

Dr. Lothar Wenzel

Digitale Signalverarbeitung ist keine Hexerei

Teil 9: Anwendungen zum Kapitel Systeme

Die letzten drei Teile dieser Serie befaßten sich mit den Grundlagen der Systembetrachtungen. Dabei wurden zunächst Systeme definiert und typische Fälle wie Rückkopplung, PID-Regler, Abtastung, Laplace-Transformation, Signalrekonstruktion sowie Datenfitting und Vorhersageprobleme erörtert. Dieser letzte Teil des Kapitels Systeme zeigt anhand von einigen Beispielen anschaulich die Anwendungsmöglichkeiten dieser theoretischen Grundlagen.

In diesem Abschnitt der Serie geht es wieder um Anwendungen bereits vorgestellter Ideen und Theorien, die gerade das Gebäude der digitalen Signalverarbeitung ausmachen. Im Mittelpunkt stehen Fragen im Zusammenhang mit Systemen, Regelungen, Abtastungen und insbesondere auch Fittingstrategien. Bereits bei den hier vorzustellenden, relativ einfachen Anwendungen zeigt sich eine generelle Tendenz des Gebietes der digitalen Signalverarbeitung: Erst die Kombination unterschiedlicher Teilthemen führt zum Erfolg. So zeichnet sich das erste Beispiel durch eine gelungene Kombination von zweidimensionaler Abtastung, geeigneter Überwachungsstrategie und im Verborgenen arbeitenden Faltungsoperatoren aus.

Tracking – Verfolgung von Objekten

Genauer gesagt geht es um das Tracking von Objekten in gegebenen Bildsequenzen. Unter den vielen möglichen Auf-

gabenstellungen aus der Familie der Trackingprobleme sei eine herausgegriffen, die sozusagen als der Standardfall gelten kann: Eine Kamera sei fest montiert und beobachtet eine natürliche Szene. Ganz am Anfang werde ein Objekt im Bild als Target (gebräuchlich ist hierfür zuweilen auch der Begriff Template) fixiert. Im Regelfall geschieht das durch Definition eines Rechteckes, das vollkommen das zur Debatte stehende Objekt umschließt. Derartige Rechtecke sind normalerweise eher klein gegenüber dem ganzen von der Kamera erfaßten Ursprungsbild. Eine typische Größe eines Templates bei einem originalen Bildformat von 512×512 Bildpunkten könnte beispielsweise 64×64 sein. Ziel ist es nun, das sich bewegende Objekt zu verfolgen, indem das umgebende Rechteck passend mitgeführt wird. Üblicherweise ist man an einem Echtzeit-Verhalten interessiert, was die Aufgabenstellung nicht gerade erleichtert.

Der Anwendungsbereich hierfür ist breit und reicht von militärischen Projekten über zivile Überwachungsaufgaben bis hin zur sicheren und korrekten Erkennung von Kopf- oder Augenbewegungen im Sinne von neuartigen Kontrollinstrumenten. Was macht diese Aufgabenstellung in voller Allgemeinheit so kompliziert?

Dafür gibt es mehr als einen einzigen Grund. Zum einen kann sich die Objektgestalt fortlaufend ändern. Eine mögliche Ursache hierfür sind zum Beispiel instabile Beleuchtungsverhältnisse. Zusätzlich kann das Objekt sowohl in der Projektionsebene als auch – und das wiegt meist schwerer – im dreidimensionalen Raum mannigfaltige Bewegungen ausführen. Gewiß sieht ein Buch anders aus, wenn man statt des Deckblattes den Buchrücken zur Ansicht angeboten bekommt. Es kommt hinzu, daß Objekte am Bildrand anders projiziert werden als in Fokusposition. Schließlich ist auch nicht zu vergessen, daß nicht selten Templates durch andere Objekte ständig oder kurzzeitig abgedeckt werden können.

Zunächst soll die Annahme gelten, daß die einzigen zulässigen Bewegungen des Objektes aus Verschiebungen parallel zu den Bildrändern bestehen können. Anders gesagt: Das Template führt keine Drehbewegungen irgendwelcher Art auch immer aus. Damit ist keineswegs ge-

meint, daß das Objekt sich nur auf geradlinigen Bahnen bewegen kann. Ganz im Gegenteil, wie *Bild 1* belegt, sind zahlreiche Trajektorien denkbar. Diese Bewegungsform ist vielleicht etwas wirklichkeitsfremd und man muß schon etwas länger nachdenken, bis man ein geeignetes Beispiel dafür in der realen Welt findet. Ein Bergsteiger, der eine steile Wand erklimmt, entspricht mit Abstrichen diesem Bild relativ gut. Natürlich sind auch Güter, die auf Förderbändern transportiert werden, eine naheliegende Anwendung dieses Prinzips.

Originalbild

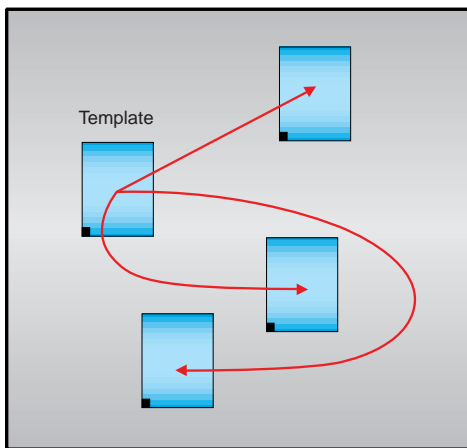


Bild 1. Gezeigt sind einige denkbare Trajektorien beim formulierten Tracking-Problem. Nur die Orientierung des Templates ist aufrechtzuerhalten (kleines Quadrat immer in der linken unteren Ecke). Ansonsten sind zahlreiche Bewegungen zulässig.

In jedem Fall wird man von Bild zu Bild nur kleinere und fast unbedeutende Bewegungen des Objektes feststellen. Ein erfolgreiches Tracking setzt demnach voraus, daß man diese Mikrobewegungen sicher und vor allen Dingen auch sehr schnell detektieren kann. Reißt der Faden, ist er nur schwer wieder aufzunehmen, weil man ständig gegen den Zeitdruck der nachfolgenden Szene anrennen muß. Seitens der Bildverarbeitung stützt man sich hierbei auf einen sehr einfachen Ablauf, der als wesentliche Komponente Gradienten in x- und y-Richtung enthält (*Bild 2*). Man nutzt im Regelfall hierbei den sogenannten Prewitt-Operator, weil dieser sich in der Praxis als hochgeeignet herausgestellt hat. Aus signaltheoretischer Sicht ist der Prewitt-Operator ein eher einfach aufgebautes zweidimensionales Filter, das extrem gut auf die Änderungen in der jeweiligen Vorzugsrichtung anspricht.

An dieser Stelle setzt die Regelungskomponente der Lösung an. Folgende Schleifenstruktur muß ständig durchlaufen werden.

- (a) Man ermittle die Differenz zwischen den Rechteckinhalten zweier unmittelbar benachbarter Rahmen (Frames)!
- (b) Man bestimme einen Schätzwert für die die Veränderungen auslösende Objektbewegung!
- (c) Man verschiebe (das ist das eigentliche Tracking) das Rechteck dementsprechend!

Die Entwicklung eines hierfür geeigneten Algorithmus stützt sich vollkommen auf eine geeignete Fittingroutine. Führt man all diese Operationen exakt aus, ergibt sich ein zweistufiges Verfahren, das das formulierte Tracking-Problem zufriedenstellend löst.

Phase A: Vorbereitung

- (1) Man bestimme die Zielregion, die im weiteren verfolgt werden soll!
- (2) Es ist das Referenzbild aufzunehmen und geeignet abzuspeichern!
- (3) Man berechne eine bestimmte Matrix, die sich einerseits auf das Template und andererseits auf die beiden Prewitt-Ableitungen stützt!

Phase B: Eigentliches Tracking

- (4) Man nutze die gerade ermittelten Bewegungsparameter, um das Objekt einen kleinen Schritt weiter zu verfolgen! Hierzu sind gegebenenfalls bestimmte Transformationen auszuführen.
- (5) Es ist die Differenz zwischen vorhergesagter und real erreichter Position zu ermitteln. Das geschieht punktweise, d.h., es geht im wesentlichen um die Differenz zweier ausreichend benachbarter Bilder.
- (6) Man bestimme hierauf basierend die neuen Bewegungsparameter! Anschließend gehe man zum Punkt (4) und wiederhole den Ablauf!

Wichtig sind noch folgende Bemerkungen. Phase A wird nur ganz zu Beginn der Objektverfolgung ausgeführt. Das ist mit Abstand der rechenaufwendigste Teil des gesamten Algorithmus. Phase B erfordert Schleifendurchläufe von nur einigen wenigen Millisekunden, woraus die Eignung für reale Anwendungsfälle deutlich wird. Läßt man kompliziertere Bewegungsarten des Objektes zu, ändert das kaum die letzten Bemerkungen, d.h., das eigentliche Tracking bleibt ausreichend schnell. Das trifft auch dann zu, wenn man die unter (a) bis (c) formulierte Reglerstruktur durch bessere und leistungsstärkere Modelle ersetzt. Insgesamt

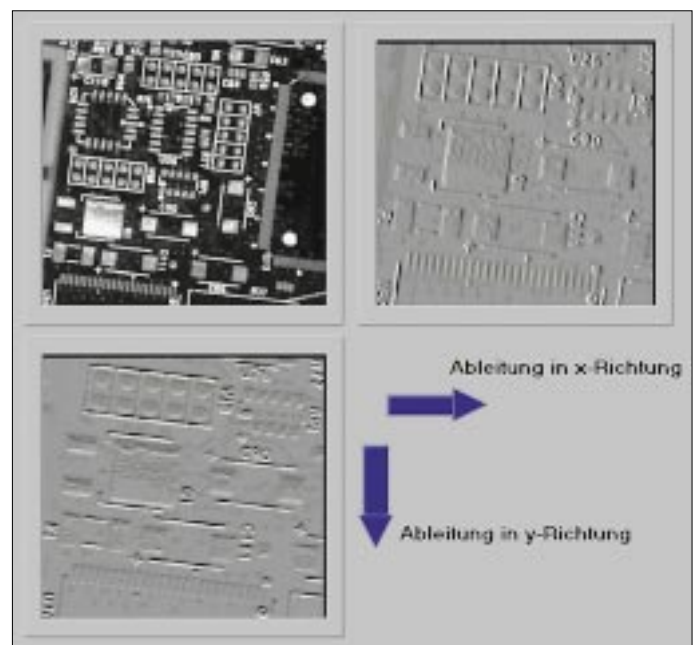


Bild 2. Der Prewitt-Operator realisiert eine Berechnung, die bei eindimensionalen Signalen der Ableitung entspricht. Da man sowohl in x- als auch in y-Richtung differenzieren kann, resultieren hieraus zwei verschiedene Bilder. Man sieht deutlich die unterschiedlichen Effekte in beiden Richtungen.

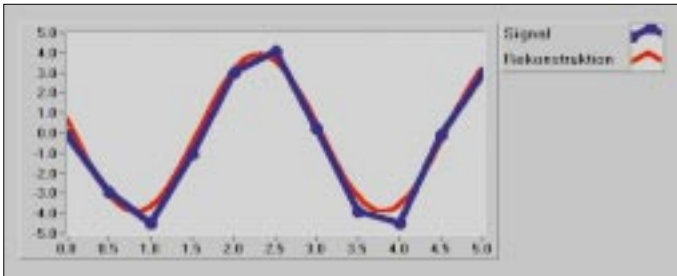


Bild 3. Es liegen nur wenige und darüber hinaus gestörte Daten der Sinusschwingung $4,23 \cdot \sin(2,08t + 3,10)$ vor. Trotzdem hat der Algorithmus keinerlei Probleme, die wahren Gegebenheiten aufzuklären. Das Rekonstruktionsergebnis ist $3,91 \cdot \sin(2,12t + 2,96)$. Stehen mehr Daten zur Verfügung, nimmt die Qualität des Resultates weiter zu.

gesehen ist der gerade vorgestellte Algorithmus zur Objektverfolgung eine wohlgelungene Mixtur aus unterschiedlichsten Teildisziplinen des Gebietes der digitalen Signalverarbeitung.

Energie-Inhalte von Signalen

Angenommen, man will die Funktionalität eines Digitalvoltmeters mit Hilfe einer A/D-Umsetzercarte und eines Computers nachstellen. Eine wichtige Teilaufgabe besteht dann darin, von einem sinusförmigen Wechselstromsignal den Energie-Inhalt und die Frequenz möglichst exakt zu bestimmen. Problematisch daran ist, daß die Frequenz von einigen Hz bis zu mehreren 100 kHz variieren kann. Klassisch würde man die Energiebestimmung höchstwahrscheinlich so vornehmen, daß man mit extrem hoher Abtastrate (etwa 1 MHz) arbeitet und den Energie-Inhalt E mit Hilfe der Gl. (1) abschätzt.

$$E = \frac{1}{N} \sum_{k=1}^N x[k]^2 \quad (1)$$

Hierbei besteht das Signal aus genau den N Werten $x[1]$, $x[2]$, ..., $x[N]$, und die Summation bezieht sich gerade auf

diese Datenmenge. Nun sind allerdings die Signale nie ganz störungsfrei, und eine A/D-Umsetzercarte arbeitet nicht absolut perfekt. Selbst wenn man die statistischen Eigenschaften dieser Störungen genauestens kennt, bleibt immer noch ein Problem offen. Zwar konvergiert die obige Abschätzung für immer größer werdende N gegen den wahren Energie-Inhalt E , jedoch muß zur Erreichung einer ausreichend hohen Genauigkeit N in der Größenordnung von Millionen Abtastungen pro Sekunde liegen. Das hätte zur Folge, daß man für die Ermittlung eines guten Schätzwertes für den Energie-Inhalt gemäß obiger Formel eine Zeitspanne von einer oder sogar mehreren Sekunden investieren müßte, um die übliche Genauigkeit zu erzielen. Dieser Wert ist für Digitalvoltmeter aber ziemlich groß, so daß man nach besseren Verfahren suchen muß.

Die Situation ändert sich jedoch schlagartig, falls man das verfügbare Wissen voll zum Einsatz bringen kann. So liegt ja gemäß Annahme ein sinusförmiges Signal vor, das sich in allgemeiner Form mit folgender Gleichung beschreiben läßt:

$$x(t) = A \cdot \sin(a + ft) \quad (2)$$

Fittet man basierend auf dieser Modellvorstellung (Bild 3) und basierend auf einigen wenigen Meßwerten, so läßt sich bereits mit mehreren einzelnen Daten eine ausgezeichnete Qualität erreichen. Nebenbei bekommt man noch Phasen- und Frequenzinformationen mitgeliefert, was allein unter Zuhilfenahme der ursprünglichen Annäherung für E nicht gelingt.

Die zugrundeliegende Methodik stammt von Prony. Man geht dabei davon aus, daß ein Signal sich als Summe von genau zwei komplexwertigen Exponentialfunktionen $A_1 \exp(ja_1 t)$ und $A_2 \exp(ja_2 t)$ darstellen läßt. Beide besitzen dementsprechend Amplituden- und Phaseninformationen. Im vorliegenden Fall ergibt sich ja ein reellwertiges Signal, woraus einige Beziehungen zwischen den A - und a -Parametern folgen. Bis auf komplex-konjugierte Unterschiede sind diese nämlich jeweils identisch. Hierdurch vereinfachen sich die Verhältnisse enorm, und die Ermittlung der unbekanntenen Parameter reduziert sich auf einige wenige elementare Operationen.

Besitzt man darüber hinaus noch weitergehende Informationen über den statistischen Charakter der Störungen, so läßt sich die Genauigkeit der Aussage weiter steigern. Stichworte hierzu sind Wiener- bzw. Kalman-Filter.

Ähnlich würden die Dinge liegen, wenn rechteckige oder andere wohldefinierte Signalformen vorlägen. Verfügbares Zusatzwissen läßt sich stets in eine quantitative Verbesserung der Ergebnisse ummünzen. Der Nachteil von Ansätzen wie dem der einfachen Abschätzung des Energie-Inhaltes E basierend auf einer Mittelwertbildung liegt gerade darin, daß man keinerlei Angaben zum Charakter der zugrundeliegenden Signale einarbeitet. Solche allgemeinen Fragestellungen können dann naturgemäß nur höchst ineffiziente Implementierungen nach sich ziehen.

Abtastung in n Dimensionen

Zuweilen liegen nicht die gewohnten eindimensionalen und zeitabhängigen Signale vor, vielmehr bestimmen zahlreiche unabhängige Größen das Signalgeschehen. Ein typisches Beispiel hierfür sind chemische Prozesse, bei denen konkrete Temperatur-, Druck-, Konzentrations- und Durchflußraten zu einem einzigen Resultat – in diesem Fall der Ausbeute – führen. In der Technik, Wissenschaft oder auch insbesondere in der Finanzwelt findet man weitere Beispiele der geschilderten Art. Gemeinsam ist all diesen Anwendungen, daß die Anzahl der unabhängigen Parame-

ter groß und zuweilen sogar riesig sein kann. In Spezialfällen kann man es auf Zahlen jenseits von 100 und teilweise sogar deutlich jenseits der 1000 bringen. In der Sprache der digitalen Signalverarbeitung heißt das nichts anderes, als daß man hochdimensionale Strukturen zu behandeln hat.

Nun zeigt eine einfache Abschätzung, daß die in der ein- oder zweidimensionalen Signalverarbeitungswelt üblichen Vorstellungen über Abtastungen bei höherer Dimensionszahl faktisch wertlos werden. Beschreibt man etwa jede Dimension mit Hilfe der gewiß nicht großen diskreten Teilmenge bestehend aus jeweils 100 Elementen, so müßte man bei Vorlage von 10 Dimensionen insgesamt $100^{10} = 10^{20}$ Abtastungen vornehmen. Da hinter jeder einzelnen

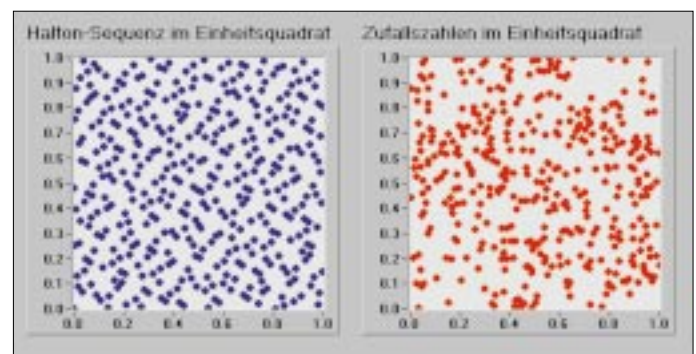


Bild 4. Darstellung von Quasizufalls- und im üblichen Sinne zufälligen Sequenzen in zweidimensionalen Fall. Die rechte Struktur (rein zufällige Folge) ist wesentlich inhomogener verteilt. Es zeigen sich einige Lücken und Anhäufungen im Datenmaterial.

Quasizufalls-Sequenzen nach Halton

Haltons Methode stützt sich auf Primzahlen. Die produzierten Zahlenfolgen sind zu jedem Augenblick perfekt auf dem Einheitsquadrat verteilt.

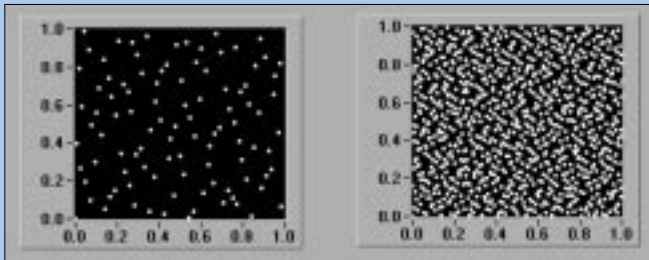
Beispiel:

1. Primzahl: 2
2. Primzahl: 3
- $n = 11$

1. Schritt: Konvertierung
- $n = 11$ in 2-er Darstellung: 1011
- $n = 11$ in 3-er Darstellung: 102

2. Schritt: Reihenfolge umdrehen
- 2-er Darstellung umdrehen und mit Komma versehen: 0,1101
- 3-er Darstellung umdrehen und mit Komma versehen: 0,201

3. Schritt: Interpretation
- $x = 1*(1/2)+1*(1/4)+0*$
 $(1/8)+1*(1/16) = 13/16 = 0,8125$
- $y = 2*(1/3)+0*(1/9)+1*$
 $(1/27) = 19/27 = 0,7037037037037$



Die ersten 100 (links) sowie die ersten 1000 Quasizufalls-Zahlen, entsprechend Haltons Idee aus den Primzahlen 2 und 3 ermittelt.

Wertermittlung physikalische Prozesse stehen, die zumeist nicht übermäßig schnell und billig zu bekommen sind, sehen die Dinge mehr als schlecht aus. Ist man auf der anderen Seite beispielhaft an dem Mittelwert der Ergebnisse bei allen möglichen Eingabeparametern interessiert, ist es nur schwer einzusehen, daß es überhaupt eine zufriedenstellende Lösung hierfür geben soll.

Trotz dieser eher düsteren Prognosen gibt es einen Ausweg aus diesem Problem, der mit fast todsicherer Genauigkeit zum Erfolg führt. Es ist klar, daß man wegen der riesigen Zahl an Varianten unmöglich alle Parameterkonstellationen austesten kann. Auf den ersten Blick gibt es dann eigentlich nur zwei Auswege. Man kann einige wenige Parameterkombinationen auf zufälliger Basis zusammenstellen, oder es steht der Weg frei, mit regulären Mustern im Parameterraum zu agieren. Solche Muster könnten Ausdünnungen des ursprünglichen Rasters sein.

Es läßt sich nun nachweisen, daß beide Verfahren in der Praxis versagen. Die Schwachstelle der ersten Methodik liegt darin begründet, daß man zwar immer neue, auf zufälliger Basis ermittelte Parameterkonstellationen zusammenstellen und somit neue Schätzwerte für den eigentlich gesuchten Mittelwert berechnen kann, jedoch ist die Konvergenzgeschwindigkeit so träge, daß diese Approximationen praktisch wertlos sind. Mit wachsender Dimensionszahl nimmt die Qualität dieser Schätzwerte sogar dramatisch ab. Die andere Methode der groberen Raster ist nicht viel besser, weil man selbst dabei jeden vernünftigen Rah-

men bezüglich der erforderlichen Rechenzeit sprengt.

Der überraschende Ausweg liegt in einer Mischung von zufälligen und deterministischen Anteilen. *Bild 4* zeigt den Unterschied zwischen Quasizufalls- und im üblichen Sinne zufälligen Sequenzen für den eher untypischen Fall eines zweidimensionalen Signals. Die links dargestellte Punktwolke sieht im Einheitsquadrat sowohl zufällig als auch exzellent verteilt aus. Man kann Quasizufalls-Folgen Schritt für Schritt aufbauen, ohne bereits einmal ermittelte Werte über Bord werfen zu müssen. Darüber hinaus sind Schätzwerte für die Mittelwerte ungleich besser zu ermitteln, falls man sich anstatt von rein zufällig gewählten Parameterkonstellationen auf Quasizufalls-Sequenzen stützt. Die Qualitätsunterschiede sind so prägnant, daß im Regelfall und insbesondere bei hoher Dimensionszahl nur letztere Methodik praktisch einsetzbar ist (*Bild 5*). Bleibt die Frage zu beantworten, wie man sich solche Folgen im n-dimensionalen Raum beschaffen kann.

Quasizufalls-Sequenzen

Es gibt zahlreiche unterschiedliche Verfahren zur Generierung von Quasizufalls-Folgen. Der mathematische Schwierigkeitsgrad ist unterschiedlich kompliziert, die entscheidenden Eigenschaften sind jedoch bei allen Methoden sehr ähnlich. Einige der Konstruktionsschemen hierfür sind bereits mehr als 30 Jahre alt und wurden seinerzeit mit einer gänzlich anderen Zielsetzung entwickelt. Anwendungen in der Signaltheorie sind relativ neu. Um die Dinge hier zu vereinfachen, sei der Fall des zweidimensionalen Signalraumes herausgegriffen. Das Konstruktionsschema geht auf Halton zurück.

Hier sei der Fall des Einheitsquadrates in den Mittelpunkt gerückt. Der *Kasten* erläutert detailliert eine Berechnungsvorschrift für solch eine Folge. Ausgangspunkt sind zwei Primzahlen, z.B. 2 und 3. Es sei das n-te Element der Quasizufalls-Zahlenfolge im Quadrat zu berechnen. Man notiere dazu n sowohl in der Zahlendarstellung zur Basis 2 als auch zur Basis 3. Daran anschließend drehe man die so ermittelte Darstellung einfach um und setze das Komma in der entsprechenden Zahlendarstellung davor. Was entsteht, sind zwei Zahlen jeweils zwischen 0 und 1, d.h. ein Punkt im Quadrat. Es sind genau die x- bzw. y-Komponenten des n-ten Punktes der Quasizufalls-Zahlenfolge.

Das Programm „HALTON“ [9] ist ein kleines Experimentierwerkzeug zum besseren Kennenlernen der Quasizufalls-Sequenzen. Man hat die Wahl zwischen verschiedenen Primzahlkombinationen (jeweils 2, 3, 5 und 7). Fallen

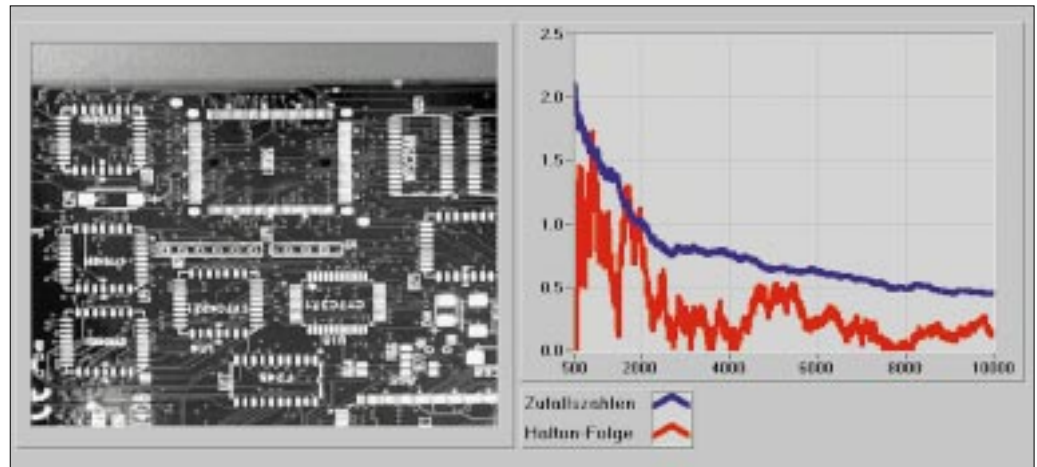


Bild 5. Quasizufalls-Sequenzen sind gewöhnlichen Zufallszahlenfolgen in vielen Fällen qualitativ deutlich überlegen. Im Beispielfall gilt es, den mittleren Grauwert basierend auf nur relativ wenigen Pixelwerten einer Vorlage (links) abzuschätzen. Die Bildgröße ist 300×388 , d.h., insgesamt liegen 116 400 Punkte vor. Der rechten Grafik läßt sich entnehmen, daß man mit etwa 2000 Testpixeln und basierend auf Halton-Sequenzen bereits eine Genauigkeit von 0,5 erreicht hat. Reine Zufallszahlenfolgen benötigen im Mittel dafür jedoch 8000 Pixel.

beide Werte zusammen, ergibt sich allerdings alles andere als eine Zufalls-Zahlenfolge, so daß man diese Situation vermeiden sollte. Neben der Halton-Sequenz als Bild werden auch gewöhnliche Zufalls-Zahlenfolgen mit ausgegeben. Das gestattet eine Begutachtung der Qualität beider Verfahren.

Ist man am höherdimensionalen Fall interessiert, muß man einfach mit entsprechend vielen verschiedenen Primzahlen starten. Jede korrekte Wahl eines Satzes von Primzahlen führt automatisch zu Quasizufalls-Zahlenfolgen, die in der oben beschriebenen Weise direkt einsetzbar sind. Es gibt folglich jede Menge derartiger Zahlenfolgen. Alle weisen dieselbe prinzipielle Qualitätseigenschaft auf.

Im nächsten Teil der Serie startet das Kapitel „Fortgeschrittene Techniken“, in dem es zunächst um die sehr modernen Teilgebiete Wavelets und Joint Time-Frequency Analysis geht. Beide Richtungen haben offengelegt, daß bei der Fourier-Transformation noch lange nicht Schluß ist. *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] Die im Text angegebenen Programme stehen zum Download bereit unter: www.natinst.com/germany