

VO 3D Vision Zusammenfassung SS05

1. Einleitung

Möglichkeiten aus 2D Bildern 3D Daten zu bekommen:

- Aus 2 oder mehr Bildern: Stereo Range Finder (nicht immer anwendbar, nicht echte CV), Fusion von 3D Daten, Motion
- Objektannahmen (a priori Wissen): Shape from Shading, Shape from Texture

Ziele von CV

- Automation
- Inspektion
- Fernerkundung
- Mensch-Maschine Schnittstelle
- Hilfestellung für Behinderte

2. Bildaufnahme

Lochkamera - Perspektivische Projektion

- Nicht linear!
- Häufig: Modellierung mit homogenen Koordinaten $(X, Y, Z, W) \rightarrow (wx, wy, wz, w)$
- Punkt wird durch eine Linie im 4D Raum durch den Ursprung abgebildet
- Oder: $x = f/Z * X$ $y = f/Z * Y$
- Parallele Linien treffen sich im Fluchtpunkt der Richtung
- Winkel bleiben nicht erhalten

→ sehr genau für Szenen

→ used in structure from Motion

Orthographische- od. Normalprojektion

- Jeder 3D Punkt wird durch einen Strahl normal zur Bildebene abgebildet
- $x = X, y = Y$

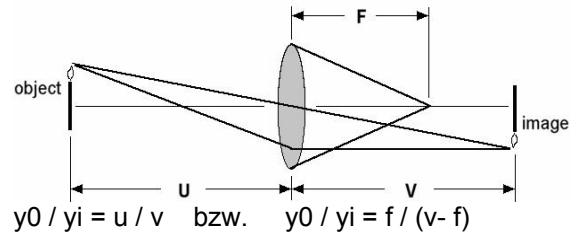
weak perspective (scaled orthographic projection)

- nur perspektivischer Effekt
- 3D Punkte in Gruppen ähnlicher depth; dann durch Gruppendepth dividieren
- $(x, y, z) \rightarrow s(x, y) \dots s$ konstant
- Parallele Linien bleiben parallel

→ einfache Mathematik, für kleine und nahe Objekte

→ für recognition

Simple Lens Parameters



- principal axis
- front/back focal length (ab Krümmung)
- focal length (ab axis)
- optical axis

Tiefenschärfe – Apertur

Apertur: etwa: Öffnung der Linse, (abhängig von Dicke und Krümmung), meist kreisförmig.

- je größer die Apertur, desto größer die unscharfen Kreise außerhalb des Tiefenschärfebereichs
- **je kleiner die A., desto schärfer das Gesamtbild.**
- Je mehr Tiefenschärfe, desto weniger Lichteinfall.
- Größere Apertur → geringerer Tiefenschärfebereich
- Verändern der Fokallänge verändert die Apertur

Tiefenschärfe: Überfokussierung – Fokussierung – Unterfokussierung

Begrenzung der Lichtstrahlen: Blende

Blendenzahl $k = f / D$ (Linsenhöhe)

Radianz - Menge des Lichts, das von einem Punkt reflektiert wird.

Irradianz - Menge des Lichts das auf diesen Punkt fällt.

Radiosity – total emittierende Energie einer Oberfläche pro Einheit (Lambertian = gleichmäßig nach Winkel, diffuse Objekte (Samt) wahllos, Spiegel - Normalreflexion)

Ladungsmodelle

CCD – Kameras arbeiten integrativ

CMOS Kameras nichtlinear

Farbe

- 3 CCD: Prisma + 3 Filter
- 1 CCD: Farbfilter am Chip streifenweise (nur $\frac{1}{3}$ horiz Auflösung)
- Mono: 3 Bilder, schwer ausrichtbar: affine Trans - Unschärfephänomene
- RGB oder YUV

3. Kalibrierung

Pinhole Camera Model

- „Klassisch“: 2D \leftrightarrow 3D Correspondence
- „Selbstkalibrierung“: 2D \leftrightarrow 2D kalibrierung

Kamerakalibrierung

korrekte 3D Info aus 2D – Relation Objekt- und Bildpunkt notwendig.

Geometrische Kalibrierung: Welt- und Bildkoordinatensystem, Eichkörper, Projektionszentrum, Skalierung in x und y Richtung.

Homogene Koordinaten erlauben die Projektion als lineare Transformation.

Kameraparameter

Extrinsische Parameter beschreiben Ort und Orientierung des Kamera Referenzrahmens zum „Weltraum“.

- 3 Eulersche Winkel: yaw, pitch, tilt
- 3 Translationsvektorkomponenten (verschiebungsvektor Objektkoordinatensys durch Proj.zentrum)

Intrinistische Parameter beschreiben Pixel Koordinaten eines Bildpunktes in Bezug auf Koordinaten im Kamera Referenzrahmen.

- Fokale Länge f: Abstand Bildebene – Projektionszentrum
- Verzerrungskoeffizient k, Linsenverz.
- Skalierungsfaktor s, Abtastfaktor in x-Richtung (Kamera AD Wandler)
- Bildhauptpunkt C_x, C_y . Schnittpunkt optische Achse mit Bildebene

Geometrische Abweichungen

- sphärische Abweichung
- Verzerrung
- Verkrümmung
- Chromatische Abweichung (Prisma)

Können durch mehrere kombinierte Linsen reduziert werden.

Linsenverzerrungen

- **radiale Verzerrung**: Verschieben eines Punkte radial zum Mittelpunkt des Bildhauptpunktes (Tonnen- bzw. Kissenverzerrung)
→ radiale Kalibrierung
- **tangentiale Verzerrung**: Verschiebung tangential zur radialen Verzerr. (meist vernachlässigt)

Bestimmung des Bildhauptpunktes

- rechnerische Modellanpassung (radiale Verzerrung)
- direkte optische Methode: Laser wird durch Linse auf Sensor projiziert. Laser sehr genau, Ausrichtungsproblem

- Variation der fokalen Länge
 - Variation der Bildebene
 - Verwendung zweier Linsen
 pro: leicht zu berechnen
con: ungenau durch Verwackeln

lineare Kamerakalibrierung mit Target

- Calibration target: auf Tsai Grid (Schachbrett im rechten Winkel)
- Ecken bzw. Koordinaten bekannt
- Werden im Bild gesucht
- mit Formeln (Koos \leftrightarrow Koos) bekommen wir die Kameraparameter

lineare Kamkalibr. mit Image Processing

- Canny Edge Detection
- Straight line fitting für die Ecken
- Intersections für Ecken
- Matchen der Bild- und Schachbrettkoordinaten

Nonlineare Methoden

Probleme linearer Kalibr.: Keine Linsenverzerrung und viele freie Parameter

Tsai: parametrisiere \square in Bezug auf Rotation, Translation, Intrinistische Params und Linsenverzerrungen.

Kalibrierung aus 2D Bewegung

- Structure from Motion
Track Points in verschiedenen, bekannten 3D Positionen der Kamera, intrinistische Params. vorher zu klären
- Self Calibration
Lösung für alles

Grauwertkalibrierung (radiometrische)

Graue Fläche aufnehmen → Koeffizienten für jeden Pixel

verschiedene Grautöne → Gammakorrektur

4. Merkmalsextraktion

Kanten

Edgel, Edgels

Diskontinuitäten der Helligkeit können Diskontin. der Oberflächennormale, der Tiefe, der Textur und der Beleuchtung bedeuten.

Kantenoperatoren

- Roberts
- Prewitt
- Sobel
- Marr- Hildreth Operator (Mexican Hat)
- Canny

Glättung

Mittelwert, Rauschunterdrückung, hohe Bildfrequenzen werden eliminiert.

5. Active Range Finder

Laufzeit Range Finder

- Optische Range Finder
- Ultraschall Range Finder

Triangulations Range Finder

- Spot Projektoren
- Moiré Range Finder
- Lichtschnitt Range Finder
- Streifenprojektoren

Vorteile

- kein Kontakt
- schnell 100 -10 M points/s
- texturunabhängig
- lichtunabhängig
- BYOL „bring your own light“
- BYOT „brin your own targets“

Herausforderungen

- Lasers speckles
- Design an Objektgröße angepasst
- endliche Projektor Power
- CCD Kameras für Bilder, nicht 3D Messungen gebaut
- Objektmaterial interagiert (ev.)

time based – Laufzeit Range Finder

Laser Radar System
zB Riegel Scanner

Ultraschall (IR) Range Finder

für Kameras (Autofocus Spot Messungen)

Vorteil:

- beleuchtungsunabhängig
- langsame Geschwindigkeit des Messstrahles

Nachteile:

- schlechte Auflösung
- geringe Genauigkeit
- Ablenkeinheit notwendig

Spez: 5cm – 5m ± 3mm

Anwendungen Füllmessungen, Car Parking Radar, Unterwassermessungen

Triangulation

WSW Satz

$$Z = b \sin \alpha \sin \beta / \sin(180 - \alpha - \beta)$$

Pattern Projection

Muster statt Laserstrahl

Szene muss still sein

Richtung klar, Muster eindeutig („binary codes“)

Moiré Range Finder

Interferenz zweier Muster die übereinander projiziert werden. Projizieren einer Liniensstruktur und beobachten der Linienstruktur durch ein weiteres Gitter

- Projection Moiré
- Shadow Moiré
- Digital Moiré with Reference
- Multiple-Frame Phase-Shifted Moiré

Binärcodierte Lichtschnitttechnik

zeitlich aufeinanderfolgender, binary codierter Lichtmuster.

Vorteil: wesentlich weniger Aufnahmen für die gleiche Anzahl von zu vermessenden Objektpunkten ($\log n$). Jede Lichtebeine durch Grauwert zuordenbar.

Tiefenberechnung mittels Look up Table, für jede Lichtebeine wird die Triangulation vorausberechnet.

Bit Plane Stack – gespeichert, wann Pixel beleuchtet und wann nicht

Spezifikationen für 3D Scanner

- Genauigkeit
- Reproduzierbarkeit
- Unsicherheit und Vorhersagbarkeit
- Systematische Fehler
- Zufallsfehler
- Auflösung

6. Shape from Monocular Images

Shape from Shading SfS

Oberflächenreflexion von untexturierten Objekten enthalten Tiefeninformation.

Flächengrenzen spielen für die Interpretation aber auch große Rolle!

Formrekonstruktion

Normalprojektion, Objekt ist nahe der optischen Achse

Reflectance Map

- Reflexionseigenschaften wichtig für SFS
- Beleuchtung ist eine Funktion der Richtung, nicht der Entfernung
 - Beleuchtung weit entfernt
 - keine Zweitbeleuchtung
 - keine Schlagschatten
- Lambertsche Oberflächen: Helligkeit hängt nur von Beleuchtungsrichtung ab, nicht von Beobachtungswinkel
- → **2D Plot des Gradientenraums (p,q) der normalisierten Bildhelligkeit einer Oberfläche als Funktion der Oberflächenorientierung**
- Nachteil: keine komplett matten Oberflächen, perfekter Spiegel reflektiert nur in eine spezielle Richtung, sonst keine Reflexion. Realität: Kombination aus beiden → gewichteter Durchschnitt der diffusen und spiegelnden Komponenten einer Oberfläche.
- müssen experimentell bestimmt werden (→ Normform), allgemeines SfS nicht möglich.

Formbestimmung

- Streifenmethode
 - für jeden Helligkeitswert des Bildes
 - Refl. Map. schränkt Orientierung ein
 - Streifen gl. Helligkt. → Höhenlinien
 - Startpunkt mit bekannter Normale (zB höchste Helligkt, Steigung = 0)
 - kleine Bewegung in Richtung stärkste Helligkeitsänderung
 - kleine Bewegung in Richtung stärkste Steigung
 - bis alles von Isolinen erfasst ist
 - Nachteil: Fehler kumulieren, keine stabile Lösung
- Photometrisches Stereo
 - 2 Bilder, gleiche Geometrie, unterschiedl. Beleuchtungsrichtung.
 - 2 Kandidaten, Schnittpunkte im Gradientenraum = Normale

- eindeutige Lösung durch dritte Lichtquelle.
- praktische Anwendg.: farbiges Licht.
- Polarisiertes Licht
 - auch für spiegelnde Oberflächen
 - genaue Normale, aber nur ungefähre Position

Voraussetzung für SfS

- Oberflächen mit konstanter Albedo → rotationsinvariant
- Orthographische Projektion
- ferne und kalibrierte Beleuchtungsquellen
- keine Schlagschatten
- keine Reflexionsbeleuchtung - Interreflexion

Shape from Texture SfT

Wiederholung des Basismusters muss weder deterministisch noch regulär erfolgen, nur statistisch gleichverteilt.

Statistische Texturanalyse

- für natürliche Texturen
- eher für Klassifikation als zur Formbestimmung, nur Überprüfung, ob Muster da ist oder nicht

Strukturelle Texturanalyse

- für deterministische Texturen (unnatürliche)
- Texturelement TEXEL, Charakterisierung durch invariante Eigenschaften

Shape from Textel

- geht von Verzerrung der Textel aus
- Textel müssen eindeutig identifizierbar sein und nicht überlappen
- alle Textel haben die selbe räumliche Ausdehnung
- Textel sind klein, planar und eindeutige Oberflächennormale

Geometrische Szeneneigenschaften

Mensch: Parallele Linien begrenzen Fläche, Fluchtpunkt ermöglicht Orientierung

Shape from parallel Lines

Aggregationstransformation

8. Stereo Vision

Stereo Correspondence

Epipolare Anordnung (Kameras in einer Ebene (Epipolarebene)) ermöglicht 1D Problem. Suche entlang der **konjugierten Epipolarlinie**.

Mensch:

Random Dot Stereogramme**Autostereogramme**

Durch richtige Konvergenz der Augen werden verschiedene Ebenen erzeugt, Bildtiefen von Konvergenz abhängig

Disparität

$$-Z = f \cdot B / (x_1 - x_2)$$

„Abstand zum Linsenzentrum = (fokale Länge mal Abstand der Linsen) durch Bildpunktdifferenz“ [kein Minus!]

Jeder Punkt, der auf der einen Epipolarlinie liegt, muss auch auf der anderen liegen.

Registrierung der Stereobilder

- Transformation in eine Bildebene
- Epipolarlinien = Zeilen der Bilder
- normal: Kameras so kalibriert
pro: einfach zum Rechnen
contra: geringer Abstand Projektionszentren
- allgemein:
pro: je größer Abstand der Kameras, desto genauerer Messwert
contra: größere Verdeckungen (Abschattungen)

Einschränkungen

- Liniensegmente besitzen in beiden Bildern gleiche Orientierung und Länge
- korresp. Punkte besitzen gleiche Intensität und Farbwerte
- Disparity ändert sich stetig
- korresp. Punkte werden nur ab einem gewissen Abstand bestimmt.

Korrespondenzanalyse

- Flächenbasierte Auswertung: Intensitätsvergleich aller Pixel; Korr. für jedes Pixel
- Merkmalsbasierte Auswertung: Vergleich Merkmale, Korr. auf Grund von ausgewählten Eigenschaften; nicht alle Pixel. Genauer, da im Subpixelbereich. (Kanten)

Korrespondenzsuche

- Problem: Punkt existiert nicht oder ist uneindeutig.
Lösung: Schwellwert
- Hängt stark vom verwendeten Fenster ab.
- Adaptive Matching (variable Fenstergröße)
- Hierarchisches Matching (Gaußpyramiden, in levels)
- Multistage Matching (Multiresolution)

9. Motion

Voraussetzung: Wissen um Bewegungsrichtung und Geschwindigkeit.

Bewegungsfeld

Durch Vektoren gekennzeichnet, die die Bewegung der korrespondierenden Szenepunkte repräsentieren.

Kamera, die nicht rotiert:

- Vektoren zeigen radial zu oder von einem Fokus.
- **FOE: Focus of Expansion** (Vorwärts)
- **FOC: Focus of Contraction** (Rückwärts)
- Punkt, wo der Bewegungsvektor der Kamera die Bildebene schneidet.

Bewegungsvektor

- Länge ist verkehrt proportional zur Entfernung des Punktes
- proportional um Sinus zwischen Richtung des Punktes und Bewegungsrichtung
- Wenn Bewegungsrichtung bekannt; Distanz kann in Auflösung v (Skalierungsfaktor) bestimmt werden.
- unendliche Entfernung, keine Bewegung

Analogie zu Stereo: Basislinie – Bewegung, jede Projektion entlang Epipolarlinie.

Sonderfall: FoC und FoE sind im unendlichen; zB. parallele Vektoren (aus fahrendem Fenster schauen). → Rigid Body Annahme

Bewegungsfeldbestimmung

Problem: Bestimmung der korresp. Punkte

- dünn besetztes Vektorfeld
- wie Stereo, nur Bewegungsrichtung nicht bekannt
- Epipolarlinie zu Beginn nicht bekannt

Voraussetzung: hohe temporäre Abtastung = geringe Unterschiede und entweder unveränderte Intensitäten oder Kanten in beiden Bildern. (**Intensity Flow** (nur Lambert'sche Flächen) od. **Edge Flow**)

2 Strategien:

Spatiotemporäre Ableitung

Intensitäten ändern sich zeitlich nicht, Grauwertänderungen sind überall stetig differenzierbar

Spatiotemporäre Kohärenz

Intensität und Kanten bleiben erhalten

- Kante beschreibt Fläche
- Punkt – Kurve
- Fläche – Volumen

Apertur Problem

zB Kanten

Bewegungsbetrag nur in Richtung Gradienten, nicht in Richtung Kante, da Beobachtungsfenster begrenzt (Größe = Apertur).

[Man sieht nicht ganze Kante – zB Horizont]

10 Kombination & 3D Objektrepräsentation

Kombination der 3D Verfahren

mehr Daten, mehr Genauigkeit

Nachteile:

- Shading: nur relative Tiefenwerte, texturabhängig
- Stereo: texturabhängig, Rechenzeit
- Lichtschnitt: durch Anzahl der Linien begrenzt, kein bildhafte Information

Vorteile:

- Shading: Tiefeninfo für alle Punkte
- Stereo: Tiefeninfo entlang von Kanten genau
- Lichtschnitt: schnelles & genaues Tiefenraster

Kombinationen:

Lichtschnitt und Shading

„Gap – Filling“ inside Structured Light Information with Shading

Stereo und Lichtschnitt

„Gap – Filling“ inside Structured Light Information with Stereo

3D Beschreibung und Form Repräsentation

- Kantenmodell
- Flächenmodell
- Volumsmodell
- Parametermodell

Kantenrepräsentation - 3D Kurven

Splinestützpunkte bzw. Datenpunkte (bei Fehlern = Ausgleichskurve)

3D Oberflächendarstellung (Patches)

patches durch 4 Kurven begrenzt
oder: Dreiecksvermaschung

Symmetrieachsentransformation (SAT)

auch: Blum Transformation oder Medial Axis Transformation

Formprimitiv: Kreis 2D oder Kugel 3D

Objekt wird durch maximal eingeschriebene, logisch verknüpfte Kreise repräsentiert. Nachteil: kleine Änderung des Objekts, große Änderung der Repr.

Generalisierte Zylinder

Teilschnitt entlang Kurve. Cross-Sections ändern Teilschnitt. Aber: uneindeutige Repr. noch allgemeiner: Super-Quadrics (alle 2D Formen als Teilschnitt)

Visuelles Potential

Menge aller möglichen Ansichten in einem Graphen

Polyhedra

Trihedral Vertices

11. Anwendungen

Medizintechnik

Computertomographie

Schichtweise Aufnahme: Bestrahlung - Projektionsebene – CCD Kamera – digitales Bildsignal – Videorecorder

Erzeugung mehrerer Schichten aus 3 Einzelbildern möglich → weniger Bestrahlung als bei Einzelschichtaufnahme

Magnetresonanz

Alle Atomkerne ungerader Anz von Neutronen oder Protonen besitzen Spin → elektrische Ladung erzeugt Magnetfeld

Nuklearmedizin

Radioaktive Testlösung an Patienten verabreicht, Verteilung der radioaktiven Teilchen ermöglicht Studium einzelner Schichten

Bsp: radioaktiver Sauerstoff zur Überprüfung der Hirnaktivitätszonen

Ergebnis: Teilschnittbilder der radioaktiven Verteilung

Ultraschall Tomographie

Ungenauer als CT und MRI (1mm), aber schneller. Aufnahme von bewegten Objekten möglich.

Aufnahmearten: händisch bewegt, mechanisch bewegt, 2D Aufnahmebild

Fischgrößenbestimmung

- Fische von der Seite stereo aufgenommen.
- Kantendetektion
- Merkmalssuche
- Größenbestimmung

Kohlenprofilmesssystem mittels Stereo

Überwachung der Winderosion von offenen Kohlentransportern.

Automatische Auswertung bewegter Wagons auf Grund von Korrespondenzen

Probleme:

- großer Aufnahmebereich
- bewegte Wagons
- große Kamerabstände
- feiner Kohlenstaub

Autom. Klassifikation von Frontplatten

2 Laser, 3 Kameras über Förderband erkennen Küchenfrontplatten

Computer Aided Classification of Ceramics

Shape from Video

3D – ShapeSnatcher

12. Industrielle BV

Anforderungen

Modularität
Flexibilität
Bedienerfreundlichkeit
Wartungsfreundlichkeit
Erfüllung firmeninterner Normen

- Prüfvorgänge
- Bewegungssteuerung und -regelung
- Koordinatenmesstechnik

Automatische Visuelle Inspektion

nur Einzellösungen, sehr teuer. Müssen im Takt sein, ISO gerecht etc...

Ausbaustand:

manuell → halbautomatisch → vollautomatisch

Schraubenprüfsystem

statistischer Ausschuss, darf für automatische Montage nicht eingesetzt werden

Kabelprüfsystem

Vermessung der Isolierung Abschirmung, Innenleiter, Sauberkeit, überstehende Drähte, Qualität der Verzinnung

Prüfsystem

- 3 Kameras in 120° Winkel
- Beleuchtung durch IR Dioden
- Durchlichtverfahren
- Anpassung durch Teaching

Identifikationssysteme

Barcodes

Postsortierung

Teileidentifikation