**Höchstens 110 Wörter in einer Minute**

**Insgesammt also 2200 Wörter**

**Bei 20 Folien nicht mehr als 110 Wörter pro Folie**

**Folie 1: (70) 1min**

Das Thema meiner Abeit war die Szenenrekonstruktion und Kamerakalibrierung aus heterogenen stereoskopischen Bildquellen

Die Szenenrekonstruktion und die Kamerakalibrierung sind Themenbereiche der sogenannten Computer Vision

Die Computer Vision ist ein Fachbereich der Computer Science mit dem Fokus auf die Entwicklung von künstlicher Intelligenz, die ein visuelles Verständnis ihrer Umgebung besitzen.

Folglich wird in der Computer Vision der Weg von visuellen Eindrücken oder Bildern aus der Realität in den Rechner beschrieben .

**Folie 2: (57) 1min**

Der Mensch ist mit der Fähigkeit ausgestattet, gesehene Bilder zu verarbeiten und kann die ihn umgebene Welt verstehen.

Maschinen, die eine ähnliche Fähigkeit besitzen, wären somit ebenfalls in der Lage Entscheidungen auf Grund von visuellen Eindrücken zu fällen.

Das entwickeln solcher Maschinen ist Forschungsmittelpunkte von aktuellen Anwendungsbereichen wie dem Autonomen Fahren, Motion- Capturing, Bewegungserkennungen oder Service Robotern.

**Folie 3: (115) 1min**

* Solche Anwendugen basieren auf der sogenannten Szenenrekonstruktion
* Bei der Szenenrekonstruktion wird die Tiefeninformation der durch Kameras aufgenommenen Umgebung rekonstruktiert
* Da die Bildtiefe aus einem Bild nicht erkennabr ist, wurde in dieser Arbeit mit stereoskopischen Aufnahmen gearbeitet
* Das weitläfuigte Ziel welches mit dem in dieser Masterarbeit erarbeiteten Ansatz verfolgt werden soll, ist es aus 2D Informationen von Kameras unterschiedlicher Auflösung und Art (Infrarot und RGB) eine 3D Szene bzw 3D-Szenenpunkte rekonstruieren zu können
* Anwendung findet dies beispielsweise bei Drohenen welche oft zwei verschiedene Kameras haben.
  + eine hochauflösende RGB Kamera und infrarotkamera für Tiefenmessungen mit meist deutlich niedrigeren Auflösung
  + Könnten die Szene instantan mit beiden Kamerabildern rekonstruiert werden, wird die nachbearbeitung und das von hand mappen erleichtert

**Folie 4: (86) 1min (Hier war ich bei 3 einahlb minuten)**

Die Hauptthemen mit denen sich in der Masterthesis beschäftigt wurden enthalten vier wichtige Grundbausteine, welche im folgenden erläutert werden

Grundlagen:

* Diese beinhalten das verwendete Kamermodell
* Grundlegendes Verständnis der Fundamental Matrix und der essentiellen Matrix, welche wichtige Instrumente in der Szenenrekonstruktion von Stereobilpaaren darstellen

Synthetische Rekonstruktion:

* Zuerst wurde der Grundriss des Algorithmus an einer synthetisch erstellten Szene implementiert

Reale Rekonstruktion:

* Danach wurde der Algorithmus für die Anwendug auf ein reales Stereobildpaar entsprechend erweitert

Auswirkung unterschiedlicher Auflösungen:

* Danach wurde getestet, welche Auswirkungen unterschiedliche Kameraauflösungen auf den Algorithmus haben
* **Folie 5 (150) 1 min**

Zunächst zu den Grundlagen:

* Zuerst wird geklärt wie ein 3D-Punkt im raum auf einen 2D-Punkt auf dem Sensor abgebildet wird
* Wir betrachten den Weg eines Punktes in eine Kamera (C) genauer

(Nächstes Bild)

* Ein 3D Punkt durchläuft insgesammt 3 verschiedene Koordinatensysteme bevor er auf dem Sensor projiziert ist.
* Der Punkt ist zunächst in Weltkoordianten gegeben (O)
* C beschreibt das Kamerazentrum und ist gleichzeitig der Urspurng des Kamerakoordiantensystems C in welches der Punkt zunächst transformiert wird
* Zeta beschreibt die Brennweite , welche in eine Abbildungsmatrix eingebaut wird und den 3D-Kamerakoordinatenpunkt auf die 2D-Bildebene projizert.
* Der Schnittpunkt der Hauptachse (Achse Zeta) mit der Bildebene wird Hauptpunkt genannt. Der Hauptpunkt ist Urspung des 2D- Bildebenenkoordintansystems
* Mit einer letzten Transformation wird der 2D-Bildebenenpunkt in einen 2D Sensorkoordinatenpunkt tranformiert. Dessen Ursprung sich meist in eienr Ecke der Bildebene befindet.
* Die Projektion eines 3D-Objektpunktes im Raum in einen 2D Punkt auf den Sensor der Kamera wird durch die sogenannten Projektionsmatrix P beschrieben

**Folie 6: (125) 1 min**

* Die Projektionsmatrix P setzt sich zusammen aus den extrinsischen Kameraparametern welche durch eine Rotationsmatrix R und eine Translation C besteht.
* Des Weiteren beinhaltet P noch die intrinsischen Kameraparameter, welche in der Matrix K zusammengefasst werden.
* Diese beinhaltet die Brennweite zeta und wird in der Matrix K\_0 fesgehalten.

(Anhand des bildes erklären, danach nächstes Bild)

* Matrix T beinhaltet eine Transformation der Matrixeinheiten in die Pixeleinheiten des Sensorkoordiantensystemsmit k\_x und k\_y und die Verschiebung des Ursprungs in eine der Ecken
* Zusammen ergeben sie die Kameramatrix K
* ÜBERLEITUNG:
* Somit ist also geklärt wie ein Punkt im Raum auf einen Sensor projiziert wird.
* In unserem Beispiel wird ein Punkt in zwei Kameras projiziert.
* Die Aufgabe der Szenenrekonstruktion besteht dann darin, aus den zwei bildpunkten den gemeinsamen 3D-Objektpunkt rück zu projizieren. Hierzu bedienen wir uns der sogenannten Epipolargeometrie

**Folie 7: (227) 2 min (8min insgesammt)**

* Die Epipolargeometrie beschreibt die Beziehung von Bildpunkten und Objektpunkten in Stereoskopischen Systemen
* Wir sehen den 3D-Objektpunkt M welcher auf die Punkte m\_tau und m‘\_tau‘ auf die entsprechenden Bildebenen projitziert wird
* Man nennt m\_tau und m‘\_tau‘ auch zueinander korrespondierende Punkte, da sie den selben Objektpunkt abbilden
* C und C‘ sind die jeweiligen Projektionszentren, sie werden durch die sogenannten Basislinie verbunden
* Die Punkte an welcher die Basislinie die Bildebenen schneidet werden als Epipole e und e‘ bezeichnet. Sie sind die jeweiligen Abbildungen der Projektionszentren in den jeweiligen Kameras
* Durch die Epipole verlaufen alle Epipolarlinien l und l‘ eines Bildes
* Die Epipolarlinien enstehen indem eine Gerade durch einen Bildpunkt und den Epipol e gezogen wird
* Die Vekoren M,C und C‘ bilden die sogenannte Epipolarebene
* Die Ebene fässt die geometrischen Bedingungen einer Abbildung eines Objektpunktes in zwei Kameras zusammen. (Schnittpunkt epipolarleinie etc)
* NÄCHSTES BILD
* Die Abbildungsvorschrift der Epipolargeometrie besagt folgendendes:
* ein Punkt auf einer Bildebene wird auf eine Linie auf der anderen Bildebene abgebildet
* Sprich: Ein Bildpunkt mi auf der Bildebene I und C bilden eine Gerade.
* Alle Punkte auf der Geraden sind mögliche Ursprungspunkte von mi. Dies ist durch die drei möglichen Punkte M1,M2, M3 dargestellt.
* Jeder dieser Punkte wird nun wiederum auf I‘ projiziert.
* Die so entstandenen Punkte liegen alle auf der zu mi korrespondierenden Epipolarlinie l‘.
* Die Abbildungsvorschrift ist dementsprechend nicht eindeutig
* Diese Abbildungsvorschrift wird in der Fundamentalmatrix und der essentiellen Matrix ausgedrückt.

**Folie 8: (121) 1 min (9.30 min insgesammt) (hier war ich schon bei 12 min ....)**

* die Fundamentalmatrix F ist eine singuläre 3x3-Matrix.
  + Man kann grob mathematisch behaupten dass die Projektionsmatrizen P und P‘ beider Kameras in F vereint sind.
* Die essentielle Matrix ist auch eine 3x3 Matrix, sie entsteht indem die intrinsischen Kameraparameter aus der Fundementalmatrix rausgezogen werden.
* Die in rot eingerahmten Gleichungen auf der Folie werden als die sogennannten epipolaren Bedinungen bezeichnet
  + Sie geben Auskuft darüber ob die Punkte m und m‘ zweier Bildebenen die Abbildungsvorschrift erfüllen 🡪 Befindet sich m‘ auf der zu m korrespondierenden Epipolarlinie, so ist das Ergebnis dieser Bedingungen null

Die Vektoren welche F und F^T Null werden lassen bilden die Epipole.

Die Gleichungen unten rechts sagen aus das wenn F bekannt ist die jeweiligen Eipolarlinien zu Bildpunkten bestimmt werden können.

**Folie 9: (55) 1 min (10.30 min insgesammt)**

* Aufgabe war es nun einen Szenenrekonstruktionsalgorithmus zu zu implementieren, welcher aus zwei gegebenen Bildpunkten eine Rückprojektion zum 3D-Objektpunkt macht.
* Gleichzeitig soll der Algorithmus die extrinsischen Kameraparameter ausgeben
* Die intrinsischen Kameraparameter werden als bekannt vorrausgesetzt
* Der Algorithmus wurde zuerst anhand eines synthetischen Beispiels implementiert
* Der Quader wurde mathematisch auf die Bildebenen zweier beliebig platzierter Kameras projiziert

**Folie 10: (144) ca 1.30 min (12. 00 min insgesamt)**

* Eine Szenenrekonstruktion ist in 3 Bereiche eingeteilt
* Die Bildaufnahme:
  + Wie eben beschrieben wurde
  + Durch diese Berechnung sind extrinsische und intrinssiche Kameraparameter bereits bekannt
* Als zweites folgt die Bildverarbeitung
  + Die jeweiligen korrespondierenden Punkte sind durch die Abbildungsberechnung bekannt
* Bildanalyse:
  + Hier beginnt der eigentliche Algorithmus
  + Zunächst muss die Abbidlungsvorschrift die Fundamentalmatrix bestimmt werden
    - Dies geschieht über den sogenannten Acht-Punkte-Algorithmus
    - Es müssen insgesammt 8 oder mehr korrespondierende Punktepaare bekannt sein
    - Die Fundamentalmatrix wird dann über ein Gleichungssystem aus diesen 8 oder mehr Punkten berechnet
    - Diese Gleichungen werden in eine Koeffizientenmatrix eingetragen
    - Handelt es sich um ein nicht überbestimmtes System so wird der jenige Vektor gesucht, welcher mit der die Koeffizientenmatrix verrechnet 0 ergibt
    - Bei überbestimmten Fällen wird derjenige Vektor gesucht, welcher mit der Koeffizientenmatrix verrechnet minimal ist.
    - Diese 9-Vektoren ergeben die Einträge der Fundamentalmatrix
* Für das Verfahren zur Bestimmung der extrinsischen Kameraparameter wird die essentielle Matrix benötigt

(Nächste Folie einblenden)

* **Folie 11: (200) 2 min (14.00 min insgesammt)**
* Da die essentielle Matrix nur die extrinsischen Kameraparameter und keine intrinsischen mehr beinhaltet bietet sich diese an
  + Die extrinsischen Kameraparameter werden bezüglich einer Kamera zur anderen bestimmt
* Bei der Bestimmung der extrinssichen Kameraparameter durch Zerlegung der essentielle Matrix entstehen insgesammt 4 verschiedene mögliche Lösungen für die extrinsischen Kamerparameter
* Diese Vier Lösungen sind bis auf eine Skaleninvarianz genau bestimmt
  + Die Skaleninvarianz bedeutet, dass es bei der Rekonstruktion die Größe der Objekte von ihrer Originalgröße abweichen, da die translationskomponenten gleich dem normierten Richtungsvektor ist
  + (Vllt hier kurz das Geogebramodell aufzeigen)(30 sec nicht mehr!!)
* In den ersten beiden Lösungen wird C‘ um 180° gedreht
* In den anderen beiden Lösungen kommt es zu einer Umkehrung der Basislinie
* Die Richtige Lösung ist abhängig davon wie die Kameras Positioniert und die Bildebene innerhalb der Kamera platziert ist
  + Da die extrinsischen Kameraparameter bekannt sind konnte die Lösung abgeglichen werden.

(Arbeitsprozess einblenden)

* Sind die extrinsischen Kameraparameter bekannt, so fehlt nun noch ein Verfahren, mit welchem die 3D-Szenenpunkte aus den Bildpunkten rekonstruiert werden können.

(Letztes bild einblenden)

* Diese Verfahren nennen sich Triangulation. Im synthetischen Beispiel wurde eine einfache gemetrische Trinangulation durchgeführt.
* Sprich es wurden die Punkte alle in das gleiche Koordinatensystem überführt und danach wurden Geraden durch die Projektionszentren und und die Bildpunkte gelegt. Die Schnittpunkte der Geraden der korrespondierenden Punkte ergaben den Objektpunk.

**Folie 12: (203) 2.30 min (16.30 min insgesamt)**

* Der entwickelte Szenenrekonstruktionsalgorithmus wurde dann auf ein reales Beispiel angewandt
* Es wurde eine Szene aufgebaut und zwei Kameras wurden vor die Szene platziert.

Als erstes kommt die Bildaufnahme:

* Die beiden Kameras wurden in einem externen Programm einzeln kalibriert um so die intrinsnschen Kameraparameter zu erhalten.
* Danach wurden Bilder der Szene aufgenommen

Bildverarbeitung:

* Für die Detektion von markanten Punkten in beiden Bildern wurde ein bereits vorhandener Algorithmus verwendet. Das Ergebnis sind die in der Stereoaufnahme gelb bezifferten Stellen.
* Somit sind korrespondierende Punkte gefunden

Einschub Bildfehler:

* Anders als beim synthetischen Beispiel muss beim arbeiten mit realen Stereobildpaaren mit einer Fehleranfälligkeit der Bilddaten gerechnet werden
* Es muss von Bildfehlern wie Bildrauschen ausgegangen werden und damiteinhergehend auch mit Fehlern bei der eingägnlichen Korrespondenzanalyse (auf Fehler hinweisen im Bild)
* Liegt beispielsweise bei der Bestimmung von korrespondierenden Punkten eine detektierte Ecke zwischen zwei Pixel, so kann optisches Rauschen, dazu führen dass diese Ecke in gleichen Bildern an verschiedenen Pixel erkannt wird.
  + Wie im Bild darunter zu sehen ist, bedeutet das, dass die detektierten korrespondierenden Punkte nicht auf den jeweiligen epipolarlinien des anderen Bildes liegen.

🡪 Das führt dazu, dass die epipolaren Bedinungen nicht mehr erfüllt werden (FORMEL ERSCHEINEN LASSEN)

* diese ungenauigkeiten können auswirkungen auf die zwischenresultate des Algorithmus haben
  + Diese müssen aufgespürt und so minimal wie möglich gehalten werden.

**Folie 13: (99) 1 min (17.30 min insgesamt)**

* Zurück zum Algorithmus
* Bei der Bildanalyse startet der Szenenenrekonstruktionsalgorithmus
* Der Unterschied zum synthetischen Beispiel ist, dass die eingehenden Bilddaten aus der Korrespondenzanalyse normiert werden.
  + Die eingehenden Bilddaten werden so normiert, dass sich ihr durschnittlicher Abstand zum Ursprung bei sqrt(2) befindet. (Am Bild zeigen)
  + Die Normierung der Bilddaten führt dazu, dass sich die Bilddaten in einem kleineren Zahlenbereich befinden als zuvor, was dazu führt, dass die Fundamentalmatrix weniger stark auf Fehler in den korrespondierenden Punkten reagiert und somit der Algorithmus im allgemeinen stabiler läuft
  + (Ist die Fundamentalmatrix aus den normierten Koordinaten bestimmt worden, wird sie wieder denormalisiert)
  + Bild wieder ausblenden !!!
* Bevor die extrinsischen Kameraparameter bestimmt werden können muss überprüft werden, ob die Fundamentalmatrix auch Singulär ist

**Folie 14: (139) 1.15 min (18.45 min insgesamt)**

* F ist nur dann eine gültige F wenn sie eine Singuläre Matrix
* Durch die Ungenauen Korrespondierenen Punkten ist das meist nicht der Fall
* Das Resultat ist, dass Aus dieser F keine eindeutigen Epipole bestimmt werden können was dazu führt, dass die Epipolarlinien durch keinen gemeinsamen Punkt gehen.
* Da aber spätestens für das hier benutzte Verfahren der Triangulierung eindeutige Epipole gefordert werden, wird F so manipuliert, dass sie wieder eine Singuläre Matrix .
* Das erbenis ist in der Abbildung zu sehen, die Epipolarlinien verlaufen wieder durch einen gemeinsamem Epipol
* Aus der Fundamentalmatrix wird dann wieder die essentielle Matrix abgeleitet
  + Ist die Fundamentalmatrix Singulär ist auch die essentielle Matrix singukär sofern sie aus F berechnet wurde.
  + Aomit können wie im synthetischen Beispiel die extrinssichen Kameraparameter bestimmt werden.
* Zuletzt folgt eine Triangulierung der Szenenpunkte auf welches ich gleich noch zu sprechen komme.

**Folie 15: (182)**

* Als letztes folgt noch die Triangulierung.
* Aufgrund der Ungeanuigkeit der korrespondierenden Punkte ist es nicht möglich durch einfache Rückprojektion die 3D-Objektpunkte zu rekonstruieren.
* Liegen die korrespondierenden Bildpunkte nicht direkt auf den jeweiligen Epipolarlinien der beiden Punkte, so ist die Epipolare Bedingung nicht erfüllt.
  + Die Rückprijezierten Punkte treffen sich nicht im Raum sondern liegen windschief zueinander im Raum

(NÄCHSTES BILD EINBLENDEN)

* Um die Triangulation der Punkte trotzdem durchführen zu können werden zwei Punkte m und m‘ gesucht, welche möglichst nahe an den Ursprünglichen Punkten liegen und gleichzeitig die Epipolaren Bedingugnen erfüllen.
* Der nächste Punkt zu m und m‘ liegt jeweils am Fuße des Lots auf der entsprechenden Epipolarlinie.
* Jedes Paar l und l‘ würde das erfüllen aber es werden die gesucht die d minimiert
* Die Minimierung wird über die sogennaten Samspson-Approximation vollzogen.
* Für dieses Verfahren muss F bekannt sein und die Epipole eine eindeutige Lösung erbringen
  + (Umriss des Verfahrens in die Zusatzfolien)
* Die neuen Punkte, welche die epipolare Bedingung wieder erfüllen, können dann über ein Triangulierungsverfahren nach Wahl Rückprojiziert werden
  + Im Algorithmus wurde ein lineares Triangulierungsverfahren nach Hartley and Zisserman verwendet.

Im fogenden wurde der Algororithmus noch auf Bilder unterschiedlicher Auflösung angewandt

**Folie 16: (96)**

* Was bedeuteten unterschiedliche Auflösungen?
* Ein Sensor hat eine maximale Auflösung.
  + Die maximale Anzahl der Sensorelemente auf einem Sensor beschreibt die maximale Auflösung.
* Die Anzahl und Größe der Einzelnen Sensorelemente variiert mit den Größen der Sensorchips
* Wird eine Auflösung kleiner der maximalen Auflösung eingestellt, desto geringer wird die Anzahl der Pixel.
* Der Prozess, welcher hier stattndet, gehört zu den Nachbarschaftsoperationen.
* Eine Veranderung der Auosung kann auch eine Anderung der Seitenverhaltnisse mit einschließen. Andert sich das Seitenverhaltnis so wird der Bereich der lichtempndlichen Flache auf dem Sensor beschrankt
* Wird nun die Auflösung einer Kamera verändert, welche Auswirkungen hat das auf den Szenenrekonstruktionsalgorithmus???

**Folie 17: (182)**

* Zunächst kann festgehalten werden, dass eine Veränderung der Auflösung eine Auswirkung auf die Skalierung der Sensorkoordinaten hat, da sich das Sensorkoordintansystem an der Beschaffenheit der Sensorelemente orientiert.
* Alle anderen Koordinatensysteme bleiben gleich
* Mit der Auflösung ändert sich die Anzahl und die Größe der Pixel
* Der Ort auf dem Sensor an welchem der bildpunkt abgebildet wird bleibt jedoch der selbe (Hier zweites Bild einblenden.)
* Eine Skalierung des Sensorkoordinatensystems bedeutet, dass sich die Brennweite in Pixeleinheiten gegeben ändert, jedoch ändert sich nicht die effektive Brennweite in Millimeter.
* Eine propotionale änderung der Kamerauflösung hat zur Folge, dass es so wirkt als wäre die Brennweite verdoppelt worden.
  + Würde bedeuten, dass sich die Kamera von der Bildebene entfernt hat🡪 hat sie aber effektiv nicht 🡪 dennoch verändert sich durch die skalierung der Pixel effektiv die Bildgröße
* Bei der Bestimmung der extrinsischen Kameraparameter wird mit der essentiellen Matrix gearbeitet, da hier die Kameramatrizen K und K‘ rausfallen, haben diese keine Wikrung auf das Ergebnis

**(Folie 18: (32)**

* Um die Aufgestellte Therorie zu überprüfen, wurden die Kameramatrizen modifiziert
* K fliegt immer raus und R äbndert sich effektiv
* Daraus folgt dass auch die Rekonstruierte Szene immer die gleichen Ergebnisse geliefert hat.

**Folie 19: (188) (Ausblick)**

* Im folgenden werden noch einige Ausblicke vorgestellt.
* Diese beinhalten kurz die Algorithmen, welche neben dem Hauptalgorithmus noch entstanden sind
* Viele Programme setzten ihr Hauptaugenmerk auf eine schnelle und effiziente Szenenrekonstruktion ohne Kamerakalibrierungs Zwischenschritt
* Rektifizierte Bilder zeichnen sich durch ins unendliche projizierte Epipole aus
* Sind die Epipole ins unendliche projiziert so sind die Epipolarlinien parallel zueinander
* Die Epipolarlinien werden somit zu einheitlichen Scanlinien, welche die erweiterete korrespondenzanalyse auf eine eindimensionale Suche entlang dieser Linie beschränkt

(Nächstes Bild)

* Der horizintale Versatz zwischen den korrespondierenden Punkten gibt ein Maß für die Tiefe des jeweiligen Punktes
* Anhand dieser Tiefe kann eine Tiefenkarte erstellt werden
* Die meisten Rektifizierungsalgorithmen setzten zwei bilder gleicher Auflösung vorraus
* Um dem entgegenzuwirken wurde ein Rektifizierungsalgorithmus implementiert welcher unterschiedliche Kamerauflösugnen nicht von vorn hinein ausschließt und es wurde analysiert was die Auswirkungen sind
* Diese Beispiele zeigen, dass das hier vorgestellte Rektizierungsverfahren auf Bilder verschiedener Auosungen mit dem selben Seitenverhaltnis der Pixel anwendbar ist. Jedoch konnen Bilder mit verschiedenen Seitenverhaltnissen mit diesem Verfahren nicht rektiziert werden. Eine mogliche Losung ware eine Erweiterung des Models, welche die Bildlangen bestimmt und schlielich die relativen Seitenverh altnisse ausgleicht.
* Dies gilt weiter zu verfolgen

**Folie 21:**

* Des Weiteren gilt es noch eine eigene einzelkamerakalibrierung zu implementieren
* Hierzu soll ein ebenfalls in der Thesis entwickelter Algorithmus zur Sortierung von Schachbretteckpuntken helfen, welcher auch bei stark verzerrten Schachbrettern funktioniert.