

# Verfahrensvergleich zur automatischen Erkennung in Metall geschlagener Zeichen

Zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

von der Fakultät für  
Elektrotechnik  
der Universität Fridericiana Karlsruhe

genehmigte

DISSERTATION

von

Dipl.-Ing. Uwe Stilla

aus Köln

Tag der mündlichen Prüfung : 22. Juni 1993  
Hauptreferent : Prof. Dr.-Ing. H. Kazmierczak  
Korreferent : Prof. Dr.-Ing. K. Meyer-Waarden



## Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter im *Institut für Biokybernetik und Biomedizinische Technik (IBB)* der Universität Karlsruhe unter Betreuung von Herrn Prof. Dr.-Ing. K. Meyer-Waarden und im *Forschungsinstitut für Informationsverarbeitung und Mustererkennung (FIM)* in Ettlingen unter Betreuung von Herrn Prof. Dr.-Ing. H. Kazmierczak. Die aus dem Bereich der Automatisierungstechnik entlehnte Thematik der Arbeit, entstand im Rahmen einer Kooperation des *IBB* mit der *Siemens AG*.

Herrn Prof. Dr.-Ing. H. Kazmierczak danke ich für die wissenschaftliche Betreuung der Arbeit, die Übernahme des Hauptreferates und die hervorragenden Arbeitsmöglichkeiten.

Herrn Prof. Dr.-Ing. K. Meyer-Waarden danke ich für die Unterstützung der Arbeit und die Übernahme des Korreferates.

Mein Dank gilt auch Herrn Dr. Schorb von der *Siemens AG*, der mir bei der Auswahl und Aufnahme von Datenmaterial mit seinen Erfahrungen zur Seite stand, Herrn Dipl.-Ing. K. Lütjen vom *FIM*, der mich durch seine anregenden und konstruktiven Diskussionen bereichert hat und Frau B. Euscher für die Anfertigung der Zeichnungen.

Weiterhin gilt mein Dank auch allen Studenten des *IBB* und Kollegen des *FIM*, die auf unterschiedliche Weise zum Gelingen der Arbeit beigetragen haben.

Herzlich bedanke ich mich für die private Unterstützung von Frau G. Baumgartner, Herrn P. Ambros und Frau A. Schlinkert. Nicht zuletzt möchte ich mich auch bei meinen Eltern bedanken, die mich auf meinem Ausbildungsweg stets gefördert haben.

## Zusammenfassung

Das Erscheinungsbild von in Metall geschlagen Zeichen kann in Abhängigkeit von der Beleuchtung stark variieren. Eine weitgehend beleuchtungsunabhängige Erkennung solcher Zeichen erfordert ein robustes Verhalten des Erkennungssystems. In dieser Arbeit werden drei Verfahren vorgestellt, mit denen eine Erkennung von Schlagziffern möglich ist.

Beim Verfahren I wird ein rein numerischer Ansatz verfolgt, bei dem der Mustervergleich auf der Korrelation mit Referenzmasken basiert. Beim Verfahren II wird ein symbolisch-numerischer Ansatz verfolgt, bei dem eine Liniendarstellung extrahiert und mit Referenzliniendarstellungen verglichen wird. Beim Verfahren III wird ein syntaktischer Ansatz verfolgt, bei dem Primitivobjekte extrahiert und modellgestützt Referenzstrukturen generiert werden.

In industriellen Anwendungen ist neben den Störungen durch Beleuchtungsvariationen auch mit ungenauer Positionierung und Ausrichtung von Werkstücken zu rechnen. Für die vorgestellten Verfahren wird die Empfindlichkeit bezüglich einer Verdrehung der Mustervorlage untersucht und an Beispielen dargestellt.

Der Operationsaufwand der Verfahren wird untersucht und an einigen Berechnungsbeispielen gegenübergestellt. Abschließend werden die Ergebnisse diskutiert und die Verfahren bewertet.

# Inhaltsverzeichnis

<b>Vorwort</b>	<b>i</b>
<b>Zusammenfassung</b>	<b>ii</b>
<b>Inhaltsverzeichnis</b>	<b>iii</b>
<b>Verzeichnis der Abbildungen</b>	<b>vi</b>
<b>Verzeichnis der Tabellen</b>	<b>ix</b>
<b>Formelzeichen und Abkürzungen</b>	<b>x</b>
<b>1. Einleitung</b>	<b>1</b>
1.1 Informationsverarbeitung durch das visuelle System . . . . .	2
1.2 Zielsetzung der Arbeit . . . . .	5
1.3 Übersicht . . . . .	6
<b>2. Experimentelle Versuchsbedingungen</b>	<b>7</b>
2.1 Bildaufnahme . . . . .	7
2.1.1 Zeichenvorlage . . . . .	8
2.1.2 Beleuchtung . . . . .	9
2.2 Beschreibung der Bildaufnahme durch ein Modell (Bildsynthese) . . . . .	10
2.3 Aufgenommenes Datenmaterial . . . . .	12
<b>3. Grundlagen</b>	<b>15</b>
3.1 Numerische Klassifikation . . . . .	15
3.2 Syntaktische Klassifikation . . . . .	17
3.3 Veröffentlichte Beiträge zur Erkennung von Schlagzeichen . . . . .	19
3.3.1 Objektangepaßte Filterung zur Konturfindung und Klassifikation . . . . .	19
3.3.2 Gradienten-Matching mit dünnen Referenzmustern . . . . .	20

3.3.3	Erkennung mit einem verteilten assoziativen Speicher . . . . .	20
3.4	Liniendarstellungen von Zeichen . . . . .	22
3.4.1	Linienfindung durch Segmentierung des Grauwertbildes . . . . .	22
3.4.2	Linienfindung durch differenzierende Maskenoperation . . . . .	25
3.4.3	Problem bei der Linienfindung . . . . .	28
3.5	Normierte maskierte Korrelation . . . . .	30
<b>4.</b>	<b>Verfahren zur Zeichenerkennung</b>	<b>33</b>
4.1	Ein numerisches Verfahren zur Zeichenerkennung (Verfahren I) . . . . .	33
4.1.1	Generierung der Referenzen . . . . .	34
4.1.2	Korrelation . . . . .	35
4.1.3	Suche der Korrelationsspitzen . . . . .	36
4.1.4	Klassifikation . . . . .	38
4.1.5	Beispiel . . . . .	39
4.2	Der elastische Linienvergleich (Verfahren II) . . . . .	41
4.2.1	Bestimmung der Distanzmatrizen . . . . .	43
4.2.2	Vergleich der Liniendarstellungen . . . . .	45
4.2.3	Beispiel . . . . .	47
4.2.4	Konturpunktdetektion . . . . .	47
4.3	Ein syntaktisches Verfahren zur Zeichenerkennung (Verfahren III) . . . . .	55
4.3.1	Zeichenerkennung mit dem Produktionssystem BPI . . . . .	56
4.3.2	Konzept des Schriftzeichens . . . . .	62
4.3.3	Generierung der Primitivobjekte . . . . .	63
4.3.4	Beispiel . . . . .	63
4.4	Ergebnisse der Verfahren I, II und III für Bildbeispiele . . . . .	67
4.5	Empfindlichkeit der Verfahren I, II und III gegenüber Verdrehung . . . . .	68
<b>5.</b>	<b>Einschränkung des Suchbereiches</b>	<b>75</b>
5.1	Segmentierung einer Schriftzeile . . . . .	75
5.2	Segmentierung einzelner Zeichen einer Schriftzeile . . . . .	77

<b>6. Abschätzung des Berechnungsaufwands</b>	<b>79</b>
6.1 Definitionen . . . . .	79
6.2 Berechnungsaufwand zur Korrelation . . . . .	81
6.2.1 Berechnung des Korrelationskoeffizienten im Ortsbereich . . . . .	82
6.2.2 Berechnung des Korrelationskoeffizienten im Ortsfrequenzbereich (FFT) . . . . .	84
6.3 Berechnungsaufwand zum Linienvergleich . . . . .	86
6.4 Berechnungsaufwand zur Einschränkung des Suchbereiches . . . . .	92
6.4.1 Einschränkung des Suchbereiches auf eine Schriftzeile . . . . .	92
6.4.2 Berechnungsaufwand zur Vereinzelung von Zeichen . . . . .	92
6.5 Bestimmung des relativen Operationsaufwandes . . . . .	93
6.6 Berechnungsbeispiele . . . . .	94
<b>7. Diskussion</b>	<b>101</b>
<b>Literaturverzeichnis</b>	<b>105</b>
<b>A. Anhang</b>	<b>113</b>

## Verzeichnis der Abbildungen

2.1	Schematische Darstellung der Bildaufnahme . . . . .	7
2.2	Oberflächenverformung durch einen Keil . . . . .	9
2.3	Erscheinungsbilder der Ziffer fünf bei unterschiedlichen Beleuchtungssituationen . . . . .	12
2.4	Grauwertgebirge der Ziffer fünf . . . . .	13
2.5	Grauwertbild mit geringem Kontrast . . . . .	14
2.6	Grauwertbilder mit dem Grauwertprofil einer Bildspalte . . . . .	14
3.1	Beispiel zur Überführung eines ikonischen Bildes in seine Konturlinie . . .	22
3.2	Konturlinien bei unterschiedlichen Schwellwerten . . . . .	23
3.3	Beispiel zur Überführung eines ikonischen Bildes in seine Skelettlinie . . .	23
3.4	Gegenüberstellung der Ergebnisse verschiedener Verdünnungsoperationen	24
3.5	Gegenüberstellung von verdünnten Schriftzeichen bei unterschiedlichen Segmentierungsergebnissen . . . . .	24
3.6	Beispiel zur Bestimmung der Konturlinie mit dem Sobel-Operator . . . .	26
3.7	Beispiel zur Bestimmung der Konturlinie durch DOG-Operation und Nullstellensuche . . . . .	27
3.8	Konturlinien (Zero-crossings) bei unterschiedlicher Wahl der Filtermasken	27
3.9	Problematik bei der Segmentation durch Schwellwertoperation auf das Grauwertbild . . . . .	28
3.10	Problematik bei der Segmentation durch Schwellwertoperation auf das Gradientenbild . . . . .	28
3.11	Problematik bei der Bestimmung von Konturlinien (Zero-crossings) bei unterschiedlicher Wahl der Filtermasken . . . . .	29
4.1	Übersicht zu den implementierten Verfahren zur Zeichenerkennung . . . .	33
4.2	Struktur des Erkennungssystems (Verfahren I) . . . . .	34
4.3	Generierung der Referenzdarstellung (Template) für die Ziffer fünf . . . .	35
4.4	Korrelation eines Kreisringes mit der entsprechenden Referenz . . . . .	36

4.5	Definition der Spitzenumgebung . . . . .	37
4.6	Korrelation eines strichförmigen Musters mit der entsprechenden Referenz	38
4.7	Korrelation einer Schriftzeile (zehn Ziffern) mit dem Template fünf . . . .	39
4.8	Korrelation der Ziffer fünf mit zehn Templates . . . . .	40
4.9	Gegenüberstellung der Korrelationsmaxima . . . . .	40
4.10	Struktur des Erkennungssystems (Verfahren II) . . . . .	41
4.11	Beispiele für Distanzmatrizen von Liniendarstellungen . . . . .	42
4.12	Schrittweise Überführung der ikonischen Darstellung in eine Liniendarstellung . . . . .	48
4.13	Liniendarstellung und Distanzmatrizen . . . . .	50
4.14	Linienvergleich von verschiedenen Mustern einer Schriftzeile mit einer Referenz . . . . .	51
4.15	Ergebnis der Linienvergleiche bezüglich der Referenzen . . . . .	52
4.16	Gegenüberstellung der Größen $k_R$ . . . . .	52
4.17	Ergebnis der Linienvergleiche bezüglich der Muster . . . . .	53
4.18	Beispiele der Punkt-zu-Punkt-Zuordnungen . . . . .	54
4.19	Struktur des Erkennungssystems (Verfahren III) . . . . .	55
4.20	Beispiele zur Aufteilung der Ziffer fünf in Liniensegmente . . . . .	56
4.21	Schematischer Aufbau des Assoziativspeichers . . . . .	58
4.22	Übersicht zum Datenfluß im Blackboard-System . . . . .	60
4.23	Aufbau von Objekten aus Primitivobjekten und Teilobjekten . . . . .	61
4.24	Konzeptgraph der Ziffer fünf . . . . .	63
4.25	Maskenaufteilung am Beispiel der Ziffer fünf . . . . .	64
4.26	Korrelation der Ziffer fünf mit Teilmasken . . . . .	65
4.27	Ableitungsgraph des Zielobjektes Z5 . . . . .	65
4.28	Symbolische Darstellung der Teilobjekte in einer Schriftzeile . . . . .	66
4.29	Symbolische Darstellung der Teilobjekte beim Muster fünf . . . . .	67
4.30	Korrelation desselben Musters mit verschiedenen Referenzen (Vollmasken)	69
4.31	Korrelation verschiedener Muster mit derselben Referenz (Vollmasken) . .	70

4.32	Linienvergleich desselben Musters mit verschiedenen Referenzen . . . . .	71
4.33	Linienvergleich verschiedener Muster mit derselben Referenz . . . . .	72
4.34	Korrelation desselben Musters mit verschiedenen Referenzen (Teilmasken)	73
4.35	Korrelation verschiedener Muster mit derselben Referenz (Teilmasken) . .	74
5.1	Segmentierungsstufen zur Einschränkung des Suchbereichs . . . . .	75
5.2	Segmentierung einer Schriftzeile . . . . .	76
5.3	Segmentierung einzelner Zeichen . . . . .	78
6.1	Definition der Bildbereiche . . . . .	80
6.2	Darstellung der Suchbereiche $S_1$ bis $S_5$ im Bildbereich B . . . . .	95
6.3	Graphische Darstellung der Aufwandswerte . . . . .	100
A.1	Beispiel zum Erscheinungsbild geprägter Zeichen . . . . .	113
A.2	Beispiel zum Erscheinungsbild gefräster Zeichen . . . . .	113
A.3	Beispiel zum Erscheinungsbild von Hand geschlagener Zeichen . . . . .	114
A.4	Verteilung des mittleren Grauwertes der Bildspalten . . . . .	114
A.5	Konzeptgraphen der Ziffern eins bis null . . . . .	117
A.6	Verwendete Teilmasken . . . . .	118
A.7	Variationen der Teilmaske ES5 mit resultierenden Korrelationsfeldern . .	119

## Verzeichnis der Tabellen

4.1	Kompatibilitätsmatrix der Attribute $a_G$ und $a_R$ . . . . .	49
4.2	Tabelle der Produktionen zur Synthese der Ziffer fünf . . . . .	63
4.3	Zusammenstellung der erkannten Muster . . . . .	67
6.1	Größe der Bildbereiche . . . . .	80
6.2	Operationsaufwand für Korrelation einer n-wertigen Referenz mittels Maskenoperation . . . . .	83
6.3	Operationsaufwand für Korrelation einer zweiwertigen Referenz mittels Maskenoperation . . . . .	84
6.4	Operationsaufwand für Korrelation einer n-wertigen Referenz mittels FFT	86
6.5	Typische Anzahl der Konturpunkte einer Referenz . . . . .	86
6.6	Operationsaufwand für den Linienvergleich . . . . .	91
6.7	Zusammenstellung der geometrischen Größen der Suchbereiche . . . . .	95
6.8	Größe des Positionierungsbereiches und Zahl der Korrelationswerte . . . .	96
6.9	Operationsaufwand zur Einschränkung des Suchbereiches . . . . .	97
6.10	Operationsaufwand . . . . .	97
6.11	Operationsaufwand für Verfahren II . . . . .	98
6.12	Operationsaufwand für Verfahren I, II, III . . . . .	99
7.1	Bewertung der Verfahren I, II, III . . . . .	102

## Formelzeichen und Abkürzungen

$A$	Attributmenge
$A_w$	Fensterfläche
$A_{\oplus}, A_{\ominus}, A_{\circ}$	Teilflächen der Referenz (Template)
$B$	Grauwertbild
$C$	Zeiger
$D$	Distanzmatrix
$E$	Grammatik
$G$	Teilbild des Grauwertbildes $B$ , Überdeckungsbereich
$G_{ij}$	Überdeckungsbereich in $B$ durch $R$ an der Position $(i,j)$
$H_y$	Sobel Operator Maske
$I$	Intensität
$K_k$	Korrelationsfeld
$L, L_G, L_R$	Liniendarstellung, - des Musters, - der Referenz
$M_{i,j}$	Menge der Punkte von $G_{ij}$
$M_r$	Menge der Punkte von $R$
$O_{Appr}$	Operationsaufw. zur Konturapproximation
$O_{Comp}$	Operationsaufw. zum Linienvergleich
$O_{Hara}$	Operationsaufw. zur Konturpunktdetektion (Haralik)
$O_{K1}$	Operationsaufw. der Maskenkorrelation mit $n$ -wertiger Referenz
$O_{K2}$	Operationsaufw. der Maskenkorrelation mit 2-wertiger Referenz
$O_{K3}$	Operationsaufw. der Maskenkorrelation mittels FFT
$O_M$	Operationsaufw. für den Linienvergleich
$O_{Ske}$	Operationsaufw. zur Verdünnung
$O_{SS}$	Operationsaufw. zur Segementierung der Zeichen
$O_{SZ}$	Operationsaufw. zur Segmentierung einer Zeile
$O_{Trans}$	Operationsaufw. zur Transformation
$O_a$	relativer Operationsaufw. bezogen auf $O_{add}$
$O_{add}, O_{mul}, O_{div}, O_{sqr t}$	Operationsaufw. für Addition, Multiplikation, Division, Radizieren
$O_{del}$	Operationsaufw. für Verdünnung
$O_{hess}$	Operationsaufw. für Hessesche Normalform
$O_k$	Operationsaufw. zur Bestimmung eines Ähnlichkeitswertes
$P$	Positionierungsbereich
$P, P_u$	Menge der Produktionen, Produktion $u$
$P_i$	Punktmenge einer Liniendarstellung
$P_R, P_G$	Punktmenge der Liniendarstellung der Referenz, des Musters
$P_R^*$	akzeptierte Punktmenge der Referenzliniendarstellung

$P_G^*$	akzeptierte Punktmenge der Musterliniendarstellung
$R$	Referenzbild
$R_b$	Binärmaske der Referenz $R$
$S$	Suchbereich
$S$	Zielobjekt (Startsymbol)
$U_\omega$	Umgebung $\omega$
$V, W, X, Y, Z$	Objekttypen der Objekte $v, w, x, y, z$
$V_n$	Menge der nichtterminalen Symbole (Teilobjekte und Zielobjekt)
$V_t$	Menge der terminalen Symbole (Primitivobjekte)
$a$	Attribut
$cov$	Kovarianz
$c_{kr}, c_{kh}$	relative Konturpunktzahlen
$c_x(z), c_y(z)$	Zeiger in x-Richtung, bzw in y-Richtung
$d$	Abstandsmaß
$d_{appr}$	Approximationsgüte
$d_{dist}$	Distanzwert
$d_{P_i}$	Abstand des Punktes $p_i$
$d_k$	Kreuzkorrelation, normiert
$d_{max}$	maximaler Distanzwert
$d_{pmax}$	Schwellwert für Distanz eines Punktes
$e = e_1, \dots, e_{nr}$	Entscheidungsvektor
$f$	Brennweite
$f(x, y)$	Approximationsfunktion
$g$	Grauwert
$g(x, y)$	Grauwertfunktion
$g_b$	Gradientenbetrag
$g_{bmin}$	Schwellwert für den Gradientenbetrag
$g_{hy}(x, y)$	Grauwertgradient
$\bar{g}$	Mittelwert der Grauwerte
$\bar{g}_\oplus, \bar{g}_\ominus$	Mittelwert der zu $A_\oplus, A_\ominus$ zugehörigen Grauwerte
$g_{sx}$	Summe der Grauwertdifferenzbeträge in x-Richtung
$g_{ssx}$	Summe von $g_{sx}$
$g_{th}$	Schwellwert
$i, j$	Laufvariablen
$k_R, k_G$	Ähnlichkeitsmaß (quantitativ)
$k_d$	Ähnlichkeitsmaß (qualitativ)
$k_k$	Korrelationskoeffizient
$k_{kT}, k_{k\theta}$	Schwellwerte

$k_1, \dots, k_{10}$	Konstanten
$l_i$	Linie mit Index $i$
$m, n$	Anzahl
$n_{FFT}$	Anzahl durchzuführender FFTs
$n_{bx}, n_{by}, n_{bxy}$	Anzahl der Bildpunkte von B in x-, bzw. y-Richtung, der Bildfläche
$n_{cd}$	Anzahl der Distanzberechnungen
$n_{cl}$	Anzahl der durchgeführten Linienberechnungen
$n_k$	Anzahl der Konturpunkte
$n_{kg}$	Anzahl der Konturpunkte (nach Transformation)
$n_{kh}$	Anzahl der Konturpunkte (nach Konturpunktdetektion)
$n_{kr}$	Anzahl der Konturpunkte der Liniendarstellung (Referenz)
$n_{ks}$	Anzahl der Konturpunkte der Skelettdarstellung
$n_l$	Anzahl der Linien
$n_{px}, n_{py}, n_{pxy}$	Anzahl der Bildpunkte von P in x-, bzw. y-Richtung, der Bildfläche
$n_q, n_{qx}, n_{qy}, n_{qxy}$	Anzahl der zu transformierenden Werte
$n_r$	Anzahl der Referenzmasken
$n_{rx}, n_{ry}, n_{rxy}$	Anzahl der Bildpunkte von R in x-, bzw. y-Richtung, der Bildfläche
$n_s$	Länge einer Bildspalte oder Bildzeile
$n_{sx}, n_{sy}, n_{sxy}$	Anzahl der Bildpunkte von S in x-, bzw. y-Richtung, der Bildfläche
$n_{r\oplus}, n_{r\ominus}, n_{r\circ}$	Anzahl der Bildpunkte der Teilfläche $A_{\oplus}, A_{\ominus}, A_{\circ}$
$n_{r\odot}$	Summe der Bildpunkte der Teilflächen $A_{\oplus}$ und $A_{\ominus}$
$n_z$	Anzahl der im Suchbereich dargestellten Zeichen
$n_{zy}$	Schriftzeichenhöhe
$\bar{n}$	Oberflächennormale
$p$	Oberflächenpunkt
$r, r_i, r(u, v)$	Bildpunkt der Referenz R an der Stelle $i$ bzw. $u, v$
$\bar{r}$	Mittelwert aller Grauwerte der Referenz
$r_b(x, y, z)$	Ortsvektor für Bildpunkt B
$r_p(x, y, z)$	Ortsvektor für Oberflächenpunkt P
$u_i$	Vektoren $u_0$ : Nullvektor
$x, y, z, u, v$	Koordinaten
$y_{max}$	Position des Maximums von $g_{ssz}(y)$
$y_{st}, y_{sb}$	Position der oberen, bzw. unteren Suchbereichsgrenze
$y_{zt}, y_{zb}$	Position des oberen, bzw. unteren Schriftzeilenrandes
$z_h$	Zeichenhöhe
$\alpha, \beta, \gamma$	Winkel der Bildaufnahme
$\varphi$	Verdrehungswinkel der Vorlage gegenüber der Referenz

$\delta_\omega$	Differenz der Korrelationsbeträge
$\phi$	Orientierung einer Linie
$\phi_b$	Lichtstrom in der Bildebene
$\phi_d$	Winkelabstand
$\phi_e$	einfallender Lichtstrom
$\rho(\alpha, \beta, \gamma)$	Reflexionsfunktion
$\rho_\oplus, \rho_\ominus$	Wert eines Bildpunktes der Referenz aus der Fläche $A_\oplus, A_\ominus$
$\sigma$	Standartabweichung der Gaußverteilung
$\sigma_g, \sigma_r$	Standartabweichung der Intensitäten von Muster bzw. Referenz
$\sigma_\oplus^2, \sigma_\ominus^2$	Varianz der zu den Flächen $A_\oplus, A_\ominus$ entsprechenden Grauwerte
$\odot$	geometrische Relation
$\xrightarrow{u}$	objektspezifische Generierungsfunktion für Objekt u
BPI	Blackboard oriented production system for image understanding
DFT	Diskrete Fourier Transformation
DOG	Difference of Gaussian
EDT	Euklidische Distanztransformation
EEDT	Erweiterte Euklidische Distanztransformation
FFT	Fast Fourier Transformation
IDFT	Inverse Diskrete Fourier Transformation
LOG	Laplace of Gaussian
OCR	Optical Character Recognition
OCR-A, OCR-B	Schriftsätze (Fonts)
SD	Spitzen-Detektor
TPL	Template



# 1. Einleitung

Im Rahmen der Automatisierung industrieller Fertigungsprozesse nimmt die Materialflußsteuerung eine zentrale Stellung ein. Einzelteile müssen in der richtigen Reihenfolge angeliefert und gemäß einem Montageplan in vorbestimmter Art und Weise montiert werden. Um eine flexible Fertigung zu ermöglichen, wird dem "entstehenden" Objekt der Montageplan mitgegeben. Die Unterscheidung der Grundbauteile wird durch eine Kennzeichnung ermöglicht. Zur Fließbandfertigung eines Kraftfahrzeuges mit individueller Ausstattung kennzeichnet man beispielsweise die Karosserie des Fahrzeugs mit eingepprägten Schriftzeichen. Aber auch zur Identifizierung von äußerlich gleich aussehenden, in ihrer Funktion aber unterschiedlichen Teilen, wird häufig eine Markierung durch alphanumerische Zeichen vorgenommen.

Eine gängige Art der Kennzeichnung ist das Stanzen, bzw. Prägen, Gravieren oder Schlagen von Schriftzeichen, da sie von nachfolgenden Fertigungsschritten wie z.B. einer Überlackierung nicht verfälscht wird und außerdem an den Menschen als Kontrollinstanz angepaßt ist. Für viele Anwendungen scheidet eine Beschriftung durch Farbauftrag aus, da sie durch Korrosion und mechanische Beschädigungen unlesbar werden kann oder keine ausreichende Fälschungssicherheit bietet. Dies ist z.B. bei Fahrgestellnummern von Kraftfahrzeugen gefordert. Soll die Kontrollfunktion durch maschinelles Zeichenlesen automatisiert werden, ergeben sich oft große Probleme.

Die optische Erkennung von Schriftzeichen, kurz als OCR (**O**ptical **C**haracter **R**ecognition) bezeichnet, ist keine neue Aufgabenstellung in der Automatisierungstechnik. Erste OCR-Maschinen wurden bereits zu Beginn der 60er Jahre [78, 66] kommerziell eingesetzt. Seitdem wurden viele Beiträge veröffentlicht, die sich mit OCR-Problemen beschäftigen. Stand zunächst die Erkennung von Schriftzeichen in gedruckter Form im Vordergrund, so wandte man sich im Laufe der Zeit und mit zunehmender Leistungsfähigkeit der Rechner den Problemen handgeschriebener Zeichen zu. Ungezwungene handgeschriebene Zeichen stellen gegenüber den festen Schriftsätzen gedruckter Zeichen (Fonts) durch die große Variationsbreite der Zeichenform deutlich höhere Anforderungen an das Zeichenerkennungssystem (siehe z.B. [16]). Die Erkennung handgeschriebener Zahlen findet z.B. beim automatischen Lesen von Bankbelegen oder Postleitzahlen (ZIP-Code) Anwendung. Besondere Aufmerksamkeit hat man in den letzten Jahren auch den typischen Problemen der Erkennung von Schriftzeichen unterschiedlicher Sprachen ( z.B. Chinesisch, Japanisch, Arabisch) oder von handgeschriebenen mathematischen Ausdrücken gewidmet. Eine Zusammenstellung der Literatur zur Zeichenerkennung findet man in den Beiträgen von Rosenfeld [74, 75, 76].

Zur Erkennung von gedruckten Zeichen aus Videobildern existieren ausgereifte Verfah-

ren. Bei diesen wird jedoch vorausgesetzt, daß sich die Zeichen deutlich vom Hintergrund abheben. Diese Voraussetzung ist in der Regel bei Dokumenten als Vorlage gegeben, da sich Schrift (z.B. Tinte, Toner) und Hintergrund (z.B. weißes Papier) durch die unterschiedlichen Reflexionseigenschaften der Materialien deutlich unterscheiden. Bei guter Druckqualität und hohem Kontrast lassen sich die zur Schrift zugehörigen Bildpunkte von der Menge der Bildpunkte des Hintergrunds durch eine einfache Schwellwertoperation trennen.

Diese Bedingung ist bei den durch mechanische Bearbeitung aufgetragenen Zeichen in der Regel nicht gegeben. Es ist problematisch, die dreidimensionale Zeichenstruktur eindeutig und reproduzierbar in die zweidimensionale Grauwertverteilung des Videobildes zu überführen. Im Gegensatz zu einer Schwärzungsverteilung werden die Zeichen aufgrund der unterschiedlichen Mikrostruktur nur eines Materials, unterschiedlicher Orientierung der Oberflächennormalen und des Schattenwurfs wahrgenommen. Das Erscheinungsbild hängt stark von der Beleuchtung ab, wobei die aufgenommenen Grauwertbilder sehr kontrastarm sein können.

Zur Optimierung der Erkennungssicherheit haben sich für das Lesen von Klarschrift (z.B. bei Bankbelegen) speziell für die Zeichenerkennung entworfene und standardisierte Schriftsätze (OCR-A, OCR-B) [62, 63, 64, 65] durchgesetzt. Solche für die Bildauswertung spezifizierte Schriftsätze findet man bei Prägestempeln oder Schlagzeichen i.allg. nicht.

## 1.1 Informationsverarbeitung durch das visuelle System

Ersetzt man bei optischen Inspektionsaufgaben den menschlichen Entscheidungsträger durch einen Automaten mit einem Bildaufnahmesystem, so liegt es nahe sich zu verdeutlichen, welche Wahrnehmungsfähigkeiten der Mensch mit seinem visuellen System besitzt. Ein allgemeiner Vergleich der Erkennungsleistung von Mensch und Maschine ist jedoch schwierig, da durch die natürlich bedingten Unterschiede der "Hardware", wie z.B. der Sensorik, auch unterschiedliche Informationen der Szene entnommen werden können. Beispielsweise läßt sich durch eine Infrarot-Kamera Information aus dem nicht sichtbaren Bereich des Lichtes aufnehmen. Andererseits besitzt der Mensch durch die Mobilität und die Integrierung des Handlungsablaufs in die Erkennung die Möglichkeit, wertvolle 3D-Informationen aufzunehmen und sich an die Umgebungsbedingungen zu adaptieren.

Eine Abgrenzung der Information zur Objekterkennung, die in einer Szene zur Verfügung steht, ist dann gegeben, wenn man nicht von der Szene selbst, sondern von einem

digitalisierten Grauwertbild der Szene ausgeht. Damit ist nicht zu bestimmen, mit welchem Stempel aus einer Auswahl eine Prägung ausgeführt wurde, sondern, ob das aus einer Grauwertdarstellung vom Menschen "sicher" erkannte Zeichen auch von einer Maschine erkannt wird.

Wechselnde Beleuchtungsverhältnisse, Translations-, Rotations-, Dilatations- und Deformationsbewegungen erschweren das Problem der automatischen Objekterkennung. Dem Menschen dagegen erscheint es i.allg. einfach, Schriftzeichen auf einem Videobild unabhängig von der Beleuchtung, der Lage auf dem Bild, der Zeichenorientierung, der Zeichengröße und der Art des Zeichensatzes wahrzunehmen. Die Prinzipien der menschlichen Wahrnehmung mit ihren Konstanz- und Invarianzeigenschaften sind wegen der vorbildlichen Funktion von besonderem technischen Interesse. Die komplexen Vorgänge der Erkennung sind jedoch schwierig nachzuvollziehen.

Durch die Anwendung systemtheoretischer Methoden auf natürliche neuronale Netzwerke (NNN), insbesondere die Erweiterung der Beschreibung von zeit- auf ortsabhängige Vorgänge, konnte das Verstehen neuronaler Filteroperationen erleichtert werden [47]. Beispielsweise findet in der Retina neben der Transformation der zweidimensionalen Leuchtdichteverteilung in ein elektrisches Impulsmuster auch eine Kontrastverschärfung des Bildes statt. Diese von den Ganglienzellen durchgeführte Bildvorverarbeitung läßt sich als Filteroperation zur Kantenverstärkung auffassen. Das Empfindlichkeitsprofil im rezeptiven Feld einer Ganglienzelle kann durch die Differenz zweier Gaussverteilungen (DOG) beschrieben werden.

Einen wichtigen Beitrag zum Verständnis der Informationsverarbeitung durch das visuelle System erbrachten Hubel und Wiesel [43] mit ihren Untersuchungen der funktionellen Eigenschaften von Nervenzellen im primären visuellen Kortex. Der aus einzelnen Säulen modular aufgebaute visuelle Kortex verarbeitet mit jeder Säule einen kleinen Ausschnitt des visuellen Feldes. Die Neurone einer Säule wurden klassifiziert in einfache Zellen (simple cells), komplexe Zellen (complex cells) und hyperkomplexe Zellen (hypercomplex cells). Komplexe Zellen zeigen gegenüber einfachen Zellen in ihrer Ausgangserregung eine Unabhängigkeit der Reizposition im rezeptiven Feld. Hyperkomplexe Zellen zeigen gegenüber den komplexen Zellen eine Abhängigkeit der örtlichen Ausdehnung des Reizes. Durch Einzelableitung der Aktivität von Nervenzellen konnte nachgewiesen werden, daß einzelne Zellgruppen ausschließlich auf bestimmte Reizmerkmale reagieren. Alle abgeleiteten Neurone innerhalb einer Säule reagieren maximal auf eine bestimmte Orientierung eines Lichtreizes in Form von einer Kante, einer Linie oder eines Spaltes. Bezüglich der stationären optimalen Reizstruktur werden zur Extraktion von Merkmalen funktionell Kantendetektoren, Liniendetektoren und Spaltdetektoren mit unterschiedlicher Breite unterschieden.

Aufbauend auf neurophysiologischen Erkenntnissen entwickelten Marr und Hildreth [53] ein Modell zur Kantendetektion, in dessen Mittelpunkt der auf die Gaussfunktion angewendete Laplace-Operator (LOG-Operator) steht. Bei geeigneter Parameterwahl nähert ein DOG-Profil die zweite Ableitung einer Gaussverteilung (LOG) an. Intensitätsschwankungen im Bild werden durch LOG-Filterung mit anschließender Nullstellenbestimmung detektiert und durch in Linien angeordnete Punktmengen (zero-crossings) für verschiedene Skalierungen in mehreren Kanälen dargestellt (siehe auch Kap. 3.4.2). Ein Bild läßt sich durch Segmente der zero-crossings in Form von orientierten Primitivobjekten repräsentieren. Durch Kombination der Segmente verschiedener Kanäle wird eine Bildbeschreibung erreicht, die als "raw primal sketch" bezeichnet wird. Als Beschreibungselemente (Objekte) werden Kanten, Striche und Kleckse (blobs) verwendet. Eine hierarchische Beschreibung des Bildes ("full primal sketch") entsteht durch aktive Konstruktion mittels Auswahl, Gruppierung und Zusammenfassung von Objekten zu abstrakteren Objekten [54].

Wesentliche Beiträge zum Verständnis der Wahrnehmung lieferten aber nicht nur Neurophysiologen, sondern auch Psychologen, beispielsweise durch die Gruppierungsregeln aus der Gestalttheorie. Liegen bestimmte geometrische Relationen von Objekten vor, so besteht eine Tendenz des Wahrnehmungssystems, Objekte auf Grund von Nähe, Symmetrie, Fortsetzung oder Ähnlichkeit zu Gruppen zusammenzufassen. Weiterhin zeigen Versuche mit optischen Täuschungen, daß zwischen der subjektiven Wahrnehmung und den zugrundeliegenden objektiven Meßwerten eines Bildes erhebliche Differenzen bestehen können.

Untersuchungen zur rotationsinvarianten Erkennung wurden von Cooper und Shepard [9, 10] durchgeführt. Ihre Ergebnisse deuten darauf hin, daß verdrehte identische Objekte zum Bildvergleich mental gedreht werden. Dabei zeigte sich, daß die benötigten Reaktionszeiten zur Wiedererkennung bekannter Objekte proportional mit dem Verdrehungswinkel ansteigen. Dieser lineare Zusammenhang mit der Winkeldifferenz war unabhängig von der Komplexität der verwendeten Muster.

Besondere Bemühungen Adaptions- und Lernvorgänge zu untersuchen, sind derzeit im Bereich Künstlicher Neuronaler Netze (ANN) zu erkennen. In Zusammenhang mit der Zeichenerkennung ist aus der Vielzahl der veröffentlichten Beiträge, die sich mit ANN-Modellen zur Mustererkennung beschäftigen, das Neocognitron [24, 25] von Fukushima hervorzuheben. Bei diesem mehrschichtigen Modell wurden beim strukturellen Aufbau in besonderem Maße neurophysiologische Erkenntnisse des visuellen Systems berücksichtigt. Die charakteristische Struktur der Merkmalsextraktion ist in Analogie zu den von Hubel und Wiesel klassifizierten Nervenzellen mit ihren reizspezifischen Reaktionen entworfen worden. Als Eingangsmuster verwendetet Fukushima Binärbilder von hand-

geschriebenen Ziffern. Das Netz zeigt ein invariantes Verhalten gegenüber Verschiebungen, Größenänderungen und Deformationen der Eingangsmuster.

Übersichtliche Einführungen in die Bildverarbeitung unter dem Aspekt *Mensch und Maschine* geben z.B. Frisby [20] und Levine [48]

## 1.2 Zielsetzung der Arbeit

Zur Erkennung von Schriftzeichen existieren bereits zahlreiche Verfahren. Ein Leistungsvergleich dieser Verfahren wäre prinzipiell durch Bestimmung der Erkennungswahrscheinlichkeiten möglich. Dazu wird eine ausgewählte Datensammlung mit den Verfahren bearbeitet und festgestellt, wieviele Zeichen richtig oder falsch erkannt bzw. zurückgewiesen werden. Zur Durchführung solcher Experimente mit Experimentalsystemen ist jedoch ein hoher Zeitaufwand verbunden, entsprechender Speicherplatz notwendig und die Verfügbarkeit der Algorithmen vorausgesetzt. Selbst wenn die Algorithmen der Verfahren vorliegen, ist das verfahrensspezifisches Vorwissen (z.B. Parametereinstellungen), das der Anwender oftmals aus der Kenntnis vorliegender Daten mit einbringt, nicht zu unterschätzen. Dies erschwert natürlich einen objektiven Vergleich.

Ein direkter Vergleich von Ergebnissen aus Veröffentlichungen (z.B. [12, 18, 80]) ist wegen der unterschiedlichen Aufnahmebedingungen, Art der Zeichenvorlagen und Variation (Art und Größe) der Störungen meist nicht möglich. Aus den oben genannten Gründen wird auf einen Leistungsvergleich durch Erkennungswahrscheinlichkeiten verschiedener Verfahren verzichtet.

Ziel der vorliegenden Arbeit ist die Auswahl und Darstellung von drei Verfahren, die sich zur Erkennung in Metall geschlagener Zeichen eignen. Die Untersuchung der implementierten Verfahren wird an demselben Datenmaterial vorgenommen. Es wird davon ausgegangen, daß das Erscheinungsbild der Zeichen durch Änderung der Reflexionsbedingungen stark variieren kann. Weiterhin soll berücksichtigt werden, daß im Bereich der industriellen Fertigung nicht immer eine exakte Positionierung und Ausrichtung von Werkstücken gewährleistet werden kann. Es soll exemplarisch die Empfindlichkeit der Verfahren an Größen dargestellt werden, die zur Entscheidung bei der Klassifikation herangezogen werden.

Da in zeitkritischen Anwendungen der durch Rechnerleistung aufzubringende Operationsaufwand eine Rolle spielt, ist auch der notwendige Berechnungsaufwand der Verfahren zu berücksichtigen. Der Aufwand der einzelnen Verfahren soll in Abhängigkeit der untersuchten Bildbereiche bestimmt, an konkreten Beispielen abgeschätzt und vergleichend gegenübergestellt werden.

## 1.3 Übersicht

Im Anschluß an die Einleitung werden in **Kapitel 2** zunächst die experimentellen Versuchsbedingungen zur Bildaufnahme mit dem Bildaufnahmesystem, der Zeichenvorlage und der Beleuchtung beschrieben. Prinzipiell kann aus einer analytischen Beschreibung der Bildaufnahme ein synthetisches Bild bestimmt werden. Auf die Vorgehensweise zur synthetischen Bestimmung des Bildes und auf die Problematik mit den Aufnahmeparametern in der Praxis einen Mustervergleich durchzuführen, wird kurz eingegangen. Das aufgenommene Datenmaterial wird in Auszügen in verschiedenen Darstellungen vorgestellt.

In **Kapitel 3** werden die Grundlagen erläutert, die eine Einteilung der verschiedenen Verfahren zur Mustererkennung ermöglichen. Es werden drei veröffentlichte Beiträge vorgestellt, die sich mit der Erkennung von in Metall geprägten Zeichen beschäftigen. Im Anschluß werden die prinzipiellen Probleme, die bei der Überführung des ikonischen Bildes in eine Liniendarstellung entstehen, an Beispielen dargestellt und es wird auf die Grundlagen eines Mustervergleichs mittels Korrelation eingegangen.

In **Kapitel 4** werden drei Verfahren zur Erkennung in Metall geschlagener Zeichen vorgestellt. Das erste Verfahren ist ein rein numerisches Verfahren und basiert auf einem direkten Mustervergleich des ikonischen Bildes mit einer Maske mittels der Korrelation. Im zweiten Verfahren wird aus dem ikonischen Bild eine Liniendarstellung gewonnen und mit einer Referenzliniendarstellung verglichen. Im dritten Verfahren wird ein syntaktischer Ansatz verfolgt und das Bild durch eine Menge von Objektprimitiven beschrieben. Die Erkennung erfolgt durch eine modellgestützte Analyse mit einem Produktionssystem. Für die drei Verfahren werden Ergebnisse der Erkennung einer Mustervorlage bei unterschiedlicher Beleuchtung gegenübergestellt. Weiterhin wird an Beispielen die Empfindlichkeit der Verfahren bezüglich einer Verdrehung der Mustervorlage dargestellt.

In **Kapitel 5** wird ein Verfahren zur Einschränkung des Suchbereiches beschrieben, mit dem in einem Bild Regionen von Schriftzeilen und Schriftzeichen segmentiert werden können.

In **Kapitel 6** wird der Berechnungsaufwand der in Kapitel 4 und Kapitel 5 dargestellten Verfahren untersucht. Für jedes Verfahren wird der relative Berechnungsaufwand unter Annahme der Realisierung mit einem konkreten Prozessor bestimmt und für fünf Anwendungsbeispiele die aufzubringende Operationsarbeit geschätzt.

Abschließend werden in **Kapitel 7** die Ergebnisse diskutiert und die Verfahren bewertet.

## 2. Experimentelle Versuchsbedingungen

### 2.1 Bildaufnahme

Zur Erprobung und Beurteilung von Erkennungsverfahren wurde von einer Zeichenvorlage Bildmaterial unter verschiedener Beleuchtung aufgenommen. Abb. 2.1 zeigt den Versuchsaufbau mit Beleuchtungseinrichtung, Zeichenvorlage und Bildaufnahmesystem, das im wesentlichen aus Kamera, Bildspeicher und Rechner besteht.

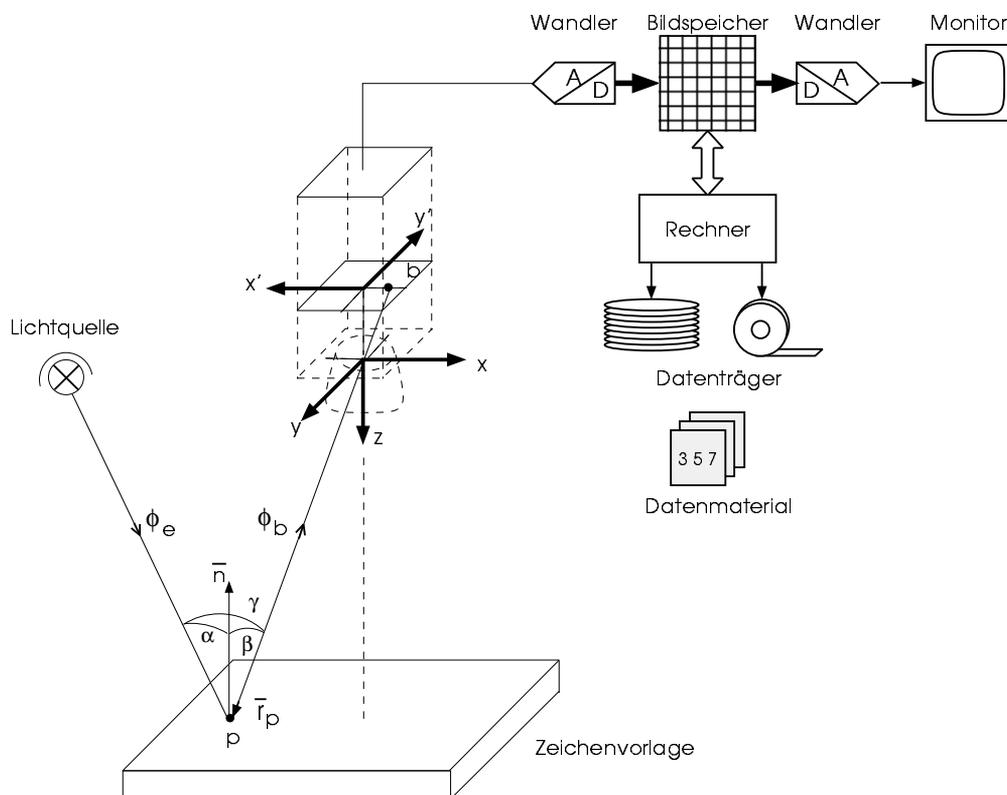


Abb. 2.1: Schematische Darstellung der Bildaufnahme

Das Bild einer Szene wird vom Menschen durch die Empfindung einer örtlichen Verteilung von Helligkeitswerten wahrgenommen. Die Helligkeit ist eine subjektive Erfahrung [47]. Die im Zusammenhang mit der Helligkeit stehende physikalisch meßbare Größe ist die Leuchtdichte. Das Aufnahmesystem mit Kamera, wie auch unser visuelles System, erhalten die gesamte Information über die Szene in Form der Leuchtdichtever-

teilung in der Bildebene. Ein Rückschluß auf die Objekte der Szene wird durch die unterschiedliche vom Ort abhängige Reflexion des einfallenden Lichtes ermöglicht. Der Reflexionsgrad hängt von der Mikrostruktur der Oberfläche ab und wird üblicherweise als Funktion von drei Winkeln  $(\alpha, \beta, \gamma)$  beschrieben (Abb. 2.1), die durch die lokale Flächennormale  $\bar{n}$ , den einfallenden Lichtstrahl und den emittierten Lichtstrahl bestimmt werden [41]. Die aufgenommene Intensitätsverteilung hängt jedoch nicht nur von diesen oben genannten Winkeln  $(\alpha, \beta, \gamma)$ , sondern auch von der Beleuchtung selbst ab.

Ein Oberflächenpunkt  $p$  sei durch den Ortsvektor  $\bar{r}_p(x, y, z)$  eines Koordinatensystems beschrieben, dessen Ursprung im Linsenzentrum liegt (Abb. 2.1). Im Abstand  $f$  auf der  $z$ -Achse liegt eine Bildebene, auf der der entsprechende Bildpunkt  $b$  durch  $\bar{r}_b(x', y', f)$  bestimmt ist. Bezeichnet  $\phi_e(x, y, z)$  den einfallenden Lichtstrom im Punkt  $p$  und  $\rho(\alpha, \beta, \gamma)$  den winkelabhängigen Reflexionsgrad, so ergibt sich für den Lichtstrom  $\phi_b(x', y')$  in der Bildebene die allgemeine Beziehung [91]

$$\phi_b(\bar{r}_b) = \rho(\alpha, \beta, \gamma) \cdot \phi_e(\bar{r}_p) \quad (2 - 1)$$

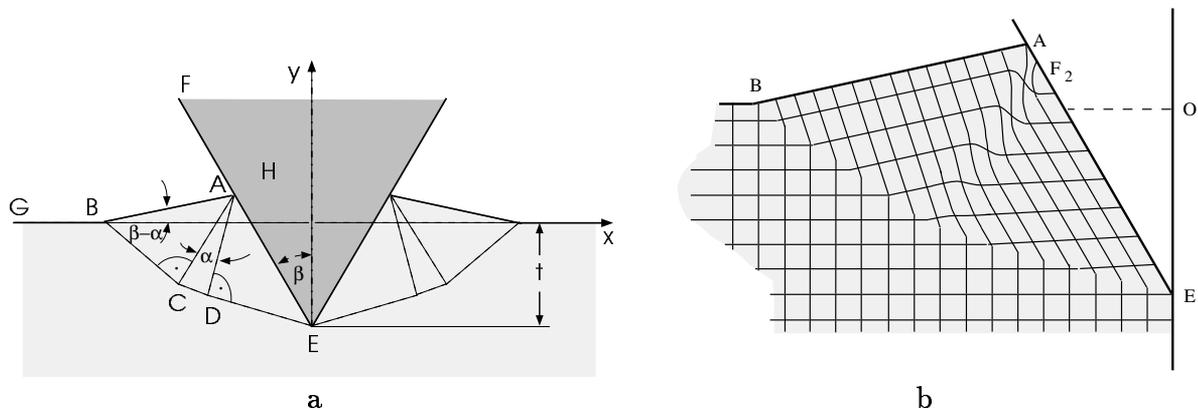
Der reflektierte Lichtstrom hängt also sowohl von dem Reflexionsgrad als auch von dem ortsabhängigen einfallenden Lichtstrom ab. Die Abhängigkeit des Reflexionsgrades von der Wellenlänge (siehe z.B. [33]) des Lichtes ist dabei noch nicht berücksichtigt.

### 2.1.1 Zeichenvorlage

Als Zeichenträger dienten metallische Werkstücke (blankgezogener Flachstahl ST37). In das Vorlagenmaterial wurden mit Schlagstempeln manuell Zeichen eingebracht. Beim Einschlagen der Stempel finden Fließvorgänge unterhalb des Stempels statt, die ein Aufwölben der Oberfläche in der näheren Umgebung der Eindruckstelle bewirken.

Erste theoretische Beschreibungen zu den Fließvorgängen beim Eindringen eines **flachen starren Stempels** in einen halbbunendlichen Körper wurden von Prantl [72] durchgeführt. Die Beschreibungen [38] zum Eindringen eines **geschmierten Keils** in einen halbbunendlichen Körper kommen aufgrund der Stempelform von Schlagziffern den vorliegenden Bedingungen näher. Eine ausführliche Diskussion der Spannungsverteilung dieser Modelle findet sich bei Prager und Hodge [71].

Einen Eindruck von der auftretenden Formänderung vermittelt Abb. 2.2. In Abb. 2.2a stellt EF die linke Flanke eines Keils (Öffnungswinkel  $2\beta$  dar, der um die Tiefe  $t$  eingedrungen ist. BG ist die noch unverformte Körperoberfläche und AB die Kontur einer Lippe, die hochgedrückt wird. Das Volumen der hochgestauchten Lippe ist dabei gleich dem durch den Keil verdrängten Rauminhalt. Abb. 2.2b zeigt die in [38] berechnete



**Abb. 2.2:** Oberflächenverformung durch Eindringen eines geschmierten Keils in einen halb-unendlichen Körper.  
a) Spannungsfeld(nach [71]), b) Das verformte Liniennetz (nach [38, 71]).

Verformung der ursprünglich quadratischen Maschen eines Gitternetzes nach dem Eindringen eines geschmierten Keils in einen halbunendlichen Körper.

Um eine Verformung der gesamten Oberfläche zu vermeiden, wurde bei den Werkstücken auf ausreichende Dicke geachtet. Die Schriftzeichen wurden wie in vielen typischen Anwendungen in einer Zeile angeordnet. Der Zeichenvorrat beschränkte sich auf zehn Ziffern eines Schriftsatzes mit einer Zeichenhöhe von 6 mm.

### 2.1.2 Beleuchtung

Natürliches Licht variiert so stark, daß es als Beleuchtungsquelle zur automatischen Inspektion in der industriellen Fertigung ausscheidet. Zum Aufbau einer Beleuchtungseinrichtung mit Kunstlicht macht Batchelor [2] eine Reihe von Vorschlägen. Die Möglichkeiten, spezielle Einrichtungen zur Beleuchtung einzusetzen, sind vom Umfeld abhängig, und eine Entscheidung kann erst bei genauer Kenntnis der Umgebungsbedingungen getroffen werden. Um die nachfolgende Bildauswertung nicht an eine spezielle Beleuchtungssituation zu adaptieren, wurde eine einfache Beleuchtungsausrüstung mit handelsüblichen Beleuchtungsmitteln gewählt. Es wurden keine besonderen Hilfsmittel [37] zur Regulierung, Ausrichtung oder Manipulation des Lichtes verwandt.

Sind bei einer speziellen Realisierung eines Inspektionssystems Optimierungen notwendig, wird man aus Kostengründen zunächst eine Verbesserung der Beleuchtung vornehmen, ehe man in die Verbesserung der nachfolgenden Bildverarbeitung investiert.

Unter dem Gesichtspunkt, daß zur Zeichenerkennung die Reflexionseigenschaft des Objektes charakteristisch ist und nicht die Beleuchtung, wird man mit der Beleuchtungs-

einrichtung eine vom Ort unabhängige Ausleuchtung anstreben. Von dieser Idealvorstellung kann jedoch generell nicht ausgegangen werden. Die Beleuchtungskomponente ändert sich mit dem Ort, wobei die Amplitudenänderung der Beleuchtung geringer angenommen wird als die der Reflexion. Betrachtet man nun die Beleuchtungsänderung als Störung, so ist eine Störungsreduktion durch ein lineares Filter nicht möglich, da nach Gl. 2-1 Signal und Störung nicht additiv überlagert, sondern multiplikativ miteinander verknüpft sind.

Eine Korrektur der Beleuchtungsdynamik kann direkt durch Division mit der Beleuchtungsfunktion oder durch ein homomorphes System vermindert werden. Bei einem homomorphen System transformiert man die nichtadditive Verknüpfung in eine additive, um eine lineare Filterung des erhaltenen Signals durchführen zu können. Das Ergebnis wird wieder der entsprechenden inversen Transformation unterworfen. Bei der vorliegenden multiplikativen Verknüpfung von Beleuchtungs- und Reflexionsfunktion können die erforderlichen Transformationen durch Logarithmieren und Exponentiation durchgeführt werden.

In den weiteren Betrachtungen wird davon ausgegangen, daß die Beleuchtungsfunktion in regionalen Bildbereichen (Bereich eines Schriftzeichens) als konstant angenommen werden kann.

## 2.2 Beschreibung der Bildaufnahme durch ein Modell (Bildsynthese)

Wird ein Objekt mit vollständig bekannten Oberflächeneigenschaften (*Position, Orientierung, spektrale Reflexion, Absorption, Diffusion, ...*) durch eine oder mehrere Lichtquellen mit bekannten Eigenschaften (*Position, spektrale Verteilung der Energie, Fokussierung des Lichtstrahls, ...*) beleuchtet, so kann das Bild, das von einer Kamera mit bekannten Eigenschaften (*Position, Orientierung, Brennweite, Auflösung, Verzeichnung, ...*) aufgenommen wird, berechnet werden [60].

In der Computergraphik sind durch die unterschiedlichen Annahmen von idealisierten Eigenschaften von Objektoberfläche, Beleuchtung und Kamera eine Reihe von Modellen zur Bildsynthese entwickelt worden, die synthetische Bilder realitätsnah erscheinen lassen. Die Modellvorstellungen zur Berechnung eines Bildes lassen sich neben der geometrischen Szenenbeschreibung in die Teilmodelle *Beleuchtungsmodell, Reflexionsmodell* und *Kameramodell* gliedern.

- In einfachen **Beleuchtungsmodellen** geht man von der Vorstellung einer idealen diffusen Strahlungsquelle aus, die gleiche Strahlungsdichte in alle Richtungen aussendet, oder man stellt sich eine Punktquelle mit unendlich großem Abstand vor,

die eine parallele Strahlung aussendet. Komplexe Lichtquellen und Beleuchtungsmodelle [52] berücksichtigen die Lichtquellengeometrie, die ausgesendete spektrale Verteilung und die räumliche Verteilung der Beleuchtungsintensität.

- In einfachen **Reflexionsmodellen** nimmt man die Objektoberfläche als *ideal matt* oder *ideal spiegelnd* an. Eine **ideal matte** Oberfläche (Lambertsche Oberfläche), die von allen Seiten gleich hell erscheint, wird durch die einfache Reflexionsfunktion  $\rho(\alpha, \beta, \gamma) = \cos \beta$  beschrieben (siehe Gl. 2-1). Für einen **idealen Spiegel** nimmt die Reflexionsfunktion  $\rho = 1$  an, wenn  $\alpha = \beta$  und  $\gamma = \alpha + \beta$  sind. Reale Oberflächen sind jedoch weder ideal matt noch ideal spiegelnd. Um sich an die realen Verhältnissen anzunähern, kombiniert man einzelne Reflexions- und Beleuchtungsmodelle durch gewichtete Summierung der einzelnen Intensitätsbeiträge [69, 8]. Dabei wird oftmals noch ein ambienter Intensitätsanteil berücksichtigt, der aus dem reflektierten Umgebungslicht stammt. Zusammenstellungen von verschiedenen Modellvorstellungen finden sich in den Arbeiten von [31, 52].
- In einfachen **Kameramodellen** [14] stellt man sich eine orthographische Projektion oder perspektivische Projektion durch eine Lochkamera auf eine Bildebene vor. Aufwendigere Kameramodelle berücksichtigen z.B. radiometrische oder geometrische Eigenschaften einer Linse oder eines Linsensystems. Ein Verfahren zur Bestimmung unbekannter Kameraparameter ist z.B. in [86] beschrieben

Prinzipiell läßt sich ein Mustervergleich durchführen, indem man das vorraussichtliche Erscheinungsbild jedes Zeichens mit einem parametrisierten Modell berechnet und mit dem aufgenommenen Bild vergleicht. Oder man beschreitet den umgekehrten Weg und löst das inverse Problem durch ein "Shape-from-Shading"-Verfahren [42], indem man aus einem oder mehreren Bildern die Orientierung von Objektoberflächen oder den geometrischen Verlauf der Oberfläche bestimmt.

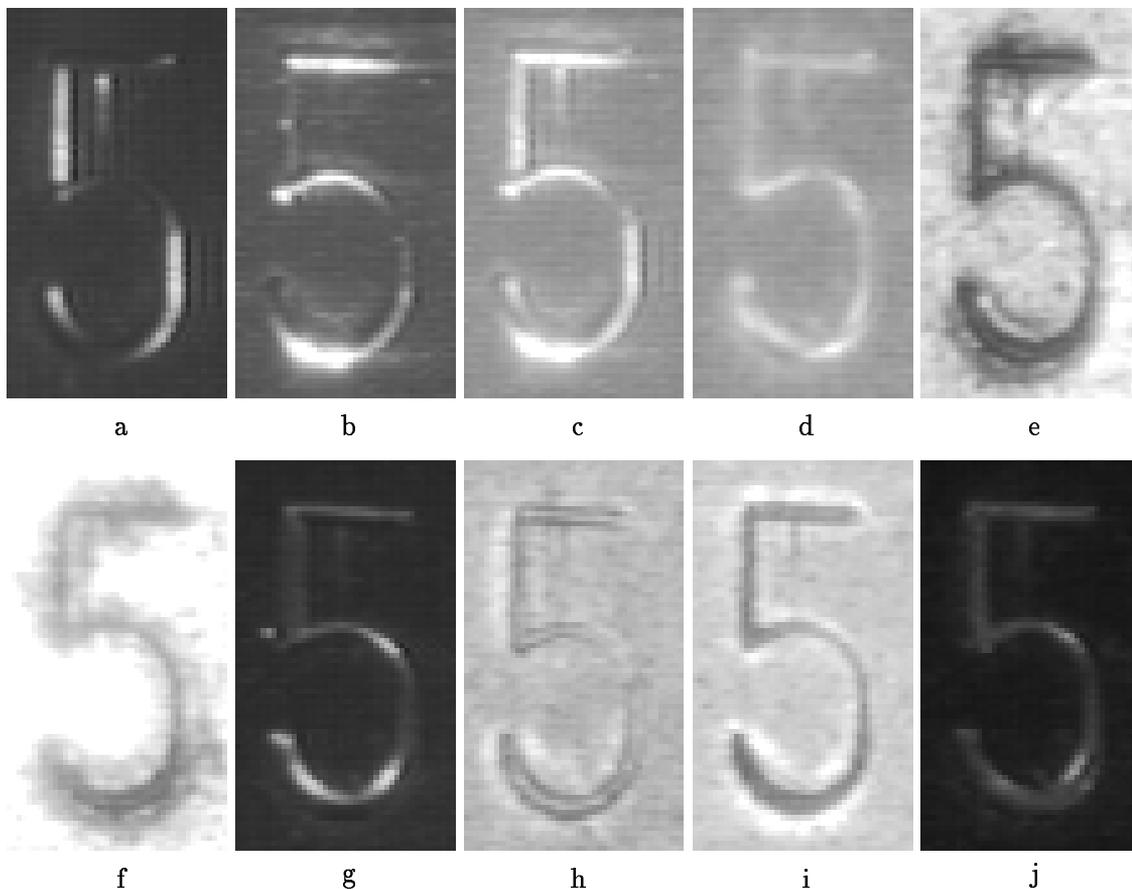
In industriellen Anwendungen sind jedoch nicht alle parametrisierten Eigenschaften bekannt, nicht stabil oder die Berechnung eines entsprechenden synthetischen Bildes ist zu rechenintensiv. So besteht manchmal Unkenntnis über die genaue zu erwartende Zeichenposition auf Werkstücken, die Eigenschaften der Beleuchtung oder die Reflexionsfunktion der Werkstücke. Bei metallischen Werkstücken können Oxidationsvorgänge die Oberflächeneigenschaften verändern und zu einem sehr unterschiedlichen Reflexionsverhalten führen. Durch Verschmutzung und Abnutzung des Stempels kann die 3-D-Geometrie des Stempelabdrucks variieren.

Am Beispiel von in Metall geprägten Zeichen versuchte Schmidt [80] durch Umkehrung der optischen Abbildung, die durch ein Beleuchtungs- und Reflexionsmodell beschrieben wird, vom Grauwertbild auf die Objektkonturen zu schließen. Das numerische Verfah-

ren zur Lösung des aufgestellten quasilinearen partiellen Differentialgleichungssystems erwies sich jedoch als sehr empfindlich gegenüber Störungen.

### 2.3 Aufgenommenes Datenmaterial

Die in Kap. 2.1.1 beschriebene Zeichenvorlage wurde unter verschiedenen Beleuchtungssituationen aufgenommen. Ein abgetastetes Bild  $B = [g(x, y)]$  mit  $x = 1, \dots, 512$ ,  $y = 1, \dots, 512$  besitzt eine Helligkeitsauflösung von 256 Grauwertstufen (8 Bit). Die Aufnahmen mit unterschiedlichem Erscheinungsbild derselben Vorlage sind an einigen Beispielen am Bildausschnitt der Ziffer fünf in Abb. 2.3 gegenübergestellt. Wie aus der



**Abb. 2.3:** Gegenüberstellung der Erscheinungsbilder der Ziffer fünf bei unterschiedlichen Beleuchtungssituationen

Abbildung ersichtlich, kann die Schlagziffer abhängig von der Beleuchtung sowohl durch einen hellen (Zeichen-)Vordergrund gegenüber einem dunkleren Hintergrund (Abb. 2.3j) wie auch umgekehrt (Abb. 2.3i) mit unterschiedlichem Kontrast erscheinen. Dabei können Teile des Vorder- oder Hintergrundes Glanzflecken aufweisen (Abb. 2.3c,f) und Be-

reiche des Zeichen wegen zu geringen Kontrasts unsichtbar werden. In Abb. 2.3a sind horizontale Teile und in Abb. 2.3b vertikale Teile des Zeichenvordergrundes kaum zu erkennen. Die unterschiedlichen Intensitäten des Zeichenvordergrundes von Abb. 2.3g sind in der Darstellung Abb. 2.4b in der Form eines Grauwertgebirges deutlich erkennbar. Die Intensität entlang des unteren Bogens variiert stark. Ein Vergleich der beiden Grauwertgebirge in Abb. 2.4 zeigt, daß die "Strichdicke" des Zeichenvordergrundes mit unterschiedlicher Breite auftreten kann. In Abb. 2.4a weist der Zeichenvordergrund eine größere Breite auf als in Abb. 2.4b.

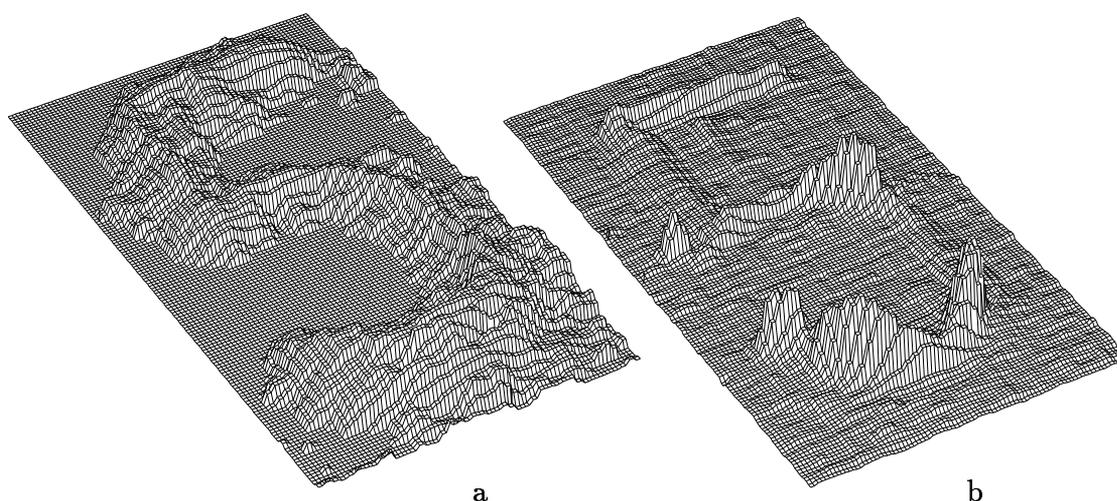


Abb. 2.4: Grauwertgebirge der Ziffer fünf (Der Grauwert ist als Höhe über dem Ort dargestellt).  
a) Muster Abb. 2.3f, b) Muster Abb. 2.3g  
(zur Gegenüberstellung zu a) wurde bei b) die Grauwertskalierung invertiert)

In Abb. 2.5 ist eine Aufnahme mit geringem Kontrast dargestellt. Die Störungen im Hintergrund (siehe z.B. im oberen Winkel) besitzen ähnliche Intensitäten wie die des Vordergrundes.

Beim Zeichenvordergrund lassen sich zwei typische Erscheinungsformen unterscheiden (Abb. 2.6). Der Vordergrund erscheint dunkel ausgefüllt in **flächiger** Darstellung (Abb. 2.6a) oder durch die Kontur in **linienhafter** Darstellung (Abb. 2.6b). Rechts neben den Zeichen sind in Abb. 2.6 jeweils die Grauwertprofile entlang einer ausgeblendeten Bildspalte (weiß im Grauwertbild) aufgetragen. Zur Normierung des Profils wurden eine schwarze Bildzeile (oben) und eine weiße Bildzeile (unten) mit den Grauwerten 0 und 255 eingeblendet. Die drei Durchtrittsbereiche der Bildspalte durch den Vordergrund zeichnen sich durch niedrige Werte aus, wobei in Abb. 2.6a an jedem Durchtritt eine einzelne ausgeprägte Spitze und in Abb. 2.6b eine "Doppelspitze" auftritt.

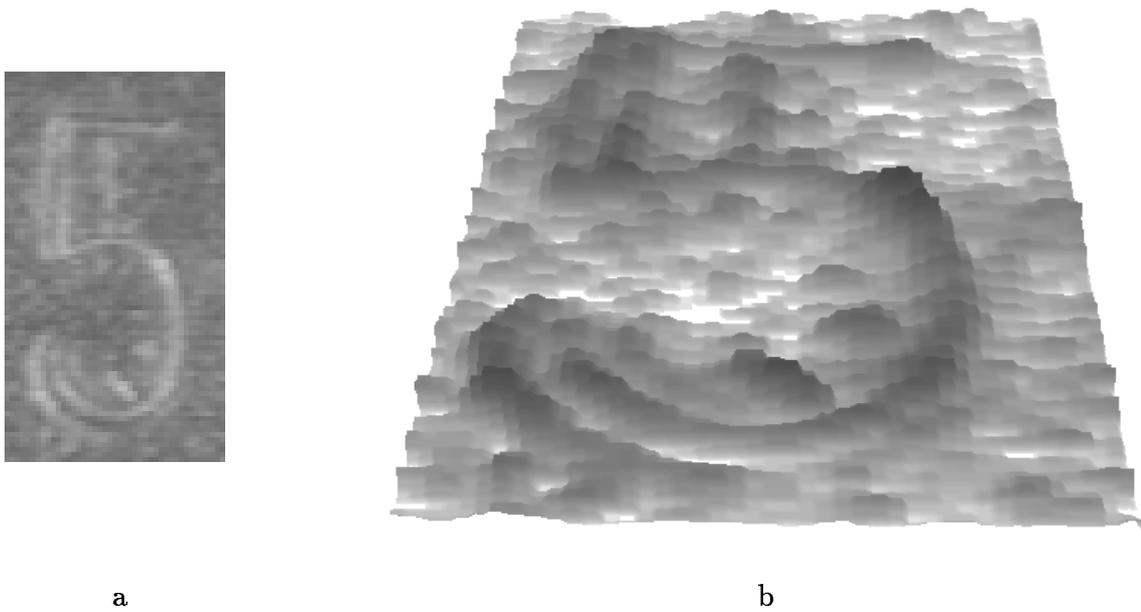


Abb. 2.5: Grauwertbild mit geringem Kontrast. a) Grauwertbild, b) Grauwertgebirge mit überlagertem Grauwertbild (die Grauwertskalierung wurde invertiert)

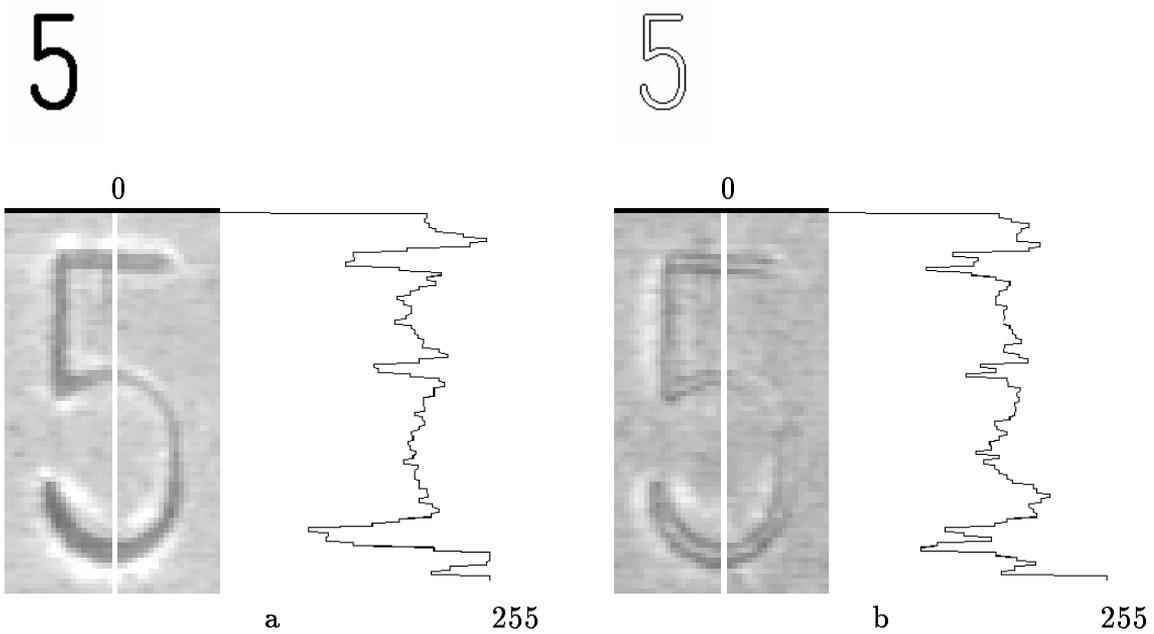


Abb. 2.6: Grauwertbilder mit dem Grauwertprofil der in weiß ausgeblendeten Bildspalte.  
 a) Zeichenvordergrund in flächiger Darstellung,  
 b) Zeichenvordergrund in linienhafter Darstellung

### 3. Grundlagen

Die automatische Schriftzeichenerkennung ist ein spezielles Problem der Mustererkennung. Dabei werden die in einem Bild vorliegenden Muster bestimmten Zeichenklassen (Musterklassen), z.B. den Ziffern null bis neun, zugewiesen. Die Bedeutungszuweisung bei dieser Überführung des Musters von der numerischen in die symbolische Repräsentation bezeichnet man als Klassifikation. Zur Vorgehensweise lassen sich *numerische* und *syntaktische* (nichtnumerische) Klassifikationsverfahren unterscheiden.

Bei der **numerischen Klassifikation** werden die Werte bestimmter aus dem Muster extrahierter Merkmale gemessen und als Zahlenwerte den Komponenten eines Vektors zugeordnet. Aufgrund dieses Merkmalvektors wird die Entscheidung, welcher Klasse das Muster angehört, als Ganzes getroffen.

Zur **syntaktischen Klassifikation** wird der Bildinhalt durch eine Folge von Symbolen beschrieben, die einem bestimmten Merkmal entsprechen oder die Relation zwischen Merkmalen bezeichnen. Durch einen formal erfaßten Mechanismus des schrittweisen Zerlegens und/oder Aufbaus des Musters kann die Klassenzugehörigkeit geprüft werden. Dieser mehrstufige (sequentielle) Erkennungsvorgang ist typisch für die syntaktische Klassifikation gegenüber dem meist einstufigen (parallelen) Erkennungsvorgang bei numerischen Klassifikationsverfahren.

#### 3.1 Numerische Klassifikation

Die  $n$  Komponenten des Merkmalsvektors lassen sich auch als Achsen eines  $n$ -dimensionalen Merkmalsraums auffassen. Mit dieser geometrischen Deutung bildet der vom Merkmalsvektor erfaßte Bildinhalt eines Musters einen Punkt im Merkmalsraum. Bei geeigneter Wahl der Merkmale werden die Muster der gleichen Klasse kompakt in Form einer Punktwolke (Cluster) abgebildet. Die Cluster unterschiedlicher Klassen sollten dabei soweit auseinanderliegen, daß sie sich nicht durchdringen. Die Streuung der Punkte innerhalb des Clusters wird durch den Intraklassenabstand beschrieben; der Abstand der Punktwolken zweier Klassen zueinander durch den Interklassenabstand. Diese Maße geben den mittleren quadratischen Abstand aller Punkte innerhalb einer Klasse bzw. zwischen unterschiedlichen Klassen an. Ziel von Optimierungen bei der Merkmalsauswahl ist die Minimierung der Intraklassenabstände bei gleichzeitiger Maximierung der Interklassenabstände.

Um ein unbekanntes Muster zu klassifizieren, werden bestimmte Klasseneigenschaften gespeichert, aus denen auf Form und Lage des Clusters geschlossen werden kann. Ab-

hängig von der Beschreibung der Klasseneigenschaften lassen sich nach der Vorgehensweise *statistische* Klassifikation, *verteilungsfreie* Klassifikation und *nichtparametrische* Klassifikation unterscheiden [60, 58]:

- Bei der **statistischen Klassifikation** wird der Merkmalsvektor eines Musters als Wert einer Zufallsvariablen aufgefaßt und die Kenntnisse über die Musterklassen werden in den bedingten Wahrscheinlichkeits-Verteilungsdichten zusammengefaßt. Durch Einführung einer Kostenfunktion wird das Klassifikationsergebnis in Abhängigkeit von der tatsächlichen Klassenzugehörigkeit bewertet. Durch Minimierung des Erwartungswertes der Kostenfunktion, des sog. Risikos, wird eine Klassifikationsregel abgeleitet (Bayes-Klassifikator), die den Merkmalsvektor auf die Musterklassen im Sinne der statistischen Entscheidungstheorie optimal abbildet.
- Die **verteilungsfreien Klassifikationsverfahren** benötigen keine Annahmen über die bedingten Verteilungsdichten. Die Eigenschaften der Klassen werden durch Trennfunktionen (Diskriminantenfunktionen) aus einer vorgegebenen parametrischen Familie von Funktionen beschrieben. Die Trennfunktion wird gegenüber einer vorgegebenen idealen Trennfunktion mit einem Gütemaß, wie dem mittleren quadratischen Fehler, angepaßt. Geometrisch gedeutet, sucht man eine Trennebene (Hyperfläche) im Merkmalsraum, die die Cluster zweier Klassen trennt. Zur Unterscheidung mehrerer Klassen benötigt man mehrere Trennebenen bzw. Diskriminantenfunktionen.
- Bei den **nichtparametrischen Klassifikatoren** liegt ebenfalls ein statistisches Modell zugrunde, jedoch verzichtet man auf die formelmäßige Erfassung der durch bestimmte Parameter gekennzeichneten Verteilungsdichten. So können z.B. in gegebenen Raumelementen die Verteilungsdichten geschätzt werden. Es kann auch die gesamte klassifizierte Stichprobe gespeichert und von einem neuen Mustervektor zu allen der Abstand berechnet werden. Das neue Muster wird dann der Klasse mit dem geringsten Abstand zugeordnet (Nächster-Nachbar-Regel).

Durch vereinfachte Annahme bestimmter Voraussetzungen lassen sich aus den Gruppen der Klassifikatoren weitere numerische Klassifikatoren ableiten [58].

Das prinzipielle Problem bei der numerischen Klassifikation besteht in der Extraktion geeigneter Merkmale, mit denen Merkmalsvektoren einer Klasse kompakt abgebildet werden und gleichzeitig einen großen Abstand zu allen anderen Klassen erreichen [45].

Bei der Anwendung statistischer Klassifikatoren stellt die Ermittlung der erforderlichen statistischen Größen eines der Hauptprobleme dar. Da bei konkreten Problemen die

Dichten in der Regel nicht gegeben sind, müssen sie mit Hilfe einer Stichprobe geschätzt werden. Die Approximation der Dichten ist dabei mehr oder weniger nur unvollkommen möglich. Zur Schätzung der Kovarianzmatrix sind mehrere Tausend Muster je Klasse notwendig [58]. Weiterhin ist die Anwendung der statistischen Klassifikation praktisch nur sinnvoll, wenn die Merkmale klassenweise angenähert normalverteilt oder statistisch unabhängig sind.

Da die hier verwendete Stichprobe vergleichsweise klein ist, werden die allgemeinen numerischen Ansätze nicht verfolgt.

Beim **Template Matching** wird ohne besondere Bestimmung eines Merkmalsvektors ein direkter Mustervergleich zwischen Template (Maske) und Bild bzw. Bildteilen durchgeführt. Um zu entscheiden, wo und ob das Muster des Templates im Bild enthalten ist, berechnet man ein Abstandsmaß. Für jede Klasse ist mindestens ein Template zu entwerfen. Faßt man das Template als Filter eines linearen Systems auf, erhält man durch Maximierung des Signal-Rausch-Verhältnisses ein signalangepaßtes Filter (matched filter).

### 3.2 Syntaktische Klassifikation

Syntaktische Verfahren erlauben in besonderem Maße die Beschreibung und Analyse struktureller Zusammenhänge des Musters [27, 21]. Voraussetzung für die symbolische Beschreibung eines in ikonischer Form vorliegenden Musters ist die Zerlegung der Objektstruktur in Primitivobjekte (Primitiveelemente, terminale Symbole). Durch schrittweises Zusammenfassen von Objekten und deren Relationen zu komplexeren Objekten, werden mit der Komprimierung der Musterbeschreibung verschiedene Abstraktionsstufen erreicht.

Die strukturellen Zusammenhänge des Musters bzgl. seiner Objekte lassen sich durch einen Graphen verdeutlichen, wobei die Objekte durch Knoten und die Beziehungen zwischen den Objekten durch Kanten repräsentiert werden.

Der Aufbau von komplexen Objekten aus weniger komplexen Objekten kann mit einem Ersetzungssystem beschrieben werden, das in Anlehnung an die formalen Sprachen durch eine Grammatik bestimmt ist [77]. Eine formale Grammatik  $E$  ist durch ein Vier-tupel definiert:

$$E = (V_t, V_n, S, P) \quad (3 - 1)$$

Dabei bezeichnet  $V_t$  die Menge der terminalen Symbole (Primitivobjekte),  $V_n$  die Menge der nichtterminalen Symbole (Teilobjekte),  $S$  das Startsymbol (Referenzobjekt) und  $P$  eine Menge von Produktionen (Ersetzungsregeln). Die Produktionen geben an,

wie sich terminale und nichtterminale Symbole in andere nichtterminale Symbole überführen lassen.

Umgekehrt lassen sich aus einem Startsymbol alle erlaubten terminalen Symbolkombinationen ableiten. Eine solche Symbolkombination bezeichnet man auch als Satz, die Menge aller Sätze der Grammatik  $E$  als Sprache  $L(E)$ . Der Begriff der formalen Sprache  $L_i$  entspricht bei den numerischen Verfahren der Klasse  $z_i$ , d.h. für jede Klasse gibt es eine Sprache  $L$  der Grammatik  $E_i$ .

Die Klassifikation besteht in der Prüfung, ob die Symbolkombination des Musters als Satz in der Sprache  $L(E)$  enthalten bzw. durch die Grammatik erzeugbar ist. Die zur Prüfung durchgeführte Zergliederung des Startsymbols in terminale Konstellationen wird als Parsen bezeichnet. Bei einer Erkennung durch Bildsynthese wird überprüft, ob ein Startsymbol (Klasse) zu einer unbekanntem Konstellation von Terminalen gehört (**top down**). Bei einem Erkennungsvorgang durch Konstruktion wird geprüft, ob durch Produktionsregeln aus den Terminalen auf ein Startsymbol geschlossen werden kann (**bottom up**).

Je nach Einschränkung des Regelsystems (Chomsky-Hierarchie) unterscheidet man verschiedene Typen von Grammatiken, die regulären, kontextfreien, kontextsensitiven und allgemeinen Grammatiken. Die Eigenschaften dieser Grammatiken sind in der Literatur ausführlich behandelt [77, 34]. Ein besonderer Typ der Grammatik ist die attributierte Grammatik. Sie ordnet den Symbolen (Terminale und Nichtterminale) neben der Kennzeichnung des Typs zusätzlich noch einen Attributvektor zu, der bestimmte Eigenschaften wie z.B. Länge, Orientierung, Güte usw. beschreibt. Da zur Auswertung dieser Eigenschaften die Regeln neben den Vorschriften zur Transformation von Symbolen auch Vorschriften zur Behandlung von Parametern enthalten müssen, läßt sich diese Grammatik als Kombination von numerischen und syntaktischen Verfahren auffassen.

Durch die Verwendung von Parametern kann die Komplexität der Sprache reduziert werden. Es muß jedoch der Extremfall vermieden werden, daß die gesamte Information in den Parametern steckt. Nach [58] besteht die Kunst darin, einen geeigneten Kompromiß zwischen der Information, die in den Symbolen enthaltenen ist und der Information, die in den Parameter enthaltenen ist, zu finden.

## 3.3 Veröffentlichte Beiträge zur Erkennung von Schlagzeichen

### 3.3.1 Objektangepaßte Filterung zur Konturfindung und Klassifikation

Schmidt [80] untersucht verschiedene Konturfindungsverfahren, die eine Zeichenerkennung unter dem störenden Einfluß von Reflexion und Schatten ermöglichen sollen. Als Vorlage werden in Aluminiumblech geprägte Ziffern verwendet, die auf der Vorderseite erhaben sind und sich bei senkrechter Beleuchtung durch deutlich erkennbare Konturlinien im Bild äußern. Bei den untersuchten Objekten handelt es sich um einzelne Ziffern, die bereits aus unterschiedlich beleuchteten Zifferngruppen getrennt wurden.

Der Schwerpunkt der Arbeit liegt in der Untersuchung und Verbesserung von Konturfindungsverfahren durch objektangepaßte Filter (siehe auch [79]). Mit der A-priori-Information der idealen Kontur und einem repräsentativen ikonischen Muster wird für jede Objektklasse (Ziffer) eine Filtermaske entworfen. Die Filter sind in einer Filterbank parallel angeordnet, wobei die Filtergröße jedes Filters der Bildgröße des Zeichens entspricht. Zur Klassifikation mit  $n$  Objektklassen werden durch Transformation (FFT), Filterung, Rücktransformation (IFFT) und Schwellwertbildung  $n$  Konturbilder erzeugt. Zur Klassifikation wird vorgeschlagen, aus den  $n$  Konturbildern Konturmerkmale zu extrahieren und mit einer Referenzkontur zu vergleichen. Die beste Übereinstimmung ist dann für das Konturbild des angepaßten Zweiges der Filterbank zu erwarten. Zur Bewertung der Ergebnisse benutzt Schmidt ein Gütemaß, das aus dem Vergleich der Konturbilder mit der Referenzkontur bestimmt wird.

Bei dem im Kap. 4.1 beschriebenen Verfahren I wird in ähnlicher Weise wie bei dem Verfahren von Schmidt eine Filterbank eingesetzt. Beim Entwurf einer Maske auf der Basis eines ikonischen Musters wird jedoch keine Adaption an die konkrete Beleuchtungssituation der Vorlage vorgenommen. Im Gegensatz zu der oben beschriebenen Vorgehensweise liegen als Ergebnis in den einzelnen Zweigen keine Konturbilder vor, sondern Ähnlichkeitswerte (Korrelationswerte), die eine unmittelbare Klassenzuweisung ermöglichen.

An einem Beispiel geht Schmidt auch auf die Möglichkeit der unmittelbaren Diskriminierung des Anpassungsfalles von den Fehlanpassungsfällen ein. Er benutzt dazu Merkmale, die aus den Histogrammen der gefilterten und rücktransformierten Grauwertbilder gewonnen werden und setzt einen statistischen Klassifikator ein.

### 3.3.2 Gradienten-Matching mit dünnen Referenzmustern

Evers und Kämmer [18] beschreiben zur Erkennung von in Metall geschlagenen Zeichen ein Verfahren auf der Basis von einem Gradienten-Matching. Der Klassifikation geht eine Segmentierung des Bildes in Zeichenketten und einzelnen Zeichen voraus. Das in Kap. 5.1 beschriebene Verfahren der Zeilensegmentierung entspricht dem von Evers und Kämmer beschriebenen Verfahren zur Lagebestimmung einer Schriftzeile.

Zur Klassifikation werden Referenzbilder der Schlagstempel benötigt, die in Form von binarisierten Konturbildern gespeichert sind. Jedem Konturpunkt der Referenz weist man durch Untersuchung der direkten Nachbarschaft eine Richtungsklasse zu. Für jede Richtungsklasse wird eine Gewichtsmatrix bestimmt, in der an der korrespondierenden Position eines Konturpunktes ein Richtungsähnlichkeitswert eingetragen wird. Den Punkten in der Nachbarschaft zu dieser Position wird ein Wert zugewiesen, der durch eine Abstandsfunktion bestimmt ist.

Zum Vergleich von Muster und Referenz wird das ikonische Bild des Musters in ein Gradientenbild überführt und durch eine Schwellwertoperation die Konturpunkte bestimmt. Mit der zu jedem Konturpunkt gespeicherten Richtung des Gradienten erfolgt eine Zuweisung des Konturpunktes zu der der Richtung entsprechenden Gewichtsmatrix. Die durch die Konturpunkte markierten Gewichtswerte in den Gewichtsmatrizen werden aufsummiert, so daß man für jede Referenz einen Ähnlichkeitswert erhält. Dem Muster wird die Klasse der Referenz zugewiesen, die den maximalen Ähnlichkeitswert besitzt.

Ähnlich wie bei diesem Verfahren, basiert auch bei dem in Kap. 4.2 beschriebenen Verfahren II die Erkennung auf den Vergleichen der Musterkontur mit den Referenzkonturen. Die Beschreibung des Musters wird dabei jedoch in eine symbolische Darstellung in Form von Linienstücken überführt. Zur Vermeidung von Störung durch geringe Verschiebung der Konturen gegeneinander sowie zusätzliche Störkonturen, werden die Ähnlichkeitswerte nicht unmittelbar aus den Gewichtsmatrizen (Distanzmatrizen) gewonnen, sondern über eine Punkt-zu-Punkt-Zuordnung zwischen Referenz und Muster.

### 3.3.3 Erkennung mit einem verteilten assoziativen Speicher

Dehesa et al. [12] beschreiben ein Verfahren zur Erkennung von in Metall geschlagenen Zeichen unter Verwendung eines verteilten assoziativen Speichers nach Kohonen [46]. Es wird davon ausgegangen, daß die zu erkennenden Zeichen vereinzelt vorliegen. In einer dreistufigen Vorverarbeitung wird eine Hochpaßfilterung, eine shiftinvariante Transformation und eine rotationsinvariante Transformation durchgeführt. Zur Hochpaßfil-

terung werden der Gradientenoperator nach Roberts, der Laplace-Operator und der DOG-Operator vorgeschlagen. Durch die Fast Fourier Transformation oder die Transformation nach Wagh und Kanetar wird die Shiftinvarianz, durch komplexe Logarithmen oder Konzentrische Abtastung wird die Rotationsinvarianz erzielt. In einer weiteren Stufe wird eine vom Bildmaterial abhängige Informationsreduktion durchgeführt.

Bei dem von Kohonen bezeichneten "associative mapping" spannen die Merkmalvektoren der Referenzen eine orthogonale Basis im Merkmalsraum auf. Durch eine orthogonale Projektion des Merkmalvektors des Musters in diese Basis, läßt sich das Muster durch eine Linearkombination der Basisvektoren beschreiben. Jeder Koeffizient der Linearkombination bildet das Maß für die Ähnlichkeit mit der entsprechenden Referenz. Das Muster wird der zum maximalen Koeffizienten zugehörigen Referenz zugewiesen.

Die vorgestellten drei Arbeiten [80, 18, 12] gehen jeweils davon aus, daß vor der eigentlichen Erkennung in der vorgeschalteten Stufe eine Zeichenvereinzelung durchgeführt wird. Bei den in Kap. 4 dargestellten Verfahren wird von solch einer zweistufigen Vorgehensweise nicht generell ausgegangen. Die Segmentierung einer Schriftzeile oder Segmentierung von Zeichen (Vereinzelung) wird vielmehr als Einschränkung des Suchraumes verstanden.

Vergleichende Aussagen über die Empfindlichkeit der Verfahren bezüglich Beleuchtungsänderungen oder anderer Störungen können nicht getroffen werden.

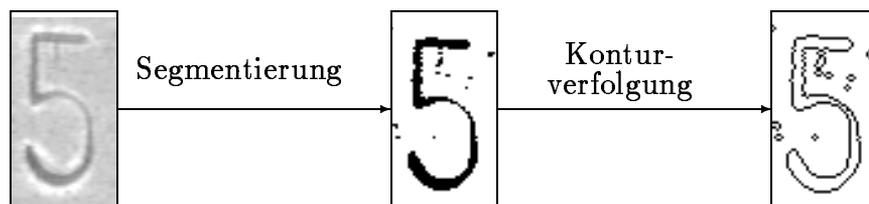
### 3.4 Liniendarstellungen von Zeichen

Ein in einem ikonischen Bild (Grauwertbild) dargestelltes Schriftzeichen läßt sich ohne wesentlichen Informationsverlust auch durch ein Linienmuster wiedergeben. Zur Darstellung wird häufig die Konturlinie oder die Mittelachse bzw. Skelettlinie des Zeichens verwendet. Unter einem Skelett versteht man die ein Bildpunkt dicke Liniendarstellung von flächenhaften Objekten; beim vorliegenden Datenmaterial also Zeichen bzw. Zeichenteile.

Ein naheliegender Ansatz zur Zeichenerkennung ist daher, zunächst das ikonische Bild in ein Linienmuster zu überführen und anschließend eine Analyse des Linienmusters vorzunehmen. Auf dem Gebiet der Zeichenerkennung wurden zur Analyse von linienhaften Mustern viele Beiträge veröffentlicht (siehe z.B. [90]) und ausgereifte Verfahren entwickelt. Im folgenden werden einige einfache Vorgehensweisen zur Linienfindung an dem aufgenommenen Datenmaterial dargestellt und typische Probleme aufgezeigt.

#### 3.4.1 Linienfindung durch Segmentierung des Grauwertbildes

Eine einfache Vorgehensweise zur Überführung eines in ikonischer Form dargestellten Zeichens in seine Konturlinien ergibt sich durch die Segmentierung mittels einer Schwellwertoperation und anschließender Verfolgung des Segmentrandes (Abb. 3.1).



**Abb. 3.1:** Beispiel zur Überführung eines ikonischen Bildes in seine Konturlinie durch Segmentierung (Schwellwertoperation) und Konturverfolgung.

Mit einer Variation des Schwellwertes variiert auch die Fläche bzw. Dicke des segmentierten Zeichenvordergrundes. In Abb. 3.2 sind die aus dieser Variation resultierenden unterschiedlichen Konturlinien für fünf Schwellwerte nebeneinander dargestellt.

Da die "Strichdicke" eines Schriftzeichens i.allg. keine strukturelevante Information trägt, überführt man das segmentierte Bild zur Zeichenerkennung oftmals in seine Skelettlinie (Abb. 3.3). Verfahren zur Verdünnung sind in der Literatur ausführlich diskutiert; grundlegende Beschreibungen findet man z.B. bei Pavlidis [67, 68].

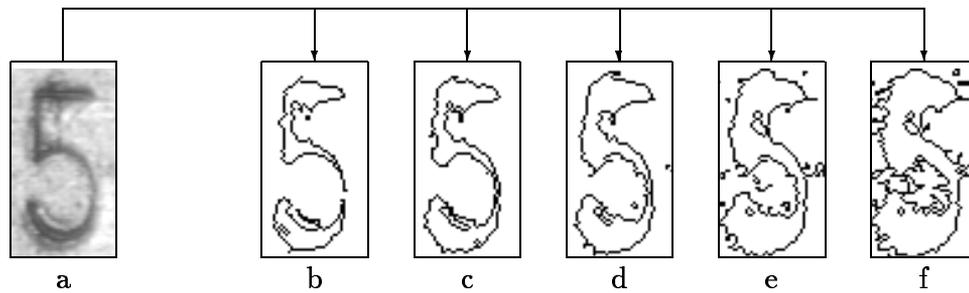


Abb. 3.2: Konturlinien bei unterschiedlichen Schwellwerten.  
a) Ikonisches Bild, b-f) Konturlinien von a) durch Segmentierung  
mit den Schwellwerten  $g_{th} = \{148, 162, 176, 190, 197\}$

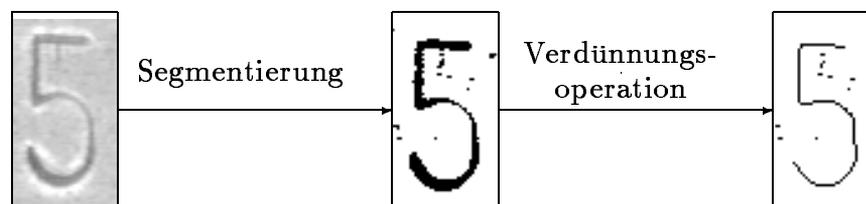


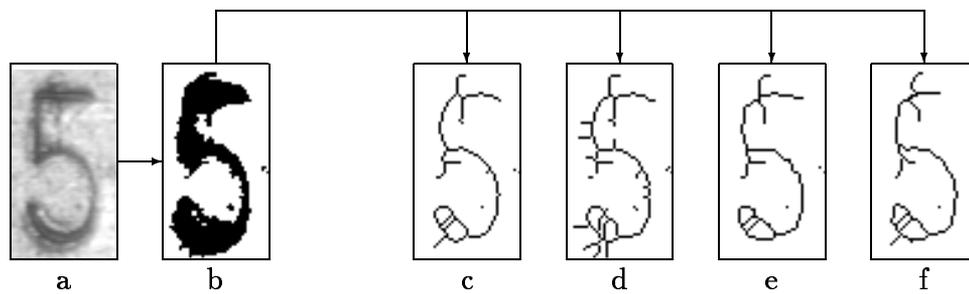
Abb. 3.3: Beispiel zur Überführung eines ikonischen Bildes in seine Skelettlinie durch Segmentierung (Schwellwertoperation) und Verdünnung.

An die Verdünnungsoperation wird die topologische Forderung gestellt, die Verbindung der Originalfigur zu erhalten. Ohne diese Forderung könnten Verbindungen einer acht sonst so aufgetrennt werden, daß das Skelett von dem einer drei nicht zu unterscheiden ist.

Im allgemeinen erhofft man sich mit der durch die Verdünnungsoperation einhergehenden Datenreduktion eine Vereinfachung der anschließenden Analyse. Im ungestörten Fall verliert man keine der für die Strukturanalyse relevanten Informationen.

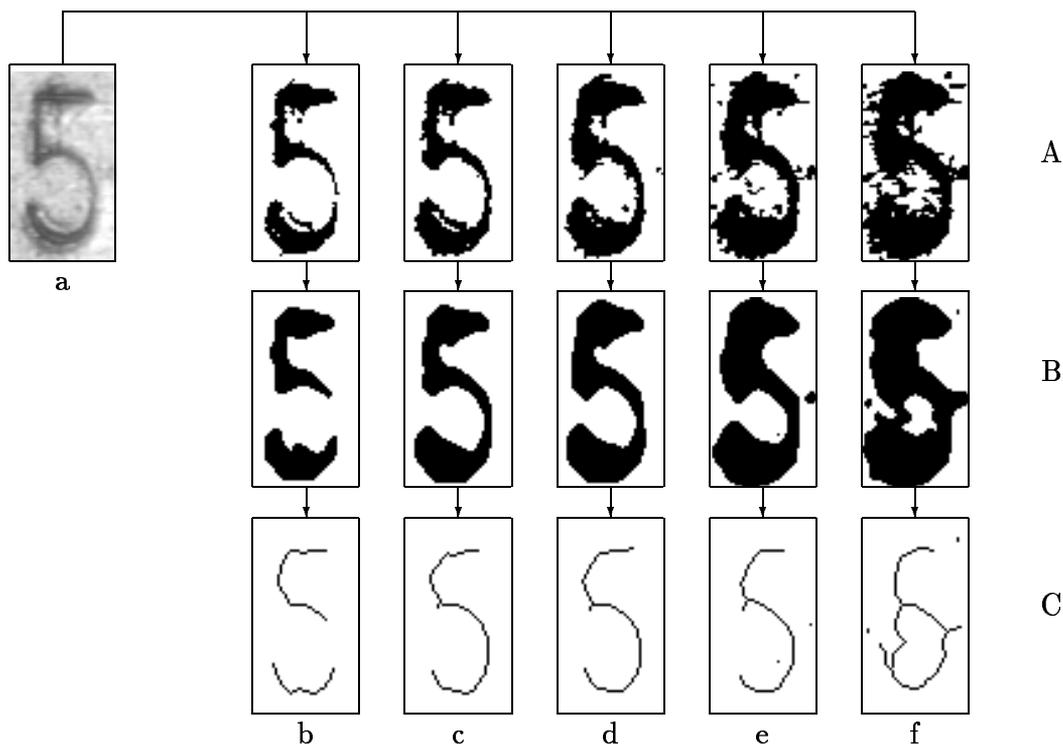
Das in Abb. 3.3 dargestellte segmentierte Bild der Ziffer fünf, wie auch die Skelettlinie, weisen nur geringe Störungen auf. Solch ein Ergebnis läßt sich nicht bei allen Aufnahmen durch Wahl eines günstigen Schwellwertes erzielen. Wie die Konturlinien aus Abb. 3.2 zeigen, kann der Zeichenrand sehr rauh sein und der Zeichenvordergrund Löcher beinhalten. Ermittelt man von derart gestörten segmentierten Bildern die Skelettlinie, so weist auch das Skelett Störungen auf (siehe Abb. 3.4).

Das segmentierte Bild in Abb. 3.4b (vergl. mit Abb. 3.2d), wurde mit den Verfahren nach Hilditch [35], Tsuruoka [87], Deutsch [13] und Tamura [82] bearbeitet. Die Ergebnisse der vier Verfahren zur Verdünnung sind in Abb. 3.4c-f dargestellt. Eine vergleichende Gegenüberstellung dieser Verfahren findet man in [83].



**Abb. 3.4:** Gegenüberstellung der Ergebnisse verschiedener Verdünnungsoperationen.  
 a) Ikonisches Bild, b) Segmentiertes Bild, c-f) Ergebnisse der Verdünnungsoperation von b) mit den Verfahren von Hilditch(c), Tsuruoka(d), Deutsch(e) und Tamura(f)

Die Ergebnisse zeigen, daß die Verfahren auf Störungen der Kontur unterschiedlich stark und auf Einschlüsse ähnlich reagieren. Eine Möglichkeit zur Reduktion der Störung des segmentierten Bildes bieten z.B. Füllalgorithmen oder morphologische Operationen wie Dilatation und Erosion. In Abb. 3.5 wurden die segmentierten Bilder (Abb. 3.5, Zeile A) durch abwechselnde Dilatation und Erosion geglättet (Abb. 3.5, Zeile B) und die Skelettlinien [35] bestimmt (Abb. 3.5, Zeile C).



**Abb. 3.5:** Gegenüberstellung von verdünnten Schriftzeichen bei unterschiedlichen Segmentierungsergebnissen. a) Ikonisches Bild, b)- f) Ergebnisse mit den Schwellwerten  $g_{th} = \{148,162,176,190,197\}$   
 Zeile A) Segmentiertes Bild, Zeile B) Ergebnis der Glättung durch Erosion und Dilatation, Zeile C) Ergebnis der Verdünnungsoperation

Bei der Segmentation können Störungen in Form von Unterbrechungen (Abb. 3.5b) oder zusätzlichen Verbindungen (Abb. 3.5f) auftreten, die von der nachfolgenden Glättung noch verstärkt werden und zu unbrauchbaren oder falschen Skelettlinien führen. Läßt man Störungen zu, so kann aus dem Skelett von Abb. 3.5Cb sowohl auf eine fünf mit einer Unterbrechung als auch auf eine sechs mit zwei Unterbrechungen geschlossen werden. Betrachtet man nur das Skelett in Abb. 3.5Cf, so würde ein unvoreingenommener Beobachter eher auf eine sechs als auf eine fünf schließen. Um Fehlinterpretationen zu vermeiden, muß zur Segmentation ein "richtiger" Schwellwert im regionalen Bereich des Zeichens gefunden werden. Dieser Wert hängt von den Aufnahmebedingungen ab und kann bei Helligkeitsschwankungen im globalen Bildbereich recht unterschiedliche Werte annehmen (sich z.B. Abb. A.4).

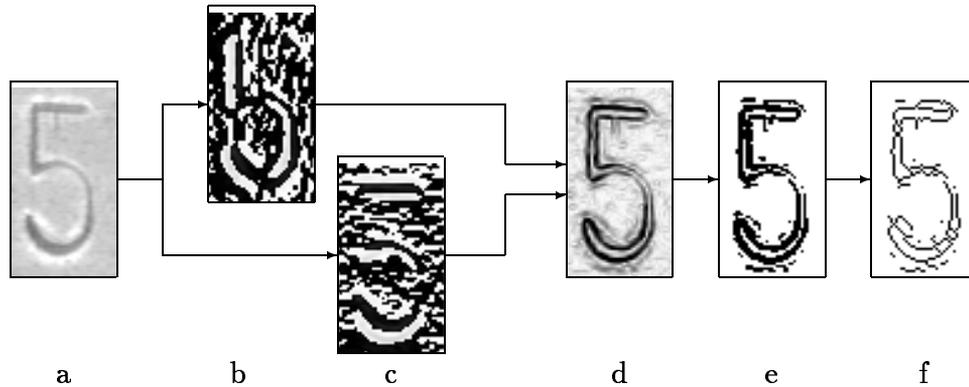
### 3.4.2 Linienfindung durch differenzierende Maskenoperation

Objektkanten in Grauwertbildern zeichnen sich oftmals durch lokale signifikante Änderungen des Grauwertes aus. Aufgrund der beleuchtungs- und vorlagenabhängigen Variation des mittleren Grauwertes von Zeichenvorder- und Zeichenhintergrund erscheint es sinnvoll, zur Detektion von Konturpunkten nicht den Grauwert selber, sondern die Grauwertänderung zu untersuchen. Zur Konturfindung werden daher häufig differenzierende Kantenoperatoren benutzt, die in Form von einer gleitenden lokalen Faltung mit Operatormasken realisiert werden. Man unterscheidet bei den differenzierenden Maskenoperationen zwischen Operationen auf der Basis des **Gradienten**, der die erste Ableitung des Bildes wiedergibt und Operationen auf der Basis des **Laplace-Operators**, der die zweite Ableitung wiedergibt.

#### Gradientenoperator

Bekannte Operatoren [48] zur näherungsweise Berechnung des Gradienten sind z.B. der Roberts-Operator, der Prewitt-Operator und der Sobel-Operator. Abb. 3.6 zeigt ein Bild zur Bestimmung der Kontur mit dem Sobel-Operator.

Die Faltung mit den beiden Sobel-Operatormasken liefert die x-Komponenten (Abb. 3.6b) und die y-Komponenten (Abb. 3.6c) des Gradienten, aus denen man den Betrag des Gradienten (Abb. 3.6d) bestimmt. Nach dieser "konturanhebenden" Operation läßt sich wie in dem vorherigen Beispiel durch eine Schwellwertoperation das Betragsbild des Gradienten segmentieren (Abb. 3.6e) und verdünnen (Abb. 3.6f). Der Schwellwert für die Segmentierung ist so einzustellen, daß die Konturlinien zusammenhängend sind und wenig Störungen aus dem Zeichenhintergrund auftreten. Man kann



**Abb. 3.6:** Beispiel zur Bestimmung der Konturlinie mit dem Sobel-Operator.  
a) Ikonisches Bild, b) x-Komponente des Gradienten, c) y-Komponente des Gradienten,  
d) Betrag des Gradienten e) Segmentiertes Bild, f) Skelett

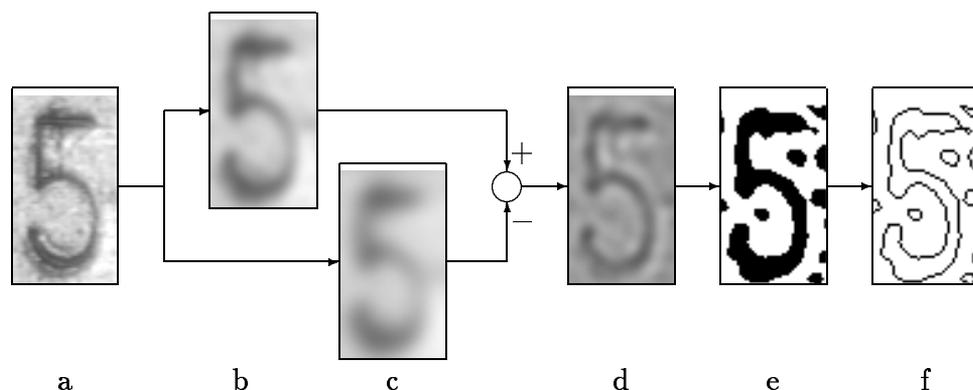
bei dieser Vorgehensweise jedoch generell nicht von geschlossenen Konturlinien ausgehen. Eine Zusammenstellung von Verfahren zur Konturaufbesserung, wie z.B. spezielle Verfahren zur Verdünnung von Gradientenbildern und Verfahren der Konturpunktverketzung findet man in [3].

### Laplace-Operator

Während bei den Gradientenoperatoren in der 1. Ableitung des Bildes die Maxima Indikatoren für Konturpunkte darstellen, sind es beim Laplace-Operator in der 2. Ableitung des Bildes die Nulldurchgänge (zero-crossing). Wegen der Rauschempfindlichkeit findet der Laplace-Operator ohne Verbindung mit glättenden Operatoren wenig Verwendung.

Marr und Hildreth [53] setzten den Laplace-Operator in Verbindung mit einem Gaußfilter zur Glättung ein. Der daraus resultierende Operator ist als  $\nabla^2 G$ -Operator oder LOG-Operator (Laplace Of Gaussian) bekannt. Der LOG-Operator läßt sich durch die Differenz zweier zweidimensionaler Gauß-Funktionen (DOG) annähern [53], deren Standardabweichungen  $\sigma_i$  und  $\sigma_e$  ein Verhältnis von 1.6 besitzen. Abb. 3.7 zeigt ein Beispiel zur Konturbestimmung durch Nullstellenbestimmung eines mit dem angenäherten  $\nabla^2 G$ -Operators (DOG) bearbeiteten Bildes.

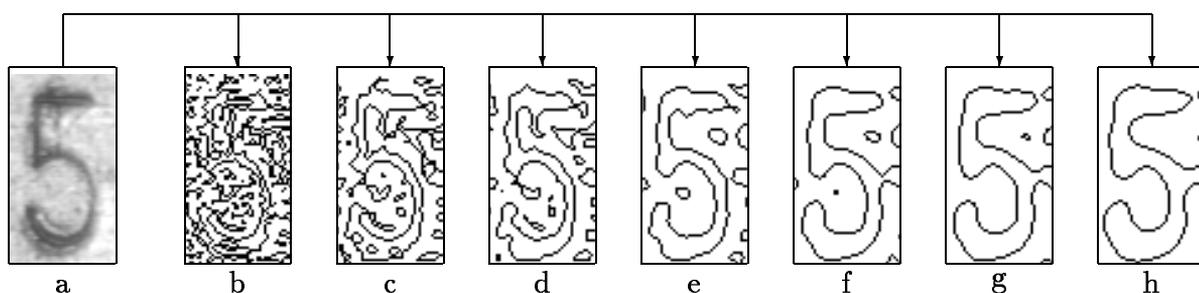
Die Zwischenergebnisse, die man durch Anwendung von zwei Tiefpaß-Operatoren mit gaußförmigen Profil ( $\sigma_e = 2.25, \sigma_i = 3.75$ ) erhält, sind in Abb. 3.7b und Abb. 3.7c dargestellt. Das Differenzbild (Abb. 3.7c) entspricht dem mit dem DOG-Operator bearbeiteten Eingangsbild (Abb. 3.7d). Der Intensitätswert 0 wird durch den Grauwert 148 dargestellt. Die Segmentation (Abb. 3.7e) wird aufgrund des Vorzeichens vor-



**Abb. 3.7:** Beispiel zur Bestimmung der Konturlinie durch DOG-Operation und Nullstellensuche.  
 a) Ikonisches Bild b) Gaußgefiltertes Bild ( $\sigma_e=2.25$ ) c) Gaußgefiltertes Bild ( $\sigma_i=3.75$ )  
 d) Ergebnis der angenäherten  $\nabla^2 G$ -Operation e) Segmentiertes Bild ( $g_{th}=148$ )  
 f) Konturlinie von e)

genommen. Durch Verfolgung des Segmentrandes bestimmt man die Nulldurchgänge (Abb. 3.7f).

Da ein einzelner Operator nicht auf alle Intensitätsänderungen im Bereich verschiedener Skalierungen optimal reagieren kann, verwendeten Marr und Hildreth den Laplace-Operator mit Gaußfunktionen unterschiedlicher Größe [53, 36]. Abb. 3.8 zeigt eine Gegenüberstellung der Nulldurchgänge bei Verwendung der Gaußfunktion mit unterschiedlichen  $\sigma$ .



**Abb. 3.8:** Konturlinien (Zero-crossings) des Beispiels von Abb. 3.7 bei unterschiedlicher Wahl der Filtermasken. a) Ikonisches Bild, b-h) Ergebnisse für  $\sigma_e = 0.75, 1.25, 1.75, 2.25, 2.75, 3.25, 3.75$

Die Beispiele von Abb. 3.5, 3.6 und 3.7 zeigen, daß die Überführung des ikonischen Bildes in ein für eine anschließende Analyse geeignetes Linienmuster bei entsprechender Wahl der Parameter möglich ist.

### 3.4.3 Problem bei der Linienfindung

Die hohe Qualität der gezeigten Zwischenergebnisse läßt sich jedoch mit den vorgestellten Verfahren nicht für alle Aufnahmen erzielen. Zum Vergleich wurde für die Beispiele die Aufnahme aus Abb. 2.3h als Eingangsbild verwendet. Die typischen Probleme sind in den Abb. 3.9, 3.10, 3.11 dargestellt.

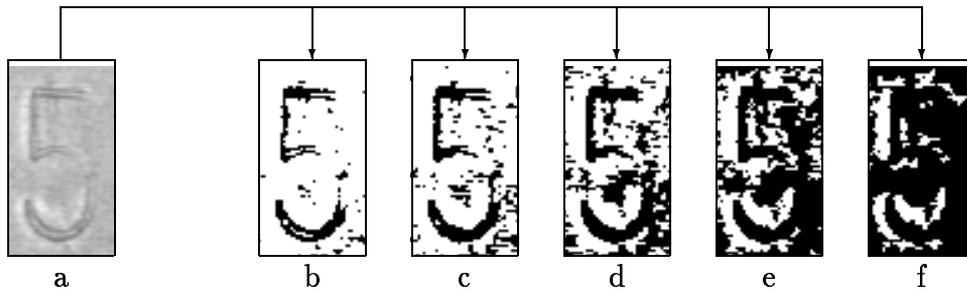


Abb. 3.9: Problematik bei der Segmentation durch Schwellwertoperation auf das Grauwertbild.  
a) Ikonisches Bild b-f) Segmentationsergebnisse durch Variation des Schwellwertes

Abb. 3.9 zeigt, daß sich keine Schwelle finden läßt, mit der es möglich ist, den Zeichenvordergrund vom Zeichenhintergrund zu trennen bei gleichzeitiger Forderung nach einem zusammenhängenden Zeichenvordergrund.

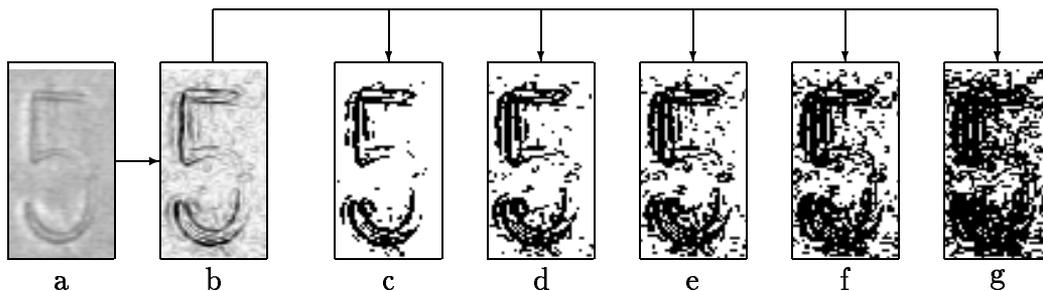
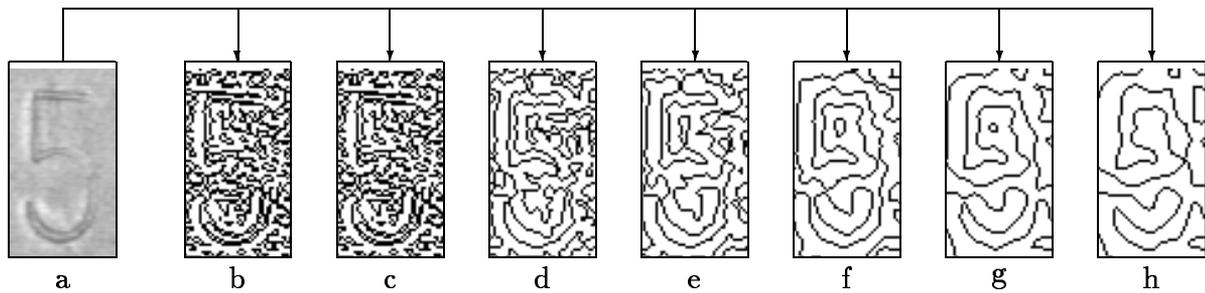


Abb. 3.10: Problematik bei der Segmentation durch Schwellwertoperation auf das Gradientenbild.  
a) Ikonisches Bild, b) Betrag des Gradienten,  
c-g) Segmentationsergebnisse durch Variation des Schwellwertes (vgl. Abb. 3.6e)

In Abb. 3.10b ist das Betragsbild des Gradienten dargestellt. Die Segmentierung mit unterschiedlichen Schwellen zeigt, daß weitgehend zusammenhängende Segmente bei geringer Hintergrundstörung durch Variation des Schwellwertes nicht gefunden werden können.

Auch in Abb. 3.11 ist erkennbar, daß eine zusammenhängende Zeichenkonturlinie bei geringer Zahl der Störkonturen durch Variation der Größe der Gaußverteilung nicht gefunden werden kann.



**Abb. 3.11:** Problematik bei der Bestimmung von Konturlinien (Zero-crossings) bei unterschiedlicher Wahl der Filtermasken. a) Ikonisches Bild, b-h) Ergebnisse für  $\sigma_e = 0.75, 1.25, 1.75, 2.25, 2.75, 3.25, 3.75$  (vgl. Abb. 3.8)

Mit den angesprochenen Verfahren war es nicht möglich, das vorliegende Datenmaterial in eine bildhafte Liniendarstellung zu überführen, die eine nachfolgende Analyse generell vereinfacht. Aus den letzten drei Bildbeispielen wird deutlich, daß man generell nicht von zusammenhängenden Segmenten, Konturlinien oder Skelettlinien ausgehen kann und daß der Zeichenhintergrund nicht frei von störenden Linien ist. Nachgeschaltete Verfahren müssen dementsprechend ein robustes Verhalten gegenüber diesen Störungen aufweisen. In Kap. 4.2 wird ein Verfahren dargestellt, mit dem eine Erkennung ausgehend von einer Liniendarstellung durchgeführt wird.

Es besteht auch die Möglichkeit, nicht nur von einer Liniendarstellung auszugehen, sondern eine ganze Sequenz von Liniendarstellungen, die durch Variation eines oder mehrerer Parameter erzeugt wird, zu Erkennung heranzuziehen. Ein Beispiel zur Auswertung einer Sequenz von Liniendarstellungen, die durch ein Mehrschwellenverfahren (entsprechend Abb. 3.2) erzeugt und mit dem in Kap. 4.3.1 beschriebenen System analysiert wird, ist in [81] angegeben.

Bei den hier dargestellten Vorgehensweisen werden in die Vorverarbeitung keine besonderen Informationen über die Struktur der zu analysierenden Objekte eingebracht. Auf einen Ansatz, bei dem objekt- und beleuchtungsbezogene Informationen in die Linienfindung eingebracht werden, ist in Kap. 3.3.1 hingewiesen.

Die Kenntnis über die Struktur der zu erwartenden Muster wird bei den in Kap. 3.5 und Kap. 4.3 vorgestellten Verfahren in unmittelbarer Form genutzt, indem durch Korrelation einer Maske mit dem ikonischen Bild ein direkter Mustervergleich durchgeführt wird.

### 3.5 Normierte maskierte Korrelation

Um die Ähnlichkeit bzw. Unähnlichkeit zwischen einem Muster G und einer Referenz R anzugeben, benutzt man ein Abstands- oder Fehlermaß. Dies erhält man z.B. durch paarweisen Vergleich der m Elemente von G mit R und Mittelung der absoluten Abweichungen.

$$d = \frac{1}{m} \sum_i |g_i - r_i| \quad (3-2)$$

Sollen große Abweichungen stärker gewichtet werden, wird häufig der *mittlere quadratische Fehler* als Maß verwendet:

$$d = \frac{1}{m} \sum_i (g_i - r_i)^2 \quad (3-3)$$

Dieses Abstandsmaß d läßt sich in drei Komponenten zerlegen:

$$d = \frac{1}{m} d_I - \frac{2}{m} d_{II} + \frac{1}{m} d_{III} \quad (3-4)$$

bzw.

$$d = \frac{1}{m} \sum_i g_i^2 - \frac{2}{m} \sum_i g_i \cdot r_i + \frac{1}{m} \sum_i r_i^2 \quad (3-5)$$

Benutzt man dieses Maß um das Muster G einem von verschiedenen Referenzmustern zuzuordnen, so bleibt der Term  $d_I$  konstant. Sucht man dagegen in einem Bild B, das größer als R ist, die Position des von R überdeckten Teilbildes G in B mit der besten Übereinstimmung, so kann man den Term  $d_{III}$  als konstant betrachten. Statt nun die Position von R in B durch Minimierung von d zu suchen, läßt sich dies auch durch Maximierung von  $d_{II}$  erzielen. Der Term  $\frac{1}{m} d_{II}$  wird auch als *Kreuzkorrelation* bezeichnet. Da jedoch hohe Intensitäten auch hohe Ähnlichkeiten vortäuschen, normiert man das Maß und erhält die *normierte Kreuzkorrelation*  $d_k$

$$d_k = \frac{\sum_i g_i \cdot r_i}{\sqrt{\sum_i g_i^2 \cdot \sum_i r_i^2}} \quad (3-6)$$

Die Schwarzsche Ungleichung

$$\left( \sum_i g_i \cdot r_i \right)^2 \leq \left( \sum_i g_i^2 \right) \cdot \left( \sum_i r_i^2 \right) \quad (3-7)$$

läßt erkennen, daß für positive Werte von  $g_i$  und  $r_i$  die *normierte Kreuzkorrelation* Werte zwischen  $0 < d_k \leq 1$  annimmt.

Faßt man  $R$  und  $G$  als zweidimensionale Bildmuster  $R = (r(u, v))$  und  $G = (g(x, y))$  mit  $(u \leq x, v \leq y)$  auf, läßt sich für jede Plazierung  $(i, j)$  des Referenzmusters im Bild  $B$  zu dem Teilbild  $G_{ij}$  das Maß für die Ähnlichkeit  $d_k(i, j)$  berechnen. Bezeichnet  $M_{ij}$  die Menge aller Punkte  $(x, y)$ , für die  $(x-i, y-j)$  zum Gültigkeitsbereich des Referenzmusters gehört und  $M_r$  selbst die Menge aller gültigen Punkte  $(u, v)$  des Referenzmusters, so erhält man

$$d_k(i, j) = \frac{\sum_{(x,y) \in M_{ij}} g(x, y) \cdot r(x-i, y-j)}{\sqrt{\sum_{(x,y) \in M_{ij}} (g(x, y))^2 \cdot \sum_{(u,v) \in M_r} (r(u, v))^2}} \quad (3-8)$$

Das Ergebnisfeld von  $d_k(i, j)$  weist allerdings eine hohe Gleichkomponente auf und ist nicht unabhängig gegenüber Störungen von  $B$  durch Überlagerung einer Gleichkomponente. Befreit man die Elemente  $g_i$  und  $r_i$  von den jeweiligen Mittelwerten  $\bar{g}$  und  $\bar{r}$ , erhält man statt der Korrelation  $\frac{1}{m} \sum_i g_i r_i$  die Kovarianz  $\frac{1}{m} \sum_i (g_i - \bar{g})(r_i - \bar{r})$ . Bei gleichem Normierungsansatz von Gl. 3-6 ergibt sich der *normierte Korrelationskoeffizient*  $k_k$ , der auch als empirischer Korrelationskoeffizient bezeichnet wird [89, 6].

$$k_k = \frac{\sum_i (g_i - \bar{g}) \cdot (r_i - \bar{r})}{\sqrt{\sum_i (g_i - \bar{g})^2 \cdot \sum_i (r_i - \bar{r})^2}} \quad (3-9)$$

bzw. mit  $\bar{g}_{ij}$  als Mittelwert des Teilbildes  $G_{ij}$

$$k_k(i, j) = \frac{\sum_{(x,y) \in M_{ij}} (g(x, y) - \bar{g}_{ij}) \cdot (r(x-i, y-j) - \bar{r})}{\sqrt{\sum_{(x,y) \in M_{ij}} (g(x, y) - \bar{g}_{ij})^2 \cdot \sum_{(u,v) \in M_r} (r(u, v) - \bar{r})^2}} \quad (3-10)$$

Der Korrelationskoeffizient, der nur definiert ist, wenn sowohl die Varianz der Referenz als auch die Varianz des aktuellen Bildes größer Null ist, hat einen Wertebereich von -1 bis +1. Für den Fall, daß  $k_k = +1$  oder  $k_k = -1$  wird, liegt eine vollkommene Korrelation vor. Im ersten Fall sind  $R$  und  $G_{ij}$  identisch, im zweiten Fall ist  $G_{ij}$  das Komplement (Negativ) von  $R$ . Für  $k_k = 0$  sind  $R$  und  $G_{ij}$  unkorreliert, also unähnlich. Faßt man  $R$  und  $G_{ij}$  als Punkte im  $m$ -dimensionalen Raum auf, läßt sich  $k_k$  auch als Winkelabstand  $\phi_d$  zwischen diesen beiden Vektoren interpretieren, denn es gilt für die Vektoren  $a$  und  $b$

$$\cos \phi_d = \frac{ab}{\sqrt{a^2 b^2}} = \frac{ab}{|a||b|} \quad (3-11)$$

Nach Umformung von Gl. 3-9 (siehe Gl. A-2) erhält man

$$k_k = \frac{\overline{g \cdot r} - \bar{g} \cdot \bar{r}}{\sqrt{(\overline{g^2} - \bar{g}^2)(\overline{r^2} - \bar{r}^2)}} = \frac{\overline{g \cdot r} - \bar{g} \cdot \bar{r}}{\sqrt{\sigma_g^2 \cdot \sigma_r^2}} = \frac{\overline{g \cdot r} - \bar{g} \cdot \bar{r}}{\sigma_g \cdot \sigma_r} \quad (3-12)$$

Die Werte  $\overline{r}, \overline{r^2}$  und  $\bar{r}^2$  sind für jede Referenz konstant und brauchen nur einmal berechnet zu werden. Das Ähnlichkeitsmaß  $k_k$  ist gegenüber einer linearen Transformation  $k_k(\alpha g + \beta, \gamma r + \delta) = \text{sgn}(\alpha, \gamma) k_k(g, r)$  invariant. D. h. man kann durch eine Transformation von  $r$  mit  $\gamma$  und  $\delta$ ,  $r$  so normieren, daß  $\bar{r}$  und  $\sigma_r = 1$  wird. Dadurch vereinfacht sich die Berechnung von  $k_k$  zu

$$k_k = \frac{\overline{g \cdot r_n}}{\sigma_g} \quad (3-13)$$

bzw.

$$k_k(i, j) = \frac{\frac{1}{m_r} \sum_{(x,y) \in M_{ij}} g(x, y) \cdot r(x-i, y-j)}{\sqrt{\frac{1}{m_r} \sum_{(x,y) \in M_{ij}} g(x, y)^2 - \left(\frac{1}{m_r} \sum_{(x,y) \in M_{ij}} g(x, y)\right)^2}} \quad (3-14)$$

Ein Sonderfall ergibt sich, wenn die Referenz  $R_i$  nur zwei Werte annimmt ( $r \in \{\rho_{\oplus}, \rho_{\ominus}\}$ ).

$$k_k = \frac{\overline{g \cdot r} - \bar{g} \cdot \bar{r}}{\sqrt{(\overline{g^2} - \bar{g}^2)(\overline{r^2} - \bar{r}^2)}} = \frac{\overline{g \cdot r} - \bar{g} \cdot \bar{r}}{\sqrt{(\overline{g \cdot g} - \bar{g} \cdot \bar{g})(\overline{r \cdot r} - \bar{r} \cdot \bar{r})}} \quad (3-15)$$

Bildet man zur Mittelwertberechnung getrennte Summen für die Elemente  $r$ , die den Wert  $\rho_{\oplus}$  und  $\rho_{\ominus}$  tragen sowie für die entsprechenden Elemente  $g$  (siehe Anhang Gl. A-4), so erhält man für den Zähler und den Nenner

$$\overline{g \cdot r} - \bar{g} \cdot \bar{r} = \frac{m_{\oplus}}{m} \frac{m_{\ominus}}{m} (\overline{g_{\oplus}} - \overline{g_{\ominus}}) (\rho_{\oplus} - \rho_{\ominus}) \quad (3-16)$$

$$\overline{r \cdot r} - \bar{r} \cdot \bar{r} = \frac{m_{\oplus}}{m} \frac{m_{\ominus}}{m} (\rho_{\oplus}^2 - \rho_{\ominus}^2) (\rho_{\oplus} - \rho_{\ominus}) \quad (3-17)$$

$$\overline{g \cdot g} - \bar{g} \cdot \bar{g} = \frac{m_{\oplus}}{m} \sigma_{\oplus}^2 + \frac{m_{\ominus}}{m} \sigma_{\ominus}^2 + \frac{m_{\oplus}}{m} \frac{m_{\oplus}}{m} (\overline{g_{\oplus}} - \overline{g_{\ominus}})^2 \quad (3-18)$$

Setzt man diese Ausdrücke in Gl. 3-15 ein, so kann  $k_k$  in folgender Form dargestellt werden:

$$k_k = \text{sgn}(\rho_{\oplus} - \rho_{\ominus}) \cdot \text{sgn}(\overline{g_{\oplus}} - \overline{g_{\ominus}}) \frac{1}{\sqrt{\frac{m}{m_{\ominus}} \left(\frac{\sigma_{\oplus}}{\overline{g_{\oplus}} - \overline{g_{\ominus}}}\right)^2 + \frac{m}{m_{\oplus}} \left(\frac{\sigma_{\ominus}}{\overline{g_{\oplus}} - \overline{g_{\ominus}}}\right)^2 + 1}} \quad (3-19)$$

Werden von den Bildmustern  $R$  und  $G_{ij}$  nicht alle Bildpunktpaare zur Bestimmung der Korrelation herangezogen, sondern bestimmte Bildbereiche durch eine entsprechende Maske zur Referenz oder durch Nullsetzen von Referenzwerten ausgeblendet, so spricht man von einer maskierten Korrelation.

## 4. Verfahren zur Zeichenerkennung

Im folgenden Kapitel werden drei Verfahren (Abb. 4.1) zur Erkennung in Metall geschlagener Zeichen vorgestellt. Das erste Verfahren ist ein rein numerisches Verfahren und basiert auf einem direkten Mustervergleich im ikonischen Bild mittels Korrelation. Im zweiten Verfahren wird aus dem ikonischen Bild eine Liniendarstellung gewonnen und mit einer Referenzliniendarstellung verglichen. Im dritten Verfahren wird ein syntaktischer Ansatz verfolgt und das Bild durch eine Menge von Objektprimitiven beschrieben. Die Erkennung erfolgt durch eine modellgestützte Analyse mit einem Produktionssystem.

Zur Aufgabe, Zeichen in einem Grauwertbild zu erkennen, gehört die **Positionsbestimmung** der Zeichen und die **Zuweisung** des Zeichens zu einer Klasse.

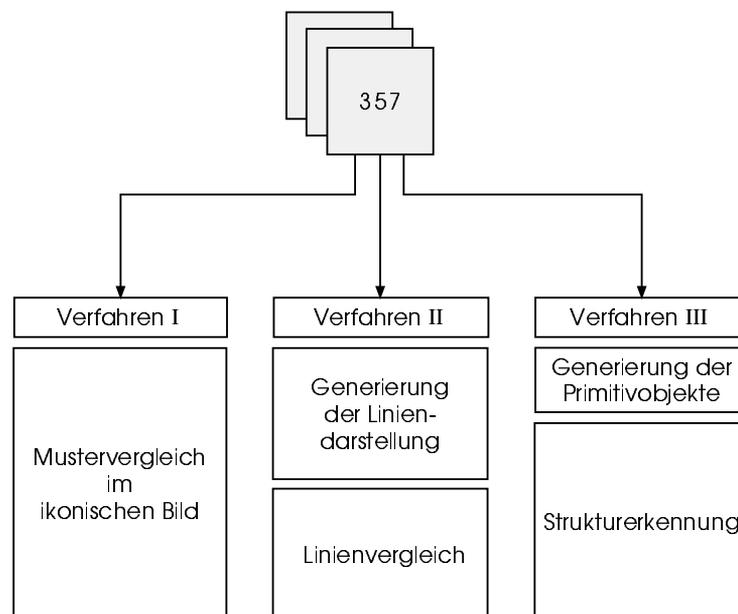


Abb. 4.1: Übersicht zu den implementierten Verfahren zur Zeichenerkennung

### 4.1 Ein numerisches Verfahren zur Zeichenerkennung (Verfahren I)

Das Erkennungssystem des Verfahrens I ist dreistufig aufgebaut (siehe Abb. 4.2). Es besteht aus einer Korrelationsstufe, einer Spitzen-Detektionsstufe und einem Maximumdetektor. Als Referenzen werden manuell erstellte Masken (Templates) der zu erkennenden Zeichen verwendet. Die Bildvorlage B wird mit jeder Referenzmaske (TPL 1

bis TPL  $n_r$  ) korreliert, wobei sich die parallele Anordnung der Korrelation mit der Struktur einer Filterbank vergleichen läßt. In den aus der Korrelationsstufe errechneten Korrelationsfeldern  $K_{k1}$  bis  $K_{knr}$  werden in der nachfolgenden Stufe von den Spitzen-Detektoren (SD 1 bis SD  $n_r$ ) hohe Korrelationswerte mit einer bestimmten Umgebung gesucht. Die Auswertung der Spitzen in den Spitzenmengen  $P_1$  bis  $P_{nr}$  und damit die eigentliche Klassenzuweisung erfolgt in der Klassifikationsstufe durch einen Maximum-detektor.

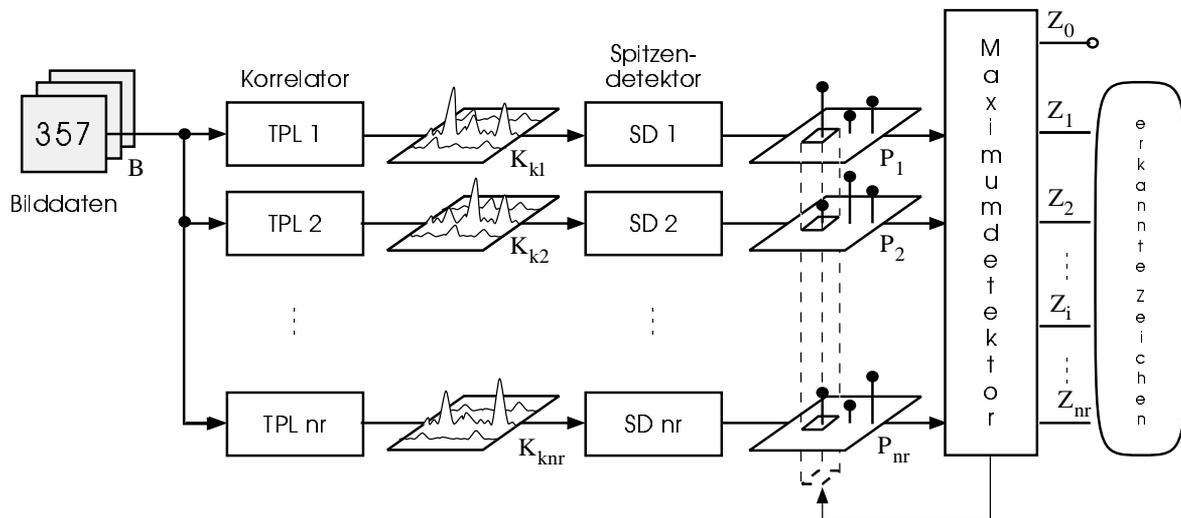


Abb. 4.2: Struktur des Erkennungssystems (Verfahren I)

#### 4.1.1 Generierung der Referenzen

Abbildung 4.3 zeigt beispielhaft die schrittweise Erstellung eines Templates für die Ziffer fünf. Mit einer Art "Zeichenprogramm" werden in einer dem Grauwertbild (Abb. 4.3a) überlagerten Bit-Ebene manuell Stützpunkte für das Skelett der Ziffer (Abb. 4.3b) gesetzt, die dann durch ein Polygon verbunden werden (Abb. 4.3c). Die Liniendicke der anfangs ein Bildpunkt starken Skelettlinie wird jeweils um ein Bildpunkt erhöht, bis die Ziffer verdeckt ist (Abb. 4.3d). Die entstandene Fläche  $A_{\oplus}$  wird dem Zeichenvordergrund zugeordnet. Durch weitere Verdickung der Linie wird eine Fläche  $A_{\circ}$  um die Fläche  $A_{\oplus}$  gelegt, die den Übergangsbereich (siehe auch Zeichenvorlage in Abb. 2.2) zwischen Zeichenvordergrund und Zeichenhintergrund repräsentiert. Wegen der unsicheren Zuordnung der Bildpunkte dieses Bereiches dient diese Fläche (Don't-Care-Bereich) zur Maskierung von Bildpunkten, die bei der Auswertung nicht berücksichtig

sichtigt werden. Um die Fläche  $A_o$  wird durch weitere Verdickung eine Fläche  $A_\ominus$  angelegt (Abb. 4.3e), die dem Zeichenhintergrund zugeordnet wird. Ein Template ist somit aufgeteilt in die drei Flächen  $A_\oplus$  (schwarz),  $A_o$  (transparent) und  $A_\ominus$  (weiß) (siehe Abb. 4.3f), wobei die zur jeweiligen Fläche zugehörige Anzahl von Bildpunkten durch  $n_{r\oplus}$ ,  $n_{r_o}$ ,  $n_{r\ominus}$  bezeichnet wird.

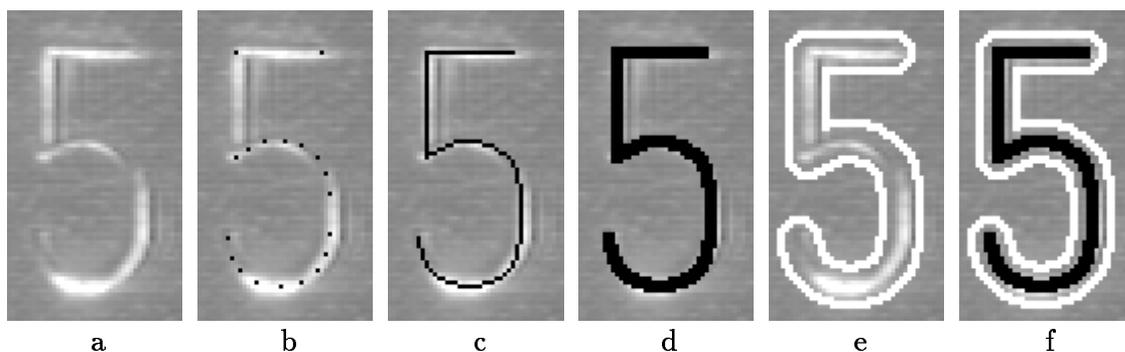


Abb. 4.3: Generierung der Referenzdarstellung (Template) für die Ziffer fünf auf unterlegtem Grauwertbild. a-f) siehe Text

#### 4.1.2 Korrelation

Die Korrelation wird in normierter und maskierter Form durchgeführt (siehe Kap. 3.5). Für jede Position des Templates auf der Bildvorlage wird ein Korrelationswert  $k_k$  berechnet, wobei sich die Menge der Korrelationswerte  $K_k$  wieder als Bild darstellen läßt. Zur Berechnung des Korrelationskoeffizienten werden die Bildpunkte der Vorlage herangezogen, die durch die Flächen  $A_\oplus$  und  $A_\ominus$  maskiert werden. An den Korrelationspositionen, an denen Vorlage und Referenz übereinstimmen, findet man hohe positive Korrelationswerte, bzw. an den Positionen, an denen das Negativ der Referenz vorliegt, hohe negative Korrelationswerte. Da das Vorzeichen der Grauwertdifferenz zwischen Zeichen und Hintergrund von der Beleuchtung abhängt und für die Erkennung nicht ausschlaggebend ist, genügt es, die Beträge der Korrelation zu untersuchen.

Ein Beispiel zur Korrelation eines Kreisringes mit dem entsprechenden Template ist in Abbildung 4.4 zu sehen. Die oberen Bilder zeigen das Muster (Abb. 4.4a), die Referenz (Abb. 4.4b) und das Korrelationsergebnis (Abb. 4.4c). Je dunkler der Grauton eines Bildpunktes des Korrelationsbildes ist, desto höher ist der Betrag des Korrelationskoeffizienten für die entsprechende Position der Referenz auf der Mustervorlage. Unter Vorlagenmuster und Korrelationsergebnis sind zur Veranschaulichung jeweils die Grauwertbilder in Form eines Gebirges dargestellt (Abb. 4.4d,e). Die Spitze in der Mitte von Abb. 4.4e gibt die Position der Referenz mit der besten Übereinstimmung an.

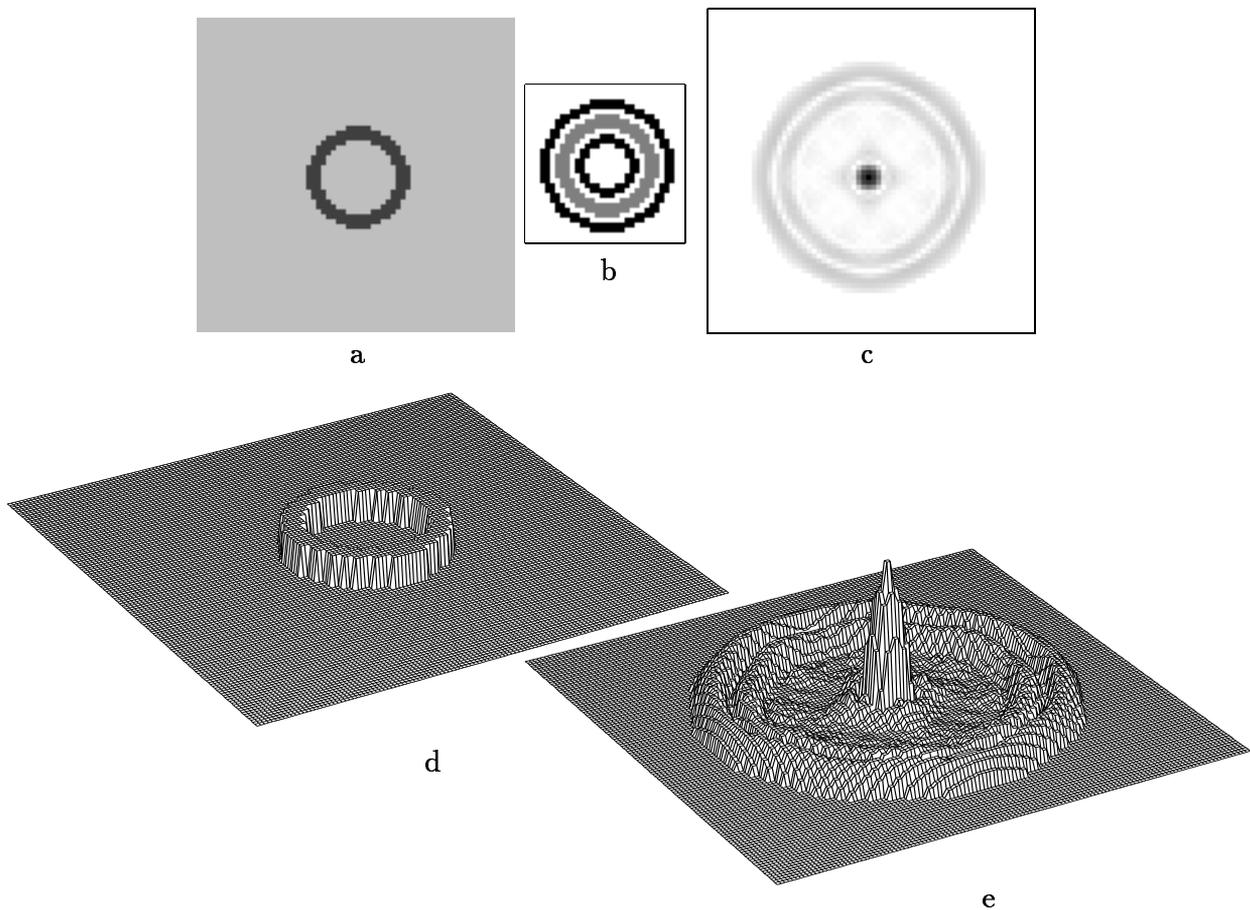


Abb. 4.4: Korrelation eines Kreisringes mit der entsprechenden Referenz.  
 a) Muster (Ausschnitt), b) Referenz, c) Korrelationsfeld  $|K_k|$  (Ausschnitt),  
 d) Grauwertgebirge von a, e) Grauwertgebirge von c

### 4.1.3 Suche der Korrelationsspitzen

Zur Positionsbestimmung eines Zeichens im Grauwertbild werden im Korrelationsfeld Spitzen gesucht. Dies geschieht durch Auswertung des **Betrages** des zu jeder Maskenposition ermittelten Korrelationskoeffizienten und der **Form** des Korrelationsgebirges.

Zunächst werden alle Maskenpositionen  $(i,j)$  durchlaufen und geprüft, ob der Betrag des Korrelationskoeffizienten  $k_k$  über einem vorgegebenen Schwellwert  $k_{kT}$  liegt.

$$|k_k(i, j)| \geq k_{kT} \quad (4-1)$$

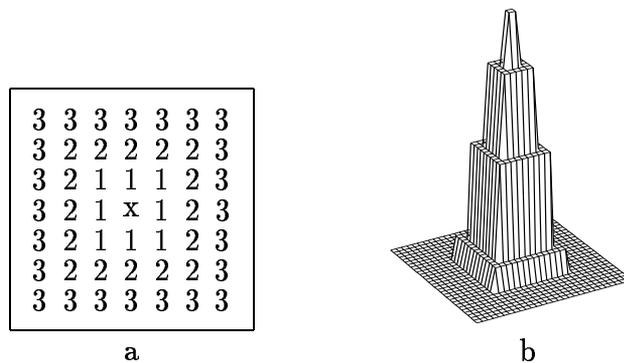
Es werden nur diejenigen Bildpunkte weiter untersucht, die dem Schwellwertkriterium (Gl. 4-1) genügen.

Zur weiteren Auswertung wird die topologische Form des Korrelationsgebirges in der Umgebung der Spitzen herangezogen. Zur Untersuchung der Umgebung wird eine stufenförmig einhüllende Funktion vorgegeben, die fordert, daß die Beträge der Korrela-

tionskoeffizienten der Umgebung  $U_w$  um mindestens den Wert  $\delta_w$  gegenüber  $|k_k(i, j)|$  abfallen. Es wird eine Folge von Umgebungen  $U_1 \dots U_n$  mit den Differenzen  $\delta_1 \dots \delta_n$  angegeben.

$$\begin{aligned}
 (x, y) \in M_{U_1} \quad |k_k(i, j)| &\geq |k_k(x, y)| + \delta_1 \\
 (x, y) \in M_{U_2} \quad |k_k(i, j)| &\geq |k_k(x, y)| + \delta_2 \\
 &\vdots \\
 (x, y) \in M_{U_n} \quad |k_k(i, j)| &\geq |k_k(x, y)| + \delta_n
 \end{aligned}$$

Ein Beispiel für die Anordnung für drei Punktumgebungen  $U_1, U_2$  und  $U_3$  um den Bildpunkt  $x$  ist durch den Index von  $U$  in Abb. 4.5a dargestellt. Die graphische Darstellung einer einhüllenden Funktion ist beispielhaft für die Werte  $\delta_1 = 0.2 k_k$ ,  $\delta_2 = 0.5 k_k$ ,  $\delta_3 = 0.9 k_k$  in Abb. 4.5b dargestellt. Diese Werte eignen sich zur Detektion der in Abb. 4.4c,e dargestellte Spitze. Verwendet man statt dem hier dargestellten ungestörten (synthetischen) Muster ein gestörtes Muster, so ist eine breitere und weniger steil verlaufende Korrelationsspitze zu erwarten.



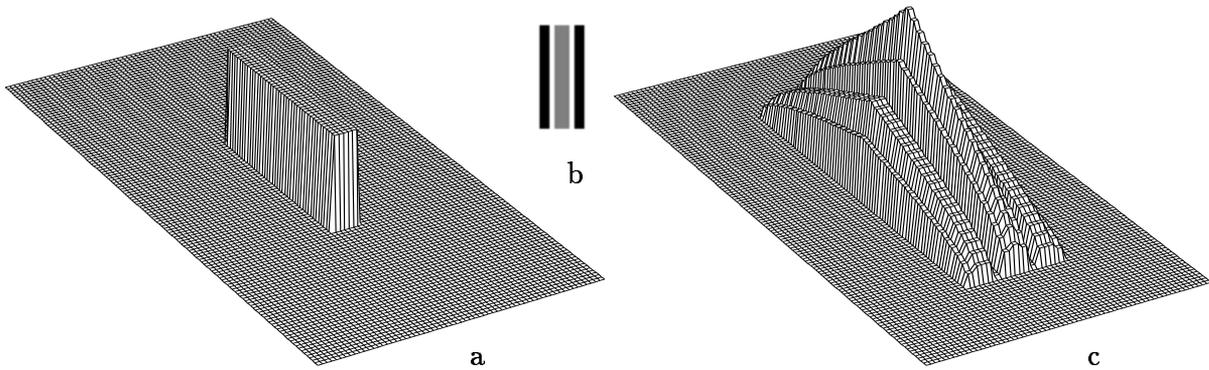
**Abb. 4.5:** Definition der Spitzenumgebung.

a) Beispiel für drei Umgebungen um den Mittelpunkt  $x$ , b) Einhüllende Funktion

Genügt ein Bildpunkt sowohl dem Schwellwertkriterium als auch dem Deltakriterium, wird er als Spitze bezeichnet.

Auch in ungestörten Bildvorlagen erhält man nicht immer eine solche deutliche Ausprägung der Spitzen wie in Abb. 4.4e. Korreliert man beispielsweise eine Referenz mit linienhafter Struktur (siehe Abb. 4.6b) mit der entsprechenden ungestörten Mustervorlage, so erkennt man in Abb. 4.6c, daß die Spitzen eine andere typische Umgebung besitzen. Ist der Balken des Musters (Abb. 4.6a) länger als die Referenz (Abb. 4.6b), entsteht nicht mehr ein einzelnes Maximum wie in Abb. 4.6c dargestellt, sondern ein sogenannter "Korrelationsrücken".

Für die Spitzen-Detektoren  $SD_i$   $i = 1 \dots n_r$ , sind die Umgebungen und die Formen der einhüllenden Funktionen entsprechend den Referenzen anzupassen.



**Abb. 4.6:** Korrelation eines strichförmigen Musters mit der entsprechenden Referenz  
a) Muster (Ausschnitt), b) Referenz, c) Korrelationsfeld  $|K_k|$  (Ausschnitt)

#### 4.1.4 Klassifikation

Durch die Klassifizierung werden den vorliegenden Mustern die Zeichenklassen  $Z \in \{z_1, z_2, \dots, z_{nr}\}$  zugewiesen. Die Zuordnung der Klasse zu einem Muster erfolgt durch einen binären Entscheidungsvektor  $e = (e_0, e_1, e_2, \dots, e_{nr})$ , wobei  $e_i=1$  und  $e_j=0$  für  $j \neq i$  die Entscheidung zugunsten der Musterklasse  $z_i$  bedeutet. Da nicht sichergestellt ist, daß es sich bei jeder detektierten Spitze um ein Zeichen handelt, wird die Entscheidung  $e_0$  für eine Rückweisung eingeführt.

Zur Klassifikation werden die Spitzen untersucht. Um deren Bildpunkt wird eine Fensterfläche  $A_w$  gelegt und untersucht, ob andere Spitzen als konkurrierende Lösungen vorhanden sind. Die Größe der Fensterfläche  $A_w$  orientiert sich dabei an Höhe und Breite der Zeichen. Für die in der Fensterfläche gefundenen Spitzen wird die Entscheidung  $e_i$  durch Zuweisung des Maximums getroffen.

$$e_i = 1 \quad \text{wenn} \quad e_i = \max_j e_j$$

$$e_i = 0 \quad \text{für} \quad j \neq i \quad j = 1 \dots n_r$$

Zur Zurückweisung wird ein heuristisches Kriterium verwendet. Man ermittelt den größten und zweitgrößten Korrelationswert  $k_k$  der in der Fensterfläche gefundenen Spitzen.

$$k_{k_{j1}} = \max_i k_{k_i}$$

$$k_{k_{j2}} = \max_{j \neq j_1} k_{k_j}$$

Liegt die relative Differenz von größten und zweitgrößten Korrelationswert unterhalb einer Schwelle  $k_{k\ominus}$

$$\frac{k_{k_{j1}} - k_{k_{j2}}}{k_{k_{j1}}} < k_{k\ominus} \quad (4 - 2)$$

wird das Muster zurückgewiesen und die Entscheidung  $e_0$  getroffen.

#### 4.1.5 Beispiel

Ein Beispiel zur Korrelation mit dem aufgenommenen Bildmaterial ist in den Abbildungen 4.7 und 4.8 illustriert. Abb. 4.7 zeigt eine Schriftzeile in einem Bildausschnitt, die Referenz der Ziffer fünf und das zugehörige Korrelationsergebnis.

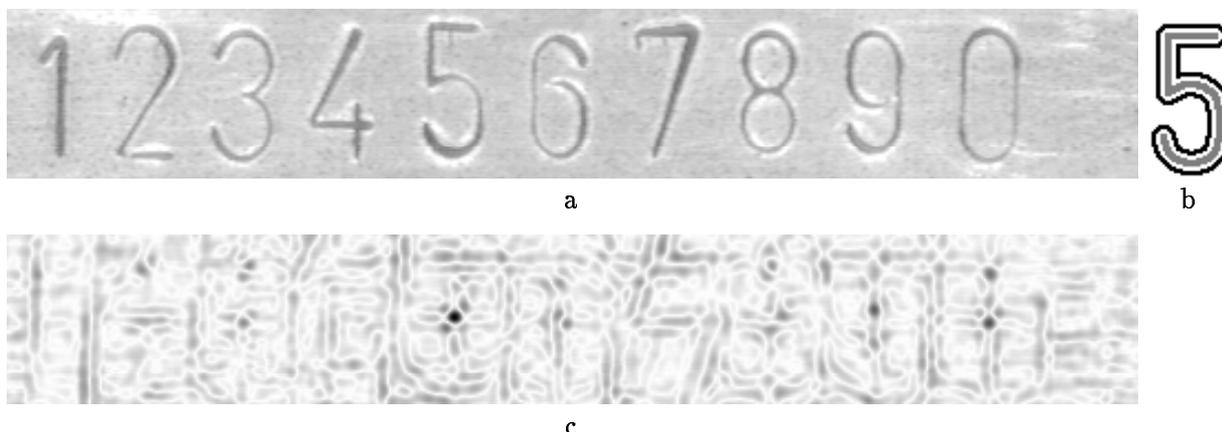


Abb. 4.7: Korrelation einer Schriftzeile (zehn Ziffern) mit dem Template fünf.  
a) Muster (Ausschnitt), b) Template, c) Korrelationsfeld  $|K_{k5}|$  (Ausschnitt)

Abb. 4.8 zeigt einen Bildausschnitt mit der Ziffer fünf, die verwendeten Referenzen und die zum Bildausschnitt zugehörigen Korrelationsergebnisse. Ausgeprägte Spitzen sind neben der Ziffer fünf auch bei den Ziffern drei, sechs, neun und null zu finden. Die gefundenen Maxima zu jedem Template sind in Abb. 4.9 gegenübergestellt.



Abb. 4.8: Korrelation der Ziffer fünf mit zehn Templates.  
 Aa) Muster (Ausschnitt), Zeile A) Template TPL 1 bis TPL 0,  
 Zeile B) Korrelationsfelder  $|K_{k1}|$  bis  $|K_{k0}|$  (Ausschnitt)

$|k_k|$

Referenzziffer

Abb. 4.9: Gegenüberstellung der Maxima  $|K_{k1}|$  bis  $|K_{k0}|$  aus dem Beispiel von Abb. 4.8.

## 4.2 Der elastische Linienvergleich (Verfahren II)

Häufig verwendet man bei Vergleichsverfahren nicht das ikonische Bild selbst, wie beispielsweise bei der Grauwertkorrelation, sondern eine symbolische Repräsentation davon. Dies kann z.B. eine linienhafte Darstellung des Bildinhaltes sein, die über eine Segmentation (Kap. 3.4) in Form von Konturlinien (Abb. 3.1) oder Skelette (Abb. 3.4.1) oder durch einen Liniendetektor (Kap. 4.1.2, Abb. 4.6) gewonnen wird.

Die Liniendarstellung des Bildes kann mit der Liniendarstellung eines anderen Bildes oder mit einem Modell verglichen werden. In dieser Untersuchung wird von einem Modell-Bildvergleich ausgegangen, d.h. die aus einem Muster extrahierte Liniendarstellung wird mit einer gespeicherten Referenzliniendarstellung verglichen.

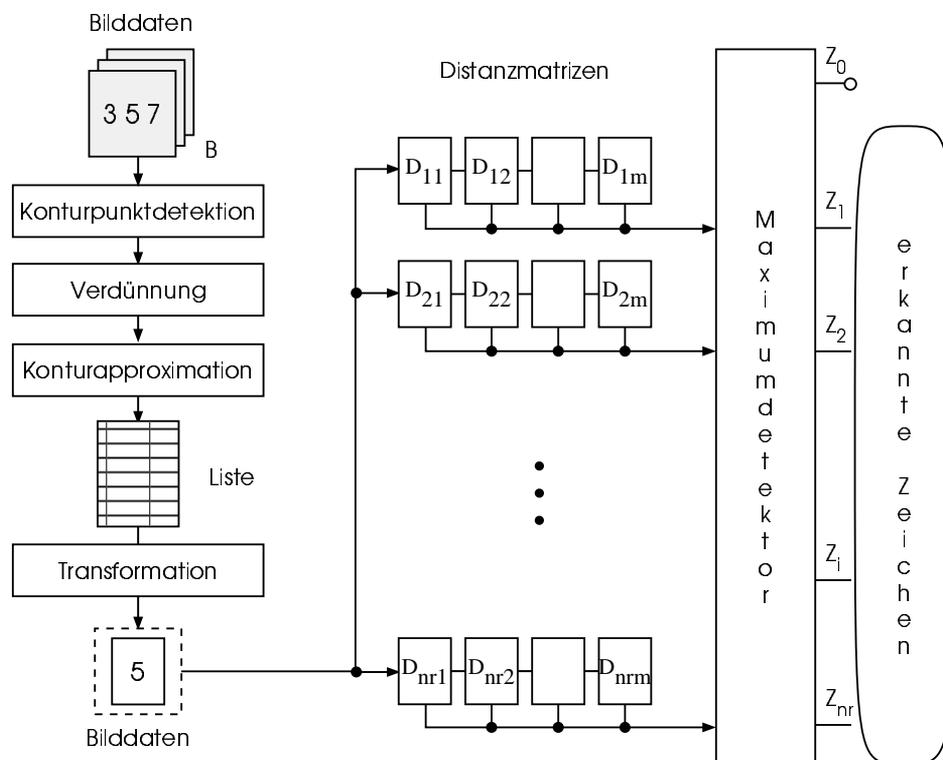


Abb. 4.10: Struktur des Erkennungssystems (Verfahren II)

Die Struktur des Erkennungssystems ist in Abb. 4.10 dargestellt. Die Bilddaten werden durch Konturpunktdetektion, Verdünnung und Konturapproximation in eine linienhafte Darstellung überführt und in einer Liste gespeichert. Für einen Vergleich eines Bildausschnittes mit verschiedenen Referenzliniendarstellungen werden relevante Linien der Liste selektiert, transformiert und in eine Bildmatrix projiziert. Der Vergleich dieser ikonischen Darstellung mit  $n_r$  Referenzdarstellungen wird mittels Distanzmatrizen durchgeführt, wobei für jede Referenz  $m$  Distanzmatrizen  $D$  existieren. Für die  $n_r$  Zweige

werden qualitative und quantitative Maßen bestimmt und in einem Maximumdetektor ausgewertet.

Das in dieser Arbeit eingesetzte Vergleichsverfahren [28] baut auf den Verfahren von Tasto und Block [85] und dem sehr ähnlichen Chamfer-Matching-Verfahren von Barrow et.al. [4] auf. Bei beiden Verfahren wird davon ausgegangen, daß die Referenzliniendarstellung und die Musterliniendarstellung sich durch eine Translation mit geringen Abweichungen zur Deckung bringen lassen. Zur Durchführung des Vergleiches wird auf die Musterliniendarstellung eine euklidische Distanz-Transformation (EDT) angewendet. Die Referenzliniendarstellung, die nur relevante Linien enthalten darf, wird dieser

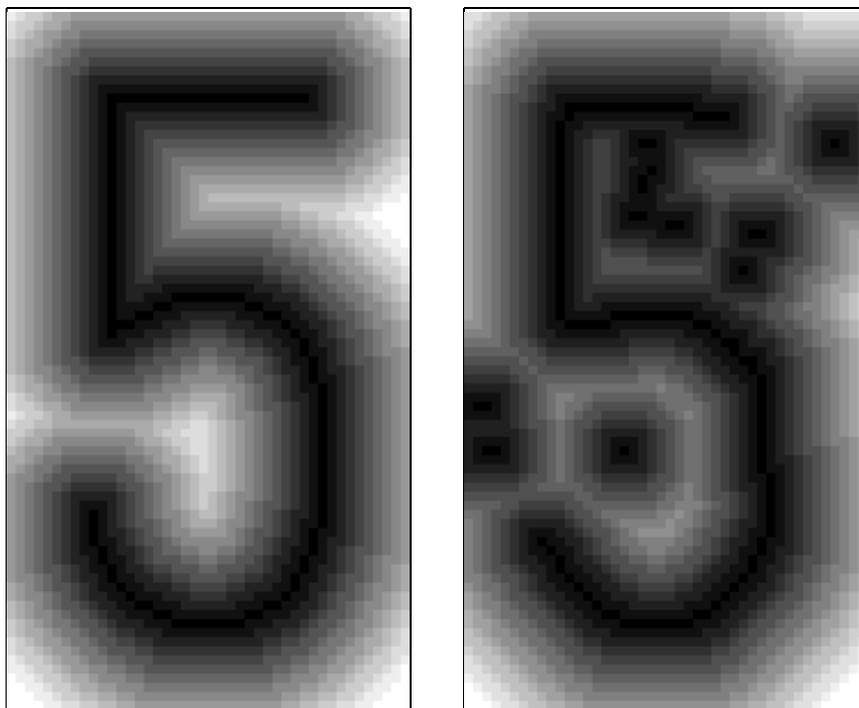


Abb. 4.11: Beispiele für Distanzmatrizen von Liniendarstellungen. Die Distanzwerte sind durch den Grauwert codiert. Die initiierende Punktmenge ist schwarz dargestellt; je höher der Distanzwert, desto heller der Grauton.  
a) Referenzliniendarstellung, b) Musterliniendarstellung mit Störungen

Distanzmatrix überlagert und die Übereinstimmung durch die Summe der Distanzwerte in den Punkten der Referenz gemessen. Dieses Maß zur Güte der Übereinstimmung läßt sich anschaulich durch die Differenzfläche zwischen den Linien darstellen. Sind die beiden Darstellungen identisch oder bildet die Referenz eine Untermenge des Musters, so ist die Distanzsumme Null.

Um typische Problemfälle für den Vergleich auszuschließen, müssen die Liniendarstellungen bestimmte Voraussetzungen erfüllen. Diese Voraussetzungen werden verletzt,

wenn durch Störungen die Liniendarstellungen unvollständig sind, zusätzliche Linien entstehen oder Kanten durch Konturlinien ungenau wiedergegeben werden. Insbesondere ergeben sich bei gestörten Mustern Probleme, wenn man vor dem Vergleich Referenz- und Musterliniendarstellung vertauscht und die EDT statt auf die Musterliniendarstellung auf die Referenzliniendarstellung anwendet. Dies läßt jedoch gerade bezüglich des Rechenaufwandes einen besonders effizienten Vergleich zu, da für feste Referenzen die Distanzmatrizen vorab berechnet werden können. Ein weiterer Nachteil der genannten Verfahren ist, daß einander entsprechende Teile der Liniendarstellungen nicht explizit angegeben werden können.

Das weiterentwickelte Verfahren von Groch [28], bei dem, aufbauend auf die EDT, die sogenannte erweiterte euklidische Distanztransformation (EEDT) eingesetzt wird, zeigt gegenüber den genannten Schwierigkeiten ein tolerantes Verhalten. Bei diesem Verfahren liefert die EEDT an jeder Position einer Distanzmatrix nicht nur wie üblich einen Distanzwert, der lediglich die Entfernung zu einem nächstgelegenen Objektpunkt angibt, sondern es werden die nächstgelegenen Objektpunkte, d.h. die Punkte, die den betreffenden Distanzwert erzeugt haben, erfaßt. Zwischen zwei Liniendarstellungen kann somit nicht nur eine Bewertung, sondern auch eine Punkt-zu-Punkt-Zuordnung für entsprechende Teile der Liniendarstellungen angegeben werden. Versieht man Linien bzw. Linienpunkte mit Attributen, so wird mit der Attributierung eine Möglichkeit gegeben, zu verhindern, daß unpassende Teile verglichen werden. Mit den Attributen können geometrische Eigenschaften (z.B. Orientierung) oder radiometrische Eigenschaften (z.B. mittlerer Kontrast) erfaßt werden.

Prinzipiell liefert der Linienvergleich nicht die geometrische Transformation um Liniendarstellungen zur Deckung zu bringen. Zur Bestimmung der Überdeckung müßten in einem Suchbereich alle Möglichkeiten der Translation, Rotation und Skalierung ausgewertet werden. Mit dem Verfahren können jedoch Hypothesen für die Transformation überprüft und bewertet werden. Sind beispielsweise zwei identische Liniendarstellungen geringfügig gegeneinander verschoben, so kann aus den Punkt-zu-Punkt-Zuordnungen ein Verschiebevektor abgeleitet werden.

#### 4.2.1 Bestimmung der Distanzmatrizen

Herkömmliche diskrete Distanztransformationen bestimmen eine Distanzmatrix, bei der die kürzeste Entfernung (entsprechend einer Metrik) zu der initiierenden Punktmenge  $P_R$  gebildet wird. Die Positionen der Referenzpunkte tragen die Distanz Null. Außer an diesen Punkten ist an keiner Position bekannt, wo die den Distanzwert verursachenden Punkte liegen.

Zur Berechnung der Distanzmatrix bietet sich der sequentielle Algorithmus nach Danielsen [11] oder der parallele iterative Algorithmus nach Yamada [92] an. Im folgenden wird der kompakt formulierte parallele Ansatz nach Yamada skizziert. Bei diesem Ansatz wird mit einer Iteration eine Sequenz von Distanzmatrizen generiert, wobei die Distanzwerte an den Matrixpositionen jeweils aus den Werten der Nachbarschaft der zuvor berechneten Distanzmatrix abgeleitet werden.

Eine Matrixposition wird mit  $z = (x, y)$  bezeichnet. Die Nachbarschaft einer Matrixposition wird durch den Vektor  $u_i$  mit  $i = 0, 1, 2, \dots, 8$  erfaßt.  $u_0$  stellt den Nullvektor dar, der auf den Matrixpunkt selber zeigt. Die Vektoren  $u_1$  bis  $u_8$  zeigen beginnend mit der rechten Position gegen den Uhrzeigersinn auf die Positionen der um den Mittelpunkt angelagerten Nachbarn.

Zur Bestimmung des Distanzmaßes werden die zweielementigen Vektoren

$$c(z) = (c_x(z), c_y(z))$$

eingeführt, die im folgenden mit Zeiger bezeichnet werden. Diese Zeiger geben die Anzahl vorzeichenbehafteter Rasterschritte in x- und in y-Richtung zum jeweiligen initierenden Punkt an. Der Betrag eines Zeigers ist definiert als:

$$|c(z)| = \sqrt{c_x^2 + c_y^2}$$

**Anfangsbedingung:** Die Anfangswerte für den Iterationsschritt  $t=0$  sind gegeben durch

$$c_1^0(z) = \begin{cases} (0, 0) & \text{für } z \in P_R \\ (\mathbb{N}', \mathbb{N}') & \text{sonst} \end{cases} \quad (4-3)$$

wobei  $\mathbb{N}'$  eine hinreichend große natürliche Zahl darstellt.

**Iteration:** Für  $t = 1, 2, 3, \dots$  lautet die Iterationsvorschrift:

$$c^t(z) = c^{t-1}(z + u_{i'}) + u_{i'} \quad (4-4)$$

wobei  $i'$  die Gleichung erfüllt

$$|c^t(z + u_{i'}) + u_{i'}| = \min_{i=0, \dots, 8} \{ |c^{t-1}(z + u_i) + u_i| \} \quad (4-5)$$

**Abbruchbedingung:** Die Abbruchbedingung für die Iteration lautet:

$$|c^t(z)| = |c^{t-1}(z)| \quad \text{für alle } z \quad (4-6)$$

Das distanztransformierte Bild  $D(z)$  ergibt sich dann aus den ganzzahlig gerundeten Werten  $d(z)$  mit

$$d(z) = |c^T(z)| = \sqrt{(c_x^T(z))^2 + (c_y^T(z))^2} \quad T : \text{Endwert von } t$$

Bei dieser Vorgehensweise werden zur EDT die Zeiger lediglich zur Bestimmung des Distanzwertes der Distanzmatrix  $D(z)$  benötigt und danach wieder verworfen.

Um nun eine Zuordnung des Distanzwertes zu dem jeweiligen initiierenden Punkt zu erhalten, speichert man die Zeiger in einer Matrix. Die Gleichung 4-5 kann für mehr als ein  $i$  erfüllt sein. Von betragsminimalen Zeigern, die auf denselben initiierenden Punkt hinweisen, wird der Zeiger mit dem Minimum von  $i$  gewählt. Weisen betragsminimale Zeiger auf unterschiedliche initiierende Punkte, so werden zu jeder Position  $z$  zusätzliche Zeiger gespeichert. Die Menge der gespeicherten Zeiger  $c(z)$  wird mit  $C(z)$  bezeichnet.

#### 4.2.2 Vergleich der Liniendarstellungen

$L_G$  sei die Liniendarstellung eines Musters, das mit der Liniendarstellung einer Referenz  $L_R$  zu vergleichen ist. Eine Liniendarstellung  $L$  besteht aus diskreten Linien  $l_i$ , wobei  $l_i$  aus der Punktmenge  $P_i$  mit den Punkten  $p_{ij}$  besteht.

Zu einer Liniendarstellung wird eine Menge von Attributen  $a_1, \dots, a_N$  mit den zugehörigen Mengen der Attributwerte  $A_1, \dots, A_N$  definiert. Jeder Linie  $l_i$  der Liniendarstellung wird eine Attributwertekombination zugewiesen.

Zum Vergleich werden für jede Attributwertekombination die Untermengen  $P_{Rk}$  und  $P_{Gk}$  von  $P_R$  und  $P_G$  gebildet, die diese Attributwertekombination besitzen. Für jede Untermenge  $P_{Rk} \subset P_R$  wird eine EDT durchgeführt, wobei die Menge  $P_{Rk}$  die initiierende Punktmenge bildet. Eine Untermenge  $P_{Rk}$  wird mit all den Untermengen  $P_{Gk}$  verglichen, deren Attributwerte kompatibel sind. Liegen die Attributwertemengen geordnet vor, so lassen sich z.B. zwei Attribute als kompatibel definieren, wenn sie gleich oder benachbart sind. Für den Vergleich werden für jeden Punkt  $p_i$  aus  $P_{Gk}$  die Menge seiner Initiatorpunkte  $P_{Ri}(p_i)$  bestimmt.

$$P_{Ri}(p_i) = \{ p_i + c(p_i) \mid c(p_i) \in C(p_i) \} \quad (4-7)$$

Ein diskreter Punkt  $p_i$  wird als akzeptiert in  $P_G$  markiert, wenn es mindestens einen Initiatorpunkt in  $P_R$  gibt, der verträgliche Attributwertekombinationen der Punktattribute trägt. Die Menge der Initiatorpunkte, die eine solche verträgliche Attributwertekombination tragen, wird als akzeptiert in  $P_R$  markiert.

Bei gestörten Bildvorlagen ist davon auszugehen, daß auch nicht relevante Linien zum Vergleich herangezogen werden. Da sie keine direkte Entsprechung in der Referenzliniendarstellung besitzen, weisen sie hohe Distanzwerte auf. Zur Vermeidung solcher Störungen werden Punkte, für deren Distanzwert  $d(p_i) > d_{pmax}$  gilt, ausgesondert.

Die Anzahl der akzeptierten Punkte in den Liniendarstellungen  $L_R$  und  $L_G$  werden mit  $P_R^*$  und  $P_G^*$  bezeichnet. Zur Beurteilung des Ergebnisses werden das qualitative Maß  $k_d$  und die quantitativen Maße  $k_R$  und  $k_G$  berechnet.

Die qualitative Bewertung ist wie folgt definiert:

$$k_d = \frac{\sum_{p_i \in P_G^*} d(p_i)}{P_G^*} \quad (4-8)$$

Durch  $k_d$  wird die Güte der Übereinstimmung erfaßt. Bei vollständiger Übereinstimmung der Liniendarstellungen (oder  $P_G \in P_R$ ) wird die Bewertung  $k_d$  Null; bei starker Abweichung wird  $k_d$  groß. Die quantitativen Bewertungen  $k_R$  und  $k_G$  sind wie folgt definiert:

$$k_R = \frac{P_R^*}{P_R} \quad k_G = \frac{P_G^*}{P_G} \quad (4-9)$$

Die beiden Größen  $k_R$  und  $k_G$  bewerten, wieviele Punkte der Liniendarstellung bezüglich der Referenz bzw. bezüglich des untersuchten Musters akzeptiert wurden. Wird eine Liniendarstellung vollständig akzeptiert, nimmt die entsprechende Größe den Wert eins an; je weniger Punkte akzeptiert werden, um so kleiner wird die entsprechende Größe.

Für ein Muster wird eine Entscheidung getroffen, wenn folgende Bedingungen alle erfüllt sind:

$$\min_i k_{di} < k_{dmin} \quad (4-10)$$

$$\max_i k_{Ri} > k_{Rmin} \quad (4-11)$$

$$\max_i k_{Gi} > k_{Gmin} \quad (4-12)$$

Es werden die Referenzen mit größtem und zweitgrößtem Wert für  $k_R$  ermittelt.

$$k_{R_{j1}} = \max_i k_{R_i} \quad (4-13)$$

$$k_{R_{j2}} = \max_{j \neq j1} k_{R_j} \quad (4-14)$$

Ist die Bedingung für den Schwellwert  $k_{R\ominus}$

$$k_{R_{j1}} - k_{R_{j2}} > k_{R\ominus} \quad (4-15)$$

erfüllt, wird das Muster der Klasse zugeordnet, die das Maximum von  $k_{R_i}$  aufweist. Ist die Bedingung nicht erfüllt, wird das Muster der Klasse der Referenzen  $R_{j1}$  und  $R_{j2}$  zugeordnet, die den größten Wert  $k_G$  besitzen.

### 4.2.3 Beispiel

Eine Realisation des Vergleiches zweier Liniendarstellungen soll an einem Beispiel dargestellt werden. Verglichen wird eine Liniendarstellung der Ziffer fünf mit der Liniendarstellung der entsprechenden Referenz. Das ikonische Bild (Abb. 4.12a) wird durch die Verarbeitungsschritte Konturpunktdetektion, Verdünnung, Konturpunktverkettung und Konturapproximation in die Liniendarstellung (Abb. 4.12d) überführt.

Für Vergleiche mit verschiedenen Lagepositionen wird ein Ausschnitt des Musters mit der Referenz verglichen. Verschiebt man den Ausschnitt, so geht man nicht von der ikonischen, sondern von der symbolischen Darstellung aus.

### 4.2.4 Konturpunktdetektion

Die Detektion der Konturpunkte wird mit dem Verfahren nach Haralick [32] durchgeführt. Das Bild wird in Ausschnitte zerlegt und in jedem Ausschnitt eine Approximation der diskreten Grauwertfunktion mit Hilfe orthogonaler Polynome durchgeführt. Bei einem Ansatz mit kubischen Polynomen lautet die kontinuierliche Approximationsfunktion  $f(x, y)$  :

$$f(x, y) = k_1 + k_2x + k_3y + k_4x^2 + k_5xy + k_6y^2 + k_7x^3 + k_8x^2y + k_9xy^2 + k_{10}y^3 \quad (4 - 16)$$

Die Polynom-Koeffizienten werden durch Maskenoperationen direkt aus dem Bild ermittelt [3, 32]. Hat die Funktion  $f(x, y)$  in der Richtung des steilsten Anstiegs in einer kleinen Umgebung des Ursprungs einen Wendepunkt, so wird der zu untersuchende Bildpunkt als Kantenpunkt markiert. Abb. 4.12b zeigt die detektierten Konturpunkte des Grauwertbildes von Abb. 4.12a.

### Verdünnung und Konturpunktverkettung

Die in der oben beschriebenen Weise detektierten Konturpunkte können zu Kontursegmenten mit einer Strichstärke von mehr als einem Bildpunkt führen. Daher wird eine Verdünnungsoperation mit dem Verfahren nach Deutsch [13] angeschlossen (siehe Abb. 4.13c). Dabei wird für jeden Konturpunkt des Binärbildes durch Untersuchung der 8er-Nachbarschaft geprüft, ob der Punkt gelöscht werden kann.

Für die nachfolgende Konturapproximation erfolgt eine Untersuchung der Konturpunkte, ob es sich um Endpunkte, Verbindungspunkte oder Kreuzungspunkte handelt. Die durch End- und Kreuzungspunkte begrenzten Konturpunkte werden in Ketten zusammengefaßt und geordnet in Listen eingetragen.

## Konturapproximation

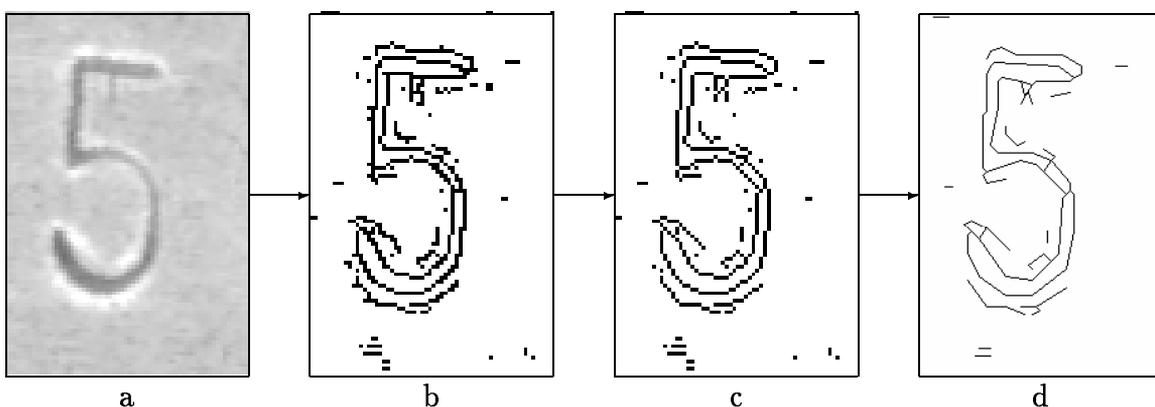
Ein Laufzeitvergleich verschiedener Verfahren der Approximation von ebenen Kurven durch Geradenstücke ist in [15] angegeben.

Zur Überführung der Konturpunktmenge in eine Liniendarstellung wird ein Splitting-Verfahren nach [14, 73] verwendet. Für jede Konturpunktmenge wird eine iterative Polygonapproximation durchgeführt.

Durch die Endpunkte der Konturpunktmenge  $p_b, p_e$  wird eine Gerade gelegt und von allen Punkten der Menge der Distanzwert  $d_{dist}$  zur Geraden berechnet. Liegt der maximale Distanzwert  $d_{max}$  unterhalb eines Schwellwertes  $d_{appr}$

$$d_{max} = \max\{d_{dist}\} \leq d_{appr}, \quad (4 - 17)$$

so bildet die Strecke  $\overline{p_b p_e}$  das Approximationspolygon der Konturpunktmenge. Ist die Bedingung Gl. 4-17 nicht erfüllt, bildet der Konturpunkt  $p(d_{max})$  einen neuen Scheitelpunkt  $p_s$  im Approximationspolygon. Durch die Punkte  $p_b, p_s$  und  $p_s, p_e$  werden erneut Geraden gelegt, die Bedingung Gl. 4-17 geprüft und gegebenenfalls weitere Scheitelpunkte bestimmt. Das Ersetzen einer Strecke durch zwei Strecken wird solange wiederholt, bis für alle Punkte der Menge Gl. 4-17 erfüllt ist. Der Splitting-Vorgang läßt sich durch einen Binärbaum beschreiben [73], bei dem das Splitting durch jeweils zwei Kanten und die entstandenen Strecken durch Knoten dargestellt werden. Für den Fall, daß alle Strecken des abschließenden Approximationspolygons, die den Terminalen des Baumes entsprechen, durch gleich viele Splittvorgänge entstanden sind, erhält man einen ausgewogenen Binärbaum (balanced tree).



**Abb. 4.12:** Schrittweise Überführung der ikonischen Darstellung in eine Liniendarstellung.  
a) Muster der Ziffer fünf, b) Menge der detektierten Konturpunkte von a,  
c) Konturpunkte nach Verdünnungsoperation, d) Liniendarstellung von c

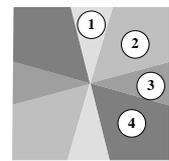
## Attributierung

Im vorliegenden Beispiel wird jeder Linie der Liniendarstellung  $L_G$  ein Attribut  $a_G$  zugewiesen, das eine Richtungsinformation trägt. Mit der Attributwertemenge  $A$

$$A = \{1, 2, 3, 4\}$$

des Attributes  $a_G$  werden 4 Hauptorientierungen unterschieden, die durch die folgenden Intervalle der Orientierung  $\phi$  einer Linie bestimmt sind:

$$a_G = \begin{cases} 1 & \text{wenn } 0 \leq \phi \leq 15 \\ 2 & \text{wenn } 15 < \phi < 75 \\ 3 & \text{wenn } 75 \leq \phi \leq 105 \\ 4 & \text{wenn } 105 < \phi < 165 \\ 1 & \text{wenn } 165 \leq \phi \leq 180 \end{cases}$$



Den Linien der Liniendarstellung  $L_R$  wird in gleicher Weise das Attribut  $a_G$  mit der Attributwertemenge  $A$  zugeordnet. Zwischen den Attributen  $a_R$  und  $a_G$  werden kompatible Attributwertekombinationen definiert, die in Tab. 4.1 aufgeführt und durch ein Pluszeichen (+) gekennzeichnet sind.

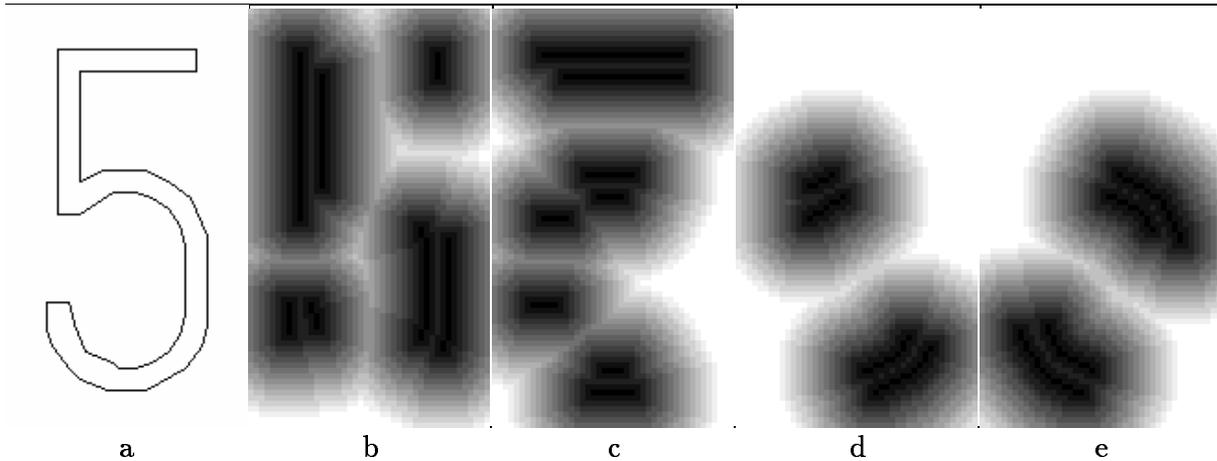
$a_R \ a_G$	1	2	3	4
1	+	+		+
2	+	+	+	
3		+	+	+
4	+		+	+

Tabelle 4.1: Kompatibilitätsmatrix der Attribute  $a_G$  und  $a_R$

Mit den vier Attributwerten von  $a_R$  ergeben sich auch vier Teilmengen der Liniendarstellung  $L_R$  mit den entsprechenden Punktteilmengen. Für jede dieser Punktteilmengen wird eine EEDT durchgeführt. In Abb. 4.13 sind die vier Distanzmatrizen graphisch dargestellt.

## Vergleich

Das Ergebnis des Vergleichs einer Liniendarstellung der Ziffer fünf mit den Liniendarstellungen verschiedener Muster ist in Abb. 4.14 dargestellt. In Zeile B sind die ak-



**Abb. 4.13:** Liniendarstellung und Distanzmatrizen. a) Referenzliniendarstellung der fünf, b) Distanzmatrix  $D_{51}$ , c) Distanzmatrix  $D_{52}$ , d) Distanzmatrix  $D_{53}$ , e) Distanzmatrix  $D_{54}$  (vgl. Abb. 4.11)

zeptierten Punkte  $P^*$  der Musterliniendarstellung und in Zeile C die zurückgewiesenen Punkte eingetragen.

Das Ergebnis des Vergleichs einer Liniendarstellung der Ziffer fünf mit den zehn Referenzziffern ist in Abb. 4.15 dargestellt. Neben der Liniendarstellung des untersuchten Musters ist in Zeile A die jeweils für die entsprechende Referenz akzeptierte Punktmenge  $P_R^*$  dargestellt. Darunter sind jeweils in Zeile B die zurückgewiesenen Punktmenge von  $P$  dargestellt.

Die nach Gl. 4-9 berechneten Abstandswerte  $k_R$  der in Abb. 4.15 dargestellten Ergebnisse sind in Form eines Balkendiagramms in Abb. 4.16 gegenübergestellt. Den höchsten Wert für  $k_R$  erhält man für die Ziffer fünf, gefolgt von den Werten der Ziffern sechs und drei.

Die Punkt-zu-Punkt-Zuordnung der Liniendarstellungen läßt sich bei einem nicht deckungsgleichen Linienvergleich anschaulich durch Verbindungslinien darstellen. Abb. 4.18a zeigt beispielhaft das Ergebnis eines solchen Vergleiches, bei dem die Lage des Musters gegenüber der optimalen Lage zur Referenz um jeweils vier Bildpunkte nach links und oben verschoben und um vier Grad verdreht wurde. Die schwarz eingetragene Punktmenge der Referenz (initiierende Punktmenge der Distanzmatrizen) ist die Punktmenge des Musters unterlegt. Korrespondierende Punkte der Darstellungen sind durch dünne Linien verbunden. Abb. 4.18a zeigt das Ergebnis bei Verwendung der in Abb. 4.13a dargestellten Referenz. Zur Erweiterung ist in Abb. 4.18b gezeigt, daß die gleiche Musterliniendarstellung auch mit einer Referenzliniendarstellung der Ziffer fünf in Form einer Skelettlinie verglichen werden kann. Nimmt man an, daß das Mini-

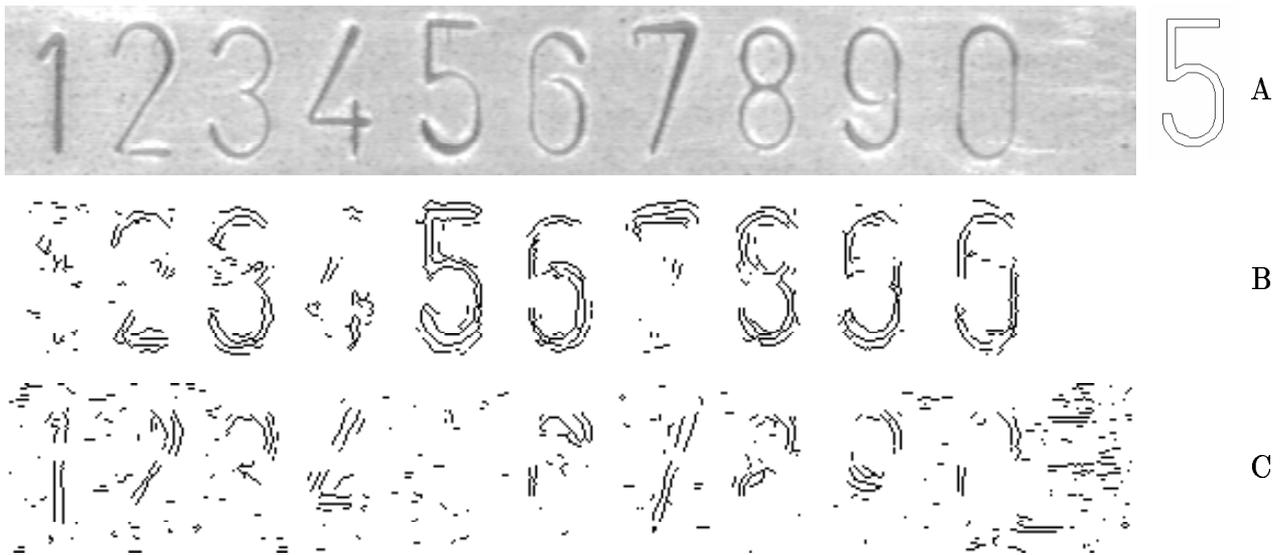


Abb. 4.14: Linienvergleich von verschiedenen Mustern (Ausschnitte aus der Musterliniendarstellung) einer Schriftzeile mit einer Referenz.  
 Zeile A) Muster (Ausschnitt) und Referenzliniendarstellung  
 Zeile B) Akzeptierte Punkte  $P_G^*$  der Musterliniendarstellung  
 Zeile C) Zurückgewiesene Punkte von  $P_G$

zum mittleren Abstand  $k_d$  bei einer zentrierten Lage von der Skelettlinie der Referenz und der Konturlinie des Musters vorliegt, so wird  $k_d$  im Gegensatz zur gleichen Beschreibung von Muster und Referenz nicht den Wert Null annehmen.

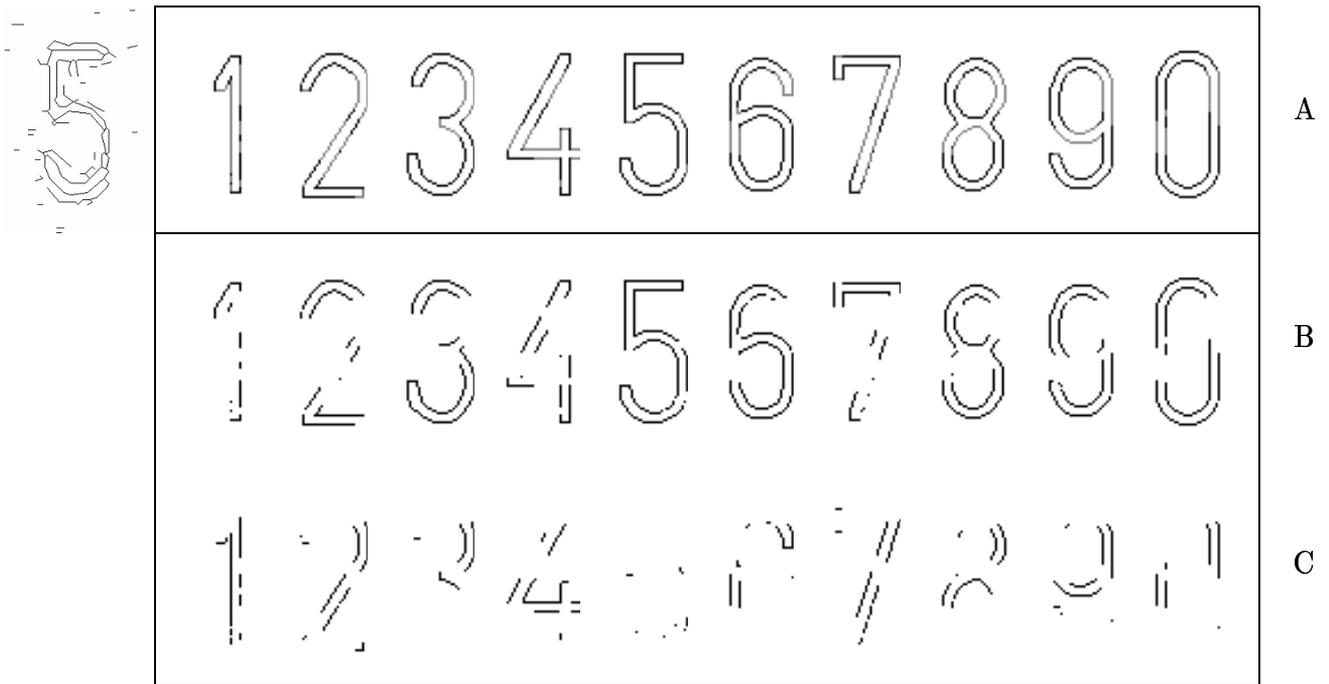


Abb. 4.15: Ergebnis der Linienvergleiche bezüglich der Referenzen.  
 Zeile A) Musterliniendarstellung und Referenzliniendarstellungen.  
 Zeile B) Akzeptierte Punkte  $P_R^*$  der Referenzliniendarstellung  
 Zeile C) Zurückgewiesene Punkte von  $P_R$

$k_R$

Referenzziffer

Abb. 4.16: Gegenüberstellung der Größen  $k_R$  für das in Abb. 4.15 dargestellte Beispiel

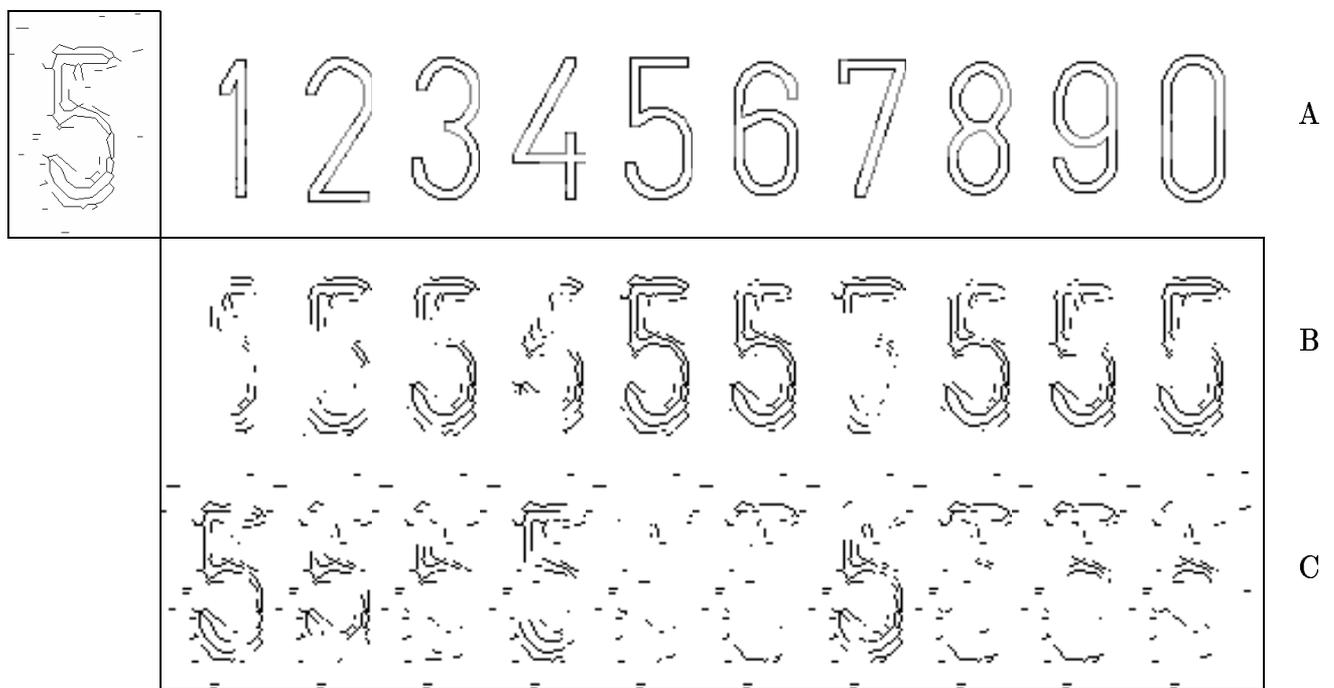


Abb. 4.17: Ergebnis der Linienvergleiche bezüglich der Muster  
 Zeile A) Musterliniendarstellung und Referenzliniendarstellungen  
 Zeile B) Akzeptierte Punkte  $P_G^*$  der Musterliniendarstellung  
 Zeile C) Zurückgewiesene Punkte von  $P_G$

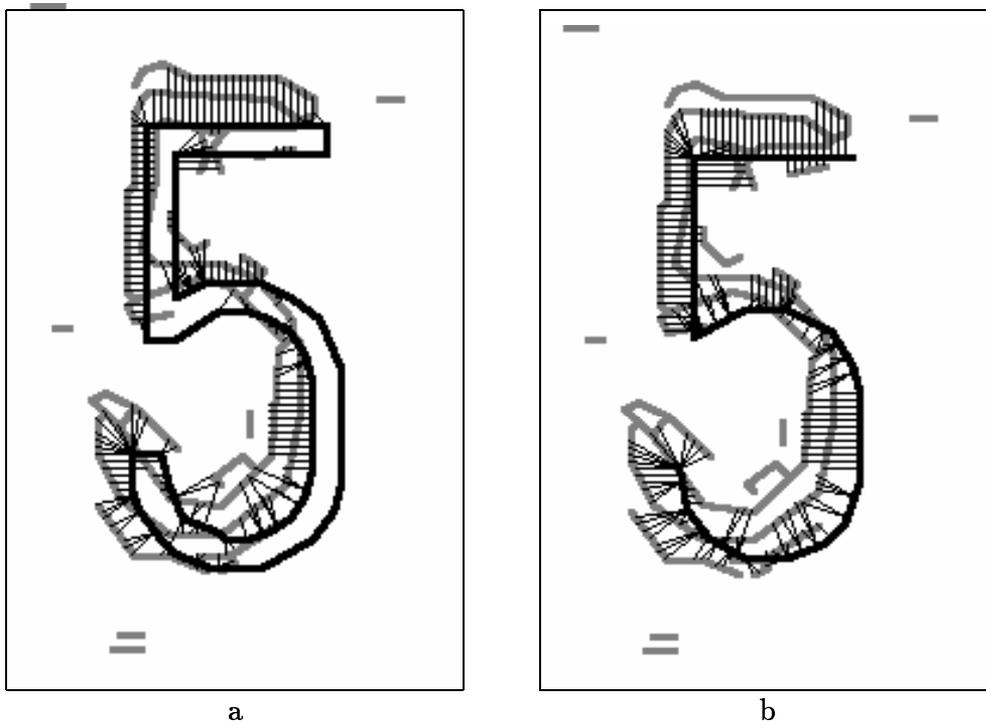


Abb. 4.18: Beispiele der Punkt-zu-Punkt-Zuordnungen.  
 Die Linien der Referenz sind schwarz dargestellt;  
 die Linien des Musters (siehe Abb. 4.12) sind grau unterlegt;  
 korrespondierende Punkte sind durch dünne Linien verbunden.  
 a) Referenz als Konturlinie, b) Referenz als Skelettlinie

### 4.3 Ein syntaktisches Verfahren zur Zeichenerkennung (Verfahren III)

Im Verfahren III wird zur Zeichenerkennung eine syntaktische Bildanalyse eingesetzt. Abb. 4.19 zeigt schematisch die Struktur des Erkennungssystems. Das in ikonischer Form vorliegende Bild  $B$  wird durch Korrelation und Suche nach Korrelationspitzen (siehe Kap. 4.1.2 und Kap. 4.1.3) in eine Menge von attributierten Primitivobjekten  $V_t$  überführt, die eine symbolische Repräsentation des Bildes darstellen. Für jeden Objekttyp der Primitivobjekte existiert ein Korrelationszweig. Die syntaktische Bildanalyse erfolgt in Form eines Synthesevorgangs, bei dem ausgehend von den Primitivobjekten über Teilobjekte die Zielobjekte zusammengesetzt werden. Welche Objekte sich wie zu komplexeren Objekten zusammensetzen lassen, wird durch die Produktionen einer attributierte Grammatik beschrieben.

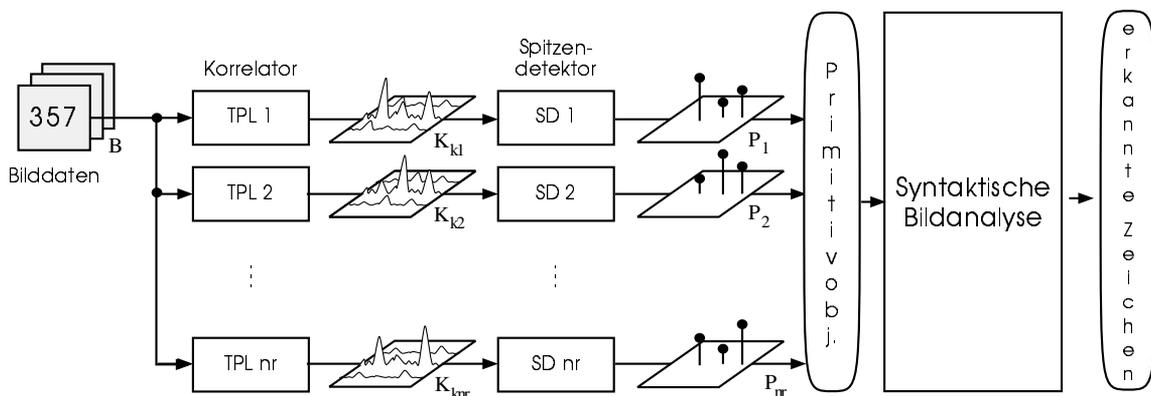


Abb. 4.19: Struktur des Erkennungssystems (Verfahren III)

Zur Aufstellung einer Grammatik bedarf es einer Modellvorstellung (Konzept) von Referenzobjekten und Objektprimitiven. Ausgehend von den bildlichen Mustern nimmt man eine Aufteilung des Musters in primitive Graphikelemente vor. Beispiele zur Aufteilung der Ziffer fünf in Liniensegmente sind in Abb. 4.20 dargestellt. Allgemeine systematische Ansätze zur Aufteilung in primitive Graphikelemente sind nicht bekannt. Bei der Auswahl der Elemente ist zu beachten, daß signifikante Unterschiede zwischen den Schriftzeichen erfaßt werden. Weiterhin wird man versuchen, die Gesamtzahl der Elemente, die alle auftretenden Klassen (Ziffern) beschreiben, möglichst gering zu halten. Beschreibt jedoch das Primitivobjekt bereits das Schriftzeichen vollständig, wird eine strukturelle Beschreibung bedeutungslos.

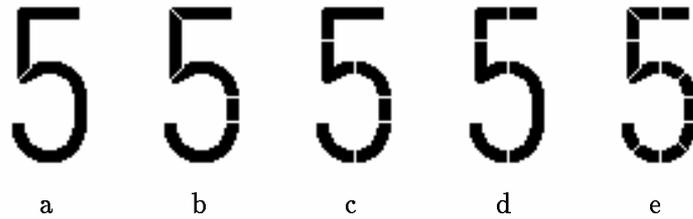


Abb. 4.20: Beispiele zur Aufteilung der Ziffer fünf in Liniensegmente.  
 a) zwei Segmente, b) fünf Segmente, c) sechs Segmente,  
 d) fünf Segmente, e) zwölf Segmente

### 4.3.1 Zeichenerkennung mit dem Produktionssystem BPI

Das von Lütjen [50] vorgestellte Produktionssystem BPI (blackboard oriented production system for image understanding) baut auf der Architektur allgemeiner Produktionssysteme und der Blackboard-Architektur auf.

#### Produktionssysteme

Produktionssysteme wurden erstmals von Post [70] vorgestellt. Sie sind durch drei Komponenten charakterisiert [7, 60], nämlich einer Menge von Produktionen, eine Datenbasis und eine Systemsteuerung. Eine Produktion ist eine Anweisung mit einem Bedingungsteil und einem Aktionsteil. Die Systemsteuerung überprüft die Datenbasis, ob der Bedingungsteil einer Produktion erfüllt ist und initiiert gegebenenfalls die Ausführung der Aktion. Durch die Ausführung ändert sich i. allg. der Zustand der Datenbasis.

Die Produktionen einer Grammatik (siehe Gl. 3-1) geben an, wie eine gegebene Menge von Primitivobjekten oder Teilobjekten in andere Teilobjekte oder in ein Zielobjekt (Referenzobjekt) überführt wird. Für die Generierung jedes Objektes existieren eine oder mehrere Produktionen.

In der formalen Darstellung einer Produktion wird angegeben, daß sich aus Objekten  $x$  und  $y$  mit einem bestimmten Prädikat ein Objekt  $z$  "produzieren" läßt. Das Prädikat berücksichtigt den Typ von  $x, y$ , der als spezielles Attribut vorausgesetzt wird, und eine topologische oder geometrische Relation  $\odot$  zwischen  $x$  und  $y$ , die andere geeignete Attribute berücksichtigt.

Für eine Produktionen  $P_u$  aus der Menge  $P = \{P_1, P_2, \dots, P_n\}$  wird folgende Schreibweise verwendet [81]:

$$P_u : X \wedge Y \odot \xrightarrow{u} Z$$

Dabei gilt, daß  $P_u$  ausgeführt werden soll, wenn

$$( \text{Typ}(x) = X \wedge \text{Typ}(y) = Y ) \wedge \odot(x, y)$$

wahr ist. In diesem Fall wird das neue Objekt  $z$  vom Typ  $Z$  mit seinen Attributwerten durch die Funktion  $\xrightarrow{u}(x, y)$  generiert.

Liegt beispielsweise eine Produktion  $P_1$  vor, die wie folgt beschrieben ist,

$$P_1 : \text{LINIE} \wedge \text{LINIE} \text{ -- } \xrightarrow{1} \text{LINIE}$$

so bedeutet dies: Liegen zwei Objekte vom Typ LINIE vor und liegen diese Objekte in der Verlängerung zueinander, angedeutet durch das Zeichen "--" für kollinear, so wird daraus ein neues Objekt vom Typ LINIE mit der speziellen Funktion  $\xrightarrow{1}$  erzeugt. Durch wiederholtes Anwenden von Produktionen lassen sich komplexere Objekte aufbauen.

### Blackboardarchitektur

Die Blackboardarchitektur wurde erstmals von Newell [57] vorgestellt und im Spracherkennungssystem Hearsay-II [49] implementiert. Eine formale Beschreibung verschiedener Implementierungen von Blackboardsystemen findet man in [88]. Die grundlegende Architektur und Arbeitsweise des Blackboardsystems kann anschaulich durch eine Metapher dargestellt werden:

Man stelle sich vor, daß mehrere Spezialisten mit unterschiedlichem Wissen zur kooperativen Lösung eines Problems benötigt werden. Die Spezialisten sind alle an einer Schultafel (Blackbord) versammelt, um Hypothesen und Zwischenergebnisse zu diskutieren. Sobald durch einen Spezialisten Daten auf das Blackboard geschrieben werden, können andere Spezialisten entscheiden, ob sie die Information gebrauchen können um zum Problemlösungsprozeß beizutragen. Diese Spezialisten wiederum schreiben ihre eigenen Ergebnisse auf das Blackboard um den Problemlösungsprozeß fortzuführen. Der Prozeß verläuft so lange, bis eine Lösung für das Problem gefunden wird oder bis kein Spezialist mehr einen Beitrag leisten kann. [88]

In einem Blackboardsystem übernehmen sogenannte Wissensquellen (knowledge sources) die Rollen von Spezialisten und eine globale Datenbasis wird als Blackboard benutzt. Eine dem System beigefügte Kontrolleinheit übernimmt die Rolle eines Diskussionsleiters.

## Aufbau des BPI-Systems

Die Wissensquellen des BPI-Systems, die auch als Produktionen eines Produktionssystems interpretiert werden können, haben Lese- und Schreibzugriff auf eine globale Datenbasis (Blackboard). Die formale Beschreibung des Bildauswertemodells läßt sich aus dem Modell des Ersetzungssystems ableiten. Der Inferenzmechanismus kann dann auf ein Erkennungsverfahren zurückgeführt werden, das eine spezielle Realisierungsform des "bottom up" Parsers darstellt.

Das BPI-System läßt sich in die Gruppen *Blackboard*, *Warteschlange*, *Steuermodul* und *Verarbeitungsmodul* einteilen (siehe Abb. 4.22).

Das **Blackboard** ist eine gemeinsame Datenbasis, auf die alle Teilprogramme des Verarbeitungsmoduls zentral zugreifen. Alle Objekte sind mit ihren Attributwerten in der für die Bildauswertung optimierten Datenbasis eingetragen. Ein schneller inhaltsadressierter Zugriff auf Objektmengen wird durch eine spezielle Hardwarerealisierung eines Assoziativspeichers ermöglicht, der die Datenbasis in invertierter Form bereitstellt. Das Speicherkonzept wird in Abb. 4.21 verdeutlicht.

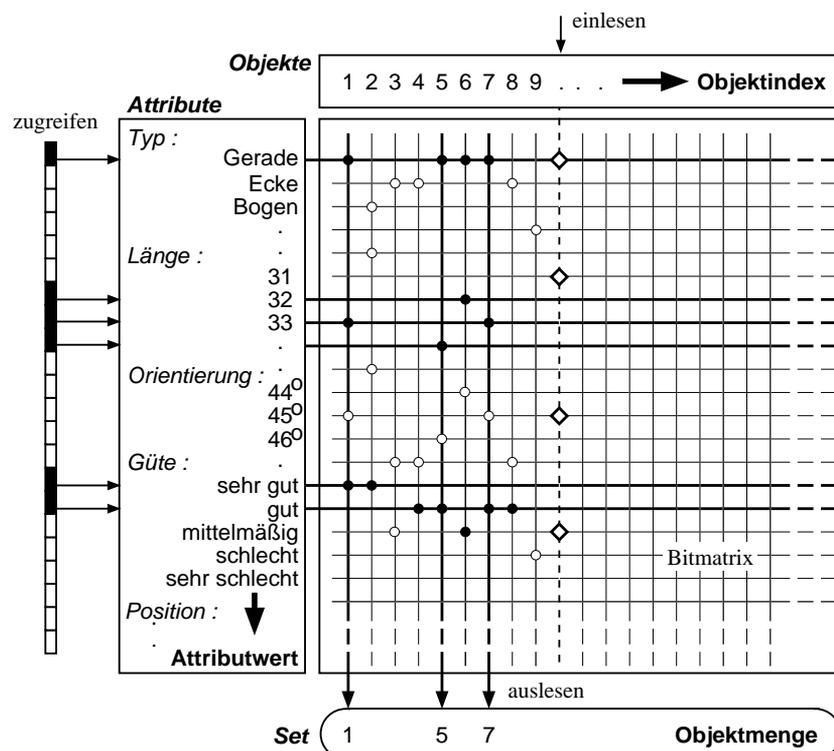


Abb. 4.21: Schematischer Aufbau des Assoziativspeichers

Die Zuordnung von Attributwerten zu Objekten erfolgt über eine Bitmatrix, wobei in

den Spalten die Objekte (gekennzeichnet durch den Index) und in den Zeilen die einzelnen Attributwerte (z.B. Länge=31, Länge=32, Länge=33 usw.) eingetragen sind. Das heißt, jedem Attributwert ist ein Bitfeld zugeordnet, das für jedes Objekt eine Bitposition hat. In diesen Bitfeldern wird also vermerkt, welche Objekte den zum Bitfeld gehörigen Attributwert besitzen. Durch Angabe von Attributwerten, Attributintervallen und weiteren Mengenoperationen lassen sich so gezielte Objektmengen (z.B. alle Objekte des Typs: *Gerade*, mit der Länge > 31 und der Bewertung *gut* bis *sehr gut*) effektiv auslesen und in einer Menge sammeln. Ein Objekt  $x$  vom Typ  $X$  wird im folgenden kurz mit Objekt  $X$  bezeichnet.

In der **Warteschlange** sind alle noch nicht bearbeiteten Elemente eingetragen und zur Verarbeitung nach Prioritätsklassen geordnet. Die Elemente können Objekte oder Verarbeitungselemente sein. Ein **Verarbeitungselement** besteht aus einem Objekt  $X$  und einer Hypothese, die für das Objekt annimmt "Teil von" einem Objekt  $Z$  einer höheren Abstraktionsstufe zu sein.

Das **Steuermodul** (Dispatcher) steuert den Datenfluß. Den Primitiv- und Teilobjekten werden Hypothesen (Regeln) zugewiesen und als Verarbeitungselemente werden sie wieder in die Warteschlange eingetragen. Die Verarbeitungselemente werden zur Überprüfung der Hypothese dem Verarbeitungsmodul übergeben.

Im **Verarbeitungsmodul** wird die Hypothese des Verarbeitungselementes geprüft, indem die Kompatibilität zu anderen Objekten untersucht wird. Diese Untersuchung wird in vier Schritten (V1-V4) durchgeführt (Abb. 4.22). Dazu werden Suchbereiche aufgespannt (V1) und die notwendigen Objekte zur Bestätigung der Hypothese im Blackboard gesucht (V2). Ein Suchbereich kann z.B. eine Suchfläche um den Endpunkt einer Linie sein, in der der Endpunkt einer Anschlußlinie gesucht wird. Kann die Kompatibilität unter Berücksichtigung von Toleranzen bestätigt werden (V3), wird ein neues Objekt erzeugt (V4) und bewertet. Anderenfalls wird die Hypothese verworfen.

Die verwendeten Produktionen sind durch Verarbeitungsmodule realisiert, die die geometrische Relation  $\odot$  überprüfen (V1-V3) und bei Erfüllen der Relation die objektspezifische Funktion  $\xrightarrow{x}$  durchführen (V4). Erläuterungen zur allgemeinen Struktur der Überprüfungsmodule sowie zur Überprüfung spezieller geometrischer Relationen findet man in [81].

### Ablaufzyklus

Zunächst werden alle Primitivobjekte, die bestimmten Mindestanforderungen (z.B. Mindestgüte) gerecht werden, in die Datenbasis eingetragen. Zusätzlich erfolgt ein Ein-

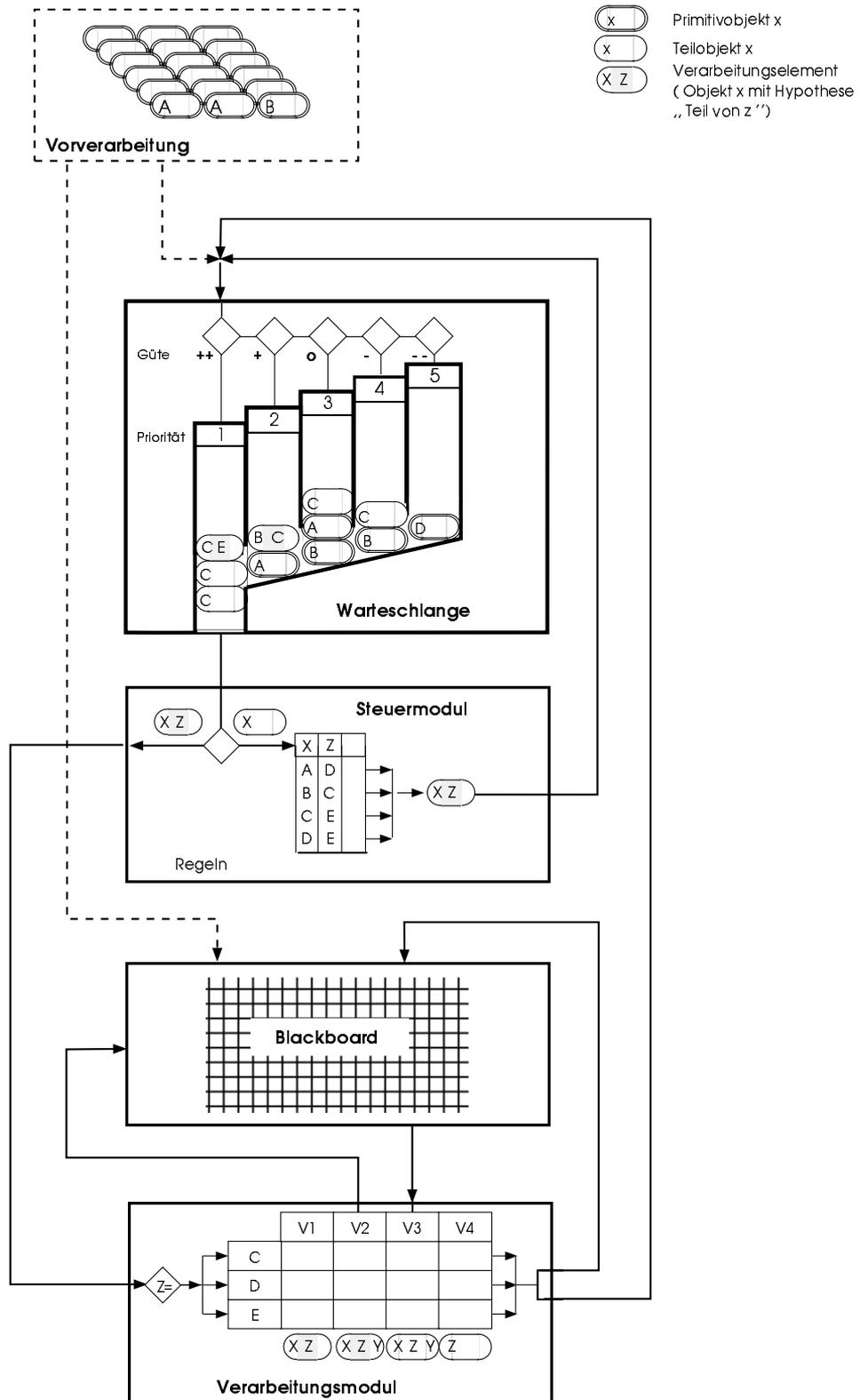


Abb. 4.22: Übersicht zum Datenfluß im Blackboard-System

trag in die prioritätsgeordnete Warteschlange, wobei sich die Stellung des Eintrags nach der Bewertung des Objektes richtet. Nach diesem Ladevorgang von Blackboard und Warteschlange beginnt eine zyklische Verarbeitung.

Der Dispatcher nimmt das erste Element aus der Warteschlange und prüft, ob es sich um ein *Verarbeitungselement* oder um ein *Objekt ohne Hypothese* handelt. Da noch keine Verarbeitung stattgefunden hat, wird es sich bei den ersten Elementen um Primitivobjekte ohne Hypothese handeln. Mit dem Objekt und einer objekttypischen "Teil von" - Hypothese wird ein Verarbeitungselement gebildet und entsprechend der Bewertung in die Warteschlange eingetragen. Liegen zu einem Objekttyp mehrere Hypothesen vor, z.B. "V ist Teil von X" und "V ist Teil von Y", so werden mehrfach Verarbeitungselemente eingetragen. Mit jeder Elemententnahme aus der Warteschlange rücken Objekte und Verarbeitungselemente vor.

Wird nun ein **Verarbeitungselement** aus der Warteschlange entnommen, so "triggert" dieses Element abhängig von der Hypothese ein Überprüfungsprogramm im Verarbeitungsmodul. Dieses sucht zur Bestätigung der Hypothese ein oder mehrere Objekte, mit dem oder denen das hypothetische Objekt aufgebaut werden kann.

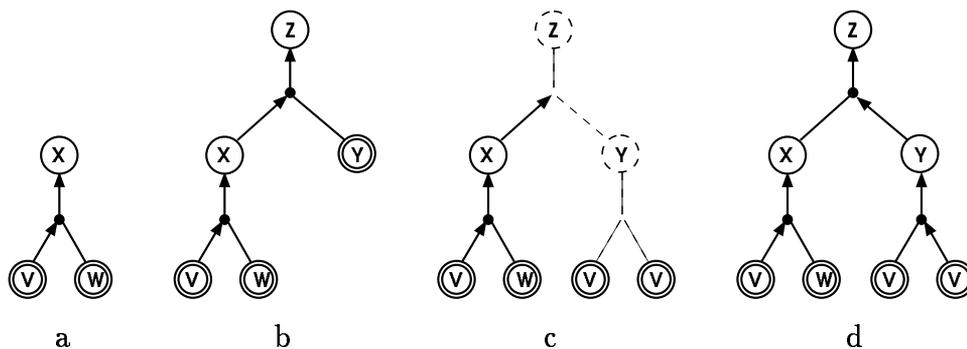


Abb. 4.23: Aufbau von Objekten aus Primitivobjekten © und Teilobjekten ○.  
a)-d) siehe Text

Besitzt also Objekt V die Hypothese "Teil von Objekt X", das aus den beiden Objektteilen V und W besteht, so wird zu V das Gegenstück W gesucht. Ist das Gegenstück W im Blackboard vorhanden, wird das Objekt X generiert und die zugehörigen Attribute werden aus den Vorgängerobjekten V und W berechnet (Abb. 4.23a). In Abb. 4.23 ist das triggernde Objekt mit dem vom Objekt ausgehenden Pfeil, die erfolgreiche Suche durch einen Knoten und die Erzeugung des neuen Objektes mit einem Pfeil gekennzeichnet. Dieses neu generierte Objekt wird in die Warteschlange und in das Blackboard eingetragen. Durch den Eintrag in die Warteschlange kann Objekt X selbst wieder triggern, um ein noch komplexeres Objekt Z aufzubauen (Abb. 4.23b)

und durch den Eintrag im Blackboard kann Objekt  $X$  auch von Objekt  $Y$  gefunden werden.

Das gesuchte Gegenstück  $Y$  ist jedoch nicht immer Primitivobjekt, wie in Abb. 4.23b dargestellt, sondern kann sich als Teilobjekt z.B. aus zwei Primitivobjekten  $V$  zusammensetzen (Abb. 4.23c). Wurden diese Teilobjekte schlecht bewertet, wird zum Zeitpunkt der Suche von Objekt  $X$  das Objekt  $Y$  noch nicht existieren (Abb. 4.23c) und somit auch  $Z$  nicht aufgebaut werden. Später wird jedoch Objekt  $Y$  triggern, um mit dem im Blackboard gefundenen  $X$  das Objekt  $Z$  zu generieren (Abb. 4.23d).

Auf diese Art und Weise ist ein parallel ablaufender Syntheseprozess möglich; die jeweils am besten bewerteten Teilobjekte werden weiter zusammengesetzt, bis vorgegebene Referenzobjekte gefunden sind. Das Zusammensetzen der Objekte des Blackboards läßt sich anschaulich mit einem Puzzle vergleichen [61].

### 4.3.2 Konzept des Schriftzeichens

Das Konzept der Referenz, das in Form einer Grammatik formuliert werden kann, wird am Beispiel der Ziffer fünf dargestellt. Die Grammatik wird nach Gl. 3-1 durch die Mengen  $V_t$ ,  $V_n$ ,  $S$  und  $P$  wie folgt beschrieben:

$$\begin{aligned} V_t &= \{ ER5, ES5, BOR, BUR, BUL \} \\ V_n &= \{ BUU, BUUOR, RS5, Z5 \} \\ S &= \{ Z5 \} \\ P &= \{ P_1, P_2, P_3, P_4 \} \end{aligned}$$

Eine Aufteilung der Ziffer in Segmente wurde nach den in Abb. 4.20c dargestellten Schnitten vorgenommen, wobei zur Beschreibung die fünf Primitivobjekte ER5 (Ecke mit rechtem Winkel), ES5 (Ecke mit spitzen Winkel), BOR (Bogen oben rechts), BUR (Bogen unten rechts) und BUL (Bogen unten links) herangezogen werden. Ausgehend von diesen Primitivobjekten werden durch die Produktionen die Objekte der Menge  $V_n$  erzeugt. Das Objekt  $Z5$  stellt dabei das Zielobjekt der Menge  $S$  dar. Die Produktionen  $P_1$  bis  $P_4$  sind in Tab. 4.2 aufgeführt.

Für jedes Referenzobjekt kann ein Konzeptgraph angegeben werden, der beschreibt, mit welchen Teilobjekten das Referenzobjekt aufgebaut wird. Die Knoten des Graphen entsprechen den Objekten und die Kanten des Graphen geben die Teil\_von-Relation wieder. Ein Beispiel für den Konzeptgraph der Ziffer fünf ist in Abb. 4.24 zu sehen.

Die Konzeptgraphen der zehn Ziffern sind in Abb. A.5 zusammengestellt.

P-Nr.	Objektverbindung	$\odot$	$\xrightarrow{x}$	aufzubauendes Obj.
$P_1$	BUL $\wedge$ BUR	$\sqcup$	$\xrightarrow{1}$	BUU
$P_2$	BUU $\wedge$ BOR	$\sqcap$	$\xrightarrow{2}$	BUUOR
$P_3$	ER5 $\wedge$ ES5	$\sqsubset$	$\xrightarrow{3}$	RS5
$P_4$	RS5 $\wedge$ BUUOR	$\natural$	$\xrightarrow{4}$	Z5

$\sqcup$	untere Anschlußpunkte benachbart
$\sqcap$	rechte Anschlußpunkte benachbart
$\sqsubset$	linke Anschlußpunkte benachbart
$\natural$	unterer Anschlußpunkt mit oberem Anschlußpunkt benachbart

Tabelle 4.2: Tabelle der Produktionen zur Synthese der Ziffer fünf

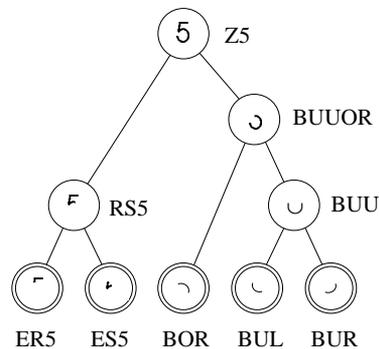


Abb. 4.24: Konzeptgraph der Ziffer fünf

### 4.3.3 Generierung der Primitivobjekte

Die Primitivobjekte werden durch Korrelation eines Satzes von Referenzen (Teilmasken) mit dem Muster gewonnen. Die Referenzen sind Teile der in Abb. 4.8 dargestellten Referenzen (Vollmasken). Die Lage der Schnitte zur Aufteilung der Vollmaske ist am Beispiel der Ziffer fünf in Abb. 4.25 dargestellt.

Die Berechnung des Korrelationskoeffizienten  $k_k$  und die Detektion der Spitzen erfolgt wie in Kap. 4.1.2 und 4.1.3 beschrieben.

### 4.3.4 Beispiel

Die einzelnen Verfahrensschritte der Erkennung sollen am Beispiel der Ziffer fünf erläutert werden. Zur Generierung der Primitivobjekte ER5, ES5, BOR, BUL und BUR

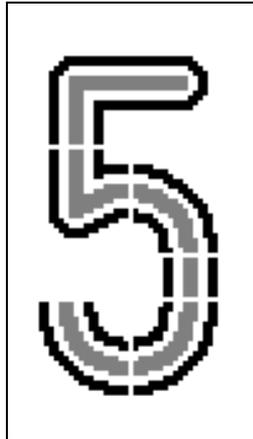


Abb. 4.25: Maskenaufteilung am Beispiel der Ziffer fünf (siehe Abb. 4.3)

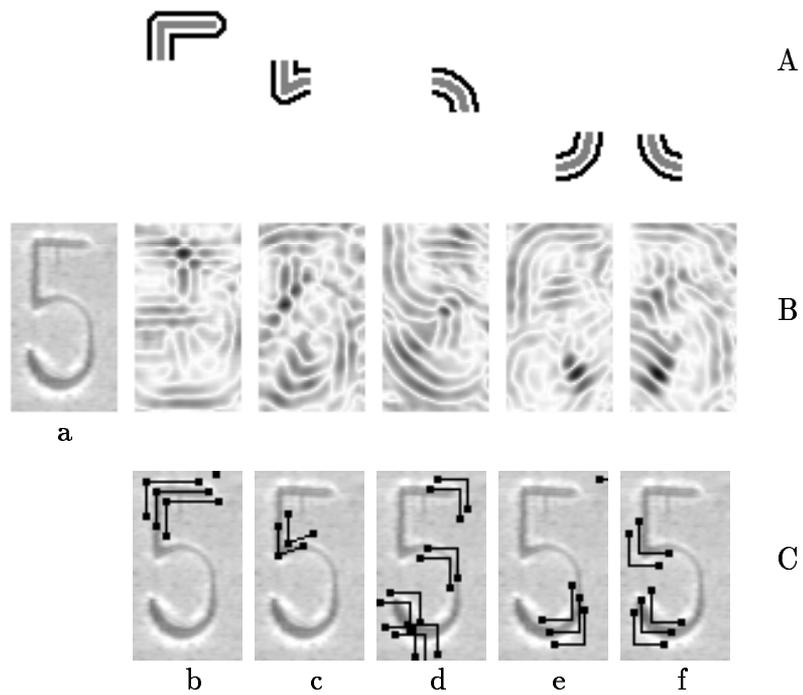
wird das Muster, von dem in Abb. 4.26Ba ein Ausschnitt abgebildet ist, mit den in Abb. 4.26A dargestellten Teilmasken korreliert.

Der entsprechende Ausschnitt des jeweiligen Korrelationsfeldes ist unterhalb der Masken in Abb. 4.26B dargestellt. Die typische Struktur des Korrelationsfeldes bleibt auch weitgehend erhalten, wenn die Teilmasken durch kleine Variationen verändert werden. Abb. A.7 zeigt am Beispiel der Teilmaske ES5 die Auswirkung einer Variation der Maskenaufteilung auf das Korrelationsfeld. Vom Korrelationsfeld ist jeweils ein Ausschnitt um den Punkt dargestellt, für den der höchste Korrelationswert erwartet wird.

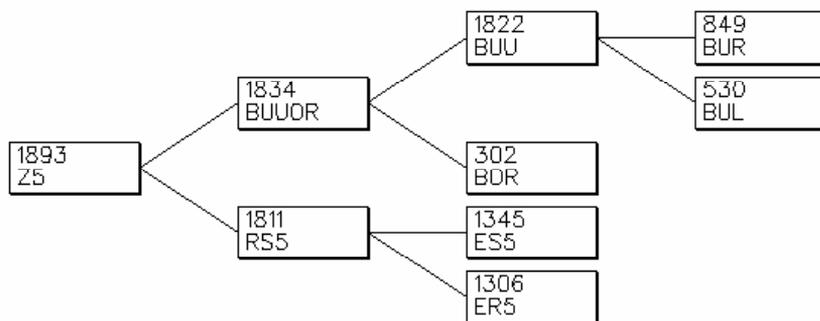
Durch die Untersuchung des Korrelationsfeldes mit Spitzendetektoren (Abb. 4.19) werden die Positionen der Primitivobjekte bestimmt. In Abb. 4.26C sind die detektierten Primitivobjekte für den unterlegten Ausschnitt des Musters symbolisch dargestellt.

Alle Primitivobjekte werden ins Blackboard eingetragen. Neben den Attributen zum Typ und der Geometrie besitzt jedes Primitivobjekt ein Bewertungsattribut, daß aus  $|k_k|$  und  $|\overline{g_\oplus} - \overline{g_\ominus}|$  (siehe Gl. 3-19) bestimmt wird. Zur Bewertung von Teilobjekten wird die Erfüllung der erwarteten geometrischen Relation und die Bewertung der Vorgängerobjekte in gewichteter Form berücksichtigt. Die heuristisch bestimmten Bewertungsfunktionen sind spezifisch für den jeweiligen Objekttyp.

Durch das Anwenden der Produktionen werden ausgehend von den Primitivobjekten Teilobjekte und Zielobjekte zusammengesetzt. Der Konstruktionsablauf läßt sich durch den Ableitungsgraphen, der die konkreten Zusammensetzungen dokumentiert, graphisch darstellen. Den Ableitungsgraphen des Zielobjektes Z5 mit dem Objektindex 1893 zeigt Abb. 4.27.



**Abb. 4.26:** Korrelation der Ziffer fünf mit Teilmasken.  
 Zeile A) Darstellung der Teilmasken (vgl. Abb. 4.25), Ba) Muster (Ausschnitt),  
 Bb) - Bf) Korrelationsfelder  $|K_{kER5}|$ ,  $|K_{kES5}|$ ,  $|K_{kBOR}|$ ,  $|K_{kBUR}|$ ,  $|K_{kBUL}|$  (Ausschnitte)  
 Zeile C) Symbolische Darstellung der detektierten Primitivobjekte aus Ab - Af  
 auf dem unterlegten Grauwertbild von Ba



**Abb. 4.27:** Ableitungsgraph des Zielobjektes Z5 (Objektindex: 1893)

Zum Aufbau von Z5 (1893) wurden die Teilobjekte RS5 (1811), BUU (1822) und BUUOR (1834) erzeugt. Die Generierungsreihenfolge läßt sich am Objektindex ablesen, der bei der Generierung eines Objektes in aufsteigender Folge vergeben wird. Neben den im Ableitungsgraphen benannten Teilobjekten werden noch weitere Teilobjekte vom gleichen Typ generiert.

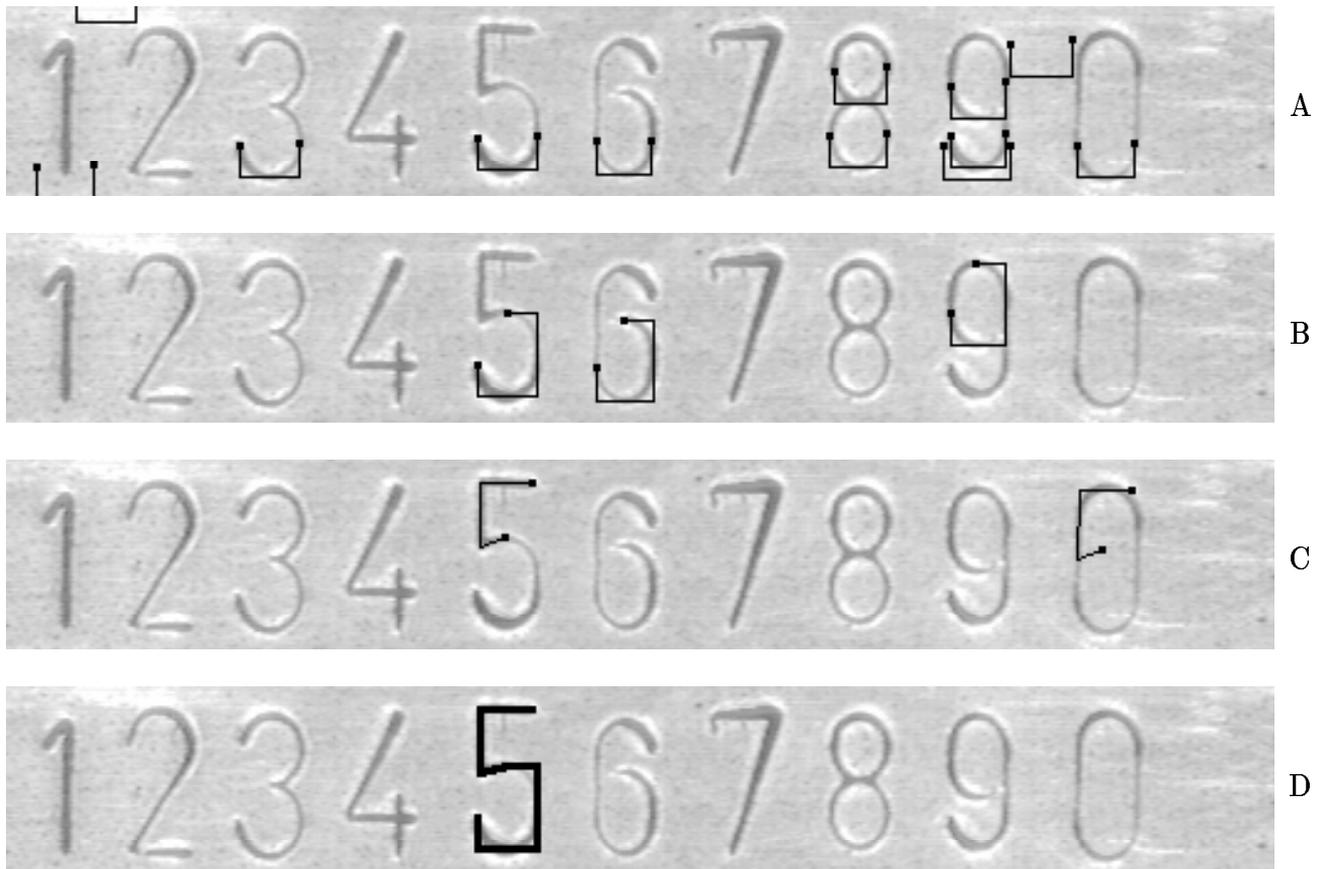


Abb. 4.28: Symbolische Darstellung der Teilobjekte in einer Schriftzeile  
 A) Teilobjekte BUU, B) Teilobjekte BUUOR, C) Teilobjekte RS5, D) Zielobjekt Z5

In Abb. 4.28 sind alle Teilobjekte vom Typ BUU, RS5 und BUUOR dargestellt, deren Koordinaten in den Bereich der unterlegten Schriftzeile fallen. Man sieht, daß die Teilobjekte BUUOR (Abb. 4.28B) nicht nur bei der Ziffer fünf, sondern auch bei der Ziffer sechs und neun auftreten und daß das Teilobjekt RS5 (Abb. 4.28C) fälschlicherweise bei der Ziffer null auftritt. Da zum Aufbau des Zielobjektes Z5 sowohl das Teilobjekt BUUOR und das Teilobjekt RS5 benötigt wird, entsteht nur an der Stelle der Ziffer fünf das Zielobjekt Z5.

Bestimmte Objekte wie z.B. das Teilobjekt BUU (Abb. 4.28A), die auch bei der Ziffer

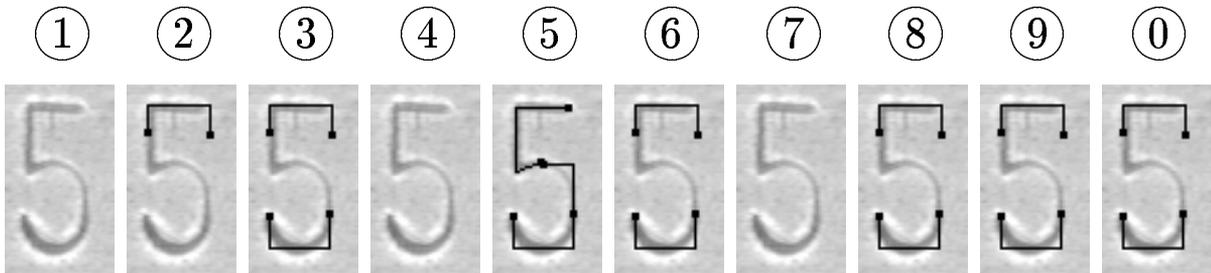


Abb. 4.29: Symbolische Darstellung der Teilobjekte beim Muster fünf, die beim Aufbau der Zielobjekte Z1 bis Z0 entstehen

drei und sechs auftreten, werden zum Aufbau der Zielobjekte Z3 und Z6 benötigt. Die Teilobjekte, die beim Aufbau der Zielobjekte Z1 bis Z0 bei dem Muster der Ziffer fünf entstehen, sind in Abb. 4.29 zusammengestellt.

#### 4.4 Ergebnisse der Verfahren I, II und III für Bildbeispiele

In den vorhergehenden Kapiteln wurden an den gezeigten Beispielen bereits Ergebnisse der Erkennung dargestellt. Eine Zusammenstellung der Ergebnisse der Erkennung durch die Verfahren I, II und III der in Abb. 2.3 dargestellten Muster ist in Tab. 4.3 aufgeführt. Erkannte Muster sind durch ein "+"-Zeichen, nicht erkannte Muster sind durch ein "o"-Zeichen und falsch zugewiesene Muster durch ein "-"-Zeichen gekennzeichnet.

Abb.	2.3a	2.3b	2.3c	2.3d	2.3e	2.3f	2.3g	2.3h	2.3i	2.3j
Verfahren I	-	+	+	+	+	+	+	+	+	+
Verfahren II	o	+	+	+	-	+	+	+	+	+
Verfahren III	o	o	+	+	+	+	+	o	+	+

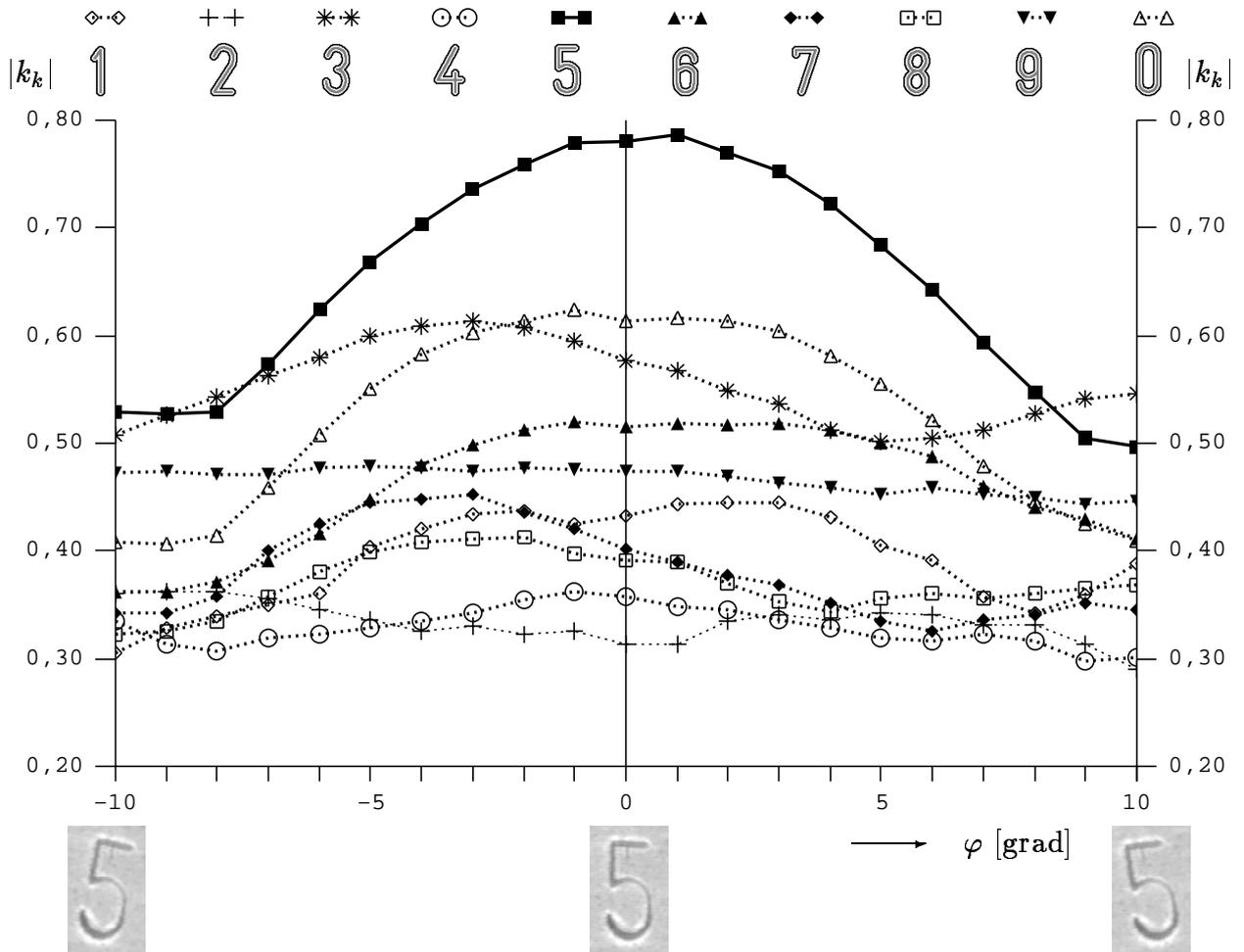
Tabelle 4.3: Zusammenstellung der erkannten Muster aus Abb. 2.3

## 4.5 Empfindlichkeit der Verfahren I, II und III gegenüber Verdrehung

In der industriellen Anwendung kann nicht immer sichergestellt werden, daß Werkstücke bei der Bildaufnahme exakt die gleiche Ausrichtung aufweisen. Es muß daher eine bestimmte Verdrehung der Vorlage gegenüber einer Referenzlage toleriert werden. Der Einfluß einer Verdrehung des Musters gegenüber der Referenz auf die der Erkennung zugrunde liegenden Größen soll an den folgenden Beispielen demonstriert werden. Es wurden einerseits die Abhängigkeiten der Verdrehung desselben Musters gegenüber verschiedenen Referenzen und andererseits die Abhängigkeiten der Verdrehung verschiedener Muster gegenüber derselben Referenz aufgenommen. Die Ergebnisse sind für das Verfahren I in Abb. 4.30 und Abb. 4.31, für das Verfahren II in Abb. 4.32 und Abb. 4.33 und für das Verfahren III in Abb. 4.34 und Abb. 4.35 in Form von Diagrammen dargestellt.

In Abb. 4.30 ist der Betrag des Korrelationskoeffizienten  $|k_k|$  für die Korrelation desselben Musters (Abb. 4.30 unten) mit den zehn Referenzen (Vollmasken)(Abb. 4.30 oben) dargestellt. Für jede Referenz wurde das Maximum von  $|k_k|$  in einer Fensterfläche  $A_w$  des Korrelationsfeldes  $|K_k|$  bestimmt. Die Fensterfläche  $A_w$  (Größe 48\*80 Bildpunkte) entspricht der Größe des Bildausschnittes der Ziffer fünf in Abb. 4.30 unten. Das Muster wurde um bis zu  $\varphi = \pm 10^\circ$  verdreht.

Den höchsten Wert von  $|k_k|$  für die  $0^\circ$ -Lage des Musters erhält man für die entsprechende Referenz der Ziffer fünf. Die Werte von  $|k_k|$  der konkurrierenden Referenzen liegen niedriger. Mit zunehmender Verdrehung aus dieser Lage in positiver oder negativer Winkelrichtung fällt der Wert von  $|k_k|$  ab, wobei der Kurvenverlauf (durchgezogene Linie in Abb. 4.30) eine glockenförmige Form zeigt. Eine Erkennung des Musters der Ziffer fünf an dem höchsten Korrelationswert ist nach Abb. 4.30 für  $-7^\circ \leq \varphi \leq +7^\circ$  möglich. Da bei den Verdrehungswinkeln von  $\varphi = -7^\circ$  und  $\varphi = +7^\circ$  der Abstand von  $|k_k|$  der Referenz der Ziffer fünf zu den Werten  $|k_k|$  der konkurrierenden Referenzen sehr klein ist, würde man die erlaubten Toleranzen niedriger ansetzen (z.B  $\Delta\varphi = \pm 5^\circ$ ).



**Abb. 4.30:** Korrelation desselben Musters mit verschiedenen Referenzen (Vollmasken).  
 Maximaler Betrag des Korrelationskoeffizienten  $k_k$  in Abhängigkeit  
 vom Verdrehungswinkel  $\varphi$  des aufgenommenen Musters

In Abb. 4.31 ist der Betrag des Korrelationskoeffizienten  $|k_k|$  für die Korrelation verschiedener Muster (Abb. 4.31 oben) mit derselben Referenz (Abb. 4.31 unten) dargestellt. Es sind die Korrelationswerte  $|k_k|$  für Verdrehungswinkel zwischen  $\varphi = 0^\circ$  und  $\varphi = 10^\circ$  eingetragen. Der typische Verlauf von  $|k_k(\varphi)|$  der Referenz der Ziffer fünf aus Abb. 4.30 zeigt sich auch bei unterschiedlichen Mustern.

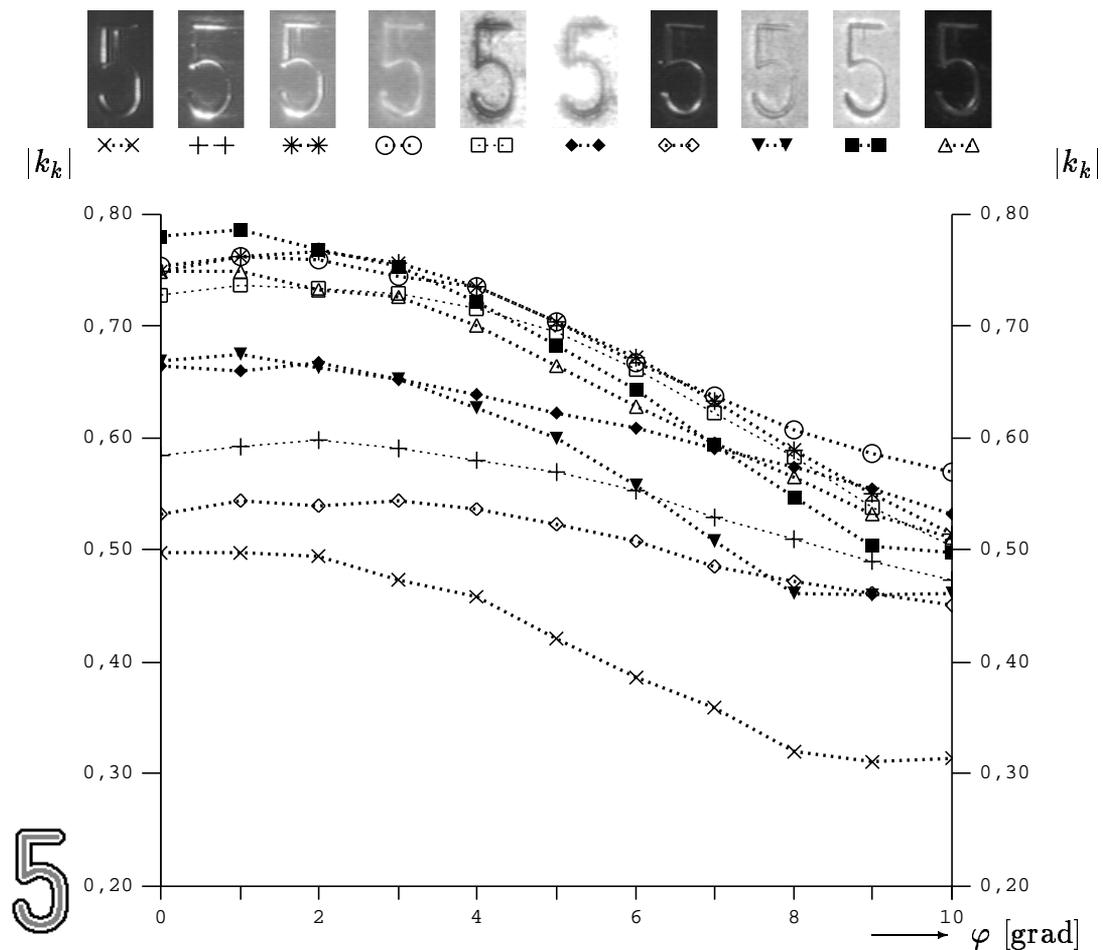
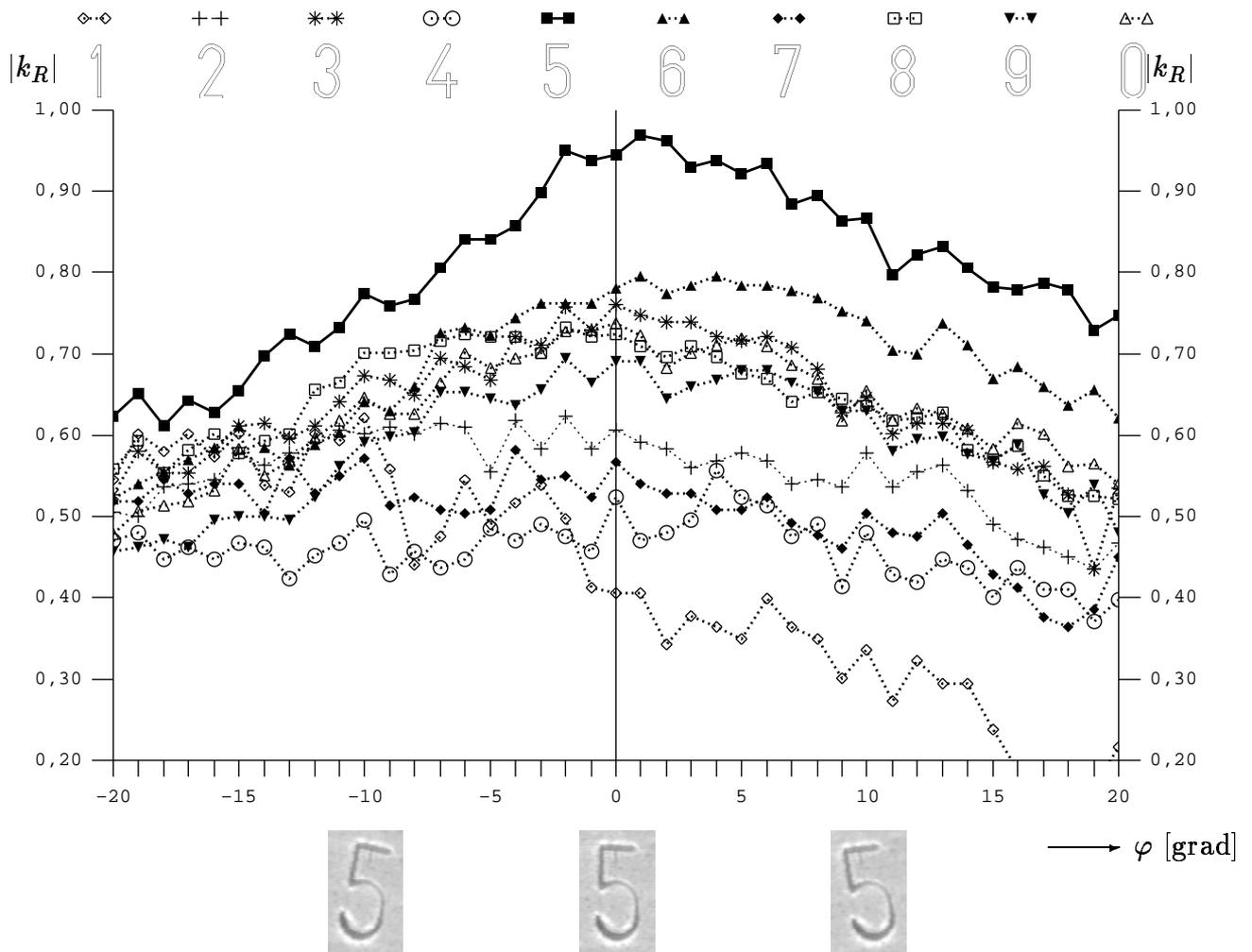


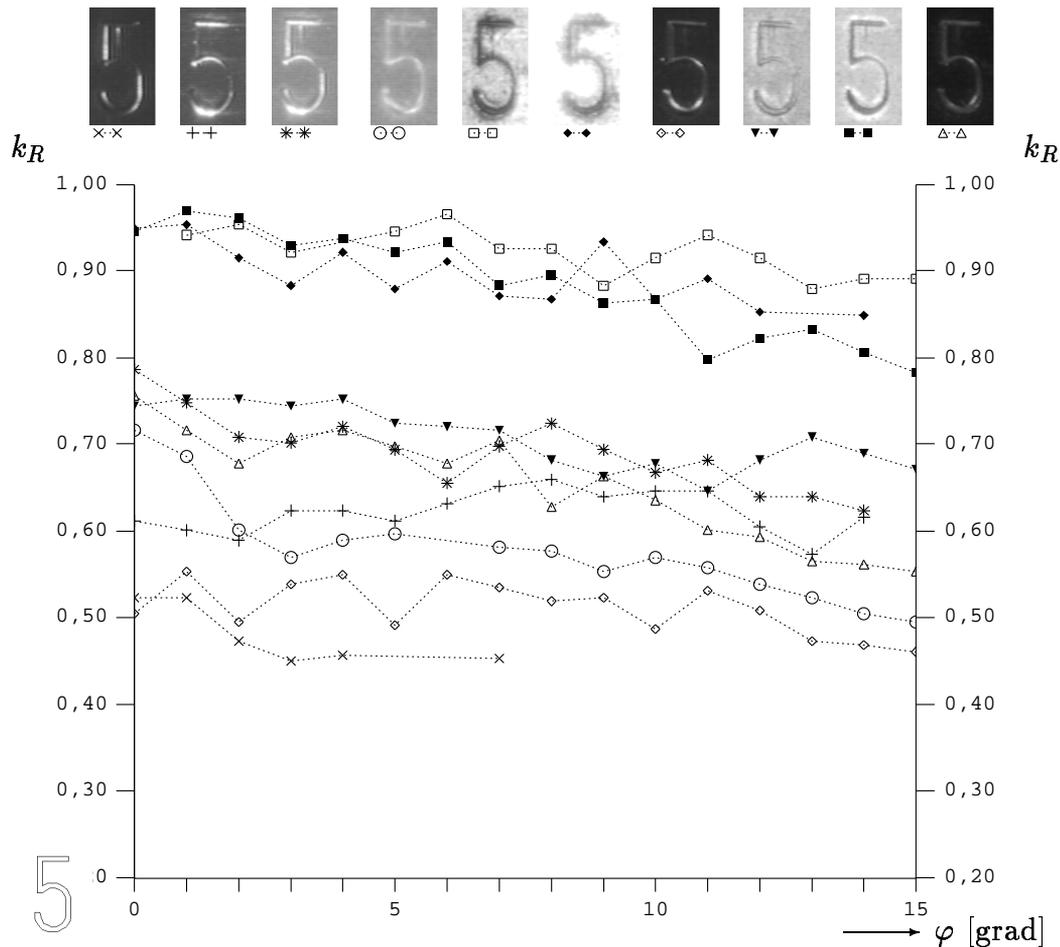
Abb. 4.31: Korrelation verschiedener Muster mit derselben Referenz (Vollmasken).  
Maximaler Betrag des Korrelationskoeffizienten  $k_k$  in Abhängigkeit vom Verdrehungswinkel  $\varphi$  des aufgenommenen Musters

In Abb. 4.32 ist das Ähnlichkeitsmaß  $k_R$  für den Linienvergleich desselben Musters (Abb. 4.32 unten) mit verschiedenen Referenzen (Abb. 4.32 oben) dargestellt. Das Muster wurde um bis zu  $\varphi = \pm 20^\circ$  verdreht. Das Ähnlichkeitsmaß  $k_R$  der entsprechenden Referenz besitzt im dargestellten Beispiel für alle Verdrehungswinkel von  $\varphi = -20^\circ$  bis  $\varphi = +20^\circ$  den maximalen Wert von allen verglichenen Referenzen.



**Abb. 4.32:** Linienvergleich desselben Musters mit verschiedenen Referenzen. Ähnlichkeitsmaß  $k_R$  in Abhängigkeit vom Verdrehungswinkel  $\varphi$  des aufgenommenen Musters

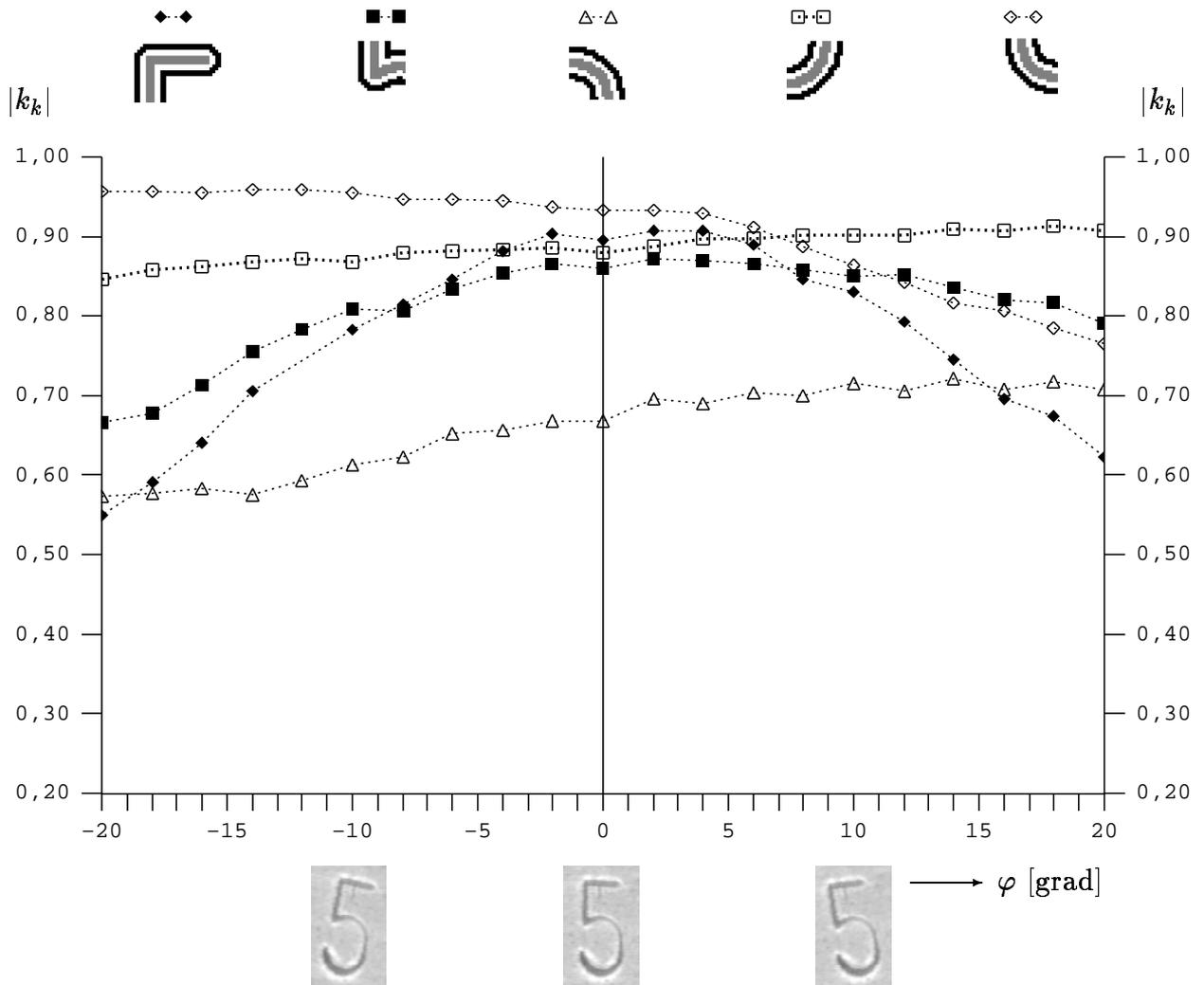
In Abb. 4.33 ist das Ähnlichkeitsmaß  $k_R$  für den Linienvergleich verschiedener Muster (Abb. 4.33 oben) mit derselben Referenz (Abb. 4.33 unten) dargestellt.



**Abb. 4.33:** Linienvergleich verschiedener Muster mit derselben Referenz. Ähnlichkeitsmaß  $k_R$  in Abhängigkeit vom Verdrehungswinkel  $\varphi$  des aufgenommenen Musters

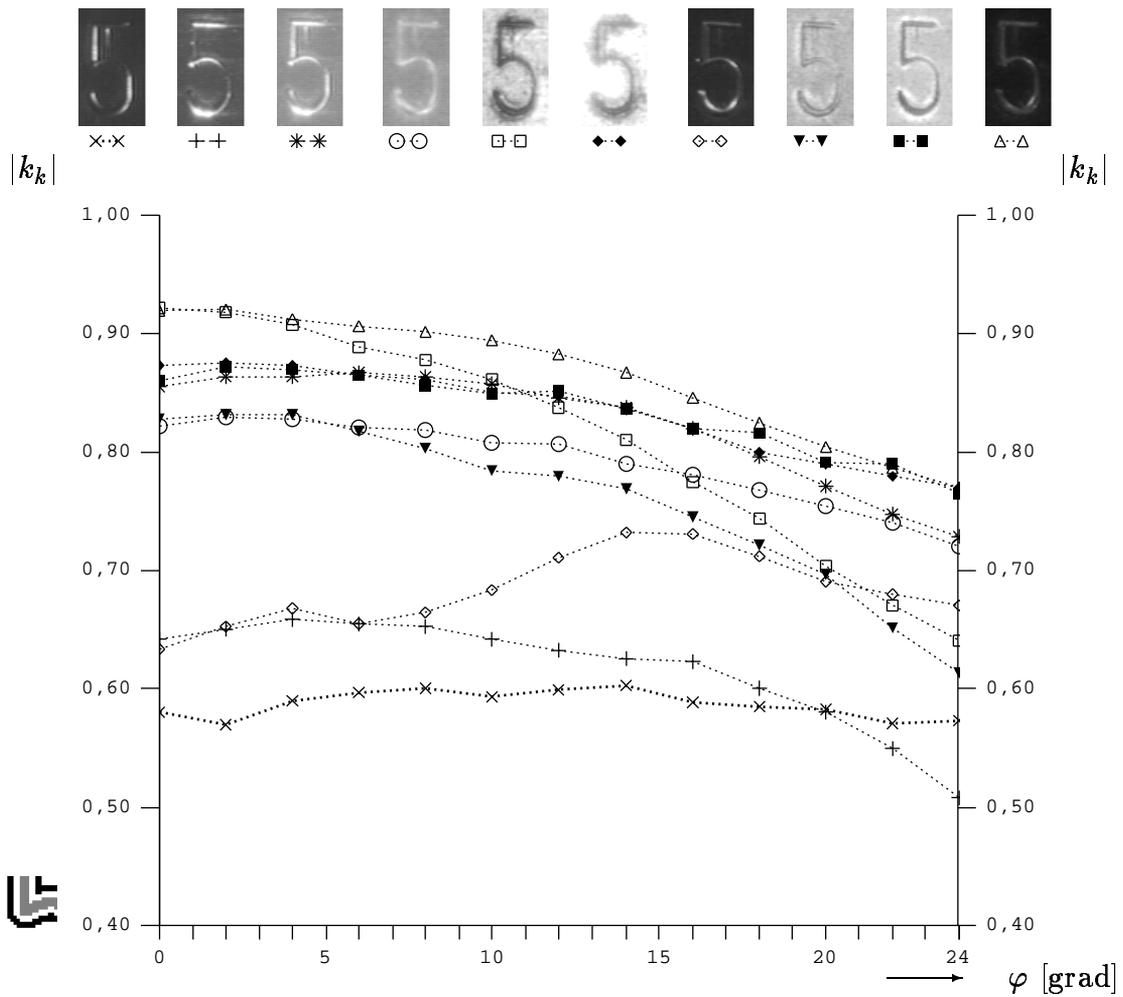
In Abb. 4.34 ist der Betrag des Korrelationskoeffizienten  $|k_k|$  für die Korrelation desselben Musters (Abb. 4.34 unten) mit den fünf Referenzen (Teilmasken der Ziffer fünf)(Abb. 4.34 oben) dargestellt. Für jede Referenz wurde das Maximum von  $|K_k|$  in einer Fensterfläche (siehe Abb. 4.26Bb-Bf) in der  $O^o$ -Lage des Musters markiert. Um die fünf markierten Punkte wurde im Muster jeweils eine Fensterfläche von  $9 \times 9$  Bildpunkten gelegt. Für die Korrelation mit verdrehtem Muster wurde für jede Referenz das Maximum von  $|K_k|$  in dem zur Referenz zugehörigen Fenster bestimmt. Das heißt, daß eine Verschiebung des Korrelationsmaximums von maximal  $\pm 4$  Bildpunkten in x- und y-Richtung zugelassen wurde.

Das Muster wurde um bis zu  $\varphi = \pm 20^\circ$  verdreht. Im Mittel weist die Korrelation mit den Teilmasken höhere Maxima der Beträge des Korrelationskoeffizienten  $k_k$  auf als bei der Korrelation mit Vollmasken (siehe Abb. 4.30). Die Korrelationswerte der drei Korrelationsmasken für die Viertelkreisbögen (Abb. 4.34 rechts oben) zeigen eine weitgehende Rotationsinvarianz.



**Abb. 4.34:** Korrelation desselben Musters mit verschiedenen Referenzen (Teilmasken).  
Maximaler Betrag des Korrelationskoeffizienten  $k_k$  in Abhängigkeit vom Verdrehungswinkel des aufgenommenen Musters

In Abb. 4.35 ist der Betrag des Korrelationskoeffizienten  $|k_k|$  für die Korrelation verschiedener Muster (Abb. 4.35 oben) mit derselben Referenz (Abb. 4.35 unten) dargestellt. Es sind die Korrelationswerte  $|k_k|$  für einen Verdrehungswinkel zwischen  $\varphi = 0^\circ$  bis  $\varphi = 20^\circ$  eingetragen. Verdrehungen bis zu  $\varphi = 10^\circ$  zeigen nur geringe Veränderungen von  $|k_k|$ .



**Abb. 4.35:** Korrelation verschiedener Muster mit derselben Referenz (Teilmasken).  
Maximaler Betrag des Korrelationskoeffizienten  $k_k$  in Abhängigkeit vom Verdrehungswinkel  $\varphi$  des aufgenommenen Musters

## 5. Einschränkung des Suchbereiches

Sind die Positionen der im Bild auftretenden Zeichenketten vorher ungefähr bekannt, so läßt sich der zu untersuchende Bereich im Bild  $B$  auf einen Suchbereich  $S$  entsprechend einschränken. Können keine genauen Angaben über die zu erwartenden Positionen von Zeichen getroffen werden, so läßt sich unter bestimmten Bedingungen die Lokalisierung eines Zeichens in einer Vorverarbeitungsstufe durchführen [18, 30].

Sind Zeichen in einer Kette angeordnet, so kann eine Segmentierung der Zeichen oder anders gesagt eine Vereinzelung von Zeichen zweistufig erfolgen (siehe Abb. 5.1). In der ersten Stufe wird durch eine Zeilensegmentierung eine Schriftzeile vom Hintergrund getrennt und in der zweiten Stufe durch eine Zeichensegmentierung die einzelnen Zeichen der Schriftzeile getrennt.

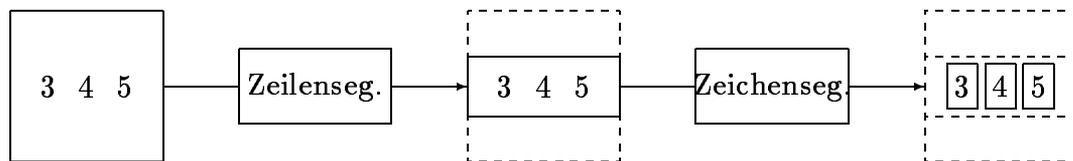


Abb. 5.1: Segmentierungsstufen zur Einschränkung des Suchbereichs

Die Vorgehensweisen zur Positionsbestimmung einer Zeichenkette und zur Positionsbestimmung einzelner Zeichen werden in Kap. 5.1 und Kap. 5.2 beschrieben.

### 5.1 Segmentierung einer Schriftzeile

Zur Segmentierung einer Schriftzeile wird ein einfaches Verfahren, das von Evers und Kämmerer [18] vorgeschlagen wurde, an einem Beispiel dargestellt. Dem Verfahren sind die Annahmen zugrunde gelegt, daß die Schriftzeichen annähernd in einer Reihe horizontal angeordnet vorliegen und die Schriftzeichenhöhe  $n_{zy}$  bekannt ist. Die Positionsbestimmung der Schriftzeile basiert auf der Auswertung der horizontalen Komponenten der Grauwertgradienten. Für jede Bildzeile werden die Differenzbeträge der Grauwerte horizontal benachbarter Bildpunkte summiert, so daß man die Summenverteilung  $g_{sx}(y)$  erhält

$$g_{sx}(y) = \sum_{x=1}^{n_{bx}-1} |g(x, y) - g(x+1, y)| \quad \text{für} \quad y = 1, \dots, n_{by} \quad (5-1)$$

Die Summenverteilung  $g_{sx}(y)$  des in Abb. 5.2a dargestellten Grauwertbildes ist in Abb. 5.2b als Profil dargestellt. Im Bereich der Schriftzeile nimmt  $g_{sx}(y)$  höhere Werte an als im Hintergrund. Zur Positionsbestimmung der Schriftzeile schiebt man ein Integrationsfenster mit der aus der Zeichenhöhe abgeleiteten Breite über das Profil und erhält die geglättete Summenverteilung  $g_{ssx}$  (Abb. 5.2c).

$$g_{ssx}(y) = \sum_{y=y'}^{y'+n_{zy}} g_{sx}(y) \quad \text{für} \quad y' = 1, \dots, n_{by} - n_{zy} \quad (5-2)$$

An der Position, an der das Integrationsfenster die Schriftzeile überdeckt, weist die Summenverteilung das lokale Maximum  $g_{ssx}(y_{max})$  auf.

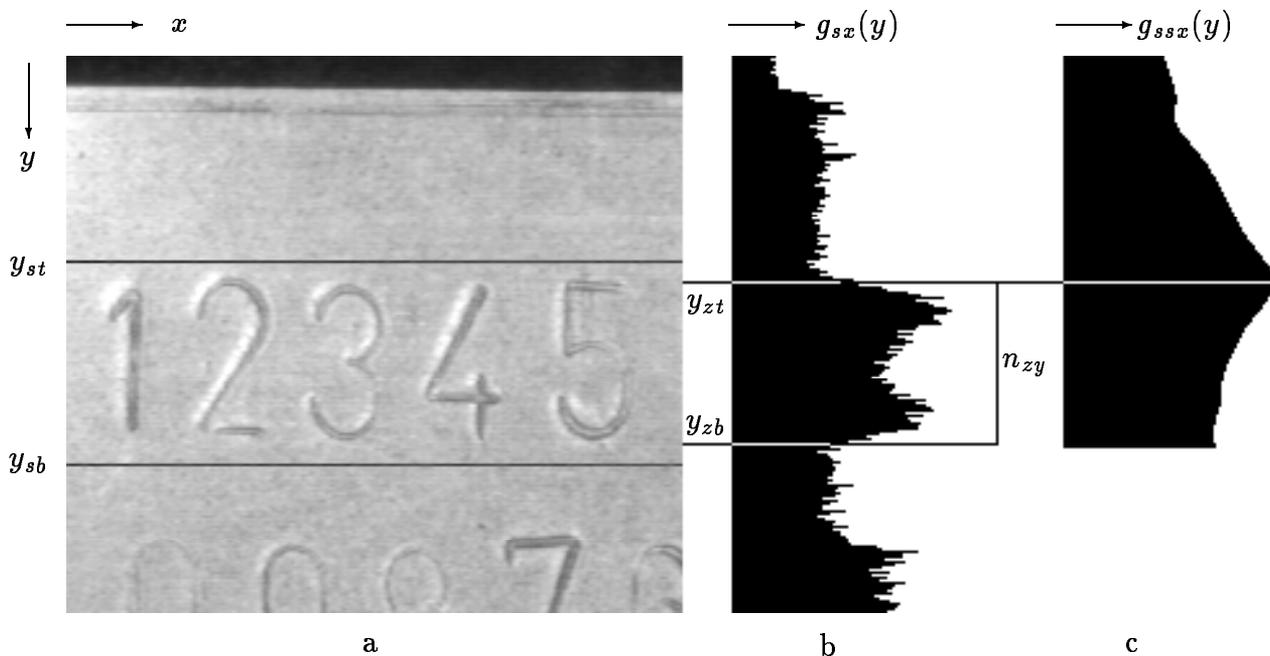


Abb. 5.2: Segmentierung einer Schriftzeile.

a) Grauwertbild, b) Summenprofil  $g_{sx}(y)$  nach Gl. 5-1, c) Summenprofil  $g_{ssx}(y)$  nach Gl. 5-2

Durch Bestimmung von  $y_{max}$  erhält man die Position des oberen Schriftzeilenrandes  $y_{zt}$  und die Position des unteren Schriftzeilenrandes aus  $y_{zb} = y_{max} + n_{zy} - 1$ . Zur Berücksichtigung einer leichten Verdrehung und Ungenauigkeiten bei der Positionierung der Schriftzeile fügt man oben und unten noch einige Zeilen hinzu, und erhält  $y_{st}$  und  $y_{sb}$  als Suchbereichsgrenzen.

Eine Kontrastverstärkung des Profils ist möglich, indem man kleine Differenzbeträge durch Vergleich mit einem Schwellwert detektiert und bei der Summation (Gl. 5-1) gleich zu null setzt.

Beim Verfahren wird vorausgesetzt, daß die Zeichenränder im Vergleich zum Hintergrund hohe Gradientenbeträge aufweisen und in Bildzeilen einer Schriftzeile häufiger hohe Gradientenbeträge auftreten als in Bildzeilen außerhalb der Schriftzeile. Hohe Gradientenbeträge können jedoch auch durch im Bild erfaßte Werkstückkonturen (z.B. Bohrungen) oder Störkonturen (z.B. Kratzer) auftreten. Auch Glanzflecken oder körnige Oberflächen können das Signal-Stör-Verhältnis verschlechtern.

## 5.2 Segmentierung einzelner Zeichen einer Schriftzeile

Zur Vereinzelung einzelner Zeichen wird vorausgesetzt, daß bereits eine Schriftzeile segmentiert wurde. Ähnlich wie bei der Vorgehensweise zur Segmentierung der Schriftzeile wird auch hier der Grauwertgradient ausgewertet. Statt der horizontalen Komponenten der Grauwertgradienten werden nun die vertikalen Komponenten ausgewertet. Da die Zahl der von einer Spalte erfaßten Zeichenkonturen geringer ist als die von einer Bildzeile, nimmt man zur Verbesserung des Signal-Stör-Verhältnisses noch eine Glättung vor. Statt den von Evers und Kämmerer benutzten vier Kompaßgradientenmasken wird an dieser Stelle nur eine Maske eingesetzt. Diese Richtungsmaske entspricht der vertikalen Sobel-Operator-Maske  $H_y$ .

$$H_y : \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 0 & 0 & 0 \\ \hline -1 & -2 & -1 \\ \hline \end{array}$$

Die vertikale Komponente des Grauwertgradienten  $g_{hy}(x, y)$  berechnet aus:

$$g_{hy}(x, y) = \sum_{u=1}^3 \sum_{v=1}^3 g(x-2+u, y-2+v) \cdot h_y(u, v) \quad \text{mit} \quad \begin{array}{l} x = 2, \dots, n_{bx} - 1, \\ y = y_{st} + 1, \dots, y_{sb} - 1 \end{array} \quad (5-3)$$

Als Bewertungsmaß  $g_{sy}(x)$  wird analog zu Gl. 5-1 für jede Spalte die Summe der Beiträge der Gradienten nach Gl. 5-3 gebildet.

$$g_{sy}(x) = \sum_{y=y_{st}+1}^{y_{sb}-1} |g_{hy}(x, y)| \quad \text{mit} \quad x = 2, \dots, n_{bx} - 1 \quad (5-4)$$

Die Summenverteilung  $g_{sy}(x)$  ist in Abb. 5.3b als Profil dargestellt.

Zur Auswertung wird das Profil geglättet:

$$g'_{sgy}(x) = \frac{1}{3}(g_{sy}(x-1) + g_{sy}(x) + g_{sy}(x+1)) \quad \text{für} \quad x = 3, \dots, n_{bx} - 2 \quad (5-5)$$

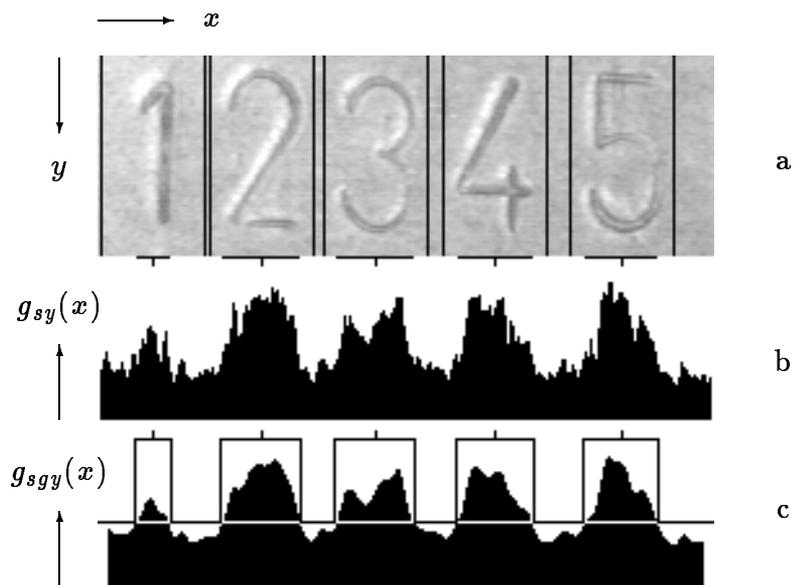


Abb. 5.3: Segmentierung einzelner Zeichen

a) Grauwertbild (Schriftzeile) b) Summenprofil  $g_{sy}(x)$  c) geglättetes Summenprofil  $g_{sgy}(x)$

Die Verteilung  $g_{sgy}$ , die man durch dreimalige Glättung in dieser Weise erhält, ist in Abb. 5.3c dargestellt. Durch eine Schwellwertoperation werden  $n_z$  Segmente bestimmt. Der Mittelpunkt dieser Segmente deutet auf die Mitte des Zeichens hin und dient zur Positionierung von  $n_z$  Suchbereichen mit der Breite  $n_{sx}$ . Die Breite der Suchbereiche leitet sich aus der maximalen Zeichenbreite ab, wobei aufgrund möglicher Positionierungsungenauigkeiten noch einige Spalten hinzugezählt werden (siehe Abb. 5.3).

## 6. Abschätzung des Berechnungsaufwands

Im folgenden Kapitel wird der Berechnungsaufwand für die in Kap. 4.1, Kap. 4.2 und Kap. 4.3 beschriebenen Verfahren (I bis III) abgeschätzt. Es wird der Aufwand betrachtet, der in einem Suchbereich zur Bereitstellung von Ähnlichkeitswerten zu verschiedenen Referenzen (Verfahren I und II) bzw. zur Generierung der Primitivobjekte (Verfahren III) notwendig ist. Da zur Bestimmung des Ähnlichkeitsmaßes nach Verfahren I wie auch zur Bestimmung der Primitivobjekte bei Verfahren III in gleicher Weise ein maskengesteuertes Korrelationsverfahren eingesetzt wird, bedarf es einer Unterscheidung zur Bestimmung des Berechnungsaufwandes erst bei einem konkreten Beispiel (Kap. 6.6).

Bei der Aufwandsabschätzung werden Vergleiche, Adressierungsrechnungen und Umspeicherzyklen nicht berücksichtigt.

Nach der Definition der Bezeichnungen von Bildbereichen (Kap. 6.1) wird der Aufwand für die Korrelation (Kap. 6.2) und der Aufwand für den Linienvergleich (Kap. 6.3) abgeschätzt. Der Aufwandsabschätzung zur Einschränkung des Suchbereiches wird in Kap. 6.4 nachgegangen. Eine Gegenüberstellung dieser Aufwandsabschätzungen auf der Basis des relativen Operationsaufwands erfolgt in Kap. 6.5. Ein Vergleich der Aufwandswerte wird an konkreten Beispielen in Kap. 6.6 vorgenommen.

### 6.1 Definitionen

In manchen Anwendungen kann aufgrund technischer Randbedingungen sichergestellt werden, daß in bestimmten Bildbereichen des Grauwertbildes keine Schriftzeichen auftreten, so daß der Bildbereich B auf einen Suchbereich S eingeschränkt werden kann (Abb. 6.1). In diesem Suchbereich seien dann  $n_z$  Schriftzeichen enthalten. Von der Abbildung einer Referenz R in S wird ein Bereich G überdeckt. Durch alle möglichen Verschiebepositionen der Referenz R im Suchbereich S wird mit einem Bezugspunkt (z.B. Mittelpunkt) der Referenz der Positionierungsbereich P aufgespannt. Für jede Positionierungsmöglichkeit ist ein Ähnlichkeitswert (z.B. Korrelationswert) mit einer Referenzdarstellung zu errechnen.

Die Zahl der Positionierungsmöglichkeiten  $n_{pxy}$  der Referenz R in einem Suchbereich S ergibt sich aus

$$n_{pxy} = n_{px} \cdot n_{py} = (n_{sx} - n_{rx} + 1) \cdot (n_{sy} - n_{ry} + 1) \quad (6 - 1)$$

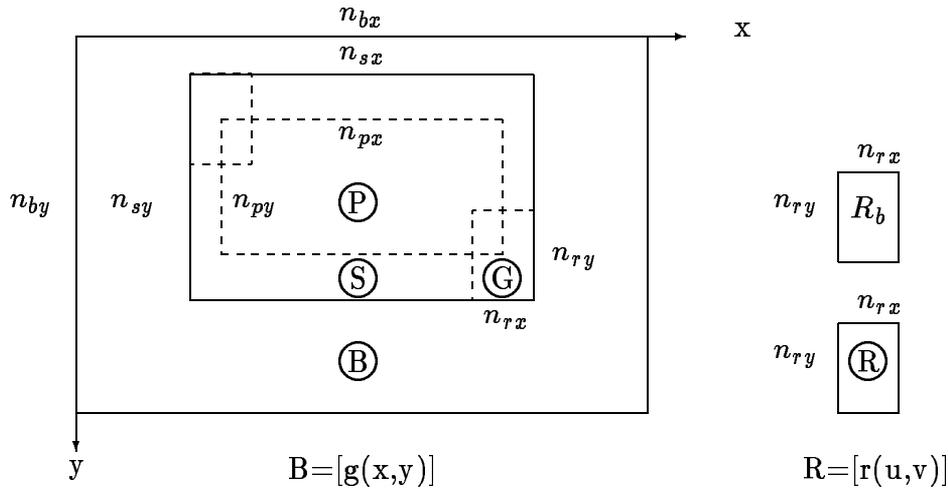


Abb. 6.1: Definition der Bildbereiche

Symb.	Bildbereich	Bildgröße
$B$	Grauwertbild	$n_{bxy} = n_{bx} \cdot n_{by}$
$R$	Referenz, Template, Distanzmatrix	$n_{rxy} = n_{rx} \cdot n_{ry}$
$R_b$	Binärmaske der Referenz	$n_{rxy} = n_{rx} \cdot n_{ry}$
$G$	Überdeckungsbereich von R	$n_{rxy} = n_{rx} \cdot n_{ry}$
$S$	Suchbereich, Region Of Interest (ROI)	$n_{sxy} = n_{sx} \cdot n_{sy}$
$P$	Positionierungsbereich	$n_{pxy} = n_{px} \cdot n_{py}$

Tabelle 6.1: Größe der Bildbereiche (siehe Abb. 6.1)

Bezeichnet  $n_K$  die Zahl der zu bestimmenden Ähnlichkeitswerte für  $n_r$  gleichgroße Referenzen, so sind  $n_K = n_r \cdot n_{pxy}$  Ähnlichkeitswerte zu bestimmen.

Eine Referenz  $R$  besitzt  $n_{rxy}$  Bildpunkten. Bei der in Kap. 4.1 dargestellten Vorgehensweise zur Korrelation können die Bildpunkte die Werte  $r(u, v) \in \{0, \rho_{\oplus}, \rho_{\ominus}\}$  annehmen. Es besitzen  $n_{r\oplus}$  Bildpunkte den Wert  $\rho_{\oplus}$ ,  $n_{r\ominus}$  Bildpunkte den Wert  $\rho_{\ominus}$  und  $n_{r0}$  Bildpunkte den Wert null, so daß gilt:

$$n_{rxy} = n_{r\oplus} + n_{r\ominus} + n_{r0} \quad (6 - 2)$$

Die Anzahl der Elemente mit  $r(u, v) \neq 0$  sei  $n_{r\odot} = n_{r\oplus} + n_{r\ominus}$ .

Da die Bildpunkte der Referenz mit dem Wert null zur Berechnung des Korrelationswertes nicht beitragen, wird zur Maskierung eine Binärmaske  $R_b$  mit der Größe der Re-

ferenz R definiert, für die gilt

$$r_b(u, v) = \begin{cases} 0 & \text{wenn } r(u, v) = 0 \\ 1 & \text{sonst} \end{cases} \quad (6-3)$$

Der zur Bestimmung eines Ähnlichkeitswertes durchzuführende Operationsaufwand  $O_k$  wird in Anteile aufgeteilt, die durch Additionen und Subtraktionen ( $O_{add}$ ), Multiplikationen ( $O_{mul}$ ), Divisionen ( $O_{div}$ ) und durch die Bestimmung der Wurzel ( $O_{sqrt}$ ) entstehen.

## 6.2 Berechnungsaufwand zur Korrelation

Wie bei der Faltung eines Grauwertbildes mit einer Filtermatrix lassen sich bei der Berechnung der Korrelation prinzipiell zwei Vorgehensweisen unterscheiden:

- Direkte Berechnung im Ortsbereich
- Indirekte Berechnung im Ortsfrequenzbereich

Bei der **direkten** Berechnung werden die Operationen in Form einer gleitenden Maskenoperation durchgeführt. Bei der **indirekten** Berechnung wird die diskrete Fourier-Transformation (DFT) von Muster G und Referenz R durchgeführt, punktweise das Produkt der Transformatierten gebildet und anschließend die inverse Transformation (IDFT) angewendet. Die zweidimensionale  $DFT_{xy}$  wird dabei aus zwei eindimensionalen  $DFT_y$  und  $DFT_x$  in Zeilen- und Spaltenrichtung gebildet.

Zur Berechnung der diskreten Fourier-Transformierten verwendet man häufig einen speziellen Algorithmus, der wegen der geringen erforderlichen Anzahl von Operationen und der damit verkürzten Transformationszeit als schnelle Fourier-Transformation (FFT) bezeichnet wird. Zur Ausnutzung der optimalen Effizienz des FFT-Algorithmus wird vorausgesetzt, daß die Zahl der zu transformierenden Werte  $n_q$  eine Potenz von zwei bildet. Bildet die Länge  $n_s$  einer zu transformierenden Bildzeile bzw. Bildspalte keine Potenz von zwei, so wird dies durch Auffüllen von Nullen (oder einer nach Null abklingenden Folge von Werten) auf die Länge  $n_q$  erreicht.

$$n_q = 2^\alpha \quad \text{mit } \alpha \in [ld\ n_s, ld\ n_s + 1), \quad \alpha \in \mathbb{IN} \quad (6-4)$$

Das transformierte Bild besitzt dann  $n_{qxy}$  Bildpunkte, mit

$$n_{qxy} = n_{qx} \cdot n_{qy} \quad (6-5)$$

$$= 2^{ld\ n_{qx} + ld\ n_{qy}} \quad (6-6)$$

Mit Gl. 6-4 wird der obere Wert abgeschätzt durch

$$n_{qxy} < 2^{ld\ n_{sx}+1+ld\ n_{sy}+1} \quad (6-7)$$

$$< 2^2 \cdot 2^{ld\ n_{sx}} \cdot 2^{ld\ n_{sy}} \quad (6-8)$$

$$< 4n_{pxy} \quad (6-9)$$

### 6.2.1 Berechnung des Korrelationskoeffizienten im Ortsbereich

In Kap. 3.5 ist die Gleichung für den normierten Korrelationskoeffizienten  $k_k$  dargestellt (Gl. 3-10). Durch die Normierung der Referenz ergibt sich aus Gl. 3-10 für den allgemeinen Fall ( $r_i \in \mathbb{IN}$ ) Gl. 3-13 und für den Sonderfall, daß die maskierte Referenz nur zwei Werte annimmt, Gl. 3-19. Zur Ermittlung der Aufwandsreduktion durch Benutzung von Gl. 3-19 gegenüber Gl. 3-13 wird der Berechnungsaufwand für beide Gleichungen bestimmt.

#### a) Maskierte Referenz kann mehr als zwei Werte annehmen ( $r_i \in \mathbb{IN}$ )

Berechnung des Korrelationskoeffizienten nach Gl. 3-13

$$k_k = \frac{\overline{g \cdot r}}{\sigma_g} = \frac{\overline{g \cdot r}}{\sqrt{g^2 - \bar{g}^2}} = \frac{\overline{g \cdot r}}{\sqrt{\overline{g \cdot g} - \bar{g} \cdot \bar{g}}} \quad (6-10)$$

mit dem Mittelwert  $\bar{g}$

$$\bar{g} = c \sum_i^n g_i \quad \text{mit} \quad c = \frac{1}{n} \quad (6-11)$$

Für die Berechnung eines Korrelationswertes nach Gl. 6-10 ist der Aufwand  $O_{k1}$  nötig

$$O_{k1} = (3n_{r\ominus} - 2)O_{add} + (2n_{r\ominus} + 2)O_{mul} + 1O_{div} + 1O_{sqr} \quad (6-12)$$

und für die Berechnung aller Korrelationswerte mit  $n_r$  Referenzen ist der Gesamtaufwand  $O_K$  nötig

$$O_{K1} = n_r n_{pxy} (3n_{r\ominus} - 2)O_{add} + n_r n_{pxy} (2n_{r\ominus} + 2)O_{mul} + n_r n_{pxy} O_{div} + n_r n_{pxy} O_{sqr} \quad (6-13)$$

#### b) Maskierte Referenz nimmt nur zwei Werte an ( $r_i \in \{\rho_{\oplus}, \rho_{\ominus}\}$ )

Der Korrelationswert  $k_k$  für eine Referenz mit den Werten  $\rho_{\oplus}$  und  $\rho_{\ominus}$  berechnet sich wie folgt:

Operation		add	mul	div	sqrt
$\bar{g}$		$n_{r\ominus} - 1$	1		
$\bar{g} \cdot \bar{g}$	+		1		
$\overline{g \cdot g}$	+	$n_{r\ominus} - 1$	$n_{r\ominus}$		
$\overline{g \cdot r}$	+	$n_{r\ominus} - 1$	$n_{r\ominus}$		
$k_k$	+	1		1	1
Summe	=	$3n_{r\ominus} - 2$	$2n_{r\ominus} + 2$	1	1

Tabelle 6.2: Operationsaufwand für Korrelation einer n-wertigen Referenz mittels Maskenoperation ( $k_k$  nach Gl. 6-10)

$$k_k = \text{sgn}(\rho_{\oplus} - \rho_{\ominus}) \cdot \text{sgn}(\bar{g}_{\oplus} - \bar{g}_{\ominus}) \frac{1}{\sqrt{\frac{m}{m_{\ominus}} \left( \frac{\sigma_{\oplus}}{\bar{g}_{\oplus} - \bar{g}_{\ominus}} \right)^2 + \frac{m}{m_{\oplus}} \left( \frac{\sigma_{\ominus}}{\bar{g}_{\oplus} - \bar{g}_{\ominus}} \right)^2 + 1}} \quad (6-14)$$

Terme, die typisch für die Referenz, nicht jedoch für den untersuchten Bereich des Grauwertbildes sind, werden vor dem Erkennungsvorgang berechnet und als Konstante ( $c_1, c_2, c_3$ ) geführt. Ersetzt man die Differenz der Mittelwerte durch

$$\Delta \bar{g}_{\ominus} = \bar{g}_{\oplus} - \bar{g}_{\ominus} \quad (6-15)$$

so erhält man

$$k_k = c_1 \cdot \text{sgn}(\Delta \bar{g}_{\ominus}) \frac{1}{\sqrt{c_2 \frac{\sigma_{\oplus}^2}{\Delta \bar{g}_{\ominus}^2} + c_3 \frac{\sigma_{\ominus}^2}{\Delta \bar{g}_{\ominus}^2} + 1}} \quad (6-16)$$

Für die Ausdrücke  $\Delta \bar{g}_{\ominus}, \sigma_{\oplus}^2, \sigma_{\ominus}^2$  sind die Mittelwerte  $g_{\oplus}, g_{\ominus}$  zu bestimmen. Der Berechnungsaufwand für die Mittelwerte, sowie der zusätzliche Berechnungsaufwand daraus abgeleiteter Terme sind in der Tab. 6.3 angegeben. Der Aufwand für Bestimmung und Setzen des Vorzeichens  $\text{sgn}(\Delta \bar{g}_{\ominus})$  wird vernachlässigt.

Zur Bestimmung eines Korrelationswertes mit dem Operationsaufwand für  $O_{add}, O_{mul}, O_{div}, O_{sqrt}$  ergibt sich mit  $n_{r\ominus} = n_{r\oplus} + n_{r\ominus}$

$$O_{k2} = (2n_{r\ominus} + 1)O_{add} + (n_{r\ominus} + 10)O_{mul} + 3O_{div} + 1O_{sqrt} \quad (6-17)$$

Für die Bestimmung aller Korrelationswerte mit  $n_r$  Referenzen ergibt sich der Aufwand  $O_{K2}$  als

$$O_{K2} = n_r n_{pxy} (2n_{r\ominus} + 1) O_{add} + n_r n_{pxy} (n_{r\ominus} + 10) O_{mul} + 3n_r n_{pxy} O_{div} + n_r n_{pxy} O_{sqrt} \quad (6-18)$$

Operation		add	mul	div	sqrt
$\overline{g_{\oplus}}$		$n_{r_{\oplus}} - 1$	1		
$\overline{g_{\ominus}}$		$n_{r_{\ominus}} - 1$	1		
$\Delta \overline{g_{\ominus}^2}$	+	1	1		
$\sigma_{\oplus}^2$	+	$n_{r_{\oplus}}$	$n_{r_{\oplus}} + 2$		
$\sigma_{\ominus}^2$	+	$n_{r_{\ominus}}$	$n_{r_{\ominus}} + 2$		
$k_k$	+	2	3	3	1
Summe	=	$2(n_{r_{\oplus}} + n_{r_{\ominus}}) + 1$	$n_{r_{\oplus}} + n_{r_{\ominus}} + 10$	3	1

**Tabelle 6.3:** Operationsaufwand für Korrelation einer zweiwertigen Referenz mittels Maskenoperation ( $k_k$  nach Gl. 6-16)

### 6.2.2 Berechnung des Korrelationskoeffizienten im Ortsfrequenzbereich (FFT)

Zur Berechnung der Kreuzkorrelation zweier Bilder multipliziert man im Frequenzbereich die Transformaten des einen Bildes mit der konjugiert komplexen der Transformaten des anderen Bildes und führt für das Produkt die Rücktransformation durch. Faltung und Korrelation unterscheiden sich nur dadurch, daß im Ortsbereich eine Bildfunktion gespiegelt bzw. im Frequenzbereich konjugiert komplex darzustellen ist. Die Berechnung der Mittelwerte der gemischten Produkte  $g \cdot r$  läßt sich als Faltungsoperation  $\otimes$  von  $g$  mit  $r'$  auffassen

$$\overline{g \cdot r} = g \otimes r' \quad \text{mit} \quad r'(u, v) = \frac{1}{n_{r_{\ominus}}} \cdot r(u, v) \quad (6-19)$$

wobei  $n_{r_{\ominus}}$  die Anzahl der maskierten Elemente angibt. Der Mittelwert der maskierten Elemente von  $\overline{g}$  ergibt sich durch

$$\overline{g} = g \otimes r'_b \quad \text{mit} \quad r'_b(u, v) = \frac{1}{n_{r_{\ominus}}} \cdot r_b(u, v) \quad (6-20)$$

bzw. für

$$\overline{g^2} = g^2 \otimes r'_b \quad \text{mit} \quad r'_b(u, v) = \frac{1}{n_{r_{\ominus}}} \cdot r_b(u, v) \quad (6-21)$$

Zur Berechnung des Korrelationskoeffizienten nach Gl. 3-13

$$k_k = \frac{\overline{g \cdot r}}{\sigma_g} = \frac{\overline{g \cdot r}}{\sqrt{\overline{g^2} - \overline{g}^2}} \quad (6-22)$$

werden die Transformationen nach Tab. 6.4 benötigt.

Da die Transformationen für  $r$  und  $r_b$  vor dem Erkennungsvorgang durchgeführt werden können, sind insgesamt fünf Transformationen vorzunehmen.

Ein zu transformierendes Bild besitzt  $n_{qxy} = n_{qx} \cdot n_{qy}$  Bildpunkte. Die Transformation  $DFT_x$  der  $n_{qx}$  Spalten benötigen den Operationsaufwand

$$O_{FFTX} = n_{qx}((n_{qy} \text{ ld } n_{qy})O_{jadd} + \frac{n_{qy}}{2}(\text{ld } n_{qy} - 1)O_{jmul}) \quad (6-23)$$

Die anschließende Transformation  $DFT_y$  der  $n_{qy}$  Zeilen benötigt

$$O_{FFTY} = n_{qy}((n_{qx} \text{ ld } n_{qx})O_{jadd} + \frac{n_{qx}}{2}(\text{ld } n_{qx} - 1)O_{jmul}) \quad (6-24)$$

Der Gesamtaufwand  $O_{FFTXY} = O_{FFTX} + O_{FFTY}$  ergibt sich aus

$$O_{FFTXY} = (n_{qxy} \text{ ld } n_{qxy})O_{jadd} + \frac{n_{qxy}}{2}(\text{ld } n_{qxy} - 2)O_{jmul} \quad (6-25)$$

Anstatt der Faltung im Ortsbereich wird eine entsprechende Multiplikation im Ortsfrequenzbereich durchgeführt, die  $n_{qxy}$  komplexe Multiplikationen ( $O_{jmul}$ ) benötigt.

Da Eingangsbild und Referenz reellwertig sind, kann die Symmetrie zwischen den komplexen und konjugiert komplexen Werten im Ortsfrequenzbereich ausgenutzt werden, so daß statt der  $n_{qxy}$  komplexen Multiplikationen etwa nur  $\frac{n_{qxy}}{2}$  komplexe Multiplikationen  $O_{jmul}$  auszuführen sind.

Der Aufwand für die komplexen Operationen  $O_{jadd}, O_{jmul}$  ist in Tab. 6.4 durch den Aufwand der reellen Operationen  $O_{add}, O_{mul}$  ersetzt:

$$\begin{aligned} O_{jadd} &= 2 \cdot O_{add} \\ O_{jmul} &= 4 \cdot O_{mul} + 2 \cdot O_{add} \end{aligned}$$

Der Aufwand für die Korrelation mit nur einer Referenzmaske beträgt

$$\begin{aligned} O_{K3(n_r=1)} &= n_{qxy}(15 \text{ ld } n_{qxy} - 6) O_{add} + \\ & n_{qxy}(10 \text{ ld } n_{qxy} - 13) O_{mul} + n_{qxy}O_{div} + n_{qxy}O_{sqr t} \end{aligned} \quad (6-26)$$

Durch die Korrelation mehrerer Referenzmasken mit demselben Grauwertbild brauchen die Transformationen  $g, g^2$  jeweils nur einmal durchgeführt werden. Für  $n_r$  Referenzmasken sind

$$n_{FFT} = 3 \cdot n_r + 2 \quad (6-27)$$

Transformationen auszuführen, so daß sich ein Gesamtaufwand  $O_{K3}$  ergibt mit

$$\begin{aligned} O_{K3} &= (n_r n_{qxy}(9 \text{ ld } n_{qxy} - 2) + n_{qxy}(6 \text{ ld } n_{qxy} - 4)) O_{add} + \\ & (n_r n_{qxy}(6 \text{ ld } n_{qxy} - 6) + n_{qxy}(4 \text{ ld } n_{qxy} - 7)) O_{mul} + \\ & n_r n_{qxy} O_{div} + n_r n_{qxy} O_{sqr t} \end{aligned} \quad (6-28)$$

OP	FFT	OP	add	mul	div	sqrt
	$\Rightarrow$	<b>r</b>				
	$\Rightarrow$	<b>r<sub>b</sub></b>				
$g^2$	$\Rightarrow$	<b>g</b>	$n_{qxy}(3 \text{ld } n_{qxy} - 2)$	$n_{qxy}(2 \text{ld } n_{qxy} - 4)$		
	$\Rightarrow$	<b>g<sup>2</sup></b>	$n_{qxy}(3 \text{ld } n_{qxy} - 2)$	$n_{qxy}(2 \text{ld } n_{qxy} - 4)$		
		<b>g · r</b>	$n_{qxy}$	$2n_{qxy}$		
		<b>g · r<sub>b</sub></b>	$n_{qxy}$	$2n_{qxy}$		
		<b>g<sup>2</sup> · r<sub>b</sub></b>	$n_{qxy}$	$2n_{qxy}$		
$\overline{g \cdot r}$	$\Leftarrow$		$n_{qxy}(3 \text{ld } n_{qxy} - 2)$	$n_{qxy}(2 \text{ld } n_{qxy} - 4)$		
$\overline{g}$	$\Leftarrow$		$n_{qxy}(3 \text{ld } n_{qxy} - 2)$	$n_{qxy}(2 \text{ld } n_{qxy} - 4)$		
$\overline{g^2}$	$\Leftarrow$		$n_{qxy}(3 \text{ld } n_{qxy} - 2)$	$n_{qxy}(2 \text{ld } n_{qxy} - 4)$		
$K_k$			$n_{qxy}$		$n_{qxy}$	$n_{qxy}$

Tabelle 6.4: Operationsaufwand für Korrelation einer n-wertigen Referenz mittels FFT ( $k_k$  nach Gl. 6-22)

### 6.3 Berechnungsaufwand zum Linienvergleich

Da der Aufwand zur Überführung des Grauwertbildes in eine Liniendarstellung nicht nur von den Dimensionen des untersuchten Bildbereiches, sondern auch von dem Bildinhalt (z.B. dargestelltes Zeichen, Art und Intensität von Störungen, eingestellte Parameter der Konturdetektion) abhängt, wird zunächst eine Abschätzung der Anzahl detektierter Konturpunkte  $n_{kh}$  vorgenommen.

Im ungestörten Fall entspricht  $n_{kh}$  eines Zeichens der Zahl der Konturpunkte  $n_{kr}$  der entsprechenden Referenz. Bezieht man  $n_{kr}$  auf die Zeichenhöhe  $n_{zy}$ , so erhält man für die verwendeten Referenzfiguren eins bis null die in Tab. 6.5 angegebenen Werte  $c_{kr}$ . Der arithmetische Mittelwert von  $c_{kr}$  für die zehn Ziffern eins bis null beträgt  $\overline{c_{kr}} = 3.6$ .

Zeichen z	1	2	3	4	5	6	7	8	9	0	$\bar{z}$
$c_{kr} = n_{kr}/n_{zy}$	2.3	3.5	3.6	3.3	4.1	4.3	3.0	4.0	3.7	4.4	3.6

Tabelle 6.5: Typische Anzahl der Konturpunkte einer Referenz.  
( $c_{kr}$  ist die Anzahl der Konturpunkte der Referenz,  
bezogen auf die in Bildpunkte gemessene Zeichenhöhe  $n_{zy}$ )

Als Abschätzung für den gestörten Fall wird angenommen, daß  $n_{kh} \approx 2 n_{kr}$  entspricht, wie es z.B. bei der Darstellung in Abb. 4.12b zutrifft. Bezieht man  $n_{kh}$  auf die Zeichenhöhe  $n_{zy}$  und bildet wie in Tab. 6.5 den Mittelwert von  $c_{kh}$ , so ergibt sich für ein Bild, in dem  $n_z$  Zeichen dargestellt sind

$$n_{kh} = \overline{c_{kh}} n_{zy} n_z = 2 \overline{c_{kr}} n_{zy} n_z = 7.2 n_{zy} n_z \quad (6 - 29)$$

Im folgenden wird nun der Operationsaufwand für die Vorverarbeitungstufen geschätzt.

**Konturpunkt detektion nach Haralik [32]** Bei dem gewählten Verfahren wird für jeden zu untersuchenden Bildpunkt in einem Unterausschnitt der Größe  $5 \times 5$  geprüft, ob es sich um einen Wendepunkt der Approximationsfunktion handelt (Kap. 4.2.4). Zur Bestimmung der Funktion nach Gl 4-16 sind die Koeffizienten  $k_1$  bis  $k_{10}$  für jeden Bildpunkt zu berechnen. Zur Störungsreduktion wird gefordert, daß der Gradientenbetrag  $g_b$  der Funktion über einem Mindestbetrag  $g_{bmin}$  liegt.

$$g_b^2 = k_2^2 + k_3^2 \geq g_{bmin}^2 \quad (6 - 30)$$

Dazu sind für jeden Bildpunkt die Koeffizienten  $k_2$ ,  $k_3$  und  $g_b$  zu berechnen. Der Operationsaufwand für diese Berechnung beträgt für jeden Positionierungspunkt

$$O_{hara1} = 39 O_{add} + 14 O_{mul} \quad (6 - 31)$$

Ist die Bedingung aus Gl. 6-30 erfüllt, werden die Koeffizienten  $k_1, k_3, \dots, k_{10}$  berechnet und durch eine Kurvendiskussion wird geprüft, ob sich in der Nähe des betrachteten Bildpunktes ein Wendepunkt befindet. Für diese Prüfung wird ein maximaler zusätzlicher Aufwand

$$O_{hara2} = 137 O_{add} + 35 O_{mul} \quad (6 - 32)$$

benötigt. Nimmt man an, daß für alle detektierten Konturpunkte  $n_{kh}$  der Aufwand  $O_{hara1} + O_{hara2}$  und für alle anderen Bildpunkte  $n_{pxy} - n_{kh}$  lediglich der Aufwand  $O_{hara1}$  durchgeführt wurde, so ergibt sich für einen Positionierungsbe-  
reich P eines Unterausschnittes ein Gesamtaufwand von:

$$O_{Hara} = n_{pxy} O_{hara1} + n_{kh} O_{hara2} \quad (6 - 33)$$

In einem Suchbereich  $S$  mit  $n_{sxy}$  Bildpunkten gibt es für einen Unterausschnitt  $5 \times 5$  (Maske)  $n_{pxy} = (n_{sx} - 4)(n_{sy} - 4)$  Positionierungsmöglichkeiten. Da davon ausgegangen werden kann, daß  $n_{sx} \gg 4$  und  $n_{sy} \gg 4$  ist, wird näherungsweise  $n_{pxy} = n_{sxy}$  gesetzt.

$$O_{hara} = (39 n_{sxy} + 137 n_{kh}) O_{add} + (14 n_{sxy} + 35 n_{kh}) O_{mul} \quad (6 - 34)$$

Setzt man Gl. 6-29 ein, so erhält man

$$O_{Hara} \approx (39 n_{sxy} + 986 n_{zy} n_z) O_{add} + (14 n_{sxy} + 252 n_{zy} n_z) O_{mul} \quad (6 - 35)$$

**Verdünnung.** Bei der Verdünnungsoperation des Binärbildes wird durch Untersuchung der 8er-Nachbarschaft geprüft, ob ein gesetzter Punkt (Konturpunkt) gelöscht werden kann. Diese Untersuchung besteht im wesentlichen aus Vergleichsoperationen.

Bei dem in *SPIDER* [84] implementierten Algorithmus werden die bei [13] aufgeführten fünf Bedingungen zum Löschen eines Konturpunktes geprüft. Zählt man die zur Bestimmung von Vergleichsgrößen durchgeführten Additionen sowie die durch Additionen ersetzten logischen Operationen, so fällt für die Überprüfung des Löschens eines gesetzten Bildpunktes ein maximaler Operationsaufwand von

$$O_{del} = 23 O_{add} \quad (6 - 36)$$

an. Die Rückweisung einer Löschung erfolgt, sobald eine Bedingung verletzt ist. Das Löschen von Konturpunkten findet iterativ durch ein "Abschälen" des Kontursegmentes statt.

Die Zahl der Bildpunkte, deren Nachbarschaft zu untersuchen ist, entspricht zu Beginn der Iteration der Zahl der Konturpunkte  $n_{kh}$ . Die Zahl zu löschender Konturpunkte wird durch die "Dicke" der erhaltenen Kontursegmente bestimmt. Nach der Verdünnung gibt es  $n_{ks}$  Konturpunkte, die aufgrund experimenteller Untersuchungen mit

$$n_{ks} = 0.8 n_{kh} \quad (6 - 37)$$

angenommen werden. In dem Beispiel von Abb. 4.12c wurden von 509 Konturpunkten 92 gelöscht, woraus sich eine Reduktionsfaktor von  $n_{ks}/n_{kh} = 0.82$  ergibt. Zur Abschätzung des Aufwandes wird angenommen, daß für alle detektierten Konturpunkte der maximale Operationsaufwand  $O_{del}$  nötig ist. Für die Verdünnungsoperation ergibt sich dann ein Gesamtaufwand  $O_{ske}$  von

$$O_{ske} = \overline{c_{kh}} n_{zy} n_z O_{del} \approx 145 n_{zy} n_z O_{add} \quad (6 - 38)$$

**Konturapproximation.** Im Hinblick auf einfache Distanzberechnungen wird die durch End- oder Scheitelpunkte verlaufende Gerade durch die Hessesche Normalform dargestellt. Der Aufwand für die Berechnung der Parameter beträgt

$$O_{hess} = 4 O_{add} + 7 O_{mul} + 1 O_{div} + 1 O_{sgrt} \quad (6 - 39)$$

Die Abstandsberechnung der Konturpunkte zu dieser Geraden erfordert den Aufwand

$$O_{dis} = 2 O_{add} + 2 O_{mul} \quad (6 - 40)$$

Der Aufwand für die Konturapproximation  $O_{appr}$  ergibt sich dann aus der Anzahl der durchgeführten Geradenberechnungen  $n_{cl}$  und der Anzahl der durchgeführten Abstandsberechnungen  $n_{cd}$

$$O_{appr} = n_{cl} \cdot O_{hess} + n_{cd} \cdot O_{dis} \quad (6 - 41)$$

Von Ramer [73] wurde unter der Annahme, daß sich das Approximationspolygon durch einen ausgewogenen Binärbaum beschreiben läßt und somit die Anzahl der Geradenstücke  $n_l$  eine Potenz von zwei darstellt, die Zahl der durchzuführenden Distanzberechnungen  $n_{cd}$  für eine offene Konturpunktkette mit  $n_k + 1$  Punkten bestimmt:

$$n_{cd} = n_k \cdot (1 + \log_2 n_l) - 2 n_l + 1 \quad (6 - 42)$$

Die Wert von  $n_{cl}$  beträgt dabei

$$n_{cl} = 2 n_l - 1 \quad (6 - 43)$$

In unausgewogenen Bäumen kann die Zahl  $n_{cd}$  größer oder kleiner sein. Die Grenzen von  $n_{cd}$  sind:

$$n_k - 1 \leq n_{cd} \leq (n_k - 1) n_l \quad (6 - 44)$$

Die tatsächliche Zahl der zu berechnenden Geraden  $n_{cl}$  wird durch die Form des Linienmusters mit den Störungen und durch die Approximationsgüte  $d_{appr}$  bestimmt. Die Anzahl der Abstandsberechnungen zu einer Approximationsgeraden wird durch die Anzahl der Konturpunkte einer Teilkette bestimmt, die wiederum von Rasterung bzw. Zeichenhöhe  $n_{zy}$  abhängt. Experimentelle Untersuchungen der Polygonapproximation mit ungestörten Mustern, den benutzten Referenzliniendarstellungen, haben bei verschiedenen Rasterungen bei Kenntnis von  $n_l$  und  $n_k$  gezeigt, daß die Werte nach Gl. 6-43 und Gl. 6-42 gute Näherungen bilden.

Im Beispiel von Abb. 4.12c wurde die Konturpunktmenge (417 Punkte) mit der Approximationsgüte von  $d_{appr} = 0.016 n_{zy}$  ( $d_{appr}$  entspricht 1 Bildpunkt) in 86 Linien überführt, wobei ein Aufwand von  $O_{appr}(\text{Abb. 4.12c}) = 115 O_{hess} + 891 O_{dis}$  entstand. Für die weiteren Abschätzungen gestörter Muster werden für die Ziffern eins bis null bei Annahme von  $d_{appr} = 0.016 n_{zy}$  pro Zeichen im Mittel für  $n_{cl} = 100$  und für  $n_{cd} = 100 n_{zy}$  angesetzt.

$$O_{Appr} = (200 n_{zy} + 400) n_z O_{add} + (200 n_{zy} + 700) n_z O_{mul} + 100 n_z O_{div} + 100 n_z O_{sqrt} \quad (6 - 45)$$

**Attributierung.** Den Linienstücken werden Attributwerte zugewiesen. Mit den zu den Linien gespeicherten Projektionen auf die Koordinatenachsen  $\Delta x$ ,  $\Delta y$  werden die zugehörige Steigung bzw. der Winkel aus einer Tabelle entnommen und das attributierte Objekt Linie in eine dem Attributwert zugeordnete Liste eingetragen. Da die Berechnung des Winkels durch einen Tabellenzugriff ersetzt wird, wird kein Operationsaufwand angesetzt.

**Transformation.** Für den Vergleich werden Linien in die Distanzmatrizen projiziert. Sind die Dimensionen des Suchbereiches  $n_{sx}, n_{sy}$  größer als die Dimensionen der Distanzmatrizen  $n_{rx}, n_{ry}$ , so sind die im Überdeckungsbereich liegenden Linien zu selektieren und deren Koordinatenwerte mit einem entsprechenden Offset zu versehen. Die Anzahl der Positionierungsmöglichkeiten bestimmt sich nach Gl. 6-1.

Zur Projektion in die Distanzmatrizen wird zu jeder Linie die entsprechende Punktmenge im gerasterten Bild bzw. in der Distanzmatrix generiert.

Ausgehend vom transformierten Endpunkt wird inkremental entlang einer Koordinatenrichtung durch Addition mit dem Steigungswert  $\Delta y/\Delta x$  bzw. dem reziproken Wert die orthogonale Koordinate bestimmt. Werden  $n_{kg}$  Punkte generiert, so fallen bei langen Linien ungefähr  $n_{kg}$  Additionen an.

Die Zahl der generierten Punkte  $n_{kg}$  wird angenommen als

$$n_{kg} = 0.8 n_{ks} \quad (6 - 46)$$

Nimmt man an, daß es sich bei jeder selektierten Linienmenge um die Darstellung eines Zeichens handelt, beträgt  $n_{kg}$  unter Berücksichtigung von Gl. 6-37 und Gl. 6-29:

$$n_{kg} = 0.8 n_{ks} = 0.64 n_{kh} = 4.6 n_{zy} \quad (6 - 47)$$

Der Operationsaufwand  $O_{trans}$  beträgt

$$O_{trans} = 4.6 n_{pxy} n_{zy} O_{add} \quad (6 - 48)$$

### Linienvergleich

Zur Bestimmung der Abstandsmaße  $k_d$  (Gl. 4-8) und  $k_R$  (Gl. 4-8) ist die Summe der Abstände  $d_{p_i}$  und die Anzahl der Punkte  $n_{kg}^*$ ,  $n_{kr}^*$  zu bilden, so daß für den Vergleich mit einer entsprechenden Referenzdarstellung der Operationsaufwand

$$O_{comp(n_r=1)} \approx (2 n_{kg}^* + n_{kr}^*) O_{add} + 2 O_{div} \quad (6 - 49)$$

anfällt.

Die Zahl der Additionen zur Berechnung der Größen  $k_d$ ,  $k_R$  wird durch die Anzahl der akzeptierten Punkte  $n_{kg}^*$ ,  $n_{kr}^*$  bestimmt, wobei  $0 \leq n_{kg}^* \leq n_{kg}$  und  $0 \leq n_{kr}^* \leq n_{kr}$  gilt. Wird das Linienmuster mit seiner entsprechenden Referenz verglichen, so ist ein hoher Wert für  $n_{kg}^*$  zu erwarten; wird es dagegen mit einer konkurrierenden Referenz verglichen, so ist ein niedriger Wert für  $n_{kg}^*$  zu erwarten. Es wird angenommen, daß für einen Vergleich eines optimal positionierten Musters mit der entsprechenden Referenz  $n_{gk}^* = 0.8 n_{kg}$  akzeptiert werden und bei den  $n_r - 1$  Vergleichen mit konkurrierenden Referenzen im Mittel  $n_{gk}^* = 0.4 n_{kg}$  akzeptiert werden.

Im Beispiel Abb. 4.12, bei dem die Referenz um jeweils vier Bildpunkte nach links und nach oben aus der optimalen Lage verschoben und um  $4^\circ$  verdreht wurde, beträgt  $n_{gk+}^* = 0.63$  und  $\overline{n_{gk-}^*} = 0.48$ . Wird nun eine Liniendarstellung mit einer entsprechenden Referenzdarstellung und  $n_r - 1$  nicht entsprechenden Referenzdarstellungen verglichen, so beträgt der Operationsaufwand  $O_{comp}$

$$O_{comp} \approx (3.7 n_r n_{zy} + 1 n_{zy}) O_{add} + 2 n_r O_{div} \quad (6 - 50)$$

Für  $n_{pxy}$  Positionierungen besteht der Gesamtaufwand  $O_{Comp}$

$$O_{Comp} = n_{pxy} O_{comp} \quad (6 - 51)$$

Der Gesamtaufwand  $O_M$  ist in Tab. 6.6 zusammengestellt. Um das ikonische Bild des Suchbereiches S in eine Liniendarstellung zu überführen, ist der Operationsaufwand  $O_{Hara} + O_{Ske} + O_{Appr}$  nötig.

$O$	$O_{add}$	$O_{mul}$	$O_{div}$	$O_{sqrt}$
$O_{Hara}$	$39 n_{sxy} + 986 n_{zy} n_z$	$14 n_{sxy} + 252 n_{zy} n_z$		
$O_{Ske}$	$145 n_{zy} n_z$			
$O_{Appr}$	$(200 n_{zy} + 400) n_z$	$(200 n_{zy} + 700) n_z$	$100 n_z$	$100 n_z$
$O_{Trans}$	$4.6 n_{zy} n_{pxy}$			
$O_{Comp}$	$(3.7 n_r + 1) n_{zy} n_{pxy}$		$2 n_r n_{pxy}$	

Tabelle 6.6: Operationsaufwand für den Linienvergleich

$$\begin{aligned}
 O_M = & [39 n_{sxy} + 1331 n_{zy} n_z + (3.7 n_r + 5.6) n_{zy} n_{pxy} + 400 n_z] O_{add} + \\
 & [14 n_{sxy} + (452 n_{zy} + 700) n_z] O_{mul} + \\
 & [2 n_r n_{pxy} + 100 n_z] O_{div} + 100 n_z O_{sqrt}
 \end{aligned} \quad (6-52)$$

## 6.4 Berechnungsaufwand zur Einschränkung des Suchbereiches

### 6.4.1 Einschränkung des Suchbereiches auf eine Schriftzeile

Der Aufwand zur Einschränkung des Suchbereiches durch Segmentierung einer Schriftzeile wird nach Gl. 5-1 und Gl. 5-2 bestimmt. Zur Bestimmung des Profils nach Gl. 5-1 sind  $n_{bx} - 1$  Differenzen zu bilden und diese durch  $n_{bx} - 2$  Additionen zu summieren. Führt man diese Operation für  $n_{by}$  Bildzeilen durch, ergibt sich ein Aufwand von

$$O_{s1} = ((n_{bx} - 1) + (n_{bx} - 2)) \cdot n_{by} \cdot O_{add} = (2n_{bx} - 3)n_{by} \cdot O_{add} \quad (6 - 53)$$

$$O_{s1} \approx 2 \cdot n_{bx} \cdot n_{by} \cdot O_{add} \quad (6 - 54)$$

Die für das zu verschiebende Integrationsfenster durchzuführenden Summationen benötigen einen Aufwand von  $O_{s2}$

$$O_{s2} = (n_{by} - n_{zy} + 1) \cdot (n_{zy} - 1) \cdot O_{add} \quad (6 - 55)$$

$$O_{s2} \approx n_{by} \cdot n_{zy} \cdot O_{add} \quad (6 - 56)$$

Vernachlässigt man die Vergleichsoperationen zur Bestimmung des Maximums, so ergibt sich der Gesamtaufwand zur Segmentierung einer Zeile

$$O_{SZ} \approx n_{by}(2n_{bx} + n_{zy})O_{add} \quad (6 - 57)$$

Durch die in Kap. 5.1 beschriebene rechenzeitgünstige Vorgehensweise kann der Berechnungsaufwand gering gehalten werden. Würde man vergleichsweise die Berechnung des Gradienten mit einer entsprechenden Sobel-Operatormaske (3 x 3) durchführen, so ist der Operationsaufwand  $O_{add}$  für jede Differenzbildung in Gl. 5-1 durch den Operationsaufwand  $9O_{mul} + 8O_{add}$  zu ersetzen.

### 6.4.2 Berechnungsaufwand zur Vereinzelung von Zeichen

Der Berechnungsaufwand für die Faltung mit der Sobel-Operator-Maske  $H_y$  beträgt nach Gl. 5-3

$$O_{s3} = (n_{bx} - 2) \cdot (n_{sy} - 2) \cdot (9O_{mul} + 8O_{add}) \quad (6 - 58)$$

$$O_{s3} \approx n_{bx} \cdot n_{sy} \cdot (9O_{mul} + 8O_{add}) \quad (6 - 59)$$

Für die Bildung der Summenverteilung nach Gl. 5-4 sind anzusetzen:

$$O_{s4} = (n_{bx} - 2) \cdot (n_{sy} - 1) \cdot O_{add} \quad (6 - 60)$$

$$O_{s4} \approx n_{bx} \cdot n_{sy} \cdot O_{add} \quad (6 - 61)$$

Die dreimalige Glättung nach Gl. 5-5 benötigt unter Vernachlässigung der Normierung mit dem Faktor  $\frac{1}{3}$  folgenden Operationsaufwand:

$$O_{s5} = (n_{bx} - 4 + n_{bx} - 6 + n_{bx} - 8) \cdot 2 \cdot O_{add} = (6n_{bx} - 36) \cdot O_{add} \quad (6 - 62)$$

Der Gesamtaufwand zur Vereinzelung der Zeichen beträgt näherungsweise

$$O_{SS} \approx 9 \cdot n_{bx} \cdot n_{sy} (O_{add} + O_{mul}) \quad (6 - 63)$$

## 6.5 Bestimmung des relativen Operationsaufwandes

Um den Operationsaufwand, den die verschiedenen Berechnungswege erfordern, gegenüberzustellen, werden die unterschiedlichen Operationen auf den Operationsaufwand  $O_a$  einer Addition bezogen und entsprechend ihrer relativen Ausführungszeiten gewichtet. Legt man z.B. die Operationszeiten eines IMS T800 Prozessors <sup>1</sup> zugrunde, lassen sich für die floating point Operationen (double precision) folgende Operationsaufwände in Bezug setzen

$$\begin{aligned} O_{add} &= O_a \\ O_{mul} &= 3 \cdot O_a \\ O_{div} &= 5 \cdot O_a \\ O_{sqrt} &= 27 \cdot O_a \end{aligned}$$

Wegen dem vergleichsweise hohen Operationsaufwand für das Wurzelziehen  $O_{sqrt}$ , wird man versuchen eine häufige Nutzung der Wurzelfunktion zu vermeiden. Oftmals können auch die entsprechenden quadratischen Größen zu einer Entscheidung herangezogen werden oder die Wurzeln in einer Tabelle gespeichert werden. Bei der vorliegenden Abschätzung wurden diese Optimierungen nicht vorgenommen.

Es ergeben sich für den Aufwand  $O_{K1}$  (direkte Korrelation, Maske n-wertig) nach Gl. 6-13,  $O_{K2}$  (direkte Korrelation, Maske zweiwertig) nach Gl. 6-18,  $O_{K3}$  (indirekte Korrelation, Maske n-wertig) nach Gl. 6-28 und  $O_M$  (Linienvergleich) nach Gl. 6-52 folgende Größen:

$$O_{K1} = n_r \cdot n_{pxy} \cdot (9n_{r\odot} + 36)O_a \quad (6 - 64)$$

$$O_{K2} = n_r \cdot n_{pxy} \cdot (5n_{r\odot} + 73)O_a \quad (6 - 65)$$

---

<sup>1</sup>32 Bit Transputer mit floating point unit

$$O_{K3} = (n_r \cdot n_{qxy}(27 \text{ld } n_{qxy} + 13) + n_{qxy}(18 \text{ld } n_{qxy} - 25))O_a \quad (6 - 66)$$

$$O_M = (81n_{sxy} + ((5.6 + 3.7n_r)n_{zy} + 10n_r)n_{pxy} + (2687n_{zy} + 5700)n_z)O_a \quad (6 - 67)$$

Die Beziehung zwischen  $O_{K1}$  und  $O_{K2}$  läßt sich für die Annahme  $n_{r\odot} > 100$  näherungsweise durch

$$\frac{O_{K1}}{O_{K2}} \approx 1.8 \quad (6 - 68)$$

ausdrücken.

Der Aufwand für die Korrelation bzw. den Linienvergleich mit nur einer Referenz beträgt für  $n_{r\odot} > 100$  und  $\text{ld } n_{qxy} > 10$  näherungsweise

$$O_{K2(n_r=1)} \approx 5 n_{pxy} n_{r\odot} O_a \quad (6 - 69)$$

$$O_{K3(n_r=1)} \approx n_{qxy} 45 \text{ld } n_{qxy} O_a \quad (6 - 70)$$

$$O_{M(n_r=1, n_z=1)} \approx (81n_{sxy} + (10n_{pxy} + 2687)n_{zy})O_a \quad (6 - 71)$$

## 6.6 Berechnungsbeispiele

Im folgenden wird der Berechnungsaufwand an verschiedenen Beispielen durch Annahme konkreter Randbedingungen bestimmt und gegenübergestellt. Es wird von einem aufgenommenen Grauwertbild B der Größe  $n_{bx} = 512$  und  $n_{by} = 512$  ausgegangen. Mit der Einschränkung des Bildbereiches B auf den Suchbereich S werden zur Berechnung, ausgehend von den beispielhaft ausgewählten Suchbereichen  $S_1$  bis  $S_5$  (Abb. 6.2), die Fälle a-e unterschieden:

- a) Der Bildbereich B kann nicht eingeschränkt werden. Der Suchbereich  $S_1$  entspricht dem Bildbereich B
- b) Mit Hilfe von Vorwissen über den Auftrittsort von Zeichen kann der Bildbereich B auf den Suchbereich  $S_2$  mit der Größe  $n_{sx} = 256$  und  $n_{sy} = 128$  eingeschränkt werden.
- c) Wie in Fall b), jedoch wird der Bildbereich B auf  $S_3$  mit der Größe  $n_{sx} = 300$  und  $n_{sy} = 100$  eingeschränkt.
- d) Es wird angenommen, daß in einer Vorverarbeitungsstufe eine Zeilensegmentierung durchgeführt wurde und der resultierende Suchbereich  $S_4$  auf eine Größe von  $n_{sx} = 512$  und  $n_{sy} = 80$  eingeschränkt ist.

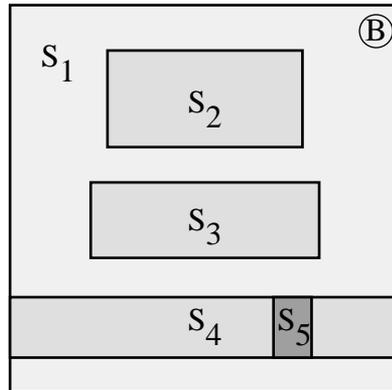


Abb. 6.2: Darstellung der Suchbereiche  $S_1$  bis  $S_5$  im Bildbereich B

- e) Wie in Fall d), jedoch wurde zusätzlich zur Zeilensegmentierung eine Zeichenvereinzelung durchgeführt, so daß der Suchbereich  $S_5$  für ein Zeichen auf die Größe  $n_{sx} = 44$  und  $n_{sy} = 80$  eingeschränkt wird.

Es wird angenommen, daß sich im Suchbereich  $n_z$  Zeichen befinden, die eine Zeichenhöhe  $n_{zy} = 70$  Bildpunkte besitzen. In Tab. 6.7 sind die geometrischen Größen der Suchbereiche zusammengestellt (siehe Tab. 6.1 und Gl. 6-9). Da für die Korrelation

Fall	$n_z$	S	$n_{sx}$	$n_{sy}$	$n_{qx}$	$n_{qy}$	$n_{sxy}$	$n_{qxy}$
a)	5	$S_1 = B$	512	512	512	512	262.144	262.144
b)	5	$S_2 < B$	256	128	256	128	32.768	32.768
c)	5	$S_3 < B$	300	100	512	128	30.000	65.536
d)	5	$S_4 < B$	512	80	512	128	40.960	65.536
e)	1	$S_5 < B$	44	80	64	128	3.520	8.192

Tabelle 6.7: Zusammenstellung der geometrischen Größen der Suchbereiche für die Fälle a bis e (vgl. Sbb. 6.2)

bei den Verfahren I und III eine unterschiedliche Zahl von Masken verschiedener Größe Verwendung finden, wird der konkrete Berechnungsaufwand in den Fällen A und B unterschieden. Im Fall A wird jeweils das gesamte Zeichen durch eine Referenzmaske erfaßt. Die mittlere Größe der 10 Referenzmasken beträgt  $n_{rx} = 36$  und  $n_{ry} = 70$ , wobei im Mittel  $n_{r\odot} = 928$  Bildpunkte zur Abstandsberechnung herangezogen werden. Im Fall B werden durch die Referenzmasken Teile von Zeichen erfaßt. Die mittlere Größe der  $n_r = 15$  Referenzmasken beträgt  $n_{rx} = 24$  und  $n_{ry} = 28$ , wobei im Mittel  $n_{r\odot} = 250$

Bildpunkte zur Abstandsberechnung herangezogen werden. Die sich aus den Suchbereichen und Maskengrößen ergebenden Positionierungsbereiche  $n_{pxy}$  (Gl. 6-1) sowie die Gesamtzahl der zu berechnenden Korrelationswerte  $n_k$  sind in der Tab. 6.8 zusammengestellt.

Fall	$n_{rx}$	$n_{ry}$	$n_{r\ominus}$	$n_{pxy}$	$n_r$	$n_k$	
A	a)	36	70	928	211.311	10	2.113.110
	b)	.	.	.	13.039	.	130.390
	c)	.	.	.	8.215	.	82.150
	d)	.	.	.	5.247	.	52.470
	e)	.	.	.	99	.	990
B	a)	24	28	250	237.165	15	3.557.475
	b)	.	.	.	23.533	.	352.995
	c)	.	.	.	20.221	.	303.315
	d)	.	.	.	25.917	.	388.755
	e)	.	.	.	1.113	.	1.113

**Tabelle 6.8:** Größe des Positionierungsbereiches  $n_{pxy}$  und Zahl der Korrelationswerte  $n_k$  für die Fälle Aa - Be.

Für dieselben Suchbereiche weisen die Positionierungsbereiche  $n_{pxy}$  im Fall B wegen der kleineren Maske gegenüber Fall A einen größeren Wert auf (vgl. in Tab. 6.8 z.B. Ab mit Bb). Bei annähernd gleich großen Suchbereichen  $n_{sxy}$ , wie z.B im Fall b und c (Tab. 6.7) können die die entsprechenden Größen  $n_{pxy}$  der Positionierungsbereiche aufgrund der unterschiedlichen Seitenverhältnisse von  $S_2$  und  $S_3$  stark voneinander abweichen (vgl. in Tab. 6.8 Ab und Ac ). Vergleicht man die Suchbereiche und Positionierungsbereiche im Fall c mit denen von Fall d, so führt im Fall d der größere Suchbereich  $S_4$  im Fall B auch zu einem größeren Positionierungsbereich, im Fall A jedoch zu einem kleineren Positionierungsbereich als bei Fall c.

### Operationsaufwand zur Einschränkung des Suchbereichs

Ein Einschränkung des Suchbereiches durch eine Vorverarbeitungsstufe wird in den Fällen d und e angenommen. Im Fall d ergibt sich zur Segmentierung einer Zeile mit  $n_{zy} = 70$  und den in Tab. 6.7 angegebenen Werten der Operationsaufwand nach Gl. 6-57 zu  $O_{SZ} \approx 560000O_a$  (Tab. 6.9). Durch die Einschränkung auf den Suchbereich  $S_4$  reduziert man B um 221184 Bildpunkte auf ungefähr ein Sechstel der ursprünglichen Bildpunkte.

Im Fall e wird in der Vorverarbeitungsstufe zusätzlich zur Zeilensegmentation eine Zeichenvereinzelung vorgenommen, die den Suchbereich von  $S_4$  auf  $5 \cdot S_5$  einschränkt. Der dazu nötige Operationsaufwand beträgt nach Gl. 6-63  $O_{SS} = 1474560O_a$ . Der Suchbereich  $S_4$  verringert sich dabei um 23360 Bildpunkte auf ungefähr die Hälfte der ursprünglichen Bildpunkte. Im Vergleich zur Zeichenvereinzelung werden bei der Zeilensegmentation ca. zehnmal mehr Bildpunkte reduziert.

Fall	$O_{ZS}$	$O_{SS}$	$n_{bxy} - n_{sxy}$
d)	0.56		221184
e)	0.56	1.47	244544

Tabelle 6.9: Operationsaufwand  $O_{ZS}$ ,  $O_{SS}$  (in Millionen  $O_a$ ) zur Einschränkung des Suchbereiches

### Operationsaufwand für die Korrelation

Der für die Korrelation aufzubringende Operationsaufwand für die Fälle Aa - Be ist für die Berechnung mittels Maskenoperation nach Gl. 6-65 ( $O_{K2}$ ) und mittels FFT nach Gl. 6-66 ( $O_{K3}$ ) in Tab. 6.10 dargestellt.

Fall	A (Verfahren I)		B (Verfahren III)	
	MskOp $O_{K2}$	FFT $O_{K3}$	MskOp $O_{K2}$	FFT $O_{K3}$
a)	9.959	1384	4.707	2.037
b)	615	145	468	213
c)	387	308	401	454
d)	247	308	514	454
e)	5	31	22	46

Tabelle 6.10: Operationsaufwand  $O_a$  (in Millionen) für die Berechnung der  $n_k$  Korrelationswerte für  $O_{K2}$  nach Gl. 6-65 und für  $O_{K3}$  nach Gl. 6-66

Der Vorteil des geringeren Operationsaufwandes der FFT gegenüber der Maskenoperation kommt dann zum Tragen, wenn die Anzahl  $n_q$  identisch mit der Anzahl  $n_s$  ist (Fall a, b) und wenn große Masken verwendet werden (Fall A). Sind die Größen  $n_q$  und  $n_s$  nicht identisch, so kann in Abhängigkeit von Maskengröße und Anzahl (Fall A, B) sowie von der Geometrie des Suchbereiches (Fall c, d) bei gleichen Suchbereichen sowohl

die Berechnung mittels FFT als auch die Berechnung mittels Maskenoperation günstiger sein. Die Berechnung der Korrelation mittels FFT wird dann ungünstiger, wenn sehr kleine Suchbereiche (Fall e) berücksichtigt werden.

### Operationsaufwand für den Linienvergleich

In Tab. 6.11 ist der Operationsaufwand für die einzelnen Vorverarbeitungsstufen (Kap. 6.3) und Fälle a - e (Tab. 6.7) zusammengestellt.

	$O_{Hara}$	$O_{Ske}$	$O_{Appr}$	$O_{Trans}$	$O_{Comp}$	$O_M$
a)	21.79	0.05	0.28	62.21	535.04	619
b)	3.21	0.05	0.28	3.84	33.01	40
c)	2.99	0.05	0.28	2.42	20.80	27
d)	3.88	0.05	0.28	1.54	13.29	19
e)	0.40	0.01	0.23	0.03	0.25	0.9

Tabelle 6.11: Operationsaufwand  $O_M$  (in Millionen  $O_a$ ) für Verfahren II

### Vergleich des Operationsaufwandes der Verfahren I , II, III

Der Aufwand für den Spitzen-Detektor bei Verfahren I und III wird nicht berücksichtigt, da zur Detektion im wesentlichen Vergleichsoperationen benötigt werden. Zudem wird die Anzahl der Bildpunkte, die dem Schwellwertkriterium nach Gl. 4-1 genügen, als gering angenommen. Ebenso bleibt für Verfahren I und II der Aufwand des Maximumdetektors für die Zuweisung einer Zeichenklasse bzw. Rückweisung unberücksichtigt.

Bei syntaktischen Verfahren kann i. allg. für das Parsen, d.h. für die Suche nach kompatiblen Objekten bei einer großen Anzahl von Primitivobjekten ein erheblicher Zeitaufwand anfallen. Da zum einen im Verfahren III alle Zwischenergebnisse (Teilobjekte) gespeichert werden und zum andern für die Suche nach Objekten von einem assoziativen Zugriff (siehe Abb. 4.21) ausgegangen wird, wird der vom Bilddatenmaterial abhängige Aufwand für die syntaktische Analyse nicht in Rechnung gestellt.

Allgemein wird man für einen großen Suchbereich sowohl für Verfahren I als auch für Verfahren III die Berechnung der Korrelation mittels FFT statt der Berechnung mittels

Maskenoperation vorziehen. Aus Gl. 6-28 ist ersichtlich, daß sich der Operationsaufwand der beiden Verfahren für einen gleich großen Suchbereich nur bei unterschiedlicher Größe von  $n_r$  unterscheidet. In den betrachteten Beispielen (Tab. 6.8) ist die Zahl der Referenzen  $n_r$  bei der Korrelation mit Teilmasken, so wie sie im Verfahren III verwendet werden, höher ( $n_r = 15$ ) als bei der Korrelation mit Vollmasken, wie sie im Verfahren I verwendet werden ( $n_r = 10$ ). Verfahren I ist somit für große Suchbereiche bezüglich des Operationsaufwandes immer günstiger als Verfahren III.

Dieser Vorteil könnte erst dann aufgehoben werden, wenn zur Erkennung einer Menge unterschiedlicher Zeichen die Zahl der sie beschreibenden Typen von Primitivobjekten bzw. Teilmasken geringer ist, als die Zahl der unterscheidbaren Zeichen bzw. Vollmasken.

Der Gesamtaufwand für die drei Verfahren ist in Tab. 6.12 gegenübergestellt und in Form eines Säulendiagramms in Abb. 6.3 dargestellt.

Fall	Verfahren I	Verfahren II	Verfahren III
a)	1.384	619	2.037
b)	145	40	213
c)	308	27	401
d)	247	19	454
e)	5	0,9	22

Tabelle 6.12: Operationsaufwand (in Millionen  $O_a$ ) für Verfahren I, II, III.  
Für Verfahren I und III wurde jeweils das Minimum aus Tab. 6.10 gewählt

Abb. 6.3: Graphische Darstellung der Aufwandswerte aus Tab. 6.10

## 7. Diskussion

In Kap. 4 wurden drei Verfahren vorgestellt, die es ermöglichen, in Metall geschlagene Zeichen, die in Form eines Grauwertbildes aufgenommen wurden, zu erkennen.

Bei der Entwicklung und Untersuchung der Verfahren wurde berücksichtigt, daß sich die Beleuchtung und damit das Erscheinungsbild der Zeichen ändern kann. Bezüglich der Änderungen der Beleuchtung zeigte Verfahren I das toleranteste Verhalten. Bei Verfahren II kann es bei einer fleckigen Verteilung von Glanzstellen im Hintergrund leicht zu Fehlzuweisungen kommen. Da bei der Konturpunktdetektion keine für die Zeichen spezifische Information (wie z.B. geometrische Information) eingeht, werden die Konturen von Störungen, deren Anzahl sehr hoch sein kann, in gleicher Weise wie die Objektkonturen weiterverarbeitet. Verfahren III reagiert sehr empfindlich, wenn Teile eines Zeichens nicht sichtbar sind bzw. nicht detektiert werden können. Da der Analyseprozess von den Primitivobjekten ausgeht, kann die Objektstruktur eines Zeichens nicht vollständig aufgebaut werden, wenn für den Aufbau von Teilstrukturen Primitivobjekte fehlen. Aus diesem Grunde wird man Schwelle (Gl. 4-1) und Anforderung zur Detektion der Primitivobjekte (Kap. 4.1.3) niedrig ansetzen, so daß sehr viele Objektprimitive generiert werden. Durch die Überprüfung der geometrischen Relationen zwischen den Objekten filtert man nichtentsprechende Objekte heraus.

Verfahren I erwies sich als empfindlich bezüglich der Rotation der Bildvorlage aus der Referenzlage. Dieses Verhalten ist durch die mit einer starren geometrischen Form durchgeführte Ähnlichkeitsmessung erklärbar. Ein ähnlich empfindliches Verhalten ist auch bei perspektivischen Verzerrungen zu erwarten. Durch die geometrischen Toleranzbereiche bei der Zuordnung von Teilen des Musters zur Referenz (Verfahren II) und der Zusammensetzung von Musterteilen (Verfahren III) zeigen die Verfahren II und III gegenüber Verfahren I ein unempfindlicheres Verhalten bei einer Rotation der Vorlage.

Der in Kap. 6 bestimmte Operationsaufwand kann aufgrund der getroffenen Annahmen nur als grobe Abschätzung für den Vergleich der Verfahren aufgefaßt werden. Mit den Ergebnissen aus Kap. 6.6 und Abb. 6.3 wird Verfahren III als das Verfahren mit dem höchsten Operationsaufwand und Verfahren II als Verfahren mit dem geringsten Operationsaufwand eingestuft. Zu berücksichtigen ist, daß der für das Verfahren II bestimmte Operationsaufwand im Gegensatz zu Verfahren I und III vom Bildinhalt abhängt und der Operationsaufwand für Verfahren III wegen der Vernachlässigung des Aufwandes für die syntaktische Analyse als unterer Grenzwert anzusehen ist.

Aus den Abb. 4.2 und Abb. 4.19 ist ersichtlich, daß sich die parallele Struktur des Erkennungsystems in einfacher Weise in das Konzept einer zeitlich parallelen Bearbeitung

durch mehrere Prozessoren überführen läßt, indem z.B. für jeden Korrelationszweig ein Prozessor eingesetzt wird. Aus Abb. 4.10 ist ersichtlich, daß eine Parallelisierung in einfacher Weise auch für den Vergleich mit den Distanzmatrizen möglich ist.

Es wird für den Operationsaufwand  $O_a$  unter Berücksichtigung der Ausführungszeit für eine Addition von 10 Taktzyklen bei einer Taktfrequenz von 20 MHz (= 500 ns), sowie unberücksichtigter Lade-, Speicher-, und Vergleichsoperationen ein grober Richtwert von 1  $\mu$ s angenommen. Dann entspricht ein Operationsaufwand von 100 Mio  $O_a$  für einen Prozessor einer Ausführungszeit von 100 Sekunden. Geht man davon aus, daß die parallele Struktur des Verfahrens I für zehn Referenzen durch zehn parallel angeordnete Prozessoren oder Prozessorgruppen realisiert wird, reduziert sich die Ausführungszeit auf 10 Sekunden. Schätzt man weiterhin ab, daß man eine Suchfläche von 256 \* 128 Bildpunkte in acht quadratische Suchflächen aufteilt und jedes Quadrat von einem Prozessor bearbeitet wird, dann ergibt sich für die zu bearbeitende Suchfläche ohne Berücksichtigung des Transpositionsaufwandes eine Ausführungszeit von 1,25 Sekunden. Befinden sich in diesem Suchbereich fünf Zeichen, so ergibt sich bei dem Einsatz von 80 Prozessoren eine Bearbeitungsrate von vier Zeichen pro Sekunde. Dieser Überschlag läßt erkennen, daß Realisierungen auf der Basis der Korrelation teuer sind bzw. geringe Bearbeitungsraten besitzen.

Eine starke Reduktion des Operationsaufwandes kann erreicht werden, wenn in eine dem Erkennungsprozeß vorgeschalteten Stufe, wie in Kap. 5 beschrieben, eine Einschränkung des Suchbereiches erreicht wird. Ausgehend von einem quadratischen Grauwertbild ist der Gewinn bei einer Schriftzeilensegmentierung verglichen zur Segmentierung einzelner Schriftzeichen hoch.

In Tab. 7.1 sind die Bewertungen der Verfahren zusammenfassend gegenübergestellt.

+ hoch, o mittel, - gering	Verfahren I	Verfahren II	Verfahren III
Beleuchtungsunempfindlichkeit	+	o	o
Rotationsunempfindlichkeit	-	o	o
Vorteil bzgl. Operationsaufwand	o	+	-
Parallelisierbarkeit der Struktur	+	o	+

Tabelle 7.1: Bewertung der Verfahren I, II, III

Durch den Einsatz von a-priori-Wissen ergeben sich Möglichkeiten zur Erhöhung der Erkennungssicherheit und Reduktion des Berechnungsaufwandes. So läßt sich z.B.

durch Einführung von Prüfzeichen und Auswertung dieser Zeichen die Erkennungssicherheit erhöhen.

Eine Reduktion des Berechnungsaufwandes kann erfolgen, wenn ein Zeichensatz für bestimmte Produktionsgruppen oder für bestimmte Zeichenpositionen beschränkt wird oder wenn Zeichen an bestimmten Positionen der Zeichenkette vorher bekannt sind, wie z.B bei Fahrgestellnummern von Kraftfahrzeugen.



## Literaturverzeichnis

- [1] Aho AV, Ullmann JD (1972) The theory of parsing, translation, and compiling, Vol. I: Parsing. London: Prentice-Hall
- [2] Batchelor BG (1985) Lighting and viewing techniques. In: Batchelor BG, Hill DA, Hodgson DC (eds) Automated visual inspection. Amsterdam: North Holland
- [3] Bässmann H, Besslich P (1989) Konturorientierte Verfahren in der digitalen Bildverarbeitung Berlin: Springer
- [4] Barrow HG, Tenenbaum JM, Bolles RC, Wolf HC (1977) Parametric correspondence and chamfer matching: Two new techniques for image matching. 5th International Joint Conference on Pattern Recognition, 659-663
- [5] Bohner M (1974) Ein Verfahren zur quantitativen Bestimmung der Unterscheidbarkeit gedruckter Schriftzeichen. Karlsruhe: Universität, Fak. f. Elektrotechnik, Diss.
- [6] Bronstein IN, Semendjajew KA (1980) Taschenbuch der Mathematik. Thun: Harri Deutsch
- [7] Bunke H (1985) Modellgesteuerte Bildanalyse. Stuttgart: Teubner
- [8] Cook RL, Torrance KE (1982) A reflection model for Computer graphics. *ACM Trans Graph*, **1**: 7-24
- [9] Cooper LA (1975) Mental rotation of random two-dimensional shapes. *Cognitive Psychology*, **7**: 20-43
- [10] Cooper LA, Shepard RN (1975) Transformation on representation of objects in space. In: Carterette EC, Friedman MP (eds) Handbook of perception. New York: Academic Press
- [11] Danielsson PE (1980) Euclidean distance mapping. *Computer Graphics and Image Processing*, **14**: 227-248
- [12] Dehesea M, Sachs C, Liedtke CE (1989) Recognition of engraved characters using distributed associative memories. SCIA Scandinavian Conf. on Image Analysis, Oulo, Finnland, 112-119
- [13] Deutsch ES (1972) Thinning algorithms on rectangular, hexagonal, and triangular arrays. *CACM*, **15**: 827-837

- [14] Duda RO, Hart PE (1973) Pattern classification and scene analysis. New York: Wiley, 328-339
- [15] Dunham JG (1986) Optimum uniform piecewise linear approximation of planar curves. *IEEE PAMI*, **8**: 67-75
- [16] Duerr B, Haettich W, Tropf H, Winkler G (1980) A combination of statistical and syntactical pattern recognition applied to classification of unconstrained handwritten numerals. *Pattern Recognition*, **12**: 189-199
- [17] Englemore R, Morgan T (eds)(1988) Blackboard systems. Wokingham: Addison-Wesley
- [18] Evers C, Kämmerer B (1989) Recognition of punched characters on workpieces by matching of gradients with thin reference patterns. SCIA'89, 577-584
- [19] Fieß HJ (1989) Erkennung von in Metall geprägten Zeichen mit einem konnektionistischen Ansatz. Karlsruhe: Universität, Fak. f. Elektrotechnik, Inst. f. Biokybernetik und Biomed. Technik, Diplomarbeit
- [20] Frisby JP (1979) Seeing, Brain and Mind. Oxford: Oxford University Press
- [21] Fu KS (1982) Syntactic pattern recognition and applications. New Jersey: Prentice-Hall
- [22] Füger H, Lütjen K, Jurkiewicz K (1990) Kontextsensitive Bildanalyse in Luftbildern. In: Großkopf RE (ed) Mustererkennung 1990, 12. DAGM-Symposium. Berlin: Springer, 585-592
- [23] Füger H, Jurkiewicz K, Lütjen K, Stilla U (1992) Ein wissensbasiertes System für die automatische Bildanalyse. In: Fritz LW, Lucas JR (eds) International archives of photogrammetry and remote sensing, Int. soc. f. photogram. a rem. sen. (ISPRS), XVIIth Congress, Washington. Bethesda, MD: ASPRS, Vol XXIX, Commission III, 167-172
- [24] Fukushima K (1980) Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, **36**: 193-202
- [25] Fukushima K (1988) Neocognitron: A hierarchical neuronal network capable of visual pattern recognition. *IEEE Computers*, **3**
- [26] Gerlach H (1979) Digitale Bildfolgenauswertung zum Wiederfinden von Objekten in natürlicher Umgebung. Karlsruhe: Diss.

- [27] Gonzalez RC, Thomason MG (1978) Syntactic pattern recognition. Reading: Addison-Wesley
- [28] Groch WD (1988) The comparison of networks of lines using extended distance transformations; an important tool for matching. Ettlingen: FIM, interner Bericht
- [29] Hahn M (1990) Erkennung von in Metall geprägten Zeichen durch Skelettbildung und syntaktische Bildanalyse. Karlsruhe: Universität, Fak. f. Elektrotechnik, Inst. f. Biokybernetik und Biomed. Technik, Diplomarbeit
- [30] Hänisch N (1990) Erkennung von in Metall geprägten Zeichen durch ein mehrstufiges numerisches Verfahren. Karlsruhe: Universität, Fak. f. Elektrotechnik, Inst. f. Biokybernetik und Biomed. Technik, Diplomarbeit
- [31] Hall R (1986) A characterization of illumination models and shading techniques. *The visual computer*, **2**: 268-277
- [32] Haralick RM (1984) Digital step edges from zero crossing of second directional derivatives *PAMI*, **6**: 58-68
- [33] Hecht E, Zajac A (1974) Optics. Reading: Addison-Wesley
- [34] Herschel R (1974) Einführung in die Theorie der Automaten, Sprachen und Algorithmen. München: Oldenbourg
- [35] Hilditch CJ (1969) Linear skeleton from square cupboards. *Machine Intelligence*, **6**: 403-420
- [36] Hildreth EC (1983) The detection of intensity changes by computer an biological vision systems. *Computer Vision, Graphics, and Image Processing*, **22**: 1-27
- [37] Hill DA (1985) Illumination equipment. In: Batchelor BG, Hill DA, Hodgson DC (eds) Automated visual inspection. Amsterdam: North Holland
- [38] Hill R, Lee EH, Tupper SJ (1947) The theory of wedge indentation of ductile materials. *Proc. Roy. Soc. London (A)*, **188**: 273-289
- [39] Holder S, Dengler J (1987) Erkennung eines eingeschränkten Zeichensatzes mit größen- und fontinvarianten Merkmalen. In: Paulus E (ed) Mustererkennung 1987, 9. DAGM-Symposium. Berlin: Springer, 89
- [40] Holder S, Dengler J (1988) Font- and Sizeinvariant Character Recognition with Greyvalue Image Features. Pattern Recognition, 9th Int. Conf., Proc. 252-254
- [41] Horn BKP (1977) Understanding image intensities. *Artificial Intelligence*, **8**: 201-231

- [42] Horn BKP (1986) Robot Vision. Cambridge: MIT Press
- [43] Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.*, **195**: 215-243
- [44] Hull JJ, Srihari SN, Cohen E, Kuan L, Cullen P, Palumbo P (1988) A blackboard-based approach to handwritten ZIP code recognition. Pattern Recognition, 9th Int. Conf., Proc. 111-113
- [45] Kazmierczak H (ed)(1980) Erfassung und maschinelle Verarbeitung von Bilddaten: Grundlagen und Anwendungen. Wien: Springer
- [46] Kohonen T (1984) Self-organization and associative memory. Berlin: Springer
- [47] Korn A (1982) Bildverarbeitung durch das visuelle System. Fachberichte Messen, Steuern, Regeln, Bd. 8, Berlin: Springer
- [48] Levine MD (1985) Vision in man and machine. New York: McGraw-Hill
- [49] Lesser VR, Fennell RD, Erman LD, Reddy DR (1975) Organization of the Hearsay-II speech understanding system. *IEEE Transaction on Acoustic, Speech, and Signal Processing* **23**: 11-24
- [50] Lütjen K (1986) BPI: Ein Blackboard-basiertes Produktionssystem für die automatische Bildauswertung. In: Hartmann G (ed) Mustererkennung 1986, 8. DAGM-Symposium. Berlin: Springer, 164-168
- [51] Lütjen K, Füger H, Greif HJ, Jurkiewicz K (1987) Auswahlverfahren für die wissensbasierte Bildauswertung mit dem blackboard-orientierten Produktionssystem BPI. In: Paulus E (ed) Mustererkennung 1987, 9. DAGM-Symposium. Berlin: Springer, 290-294
- [52] Magnenat-Thalmann N, Thalmann D (1987) Image Synthesis: Theory and Practice Tokyo: Springer
- [53] Marr D, Hildreth E (1980) Theory of edge detection. *Proc R Soc Lond*, **207**: 187-217
- [54] Marr D (1982) Vision. San Francisco: Freeman
- [55] Martinek D (1989) Automatische Merkmalextraktion aus Videobildern zur visuellen Erkennung geprägter Ziffern unter verschiedenen Beleuchtungsbedingungen. Karlsruhe: Universität, Fak. f. Elektrotechnik, Inst. f. Biokybernetik und Biomed. Technik, Diplomarbeit

- [56] Nehl E (1988) Eindrückverfahren. In: Lange K (ed) Umformtechnik: Handbuch für Industrie und Wissenschaft, Bd.2 Berlin: Springer, 575-604
- [57] Newell A (1962) Some problems of basic organization in problem-solving programs. In: Yovits MC, Jacobi GT, Goldstein GD (eds) Proceedings second conference on self-organizing systems, Spartan Books, 393-423
- [58] Niemann H (1983) Klassifikation von Mustern. Berlin: Springer
- [59] Niemann H, Bunke H (1987) Künstliche Intelligenz in Bild und Sprachanalyse. Stuttgart: Teubner
- [60] Niemann H, (1990) Pattern analysis and understanding. Springer: Berlin
- [61] Nii HP (1986) Blackboard systems. *AI Magazine*, **7**: 82-106
- [62] Norm DIN 66008 (1978) Schrift A für die maschinelle optische Zeichenerkennung: Zeichen und Nennmaße.
- [63] Norm DIN 66009 (1977) Schrift B für die maschinelle optische Zeichenerkennung: Zeichen, Nennmaße und Anordnung auf dem Zeichenträger.
- [64] Norm DIN 66223 (1983) Schriften für die maschinelle optische Zeichenerkennung: Zeichenträger, gedruckte Zeichen, Anforderung und Prüfung.
- [65] Norm ISO 1831 (1980) Printing specification for optical character recognition.
- [66] International Business Forms Industries (1969) Optical character recognition and the years ahead. Illinois, Elmhurst: The Business Press
- [67] Pavlidis T (1977) Structural Pattern Recognition. Berlin: Springer
- [68] Pavlidis T (1982) Algorithms for Graphics and Image Processing. Berlin: Springer
- [69] Phong BT (1975) Illumination for Computer generated pictures. *Commun ACM*, **18**: 311-317
- [70] Post E (1943) Formal reductions of the general combinatorial problem. *American Journal of Mathematics*, **65**: 197-268
- [71] Prager W, Hodge PG (1954) Theorie ideal plastischer Körper. Wien: Springer
- [72] Prandtl L (1920) Über die Härte plastischer Körper. Göttinger Nachr., math.-phys. Kl., 74-85
- [73] Ramer U (1972) An iterative procedure for the polygonal approximation of plane curves. *CGIP*, **1**: 244-256

- [74] Rosenfeld A (1988) Image analysis and computer vision: 1987. *CVGIP*, **42**: 234-293
- [75] Rosenfeld A (1989) Image analysis and computer vision: 1988. *CVGIP*, **46**: 196-264
- [76] Rosenfeld A (1990) Image analysis and computer vision: 1989. *CVGIP*, **50**: 188-240
- [77] Salomaa AK (1973) Formal languages. New York: Academic Press
- [78] Schantz HF (1982) The history of OCR. Manchester: Recognition technologies users association.
- [79] Schmidt R (1987) Contour extraction of objects with shadows and reflections by object-adapted filtering. In: Meyer-Ebrecht (ed) ASST'87, 6. Aachener Symposium für Signaltheorie. 139-143
- [80] Schmidt R (1988) Verfahren zur Konturfindung bei Objekten mit spiegelnden Reflexionseigenschaften. Karlsruhe: Universität, Fak. f. Elektrotechnik, Diss.
- [81] Stilla U, Jurkiewicz K (1991) Objektklassifikation mit einem blackboardorientierten Inferenzmechanismus. Ettlingen: FIM/FGAN, FIM Bericht 230
- [82] Tamura H (1975) Further considerations on line thinning schemes. Paper of IECEJ Techn. Group: PRL 75-66 (siehe [83, 84])
- [83] Tamura H (1978) A comparison of the line thinning algorithms from digital geometry viewpoint. *IJCPR*: 715-719
- [84] Tanaka Y (ed) (1983) SPIDER: Subroutine package for image data enhancement and recognition. Tokio: Joint System Development Corp.
- [85] Tasto M, Block U (1977) Locating objects in complex scenes using a spatial distance measure. 3th International Joint Conference on Pattern Recognition, 336-340
- [86] Tsai RY (1986) An efficient and accurate camera calibration technique for 3D machine vision. *CVPR'86*, IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 364-374
- [87] Tsuruoka et al. (1983) Thinning algorithms for digital pictures and their application to handprinted characters recognition. *IECE Trans.*, **J66-D**: 525-532
- [88] Velthuisen H (1992) The nature and applicability of the blackboard architecture. Leidschendam: PTT Research

- [89] Waerden van der BL (1965) *Mathematische Statistik*. Berlin: Springer
- [90] Wang PSP (ed)(1991) *Character & handwriting recognition: Expanding frontiers*. Singapur: World Scientific
- [91] Winston PH (1975) *The psychology of computer vision*. New York: McGraw-Hill
- [92] Yamada H (1984) Complete euclidean distance transformation by parallel operation. 7th International Joint Conference on Pattern Recognition, 69-71



## A. Anhang

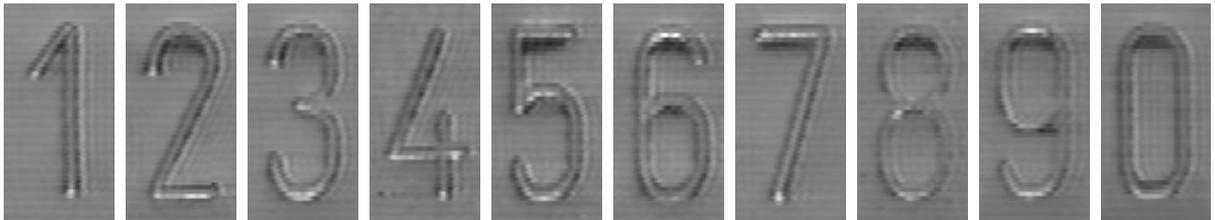


Abb. A.1: Beispiel zum Erscheinungsbild geprägter Zeichen



Abb. A.2: Beispiel zum Erscheinungsbild gefräster Zeichen



Abb. A.3: Beispiel zum Erscheinungsbild von Hand geschlagener Zeichen

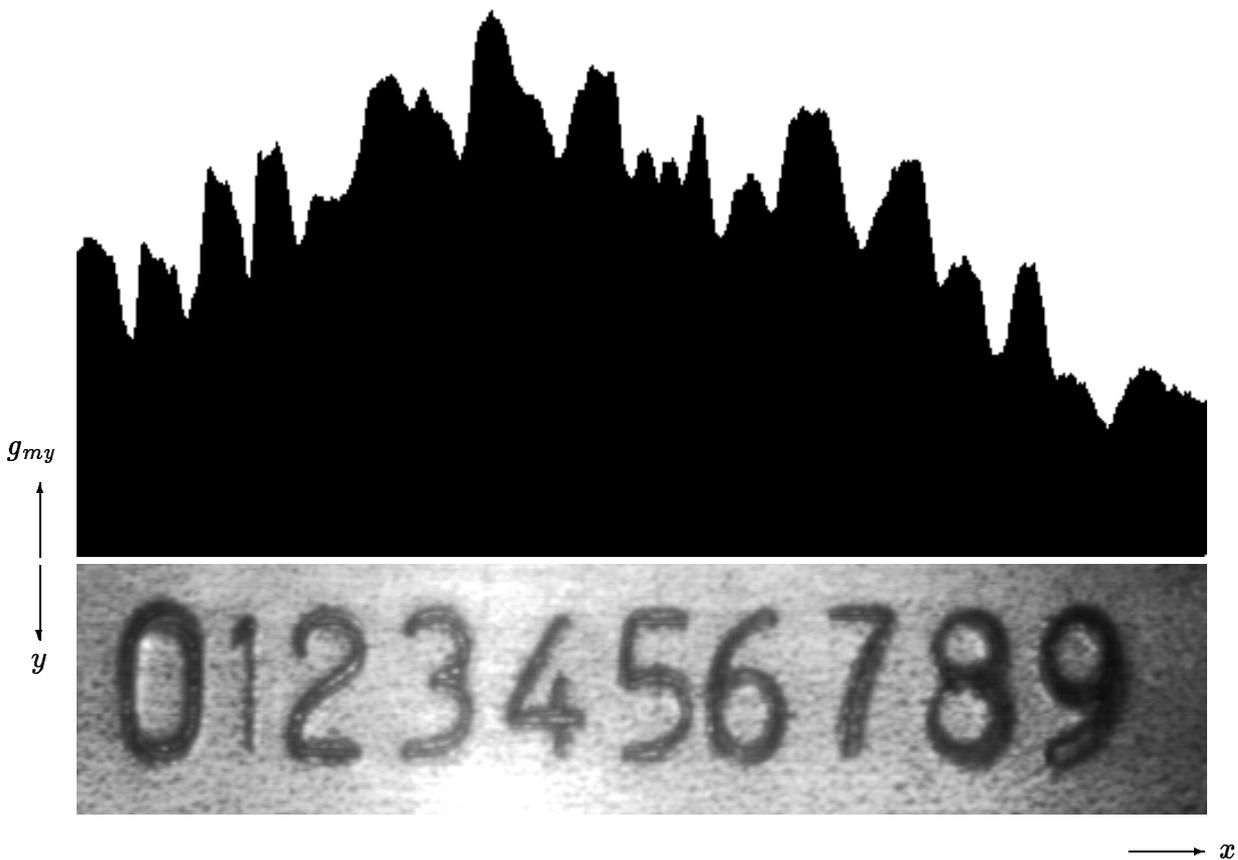


Abb. A.4: Verteilung des mittleren Grauwertes  $g_{my}$  der Bildspalten.

$$g_{my}(x) = \frac{1}{n_{sy}} \sum_{y=1}^{n_{sy}} g(x,y)$$

$$\begin{aligned}
\frac{1}{m} \sum_i (g_i - \bar{g})(r_i - \bar{r}) &= \frac{1}{m} \sum_i g_i r_i - \bar{g} \frac{1}{m} \sum_i r_i - \bar{r} \frac{1}{m} \sum_i g_i + \bar{g} \cdot \bar{r} \\
&= \overline{g \cdot r} - \bar{g} \cdot \bar{r} - \bar{r} \cdot \bar{g} + \bar{g} \cdot \bar{r} \\
&= \overline{g \cdot r} - \bar{g} \cdot \bar{r}
\end{aligned} \tag{A-1}$$

$$\begin{aligned}
\sigma_r^2 &= \frac{1}{m} \sum_i (r_i - \bar{r})^2 \\
&= \frac{1}{m} \sum_i r_i^2 - \bar{r}^2 \\
&= \overline{r^2} - \bar{r}^2
\end{aligned} \tag{A-2}$$

$$\begin{aligned}
m^2(\overline{g \cdot r} - \bar{g} \cdot \bar{r}) &= m \sum_{i \in M} g_i r_i - \sum_{i \in M} g_i \sum_{i \in M} r_i \\
&= m m_{\oplus} \frac{1}{m_{\oplus}} \sum_{i \in M_{\oplus}} g_i r_i + m m_{\ominus} \frac{1}{m_{\ominus}} \sum_{i \in M_{\ominus}} g_i r_i - \\
&\quad \left( m_{\oplus} \frac{1}{m_{\oplus}} \sum_{i \in M_{\oplus}} g_i + m_{\ominus} \frac{1}{m_{\ominus}} \sum_{i \in M_{\ominus}} g_i \right) \left( m_{\oplus} \frac{1}{m_{\oplus}} \sum_{i \in M_{\oplus}} r_i + m_{\ominus} \frac{1}{m_{\ominus}} \sum_{i \in M_{\ominus}} r_i \right) \\
&= m m_{\oplus} \overline{g_{\oplus} r_{\oplus}} + m m_{\ominus} \overline{g_{\ominus} r_{\ominus}} - \\
&\quad m_{\oplus} m_{\oplus} \overline{g_{\oplus}} \cdot \bar{r}_{\oplus} - m_{\ominus} m_{\oplus} \overline{g_{\ominus}} \cdot \bar{r}_{\oplus} - \\
&\quad m_{\oplus} m_{\ominus} \overline{g_{\oplus}} \cdot \bar{r}_{\ominus} - m_{\ominus} m_{\ominus} \overline{g_{\ominus}} \cdot \bar{r}_{\ominus} \\
&= m m_{\oplus} (\overline{g_{\oplus} r_{\oplus}} - \overline{g_{\oplus}} \cdot \bar{r}_{\oplus}) + m m_{\ominus} (\overline{g_{\ominus} r_{\ominus}} - \overline{g_{\ominus}} \cdot \bar{r}_{\ominus}) + \\
&\quad m_{\oplus} m_{\oplus} \overline{g_{\oplus}} \cdot \bar{r}_{\oplus} - m_{\ominus} m_{\oplus} \overline{g_{\ominus}} \cdot \bar{r}_{\oplus} - m_{\oplus} m_{\ominus} \overline{g_{\oplus}} \cdot \bar{r}_{\ominus} + m_{\ominus} m_{\oplus} \overline{g_{\ominus}} \cdot \bar{r}_{\ominus} \\
&= m m_{\oplus} (\overline{g_{\oplus} r_{\oplus}} - \overline{g_{\oplus}} \cdot \bar{r}_{\oplus}) + m m_{\ominus} (\overline{g_{\ominus} r_{\ominus}} - \overline{g_{\ominus}} \cdot \bar{r}_{\ominus}) + m_{\oplus} m_{\ominus} (\overline{g_{\oplus}} - \overline{g_{\ominus}}) (\bar{r}_{\oplus} - \bar{r}_{\ominus}) \\
\text{cov}(g, r) &= \frac{m_{\oplus}}{m} \text{cov}_{\oplus}(g, r) + \frac{m_{\ominus}}{m} \text{cov}_{\ominus}(g, r) + \frac{m_{\oplus}}{m} \frac{m_{\ominus}}{m} (\overline{g_{\oplus}} - \overline{g_{\ominus}}) (\bar{r}_{\oplus} - \bar{r}_{\ominus}) \\
&= \frac{m_{\oplus}}{m} \text{cov}_{\oplus}(g, r) + \frac{m_{\ominus}}{m} + \frac{m_{\oplus}}{m} \frac{m_{\ominus}}{m} \text{cov}_{\ominus}(g, r) \frac{m_{\oplus}}{m} (\overline{g_{\oplus}} - \overline{g_{\ominus}}) (\bar{r}_{\oplus} - \bar{r}_{\ominus})
\end{aligned} \tag{A-4}$$

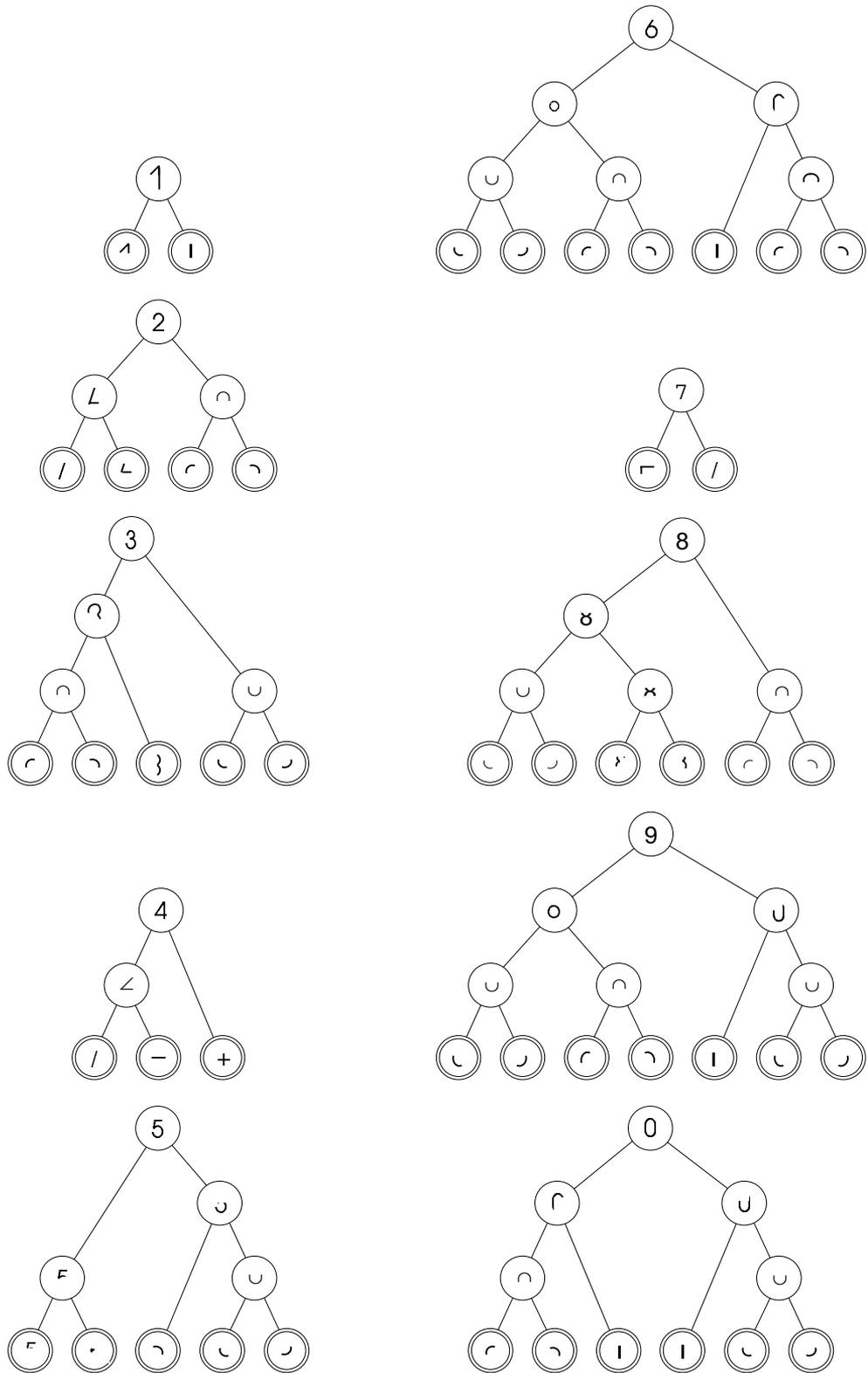
$$\overline{g \cdot r} - \bar{g} \cdot \bar{r} = \frac{m_{\oplus}}{m} \frac{m_{\ominus}}{m} (\overline{g_{\oplus}} - \overline{g_{\ominus}}) (\rho_{\oplus} - \rho_{\ominus}) \tag{A-5}$$

$$\overline{r \cdot r} - \bar{r} \cdot \bar{r} = \frac{m_{\oplus}}{m} \frac{m_{\ominus}}{m} (\rho_{r_{\oplus}} - \rho_{r_{\ominus}}) (\rho_{\oplus} - \rho_{\ominus}) \tag{A-6}$$

$$\overline{g \cdot g} - \bar{g} \cdot \bar{g} = \frac{m_{\oplus}}{m} \sigma_{\oplus}^2 + \frac{m_{\ominus}}{m} \sigma_{\ominus}^2 + \frac{m_{\oplus}}{m} \frac{m_{\ominus}}{m} (\overline{g_{\oplus}} - \overline{g_{\ominus}})^2 \tag{A-7}$$

$$\begin{aligned}
k_r &= \frac{\sum_i [g_i - \bar{g}] \cdot [r_i - \bar{r}]}{\sqrt{\sum_i [g_i - \bar{g}]^2 \cdot \sum_i [r_i - \bar{r}]^2}} \\
&= \frac{\frac{m_{\oplus}}{m} \frac{m_{\ominus}}{m} (\overline{g_{\oplus}} - \overline{g_{\ominus}}) (\rho_{\oplus} - \rho_{\ominus})}{\sqrt{\left( \frac{m_{\oplus}}{m} \sigma_{g_{\oplus}}^2 + \frac{m_{\ominus}}{m} \sigma_{g_{\ominus}}^2 + \frac{m_{\oplus}}{m} \frac{m_{\ominus}}{m} (\overline{g_{\oplus}} - \overline{g_{\ominus}})^2 \right) \cdot \frac{m_{\oplus}}{m} \frac{m_{\ominus}}{m} (\rho_{\oplus} - \rho_{\ominus})^2}}
\end{aligned} \tag{A-8}$$

$$\begin{aligned}
&= \frac{\rho_{\oplus} - \rho_{\ominus}}{|\rho_{\oplus} - \rho_{\ominus}|} \cdot \frac{\overline{g_{\oplus}} - \overline{g_{\ominus}}}{\sqrt{\left(\frac{m}{m_{\oplus}}\sigma_{g+}^2 + \frac{m}{m_{\ominus}}\sigma_{g-}^2 + (\overline{g_{\oplus}} - \overline{g_{\ominus}})^2\right)}} \\
&= \operatorname{sgn}(\rho_{\oplus} - \rho_{\ominus}) \cdot \operatorname{sgn}(\overline{g_{\oplus}} - \overline{g_{\ominus}}) \frac{1}{\sqrt{\frac{m}{m_{\oplus}}\left(\frac{\sigma_{\oplus}}{\overline{g_{\oplus}} - \overline{g_{\ominus}}}\right)^2 + \frac{m}{m_{\ominus}}\left(\frac{\sigma_{\ominus}}{\overline{g_{\oplus}} - \overline{g_{\ominus}}}\right)^2 + 1}}
\end{aligned} \tag{A-9}$$



**Abb. A.5:** Konzeptgraphen der Ziffern eins bis null

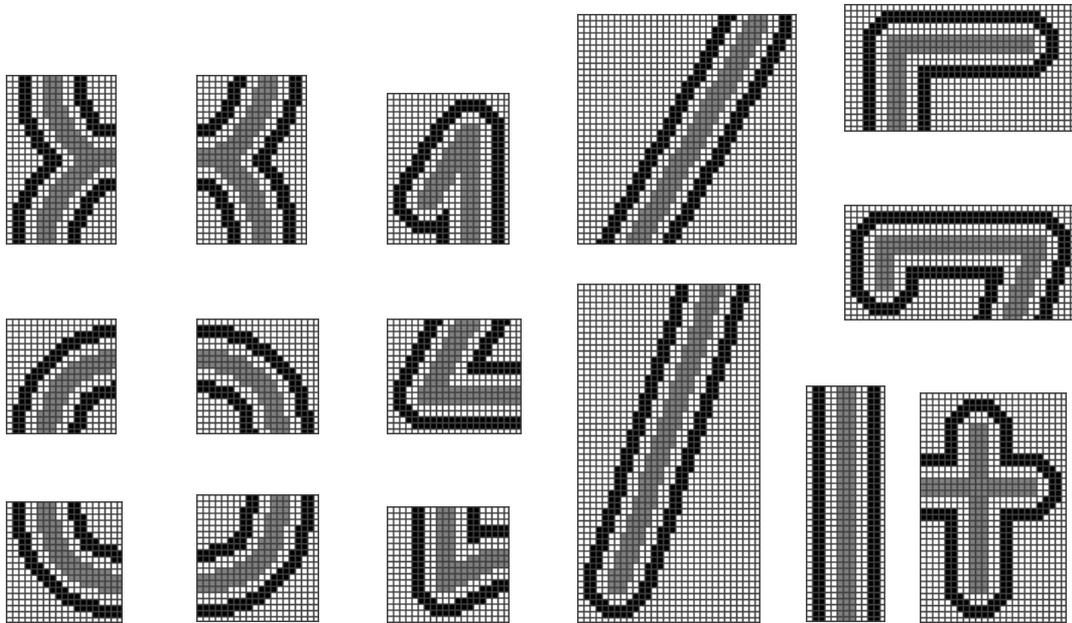


Abb. A.6: Verwendete Teilmasken

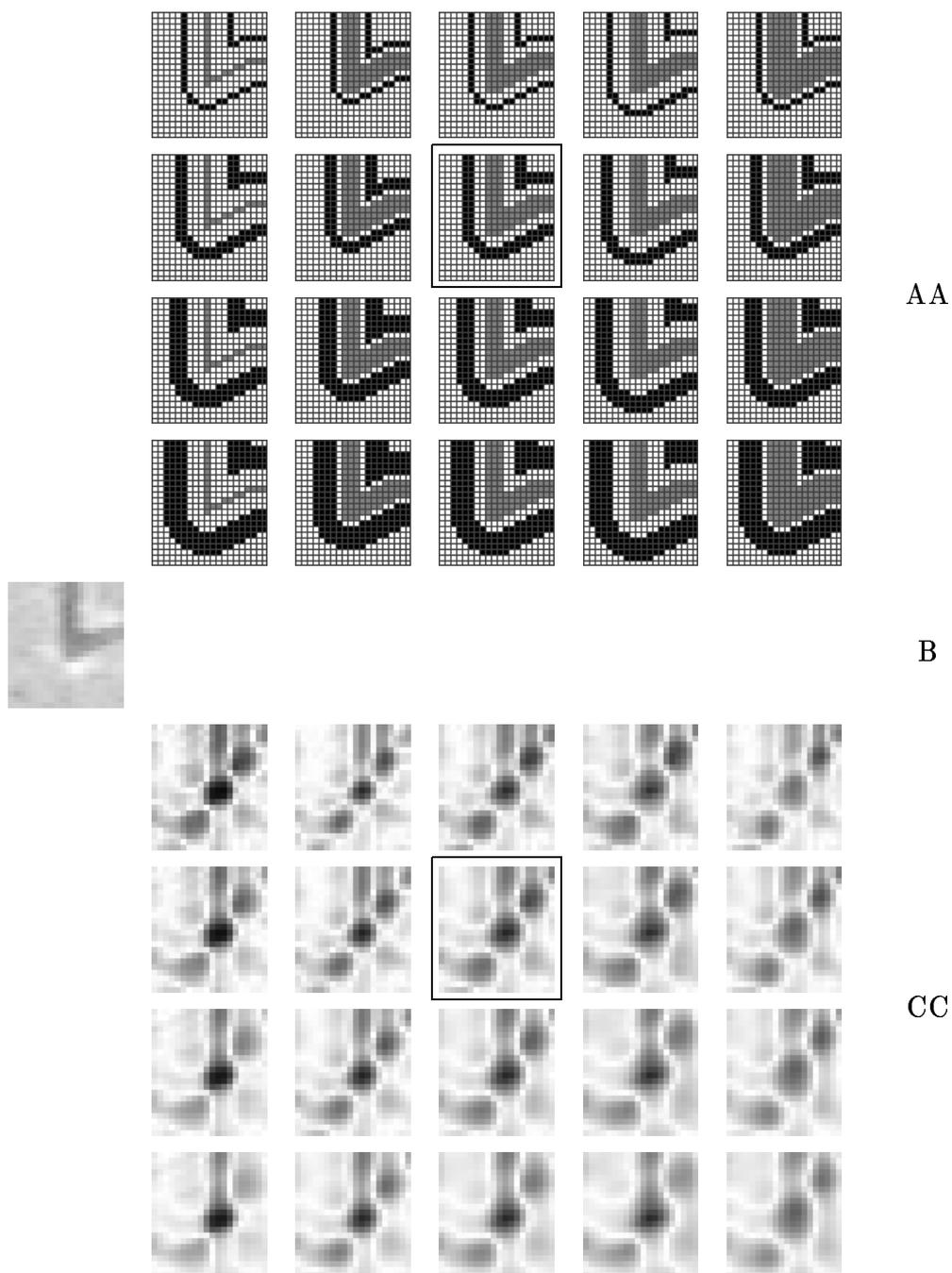


Abb. A.7: Variationen der Teilmaske ES5 mit resultierenden Korrelationsfeldern.

AA) Variation der Fläche  $A_{\oplus}$ ,  $A_{\circ}$ ,  $A_{\ominus}$  (siehe Kap. 4.1.1).

Die Teilmaske aus Abb. 4.26Ac ist eingerahmt,

B) Muster (Ausschnitt) (vgl. Abb. 4.26Ba),

CC) Korrelationsfelder entsprechend Block AA



## Lebenslauf

Uwe Stilla

geboren am	31. Januar 1957
in	Köln
Familienstand	ledig
1963 - 1967	Grundschule Köln und Mayen
1967 - 1973	Staatl. Realschule Mayen - Mittlere Reife
1973 - 1975	Fachoberschule Andernach - Fachhochschulreife
1976 - 1980	Studium der Nachrichtentechnik an der Gesamthochschule Paderborn / Abt. Meschede, Diplom - allgemeine Hochschulreife
1980 - 1987	Studium der Elektrotechnik an der Universität Karlsruhe - Diplom
1987 - 1990	Wissenschaftlicher Mitarbeiter am Institut für Biokybernetik und Biomedizinische Technik (Prof. Dr.-Ing. K. Meyer-Waarden)
seit WS 1989	Lehrbeauftragter für die Vorlesung Biokybernetik des Studiengangs Medizinische Informatik der Universität Heidelberg / Fachhochschule Heilbronn
seit 1.1.1991	Wissenschaftlicher Angestellter bei Prof. Dr.-Ing. H. Kazmierczak am Forschungsinstitut für Informationsverarbeitung und Mustererkennung
22.06.1993	Promotionsprüfung