

Informationstheoretische Grenzen optischer 3D-Sensoren

Der Technischen Fakultät der
Universität Erlangen-Nürnberg
zur Erlangung des Grades

DOKTOR-INGENIEUR

vorgelegt von
Christoph Wagner
Erlangen, 2003

Als Dissertation genehmigt von
der Technischen Fakultät der
Universität Erlangen-Nürnberg

Tag der Einreichung: 24.01.2003
Tag der Promotion: 14.03.2003
Dekan: Prof. Dr. A. Winnacker
1. Berichterstatter: Prof. Dr. P. Steffen
2. Berichterstatter: Prof. Dr. G. Häusler

Inhaltsverzeichnis

Abstract	3
1 Aufgabenstellung und Zusammenfassung	5
2 Einleitung	7
3 Optische Grundlagen	9
3.1 Beleuchtung	9
3.1.1 Kohärenz der Beleuchtung	10
3.2 Objekt	11
3.2.1 Polarisations-eigenschaften	12
3.3 Beobachtung	12
3.4 Kohärentes Rauschen durch Speckle	13
3.5 Photonenrauschen	15
4 Informationstheoretische Grundlagen	17
4.1 Das Übertragungsmodell	18
4.2 Ein Maß für die Information	18
4.3 Mittlere Transinformation und Entropie	19
4.4 Transinformation bei stetigen Zufallsvariablen	20
4.5 Differentielle Entropie	20
4.6 Kanalkapazität	21
4.7 Codierung	22
5 Informationstheoretisches Sensormodell	23
5.1 Der Kanal	24
5.1.1 Kapazität des optischen Kanals	27
5.1.2 Kapazität des elektronischen Kanals	29
5.1.3 Kapazität des diskreten Kanals	31
5.1.4 Gesamtkanal	33
5.2 Realisierte Transinformation	35

5.3	Der Informationswirkungsgrad	36
5.4	Eigenschaften von Codierung und Quelle	38
5.5	Möglichkeiten der Codierung	44
5.6	Single Shot Sensoren	46
6	Optimierter optischer 3D-Sensor	51
6.1	Photometrisches Stereo	53
6.1.1	Vereinfachtes photometrisches Stereo	56
6.2	Messunsicherheit	57
6.2.1	Messunsicherheit zwischen benachbarten Pixeln	59
6.2.2	Gesamte Messunsicherheit	61
6.3	Integration der Neigungsdaten	62
6.4	Single Shot Sensor für einfarbige Textur	64
6.5	Photometrische Deflektometrie	66
7	Messergebnisse	69
7.1	Makroskopisches Messfeld	69
7.2	Informationswirkungsgrad	76
7.3	Erweiterter Informationswirkungsgrad	77
7.4	Formtreue	78
7.5	Messung mit Textur	80
7.6	Mikroskopisches Messfeld	80
7.7	Messunsicherheit	82
8	Zusammenfassung und Ausblick	87
	Danksagung	93
	Lebenslauf	94

Abstract

Most known optical range sensors require a large amount of 2D raw data from which the 3D data are decoded, associated with considerable “cost”. The cost arises from expensive hardware as well as from the time necessary to acquire the images. This contribution deals with the question of how to acquire a maximum of shape information with a minimum of image raw data. Information theory according to Shannon is an appropriate means to develop a sensor model. Using the term channel capacity, the model describes the maximum information that a sensor can provide. Besides the channel, the properties of the source of information (the shape of the object) are investigated. Typical sources show a large amount of redundancy which should be reduced by proper optical coding. Several options for analogue optical coding are discussed. Photometric stereo is an appropriate method to acquire shape data of low measurement uncertainty with a minimum of raw data. Based on this method an improvement called photometric deflectometry is introduced, which is able to measure not only diffusely reflecting surfaces, but also specular object surfaces. A large illumination aperture minimizes coherent noise and results in a measurement uncertainty of less than one micron. With the help of a color camera a single shot sensor is implemented, which can record the raw data simultaneously.

Kapitel 1

Aufgabenstellung und Zusammenfassung

Die meisten optischen 3D-Sensoren erfordern eine erhebliche Menge an zweidimensionalen Rohdaten (Kameraaufnahmen), aus denen anschließend die dreidimensionalen Formdaten gewonnen werden. Für diese Rohdaten ist ein hoher Preis zu zahlen. Dieser geht zurück auf teure Kameras und andere Hardware, aber auch auf die Tatsache, dass viel Zeit nötig ist, die Daten aufzunehmen. Diese Arbeit beschäftigt sich deshalb mit der Frage, wie man mit einem Minimum an Rohdaten 3D-Formdaten geringer Messunsicherheit gewinnen kann.

Die Informationstheorie nach Shannon ist ein geeignetes Mittel, diese Frage zu beantworten. Auf ihrer Grundlage wird ein informationstheoretisches Sensormodell entwickelt. Dieses Modell beschreibt mit dem Begriff der Kanalkapazität, wie viel Information ein Sensor maximal liefern kann. Neben dem Kanal werden auch die Eigenschaften der Informationsquelle, nämlich der dreidimensionalen Form des Objektes, untersucht. Die Untersuchungen zeigen, dass das Formsignal typischer Objekte ein hohes Maß an Redundanz enthält. Es ist ungünstig, ein redundantes Signal zu übertragen, da hierfür unnötig viel Kanalkapazität zur Verfügung gestellt werden muss. Daher stellt sich die Frage, ob es optische Codierungsmöglichkeiten gibt, die Redundanz des Formsignals wesentlich zu reduzieren.

Hierzu werden verschiedene Ansätze zur optischen Codierung diskutiert. Das Sensorprinzip des photometrischen Stereo erweist sich als vorteilhafte Methode und es wird gezeigt, dass 3D-Formdaten geringer Messunsicherheit mit einem Minimum an Rohdaten gewonnen werden können. Mit der Erweiterung des Verfahrens zur photometrischen Deflektometrie können neben diffus reflektierenden auch spiegelnde Objekte erfasst werden. Eine große Beleuchtungsapertur minimiert das kohärente Rauschen und ermöglicht eine

6 KAPITEL 1. AUFGABENSTELLUNG UND ZUSAMMENFASSUNG

Messunsicherheit von weniger als $1\ \mu\text{m}$. Mit Hilfe einer Farbkamera kann ein Single Shot Sensor realisiert werden, der alle Rohdaten gleichzeitig aufnimmt.

Kapitel 2

Einleitung

Eine wichtige Fragestellung in der optischen 3D-Messtechnik, wie auch in vielen anderen Bereichen der Wissenschaft, ist die Frage nach den fundamentalen Grenzen. Diese Frage ist aus zwei Gründen wichtig. Zum einen gewinnt man ein tieferes theoretisches Verständnis der Zusammenhänge, zum anderen helfen diese Grenzen in der Praxis weiter. Es ist unmöglich, eine fundamentale Grenze zu überschreiten, auch wenn ein immenser technologischer Aufwand betrieben wird. Die Technologie sollte dagegen versuchen, sich der Grenze mit vertretbarem Aufwand zu nähern.

In dieser Arbeit wird deshalb die Frage untersucht, wie viel zweidimensionale Rohdaten einer Kamera benötigt werden, um die Form von Objektoberflächen im Dreidimensionalen zu ermitteln. Diese informationstheoretische Fragestellung ist umso wichtiger, als die meisten optischen 3D-Sensoren ein großes Maß an 2D-Rohdaten benötigen, um schließlich dreidimensionale Daten zu gewinnen. Für diese Rohdaten ist ein hoher Preis zu zahlen. Dieser geht zurück auf teure Kameras und andere Hardware, aber auch auf die Tatsache, dass viel Zeit nötig ist, die Daten aufzunehmen. Es wird diskutiert, wie man mit weniger Rohdaten auskommen kann und dennoch 3D-Daten geringer Messunsicherheit gewinnt. Es wird ein Sensorprinzip vorgeschlagen, das auf diesen Überlegungen aufbaut.

Offensichtlich ist der Bedarf an Rohdaten verschiedener 3D-Sensoren ganz unterschiedlich. Stereoskopische Verfahren [1] benötigen lediglich zwei Rohbilder, um daraus 3D-Daten zu erzeugen. Aktive Triangulation mit Hilfe von Streifenprojektion [2] benötigt mindestens drei Rohbilder. Laserlichtschnittverfahren [3] benötigen typischerweise mehrere hundert Bilder. Die Weißlichtinterferometrie an optisch rauen Oberflächen erzielt sehr geringe Messunsicherheiten, dafür müssen jedoch ebenfalls mehrere hundert Bilder aufgenommen werden [4]-[7]. Diese Sammlung an Beispielen wirft Fragen auf.

1. Wie viel Rohdaten benötigt man minimal, um 3D-Daten einer vorge-

gebenen, geringen Messunsicherheit zu erhalten?

2. Wie nutzt man eine vorgegebene Menge an Rohdaten am besten, um eine möglichst geringe Messunsicherheit der 3D-Daten zu erzielen?

Die Informationstheorie stellt ein wertvolles Hilfsmittel bei der Beantwortung dieser Fragen dar. Kerngedanke dieser Arbeit ist es, den idealisierten optischen 3D-Sensor als Kommunikationssystem zu betrachten. Ein solches System besteht aus fünf Komponenten. Zuerst ist die Quelle der Nachricht zu nennen. Daran schließt sich der Codierer an, welcher die Nachricht geeignet aufbereitet, damit sie über den Kanal übertragen werden kann. Anschließend ist es Aufgabe des Decodierers, die ursprüngliche Nachricht wiederherzustellen und an die Senke weiterzuleiten. Dieses Kommunikationssystem wiederum wird mit Hilfe der Informationstheorie modelliert. Die Informationstheorie nach Shannon [8]-[16] definiert zunächst ein mathematisches Maß für die Information. Besonders wichtig ist auch der Begriff der Kanalkapazität. Sie gibt an, wie viel Information der Kanal maximal weitergeben kann. Diese allgemeinen Begriffe werden speziell für optische 3D-Sensoren konkretisiert. Die physikalischen Zusammenhänge im Sensor werden so aufbereitet, dass sie informationstheoretisch interpretiert werden können. Es wird der Begriff des Informationswirkungsgrades eingeführt, mit dessen Hilfe verschiedene Sensorprinzipien diskutiert und bewertet werden.

Die Erkenntnisse, die aus dem Modell gewonnen werden, zeigen den Weg zu einem Sensorprinzip mit sehr geringem Bedarf an Rohdaten bei gleichzeitig geringer Messunsicherheit. Ausgehend vom Messprinzip des photometrischen Stereo wird ein verbesserter Sensor entwickelt. Mit der Erweiterung des Verfahrens zur photometrischen Deflektometrie können neben diffus reflektierenden auch spiegelnde Objekte erfasst werden. Eine besonders große Beleuchtungsapertur minimiert das kohärente Rauschen und ermöglicht eine Messunsicherheit von weniger als $1 \mu\text{m}$. Mit Hilfe einer Farbkamera wird ein Single Shot Sensor realisiert, der alle Rohdaten gleichzeitig aufnimmt.

Die Arbeit ist wie folgt gegliedert: In Kapitel 3 werden zunächst die Grundlagen für optische 3D-Sensoren umrissen. Darauf folgen für den Fortgang der Arbeit relevante Teile der Informationstheorie in Kapitel 4. Auf diesem Hintergrund wird in Kapitel 5 das informationstheoretische Modell eingeführt. Schlussfolgerungen daraus führen zu einem optimierten Sensor in Kapitel 6, dessen Messergebnisse in Kapitel 7 dargestellt werden.

Kapitel 3

Optische Grundlagen

Das vorliegende Kapitel fasst die optischen Grundlagen der 3D-Messtechnik zusammen, die für die spätere informationstheoretische Modellierung und die Beschreibung des optimierten Sensors relevant sind. Es gibt eine Vielzahl von Sensorprinzipien, die sich in drei große Kategorien ordnen lassen, die Triangulation, die Interferometrie an optisch glatten Oberflächen und die Interferometrie an optisch rauen Oberflächen [7].

Auch wenn diese Kategorien in ihrer Funktionsweise sehr unterschiedlich sein mögen, so haben sie doch eines gemeinsam. In allen Fällen gibt es eine Beleuchtungseinheit, eine Beobachtungseinheit und das zu vermessende Objekt. Diese einheitliche Gliederung ist wichtig, da das in Kapitel 5 beschriebene informationstheoretische Modell auf diese Gliederung zurückgreift. Damit ist das Modell also grundsätzlich auf alle Arten von bekannten optischen 3D-Sensoren anwendbar. Die folgenden Abschnitte sind gegliedert in die optischen Grundlagen bezüglich der Beleuchtung, des Objektes und der Beobachtung. Daran schließen sich zwei Abschnitte über Rauschquellen bei optischen 3D-Sensoren an.

3.1 Beleuchtung

Ein Kernbestandteil eines optische Sensors ist die Beleuchtungseinheit. Die Art und Strukturierung der Beleuchtung spiegelt die physikalische Grundidee des Sensors wider, die häufig auch in seinem Namen zu finden ist. So verwendet z.B. ein Laserpunktsensor einen punktförmigen Lichtfleck. Lichtschnittsensoren [3] verwenden linienartige Beleuchtungsmuster und die Streifenprojektion oder phasenmessende Triangulation [2] arbeitet mit flächigen Streifenmustern. Neben der räumlichen Strukturierung wird die Beleuchtung von ihren Kohärenzeigenschaften bestimmt. Sie haben maßgebenden Einfluss

auf das Rauschen und die Messunsicherheit des Verfahrens.

3.1.1 Kohärenz der Beleuchtung

Die Kohärenz der Beleuchtung beschreibt, ob das einfallende Wellenfeld in der Lage ist, Interferenzen zu erzeugen [17], [18]. Dazu betrachtet man das komplexe analytische Signal $\vec{u}(P_1, t + \tau)$ des elektromagnetischen Feldes am Ort P_1 zur Zeit $t + \tau$ und setzt es in Beziehung mit dem Signal $\vec{u}(P_2, t)$ am Ort P_2 zur Zeit t . Die wechselseitige komplexe Kohärenzfunktion ist definiert als

$$\begin{aligned}\Gamma_{12}(\tau) &= \langle \vec{u}(P_1, t + \tau) \vec{u}^*(P_2, t) \rangle \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{+T/2} \vec{u}(P_1, t + \tau) \vec{u}^*(P_2, t) dt\end{aligned}\quad (3.1)$$

Betrachtet man die Interferenzfähigkeit zu verschiedenen Zeiten mit der Zeitdifferenz τ am selben Ort P , handelt es sich um zeitliche Kohärenz.

$$\Gamma(\tau) = \langle \vec{u}(P, t + \tau) \vec{u}^*(P, t) \rangle \quad (3.2)$$

Die räumliche Kohärenz wiederum beschreibt die Interferenzfähigkeit an verschiedenen Orten P_1 und P_2 zur gleichen Zeit ($\tau = 0$).

$$\Gamma_{12}(\tau = 0) = \langle \vec{u}(P_1, t) \vec{u}^*(P_2, t) \rangle \quad (3.3)$$

Für Aufgaben der Praxis ist es sinnvoll, die wechselseitige Kohärenzfunktion $\Gamma_{12}(\tau)$ zu normieren. Diese Größe wird komplexer Kohärenzgrad $\gamma_{12}(\tau)$ genannt.

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{\Gamma_{11}(\tau = 0) \Gamma_{22}(\tau = 0)}} \quad (3.4)$$

Im Experiment beschreibt der komplexe Kohärenzgrad den Kontrast ν der Interferenzerscheinung. Für gleiche Intensität von $\vec{u}(P_1, t)$ und $\vec{u}(P_2, t)$ gilt

$$\nu = |\gamma_{12}(\tau)| \quad (3.5)$$

Der Kontrast von Interferenzerscheinungen wird eine wichtige Rolle spielen beim informationstheoretischen Modell. Der Kontrast ist ein Maß für das Rauschen und begrenzt die Menge an Information, die über ein Objekt gewonnen werden kann.

Alternativ kann man die zeitliche Kohärenz auch mit dem Begriff der Kohärenzzeit τ_c oder der Kohärenzlänge l_c beschreiben.

$$\begin{aligned}\tau_c &= \int_{-\infty}^{+\infty} |\gamma(\tau)|^2 d\tau \\ l_c &= c \cdot \tau_c\end{aligned}\quad (3.6)$$

Hier bezeichnet c die Lichtgeschwindigkeit. Anschaulich steht die Kohärenzzeit τ_c für die maximale zeitliche Verzögerung, bei der ein Wellenfeld noch mit sich selbst interferieren kann. Entsprechend kann man die Kohärenzlänge l_c als diejenige Distanz in Ausbreitungsrichtung des Wellenfeldes verstehen, bei der gerade noch Interferenzen beobachtbar sind.

Darüber hinaus gilt folgender Zusammenhang zwischen der Kohärenzzeit τ_c , der Kohärenzlänge l_c und der zeitlichen Bandbreite B_t des Wellenfeldes.

$$B_t \approx \frac{1}{\tau_c} = \frac{c}{l_c} \quad (3.7)$$

Man kann also die zeitlichen Kohärenzeigenschaften eines Wellenfeldes anhand einer ganzen Reihe von Größen beschreiben, von der wechselseitigen Kohärenzfunktion über den komplexen Kohärenzgrad, den Interferenzkontrast bis hin zur Kohärenzzeit, Kohärenzlänge und zeitlichen Bandbreite. Je nachdem, welche dieser Größen für eine Beschreibung am aussagekräftigsten ist, wird man die eine oder andere von ihnen wählen.

3.2 Objekt

Im folgenden soll davon ausgegangen werden, dass alle betrachteten Objekte Körper mit einer undurchsichtigen Oberfläche sind. Für die Optik ist damit nur die Oberfläche relevant, das Innere des Objekts bleibt verborgen. Man unterscheidet zwischen im optischen Sinn glatten und rauen Oberflächen. Optisch glatte Oberflächen weisen Unebenheiten auf, die kleiner sind als eine halbe Wellenlänge des verwendeten Lichts, optisch raue dagegen Unebenheiten größer als die halbe Wellenlänge. Bei optisch glatten Oberflächen sind Interferenzen und kohärente Beleuchtung oft sehr nützlich. Sie geben Aufschluss über die dreidimensionale Form des Objektes in Bruchteilen der Wellenlänge.

Bei den meisten Objekten allerdings handelt es sich um optisch raue Gegenstände. Bei ihnen sind Interferenzen häufig störend. Rauigkeiten von mehr als einer halben Wellenlänge führen zu Speckles, das sind Interferenzerscheinungen, die keinen Aufschluss mehr über die Form des Objektes in Bruchteilen einer Wellenlänge zulassen. Die Überlagerung von gestreuten Elementarwellen erfolgt ungeordnet und kann als statistisches Phänomen betrachtet werden. Bei rauen Objekten erzeugt also das Zusammenspiel aus Kohärenz der Beleuchtung und Rauigkeit des Objektes ein statistisches Rauschsignal, das Specklerauschen. Seine Größe wird durch den Specklekontrast charakterisiert, der wiederum direkt in das informationstheoretische Modell eingehen wird.

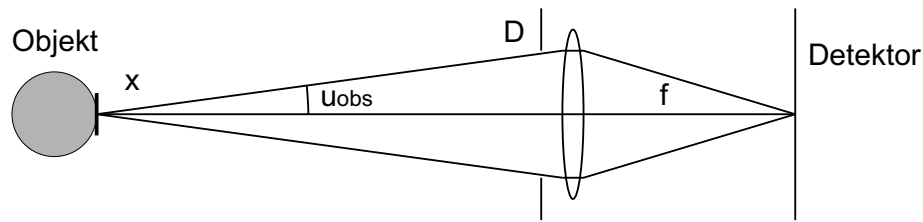


Abbildung 3.1: Beugungsbegrenztes optisches Beobachtungssystem bei der Wellenlänge λ mit weit entferntem Objekt

3.2.1 Polarisations-eigenschaften

Beim Objekt spielen auch Polarisations-eigenschaften eine Rolle. Metalle erhalten in der Regel die Polarisationsrichtung des einfallenden Lichtes, während die meisten dielektrische Körper die Polarisation nicht erhalten. Bei nicht polarisationserhaltenden Körpern findet eine Aufteilung auf nicht interferenzfähige Polarisationsrichtungen statt, was wiederum den Kontrast der Interferenzen senkt. Bei polarisationserhaltenden Objekten ist keine Kontrastreduktion zu erwarten.

3.3 Beobachtung

Als dritter Baustein eines optischen Sensors wird nun die Beobachtungseinheit betrachtet (siehe Abbildung 3.1). Ein wichtiges Merkmal für die Beobachtung ist die Beobachtungsapertur $\sin u_{obs}$. Sie legt die beugungsbegrenzte laterale Auflösung auf dem Objekt fest. Nach Abbé ist die minimal auflösbare Distanz

$$\delta x = \frac{\lambda}{\sin u_{obs}} \quad (3.8)$$

Hier ist λ die Wellenlänge. Man kann die Eigenschaften der Beobachtung gemäß der Fourier-Optik auch im Ortsfrequenzraum anstelle des Ortsraums darstellen [19]. Als Bandbreite des Beobachtungssystems $B_x \cdot B_y$ erhält man

$$B_x \cdot B_y = \frac{\pi D^2}{16 \lambda^2 f^2} \quad (3.9)$$

Hier ist D der Durchmesser der Blende und f die Brennweite des optischen Systems.

Bei elektronisch arbeitenden Kameras sind darüber hinaus die Größe $\Delta x'$ bzw. $\Delta y'$ der Bildelemente (Pixel) und die Gesamtausdehnung $\Delta X'$ bzw. $\Delta Y'$ des Detektors von Bedeutung (siehe Abb. 3.2). Ihnen entspricht auf

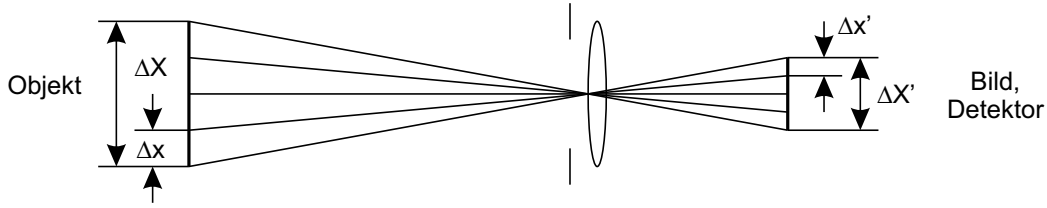


Abbildung 3.2: Abbildungssystem mit elektronischer Detektor, dessen Pixel die Ausdehnung $\Delta x' \cdot \Delta y'$ besitzen. Zur Vereinfachung ist nur die x -Koordinatenrichtung dargestellt

dem Objekt die objektseitige Pixelgröße Δx bzw. Δy und die Größe des Messfeldes ΔX bzw. ΔY . (Bemerkung: Alle objektseitigen Größen werden ohne Strich dargestellt, alle bildseitigen mit Strich.)

3.4 Kohärentes Rauschen durch Speckle

Im Zusammenhang mit den rauen Objekten wurde bereits erwähnt, dass bei der Wechselwirkung von kohärenter Beleuchtung und rauer Oberfläche Speckle entstehen [20], [21], [22]. Jeder Punkt einer kohärent beleuchteten rauen Oberfläche kann als Ausgangspunkt einer Elementarwelle mit zufälliger Phasenlage gesehen werden. Diese Elementarwellen überlagern sich und erzeugen in einem Abstand von der Oberfläche ein granulationsartiges Interferenzmuster, das als Speckle bezeichnet wird. Objektive Speckle entstehen, wenn sich das gestreute Licht im Raum frei ausbreitet, subjektive Speckle entstehen, wenn ein abbildendes System beteiligt ist. Bei der optischen 3D-Messtechnik handelt es sich meist um subjektive Speckle, da in der Regel ein abbildendes System, die Beobachtung, beteiligt ist. Der Specklekontrast ist [23]

$$c_{\text{speckle}} = \frac{\sigma_I}{\bar{I}} \quad (3.10)$$

Hier ist \bar{I} die mittlere Intensität und σ_I deren Standardabweichung. Der Specklekontrast hat die Bedeutung eines inversen Signal-Rausch-Verhältnis. In der Sprache der Nachrichtentechnik charakterisiert \bar{I}^2 die Signalleistung und σ_I^2 die Rauschleistung. Der Wunsch nach einem hohen Signal-Rausch-Verhältnis in der Nachrichtentechnik ist gleichbedeutend mit dem Wunsch nach einem geringen Specklekontrast in der optischen 3D-Messtechnik.

Der Specklekontrast wird von drei Faktoren bestimmt, dem Beitrag der räumlichen Kohärenz c_{spat} , dem Beitrag der zeitlichen Kohärenz c_{temp} und dem polarisationsabhängigen Beitrag c_{pol} [24]. Bei elektronischen Kameras

kommt ein vierter Beitrag hinzu. Fallen mehrere Speckle auf einen Pixel der Kamera, so findet eine Mittelung der Intensitäten statt. Der von der Kamera wahrgenommene Specklekontrast reduziert sich entsprechend dem Beitrag von c_{pixel} . Der Specklekontrast berechnet sich zu

$$c_{speckle} = c_{spat} \cdot c_{temp} \cdot c_{pol} \cdot c_{pixel} \quad (3.11)$$

Der Beitrag der räumlichen Kohärenz ist Eins (maximale Kohärenz), solange die Beleuchtungsapertur $\sin u_{ill}$ kleiner als die Beobachtungsapertur $\sin u_{obs}$ ist. Eine Verminderung auf Werte kleiner Eins wird erreicht, wenn die Beleuchtungsapertur größer ist als die Beobachtungsapertur [6].

$$c_{spat} = \min \left(\frac{\sin u_{obs}}{\sin u_{ill}}, 1 \right) \quad (3.12)$$

Über die Größe von Beleuchtungs- und Beobachtungsapertur kann man daher den Specklekontrast wirksam reduzieren und das kohärente Rauschen minimieren.

Einen geringeren Einfluss hat die zeitliche Kohärenz. Ist die Kohärenzlänge l_c (siehe Gl. 3.6) der Beleuchtungsquelle kleiner als die Rauigkeit des Objektes (charakterisiert durch die Standardabweichung σ_o), so kann über den Beitrag c_{temp} mit Werten kleiner Eins eine weitere Kontrastreduktion erreicht werden [25].

$$c_{temp} = \left(1 + \left(\frac{4\sigma_o}{l_c} \right)^2 \right)^{-1/4} \quad (3.13)$$

Außer der Rauigkeit des Objektes kann auch die Beleuchtung des Objektes fernab der spekularen Richtung [47] für Gangunterschiede sorgen (siehe Abb. 3.3). Wird das Objekt (senkrecht zur optischen Achse angenommen) unter dem Winkel θ beleuchtet, so ergibt sich innerhalb der auflösbaren Distanz $\delta x = \lambda / \sin u_{obs}$ ein Gangunterschied von

$$\Delta s = \delta x \cdot \sin \theta \quad (3.14)$$

Dieser Gangunterschied wird bei Lampalzer [47] als Keilhöhe bezeichnet. Dominieren diese Gangunterschiede, so erhält man für den Beitrag der zeitlichen Kohärenz

$$c_{temp} = \left(1 + \left(\frac{4\Delta s}{l_c} \right)^2 \right)^{-1/4} \quad (3.15)$$

Darüber hinaus gehen auch die Polarisationsseigenschaften des Objektes mit

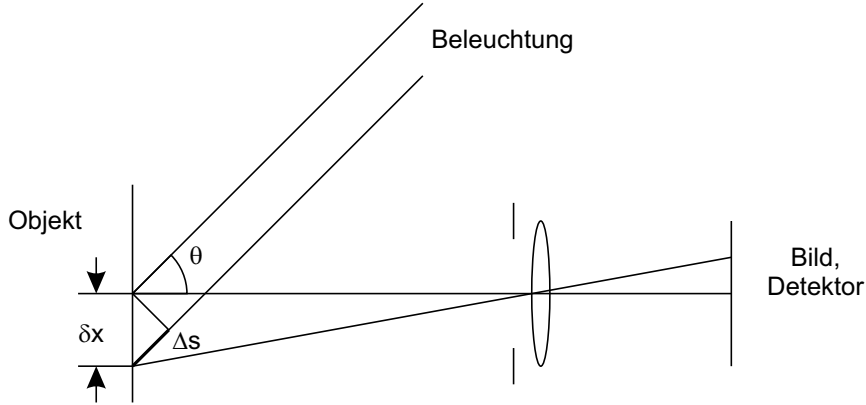


Abbildung 3.3: Gangunterschiede innerhalb der auflösbaren Distanz δx bei Beleuchtung unter dem Winkel θ fernab der spekularen Richtung

ein. Ein depolarisierendes Objekt reduziert den Specklekontrast um den Faktor $\frac{1}{\sqrt{2}}$ [24].

Der Beitrag der Pixel c_{pixel} ist Eins, solange die objektseitige Specklegröße die objektseitige Pixelgröße Δx übersteigt. Sind die Speckle kleiner als die Pixel, erreicht man eine Kontrastreduktion.

$$c_{pixel} = \min \left(\frac{\lambda}{\Delta x \cdot \sin u_{obs}}, 1 \right) \quad (3.16)$$

Für den Fall, dass $c_{pixel} = 1$ ist, gilt für den Specklekontrast

$$c_{speckle} = \min \left(\frac{\sin u_{obs}}{\sin u_{ill}}, 1 \right) \cdot c_{temp} \cdot c_{pol} \quad (3.17)$$

Für den Fall der Kontrastreduktion kann man c_{spat} und c_{pixel} zusammenfassen und erhält

$$c_{speckle} = \frac{\lambda}{\Delta x \cdot \max(\sin u_{ill}, \sin u_{obs})} \cdot c_{temp} \cdot c_{pol} \quad (3.18)$$

Hier ist nicht mehr das Verhältnis von Beobachtungs- und Beleuchtungsapertur maßgeblich, sondern die größere von beiden. Daneben geht das Verhältnis aus objektseitiger Pixelgröße und der Wellenlänge ein.

3.5 Photonenrauschen

Neben dem kohärenten Rauschen tritt bei optischen Detektoren auch das Photonenrauschen in Erscheinung [31]. Das einfallende Wellenfeld wird nicht

als kontinuierlicher Energiestrom registriert, sondern als Folge von einzelnen Photonen. Die Anzahl der Photonen, die pro Zeiteinheit und Flächeneinheit am Detektor registriert wird, schwankt statistisch. Treffen im Mittel \bar{n} Photonen ein, so ist die Standardabweichung $\sigma_n = \sqrt{\bar{n}}$.

Zusammenfassend kann man sagen, dass sowohl das kohärente Rauschen als auch das Photonenrauschen die Fähigkeit eines 3D-Sensors begrenzen, Forminformation des Objektes zu übertragen. Daher werden Rauschquellen bei der informationstheoretischen Modellierung eine große Rolle spielen.

Kapitel 4

Informationstheoretische Grundlagen

In der Einleitung wurde die Frage gestellt, wie viel Rohdaten einer Kamera benötigt werden, um die dreidimensionale Form eines Objektes zu ermitteln. Dies ist eine Abwägung zwischen Nutzen (3D-Form) und Aufwand (Rohdaten). Wie aber kann man sowohl den Nutzen als auch den Aufwand quantifizieren? Die Informationstheorie nach Shannon stellt einen geeigneten Rahmen zur Verfügung und definiert Größen wie Information und Kanalkapazität [8]-[16].

Die Informationstheorie wird hauptsächlich zur effizienten Übertragung von Nachrichten angewandt. Damit können beispielsweise Mobilfunksysteme oder auch Systeme der optischen Nachrichtentechnik beschrieben werden. Genauso gut lässt sie sich aber auch auf die Bildentstehung in der Optik anwenden [26], [27], [28]. Weniger gebräuchlich ist die Anwendung der Informationstheorie im Bereich der Messtechnik, insbesondere der optischen 3D-Messtechnik. Es gibt hier keine Kommunikation zwischen zwei Personen. Der Ursprung einer Nachricht ist vielmehr ein Objekt. Die Information ist gegeben durch seine Form, für die sich eine Person interessiert. Dennoch ist auch hier die Informationstheorie zur Beschreibung des Modells gut geeignet.

Im folgenden sollen die Grundgedanken der Informationstheorie herausgestellt werden, so fern sie für die spätere Modellbildung relevant sind. Im Übertragungsmodell werden zunächst die Bestandteile eines Nachrichtensystems wie Quelle, Codierung, Kanal, Decodierung und Senke vorgestellt. Daran schließt ein Abschnitt über das Maß für die Information an. Am einfachsten lässt sich die Idee dafür bei der diskreten Nachrichtenübertragung illustrieren. In der optischen 3D-Messtechnik handelt es sich allerdings um analoge Messgrößen. Der Informationsbegriff wird daher für solche Größen ebenfalls erläutert. Die Überlegungen münden in den Begriff der Kanalkapa-

zität, die für das Sensormodell eine zentrale Rolle spielt.

4.1 Das Übertragungsmodell

In der Informationstheorie stützt man sich auf das Übertragungsmodell für Kommunikationssysteme (siehe Abb. 4.1) [29]. Ein Kommunikationssystem kann aufgeteilt werden in die Quelle der Nachricht, den Codierer, den Kanal, den Decodierer und die Senke. Die Funktion der Einheiten wird am Beispiel einer Rundfunkübertragung deutlich. Der Radiosprecher stellt die Quelle der Nachricht dar. Er drückt die Nachricht in Form eines akustischen Signals aus, welches zunächst im Mikrophon in ein elektrisches gewandelt wird. Dieses Signal ist aber nicht geeignet, um direkt übertragen zu werden. Erst wenn es durch Modulation in das Antennensignal verwandelt wird (Codierung), kann es über die Funkstrecke (Kanal) übertragen werden. Im Radio des Zuhörers werden die Antennensignale demoduliert (Decodierung) und wieder in akustische Signale verwandelt, die dann den Zuhörer erreichen (Senke). Dies ist ein Beispiel für eine Übertragung mit analogen Signalen. Im Gegensatz dazu verwendet man für die Übertragung im Mobilfunk digitale Signale, also Signale, bei denen nur bestimmte Wellenformen oder Symbole zugelassen sind. Die Menge der zulässigen Symbole bezeichnet man als Alphabet. Mit diesen Begriffen wird nun ein Maß für die Information festgelegt.



Abbildung 4.1: Übertragungsmodell für Kommunikationssysteme

4.2 Ein Maß für die Information

Um Information zu messen, wird die Nachrichtenübertragung weiter abstrahiert. Das gesendete Signal wird durch eine diskrete Zufallsvariable A beschrieben, die die Werte a_i , $i = 1, 2, \dots, n$ annehmen kann. Die Symbole a_i könnten z.B. die Werte 0 und 1 sein (binäre Symbole). Das empfangene Signal wird entsprechend durch die Zufallsvariable B mit den Werten

b_j , $j = 1, 2, \dots, m$ beschrieben. Der Empfänger wird nun versuchen, vom empfangenen Signal B auf das gesendete Signal A zurückzuschließen. Man könnte beispielsweise versuchen, durch das Werfen einer Münze auf das Wetter des kommenden Tages zu schließen. Bei Kopf nimmt man gutes Wetter an, bei Zahl schlechtes Wetter. In diesem Beispiel ist offensichtlich, dass die Münze keinen Rückschluss auf das Wetter zulässt, da beide Ereignisse statistisch unabhängig sind. Für statistisch unabhängige Signale A und B fordert man deshalb, dass die Information Null sein soll. Wenn das Ereignis a_i aber zwangsläufig das Ereignis b_j nach sich zieht, soll die Information maximal sein. Darüber hinaus fordert man, dass sich die Information additiv verhalten soll, wenn statt einem Symbol mehrere Symbole übermittelt werden. Das Informationsmaß nach Shannon erfüllt diese Kriterien [30].

Definition: Die Transinformation, die das Ereignis $B = b_j$ über das Ereignis $A = a_i$ liefert, ist

$$I(a_i; b_j) = \log_2 \frac{P(a_i|b_j)}{P(a_i)} = \log_2 \frac{P(a_i, b_j)}{P(a_i)P(b_j)} \quad (4.1)$$

Die Einheit der Transinformation ist 1 *bit*. Die Definition enthält das Verhältnis der bedingten Wahrscheinlichkeit $P(a_i|b_j)$ zur Wahrscheinlichkeit $P(a_i)$. Ist das Ereignis a_i statistisch unabhängig vom Ereignis b_j , so sind beide Wahrscheinlichkeiten gleich, das Verhältnis ist Eins und damit die Transinformation Null. Folgt aber auf das Symbol a_i mit Sicherheit das Ereignis b_j ($P(a_i|b_j) = 1$), so ist die Transinformation maximal. Alternativ kann die Transinformation auch anhand der Verbundwahrscheinlichkeit $P(a_i, b_j)$ und der Wahrscheinlichkeiten $P(a_i)$, $P(b_j)$ angegeben werden.

Neben der Transinformation zwischen zwei Ereignissen kann man auch die Information eines Ereignisses als solches definieren.

Definition: Die Information des Ereignisses $A = a_i$ ist

$$I(a_i) = -\log_2 P(a_i) \quad (4.2)$$

Dies ist ein Spezialfall der Transinformation, wenn nämlich das Ereignis a_i mit Sicherheit das Ereignis b_j nach sich zieht.

4.3 Mittlere Transinformation und Entropie

Bisher wurde die Transinformation für zwei bestimmte Symbole $A = a_i$ und $B = b_j$ betrachtet. Will man eine Aussage über alle Symbole des Eingangsalphabets und des Ausgangsalphabets machen, so berechnet man die mittlere

Transinformation für alle Kombinationen von Eingangs- und Ausgangssymbolen. Dabei wird die jeweilige Transinformation mit der Wahrscheinlichkeit dieser Symbolkombination gewichtet.

Definition: Die mittlere Transinformation zwischen den Zufallsvariablen A und B ist

$$I(A; B) = \sum_{i=1}^n \sum_{j=1}^m P(a_i, b_j) \log_2 \frac{P(a_i, b_j)}{P(a_i)P(b_j)} \quad (4.3)$$

Analog definiert man die mittlere Information $H(A)$, die auch Entropie heißt.

Definition: Die Entropie der Zufallsvariablen A ist

$$H(A) = - \sum_{i=1}^n P(a_i) \log_2 P(a_i) \quad (4.4)$$

Alle bisher definierten Größen gelten für diskrete Zufallsvariablen. In der optischen 3D-Messtechnik spielen aber kontinuierliche Variable die weitaus wichtigere Rolle. Deshalb werden im folgenden stetige Zufallsvariable betrachtet.

4.4 Transinformation bei stetigen Zufallsvariablen

Die Definition der mittleren Transinformation kann auf stetige Zufallsvariable $(p(a), p(b), p(a, b))$ übertragen werden. Anstelle der Wahrscheinlichkeiten werden Wahrscheinlichkeitsdichten, anstelle der Summe eine Integration geschrieben.

Definition: Die mittlere Transinformation zwischen den stetigen Zufallsvariablen A und B ist

$$I(A; B) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(a, b) \log_2 \frac{p(a, b)}{p(a)p(b)} da db \quad (4.5)$$

4.5 Differentielle Entropie

Der Begriff der Entropie lässt sich nicht ohne weiteres auf stetige Zufallsvariable übertragen. Die Schwierigkeit liegt darin, dass man zur exakten Darstellung von stetigen Zufallsvariablen unendlich viele diskrete Werte benötigt.

Deshalb ist der Informationsgehalt einer stetigen Zufallsvariablen unendlich und auch ihre Entropie unendlich. Stattdessen definiert man den Begriff der differentiellen Entropie.

Definition: Die differentielle Entropie einer stetigen Zufallsvariablen A ist

$$h(A) = - \int_{-\infty}^{\infty} p(a) \log_2 p(a) da \quad (4.6)$$

Zurück zur Transinformation stetiger Zufallsvariablen: In der Nachrichtenübertragung und auch der Messtechnik möchte man die mittlere Transinformation möglichst groß machen. Aus einem empfangenen Signal möchte man möglichst genaue Rückschlüsse ziehen, welches Signal gesendet wurde. Shannon zeigt, dass es bei der Übertragung von Information eine obere Grenze gibt, die Kanalkapazität.

4.6 Kanalkapazität

Die Kanalkapazität gibt an, welche Transinformation unter gegebenen Randbedingungen maximal übertragen werden kann. Anders ausgedrückt, wenn die Transinformation niedriger als die Kanalkapazität ist, so ist eine fehlerfreie Übertragung möglich. Versucht man die Transinformation größer zu machen, so sind Übertragungsfehler unvermeidlich.

Die Eigenschaften des Kanals sind für diese Beschränkung verantwortlich. An dieser Stelle soll ein Kanal angenommen werden, der kontinuierlich bezüglich der Zeit und bezüglich der Signale ist. Störungen des Kanals sind als additives weißes Gauss'sches Rauschen zu beschreiben. Diesen Kanal nennt man daher zeitkontinuierlichen additiven weißen Rauschkanal bzw. „Additive White Gaussian Noise Channel“ (AWGN-Kanal). Man nimmt an, dass der Kanal nur Frequenzen innerhalb einer bestimmten zeitlichen Bandbreite B_t übertragen kann. Die Kanalkapazität ist dann [32]

$$C = \Delta T \cdot B_t \cdot \log_2 \left(1 + \frac{S_{in}}{N} \right) \quad (4.7)$$

Hier ist S_{in}/N das Signal-Rauschleistungsverhältnis des Signals am Eingang des Kanals und ΔT die Dauer der Übertragung. Für die vorliegende Arbeit ist die Kanalkapazität von großer Bedeutung. Sie ist ein Maß für den Aufwand, den man in der optischen 3D-Messtechnik treiben muss. Auf der anderen Seite beschreibt die Transinformation den Nutzen, den man daraus ziehen kann. Mit diesen Größen ist es möglich, ein Verhältnis aus Nutzen und Aufwand

anzugeben. Wie erreicht man ein Verhältnis von Nutzen zu Aufwand, das möglichst nahe bei Eins liegt? Es liegt an der Codierung, die Kapazität des Kanals möglichst gut auszunutzen.

4.7 Codierung

Die Codierung hat bei der Nachrichtenübertragung zweierlei Funktion. Zum einen soll sie Redundanz der Quelle eliminieren (Quellcodierung), zum anderen aber wieder gezielt Redundanz hinzufügen, damit Übertragungsfehler ausgeschlossen werden (Kanalcodierung). Im Bereich der digitalen Nachrichtenübertragung gibt es eine Vielzahl von Strategien zur Quell- und Kanalcodierung. In der analog arbeitenden optischen 3D-Messtechnik stehen dagegen nur ausgewählte Codierungsverfahren zur Verfügung. In digitaler Technik lassen sich praktisch beliebige Rechenvorschriften implementieren, was im analogen Bereich sehr schwierig ist. Daher muss es das Ziel sein, mit den begrenzten Möglichkeiten dennoch ein effektives Codierungsverfahren zu finden. Dieses ist an die Eigenschaften der Quelle (typische 3D-Formen) anzupassen.

Nach diesen allgemeinen Grundlagen der Informationstheorie soll im folgenden Kapitel ein informationstheoretisches Modell entwickelt werden, das genau auf die Gegebenheiten in der optischen 3D-Messtechnik zugeschnitten ist.

Kapitel 5

Ein Informationstheoretisches Sensormodell

Ein optischer 3D-Sensor kann als Kommunikationssystem betrachtet werden (siehe Abb. 5.1), wie es im vorigen Kapitel beschrieben wurde. Die Quelle des Systems ist das Formsignal $z(x, y)$ des Objektes. Genauer ausgedrückt ist $z(x, y)$ das Formsignal der Objektoberfläche. Die Form wird durch eine geeignete Beleuchtungseinheit codiert. Es gibt viele Arten der Beleuchtung, die sich in ihrer räumlichen und zeitlichen Struktur, aber auch ihren Kohärenzeigenschaften unterscheiden. So verwendet z.B. ein Lichtschnittsensor ein linienartiges Lichtmuster zur Codierung. Bei der Streifenprojektion wiederum werden flächige Lichtmuster gebraucht, insbesondere Sinusmuster. Meist werden mehrere verschiedene Sinusmuster nacheinander projiziert, was sich in der zeitlichen Struktur der Codierung niederschlägt. Wie auch immer die Beleuchtung gewählt wird, sollen alle ihre Eigenschaften in der Codierungseinheit zusammengefasst werden. Die Lichtwelle, die nun vom Objekt gestreut oder reflektiert wird, ist als codiert zu betrachten.

Nach der Codierung durchläuft das Signal den Kanal. Der Kanal ist im

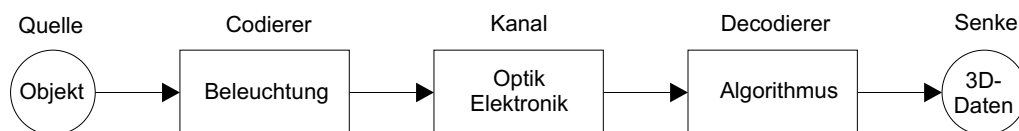


Abbildung 5.1: Übertragungsmodell für optische 3D-Sensoren

Sinn der Informationstheorie das Element, das die Übertragung zum einen ermöglicht, zum anderen aber durch Rauschen beschränkt. Im Fall des optischen 3D-Sensors ist dies die Beobachtungseinheit, bestehend aus dem leeren Raum, den optischen und elektronischen Komponenten. Die maßgeblichen Rauschquellen sind durch das kohärente Specklerauschen, das Photonenrauschen und das Quantisierungsrauschen gegeben. (Bei digitalen Kameras findet die Quantisierung in der Kamera statt, bei analogen Kameras in einer Bildspeicherkarte). Am Ende des Kanals liegen die Rohdaten in Form von zweidimensionalen Kamerabildern vor.

Anschließend gelangen die Rohdaten zur Decodierung. Diese ist das Gegenstück der Codierung, in der Regel ein PC. Entsprechend der Beleuchtung schließt der Decoder auf die gesuchte Form des Objektes zurück. Dies gelingt nicht perfekt, anstelle des wahren Formsignals $z(x, y)$ erhält man das gemessene Formsignal $z_m(x, y)$. Dieses Signal wird an die Senke, einen menschlichen Beobachter, weitergegeben.

Bei der Modellbildung wird zunächst auf den Kanal und seine Kapazität eingegangen. Die Kapazität gibt an, wie viel Information maximal übertragen werden kann. Daran schließt sich ein Abschnitt an über die Eigenschaften der Quelle und ein Ansatz zur Codierung. Die Codierung bestimmt, wie viel Information in Wirklichkeit übertragen wird. Mit der Kanalkapazität einerseits (wie viel Information kann im besten Fall übertragen werden) und der realisierten Transinformation andererseits kann man den Aufwand eines 3D-Sensors dem Nutzen gegenüberstellen. Die Kanalkapazität ist das Maß für den Aufwand, die realisierte Transinformation das Maß für den Nutzen.

5.1 Der Kanal

Bei der optischen 3D-Formerfassung stellt die Beobachtungseinheit den Kanal des Kommunikationssystems dar. Die Beobachtungseinheit ist durch eine Videokamera, eine Digitalkamera oder einen ähnlichen Detektor gegeben. Sie besteht aus dem leeren Raum, optischen Komponenten (Objektiv der Kamera), analogen elektronischen Komponenten (der Detektorchip, Verstärker etc.) und digitalen elektronischen Komponenten (Analog/Digital-Wandler), die nacheinander durchlaufen werden. Dementsprechend wird auch der Gesamtkanal in zugehörige Einheiten aufgeteilt, den optischen (Unter)-Kanal, den elektronischen (Unter)-Kanal und den diskreten (Unter)-Kanal. Die Rauschquellen der jeweiligen Komponente werden dem entsprechenden (Unter)-Kanal zugeordnet. So ist der optische Kanal für kohärentes Rauschen verantwortlich, der elektronische Kanal für das Photonenrauschen an der Schnittstelle der Optik zur Elektronik und der diskrete Kanal für das

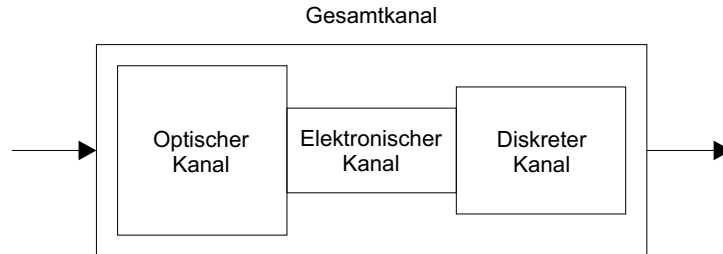


Abbildung 5.2: Struktur des Gesamtkanals

Diskretisierungsrauschen.

Ausgehend von der allgemein üblichen Darstellung der Kanalkapazität in der Nachrichtentechnik wird eine leicht veränderte Darstellung für die optische 3D-Messtechnik abgeleitet. Sie trägt dem Umstand Rechnung, dass in der optischen Messtechnik nur das Signal am Ausgang des Gesamtkanals bekannt ist und nicht an dessen Eingang. Darüber hinaus ist das Signal über der Zeit durch ein Signal über einer zweidimensionalen Ortsvariablen (die Ausdehnung des Detektors) und der Zeit zu ersetzen.

Die Kapazität des zeitkontinuierlichen Kanals mit additivem weißem Gaußschen Rauschen (AWGN-Kanal) ist [32]

$$C = \Delta T \cdot B_t \cdot \log_2 \left(1 + \frac{S_{in}}{N} \right) \quad (5.1)$$

Hier ist B_t die zeitliche Bandbreite des Signals, S_{in} die mittlere Leistung des Signals am Eingang des Kanals und N die mittlere Leistung des Kanalrauschens. Die Kapazität wird in der Einheit *bit* gemessen. In der Messtechnik ist der Eingang des Kanals nicht zugänglich. Information über die interessierende Größe bekommt man nur durch Messung, wobei immer der Kanal durchlaufen werden muss. Man kennt nur die Größe des Ausgangssignals, das sich aus der Addition des Eingangssignals und des Rauschens ergibt. Nimmt man Eingangssignal und Rauschen als statistisch unabhängig voneinander an, so ist die Leistung am Ausgang die Summe der Eingangsleistung und der Rauschleistung.

$$S_{out} = N + S_{in} \quad (5.2)$$

Damit erhält man als erste Modifikation

$$C = \Delta T \cdot B_t \cdot \log_2 \left(\frac{S_{out}}{N} \right) \quad (5.3)$$

Die zweite Modifikation betrifft die Tatsache, dass in der Nachrichtentechnik häufig zeitliche Abhängigkeiten eine Rolle spielen, in der Optik aber räumliche und zeitliche Abhängigkeiten. Erweitert man die Darstellung der Kanalkapazität um zwei räumliche Dimensionen, wobei $\Delta X'$ und $\Delta Y'$ Ortsintervalle darstellen und B_x bzw. B_y die entsprechende Ortsbandbreiten, so erhält man

$$C = 4 \cdot \Delta X' \cdot B_x \cdot \Delta Y' \cdot B_y \cdot \Delta T \cdot B_t \cdot \log_2 \left(\frac{S_{out}}{N} \right) \quad (5.4)$$

Es folgt ein kurzer Einschub, der erläutert, wie Gl. 5.4 und insbesondere der Faktor 4 zustande kommen.

Ein zeit- und ortskontinuierliches Signal $s_{out}(x, y, t)$ der Bandbreite B_x , B_y und B_t kann entsprechend dem Abtasttheorem vollständig durch eine Folge von diskreten Werten repräsentiert werden. Im Ortsintervall $\Delta X'$, $\Delta Y'$ und dem Zeitintervall ΔT sind dies

$$M_x \cdot M_y \cdot M_t = (2\Delta X' \cdot B_x) \cdot (2\Delta Y' \cdot B_y) \cdot (2\Delta T \cdot B_t) \quad (5.5)$$

Werte [32]. Der bandbegrenzte zeit- und ortskontinuierliche AWGN-Kanal kann also durch einen äquivalenten zeit- und ortsdiskreten Kanal ersetzt werden. Die Kapazität eines solchen Kanals ist aber bekannt und beträgt [33]

$$C = \frac{1}{2} \cdot M_x \cdot M_y \cdot M_t \cdot \log_2 \left(\frac{S_{out}}{N} \right) \quad (5.6)$$

Ersetzen von M_x , M_y und M_t liefert

$$C = \frac{1}{2} \cdot (2\Delta X' \cdot B_x) \cdot (2\Delta Y' \cdot B_y) \cdot (2\Delta T \cdot B_t) \cdot \log_2 \left(\frac{S_{out}}{N} \right) \quad (5.7)$$

und wird vereinfacht zu

$$C = 4 \cdot \Delta X' \cdot B_x \cdot \Delta Y' \cdot B_y \cdot \Delta T \cdot B_t \cdot \log_2 \left(\frac{S_{out}}{N} \right) \quad (5.8)$$

Aus dieser allgemeinen Darstellung der Kanalkapazität wird im folgenden die Kapazität des optischen, des elektronischen und des diskreten Kanals hergeleitet.

5.1.1 Kapazität des optischen Kanals

Beim optischen Kanal geht man davon aus, dass es sich um ein beugungsbegrenztes optisches Abbildungssystem der Brennweite f , eine kreisrunde Apertur mit dem Durchmesser D und ein Bildfeld der Größe $\Delta X' \cdot \Delta Y'$ handelt (siehe Abb. 5.3). Man kann zeigen, dass ein solches System nur Orts-

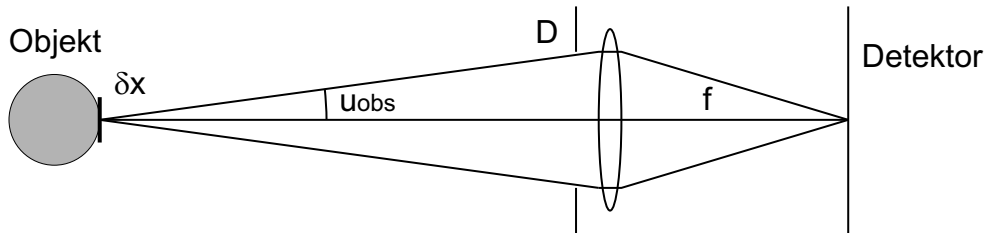


Abbildung 5.3: Beugungsbegrenztes optisches Beobachtungssystem mit weit entferntem Objekt

frequenzen einer bestimmten Bandbreite übertragen kann (siehe Kapitel 3). Der Bereich der Frequenzen, die übertragen wird, schließt im Frequenzraum eine kreisförmige Fläche der Größe

$$B_x \cdot B_y = \frac{\pi D^2}{16\lambda^2 f^2} \quad (5.9)$$

ein. Die zeitliche Bandbreite erhält man laut Gl. 3.7 zu

$$B_t \approx \frac{c}{l_c} \quad (5.10)$$

Hier ist c die Lichtgeschwindigkeit und l_c die Kohärenzlänge. Im Vergleich zur Bandbreite des elektronischen Kanals, der anschließend diskutiert wird, ist die zeitliche Bandbreite des optischen Kanals immens. Mit einer typischen Kohärenzlänge von Weißlichtquellen von wenigen Mikrometern erhält man Bandbreiten im Bereich von 10^{14} Hz.

Für das Rauschen des optischen Kanals ist (bei optisch rauen Objekten) kohärentes Specklerauschen verantwortlich. In der Optik wird das Ausmaß des Specklerauschens durch den Specklekontrast $c_{speckle}$ gekennzeichnet (siehe Gl. 3.10).

$$c_{speckle} = \frac{\sigma_I}{\bar{I}} \quad (5.11)$$

In der Nachrichtentechnik verwendet man stattdessen den Begriff Signal-Rausch-Verhältnis (SNR). Für ein großes Signal-Rausch-Verhältnis bzw. eine kleinen Specklekontrast lassen sich die beiden Darstellungen ineinander

umrechnen.

$$\frac{S_{opt}}{N_{opt}} = \frac{\bar{I}^2}{\sigma_I^2} \approx \frac{\bar{I}^2}{\sigma_I^2} = \frac{1}{c_{speckle}^2} \quad (5.12)$$

Hier ist σ_I die Standardabweichung der Intensität verursacht durch Speckle und \bar{I} die mittlere Intensität.

Der Specklekontrast wiederum setzt sich aus einem Beitrag der räumlichen und der zeitlichen Kohärenz sowie einem polarisationsabhängigen Beitrag und einem Mittelungseffekt der Pixel zusammen (siehe Gl. 3.11).

$$c_{speckle} = c_{spat} \cdot c_{temp} \cdot c_{pol} \cdot c_{pixel} \quad (5.13)$$

Großen Einfluss hat hierbei der Beitrag der räumlichen Kohärenz und der Pixel. Mit großer Beleuchtungsapertur bzw. Beobachtungsapertur, kann man eine Reduktion des Specklekontrastes auf Werte sehr viel kleiner als Eins erreichen und damit ein gutes Signal-Rausch-Verhältnis realisieren (Gl. 3.18).

$$c_{spat} \cdot c_{pixel} = \frac{\lambda}{\Delta x \cdot \max(\sin u_{ill}, \sin u_{obs})} \quad (5.14)$$

Der schlechteste Fall ist die Beleuchtung mit einer Laserlichtquelle, die einer punktförmigen Lichtquelle und hoher räumlicher Kohärenz entspricht und die Beobachtung mit kleiner Apertur. Demgegenüber erlauben ausgedehnte Lichtquellen eine wesentliche Reduktion des Kontrastes. Die Kapazität des optischen Kanals ist nun nach Gl. 5.4

$$C_{opt} = 4 \cdot \Delta X' \cdot B_x \cdot \Delta Y' \cdot B_y \cdot \Delta T \cdot B_t \cdot \log_2 \left(\frac{S_{opt}}{N_{opt}} \right) \quad (5.15)$$

Setz man für das Signal-Rausch-Verhältnis Gl. 5.12 ein, so kann man die optische Kanalkapazität mit Hilfe des Specklekontrastes ausdrücken

$$C_{opt} = 4 \cdot \Delta X' \cdot B_x \cdot \Delta Y' \cdot B_y \cdot \Delta T \cdot B_t \cdot \log_2 \left(\frac{1}{c_{speckle}^2} \right) \quad (5.16)$$

$$C_{opt} = -8 \cdot \Delta X' \cdot B_x \cdot \Delta Y' \cdot B_y \cdot \Delta T \cdot B_t \cdot \log_2 (c_{speckle}) \quad (5.17)$$

Abb. 5.4 zeigt das Bild einer Münze, einmal mit räumlich kohärenter, Abb. 5.5 mit räumlich inkohärenter Beleuchtung. Gleichung 5.17 enthält eine Bauanleitung für einen guten optischen Sensor, dessen optischer Kanal ein Maximum an Information übertragen kann: eine große Detektorfläche zusammen mit einer großen Beobachtungsapertur und räumlich inkohärenter Beleuchtung. Als nächstes wird der elektronische Kanal betrachtet.

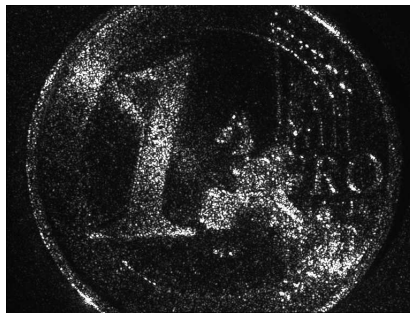


Abbildung 5.4: Bild einer Münze bei räumlich kohärenter Beleuchtung



Abbildung 5.5: Bild einer Münze bei räumlich inkohärenter Beleuchtung

5.1.2 Kapazität des elektronischen Kanals

Der elektronische Kanal wird gebildet vom Detektorchip und den analogen elektronischen Komponenten. Man geht davon aus, dass der Detektor $M_x \cdot M_y$ Bildelemente (Pixel) der Größe $\Delta x' \cdot \Delta y'$ besitzt. Die gesamte Ausdehnung des Detektors beträgt damit

$$\Delta X' = M_x \cdot \Delta x' \quad (5.18)$$

$$\Delta Y' = M_y \cdot \Delta y' \quad (5.19)$$

und stimmt mit der Größe des Bildfeldes der optischen Abbildung überein. Er nimmt eine Serie von M_t Bildern im zeitlichen Abstand Δt auf. Die Zeitspanne für die Bilderserie beträgt

$$\Delta T = M_t \cdot \Delta t \quad (5.20)$$

Da der elektronische Kanal diskret bezüglich der Ortsvariablen x und y ist, muss das Abtasttheorem beachtet werden. Dasselbe gilt auch für die Zeit t , denn nur zu diskreten Zeitpunkten werden Bilder aufgenommen. Die Abtastfrequenz beträgt

$$f_x = \frac{1}{\Delta x'} = \frac{M_x}{\Delta X'} \quad (5.21)$$

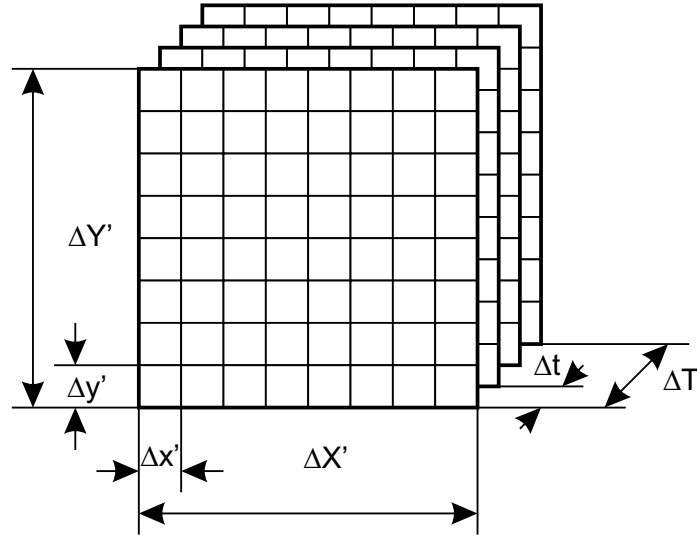


Abbildung 5.6: Detektor mit $M_x \cdot M_y$ Bildelementen der Größe $\Delta x' \cdot \Delta y'$ und einer Gesamtausdehnung von $\Delta X' \cdot \Delta Y'$. Der Detektor nimmt M_t Bilder im zeitlichen Abstand von Δt auf.

$$f_y = \frac{1}{\Delta y'} = \frac{M_y}{\Delta Y'} \quad (5.22)$$

$$f_t = \frac{1}{\Delta t} = \frac{M_t}{\Delta T} \quad (5.23)$$

Die Bandbreite des optischen Signals darf gemäß dem Abtasttheorem nicht größer sein als die halbe Abtastfrequenz des Detektors. (Anmerkung: Streng genommen kann ein Signal der endlichen Ausdehnung $\Delta X'$, $\Delta Y'$ und ΔT nicht bandbegrenzt sein. Als Abhilfe stelle man sich die periodische Fortsetzung in x - y - und t -Richtung vor.)

$$B_x = \frac{M_x}{2\Delta X'} \quad (5.24)$$

$$B_y = \frac{M_y}{2\Delta Y'} \quad (5.25)$$

$$B_t = \frac{M_t}{2\Delta T} \quad (5.26)$$

Dies stellt eine Einschränkung für die optische Bandbreite dar. Sie sollte wie bereits gesagt zwar möglichst groß sein, andererseits darf aber das Abtasttheorem nicht verletzt werden. In der Praxis löst man das Problem so, dass man die Apertur auf einen zulässigen Wert schließt oder die Kameraoptik

leicht defokussiert. Dies hat den Vorteil, dass immer noch genügend Licht zum Detektor gelangt, die Bandbreite aber trotzdem nicht zu groß ist.

Als Rauschquellen des elektronischen Kanals kommen verschiedene Effekte in Frage. Zum einen gibt es technologieabhängiges Rauschen von Kameras wie thermisches Rauschen, Resetrauschen etc. Diese Rauschquellen sind bei heutigen Kameras bereits sehr klein und werden deshalb nicht weiter betrachtet. Daneben gibt es aber eine andere Rauschquelle, die grundsätzlicher Natur ist. Das Photonenrauschen hängt allein davon ab, wie viele Photonen auf ein Pixel treffen. Für eine mittlere Anzahl von \bar{n} Photonen ist die Standardabweichung der Anzahl

$$\sigma_n = \sqrt{\bar{n}} \quad (5.27)$$

(siehe Abschnitt 3.5) und das Signal-Rauschverhältnis berechnet sich zu

$$\frac{S_{el}}{N_{el}} = \frac{\bar{n}^2}{\sigma_n^2} \approx \frac{\bar{n}^2}{\bar{n}} = \bar{n} \quad (5.28)$$

Setzt man die Bandbreite und das SNR in die Gleichung für die Kanalkapazität ein, so erhält man

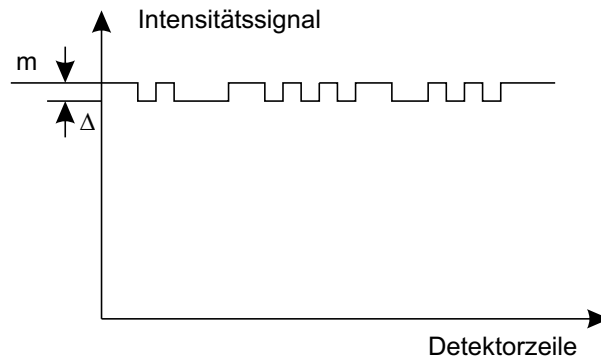
$$C_{el} = M_x \cdot M_y \cdot M_t \cdot \log_2(\sqrt{\bar{n}}) \quad (5.29)$$

Die Kapazität des elektronischen Kanals wird so groß wie möglich, wenn die Kamera maximal angesteuert wird. Natürlich ist darauf zu achten, dass die Kamera nicht übersteuert wird. Dazu ist eine hohe Full-Well-Capacity n_{max} wünschenswert.

Oben wurde bereits erwähnt, dass die Kanalkapazität ein Ausdruck des Aufwands ist, den man in der 3D-Messtechnik treiben muss. Dies trifft insbesondere auf den elektronischen Kanal zu. Der Preis einer Kamera steigt steil an mit der Zahl der Pixel und mit dem Signal-Rausch-Verhältnis. Deshalb wird man darauf achten, diese kostbare Resource möglichst effizient zu nutzen. Zunächst aber zurück zum dritten Teilstück des Kanals, dem diskreten Kanal.

5.1.3 Kapazität des diskreten Kanals

Der diskrete Kanal beschreibt den Einfluss der Digitalisierung auf das Kamerasignal. (Der Kanal wird für diese Herleitung als kontinuierlich bezüglich der Werte betrachtet). Das analoge Signal des Detektorchips wird mit Hilfe eines A/D-Wandlers in diskrete Werte quantisiert, wobei die Quantisierungsstufen Δ betragen (siehe Abb 5.7). Bei voller Aussteuerung erreicht der

Abbildung 5.7: Quantisierung des Detektorsignals in Stufen der Höhe Δ

A/D-Wandler den Wert m_{max} . Die Fehler der Diskretisierung äußern sich im Diskretisierungsrauschen mit der mittleren Leistung

$$N_{dis} = \sigma_{dis}^2 = \frac{1}{12}\Delta^2 \quad (5.30)$$

Wird der A/D-Wandler bis zum Wert m angesteuert, so ist die mittlere Leistung des Signals

$$S_{dis} = m^2 \cdot \Delta^2 \quad (5.31)$$

und das Signal-Rausch-Verhältnis

$$\frac{S_{dis}}{N_{dis}} = 12m^2 \quad (5.32)$$

Die Kapazität des diskreten Kanals ist damit

$$C_{dis} = \frac{1}{2} \cdot M_x \cdot M_y \cdot M_t \cdot \log_2 \left(\frac{S_{dis}}{N_{dis}} \right) \quad (5.33)$$

$$C_{dis} = M_x \cdot M_y \cdot M_t \cdot \log_2 \left(\sqrt{12} \cdot m \right) \quad (5.34)$$

Die Erfahrung mit gängigen Kameras lehrt, dass das Quantisierungsrauschen bereits bei einer 8-Bit-Quantisierung und guter Aussteuerung kleiner ist als das Photonenrauschen und somit völlig ausreicht. Erst bei kleiner Aussteuerung bringt die Quantisierung mit 10 oder auch 12 Bit Vorteile. Für eine hohe elektronische Kanalkapazität sollten geringe Aussteuerungen aber ohnehin vermieden werden.

Hiermit sind die drei Bestandteile des Kanals, optischer, elektronischer und diskreter Kanal erläutert. Im nächsten Abschnitt geht es darum, wie diese Bestandteil im Gesamtkanal zusammenwirken.

5.1.4 Gesamtkanal

Die einzelnen Komponenten des Gesamtkanals (optischer Kanal, elektronischer Kanal und diskreter Kanal) werden von einem Signal nacheinander durchlaufen. Die Kapazität ist also niemals größer als das Minimum der Einzelkapazitäten. Die genaue Kapazität des Gesamtkanals erhält man, wenn man alle Rauschsignale addiert und zu einem gesamten Rauschsignal zusammenfasst. Da die drei Rauschsignale als statistisch unabhängig voneinander angenommen werden können, ist die mittlere Leistung des gesamten Rauschens die Summe der jeweiligen Leistungen.

Das Rauschen der einzelnen Teilkanäle wurde bisher in unterschiedlichen Einheiten dargestellt, beim optischen Kanal in der Einheit der Intensität, beim elektronischen Kanal in Anzahl der Photonen und beim diskreten Kanal in Quantisierungsstufen. Um das Rauschen addieren zu können, wird es für alle Teilkanäle in Quantisierungsstufen angegeben. Man erhält

$$\sigma_{opt} = c_{speckle} \cdot m \cdot \Delta \quad (5.35)$$

$$\sigma_{el} = \sqrt{\frac{m \cdot m_{max}}{n_{max}}} \cdot \Delta \quad (5.36)$$

$$\sigma_{dis} = \frac{1}{\sqrt{12}} \cdot \Delta \quad (5.37)$$

Die Gesamtleistung des Rauschens ist dann

$$N = \sigma_{opt}^2 + \sigma_{el}^2 + \sigma_{dis}^2 \quad (5.38)$$

Die mittlere Leistung des Signals wird ebenfalls in Quantisierungsstufen ausgedrückt und beträgt

$$S = m^2 \cdot \Delta^2 \quad (5.39)$$

Damit ist das Signal-Rausch-Verhältnis des Gesamtkanals

$$\frac{S}{N} = \frac{m^2 \cdot \Delta^2}{\sigma_{opt}^2 + \sigma_{el}^2 + \sigma_{dis}^2} \quad (5.40)$$

und die Kapazität

$$C = \frac{1}{2} \cdot M_x \cdot M_y \cdot M_t \cdot \log_2 \left(\frac{S}{N} \right) \quad (5.41)$$

Ein Beispiel: Eine Kamera mit $M_x \cdot M_y = 1000 \cdot 1000$ Pixeln nimmt $M_t = 1$ Bilder auf. Sie besitzt eine Full-Well-Capacity von 20000 Elektronen entsprechend $n_{max} = 40000$ Photonen (Quantenwirkungsgrad 0,5). Das Signal wird

mit 8 *bit* entsprechend $m_{max} = 256$ Stufen quantisiert. Die Kamera ist voll angesteuert ($m = m_{max}$) und nimmt ein einziges Bild auf. Der Specklekontrast bei der Aufnahme beträgt 0,005.

Mit diesen Angaben berechnet man die Standardabweichung des optische Rauschens zu 1,28, das elektronische ebenfalls zu 1,28 und das diskrete zu 0,29 Quantisierungsstufen. Das Gesamtrauschen liegt bei 1,83 Quantisierungsstufen, das Signal-Rausch-Verhältnis bei 140^2 . Für die gesamte Kanalkapazität ergibt sich ein Wert von 7,1 *Mbit* entsprechend 7,1 *bit/Pixel*.

In diesem Beispiel stehen das optische Rauschen und das elektronische Rauschen im Vordergrund, das Quantisierungsrauschen tritt dahinter zurück. Diese Verhältnisse hat man immer dann, wenn genügend Licht zur Verfügung steht und die Kamera gut angesteuert ist. Wo immer möglich sollte man die Kamera gut aussteuern, denn dann hat bereits eine preisgünstige Kamera mit 8-bit Quantisierung nur noch vernachlässigbares Quantisierungsrauschen. Nur wenn eine geringe Aussteuerung unvermeidlich ist, sollte man aufwändigere Kameras mit 10- bzw. 12-bit Quantisierung in Erwägung ziehen.

Zurück zum optischen und elektronischen Rauschen. Das kohärente optische Specklerauschen gewinnt immer dann an Bedeutung, wenn es sich um kleine Messfelder und kleine Aperturen handelt (siehe Gl. 5.14). Wenn die objektseitige Pixelgröße Δx beispielsweise nur 100 Wellenlängen oder weniger beträgt, muss mit Specklerauschen gerechnet werden.

Bei sehr geringem kohärenten Rauschen bleibt als einzige bedeutende Rauschquelle das Photonenrauschen des elektronischen Kanals. Von einer Kamera mit einer Full-Well-Capacity von 20000 Elektronen – entsprechend 40000 Photonen (Quantenwirkungsgrad 0,5) – ist ein SNR von maximal 200^2 zu erwarten. Dies entspricht 7,6 bit pro Pixel. Eine höhere Full-Well-Capacity ist oft mit einem wesentlich höheren Preis der Kamera verbunden.

Welche Bedeutung hat die Kapazität des Gesamtkanals? Sie gibt die Information an, die maximal übertragen werden kann. Die Kanalkapazität ist damit auch Maß für den Aufwand, den man treiben muss. Insbesondere gilt dies für den elektronischen Kanal. Je mehr Bildelemente $M_x \cdot M_y$ eine Kamera hat und je höher die Bildfrequenz B_t , umso höher ist zwangsläufig ihr Preis. Dasselbe gilt für eine große Full-Well-Capacity n_{max} . Lediglich die Kapazität des optischen Kanals ist vergleichsweise billig. Ein Objektiv, das genau eine Bandbreite gemäß dem Samplingtheorem aufweist, ist meist um ein Vielfaches günstiger als der Detektor. Etwas teurer kann eine große Beleuchtungsapertur werden, die für ein geringes kohärentes Rauschen des optischen Kanals nötig ist.

Demgegenüber steht die Information $I(Z(x, y); Z_m(x, y))$, welche in Wirklichkeit übertragen wird. Hier ist $Z(x, y)$ der Zufallsprozess, der die wah-

re Form beschreibt und $Z_m(x, y)$ der Prozess, der die gemessene Form beschreibt. Man wünscht sich, dass die Transinformation der Kanalkapazität möglichst nahe kommt. Anders ausgedrückt: Bei gegebener Kanalkapazität soll die Transinformation maximiert werden, bzw. soll bei gegebener Transinformation die Kanalkapazität minimiert werden. Dazu ist es sinnvoll, den Informationswirkungsgrad η einzuführen. Er stellt das Verhältnis aus Transinformation und Kanalkapazität dar. Sein Wert liegt im besten Fall bei Eins, im schlechtesten bei Null.

$$\eta = \frac{I(Z; Z_m)}{C} \quad (5.42)$$

Mit seiner Hilfe kann man verschiedene optische 3D-Verfahren daraufhin untersuchen, wie effizient sie mit den zur Verfügung gestellten Ressourcen umgehen. Die Kanalkapazität im Nenner des Wirkungsgrads wurde bereits angegeben. Um die Information, die in Wirklichkeit übertragen wird, soll es im nächsten Abschnitt gehen.

5.2 Realisierte Transinformation

Die Kapazität des Gesamtkanals ist ein Maß dafür, wie viel Information ein optischer Sensor maximal übertragen kann. Die Information $I(Z, Z_m)$, die in Wirklichkeit übertragen wird, liegt darunter. Es stellt sich die Frage, wie man sie bestimmt. Die Beziehung

$$I(Z; Z_m) = h(Z_m) - h(N) \quad (5.43)$$

beschreibt die Transinformation als Differenz der differentiellen Entropien der gemessenen Größe Z_m und des Rauschens N [34]. Differentielle Entropien (kontinuierliche Signale) sind schwer zu bestimmen. Deshalb quantisiert man sowohl den Messwert Z_m als auch das Rauschen N . Man kann die Transinformation als Differenz von zwei gewöhnlichen Entropien (bei diskreten Signalen) nähern.

$$I(Z; Z_m) \approx H(Z_{mq}) - H(N_q) \quad (5.44)$$

Hier ist Z_{mq} die quantisierte gemessene Form und N_q das quantisierte Rauschen. Quantisiert man gerade so fein wie die Größe des Rauschens, so ist $H(N_q) \approx 0$. Man kann sich dabei an der Standardabweichung des Rauschens orientieren und erhält eine Schätzung der Transinformation.

$$I(Z; Z_m) \approx H(Z_{mq}) \quad (5.45)$$

Meist kennt man auch nicht den gesamten Prozess Z_{mq} , sondern nur eine konkrete Realisierung z_{mq} . Für diese Realisierung kann man dann die Transinformation berechnen

$$I(z; z_m) \approx H(z_{mq}) \quad (5.46)$$

Diese Informationsgröße wiederum kann man aus einem gemessenen Datensatz mit Hilfe gängiger verlustfreier Kompressionsprogramme nach oben abschätzen. Wenn sich ein Datensatz beispielsweise verlustfrei von 500kB auf 250kB komprimieren lässt, liegt die Information bei weniger als 250kB.

Wenn man einen optische 3D-Sensor bewerten will, geht man wie folgt vor. Man wählt ein typisches Objekt, das gemessen werden soll und nimmt den gemessenen Datensatz z_m auf. Die Messung wird so quantisiert, dass eine Quantisierungsstufe der Standardabweichung des Rauschens entspricht. Die resultierende Datei wird komprimiert. Die Größe der Datei ergibt eine Abschätzung der Transinformation $I(z; z_m)$. Aus der Transinformation und der Kanalkapazität kann nun der Informationswirkungsgrad η berechnet werden.

5.3 Der Informationswirkungsgrad

Der Informationswirkungsgrad für ein bestimmtes Sensorprinzip ist nach Gl. 5.42

$$\eta = \frac{I(Z; Z_m)}{C} \quad (5.47)$$

Misst man den Wirkungsgrad für verschiedene Sensorprinzipien, erhält man einen Anhaltspunkt für die Effizienz des Verfahrens. Solche Versuche wurden durchgeführt für einen Sensor zur Streifenprojektion (siehe Abschnitt 7.2). Für ein Weißlichtinterferometer (Kohärenzradar, kurz KORAD) und ein Stereoverfahren wurden Abschätzungen vorgenommen. In allen Fällen wurde dieselbe Kamera zu Grunde gelegt. Das Signal-Rausch-Verhältnis der Kamera liegt bei $S/N = 97^2$ entsprechend 6,6 bit pro Pixel. Das Objekt wurde vermessen und entsprechend der Höhe des Rauschens quantisiert.

Zunächst soll ein Verfahren der Streifenprojektion betrachtet werden. Bei der phasenmessenden Triangulation (PMT) [35] werden für eine 3D-Aufnahme $M_t = 8$ Bilder aufgenommen. Es wurde der Kopf einer Schaufensterpuppe vermessen. Die Standardabweichung des Rauschens liegt bei ca. $114\mu m$. Die Formdaten wurden auf $114\mu m$ quantisiert und ergeben kom-

primiert $6,1\text{bit}/\text{Pixel}$. Der Wirkungsgrad berechnet sich zu

$$\eta_{PMT} = \frac{H(Z; Z_m)}{\frac{1}{2} \cdot M_x \cdot M_y \cdot M_t \cdot \log_2 \left(\frac{S}{N} \right)} \quad (5.48)$$

$$\eta_{PMT} < \frac{6,1\text{bit}}{8 \cdot 6,6\text{bit}} \quad (5.49)$$

$$\eta_{PMT} < 12\% \approx \frac{1}{8} \quad (5.50)$$

Die Kapazität des Kanals wird also nur zu etwa $1/8$ genutzt. Es wird zwar eine Kapazität mit dem Wert von $M_t = 8$ Bildern bereitgestellt, dafür bekommt man aber nur Information im Wert eines einzigen Bildes, während ein Äquivalent von 7 Bildern ungenutzt bleibt. Es wird deutlich, dass sowohl Bedarf als auch Spielraum für Verbesserungen bleibt.

Bei anderen Verfahren liegt der Wirkungsgrad noch weitaus niedriger, so z.B. bei den stereoskopisch arbeitenden Verfahren: Ein binokulares Stereoverfahren mit $M_t = 2$ liefert Formdaten des selben Objektes mit gleicher Standardabweichung wie die phasenmessende Triangulation, allerdings können lateral keine feinen Details aufgelöst werden [1]. Das Stereoverfahren arbeitet mit einem Korrelationsverfahren mit einer typischen Größe der Korrelationsbereiche von $7 \cdot 7\text{Pixel}^2$. Damit ist die realisierte Transinformation um den Faktor 49 niedriger.

$$\eta_{stereo} = \frac{H(Z; Z_m)}{\frac{1}{2} \cdot M_x \cdot M_y \cdot M_t \cdot \log_2 \left(\frac{S}{N} \right)} \quad (5.51)$$

$$\eta_{stereo} < \frac{0,12\text{bit}}{2 \cdot 6,6\text{bit}} \quad (5.52)$$

$$\eta_{stereo} < 0,94\% \quad (5.53)$$

Binokulare Stereoverfahren mögen auf den ersten Blick sehr vorteilhaft erscheinen, da sie nur $M_t = 2$ Bilder benötigen. Betrachtet man aber den Wirkungsgrad von weniger als $1/100$, so wird dieser Vorteil mehr als relativiert.

Es soll hier noch ein drittes Verfahren aus dem Gebiet der Interferometrie betrachtet werden. Ein Weißlichtinterferometer [4] vermessen ein optisch rauhes Objekt über einen Bereich von 1mm in der Tiefe. Dazu wird alle 500nm (bei schnellem Verfahren) ein Interferenzbild aufgenommen ($M_t = 2000$). Die Formdaten werden auf $1\mu\text{m}$ quantisiert. Dies ergibt eine realisierte Transin-

formation von maximal $10 \text{ bit}/\text{Pixel}$.

$$\eta_{KORAD} = \frac{H(Z; Z_m)}{\frac{1}{2} \cdot M_x \cdot M_y \cdot M_t \cdot \log_2\left(\frac{S}{N}\right)} \quad (5.54)$$

$$\eta_{KORAD} < \frac{10 \text{ bit}}{2000 \cdot 6,6 \text{ bit}} \quad (5.55)$$

$$\eta_{KORAD} < 0,076\% \quad (5.56)$$

Dieses Verfahren ist in der Lage, Höhenstrukturen im Bereich weniger Mikrometer zu erfassen, dafür ist allerdings ein extrem hoher Preis im Sinn des Informationswirkungsgrades zu zahlen.

Am Wirkungsgrad der einzelnen Verfahren wird deutlich, dass eine Codierung nötig ist, um die Effizienz zu steigern. Was ist der Grund dafür, dass sehr viel weniger Information übertragen wird, als es die Kanalkapazität grundsätzlich zulässt?

Ein Grund liegt darin, dass bereits das Objektsignal z Redundanz enthält. Bei redundanten Signalen kann die Kanalkapazität nicht erreicht werden.

Ein weiterer Grund ist darin zu sehen, dass die Codierverfahren zusätzlich Redundanz hinzufügen. Beim Weißlichtinterferometer z.B. sind die Bilder einer Serie weitgehend gleich, was hoher Redundanz entspricht. Lediglich an Stellen, an denen sich Interferenzen zeigen, gibt es Änderungen.

Darüber hinaus gibt es aber noch einen dritten Grund. Es ist bisher kein Verfahren bekannt, das nur das Formsignal z codieren würde. Alle Verfahren übermitteln auch die Textur des Objektes. Die Reflektivität der Oberfläche spielt eine Rolle, wie immer man die Beleuchtung wählt. Dies ist bei Verfahren der Streifenprojektion der Fall, aber auch bei interferometrischen Verfahren. In Wahrheit hat der Kanal also nicht nur die Form des Objektes, sondern auch noch seine Textur zu übertragen. Daneben hängt die Intensität in einem Kamerabild auch vom Umgebungslicht ab. Bei manchen Verfahren wird deshalb auch die Helligkeit des Umgebungslichtes auf dem Objekt übertragen. Um diese Eigenschaften der Codierung und der Quelle geht es in dem nun folgenden Abschnitt.

5.4 Eigenschaften von Codierung und Quelle

Im Kommunikationssystem ist die Codierung dafür zuständig, das Signal der Quelle so aufzubereiten, dass es über den Kanal übertragen werden kann. Ziel ist es dabei, die übertragene Information bis zur Kanalkapazität zu steigern. Im Fall der optischen 3D-Sensoren ist die Codierung durch die Beleuchtungseinheit gegeben. Bei Verfahren der aktiven Triangulation wird beispielsweise ein Lichtpunkt (Laserpunktsensor), eine Linie (Laserlichtschnittverfahren)

oder ein flächiges Muster (Streifenprojektion und phasenmessende Triangulation) aufprojiziert. Verfahren der passiven Triangulation (passives Stereo) verwenden das Umgebungslicht zur Codierung. Die Verfahren der Interferometrie (an glatten und rauen Oberflächen) benutzen eine kohärente oder teilkohärente Beleuchtung zur Codierung. Die reflektierte oder gestreute Beleuchtungswelle ist als das codierte Signal zu betrachten. All diesen Verfahren ist gemeinsam, dass der Reflexionskoeffizient der Objektoberfläche bei der Codierung eingeht. Ob man möchte oder nicht, wird damit neben der dreidimensionalen Form des Objektes immer dessen Textur über den Kanal übertragen. Die Kapazität steht also nicht ungeteilt zur Verfügung, sondern wird mehrfach in Anspruch genommen. Damit nicht genug. Viele Verfahren der optischen 3D-Messtechnik sind dafür ausgelegt, auch ein erhebliches Maß an Umgebungslicht zuzulassen. Das Objekt kann nicht trennen zwischen aktiver Beleuchtung und Umgebungslicht, und das Umgebungslicht wird daher zwangsläufig in das codierte Signal eingehen. Man kann also zusammenfassen, dass die Resource Kanal sowohl für die 3D-Form des Objektes, als auch für die Textur und das Umgebungslicht in Anspruch genommen wird. Damit steht für die Übertragung der 3D-Form nur noch ein Teil der Kanalkapazität zur Verfügung.

Ausnahmen sind denkbar, z.B. wenn der Reflexionskoeffizient über der gesamten Objektoberfläche a priori bekannt ist oder wenn das Umgebungslicht vernachlässigt werden kann. Im allgemeinen liegt aber eine geteilte Nutzung des Kanals vor. Die Konsequenz daraus ist, dass der Informationswirkungsgrad in seiner bisherigen Form prinzipiell nicht den Wert Eins annehmen kann. Wenn man für eine grobe Abschätzung annimmt, dass 3D-Signal, Textursignal und das Signal des Umgebungslichtes jeweils gleich viel Information übertragen, folgt daraus, dass der Informationswirkungsgrad nicht größer als $1/3$ werden kann. Anders ausgedrückt, die Kanalkapazität (Aufwand) muss dreimal größer sein als der Nutzen (3D-Transinformation). Interessanterweise taucht dieser Faktor Drei bei den unterschiedlichsten Verfahren wieder auf:

1. Bei der Streifenprojektion mit zeitlichem Phasenschieben sind mindestens drei Aufnahmen nötig, um ein 3D-Messung mit identischer Bandbreite zu erhalten.
2. Beim räumlichen Phasenschieben nach Takeda [38], [39] erhält man mit einer Kameraaufnahme eine 3D-Messung mit einem Drittel der Bandbreite.
3. Auch die Interferometrie ist davon nicht ausgenommen. Beim zeitlichen Phasenschieben sind mindestens drei Aufnahmen nötig für eine 3D-Aufnahme gleicher Bandbreite.

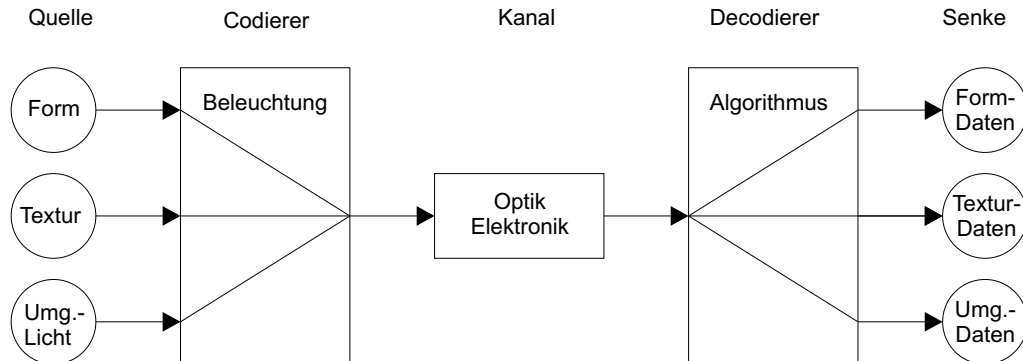


Abbildung 5.8: Erweitertes Übertragungsmodell für optische 3D-Sensoren, welches Form, Textur und Umgebungslicht berücksichtigt

Bei allen Verfahren taucht der Faktor Drei wieder auf, wenn die Textur und das Umgebungslicht unbekannt sind. Diesen Sachverhalt kann man bei der Darstellung des Kommunikationssystems berücksichtigen (siehe Abb. 5.8). Neben der ursprünglichen Quelle der Information sind die Textur und das Umgebungslicht mit aufgenommen. Alle drei gehen in die Codierung ein, was durch die entsprechenden Pfeile veranschaulicht wird.

Es stellt sich die Frage, ob die Textur und das Umgebungslicht Teil der Informationsquelle sind oder nicht. In der 3D-Messtechnik interessiert man sich primär für die Form des Objektes. Erkenntnisse über die Textur und das Umgebungslicht sind ein Nebenprodukt der Messung. Auf der anderen Seite wiederum interessiert man sich in der Bildverarbeitung hauptsächlich für die Textur von Objekten. Hier hat man wieder dieselbe Schwierigkeit. Auch die Textur kann nicht isoliert von der 3D-Form und dem Umgebungslicht übermittelt werden. Schattierungseffekte von Oberflächen mit Relief täuschen Textur vor, die in Wirklichkeit aus der 3D-Form herrührt. Es wird deutlich, dass die Interessen der 3D-Formerfassung und der Bildverarbeitung verknüpft sind. Optische 3D-Messtechnik liefert als Nebenprodukt die Textur und die Bildverarbeitung bekommt nur exakte Texturdaten, wenn nebenbei die 3D-Form abfällt. Man kann aus dieser Not eine Tugend machen. Man betrachtet die Fragestellung nach 3D-Form und nach der Textur als Einheit. Das Interesse richtet sich dann sowohl auf die Textur als auch auf die 3D-Form und ein Sensor kann beide Fragestellungen gleichermaßen lösen. In diesem Sinn kann man dann die 3D-Form und die Textur als Quelle von Information betrachten.

ähnlich kann man auch beim Umgebungslicht argumentieren. Es spielt vor allem dann eine Rolle, wenn man nicht nur an Eigenschaften eines Objektes, sondern an einer ganzen Szene interessiert ist. Dies ist der Fall, wenn man ein Objekt realistisch darstellen will. Eine realistische Darstellung erhält man, wenn das Umgebungslicht bekannt ist. Eine Szene im Freien erscheint völlig anders bei Sonnenaufgang, als dies um die Mittagszeit der Fall ist. Ebenso wirkt eine Szenerie völlig anders an einem sonnigen Tag als bei bedecktem Himmel. Insofern kann auch das Umgebungslicht von Interesse sein, man bezieht seine Transinformation beim Nutzen ein. Ein entsprechend modifizierter Informationswirkungsgrad, der im weiteren als erweiterter Informationswirkungsgrad bezeichnet wird, lautet dann

$$\eta_{ext} = \frac{I(Z; Z_m) + H(T; T_m) + H(U; U_m)}{C} \quad (5.57)$$

Hier ist T die Textur, U das Umgebungslicht bzw. T_m und U_m die gemessenen Werte.

Je mehr man (wie beim Umgebungslicht) Wert legt auf die realistische Darstellung eines Objektes, umso mehr spielt auch Farbe eine Rolle. Versteht man Textur und Umgebungslicht nicht nur als Grauwert, sondern als drei unabhängige Farbkomponenten, so teilt sich die Quelle der Texturinformation und der Umgebungslichinformation jeweils in drei Komponenten auf. Insgesamt hat man dann 7 Quellen der Information, eine für die 3D-Form, drei für die Textur und weitere drei für das Umgebungslicht. Bezieht man die Farbe mit ein, so sind demnach sehr viel höhere Anforderungen an die Kapazität des Kanals zu stellen.

Es ist zu bemerken, dass auch in der Kunst 3D-Form, Textur und Umgebungslicht wiederzufinden sind. In der Bildhauerei steht die 3D-Form im Vordergrund, während die Textur dahinter zurücktritt. So hat eine Skulptur im klassische Stil meist eine einheitlich weiße Textur und die Form tritt deutlich hervor. Bei einem Gemälde wiederum liegt die Betonung auf der Textur. Bei der Malerei spricht man von der Lokalfarbe, als einer Eigenschaft des Objektes. Auf der anderen Seite ist die Form des Gemäldes wiederum flach. Erst zu einem relativ späten Zeitpunkt der Kunstgeschichte hat das Umgebungslicht an Bedeutung gewonnen. Die Werke des Impressionismus legen hohen Wert auf das Umgebungslicht. Die Farbe ist hier nicht allein Eigenschaft des Objektes, sondern wird ganz wesentlich von den Eigenschaften des Umgebungslichtes bestimmt.

Zurück zum Kommunikationssystem. Nach der Codierung werden als nächstes die Eigenschaften der Quelle betrachtet. Auch sie stellen einen Grund dafür da, warum die meisten optischen 3D-Messverfahren einen geringen Informationswirkungsgrad besitzen. In der Informationstheorie ist be-

kannt, dass die Kanalkapazität nur erreicht werden kann, wenn das Signal am Eingang des Kanals unkorreliert ist. Genauer gesagt muss das Signal zum Zeitpunkt t_1 statistisch unabhängig sein vom Signal zum Zeitpunkt t_2 mit $t_1 \neq t_2$ und auch allen weiteren denkbaren Zeitpunkten. Was für die zeitliche Unabhängigkeit gilt, ist ebenso für die räumliche Unabhängigkeit richtig. Das Formsignal z an der Stelle (x_1, y_1) muss statistisch unabhängig sein vom Formsignal an der Stelle (x_2, y_2) . Die Erfahrung lehrt aber, dass dies bei den meisten Objekten nicht der Fall ist. Misst man beispielsweise die Höhe einer Tischoberfläche über dem Fußboden, so sind die Werte z sehr stark korreliert, da es sich näherungsweise um eine ebene Fläche handelt. Viele andere technische Objekte sind ebenfalls von ebenen Flächen begrenzt. Auch zylindrische und kugelförmige Oberflächen sind häufig anzutreffen, die ebenfalls stark korreliert sind. Bei natürlichen Objekten wie der Erdoberfläche, Pflanzen, Tieren und dem Mensch fällt eine Aussage nicht mehr ganz so leicht. Die Erdoberfläche hat auf einer großen Skala Kugelform, ist also stark korreliert. Je mehr man in kleinere Dimensionen vorstößt, umso mehr spielen Unebenheiten eine Rolle. So ist die Oberfläche eines Felsblocks keine Ebene. Doch selbst bei diesen unregelmäßigen Freiformflächen ist man von der statistischen Unabhängigkeit noch weit entfernt. Allein die Tatsache, dass man eine zusammenhängende Oberfläche im Raum angeben kann, ist schon ein Indiz dafür, dass benachbarte Höhenwerte nicht statistisch unabhängig voneinander sein können. Untersucht man das Leistungsdichtespektrum Φ_{zz} von technischen, aber auch natürlichen Objekten, so stellt man eine Gemeinsamkeit fest. Tiefe Frequenzen sind im Verhältnis zu den hohen Frequenzen überproportional vertreten. Eine schematische Darstellung bezüglich der Ortsfrequenz f_x (mit $f_y = 0$) ist in Abb. 5.9 dargestellt. Abbildung 5.10 zeigt das gemessene zweidimensionale Spektrum einer Münze. Die jeweiligen Frequenzanzeile sind grautoncodiert dargestellt. Die tiefen Frequenzen definieren die Grundform, die hohen definieren feine Strukturen und Rauigkeit. Es besteht ein Zusammenhang zwischen der Autokorrelationsfunktion $\varphi_{zz}(x, y)$ eines Signals und dessen Leistungsdichtespektrum $\Phi_{zz}(f_x, f_y)$. Das Leistungsdichtespektrum ist die Fourier-Transformierte der Autokorrelationsfunktion.

$$\Phi_{zz}(f_x, f_y) = F\{\varphi_{zz}(x, y)\} \quad (5.58)$$

Fordert man ein unkorreliertes Signal

$$\varphi_{zz}(x, y) = \text{const} \cdot \delta(x, y) \quad (5.59)$$

so muss das Leistungsdichtespektrum eine konstante Funktion sein [36].

$$\Phi_{zz}(f_x, f_y) = \text{const} \quad (5.60)$$

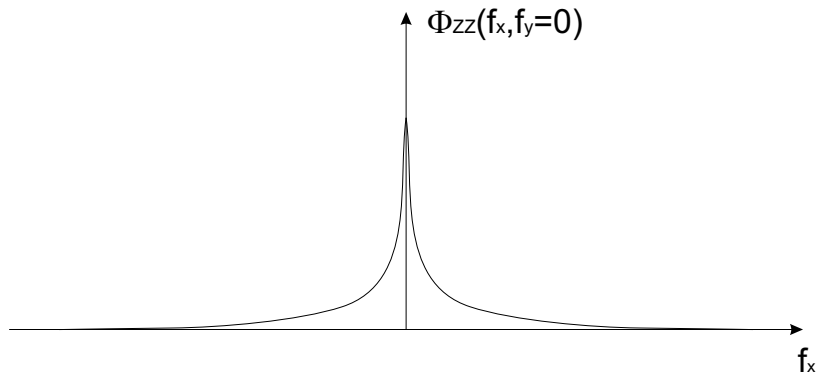


Abbildung 5.9: Leistungsdichtespektrum $\Phi_{zz}(f_x, f_y = 0)$ des Formsignals $z(x, y)$ einer typischen Objektoberfläche (schematisch)

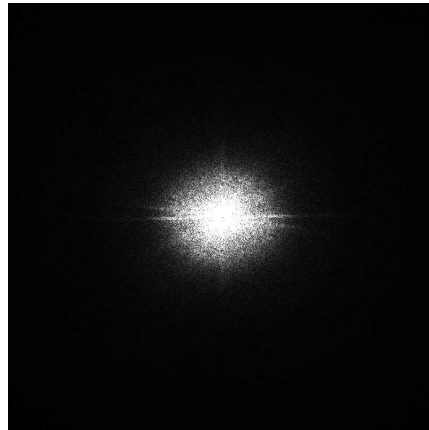


Abbildung 5.10: Leistungsdichtespektrum $\Phi_{zz}(f_x, f_y)$ des Formsignals $z(x, y)$ einer Münze (gemessen)

Ein solches Spektrum charakterisiert einen weißen Rauschprozess. Spätestens hier wird deutlich, dass die Form der meisten Objekte korreliert sein muss, weil sie sich nicht durch ein weißes Rauschen beschreiben lassen. Für gewöhnlich hat man im Spektrum hohe Werte bei den tiefen Frequenzen und niedrige Werte bei den hohen Frequenzen. Insbesondere dominiert der Gleichanteil. Misst man beispielsweise mit einem Sensor die Entfernung z zur Oberfläche eines Objektes, so ist der Gleichanteil durch den mittleren Arbeitsabstand gegeben. Es ist anschaulich klar, dass man nicht für jeden Messpunkt auf dem Objekt den mittleren Arbeitsabstand neu über den Kanal zu übermitteln braucht. Wenn man dies trotzdem tut, enthält das Signal am Eingang des Kanals Redundanz und die Kanalkapazität kann nicht ausgeschöpft werden. Das Ziel eines geeigneten Codierungsverfahrens muss es daher sein, niedrige Frequenzen und insbesondere den Gleichanteil zu dämpfen, während hohe Frequenzanteile angehoben werden. Nach der Codierung strebt man einen möglichst konstanten Frequenzverlauf an. Es muss diskutiert werden, welche Möglichkeiten man zur Codierung hat und inwieweit sie diesem Ziel nahekommen.

5.5 Möglichkeiten der Codierung

In der digitalen Nachrichtentechnik hat man ein weites Feld von Möglichkeiten, ein Signal zu codieren. Jede Codierungsvorschrift, die in eine digitale elektronische Schaltung oder ein Programm für einen Mikroprozessor übersetzt werden kann, ist grundsätzlich zur Codierung geeignet. Arithmetische und logische Operationen können schnell und ohne Verfälschung des Ergebnis durchgeführt werden. Im Bereich der optischen 3D-Messtechnik sind die Verhältnisse anders. Dort handelt es sich um analoge Signale. Diese Signale sollen codiert werden, über den optischen und elektronischen Kanal übermittelt und erst dann diskretisiert werden. Eine Diskretisierung vor dem Kanal ist nicht möglich. Bei der Manipulation von analogen Signalen sind die Möglichkeiten beschränkt [37].

Kohärente Filterverfahren der Fourier-Optik kommen in Frage an optisch glatten Oberflächen. Beim 4f-Aufbau erzeugt die erste Linse die Fourier-Transformierte eines einfallenden optischen Signals. Diese kann mit Hilfe von Filtern, die auf Amplitude und Phase wirken, manipuliert werden. Anschließend erzeugt die zweite Linse die Fourier-Rücktransformation und man erhält das gefilterte optische Signal. Anders als bei den digitalen Verfahren kommt es bei den analogen zu Verfälschungen. Analoge Filteroperation mit Hilfe der Fourier-Optik erzeugen Signale, die nur auf einige Prozent genau sind. Weit schwerer wiegt aber noch die Tatsache, dass diese Technik auf optisch glat-

te Oberflächen beschränkt ist. Die weitaus häufiger vorkommenden optisch rauen Oberflächen lassen keine kohärente Filterung zu, da dort die Phase der rückgestreuten Lichtwellen unvorhersehbar ist.

Schaut man sich bei den räumlich inkohärenten Verfahren um, so ist als wichtigstes die aktive Triangulation zu nennen. Hier wird die das Formsignal z mit Hilfe der Intensität der Beleuchtung codiert. Die Codierung erfolgt unabhängig von Nachbarpunkten der Objektoberfläche immer gleich. Das bedeutet, dass keine Möglichkeit zur Filterung besteht, denn Filterung bezieht Nachbarpunkte ein. Ebenso verhält es sich bei der passiven Triangulation. Somit scheidet neben den interferometrischen Verfahren auch die Triangulationsverfahren aus.

Bleiben dann überhaupt noch Verfahren übrig? Es gibt eine Gruppe von Verfahren, die nicht das Formsignal $z(x, y)$ an sich, sondern dessen Ableitung nach dem Ort x bzw. y messen. Die Ableitung ist eine Filteroperation, die die gewünschten Eigenschaften besitzt. Hohe Frequenzanteile werden verstärkt, während niedrige gedämpft werden. Insbesondere wird der Gleichanteil völlig eliminiert. Welches sind die differenzierenden Verfahren? Das sind alle Verfahren, in denen die Neigung der Oberfläche eine Rolle spielt. Mit dem Reflexionsgesetz an spiegelnden Oberflächen hat man einen Zusammenhang, der von der Neigung abhängt. Die zugehörigen Verfahren sind die deflektometrischen Verfahren. Sie messen die Neigung von optisch glatten Objekten mit Hilfe des Reflexionsgesetz. Ebenso existiert ein Verfahren für optisch raue Oberflächen. Streuende Oberflächen und insbesondere der Lambert-Streuer zeigen bei gerichteter Beleuchtung eine Abhängigkeit der gestreuten Intensität von der Neigung der Oberfläche. Der Winkel zwischen einfallender Beleuchtung und Normalenvektor der Oberfläche bestimmen, welche Intensität gestreut wird. Senkrechter Einfall auf die Oberfläche bewirkt maximale Intensität, während bei streifendem Einfall kein Licht mehr gestreut wird. Verfahren, die diese Gesetzmäßigkeit ausnutzen, kennt man unter dem Namen „Shape from Shading“ bzw. „photometrisches Stereo“. Von der Schattierung einer Objektoberfläche wird auf deren Form zurückgeschlossen.

Die genannten Verfahren sind also besonders zur Codierung geeignet, da sie für ein möglichst konstantes Spektrum sorgen und damit Redundanz der Objektform z eliminieren. Welches Potenzial besitzen diese Verfahren? Am besten lässt sich dies mit dem Informationswirkungsgrad η abschätzen.

Der Kopf einer Schaufensterpuppe, der bereits mit dem Verfahren der phasenmessenden Triangulation gemessen wurde, ist mit Verfahren des photometrischen Stereo erneut vermessen worden (siehe Abschnitt 7.2). Die Standardabweichung des Rauschens liegt bei ca. $44 \mu m$. Die Formdaten werden auf $44 \mu m$ quantisiert und ergeben komprimiert 9.9 bit/Pixel. Dieser Wert liegt über dem Wert der Phasenmessenden Triangulation (6,1 bit) und wird

mit 3 anstelle von 8 Aufnahmen erzielt. Damit ergibt sich für den Informationswirkungsgrad

$$\eta_{PhotoStereo} = \frac{H(Z; Z_m)}{\frac{1}{2} \cdot M_x \cdot M_y \cdot M_t \cdot \log_2 \left(\frac{S}{N} \right)} \quad (5.61)$$

$$\eta_{PhotoStereo} = \frac{9,9 \text{ bit}}{3 \cdot 6,6 \text{ bit}} \quad (5.62)$$

$$\eta_{PhotoStereo} = 50\% \quad (5.63)$$

Damit liegt der Informationswirkungsgrad etwa viermal höher als bei der phasenmessenden Triangulation. Es wird deutlich, dass eine geeignete Codierung den Informationswirkungsgrad wesentlich verbessern kann. Bei der analogen Codierung von optischen Signalen gibt es leider nur wenige Möglichkeiten. Das Verfahren der Phasenmessenden Deflektometrie sowie des photometrischen Stereo haben sich als gut geeignet erwiesen. Möglicherweise gibt es noch andere, unbekanntere Verfahren zur analogen optischen Codierung, die einen noch höhere Informationswirkungsgrad erreichen.

5.6 Single Shot Sensoren

Aus den Anfängen der Fotografie weiß man, dass damals Belichtungszeiten von vielen Sekunden üblich waren. Wollte man Personen fotografieren, so gab es die Schwierigkeit, während dieser Zeit stillzuhalten. Man behalf sich damit, dass der Kopf der Person mit einer Art Ständer fixiert wurde. Viele optische 3D-Sensoren erinnern zurück an diese Zeit. Messzeiten von einigen, manchmal sogar vielen Sekunden sind üblich. Will man Personen aufnehmen, so ist man häufig gezwungen, ähnliche Maßnahmen wie damals zu ergreifen. Die Person setzt sich beispielsweise auf einen speziellen Stuhl, mit dem der Kopf gehalten werden kann.

Die Entwicklung in der Fotografie ging weiter. Empfindlichere Fotoemulsionen und Blitzlampen halfen die Aufnahmezeit auf Bruchteile von Sekunden zu reduzieren. Es stellt sich die Frage, ob dies auch bei optischen 3D-Sensoren möglich ist.

Die meisten optischen 3D-Sensoren benötigen mehrere Kameraaufnahmen ($M_t > 1$). Typische Bildwiederholfrquenzen von elektronischen Kameras liegen bei 25 oder 30 Bildern pro Sekunde und reichen bis zu Tausenden Bildern pro Sekunde bei hochspezialisierten Kameras. Für die weitverbreitete kommerzielle Anwendung von optischen 3D-Sensoren sind solche Kameras vorerst aber meist zu teuer. Zudem kann bei sehr kurzen Belichtungszeiten die Kamera häufig nicht ausreichend angesteuert werden, da nicht genügend

Lichtenergie zur Verfügung steht. Bleibt also die Frage, was man mit einer Bildwiederholrate von etwa 25 bis 30 Bildern in der Sekunde erreichen kann. Nehmen wir an, ein Verfahren wie die phasenmessende Triangulation arbeite mit $M_t = 8$ Bildern und einer Rate von 25 Bildern in der Sekunde. Die gesamte Aufnahmezeit beträgt dann 320ms oder etwa $1/3$ Sekunde. Von der Fotografie weiß man, dass Aufnahmen mit einer solchen Belichtungszeit als sehr kritisch einzustufen sind. Schon die geringste Bewegung einer Person erzeugt Bewegungsunschärfe. Dasselbe ist der Fall bei der optischen 3D-Formfassung. Zugegeben, $1/3$ Sekunde ist ein großer Fortschritt gegenüber Messzeiten von mehreren Zehn Sekunden, die immer noch üblich sind. Dennoch ist mit dieser Belichtungszeit noch kein wirklicher Durchbruch erreicht. Selbst wenn man mit einer Serie von nur zwei Bildern auskommen könnte, bleibt eine Messzeit von immerhin ca. $1/12$ Sekunde. Der wirkliche Durchbruch liegt beim Übergang von zwei auf nur ein einziges Kamerabild. Man könnte argumentieren, dass für ein Kamerabild auch noch $1/25$ Sekunde nötig ist. Dies ist nur bedingt richtig. Die Zeit zum Auslesen eines Kameradetektors ist beispielsweise $1/25$ Sekunde, die Belichtungszeit kann aber deutlich niedriger liegen. Gängige elektronische Kameras ermöglichen problemlos Belichtungszeiten von weniger als $1/1000\text{ s}$. Daneben kann man auch durch ein Blitzlicht Belichtungszeiten von weniger als $1/10000\text{ s}$ erreichen. Der Durchbruch bei der Aufnahmezeit liegt also beim Übergang von zwei auf eine einzige Kameraaufnahme. Ein solcher Sensor wird im folgenden als Single Shot Sensor bezeichnet. Ist es möglich, mit der Kanalkapazität einer einzigen Aufnahme ($M_t = 1$) 3D-Messungen mit akzeptabler Transinformation zu gewinnen?

Dazu zurück zum erweiterten Informationswirkungsgrad η_{ext} . Man nimmt an, dass die Form, die Textur und das Umgebungslicht unbekannt sind. Der Kanal muss also alle drei Signale codieren und übermitteln. Zur Vereinfachung wird angenommen, dass die jeweilige Transinformation gleich groß ist. Dann kann man schreiben

$$\eta_{ext} = \frac{H(Z; Z_m) + H(T; T_m) + H(U; U_m)}{C} \quad (5.64)$$

$$\eta_{ext} = \frac{3 \cdot H(Z; Z_m)}{C} \quad (5.65)$$

Der erweiterte Informationswirkungsgrad ist kleiner Eins und es gilt

$$H(Z; Z_m) < \frac{1}{3} \cdot C \quad (5.66)$$

Bei einem Signal-Rausch-Verhältnis von $S/N = 97^2$ entfallen dann lediglich etwa $2,2\text{ bit}$ auf jedes Pixel. Dieser Wert reicht im allgemeinen nicht für eine

akzeptable 3D-Messung aus, insbesondere wenn man bei der Messung dieselbe Ortsbandbreite ($B_x \cdot B_y$) fordert wie die der Kameraaufnahme. Verzichtet man aber, wie beim räumlichen Phasenschieben nach Takeda [38], [39], auf einen Faktor 3 bei der Ortsbandbreite, so entfallen maximal $6,6 \text{ bit}$ auf jedes Pixel, was akzeptable Messungen möglich macht. Bereits an diesem Beispiel wird deutlich, dass ein Single Shot Sensor im engeren Sinne des Wortes (eine einzige Kameraaufnahme, keine Einschränkung beim Ortsbandbreitenprodukt, bei der Textur und beim Umgebungslicht) Formmessungen liefert, die im allgemeinen nicht brauchbar sind. Wenn man jedoch die Ansprüche lockert, sind Single Shot Sensoren im weiteren Sinn möglich. Die interessante Frage im Einzelfall ist also, auf welchen Anspruch man am ehesten verzichten kann.

Ein weiteres Beispiel: Bei manchen Objekten kann man davon ausgehen, dass der Reflexionsgrad konstant und die Textur also bekannt ist. Umgebungslicht wird bei der Messung nicht zugelassen. Man kann also auf eine Messung der Textur und des Umgebungslichtes verzichten. Es gibt ein Verfahren, das zur Gruppe Shape from Shading gehört [40], das genau diese Annahmen trifft. Die Intensitäten in einem einzelnen Bild geben Aufschluss darüber, wie stark eine Objektoberfläche geneigt ist. Aus der Neigung kann dann wiederum auf die dreidimensionale Form zurückgeschlossen werden.

Zugeständnisse kann man auch bei der Anzahl der Aufnahmen machen. Man lässt mehrere Aufnahmen zu unter der Bedingung, dass die Aufnahmen gleichzeitig erfolgen. Ein Beispiel dafür sind die Stereoverfahren mit zwei oder mehr Kameras.

Der Einsatz von mehreren Kameras für ein Single Shot Verfahren im weiteren Sinn scheint kein allzu großes Zugeständnis zu sein. Der Ansatz der Parallelisierung ist deshalb ein ausichtsreicher Kandidat für einen Single Shot Sensor. Leider liegen aber die Wirkungsgrade gerade bei den Stereoverfahren sehr niedrig ($\eta < 1\%$).

Eine Parallelisierung der Bildaufnahme kann auch erreicht werden, indem man die drei Kanäle einer Farbkamera verwendet. Beispielsweise kann man das Licht einer Weißlichtquelle spektral zerlegen und damit ein Objekt ausleuchten. Hier wird eine Codierung also mit Hilfe der Wellenlänge erreicht [41]. Leider tauchen auch bei diesem Verfahren Schwierigkeiten auf. Zum einen verliert man bei der spektralen Zerlegung einen Großteil der Lichtenergie. Zum zweiten findet bei der Codierung keine Reduktion von Redundanz statt, so dass die realisierte Transinformation bei weitem nicht an die Kanalkapazität heranreicht.

Man kann also festhalten: Mit den heute üblichen Signal-Rausch-Verhältnissen bei elektronischen Kameras kann es keinen Single Shot Sensor im engeren Sinn (eine einzige Aufnahme, volle Bandbreite des Formsignals,

Messung der Textur und des Umgebungslichtes) geben, der akzeptable Messungen produziert. Erst wenn man Teile dieser Forderungen aufgibt, kommt man zu einem Single Shot Sensor im weiteren Sinn, der akzeptable Messungen liefert.

Im folgenden Kapitel soll ein Sensor beschrieben werden, der einen hohen Informationswirkungsgrad realisieren kann. Dieser Sensor wird so modifiziert, dass ein Single Shot Sensor im erweiterten Sinn entsteht.

Kapitel 6

Ein optimierter optischer 3D-Sensor

Die Erkenntnisse des informationstheoretischen Sensormodells zeigen den Weg zu optimierten optischen 3D-Sensoren. Der Informationswirkungsgrad und der erweiterte Informationswirkungsgrad stellen Werkzeuge dar, mit deren Hilfe verschiedene optische Sensorprinzipien verglichen werden können. Gute Sensoren im informationstheoretischen Sinne zeichnen sich durch einen hohen Wirkungsgrad aus. Dieser wiederum wird maßgeblich durch die Eigenschaften der Quelle und der Codierung bestimmt. Es hat sich herausgestellt, dass das Signal z der Quelle bereits Redundanz enthält. Bei der Codierung ist darauf zu achten, dass diese Redundanz möglichst gut eliminiert und nicht etwa zusätzlich unnötig Redundanz hinzugefügt wird. Darüber hinaus muss man genau darauf achten, ob neben der Form des Objektes auch dessen Textur und das Umgebungslicht unbekannt sind. Je nach den Gegebenheiten mag einmal die eine, einmal die andere Codierungsmethode vorteilhaft sein.

Dieses Kapitel beschreibt in einem ersten Schritt einen Sensor mit hohem Wirkungsgrad, der neben der Form auch die Textur eines Messobjektes und das Umgebungslicht ermitteln kann und für allgemeine Zwecke eingesetzt werden kann.

Im zweiten Schritt liegt der Schwerpunkt auf einer extrem schnellen Datenaufnahme. Unter der Voraussetzung, dass kein Umgebungslicht und nur einfarbige Textur zugelassen wird, kann ein spezialisierter Single Shot Sensor im erweiterten Sinn beschrieben werden.

Zunächst zum Sensor für allgemeine Zwecke. Bei der Betrachtung von Quelle und Codierung hat sich gezeigt, dass für einen hohen Informationswirkungsgrad eine analoge optische Codierung benötigt wird, welche die hohen Frequenzen des Quellsignals z anhebt. Als mögliche Verfahren bieten sich die differenzierenden Verfahren wie die Deflektometrie und Verfahren aus der

Gruppe Shape from Shading an [42]. Andere Verfahren, die eine inkohärente analoge optische Codierung vornehmen, sind bislang nicht bekannt. Häusler [43] geht in seinem Beitrag ausführlich auf die Funktionsweise eines deflektometrischen Verfahrens ein. Die sogenannte phasenmessende Deflektometrie ermöglicht die Vermessung optisch glatter Freiformflächen mit äußerst geringer Messunsicherheit im Nanometerbereich. Dennoch benötigt dieses Verfahren nur $M_t = 8$ Aufnahmen. Dies ist ein gutes Beispiel für ein Verfahren, das einen hohen Informationswirkungsgrad aufweist. Zur Funktion des Verfahrens sei auf die entsprechende Veröffentlichung verwiesen [43].

Die besondere Rolle, welche die deflektometrischen Verfahren für optisch glatte Objekte spielen, wird bei den rauen Objekten von Verfahren des Shape from Shading übernommen. Bei Shape from Shading geht es darum, aus der Variation der Helligkeit in einem Bild auf die Form des abgebildeten Objektes zu schließen. Betrachtet man beispielsweise die Fotografie eines Gesichtes, so stellt man Variationen der Helligkeit fest, obwohl man annehmen kann, dass sich der Reflexionskoeffizient der Haut kaum ändert. Vielmehr entstehen diese Variationen dadurch, dass bestimmte Teile der Oberfläche so orientiert sind, dass sie mehr Licht zur Kamera strahlen als andere. Fällt das Licht einer Quelle senkrecht auf die Oberfläche, so ist die Helligkeit maximal, bei streifendem Einfall ist sie minimal. Ein anschauliches Beispiel bietet auch die Beleuchtung der Erdoberfläche durch die Sonne. Zur Mittagszeit im Sommer scheint die Sonne nahezu senkrecht auf die Erdoberfläche, was große Helligkeit bewirkt. Bei Sonnenuntergang streift das Licht gerade die Oberfläche und man hat geringe Helligkeit. Gelegentlich wird Shape from Shading mit dem Schattenwurf eines Gegenstandes verwechselt. Obwohl man auch aus der Schattenlinie eines Gegenstandes auf dessen Form schließen kann (Shape from Silhouette), ist dies bei Shape from Shading nicht gemeint. Shape from Shading gibt einen Hinweis auf die Neigung über die gesamte Oberfläche, nicht nur entlang einer Konturlinie. Bei Shape from Shading gibt es sehr unterschiedliche Fragestellungen. Eine gute Übersicht bieten Horn und Brooks [42]. Eine Gruppe von Verfahren beschäftigt sich damit, die Form einer als untexturiert angenommenen Oberfläche zu ermitteln. Andere Verfahren bestimmen neben der Form des Objektes auch die Position der Lichtquelle (Source from Shading). Beim photometrischen Stereoverfahren wiederum ist die Position der Lichtquellen a priori bekannt und Objekttextur ist zugelassen. Dies ist im Blick auf einen allgemein verwendbaren Sensor besonders interessant, denn oft ist die Textur nicht bekannt. Im weiteren geht es deshalb um ein photometrisches Stereoverfahren.

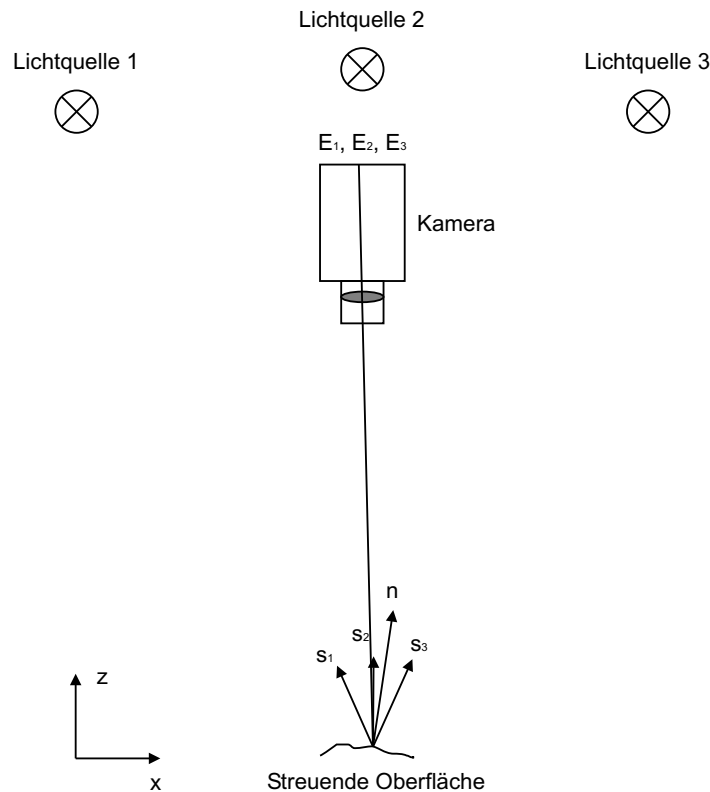


Abbildung 6.1: Funktionsprinzip des photometrischen Stereoverfahrens an streuenden Oberflächen

6.1 Photometrisches Stereo

Das Verfahren des photometrischen Stereo wird beispielsweise von Woodham beschrieben [44] und wird hier wiedergegeben. Man geht davon aus, dass ein zu vermessendes Objekt nacheinander von drei punktförmigen Lichtquellen beleuchtet wird. Im einfachsten Fall kann man annehmen, dass sich die Lichtquellen in großer Entfernung befinden und so die Beleuchtungsrichtung über dem Objekt für jede Quelle konstant bleibt. Besonders einfache Verhältnisse liegen vor, wenn die Oberfläche als Lambert-Streuer dargestellt werden kann. Beim Lambert-Streuer hängt die gestreute Intensität des Objektes nur von der Beleuchtungsrichtung und der Neigung des Objektes, nicht aber von der Beobachtungsrichtung ab. Für jede der drei Beleuchtungsquellen nimmt eine Kamera ein eigenes Bild auf. Die Position des Objektes und der Kamera bleiben dabei fest.

Zur mathematischen Darstellung des photometrischen Stereoverfahrens

benötigt man die drei Beleuchtungsrichtungen und den Normalenvektor \vec{n} der Objektoberfläche (siehe Abb. 6.1). Die Beleuchtungsrichtungen werden durch die Vektoren \vec{s}_1 , \vec{s}_2 und \vec{s}_3 beschrieben. Sie weisen von der Objektoberfläche zur jeweiligen Lichtquelle.

$$\begin{aligned}\vec{s}_1 &= (s_{11}, s_{12}, s_{13})^T \\ \vec{s}_2 &= (s_{21}, s_{22}, s_{23})^T \\ \vec{s}_3 &= (s_{31}, s_{32}, s_{33})^T\end{aligned}\tag{6.1}$$

Da sich die Lichtquellen in großer Entfernung befinden, bleiben diese Vektoren für alle Oberflächenpunkte näherungsweise konstant. Der Normalenvektor $\vec{n}(x, y)$ dagegen variiert entsprechend der Form der Oberfläche.

$$\vec{n}(x, y) = (n_x(x, y), n_y(x, y), n_z(x, y))^T\tag{6.2}$$

Es wird angenommen, dass die Oberfläche $z(x, y)$ differenzierbar ist und der Normalenvektor überall existiert. Die Kamera nimmt die Bilder E_1 , E_2 und E_3 des Objektes auf, jeweils ein Bild mit Beleuchtungsquelle 1, 2 und 3. Die Kamera befinde sich in großer Entfernung auf der z -Achse, und jedes Bildelement nimmt die einfallende Beleuchtungsstärke $E_i(x, y)$ ($i = 1, 2, 3$) auf, die den Koordinaten (x, y) der Oberfläche zugeordnet werden kann. Nach dem Lambert-Gesetz variiert die gestreute Leuchtdichte mit dem Kosinus zwischen der Beleuchtungsrichtung \vec{s}_i und dem Normalenvektor \vec{n} . Alternativ kann man die gestreute Leuchtdichte auch über das Skalarprodukt von Beleuchtungsrichtung und Normalenrichtung ausdrücken. Diese Darstellung hat den Vorteil, dass die Zusammenhänge linear dargestellt werden können. Neben der Neigung der Oberfläche und der Beleuchtungsrichtung hängt die gestreute Leuchtdichte auch vom lokalen Reflexionskoeffizienten $\rho(x, y)$ der Oberfläche (Textur) und der Beleuchtungsstärke der Quellen und den Parametern der Kameraoptik ab. Alle konstanten Faktoren wie die Beleuchtungsstärke der Quelle und die Parameter der Kameraoptik werden in der Länge des Beleuchtungsvektors zusammengefasst. Damit kann man schreiben

$$\begin{aligned}E_1 &= \rho \cdot \vec{s}_1 \cdot \vec{n} \\ E_2 &= \rho \cdot \vec{s}_2 \cdot \vec{n} \\ E_3 &= \rho \cdot \vec{s}_3 \cdot \vec{n}\end{aligned}\tag{6.3}$$

Diese drei Gleichungen kann man zu einer einzigen in Matrixschreibweise zusammenfassen, wenn man folgende Schreibweise für die Bilder einführt.

$$\vec{E} = (E_1, E_2, E_3)^T\tag{6.4}$$

Die Beleuchtungsvektoren bilden die Zeilen der Beleuchtungsmatrix

$$S = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix} \quad (6.5)$$

Damit kann man schreiben

$$\vec{E} = \rho S \vec{n} \quad (6.6)$$

Diese Gleichung beschreibt, wie der Codierer des Kommunikationssystems (die Beleuchtungseinheit) die Form des Objektes in Intensitätswerte übersetzt. Die Neigung, gegeben durch den Normalenvektor \vec{n} , wird multipliziert mit der Beleuchtungsmatrix S und anschließend mit dem skalaren Reflexionskoeffizient ρ . Sowohl die Form als auch die Textur $\rho(x, y)$ werden über den Kanal übertragen. Der Decodierer wiederum rekonstruiert aus den gemessenen Beleuchtungsstärken die Neigung bzw. die Form des Objektes. Aufgelöst nach \vec{n} bzw. ρ erhält man

$$\vec{n} = \frac{1}{\rho} S^{-1} \vec{E} \quad (6.7)$$

mit

$$\rho = \left\| S^{-1} \vec{E} \right\| \quad (6.8)$$

Die Beleuchtungsmatrix S lässt sich immer dann invertieren, wenn die Beleuchtungsvektoren linear unabhängig sind, d. h. wenn das Objekt und die drei Beleuchtungsquellen nicht in einer Ebene liegen.

Anstelle des Normalenvektors kann man die Neigung der Oberfläche auch durch deren partielle Ableitungen nach x und y darstellen. Diese Darstellung ist vorteilhaft für die spätere Integration des Formsignals $z(x, y)$.

$$p = \frac{\partial z}{\partial x} = -\frac{n_x}{n_z} \quad (6.9)$$

$$q = \frac{\partial z}{\partial y} = -\frac{n_y}{n_z} \quad (6.10)$$

Diese Beschreibung des photometrischen Stereo gilt für Lambert-Streuer und weit entfernte Beleuchtungsquellen in drei Dimensionen. Für die spätere Abschätzung der Messunsicherheit ist allerdings eine weitere Vereinfachung auf nur zwei Dimensionen, beispielsweise auf x und z sinnvoll. Eine solche Beschreibung wird im folgenden angegeben.

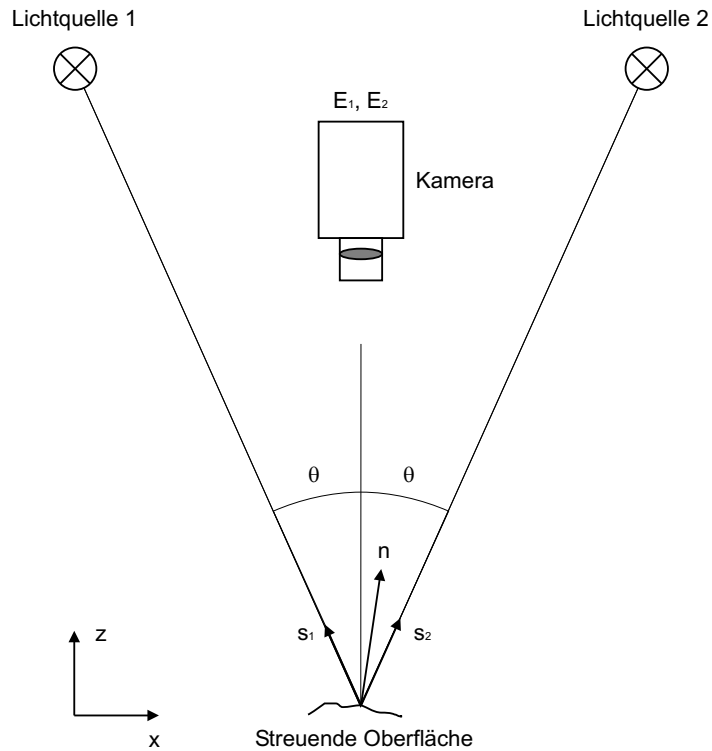


Abbildung 6.2: Vereinfachtes photometrisches Stereoverfahren an streuenden Oberflächen

6.1.1 Vereinfachtes photometrisches Stereo

Beim vereinfachten photometrischen Stereoverfahren werden anstelle von drei nur zwei Dimensionen betrachtet, eine Dimension senkrecht zur Beobachtungsrichtung, z.B. x und die Dimension in Beobachtungsrichtung z . Anstelle von drei Beleuchtungsquellen sind dann nur noch zwei nötig (Abb. 6.2). Reduziert auf die Dimensionen x und z kann man Gleichung 6.6 wie folgt schreiben.

$$\begin{pmatrix} E_1 \\ E_2 \end{pmatrix} = \rho \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} \begin{pmatrix} n_x \\ n_z \end{pmatrix} \quad (6.11)$$

Weiter geht man davon aus, dass die Lichtquellenvektoren symmetrisch zur z -Achse sind und mit ihr den Winkel θ einschließen. Dies stellt keine wesentliche Einschränkung dar, weil man ohnehin in der Praxis versuchen wird, einen symmetrischen Aufbau zu realisieren. Die Beleuchtungsvektoren haben beide

den Betrag E_0 .

$$\begin{aligned}\vec{s}_1 &= E_0 \cdot (-\sin \theta, \cos \theta)^T \\ \vec{s}_2 &= E_0 \cdot (\sin \theta, \cos \theta)^T\end{aligned}\quad (6.12)$$

Einsetzen der Beleuchtungsmatrix liefert

$$\begin{pmatrix} E_1 \\ E_2 \end{pmatrix} = \rho \cdot E_0 \begin{pmatrix} -\sin \theta & \cos \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} n_x \\ n_z \end{pmatrix}\quad (6.13)$$

Das Verfahren ist umso empfindlicher für Änderungen der Normalen, je größer der Winkel θ ist. Dies ist auch anschaulich klar, denn streifende Beleuchtung $\theta = 90^\circ$ macht jede kleinste Unebenheit sichtbar. Aufgelöst nach dem Normalenvektor ergibt sich

$$\begin{pmatrix} n_x \\ n_z \end{pmatrix} = \frac{1}{2\rho E_0 \sin \theta \cos \theta} \begin{pmatrix} -\cos \theta & \cos \theta \\ \sin \theta & \sin \theta \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \end{pmatrix}\quad (6.14)$$

Die Neigung p ergibt sich dann zu

$$p = \frac{\partial z}{\partial x} = -\frac{n_x}{n_z} = \cot \theta \cdot \frac{E_1 - E_2}{E_1 + E_2}\quad (6.15)$$

Die Neigung ist also der Kontrast der Beleuchtungsstärken E_1 und E_2 multipliziert mit einer Konstanten, die den Ort der Beleuchtungsquellen berücksichtigt. Den zulässigen Wertebereich für p erhält man aus der Bedingung, dass sowohl E_1 als auch E_2 nicht negativ sein dürfen (im Grenzfall streift die Beleuchtung die Oberfläche).

$$\begin{aligned}p_{min} &= -\cot \theta \\ p_{max} &= \cot \theta\end{aligned}\quad (6.16)$$

Für die Praxis bedeutet dies, dass stark geneigte Objekte mit kleinem Winkel θ vermessen werden müssen, schwach geneigte dagegen können mit großem Winkel θ und entsprechend höherer Empfindlichkeit erfasst werden.

6.2 Messunsicherheit

Mit dem vereinfachten photometrischen Stereoverfahren ist es nun möglich, die Messunsicherheit zu untersuchen. Für eine Messung werden zunächst die Beleuchtungsstärken E_1 und E_2 mit einer Kamera aufgenommen. Die Beleuchtungsstärken sind aber durch Rauschen des Übertragungskanals (kohärentes Rauschen, Photonenrauschen und Diskretisierungsrauschen, siehe Kapitel 5.1) verfälscht. Anschließend werden die Beleuchtungsstärken decodiert

und man erhält die Neigung p als Funktion der Beleuchtungsstärken.

$$p = f(E_1, E_2) = \cot \theta \cdot \frac{E_1 - E_2}{E_1 + E_2} \quad (6.17)$$

Gemäß dem Gauss'schen Fehlerfortpflanzungsgesetz wirkt sich das Rauschen der Beleuchtungsstärken $\sigma_{E_1}^2$ und $\sigma_{E_2}^2$ auf das Rauschen der Neigungsdaten σ_p^2 aus.

$$\sigma_p^2 = \left(\frac{\partial f}{\partial E_1} \right)^2 \sigma_{E_1}^2 + \left(\frac{\partial f}{\partial E_2} \right)^2 \sigma_{E_2}^2 \quad (6.18)$$

Die partiellen Ableitungen der Funktion $f(E_1, E_2)$ sind

$$\frac{\partial f}{\partial E_1} = \cot \theta \cdot \frac{2E_2}{(E_1 + E_2)^2} \quad (6.19)$$

$$\frac{\partial f}{\partial E_2} = \cot \theta \cdot \frac{-2E_1}{(E_1 + E_2)^2} \quad (6.20)$$

Das Signal-Rausch-Verhältnis S/N des Gesamtkanals (siehe Gl. 5.40) gibt das Signal-Rausch-Verhältnis der Beleuchtungsstärken E_1 und E_2 an.

$$\begin{aligned} \frac{E_1^2}{\sigma_{E_1}^2} &= \frac{S}{N} \\ \frac{E_2^2}{\sigma_{E_2}^2} &= \frac{S}{N} \end{aligned} \quad (6.21)$$

Für die Rauschleistungen gilt also

$$\begin{aligned} \sigma_{E_1}^2 &= \frac{1}{S/N} \cdot E_1^2 \\ \sigma_{E_2}^2 &= \frac{1}{S/N} \cdot E_2^2 \end{aligned} \quad (6.22)$$

Damit erhält man für die Standardabweichung der Neigung

$$\sigma_p = \frac{\cot \theta}{\sqrt{2} \cdot S/N} \cdot \frac{4E_1 E_2}{(E_1 + E_2)^2} \quad (6.23)$$

Drückt man die Beleuchtungsstärken E_1 und E_2 mit Hilfe der Neigung

$$p = \cot \theta \cdot \frac{E_1 - E_2}{E_1 + E_2} \quad (6.24)$$

aus, so erhält man

$$\sigma_p = \frac{\cot \theta}{\sqrt{2} \cdot S/N} \cdot (1 - p^2 \cdot \tan^2 \theta) \quad (6.25)$$

Somit geht das Signal-Rausch-Verhältnis S/N des Gesamtkanals direkt in das Rauschen der Neigungsdaten ein. Je nach Neigung ist der Einfluss aber unterschiedlich groß. Trägt man die Standardabweichung in Abhängigkeit der Neigung auf, ergibt sich eine nach unten geöffnete Parabel mit dem Maximum für $p = 0$ und den Minima Null am Rand des Wertebereichs von p .

$$\begin{aligned}\sigma_{p,max} &= \frac{\cot \theta}{\sqrt{2 \cdot S/N}} \\ \sigma_{p,min} &= 0\end{aligned}\tag{6.26}$$

Das maximale Rauschen der Neigungsdaten ist also im wesentlichen so groß wie das Rauschen der Beleuchtungsstärken. Das Signal-Rausch-Verhältnis der Beleuchtungsstärke wird lediglich durch den Geometriefaktor $\cot \theta$ und einen Faktor $1/\sqrt{2}$ umgerechnet. Um möglichst geringes Rauschen zu erhalten, sollte ein großer Winkel θ (große Empfindlichkeit) gewählt werden. Dabei darf $\cot \theta$ aber nicht kleiner als die maximale Neigung des Objektes sein.

6.2.1 Messunsicherheit zwischen benachbarten Pixeln

Bislang wurde nur die Messunsicherheit der Neigung $p = \partial z / \partial x$ betrachtet. Letztendlich ist aber nicht die Neigung, sondern das integrierte Formsignal $z(x)$ von Interesse. Wie wirkt sich das Rauschen der Neigungswerte auf das Rauschen des Formsignals aus? Für eine vereinfachte Berechnung werden zunächst nur benachbarte Bildelemente der Kamera und deren zugehörige Höhenwerte z betrachtet. Die Höhendifferenz zwischen benachbarten Pixeln beträgt

$$\Delta z = p \cdot \Delta x\tag{6.27}$$

Δx bezeichnet die objektseitige Pixelgröße (siehe Abb. 3.2). Für die Standardabweichungen gilt entsprechend

$$\sigma_{\Delta z} = \sigma_p \cdot \Delta x\tag{6.28}$$

Im ungünstigsten Fall für $p = 0$ ergibt sich die statistische Messunsicherheit zwischen benachbarten Pixeln zu

$$\sigma_{\Delta z} = \frac{\cot \theta}{\sqrt{2 \cdot S/N}} \cdot \Delta x\tag{6.29}$$

Man könnte annehmen, dass die Messunsicherheit immer kleiner wird, wenn man die objektseitige Pixelgröße verkleinert und das Objekt feiner auflöst. Dies ist nicht der Fall, denn das Signal-Rausch-Verhältnis des Kanals und die objektseitige Pixelgröße sind gekoppelt. Für kleine Δx dominiert das

Specklerauschen (siehe Abschnitt 5.1.4) und das Signal-Rausch-Verhältnis ist gegeben durch

$$\frac{S}{N} = \frac{S_{opt}}{N_{opt}} = \frac{1}{c_{speckle}^2} \quad (6.30)$$

(siehe Gleichung 5.12). Der Specklekontrast wiederum ist

$$c_{speckle} = \frac{\lambda}{\Delta x \cdot \max(\sin u_{ill}, \sin u_{obs})} \cdot c_{temp} \cdot c_{pol} \quad (6.31)$$

(siehe Gleichung 3.18). Im Nenner des Specklekontrasts taucht die objektseitige Pixelgröße Δx wieder auf. Einsetzen des Specklekontrastes liefert

$$\sigma_{\Delta z} = \frac{\cot \theta}{\sqrt{2}} \cdot \frac{\lambda}{\max(\sin u_{ill}, \sin u_{obs})} \cdot c_{temp} \cdot c_{pol} \quad (6.32)$$

Das Signal-Rausch-Verhältnis und die objektseitige Pixelgröße sind genau so gekoppelt, dass sich ihr Einfluss auf die Messunsicherheit (für kleines Δx) gerade weghebt. Unabhängig, wie klein Δx ist, bleibt doch die Messunsicherheit gleich. Für große Beleuchtungs- oder Beobachtungsaperturen (z.B. 0,5) und einen Beleuchtungswinkel von $\theta = 45^\circ$ werden Messunsicherheiten in der Größenordnung der Wellenlänge λ erreicht.

$$\begin{aligned} \sigma_{\Delta z} &= \frac{\cot 45^\circ}{\sqrt{2}} \cdot \frac{\lambda}{0.5} \cdot c_{temp} \cdot c_{pol} \\ \sigma_{\Delta z} &= \sqrt{2} \cdot \lambda \cdot c_{temp} \cdot c_{pol} \end{aligned} \quad (6.33)$$

Bei schwach geneigten Objekten (kleines p_{max}) kann man entsprechend Gl. 6.16 den Beleuchtungswinkel θ größer wählen:

$$\begin{aligned} p_{min} &= -\cot \theta \\ p_{max} &= \cot \theta \end{aligned} \quad (6.34)$$

Damit lassen sich noch geringere Messunsicherheiten erreichen, die in dem Bereich liegen, die sonst von interferometrischen Verfahren erreicht werden. Bei fast streifender Beleuchtung $\theta = 85^\circ$ erhält man

$$\begin{aligned} \sigma_{\Delta z} &= \frac{\cot 85^\circ}{\sqrt{2}} \cdot \frac{\lambda}{0.5} \cdot c_{temp} \cdot c_{pol} \\ \sigma_{\Delta z} &= 0,12 \cdot \lambda \cdot c_{temp} \cdot c_{pol} \end{aligned} \quad (6.35)$$

Im Gegensatz zu den interferometrischen Verfahren hat man beim photometrischen Stereo aber den Vorteil, dass man mit räumlich inkohärenter Beleuchtung arbeiten kann und daher eine sehr viel geringere laterale Messunsicherheit erzielt. Ein weiterer Vorteil liegt darin, dass beim photometrischen

Stereo nur drei Aufnahmen gebraucht werden, während in der Weißlichtinterferometrie je nach Ausdehnung des Messfelds mehrere hundert Aufnahmen nötig sind.

Formt man Gleichung 6.32 ein wenig um, so erkennt man die Verwandtschaft des photometrischen Stereoverfahrens zu den Triangulationsverfahren, insbesondere der phasenmessenden Triangulation.

$$\sigma_{\Delta z} = \frac{1}{\sqrt{2} \tan \theta} \cdot \frac{\lambda}{\max(\sin u_{ill}, \sin u_{obs})} \cdot c_{temp} \cdot c_{pol} \quad (6.36)$$

Bei der phasenmessenden Triangulationsverfahren ist die Messunsicherheit nach Lampalzer [45]

$$\sigma_{\Delta z, PMT} = \frac{1}{2\pi \cdot \tan \theta} \cdot \frac{\lambda}{\max(\sin u_{ill}, \sin u_{obs})} \cdot c_{temp} \cdot c_{pol} \quad (6.37)$$

Der Beleuchtungswinkel θ beim photometrischen Stereo entspricht dem Triangulationswinkel bei der Triangulation. Es gibt allerdings auch Unterschiede. Bei den Triangulationsverfahren wird eine longitudinale Ortsänderung Δz in eine laterale Ortsänderung Δx übersetzt. Beim photometrischen Stereo taucht eine longitudinale Ortsänderung nicht auf. Hier wird vielmehr eine Neigung in eine Intensität übersetzt.

6.2.2 Gesamte Messunsicherheit

Bisher wurde nur die Messunsicherheit zwischen benachbarten Pixeln untersucht. Entfernt man sich nun, statt einem Pixel, k Pixel vom Ausgangspixel, so ergibt sich eine größere Unsicherheit. Dabei entsteht der Höhenwert Δz_k durch Summation aus Höhenänderungen für benachbarte Pixel Δz_i .

$$\Delta z_k = \sum_{i=1}^k \Delta z_i \quad (6.38)$$

Da die Fehler benachbarter Pixel als statistisch unabhängig angenommen werden können (solange die Specklegröße kleiner als die Pixelgröße ist), ergibt sich

$$\sigma_{\Delta z, k} = \sqrt{k} \cdot \sigma_{\Delta z} \quad (6.39)$$

Für typische Werte von $k < 1000$ Pixel in einer Zeile des Detektors sind damit Messunsicherheiten zu erwarten, die um den Faktor 32 höher liegen als die von benachbarten Pixeln. Dieser Nachteil lässt sich reduzieren, indem man das photometrische Stereoverfahren mit einem binokularen Stereoverfahren kombiniert. Beide Verfahren ergänzen sich sehr gut. Das photometrische Stereoverfahren sorgt für eine geringe lokale Messunsicherheit (hohe

Ortsfrequenzen), während das biokulare Stereoverfahren die globale Messunsicherheit minimiert (kleine Ortsfrequenzen). Angenommen, das biokulare Stereoverfahren liefert alle $k = 10$ Pixel einen rauschfrei angenommenen Höhenwert, so erhält man

$$\sigma_{\Delta z, k} = 3,2 \cdot \sigma_{\Delta z} \quad (6.40)$$

Dieser Wert erhöht sich dann noch durch die Messunsicherheit des binokularen Stereoverfahrens.

Zusammenfassend lässt sich feststellen, dass das Verfahren des photometrischen Stereo in der Lage ist, lokale Höhenunterschiede in der Größenordnung einer Wellenlänge zu detektieren, bei streifender Beleuchtung sogar noch darunter (siehe auch Messungen im Abschnitt 7.7). Globale Höhenunterschiede auf dem Objekt sind mit einer größeren Messunsicherheit behaftet. Abhilfe schafft die Kombination des photometrischen Stereoverfahrens mit einem binokularen Stereoverfahren, das in der Lage ist, globale Höhenunterschiede mit geringer Messunsicherheit zu erfassen.

Nach der Betrachtung der Messunsicherheit anhand der vereinfachten Beschreibung $z(x)$ wird nun wieder die allgemeine Beschreibung $z(x, y)$ benötigt. Es geht darum, wie aus der Neigung $p = \partial z / \partial x$ und $q = \partial z / \partial y$ das Formsignal $z(x, y)$ integriert wird.

6.3 Integration der Neigungsdaten

Bereits die Integration einer Ableitung $p = dz/dx$, die von einer einzigen Variablen abhängt, kann Schwierigkeit bereiten, umso mehr die Integration einer partiellen Ableitungen $p = \partial z / \partial x$ und $q = \partial z / \partial y$, die von zwei Variablen abhängen. In zwei Dimensionen gibt es unendlich viele Möglichkeiten, von einem Ausgangspunkt (x_0, y_0) zu einem Zielpunkt (x, y) zu integrieren. Zudem sind in unserem Fall die partiellen Ableitungen p und q durch die Messung rauschbehaftet. Unterschiedliche Integrationswege können deshalb zu unterschiedlichen Ergebnissen führen. Um dieses Problem zu lösen, benutzt man Begriffe aus der Feldtheorie (Gradient, Rotation und Divergenz), die sich unter anderem mit solchen Problemen beschäftigt.

Man nimmt an, dass die Oberfläche des Objektes $z(x, y) = z(\vec{x})$ stetig differenzierbar ist. Die Neigung ist gegeben durch

$$\vec{f} = \nabla z = \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y} \right) \quad (6.41)$$

Bei einem idealen Sensor ohne Messunsicherheit könnte man \vec{f} entlang eines Pfades K integrieren. Dazu muss man die Höhe $z(\vec{x}_0)$ an einem Punkt der

Oberfläche kennen.

$$z(\vec{x}) = z(\vec{x}_0) + \int_K \vec{f} d\vec{s} \quad (6.42)$$

Da \vec{f} als der Gradient von z berechnet wurde, ergibt sich für alle Pfade K von \vec{x}_0 nach \vec{x} derselbe Wert des Integrals (\vec{f} ist ein konservatives Vektorfeld). Bei realen Sensoren ist die Situation anders. Aufgrund der Messunsicherheit ist das gemessene Feld

$$\vec{g} = \begin{pmatrix} p \\ q \end{pmatrix} \quad (6.43)$$

im allgemeinen nicht konservativ. Daher können unterschiedliche Pfade K unterschiedliche Ergebnisse liefern. Es kann gezeigt werden, dass ein Vektorfeld konservativ ist auf einem einfach zusammenhängenden Gebiet, genau dann wenn seine Rotation Null ist. Das bedeutet, man muss \vec{g} so modifizieren, dass seine Rotation Null ist. Es ist bekannt, dass die Divergenz eines Vektorfelds \vec{g} dessen Rotation eliminiert.

$$\nabla \vec{g} = \Delta z := \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} \quad (6.44)$$

Das Problem wurde auf eine andere Ebene verlagert. Anstelle einer Ableitung erster Ordnung muss nun eine Ableitung zweiter Ordnung integriert werden. Eine unmittelbare numerische Integration von Δz ist nicht möglich, da nur die Summe der partiellen Ableitungen $\frac{\partial^2 z}{\partial x^2}$ und $\frac{\partial^2 z}{\partial y^2}$ gegeben ist. Stattdessen kann man eine Integration im Frequenzbereich durchführen. Die Fourier-Transformierte von ∇z ist

$$F\{\nabla z\} = F\{z\} \cdot (j2\pi f_x, j2\pi f_y) \quad (6.45)$$

und die Fourier-Transformierte von Δz ist

$$F\{\Delta z\} = -4\pi^2(f_x^2 + f_y^2) \cdot F\{z\} \quad (6.46)$$

Andererseits ist

$$F\{\Delta z\} = F\{\nabla \vec{g}\} \quad (6.47)$$

Gleichsetzen von Gl. 6.46 und 6.47 sowie Auflösen nach $F\{z\}$ liefert

$$F\{z\} = -\frac{1}{4\pi^2(f_x^2 + f_y^2)} \cdot F\{\nabla \vec{g}\} \quad (6.48)$$

In anderen Worten, man muss die Divergenz der Neigungsdaten berechnen, das Ergebnis Fourier-transformieren und mit einer Funktion von Typ $\frac{1}{f^2}$ filtern. Die inverse Fourier-Transformierte liefert die Höhe $z(x, y)$. Das Ergebnis ist dasselbe wie das von Frankot and Chellapa [46], doch ist der Ansatz basierend auf Vektorfeldern ein anderer als der von ihnen vorgeschlagene, der einen quadratischen Fehler minimiert.

6.4 Single Shot Sensor für einfarbige Textur

Auf die Beschreibung eines Sensors für allgemeine Zwecke (Textur ist zugelassen) folgen Überlegungen zu einem Single Shot Sensor. In Kapitel 5.6 wurde bereits dargestellt, dass es einen Single Shot Sensor im engeren Sinne (eine einzige Aufnahme, volle Bandbreite, Textur und Umgebungslicht unbekannt) nicht geben kann. Rückt man aber von diesen sehr harten Forderungen ein wenig ab, so sind Single Shot Sensoren im erweiterten Sinn möglich. Je nach Anwendung kann man eher auf die eine oder die andere Forderung verzichten. An dieser Stelle soll ein Single Shot Sensor mit 3 Aufnahmen (Einschränkung 1) beschrieben werden, der ein Formsignal z_m mit der vollen Bandbreite des Detektors $B_x \cdot B_y$ liefert. Objekttextur wird zugelassen, allerdings darf die Objekttextur nur einfarbig sein (Einschränkung 2). Dies hängt damit zusammen, dass die drei Bilder aus dem Rot- Grün- und Blauanteil eines Farbbildes gewonnen werden sollen. Umgebungslicht wird nicht zugelassen (Einschränkung 3). Die Grundidee bei diesem Sensor ist wie folgt. Eine Farbkamera kann gleichzeitig 3 Bilder aufnehmen, was genau der Anzahl von Aufnahmen beim photometrischen Stereo entspricht. Darüber hinaus hat eine Farbkamera, anders als drei einzelne S/W-Kameras die Eigenschaft, dass alle drei Bilder vom selben Ort aus aufgenommen werden, wie dies für das photometrische Stereoverfahren nötig ist. Die verschiedenen Beleuchtungsrichtungen werden in den drei Grundfarben Rot, Grün und Blau gekennzeichnet. Damit ist es möglich, alle drei Beleuchtungen simultan zu betreiben. In der Farbkamera werden anschließend der Anteil jeder Beleuchtungsrichtung anhand der Farbe separiert. Dadurch hat man ein sequentielles photometrisches Stereoverfahren in ein zeitparalleles umgewandelt. Es ist vorteilhaft, eine 3-Chip-Farbkamera einzusetzen, da bei ihr jeder Farbkanal mit der vollen Bandbreite erfasst werden kann.

Bei 1-Chip-Farbkameras werden die Farbkanäle nur mit reduzierter Bandbreite abgetastet. Beim häufig verwendeten Bayer-Muster sind 50% der Pixel für grünes Licht und jeweils 25% für rotes bzw. blaues Licht empfindlich. Dies würde die Bandbreite $B_x \cdot B_y$ auf 50% bzw. 25% reduzieren. Dazu kommen Schwierigkeiten bei der Interpolation, weil die drei Farbkanäle nicht am selben Ort abgetastet werden. Wir betrachten also hier eine 3-CCD-Kamera. Genau wie das sequentielle photometrische Stereo kann man auch das farbcodierte parallele photometrische Stereo durch drei Gleichungen beschreiben (siehe Gleichung 6.3).

$$\begin{aligned}
 E_R &= \rho_R \cdot \vec{s}_R \cdot \vec{n} \\
 E_G &= \rho_G \cdot \vec{s}_G \cdot \vec{n} \\
 E_B &= \rho_B \cdot \vec{s}_B \cdot \vec{n}
 \end{aligned} \tag{6.49}$$

Anstelle der Indizes 1, 2 und 3 werden hier die Indizes R, G und B für die drei Grundfarben verwendet. Im Unterschied zum sequentiellen Verfahren fällt auf, dass hier statt einem Reflexionskoeffizienten (Textur) die drei Reflexionskoeffizienten ρ_R , ρ_G und ρ_B auftauchen. Drei Gleichungen genügen nicht, um diese drei Reflexionskoeffizienten und den Normalenvektor zu bestimmen. Nur unter der Annahme, dass alle drei Reflexionskoeffizienten für einen festen Punkt der Oberfläche sich proportional zueinander verhalten (bekannter Proportionalitätsfaktor), kann man Normale und Textur bestimmen.

$$\begin{aligned} E_R &= \text{const}_R \cdot \rho \cdot \vec{s}_R \cdot \vec{n} \\ E_G &= \text{const}_G \cdot \rho \cdot \vec{s}_G \cdot \vec{n} \\ E_B &= \text{const}_B \cdot \rho \cdot \vec{s}_B \cdot \vec{n} \end{aligned} \quad (6.50)$$

In der Praxis bedeutet dies, das Objekt darf zwar Textur besitzen, muss aber einfarbig sein in dem Sinn, dass das Verhältnis der Reflexionskoeffizienten überall gleich ist.

Man könnte argumentieren, dass diese Einschränkung beträchtlich ist. Andererseits ist bei vielen industriell hergestellten Objekten die Farbe im voraus bekannt. Mit diesem Wissen kann man dann einen Single Shot Sensor zur Messung der Form angeben, der auf diese Gegebenheiten zugeschnitten ist. Insbesondere sind alle Gegenstände mit metallisch glänzenden Oberfläche und viele Gegenstände aus Kunststoff einfarbig. Wo dies nicht der Fall ist, wird man einen anderen Weg gehen müssen. Beispielsweise ist das räumliche Phasenschieben nach Takeda [38], [39] nicht durch mehrfarbige Textur eingeschränkt. Stattdessen muss man sicherstellen, dass die Objektform z und die Objekttextur t nur ein Drittel der Bandbreite des Detektors beanspruchen.

Zurück zum Single Shot Sensor für einfarbige Textur. Wie bereits erwähnt ist dieses Verfahren sehr gut geeignet für metallisch glänzende und damit einfarbige Oberflächen. Hierbei taucht die Schwierigkeit auf, dass es sich nicht um diffus reflektierende Oberflächen mit Lambert-Charakteristik handelt. Vielmehr beobachtet man je nach Rauigkeit der Oberfläche mehr oder weniger gerichtete Reflexion und Glanz. Auch die Deflektometrie hätte Schwierigkeiten mit solchen Objekten, da sie wiederum für eine exakte optische Abbildung nicht eben genug sind. Damit bewegt man sich in einem Übergangsbereich zwischen den Lambert-Streuern und den optisch glatten Oberflächen, die mit bestehenden Verfahren sehr schwer zu vermessen sind. Aus diesem Grund wird im folgenden Abschnitt ein Verfahren beschrieben, das eine Brücke zwischen dem photometrischen Stereoverfahren und den deflektometrischen Verfahren schlägt.

6.5 Photometrische Deflektometrie

Bei glänzenden Oberflächen sollte die Lichtquelle nicht direkt das Objekt beleuchten, wie dies bei matten Objekten der Fall ist. Insbesondere sind punktförmige Lichtquellen zu vermeiden. Ansonsten erhält man lediglich einen hellen Reflex der Lichtquelle an einigen wenigen Punkten der Oberfläche, während die übrigen Bereiche dunkel bleiben. Die Annahme eines Lambert-Streuers stellt keine geeignete Beschreibung für das Objekt dar. Mit Hilfe eines durchscheinenden Streuschirms zwischen Lichtquelle und Objekt kann man aber dieses neue Problem auf das bisherige zurückführen. Wenn man diesen Streuschirm (Lambert-Streuer) in der Form einer Halbkugel mit dem Objekt im Mittelpunkt wählt, ergeben sich besondere Vorteile. Damit lassen sich glänzende Objekte so einfach vermessen, als ob sie matt wären. (Für eine einfache Darstellung wird angenommen, dass das Objekt sehr viel kleiner als der Radius der Kugel ist). Dieses Verfahren soll im weiteren photometrische Deflektometrie genannt werden.

Die Kamera betrachtet aus der Beobachtungsrichtung \vec{b} über ein spiegelndes Flächenelement mit Normalenvektor \vec{n} die Innenseite der Halbkugel (siehe Abb. 6.3). Die Kugel wird aus drei Richtungen beleuchtet und die Beleuchtungsstärke E_i für jeden Punkt der Kamera aufgenommen. Den Normalenvektor der Halbkugel \vec{m} erhält man nach der Berechnungsvorschrift für streuende Objekte.

$$\vec{m} = \frac{1}{\rho} S^{-1} \vec{E} \quad (6.51)$$

mit dem Reflexionskoeffizienten ρ

$$\rho = \left\| S^{-1} \vec{E} \right\| \quad (6.52)$$

Die Form der Kugel hat eine besondere Eigenschaft. Bei ihr ist die Richtung der Normalen \vec{m} gleichzeitig auch die Richtung des Radius \vec{r} und für kleine Objekte ist diese wiederum die Richtung vom Objekt zur Kugel. Die radiale Richtung \vec{r} , die Beobachtungsrichtung \vec{b} und die Normale des Messobjektes \vec{n} genügen dem Reflexionsgesetz. Daher kann man schreiben

$$\vec{n} = \text{const} \cdot (\vec{m} + \vec{b}) \quad (6.53)$$

wobei \vec{m} , \vec{n} und \vec{b} jeweils Einheitsvektoren sind. Man kann also mit Hilfe der streuenden Kugel glänzende Objekte genauso einfach vermessen, als ob sie eine matte Oberfläche hätten. Für die Neigung der Oberfläche ausgedrückt durch die partiellen Ableitungen p und q erhält man

$$p = \frac{\partial z}{\partial x} = -\frac{n_x}{n_z} \quad (6.54)$$

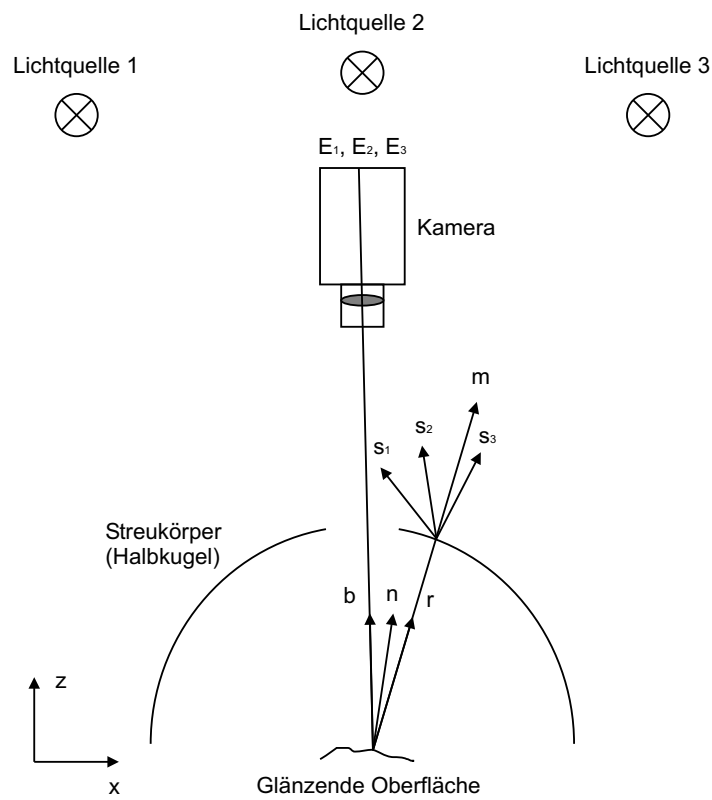


Abbildung 6.3: Funktionsprinzip der photometrischen Deflektometrie an glänzenden Oberflächen

$$q = \frac{\partial z}{\partial y} = -\frac{n_y}{n_z} \quad (6.55)$$

Das Verfahren der photometrischen Deflektometrie lässt sich wie folgt zusammenfassen. Das Vorgehen ist ähnlich wie bei matten Oberflächen, es wird lediglich ein Streukörper in Form einer Halbkugel zwischen Lichtquellen und Objekt gebracht. Der Streukörper wird aus drei Richtungen beleuchtet und jeweils ein Bild des Objektes aufgenommen. Diese Bilder werden verarbeitet, wie es bei matten Oberflächen üblich ist. Man erhält für jedes lokale Flächenstück des Objektes den lokalen Normalenvektor \vec{m} derjenigen Streufläche, welche Licht in die Kamera spiegelt. Mit Hilfe des Reflexionsgesetzes erhält man daraus den lokalen Normalenvektor \vec{n} und die Neigung des Objektes p und q .

Dieses Verfahren kann man einerseits zu den deflektometrischen Verfahren rechnen, da es auf dem Reflexionsgesetz beruht. Zum anderen lehnt es sich stark an das Verfahren des photometrischen Stereo an. Diese Kombination ist vorteilhaft, da die Vorteile des jeweiligen Verfahrens genutzt, die Defizite aber eliminiert werden. Das Defizit anderer deflektometrischer Verfahren ist, dass sie sich nicht auf teilweise spiegelnden, optisch rauen Oberflächen anwenden lassen. Ebensowenig könnte man allein mit photometrischem Stereo eine solche Oberfläche vermessen. Die Kombination ist dazu in der Lage.

Dadurch, dass die Helligkeitsverteilung über der streuenden Kugel keine feinen Strukturen enthält, ist es nicht störend, dass die Oberfläche keine exakte optische Abbildung der Kugel liefert. Vielmehr reflektiert jedes Flächenelement Licht von einem ganzen Areal der Kugel. Hätte man feine Muster auf den Streukörper projiziert, würde der Mittelungseffekt den Kontrast dieser Streifen stark mindern, so dass keine brauchbare Struktur mehr sichtbar ist. Bei Beleuchtung entsprechend dem photometrischen Stereoverfahren ist es zulässig, über große Bereiche des Streukörpers zu mitteln, ohne dass eine wesentliche Kontrastreduktion zwischen den hellen und dunklen Bereichen des Streukörpers eintritt.

Die photometrische Deflektometrie lässt sich ohne Schwierigkeiten in den bereits beschriebenen Single Shot Sensor integrieren. Die Kombination besteht aus einer Beleuchtungseinrichtung in den drei Grundfarben, dem Streukörper und einer Farbkamera. Metallisch spiegelnde Objekte haben die Eigenschaften, dass ihre Textur einfarbig ist. Damit sind die Voraussetzungen für den Single Shot Sensor erfüllt.

Im folgenden werden einige Messergebnisse vorgestellt, die mit dem allgemeinen photometrischen Verfahren, der photometrischen Deflektometrie und dem Single Shot Sensor erzielt wurden.

Kapitel 7

Messergebnisse

Die vorangegangenen Überlegungen für einen optimierten Sensor (siehe Kapitel 6) werden an praktischen Messungen verifiziert. Dazu werden zwei Versuchsaufbauten verwendet, ein makroskopischer Aufbau für ein Messfeld von ca. 200 mm und ein mikroskopischer Aufbau mit einem Feld von wenigen Millimetern. Der Aufbau mit großem Messfeld bietet die Möglichkeit, einen direkten Vergleich zu ziehen zwischen dem vorgeschlagenen Verfahren des photometrischen Stereo und einem alternativen Verfahren, der phasenmessenden Triangulation. Der kombinierte Aufbau ist so gestaltet, dass mit derselben Kamera sowohl das eine als auch das andere Verfahren ausgeführt werden kann. Die Korrespondenz und Vergleichbarkeit der Messungen ist dadurch sichergestellt. Das photometrische Stereoverfahren und die phasenmessende Triangulation werden hinsichtlich ihres Informationswirkungsgrades verglichen. Darüber hinaus werden die beiden dreidimensionalen Datensätze hinsichtlich ihrer Formtreue verglichen.

Der kleinere, mikroskopische Messaufbau wird dazu genutzt, die untere Grenze der Messunsicherheit zu ermitteln, und mit den theoretischen Werten aus Abschnitt 6.2 zu vergleichen. Durch Einfügen eines kugelförmigen Streukörpers kann mit diesem Aufbau auch das Verfahren der photometrischen Deflektometrie ausgeführt werden (glänzende Oberflächen).

Der Aufbau kann mit Hilfe weniger Änderungen zum Single Shot Sensor für einfarbige Textur umgerüstet werden. Auch hierfür wird ein Messbeispiel gezeigt.

7.1 Makroskopisches Messfeld

Der Versuchsaufbau für große Messfelder ist in Abb. 7.1 und 7.2 dargestellt. Um die Kamera (Farbkamera vom Typ Sony XC-003P, 3-CCD, 1/3-Zoll-

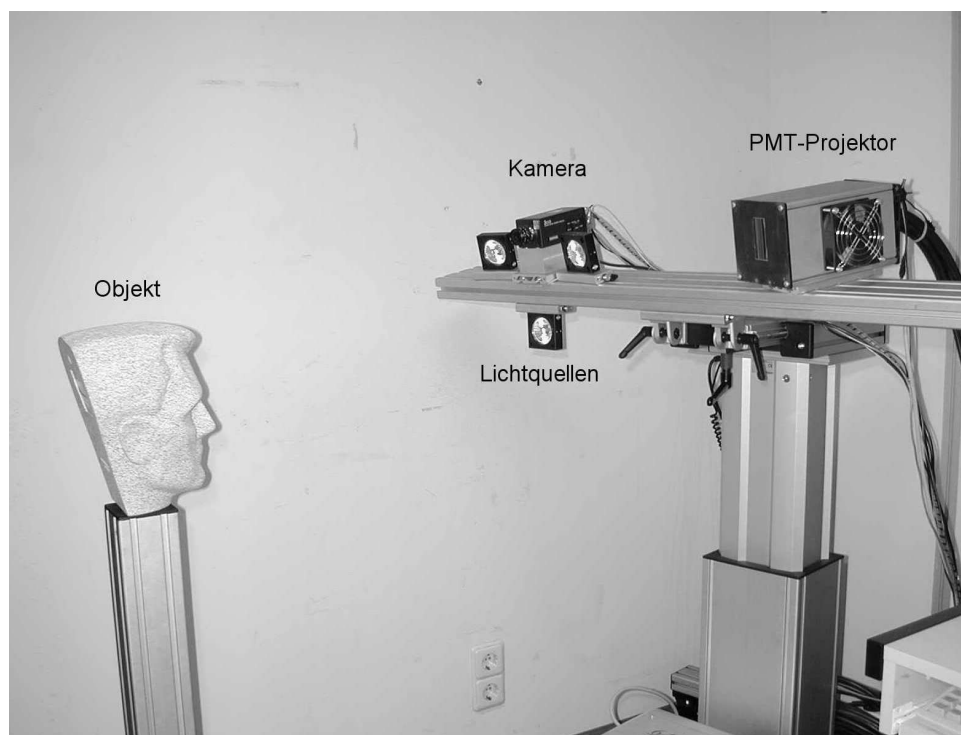


Abbildung 7.1: Kombiniertes Aufbau für photometrisches Stereo und PMT



Abbildung 7.2: Aufbau für photometrisches Stereo



Abbildung 7.3: Kamerabild des Objektes



Abbildung 7.4: Grautoncodierte Neigung der Oberfläche $p = \partial z / \partial x$



Abbildung 7.5: Grautoncodierte Neigung der Oberfläche $q = \partial z / \partial y$

Abbildung 7.6: Grautoncodierte Höhe z der Oberfläche

Chip) sind drei Halogenlampen mit Reflektor angeordnet. Die Lichtquellen befinden sich nahe der optischen Achse der Kamera, um große Neigungen messen zu können (zulässiger Messbereich der Neigung siehe Gleichung 6.16). Das Objekt (im Versuch der Kopf einer Schaufensterpuppe, siehe Abb. 7.1) wird nacheinander von den drei Lichtquellen beleuchtet, wobei die Kamera jeweils ein Bild aufnimmt. Die Bilder werden entsprechend dem photometrischen Stereoalgorithmus ausgewertet. Zunächst erhält man die Neigung der Oberfläche $p = \partial z / \partial x$ (Abb. 7.4) und $q = \partial z / \partial y$ (Abb. 7.5), die anschließend zum Formsignal $z(x, y)$ (Abb. 7.6) aufintegriert wird. Diese Formdaten können mit Hilfe eines Visualisierungsprogramms (3D-Viewer, Fa. 3D-Shape) dargestellt werden. Abb. 7.7 zeigt eine solche gerenderte Darstellung. Alternativ ist es möglich, mit dem Aufbau aus Abb. 7.1 auch phasenmessende Triangulation (PMT) zu betreiben. Hierzu wird der Projektor zusammen mit der Farbkamera benutzt. Eine gerenderte Darstellung der Messung zeigt Abb. 7.8. Der Vergleich von Abb. 7.7 und 7.8 zeigt bereits rein subjektiv ein signifikant höheres Rauschen bei der phasenmessenden Triangulation. Dieser Eindruck bestätigt sich auch später beim quantitativen Vergleich der Messunsicherheiten (siehe Tabelle 7.1). Darüber hinaus ist bemerkenswert, dass es bei der phasenmessenden Triangulation große Bereiche mit ungültigen Werten auf der rechten Seite des Gesichts aufgrund von Abschattungen gibt. Dies rührt daher, dass diese Stellen aufgrund des größeren Triangulations-



Abbildung 7.7: 3D-Ansicht des Objektes gemessen mit photometrischem Stereo



Abbildung 7.8: 3D-Ansicht des Objektes gemessen mit PMT

winkels θ (Winkel zwischen Beleuchtungs- und Beobachtungsrichtung) im Schatten lagen. Beim photometrischen Stereo dagegen ist die gesamte rechte Seite vollständig, einschließlich des rechten Nasenflügels. Dies hängt mit dem kleineren Triangulationswinkel θ im Versuch zusammen. Auf diesem Hintergrund ist es umso erstaunlicher, dass dieses Verfahren dennoch die kleinere Messunsicherheit zeigt. Im folgenden werden die Messungen des photometrischen Stereoverfahrens und der phasenmessenden Triangulation bezüglich ihres Informationswirkungsgrades verglichen.

7.2 Informationswirkungsgrad

Der Informationswirkungsgrad stellt ein Maß dafür dar, wie gut ein Sensor die zur Verfügung gestellte Kanalkapazität ausnutzt (siehe Abschnitt 5.3). Der Wirkungsgrad von photometrischem Stereo und phasenmessender Triangulation wird anhand der bereits gewonnenen Messdaten verglichen. Entsprechend der Festlegung des Informationswirkungsgrades muss dazu das Verhältnis aus realisierter Transinformation und Kanalkapazität gebildet werden.

$$\eta = \frac{I(Z; Z_m)}{C} \quad (7.1)$$

Die realisierte Transinformation wird bestimmt, indem man die Standardabweichung σ_z des Rauschens bestimmt (Messung an einer Referenzebene), die Formdaten $z(x, y)$ quantisiert (Höhe der Stufen gleich der Standardabweichung) und anschließend verlustfrei komprimiert. Die Größe der komprimierten Datei stellt dann eine Schätzung der realisierten Transinformation dar. Die Kapazität C des Kanals ergibt sich aus der Anzahl der Bildelemente $M_x \cdot M_y$, der Anzahl der Aufnahmen M_t und dem Signal-Rausch-Verhältnis S/N der Kamera.

$$C = \frac{1}{2} \cdot M_x \cdot M_y \cdot M_t \cdot \log_2 \left(\frac{S}{N} \right) \quad (7.2)$$

Damit gilt für den Informationswirkungsgrad

$$\eta = \frac{I(Z; Z_m)}{\frac{1}{2} \cdot M_x \cdot M_y \cdot M_t \cdot \log_2 \left(\frac{S}{N} \right)} \quad (7.3)$$

Formt man diese Gleichung geringfügig um, so erhält man eine Darstellung, die in der Praxis besonders vorteilhaft ist.

$$\eta = \frac{I(Z; Z_m)/(M_x \cdot M_y)}{\frac{1}{2} \cdot M_t \cdot \log_2 \left(\frac{S}{N} \right)} \quad (7.4)$$

Im Zähler erscheint die realisierte Transinformation pro Pixel, im Nenner die Kanalkapazität pro Pixel. Die jeweiligen Werte für photometrisches Stereo und phasenmessende Triangulation sind in Tabelle 7.1 dargestellt.

Verfahren	σ_z	Transinf./Pixel	S/N	Kap./Pixel	η
PMT	$114 \mu m$	6,1 bit/Pixel	97^2	$8 \cdot 6,6 \text{ bit/Pixel}$	12%
Phot. Stereo	$44 \mu m$	9,9 bit/Pixel	97^2	$3 \cdot 6,6 \text{ bit/Pixel}$	50%

Tabelle 7.1: Vergleich des Informationswirkungsgrad η für phasenmessende Triangulation und photometrisches Stereo

Die beiden Verfahren unterscheiden sich bezüglich des Rauschens. Bei der phasenmessenden Triangulation ergibt sich im beschriebenen Aufbau eine Messunsicherheit von $\sigma_z = 114 \mu m$, beim photometrischen Stereo liegt sie mit $44 \mu m$ wesentlich darunter. Das geringere Rauschen kommt auch in der realisierten Transinformation von knapp 10 bit , gegenüber rund 6 bit bei der PMT zum Ausdruck. Das Signal-Rausch-Verhältnis des Kanals ist bei beiden Verfahren gleich (97^2), da dieselbe Kamera verwendet wurde. Mit einer Folge von $M_t = 8$ Bildern bei der phasenmessenden Triangulation und nur $M_t = 3$ Bildern beim photometrischen Stereo ergibt sich die Kanalkapazität zu $8 \cdot 6,6 \text{ bit/Pixel}$ bzw. $3 \cdot 6,6 \text{ bit/Pixel}$. Man erhält für den Informationswirkungsgrad η einen Wert von 12% gegenüber 50% beim photometrischen Stereo. Im direkten Vergleich liefert das photometrische Stereo also rund viermal mehr Information, wenn dieselbe Kanalkapazität zur Verfügung gestellt wird. Dies stellt einen großen Fortschritt dar, dennoch ist die Frage berechtigt, warum anstatt 50% Wirkungsgrad nicht ein Wert von nahe 100% beim photometrischen Stereo erreicht wird. Diese Frage führt hin zum erweiterten Informationswirkungsgrad.

7.3 Erweiterter Informationswirkungsgrad

In Abschnitt 5.4 wurde bereits gezeigt, dass sich in der optischen 3D-Messtechnik Form, Textur und Umgebungslicht nicht bereits vor dem Kanal, sondern erst bei der Decodierung trennen lassen. Daher ist es sinnvoll, bei der realisierten Transinformation nicht nur das Formsignal $z(x, y)$, sondern auch die Textur $t(x, y)$ und gegebenenfalls das Umgebungslicht $u(x, y)$ zu berücksichtigen (siehe Gl. 5.57).

$$\eta_{ext} = \frac{I(Z; Z_m) + I(T; T_m) + I(U; U_m)}{C} \quad (7.5)$$

Setzt man die Kanalkapazität C ein, so erhält man

$$\eta_{ext} = \frac{I(Z; Z_m) + I(T; T_m) + I(U; U_m)}{\frac{1}{2} \cdot M_x \cdot M_y \cdot M_t \cdot \log_2 \left(\frac{S}{N} \right)} \quad (7.6)$$

Für die Praxis bezieht man die realisierte Transinformation und Kanalkapazität wieder auf die Zahl der Pixel.

$$\eta_{ext} = \frac{(I(z; z_m) + I(t; t_m) + I(u; u_m)) / (M_x \cdot M_y)}{\frac{1}{2} \cdot M_t \cdot \log_2 \left(\frac{S}{N} \right)} \quad (7.7)$$

Für die vorliegende Messung (selber Aufbau, selbes Objekt, siehe Abb. 7.3) wird die Form und die Textur, nicht aber das Umgebungslicht einbezogen, da in einem abgedunkelten Raum gemessen wurde. Tabelle 7.2 zeigt einen Vergleich der beiden Messverfahren in Bezug auf diesen erweiterten Wirkungsgrad.

Verfahren	Transinf. Form	Transinf. Textur	Kap./Pixel	η_{ext}
PMT	6,1 bit/Pixel	4,8 bit/Pixel	8 · 6,6 bit/Pixel	21%
Phot. Stereo	9,9 bit/Pixel	4,8 bit/Pixel	3 · 6,6 bit/Pixel	74%

Tabelle 7.2: Vergleich des Informationswirkungsgrad η für phasenmessende Triangulation und photometrisches Stereo

An der Transinformation der Form ändert sich nichts im Vergleich zum einfachen Informationswirkungsgrad. Hinzu kommt allerdings die Transinformation des Texturbildes, bei der phasenmessenden Triangulation und beim photometrischen Stereo jeweils 4,8 bit/Pixel (siehe Abb. 7.3)). Der erweiterte Informationswirkungsgrad ergibt sich demnach zu 21% bzw. 74%. Mit dem Verfahren des photometrischen Stereo lassen sich rund drei Viertel der zur Verfügung gestellten Kanalkapazität zur Messung der Form und der Textur nutzen, während ein Viertel ungenutzt bleibt. Dies liegt daran, dass die Codierung durch Differenzieren einen großen Teil der Redundanz, jedoch nicht sämtliche Redundanz entfernen kann. In Anbetracht der Tatsache aber, dass es derzeit nur ein einziges analoges Codierungsverfahren gibt, ist es geradezu ein Glücksfall, dass überhaupt die Redundanz wesentlich reduziert werden kann. So lässt sich auch die geringe noch ungenutzte Kapazität akzeptieren.

7.4 Formtreue

Das Verfahren des photometrischen Stereo ist in der Lage, lokale Höhenunterschiede mit geringer Messunsicherheit zu detektieren. Globale Höhen-

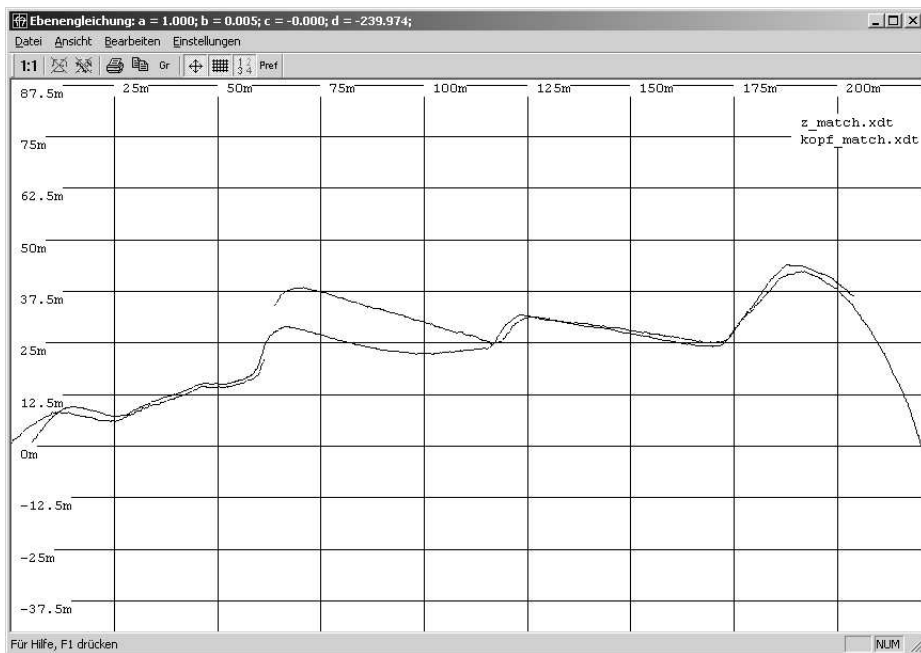


Abbildung 7.9: Vertikaler Schnitt durch photometrische Messung und PMT-Messung entlang der Nase

unterschiede auf dem Objekt sind mit einer größeren Messunsicherheit behaftet. Es stellt sich die Frage, wie groß diese Abweichungen in der Praxis sind. Zu diesem Zweck werden die photometrische Messung und die PMT-Messung mit Hilfe eines Softwareprogrammes (Slim^{3D}, Fa. 3D-Shape) registriert und verglichen. Abbildung 7.9 zeigt einen senkrechten Schnitt. Die Messungen zeigen gute Übereinstimmung mit Ausnahme der Nase. Dies rührt daher, dass beim photometrischen Stereo ein differenzierbares Formsignal $z(x, y)$ vorausgesetzt wird, an der Nasenspitze aber ein Sprung vorliegt.

Abhilfe könnte die Kombination des photometrischen Stereoverfahrens mit einem binokularen Stereoverfahren bringen, das in der Lage ist, solche Höhenunterschiede korrekt wiederzugeben. Darüber hinaus kann man speziell zur Vermessung von Gesichtern die Beleuchtungs- und Beobachtungsrichtungen so wählen, dass die Nasenpartie leicht von unten beleuchtet und beobachtet wird. Damit wird aus der Stufe eine mehr oder weniger steile Fläche.

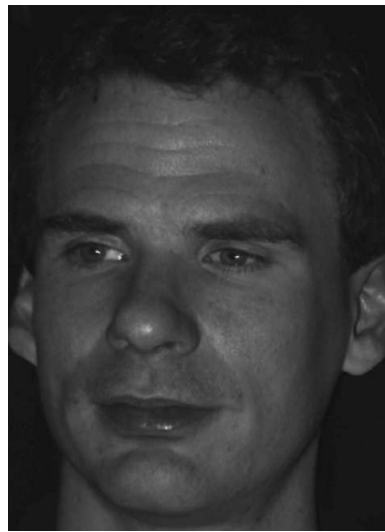


Abbildung 7.10: 3D-Messung Gesicht, Ansicht 1

7.5 Messung mit Textur

Beim erweiterten Informationswirkungsgrad wurde festgestellt, dass neben der Form auch die Textur eines Objektes Information enthält. Im folgenden wird ein 3D-Messung eines Gesichtes einschließlich Textur gezeigt. Abbildung 7.10 und 7.11 zeigt das Gesicht aus zwei verschiedenen Blickwinkeln, wie sie anhand der 3D-Daten und der Textur vom Viewer rekonstruiert wurde. Es ist anzumerken, dass das Gesicht nicht aus den gezeigten Blickwinkeln aufgenommen ist und diese Ansichten virtuell am Rechner entstanden sind. Erst die Kenntnis von 3D-Form und Textur macht solche virtuellen Ansichten möglich.

7.6 Mikroskopisches Messfeld

Mit dem photometrischen Stereoverfahren lassen sich makroskopische Messfelder (Beispiel Gesicht), genauso gut aber auch mikroskopisch kleine Messfelder aufnehmen. Abbildung 7.12 zeigt einen solchen Aufbau. Er besteht aus einem Mikroskop mit einer CCD-Kamera (Sony XC-ST50 CE, 1/2-Zoll-Chip, Schwarz-Weiß) und der Beleuchtung. Das Messfeld hat je nach gewählter Vergrößerung eine laterale Ausdehnung von ca. 4 mm bis zu 25 mm. Besonders bei diesen kleinen Messfeldern kann kohärentes Specklerauschen eine bedeutende Rolle spielen. Es wurde daher auf räumlich „inkohärente“ Beleuchtung mit einer großen Apertur geachtet. Zu diesem Zweck wurde ein halbku-



Abbildung 7.11: 3D-Messung Gesicht, Ansicht 2



Abbildung 7.12: Mikroskopischer Aufbau für photometrisches Stereo

gelförmiger Streukörper zwischen Lichtquelle und Objekt eingefügt. Mit dieser Anordnung können sowohl diffus streuende als auch glänzende Oberflächen (Metalle, Glas etc.) vermessen werden. Zu diesem Zweck kommt das Verfahren der photometrischen Deflektometrie zum Einsatz (siehe Kapitel 6.5). Abbildung 7.13 zeigt Ergebnisse einer Messung an einer Euro-Münze mit diesem Verfahren. Dargestellt ist eine gerenderte 3D-Ansicht, wie sie mit Hilfe eines 3D-Viewers aus den Daten erzeugt werden kann. Es fällt auf, dass die Messung sehr geringes Rauschen zeigt. Zu bemerken ist auch, dass für diese Darstellung keinerlei Glättungsfilter eingesetzt wurden. Abbildung 7.14 zeigt eine weitere Messung an der Münze, diesmal mit dem Single Shot Sensor. Die S/W CCD-Kamera ist hierfür durch eine Farbkamera zu ersetzen. Die Beleuchtungsquellen sind mit Hilfe von Farbfiltern in den Farben Rot, Grün und Blau kodiert. Diese Messung macht es möglich, mit nur einer einzigen Aufnahme (1 Farbaufnahme, 3 Farbkanäle) gleichzeitig alle Daten aufzunehmen, die für eine 3D-Messung nötig sind (siehe Kapitel 6.4). Leider zeigt diese Messung ein höheres Rauschen als die Messung mit der S/W-Kamera. Dies liegt hauptsächlich daran, dass der verwendete Farbsensor eine geringere Full-Well-Capacity und damit ein höheres Photonenrauschen aufweist. Darüber hinaus ist der Abgleich der Helligkeit der drei Lichtquellen mit den Kanälen der Farbkamera problematischer als bei der S/W-Kamera. Erschwerend kommt bei der verwendeten 3-Chip-Kamera hinzu, dass sich die Farbempfindlichkeit aufgrund des dichroitischen Teilerspiegels über dem Bildfeld nicht konstant ist. Aus diesem Grund beobachtet man fälschlicherweise eine Durchbiegung der Münzenoberfläche. Abhilfe könnte eine Kalibrierung schaffen, die diese Eigenschaften der Kamera mit einbezieht. Zusammenfassend kann man sagen, dass der Sensor mit S/W-Kamera vorzugsweise für zulässige Aufnahmezeiten über 100 ms eingesetzt wird. Sind sehr viel geringere Aufnahmezeiten gefordert, so ist dem Single Shot Sensor der Vorzug zu geben.

7.7 Messunsicherheit

In Kapitel 6.2.1 ist die minimale Messunsicherheit angegeben, die man von einem photometrischen Stereosensor erwarten kann. Die Messunsicherheit zwischen benachbarten Pixeln beträgt nach Gleichung 6.32

$$\sigma_{\Delta z} = \frac{\cot \theta}{\sqrt{2}} \cdot \frac{\lambda}{\max(\sin u_{ill}, \sin u_{obs})} \cdot c_{temp} \cdot c_{pol} \quad (7.8)$$

Um die Messunsicherheit im Experiment zu bestimmen, wird eine ebene, sehr feinkörnige goldbedampfte Mattscheibe vermessen (Messfeld 3,9 mm).

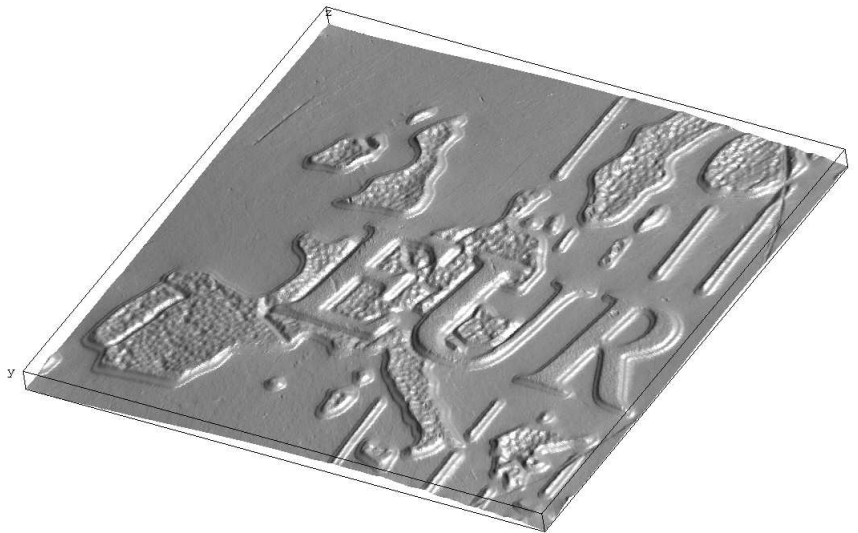


Abbildung 7.13: 3D-Messung von Münze, Sequenz von Aufnahmen

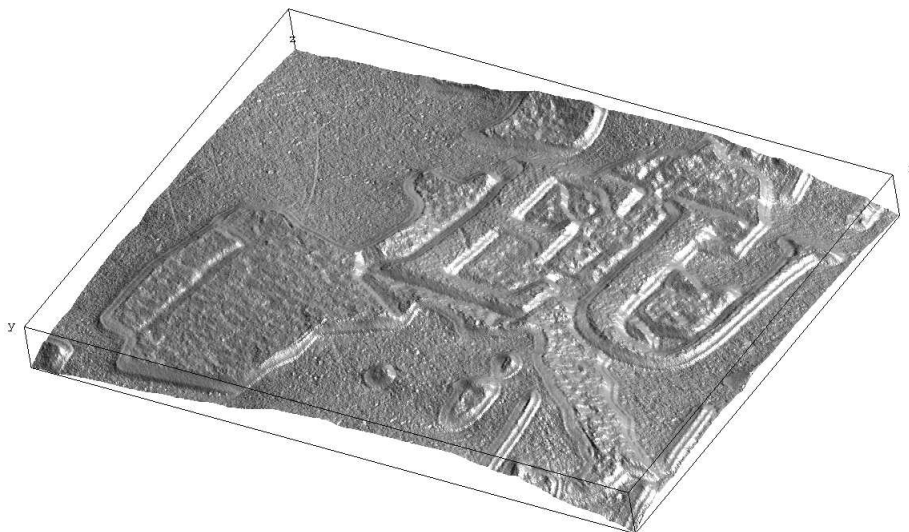


Abbildung 7.14: 3D-Messung von Münze, Single Shot

Die Parameter im Experiment sind wie folgt gewählt. Die Beleuchtungsapertur $\sin u_{ill}$ erreicht mit Hilfe eines halbkugelförmigen Streukörpers Werte von $\sin u_{ill} = 0,7$ und ist damit größer als die Beobachtungsapertur ($\sin u_{obs} = 0,087$). Der Beleuchtungswinkel θ beträgt 47° . Die mittlere Wellenlänge der Beleuchtung mit weißem Licht ist ca. $\lambda = 600nm$. Die Streuscheibe erhält die Polarisation (metallische Oberfläche), deshalb wird der Beitrag der Polarisation mit $c_{pol} = 1$ angesetzt. Der Beitrag der zeitlichen Kohärenz lässt sich bestimmen zu (siehe Gl. 3.13)

$$c_{temp} = \left(1 + \left(\frac{4\sigma_o}{l_c} \right)^2 \right)^{-1/4} \quad (7.9)$$

Hierbei ist σ_o die Rauigkeit der Oberfläche und l_c die Kohärenzlänge des Lichts. Anschaulich ist klar, dass der Specklekontrast abnimmt, wenn die Gangunterschiede an der Oberfläche aufgrund der Rauigkeit zunehmen und die Kohärenzlänge übersteigen. Ein anderer Grund für Gangunterschiede an der Oberfläche sind nach Lampalzer die Beleuchtung fernab der spekularen Reflexionsrichtung [47]. Im vorliegenden Experiment sind diese Gangunterschiede, die man als Keilhöhe bezeichnen kann [47], größer als die Rauigkeit der Oberfläche. Mit dem Beleuchtungswinkel $\theta = 47^\circ$ und einer Auflösungszelle von

$$\delta x = \frac{\lambda}{\sin u_{obs}} = \frac{0,6\mu m}{0,087} = 6,9\mu m \quad (7.10)$$

liegt die Keilhöhe bei (siehe Gl. 3.14)

$$\Delta s = \delta x \cdot \sin \theta = 5,0\mu m \quad (7.11)$$

Die Kohärenzlänge des Glühlampenlichts (sichtbarer Bereich) wird zu

$$l_c = \frac{\lambda^2}{\Delta\lambda} = \frac{(0,6\mu m)^2}{0,3\mu m} = 1,2\mu m \quad (7.12)$$

abgeschätzt. Damit ergibt sich für den Beitrag der zeitlichen Kohärenz

$$c_{temp} = \left(1 + \left(\frac{4 \cdot 5,0\mu m}{1,2\mu m} \right)^2 \right)^{-1/4} = 0,24 \quad (7.13)$$

und man erwartet eine Messunsicherheit von

$$\sigma_{\Delta z} = \frac{\cot 47^\circ}{\sqrt{2}} \cdot \frac{600nm}{0,7} \cdot 0,24 \cdot 1 = 136nm \quad (7.14)$$

Das rechnerische Ergebnis für die Messunsicherheit wurde anhand eines Ex-

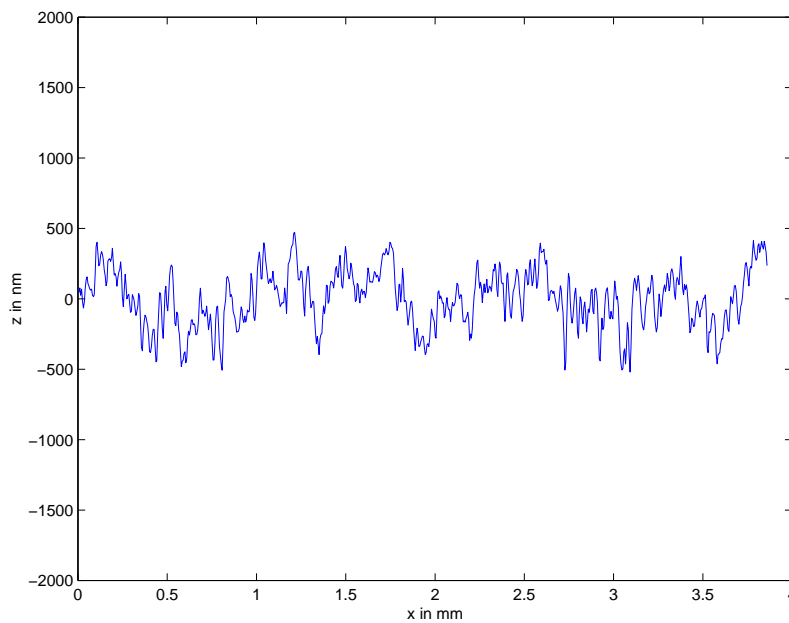


Abbildung 7.15: Schnitt durch goldbedampfte Streuscheibe

perimentes verifiziert. Abbildung 7.15 zeigt einen Schnitt der Streuscheibe. Der gemessene Wert der Messunsicherheit zwischen benachbarten Pixeln liegt bei 103 nm. Die Abweichung vom theoretischen Wert lässt sich damit erklären, dass die Überlegungen zur zeitlichen Kohärenz nicht als exakte Berechnung, sondern eher als Abschätzung zu verstehen sind. In jedem Fall ist deutlich geworden, dass es sich beim photometrischen Stereo um ein Verfahren handelt, dessen Messunsicherheit für mikroskopische Felder und geeignete Beleuchtung auf Bruchteile von Mikrometern reduziert werden kann.

Kapitel 8

Zusammenfassung und Ausblick

In der Einleitung wurde die Frage gestellt, wie viel zweidimensionale Rohdaten einer Kamera benötigt werden, um die Form von Objektoberflächen im Dreidimensionalen zu ermitteln. Dies ist eine Abwägung zwischen Aufwand (2D-Rohdaten) und Nutzen (3D-Daten). Die Informationstheorie stellt ein geeignetes Werkzeug dar, diese Frage zu beantworten. Der Ablauf einer optischen 3D-Messung wird als ein Kommunikationssystem aufgefasst, bestehend aus Quelle, Codierung, Kanal, Decodierung und Senke. Der Kanal begrenzt die Menge an Information, die ein Sensor maximal aufnehmen kann (Kanalkapazität) und ist ein Maß für den Aufwand. Dem gegenüber steht der Nutzen, also die Information, die tatsächlich übertragen wird. Der Informationswirkungsgrad wird als das Verhältnis von Nutzen zu Aufwand definiert. Mit seiner Hilfe lassen sich verschiedenste optische 3D-Sensoren miteinander vergleichen. Es stellt sich heraus, dass gängige Sensoren einen Wirkungsgrad von weniger als 12%, einige sogar weniger als 1% aufweisen. Grund dafür ist eine ungünstige Codierung der Form. Zum einen besitzt bereits die Objektform an sich Redundanz, zum anderen aber fügen diese Verfahren selbst zusätzlich Redundanz hinzu. Eine geeignete analoge optische Codierung ist in der Lage, diese Redundanz zu mindern. Verfahren, die eine optische Differenzierung des Formsignals bewirken, zeigen sich als gut geeignet, insbesondere auch das photometrischen Stereoverfahren für diffuse Oberflächen. Mit diesem an sich bekannten Verfahren wird ein Wirkungsgrad von etwa 50% realisiert. Mit lediglich drei Bildern ist es möglich, Formdaten mit geringer Messunsicherheit zu gewinnen.

Das Verfahren wird um mehrere Aspekte erweitert. Durch eine hohe Beleuchtungsapertur gelingt es, die lokale Messunsicherheit auf weniger als $1\mu\text{m}$ zu reduzieren. Mit Hilfe eines abgeleiteten Verfahrens, der photometrischen Deflektometrie, gelingt es, auch glänzende, insbesondere metallische Oberflächen zu vermessen. Eine weitere Modifikation macht den Sensor zum Single

Shot Sensor, der alle Bilddaten gleichzeitig aufnehmen kann. Die Beleuchtung des Sensors wird farblich codiert und ein einzelnes Bild einer Farbkamera liefert mit seinen drei Farbkanälen alle nötigen Rohdaten.

Für eine zukünftige Weiterentwicklung des Verfahrens scheint es aussichtsreich, ein binokulares Stereoverfahren zu integrieren. Damit können Schwächen des Verfahrens bei der globalen Messunsicherheit überwunden werden.

Literaturverzeichnis

- [1] S. Gautama, S. Lacroix und M. Devey, „On the performance of stereo matching algorithms“, in *Vision Modeling and Visualization '99*, B. Girod, H. Niemann, H. Seidel (Eds.), 59–66 (1999)
- [2] V. Srinivasan, H. Liu und M. Halioua, „Automated phase-measuring profilometry of 3-D diffuse objects“, *Appl. Opt.* **23**, 3105 (1984)
- [3] R. Dorsch, G. Häusler und J. Herrmann, „Laser triangulation: Fundamental uncertainty in distance measurement“, *Appl. Opt.* **33**, 1306–1314 (1994)
- [4] T. Dresel, G. Häusler und H. Venzke, „3-D-sensing of rough surfaces by 'coherence radar'“, *Appl. Opt.* **31**, 919–925 (1992)
- [5] G. Häusler und J. Hermann, „Physical limits of 3D-sensing“, in *Optics, Illumination, and Image Sensing for Machine Vision VII*, D. Svetkoff, ed., Proc. SPIE 1822, 150–158 (1992)
- [6] G. Häusler, „About the scaling behavior of optical range sensors“, Proc. of the 3rd Int. Workshop on Automatic Processing of Fringe Patterns, Bremen (1997)
- [7] G. Häusler, P. Ettl, M. Schenk, G. Bohn und I. Laszlo, „Limits of Optical Range Sensors and How to Exploit Them“, in *Trends in Optics and Phototonics, Ico IV*, T. Asakura, ed., (Springer Verlag Berlin, Heidelberg, New York, 1999), Vol. 74, 328–342
- [8] C. Shannon und W. Weaver, *The mathematical theory of communication* (University of Illinois Press, 1949)
- [9] R. Fano, *Transmission of information* (1961)
- [10] J. Wozencraft und I. Jacobs, *Principles of communication engineering* (John Wiley & Sons, New York, 1965)

- [11] R. Gallager, *Information theory and reliable communications* (John Wiley & Sons, New York, 1968)
- [12] A. Viterbi und J. Omura, *Principles of digital communication and coding* (McGraw Hill, New York, 1979)
- [13] R. Blahut, *Principles and practice of information theory* (Addison-Wesley, Reading Mass., 1987)
- [14] T. Cover und J. Thomas, *Elements of information theory* (John Wiley & Sons, New York, 1991)
- [15] R. Johannesson, *Informationstheorie - Grundlagen der (Tele-)Kommunikation* (Addison-Wesley Studentlitteratur, 1992)
- [16] J. Huber, *Information theory and its applications in communications engineering*, Vorlesungsskript FAU Erlangen-Nürnberg
- [17] M. Born, E. Wolf *Principles of optics* (Pergamon Press, 1987)
- [18] J. Goodman, *Statistical optics* (John Wiley & Sons, New York, 1985), 157–187
- [19] J. Goodman, *Introduction to Fourier optics* (McGraw-Hill, New York, 1996)
- [20] A. Papoulis, *Probability, random variables and stochastic processes* (McGraw-Hill, New York, 1984)
- [21] J. Dainty, *Laser Speckle and related phenomena* (Springer-Verlag, Berlin, 1984)
- [22] J. Goodman, *Statistical optics* (John Wiley & Sons, New York, 1985), 347–356
- [23] J. Goodman, *Statistical optics* (John Wiley & Sons, New York, 1985), 350
- [24] R. Lampalzer, „Physikalische Grenzen von optischen 3D-Sensoren mit phasenmessender Triangulation“, Dissertation FAU Erlangen-Nürnberg 2002, 49
- [25] H. Pedersen, „On the contrast of polychromatic speckle patterns and its dependence on the roughness“, *Opt. Acta*, **22**, no. 1 (1975), 15–24

- [26] B. Frieden, *Probability, statistical optics, and data testing* (Springer-Verlag, Berlin, Heidelberg, New-York, 2001)
- [27] R. Röhler, *Informationstheorie in der Optik* (Wissenschaftliche Verlagsgesellschaft, Stuttgart, 1967)
- [28] M. Neifeld, „Information, resolution, and space-bandwidth product“, *Optics Letters* **23**, 1477–1479 (1998)
- [29] R. Johannesson, *Informationstheorie - Grundlagen der (Tele-)Kommunikation* (Addison-Wesley Studentlitteratur, 1992), 11
- [30] B. Frieden, *Probability, statistical optics, and data testing* (Springer-Verlag, Berlin, Heidelberg, New-York, 2001), 18-20
- [31] J. Goodman, *Statistical optics* (John Wiley & Sons, New York, 1985), 465–524
- [32] R. Johannesson, *Informationstheorie - Grundlagen der (Tele-)Kommunikation* (Addison-Wesley Studentlitteratur, 1992), 235–238
- [33] R. Johannesson, *Informationstheorie - Grundlagen der (Tele-)Kommunikation* (Addison-Wesley Studentlitteratur, 1992), 231–232
- [34] J. Huber, *Information theory and its applications in communications engineering*, Vorlesungsskript FAU Erlangen-Nürnberg, 305
- [35] G. Häusler, „About the scaling behavior of optical range sensors“, Proc. of the 3rd Int. Workshop on Automatic Processing of Fringe Patterns, Bremen (1997)
- [36] R. Röhler, *Informationstheorie in der Optik* (Wissenschaftliche Verlagsgesellschaft, Stuttgart, 1967), 60–161
- [37] R. Röhler, *Informationstheorie in der Optik* (Wissenschaftliche Verlagsgesellschaft, Stuttgart, 1967), 166–167
- [38] M. Takeda, H. Ina, S. Kobayashi „Fourier transform method of fringe pattern analysis for computer-based topography and interferometry“, *JOSA* **72**, 1982, 156–160
- [39] M. Takeda, K. Mutoh „Fourier transform profilometry for the automatic measurement of 3-D object shapes“, *Appl. Opt.* **22**, No. 24, 1983, 3977–3982

- [40] B. Horn, „Obtaining shape from shading information“, in *Shape from shading*, B. Horn und M. Brooks, eds., (MIT Press, Cambridge, Massachusetts, London, 1989), 123–171
- [41] G. Häusler und D. Ritter, „Parallel three-dimensional sensing by color coded triangulation“, *Appl. Opt.* **32**, 7164–7169 (1993).
- [42] B. Horn und M. Brooks, *Shape from shading* (MIT Press, Cambridge, Massachusetts, London, 1989)
- [43] G. Häusler, „Verfahren und Vorrichtung zur Ermittlung der Form oder der Abbildungseigenschaften von spiegelnden oder transparenten Objekten“, Deutsches Patent DE 199 44 354 (1999)
- [44] R. Woodham, „Photometric method for determining surface orientation from mutiple images“, in *Shape from shading*, B. Horn und M. Brooks, eds., (MIT Press, Cambridge, Massachusetts, London, 1989), 513–532
- [45] R. Lampalzer, „Physikalische Grenzen von optischen 3D-Sensoren mit phasenmessender Triangulation“, Dissertation FAU Erlangen-Nürnberg (2002), 50
- [46] R. Frankot und R. Chellappa, „A method for enforcing integrability in shape from shading algorithms“, in *Shape from shading*, B. Horn und M. Brooks, eds., (MIT Press, Cambridge, Massachusetts, London, 1989), 89–122
- [47] R. Lampalzer, „Physikalische Grenzen von optischen 3D-Sensoren mit phasenmessender Triangulation“, Dissertation FAU Erlangen-Nürnberg (2002), 47–48

Danksagung

Zuallererst möchte ich Herrn Prof. Dr. Steffen für die Betreuung dieser Arbeit von Seiten der elektrotechnischen Fakultät danken.

Bei Herrn Prof. Dr. Leuchs bedanke ich mich für die freundliche Aufnahme am Lehrstuhl für Optik.

Mein besonderer Dank gilt Herrn Prof. Dr. Gerd Häusler für die fachkundige und interessierte Betreuung der Arbeit seitens der Physik, die zahlreichen kreativen Diskussionen und Anregungen, vor allem aber für die gute und kollegiale Arbeitsatmosphäre in seiner Arbeitsgruppe.

Bei allen meinen Kollegen und Kolleginnen möchte ich mich bedanken für die vielen fachlichen, aber auch erheiternden Diskussionen, die tiefgründigen Mensagespräche und netten Angriffe auf das Zwerchfell, die ich sehr vermisse.

Mein Dank gilt auch unserer Sekretärin Frau Dollinger, die bei allen noch so komplizierten Verwaltungsfragen Rat wusste.

Zuletzt und ganz besonders möchte ich meiner Frau Audry danken, die mich wie niemand sonst dazu ermuntert hat, Ideen zu entwickeln und umzusetzen.

Lebenslauf

Familienstand

Name: Christoph Wagner
Geburtsdatum: 05. Juli 1971
Geburtsort: Pforzheim
Familienstand: verheiratet

Schulausbildung

Sept. 1978 – Aug. 1982 Grundschule Ispringen
Sept. 1982 – Juni 1991 Bildungszentrum Königsbach, Abitur

Praktika

Sept. 1991 – Apr. 1992 Gemeindepraktikum bei Operation
Mobilisation, Paris
Mai 1992 – Jul. 1992 Grundlagenpraktikum Elektrotechnik
Apr. 1996 – Jul. 1996 Fachpraktikum Elektrotechnik
bei Firma Polytec, Waldbronn

Hochschulausbildung

Okt. 1992 – Sept. 1998 Studium der Elektrotechnik an der
Universität Karlsruhe, Diplom
Apr. 1998 – Sept. 1998 Diplomarbeit am Bremer Institut
für angewandte Strahltechnik (BIAS)
Okt. 1998 – Apr. 1999 Freier Mitarbeiter am Bremer Institut
für angewandte Strahltechnik
seit
Mai 1999 Wissenschaftlicher Angestellter am
Lehrstuhl für Optik des Physikalischen
Instituts der Universität Erlangen-Nürnberg,
Promotion