

# Lichtfeldkameras als Interaktionsmedium und Input für die Medienfassade am Campus der HAW Hamburg

Jorin Kleimann

Hochschule für angewandte Wissenschaften Hamburg  
Berliner Tor 7, 20099 Hamburg  
jorin.kleimann@haw-hamburg.de  
www.haw-hamburg.de

**Abstract.** Es soll mit dem Einsatz einer Lichtfeldkamera an einer Medienfassade am Campus Berliner Tor der HAW Hamburg der Raum vor der Medienfassade beobachtet und erfasst werden. Mit Hilfe der Lichtfeldkamera sollen die Nutzer einer Anwendung auf der Medienfassade in der Lage sein mit ihr zu interagieren. Dies soll über Gesten und menschliche Bewegungen geschehen, welche die Lichtfeldkamera auffasst und verarbeitet. Anwendungsentwicklern soll ein Interface bzw. eine API geschaffen werden, mit der die von der Lichtfeldkamera aufgezeichneten Gesten als Steuerung für die Anwendung auf der Medienfassade genutzt werden kann.

**Keywords:** light field camera, human computer interaction, gesture recognition, public display

## 1 Einleitung

### 1.1 Motivation

Lichtfeldkameras sind sowohl in der Fotografie als auch in der Sparte der Videoaufnahmen eine interessante und aktuelle Technologie. Lichtfeldkameras ermöglichen durch eine große Anzahl an sogenannten Mikrolinsen die zusätzlich zur Hauptlinse ein Objekt aus verschiedensten Winkeln aufnehmen die nachträgliche Fokussierung einzelner Bereiche eines Bildes. Außerdem kann durch die Speicherung des Lichtfeldes die Tiefe bestimmt und 3-Dimensionale aufnahmen erzeugt werden. So sind Lichtfeldkameras durch den steigenden Markt der 3D-Displays und 3D-Fernseher in den Fokus von Forschung und Entwicklung geraten.

Ein Beispiel für diese Entwicklung ist die Lichtfeldkamera der Firma Lytro [2]. Diese für den breiten Markt entwickelte Kamera ermöglicht sowohl 3D-Filmaufnahmen als auch die Möglichkeit Bilder nachträglich zu fokussieren. Durch den nachträglichen Fokus können Fotografen sogenannte "lebendige Bilder" erstellen, in denen man per Mausklick den Schärfebereich definieren kann.

Auch Unternehmen wie Raytrix [3], die ihren Fokus mehr auf den professionellen Markt legt und mit ihren Lichtfeldkameras in der Lage ist 3D Aufnahmen in hoher Qualität zu machen, definieren den Trend Lichtfeldkameras.

Medienfassaden oder auch öffentliche Displays werden heute vermehrt eingesetzt um öffentliche Orte und Plätze im digitalen Zeitalter neu zu erfinden. Dabei bieten solcherlei Displays Anwendungsspielraum für soziale Interaktionen, als Informationstafeln oder für interaktive Spiele.

Die Verwendung einer Lichtfeldkamera zur Interaktion mit einer Medienfassade oder einem Display im öffentlichen Raum soll dazu dienen, den Raum vor der Medienfassade zu erfassen und etwaige Interaktionen der Nutzer mit der Medienfassade für spezielle Anwendungen zu erkennen.

## 1.2 Problemstellung

Am Campus Berliner Tor der Hochschule für angewandte Wissenschaften wurde im Rahmen des “Pimp my Campus”-Projektes eine Medienfassade errichtet, mit der über verschiedenste Arten interagiert werden soll. Dabei dient eine Lichtfeldkamera sowohl zur Erkennung des Raumes vor der Medienfassade, als auch zur Erkennung von Gestiken und Bewegung der Menschen vor der Fassade. Die Medienfassade am Campus Berliner Tor der HAW Hamburg ist aus technischer Sicht nur ein einfaches, nicht hochauflösendes, dafür aber sehr großes Display, das an einer der am meist frequentierten Orte angebracht ist. An der Medienfassade ist derzeit ein herkömmlicher Computer angeschlossen.

Auf lange Sicht soll mit Hilfe der Lichtfeldkamera Anwendungsentwicklern, die beispielsweise Anwendungen für die Medienfassade entwickeln wollen, die Möglichkeit gegeben werden, die Lichtfeldkamera als Steuereinheit für ihre Anwendungen zu nutzen. Dies kann über eine API geschehen, die die Anwendungsentwickler nutzen.

## 1.3 Grundsätzliche Idee

Die grundsätzliche Idee hinter meiner Arbeit ist es, die an der Hochschule vorhandenen Lichtfeldkamera der Firma Raytrix zu benutzen um damit den Raum und die Interaktionen vor der Medienfassade aufzunehmen bzw. aufzuzeichnen und sie weiterzugeben. Der Einsatz der Lichtfeldkamera für die Medienfassade ist aus dem Entwurf für die Medienfassade heraus entstanden. Grundsätzlich soll es nach der Installation und vollständigen Implementierung der Kamera für Anwendungsentwickler möglich sein, die von mir definierten Gesten und Steuerungsmöglichkeiten (die dann mit Hilfe der Kamera aufgenommen werden) als Steuerung für ihre Anwendungen zu nutzen.

## 2 Lichtfeldkameras und plenoptische Funktion

Die Funktion von Lichtfeldkameras, in einigen Publikationen auch plenoptische Kameras genannt, basiert auf der Idee der plenoptischen Funktion. Die plenoptische Funktion ist eine Funktion, die in der Bildverarbeitung genutzt wird, um ein Bild aus jedem Blickwinkel und jeder beliebigen Position zu beschreiben.

Dabei wird die Verteilung der Strahldichte entlang von Lichtstrahlen in einem Raum gemessen um so das Bild in seiner Tiefe beschreiben zu können [4].

Innerhalb einer Lichtfeldkamera, die auf dieser Funktion basiert, ist zusätzlich zur Hauptlinse, die auch in herkömmlichen Kameras zu finden ist, eine große Anzahl an sogenannten Mikrolinsen auf einer Glasplatte angebracht. Diese Glasplatte mit den Mikrolinsen sitzt direkt vor dem Bildsensor der Lichtfeldkamera. Im Falle moderner Lichtfeldkamas wie denen der Firma Raytrix können bis zu 20.000 Mikrolinsen auf einer solchen Glasplatte sitzen [1].

Die Hauptlinse der Lichtfeldkamera bildet bei einer Aufnahme die außenliegende (also die aufzunehmende) Szene vor der Glasplatte mit den Mikrolinsen ab. Diese dann zwischen Hauptlinse und Mikrolinsen liegende Szene wird von den einzelnen Mikrolinsen aufgenommen und auf dem Sensor abgebildet. Dieser Vorgang ist in Abbildung 1 zu erkennen.

Dabei ist es wichtig, dass der Abstand des Bildsensors zur Glasplatte gleich der Höhe einer Mikrolinse ist. Ebenso muss die Höhe der Hauptlinse gleich dem Abstand zwischen Mikrolinsen und Hauptlinse sein. Um zu verhindern, dass die Mikrobilder der Mikrolinsen überlappen, müssen die jeweiligen Abstände in folgendem Verhältnis stehen [6]:

$$\frac{f_m}{\varphi_m} = \frac{f_l}{\varphi_l} \quad (1)$$

Die Abbildung durch die Mikrolinsen erfolgt in Abhängigkeit des jeweils einfallenden Lichtfeldes. Das bedeutet, dass jeder einfallende Lichtstrahl aus der gleichen Richtung durch die Mikrolinsen auf den gleichen Koordinaten auf dem Bildsensor abgebildet wird.

Die einzelnen Mikrolinsen bilden eigene, sogenannte Mikrobilder, die zusammengefügt das komplette Bild ergeben. Aus den einzelnen Mikrobildern der Mikrolinsen entsteht dann in der Kamera das sogenannte Rohbild. In diesem Rohbild sind die einzelnen Mikrobilder zusammengefügt. Aus der Differenz benachbarter Mikrobilder für gleiche Bildpunkte lässt sich die Position im Raum bestimmen. So kann eine Tiefenkarte für das Bild erstellt werden, in denen die verschiedenen Tiefenbereiche farblich unterschieden sind.



**Fig. 1.** Aufnahme des Zwischenbildes [1]

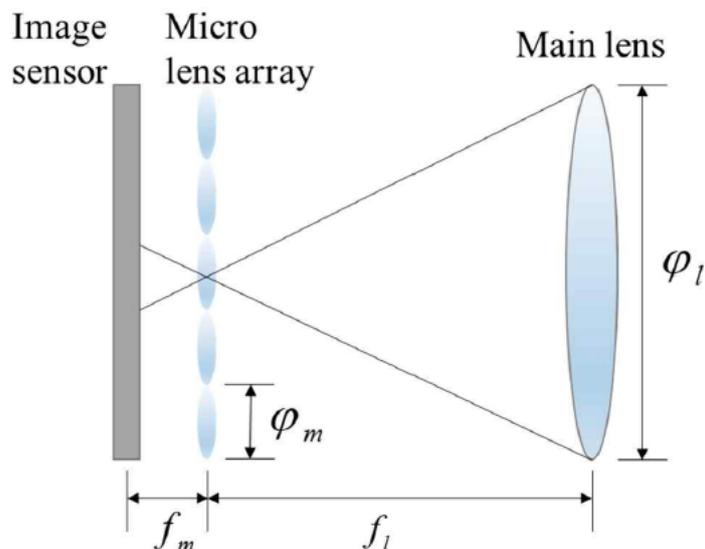


Fig. 2. Linsen einer Lichtfeldkamera [6]

### 3 Interaktion mit Medienfassade

Die Interaktion mit einer Anwendung auf der Medienfassade kann auf verschiedenste Weisen vonstatten gehen. Dabei sind Smartphones und Tablets, Controller von Spielekonsolen, Kameras, Laserpointer, oder Gesten vorstellbare Möglichkeiten der Steuerung. Diese Arbeit konzentriert sich jedoch nur auf die Interaktion mit der Medienfassade über eine an der Medienfassade angebrachte Lichtfeldkamera, die in der Lage ist, die Bewegungen vor der Medienfassade aufzuzeichnen.

Diese Bewegungen lassen sich in vielfältige mögliche Szenarien aufteilen. Ein Szenario ist die Erkennung von Gesten und Bewegungen der Menschen bzw. der Nutzer einer Anwendung oder eines Spiels auf der Medienfassade. Durch eine Bewegungserkennung ist es beispielsweise möglich ein menschliches Pong zu entwerfen, bei dem die Spieler durch hin- und herlaufen ihre Pong-Schläger auf der Medienfassade hoch und runter bewegen können. Die Lichtfeldkamera erkennt diese Bewegungen und gibt den jeweiligen Standort der Person auf dem virtuellen Spielfeld an die Anwendung weiter.

Ebenso könnten vor der Medienfassade sportliche Aktivitäten stattfinden, die auf der Medienfassade vorgemacht, erklärt, und gezeigt werden können. Die Menschen vor der Medienfassade können die Übungen nachmachen. Die Lichtfeldkamera ist mit Hilfe ihrer Tiefenerkennung dann in der Lage die Übungen zu verifizieren. Ein weiteres mit Hilfe der Lichtfeldkamera mögliches Szenario ist die Erkennung bzw. die Aufzeichnung und Abbildung von Gesichtern der Personen die sich vor der Medienfassade bewegen. Sieht man von den datenschutz-

rechtlichen Problemen einmal ab, mit denen sich bei einer Implementierung genauestens auseinandergesetzt werden müsste, ist dieses Szenario gerade für die Bereiche Werbung und soziale Interaktionen interessant. Mit Hilfe der Lichtfeldkamera könnten Gesichter aufgezeichnet (jedoch nicht zwingend gespeichert) und dann auf der Medienfassade abgebildet werden.

Die Lichtfeldkamera kann auch dabei helfen, den Raum vor der Medienfassade zu analysieren. Dabei könnten zum Beispiel folgende Fragen mit Hilfe der Lichtfeldkamera beantwortet werden: Wieviele Personen gehen an der Fassade vorbei? Wie sehen Bewegungsmuster bei verschiedenen Anzeigen aus? Wann bleiben mehr Menschen stehen?

## 4 Verwandte Arbeiten und aktuelle Forschung

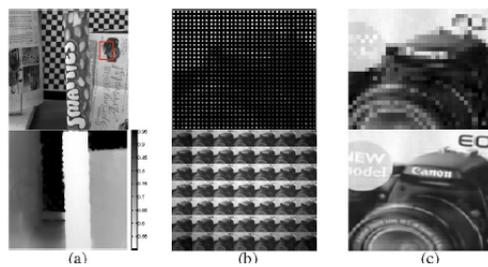
Es gibt zwar verschiedenste Forschungen und Arbeiten zu Lichtfeldkameras, zu deren Funktion und wie man mit etwaigen Problemen umgeht oder sie verhindern kann. Ebenso gibt es einige aktuelle Forschungen zur Thematik der öffentlichen Displays, in dessen Bereich die Medienfassade fällt. Jedoch gibt es wenig bis keine Forschung im Bereich der öffentlichen Displays mit Verwendung von Lichtfeldkameras. Im folgenden werde ich einige der interessanten Arbeiten aus dem Bereich Lichtfeldkameras und öffentliche Displays vorstellen.

### 4.1 Gesichtserkennung auf weite Entfernung mit Lichtfeldkameras

Eine für die Forschung in dem Bereich der Lichtfeldkamera interessante und für mich hilfreiche Arbeit ist eine empirische Studie über die Gesichtserkennung und -identifizierung mit Hilfe von Lichtfeldkameras [7]. Durch die Möglichkeit der Lichtfeldkameras Bilder nachträglich in bestimmten Bereichen zu fokussieren, lassen sich bei Gesichtern ein großer Teil des für die Gesichtserkennung relevanten Teils fokussieren. Ebenso lassen sich Bereiche eines Bildes, in denen mehrere Gesichter befinden gut nachträglich fokussieren. In einem Experiment haben die Forscher Bilder von mehreren Gesichtern in einer Szene mit einer Lichtfeldkamera aufgezeichnet, die einzelnen Gesichter nachträglich fokussiert und die Gesichter extrahiert. Diese Bilder der Gesichter wurden dann auf ihre Auflösung und Schärfe untersucht. Anschließend kommen die Autoren zu dem Schluss, dass die Lichtfeldaufnahmen vielversprechend für die Gesichtsbio-metrie sind. Auf mittlerer Entfernung erhielten die Autoren eine Identifikationsrate von 53,62%. Auch wenn diese Rate nicht sonderlich hoch erscheint, ist sie doch deutlich höher als die Gesichtserkennung mit Bildern herkömmlicher Kameras auf die gleiche Entfernung. Ein Schluss, den ich aus diesem Paper für meine zukünftige Arbeit ziehen kann ist, dass ich davon ausgehen kann, dass die Lichtfeldkamera nicht nur in der Lage ist Gesichter, sondern auch Gesten und Bewegungen auf weitere Entfernungen (10-20 Meter) erkennen kann.

#### 4.2 Die Lichtfeldkamera mit erweiterter Tiefenschärfe und hoher Auflösung

Ein weiteres Paper, welches sich mehr den Grundlagen, den Vorteilen und Problemen von Lichtfeldkameras beschäftigt ist das Paper mit dem Titel “The Light Field Camera: Extended Depth of Field, Aliasing, and Superresolution” [5]. Die Arbeit zeigt, dass Lichtfeldkameras eine höhere Tiefenschärfe besitzen und trotzdem Details in hoher Auflösung wiedergeben können. Hierzu wurde eine Methode entwickelt, welche die Nachteile von Lichtfeldkameras, wie die minimierte Auflösung die durch die Anzahl der Mikrolinsen definiert wird, oder auch trübe aufnahmen mit denen aktuelle Kameramodelle noch zu kämpfen haben, versucht zu minimieren.



**Fig. 3.** a oben: Ursprünglich mit der Lichtfeldkamera aufgenommenes Bild. b oben: aus dem Lichtfeldbild extrahiertes Objekt. c oben: detailliertes Bild des Objektes, ein Pixel pro Mikrolinse. c unten: mit der Methode aus dem Paper hochaufgelöstes Objekt

#### 4.3 3D Freihand-Gestiken für interaktive Displays im öffentlichen Raum

Diese Arbeit beschäftigt sich mit der wachsenden Verwendung von großen Displays in öffentlichen Räumen [8]. Die Autoren gehen auf die Möglichkeiten der Interaktion mit solcherlei großen Displays ein. Mit der Argumentation, dass Tastatur und Maus für öffentliche Displays, bei denen der Anwender meist weiter weg steht, eher ungeeignet sind, haben die Autoren Gestiken entwickelt, mit denen diese gesteuert werden können. Außerdem gehen die Autoren in dem veröffentlichten Paper auf verwandte Arbeiten ein. Es wird unter anderem auf die Auswirkung von öffentlichen Displays für den urbanen Raum und das urbane Leben eingegangen. Die Autoren entwickeln eine Steuerung durch Gestiken für eine digitale Stadt in 3D. Auffällig ist, dass die Gestensteuerung nicht signifikant langsamer war als das navigieren mit Maus und Tastatur. Wenn man dabei in betracht zieht, dass Maus und Tastatur für öffentliche Displays so gut wie nicht nutzbar sind, ist die Steuerung über Gesten eine gute Alternative.

## 5 Ziele des Projektes

Aus der anfangs genannten Grundidee und der Problemstellung entstehen die Ziele für das kommende Grundprojekt, das Hauptprojekt und die anschließende Masterarbeit. Dabei sollen die einzelnen Ziele als etappen gesehen werden, die im Aufwand und Schwierigkeitsgrad steigen. Die Ziele sollen aufeinander aufbauend sein. Die im kommenden Abschnitt genannten Ziele für das Hauptprojekt als auch für die Masterarbeit sind im Aufwand noch nicht abzusehen und somit noch nicht definitiv festlegt.

### 5.1 Ziel des Grundprojekts

Im Grundprojekt soll als Ziel noch nicht das Arbeiten mit der vorhandenen Lichtfeldkamera sein, sondern viel mehr das Arbeiten mit einer herkömmlichen Videokamera. So kann beispielsweise die Medienfassade durch einen normalen Monitor simuliert werden. Als Kamera könnte eine herkömmliche Webcam zum Einsatz kommen. Mit Hilfe der Webcam soll dann der Raum vor dem Monitor ausgewertet und auf menschliche Objekte untersucht werden. Dies kann beispielsweise mit schon vorhandener Software oder SDKs geschehen. Je nach absehbarem Arbeitsaufwand kann dann auch eine Software zur Gestenerkennung angewendet werden und eventuell eigene Gesten zur Steuerung entwickelt und getestet werden.

### 5.2 Ausblick Hauptprojekt und Masterarbeit

Im Hauptprojekt soll auf die Erkenntnisse aus dem Grundprojekt zurückgegriffen werden und die Lichtfeldkamera angewendet werden. Zum einen soll ihre Genauigkeit, ihre API getestet und angewendet werden und anschließend die Tiefenerkennung angewendet werden. Hierzu muss dann vorraussichtlich auch eine API entwickelt werden, welche die Anwendungsentwickler, die Anwendungen für die Medienfassade entwickeln, nutzen können. Das Hauptprojekt wird auch unter Berücksichtigung der Literatur und aktueller Forschung geschehen. Jedoch können sich die Ziele im Hauptprojekt noch leicht verschieben, wenn beispielsweise in der Implementierung der Lichtfeldkamera Schwächen entdeckt werden, die nicht eingeplant sind.

Aufgrund der fehlenden Erfahrungen und der bisher nur in der Theorie überlegten Ziele und Anwendungen für die Projekte kann ich noch keine nähere Auskunft über den engeren Themenbereich der Masterarbeit geben. Außerdem ist es noch nicht abzusehen wo das Thema Lichtfeldkamas und interaktive Displays in zwei Jahren steht.

## 6 Zusammenfassung und etwaige Probleme

Die verschiedenen vorgestellten Arbeiten haben gezeigt, dass Lichtfeldkamas ein aktuelles und ein oft bearbeitetes Thema sind. Mit der Erschließung des

Marktes für normale Konsumenten werden Lichtfeldkameras in der Zukunft mit hoher Wahrscheinlichkeit einiges an aufsehen erregen und somit hoffentlich auch einen weiteren Schub in der Forschung ergeben. Jedoch hat die Recherche auch gezeigt, dass die Verwendung von Lichtfeldkameras im Verbund mit Medienfasaden oder öffentlichen Displays so direkt noch nicht implementiert wurde. Hier steckt auch eine große Gefahr. Denn wenn man sich in ein recht neues Gebiet bewegt, ist der auf einen zukommende Arbeitsaufwand schwer einzuschätzen. Jedoch habe ich die Ziele für das Hauptprojekt bewusst noch nicht konkret gesteckt, da eben dieser Arbeitsaufwand noch nicht abzusehen ist.

## References

1. Eine kleine Revolution: Die 3D Lichtfeldkamera. [http://www.raytrix.de/index.php/newsreader/items/raytrix-in-deutsches-museum.html?file=tl\\_files/downloads/public/documents/Raytrix\\_DeutschesMuseum\\_Poster.pdf](http://www.raytrix.de/index.php/newsreader/items/raytrix-in-deutsches-museum.html?file=tl_files/downloads/public/documents/Raytrix_DeutschesMuseum_Poster.pdf)
2. Lytro ILLUM - A new way to capture your stories. <https://www.lytro.com/illum/>
3. Raytrix - 3D light field camera technology. <http://www.raytrix.de>
4. Adelson, E.H., Bergen, J.R.: The Plenoptic Function and the Elements of Early Vision. In: Computational Models of Visual Processing. pp. 3–20. MIT Press (1991)
5. Bishop, T., Favaro, P.: The Light Field Camera: Extended Depth of Field, Aliasing, and Superresolution. Pattern Analysis and Machine Intelligence, IEEE Transactions on 34(5), 972–986 (May 2012)
6. Jeong, Y., Kim, J., Yeom, J., Hong, K., Lee, B.: Handheld light field camera for intergral imaging with pixel mapping algorithm. In: 3D Imaging (IC3D), 2013 International Conference on. pp. 1–4 (Dec 2013)
7. Raghavendra, R., Raja, K.B., Yang, B., Busch, C.: Improved Face Recognition at a Distance Using Light Field Camera & Super Resolution Schemes. In: Proceedings of the 6th International Conference on Security of Information and Networks. pp. 413–416. SIN '13, ACM, New York, NY, USA (2013)
8. Ren, G., Li, C., O'Neill, E., Willis, P.: 3D Freehand Gestural Navigation for Interactive Public Displays. Computer Graphics and Applications, IEEE 33(2), 47–55 (March 2013)