

3D-Aufnahmetechnik und Kameras

Produktionen für 3D-Kino und 3D-Fernsehen stellen Kamera- und Aufnahmetechnik vor neue Herausforderungen. Neben der klassischen Stereoskopie und deren Erweiterung durch Multi-Kamera-Arrays werden auch neue Technologien wie Time-of-Flight und Plenoptic beschrieben. Die heutigen Einsatzmöglichkeiten in der Film- und TV-Produktion sowie Perspektiven für die Zukunft werden erläutert.

3D cinema and 3D TV production create new challenges in terms of camera and acquisition techniques. This article describes classic stereoscopy and its enhancement by multi-camera arrays, as well as new technologies such as time-of-flight and plenoptics. The applications of these ideas in film and TV production as well as future perspectives will be explained.

Einleitung

Bereits mit Filmkameras konnten seit fast 100 Jahren sowohl Einzelbilder als auch bewegte Bilder stereoskopisch erfasst und mit einer geeigneten Darstellung für den Menschen ein 3D-Eindruck erzeugt werden. Die Möglichkeiten der Nachbearbeitung von Filmmaterial im Sinne einer 3D-Produktion waren allerdings sehr eingeschränkt und der apparative Aufwand für einen 3D-Film war so groß, dass nur wenige 3D-Filme entstanden.

Erst mit der vollständigen Digitalisierung der gesamten Verarbeitungskette von Aufnahme über Postproduktion und Projektion erlebt

die 3D-Technik heute eine Renaissance. Mit der Einführung der digitalen Projektoren ergaben sich qualitativ brauchbare Lösungen für die Wiedergabe. Die ersten 3D-Filme der neuen digitalen Generation waren allerdings meist komplett animierte Trickfilme ohne reale Bilder, da 3D-Postproduktions- und Grafiksoftware dafür in hoher Qualität verfügbar ist.

Die 3D-Aufnahme von „real world“-Szenen mit 3D-Kameratechnik gestaltet sich dagegen aufwendiger und schwieriger und ist heute noch Gegenstand von Forschung und Entwicklung. Je nach Anforderung an die Weiterverarbeitung und Darstellung muss eine 3D-Aufnahmetechnik

- zwei komplette Ansichten einer Szene für eine Stereoskopie
 - oder neue und mehrere Ansichten für Displays
 - oder ein 2D-Bild mit einer zugehörigen Tiefeninformation für jeden Bildpunkt
 - oder ein 3D-Modell von Objekten liefern.
- Technikentwickler, Kameralente und Produzenten experimentieren heute mit verschiedenen Kameras und Anordnungen. Eine allgemein etablierte Technik – die alle Anforderungen erfüllt – gibt es noch nicht. Im Folgenden werden einige Verfahren zur Erfassung von 3D-Bildinhalten vorgestellt:

- Triangulation mit stereoskopischen Kameras,
- Multi-Kamera-Arrays als Erweiterung der Stereoskopie,
- Time-of-Flight-Prinzip zur direkten Messung der z-Komponente und
- Plenoptic, eine neue Technik mit Mikrolinsen-Arrays vor dem Sensor.

Triangulation

Triangulation ist ein aus der Messtechnik bekanntes Verfahren, mit dem durch Beobachtung eines Punktes aus verschiedenen Rich-

tungen dessen Abstand vom Beobachter bestimmt werden kann. Wird eine Szene mit einem stereoskopischen System (zwei Kameras) aufgenommen, besteht somit die prinzipielle Möglichkeit, eine Tiefenkarte zu berechnen. Für eine solche Berechnung wird ein Stereosystem typischerweise in einem Vorverarbeitungsschritt rektifiziert, das heißt in ein ideales achsparalleles System umgerechnet. Die Rektifizierung vereinfacht den anschließenden Schritt der Korrespondenzfindung (Disparitätsschätzung) erheblich, da sich das Pendant eines bestimmten Punktes in einer Kamera in exakt der gleichen Bildzeile der anderen Kamera befindet. Die gefundenen Korrespondenzen werden in Form von Disparitätskarten beschrieben, die jedem Bildpunkt einen Vektor zuordnen, der auf den entsprechenden Punkt im anderen Bild zeigt. Für rektifizierte Bildpaare reduziert sich der Disparitätsvektor auf einen Skalar.

Für ein rektifiziertes Stereosystem kann die Tiefe „Z“ eines bestimmten Bildpunktes mittels der Formel $Z = B \cdot F/D$ berechnet werden, wobei „B“ die Basisweite der Kameras, „F“ die fokale Länge und „D“ die mittels Korrespondenzanalyse gefundene Disparität beschreibt. Da die Bilder aus diskreten Bildpunkten bestehen, ergeben sich bei der Umrechnung in Tiefe diskrete Tiefenebenen. Der Abstand „ΔZ“ zwischen zwei Tiefenebenen ist insbesondere auch von der Tiefe abhängig und berechnet sich aus der Formel

$$\Delta Z = Z \cdot \Delta u / \left(\frac{B \cdot F}{Z} \pm \Delta u \right),$$

wobei „Δu“ die Genauigkeit der Disparitätsschätzung beschreibt. Bei einer Genauigkeit von einem Bildpunkt, einer Basisweite von 20 cm und einer fokalen Länge von 2200 Bildpunkten ergeben sich die in **Tabelle 1** dargestellten exemplarischen Abstände zwischen den gefundenen Tiefenebenen.

Eine Möglichkeit zur Korrespondenzfindung ist das am Fraunhofer-Institut für Nachrichtentechnik (HHI) entwickelte Verfahren HRM (Hybrid Recursive Matching) [1].

Das Verfahren kombiniert Blockrekursion mit einer pixelrekursiven Schätzung des optischen Flusses. Zwei örtliche und ein zeitlicher Kandidat werden für den blockrekursiven Schritt getestet und zum Beispiel mittels Kreuzkorrelation mit den entsprechenden Blöcken des anderen Bildes verglichen. Der Gewinner wird dann mithilfe der Pixelrekursion verschoben und das Ergebnis wieder getestet. Für jedes Pixel werden also insgesamt nur vier Blöcke verglichen, sodass das Verfahren extrem schnell ist. Durch die Einbezie-

Tabelle 1. Exemplarische Abstände zwischen den gefundenen Tiefenebenen

Z in m	2	4	6	8	10	20	40
D in pixel	220	110	73	55	44	22	11
ΔZ/Z in %	0,5 %	0,9 %	1,4 %	1,8 %	2,2 %	4,4 %	8,3 %
ΔZ in cm	0,9	3,6	8,1	14,3	22,2	87,0	333,3

hung von örtlichen wie auch zeitlichen Nachbarn wird eine gute örtlich-zeitliche Glättung erreicht.

In untexturierten Bereichen sowie bei periodischen Texturen kann es jedoch immer zu Fehlschätzungen kommen. Eine Möglichkeit, falsche Korrespondenzen zu erkennen, bietet ein trifokaler Ansatz. Dabei werden nicht nur zwei, sondern drei Kameras verwendet. Für beide Stereosysteme (Mitte-Links und Mitte-Rechts) werden zuerst unabhängige Korrespondenzschätzungen ausgeführt. Die Ergebnisse können dann miteinander verglichen werden.

Dazu werden die Disparitäten eines Stereopaars in das andere Paar projiziert und mit den dort geschätzten Disparitäten verglichen. Durch Anwendung eines Schwellwertes kann entschieden werden, ob eine bestimmte Abweichung noch akzeptiert wird oder nicht.

Im **Bild 1** werden die geschätzten Tiefen durch Farben dargestellt. Das linke Bild zeigt eine dichte Tiefenkarte vor einer Konsistenzprüfung mithilfe der dritten Kamera. Man sieht zum Beispiel im blauen Hintergrund einige grüne oder rote Fehlschätzungen, die auf weiter vorne liegende Objekte hindeuten würden. In der Mitte sind die größten Fehler schwarz markiert. Im rechten Bild sind nur noch die Bereiche übrig und damit bunt dargestellt, bei denen die Schätzung über beide Stereopaare exakt das gleiche Ergebnis geliefert hat.

Werden Tiefenkarten für zum Beispiel die Postproduktion generiert, kann es wichtiger sein, nur verlässliche Tiefen zu haben, auch wenn die Tiefenkarte nur dünn besetzt ist. Soll jedoch direkt ein 3D-Display bedient werden, das zum Beispiel „Video+Tiefe“ als Eingangsformat erwartet, sind die Lücken durch geeignete Filtertechniken zu füllen. **Bild 2** zeigt ein Beispiel für eine so gewonnene dichte Tiefenkarte.

Die Berechnung von Tiefenkarten mittels Triangulation stellt somit ein Bindeglied zur Verfügung, um rein stereoskopische Aufnah-

men, wie sie zurzeit für das 3D-Kino gemacht werden, in ein Format zu überführen, das von neueren brillenlosen Displaysystemen benötigt wird.

Mehrfach-Kamerasysteme

Die Aufnahmetechniken mit zwei und mehr Kameras beruhen auf dem Prinzip der Triangulation. Wie viele Kameras man braucht, wie sie angeordnet sein sollen und wie sie bedient werden sind Fragestellungen, mit der sich Kameraentwickler und Filmproduzenten derzeit beschäftigen. Im Folgenden werden die heute gängigsten Ansätze von Multi-Kamera-Arrays vorgestellt.

Stereokameras

Die einfachste Methode, die dem menschlichen 3D-Sehen nachempfunden ist, ist die Verwendung von zwei identischen, parallel ausgerichteten Kameras mit einem Basisabstand „B“ als Abstand zwischen den optischen Achsen der Kameras (**Bild 3a**). Diese Anordnung wird als Stereoskopie bezeichnet und liefert zwei Ansichten der Szene als zwei vollständige Bilder bzw. Filme, die auch direkt die Ansichten für das linke und rechte Auge sein können. Der 3D-Eindruck, den ein menschlicher Betrachter hat, hängt neben dem verwendeten Objektiv direkt vom Basisabstand „B“ der Kameras ab. Ein Basisabstand von 65 mm entspricht etwa dem Augenabstand und erzeugt ein 3D-Bild, das dem menschlichen Sehen gleichkommt.

Im einfachsten Fall werden die Kameras nebeneinander aufgebaut (Side-by-Side-Rig), sofern es die Baugröße des Kamerakörpers oder der Objektive zulässt. Mit vielen Kameras und Objektiven ist ein Abstand von 65 mm – oder erforderlichenfalls auch noch kleiner – bautechnisch nicht möglich. Dann kommen Lösungen mit einem Spiegel-Rig zum Einsatz: Vor

den Kameras befindet sich ein halbdurchlässiger Spiegel in 45°-Position, die Kameras erhalten entweder das durchgehende oder das gespiegelte Bild mit einem horizontalen Versatz um „B“. Mit dieser Anordnung sind alle Basisabstände bis herunter zu 0 mm einstellbar.

Der große Vorteil der Stereokamera-Anordnung kommt zum Tragen, wenn die beiden Kanäle unverändert übertragen und direkt als Ansicht für das linke und rechte Auge verwendet werden können. Falls eine weitergehende Verarbeitung der beiden Bilddaten stattfinden soll, zum Beispiel zur Berechnung von weiteren Ansichten, Zwischenansichten, Codierung als Bild plus Tiefeninformation usw., hat die 2-Kamera-Stereokamera-Anordnung ihre Grenzen (s. Triangulation).

Zurzeit sind Stereokamerasysteme (Side-by-Side- oder im Spiegel-Rig) trotz einiger Einschränkungen die meistbenutzten Systeme. Die Praxis zeigt, dass damit durchaus hochwertige 3D-Kinoproduktionen möglich sind.

Center- und Satelliten-Kameras

Diese Anordnung besteht aus einer (hochwertigen) Center-Kamera und im horizontalen Abstand links und rechts davon aus zwei oder mehr Satellitenkameras (**Bild 3b**). Die Idee hierbei ist es, eine hochwertige Kamera zu haben, die das Bild liefert. Die Satellitenkameras werden nur zur Berechnung der Tiefeninformation verwendet, die mit zwei Satelliten zuverlässiger wird. Satelliten können deshalb auch geringere technische Leistungsdaten zum Beispiel bezüglich Auflösung, Farbtreue usw. haben. Bei der Berechnung neuer Ansichten können Bildteile, die in der Center-Kamera verdeckt sind, mithilfe der Satellitenkameras ergänzt werden.

Lineares Kamera-Array

Insbesondere wenn viele Ansichten einer Szene gefordert sind, kommen Kamera-Arrays mit „N“ gleichwertigen Kameras zum Einsatz, die im horizontalen Abstand parallel oder auch konvergent ausgerichtet sind (**Bild 3c**). Hiermit sind praktisch beliebig viele Zwischenansichten berechenbar, das Problem der Objektverdeckung wird überwiegend vermieden.



Dipl.-Ing. **Hans Bloß (FKTG)** † war Leiter der Abteilung „Electronic Imaging“ am Fraunhofer IIS in Erlangen und Sprecher der Fraunhofer-Allianz „Digital Cinema“ (Nachruf auf S. 133)



Dipl.-Ing. **Ralf Tanger** ist Projektleiter in der Abteilung „Bildsignalverarbeitung“ am Fraunhofer-Institut für Nachrichtentechnik Heinrich-Hertz-Institut (HHI) in Berlin



Peter Centen (1), **Klaas Jan Damsta (2)** und **Jan van Rooy (3)** sind Mitarbeiter des R&D-Departements „Camera Division“ von Grass Valley in Breda (Niederlande) und **Valter Drazic (4)** und **Arno Schubert (5)** sind Mitarbeiter von Thomson Corporate Research in Rennes, Frankreich



Bild 1. Ergebnisse einer Tiefenschätzung für verschiedene trifokale Abweichungen



Bild 2. Füllen einer Tiefenkarte

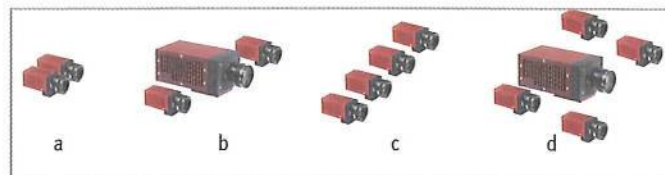


Bild 3. Verschiedene Anordnungen eines Multi-Kamera-Arrays

2-dim.-Kamera-Arrays

Der am weitesten gehende Ansatz sind „2-dim.“-Kamera-Arrays, entweder mit einer Center-Kamera und Satelliten-Kameras oder mit kreuzförmig angeordneten Kameras in horizontaler und vertikaler Richtung (Bild 3d). Damit lassen sich Tiefenkarten noch zuverlässiger errechnen, eine Verschiebung der Ansicht in vertikaler Richtung ist ebenfalls möglich. Besonders wenn in der Postproduktion noch animierte Objekte eingebracht werden sollen oder Objekte aus den Kameraaufnahmen ausgeschnitten und in andere Szenen kopiert werden, sind solche Aufnahmetechniken hilfreich. Man darf allerdings den apparativen Aufwand und die dabei entstehende Datenmenge nicht unterschätzen.

Technische Kameraanforderungen

Die technischen Anforderungen an Kameras für 3D-Produktion sind zunächst vergleichbar mit 2D-Kameras. Einige Punkte erfordern jedoch besondere Beachtung, bzw. es sind auch neue Aspekte zu berücksichtigen:

Auflösung: Der Zuschauer erwartet von 3D ein neues Seherlebnis mit hoher technischer Qualität. Geringe Auflösung (SD-Video) oder gar sichtbare Pixelraster werden wohl keine Akzeptanz finden (Ausnahme Handy-TV). Deshalb sollte man in der Produktion mindestens mit HD-Auflösung arbeiten, zumal in der Nachbearbeitung Auflösung verloren geht, da Randbereiche der Bilder, die nicht von allen Kameras erfasst werden, nicht nutzbar sind.

Eine geringe **Baugröße** spielt aus zwei Gründen eine wesentlich wichtigere Rolle als bei 2D-Kameras: Mit kleinen Kameras lassen sich auch kleine Basisabstände (ab etwa 50 mm) realisieren und damit meist die aufwendigen Spiegel-Rigs vermeiden. Multi-Kamera-Arrays (drei und mehr Kameras) können nur mit kleinen leichten Bauformen praxistauglich aufgebaut werden. Oft sind hier allerdings die Objektive der limitierende Faktor, da die Miniaturisierung guter Optiken nicht in gleichem Maß möglich ist wie bei elektronischen Komponenten.

Geometrische (Tonnen- und Kissenverzerrungen) und **farbmetrische Fehler** (chromatische Abberation) können bei 2D-Bildern bis zu einem gewissen Maß akzeptiert

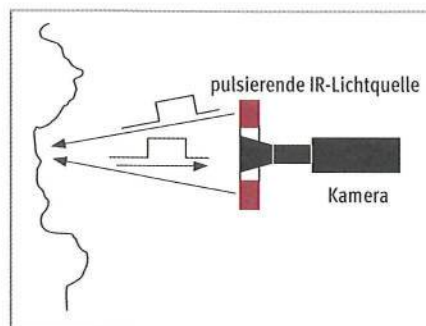


Bild 4. Kamera mit ToF-Prinzip

werden. In der 3D-Produktion, insbesondere wenn Tiefenkarten oder neue Ansichten erstellt werden müssen, sind diese Fehler nicht tolerierbar. Deshalb ist eine Korrektur dieser Fehler am besten bereits in der Kamera oder in der Postproduktion erforderlich.

Tiefenschärfe ist ein Stilmittel, das in der 2D-Spielfilmproduktion bewusst eingesetzt wird. Hier wird mit geringen Tiefenschärfen gearbeitet, um die Aufmerksamkeit der Betrachter zu führen und um „Tiefe“ in das Bild zu bringen. Ob das auch für 3D-Filme gilt, wird zurzeit vielfach diskutiert. Hier wird meist die Forderungen gestellt, dass eine 3D-Kamera zunächst durchgängig scharfe Bilder (hoher Tiefenschärfebereich) liefern soll.

Die **Bedienung der Kameras**, Einstellungen und Abgleich stellen hohe Ansprüche an den Kameramann. Idealerweise haben die Bilder aller Kameras die gleiche Helligkeit, Farbe usw. und alle Objektive sind identisch eingestellt. Alle Kameras synchron zu bedienen und identisch abzugleichen, überfordert den Kameramann am Set. Es wäre ein eigenständiger gegenseitiger automatischer und intelligenter Abgleich aller Kameras untereinander nötig, ein Feature, das zurzeit keine Kamera im Array bieten kann.

Fazit

3D-Produktionen werden derzeit mit „normalen“ 2D-Kameras gemacht. Dass das möglich ist, zeigen einige erfolgreiche Produktionen. Der Aufwand am Set ist aber relativ hoch, und viele Probleme werden in die Postproduktion verlagert. Falls 3D-Kino- und -TV-Produktionen weiter zunehmen, wovon man zurzeit ausgehen muss, entsteht ein Bedarf nach Kameras,

die auch die speziellen Anforderungen einer 3D-Produktion erfüllen. Am Fraunhofer IIS laufen derzeit Forschungsarbeiten, zum Beispiel das vom BMWi geförderte Projekt „Prime“ (www.prime3d.de), die kompakte und auch für 3D-Arrays taugliche Kameras zum Inhalt haben.

ToF-Kameras

Bei der Time-of-Flight-ToF-Methode (Bild 4) wird die Zeit, die ein Lichtimpuls von der Lichtquelle zur Szene und von der Szene zurück zur Kamera benötigt, ausgewertet. Falls die Lichtquelle an der Kamera montiert ist, wird dabei zweimal die Distanz (Z-Range) zwischen der Kamera und der Szene gemessen.

Das Spektrum der eingesetzten Lichtquelle wird üblicherweise so gewählt, dass es außerhalb des sichtbaren Lichtbereichs liegt, um Interferenzen mit der Bildaufnahme der Kameras zu vermeiden. Es handelt sich hierbei üblicherweise um IR-Lichtquellen. Die Lichtquelle kann entweder mit Impulsen angesteuert oder mit einer kontinuierlichen sinusförmigen Ansteuerung betrieben werden. Bei der sinusförmigen Ansteuerung wird die Phasenverschiebung zwischen gesendetem und empfangenem Licht ausgewertet – daher die Bezeichnung: Time-of-Flight (ToF) – und damit die Entfernung ermittelt.

Eines der ältesten Patente für ToF-Kameras kommt von Medina [2]. Unter den 32 Ansprüchen befindet sich auch die Sonne als periodische Lichtquelle!

Wenn man eine Szene mit einer 3-Sensor-Kamera aufnimmt, wird das Bild in die roten (R-), grünen (G-) und blauen (B-)Bildanteile zerlegt und aufgenommen. Wenn nun ein vierter Sensor verwendet wird, um die ToF zu messen, kann neben der „R,G,B“-Bildinformation noch die Entfernungsinformation der Szene auf Bildpunktbasis gewonnen werden.

Dadurch kann man Z-Range oder Time-of-Flight am besten als „Radar auf Pixel-by-Pixel-Basis“ bezeichnen.

Für Time-of-Flight (ToF) ist die Kamera mit einer pulsierenden IR-Lichtquelle versehen. Diese Lichtquelle sendet kurze IR-Lichtimpulse aus, die von der Szene reflektiert und von der Kamera aufgenommen werden. Durch eine synchrone Auswertung der reflektierten Lichtimpulse wird das Tiefenmuster der aufgenommenen Szene gewonnen. Um ein starkes Signal zu erhalten, sind mehrere kurze Impulse nötig, deren Dauer den messbaren Entfernungsbereich definiert. Wenn zum Beispiel bis zu 15 m gemessen werden soll, muss die Impulsdauer 100 ns oder mehr betragen. Die Zeit

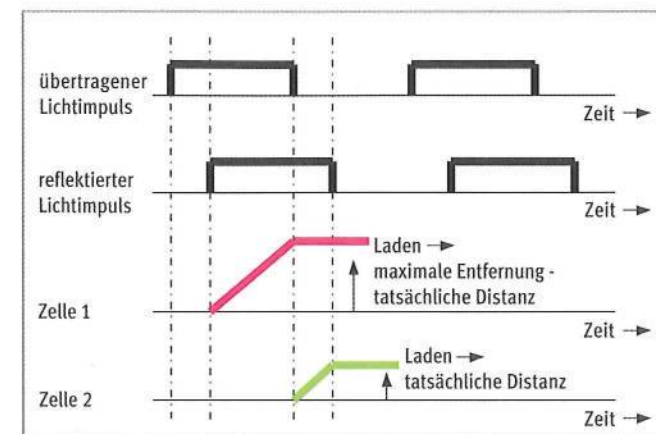


Bild 5. Zeitdiagramm der DPM-ToF-Zellen

zwischen zwei Impulsen wird so gewählt, dass der reflektierte Impuls in dasselbe Intervall fällt wie der gesendete Impuls, da sonst zweideutige Messergebnisse auftreten können.

Die Intensität „L“ des von der Kamera empfangenen Lichts (1) ist proportional zur Intensität des gesendeten Lichts „I“ geteilt durch die Distanz (R) zwischen Szene und Kamera in der vierten Potenz und multipliziert mit dem unbekannten Reflexionskoeffizienten der Szene (ρ). In einem Sensorbildpunkt wird das empfangene Licht in ein Ladungspaket umgewandelt, dessen Größe abhängig ist von der Integrationszeit des Bildpunktes und der Dauer des Lichtimpulses.

$$L \approx \frac{I}{R^4} t \quad (1)$$

Die Gleichung (1) hat zwei Unbekannte; die erste ist die Time-to-Flight und die zweite der Reflexionskoeffizient der Szene.

Schon 1998 hat die Kamera-Forschungs- und Entwicklungsabteilung von Thomson Grass Valley zusammen mit Dalsa einen Bildwandler mit eingebauter „Synchronous-Detection“ [3] entwickelt. Durch das Hinzufügen einer metallischen Schicht und Änderung der Sensor-Ansteuerung kann Synchronous-Detection ausgeführt werden. Die ToF-Messung wird dadurch ermöglicht. Synchronous-Detection verwendet dabei die bekannte Flexibilität der DPM- („Dynamic Pixel Management“) CCD-Sensoren [4], wobei hier jedoch durch das Anlegen entsprechender Ausleseimpulse zweigeteilte Bildpunkte erzeugt werden. In den letzten Jahren wurde diese Methode als „Lock-in Pixel“ bekannt [5].

Im Bild 5 sind die notwendigen zeitlichen Abläufe dargestellt. Von oben nach unten: Die erste Kurve stellt die Form und die Zeit des gesendeten Lichtimpulses dar; die zweite Kurve zeigt die entsprechenden Signale des reflektierten und empfangenen Lichtimpulses. Der empfangene Lichtimpuls ist um die Zeit verzögert, die das Licht benötigt, um den zweifachen Abstand zwischen Kamera und Objekt zurückzulegen. Die dritte Kurve zeigt den La-

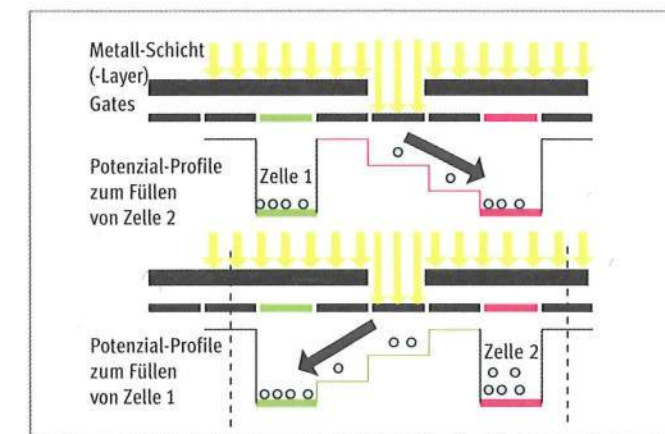


Bild 6. DPM-ToF-Bildpunkte

dungsaufbau im Bildpunkt 1, während der Lichtimpuls übertragen wird. Die vierte Kurve zeigt den Ladungsaufbau in Zelle 2 zu der Zeit, in der kein Lichtimpuls übertragen wird und der reflektierte Lichtimpuls noch nicht beendet ist.

Die ToF-Auswertung zeigt Bild 6. Sechs Register bestimmen das Verhalten der Bildpunkte. Der ToF-Bildpunkt besteht dabei aus zwei Zellen, die kohärieren. Die metallische Schicht verhindert, dass Licht direkt auf eine der beiden Zellen fallen kann. Durch ein Schalten der Register, wie in den roten und grünen Kurven im oberen und unteren Teil des Bildes dargestellt, kann Synchronous-Detection für das einfallende Licht vorgenommen werden. Die „grüne“ Zelle 2 nimmt dabei das Licht proportional mit der „aktuellen Entfernung“ auf, die „rote“ Zelle 1 das Licht proportional mit der „maximalen Entfernung – aktuellen Entfernung“.

Es lässt sich zeigen [2], dass die Summe der Ausgangssignale in Zelle 1 und Zelle 2 proportional der Menge des reflektierten Lichts ist. Teil 2 von Bild 6 demonstriert, dass das Ausgangssignal der Zelle 2 proportional der Entfernung und der Reflexion der Szene ist. Das Verhältnis zwischen dem Ausgangssignal der Zelle 2 und dem summierten Ausgang der Zellen 1 und 2 ist die „Time-of-Flight“ dividiert durch die übertragene Lichtimpulsdauer.

Da das Reflexionsverhalten der Szene ein unbekannter Faktor ist, wird dieser beim Divisionsvorgang nicht berücksichtigt. Anhand des erhaltenen Ergebnisses kann nun die aktuelle Entfernung errechnet werden. Dabei ist die ermittelte Entfernung die Summe der Distanzen von der „IR-Lichtquelle-zur-Szene“ und von der „Szene-zurück-zur-Kamera“.

Die **Vorteile** der beschriebenen Entfernungsmessung können folgendermaßen zusammengefasst werden:

- Es sind nur einige kleinere Änderungen an einem (HD-)DPM-CCD-Sensor notwendig.
- Es wird keine große Rechenleistung benö-

tigt, um die Entfernungsmessung vorzunehmen, da nur das Verhältnis zweier Werte zueinander errechnet werden muss.

- Dadurch ist das Kamerasystem sehr einfach.
- Für jeden Bildpunkt sind sowohl die Werte für R, G, B als auch die Entfernung bekannt.

Nachteile:

- Die normale Szenenbeleuchtung durch die Sonne oder auch die oftmals verwendeten Halogenlampen hat einen großen IR-Anteil im Licht, der zu Messfehlern führen kann.
- Es können Interferenzen auftreten, wenn mehrere ToF-Kameras dieselbe Szene aufnehmen.
- Der IR-Lichtimpuls kann mehrfach reflektiert werden.
- Aus Sicherheitsgründen wird die übertragene Lichtmenge limitiert. Dadurch können die gewonnenen Bilder verrauscht sein.

Plenoptic-Kameras

Hinter dem Begriff „Plenoptic“ verbirgt sich eine interessante Technik zur direkten 3D-Bildaufnahme. Die gleichzeitige Erzeugung von Aufnahmen aus mehreren Blickpunkten auf einem Sensor vermeidet komplizierte Abgleiche, Bildeinstellungen und Synchronisierungen. Zusätzlich kann eine Entfernungsinformation erzeugt werden, ohne das aufgenommene Objekt mit aktivem Licht wie Laser oder IR anzuleuchten. Die Erzielung guter Resultate mit hoher Qualität in der räumlichen Auflösung und der Entfernungsinformation ist nicht einfach.

Das Plenoptic-System ist eine Variante der Integralfotografie, die vor 100 Jahren durch den Nobelpreisträger Lippmann [6] erfunden wurde: Die einfache Anwendung eines Mikrolinsenrasters auf dem Bildaufnehmer erzeugt eine Richtungsinformation der aufgenommenen Bilder.

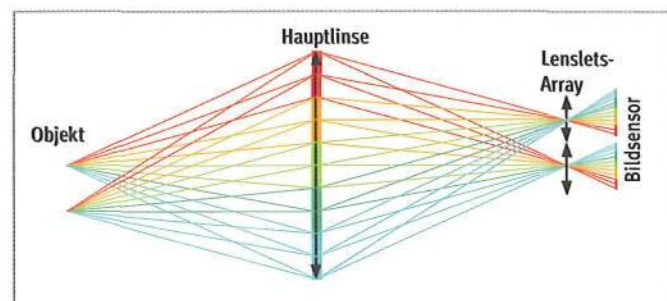


Bild 7. Basisprinzip Plenoptik-Kamera

Bild 7 illustriert das Basisprinzip von Plenoptik. Zwei Objektpunkte werden auf zwei Mikrolinsen abgebildet. Hinter den Mikrolinsen befinden sich jeweils mehrere Bildpunkte des Bildaufnehmers. Jeder Objektpunkt wird winkelabhängig auf mehreren Bildpunkten des Sensors abgebildet. Jeder zu einer spezifischen Mikrolinse gehörende Bildpunkt sieht das Objekt durch einen spezifischen Bereich der Hauptoptik. Wenn man das Sensorsignal so zerlegt, dass alle Bildpunkte, die das Objekt durch denselben Bereich des Objektivs „sehen“, ein Bild ergeben, können mehrere Abbildungen derselben Szene gewonnen werden, die sich durch eine unterschiedliche Parallaxe voneinander unterscheiden (Bild 8). Da jeder individuelle Bildpunkt unter einer Mikrolinse eine spezifische Parallaxe aufnimmt, können durch das Zerlegen und Verbinden der vergleichbaren Bildpunkte miteinander einzelne Bilder erzeugt werden. Wenn eine Mikrolinse 15 Bildpunkte überdeckt, wie im Bild 2 dargestellt, können auf diesem Weg 15 Abbildungen mit unterschiedlicher Parallaxe erzeugt werden. Die verschiedenen Abbildungen können nun entweder über ein autostereoskopisches Display wiedergegeben oder dazu benutzt werden, eine Entfernungsinformation zu generieren.

Verglichen mit anderen Formen zur Gewinnung der Tiefeninformation hat die Plenoptik-Kamera einige spezifische **Vorteile**:

- Die Tiefeninformation wird in einer Kamera gewonnen; dadurch ist eine Kalibrierung nicht notwendig und die Anwendung einfach.
- Die Plenoptik-Kamera ist ein passives System, benötigt keine Beleuchtung der Szene und kann daher auch ohne Einschränkungen für Außenaufnahmen eingesetzt werden.
- Es werden automatisch auch immer „Multi-View-Bilder“ erzeugt.
- Das Tiefenmuster kann einfach errechnet werden, Probleme der Übereinstimmung werden minimiert.

Die **Hauptnachteile** sind:

- Da die Tiefeninformation durch ein Objektiv gewonnen wird, ist damit nur ein eingeschränkter Abstand der Blickpunkte möglich; dadurch könnte die Genauigkeit der Entfernungsmessung zu gering sein.

- Die Auflösung der einzelnen Abbildungen entspricht der Sensorauflösung geteilt durch die Anzahl der verschiedenen Abbildungen. Die räumliche Auflösung der einzelnen Abbildungen ist gering.

Zwei Hauptarbeiten haben das Plenoptik-System in den letzten Jahren vorangetrieben: eine Veröffentlichung durch Adelson [7] und eine weitere Studie durch die Stanford University [8]. Thomson arbeitet im Rahmen des europäischen IST-FPr-Projekts „Osiris“ (www.osiris-project.com) an der Untersuchung der Plenoptik-Aufnahme. Die Originalität der Arbeit liegt in der Optimierung des Mikrolinsenrasters der Kamera zur Gewinnung der Entfernungsinformation für 3D-Anwendungen und der gleichzeitigen Optimierung des Systems zur Erzeugung von Bildern mit brauchbarer Auflösung.

Zusammenfassung und Ausblick

Derzeit kann die stereoskopische Aufnahme als die meistbenutzte und zuverlässigste Technik zur Erzeugung von 3D-Bildern angesehen werden. Sie ist vom Konzept her zwar einfach und es können zunächst die bereits existierenden 2D-Kameras verwendet werden, die Handhabung am Set ist aber noch sehr aufwendig und es wird viel Arbeit in die Postproduktion verlagert. Praktisch alle aktuellen heute verfügbaren 3D-Filme werden mit diesen Mehrfachkamerasystemen produziert. Es wird erwartet, dass das auch in naher Zukunft so bleiben wird und Multi-Kamera-Arrays entsprechend weiterentwickelt werden.

Laserscanner wie sie in der industriellen Messtechnik verwendet werden sind von der Bildrate her zu langsam und werden wahrscheinlich am Set keine Akzeptanz finden.

Techniken wie die vorgestellte Time-of-Flight-Methode oder Plenoptik-Systeme genügen im derzeitigen Entwicklungsstand noch nicht den Anforderungen einer Film- oder TV-Produktion. Besonders die örtliche Auflösung, Pixel und Sensorgröße sind hier noch Schwachpunkte. Sie können aber mit der zu erwartenden technischen Weiterentwicklung eine wertvolle Ergänzung zu Stereo- und Multi-Kamera-Arrays werden, besonders wenn zu den stereoskopischen Ansichten zusätzlich Kalibrierdaten oder absolute Entfernungsmessungen nötig sind.

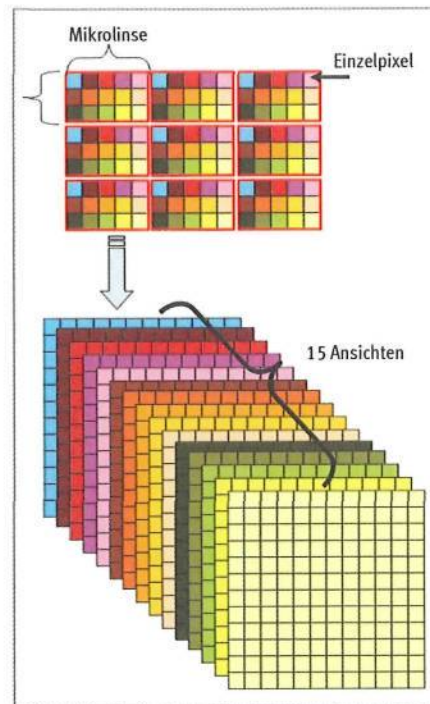


Bild 8. Prinzip des Parallax-Demultiplexings

Literatur

- [1] Atzpadin, N.; Kauff, P.; Schreier, O.: Stereo Analysis by Hybrid Recursive Matching for Real-Time Immersive Video Conferencing. IEEE Transaction on Circuits and Systems for Video Technology, Vol. 14, No. 3, S. 312–333, März 2004.
- [2] Medina, Three Dimensional Camera and Range Finder, Patent US5081530.
- [3] Stekelenburg, M.; Centen, P.: Private communication. 1998.
- [4] Centen, P.; Mierop, A.; Stoldt, H.; Theuvsen, A.; Pine, G.; Stekelenburg, M.: Dynamic Pixel Management for second generation aspect ratio switching. Montreux International Television Symposium Record, 8. bis 13. Juni 1995, S. 199–211.
- [5] Buttgen, B.; Oggier, T.; Lehmann, M.; Kaufmann, R.; Lustenberger, F.: CCD/CMOS Lock-in Pixel for Range Imaging: Challenges, Limitations and state-of-the-art. Proc. Range Imag. Res. Days, ETH, Zürich, Schweiz, September 2005, S. 21–32.
- [6] Epreuves réversibles, photographies intégrales, Académie des Sciences. www.tgeorgiev.net/Lippmann/EpreuvesReversibles.pdf
- [7] Single Lens Stereo With a plenoptic Camera. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, Nr. 2, Februar 1992.
- [8] Light Field Photography with a Hand-Held Plenoptic Camera. Ren Ng, Stanford Tech Report CTSR 2005-02.

Einige Forschungsarbeiten in diesem Bericht werden durch das Bundesministerium für Wirtschaft und Technologie BMW im Rahmen des PRIME-Projektes gefördert (Produktions- und Projektionstechniken für Immersive Medien, www.prime3d.de).

Bitte lesen Sie auch den Beitrag „3DTV: Internationale Aktivitäten und Standards“ auf S. 158–164 und den Nachruf S. 133 in dieser Ausgabe.



ANGACable

Fachmesse für Kabel, Breitband und Satellit

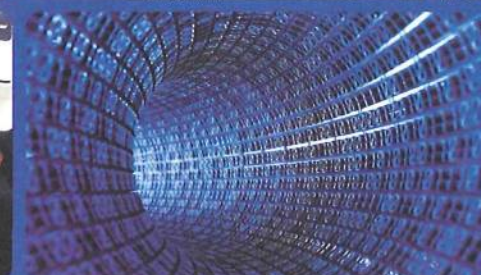
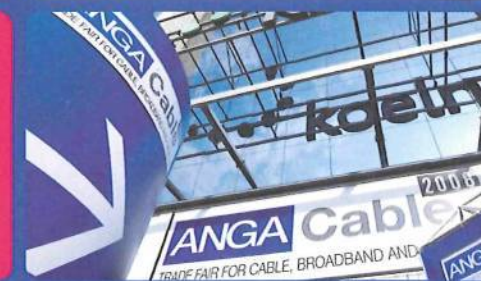


Die Fachmesse

- 369 internationale Aussteller in 2008
- 12.000 Fachbesucher aus 68 Ländern
- 22.000 qm Brutto-Ausstellungsfläche

26. bis 28. Mai 2009
Köln

www.angacable.de



Der Kongress

- 1.200 Teilnehmer aus 32 Ländern in 2008
- Strategie- und Technik-Panels zu den Themen Triple Play, Plattformen, Regulierung, Content, HDTV, IPTV, Next Generation Networks u. v. a.



ANGA Services GmbH
Sebastianstr. 189 · 53115 Bonn
Tel. +49 (0)228/96 21 890
Fax +49 (0)228/96 21 895
info@angacable.de
www.angacable.de