

# Merkmalsextraktion für eine automatische Bildsuche

## Feature Extraction for Automated Image Retrieval

Ullrich Köthe und Hans Meine

AB Kognitive Systeme, Universität Hamburg

Vogt-Kölln-Str. 30, 22527 Hamburg

Tel.: +49-40-42883-2450, Fax: +49-40-42883-2572

E-mail: {koethe,meine}@informatik.uni-hamburg.de, Internet: kogs.informatik.uni-hamburg.de

### Zusammenfassung:

Für viele Anwendungen der Bildsuche werden hochgenaue lokale Bildmerkmale benötigt. Wir beschreiben, wie man durch hinreichende Bildauflösung, Interpolation, und subpixel-genaue Kantendetektion die Messgenauigkeit für geometrische und topologische Merkmale optimieren kann.

### Abstract:

Many applications of content-based image retrieval require very accurate local image features. We describe how the measurement accuracy of geometrical and topological features can be optimized by means of appropriate image resolution, interpolation, and subpixel-accurate edge detection.

### Einleitung

Die existierenden Verfahren zur inhaltsorientierten Suche in Bilddatenbanken basieren typischerweise auf globalen Merkmalen, wie z.B. Farbhistogrammen, Texturdeskriptoren oder Richtungshistogrammen. Auf diese Weise gelingt es, Bilder aus einer Datenbank zu filtern, deren Farbcharakter, Layout oder Hauptorientierungen gut mit einem gegebenen Beispielbild übereinstimmen.

Häufig ist man aber an anderen, subtileren Ähnlichkeitsmerkmalen interessiert, die sich durch globale Beschreibungen nicht erfassen lassen. Gerade in der Kunstgeschichte und Denkmalpflege spielen solche eher lokalen Merkmale eine große Rolle. Voraussetzung für eine erfolgreiche Suche ist in diesem Fall, dass man zunächst einmal in der Lage ist, die geeigneten Merkmale überhaupt zu detektieren. Selbst dann reicht die Genauigkeit traditioneller Detektoren häufig nicht aus, um die entscheidenden Unterschiede festzustellen, obwohl sie für einen Menschen deutlich sichtbar sind.

In diesem Beitrag wollen wir Wege beschreiben, wie man die in einem Bild enthaltene Information tatsächlich extrahieren kann. Es hat sich herausgestellt, dass herkömmliche Verfahren mit den Bilddaten häufig zu sorglos umgehen und dadurch Informationen verloren gehen, die eigentlich vorhanden sind. Ein typischer Fehler ist zum Beispiel eine zu geringe Auflösung während der Bildanalyse: überraschenderweise ist es oft notwendig, die Auflösung gegenüber dem Originalbild zu *erhöhen*, um Informationsverlust zu vermeiden. Traditionell versucht man dagegen, die Auflösung schnell zu *verringern*, um die Rechenzeit zu minimieren. Ein anderes Problem entsteht durch Merkmalsmodelle und -repräsentationen ungenügender Mächtigkeit: Wenn das Bild Konfigurationen enthält, die für die gewählten Modelle zu komplex sind, können die resultierenden Beschreibungen im schlimmsten Falle ihre Aussagekraft vollkommen verlieren.

Ein interessantes Anwendungsgebiet, aus dem wir im Folgenden einige Beispiele bringen wollen, ist die Analyse historischer chinesischer Handschriften [6]. Diese sind durch neue archäologische Funde erst in den letzten Jahren ins Blickfeld des Interesses gerückt, nachdem zuvor im Wesentlichen gedruckte Dokumente bekannt waren (bekanntlich wurde der Buchdruck in China deutlich früher eingeführt als in Europa). Daher existiert in China keine wissenschaftliche Tradition der Handschriftenanalyse, und es bestehen Bedarf und Gelegenheit, europäisches Know-how zur Lösung der neuen Probleme einzubringen.

Ein wichtiges Feld der Analyse chinesischer Handschriften ist die Zuordnung einzelner Textfragmente zu individuellen Schreibern. Die frühen Manuskripte wurden oft auf Bambusleisten geschrieben, und die aufeinander folgenden Leisten (Spalten) eines Dokuments wurden mit Fäden

verbunden (Abb. 1). Da diese Fäden inzwischen verrottet sind, ist unklar, welche Leisten zu einem Dokument gehören, und in welcher Reihenfolge sie angeordnet werden müssen. Bisher wurde dies nach dem Text der Fragmente entschieden, aber es besteht die Gefahr, dass man dabei mit den Augen einer modernen Textdeutung den ursprünglichen Sinn verfälscht: Man prägt der Rekonstruktion eine bereits feststehende Deutung auf, anstatt umgekehrt den historischen Sinn aus dem Dokument zu erschließen. Diese Gefahr ist im chinesischen besonders groß, weil viele Schriftzeichen mehrdeutig sind, und die Bedeutungen sich im Laufe der Zeit gewandelt haben.

Daher ist es wünschenswert, die Textrekonstruktion anhand objektiver Merkmale vorzunehmen. Eine sinnvolle Hypothese besteht darin, Fragmente desselben Schreibers oder derselben Schule mit höherer Wahrscheinlichkeit zum selben Dokument zu gruppieren. Eigenarten eines Schreibers kann man an Strichstärken, Strichwinkeln und dergleichen festmachen. Allerdings sind solche Merkmale nur dann aussagekräftig, wenn sie mit hinreichender Genauigkeit gemessen werden. Einige hochgenaue Messverfahren werden im Folgenden vorgestellt.



Abb. 1 Chinesisches Manuskript auf Bambus

## Bildauflösung im Kontext der Objekterkennung

Offensichtlich können die interessierenden Objekte im Bild nur dann korrekt detektiert werden, wenn das Bild eine ausreichende Auflösung hat. Aber was heißt das genau? Eine Aussage dazu liefert unser geometrisches Abtasttheorem. Wir setzen voraus, dass man im Prinzip ein ideales Bild mit unendlicher Auflösung aufnehmen könnte, in dem alle Objektregionen die korrekte Topologie haben. Das heißt, die korrekte Anzahl von Objekten wurde gefunden, die Nachbarschaft zwischen Objekten ist korrekt, eventuelle „Löcher“ in Objekten sind erhalten geblieben usw. Das tatsächlich aufgenommene digitale Bild unterscheidet sich erheblich von diesem idealen Bild: durch das optische System entsteht eine unvermeidliche Unschärfe, die Auflösung ist durch die Pixel begrenzt, die Farbwerte werden durch Rauschen verfälscht und überdies quantisiert. Dennoch kann man unter bestimmten Bedingungen garantieren, dass die korrekte Objekttopologie auch aus dem digitalen Bild rekonstruiert werden kann [3].

Wir nehmen an, dass die *Kontur* der rekonstruierten Objekte durch eine Menge von Punkten gegeben ist, welche durch Linien verbunden werden. Die Punkte können z.B. Pixel sein, die als Kantenpunkte markiert wurden, oder Edgel, die von einem subpixel-genauen Kantendetektor wie dem Canny-Algorithmus stammen. Die Genauigkeit der Kantenpunkte wird durch zwei Größen  $p$  und  $q$  charakterisiert:  $p$  gibt an, wie weit ein (korrekter) Kantenpunkt im idealen Bild maximal vom nächsten detektierten Kantenpunkt entfernt ist, und  $q$  gibt an, wie weit ein detektiertes Kantenpunkt maximal vom nächstgelegenen tatsächlichen Kantenpunkt entfernt ist (Abb. 2 links).

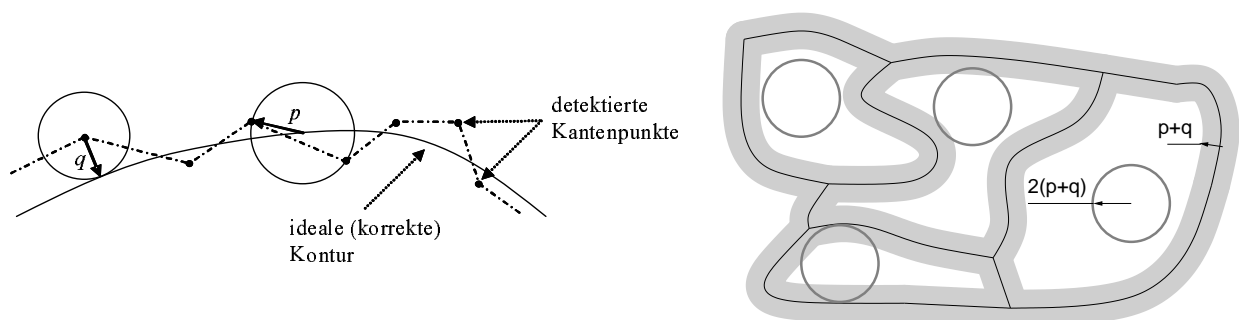


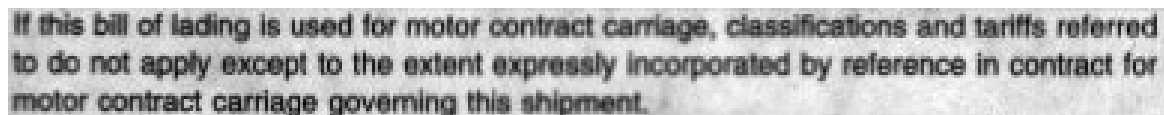
Abb. 2 Links: Bewertung der Genauigkeit einer Segmentierung; Rechts: Bedingungen für den Erhalt der Topologie bei der Segmentierung mit Genauigkeitswerten  $p$  und  $q$ .

Aus den Genauigkeitswerten  $p$  und  $q$  ergeben sich zwei Bedingungen, wie groß die im Bild enthaltenen Objekte mindestens sein müssen, damit die Messfehler sich nicht dahingehend auswirken, dass Objekte fehlerhaft detektiert werden (vgl. Abb. 2 rechts): Jede Objektregion muss mindestens einen Kreis mit dem Radius  $2(p+q)$  enthalten, und es muss möglich sein, die ideale Kontur um  $p+q$  Einheiten zu verschieben, ohne dass sich die Objekttopologie ändert (mathematisch exakt muss es möglich sein, jede Region um diesen Betrag zu erodieren, ohne dass sich die Topologie ändert). Wenn  $p$  und  $q$  beispielsweise den Wert 1 Pixel haben, muss jede Region mindestens einen Kreis mit dem Radius 4 enthalten, und die idealen Konturen müssen einen Spielraum von 2 Pixeln haben, wo sich keine andere Kontur befindet.

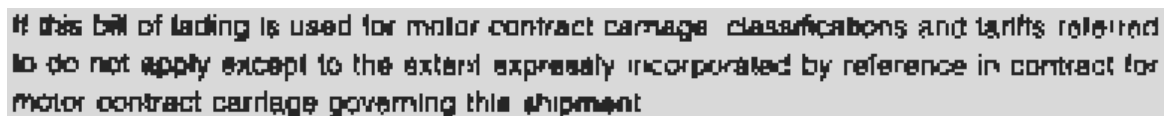
Es ist offensichtlich, dass diese Bedingungen für gegebenes  $p$  und  $q$  immer erfüllbar sind, wenn man die Pixelauflösung des Bildes hinreichend groß wählt. Andererseits will man in der Praxis mit einer möglichst geringen Auflösung auskommen, um Speicher und Rechenzeit zu sparen. Bei einem gegebenen Bild ist man daher interessiert, ein Segmentierungsverfahren mit möglichst kleinem Fehler zu verwenden. Die diesbezüglichen Unterschiede zwischen verschiedenen Segmentierungsverfahren sind in der Tat beträchtlich, aber ein Phänomen scheint allen Verfahren gemeinsam zu sein: Die optimale Genauigkeit ist nicht erreichbar, wenn man sich bei der Bildanalyse auf die Auflösung des Originalbildes beschränkt [2].

Wir wollen dieses Phänomen zunächst am Beispiel der Schwellwertbildung erläutern. Man betrachte das Originalbild in Abb. 3. Die Schrift ist sehr klein, aber für einen Menschen immer noch einwandfrei lesbar (zugunsten der Lesbarkeit verwenden wir hier lateinische Schrift, aber das Phänomen existiert analog für chinesische Schriftzeichen). Nach der Binarisierung ist der Text hingegen kaum noch erkennbar. Interpoliert man stattdessen das Originalbild mit Hilfe eines Splinerverfahrens auf die 3-fache Größe, ist die Qualität des daraus binarisierten Bildes deutlich besser. Dies erscheint zunächst überraschend, denn durch Interpolation kann man ja keine neue Information gewinnen.

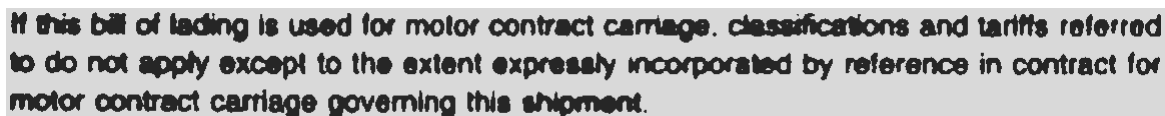
Bei genauerer Überlegung wird klar, dass dies auch nicht der Fall ist. Vielmehr wird durch die Interpolation verhindert, dass während der Analyse Information verloren geht, die im Originalbild bereits vorhanden war. Dies kann man sich sehr einfach verdeutlichen: Das ideale Bild eines gedruckten Textes hat unendliche Auflösung, aber nur zwei Farbwerte (schwarz und weiß). Das digitale Bild dagegen weist eine unvermeidliche Unschärfe auf, so dass beliebige Grauwerte zwischen schwarz und weiß vorkommen können. Dies wird oft als unerwünschter Effekt gesehen, aber bei begrenzter Auflösung ist es in Wirklichkeit eine sehr vorteilhafte Bildeigenschaft: Je heller bzw. dunkler ein Pixel ist, desto weiter muss es von der ursprünglichen Kontur entfernt gewesen sein, da in seiner Umgebung offenbar nur gleichfarbige Punkte existieren. Folglich kodieren die Grauwerte des digitalen Bildes geometrische Informationen über den Kantenerlauf. Daher ist es möglich, die Objektkanten mit einer höheren Genauigkeit als der reinen Pixelauflösung zu rekonstruieren. Binarisiert man nur bei der Originalauflösung, geht die in den Grauwerten kodierte geometrische Information verloren. Durch vorherige Interpolation wird diese Information auf mehrere Pixel verteilt und bleibt bei der Binarisierung erhalten: Wenn das digitale Originalbild zum Beispiel



If this bill of lading is used for motor contract carriage, classifications and tariffs referred to do not apply except to the extent expressly incorporated by reference in contract for motor contract carriage governing this shipment.

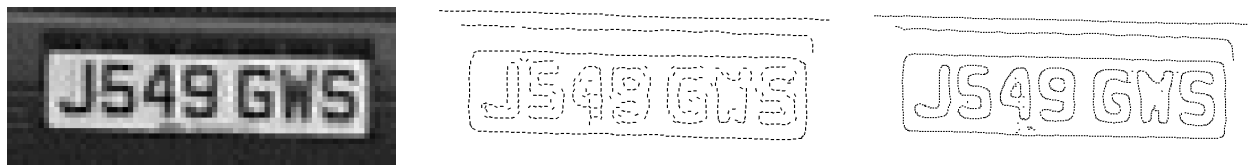


If this bill of lading is used for motor contract carriage classifications and tariffs referred to do not apply except to the extent expressly incorporated by reference in contract for motor contract carriage governing this shipment



If this bill of lading is used for motor contract carriage. classifications and tariffs referred to do not apply except to the extent expressly incorporated by reference in contract for motor contract carriage governing this shipment.

Abb. 3 Originaltext (oben) und Binarisierungen mit der ursprünglichen Auflösung (Mitte) und nach Interpolation auf die dreifache Größe (unten).



**Abb. 4** Kantendetektion bei der Originalauflösung (Mitte) und nach zweifacher Interpolation (rechts).

8 Bit pro Pixel enthält, so wird bei dreifacher Überabtastung jeder Originalpixel durch 9 neue Pixel ersetzt. Nach der Binarisierung bleiben also 9 Bit und damit in etwa soviel Information wie im Grauwertoriginal (Abb. 3 unten). Eine noch höhere Genauigkeit kann man erreichen, wenn man die Konturen mit Hilfe von Polygonen anstelle von Pixeln kodiert (Subpixel-Kanten, siehe unten).

Ein analoger Effekt kann bei der Kantendetektion beobachtet werden (Abb. 4). Die meisten Kantendetektoren berechnen zunächst einen so genannten Kantenindikator, der für jedes Pixel eine Kantenstärke enthält, also eine Zahl, die umso höher ist, je wahrscheinlicher der betreffende Pixel zu einer Kante gehört. Oft wird das Bild zunächst mit einer Reihe von Kantenfiltern gefiltert (z.B. Sobelfiltern, Gaußschen Gradienten, Gaborfiltern), und die Kantenstärke wird aus der Summe der Quadrate der einzelnen Filterantworten berechnet. In diesem Falle kann man die Notwendigkeit der Interpolation aus dem Shannonschen Abtasttheorem begründen. Wir setzen voraus, dass das Originalbild den Bedingungen des Abtasttheorems genügt, also der Pixelabstand kleiner ist als die halbe Wellenlänge der höchsten im Bild vorkommenden Frequenz (Nyquistfrequenz). In der Praxis heißt das, dass das optische System das ideale Bild vor der Abtastung hinreichend glättet (dies ist ein weiteres Argument dafür, dass eine gewisse optische Unschärfe kein unerwünschter Effekt ist). Nach einer linearen Kantenfilterung sind die Bedingungen des Abtasttheorems immer noch erfüllt, aber durch Quadrieren der Filterantworten verdoppelt sich die Nyquistfrequenz [2]. Das Kantenindikatorbild muss deshalb doppelt so dicht abgetastet werden wie das Originalbild, um Aliasing und Informationsverlust zu vermeiden. Auch hier ist es also notwendig, das Originalbild zunächst zu interpolieren.

## Kantendetektion mit Pixelgenauigkeit und Subpixelgenauigkeit

Neben den allgemeinen Überlegungen zur Bildauflösung wird die erreichbare Genauigkeit entscheidend durch die Eigenschaften des Kantendetektors bestimmt. Die Unterschiede der Werte von  $p$  und  $q$  zwischen einfachen und fortgeschrittenen Detektoren können einen Faktor 5 und mehr ausmachen. Betrachten wir zunächst die Kantendetektion mit Pixelgenauigkeit. Wir können hier zwei grundlegende Vorgehensweisen unterscheiden: (i) Man markiert Pixel als zugehörig zu verschiedenen Regionen, z.B. mit Hilfe einer Schwellwertbildung oder eines nullstellen-basierten Kantendetektors („zero-crossing detector“). Die Kante verläuft dann zwischen den Regionen, also zwischen denjenigen Pixeln, die unterschiedlichen Regionen angehören. Daher heißen die resultierenden Kanten *inter-pixel Kanten* oder *crack-edges*. (ii) Alternativ markiert man (z.B. beim Canny-Kantendetektor [1]) gewisse Pixel als Kantenpixel, und die Regionen werden von allen nicht markierten Pixeln gebildet. Man bezeichnet diese Kanten als *Pixelkanten*.

In beiden Fällen ist die mögliche Kantengenauigkeit durch das Pixelraster beschränkt. Abgesehen von anderen Störeinflüssen wie Rauschen und Quantisierung entsteht immer ein zusätzlicher Fehler durch die Rundung der Kantenposition zur nächsten Pixelkoordinate (bzw. Interpixelkoordinate). Aus der Rundung ergibt sich, dass die die Größe  $p$  (der maximale Abstand der exakten Kontur zum nächstgelegenen detektierten Konturpunkt) einen Wert von  $\sqrt{2}$  nicht unterschreiten kann. Die Größe  $q$  (der maximale Abstand der detektierten Punkte zur exakten Kontur) beträgt mindestens  $\sqrt{2}/2$  (Interpixel-Kanten) bzw.  $1/2$  (Pixelkanten) [3]. Man kann sich die Einschränkungen einer pixel-basierten Kantenrepräsentation auch dadurch plausibel machen, dass zwischen zwei Kanten stets mindestens ein Regionenpixel liegen muss. Durch diesen „Mindestabstand“ ist die Genauigkeit einer pixelbasierten Kante von vornherein begrenzt (vgl. Abb. 5 links unten).



**Abb. 5** Oben: Segmentierung verschiedener chinesischer Zeichen. Unten: Vergleich zwischen pixelgenauer und subpixelgenauer Segmentierung. Der Ausschnitt zeigt die Überlegenheit der letzteren.

Um die in einem Grauwert- oder Farbbild enthaltene geometrische Information vollständig zu extrahieren, werden Kantendetektoren mit Subpixelgenauigkeit benötigt. Auch hierbei gibt es zwei grundlegende Vorgehensweisen: (i) Man kann die Kanten zunächst mit Pixelgenauigkeit bestimmen und ihre Position danach verbessern. Dies ist der Ansatz des Canny-Verfahrens [1]. (ii) Man kann die pixelgenauen Kantendetektoren direkt zu subpixel-genauen Verfahren verallgemeinern. Dies ist sowohl bei nullstellen-basierten Methoden möglich, wie auch bei solchen auf der Basis von Kantenindikatorbildern.

Im Falle des Canny-Algorithmus detektiert man zunächst Kantenpixel, welche dadurch definiert sind, dass die beiden benachbarten Pixel in Gradientenrichtung einen niedrigeren Gradienten (=Kantenstärke) haben als der Kantenpixel. Eine subpixel-genaue Kantenposition wird dann üblicherweise durch das Anpassen einer Parabel an die Gradientenwerte dieser 3 Pixel bzw. an die 9 Pixel in der 3x3 Umgebung gewonnen: Der genaue Kantenpunkt liegt auf der Geraden, die durch die Gradientenrichtung festgelegt wird, an der Position, wo sich der Scheitel der Parabel befindet (vgl. Abb. 4). Für gerade Kanten erreicht dieses Verfahren im günstigsten Fall Werte von  $q = 0.3$  Pixel. Da jedoch pro Pixel höchstens ein Kantenpunkt detektiert wird, kann  $p$  den Wert  $\sqrt{2}/2$  nicht unterschreiten.

Eine noch höhere Genauigkeit kann man durch echte Subpixelverfahren erreichen [4,5]. Wir erläutern zunächst den subpixel Nullstellendetektor. Gegeben sei ein Bild, in dem Kanten durch Nulldurchgänge definiert sind. Man beachte, dass damit auch die Schwellwertbildung erfasst ist, weil man jedes Schwellwertproblem durch Subtrahieren des Schwellwertes von den Pixelwerten in ein Nullstellenproblem umformen kann. Über diesem diskreten Bild wird mit Hilfe der Splineinterpolation eine kontinuierliche Funktion definiert, die an den Pixelpositionen gerade die ursprünglichen Werte reproduziert, aber auch an allen anderen reellwertigen Koordinaten sinnvolle Werte liefert. Wir suchen nun die exakten Nullstellen dieser Funktion entlang der Gitterlinien. Da Splines stückweise polynomielle Funktionen sind, ist dies analytisch möglich. Ausgehend von den so gefundenen Nullstellen wird die gesamte Kontur mittels eines *Prediktor-Korrektor-Verfahrens* vervollständigt: Man berechnet zunächst im letzten bekannten Konturpunkt die Tangente, geht ein kleines Stück (z.B. 1/10 Pixel) in dieser Richtung weiter (Prediktorschritt), und sucht dann senkrecht zur Prediktorrichtung wieder die exakte Nullstellenposition (Korrektorschritt). Die Suche terminiert, wenn sie auf ein bereits bekanntes Konturstück stößt, vgl. Abb. 5 oben rechts und Abb. 6 Mitte.

Ähnlich geht man vor, um aus einem Kantenindikatorbild subpixel-genaue Kanten zu gewinnen [7,5]. Auch hier wird das Kantenindikatorbild zunächst zu einer kontinuierlichen Funktion interpoliert, die man an beliebigen reellwertigen Koordinaten abfragen kann. Auf dieser Funktion sucht man nun sämtliche Sattelpunkte, deren Kantenstärke einen gewissen Mindestwert überschreitet. Von dort aus verfolgt man die Kante in derjenigen Richtung weiter, in der die Kantenstärke am schnellsten anwächst, bis man an einem Maximum der Kantenstärke anlangt, wo die Kantenverfolgung terminiert. Man kann mathematisch beweisen, dass die gefundenen Kanten gerade die exakten, reellwertigen Wasserscheiden sind (im Unterschied zu den auf ganzzahligen Koordinaten definierten Wasserscheiden, die als pixel-genaue Kantendetektor verbreitet sind).

Im Falle der chinesischen Schriftzeichen liefern beide subpixel-genaue Detektoren sehr ähnliche Ergebnisse. Sie verbessern  $q$  auf Werte bis zu 0.15 Pixel. Wichtiger ist jedoch, dass die Kantenpunkte bei diesen Verfahren beliebig dicht gesetzt werden können, so dass sich  $p$  auf Werte um 0.2 Pixel verbessern lässt. Dies ist besonders für eine korrekte Wiedergabe gekrümmter Konturen wichtig. Man beachte, dass diese Angaben sämtlich Maximalfehler sind. Die mittleren Fehler sind deutlich niedriger. Mit diesen verfeinerten Fehlerschranken kann man ein Bild noch korrekt segmentieren, wo dies mit pixel-genaue Verfahren nicht mehr möglich ist. Bei vergleichbarer Endgenauigkeit kommt man deshalb mit einer geringeren Auflösung des Originalbildes aus.

## Merkmalsgewinnung

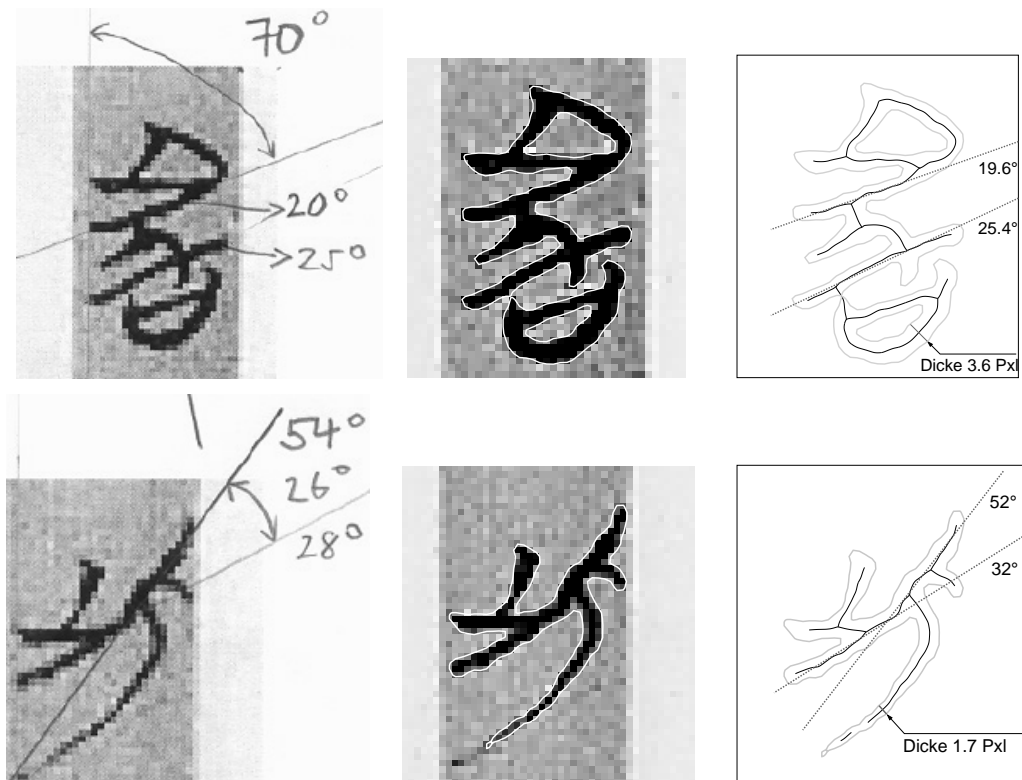
Eine hochgenaue Segmentierung bietet nun diverse Möglichkeiten, interessante Merkmale zu berechnen. Abb. 6 links zeigt Beispiele für das bisherige manuelle Vorgehen [6]. In diesem Fall sind insbesondere die Winkel der einzelnen Striche von Bedeutung, aber auch die Länge und Dicke der Striche sowie die Reihenfolge der Strichsetzung spielen eine wichtige Rolle. Wir erforschen zur Zeit, inwieweit diese Merkmale automatisch oder zumindest halbautomatisch detektiert werden können. Abb. 6 Mitte zeigt zunächst, dass die Segmentierung mit einem subpixel-genaue Schwellwertverfahren die Gewähr bietet, dass die für diese Analyse wichtigen Merkmale nicht verloren gehen. Insbesondere werden die einzelnen Teile eines Zeichens und eventuell eingeschlossene Flächen korrekt detektiert. Man beachte auch den Vergleich zwischen pixel-genaue und subpixel-genaue Segmentierung in dem Bildausschnitt Abb. 5 unten rechts – die Details der Pinselsetzung und die Strichdicke sind nur bei der subpixel-genaue Kante sichtbar.

Neben der Kontur eines Zeichens kann man auch aus der Zentralachse („Skelettlinie“) viele nützliche Merkmale ableiten, siehe Abb. 6 rechts. Es gibt viele Algorithmen zur Bestimmung der Zentralachse, aber im vorliegenden Kontext eignet sich der oben erwähnte subpixel-genaue Wasserscheidenalgorithmus besonders gut. Anstatt ihn, wie sonst üblich, auf ein Kantenindikatorbild anzuwenden, wendet wir ihn hier auf das Originalbild an. Dies ist möglich, weil der Algorithmus letztendlich nur nach Linien maximaler oder minimaler Intensität („ridges“ bzw. „valleys“) sucht, seien dies nun Orte hoher Kantenstärke, oder die Achse eines dunklen Pinselstrichs. Durch Anpassen einer Gerade an Teile einer Zentralachse lassen sich die Strichwinkel bestimmen, und unsere Messungen stimmen sehr gut mit den Ergebnissen der manuellen Analyse überein. Zeichnet man eine Linie senkrecht zur Zentralachse, kann man aus ihren Schnittpunkten mit der Kontur des Zeichens die lokale Strichbreite ermitteln.

Im Moment entwickeln wir ein interaktives System zur Bestimmung der Merkmale von chinesischen Zeichen. Es ist bereits abzusehen, dass dieses Vorgehen gegenüber der manuellen Analyse schneller ist und besser reproduzierbare Ergebnissen liefert. Die gewonnenen Merkmale sollen in eine Datenbank eingespeist werden, die die Suche nach Gemeinsamkeiten und Unterschieden zwischen Textfragmenten unterstützt. Perspektivisch sollen die interaktiven (also halbautomatischen Verfahren), soweit sinnvoll, durch vollautomatische ersetzt werden.

## Schlussfolgerungen

Wir haben an verschiedenen Beispielen gezeigt, wie man die Genauigkeit von Segmentierungsverfahren durch geeignete Wahl der Bildauflösung, Interpolation, und subpixel-genaue Kantendetek-



**Abb. 6** Links: Manuelle Merkmalsextraktion aus [6]; Mitte: Segmentierung durch subpixel-genaue Schwellwertkante; Rechts: Kontur mit Strichdicken, Zentralachse mit angepassten Geraden und deren Winkeln. Man beachte die hohe Genauigkeit, obwohl die Strichdicke unter 4 Pixeln liegt.

tion optimieren kann. Dadurch kann man unter bestimmten Bedingungen garantieren, dass wichtige Eigenschaften eines Zeichens, wie z.B. die Anzahl seiner Teile und die Anzahl der eingeschlossenen Flächen, erhalten bleiben. In einem zunächst noch halbautomatischen System kann man weitere Merkmale wie Strichrichtung und Strichdicke sehr genau bestimmen. Erst durch ihre hohe Genauigkeit werden diese Merkmale für eine inhaltsbasierte Suche interessant.

## Referenzen

- [1] J. Canny: "A Computational Approach to Edge Detection", IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6):679-698, 1986
- [2] U. Köthe: "Edge and Junction Detection with an Improved Structure Tensor", in: B. Michaelis, G. Krell (Eds.): Pattern Recognition, DAGM '03, Springer LNCS 2781, pp. 25-32, 2003.
- [3] U. Köthe, P. Stelldinger, H. Meine: "Provably Correct Edgel Linking and Subpixel Boundary Reconstruction", in: Pattern Recognition, DAGM '06, Springer LNCS 4174, pp. 81-90, 2006.
- [4] H. Meine, U. Köthe: "Image Segmentation with the Exact Watershed Transform", in: J.J. Villanueva (Ed.): VIIP 2005, Proc. 5<sup>th</sup> IASTED International Conference on Visualization, Imaging, and Image Processing, pp. 400-405, ACTA Press, 2005.
- [5] H. Meine, U. Köthe: "A New Sub-Pixel Map for Image Analysis", in: Proc. Intl. Workshop on Combinatorial Image Analysis, IWCIA '06, Springer LNCS 4040, 2006.
- [6] M. Richter: "Tentative Criteria for Discerning Individual Hands in the Guodian Manuscripts", Confucianism Resurrected: 3<sup>rd</sup> Intl. Conf. on Excavated Chinese Manuscripts, 2004.
- [7] C. Steger: "Subpixel-Precise Extraction of Watersheds", in: ICCV '99, Proc. 7<sup>th</sup> Intl. Conf. Computer Vision, vol. II, pp. 884-890, 1999.