Dr. Lothar Wenzel

Digitale Signalverarbeitung ist keine Hexerei

Teil 11: Zweidimensionale Signal- bzw. Bildverarbeitung

Die Bildverarbeitung spielt im Gesamtgebäude der digitalen Signalverarbeitung eine etwas merkwürdige Rolle. Zwar stützt man sich auf den wohlfundierten eindimensionalen Fall, und viele der dort gefundenen Resultate sind im Rahmen der Verallgemeinerung nicht ohne Reiz, jedoch begegnet man in der Bildverarbeitung auf Schritt und Tritt Heuristiken, die manchmal mehr an Kochbücher als an Wissenschaft erinnern. Als Fazit bleibt: Digitale Bildverarbeitung ist zugleich mehr und weniger als zweidimensionale Signaltheorie!

Alles fängt schon beim Begriff des Bildes an. Während man beispielhaft ohne Mühe und vollkommen korrekt stetige eindimensionale Signale definieren kann, fehlt eine Definition für ein Bild vollständig. Fest steht, daß Objekte, die als Bilder bezeichnet werden können, Struktur enthalten müssen. Da Struktur und Information eng zusammenhängen, hat man versucht, über Ansätze wie Entropie dem Begriff „Bild" näherzukommen. Eine zufriedenstellende Theorie steht noch aus, obwohl viele meinen, daß diese durchaus in Reichweite ist. So gibt es unter Umständen ein Superbild, das alle anderen in einer bestimmten Weise codiert. Der Leser möge bei der nächsten Party die Frage nach dem Wesen von Bildern zur Diskussion stellen. Je nach Blickwinkel wird er andere Antworten erhalten, die viel über die Komplexität dieser Fragestellung aussagen.

Wenn aber die allerelementarste Grundlage einer Theorie nicht als abgesichert gelten kann, so dürfte diese Verschwommenheit auch das gesamte darüberstehende Gebäude auszeichnen. In der Tat ist das so und hat insbesondere die Konsequenz, daß man den bereits genannten kochbuchartigen Stil bei vielen Anwendungen und Algorithmen der Bildverarbeitung vorfindet. Falls eine Implementierung mehr als 95 % aller denkbaren Fälle abdeckt, kann man schon zufrieden sein. Immer, wirklich immer, kann man einen konstruierten Algorithmus durch ein passendes Gegenbeispiel zu Fall bringen. Die Unschärfe in der Definition des Bildes schlägt hier einfach bei Bedarf durch.

Es gibt weitere Unterschiede zur eindimensionalen Signalverarbeitung. Im Regelfall besitzt man bei letzterer nämlich Freiheitsgrade sowohl bezüglich der eingesetzten Taktrate als auch in Hinsicht auf die Signallänge. Im Bedarfsfall kann an diesen Rädern gedreht werden. Das ist mit Regelmäßigkeit in der Bildverarbeitung nicht gegeben. Sowohl Bildformat als auch Auflösung sind durch Standards oder Hardware fast unverrückbar vorgegeben. Die Daten liegen oftmals in Form von 256 Graustufen vor. Eine Amplitudenauflösung von 8 bit ist in der eindimensionalen Signalverarbeitung jedoch eher selten akzeptabel. Farbbilder bestehen aus drei unterschiedlichen 8-bit-Kanälen, was erst recht keine Entsprechung in der eindimensionalen Welt hat. Man halte sich diese Einschränkungen immer vor Augen, wenn die Sinnfälligkeit von Verallgemeinerungen einzuschätzen ist, die ursprünglich für eindimensionale Signale entwickelt wurden.

Vorsicht bei Gefiltertem

Unerwarteten Problemen kann man schon bei linearen Filtern begegnen. Die gesamte Theorie der FIR- bzw. IIR-Filter läßt sich in sehr geradliniger Art und Weise in die

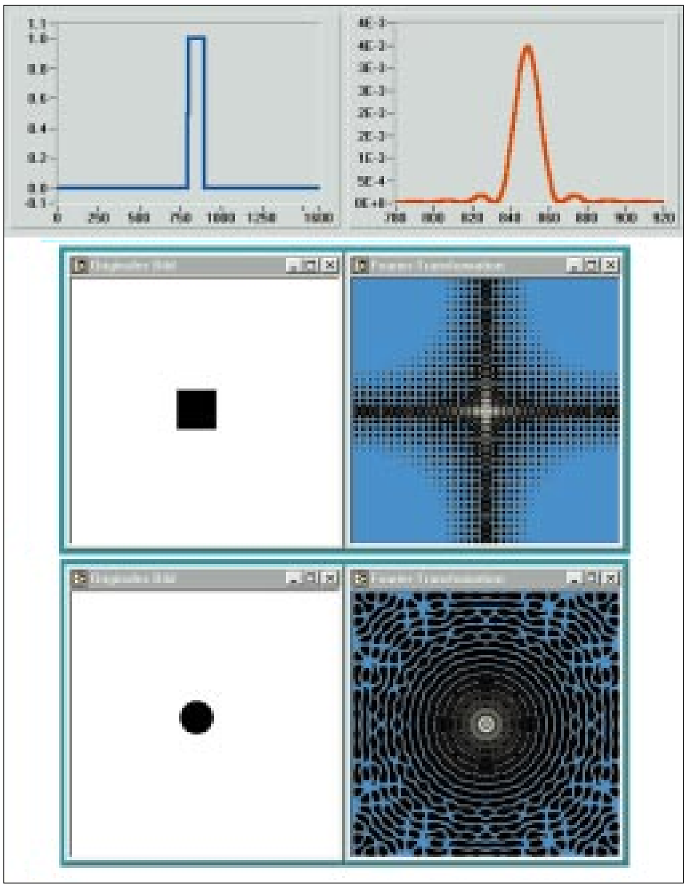

Bild 1. FIR-Tiefpaßfilter lassen sich ähnlich wie im Fall eindimensionaler Signale mit der Fourier-Transformation aufbauen. Allerdings gibt es bessere Verfahren, die darüber hinaus bestimmte Phaseneigenschaften gut konservieren. Gezeigt sind jeweils einfache Signale und deren zugeordnete Leistungsspektren. Die Ergebnisse der Fourier-Transformation kann man direkt in Filterkerne umarbeiten.

Bildverarbeitung übertragen. Als Beispiel seien FIR-Tiefpaßfilter genannt, deren Konstruktion auf der Fourier-Transformation beruht (*Bild 1*). Je nach Filterordnung kann der Kern dieser FIR-Filter unterschiedliche Ausdehnungen haben. Das ist im Fall eindimensionaler Signale vollkommen unproblematisch, weil fast ausnahmslos Signale sehr lang sind und die Ausdehnung des Filterkerns immer klein gegenüber der Signallänge ist. Bilder sind aber von Natur aus relativ kleindimensioniert.

Wenn der Kern einigermaßen groß ist, steckt man sofort in Schwierigkeiten. Diese resultieren einfach daraus, daß man den gesamten Rand von der Breite des Kerns opfern muß. Für diese Randzone liegen einfach nicht genug Werte vor, um etwas Vernünftiges zu berechnen. Leider sind die qualitativ besseren Filter gerade die mit ausgedehnteren Kernen. Somit steckt man in einem Dilemma, dem man zumeist dadurch begegnet, daß man mit geringerer Ausdehnung des Kerns arbeitet. Solche Nachbarschaften können aus 3 × 3, 5 × 5 Elementen oder im Einzelfall auch einmal bis zur Größenordnung 31 × 31 gehen. Die Ungeradzahligkeit dieser Dimensionen ist eine Folge der Symmetrieforderung (in Bildern gibt es nur im Ausnahmefall Vorzugsrichtungen). Kleinere Nachbarschaften kommen natürlich auch direkt der Abarbeitungsgeschwindigkeit zugute, die bei Bildverarbeitungsroutinen fast stets eine kritische Größe ist. *Bild 2* erläutert einige der gerade genannten Effekte anhand eines Beispiels.

Übrigens ist die Bildverarbeitung in bezug auf das Echtzeit-Verhalten der generellen Signaltheorie insofern überlegen, als man den letztlich anzustrebenden Grenzwert kennt. Fast immer ist dieser mit 30 Bilder/s definiert – was nicht heißt, daß man in jedem Fall dieses Ziel erreicht oder zumindest in die Nähe kommen kann.

Fourier-Transformation mit Einschränkungen

Ähnlich wie die Idee der linearen Filter läßt sich auch die der Fourier-Transformation auf die zwei- und übrigens auch höherdimensionalen Signale übertragen. Innerhalb der Bildverarbeitung wird man die Fourier-Transformation insbesondere bei zwei Gelegenheiten einsetzen. Der erste Anwendungsfall ist die Realisierung eines idealen Tiefpasses. Man setzt hierbei in der Fourier-Transformation alle hochfrequenten Anteile zu 0 und nutzt die gleichnamige Umkehroperation, um letztlich die Tiefpaßwirkung zu erzielen (*Bild 3*). Man umgeht hierbei vollkommen die weiter oben beschriebenen Schwierigkeiten beim Einsatz zahlreicher Filterkoeffizienten. Exzellenter Qualität des Ergebnisses steht allerdings ein unangenehmes und nur zögerliches Laufzeitverhalten entgegen. Man dürfte sich schwer tun,

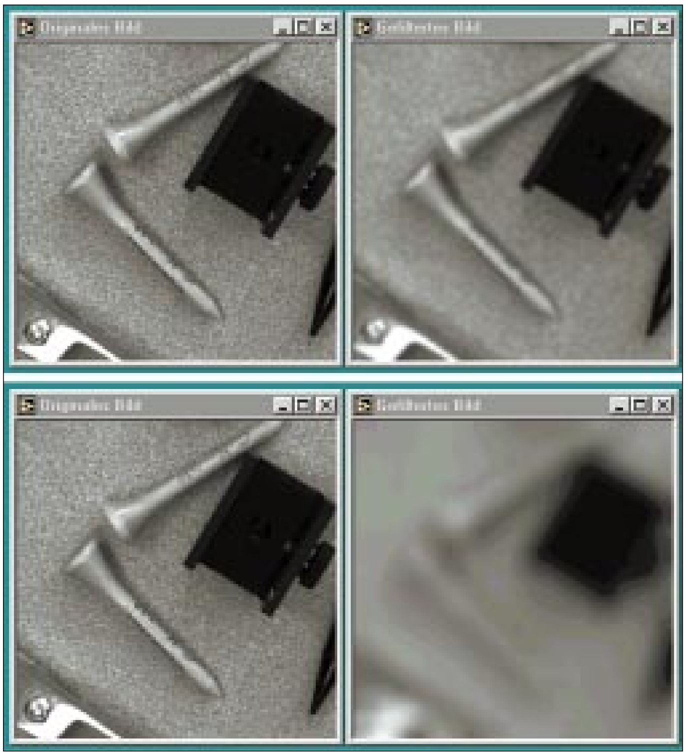

Bild 2. Die Größe des Filterkerns hat direkte Auswirkung sowohl auf die Ergebnisse als auch auf die erforderliche Rechenzeit. Oben besteht der Filterkern aus 5 × 5 = 25 Elementen, unten hingegen aus 25 × 25 = 625 Einträgen. Die Rechenzeiten sind 300 ms sowie 1400 ms (Pentium 200 MHz, IMAQ Vision, National Instruments).

selbst unter Einsatz der modernen Prozessoren die Angelegenheit in weniger als 500 ms zu erledigen, wobei Standardbildformate (z.B. 640×480) als Maßstab dienen.

Das andere Haupttätigkeitsfeld der Fourier-Transformation innerhalb der Bildverarbeitung bezieht sich auf die Kompression als Vorstufe zu einer Datenübertragung mit anschließender Rekonstruktion. Diese Methode hat viel mit der Filteridee gemein. Man ersetzt hierbei das Bild durch einige dominante Fourierkoeffizienten. Die Rekonstruktion gelingt durch Einsatz der inversen Fourier-Transformation. Allerdings wird bei aus Praxissicht wirklich interessanten Anwendungen anders vorgegangen. So stützt sich JPEG nicht auf das Gesamtbild, sondern auf kleinere Blöcke. Ferner wird nicht die Fourier-Transformation selbst eingesetzt, vielmehr als Abwandlung die Kosinus-Transformation, was den Vorteil mit sich bringt, daß an keiner Stelle komplexe Zahlen ins Spiel gebracht werden müssen. Als angenehmer Seiteneffekt kommt eine ziemlich flotte Abarbeitungsgeschwindigkeit hinzu.

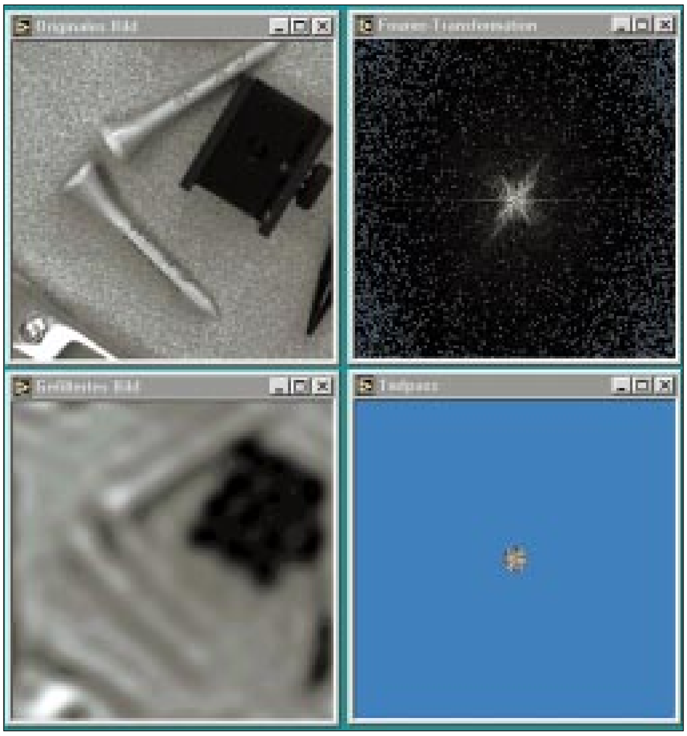

Bild 3. Realisierung eines idealen Tiefpasses, basierend auf der Fourier-Transformation. Man eliminiert hierbei alle Frequenzen jenseits einer bestimmten Schranke und transformiert das Resultat zurück in den Originalraum. Phaseneigenschaften werden hierbei allerdings nicht aufrechterhalten. Die Reihenfolge der Bilder während der Abarbeitung entspricht dem Uhrzeigersinn. Das Resultat besitzt keine höherfrequenten Anteile mehr.

Vorzugsrichtung in Bildern

Zahlreiche Bildverarbeitungsroutinen zeichnen sich dadurch aus, daß vor der eigentlich interessierenden Operation zunächst etwas an Vorbereitungsarbeit geleistet werden muß. Beispielhaft wird man eine Kantendetektion (siehe auch den abschließenden Teil 13 der Serie) erst dann vornehmen, wenn man Störungen herausgefiltert hat. Bei einem anderen Beispiel lassen sich manchmal ganze Objekte dadurch fixieren, daß man mit einer einfachen binären Schranke arbeitet, die alle Pixel unterhalb eines gegebenen Wertes weiß färbt sowie alle anderen schwarz markiert. Zur genannten Kategorie der vorbereitenden Routinen zählt ohne Zweifel auch die Karhunen-Loeve-Transformation, die manchmal auch unter dem Pseudonym Hotelling-Transformation auftritt.

Ziel ist es hierbei, die Vorzugsrichtung in einem gegebenen Bild aufzuklären. Unter der Vorzugsrichtung versteht man in diesem Zusammenhang die grundlegende Ausrichtung in einem Bild. Simpelstes Beispiel dafür ist ein liniertes Blatt Papier. Allerdings würde man zur Bestimmung der Ausrichtung der Zeilen nicht so schwere Geschütze wie das der Karhunen-Loeve-Transformation auffahren. Geht es aber etwa darum, die bildhafte Wiedergabe eines Computerboards so zu rotieren, daß man die kanonische Lage ermittelt, ist man wirklich nicht weit von realen Produktionsprozessen entfernt.

Das Verfahren der Karhunen-Loeve-Transformation läßt sich in vielerlei Hinsicht verallgemeinern. Letztlich ersetzt man ja eine Vielzahl von Vektoren einzig und allein durch zwei davon. In der folgenden Situation liegen die Dinge viel komplizierter:

Angenommen, man hat ein gegebenes Bild mit vielen weiteren zu vergleichen und den besten Treffer zu ermitteln. Zusätzlich bestehe die zuletzt genannte Sammlung von Bildern aus untereinander stark verwandten Exemplaren. Denkbar wäre beispielhaft ein Urbild und die um jeweils 1 Grad gedrehten Versionen davon. Ein Abgleich des Originals mit allen 360 Kandidaten wäre sicherlich zeitaufwendig. Nun zeigt sich aber, daß das zumeist nicht notwendig ist. Vielmehr lassen sich in der Größenordnung vier oder fünf Hauptbilder bestimmen, aus denen alle anderen durch einfachste Operationen abgeleitet werden können. Es genügt dann, die wenigen fixierten Bilder weiterzuverarbeiten, woraus enorme Rechenzeitersparnisse erwachsen. Solche Konstruktionen tragen aufgrund der starken Ähnlichkeit zu Eigenvektoren im eindimensionalen Fall deshalb wohl auch nicht zu Unrecht den Namen Eigenbilder (eigenimage).

Momente in Bildern

Man stelle sich ein weißes Blatt Papier vor und eine gegebene Auswahl von schwarz gefärbten geometrischen Objekten, z.B. Kreis, Quader, Rechteck, Halbkreis usw. Die Aufgabe bestehe nun darin, nach Aufnahme einer Figur mit Hilfe der Kamera zu entscheiden, welches Objekt man gesehen hat. Erschwerend kommt hinzu, daß das Objekt sowohl gedreht als auch skaliert sein kann. Natürlich ist auch der Ort nicht vorher bekannt. Man muß demnach Maßzahlen oder dergleichen aufspüren, die sich von solchen Bewegungen nicht beeindrucken lassen.

Es zeigt sich, daß die sogenannten Momente das Verlangte in effizienter Weise erledigen. Da es sich in der beschriebenen Situation um ein einfaches Schwarzweiß-Bild handelt, beziehen sich Momente verschiedener Ordnung nun einfach auf die Koordinaten aller schwarzen Punkte in der Vorlage. Hierbei wird ein objektabhängiges Koordinatensystem dadurch generiert, daß im ersten Schritt der Mittelpunkt (Schwerpunkt) der Figur ermittelt wird. Das geschieht durch simple Mittelwertbildung der x- und y-Komponenten aller schwarzen Punkte. Die sich daran anschließende Berechnung von Momenten höherer Ordnung ist in *Bild 4* näher erläutert. Es ist unschwer einzusehen, daß alle Momente bei einer gegebenen Figur konstant bleiben, wenn man Rotationen ausführt. Diese Eigenschaft ist äußerst angenehm und kann noch dergestalt ausgebaut werden, daß man auch Skalierungsinvarianz erreicht. Hierbei ist nur eine relativ einfache Normierungsstrategie anzuwenden.

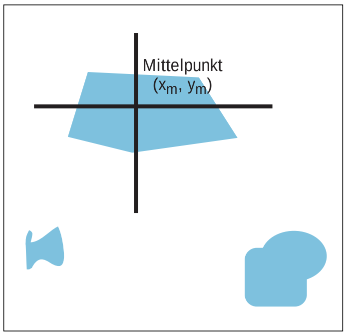

Bild 4. Einige Erläuterungen zur Bestimmung von Momenten schwarzer Objekte auf einem weißen Hintergrund. Momente erlauben eine unmittelbare physikalische Interpretation, so daß es nicht verwundert, daß mit deren Hilfe Objekte genau klassifiziert werden können.

Es zeigt sich nun, daß bereits sehr wenige Momente genügen, um Objekte eindeutig zu klassifizieren. Es läßt sich ohne weiteres ein Katalog für eine Schar von geometrischen Objekten aufbauen. Ein Vorteil dieser Methode liegt darüber hinaus darin, daß man bei Neuaufnahme einer geometrischen Figur in das Repertoire nur testen muß, ob die ausgesuchten Momente sich ausreichend gut von denen der bisherigen abgrenzen. Ist das nicht der Fall, muß ein Moment mit höherer Ordnung ins Spiel gebracht werden. Dieses Verfahren ist ausreichend schnell und besticht insbesondere durch seine Allgemeingültigkeit. Allerdings versagt es mehr oder weniger, wenn Grau- oder gar Farbwerte betrachtet werden müssen. Selbst bei Vorlage dieser Struktur läßt sich allerdings häufig über Schrankenbildung oder ähnliche Verfahren ein zugeordnetes binäres Bild generieren.

Mustererkennung – ein typisches Anwendungsbeispiel

Die Suche nach speziellen Informationen in einem Bild ist ein wichtiger und zumeist sehr kritischer Teil bei der automatisierten Bildauswertung. Es gibt eine Reihe von Suchtechniken, die das Aufspüren vorgegebener Muster in einem Bild erlauben. Die wohl bekannteste dieser Methoden beruht auf der Kreuzkorrelation von Signalen (siehe auch Teil 13 der Serie). Im vorliegenden Fall geht es naturgemäß um zweidimensionale Signale. Neuerdings gewinnen aber auch Ansätze wie neuronale Netze, Fuzzy-Logik oder das binäre Shape-Matching an Bedeutung.

Falls die Ausführungsgeschwindigkeit zum entscheidenden Faktor wird, ist zumeist eine Kombination von Linienprofilen und Kantendetektionen ein probates Mittel. Diese Technik nutzt Linienprofile zur Bestimmung des Randes bzw. der Kanten eines Objektes, das es zu untersuchen gilt. Im Interesse der Leistungssteigerung werden Annahmen über Orientierung und Form des interessierenden Objektes gemacht. Das einfachste Beispiel ist durch eine rechteckige Struktur gegeben. Linienprofile entlang der x- und y-Achse ermöglichen eine erste Orientierung. Hierdurch ist der Ort des Rechtecks schnell ermittelt. Im allgemeinen genügt bereits die genaue Kenntnis von drei Punkten, um die Orientierung des Rechtecks zu rekonstruieren. Kennt man die exakte Lage des Rechtecks, kann mit einem lokalen Koordinatensystem weitergearbeitet werden. Dieses System dient als Ausgangspunkt für die Bestimmung aller anderen zu inspizierenden Teile auf dem Originalbild.

Bild 5 zeigt eine realisierte Inspektionsaufgabe bei der Herstellung von Disketten, bei der es darum geht, ob das Label „HD“ korrekt gedruckt und angebracht wurde. Der erste Schritt besteht hierbei aus der Ermittlung der Lage der Diskette im mit einer Kamera aufgenommenen Original. Daran anschließend muß die Position des Labels bestimmt werden. Beides wird mit Hilfe zweier einfacher Routinen ausgeführt.

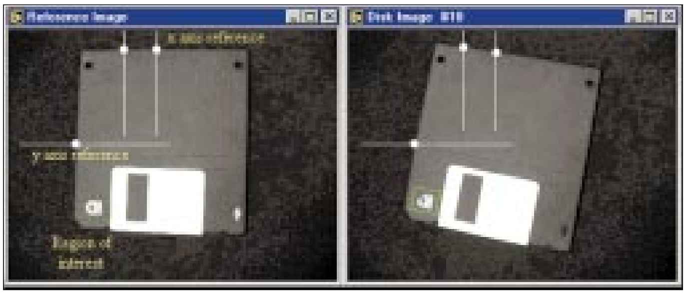

Bild 5: Inspektion des „HD“-Etiketts auf einer Diskette. Einige Linienprofile sind auf dem Referenzbild eingezeichnet. Ferner sind die Kantenpunkte auf diesen Profilen markiert. Darauf basierend gelingt der Aufbau des lokalen Koordinatensystems bezüglich des Originals. Daran anschließend kann eine „Region Of Interest“ festgelegt werden. Für jedes neue Bild muß dieser Prozeß wiederholt werden (rechts).

1. Aufbau eines Koordinatensystems relativ zu dem des Gesamtbildes. Auf dem Referenzbild (Bild 5) wird die Intensität entlang dreier Linien aufgenommen. Von diesen Linien haben zwei vertikale und eine horizontale Ausrichtung. Man sucht dann nach Kantenpunkten auf diesen Linien. Als Resultat ergeben sich zwei Punkte entlang der x-Achse und ein Punkt entlang der y-Achse des internen (Disketten) Koordinatensystems. Benutzt man diese Punkte, lassen sich sowohl der Ursprung als auch die Orientierung des gesuchten Koordinatensystems exakt berechnen. Damit ist der initiale Referenzpunkt gefunden.
2. Man bestimme eine „Region Of Interest“ (ROI) auf dem Originalbild, die garantiert das zu begutachtende Label als Teil enthält. Ausgangspunkt ist natürlich das gerade fixierte Koordinatensystem.

Charakteristisch für die allermeisten Inspektionsaufgaben ist die ständige Wiederholung ein- und desselben Zyklus (*Bild 6*). Das ist selbstverständlich auch in der vorliegenden Situation der Fall.

Die Differenz zwischen dem initialen Referenzpunkt und dem neuen entspricht gerade der Verschiebungs- und Rotationsbewegung der Diskette. Diese Information kann dazu genutzt werden, um die ROI in die korrekte Position zu bringen.

Ist die richtige Position des Etiketts erst einmal bestimmt, kann die eigentliche Überprüfung der Korrektheit der Labelanbringung mit Hilfe des sogenannten Shape-Matching erfolgen. Das Gebiet in der ROI wird, basierend auf einem passend gewählten Schrankenwert, binarisiert und mit einer vorab gespeicherten korrekten Vorlage verglichen. Werden Toleranzwerte überschritten, kommt es zur Aussonderung der gerade inspizierten Diskette. Im positiven Fall passiert die Diskette den Test.

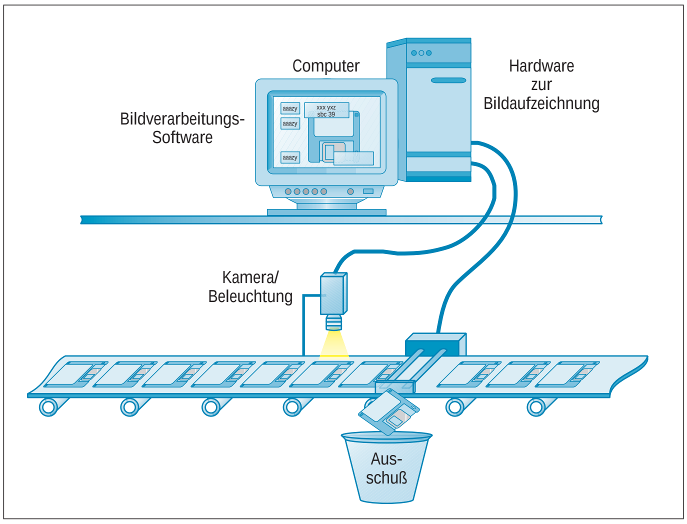

Bild 6. Ein typisches Inspektionssystem hat sich der Produktion anzupassen. Daraus resultieren zeitliche Vorgaben, die es einzuhalten gilt.

Weitere Anwendungen und Ausblick

Shape-Matching ist ein ideales Werkzeug, wenn die Form eines Objektes bekannt ist, die Orientierung und Größe jedoch Freiheitsgrade besitzen. Weitere Anwendungsfelder neben dem bereits genannten finden sich beispielhaft bei Fragestellungen im Zusammenhang mit Robotern. Eine typische Aufgabenstellung dieses Gebietes besteht darin, daß ein Roboterarm optisch unterscheidbare Teile auszusortieren und gemäß einem Vergleichskriterium zu gruppieren hat. Man benötigt in diesem Zusammenhang nachfolgende Eingabegrößen:

- die Form, nach der gesucht wird,
- ein binäres Bild, das alle potentiell zu sortierenden Teile umfaßt und
- einen Toleranzwert, der den zulässigen Unterschied zwischen Ideal und Realität definiert

Die Ausgabe besteht aus einem Bild, das einzig und allein die interessierenden Teile einzeln markiert aufzeigt. Zusätzlich wird ein Report mitgeliefert, der unter anderem die genaue Position der Einzelteile detailliert auflistet. Ferner werden bei Bedarf weitere Informationen mitgeliefert, zum Beispiel der Grad der Zuverlässigkeit der Übereinstimmung zwischen Mustervorlage und dem gefundenen Teil. Unter rauhen industriellen Bedingungen ist das nicht selten eine unerläßliche Komponente von automatisierten Inspektionssystemen.

Algorithmen wie die des binären Shape-Matching werden darüber hinaus bei der optischen Zeichenerkennung (Optical Character Recognition – OCR) beziehungsweise bei dessen Nachweis (OCV) benötigt. In diesem Fall sind die zugrundeliegenden Umrisse gerade die der verwendeten Buchstaben.

Die Forderungen an die Geschwindigkeit der automatisierten Inspektion sind zuweilen hoch. So kann man bei der Problematik der Qualitätskontrolle von Disketten ein Bild in etwa 30 ms aufbauen. Da der Produktionsprozeß nicht warten wird, ist allein dessen Taktrate maßgebend – und da kann es sehr wohl um Millisekunden gehen. Die vorgestellten Verfahren berücksichtigen das und werden äußerst schnell ausgeführt.

Im nächsten Teil dieser Serie stehen typische Anwendungen des Abschnitts „Fortgeschrittene Techniken“ im Mittelpunkt der Betrachtungen. *gs*

Literatur

[1] *Wenzel, L.:* Digitale Signalverarbeitung ist keine Hexerei, Teil 1. *Elektronik* 1998, H. 23, S. 96ff.
[2] *Wenzel, L.:* Digitale Signalverarbeitung ist keine Hexerei, Teil 2. *Elektronik* 1998, H. 25, S. 86ff.
[3] *Wenzel, L.:* Digitale Signalverarbeitung ist keine Hexerei, Teil 3. *Elektronik* 1999, H. 1, S. 52ff.
[4] *Wenzel, L.:* Digitale Signalverarbeitung ist keine Hexerei, Teil 4. *Elektronik* 1999, H. 3, S. 48ff.
[5] *Wenzel, L.:* Digitale Signalverarbeitung ist keine Hexerei, Teil 5. *Elektronik* 1999, H. 5, S. 60ff.
[6] *Wenzel, L.:* Digitale Signalverarbeitung ist keine Hexerei, Teil 6. *Elektronik* 1999, H. 7, S. 58ff.
[7] *Wenzel, L.:* Digitale Signalverarbeitung ist keine Hexerei, Teil 7. *Elektronik* 1999, H. 9, S. 62ff.
[8] *Wenzel, L.:* Digitale Signalverarbeitung ist keine Hexerei, Teil 8. *Elektronik* 1999, H. 11, S. 62ff.
[9] *Wenzel, L.:* Digitale Signalverarbeitung ist keine Hexerei, Teil 9. *Elektronik* 1999, H. 13, S. 91ff.
[10] *Wenzel, L.:* Digitale Signalverarbeitung ist keine Hexerei, Teil 10. *Elektronik* 1999, H. 15, S. 64ff.
[11] Die in den Artikeln angegebenen Programme stehen zum Download bereit unter: *www.natinst.com/germany*

Dr. Lothar Wenzel ist gebürtiger Berliner und hat Mathematik und Informatik in Greifswald und Dresden studiert. Nach Tätigkeiten in einem Kernkraftwerk und bei der BASF AG, Ludwigshafen, beschäftigt er sich z.Zt. bei National Instruments in Austin, Texas, mit dem Design von Algorithmen auf den Gebieten Simulation, Regelung, Mathematik und Bildverarbeitung.