

Seminar 3D-Rekonstruktion

Gewinnung von 3D-Informationen durch Strukturierte Beleuchtung

**Universität Mannheim
Lehrstuhl für Praktische Informatik IV
Prof. Dr. Wolfgang Effelsberg**

Simon Schütz

Betreuer: Dirk Farin

Datum: 16. Dez. 2002

Inhaltsverzeichnis

Inhaltsverzeichnis	2
Einleitung.....	3
1 Verfahren zur Berechnung von Distanzwerten	4
1.1 Projektion einfacher geometrischer Muster	4
1.1.1 Light Spot Techniken.....	4
1.1.1.1 Light Spot Projection	4
1.1.1.2 Light Spot Stereo Analysis.....	8
1.1.2 Light Stripe Projection.....	9
1.1.3 Static Light Pattern Projection	12
1.2 Projektion kodierter Muster.....	12
1.2.1 Binary Encoded Light Stripes	13
1.2.2 Colour Encoded Light Stripe Projection.....	14
1.2.3 Active Colour Stereo Analysis	15
2 Verfahren zur Berechnung der Oberflächenform.....	15
Zusammenfassung	19
Literaturverzeichnis	20

Einleitung

Die 3D-Rekonstruktion beschäftigt sich mit dem Problem Computer-Modelle aus realen 3D-Objekten zu erzeugen. Die nötigen Informationen sollen durch die Analyse von Videoaufnahmen oder Standbildern gewonnen werden. Die Videoaufnahmen und Standbilder müssen in digitaler Form vorliegen, sodass sie vom Computer verarbeitet werden können.

Die somit erhaltenen Modelle können danach in verschiedenen Anwendungen eingesetzt werden, wie etwa in der Computeranimation, der Filmproduktion oder auch der Generierung von virtuellen Welten.

Diese Arbeit beschäftigt sich mit den speziellen Möglichkeiten die sich durch den Einsatz von strukturierter Beleuchtung ergeben. Man benutzt Lichtmuster mit bekannter Struktur um die Objekte zu beleuchten (z.B. mit LCD-Projektoren oder Lasern). Beispiele hierfür wären Streifen-, Punkt- oder Ringmuster. Man probiert nun die Muster in den damit gewonnenen Standbildern wieder zu erkennen und Rückschlüsse auf die geometrische Form des zu betrachtenden Objekts zu gewinnen.

Die hier behandelten Verfahren lassen sich in zwei Gruppen aufteilen. Zum einen Verfahren, die echte Distanzwerte berechnen um das Modell zu beschreiben. Das Modell wird somit in Originalgröße bekannt sein. Ebenso lässt sich die Position relativ zur Kamera bestimmen. Zum anderen gibt es Verfahren, die die Oberflächenform berechnen, ohne sich für die wahre Größe oder Position des Objekts zu interessieren.

1 Verfahren zur Berechnung von Distanzwerten

1.1 Projektion einfacher geometrischer Muster

Die Wiedererkennung der projizierten Muster in den Aufnahmen und die damit gewonnenen Informationen sind das Ziel der strukturierten Beleuchtung. Nutzt man nur einfache Muster wie Punkte oder Linien, dann ist die Wiedererkennung recht einfach. Die folgenden Verfahren nutzen ausschließlich solche einfachen Muster.

1.1.1 Light Spot Techniken

Die „Light Spot“ oder „Lichtpunkt“ Technik verwendet einen einfachen Lichtstrahl, der einen Lichtpunkt auf dem Objekt generiert. Denkbar wäre der Einsatz eines Lasers oder eines Projektors mit vorgeschalteter Lochmaske.

1.1.1.1 Light Spot Projection

Die „Light Spot Projection“ ist eine sehr einfache Methode. Da lediglich ein Lichtstrahl verwendet wird, entsteht auch nur ein Lichtpunkt im Bild. Dieser lässt sich sehr einfach in einem Bild lokalisieren, z.B. über die Suche von extremen Helligkeitswerten. Wird der Lichtstrahl durch einen Laser erzeugt, kann man auch spezielle Kamerafilter einsetzen. Ist der Lichtpunkt im Bild lokalisiert, berechnet man einen Projektionsstrahl, der vom Projektionszentrum der Kamera durch den Lichtpunkt im Bild verläuft. Ist der Aufbau des Versuchs bekannt, so sind auch die Verläufe dieses Projektionsstrahls und des Lichtstrahls bekannt. Durch Berechnung des Schnittpunktes der beiden Strahlen erhält man dann die Position des Lichtpunktes auf der Objektoberfläche. Der Versuchsaufbau sowie die geometrischen Eigenschaften werden in Abb. 1 skizziert.

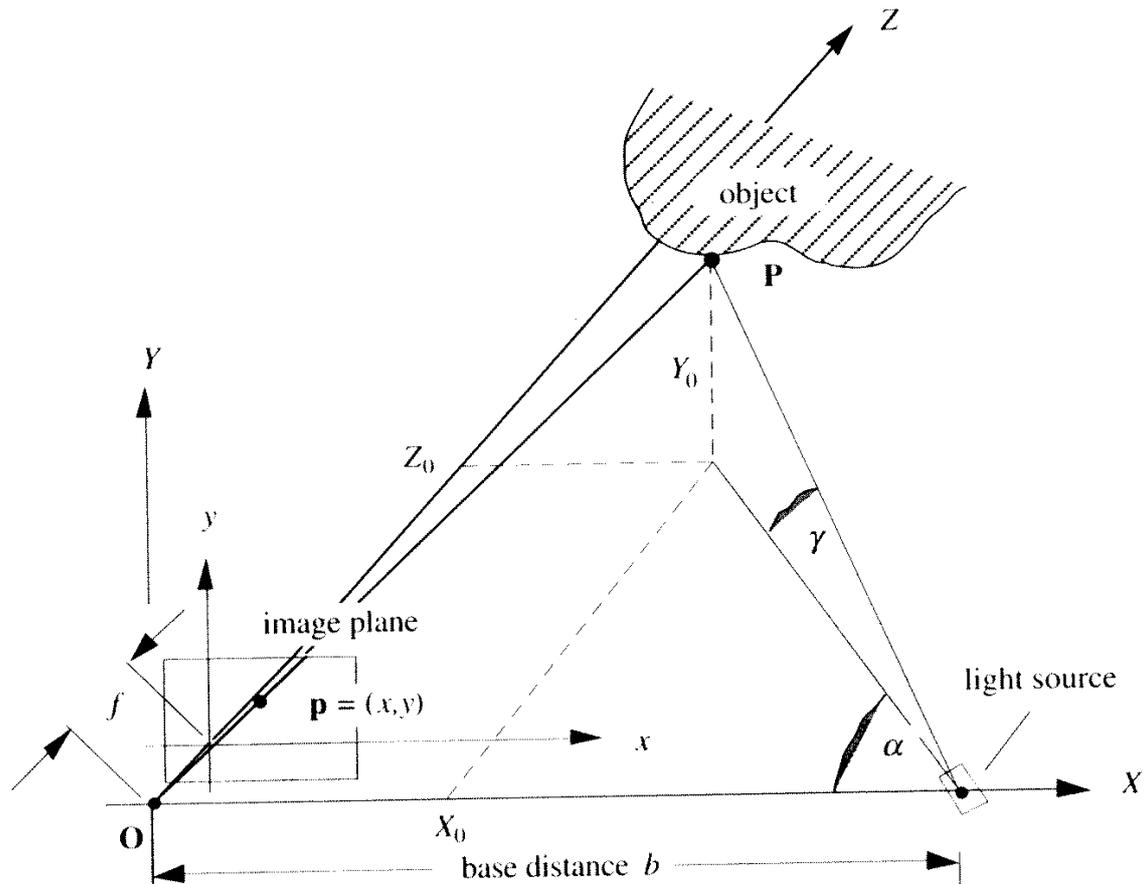


Abb. 1: Geometrie der Light Spot Projection

Für die Berechnung werden folgende Bezeichnungen verwendet:

O ist das Projektionszentrum der Kamera sowie der Nullpunkt eines gedachten Kamera-Koordinatensystems. Alle Positionsangaben im 3D-Raum erfolgen relativ zu diesem Koordinatensystem. $\mathbf{P} = (X_0, Y_0, Z_0)$ ist der Punkt in dem der Lichtstrahl das Objekt trifft. $\mathbf{p} = (x, y)$ ist der Punkt im Bild, an dem der Lichtpunkt sichtbar wird. \mathbf{b} ist die Distanz zwischen Kamera und Lichtquelle. Die Brennweite der Kamera wird mit \mathbf{f} bezeichnet und entspricht auch dem Abstand zwischen dem Kamerazentrum und der gedachten Bildfläche. Der Aufbau wird so gewählt, dass das Zentrum der Lichtquelle auf der x -Achse des Kamera-Koordinatensystems liegt. Dadurch lässt sich die Position der Lichtquelle als $(b, 0, 0)$ beschreiben. Der Winkel α ist der Winkel zwischen x -Achse des Kamera-Koordinatensystems und der orthogonalen Projektion des Lichtstrahls in die x - z -Ebene. Der Winkel γ ist der Winkel zwischen Lichtstrahl und seiner Projektion in die x - z -Ebene. Zur Berechnung der Distanzwerte wird die im englischen mit „triangulation“ (deutsch: Dreiecksbildung) bezeichnete Technik verwendet. Hierbei sucht man nach rechtwinkligen Dreiecken in der Szene, mit deren Hilfe man über die Trigonometrie die gesuchten Werte errechnen kann.

Aufgrund des Strahlensatzes lassen sich folgende Verhältnisse erkennen:

$$\frac{X_0}{x} = \frac{Y_0}{y} = \frac{Z_0}{f} \quad (\text{I.1})$$

Durch Umstellung der äußeren Terme aus (I.1) erhält man auch:

$$Z_0 = \frac{X_0}{x} \cdot f \quad (\text{I.2})$$

Mithilfe der Trigonometrie ergibt sich für den Tangens des Winkel α

$$\tan(\alpha) = \frac{Z_0}{b - X_0} \quad (\text{I.3})$$

oder durch Auflösung nach Z_0 auch

$$Z_0 = \tan(\alpha) \cdot (b - X_0) \quad (\text{I.4})$$

Aus (I.2) und (I.4) ergibt sich nun

$$\frac{X_0}{x} \cdot f = \tan(\alpha) \cdot (b - X_0) \quad (\text{I.5})$$

und durch Umformung

$$X_0 \left(\frac{f}{x} + \tan(\alpha) \right) = \tan(\alpha) \cdot b \quad (\text{I.6})$$

Multipliziert man beide Seiten von Gleichung (I.6) mit x , so ergibt sich die Berechnungsformel für X_0 , die nun nur noch aus messbaren Parametern besteht:

$$X_0 = \frac{\tan(\alpha) \cdot b \cdot x}{f + x \cdot \tan(\alpha)} \quad (\text{I.7})$$

Die Formeln für Y_0 und Z_0 lassen sich durch Ersetzung von X_0 in Gleichung (I.7) errechnen. Hierzu ersetzt man X_0 entsprechend zu (I.1) durch

$$X_0 = \frac{Y_0}{y} \cdot x \quad \text{bzw.} \quad X_0 = \frac{Z_0}{f} \cdot x$$

Löst man die entstehenden Gleichung nach Y_0 bzw. Z_0 auf, so erhält man

$$Y_0 = \frac{\tan(\alpha) \cdot b \cdot y}{f + x \cdot \tan(\alpha)} \quad (\text{I.8})$$

und

$$Z_0 = \frac{\tan(\alpha) \cdot b \cdot f}{f + x \cdot \tan(\alpha)} \quad (\text{I.9})$$

Aus den Formeln (I.7)-(I.9) wird ersichtlich, dass man die Position des Lichtpunktes $P=(X_0, Y_0, Z_0)$ im Raum errechnen kann, sobald die Geometrie des Versuchsaufbaus (α , b , f) als auch die Koordinaten des Bildpunktes $p=(x, y)$ bekannt sind.

Es werden aber auch gleich einige Probleme dieses Verfahrens offensichtlich. Für eine genaue Bestimmung der Position (X_0, Y_0, Z_0) ist erforderlich, dass die Geometrie des Versuchsaufbaus sehr genau vermessen werden muss. Speziell die Brennweite f und der Winkel α dürften hier zu Problemen führen. Des Weiteren muss der zur Erkennung des Bildpunktes p verwendete Algorithmus zuverlässig und exakt funktionieren. Es stellt sich hierbei die Frage, ob die Bestimmung evtl. auf sub-pixel Genauigkeit erfolgen müsste, da

die Entfernung des Objektes zur Kamera meist ein Vielfaches der Brennweite ist. Ein kleiner Messfehler würde folglich bei der Projektion entsprechend vergrößert.

Die Light Spot Projection speichert die gewonnenen Informationen nicht als 3D-Koordinaten, da dies nicht speichereffizient wäre. Stattdessen speichert sie für den Bildpunkt (x,y) den Abstand $d(P,O)$ des Punktes (X_0,Y_0,Z_0) zum Ursprung des Kamera-Koordinatensystems in einem sog. Range Image. Auf diese Weise wird pro Aufnahme nur ein Wert gespeichert statt eines Triples. Der Abstand $d(P,O)$ lässt sich folgendermaßen errechnen:

$$d(P,O) = \|P - O\| = \|(X_0, Y_0, Z_0)\| = \sqrt{X_0^2 + Y_0^2 + Z_0^2}$$

Umgekehrt lässt sich mithilfe eines Range Images später wieder die Position im 3D-Raum errechnen. Hierzu benötigt man die Größe der gedachten Bildfläche als auch die Brennweite der aufnehmenden Kamera. Es wird ein Strahl vom Ursprung durch den entsprechenden Bildpunkt erzeugt. Die gesuchte Position ist dann auf diesem Strahl mit der Entfernung zum Ursprung, die im Range Image gespeichert ist.

Bis hier wurde nur beschrieben, wie man **eine** Kameraaufnahme auszuwerten hat. Jede Auswertung liefert Informationen zu lediglich **einem** Punkt auf der Objektoberfläche. Um eine detaillierte Information über die Objektoberfläche zu erhalten, muss die Prozedur entsprechend oft mit verschiedenen Winkeln α und γ wiederholt werden, bis die gewünschte Auflösung erreicht ist.

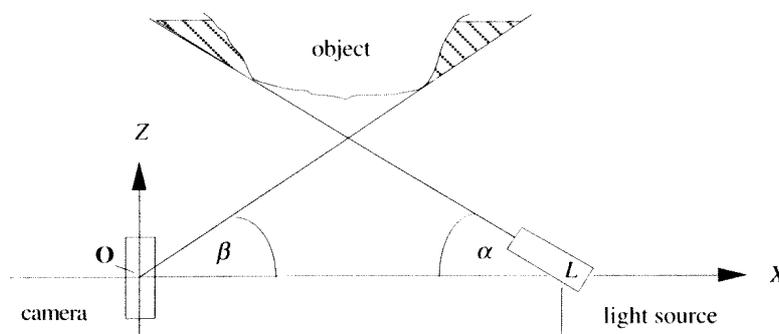


Abb. 2: Entstehung von Schatten und Verdeckungen

Vorteil der Light Spot Projection ist die sehr einfache Auswertung einer Aufnahme. Sie erreicht ihre Grenzen allerdings in folgenden Punkten:

- Die Genauigkeit hängt sehr stark von der Güte der Messung der Eingabeparameter ab.
- Eine Erhöhung der Auflösung erfordert eine entsprechende (quadratische) Erhöhung der Anzahl von Kameraaufnahmen und könnte so schnell zu Effizienzproblemen führen, wenn die Einrichtung einer Aufnahme, die Messung der Parameter oder die Erkennung des Bildpunktes nicht schnell genug

durchgeführt werden kann.

- Das Range Image kann evtl. durch Einflüsse von Schattenbildung oder Unsichtbarkeit nicht vollständig erstellt werden (vgl. Abb.2).

Die o. g. Probleme führen dazu, dass die Light Spot Projection im allgemeinen nicht ausreichend ist.

1.1.1.2 Light Spot Stereo Analysis

Wie bereits oben beschrieben, steigt und fällt die Qualität der Light Spot Projection mit der Genauigkeit der Eingabeparameter, die oft nicht leicht zu messen sind. Die Light Spot Stereo Analysis ist eine Kombination der Light Spot Projection mit der sog. Static Stereo Analysis. Der Vorteil besteht darin, dass die genaue Ausrichtung der Lichtquelle die Güte der Messwerte nicht beeinflusst.

Die Static Stereo Analysis verwendet zwei Kameras, die dieselbe Szene von verschiedenen Positionen aufnehmen. „Static“ bedeutet hierbei, dass sich die Objekte nicht bewegen, sodass beide Kameras auch wirklich die gleiche Szene aufnehmen. Denkbar ist auch, dass beide Kameras zur exakt gleichen Zeit die Bilder aufnehmen. In diesem Fall nehmen die beiden Kameras auch dann noch die gleiche Szene auf, wenn die Objekte nicht statisch sind. Ziel der Static Stereo Analysis ist nun, sich entsprechende Punkte in den beiden gewonnenen Bildern zu finden. Durch die unterschiedlichen Bildpositionen, in denen sie sichtbar werden, kann man erneut durch Triangulation die 3D-Position des Lichtpunktes auf der Objektoberfläche errechnen. Der Versuchsaufbau wird aus Abb. 3 ersichtlich.

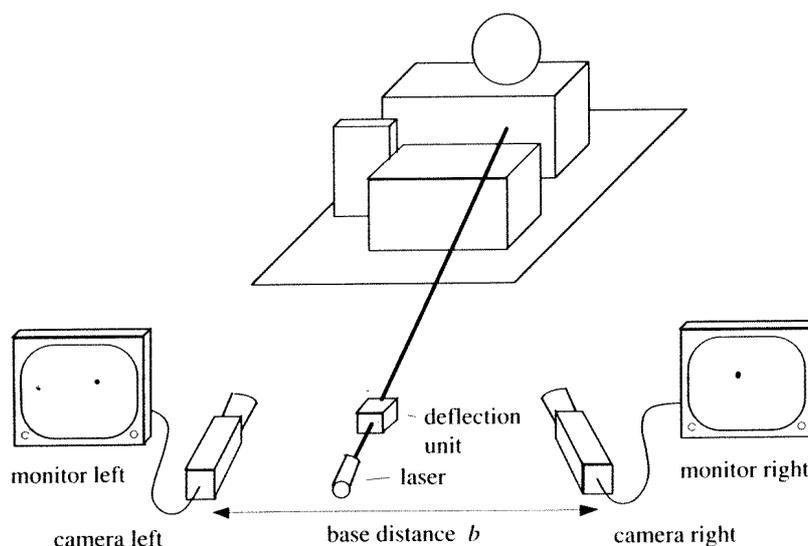


Abb. 3: Aufbau der Light Spot Stereo Analysis

Die Light Spot Technik erleichtert die Static Stereo Analysis ungemein, da sich das Finden von sich entsprechenden Stellen auf das Auffinden des vom Laser erzeugten Lichtpunktes

beschränkt. Diese können dann eindeutig einander zugewiesen werden. Der Vorteil gegenüber der Light Spot Projection besteht darin, dass die Ausrichtung des Lasers unerheblich ist, da die 3D-Information rein aus dem Versatz des Lichtpunktes im einen Bild gegenüber dem im anderen Bild errechnet wird. Natürlich muss auch noch die Ausrichtung der Kameras bekannt sein, aber diese ändert sich – im Gegensatz zur Ausrichtung des Lasers oder dessen „deflection unit“ – nicht zwischen verschiedenen Aufnahmen. Es genügt also den Kameraaufbau **einmalig** genau zu vermessen. Bei der Light Spot Projection müsste entweder eine sehr exakte „deflection unit“ verwendet werden, oder falls keine deflection unit verwendet wird, sondern der Laser selbst gedreht wird, so muss dessen Ausrichtung jedes Mal vermessen werden.

1.1.2 Light Stripe Projection

Anstatt eines Lichtstrahls wird bei der Light Stripe Projection eine Lichtfläche in die Szene projiziert, die im Bild einen Lichtstreifen generiert. Der Vorteil gegenüber den vorigen Verfahren besteht darin, dass man entlang des Streifens mehrere Berechnungen pro Bild durchführen kann anstatt nur einer.

Beim Aufbau ist darauf zu achten, dass die Lichtfläche parallel zur y -Achse des Kamerakoordinatensystems ist. Für jeden Bildpunkt auf dem Streifen bestimmt man den Abstand zu einer gedachten Referenzebene, die senkrecht zur Lichtfläche ist. Setzt man weiterhin voraus, dass sich das optische Zentrum der Kamera in der Referenzebene befindet, so lässt sich der Aufbau wie in Abb. 4 skizziert beschreiben.

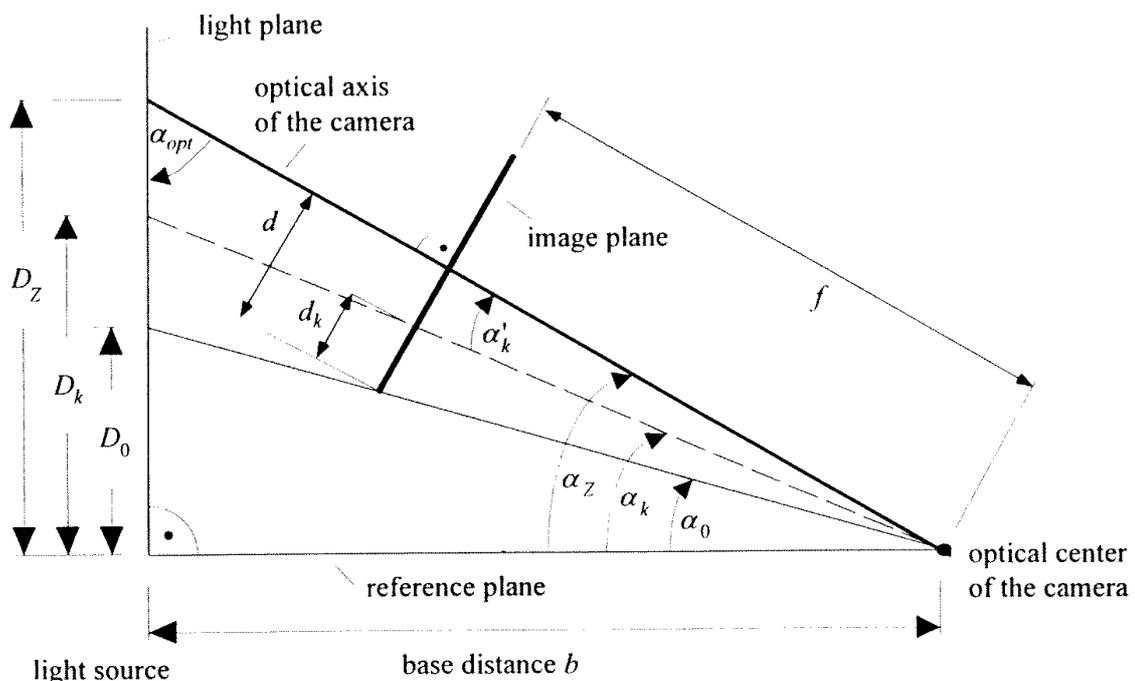


Abb. 4: Geometrie der Light Stripe Projection

Denkbar wäre bei diesem Aufbau eine Projektion des Bildpunktes vom optischen Zentrum

der Kamera auf die tatsächliche Objektoberfläche. Dies würde allerdings zu Ungenauigkeiten führen, da sich die Geometrie der Kamera, speziell die Brennweite f , als auch die Winkel nur schwer exakt bestimmen lassen. Außerdem ist das Objekt im allgemeinen um ein Vielfaches der Brennweite vom optischen Zentrum entfernt. Dies würde zu einer entsprechenden Erhöhung der Ungenauigkeit führen. Daher werden durch ein Kalibrierungsverfahren die Werte D_Z und D_0 bestimmt, anstatt sie mithilfe des Strahlensatzes zu errechnen. Man stellt ein planares Objekt parallel zur Referenzebene auf und verschiebt es so lange, bis der im Bild sichtbare Lichtstreifen genau auf der mittleren Spalte des aufgenommenen Bildes liegt. Der Abstand, den das Objekt zu diesem Zeitpunkt zur Lichtquelle hat, entspricht exakt der Distanz D_Z . Nun verschiebt man das Objekt in Richtung Lichtquelle, bis der Lichtstreifen auf der ganz linken Spalte des Bildes sichtbar ist. Der jetzige Abstand entspricht D_0 . Misst man zusätzlich zu dieser Kalibrierung noch den Abstand b der Lichtquelle zur Kamera, so kann zu jeder Spalte im Bild ein entsprechender Abstand des Punktes im 3D-Raum zur Referenzebene berechnet werden. Die entsprechende Formel lässt sich wieder durch Triangulation herleiten. Für die Winkel α_Z und α_0 gilt:

$$\alpha_Z = \arctan\left(\frac{D_Z}{b}\right) \quad (\text{II.1})$$

$$\alpha_0 = \arctan\left(\frac{D_0}{b}\right) \quad (\text{II.2})$$

Somit lassen sich diese beiden Winkel aus leicht messbaren Daten errechnen. Es ergibt sich weiterhin aus dem Strahlensatz:

$$2 \cdot d = 2 \cdot f \cdot \tan(\alpha_Z - \alpha_0) \quad (\text{II.3})$$

Unter Anwendung der allgemeinen Formel

$$\tan(x \pm y) = \frac{\tan(x) \pm \tan(y)}{1 \mp \tan(x) \cdot \tan(y)} \quad (\text{II.4})$$

sowie (II.1) und (II.2) lässt sich (II.3) folgendermaßen ausdrücken:

$$d = f \cdot \frac{(D_Z - D_0) \cdot b}{b^2 + D_Z \cdot D_0} \quad (\text{II.5})$$

Angenommen das digitale Bild hat M Spalten, dann ist die Distanz d_k zwischen Spalte k und der Spalte 0, also der mittleren Spalte, genau

$$d_k = k \cdot \frac{d}{M/2} = \frac{2kd}{M} \quad (\text{II.6})$$

Weiterhin gilt:

$$\tan(\alpha'_k) = \frac{d - d_k}{f} = \frac{d - \left(\frac{2kd}{M}\right)}{f} = \frac{d(M - 2k)}{M \cdot f} \quad (\text{II.7})$$

Unter Verwendung von (II.5) lässt sich die Brennweite f ersetzen und es folgt:

$$\tan(\alpha'_k) = \frac{(D_Z - D_0) \cdot b \cdot (M - 2k)}{M \cdot (b^2 + D_Z \cdot D_0)} \quad (\text{II.8})$$

Im letzten Schritt soll nun die Distanz D_k berechnet werden. Dies entspricht dem Abstand eines Punktes zur Referenzfläche, wenn er im Bild in Spalte k sichtbar wird.

$$D_k = b \cdot \tan(\alpha_k) = b \cdot \tan(\alpha_z - \alpha'_k) \quad (\text{II.9})$$

Entsprechend (II.8) lässt sich α'_k ersetzen:

$$D_k = b \cdot \tan\left(\alpha_z - \arctan\left(\frac{(D_z - D_0) \cdot b \cdot (M - 2k)}{M \cdot (b^2 + D_z \cdot D_0)}\right)\right) \quad (\text{II.10})$$

Löst man nun noch α_z und die (Arcus-)Tangensfunktionen auf, so ergibt sich die Formel für D_k als

$$D_k = \frac{(D_z^2 + b^2)MD_0 + 2k(D_z - D_0)b^2}{(D_z^2 + b^2)M - 2k(D_z - D_0)D_z} \quad (\text{II.11})$$

Das Ergebnis sieht recht kompliziert aus, beinhaltet aber lediglich leicht messbare Größen. Bei genauerem Hinsehen ist auch zu erkennen, dass bis auf die Spaltennummer k alle Größen Konstanten sind, die nur einmalig beim Aufbau vermessen werden müssen. Somit können bereits große Teile des Terms vorberechnet und gespeichert werden.

Die Auswertung eines Bildes ergibt also Tiefeninformationen für alle auf der Linie befindlichen Punkte. Um ein komplettes 3D-Modell zu erhalten, muss das Objekt mehrfach um einen gewissen Winkel gedreht werden, und pro Drehung ein weiteres Bild ausgewertet werden. Achtet man darauf, dass die Drehachse sowohl in der Lichtfläche als auch parallel zur Referenzebene ist, dann kann das Objekt in zylindrischen Koordinaten gespeichert werden. Zur Rekonstruktion des 3D-Profiles werden sich entsprechende Punkte auf benachbarten Linien miteinander verbunden. Ein Beispiel für dieses Verfahren ist in Abb. 5 und Abb. 6 zu sehen.

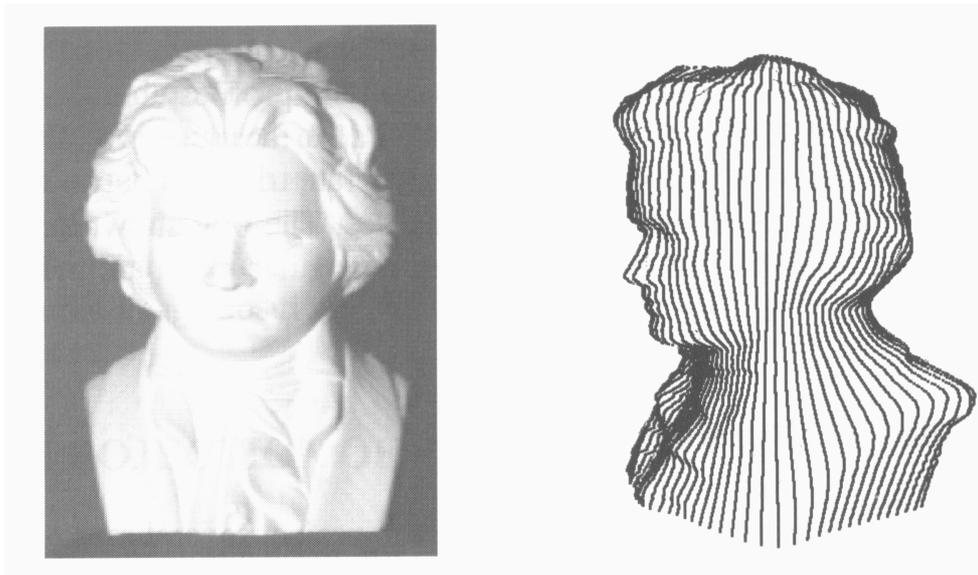


Abb. 5: links: Bild des Objekts, rechts: vermessene Profillinien

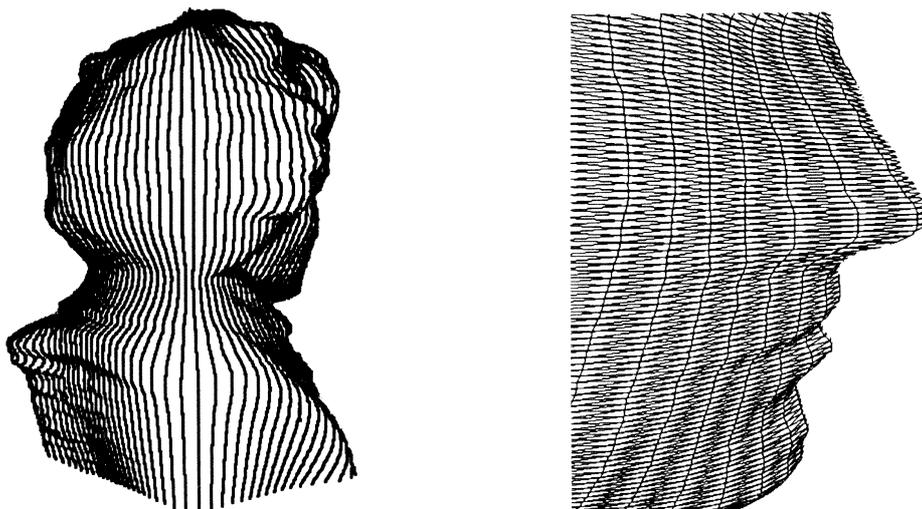


Abb. 6: Rekonstruktion des 3D-Profiles

1.1.3 Static Light Pattern Projection

Die bisher behandelten Verfahren haben alle das gleiche Problem: Sie benötigen eine hohe Anzahl von Bildern um die gesamte Oberfläche eines Objekts zu vermessen. Diese Anzahl kann verringert werden, indem man gleichzeitig mehrere Strahlen oder Linien auf das Objekt projiziert, sodass man mit einer Aufnahme mehrere Auswertungen durchführen kann. Jede einzelne Auswertung kann wie bereits in den vorigen Kapiteln beschrieben erfolgen. Gepunktete oder gestrichelte Linien, parallele Linien, konzentrische Kreise oder auch Punktmatrizen sind oft verwendete Muster.

Das Problem bei dieser Methode ist das eindeutige Auffinden von Linien oder Punkten, wenn gleich mehrere von ihnen im Bild vorhanden sind. Denkbar wäre der Einsatz von Nachbarschaftsbeziehungen. Als Beispiel könnte man bei einem Streifenmuster die Streifen von links nach rechts durchnummerieren. Dies funktioniert allerdings nicht mehr, wenn es zu Schattenwürfen oder Verdeckungen im Bild kommt, da dann der nächste sichtbare Nachbar in Wirklichkeit kein Nachbar ist.

1.2 Projektion kodierter Muster

Offensichtlich ist die genaue Wiedererkennung des Originalmusters im Bild äußerst wichtig für die Auswertung. Die bisher aufgeführten Verfahren verwenden nur einfache Muster, die dies erleichtern. Allerdings liefern sie daher auch nur ein relativ geringes Maß an Information. Man benötigt folglich viele Bilder, um ein genaues 3D-Modell zu errechnen. Wünschenswert wäre der Einsatz komplexer Muster, die mehr Informationsgehalt pro Bild generieren und trotzdem noch eindeutig erkennbar sind. Die im folgenden betrachteten Muster enthalten eine Kodierung, um diese sichere Erkennung zu gewährleisten.

1.2.1 Binary Encoded Light Stripes

Bei der Binary Encoded Light Stripes Technik zur Kodierung der einzelnen Lichtflächen (jede erzeugt einen Lichtstreifen auf dem Objekt) werden mehrere Bilder des zu betrachtenden Objektes aufgenommen. Bei jeder Aufnahme kann jede der Lichtflächen einen der beiden Zustände „an“ oder „aus“ haben. Es handelt sich daher um eine binäre Kodierung. Durch mehrere Aufnahmen erzeugt man für jede Lichtfläche einen sog. Bit-plane Stack, d.h. man speichert für jede Lichtfläche und jedes Bild den jeweiligen Zustand. Angenommen wir nehmen n Bilder auf, eine Lichtfläche kann zwei Zustände einnehmen, folglich lassen sich bis zu 2^n Lichtflächen mit den n Bildern eindeutig kodieren. Die Erzeugung des Bit-plane Stacks soll in Abb. 7 verdeutlicht werden. Eine häufig verwendete Kodierungsart ist der „Gray Code“. Der Gray Code ist eine binäre Kodierungstechnik bei der sich zwei **aufeinander folgende** Wörter in **genau einem** Bit unterscheiden.

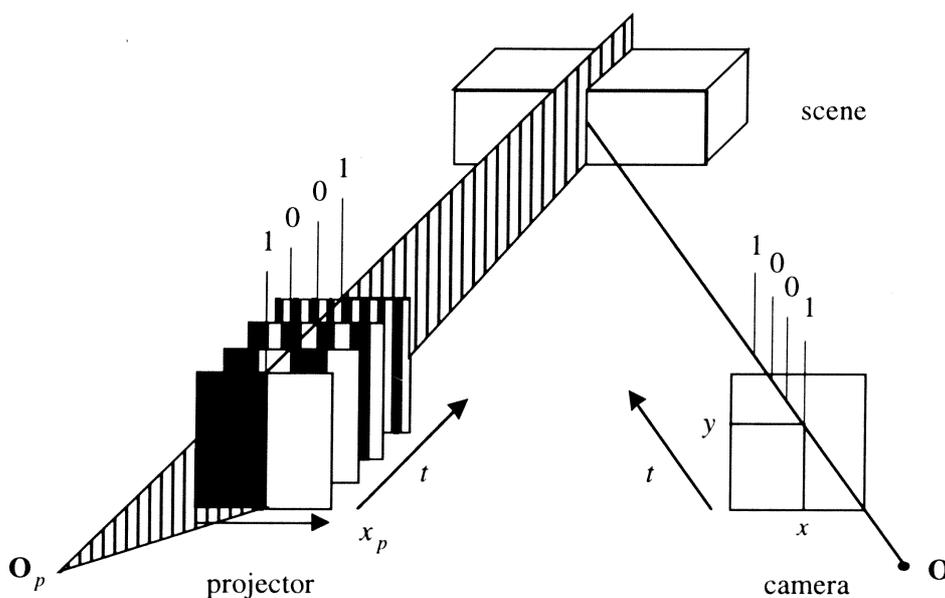


Abb. 7: Erzeugung des Bit-plane Stacks

Möchte man nun die Position eines Punktes im 3D-Raum berechnen, der in den Bildern sichtbar ist, so überprüft man in **jedem** Bild, ob der Punkt beleuchtet oder nicht beleuchtet ist. Man erhält entsprechend ein n Bit langes Codewort. Der Punkt liegt genau in der Lichtfläche, die die gleiche Codierung besitzt. Somit lässt sich jede Position des Bildes eindeutig einer Lichtfläche zuordnen. Danach kann eine Berechnung der Position des Punktes bzw. dessen Distanz zur Referenzfläche genau wie bei der Light Stripe Projection durchgeführt werden. Allerdings muss vor den Aufnahmen eine Kalibrierung erfolgen, um die genauen Ausrichtungsdaten der Lichtflächen zu erhalten. Hierzu kann man zum Beispiel zwei planare Objekte mit bekannter Orientierung beleuchten. Durch eine

Rückwärtsrechnung kann man dann die Ausrichtung der Lichtflächen erhalten.

1.2.2 Colour Encoded Light Stripe Projection

Bis auf die Static Light Pattern Projection benötigen alle bisherigen Verfahren mehrere, zu **verschiedenen Zeitpunkten** durchgeführte, Aufnahmen der **gleichen** Szene. Dies ist aber nicht möglich, falls die Szene nicht statisch ist. Die Colour Encoded Light Stripe Projection kodiert die einzelnen Lichtflächen mit verschiedenen Farbwerten, sodass mit **nur einer** Aufnahme bereits alle Lichtflächen eindeutig identifizierbar sind. Mit dieser Technik kann man auch nicht statische Objekte vermessen, vorausgesetzt die Bewegung ist noch langsam genug um eine scharfe Aufnahme mit der Kamera zu erlauben. Es soll hier allerdings auch gleich auf eine Einschränkung hingewiesen werden. Die Objekte sollten entweder eine weiße oder graue Oberfläche besitzen, da sonst die Farbwerte durch Absorption verfälscht werden. Dadurch würde die Wiedererkennung der Farbkodierung erheblich erschwert oder würde zumindest ein Mindestmaß an Vorwissen über das zu vermessende Objekt voraussetzen.

Die Kodierungsmuster werden in einzelne Subpatterns zerlegt, die auch als Colour Code bezeichnet werden. Für die Erstellung eines Kodierungsmuster müssen folgende Faktoren betrachtet werden:

- Anzahl der Lichtflächen
- Anzahl der Lichtflächen pro Subpattern (im weiteren K)
- Anzahl der Subpatterns
- Anzahl der verwendeten Farben (im weiteren L)

Benachbarte Lichtflächen eines Subpatterns sollten immer zwei verschiedene Farben haben, da sie sonst nicht unterscheidbar sind. Daraus resultiert eine Anzahl M_0 von möglichen Colour Codes von

$$M_0(K, L) = L \cdot (L - 1)^{K-1}$$

Die erste Farbe ist frei wählbar (L Möglichkeiten), jede der K-1 weiteren muss immer eine andere Farbe als der jeweilige Vorgänger sein (je L-1 Möglichkeiten).

Weiterhin ist darauf zu achten, dass die verwendeten Subpatterns so gewählt werden, dass der letzte Streifen des einen Subpatterns eine andere Farbe besitzt als der erste Streifen des nächsten Subpatterns.

Bei der Auswertung einer Aufnahme wird ein Pfad senkrecht zu den entstandenen Lichtstreifen beschriftet. Ist ein Subpattern komplett sichtbar, da es keine Verdeckungen gab, so ist die Erkennung entsprechend einfach. Probleme treten nur bei Verdeckungseffekten auf. Diese können allerdings meist durch Nachbarschaftsbeziehungen gelöst werden. Ist ein Subpattern komplett erkannt, so weiß man auch, welche Farben bzw. welches Subpattern in den angrenzenden Gebieten sichtbar

sein sollten. Sind in dem angrenzenden Subpattern einige Streifen verdeckt, so kann man aufgrund des Wissens über das Originalpattern trotzdem die sichtbaren Streifen eindeutig identifizieren.

1.2.3 Active Colour Stereo Analysis

Die Active Colour Stereo Analysis ist eine Erweiterung der Colour Encoded Light Stripe Projection. Das Problem der genauen Identifikation der Kodierungsmuster soll vereinfacht werden. Dazu wird das Bild mit zwei Kameras aus verschiedenen Positionen aufgenommen. Danach erfolgt mithilfe einer Korrespondenzanalyse die Zuordnung von Punkten im einen Bild zu Punkten im anderen Bild. Ist diese Zuordnung möglich, so kann die Positionsbestimmung des Punktes im 3D-Raum durch den relativen Versatz des Bildpunktes in den beiden Bildern errechnet werden. Hierzu ist es irrelevant, von welcher Lichtfläche der Punkt beleuchtet wurde. Somit sind Verdeckungsprobleme leichter zu lösen.

2 Verfahren zur Berechnung der Oberflächenform

Im folgenden wird ein Verfahren von S. Winkelbach und F. M. Wahl [2] vorgestellt, das die Oberflächennormalen der Objekte approximiert. Sind die Oberflächennormalen an hinreichend vielen Stellen bekannt, so kann die Oberflächenform entsprechend genau angenähert werden.

Das hier beschriebene Verfahren verwendet ein einfaches Streifenmuster, bei dem sowohl helle als auch dunkle Streifen die gleiche Breite haben. Zur Berechnung der Oberflächennormalen (im weiteren nur noch „Normalen“ genannt) werden die im Bild sichtbaren Streifenbreiten und Streifenorientierung benutzt. Sowohl aus der Breite als auch aus der Orientierung wird jeweils eine Oberflächentangente errechnet. Bildet man das Kreuzprodukt aus zwei Oberflächentangenten, so erhält man die zugehörige Normale. Das Bild wird in echten Graustufen aufgenommen um die Bestimmung von Richtung und Breite der Streifen auf sub-pixel Genauigkeit zu ermöglichen. Durch die Nutzung der lokalen Streifenbreite und -richtung entfallen Korrespondenzprobleme, da es irrelevant ist, durch welche Lichtfläche der lokale Lichtstreifen erzeugt wurde.

Die Streifenrichtung wird mithilfe einer Gradientenberechnung bestimmt. Die Bestimmung der Streifenbreite erfolgt auf sub-pixel Genauigkeit, indem man kleine Bereiche um lokale Gradientenmaxima untersucht. Auf der rechten Seite von Abb. 8 ist ein solcher Bereich abgebildet. In x-Richtung ist die Position, in y-Richtung der Graustufenwert abgetragen. Um die genaue Position der Kante zu bestimmen, wählt man a_i derart, dass die graue

Fläche rechts von a_i der weißen Fläche links von a_i entspricht. Diese Approximation führt zu einer besseren Auflösung als die eigentliche Pixelauflösung.

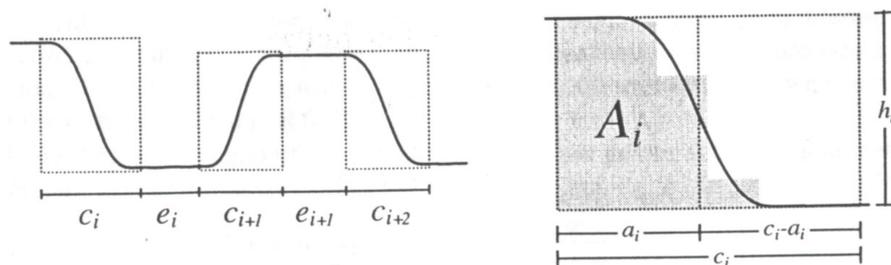


Abb. 8: links: Verlauf der Graustufenwerte, rechts: Approximation der Linienposition

Aus Abb. 8 wird ersichtlich, dass

$$\frac{a_i}{c_i} = \frac{A_i}{c_i \cdot h_i} \Leftrightarrow a_i = \frac{A_i}{h_i}$$

gilt. Die Breite des Streifens i lässt sich dann als

$$d_i' = (c_i - a_i) + e_i + a_{i+1}$$

berechnen.

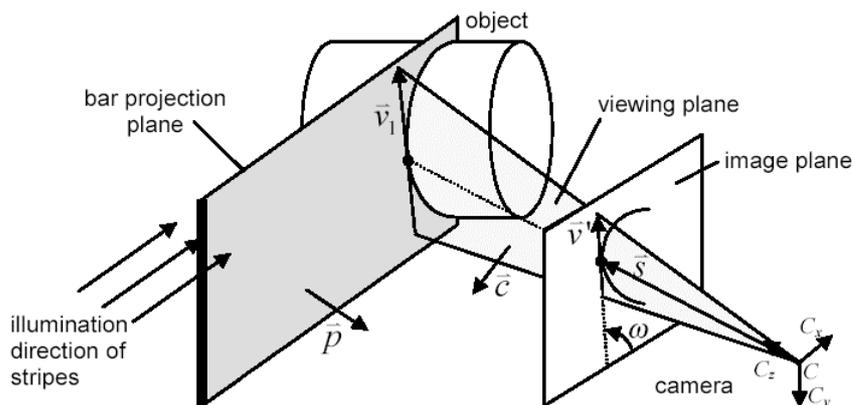


Abb. 9: Projektion der Streifenrichtung in die Bildebene

Zur Berechnung des ersten Tangentialvektors v_1 mithilfe der Streifenrichtung führen wir entsprechend zu Abb. 9 die folgenden Vektoren ein. s ist der sog. view vector, der vom Kamerazentrum zum Pixel in der Bildebene zeigt. Er liegt in der viewing plane. v' liegt in der viewing plane und in der image plane. Daher lässt sich die Normale der viewing plane als

$$c = v' \times s$$

bestimmen. Der gesuchte Tangentialvektor v_1 liegt ebenfalls in der viewing plane und steht somit senkrecht auf c . Außerdem liegt v_1 auch in der bar projection plane, der Ebene in der die Lichtstreifen projiziert werden. Deren Normale p lässt sich aus der Geometrie des Aufbaus berechnen. Unter der Annahme dass die bar projection planes der Lichtstreifen parallel sind, haben auch alle die gleiche Normale p . Der Tangentialvektor kann nun als

$$v_1 = c \times p = (v' \times s) \times p$$

bestimmt werden.

Der zweite Tangentialvektor v_2 wird über die Streifenbreite errechnet. Hierbei bedient man sich einer einfachen 2D-Beziehung zwischen der Originalbreite und der Bildbreite der Streifen. d bezeichnet hier die Originalbreite des Streifens, wie er durch die Lichtquelle erzeugt wird, d' bezeichnet die Streifenbreite wie er durch die Kamera wahrgenommen wird (vgl. Abb. 10).

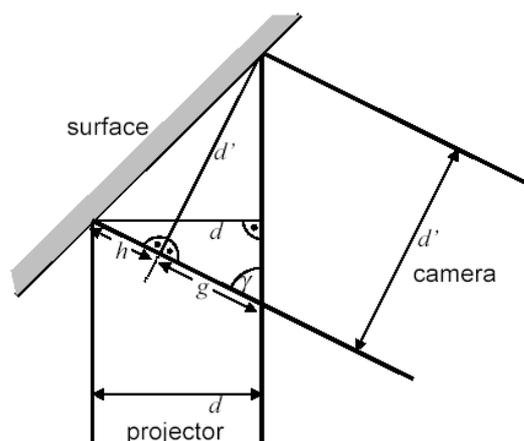


Abb. 10: Berechnung des Gradienten über das Verhältnis von Original- zu Bildbreite der Streifen

Es soll nun der Oberflächengradient $p = h/d'$ bestimmt werden. Über die Trigonometrie ergibt sich:

$$h + g = \frac{d}{\sin(\gamma)} \quad \text{und} \quad g = \frac{d'}{\tan(\gamma)}$$

Entsprechend folgt für h

$$h = \frac{d}{\sin(\gamma)} - \frac{d'}{\tan(\gamma)} = \frac{d - d' \cdot \cos(\gamma)}{\sin(\gamma)}$$

und für den Gradienten p

$$p = \frac{h}{d'} = \frac{\frac{d}{\sin(\gamma)} - \cos(\gamma)}{\sin(\gamma)}$$

Der zweite Tangentialvektor v_2 wird als

$$v_2 = \begin{pmatrix} 1 \\ 0 \\ -p \end{pmatrix}$$

definiert.

Die Oberflächennormale für den entsprechenden Bildpunkt ergibt sich nun als Kreuzprodukt der beiden Tangentialvektoren:

$$n = v_1 \times v_2$$

Experimente mit dieser Technik haben ergeben, dass das Verfahren im allgemeinen recht

gute Ergebnisse liefert. Probleme treten hauptsächlich an harten Kanten auf. Dies ist recht einsichtig, da die Oberfläche durch Integration über die Normalen angenähert wird. An einer Kante ist die eine Normale entsprechend der einen Fläche ausgerichtet, die benachbarte Normale entsprechend der angrenzenden Fläche. Dies führt bei der Integration zu einer Abrundung der Kante. Weitere Ungenauigkeiten treten an Flächen auf, welche relativ parallel zur optischen Achse der Kamera sind. Hier erscheinen die Streifen im Bild wesentlich dünner und somit ist die Bildauflösung nicht mehr ausreichend.

Zusammenfassung

Die Gewinnung von 3D-Informationen lässt sich wie oben beschrieben durch die Verwendung von strukturiertem Licht vereinfachen. Die normalerweise recht aufwendige Korrespondenzanalyse zwischen verschiedenen Bildern wird entweder vereinfacht oder entfällt vollständig. Die einfachsten Verfahren (Lightspot/-stripe Projectionen) beruhen auf primitiven geometrischen Aspekten, haben aber meist den Nachteil, dass sie viele Bilder benötigen und recht sensibel auf die Vermessung des Versuchsaufbaus sind. Die unterliegenden Algorithmen sind allerdings sehr einfach und schnell zu berechnen. Zur Erhöhung der Auflösung und zur Reduzierung der notwendigen Bilder muss zu den etwas aufwendigeren Verfahren mit Kodierung übergegangen werden. Deren Auswertungsalgorithmen sind zwar etwas aufwendiger, liefern aber auch wesentlich mehr Informationen, sodass der Mehraufwand mehr als ausgeglichen wird. Wird die Messung von beweglichen Objekten gewünscht, so muss die Kodierung entsprechend gut sein (z.B. Colour Encoded Light Stripes), um die gesamte Information aus einem einzigen Bild gewinnen zu können. Auch das von S. Winkelbach und F. M. Wahl beschriebene Verfahren ermöglicht die Messung mit nur einem einzigen Bild, stellt allerdings wesentlich geringere Anforderungen an die verwendeten Projektoren. Welches Verfahren nun wirklich verwendet werden sollte, hängt von dem jeweiligen Einsatzbereich ab. Man sollte vorab festlegen, unter welchen Bedingungen die Messungen stattfinden (z.B. bewegliche oder statische Szene, bunte oder einfarbige Objekte), wie hoch die Qualität der Messung sein sollte (angenähert oder realitätsnah) und welche Hardware zur Verfügung steht (Kamera und Projektor/Laser). Entsprechend lassen sich dann die Vor- und Nachteile der einzelnen Verfahren gegeneinander abwägen.

Literaturverzeichnis

- [1] Computer Vision
Three-Dimensional Data from Images
Reinhard Klette, Karsten Schlüns, Andreas Koschan
Springer Verlag Singapore, 1998

- [2] Shape from Single Stripe Pattern Illumination
S. Winkelbach und F. M. Wahl
2002