Folie 1:

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

Das Ziel dieser Programme ist es Maschine die Fähigkeit beizubringen räumliche objekte erkennen zu können und anhand dieser Entscheidungen zu treffen

Folie 2:

Das entwickeln solcher Programme ist Forschungsmittelpunkte von aktuellen Anwendungsbereichen wie dem Autonomen Fahren, Motion- Capturing, Bewegungserkennungen oder Service Robotern oder Aufnahmen von Drohnen mit zwei KAmeras

(Nächstes Bild)

Ein weitläufiges Ziel dieser Arbeit ist es mit Kameras unterschiedlicher Art und Auflösung eine Szenenrekonstruktion zu ermöglichen

Einsatz könnte so ein Programm besipielsweise bei Drohnen aufnahmen finden.

Diese sind häufig mit zwei unterschiedlichen Kameras ausgestattet, beispielsweise RGB und infrarot

Eine Szenenrekonstruktion aus beiden Bildern könnte die anschaffung teurer Hardware wie hochauflösende infrarot kameras, vermeiden

Folie 3:

Die Gliederung des Talks:

Grundlagen:

Einführen in Kameramodell und Fundamental und essentielle Matirix

Synthetische Rekonstruktion:

Aufbau des Algoroithmus an einem selbst erstellen Beispiels.

Reale Rekonstruktion:

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

Unterschiedlicher Auflösungen:

Danach wurde getestet, welche Auswirkungen unterschiedliche Kameraauflösungen auf den Algorithmus haben

Folie 4:

Zunächst zu den Grundlagen:

Zuerst wird geklärt wie ein 3D-Punkt im raum auf einen 2D-Punkt auf dem Sensor abgebildet wird

Vereinfacht gesprochen entsteht das Abbild eines der abgebildeten Eckpunkte indem eine Verbindungsgerade von eben diesem Punkt mit dem Kamerazetrum C gebildet wird

Der entstehende Schnittpunkt mit der Bildebene entspricht dem Bildpunkt.

Auf der Bildebene gibts es zwei verschiedene Koordinatensysteme, dass Bildebenenkoordinatensystem I in mm und das Sensorkoordinatensystem S in Pixel

Ein Bildpunkt wird in der Regel in Sensorkoordinaten gegeben

Die Projektion eines 3D-Objektpunktes im Raum in einen 2D Punkt auf den Sensor der Kamera wird durch die sogenannten Projektionsmatrix P beschrieben

Folie 5:

Ein Obkjekpunkt M wird mit P auf m projiziert

Die Projektionsmatrix P setzt sich zusammen aus den extrinsischen Kameraparametern welche durch eine Rotationsmatrix R und eine Verschiebung C des Kamerazentrums besteht.

Des Weiteren beinhaltet P noch die intrinsischen Kameraparameter, welche in der Matrix K zusammengefasst werden.

Diese beinhaltet die Brennweite, die skalierung der Koordinaten von mm in Pixel und die Verschiebung des Koordinatenursprungs in eine Ecke

ÜBERLEITUNG:

Somit ist also geklärt wie ein Punkt im Raum auf einen Sensor projiziert wird.

(Die Rückprojektion eines Bildpunktes zu einem 3D-Objektpunkt ist jedoch nicht eindeutig, da die Tiefe ab der Bildebene nicht mehr bekannt ist (Anhand des Bildes erklären))

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 6:

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 diese Punkte auch zueinander korrespondierende Punkte

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

Diese fässt die Abbildungsvorschrift zusammen

NÄCHSTES BILD

Die Abbildungsvorschrift der Epipolargeometrie besagt folgendendes:

ein Punkt auf einer Bildebene wird auf eine Linie auf der anderen Bildebene abgebildet

Was bedeutett das? 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 7:

Die Fundamentalmatrix beschreibt die Abbildungsvorschrift, dass bedeutet es gilt dass wenn mF = l‘ und F^T = l

Des weiteren stellen die Vektoren welche F und F^T Null werden lassen die Epipole dar.

die Fundamentalmatrix F ist eine singuläre 3x3-Matrix.

Die informationen von P und P‘ sind in F enthalten.

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

Folie 8:

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

Ein Quader wurde mathematisch auf die Bildebenen zweier beliebig platzierter Kameras projiziert

Für die Rekonstruktion muss nun die bekannte projektion auf die bildebenen rückwärts gemacht werden.

Das synthetische Beispiel wurde wie im Beispielbild gezeigt aufgebaut

Folie 9:

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

Die Fundamentalmatrix besitzt 9 unbekannte

Aufgrund der beschaffenheit der Bildpunkte, reichen bereits 8 Punkte um die unbekannten zu Bestimmen

Für das Verfahren zur Bestimmung der extrinsischen Kameraparameter wird die essentielle Matrix benötigt

(Nächste Folie einblenden)

Folie 10:

Für die Bestimmung der extrinsischen Kameraparameter werden die intrinsischen Kameraparameter aus F rausgezogen

Da die essentielle Matrix nur die extrinsischen Kameraparameter und keine intrinsischen mehr beinhaltet bietet sich diese für die Bestimmung der extrinsischen Kameraparameter an

Bei der Bestimmung der extrinssichen Kameraparameter durch Zerlegung der essentielle Matrix entstehen insgesammt 4 verschiedene mögliche Lösungen für die extrinsischen Kamerparameter

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.

Somit konnte getestet werden ob die Funktion der Bestimmung stimmt

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

(Arbeitsprozess einblenden)

Sind die extrinsischen Kameraparameter bekannt, so fehlt nun noch der letzte Schritt, in 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 11:

Der entwickelte Szenenrekonstruktionsalgorithmus wurde dann auf ein reales Beispiel angewandt

Es wurde eine Szene aufgebaut und zwei Kameras wurden vor die Szene platziert.

(Nächstes Bild)

Der Ablauf des Algorithmus beginnt wie im synthetischen Beispiel mit der Bildaufnahme

Hierzu gehört unter anderem die Bestimmung der intrinsischen Kameraparameter beider Kameras in einem externen Programm

Danach kommt die Bildverarbeitung:

Hier werden korrespondierende Punkte in beiden Bildern gesucht

In der Bildanalyse kommt es zu eigentlichen Szenenrekonstruktion

Dazu gehören unter anderem die Bestimmung der Fundmanetalmatrix und der essentiellen Matrix, wie auch die Bestimmung der extrinsischen Kameraparameter

Zuletzt folgt dann noch die Triangulation der 3D-Objektpunkte

Folie 12:

Nun gehen wir nocheinmal genauer auf die einzelnen Schritte ein und fangen an bei der 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:

Hier ist wichtig zu beachten , dass anders als beim synthetischen Beispiel beim arbeiten mit realen Stereobildpaaren mit einer Fehleranfälligkeit der Bilddaten gerechnet werden muss

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.

Sind die Bilder aufgenommen und die korrespondierenden Punkte bestimmt, so muss für die Bestimmung der Fundamentalmatrix die Bildpunkte zunächst normiert werden

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

Nun folgt die Bildanalyse:

Die Fundamentalmatrix wird aus den normierten Koordinaten bestimmt und danach wird die Matrix denormalisiert.

Bevor jedoch die extrinsischen Kameraparameter bestimmt werden können muss überprüft werden, ob die Fundamentalmatrix auch Singulär ist

Folie 13:

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.

Um eine richtige essentielle Matrix aus F zu Berechnen und richtige Resultate bei der rekonstruktion der extrinsischen Kameraparameter zu bekommen muss die Singularität von F durch dessen mdoifizierung erzwungen werden.

Das erbenis ist in der Abbildung zu sehen, die Epipolarlinien verlaufen wieder durch einen gemeinsamem Epipol

Wichtig: Ist die Fundamentalmatrix Singulär ist auch die essentielle Matrix singulär sofern sie aus F berechnet wurde.

Somit 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 14:

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.

Die neuen Punkte, welche die epipolare Bedingung wieder erfüllen, können dann über ein Triangulierungsverfahren nach Wahl Rückprojiziert werden

Folie 15:

Um einen Sensor mit einer schlechteren Kamera zu simulieren wurde die Auflösung einer der Kameras runtergestellt.

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.

Folie 16:

Zunächst kann festgehalten werden, dass eine Veränderung der Auflösung eine Auswirkung auf die Skalierung der Sensorkoordinaten hat, da sich die Skalierung der Achseineinheiten an den größen der Pixel orientiert.

Am Bildebenenkoordinatensystem ändert sich effektiv nichts

Der Ort auf dem Sensor an welchem der bildpunkt abgebildet wird bleibt der selbe (Hier zweites Bild einblenden.)

Eine Änderung der Kameraauflösung bewirkt eine Veränderung der resultierenden Bildgröße oder Auschnitts.

Das liegt daran, dass sich die Brennweite in Pixel gegeben mit der skalierung ändert, an der effektiven Brennweite ändert sich jedoch nichts

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

Folie 17:

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

Die Triangulation wird unabhängig von den intrinsischen Kameraparametern durchgefürt

Unterschiedliche K haben keine Auswirkung

Folie 18:

Im folgenden werden noch einige Ausblicke vorgestellt.

Diese beinhalten kurz die Algorithmen, welche neben dem Hauptalgorithmus noch entstanden sind

Zunächst die Rekonstruktion mit Rektifizierung

Viele Programme setzten ihr Hauptaugenmerk auf eine schnelle und effiziente Szenenrekonstruktion ohne Kamerakalibrierungs Zwischenschritt

Diese setzten auf eine voran geschaltene Rektifizierung der Bilder

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 dieses Versatzes 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

Folie 19:

Erklären was man sieht

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 20:

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.

Der Algorithmus basiert darauf, dass ein Startpunkt in einer Ecke des Schachbretts bestimmt wird und mit Hilfe von Richtungsvektoren werden die folgenden Nachbarpunkte gesucht und der Reihenfolge nach mit indizes versehen.