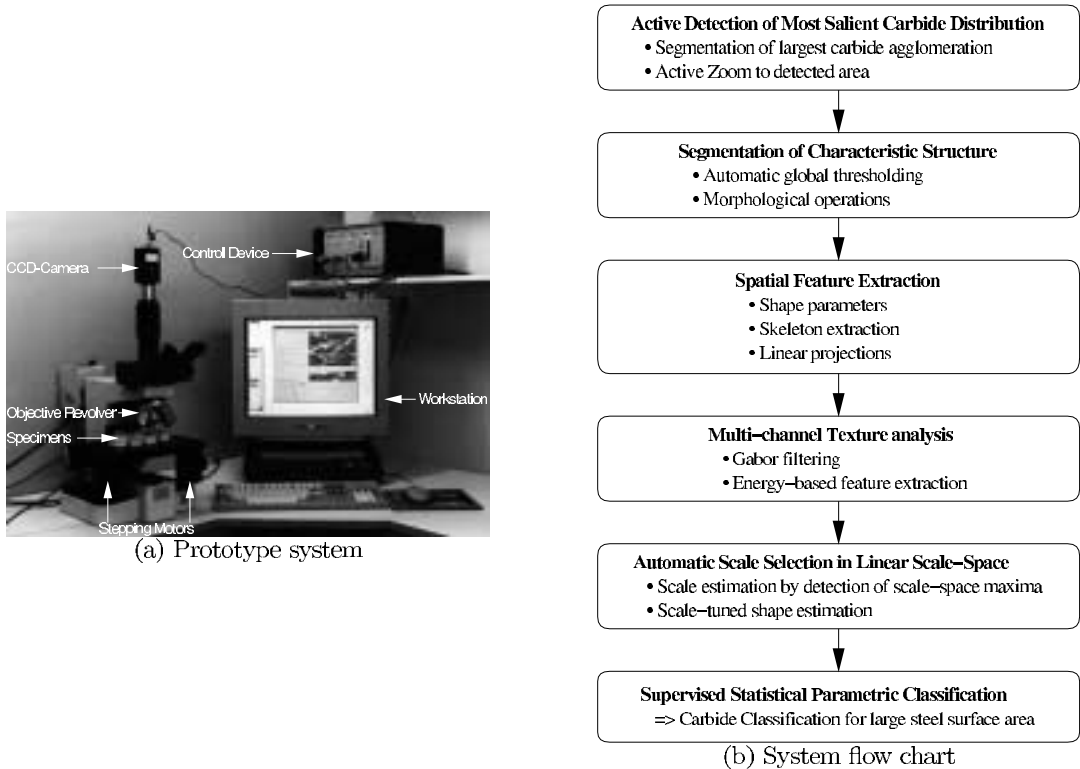(a) Prototype system

(b) System flow chart

Figure 15: Active inspection system. (a) Prototype system setup and (b) flow chart of main algorithmic modules.

where n_j denotes the number of samples in class ω_j and δ_E^2 stands for the *squared Euclidean distance*. This scheme yields a new objectively determined reference chart depicted in figure 16. In contrary to the original standard chart (figure 2) this set of representative images is defined by quantitative features with regard to the expert knowledge introduced by the manual classification.

7 Summary and conclusions

We have presented a scheme for computing scale and shape properties of textures and applied it to steel quality inspection. Besides a data-specific method to extract a characteristic structure of carbide distributions (section 3.1), which corresponds to the manually assessed information, two modules capturing the scale and the shape information of textures are described, based on multi-scale techniques.

A frequency- and orientation-selective multi-channel analysis using Gabor filters (section 3.2) is used to derive a *degree estimate of maximum Gabor energy E_{fm}*, which estimates the global dominant scale based on the energy distribution over the channels. Using the information about the global directional distribution contained in the energy distribution, a scale-adapted *type estimate of maximum Gabor energy E_{om}* is calculated. This feature estimates the shape of the texture relating the amount of horizontal to non-horizontal structures captured by the three most dominant fre-
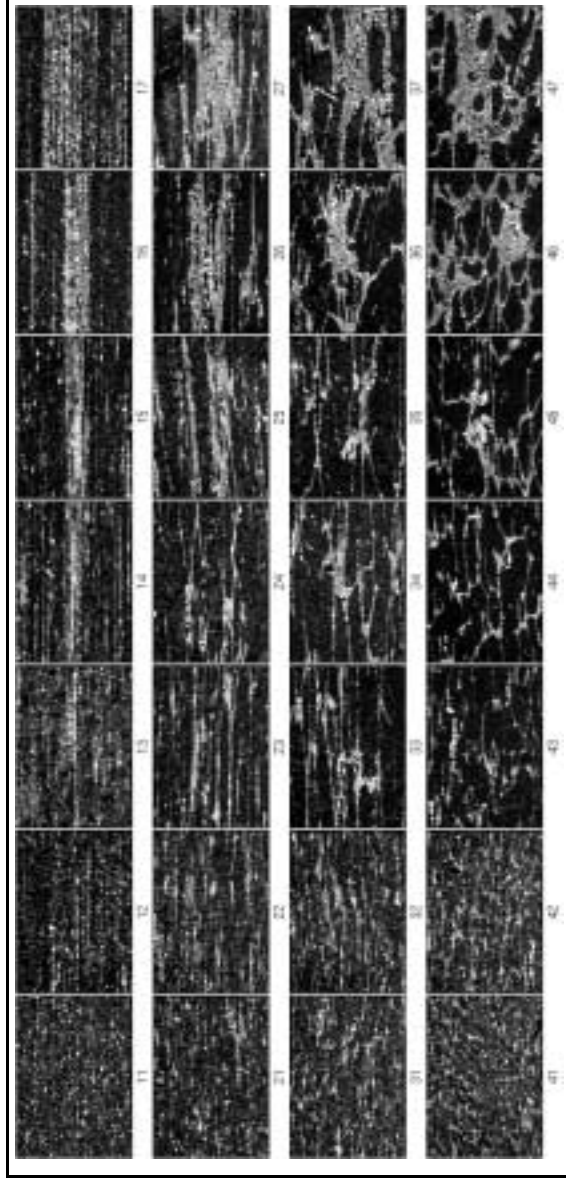
Figure 16: A quantitatively defined standard chart consisting of the 28 most representative images of the present dataset; for each reference class the image with the lowest Euclidean distance to the class mean vector is selected.

quency channels, which can be seen as a coarse scale adaptation. The features E_{fm} and E_{om} are based on the global significance of structures in the image.

On the other hand, the scale selection scheme described in section 3.3 represents a well-founded scheme to detect locally significant structures of predefined shape with respect to space and scale. Furthermore, the scale is directly related to the parameters characterizing the extensions of the structure. Therefore, the *area of the most significant blob* $A_{det}(\mathcal{S}_{ell})$ represents the extensions of the locally most significant largest structure. The multi-scale windowed second moment matrix defines a scale-space representation, which allows us to tune the local scale and the integration scale according to the characteristics of an investigated texture. The gradient expressions of the matrix only reflect variations of a size corresponding to the local scale t_l. Therefore, the *fine normalized global anisotropy* $\tilde{Q}_{det}(\gamma_l = 0.5)$ is a measure for the global ratio of translational and rotational structures of scale t_l. In the case of carbide distributions, the natural correspondence of the definition of E_{om} to the manual type classification scheme compared to $\tilde{Q}$ and the restriction to a single scale for the calculation of $\tilde{Q}$ leads to a better type classification performance of E_{om}. Nevertheless, the Gabor and scale-space features show a relatively high correlation of approx. 75% and an almost equal degree classification performance. This is mainly due to the fact that in the carbide distribution images, the locally most significant largest structure generally determines the global dominant scale as well. Thus, both approaches are convenient for the present application.

We have set up a prototype *active inspection system* including the presented classification scheme for carbide distributions and the focus-of-attention module (section 5) to detect the most salient carbide distribution. The system is capable of classifying the carbide distribution of a whole steel specimen and demonstrates the feasibility of fully-automatic steel specimen inspection concerning their carbide distribution based on the presented scale and shape estimation. The proposed automatic inspection scheme possesses a potential applicability to several related quality control tasks. Especially, in the steel producing industry, several inspection problems are based on a classification with regard to standard charts showing structures with a scaled texture nature.

Currently, the steel producing company Böhler Edelstahl is implementing the presented inspection technique on a system, which will be used in routine quality control in the near future.

Acknowledgments

References

Aloimonos, J. (1990), Purposive and Qualitative Active Vision, *in* 'Proc. DARPA Image Understanding Workshop', pp. 816–828.

Bajcsy, R. (1988), 'Active perception', *Proceedings of the IEEE* **76**(8), 996–1005.

Beiss, P. & Wähling, R. (1989), *Metal Powder Report* p. 739.

Bigün, J., Granlund, G. H. & Wiklund, J. (1991), 'Multidimensional orientation estimation with applications to texture analysis and optical flow', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**(8), 775–790.

Bovik, A. C. (1991), 'Analysis of mutlichannel narrow-band filters for image texture segmentation', *IEEE Transactions on Signal Processing* **39**(9), 2025–2043.

Bovik, A. C., Clark, M. & Geisler, W. S. (1990), 'Multichannel texture analysis using localized spatial filters', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(1), 55–73.

Chermant, J.-L., Coster, M. & Gougedon, G. (1989), 'Untersuchung der Morphologie von Al-Si-Legierungen mit Hilfe der vollautomatischen Bildanalyse', *Praktische Metallographie* **26**(8), 415–27.

Devijver, P. A. & Kittler, J. (1982), *Pattern Recognition: A statistical approach*, Pretice/Hall International.

Duin, R. P. W. (1995), Small sample size generalization, *in* 'The 9th Scandinavian Conference on Image Analysis', Uppsala, Sweden, pp. 957–964.

Eberly, D., Gardner, R., Morse, B., Pizer, S. & Scharlach, C. (1994), 'Ridges for image analysis', *Journal of Mathematical Imaging and Vision* **4**(4), 353–373.

Exner, H. E. & Hougardy, H. P. (1990), *Quantitative Image Analysis of Microstructure*, DGM-Informationsges., Oberursel.

Fischmeister, H., Paul, J. & Karagöz, S. (1988), 'Toughness relevant defects in high speed cutting steels', *Prakt. Metallographie* **25**, 28–40.

Florack, L. M. J. (1997), *Image Structure*, Vol. 10 of *Computational Imaging and Vision Series*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Florack, L. M. J., ter Haar Romeny, B. M., Koenderink, J. J. & Viergever, M. A. (1992), 'Scale and the differential structure of images', *Image and Vision Computing* **10**(6), 376–388.

Fukunaga, H. (1990), *Introduction to Statistical Pattern Recognition*, Academic Press, Boston.

Gårding, J. & Lindeberg, T. (1996), 'Direct computation of shape cues using scale-adapted spatial derivative operators', *International Journal of Computer Vision* **17**(2), 163–191.

Gonzalez, R. C. & Woods, R. E. (1992), *Digital Image Processing*, Addison-Wesley, Massachusetts.

Haralick, R. M. (1983), 'Ridges and valleys in digital images', *Computer Vision, Graphics, and Image Processing* **22**, 28–38.

Haralick, R. M. & Shapiro, L. G., eds (1992), *Computer and Robot Vision*, Vol. 1, Addison-Wesley.

Hoffbeck, J. P. & Landgrebe, D. A. (1996), 'Covariance matrix estimation and classification with limited training data', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18**(7), 763–767.

Jain, A. K. & Bhattacharjee, S. (1992), 'Text Segmentation Using Gabor Filters for Automatic Document Processing', *Machine Vision and Applications* **5**, 169–84.

Jain, A. K. & Farrokhnia, F. (1991), 'Unsupervised Texture Segmentation Using Gabor Filters', *Pattern Recognition* **24**(12), 1167–86.

Kittler, J. (1986), *Handbook of Pattern Recognition and Image Processing*, Academic Press, New York, chapter Feature selection and extraction, pp. 60–83.

Koenderink, J. J. (1984), 'The structure of images', *Biol. Cybern.* **50**, 363–370.

Koenderink, J. J. & van Doorn, A. J. (1994), 'Two-plus-one-dimensional differential geometry', *Pattern Recognition Letters* **15**(5), 439–444.

Lindeberg, T. (1993), On scale selection for differential operators, *in* K. A. Hogdra, B. Braathen & K. Heia, eds, 'Proc. 8th Scandinavian Conference on Image Analysis', Norwegian Society for Image Processing and Pattern Recognition, Norway, pp. 857–866.

Lindeberg, T. (1994), *Scale-Space Theory in Computer Vision*, Kluwer Academic Publishers.

Lindeberg, T. (1996*a*), Edge detection and ridge detection with automatic scale selection, *in* 'Proc. CVPR'96', San Francisco, California, pp. 465–470. International Journal of Computer Vision **30** (2) (Extended version to appear).

Lindeberg, T. (1996*b*), Feature detection with automatic scale selection, Technical Report ISRN KTH/NA/P-96/18-SE, Dept. of Numerical Analysis, KTH. International Journal of Computer Vision **30** (2) (in press).

Lindeberg, T. & Gårding, J. (1993), Shape from texture from a multi-scale perspective, *in* H.-H. Nagel, ed., 'Proc. 4th Int. Conf. on Computer Vision', Berlin, Germany, pp. 683–691.

Pavlidis, T. (1982), *Algorithms for Grahpics and Image Processing*, Computer Science Press, Rockville.

Pudil, P., Novovičová, J. & Kittler, J. (1994), 'Floating search methods in feature selection', *Pattern Recognition Letters* **15**, 1119–1125.

Raudys, S. J. & Jain, A. K. (1991), 'Small sample size effects in statistical pattern recognition: Recommendations for practitioners', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**(3), 252–264.

Russ, J. C. (1991), 'Computer-Aided Quantitative Microscopy', *Materials Characterization* **27**(4), 185–97.

Serra, J. (1982), *Image Analysis and Mathematical Morphology*, Vol. I, Academic Press.

Serra, J. (1988), *Image Analysis and Mathematical Morphology*, Vol. II, Academic Press.

Wiltschi, K. (1998), An Active Inspection System for the Assessment of Steel Quality, PhD thesis, Technical University Graz, Austria.

Wiltschi, K., Lindeberg, T. & Pinz, A. (1997*a*), Classification of carbide distributions using scale selection and directional information, *in* 'Proc. of ICIP 1997', Vol. III, IEEE, Santa Barbara, Californien, pp. 122–125.

Wiltschi, K., Lindeberg, T. & Pinz, A. (1997*b*), Classification of carbide distributions using scale selection and directional information, Technical Report ISRN KTH/NA/P–97/10–SE, Computational Vision and Active Perception Laboratory, KTH, Sweden.

Wiltschi, K. & Pinz, A. (1996), An active inspection system for the assessment of steel quality, *in* A. Pinz & W. Pölzleitner, eds, 'Machine Perception Applications - Proceedings of the IAPR-TC8 Workshop', Vol. 93, R. Oldenburg Wien München, pp. 127–136.

27

Wiltschi, K., Pinz, A. & Hackl, G. (1995), Feature extraction for carbide classification of high speed steel, *in* F. Solina & W. G. Kropatsch, eds, 'Proceedings of 19th ÖAGM and 1st SDVR Workshop - Visual Modules', Vol. 81, Österreichische Computer Gesellschaft, R. Oldenburg Wien München, pp. 169–176.

Wiltschi, K., Pinz, A. & Hackl, G. (1996), Image analysis of carbide distributions, *in* I. H. H. Berns, H. Hinz, ed., 'Progress in Tool Steels - Proceedings of 4th International Tooling Conference', Schürmann + Klagges KG, Bochum, pp. 213–222.

Witkin, A. P. (1983), Scale-space filtering, *in* 'Proc. 8th Int. Joint Conf. Art. Intell.', Karlsruhe, West Germany, pp. 1019–1022.

Zongker, D. & Jain, A. (1996), Algorithms for feature selection: An evaluation, *in* 'Proceedings of ICPR '96', pp. 18–22.