

# **Integration geometrischer und fotogrammetrischer Information zum Wiederfinden von Bildern**

Björn Burow

SE Mustererkennung in Bildern und 3D-Daten

Lehrstuhl Graphische Systeme BTU Cottbus

# Inhaltsübersicht

- 1. Einleitung
- 2. Methoden der Bildrückgewinnung
  - 2.1 Methode basierend auf Intensitäts-Invarianten
    - 2.1.1 Intensitätsinvarianten
    - 2.1.2 Wiedergewinnungsalgorithmus (Retrieval Algorithm)
    - 2.1.3 Halblokale Beschränkungen
  - 2.2 Methode zum Kurven - Matching
    - 2.2.1 Basis – Kurven – Matching - Algorithmus
    - 2.2.2 Weite – Basis – Linien – Matching – Algorithmus

# 1. Einleitung

- zwei Bildabstimmungsverfahren
  - Erfolg unabhängig von geometrischen u. fotogrammetrischen Beschränkungen
- erstes Verfahren
  - Nutzung lokaler Intensitätsvarianten u. halblokaler geometrischen Beschränkungen
- zweites Verfahren
  - Vergleich 3D – Kurven und Linien
    - epipolare Geometrie – und lokale fotogrammetrische – Beschränkungen
- Kombination der Verfahren möglich
  - ergänzen sich bei Anwendung der Bilddrückgewinnung aus Bilddatenbank

# 2.1 Methode basierend auf Intensitätsinvarianten

- Nachschlagetabellenmechanismus (Datenbank)
  - Merkmalsübereinstimmung u. Modell – Suche/Zuordnung und Annäherung
- größte Schwierigkeit sind Geometrieeigenschaften (Liniengruppierungen)
- alternativ nutzt man fotogrammetrische Informationen zur Charakterisierung
- Idee der Indizierung:
  - Nutzung lokaler Intensitätsinvarianten als „Bildbeschreiber“
    - Deskriptoren werden an automatisch entdeckten Interessenpunkten (interest points) berechnet
  - Basis-Idee ist Nutzung der Auto – Korrelationsfunktion, um Orte zu finden, wo sich Signal in zwei Richtungen verändert

- Ergebnisse zeigen:
  - richtige Wiedergewinnung bei:
    - teilweise Sichtbarkeit
    - kleinen Umwandlungen
    - kleinen perspektivischen Verformungen

# 2.1.1 Intensitätsinvarianten

- Nachbarschaft jedes Interessenpunktes (interest point) wird durch einem Vektor von lokalen Intensitätsinvarianten beschrieben
- um Unveränderlichkeit im Bild zu erhalten, werden differentiale Invarianten berechnet → beschränkt auf 3. Ordnung
- Vektor mit Invarianten als  $V$  bezeichnet
- Invarianten werden in maßstabgetreues Gerüst eingefügt

# 2.1.2 Wiedergewinnungs-Algorithmus (Retrieval Algorithm)

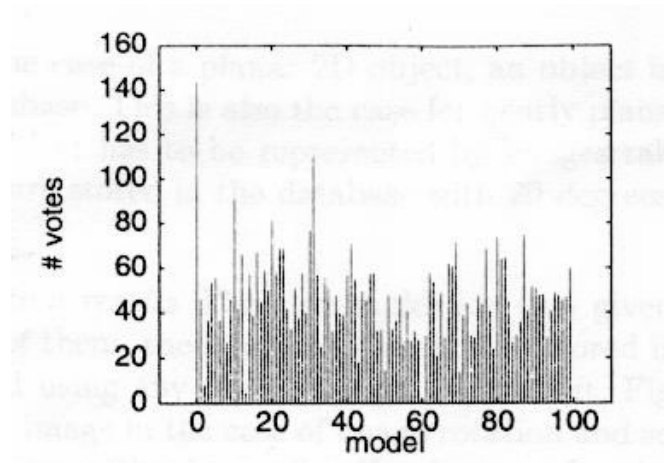
- Vektor – Vergleich
  - Mahalanobis – Distanz  $\mathbf{d}_M$  zweier invarianter Vektoren wird erzeugt, um die Zusammenhänge der markanten Punkte zu erkennen
  - die Distanz berücksichtigt die unterschiedliche Größe der Covarianzmatrix  $\mathbf{A}$  der Komponenten
    - Merkmalsvektoren sind in die Distanz eingeebnet
  - für zwei Vektoren  $\mathbf{a}$  und  $\mathbf{b}$ :
$$\mathbf{d}_M(\mathbf{b}, \mathbf{a}) = \sqrt{(\mathbf{b}-\mathbf{a})^T \mathbf{A}^{-1} (\mathbf{b}-\mathbf{a})}$$
  - Merkmale sind die Signal – Stärke und Luminanzänderungen
  - Mahalanobis – Distanz ungünstig für schnelles Indizierungsverfahren
    - bei wenig Veränderung ist ausweichen auf Euklidische Distanz  $\mathbf{d}_E$  möglich
- Bild – Datenbank
  - enthält Gruppe  $\{\mathbf{M}_k\}$  von Modellen
  - jedes Modell  $\mathbf{M}_k$  wird von Vektoren der Invarianten  $\{\mathbf{V}_j\}$  definiert
  - jeder Vektor  $\mathbf{V}_j$  hat Bindung zu Model  $\mathbf{k}$ 
    - es ergibt sich eine Datenbank  $(\mathbf{V}_j, \mathbf{k})$

- Voting – Algorithmus
  - Modellgruppe  $\mathbf{M}_k$  finden, die am besten zu vorgegebenem Bild  $\mathbf{I}$  passt
  - für Bild wird Gruppe von Vektoren  $\{\mathbf{V}_l\}$  anhand der Interessenpunkte (interest points) berechnet
  - Vergleich der Vektoren mit  $\mathbf{V}_j$ :  $\mathbf{d}_M(\mathbf{V}_l, \mathbf{V}_j) = \mathbf{d}_{l,j}$  für alle  $(l, j)$
  - Voting erfolgt, wenn Entfernung unter Schwelle  $t$
  - Summe der Modellauswahl ( $\mathbf{T}(\mathbf{k})$ ) ist Ergebnis des Voting – Algorithmus
  - meistgewählte Modell repräsentiert  $\mathbf{M}_{k^*}$  für  $\mathbf{k}^* = \arg \max_k \mathbf{T}(\mathbf{k})$



- **Multidimensionale Indizierung**

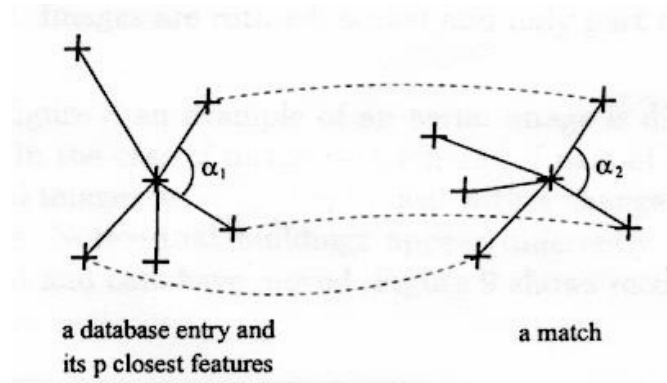
- Indizierung des Voting – Algorithmus durch **I x N**
  - **I** ist Anzahl von Merkmalen im Bild und **N** die Anzahl der Gesamtmerkmale in Datenbank
  - Indizierung notwendig, da **N** sehr groß werden kann
- Suchstruktur ist Variante von **k-d-Bäumen**



- Komplexität der Indizierung ist Reihenfolge von **I** (Merkmalen im Bild)
- sehr effiziente und schnelle Methode

## 2.1.3 Halblokale Beschränkungen

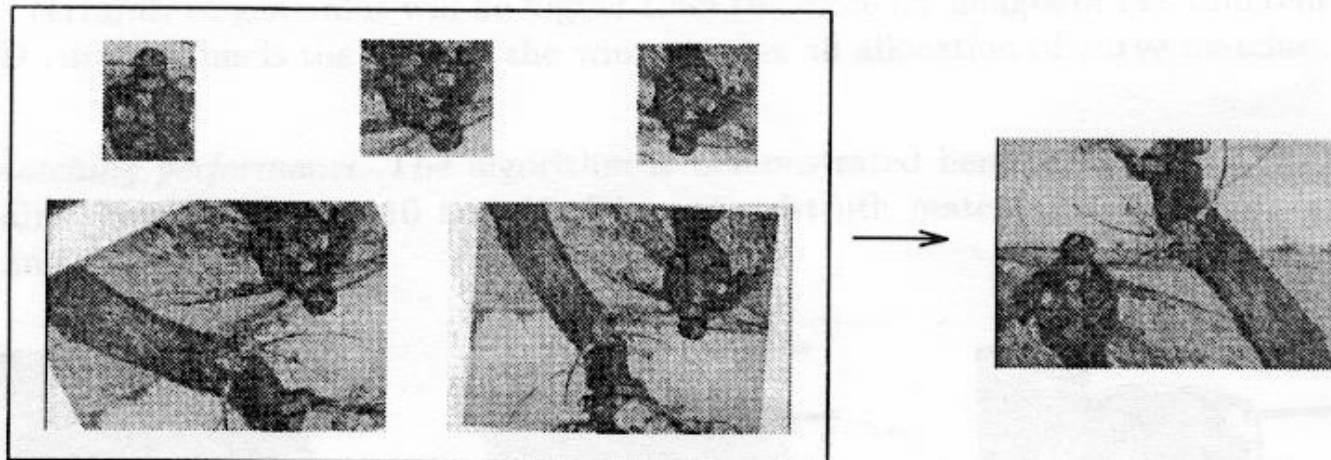
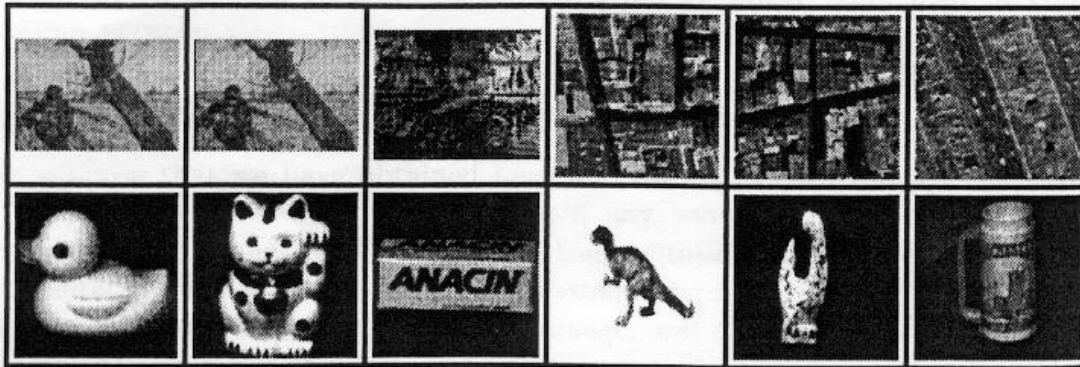
- halblokal, da von anderem Raum aus betrachtet
- Vielzahl von Modellen bzw. ähnlichen Modellen → Erhöhung der Wahrscheinlichkeit der Ähnlichkeit von Modellen
  - hinzufügen lokaler Formkonfiguration
- für jedes Merkmal (Interessenpunkt) in DB wird nächste  $p$  im Bild gewählt



- bei Notwendigkeit, dass alle  $p$  korrekt übereinstimmen, wird angenommen, dass es keine Fehl – Erkennung von Punkten gibt
  - nötig, dass 50 % der Nachbarpunkte übereinstimmen

- um Erkennungsrate zu vergrößern, werden geometrische Beschränkungen hinzugefügt
  - Winkel- und Längenverhältnisse der halblokalen Formkonfiguration muss konsistent sein

- Versuchsergebnisse



# 2.2 Methode zum Kurven - Matching

- für Zeilen- u. Kurvenabstimmung zwischen zwei perspektivischen Bildern einer von unterschiedlichen Standpunkt erworbenen 3D-Szene
- Annahme:
  - 3D – Wirkungen können nicht ignoriert werden
  - grundsätzliche Matrize „F“ für Bildpaar vorhanden
- Methode ergänzt geometrische von fotogrammetrischen Beschränkungen → Intensität der Nachbarschaft wird ergänzt
- besonders wird Ähnlichkeit der Kurven von Querbeziehungen an entsprechendem Punkt abgeschätzt

# 2.2.1 Basis – Kurven – Matching - Algorithmus

- Annahme, dass es Linien und Kurven in jedem Bild gibt
- bestimmt werden soll, welche Zeile/Kurve übereinstimmen (wenn überhaupt)
- sehr hoher Rechenaufwand, wenn für jeden Punkt im einen Bild das komplette andere Bild durchsucht wird
  - absuchen einer Linie genügt
- Nutzung der Epipolar – Geometrie

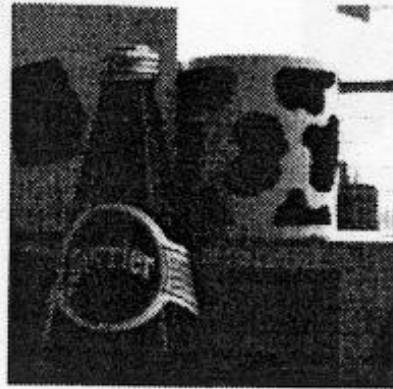
- Nutzung der photogrammetrischen Informationen zur Ermittlung von Ähnlichkeitswerten
- Berücksichtigung zweier möglichst ähnlichen Kurven  $\mathbf{c}$  und  $\mathbf{c}'$  im ersten und zweiten Bild
  - Kurven sind ähnlich, wenn sie die gleiche 3D-Kurve erzeugen
- bei Entsprechung kann ein Punkt, um Übereinstimmung auf Kurven zu deuten, bei Nutzung der Epipolar-Geometrie bestimmt werden:
  - für Bildpunkt  $\mathbf{x}$  auf Kurve ist die Epipolar-Linie im zweiten Bild  $\mathbf{l}'_e = \mathbf{F}\mathbf{x}$
  - diese Linie schneidet die Kurve  $\mathbf{c}'$  im Punkt  $\mathbf{x}'$  ähnlich zu Punkt  $\mathbf{x}$ 
    - Voraussetzung ist, dass  $\mathbf{x}$  und  $\mathbf{x}'$  Bilder des gleichen 3D – Punktes sind
  - Folglich sollten Image – Intensitätsnachbarschaften von  $\mathbf{x}$  und  $\mathbf{x}'$  ähnlich sein
  - Ähnlichkeitswerte für  $\mathbf{c}$  und  $\mathbf{c}'$  werden bestimmt, in dem man den Durchschnitt von Ähnlichkeiten von Nachbarschaften ermittelt
  - die Ähnlichkeit von Nachbarschaften wird durch Kreuz-Korrelation ermittelt

- je mehr sich die Kurven entsprechen, desto höher sind die Ähnlichkeitswerte
  - im allgemeinen sein werden die Werte höher sein als die Werte für Bilder zweier unterschiedlicher 3D-Kurven
- das bildet Basis für das Kurven - Matching





frame 11

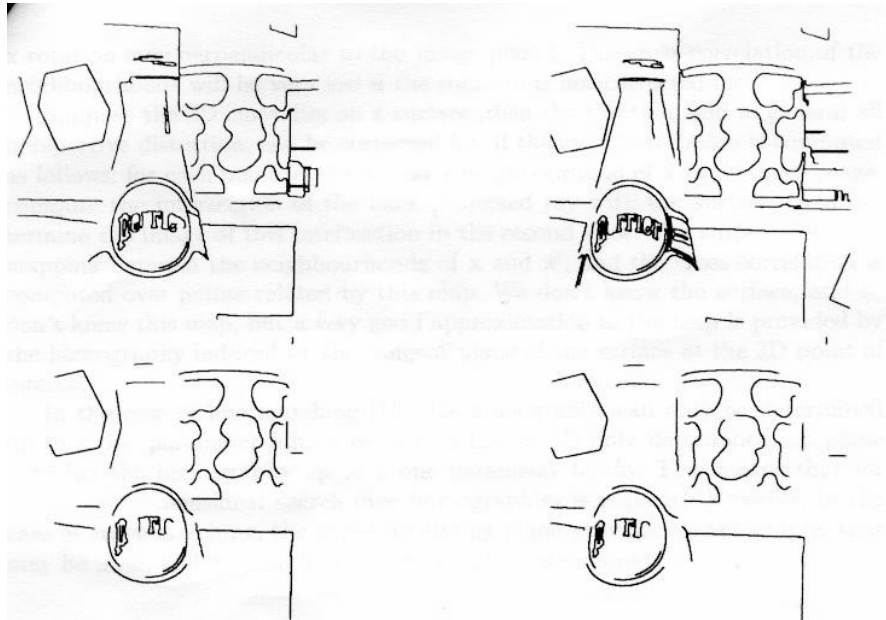


frame 15



frame 19

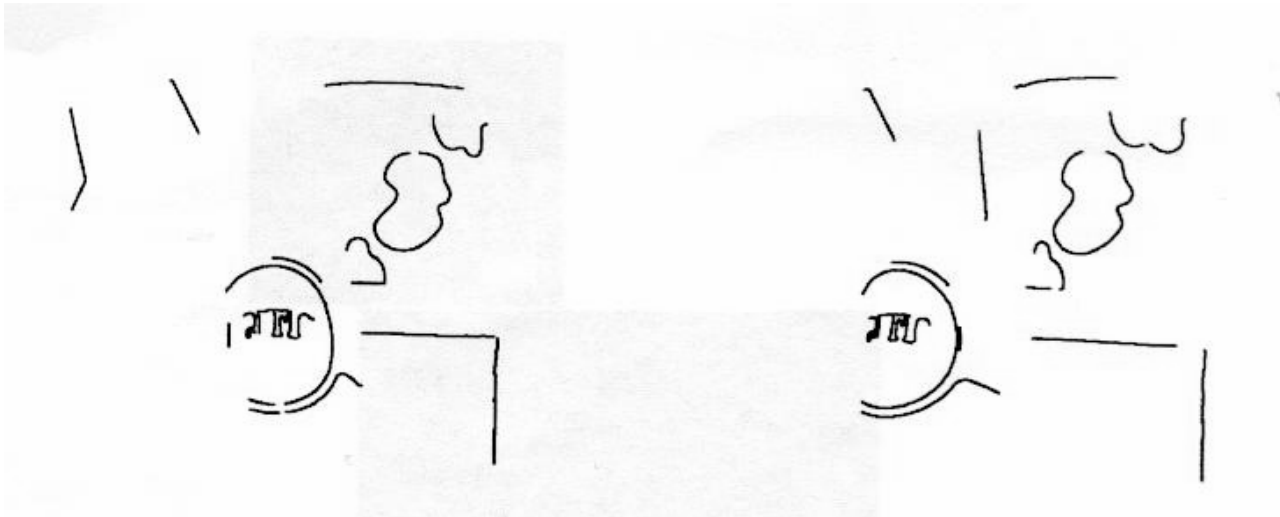
- Flaschensequenz
- ausgewählte Frames um Bilderpaare zu bilden
- Kamerabewegung sehr einheitlich, Entfernung so gut wie gleich



- Kurze – Grundlinien – Vergleich für Frame 11 und 15 einer Flaschensequenz
- obere Paar:
  - gewonnenen Umrisse (37 links, 47 rechts) der Kurven, die für Algorithmus verwendet wurden
- untere Paar:
  - zeigt nur Teile, bei denen gleiche Kanten erkannt wurden (29 Umrisse)

# 2.2.2 Weite – Basis – Linien – Matching – Algorithmus

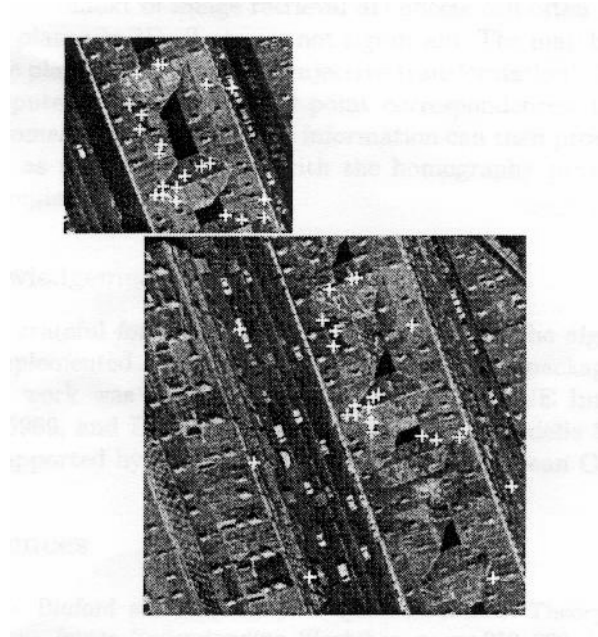
- Anwendung bei größerer Blickwinkelrotation oder breiten Basislinien zwischen den Ansichten
- Bsp.:
  - Vergleich der Frames 11 und 19 der Flaschensequenz
  - 16 ähnliche Kurvenumrisse wurden entdeckt, davon sind 14 an der richtigen Position (87,5 %)



# 3. Effizienz der Indizierung von Interessenpunkten (interest points) bei Kurvenverifizierung

- unter 2.1 wurde gezeigt, wie gegebenes Bild mit Hilfe einer Bilddatenbank wiedergewonnen werden kann
  - effizientes Verfahren durch Indizierung von Interessenpunkten (interest points)
  - Gruppenzuordnung und Bildzuordnung weiterhin notwendig
- gezeigt wird, dass Kurven - Matching ebenfalls genutzt werden kann
- Annahme, dass Bild und Datenbankbilder Ansichten unterschiedlicher Aussichtspunkte in 3D-Szene sind
- erste Überprüfung:
  - bestimmt, ob „interest points“ epipolarer Geometriebeschränkungen entsprechen
    - Ergebnis: zwischen Bild und Datenbankbildern ist grundsätzliche Matrix vorhanden

- zweite Überprüfung:
  - Zeilen/Kurven – Vergleich aus Abschnitt 2.2
  - wiedergewonnene Bild erhält Rang
  - umso höher Rang, desto größer ist Überlappung der Blickpunkte der 3D-Szene



- 11 Bilder in Datenbank mit mehr als 7 gleichen Interessenpunkten
  - mindestens notwendig, um Fundamentalmatrix zu berechnen

- Bilder werden durch Indizierung lokaler Intensitätsinvarianten und durch halblokale geometrische Beschränkungen bestimmt
- Kurven - Matching aus Abschnitt 2.2 wird auf jedes Bildpaar angewendet
- Kombination der grundsätzlichen Matrix mit den zusätzlichen geometrischen und fotogrammetrischen Beschränkungen ist beste Bildvergleich
  - Vereinigung des Indizierung von Interessenpunkten (interest points) mit dem Kurven - Matching

# 4. Zusammenfassung

- zuerst wurde gezeigt, dass Punktübereinstimmungen zwischen zwei Bildern mit Hilfe von Intensitätsmustern möglich ist (Datenbankindizierung)
- entscheidend ist Nutzung geometrischer Beschränkungen in Form lokaler Kohärenz auf Punktmustern
- Mehrdeutigkeit wird durch Einführung fotogrammetrischer Beschränkungen versucht zu beseitigen
- beide Verfahren gut für Abstimmung von Bildern einer 3D-Szene mit unterschiedlichen Blickwinkeln
- 3D – Wirkungen können je nach Anforderung auch ignoriert werden

# 5. Literaturhinweise

- Schmid, Cordelia; Zisserman, Andrew; Mohr, Roger: Integrating geometric and photometric information for image retrieval. *Lecture Notes in Computer Science* 1681 (1999), 217–233