

# Prinzip und Leistungsfähigkeit der Korrelation und Zuordnung digitaler Bilder

Wolfgang Förstner, Universität Stuttgart

## 0. Problemstellung

Seit über 25 Jahren werden Verfahren zur automatischen Stereoauswertung von Bildpaaren entwickelt. Die zentrale Aufgabe bei der Automatisierung besteht darin, zugeordnete oder homologe Punkte in beiden Bildern mit hinreichender Genauigkeit zu bestimmen. Bei gegebener Orientierung lassen sich dann vergleichsweise einfach durch Strahlenschnitt dreidimensionale Koordinaten von Objektpunkten ableiten.

Nun tragen einzelne Bildpunkte, ob in einem analogen oder digitalen Bild, nicht genügend Information - einen Helligkeits- oder Grauwert oder drei Farbwerte - um im Nachbarbild eindeutig auffindbar zu sein. Die Suche zugeordneter Punkte stützt sich daher immer auf Merkmale, die aus der näheren Umgebung der Punkte abgeleitet werden, etwa auf die Helligkeitsverteilung selbst, daraus abgeleitete Merkmale wie Kanten oder Ecken oder gar mit einer Bezeichnung versehene Merkmale wie "Signalisierter Punkt", "Baumspitze" oder "Schachtdeckel". Bei einem der Bilder kann es sich auch um ein durch die Grauwertverteilung gegebenes Objekt, etwa eine Kante oder einen signalisierten Punkt, handeln. Zuordnungsalgorithmen lassen sich dann, u. U. in leicht modifizierter Form zur Lokalisierung dieser Objekte verwenden. Die Suche zugeordneter Punkte besteht im wesentlichen in der Suche von Punkten, deren Merkmale in beiden Bildern ähnlich sind, oder kurz in der Suche ähnlicher Punkte.

Zwei grundsätzliche Fragen ergeben sich unmittelbar, unabhängig von der algorithmischen Lösung eines Verfahrens:

### 1.) Wie lassen sich Fehlzusammenordnungen vermeiden?

Da verschiedene Punkte innerhalb eines Bildes ähnliche Merkmale aufweisen, d.h. ähnlich sein können, muß man im Prinzip mit groben Fehlzusammenordnungen rechnen. Das menschliche visuelle System hat damit i. a. nur geringe Schwierigkeiten, wenn man von periodischen Mustern absieht, etwa bei karierten Stoffen oder Heizungsblenden.

### 2.) Wie genau ist die Zuordnung von Punkten möglich?

Die Genauigkeit der Parallaxenmessung in einem photogrammetrischen Auswertegerät wird - von systematischen Fehlern abgesehen - wesentlich durch die Qualität des abgebildeten Objekts und die verwendete Vergrößerung bestimmt. Bei starker Vergrößerung lassen sich in einem guten Luftbildpunkte mit einer Genauigkeit bis zu 1 µm stereoskopisch einstellen. Diese Genauigkeit wird bei schlechtem Bildmaterial, selbst bei genügender Vergrößerung nicht erreicht. Die Frage ist daher, ob automatische Systeme die in Bildern innewohnende Genauigkeit in gleicher Weise oder gar besser nutzen können wie der menschliche Betrachter.

Der vorliegende Beitrag möchte in drei Schritten diesen Fragen nachgehen. Die Diskretisierung d. i. die Auflösung des analogen Bildes in Bildelemente, und die Quantisierung, d. i. die Analog-Digital-Wandlung der Helligkeitsinformation beeinflussen die Genauigkeit der Zuordnung wesentlich. Der 1. Abschnitt zeigt einerseits, daß bei fehlerfreier Digitalisierung und idealen Daten Kanten und Kreise mit einer Genauigkeit weit unter der Bildelementgröße selbst in Binärbildern lokalisiert werden können, andererseits, inwieweit die Digitalisierung die Genauigkeit begrenzt. Da das zugrunde liegende Modell rein geometrisch ist, zufällige Fehler nur schwer zu integrieren sind und vor allem nur auf einfache Objekte anwendbar ist, werden im zweiten Abschnitt die klassische Korrelation und die damit verwandte Kleinste Quadrate Anpassung zur Objektlokalisierung und zur Zuordnung von Punkten diskutiert. Hierbei kann man die Methoden zur Beurteilung von Ausgleichsergebnissen nutzbringend verwenden. Wichtigster Punkt ist die Untersuchung der Parameter, die die Genauigkeit der Zuordnung bestimmen und die Frage nach der optimalen Größe der Bildelemente und der Bildausschnitte. Es zeigt sich, daß man eine Genauigkeit der Zuordnung ähnlich der des visuellen Systems erreichen kann. Für einen operationellen Einsatz dieser Verfahren sind aber extrem gute Näherungswerte ( $\pm 30 \mu\text{m}$ ) erforderlich, weshalb der 3. Abschnitt eine Methode beschreibt, mit der man nicht nur gute Näherungswerte erhält, sondern mit dem auch die Wahrscheinlichkeit von Fehlzusammenordnungen wesentlich vermindert wird.

## 1. Grenzen der Genauigkeit bei der Lokalisierung einfacher Objekte in digitalen Bildern

### 1.1 Digitalisierung von Bildern

Ein digitales Bild entsteht durch Abtasten eines Objekts oder des analogen Bildes eines Objekts. Vereinfacht kann man sich den Prozess folgendermaßen vorstellen: Gegeben sei eine kontinuierliche Funktion  $g(x)$ , die etwa die Dichte oder den Grauwert eines Bildes an der Stelle  $x$  darstellt (vgl. Fig. 1-1). Die Abtastung besteht in der Diskretisierung und anschließender Quantisierung. Die diskretisierte Funktion (Fig. 1-2) ist die Folge  $\{g(x_i)\}$  der Funktionswerte an den Abtastpunkten  $x_1, x_2, \dots, x_i, \dots$ . Die Abstände werden üblicherweise gleich gewählt. Der Abtastabstand sei  $\Delta x$ , also:  $x_i = i \Delta x$ . Die Funktionswerte  $g(x_i)$  sind reelle Zahlen. Die Darstellung im Rechner erfordert eine Quantisierung dieser Werte, d. h. eine Rundung entsprechend der gewählten Wortlänge

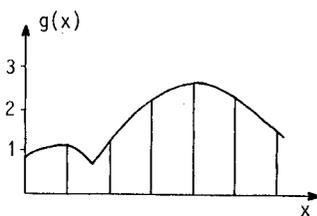


Fig. 1-1 Kontinuierliche Funktion  $g(x)$

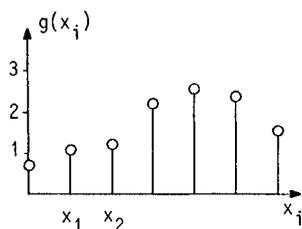


Fig. 1-2 Diskretisierte Funktion  $g(x_i)=g_i$

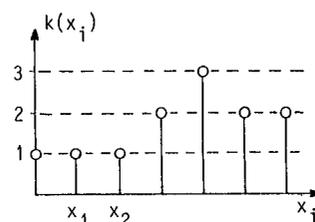


Fig. 1-3 Digitalisierte Funktion  $k(x_i)$ ,  $b = 2$  bit

$b$  in bit, üblicherweise  $b = 8$ . Damit kann der Funktionswert nur  $N = 2^b (= 256)$  verschiedene Werte annehmen:  $0, 1 \Delta g, 2 \Delta g, \dots, 2^{b-1} \Delta g$ . Hier wird normalerweise  $\Delta g = 1$  angenommen. Wir erhalten so die digitalisierte Funktion (vgl. Fig. 1-3) als eine Folge von natürlichen Zahlen  $k(x_i)=k_i$ . Offensichtlich geht bei diesem Prozess Information verloren. Dies ist einmal der Verlauf der Funktion zwischen den Stützstellen, zum anderen sind die Funktionswerte wegen der Rundungsfehler nicht rekonstruierbar.

Bei der Abtastung einer Funktion  $g(x,y)$ , also etwa eines Bildes, erhalten wir statt der Zahlenfolge ein gitterförmig angeordnetes Schema von (natürlichen) Zahlen, das sich in der Form einer Matrix im Rechner darstellen läßt. Die graphische Darstellung der Abtastpunkte in einem Gitter oder Raster kann auch so interpretiert werden, daß der Funktionswert  $k(x_i,y_i)$  an der Stelle  $(x_i,y_i)$  die Funktion in einem i. a. rechteckigen Bereich approximiert, der in der Fig. 1-4 schraffiert gekennzeichnet ist. Er wird Bildelement oder Pixel genannt (engl. picture element). Es ist auch üblich, auf die Darstellung des Abtastpunktes zu verzichten (vgl. Fig. 1-5) und die Flächenelemente zwischen den Gitterlinien als Bildelemente zu betrachten.

Fig. 1-4

Bildelement und Abtastpunkt  $(2, 1)$

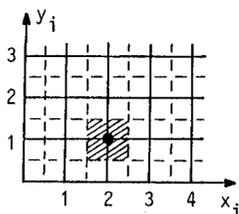
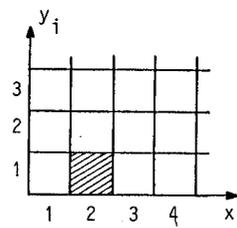


Fig. 1-5

Alternatives Koordinatensystem für die Darstellung digitaler Bilder



Bemerkungen:

- 1.) Das Bildelement ist zu unterscheiden von dem in einem Gerät realisierten und für die Abtastung notwendige Schlitzbreite oder Sensorelementgröße.
- 2.) Manchmal werden die Bildelemente auch hexagonal angeordnet. Dadurch wird ein höherer Grad der Symmetrie und eine bessere Ausnutzung der Information erreicht.

1.2 Unsicherheit bei der Lokalisierung einer Kante

Wegen des Informationsverlustes bei der Digitalisierung läßt sich die ursprüngliche Funktion nicht rekonstruieren, d. h. das gleiche digitale Bild kann verschiedene Objekte darstellen. Sogar wenn das Objekt vollständig bekannt ist und die Digitalisierung ideal erfolgt, kann die Lage des Objekts nicht exakt bestimmt werden. Wir wollen dies am Beispiel eines rampenförmigen Hell-Dunkel-Übergangs in einem Grauwertprofil, einer Kante quantitativ untersuchen.

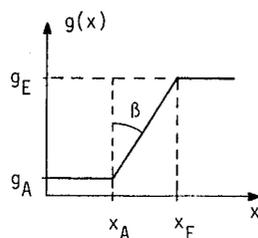


Fig. 1-6 Rampe

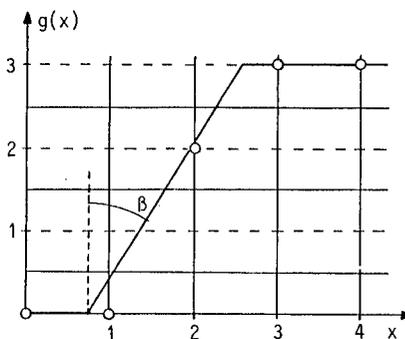


Fig. 1-7 Digitalisierte Rampe ( $x_A=0.75, \tan \beta=3/5, g_A=0, g_E=3$ ),  $o =$  Funktionswerte  $k_i$

In Fig. 1-6 ist die kontinuierliche Form einer Rampe mit dem Neigungswinkel  $\beta$  dargestellt. Fig. 1-7 zeigt die digitalisierte Rampe mit dem Parameter  $t = \tan \beta = 3/5$  für die Neigung. Dabei wurde angenommen, daß die Funktions- bzw. Grauwerte mit 2 bit durch Auf- bzw. Abrunden quantisiert wurden,  $k_i$  also nur die Werte  $0, 1, 2$  oder  $3$  annehmen kann. Der Wertebereich der Funktion ist  $n = 3$ , die Zahl der möglichen Werte von  $k_i$  ist  $N = n + 1 = 4$ . Die digitalisierte Rampe ist die Folge

{..., 0, 0, 2, 3, 3, ...}. Man kann sich den Digitalisierungsvorgang auch so vorstellen, daß an den Stellen  $x_j$  in den Höhen  $g = (2j + 1)/2$  Detektoren angebracht sind (vgl. Fig. 1-8). Es sind die entscheidenden Stellen für die Rundung der Funktionswerte. Die Zahl der Detektoren unter der Rampe ergibt dann die Folge  $\{k_i\}$  der digitalisierten Rampe. In unserem Fall haben wir  $N = 1 = n = 1$  Detektorreihen.

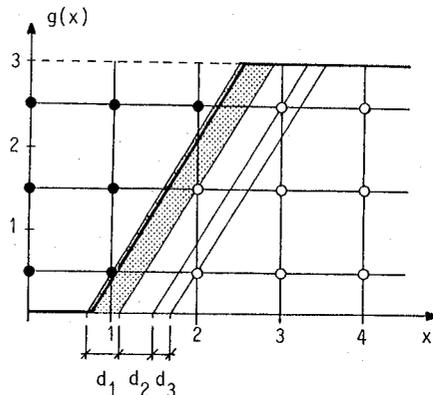


Fig. 1-8 Digitalisierte Rampe mit Detektoren und Ortselementen nach Havelock (1984)

Die Rampe kann nun innerhalb eines gewissen Bereichs nach rechts oder links verschoben werden, ohne an einen Detektor zu stoßen, d. h. ohne die Digitalisierung der Rampe zu ändern. Dieser Schlupfbereich ist im Beispiel der Fig. 1-8 das Intervall  $\bar{d}_1 = |0.7, 1.1|$  der Länge  $\bar{d}_1 = 2/5$ , in dem  $x_A$  variieren kann. Falls wir die Rampe weiter nach rechts verschieben, ändert sich der Funktionswert  $x_2$  um 1 und bleibt unverändert  $x_2 = 1$  bis  $x_A = 1.5$ . Danach ändert sich  $x_3$  um 1 bis  $x_A = 1.7$  erreicht. Wegen der Periodizität des Gitters brauchen wir nur ein Intervall der Länge 1, hier etwa von  $x_A = 0.7$  bis 1.7 untersuchen. Dieses Intervall wird in  $n = 3$  Teilbereiche  $\bar{d}_1, \bar{d}_2$  und  $\bar{d}_3$  zerlegt, für die  $\sum \bar{d}_i = 2/5 + 2/5 + 1/5 = 1$  gilt, und innerhalb derer die Digitalisierung unverändert bleibt. Diese Teilbereiche kann man als "Schlupfbereiche", "Subpixel" oder in Anlehnung an Havelock (1984) als "Ortselemente" (engl. locales) bezeichnen. Havelock hat dieses Konzept der Ortselemente auch für flächenhafte Funktionen entwickelt.

Die Ortselemente  $\bar{d}_i$  hängen bei gegebener Rampenneigung ab von

- 1.) der Zahl  $n$  der wirksamen Detektorreihen oder dem Wertebereich der Funktion, d. h. der Höhe der Rampe,
- 2.) der Neigung  $t = \tan \beta$  der Rampe und
- 3.) der Position  $x_A$  der Rampe:

$$\bar{d}_i = \bar{d}_i(n, t, x_A). \quad (1-1)$$

Zwei Beispiele sollen diese Abhängigkeit verdeutlichen. Für  $t = 0$ , d. h. eine ideale Grauwertstufe ist  $\bar{d}_1 = 1$  und alle übrigen  $\bar{d}_i = 0$ . Der maximale Schlupf ist am größten. Das gleiche gilt für  $t = 1, 2, \dots$  (vgl. Fig. 1-9). Für kleine Werte von  $\tan \beta$ ,  $t = \epsilon < 1/n$  haben  $n - 1$  Ortselemente die Länge  $\epsilon$  und ein Ortselement die Länge  $1 - (n-1)\epsilon$  (vgl. Fig. 1-10).

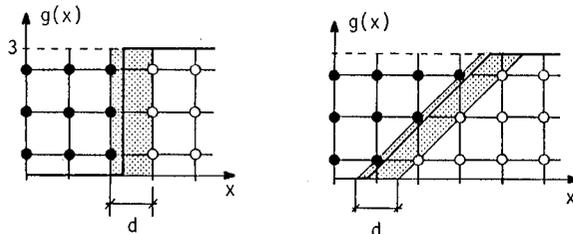


Fig. 1-9 Maximales Ortselement  $\bar{d} = \bar{d}_1 = 1$  bei  $\tan \beta = 0, +1, +2, \dots$

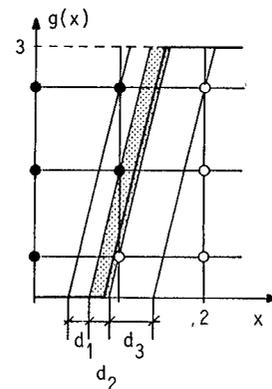


Fig. 1-10 Ortselemente bei steiler Rampe  $t = \epsilon < 1/3$ :  $\bar{d}_1 = \bar{d}_2 = \epsilon$ ,  $\bar{d}_3 = 1 - 2\epsilon$

Aus der Länge der Ortselemente kann man unmittelbar den Fehler bei der Rekonstruktion der Position der Kante ableiten. Wenn wir die Neigung der Rampe als bekannt voraussetzen, den Mittelpunkt eines zu einer Digitalisierung gehörenden Ortselementes als Schätzung  $\hat{x}_A$  für die Position  $x_A$  der Rampe wählen und annehmen, daß alle Werte  $x_A$  eines Ortselementes gleich wahrscheinlich sind, erhalten wir die Streuung

$$\sigma_{\bar{d}_i}(n, t, x_A) = \frac{1}{\sqrt{12}} \bar{d}_i \quad (1-2)$$

für die Genauigkeit dieser Schätzung. Von der Position der Kante unabhängige Aussagen über die Genauigkeit sind die Maximale Größe der Ortselemente

$$d_{\max}(n, t) = \max_i(\bar{d}_i) \quad (1-3)$$

und die mittlere Streuung

$$\sigma_d(n, t) = \sqrt{\sum_i \sigma_{d_i}^2 \bar{d}_i}, \quad (1-4)$$

die sich aus dem mit der Länge der Ortselemente gewogenen Mittel der einzelnen Streuungen ergibt. Die praktische Berechnung dieser Werte kann sich auf die von Dorst und Duin (1984) für dieses Problem entwickelte Spirographentheorie stützen.

In Fig. 1-11 und 1-12 sind für verschiedene Quantisierungen die maximalen und mittleren Fehler für die Positionierung einer Grauertrampe in Abhängigkeit von der Neigung  $t$  dargestellt. Die Funktionen werden mit wachsender Zahl von Quantisierungsstufen immer unregelmäßiger. Die Funktionen  $d_{max}$  sind Polygone, die an den Stellen  $(p/q, 1/q)$  Knicke aufweisen. Dabei sind  $p/q$  alle gekürzten Brüche mit einem Nenner  $q \leq n$ . Daraus folgt, daß die maximale Größe der Ortselemente immer größer als  $1/n$  ist (bei gegebener Neigung). Bei unbekannter Lage einer Rampe, die einen Wertebereich von  $N$  Quantisierungsstufen umfaßt, muß man also mit einer Unsicherheit von mindestens  $1/(N-1)$  bei der Positionierung rechnen. Bei 256 Graustufen ist das  $1/255 = 0.004 \Delta x$ . Der maximale Schlupf ist, unabhängig von der Zahl der Graustufen, gleich 1 und tritt, wie schon oben genannt, bei einem idealen Grauwertsprung auf (Die Funktionen sind periodisch in  $t$  mit einer Periode von  $1$ ). Abgesehen von einzelnen Spitzen bleibt der Maximalfehler aber klein, vor allem bei größerer Zahl von Stufen. Daher ist eine Rampe besser lokalisierbar als ein Sprung, eine Glättung einer Kante vor der Digitalisierung führt daher u. U. zu einer besseren Genauigkeit ihrer Lokalisierung.

Die mittleren Fehler  $\sigma_d$  zeigen ein ähnliches Bild. Mittelt man über alle Neigungen  $t$ , so erhält man mit

$$\bar{\sigma}_d^2(n) = \int_{t=0}^1 \sigma_d^2(n, t) dt \quad (1-5)$$

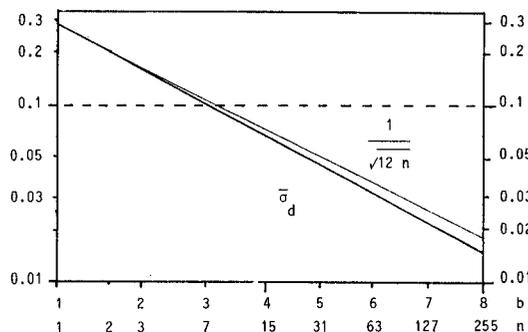
eine Streuung  $\bar{\sigma}_d(n)$  für die Positionierungsgenauigkeit in Abhängigkeit allein von der Höhe  $n$  der Rampe. Das Ergebnis der numerischen Integration ist in Fig. 1-13 in doppelt logarithmischer Skala dargestellt. Es zeigt, daß die Genauigkeit mit zunehmender Wortlänge  $b$  zunimmt. Die Streuung kann sehr gut durch

$$\bar{\sigma}_d(n) \leq \frac{1}{\sqrt{12} n} \Delta x \quad \text{mit} \quad N = n + 1 = 2^b \quad (1-6)$$

approximiert werden. Die Rundungsfehler bei der Quantisierung wirken sich demnach - unabhängig von der Neigung - unmittelbar auf die Genauigkeit der Positionierung aus. Die komplizierten geometrischen Zusammenhänge braucht man bei der praktischen Anwendung i. a. nicht berücksichtigen, sie führen aber dazu, daß die Streuung nur mit  $\sqrt{n}$  abnimmt, statt mit  $n$ , wie zu erwarten wäre. Bei 256 Graustufen ergibt sich  $\bar{\sigma}_d = 0.015 \Delta x \approx 1/60 \Delta x$ , ein extrem guter Wert, wenn man berücksichtigt, daß nur ein Grauwertprofil für die Positionierung verwendet wurde.

Fig. 1-13

Genauigkeit der Positionierung einer digitalisierten Rampe in Abhängigkeit von der Zahl  $N$  der Quantisierungsstufen,  $\Delta x = 1$



### 1.3 Genauigkeit der Positionierung von Kanten und Kreisen in Binärbildern

Die Abschätzungen für rampenartige Grauwertprofile gelten in gleicher Weise für die Positionierung von Kanten in Binärbildern mit nur zwei Graustufen (schwarz/weiß), da man sich die Detektoren in Fig. 1-8 auch in der  $xy$ -Ebene vorstellen kann. Sie zeigen dann, ob sich die Detektoren rechts oder links von der Kante befinden. Gl. 1-6 gibt also auch an, wie genau eine Kante der Längenausdehnung  $n \Delta y$ , gemessen in  $y$ -Richtung, sich in  $x$ -Richtung lokalisieren läßt. Obwohl maximale Fehler von  $1/2 \Delta x$  auftreten können, liegt die Streuung schon bei sehr kurzen Kanten ( $n \geq 8$ ) unter  $1/10 \Delta x$ .

Eine analoge Betrachtung für binär digitalisierte Kreise führt zu einem ähnlichen Ergebnis für die Bestimmung des Mittelpunktes. In Fig. 1-14 ist die Digitalisierung eines Kreises, etwa eines signalisierten Punktes oder einer Rahmenmarke dargestellt. Bildelemente, deren Mittelpunkt im Innern des Kreises liegen haben den Grauwert 1, andere den Grauwert 0. Die Schätzung des Mittelpunktes kann man auf die Randpixel stützen. Es sind die Bildelemente, von denen mindestens ein Nachbar einen anderen Grauwert hat. Die Untersuchung der Ortselemente ist hier zu kompliziert, so daß sich die Abschätzung der Positionierungsgenauigkeit leichter durch Simulation erreichen läßt. Legt man dazu mit einer kleinsten Quadrate Schätzung durch die Randpunkte einen Kreis,

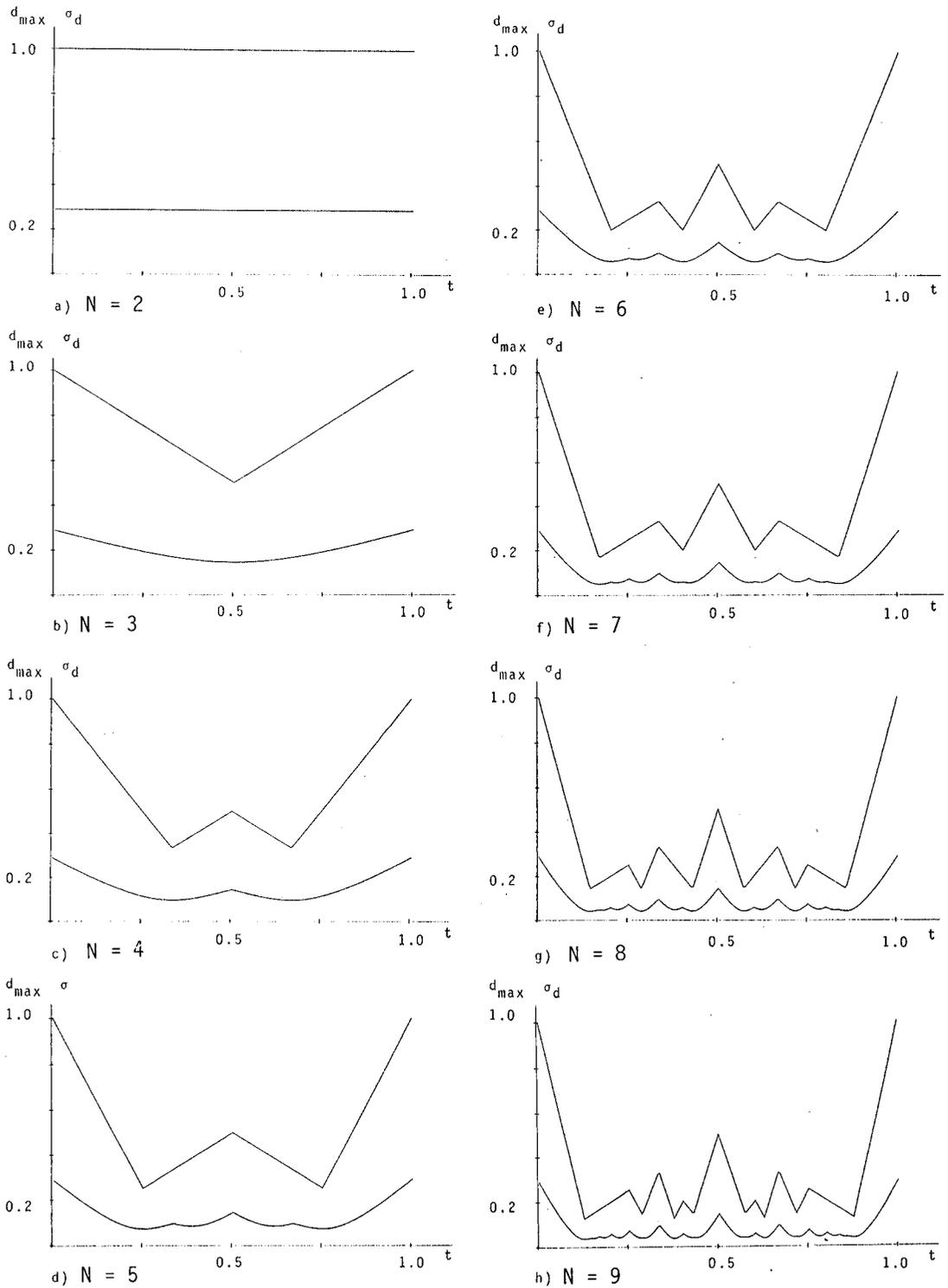


Fig. 1-11 Maximale und mittlere Fehler  $d_{max}$  und  $\sigma_d$  in Pixeleinheiten bei der Lokalisierung ideal abgetasteter Kanten in Abhängigkeit von ihrer Neigung  $t = \tan \beta$  und der Zahl  $N$  der Quantisierungsstufen

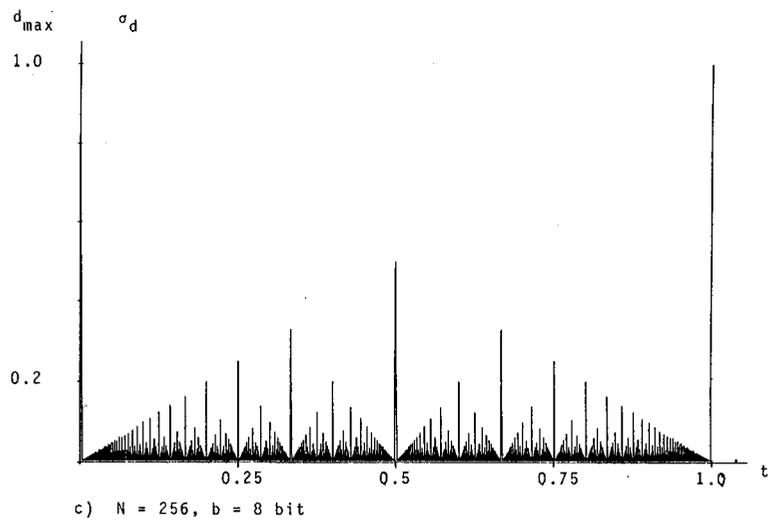
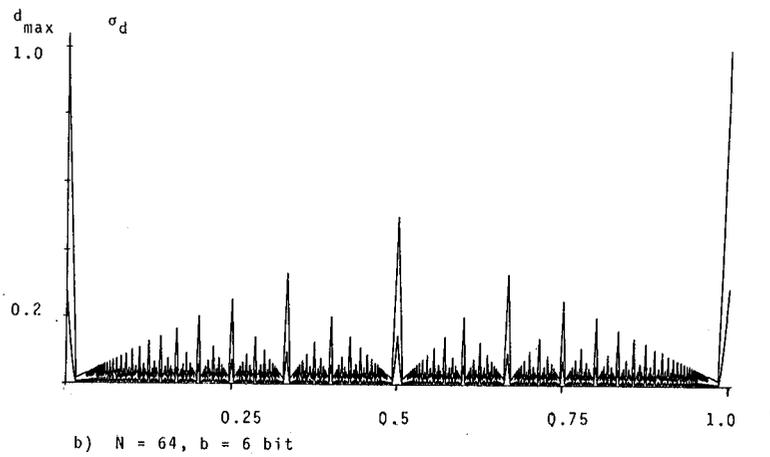
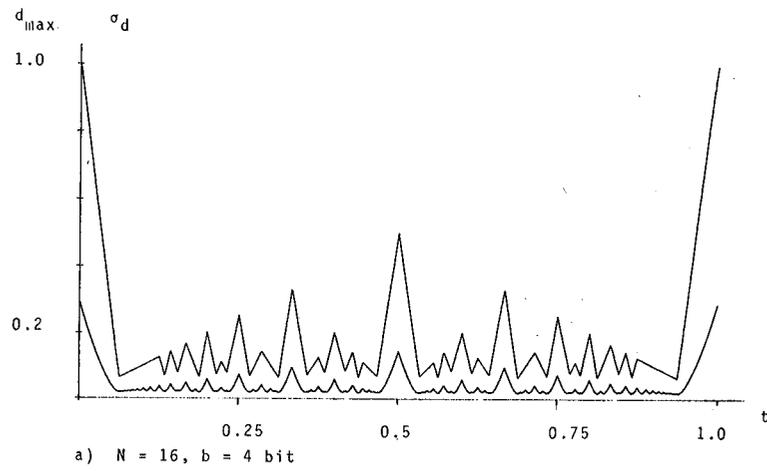


Fig. 1-12 Maximale und mittlere Fehler  $d_{max}$  und  $\sigma_d$  in Pixeleinheiten bei der Lokalisierung ideal abgetasteter Kanten in Abhängigkeit von ihrer Neigung  $t = \tan \beta$  und der Zahl  $N$  der Quantisierungsstufen, bzw. der Zahl der bits  $b$ .

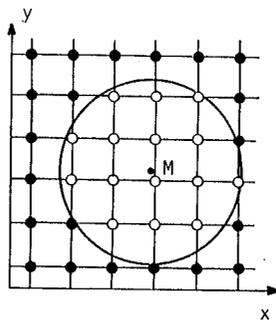


Fig. 1-14 Binär digitalisierter Kreis,  $r = 2.17 \Delta x$

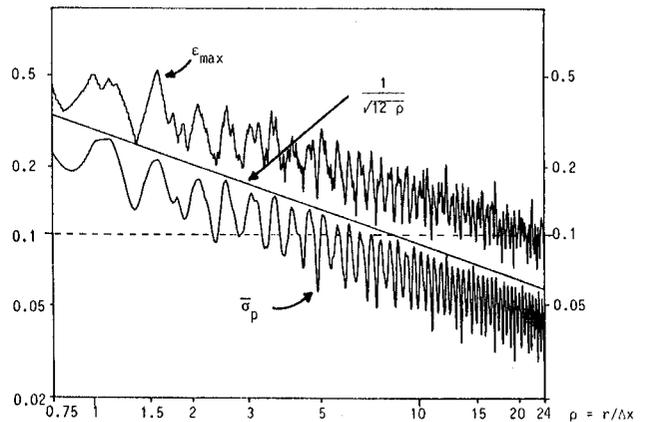


Fig. 1-15 Maximalfehler und Streuung bei der Lokalisierung des Mittelpunkte in binär digitalisierten Kreise in Abhängigkeit vom Radius,  $\Delta x = 1$

erhält man auf einfache Weise eine Schätzung für den Mittelpunkt und den Radius (Anm.: Dies ist nicht notwendigerweise die beste Lösung). Aus den Abweichungen  $\epsilon_x$  und  $\epsilon_y$  zwischen dieser Schätzung und den wahren Werten kann man durch Simulation den Maximalfehler  $\epsilon_{max}$

$$\epsilon_{max}(r) = \max_i (\sqrt{\epsilon_{x_i}^2 + \epsilon_{y_i}^2}) \quad (1-7)$$

und die Punktgenauigkeit  $\sigma_p^2$

$$\sigma_p^2(r) = \sigma_x^2 + \sigma_y^2 = 1/m (\sum_i \epsilon_{x_i}^2 + \epsilon_{y_i}^2) \quad (1-8)$$

ableiten, wenn man den Mittelpunkt  $M_i$  ( $i = 1, \dots, m$ ) innerhalb eines Pixels variiert. Das Ergebnis ist in Fig. 1-15 dargestellt. Auch hier fällt der unregelmäßige Verlauf der Funktionen  $\epsilon_{max}(r)$  und  $\sigma_p(r)$  auf. Die Maximalfehler liegen, abgesehen von sehr kleinen Kreisen ( $r < 2 \Delta x$ ) unter  $1/2 \Delta x$ . Im Gegensatz zur Rampe fallen die Maximalfehler mit wachsendem Radius. Der Punktfehler läßt sich wieder, für den gesamten untersuchten Bereich (bis  $r = 24 \Delta x$ ) durch

$$\sigma_p(\rho) \leq \frac{1}{\sqrt{12} \rho} \Delta x \quad \text{mit} \quad \rho = r / \Delta x \quad (1-9)$$

approximieren, völlig analog zu Gl. 1-6.

**Beispiel:** Rauschfreie Daten vorausgesetzt, kann man einen binär digitalisierten Punkt mit einem Durchmesser von ca. 2 Pixeln mit einer Genauigkeit von etwa  $1/4$  Pixel lokalisieren. Man muß dabei mit einem Maximalfehler unter  $1/2$  Pixel rechnen. Eine binär digitalisierte Rahmenmarke mit einem Durchmesser von 10 Pixeln kann man unter idealen Bedingungen mit einem mittleren Punktfehler von etwa  $1/6$  Pixel positionieren. Bei der Verwendung von Grautonbildern sind die Genauigkeiten wesentlich höher und erreichen ohne weiteres die Streuungen bei der Kantenpositionierung; Die Genauigkeit läßt sich also durch Erhöhung der Zahl der Graustufen wesentlich leichter verbessern als durch Verkleinern der Pixelgröße. Andererseits bildet die Pixelgröße bei der üblichen Abtastung mit 256 Graustufen die eigentliche Begrenzung in der Genauigkeit. Dieses Ergebnis hat schon Havelock (1984) gezeigt. ■

Die in diesem Abschnitt zusammengestellten Ergebnisse sollten zeigen, wo die Leistungsgrenzen der Positionierung von Objekten und damit der Zuordnung liegen. Die zugrunde gelegten Annahmen sind idealisiert und wegen der schon bei so einfachen Objekten wie einer Kante komplexen Abhängigkeiten nicht ohne weiteres verallgemeinerbar. Das Potential des Konzepts der Ortselemente von Havelock (1984) ist allerdings noch nicht ausgeschöpft.

Bei der praktischen Anwendung werden die genannten hohen Genauigkeiten wegen der unvermeidlichen zufälligen Fehler beim Abtasten der Bilder nicht erreicht. Zufällige Fehler wurden in diesem Abschnitt auch nicht angenommen. Darüber hinaus gibt die zugrunde gelegte Theorie keine Anhaltspunkte für ein praktisches Lokalisierungsverfahren. Deshalb wurden im nächsten Abschnitt die klassische Korrelation und die mit ihr verwandte Kleinste Quadrate Anpassung besprochen, die für die Zuordnung sehr viel allgemeinerer Objekte geeignet sind und auch bei der Beurteilung den Einfluß zufälliger Fehler berücksichtigen können.

## 2. Korrelation und Kleinste Quadrate Schätzung für die Zuordnung von Bildausschnitten

### 2.1 Mathematisches Modell

Wir unterstellen im folgenden, daß die Bilder perspektive Abbildungen eines dreidimensionalen Objekts sind. I. a. ist dann der geometrische Zusammenhang zwischen den Bildern wegen möglicher Verdeckungen äußerst komplex und weder stetig noch glatt, etwa an Kanten oder Rändern von Gebäuden.

Wir wollen uns für den Rest diese Beitrags auf die Auswertung von Objekten oder Objektteilen mit glatter Oberfläche beschränken. Damit wird gewährleistet, daß die perspektive Abbildung zwischen dem Objekt und den Bildern als auch die durch die Oberfläche bestimmte Abbildung zwischen den Bildern für kleine Bildausschnitte in guter Näherung durch eine affine Abbildung beschrieben werden kann. Für die praktische Anwendung bedeutet das, daß bei kantigen Objekten genügend Textur zwischen den Kanten für die Stereoauswertung vorhanden sein muß. Inwieweit die Textur für eine hinreichend gute Zuordnung ausreicht, wird in Abschnitt 2.4 behandelt.

Damit ergibt sich die Situation von Fig. 2-1. Ein Objektpunkt  $P$  bildet sich in den Bildpunkten  $P_1$  und  $P_2$  ab. Die für die Zuordnung notwendigen Umgebungen der Bildpunkte werden als Grauwertmatrix dem Rechner zur Verfügung gestellt. Die Lage der Bildausschnitte muß aus Näherungskoodinaten abgeleitet werden. Woher diese stammen, wollen wir hier nicht diskutieren. Wie genau sie sein müssen, hängt vom verwendeten Verfahren für die Zuordnung ab (s. u.). Die Grauwertmatrizen

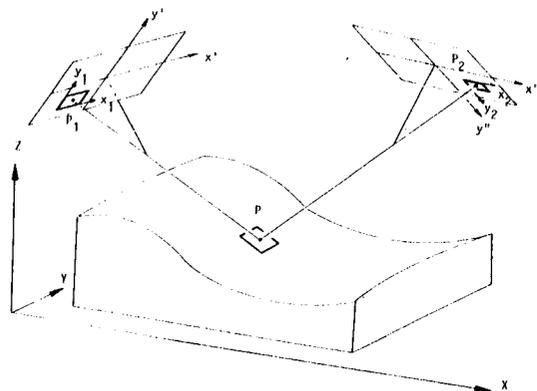


Fig. 2-1 Geometrisches Modell für die Zuordnung

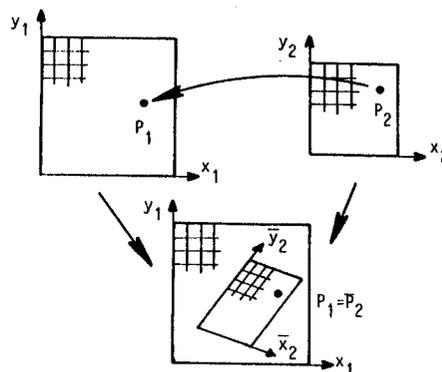


Fig. 2-2 Zur geometrischen Transformation der Bildausschnitte

enthalten die Grauwerte in einem lokalen Koordinatensystem. Es muß durch geeignete Kalibrierung auf das Bildkoordinatensystem bezogen werden (vgl. Gülch, 1985).

Die zentrale Aufgabe besteht darin, zwei zugeordnete Punkte in den beiden Bildausschnitten zu finden. Sie sind in Fig. 2-2 dargestellt. Die geometrische Transformation  $T_g: z_1 = T_g(z_2)$  der Koordinaten  $z_1 = (x_1, y_1)'$  und  $z_2 = (x_2, y_2)'$  des rechten und linken Bildausschnittes sei linear:

$$T_g: \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (2-1a)$$

bzw.

$$T_g: z_1 = A z_2 + a \quad (2-1b)$$

Falls dieses Modell streng erfüllt ist, kann man mit den Transformationsgleichungen zu jedem Bildpunkt im rechten Bildausschnitt 2 den zugehörigen Punkt im linken Bildausschnitt 1 bestimmen, insbesondere für den Mittelpunkt, den gewogenen Schwerpunkt (s. u.) oder einen anderen ausgezeichneten Punkt, etwa die Mitte eines Signals. Die Wahl diese Übertragungspunktes ist von der jeweiligen Aufgabenstellung abhängig. Man erhält so zwei Koordinatenpaare  $(x_1^i, y_1^i)$  und  $(x_2^i, y_2^i)$  von Bildpunkten im linken und rechten Bildausschnitt. Sie entsprechen, nach Berücksichtigung der (Sensor-) Kalibrierung den Bildkoordinaten in einem Stereokomparator nach dem Wegstellen der x- und y-Parallaxen und gegebenenfalls dem Einstellen eines ausgezeichneten Punktes mit der Meßmarke. Die dreidimensionalen Koordinaten des Objektpunktes erhält man in bekannter Weise durch Vorwärtsschnitt, wenn die Orientierung der Bilder bekannt ist. Außerdem liegen, falls das Modell streng erfüllt ist, alle mit Gl. 2-1 zugeordneten Punkte auf der Objektfläche, einem Ebenenstück. Aus den Parametern  $A$  der Affintransformation kann man zusätzlich den Oberflächengradient ableiten. Mit einem Zuordnungsalgorithmus, der die Parameter  $A$  und  $a$  bestimmt, lassen sich gleichzeitig Position und Neigung messen!

Neben den geometrischen Differenzen weisen die Bildausschnitte radiometrische Unterschiede auf, wegen geringen Fehlern bei der Filmentwicklung oder der Sensorkalibrierung, etc.. Wir wollen hier eine lineare radiometrische Transformation  $T_r: g_1 = T_r(g_2)$  zwischen zugeordneten Grauwerten  $g_1$  und  $g_2$  annehmen:

$$T_r: g_1 = B g_2 + b \quad (2-2)$$

Die Aufgabe besteht nun darin, die Parameter  $A$ ,  $a$ ,  $B$  und  $b$  so zu bestimmen, daß das transformierte Bild  $g_2$  dem Bild  $g_1$  möglichst ähnlich ist.

## 2.2 Korrelation für die Zuordnung von Bildausschnitten

Das klassische Verfahren der Korrelation geht von Gl. 2-2 aus. Die Grauwerte  $g_1$  und  $g_2$  werden als Realisierungen von Zufallsvariablen betrachtet. Falls die wahren Funktionen  $g_1$  und  $g_2$  linear abhängig sind, ist der Korrelationskoeffizient  $\rho_{12} = 1$ . Aus wenigstens zwei Gründen ist i. a.

$\rho_{12} < 1$ :

- a.) Die gemessenen Grauwertfunktionen sind durch den physikalischen Abbildungsvorgang und durch die Abtastung mit zufälligen Fehlern behaftet.
- b.) Die geometrische Transformation zwischen den Bildern ist nicht optimal.

zu a.)

Der Zusammenhang zwischen erreichbarem Korrelationskoeffizient  $\rho_{12}$  und der Streuung  $\sigma_n$  des Meßrauschens ist unter geringen Einschränkungen einfach herzustellen. Die Grauwertfunktionen  $g_1$  und  $g_2$  gehen aus der wahren Grauwertfunktion  $g$  durch individuelles Rauschen, Kontrast- und Helligkeitsänderung hervor:

$$\begin{aligned} g_1 &= B_1 (g + n_1) + b_1 \\ g_2 &= B_2 (g + n_2) + b_2 \end{aligned} \quad (2-3)$$

Die Streuungen  $\sigma_{n_1}$  und  $\sigma_{n_2}$  seien identisch und gleich  $\sigma_n$ . Das Signal-Rausch-Verhältnis  $SNR$  (signal to noise ratio)

$$SNR = \frac{\sigma_g}{\sigma_n} \quad (2-4)$$

ist umgekehrt proportional zur Relativgenauigkeit der Grauwertmessungen. Dann erhalten wir

$$\rho_{12} = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_n^2} = \frac{SNR^2}{SNR^2 + 1} \quad (2-5)$$

bzw.

$$SNR^2 = \frac{\rho}{1 - \rho}$$

Da der cosinus des Winkels  $\phi$  zwischen den Vektoren  $g_1$  und  $g_2$  gleich dem Korrelationskoeffizient ist, haben wir insgesamt drei äquivalente Maße für die Ähnlichkeit zweier Bildausschnitte.

Beispiele:

$$\begin{aligned} SNR = 10, & \quad \rho = 0.99, & \quad \phi = 8^\circ \\ SNR = 3.4, & \quad \rho = \sqrt{2}/2, & \quad \phi = 45^\circ \\ SNR = 1, & \quad \rho = 0.5, & \quad \phi = 60^\circ \end{aligned}$$

Die Werte  $\rho = \sqrt{2}/2 = 0.71$  und  $\rho = 0.5$  sind demnach plausible und anschaulich-kritische Werte für den empirischen Korrelationskoeffizient  $\rho_{12}$ . Bei Werten  $\rho_{12} \leq 0.71$  oder gar  $0.5$  kann man die Zuordnung der Bildausschnitte als erfolglos betrachten.

zu b.)

Das praktische Verfahren der Korrelation sucht den größten Korrelationskoeffizienten, indem die Parameter der geometrischen Transformation variiert werden. Die zum größten Korrelationskoeffizienten gehörenden Parameter sind beste erwartungstreue Schätzungen. Da der Suchraum bei der o. g. Affintransformation sehr groß ist, beschränkt man sich i. a. auf die Bestimmung der Transformationsparameter  $a_1$  und  $a_2$ , d. h. der Parallaxen. Die Verzerrungen werden so entweder ganz vernachlässigt, man nimmt also näherungsweise horizontales Gelände an ( $A = I$ ), oder der eine der beiden Bildausschnitte wird, auf der Grundlage geeigneter Näherungswerte, vorentzerrt. Auf den Einfluß der Qualität dieser Näherungswerte auf die Korrelationsgenauigkeit geht Abschnitt 2.7 ein.

Bei allgemeiner Orientierung der Bilder und insbesondere bei Messung von Einzelpunkten muß man bei diesem Verfahren mit Ungenauigkeiten rechnen, da keine Näherungswerte für die relative Entzerrung der Bildausschnitte zur Verfügung steht. Darüberhinaus bietet das Verfahren nicht unmittelbar eine Angabe über die Genauigkeit der Zuordnung im Sinne einer Standardabweichung. Das im nächsten Abschnitt beschriebene Verfahren behebt diese Nachteile.

## 2.3 Kleinste Quadrate Schätzung für die Zuordnung von Bildausschnitten

Als Ähnlichkeitsmaß zwischen den Bildausschnitten kann man auch die Quadratsumme der Grauwertdifferenzen wählen. Man erhält dann folgendes Optimierungsproblem:

$$\Phi(A, a, B, b) = \sum_{z_2} \{ (B g_2(z_2) + b) - g_1(A z_2 + a) \}^2 \rightarrow \min. \quad (2-6)$$

Alle Bildelemente  $g_2(z_2)$  im rechten Bild 2 werden radiometrisch dem linken Bild angepaßt ( $Bg_2+b$ ) und mit den durch die Transformation  $\bar{z}_2 = A z_2 + a$  zugeordneten Helligkeitswerten  $g_1(z_2)$  verglichen. Da  $g_1$  nur an diskreten, ganzzahligen Stellen definiert ist - eigentlich müßte man  $k_1(z_{1i})$  schreiben - die transformierten Koordinaten  $\bar{z}_2 = (\bar{x}_2, \bar{y}_2)$  i. a. nicht ganzzahlig sind, muß der Wert  $g_1(\bar{z}_2)$  durch geeignete, etwa bilineare Interpolation aus den umliegenden Grauwerte ermittelt werden (resampling).

Gl. 2-6 ist wegen der Funktion  $g_1$  hochgradig nichtlinear, kann aber an Näherungswerten linearisiert und dann mit Standardmethoden der Ausgleichsrechnung gelöst werden. Das Verfahren ist iterativ. Die Schätzungen sind, wenn man unkorrelierte Messungen voraussetzt, beste erwartungstreue Schätzungen für die Transformationsparameter. Aus der Inversen der Normalgleichungen und der Schätzung des Varianzfaktors  $\sigma_0^2$  läßt sich die Genauigkeit der Parameter abschätzen. Insbesondere kann man die mittleren Fehler  $\sigma_x$  und  $\sigma_y$  für ein zugeordnetes Paar von Punkten  $P_1$  und  $P_2$  ableiten. Die Streuungen stellen die Genauigkeit der x- und y-Parallaxen zwischen den Punkten dar. Fehler-, Normalgleichungen und Genauigkeitsangaben werden im folgenden Abschnitt diskutiert.

Bemerkungen:

- 1.) Das Verfahren liefert für den Fall, daß keine Verzerrungen zwischen den Bildausschnitten angenommen werden praktisch die gleichen Ergebnisse wie die Korrelation (Abschn. 2.2). Die Unterschiede sind vernachlässigbar. Daher sind die Genauigkeit von Korrelation und Kleinste Quadrate Anpassung als Zuordnungsverfahren gleich, falls man sich auf die Bestimmung von Parallaxen beschränkt und Kontrast- und Helligkeitsunterschiede der Bildausschnitte kompensiert.
- 2.) Die Lösung von Gl.2-6 kann man auch zur Lokalisierung von Objekten verwenden. Dann muß eine der beiden Funktionen vorgegeben werden. Die geschätzten Parameter beschreiben dann z. B. die geometrische und radiometrische Transformation vom Gelände ins Bild.
- 3.) Mit dem gleichen Ansatz läßt sich auch die Punktübertragung in mehrere Bilder, etwa bei der Aerotriangulation, realisieren. Nach Auswahl geeigneter Punkte, d.h. geeigneter Grauwertmatrizen, kann man die Punkte mit Gl. 2-6 in alle Nachbarbilder  $g_i$  ( $i = 2, 3, \dots$ ) übertragen (vgl. Pertl 1985).
- 4.) Das geometrische Modell Gl.2-1 geht von unbekannter relativer Orientierung der Bilder und beliebiger Geländeneigung aus, wie es etwa bei der Parallaxenmessung gerade für die relative Orientierung notwendig ist. Liegen die Parameter der relativen Orientierung vor, kann man die Bedingungen, die sich aus der Epipolargeometrie ergeben, berücksichtigen. Statt 6 geometrischer Parameter sind dann nur noch 3 Parameter, nämlich die x-Parallaxe, der Maßstab in Kernstrahlrichtung und die

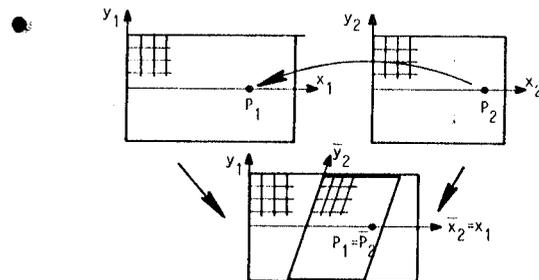


Fig. 2-3 Geometrisches Modell für die flächenhafte Zuordnung von Bildausschnitten eines Normalbildes nach der relativen Orientierung (vgl. Kreiling 1976, Berücksichtigung der Epipolargeometrie):  $x_1 = a_{11} x_2 + a_{12} y_2 + a_1$ ,  $y_1 = y_2$  (!).

Scherung (vgl. Fig. 2-3), zu bestimmen, wenn man flächenhafte Bildausschnitte zuordnet. Die 3 Parameter entsprechen der unbekannt Distanz (Höhe) des Punktes und den beiden Geländeneigungen.

5.) Die verschiedenen Verfahren unterscheiden sich neben dem mathematischen Modell vor allem darin, wie die bei der Linearisierung anfallenden Grauwertgradienten  $\partial g/\partial x$ , ggf.  $\partial g/\partial y$  bestimmt werden. Liegt die Sollform der Grauwertfunktion, wie bei der Objektlokalisierung vor, so verwendet man die daraus ableitbaren Sollgradienten. In allen anderen Fällen muß man die Gradienten aus den i. a. verrauschten Grauwertfunktionen  $g_1$  und  $g_2$ , bei der Mehrbildkorrelation (vgl. Pertl 1985) auch weiterer  $g_i$ , schätzen. Dazu ist es sinnvoll und praktisch hinreichend, wenn man das Mittel aus den Gradienten verwendet, die sich nach einer geeigneten Glättung der Grauwertfunktionen zur Rauschunterdrückung ergeben.

Mit der Darstellung der Zuordnung als Schätzproblem i. S. der Statistik stehen Verallgemeinerungen und Modifikationen aller Art zur Verfügung. Wir wollen die Eigenschaften der Zuordnung näher untersuchen. Der Übersichtlichkeit halber beschränken wir uns auf die Lokalisierung von Objekten, bei der nur die Lage des Objekts unbekannt ist.

2.4 Präzision der Objektlokalisierung

2.4.1 Schätzung der Parameter

Wir untersuchen zunächst die Präzision der Verschiebungsparameter einer Objektlokalisierung. Das mathematische Modell lautet hier, mit der als unbekannt angenommenen Funktion  $g$ :

$$\underline{g}_i(x_i, y_i) = g(x_i - x, y_i - y) + \underline{n}(x_i, y_i), \quad i = 1, 2, \dots, n \quad (2-7)$$

Gl. 2-7 gilt für alle  $n$  Bildelemente, die nicht in einem Raster angeordnet zu sein brauchen. Stochastische Größen sind unterstrichen.

Die Idee, Gl. 2-7 zu linearisieren um zu einer Schätzung  $\hat{x}$  und  $\hat{y}$  für die Verschiebungsparameter zu kommen, ist nicht neu (cf. Limb/Murphy 1975, Burckhardt/Moll 1979, Wild 1979, Meyers/Frank 1980, Thurgood/Mikhail 1982). Sie wird seit etwa 10 Jahren in der Bildfolgenanalyse (Cafforio/Rocca 1976, Schalkoff/McVey 1978, Fennema/Thompson 1979, Dinse et. al. 1981, Huang 1981) und jüngst auch in der Photogrammetrie zur Objektlokalisierung und Punktübertragung verwendet (Ackermann/Pertl 1982, Förstner 1982, Mikhail 1983, Grün 1985)

Geht man von Näherungswerten  $x_0$  und  $y_0$  aus und setzt  $\underline{\Delta g}_i = \underline{\Delta g}(x_i, y_i) = g_1(x_i, y_i) - g(x_i - x_0, y_i - y_0)$  und  $\underline{v}_i = \underline{v}(x_i, y_i) = -\underline{n}(x_i, y_i)$ , so kann man die unbekanntes Verschiebungen aus den Fehlergleichungen

$$\underline{\Delta g}_i + \underline{v}_i = g_{x,i} \hat{x} + g_{y,i} \hat{y}, \quad i = 1, 2, \dots, n \quad (2-8)$$

bestimmen. Hier ist vorübergehend  $x_0 = y_0 = 0$  gesetzt. Die Grauwertgradienten  $g_x = \partial g / \partial x$  und  $g_y = \partial g / \partial y$  sind an den entsprechenden Näherungswerten zu bestimmen. Anschaulich bedeutet Gl. 2-8 daß kleine Verschiebungen  $\hat{x}$  und  $\hat{y}$  in  $x$  und  $y$  entsprechend der Steilheit ( $g_x, g_y$ ) der Grauwertfunktion zu Änderungen  $\underline{\Delta g}$  führen. Wenn wenigstens zwei Gleichungen der Art 2-8 vorliegen, kann man im Prinzip aus den beiden Grauwertdifferenzen die Verschiebungen schätzen. Das i. a. aber stark überbestimmte Gleichungssystem 2-8 führt zu den Normalgleichungen

$$\begin{pmatrix} \Sigma g_x^2 & \Sigma g_x g_y \\ \Sigma g_y g_x & \Sigma g_y^2 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} \Sigma g_x \underline{\Delta g} \\ \Sigma g_y \underline{\Delta g} \end{pmatrix} \quad (2-9)$$

wenn man das Meßrauschen  $\underline{n}$  als unkorreliert annimmt und  $\sigma_n = \sigma_0$  setzt. In Gl. 2-9 kann man die Summen durch die entsprechenden Varianzen und Kovarianzen ersetzen, etwa  $\Sigma g_x g_y = n \sigma_{xy}$  (Mittelwert = 0 vorausgesetzt) und erhält

$$n \begin{pmatrix} \sigma_x^2 & \sigma_{g_x g_y} \\ \sigma_{g_y g_x} & \sigma_y^2 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = n \begin{pmatrix} \hat{g}_x \underline{\Delta g} \\ \hat{g}_y \underline{\Delta g} \end{pmatrix} \quad \text{oder} \quad N \hat{z} = \underline{h} \quad (2-10)$$

Die Schätzungen aus Gl. 2-10 sind beste erwartungstreue Schätzer. Mit Gl. 2-8 erhält man dann eine Schätzung  $\hat{\sigma}_n^2$  für das Meßrauschen

$$\hat{\sigma}_n^2 = \frac{\Sigma \underline{\Delta g}^2}{n - 2} \quad (2-11)$$

Schließlich ist die theoretische Kovarianzmatrix  $C_{zz}$  der unbekanntes Verschiebungen  $\underline{z}$

$$C_{zz} = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{pmatrix} = \sigma_n^2 Q = \frac{\sigma_n^2}{n} \begin{pmatrix} \sigma_{g_x}^2 & \sigma_{g_x g_y} \\ \sigma_{g_y g_x} & \sigma_{g_y}^2 \end{pmatrix}^{-1} \quad (2-12a)$$

Mit der Schätzung  $\hat{\sigma}_n^2$  aus Gl. 2-11 ergibt sich die geschätzte Kovarianzmatrix zu

$$\hat{C}_{zz} = \frac{\hat{\sigma}_n^2}{\sigma_n^2} C_{zz} \quad (2-12b)$$

#### 2.4.1 Genauigkeit der geschätzten Verschiebungen

Die Genauigkeit der geschätzten Verschiebungen hängt von drei Parametern ab

- der Streuung  $\sigma_n$  der zufälligen Fehler
- der Zahl  $n$  der verwendeten Bildelemente und
- den Streuungen  $\sigma_{g_x}$  und  $\sigma_{g_y}$  und der Korrelation der Grauwertgradienten

Um einen Vergleich zwischen den Streuungen  $\sigma_x$  bzw.  $\sigma_y$  und dem Korrelationskoeffizienten  $\rho_{12} = \rho$  als Qualitätskriterium herzustellen, nehmen wir vorübergehend an, daß die Grauwertgradienten gleiche Streuung haben und unkorreliert sind. Wir erhalten dann mit  $SNR^2 = \sigma_g^2 / \sigma_n^2 = \rho / (1 - \rho) z$ , B. für die Varianz  $\sigma_x^2$

$$\sigma_x^2 = \frac{\sigma_n^2}{\Sigma g_x^2} = \frac{1}{n} \frac{\sigma_n^2}{\sigma_x^2} = \frac{1}{n} \frac{1}{SNR^2} \frac{\sigma_g^2}{\sigma_x^2} = \frac{1}{n} \frac{1 - \rho}{\rho} \frac{\sigma_g^2}{\sigma_x^2} \quad (2-13)$$

Die Beurteilung der Zuordnung nach Gl. 2-13 berücksichtigt über den Korrelationskoeffizienten hinaus mit der Zahl der Bildelemente die Größe und mit der im Quotienten  $\sigma_{g_x}^2 / \sigma_g^2$  enthaltenen Bandbreite die Textur des verwendeten Bildausschnittes. Die Standardabweichungen bei der Zuordnung zweier abgetasteter Bildausschnitte sind um den Faktor  $\sqrt{2}$  größer.

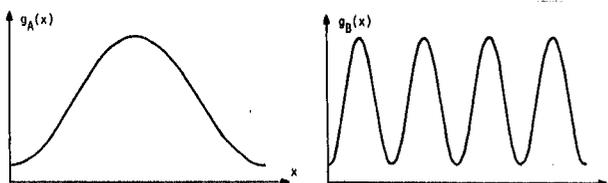
**Bemerkungen:**

- 1.) Der geschätzte Wert  $\hat{\sigma}_n$  aus Gl. 2-11 enthält alle verbleibenden Fehler des mathematischen Modells. Insbesondere enthält er evtl. nicht modellierte geometrische Verzerrungen (nichtlinearer Art) oder nicht erfaßte nichtlineare radiometrische Differenzen zwischen den Bildern. Auch echte Störungen in einem der beiden Bildausschnitte zeigen sich in der Schätzung für das Meßrauschen. Ausgenommen sind systematische Fehler, wie etwa Beleuchtungseffekte, die eine reine Verschiebung des Objekts verursachen. Sie bleiben unerkannt.
- 2.) Da die Streuungen der Verschiebungsparameter mit der Quadratwurzel der Zahl der Bildelemente fallen, kann man im Prinzip durch Vergrößern der Bildausschnitte die Genauigkeit beliebig steigern solange die Varianz  $\sigma_{g_x}^2$  der Grauwertfunktion konstant bleibt, d. h. solange durch Vergrößerung des Bildausschnittes mehr Struktur für die Zuordnung zur Verfügung steht. Einer beliebigen Steigerung der Genauigkeit sind aber wegen der i. a. nicht ebenen Oberfläche Grenzen gesetzt (vgl. Abschnitt 2.7). Doch erklärt sich daraus die im Vergleich zum Operateur geringere Genauigkeit derjenigen Verfahren, die sich bei der Zuordnung auf Bildelemente innerhalb eines Kernstrahls beschränken. Demnach sollte man auch dann eine flächenhafte Zuordnung vorsehen, wenn die Epipolargeometrie zur Verringerung des numerischen Aufwandes genutzt wird.
- 3.) Gl. 2-13 zeigt deutlich, daß die Genauigkeit nicht vom Kontrast sondern - wie zu erwarten - von der Schärfe ( $\sigma_{g_x}$ ) des abgebildeten Objekts abhängt. Zwar kann man etwa durch Verdoppeln des Kontrastes bei gleichem Rauschen die Genauigkeit verdoppeln, da dann auch die Gradienten im Bild größer werden. Aber das Objekt B in Fig. 2-4 läßt sich wegen der steileren Kanten mit 4-facher Genauigkeit lokalisieren wie das Objekt A, obwohl es den gleichen Kontrast aufweist.

Fig. 2-4

Objekte gleichen Kontrasts  
 aber verschiedener Schärfe:

$$\sigma_x(B) = \frac{1}{4} \sigma_x(A)$$



4.) Die Kovarianzmatrix spiegelt in Form der zugehörigen Fehlerellipse unmittelbar wider, ob das Objekt nach allen Richtungen gleich gut lokalisierbar ist.

**Beispiel:**

Für die in Fig. 2-5 dargestellten Punkte ergaben sich bei der Zuordnung die Werte aus Tab. 2-1. Die theoretischen Fehlerellipsen mit  $\sigma_0 = 5$  Grauwerte zeigen deutlich den Einfluß der Form und der Schärfe der Punkte auf die Form und Größe der Fehlerellipse.

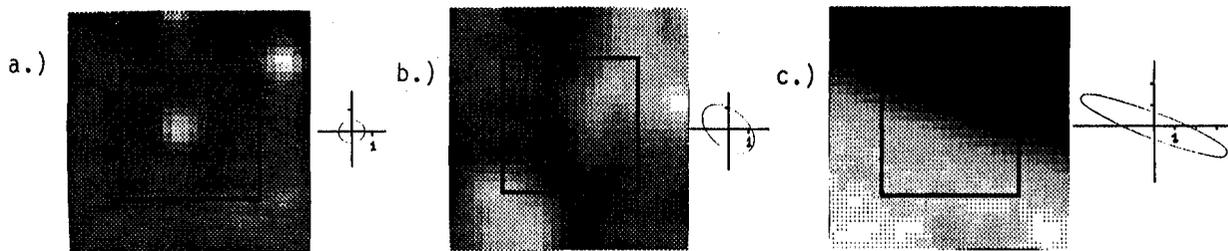


Fig. 2-5 Künstlicher (a.) und natürliche (b., c.) Punkte mit Fehlerellipsen der Parallaxenmessung,  $g_2$  nicht dargestellt,  $n = 16 \times 16$  Pixel,  $\Delta x = 20 \mu m$

	$\rho$	$\hat{\sigma}_n$  gr	$Q$   $10^{-4}(pel/gr)^2$	Streuungen	Halbachsen ( $\sigma_n = 5 gr$ )
a	.955	6.7	$\begin{vmatrix} .3552 & -.0165 \\ -.0165 & .2889 \end{vmatrix}$	$\hat{\sigma}_{-x} = .040 pel = 0.8 \mu m$ $\hat{\sigma}_y = .036 pel = 0.7 \mu m$	$a = .60 \mu m$ $b = .53 \mu m$
b	.991	4.4	$\begin{vmatrix} 1.442 & .574 \\ .574 & 1.295 \end{vmatrix}$	$\hat{\sigma}_x = .053 pel = 1.1 \mu m$ $\hat{\sigma}_y = .050 pel = 1.0 \mu m$	$a = 1.4 \mu m$ $b = 0.9 \mu m$
c	.996	6.6	$\begin{vmatrix} 2.05 & 4.35 \\ 4.35 & 11.81 \end{vmatrix}$	$\hat{\sigma}_x = .095 pel = 1.9 \mu m$ $\hat{\sigma}_y = .227 pel = 4.5 \mu m$	$a = 3.7 \mu m$ $b = 0.6 \mu m$

Tab. 2-1 Übertragungsgenauigkeit für die Punkte der Fig. 2-5, Bildelementgröße  $20 \mu m$   
 $pel = \Delta x$

Insbesondere zeigt die Fehlerellipse bei der Kante (Fig. 2-5c), daß erwartungsgemäß eine Zuordnung nur quer zur Kante mit akzeptierbarer Genauigkeit möglich ist. Da die Kante nicht ideal ist, sondern kleine Unregelmäßigkeiten aufweist, ist auch eine Zuordnung längs der Kante möglich - wenn auch mit deutlich größerer Unsicherheit. Im Extremfall kann bei Kanten die Normalgleichungsmatrix singular werden, was i. a. auf eine linienhafte Struktur im Bild hinweist. Wird dagegen die Epipolargeometrie genutzt, genügt es, wenn das Objekt eine genaue Parallaxenmessung entlang der Kernlinie zuläßt - eine wesentlich schwächere Forderung. ■

5.) Da innerhalb eines Bildes das Meßrauschen in erster Näherung konstant ist, kann man die zu erwartende Genauigkeit für die Zuordnung eines Punktes aus der Grauwertverteilung eines Bildes ableiten. Dies hat zwei wichtige Konsequenzen:

- a.) Man kann vor der eigentlichen rechenaufwendigen Zuordnung prüfen, ob eine Messung mit der geforderten Genauigkeit überhaupt möglich ist.
- b.) Man kann - falls es die Aufgabenstellung erlaubt - automatisch Punkte auswählen, die eine gute Parallaxenmessung garantieren. Wir kommen darauf im 3. Abschnitt zurück.

Im folgenden wird der Einfluß verschiedener Modellfehler auf die Genauigkeit der Zuordnung von Bildausschnitten untersucht. Hierbei zeigt sich die Leistungsfähigkeit des Kleinsten Quadrate Ansatzes.

### 2.5 Einfluß von Filtern

Filter werden in der digitalen Bildverarbeitung vor allem zur Bildverbesserung eingesetzt. Beispiele sind Filter zur Rauschunterdrückung oder zur Verstärkung von Kanten. Sie führen dazu, daß das ursprünglich unkorreliert angenommene Meßrauschen korreliert ist und gleichzeitig das Signal verändert wird. Wir wollen an drei Beispielen den Einfluß von Filtern auf die Genauigkeit der Parallaxenmessung zeigen (vgl. Förstner 1984).

a.) Das Gradientenbild  $g_x$  stellt ein Kantenbild dar. Verwendet man statt  $g_1$  und  $g_2$  nun  $g_{1,x}$  und  $g_{2,x}$  für die Zuordnung, so wird die Standardabweichung  $\sigma_x$  der Zuordnung theoretisch um den Faktor 1.6 größer. Dies stimmt mit den experimentellen Ergebnissen von Anuta (1970) überein. Der Grund liegt darin, daß bei der Gradientenbildung das Rauschen verstärkt wird. Durch zusätzliche Glättung kann dieser Effekt zwar vermindert werden, die Verwendung der Originalbildausschnitte führt - unkorreliertes Rauschen vorausgesetzt - aber auf jeden Fall zum besten Ergebnis.

b.) Sind dagegen die Meßfehler korreliert und verwendet man fälschlicherweise das o. g. Verfahren, so muß man ebenfalls mit verfälschten Ergebnissen rechnen. Wenn die Korrelation  $\rho_{ij}$  unter den Meßfehlern exponentiell mit dem Abstand der Bildelemente  $|i-j|$  abnimmt, also  $\rho_{ij} = \exp(-\rho|i-j|)$  gilt, so ist die Schätzung  $\hat{\sigma}_x$  um den Faktor  $\sqrt{(1+\rho)/(1-\rho)}$  nach beiden Richtungen unsicher. Bei  $\rho = 0.8$  erhält man einen Faktor 3. Im allgemeinen ist die Schätzung jedoch um diesen Faktor zu optimistisch. Die Schätzungen der Verschiebungen werden jedoch selbst kaum beeinflusst, sie bleiben weiterhin erwartungstreu, d. h. systematische Fehler sind nicht zu befürchten (vgl. Koch 1980, S. 164).

c.) Das Medianfilter ist ein wirksames Mittel zur Rauschunterdrückung, da es gleichzeitig Kanten erhält. Man sollte daher erwarten, daß dieses nichtlineare Filter die Zuordnungsgenauigkeit steigert, was auch zwei theoretische Arbeiten nahelegten (Geiselman 1983, Yang/Huang 1981). Die empirischen Ergebnisse von Akey und Mitchell (1984) ergaben jedoch das Gegenteil: in median-gefilterten Bildern lassen sich Objekte ungenauer lokalisieren als in Bildern, bei denen eine gleitende Mittelbildung für die Rauschunterdrückung verwendet wurde. Die Ergebnisse lassen sich durch die Instabilität des Medians in der Nähe von Kanten erklären (Gallagher 1982). Sie zeigen vor allem, wie schwierig der Einfluß nichtlinearer Filter im Einzelfall vorhersehbar ist und daß empirische Tests in jedem Falle zur Überprüfung notwendig sind.

### 2.6 Optimale Pixelgröße

Ein spezielles Filter ist die Abtastung einer kontinuierlichen Grauwertfunktion. Hier stellt sich die Frage nach der für die Genauigkeit der Punktübertragung günstigsten Schlitzbreite oder Sensorelementgröße  $\Delta s(\text{opt})$  und der Abtastfrequenz oder der Pixelgröße  $\Delta x(\text{opt})$  bei der Abtastung.

Falls die Zahl  $n$  der verwendeten Bildelemente und das Signal-Rausch-Verhältnis  $SNR$  unverändert bleibt, existiert eine optimale Schlitzbreite (Förstner 1982, 1984). Sie hängt von der Korrelationslänge  $d$  der Autokorrelationsfunktion  $R(x)$  des Objekts ab. Wenn man  $R(x) \sim \exp(-|x|/d)$  annimmt, d. h. Grauwerte in einem Abstand  $d$  nur noch zu  $1/e \approx 37\%$  korreliert sind, dann ist die optimale Schlitzbreite

$$\Delta s(\text{opt}) = 0.7 d \quad (2-14)$$

Für gute Luftbilder mit  $d = 40 \mu\text{m}$  erhält man  $\Delta s(\text{opt}) = 30 \mu\text{m}$ . Die damit erreichbare Genauigkeit für die Zuordnung ist

$$\sigma_x \approx 1.2 \frac{d}{\sqrt{N}} \frac{1}{SNR} \quad (2-15)$$

In einem guten Luftbild läßt sich damit eine Zuordnungsgenauigkeit von etwa  $1 \mu\text{m}$  erreichen, wenn

100 Bildelemente verwendet werden und das Signal Rausch Verhältnis  $SNR = 5$  beträgt. Dies entspricht etwa der vom visuellen System unter ähnlichen Bedingungen erreichbaren Parallaxenmeßgenauigkeit. Der Wert  $d = 40 \mu m$  für gute Luftbilder wurde Helava (1976) entnommen.

Die Schlitzbreite  $\Delta s$  bei der Abtastung sollte nach Castleman (1979, S. 244) etwa doppelt so groß gewählt werden wie das Abtastintervall  $\Delta x$ , um sog. Verflechtungseffekte (aliasing, z. B. Moirébildung) zu vermeiden. Daher ist  $\Delta s(opt)$  nicht die optimale Pixelgröße (s. Förstner 1982) sondern

$$\Delta x(opt) = \frac{1}{2} \Delta s(opt) = 0.35 d \quad (2-15)$$

Inwieweit dieser Unterscheidung bei der Realisierung von CCD-Kameras, bei denen immer  $\Delta s < \Delta x$  ist, etwa durch geeignete Optik, Rechnung getragen wird, soll hier nicht untersucht werden. Für gutes Luftbildmaterial erhalten wir die optimale Pixelgröße zu  $\Delta x(opt) = 15 \mu m$ .

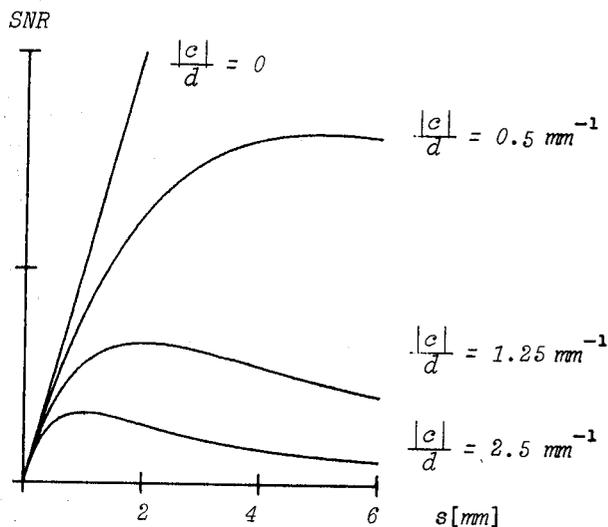
## 2.7 Einfluß nicht modellierter geometrischer Verzerrungen

Nicht modellierte geometrische Verzerrungen zwischen den Bildausschnitten können zu systematischen Fehlern führen und die Genauigkeit der Zuordnung mindern. Wir wollen hier den Einfluß nicht modellierter Maßstabsunterschiede untersuchen.

1.) Da kleine Bildausschnitte zu wenig Bildelemente enthalten und große Bildausschnitte wegen der Verzerrungen schlecht zugeordnet werden können, existiert eine bzgl. des Signal-Rausch-Verhältnisses und damit wegen Gl. 2-13b auch bzgl. der Genauigkeit günstigste Größe  $s(opt)$  der Bildausschnitte (vgl. Svedlow et. al. 1976). Fig. 2-6 zeigt den Einfluß von Maßstabsfehlern auf das

Fig. 2-6

Signal-Rausch-Verhältnis nach der Zuordnung für verschiedene Maßstabsfehler  $c$  in Abhängigkeit von der Länge  $s$  eines quadratischen Bildausschnittes (nach Svedlow et. al. 1976). Korrelationsfunktion des Objekts  $R(x) = \exp(-|x|/d)$ ,  $d$  Korrelationslänge



Signal-Rausch-Verhältnis bei quadratischer Bildausschnittgröße  $s \times s$  und stützt sich auf das in Abschnitt 2.6 benutzte Korrelationsmodell des Objekts. Der Einfluß hängt nur vom Maßstabsfehler  $c$  und dem Korrelationsabstand  $d$ , d.h. der Glattheit des Objekts ab. Offensichtlich nimmt der Einfluß des Maßstabsfehlers mit wachsendem Maßstabsunterschied zu und nimmt mit größer werdendem Korrelationsabstand  $d$  ab, da dann das Bild weniger Struktur aufweist. Die optimale Bildausschnittgröße für quadratische Ausschnitte ist

$$s(opt) = 2.5 \frac{d}{|c|} \quad (2-16)$$

Für gute Luftbilder mit  $d = 40 \mu m$  und einem Maßstabsfehler von 50 % erhält man  $s(opt) = 200 \mu m$ . Wählt man das Abtastintervall halbsogroß wie die optimale Schlitzbreite nach Gl. 2-15, so entspricht dies einem Bildausschnitt mit  $14 \times 14$  Pixeln, für  $\Delta s(opt) = \Delta x(opt)$  sind es dagegen nur  $7 \times 7$  Pixel.

2.) Den verbleibenden Einfluß der Maßstabsunterschiede auf das Ergebnis kann man eliminieren, indem man als Übertragungspunkt im 2. Bildausschnitt einen Punkt  $P_2$  wählt, dessen übertragener Punkt  $\bar{P}_2(\bar{z}_2)$  mit  $\bar{z}_2 = T_G(z_2)$  unabhängig von einer Maßstabsdifferenz ist. Man kann für eindimensionale Objekte zeigen, daß dies der mit den quadratischen Gradienten  $g_{x,i}^2$  gewogene Schwerpunkt ist (vgl. Förstner 1984):

$$x_s = \frac{\sum x_i g_{x,i}^2}{\sum g_{x,i}^2} \quad (2-17)$$

Für zweidimensionale Bildausschnitte ist der gewogene Schwerpunkt zwar nicht optimal, reduziert jedoch den Einfluß nicht modellierter geometrischer Verzerrungen. Für signalisierte Punkte mit nur flauem Hintergrund fällt der Schwerpunkt näherungsweise in die Mitte des Signals, auch wenn das Signal deutlich außerhalb der Mitte des Bildausschnitts liegt.

## 2.8 Konvergenzeigenschaften

Für die Lösung eines nichtlinearen Problems ist entscheidend, wie gut die dafür nötigen Näherungswerte sein müssen und wie schnell das Verfahren konvergiert.

Kleinste Quadrate Verfahren, die sich auf linearisierte Fehlergleichungen stützen, konvergieren i. a. quadratisch. In unserem Fall konvergiert es sogar kubisch. Man kann also mit einer raschen Konvergenz rechnen.

Dieser guten Eigenschaft steht der geringe Konvergenzbereich gegenüber. Selbst im günstigsten Fall müssen die Näherungswerte für die Verschiebungen besser als  $1.5 \Delta s(\text{opt})$  sein, wie Burkhardt und Moll zeigten (1978). Das sind nur 3 Bildelemente optimaler Größe. In guten Luftbildern müssen die Näherungswerte näher als  $50 \mu\text{m}$  am wahren Wert liegen. Entsprechendes gilt für die Parameter der Affintransformation: Näherungen für Maßstab und Scherparameter müssen in einem Bereich  $\pm 0.3$ , Näherungen für die Drehung in einem Bereich  $\pm 20^\circ$  vorliegen.

Diese ungünstige Eigenschaft ist spezifisch für alle iterativen Verfahren, die von einem Näherungspunkt aus ein globales Optimum zu erreichen suchen, aber grundsätzlich nur das nächste i. a. relative Optimum finden. Verfahren, die, wie etwa die klassische Korrelation, den gesamten Suchraum nach dem absoluten Optimum durchsuchen, finden es zwar mit entsprechendem Mehraufwand, können aber trotzdem nicht sicher sein, die richtige Lösung gefunden zu haben, etwa bei periodischen Merkmalen innerhalb der Bildausschnitte oder schon, wenn mehrere ähnliche Merkmale vorhanden sind (vgl. z. B. Fig. 2-5a).

Für einen praktischen Einsatz der in diesem Abschnitt diskutierten Verfahren zur Fein-Zuordnung von Bildpaaren benötigen wir Methoden, die uns zuverlässige und hinreichend genaue Näherungswerte für die Verschiebungen und die geometrischen Verzerrungen liefern. Eines werden wir im nächsten Abschnitt beschreiben.

## 3. Ein Algorithmus zur Zuordnung von Bildern

### 3.1 Vorbemerkungen

Eine zuverlässige Zuordnung setzt genügend Redundanz voraus. Die bisher behandelten Verfahren der Korrelation bzw. Zuordnung arbeiten mit vergleichsweise kleinen Bildausschnitten mit  $5 \times 5$  bis  $32 \times 32$  Bildelementen etwa. Obwohl formal die Überbestimmung bei der Schätzung der Parameter hoch ist, lassen sich Fehlzuordnungen nur vermeiden, wenn deutlich größere Bildausschnitte verwendet werden, etwa mit  $64 \times 64$  bis  $512 \times 512$  Bildelementen. Damit werden bei einer Pixelgröße von  $15$  oder  $20 \mu\text{m}$  Bildausschnitte von  $1 \times 1 \text{ mm}^2$  bis  $10 \times 10 \text{ mm}^2$  abgedeckt.

Die Zuordnung ist bei solch großen Bildausschnitten aus zwei Gründen sehr zuverlässig:

- a.) Die Wahrscheinlichkeit, einen sehr ähnlichen Bildausschnitt an einer anderen Stelle im selben Bild zu finden ist sehr klein, wenigstens um Größenordnungen kleiner als bei den für die Fein-Zuordnung verwendeten Bildausschnitten.
- b.) Die im Bildausschnitt erkennbaren Merkmale (Punkte) liegen auf zumindest stückweise glatten Flächen, wenn man von Waldgebieten und ähnlich in der Tiefe strukturierten Objekten absieht. Damit läßt sich die Konsistenz der Zuordnung benachbarter Merkmale über die Ähnlichkeit der Parallaxen prüfen, bzw. herstellen.

Mit der Vergrößerung der Bildausschnitte sollte es aber auch möglich sein, die Anforderung an die bereitzustellenden Näherungswerte drastisch zu vermindern. Abweichungen der Näherungswerte der Parallaxen von den wahren Werten sollten bis zu etwa  $30\%$  der Breite der Bildausschnitte möglich sein, da dann eine  $50\%$ -ige gegenseitige Flächenüberdeckung gewährleistet ist.

Es existieren bereits einige z. T. sehr leistungsfähige Verfahren zur Lösung dieser Aufgabe (Barnard/Thompson 1980, Grimson 1981, Stockmann et. al. 1982, Zimmermann/Kories 1984, Dreschler 1981, Ohta/Kanade 1984). An dem im folgenden dargestellten neuen Verfahren läßt sich das Grundprinzip, das all diesen Methoden zugrunde liegt, leicht anschaulich machen. Es basiert auf demselben geometrischen Modell, wie die Kleinste Quadrate Lösung für die Zuordnung kleiner Bildausschnitte und benutzt die Genauigkeitsmaße zur Steuerung.

### 3.2 Prinzip von Zuordnungsverfahren

Zuordnungsverfahren arbeiten in drei Stufen:

#### 1.) Interestoperator zur Auswahl markanter Punkte:

Mit Hilfe eines Interestoperators werden in beiden Bildern getrennt Punkte ausgesucht, die sich deutlich von ihrer näheren Umgebung unterscheiden und von denen man annehmen kann sie im Nachbarbild wiederzufinden. Um Fehlzuordnungen zu vermeiden ist es darüber hinaus sinnvoll, nur Punkte zu wählen, die sich auch von allen übrigen unterscheiden, d. h. interessant sind. Verfahren zur Auswahl solcher Punkte werden daher Interestoperatoren genannt. Die bekannten Interestoperatoren unterscheiden sich nur wenig, sind aber sehr verschieden motiviert.

2.) Ähnlichkeitsmaß zur vorläufigen Zuordnung:

Auf der Basis der Ähnlichkeit der Merkmale lassen sich die ausgewählten Punkte vorläufig zuordnen. Dabei können einem Punkt in einem Bild mehrere ähnliche Punkte im anderen Bild zugeordnet werden, denn der Grad der Ähnlichkeit ist i. a. nicht hinreichend für eine richtige Zuordnung.

Es ist eine große Zahl verschiedener Ähnlichkeitsmaße für diese Aufgabe in Gebrauch. Die Wahl eines Ähnlichkeitsmaßes hängt sehr von der jeweiligen Aufgabenstellung ab, insbesondere von den zu erwartenden radiometrischen und geometrischen Transformationen zwischen den Bildern. Das verwendete Ähnlichkeitsmaß sollte invariant gegen diese Transformationen sein, eine Forderung die u. U. schwer zu erfüllen ist.

3.) Konsistenz der endgültigen Zuordnung:

Das geometrische Modell des Objekts - zusammen mit der Abbildungsgeometrie - läßt sich zur endgültigen, konsistenten Zuordnung verwenden. Im Prinzip geht es darum, aus der großen Zahl der vorläufigen Zuordnungen über die Analyse der Parallaxen die richtigen Zuordnungen herauszufiltern. Da das Konsistenzmaß als solches keinen Anhalt für diese Filterung gibt, existiert eine Vielzahl von Algorithmen, die sich z. T. grundsätzlich in ihrem mathematischen Ansatz unterscheiden. Nicht alle lassen eine quantitative Beurteilung des Ergebnisses zu.

Im folgenden werden die einzelnen Stufen eines neuen Verfahrens beschrieben (vgl. Paderes et. al. 1984).

3.3 Interestoperator

Markante Punkte zeichnen sich dadurch aus, daß sie sich in allen Richtungen eindeutig lokalisieren lassen. Verwendet man als Merkmal der Punkte die Grauwerte in einer kleinen Umgebung, so läßt sich dieser Punkt eindeutig lokalisieren, wenn die Zuordnungsgenauigkeit in allen Richtungen gut ist. Sie wird repräsentiert durch die Fehlerellipse und läßt sich bei Vorgabe eines realistischen Wertes für  $\sigma_n$  aus einem Bildausschnitt ableiten (vgl. Gl. 2-12a). Damit sind die folgenden zwei Forderungen unmittelbar einleuchtend, die an die Fehlerellipse eines markanten Punktes zu stellen sind:

a.) Die Fehlerellipse muß möglichst rund sein.

Sind  $a$  und  $b$  die beiden Halbachsen, so gibt der Quotient

$$q = \frac{1 \det N}{(sp N)^2} = 1 - \frac{(a^2 - b^2)^2}{(a^2 + b^2)^2} \quad (3-1)$$

die "Rundheit" der Fehlerellipse an. Für  $a = b$  ist die Ellipse ein Kreis,  $q = 1$ . Die Determinante  $\det N$  und die Spur  $sp N$  lassen sich aus den Elementen der Normalgleichungsmatrix  $N$  berechnen, bei denen  $q$  zu klein ist, etwa  $q < 0.5$ , sind in der Richtung der großen Halbachse relativ ungenau lokalisierbar und sind daher auszuschließen, etwa Kantenpunkte. Beispielsweise ergibt sich  $q = 0.11$  für den Punkt in Fig. 2-5c.

b.) Die Fehlerellipse sollte möglichst klein sein.

Die Genauigkeit der Lokalisierung kann man durch den Punktfehler  $\sigma_p^2 = \sigma_x^2 + \sigma_y^2 = sp Q_i$  ( $Q = N^{-1}$ ,  $\sigma_0 = 1$ ) oder durch das Gewicht

$$w_i = \frac{1}{sp Q_i} = \frac{\det N_i}{sp N_i} \quad (3-2)$$

der für den Punkt  $P_i$  zu erwartenden Übertragungsgenauigkeit beschreiben. An einem markanten Punkt ist das Gewicht größer als bei allen Nachbarpunkten. Ein Beispiel für die Auswahl markanter Punkte in einem Bildpaar zeigt Fig. 3-1. Offensichtlich liegen an Kanten nur vereinzelt markante Punkte.

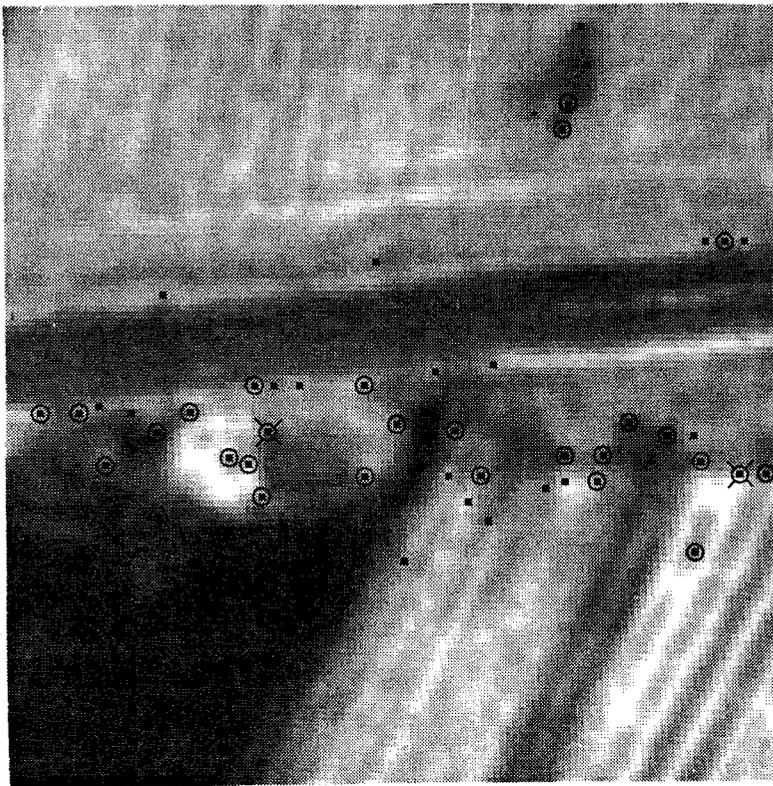
3.4 Ähnlichkeitsmaß

Die vorläufige Zuordnung der Punkte stützt sich auf die Ähnlichkeit der Merkmale der Punkte im rechten und linken Bild. Wir verwenden hier die Punktgenauigkeit  $\sigma_p$ , die sich vor der 1. Iteration einer Kleinsten Quadrate Anpassung der kleinen Bildausschnitte ergibt. Das zugehörige Gewicht der Zuordnung des  $i$ -ten Punktes im linken zum  $j$ -ten Punkt im rechten Bild ist mit  $\sigma_p^2 = \sigma_n^2 \text{tr} Q$  und Gl. 2-5

$$w_{ij} = \frac{1}{\sigma_p^2(i,j)} = \frac{n}{n} \frac{\rho_{ij}}{1 - \rho_{ij}} \frac{1}{\sigma_{g_i} \sigma_{g_j}} \sqrt{w_i w_j} \quad (3-3)$$

Hierin wurde aus Symmetriegründen ein mittlerer Wert für die Varianz der Grauwerte und der Gewichte gewählt. Der Korrelationskoeffizient mißt die Ähnlichkeit der Punkte. Er ist nicht invariant gegen geometrische Verzerrungen (s. Abschn. 2.7), erfüllt also nicht die gestellten Forderungen. Für Verzerrungen kleiner als 30 % ist er allerdings brauchbar.

Unter allen möglichen  $n_i \times n_j$  Zuordnungen werden all diejenigen ausgeschlossen ( $w_{ij} = 0$ ), bei denen  $\rho_{ij} < 0.5$  ist oder die zugehörigen x- oder y-Parallaxen zu groß sind, etwa größer als



Geometrische Transformation:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 1.011 & 0.011 \\ 0.000 & 1.001 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} + \begin{pmatrix} -1.4 \\ 5.6 \end{pmatrix}$$

Korrelationskoeffizient: 0.970

Fig. 3-1 Zugeordnete Bildausschnitte 4.8 x 4.8 mm<sup>2</sup>  
 Bildmaßstab 1 : 7 800, Pixelgröße 40 x 40 µm<sup>2</sup>  
 ■ vom Interestoperator ausgewählte Punkte  
 ⊙ vom Programm zugeordnete Punkte  
 ⊗ Punkte mit Restparallaxen > 2 Pixel

1/3 der Bildbreite  $s$ .

### 3.5 Konsistenzbedingung

Im letzten Schritt des Verfahrens müssen die falschen Zuordnungen aus der Liste der vorläufigen Zuordnungen eliminiert werden. Wenn wir auch hier annehmen, daß die geometrische Transformation eine Affintransformation ist, müssen die Koordinaten der richtigen Zuordnungen Gl. 2-1 erfüllen. Alle anderen Punkte kann man als grob falsch bezüglich dieses Modells ansehen. Die Situation ist formal der Suche grober Fehler in klassischen Ausgleichungen ähnlich. Doch unterscheiden sich beide Fälle im Prozentsatz grober Fehler: Während, etwa in der Aerotriangulation mit 1 % bis höchstens 10 % groben Fehlern zu rechnen ist, sind hier im Durchschnitt über 50 %, oft 80 % aller vorläufigen Zuordnungen falsch. Dennoch lassen sich auch hier i. a. mit Hilfe einer robusten Ausgleichung (vgl. Werner 1984, Klein/Förstner 1984) alle grob falschen Zuordnungen eliminieren. Die restlich verbleibenden kleinen Zuordnungsfehler lassen sich entweder auf der Basis der Residuen oder der anschließenden Fein-Zuordnung beseitigen.

#### Beispiel:

Das Ergebnis der automatischen Zuordnung der Bilder in Fig. 3-1 ist durch die kreisförmig gekennzeichneten Punkte dargestellt. Die visuelle Zuordnung dieser gekennzeichneten Punkte entspricht der des automatischen Verfahrens. Insgesamt wurden 29 Punkte zugeordnet. Wegen der großzügigen Toleranz von 3 Bildelementen für die Restparallaxen, sind einige Punkte nicht richtig zugeordnet. Sie haben nur einen geringen Einfluß auf die Gesamttransformation.

Die ausgewählten Punkte sind Merkmale und dienen zunächst nur der Zuordnung des Gesamtbildausschnittes. Diese Zuordnung ist richtig. Die Transformationsparameter können nun als Näherungswerte für die ggf. folgende Fein-Zuordnung dienen, sowohl für die Parallaxen als auch für die geometrischen Verzerrungen. Weiterhin kann man die ausgewählten und zugeordneten Punkte etwa für eine relative Orientierung verwenden. Ihre Verlässlichkeit ist umso höher, je größer die Ähnlichkeit ist und je kleiner die Restparallaxen sind. Auf diese Weise lassen sich zugeordnete Punkte für die Weiterverarbeitung auswählen, die ein großes Maß an Vertrauenswürdigkeit aufweisen. ■

### 3.6 Qualität der Zuordnung

Die Genauigkeit der Zuordnung hängt wesentlich von der Qualität des geometrischen Modells ab. Ist das Modell näherungsweise erfüllt, erreicht man eine Zuordnungsgenauigkeit, d. h. Meßgenauigkeit der Parallaxen für die einzelnen Punkte von etwa 0.5 Pixel. Das entspricht dem Fehler der durch die Rundung der Koordinaten der ausgewählten Punkte auf ganze Zahlen entsteht.

#### Beispiel:

Für eine relative Orientierung wurden an den 6 Standardstellen je zwei Bildausschnitte mit  $120 \times 120$  Pixeln der Größe  $40 \times 40 \mu\text{m}$  ausgewählt und zugeordnet. Um die Zahl der insgesamt 291 zugeordneten Punkte zu reduzieren, wurden in jedem Bildausschnitt die ersten 5 zugeordneten Punkte ausgewählt und mit ihnen die relative Orientierung berechnet. Die mittlere  $y$ -Parallaxe ergab sich zu  $\sigma_{py} = 21 \mu\text{m}$ . Nach Elimination von 4 der 60 Orientierungspunkte war das endgültige Ergebnis  $\sigma_{py} = 15.6 \mu\text{m}$  mit einer maximalen Parallaxe von  $42 \mu\text{m}$ . ■

Als durchgreifende Kontrolle der Gesamttransformation hat sich die Berechnung des Korrelationskoeffizienten erwiesen. Dazu werden ca. 1000 in einem Gitter angeordnete Kontrollpunkte im rechten Bild über die geschätzten Transformationsparameter Punkten im linken Bild zugeordnet und aus den korrespondierenden Grauwerten ein globaler Korrelationskoeffizient berechnet. Diese unabhängige Kontrolle ist außerordentlich empfindlich, da bei einer falschen Zuordnung des Gesamtbildausschnittes der Korrelationskoeffizient sofort unter 0.2 sinkt. Bei Korrelationskoeffizienten  $> 0.5$  kann man daher sicher sein, daß die Zuordnung der Bildausschnitte in Ordnung ist. Beispielsweise ergab sich der Korrelationskoeffizient nach der Zuordnung der Bildausschnitte in Fig. 3-1 zu  $\rho = 0.967$ .

Der Algorithmus ist in der Lage gegenseitig affin verzerrte Bildausschnitte zu verarbeiten. Da das verwendete Ähnlichkeitsmaß, der Korrelationskoeffizient, nicht invariant gegen affine Verzerrungen ist, ist das Verfahren - wie zahlreiche Versuche ergaben - auf Maßstabsunterschiede bis etwa 30 % und Rotationen bis etwa  $20^\circ$  beschränkt (vgl. Fig. 3-2).

Die Rechenzeit für die Zuordnung zweier Bildausschnitte der Größe  $128 \times 128$  Pixel beträgt ca. 2 sec. auf einer VAX 11/780. Auf dem HP 1000 Rechner am Planicom C 100 lassen sich Ausschnitte bis  $60 \times 60$  Pixel zuordnen. Die Rechenzeit dafür bleibt unter 10 sec.

## 4. Diskussion

Die Zuordnung von Bildausschnitten ist der zentrale Baustein bei der Automatisierung photogrammetrischer Aufgaben. Die Untersuchung der Leistungsfähigkeit von Zuordnungsalgorithmen hat wichtige Ergebnisse gebracht:

- 1.) Die Genauigkeit der Zuordnung wird wesentlich durch die Bildelementgröße beschränkt. Bei einer Quantisierung der Grauwerte mit 8 bit spielen die dabei auftretenden Rundungsfehler nur eine untergeordnete Rolle. Die theoretischen Streuungen für die Lokalisierung ideal abge-

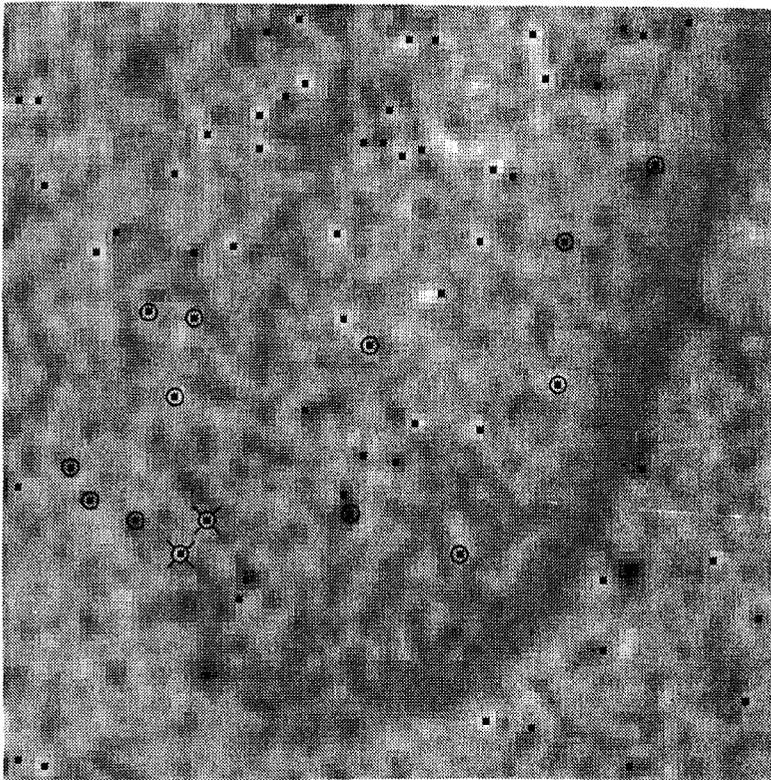
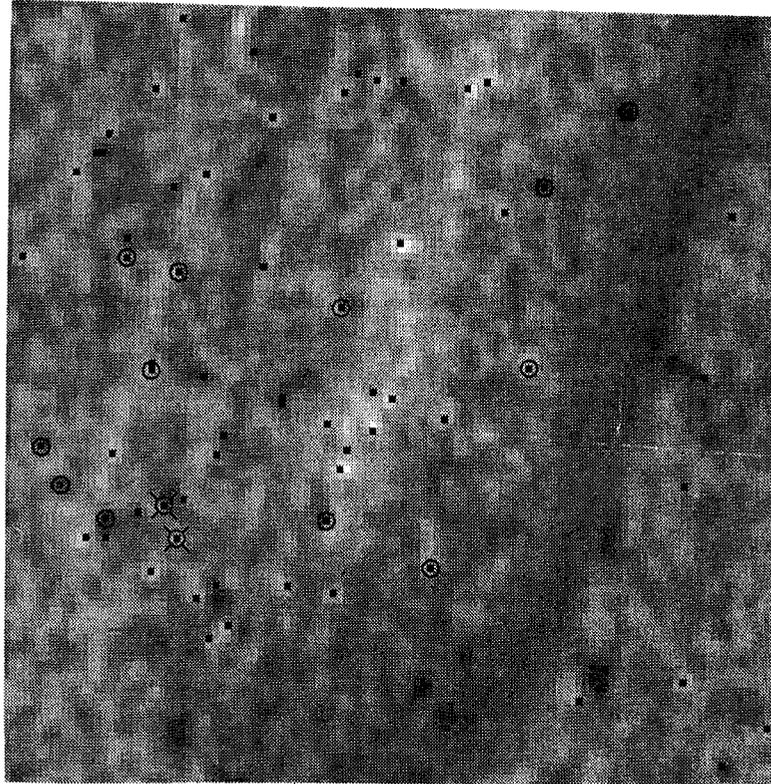


Fig. 3-2 Zugeordnete Bildausschnitte, 4.8 x 4.8 mm<sup>2</sup>  
 Bildmaßstab ca. 1 : 20, Pixelgröße 40 x 40 µm<sup>2</sup>

- vom Interestoperator ausgewählte Punkte
- ⊙ vom Programm zugeordnete Punkte
- ⊗ Punkte mit Restparallaxen > 2 Pixel



Geometrische Transformation:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.818 & -0.033 \\ 0.000 & 1.015 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} + \begin{pmatrix} 10.0 \\ -5.7 \end{pmatrix}$$

Korrelationskoeffizient: 0.697

tasteter Objekte liegen weit unter  $1/10$  der *Pixelgröße*.

- 2.) Klassische Korrelationsverfahren können durch eine Kleinste Quadrate Schätzung ersetzt werden. Dies hat weitreichende Folgerungen für die praktische Anwendung wie auch für die theoretische Durchdringung der Verfahrensschritte. Insbesondere lassen sich geometrische Verzerrungen zwischen den Bildausschnitten berücksichtigen, womit eine wesentlich bessere Zuordnungsgenauigkeit erreicht werden kann.
- 3.) Die für die Präzision wesentlichen Parameter sind die Bildausschnittgröße oder die Zahl der Bildelemente, die Relativgenauigkeit der Grauwerte oder das Signal-Rausch-Verhältnis und die Schärfe oder die Textur des Objekts.  
 Das zugrundegelegte Modell ist sehr einfach. Es ist vorläufig genügend allgemein, muß aber durch empirische Tests überprüft werden, um Hinweise für mögliche Verfeinerungen zu erhalten. Die Situation bei den Verfahren zur Zuordnung digitaler Bilder entspricht der Situation der Aerotriangulation um 1970. Damals waren die theoretischen Genauigkeitsmodelle für die Blockausgleichung die Grundlage zahlreicher empirischer Tests, die dann schließlich in das jetzt allgemein akzeptierte Verfahren der Blockausgleichung mit Selbstkalibrierung mündeten.
- 4.) Die mit Zuordnungsverfahren, wie der Korrelation oder der Kleinsten Quadrate Anpassung, erreichbaren Genauigkeiten für die Parallaxenmessung liegen bei  $1 \mu\text{m}$  und sind - unter gleichen Bedingungen - der Meßgenauigkeit eines Operateurs ähnlich. Die in den Bildern enthaltene geometrische Information läßt sich also auch mit automatischen Verfahren ausschöpfen.
- 5.) Erstmals ist es möglich, die Qualität der Messung sofort zu überprüfen. Ohne Mehraufwand steht mit den Streuungen der Parallaxen eine realistische Schätzung für die Meßgenauigkeit i. S. einer Wiederholgenauigkeit zur Verfügung. Sie kann bei Folgeschritten etwa bei der Aerotriangulation oder der Satellitenbildentzerrung weiterverwendet werden. Damit ist es, z. B. möglich, auch Punkte auf Hell-Dunkel-Kanten, die oft sehr zahlreich vorliegen, zur Verknüpfung von Bildern heranzuziehen (vgl. Paderes et. al. 1984).  
 Die Genauigkeitsangaben lassen sich aber auch für Planungszwecke verwenden, etwa zur Auswahl von Punkten, die eine gute Parallaxenmessung versprechen oder - vor allem im Nahbereich - zur Vorgabe geeigneter Zielmarken oder Oberflächenstrukturen.
- 6.) Mit den diskutierten Zuordnungsalgorithmen lassen sich nicht nur Parallaxen, d. h. Höhen, sondern auch Oberflächenneigungen messen. Sie kann man für die Bestimmung der Oberfläche nutzbringend einsetzen.
- 7.) Wegen des geringen Konvergenzbereichs der iterativen Kleinste Quadrate Anpassung sind Verfahren zur Bestimmung von Näherungswerten für die Parallaxen als auch für die geometrischen Verzerrungen erforderlich. In kupertem Gelände kann am Rand der Bilder die Bereitstellung von Näherungswerten zum Problem werden. Das vorgestellte, vergleichsweise schnelle Verfahren zur Zuordnung größerer Bildausschnitte erlaubt Abweichungen der Näherungswerte von ca.  $1 \text{ mm}$  in den Verschiebungen bzw.  $30 \%$  in den Verzerrungen. Andererseits kommen alle im Bereich der Mustererkennung entwickelten Verfahren zur Bildzuordnung für diese Aufgabe in Frage und müssen auf ihre Eignung hin geprüft werden.

Das in diesem Beitrag zugrunde gelegte geometrische Modell ist zwar für den wichtigen Bereich von Anwendungen, bei dem glatte Oberflächen vorausgesetzt werden können, brauchbar. Es läßt sich aber nicht genügend verallgemeinern, wenn die Objekte keine glatte Oberfläche aufweisen, sondern Kanten, Sprünge oder Verdeckungen. Hier sind andere Konzepte, wie sie in der Mustererkennung entwickelt werden unbedingt erforderlich.

LITERATUR

Abkürzungen:

BuL	Bildmessung und Luftbildwesen
CGIP	Computer Graphics and Image Processing
IAP	International Archives for Photogrammetry
IEEE AES	IEEE Transactions on Aerospace and Electronic Systems
IEEE TC	IEEE Transactions on Computers
IEEE COM	IEEE Transactions on Communication

IEEE GE	IEEE Transactions on Geoscience Electronics
IEEE IT	IEEE Transactions on Information Theory
IEEE PAMI	IEEE Transactions on Pattern Analysis and Machine Intelligence
IEEE SMC	IEEE Transactions on Systems Man and Cybernetics
IPS	Schriftenreihe des Instituts für Photogrammetrie Stuttgart
PhEng	Photogrammetric Engineering and Remote Sensing
PhRec	Photogrammetric Record
ZfV	Zeitschrift für Vermessungswesen

- Ackermann F., Pertl A. (1983):* Zuordnung kleiner Bildflächen durch digitale Korrelation zur Verknüpfung verschiedener oder verschiedenartiger Bilder im Anwendungsbereich Photogrammetrie und Fernerkundung, DFG-Ab-schlußbericht, 1983
- Akey M. L., Mitchell R. (1984):* Detection and Subpixel Location of Objects in Digitized Aerial Images, 7th Conf. on Pattern Recognition, Montreal 1984
- Anuta P. E. (1970):* Spatial Registration of Multispectral Digital Imagery Using Fast Fourier Transformation Techniques. IEEE GE-8, 1970, pp. 353-368
- Barnard S. T., Thompson W. W. (1980):* Disparity Analysis of Images, IEEE PAMI-2, 1980, pp. 333-340
- Burkhardt H., Moll H. (1979):* A Modified Newton-Raphson Search for the Model-Adaptive Identification of Delays, 5th IFAC-Symp. on Identification and System Parameter Estimation, Darmstadt 1979, pp. 1279-1286
- Cafforio C., Rocca F. (1979):* Tracking Moving Objects in Television Images, Signal Processing 1
- Castleman K. R. (1979):* Digital Image Processing, Prentice Hall, N. J., 1979
- Dinse Th., Enkelmann W., Nagel H. H. (1981):* Untersuchungen von Verschiebungsvektorfeldern in Bildfolgen, Informatik Fachberichte 49, Heidelberg 1981, S. 59-75
- Dorst L., Duin R. P. W. (1984):* Spirograph Theory: A Framework for Calculations on Digitized Straight Lines, IEEE PAMI-6, 1984, pp. 632-639
- Dreschler L. (1981):* Ermittlung markanter Punkte auf Bildern bewegter Objekte und Berechnung einer 3D-Beschreibung auf die-
- Fennema C. L., Thompson W. B. (1979):* Velocity Determination in Scenes Containing several Moving Objects, CGIP 9, 1979, pp. 301-315
- Förstner W. (1982):* On the Geometric Precision of Digital Correlation, IAP Vol. 24-III, Helsinki 1982, pp. 176-189
- Förstner W. (1984):* Quality Assessment of Object Location and Point Transfer Using Digital Image Correlation Techniques, IAP Vol. 25-A3a Rio de Janeiro, 1984, pp. 197-219
- Gallagher N. C. (1982):* Median Filters: Properties and Applications, 16th Asilomar Conf. on Circuits, Systems and Computers, 1982
- Geiselmann Ch. (1983):* Theoretische Untersuchungen zu Genauigkeitssteigerung der Digitalen Bildkorrelation durch Medianfilterung, Selbst. Arbeit am Inst. f. Photogr., Univ. Stuttgart, 1983
- Grimson W. E. L. (1981):* From Images to Surfaces, MIT-Press, 1981
- Grün A. (1985):* Adaptive Least Squares Correlation: A Powerful Image Matching Technique, S Afr J of Photogr., Rem. Sens. and Cart. 14(3), 1985, pp. 175-187
- Gülch E. (1985):* Instrumentelle Realisierung und Kalibrierung der digitalen Bildkorrelation am Planicomp, Vortrag Photogr. Woche 1985
- Havelock D. (1984):* Geometric Precision in Digital Images, IAP Vol. 25-A3a, Rio de Janeiro, 1984, pp. 373-380
- Helava U. V. (1976):* Digital Correlation in Photogrammetric Instruments, IAP Vol. 23-III, Helsinki 1976
- Huang T. S. (1981):* Image Sequence Analysis, Springer, NY, 1981
- Klein H., Förstner W. (1984):* Realization of Automatic Error Detection in the Block Adjustment Program PAT-M43, IAP Vol. 25-A3a, Rio de Janeiro, 1984, pp. 234-245
- Koch K.-R. (1980):* Parameterschätzung und Hypothesentests in linearen Modellen, Dümmler 1980
- Kreiling W. (1976):* Automatische Auswertung von Stereobildern durch digitale Korrelation, IAP Vol. 23-III, Helsinki, 1976
- Limb J. O., Murphy J. A. (1975):* Measuring the Speed of Moving Objects from Television Signals, IEEE COM-23, 1975, pp. 474-478
- Meyers M. H., Frank L. E. (1980):* Joint Carrier Phase and Symbol Timing Recovery for PAM Systems, IEEE COM-28, 1980, pp. 1121-1129
- Mikhail E. M. (1983):* Photogrammetric Target Location to Sub-Pixel Accuracy in Digital Images, IPS 9, 1984
- Ohta Y., Kanade T. (1984):* Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming, IEEE PAMI-7, 1985, pp. 139-154
- Paderes F. C., Mikhail E. M., Förstner W. (1984):* Rectification of Single and Multiple Frames of Satellite Scanner Imagery Using Points and Edges as Control, NASA-Symp. on Math. Pattern Recognition and Image Analysis, Houston 1984
- Pertl A. (1984):* Digital Image Correlation with the Analytical Plotter Planicomp C 100, IAP Vol. 25-A3b, Rio de Janeiro, 1984
- Pertl A. (1985):* Empirische Ergebnisse der automatischen Parallaxen- und Punktmessung, Vortrag Photogr. Woche 1985
- Schalckhoff F. J., McVey E. S. (1979):* Algorithms Development for Real Time Automatic Video Tracking Systems, Proc. 3rd Int. Computer Software and Application Conf., Chicago, 1979
- Stockmann G. C., Kopstein S., Bennett S. (1982):* Matching Images to Models for Image Registration and Object Location via Clustering, IEEE PAMI-4, 1982, pp. 229-241
- Svedlow M., McGillem C., Anuta P. (1976):* Analytical and Experimental Design and Analysis of an Optimal Processor for Image Registration, LARS, Inf. Note 090776, Purdue Univ., West Lafayette 1976
- Thurgood J. D., Mikhail E. M. (1982):* Sub-Pixel Mensuration of Photogrammetric Targets in Digital Images, School of Civil Eng., Purdue, CH-PH-82-2
- Werner H. (1984):* Automatic Gross Error Detection by Robust Estimators, IAP Vol. 25-A3b, Rio de Janeiro, 1984
- Wild E. (1979):* Interne Berichte, Inst. f. Photogr., Univ. Stuttgart, 1979
- Yang G. J., Huang T. S. (1981):* The Effect of Median Filtering on Edge Location Estimation, CGIP Vol. 15, 1981, pp. 224-245
- Zimmermann G., Kories R. (1984):* Eine Familie von Bildmerkmalen für die Bewegungsbestimmung in Bildfolgen, Inf. Fachberichte 87, Springer 1984

Summary:

The article discusses methods for finding corresponding points in digital images. The reliability and the precision of the match are of primary interest. First the limiting effect of discretization and quantization onto the precision is investigated, especially the precision of locating a ramp edge, a binary edge and a binary circle are discussed. The actual location process, digital correlation and the equivalent least squares approach, is analysed in detail, including the determination of an optimal pixel and patch size. Due to the small radius of convergence of the differential solution, finally an algorithm is presented, which is able to find corresponding points in larger image patches and thus to provide reliable approximate values for the parallaxes and the geometric distortions.

Zusammenfassung

PRINZIP UND LEISTUNGSFAHIGKEIT DER KORRELATION UND ZUORDNUNG DIGITALER BILDER

Der Artikel diskutiert die Methoden zur Suche korrespondierender Punkte in digitalen Bildern. Die Zuverlässigkeit und die Präzision der Zuordnung sind von primärem Interesse. Zuerst wird der begrenzte Einfluß der Diskretisierung und Quantisierung auf die Präzision untersucht. Im besonderen werden die Präzision der Lokalisierung rampenförmiger Kanten und von Kanten und Kreisen in Binärbildern diskutiert. Der eigentliche Lokalisierungsprozess - die digitale Korrelation und die äquivalente Kleinste Quadrate Näherung - werden im Detail analysiert, einschließlich der Bestimmung optimaler Pixel und Fenstergrößen. Aufgrund des kleinen Konvergenzradius der Kleinste Quadrate Lösung wird schließlich ein Algorithmus vorgestellt, der in der Lage ist, die zugeordneten Punkte in großen Bildausschnitten zu finden und so zuverlässige Näherungswerte für die Parallaxen und für die geometrische Entzerrung zu liefern.

Dr.-Ing. Wolfgang Förstner  
Institut für Photogrammetrie  
Universität Stuttgart  
Keplerstr. 11  
D - 7000 Stuttgart 1