



Klaus D. Tönnies

Grundlagen der Bildverarbeitung

**Unser Online-Tipp
für noch mehr Wissen ...**



... aktuelles Fachwissen rund
um die Uhr – zum Probelesen,
Downloaden oder auch auf Papier.

www.InformIT.de

verwendet werden. Soll Tiefe durch ein Stereo-Vision-Verfahren bestimmt werden, so gibt es keinen Signalgeber, sondern eine zweite Kamera (siehe Abbildung 5.2). Identifiziert man in den beiden Bildern den gleichen Punkt, so hat man wieder zwei Linien eines Dreiecks, an dessen Spitze der gesuchte Punkt liegt. Für ein Tiefenberechnungsverfahren aus Bewegung ist die Triangulation etwas komplizierter, weil sich die Form des Dreiecks aus der Verschiebung eines Punktes nach einer beliebigen, aber bekannten Bewegung zusammensetzt. Das Problem der Identifikation der Abbildung desselben Punktes in zwei oder mehreren Bildern, das in beiden Fällen zu lösen ist, wird Korrespondenzproblem genannt. Bei Active-Vision-Verfahren tritt es nicht auf, weil dort die Identifikation durch die vorab bekannte Art des abgegebenen Signals eindeutig ist.

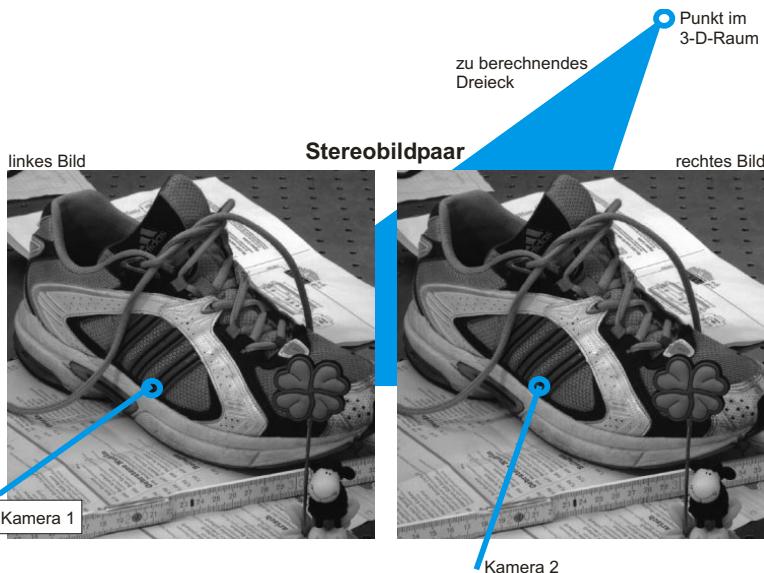


Abbildung 5.2: Beim Stereo-Vision-Versuch erfolgt die Triangulation durch Schnitt von zwei Projektionslinien von der Kamera in die Szene (Bilder von P. Engelhard, C. Rost und M. Wikert).

Verfahren der zweiten Gruppe sind weniger verbreitet. Ein bekannter Ansatz ist **Shape-from-Shading**. Information über die Objektform wird aus der Invertierung eines Beleuchtungsmodells gewonnen. Da das Beleuchtungsmodell beschreibt, wie einfallendes Licht in Abhängigkeit von der Oberflächennormalen reflektiert wird, kann man bei bekanntem Beleuchtungsmodell aus der Helligkeit an jedem Punkt auf dessen Normalenrichtung schließen. Leider ist das Problem unterbestimmt, da nur der Winkel zwischen Lichtquelle und Normale, nicht aber die Normalenrichtung selbst berechnet werden kann. Zudem setzt die Berechnung ein exakt fotometrisch kalibriertes Kamerasystem und ein genau bekanntes Beleuchtungsmodell voraus. Erstes lässt sich mit genügend technischem Aufwand erreichen, doch das Beleuchtungsmodell kann im Allgemeinen nur annähernd bestimmt werden. Zusammen mit der Mehrdeutigkeit der Lösung führt dies dazu, dass die Resultate im Vergleich zu anderen Tiefenberechnungsverfahren oft unbefriedigend sind.

Ein anderer Vertreter der zweiten Gruppe ist das so genannte **Shape-from-Texture**-Verfahren. Diesem liegt eine ähnliche Annahme zugrunde, die sich allerdings nicht auf die Beleuchtung, sondern auf die Textur bezieht. Textur ist eine Eigenschaft, die sich

aus der Abbildung der Musterung einer Objektoberfläche ergibt. Ist ein Objekt gleichmäßig texturiert und weist die Textur eine bestimmte Grundfrequenz auf (z.B. eine bekannte Anzahl von Linien pro Millimeter), dann kann man aus der Veränderung der projizierten Frequenz auf den Winkel zwischen Projektionsrichtung und Oberflächennormalenrichtung schließen. Mit zusätzlichen Glattheitsannahmen kann, wie bei Shape-from-Shading, aus dem Winkel auf eine Normalenrichtung geschlossen werden.

Von allen Methoden ist ein Active-Vision-Verfahren am einfachsten zu erläutern, weil das Korrespondenzproblem nicht existiert. Diese Methode soll nachfolgend beschrieben werden. Im einfachsten Fall wirft ein Signalgeber \bar{S} , dessen Position relativ zur Kamera bekannt ist, einen Lichtstrahl auf die Szene, so dass dieser an einem einzigen Punkt reflektiert wird (siehe Abbildung 5.3). Das Signal muss hell genug sein, um den beleuchteten Ort leicht identifizierbar zu machen. Hat dieser Ort die maximale Helligkeit f_{max} , dann muss nur nach Koordinaten (m_p, n_p) mit $f(m_p, n_p) = f_{max}$ gesucht werden.

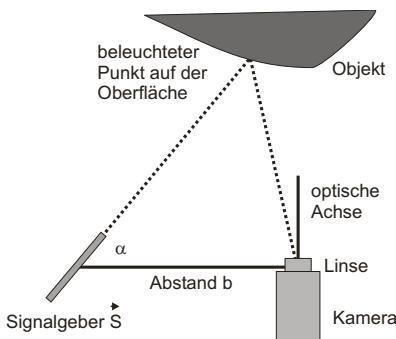


Abbildung 5.3: Active-Vision-Versuch mit strukturiertem Licht: Ein Signalgeber mit bekannter Position zur Kamera projiziert einen Lichtpunkt auf die Objektoberfläche.

Damit eine Kalibrierung der Kamera in Bezug auf ein von der Kamera unabhängiges Weltkoordinatensystem nicht benötigt wird, nehmen wir an, dass beide Koordinatensysteme identisch sind. Der Ursprung des Kamerakoordinatensystems ist der Linsenmittelpunkt, X- und Y-Achse des Systems verlaufen parallel zur X- und Y-Achse des Bildes. Die Z-Achse ist die optische Achse des Aufnahmesystems und verläuft senkrecht zum Bild durch den Linsenmittelpunkt. Der Abstand zwischen Linsenmittelpunkt \bar{L} und Bild sei f . Der Einfachheit halber nehmen wir an, dass der Koordinatenursprung des Bildkoordinatensystems in Kamerakoordinaten an der Stelle $(0,0,-f)$ liegt, d.h. es gibt positive und negative Bildkoordinaten. Die üblichen Bildkoordinaten (die ihren Ursprung in der Regel oben oder unten links im Bild haben) können leicht durch eine entsprechende Translation in diese Koordinaten überführt werden.

Der Signalgeber \bar{S} befindet sich am Punkt $(b, 0, 0)$ des Kamerakoordinatensystems, also im Abstand b auf der X-Achse (siehe Abbildung 5.4). Das Signal soll im Winkel α zur optischen Achse abgestrahlt werden. Aus diesem Winkel kann man zwei Winkel α_x zwischen der in die XZ-Ebene projizierten Signalrichtung und optischer Achse und α_y zwischen Signalrichtung und XZ-Ebene berechnen. Aus den Bildkoordinaten (m_p, n_p) mit diskretem Wertebereich seien reelle Weltkoordinaten $(x_p, y_p, -f)$ berechnet worden, indem die tatsächliche Größe eines Pixels in Millimetern für die Skalierung verwendet worden sei.

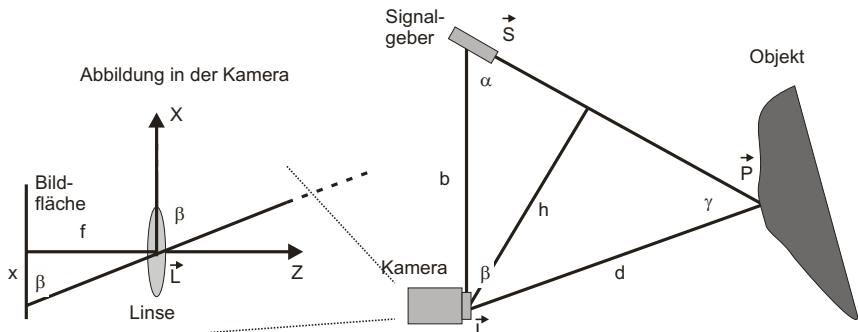


Abbildung 5.4: Berechnung der zu bestimmenden Dreiecksseite d . Das Dreieck ist durch die Seite b und die Winkel α und β eindeutig bestimmt.

Nun wollen wir den Abstand eines Punktes \vec{P} zunächst unter der Einschränkung berechnen, dass er auf einen Ort $(x_p, 0, -f)$ abgebildet wurde. Dazu müssen wir ein Dreieck bestimmen, dessen Eckpunkte der Punkt \vec{P} , der Linsenmittelpunkt und die Position des Signalgebers sind. Die erste Seite dieses Dreiecks ist die Linie $(0,0,0) - (0,0,b)$ zwischen Linse und Signalgeber. Der erste Winkel ist damit $\pi/2 - \alpha_x$. Da der Punkt auf $(x_p, 0, -f)$ abgebildet wird, muss $\alpha_y = 0$ sein und damit ist $\alpha_x = \alpha$.

Der noch fehlende Winkel β ist durch den Winkel zwischen der Linie vom Linsenmittelpunkt zum gesuchten Punkt \vec{P} und der X -Achse gegeben. Dieser Winkel findet sich ein zweites Mal als Winkel in einem Dreieck aus $(x_p, 0, -f)$, $(0,0,0)$ und $(0,0,-f)$. In diesem rechtwinkligen Dreieck sind alle Eckpunkte bekannt und man sieht, dass $\tan \beta = f/x_p$ ist. Damit ist das Dreieck vollständig bestimmt und die Koordinaten von \vec{P} können berechnet werden. Dazu wird zunächst der Abstand d zwischen \vec{P} und dem Punkt $(0,0,0)$ bestimmt. Folgende Festlegungen werden gemacht:

- Der dritte, unbekannte Winkel zwischen den Seiten (\vec{P}, \vec{L}) und (\vec{P}, \vec{S}) sei γ .
- Wir definieren eine Linie der Länge h von \vec{L} senkrecht auf die Seite (\vec{P}, \vec{S}) , die das Dreieck in zwei rechtwinklige Dreiecke teilt.

Für die Länge h gelten nun aus diesen beiden Dreiecken die folgenden Beziehungen (siehe Abbildung 5.4):

$$h = d \cdot \sin \gamma = b \cdot \sin \alpha.$$

Durch Gleichsetzung erhält man einen Ausdruck für die gesuchte Seite d :

$$d = b \cdot \frac{\sin \alpha}{\sin \gamma}$$

Da aber $\alpha + \beta + \gamma = \pi$ und $\sin(\pi - \alpha - \beta) = -\sin(\alpha + \beta)$ sind, ergibt sich schließlich

$$d = -b \cdot \frac{\sin \alpha}{\sin(\alpha + \beta)}.$$

Wenn d bekannt ist, können die Koordinaten von \vec{P} im Kamerakoordinatensystem berechnet werden:

$$\vec{P} = (d \cdot \cos \beta \quad 0 \quad d \cdot \sin \beta).$$

Falls \vec{P} nicht in der XZ -Ebene liegt, wird er auf einen Punkt $(x_p, y_p, -f)$ abgebildet. Dann wird zunächst die gleiche Berechnung für einen in die XZ -Ebene projizierten Punkt \vec{P}' durchgeführt, dessen Abbild $(x_p, 0, -f)$ wäre. Der Winkel β wird wieder basierend auf der Position $(x_p, 0, -f)$ als $\beta = \tan^{-1}(f/x_p)$ berechnet. Der erste Winkel muss jetzt der Winkel α_x sein.

Von \vec{P}' weiß man, dass er der in die XZ -Ebene projizierte Punkt \vec{P} ist. Also sind X - und Z -Koordinate von \vec{P} und \vec{P}' gleich. Die Y -Koordinate ergibt sich schließlich aus der Anwendung der Strahlensätze. Danach ist

$$\frac{y_p}{f} = \frac{Y}{d} \Leftrightarrow Y = \frac{d}{f} y_p.$$

Diese Berechnung ist einfach und beruht nur auf der Anwendung bekannter Gesetze der Geometrie. Fehler sind hauptsächlich wegen einer falschen Lagebestimmung der Kamera relativ zum Signalgeber und wegen Abweichungen von der Projektionsgeometrie zu erwarten.

Die Positionierung der Kamera wird in der Regel nicht mechanisch sichergestellt, sondern beruht auf der Kalibrierung des Systems anhand eines Kalibrierobjekts mit bekannter Position und Ausdehnung. Abweichungen von der Projektionsgeometrie stammen von Linsenverzerrungen, die ebenfalls durch Kalibrierung bestimmt und invertiert werden müssen. Eine exakte Kamerakalibrierung ist daher eine wichtige Voraussetzung für exakte Tiefenrekonstruktion. Diese Verfahren sind allerdings nicht trivial und ihre Behandlung gehört nicht in einen Einführungstext. Der Leser sei auf die entsprechende Fachliteratur verwiesen (z.B. *Computer Vision – A Modern Approach* [Forsyth 2003]).

Das genannte Verfahren ist langsam, wenn mehr als eine Tiefenbestimmung erforderlich ist. Für jeden Punkt \vec{P} muss ein neues Bild erzeugt werden, so dass der Aufwand für die Erzeugung einer Tiefenkarte von z.B. 512^2 Punkten wesentlich durch den Aufwand zur Aufnahme der ca. 250.000 Bilder verursacht wird. Durch die Projektion von senkrechten Streifen anstelle von Lichtpunkten lässt sich diese Anzahl auf 512 Bilder verringern. Da Lichtpunkte mit unterschiedlichem Y -Wert in Kamerakoordinaten auf unterschiedliche Orte im Bild abgebildet werden (wegen $y_p = (f/d)Y$), besteht keine Verwechslungsgefahr.

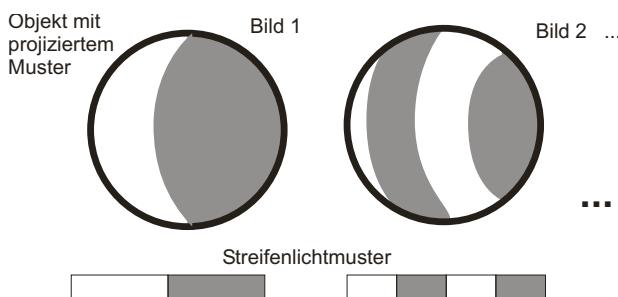


Abbildung 5.5: Bei einer codierten Streifenlichtprojektion wird eine Folge von unterschiedlichen Streifenmustern auf das Objekt projiziert. Die Hell-Dunkel-Folge über alle Bilder ordnet jedes Pixel eindeutig einem Streifen zu.

Der Aufwand kann weiter verringert werden, wenn die Signale entlang der X-Achse codiert werden. Statt eines schmalen (ideal linienförmigen) Streifens wird zunächst ein Streifen der Breite $K/2$ Pixel projiziert, wobei K die Breite der Aufnahme ist (siehe Abbildung 5.5). Damit kann zwischen zwei Bereichen „beleuchtet“(1) und „unbeleuchtet“(0) unterschieden werden. Allerdings ist die Positionsbestimmung in X-Richtung mit $\pm K/2$ sehr ungenau. Deshalb wird ein zweites Bild mit Streifenbreite $K/4$ erzeugt. Die Abfolge der Streifen ist hier „1010“. Aus dem zweiten Bild allein lässt sich nicht entnehmen, ob ein beleuchteter Streifen der erste oder der dritte Streifen ist. Zusammen mit dem ersten Bild ergeben sich jedoch vier verschiedene Codes (11, 10, 01, 00), die jeden der vier Streifen eindeutig identifizieren. Diese Codierung kann so lange fortgesetzt werden, bis die Streifenbreite genügend schmal ist. Wäre $K/512$ die Zielgröße, dann bräuchte man nun anstelle von 512 Bildern nur noch $\log_2 512 = 9$ Bilder.

Das Verfahren heißt Active Vision durch **strukturiertes Licht**. Für die Verarbeitung werden p Bilder mit unterschiedlicher Streifenbreite aufgenommen. In jedem Bild wird entschieden, welche Pixel beleuchtet und welche unbeleuchtet waren. Anschließend wird für jedes Pixel die „beleuchtet-unbeleuchtet“-Folge bestimmt. Aus dieser ergibt sich direkt die Streifennummer einer Folge von Streifen der Breite $K/2^p$.

5.2 Reduktion von Abtastfehlern

Wenn eine potentiell beliebig fein aufgelöste Szene durch eine digitale Kamera aufgenommen wird, dann geht durch die Abtastung Information zwischen den Pixeln verloren. Wichtigstes Ziel für die weitere Verarbeitung ist es, wahrnehmbare Verfälschungen durch den Informationsverlust zu verhindern. Für manche Zwecke ist es auch notwendig, die verloren gegangene Information zu ergänzen. Das ist z.B. nach manchen geometrischen Transformationen – die auf einem reellen Definitionsbereich definiert sind – der Fall.

5.2.1 Aliasing und Anti-Aliasing-Methoden

In Abschnitt 3.1.2 wurde Abtastung durch eine Multiplikation der Bildfunktion mit einer Impulsfolge modelliert. Nach Abtastung z.B. durch einen Flachbettscanner wird man ein Phänomen bemerken, das **Moiré-Muster** genannt wird (siehe Abbildung 5.6). Im abgetasteten Bild wurde Information offenbar nicht nur reduziert, sondern auch verfälscht. Den gleichen Effekt kann man auch bei Kameraaufnahmen beobachten. In Fernsehaufnahmen von sehr kleinteiligen Mustern – z.B. der Musterung eines Anzugs – können ebenfalls Moiré-Effekte wahrgenommen werden. Die Muster ändern sich mit der Bewegung des Objekts. Das gibt einen ersten Hinweis auf die mögliche Ursache. Die relative Änderung bei einem bewegten Objekt bedeutet, dass der Moiré-Effekt durch Abtastung verursacht worden sein muss. Da der Effekt nur bei kleinteiligen, also hochfrequenten Anteilen auftritt, wird die Betrachtung des Abtastvorgangs im Frequenzraum hilfreich sein.

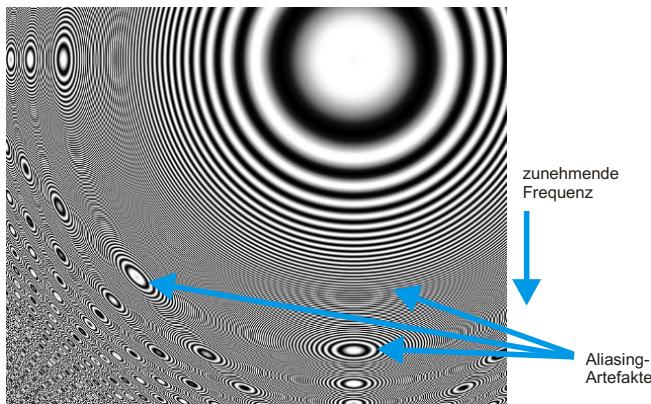


Abbildung 5.6: Moiré-Muster entstehen durch zu große Abtastintervalle in Abhängigkeit von der Frequenz des aufgenommenen Musters. Nur die konzentrischen Kreise um das Zentrum oben rechts sind tatsächlich in den Daten enthalten. Oben rechts im Bild – wo das Bild keine hochfrequenten Anteile aufweist – sind keine Moiré-Effekte zu erkennen, während sie mit steigender Frequenz stark zunehmen.

Wir haben die Abtastung einer kontinuierlichen Funktion f mit einer Abtastdichte d durch Multiplikation mit einer entsprechenden Impulsfolge δ_d als $f_d = f \cdot \delta_d$ modelliert. Diese Multiplikation entspricht einer Konvolution mit der fouriertransformierten Impulsfolge D_d im Frequenzbereich, d.h. es gilt $F_d = F^* D_d$. Die Impulsfolge erfüllt zwar nicht die Bedingungen für transformierbare Funktionen, doch kann die Fourier-Transformation durch eine Grenzwertbetrachtung bestimmt werden. Wird die Impulsfolge durch immer schmäler werdende Rechtecke einer Rechteckimpulsfolge angenähert, so ergibt sich, dass auch die fouriertransformierte Impulsfolge wieder eine Impulsfolge ist. Der Abstand zwischen den Impulsen im Frequenzraum beträgt $1/d$, so dass $D_d = \delta_{1/d}$ ist (siehe Abbildung 5.7).

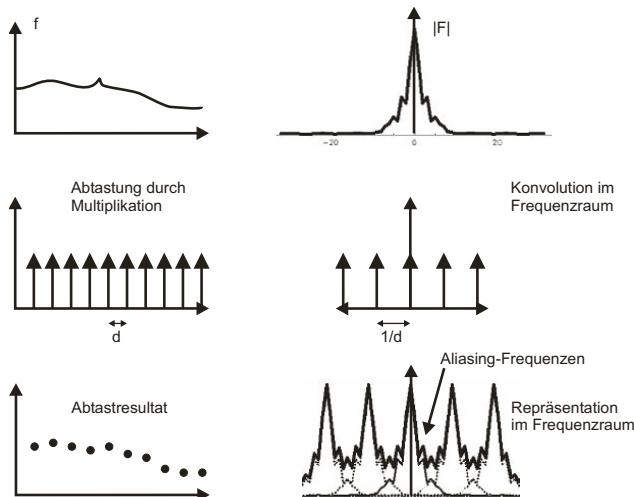


Abbildung 5.7: Abtastung in Orts- und Frequenzraum. Der Multiplikation der Funktion mit der Impulsfolge mit Abstand d zwischen Impulsen im Ortsraum entspricht eine Konvolution mit einer Impulsfolge mit Abstand $1/d$ im Frequenzraum. Das führt zur Überlagerung von Kopien der Originalfunktion, welche dann Alias-Effekte verursachen, wenn die Originalfunktion nicht bandbegrenzt ist oder die Bandbegrenzung nicht dem Shannon'schen Abtasttheorem genügt.

Mit dieser Impulsfolge wird F gefaltet. Das Ergebnis ist eine Summe von Kopien der um jeweils $1/d$ verschobenen Funktion F . Die Frequenzraumdarstellung der abgetasteten Funktion weicht also von der Originalfunktion ab. Da die Amplituden mit wachsender Frequenz für die meisten Bilder schnell gegen Null gehen, ist diese Veränderung durch die Abtastung nicht immer wahrnehmbar. Weil außerdem nur ein begrenztes Frequenzband im abgetasteten Bild repräsentierbar ist, haben die um Vielfache von $1/d$ verschobenen Maxima keinen Einfluss auf das Bild.

Gibt es aber im Bild signifikante Anteile mit hoher Frequenz, wie sie z.B. durch ein kräftiges, kleinteiliges, regelmäßiges Muster verursacht werden können, dann kann es zu sichtbaren Veränderungen von Frequenzen kommen. Sie entstehen bei der Abtastung durch Überlappung von verschobenen Kopien von F . Diese Frequenzen werden **Alias-Frequenzen** genannt (weil sie für etwas posieren, was im Originalbild nicht existiert, siehe Abbildung 5.8) und der Effekt heißt **Aliasing-Effekt**.

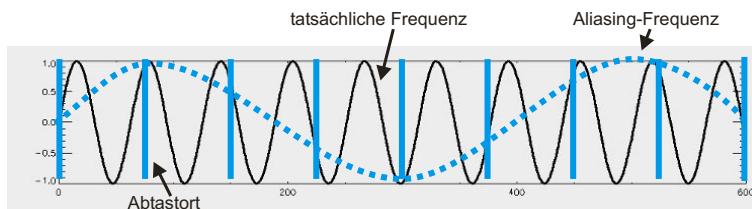


Abbildung 5.8: Darstellung der Entstehung des Aliasing-Artefakts. Die Originalfunktion (schwarz, durchgezogen) wird an den blau gekennzeichneten Orten abgetastet. Verbindet man die Funktionswerte an den Abtastorten durch eine Kurve (blau, gestrichelt), so entsteht eine andere Funktion. Diese Funktion wird durch den Betrachter wahrgenommen.

Es gibt zwei unterschiedliche Ansätze, Aliasing-Effekte bei der Generierung von Bildern zu vermeiden:

- 1** Durch Erhöhung der Abtastrate können Überlappungen von Kopien im Frequenzraum vermieden werden, falls die Originalfunktion bandbegrenzt ist. Das heißt, dass höhere Frequenzen als eine vorgegebene Grenzfrequenz nicht vorkommen.
- 2** Unregelmäßige Abtastabstände verhindern die Wahrnehmbarkeit von periodischen Signalen. Das Vorgehen besteht eigentlich aus einer zweistufigen Abtasthierarchie. Unterhalb eines bestimmten Abstands (der Grenzfrequenz) sind die Abtastorte zufällig verteilt. Oberhalb dieses Abstands ist die Summe der Abtastorte in jedem Intervall gleich. Das führt dazu, dass periodische Signale oberhalb der Grenzfrequenz wahrgenommen werden. Alle Details unterhalb der Grenzfrequenz werden als Rauschen wahrgenommen. Das entspricht der Wahrnehmung durch das Auge (Verteilung der Zellen auf der Netzhaut) oder der Aufnahme mit analogem Film (die Körnung der lichtempfindlichen Pigmente). Auch manche Algorithmen zur Generierung von Bildern (z.B. bestimmte Ray-Tracing-Algorithmen) wenden dieses Prinzip an.

Für eine digitale Kamera oder für einen Flachbettscanner ist die zuletzt genannte Lösung leider nicht anwendbar, obwohl sie für die Wahrnehmung die beste Lösung darstellt, denn wahrgenommenes Rauschen ist leichter interpretierbar als der informationsverfälschende Moiré-Effekt. Bei der Konstruktion von Bildern – und wir werden im Zusammenhang mit Bildkompression noch einmal darauf zurückkommen – kann die Strategie allerdings angewendet werden.

Für die Abtastung durch die Kamera ist dagegen der erste Weg interessanter. Es stellt sich die Frage, was die geeignete Grenzfrequenz ist. Die Antwort gibt das **Shannon'sche Abtasttheorem**, das aus der Kenntnis der Impulsfunktion im Frequenzraum abgeleitet werden kann. Wenn der Abstand zwischen zwei Impulsen $1/d$ ist, überlappen sich zwei Kopien der fouriertransformierten Funktion nicht, wenn für die höchstvorkommende Frequenz f_{max} gilt, dass $2f_{max} < 1/d$ ist. Also muss die Abtastrate d mindestens $\frac{1}{2} \cdot f_{max}$ sein, um Aliasing-Effekte zu verhindern. Die für die Abtastung eines bandbegrenzten Signals mindestens notwendige Grenzfrequenz heißt **Nyquist-Frequenz**.

5.2.2 Transformationen und Interpolation

Manchmal, etwa nach geometrischen Transformationen, ist es notwendig, den durch Abtastung verursachten Informationsverlust rückgängig zu machen. Rotation Rot_α um den Winkel α im Uhrzeigersinn um den Koordinatenursprung, Translation $Tr_{dx,dy}$ um den Vektor $(dx\ dy)$ und Skalierung Sc_s um s eines Bildpunkts mit Koordinaten (x,y) können wie folgt ausgeführt werden:

$$Rot_\alpha \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad Tr_{dx,dy} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + dx \\ y + dy \end{pmatrix}, \quad Sc_s \begin{pmatrix} x \\ y \end{pmatrix} = s \begin{pmatrix} x \\ y \end{pmatrix}.$$

Die Werte in der Matrix sind reell. Eine Transformation ist daher nur möglich, wenn die ganzzahligen Abtastkoordinaten vor der Transformation in reelle Koordinaten umgewandelt werden.

Das Resultat sind reelle Koordinaten. Nach der Transformation müsste die Originalfunktion an diesen Orten abgetastet werden. Diese Originalfunktion wird jedoch selten zur Verfügung stehen. So müssen Funktionswerte an beliebigen Orten (x,y) aus der abgetasteten Funktion interpoliert werden. Im einfachsten Fall kann man eine konstante Interpolation vornehmen. Für eine eindimensionale Funktion $f(n)$, die das Abtastergebnis einer reellen Funktion $g(x)$ ist, ist das

$$g_0(x) = f(n_x) \text{ mit } n_x = \arg \min_n (|x - n|).$$

Das bedeutet, dass für $g_0(x)$ derjenige Funktionswert $f(n)$ verwendet wird, für den $|x - n|$ minimal ist.

Wendet man das Verfahren für zweidimensionale Bilder an, so sollten die Pixelkoordinaten transformiert und anschließend sollte für jeden Abtastort nach dem nächstgelegenen Pixel gesucht werden. Der Wert dieses Pixels sollte übernommen werden. Transformiert man dagegen die Abtastorte, bildet die beiden Koordinaten auf die nächste ganze Zahl ab und schreibt den Pixelwert an diese Stelle, so kann es bei Rotationen dazu kommen, dass an manche Abtastorte keine Pixel abgebildet werden. Das liegt daran, dass die Diagonale eines Pixels größer als seine Seitenlänge ist. Das Ergebnisbild enthält dann schwarze Löcher (siehe Abbildung 5.9).

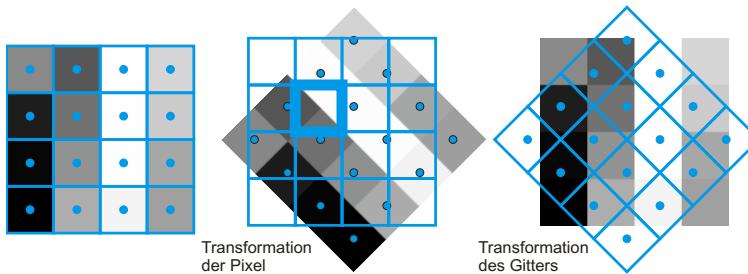


Abbildung 5.9: Falls zur Interpolation nach einer Transformation der Funktionswert des nächsten Nachbarpixels eingegeben wird, dann sollte das Gitter und nicht die Funktion transformiert werden. Sonst kann es geschehen, dass auf einen Gitterpunkt (hervorgehobenes Quadrat im mittleren Bild) kein Funktionswert abgebildet wird.

Die konstante Interpolation ist ein schnell berechenbares, einfach zu implementierendes Verfahren. Allerdings muss bedacht werden, dass es einer zweiten Abtastung mit gleicher Abtastfrequenz entspricht. Nach dem Shannon'schen Abtasttheorem kann es daher zu Aliasing-Effekten kommen, wenn nicht vorher alle Frequenzen gelöscht werden, die größer als $f_{max}/2$ sind.

Für eine genauere Interpolation kann man die ursprüngliche Funktion, falls sie beliebig häufig differenzierbar war, durch Polynome höheren Grades annähern. Jede solche Funktion lässt sich als eine Taylor-Reihe entwickeln; sie ist also durch alle ihre Ableitungen an einer beliebigen Stelle vollständig beschreibbar. Nun sind zwar Funktionswerte, nicht aber Ableitungen bekannt. Ableitungen werden daher durch Differenzen angenähert. Für die erste Ableitung einer eindimensionalen Funktion, deren in den reellen Definitionsbereich transformierte Werte an den Stellen $x_0, x_1, \dots, x_m, \dots, x_{M-1}$ definiert sind, entsteht für die Berechnung an der Stelle x zwischen den Orten x_m und x_{m+1} eine lineare Interpolationsfunktion

$$g_1(x) = \frac{x_{m+1} - x}{x_{m+1} - x_m} f(x_m) + \frac{x - x_m}{x_{m+1} - x_m} f(x_{m+1}).$$

Die lineare Interpolation ist weniger anfällig gegenüber Aliasing-Effekten. Durch die Gewichtung werden hohe Frequenzen vor der Abtastung unterdrückt.

Wenn die lineare Interpolation für zweidimensionale Bilder durchgeführt werden soll, kann man sich die Separabilität der Operation zunutze machen. Für einen Ort (x, y) , der zwischen Pixeln an $(x_m, y_n), (x_{m+1}, y_n), (x_m, y_{n+1})$ und (x_{m+1}, y_{n+1}) liegt, werden zunächst Werte an (x_m, y) und (x_{m+1}, y) interpoliert (siehe Abbildung 5.10):

$$g_1(x_m, y) = \frac{y_{n+1} - y}{y_{n+1} - y_n} f(x_m, y_n) + \frac{y - y_n}{y_{n+1} - y_n} f(x_m, y_{n+1}),$$

$$g_1(x_{m+1}, y) = \frac{y_{n+1} - y}{y_{n+1} - y_n} f(x_{m+1}, y_n) + \frac{y - y_n}{y_{n+1} - y_n} f(x_{m+1}, y_{n+1}).$$

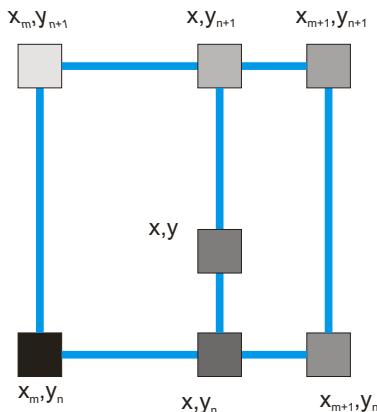


Abbildung 5.10: Die bi-lineare Interpolation zwischen Pixeln wird in zwei Schritten ausgeführt: Im ersten Schritt wird entlang der x-Richtung interpoliert. Aus diesen Werten wird anschließend der endgültige Wert durch Interpolation in y-Richtung berechnet.

Anschließend wird aus diesen beiden Werten der Wert für $g_1(x,y)$ berechnet:

$$g(x,y) = \frac{x_{m+1} - x}{x_{m+1} - x_m} g_1(x_m, y) + \frac{x - x_m}{x_{m+1} - x_m} g_1(x_{m+1}, y).$$

Anstelle von Polynomen vom Grad 1 kann man auch Polynome höheren Grades benutzen. So werden immer höhere Terme der Taylor-Reihe berücksichtigt und die tatsächliche Funktion wird besser angenähert. Allerdings wird der Fehler der Schätzung der entsprechenden Ableitungen durch Differenzen größer.

Man kann sich fragen, ob es eine für die gegebene Information optimale Interpolation gibt. Unter der Annahme, dass das Shannon'sche Abtasttheorem erfüllt ist, die abgetastete Funktion also keine höheren Frequenzen enthielt als die Grenzfrequenz, lässt sich diese Frage beantworten. Dann ist nämlich die vollständige Beschreibung der Funktion an jedem Ort durch die Wellen der fouriertransformierten Funktion gegeben. Da im Frequenzraum also alle Frequenzen der kontinuierlichen Funktion repräsentiert sind, bedeutet das für den Funktionswert an einem beliebigen Punkt (x,y) :

$$g_F(x,y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) \exp\left(i2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right)$$

Bei dieser „optimalen“ Interpolation sollte man allerdings beachten, dass die Funktion f , von der F generiert wurde, in der Regel Rauschen beinhaltet, was ebenfalls mit rekonstruiert wird. Außerdem ist die reelle Funktion selten wirklich bandbegrenzt. Ein gewisses Maß an Aliasing-Artefakten fließt daher ebenfalls mit in die Interpolation ein. Es mag aus diesen Gründen besser sein, eines der einfacheren Interpolationsverfahren zu wählen. Die dort implizit oder explizit vorgesehene Entfernung hoher Frequenzanteile wird wahrscheinlich diejenigen Anteile der Funktion unterdrücken, für die Artefakte die Information überwiegen.

5.3 Bildrestaurierung

Unter Bildrestaurierung versteht man, deterministische, bekannte, verändernde Störinflüsse rückgängig zu machen. Das zugrunde liegende Modell für die Störung ist das eines linearen Operators. Der Prozess kann als lineares Gleichungssystem $\bar{y} = A\bar{x}$ repräsentiert werden. Das gestörte Bild wird dazu in einen Vektor \bar{y} überführt, in dem alle Pixel des Bildes hintereinander gelistet sind. Das gesuchte ungestörte Bild ist ein Vektor \bar{x} , bei dem die Pixel in gleicher Weise angeordnet sind. Da die Anzahl der Pixel vor und nach der Störung gleich ist, muss die Störmatrix A quadratisch sein. Sie ist invertierbar, wenn sie vollen Rang hat, wenn also die Determinante der Matrix nicht Null ist.

Unter den Modellen für Störinflüsse unterscheidet man zwischen verschiebungsinvarianten und nicht verschiebungsinvarianten Störungen, so wie sie in *Abschnitt 3.2.1* modelliert wurden. In beiden Fällen ist die Invertierbarkeit von A nicht garantiert, aber für verschiebungsinvariante Störoperatoren gibt es einfache Methoden zur Ermittlung der Störung aus Testbildern. Es gibt zudem ein einfaches und leicht zu regularisierendes Invertierungsverfahren.

5.3.1 Restaurierung von verschiebungsinvarianten Störungen

In *Abschnitt 3.2.1* wurde gezeigt, dass sich ein verschiebungsinvarianter, linearer Operator durch die Point Spread Function (PSF) beschreiben lässt. Das Modell für die Störung ist durch eine Faltung der ungestörten Funktion f mit dem Störoperator h gegeben:

$$g(m, n) = f(m, n) * h(m, n).$$

Der Operator h ist die PSF. Kann man den Aufnahmevergang, durch den das Signal f verändert wird, in all seinen Einzelheiten beschreiben, so lässt sich der Störoperator analytisch herleiten. Das haben wir in *Kapitel 3* getan, als wir die PSF für eine Störung berechnet haben, die durch eine Bewegung der Kamera verursacht wurde.

Experimentelle Bestimmung der Point Spread Function

Oft ist es so, dass die verschiedenen Einflüsse bei der Aufnahme nicht im Einzelnen bekannt sind. Das Aufnahmesystem ist eine Black Box, in der eine wiederholbare Veränderung des Eingangssignals erfolgt. Für eine lineare, verschiebungsinvariante Veränderung muss das Verhalten der Black Box nicht im Detail bestimmt werden, um die PSF h zu ermitteln. Es genügt die Aufnahme einer Impulsfunktion. Das gemessene Aufnahmegergebnis ist die PSF, da die Konvolution mit einer Impulsfunktion die Konvolutionsfunktion selbst ergibt. Die PSF eines Aufnahmesystems wird daher auch **Impuls- oder Punktantwort (impulse response)** genannt.

Wenn auch eine Impulsfunktion selten in der Praxis generierbar ist, so kann sie doch auf verschiedene Art angenähert werden. Existiert z.B. eine Testaufnahme, in der nur an einem Ort in einer gewissen Umgebung von Null verschiedene Werte liegen, und ist diese Umgebung in der Projektion kleiner als ein Pixel, dann ist dies eine hinreichend gute Näherung an die Impulsfunktion. Die Testfunktion kann etwa generiert werden, wenn ein gleichmäßig beleuchtetes Stück Papier mit einem einzigen, kleinen schwarzen Punkt aufgenommen wird. Dass die Approximation des Impulses in diesem Fall negativ ist, spielt für die Bestimmung von h keine Rolle. Es bedeutet



Copyright

Daten, Texte, Design und Grafiken dieses eBooks, sowie die eventuell angebotenen eBook-Zusatzdaten sind urheberrechtlich geschützt. Dieses eBook stellen wir lediglich als persönliche Einzelplatz-Lizenz zur Verfügung!

Jede andere Verwendung dieses eBooks oder zugehöriger Materialien und Informationen, einschliesslich

- der Reproduktion,
- der Weitergabe,
- des Weitervertriebs,
- der Platzierung im Internet, in Intranets, in Extranets,
- der Veränderung,
- des Weiterverkaufs
- und der Veröffentlichung

bedarf der schriftlichen Genehmigung des Verlags.

Insbesondere ist die Entfernung oder Änderung des vom Verlag vergebenen Passwortschutzes ausdrücklich untersagt!

Bei Fragen zu diesem Thema wenden Sie sich bitte an: info@pearson.de

Zusatzdaten

Möglicherweise liegt dem gedruckten Buch eine CD-ROM mit Zusatzdaten bei. Die Zurverfügungstellung dieser Daten auf unseren Websites ist eine freiwillige Leistung des Verlags. Der Rechtsweg ist ausgeschlossen.

Hinweis

Dieses und viele weitere eBooks können Sie rund um die Uhr und legal auf unserer Website



herunterladen