

„A Comparative Evaluation of Template and Histogram Based 2D Tracking Algorithms“

B. Deutsch, Ch. Gräßl, F. Bajramovic, J. Denzler

Paper Review

Matthias Hillebrand

im Rahmen des Seminars
„Visuelle Überwachung“

Übersicht

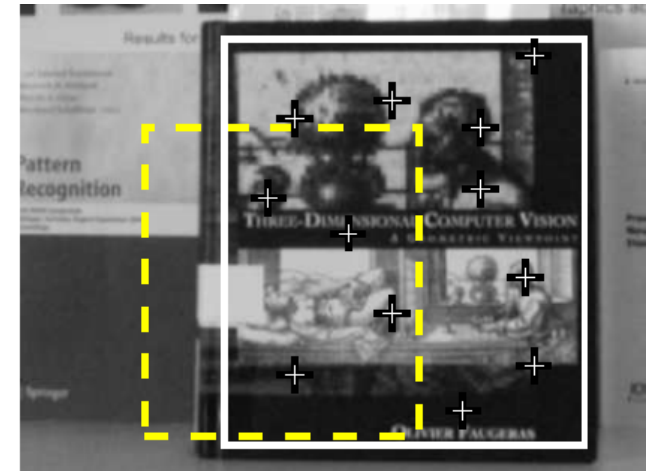
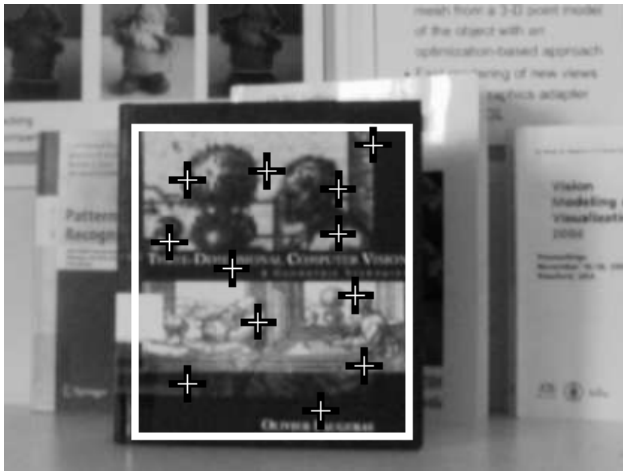
- (1) Gegenstand des Papers
- (2) Regionentracking
- (3) Histogrammtracking
- (4) Personenbewegung
- (5) Tracking-Performanz
- (6) Auswertung
- (7) Diskussion
- (8) Fazit
- (9) Literatur

Gegenstand des Papers

- Vergleich fünf verschiedener Trackingverfahren
 - zwei Regionen-basierte Tracker
 - drei Histogramm-basierte Tracker
- Personentracking auf Überwachungsvideos
 - 7 Videos von fest installierten Kameras
 - 12 Personen
 - insgesamt ca. 12.000 Bilder

Regionentracking

- Verfahren: Hager & Belhumeur, Hyperplane
- Objekt wird durch Template beschrieben
- Referenztemplate wird in der Umgebung der Position des Objekts im letzten Zeitschritt gesucht
- unter Einbeziehung möglicher Transformationen: Translation, Skalierung, Rotation, affine Transformationen etc.



Links: Initialisierung des Trackers. Mitte: Objekt nach Translation. Rechts: nach Translation und Skalierung.

Histogrammtracking

- Verfahren: Mean Shift, Trust Region, CONDENSATION
- Objekt wird durch Histogramm beschrieben
- z. B. Farb- oder Intensitätshistogramm
- Referenzhistogramm wird mit den Histogrammen einzelner Bildauschnitte in der Umgebung der Position des Objekts im letzten Zeitschritt verglichen
- Rotationsinvariant, d. h. es kann nur Translation und Skalierung betrachtet werden

Personenbewegung

- Initialisierung des Trackers: erstes Videobild, auf dem die Person komplett zu sehen ist
- Repräsentation einer Person durch rechteckigen Bildausschnitt (Ground-Truth-Daten sind für jedes Bild vorhanden)
- Zweimaliges Tracking jeder Person:
 - Translation: Bewegung des initialen Bildausschnitts mit fester Größe
 - Translation und Skalierung: Bewegung des initialen Bildausschnitts mit veränderlicher Größe



Beispielbild einer Überwachungskamera, Ground Truth und Trackingergebnis eingezeichnet

Performanz-Messung

- Vergleich der Tracking-Ergebnisse mit Ground-Truth-Daten:
 - Distanzfehler e_c : euklidischer Abstand zwischen den Mittelpunkten der Rechtecke
 - Regionenfehler e_r : nicht-überlappender Anteil der Rechtecke

$$e_r(A, B) := \frac{|A \setminus B| + |B \setminus A|}{|A| + |B|} \quad (A, B \text{ Mengen})$$

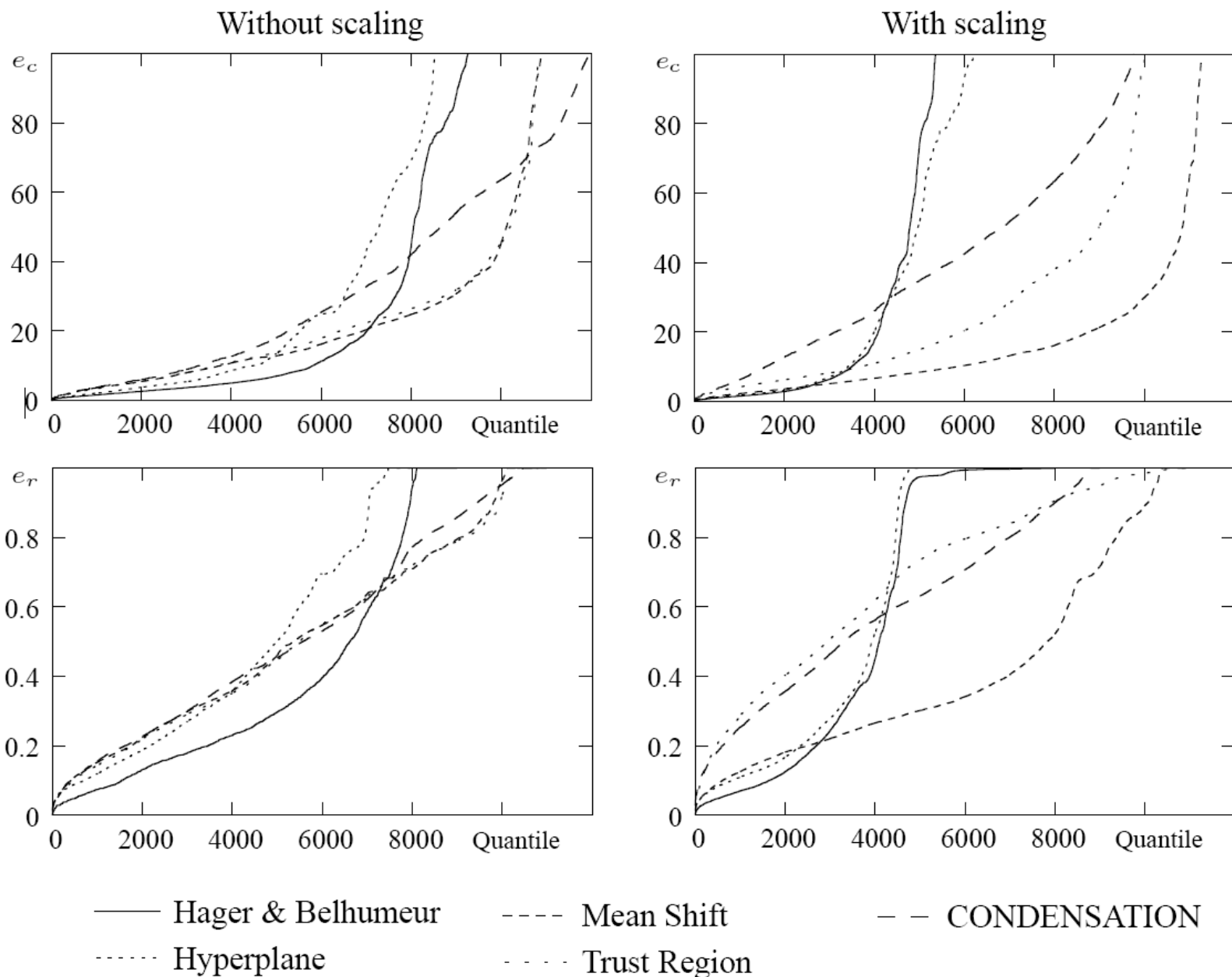
- Geschwindigkeit:
 - Zeit für Initialisierung
 - Zeit pro Bild (Translation ohne/mit Skalierung)

Auswertungs-Modus

- Getrennte Auswertung für Durchläufe ohne und mit Skalierung
- Getrennte Auswertung für Distanz- und Regionenfehler
- Für jeden Tracker:
 - Zusammenfassung der berechneten Fehler von allen 12.000 Bildern
 - Aufsteigende Sortierung
 - Grafische Präsentation

Auswertungs-Diagramm

- Oben:
Distanzfehler
- Unten:
Regionenfehler
- Links:
Translation
ohne Skalierung
- Rechts:
Translation
mit Skalierung



Auswertung

- Erkennungs-Performanz variiert stark, je nach Testverfahren und verwendetem Fehlermaß
- kein Tracker ist eindeutig der beste oder der schlechteste
- Unterschied zwischen den beiden Tracker-Klassen:
 - regionenbasierte Verfahren liefern für einen gewissen Anteil der Bilder bessere Ergebnisse
 - darüber hinaus steigt die Fehlerrate schneller
 - max. Fehler (Verlust der Person) wird eher erreicht
- histogrammbasierte Verfahren robuster
- Geschwindigkeit:
Hyperplane-Verfahren initialisiert in ~ 0.5 sec, CONDENSATION-Verfahren trackt in ~ 0.1 sec, alle anderen Zeiten $\sim 1-10$ ms

Diskussion

- Vergleich aktueller Tracking-Verfahren (1998-2004)
- kleine Testdatenbasis von nur zwölf Personen in sieben Videos
- Paper eigenständig
- Evaluation:
 - Methode von den Autoren selbst entwickelt
 - Schwäche: Bilder aller zwölf Personen in einen Topf geworfen
 - Verhalten des Trackers auf einer einzelnen Bilderfolge?
 - Durchschnittliche Dauer, bis eine Person verloren wird?

Fazit

- Tracking einer Person in einer realen Umgebung mit inhomogenem Hintergrund ist nicht trivial
- Keines der vorgestellten Verfahren kann sich im Test besonders von den anderen abheben
- Es gibt (noch) nicht das eine Verfahren, das die Antwort auf alle Fragen und Probleme ist

Literatur

- Hager, G., Belhumeur, P.:
Efficient region tracking with parametric models of geometry and illumination.
- Jurie, F., Dhome, M.:
Hyperplane approach for template matching.
- Comaniciu, D., Meer, P., Ramesh, V.:
Kernel-Based Object Tracking.
- Liu, T.L., Chen, H.T.:
Real-Time Tracking Using Trust-Region Methods.
- Perez, P., Hue, C., Vermaak, J., Gangnet, M.:
Color-Based Probabilistic Tracking.
- CAVIAR: EU funded project, IST 2001 37540,
URL: <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/> (2004)