

Detecting Pedestrians Using Patterns of Motion and Appearance

Ioannis Moutogiorgos

Universität Bielefeld
Seminar: Visuelle Überwachung
SS 2006

Übersicht

- Problembeschreibung
- Lösungsansatz
 - Fußgängererkennungssystem
- Experimentelle Evaluation
- Resultate
- Vorstellung alternativer Systeme
- Literaturangabe

Problembeschreibung

- Erfassung und Erkennung von beweglichen Objekten
- Einsatz des Systems findet im Freien statt
- Witterungsverhältnisse müssen berücksichtigt werden.
- Überlappung und Verdeckung der Objekte möglich

Abbildung 1: Beispiele erkannter Objekte mit dem dynamischen Detektor



Lösungsansatz

- Fußgängererkennungssystem
- Entwickelt von Paul Viola und Michael J. Jones
- Neuheit: Kombination von Intensitäts- und Bewegungsinformationen
- Boostingverfahren → AdaBoost
- Geringe Fehlerrate auch bei geringem Auflösungsvermögen
- Schnelle Evaluation des Systems

Fußgängererkennungssystem

- Arbeitet direkt auf den Bildern
- Extrahiert kurze Bewegungsmuster und Merkmale
- Ziel: Erfassung aller in Frage kommender Objekte
- keine weiteren Tracking-, Segmentations-, oder Registrierungsmechanismen benötigt
- Eine Vorverarbeitung der Bilder ist nicht nötig, z.B.:
 - Anpassung des Farbraums
 - Entfernung beziehungsweise Verringerung irrelevanter oder unerwünschter Bestandteile im Bild

- Feinabstimmungen und Korrekturen werden in der Trainingsphase vorgenommen
- Skalierung der Bilder im Vorfeld nötig
- AdaBoost → Erzeugung eines starken Klassifikators durch Kombination schwacher Lernalgorithmen
- Kaskadierte Architektur
- Rechteckfilter von P. Viola und Michael J. Jones werden eingesetzt
 - Erweiterung notwendig um auf Bewegungsbilder arbeiten zu können

Rechteckfilter

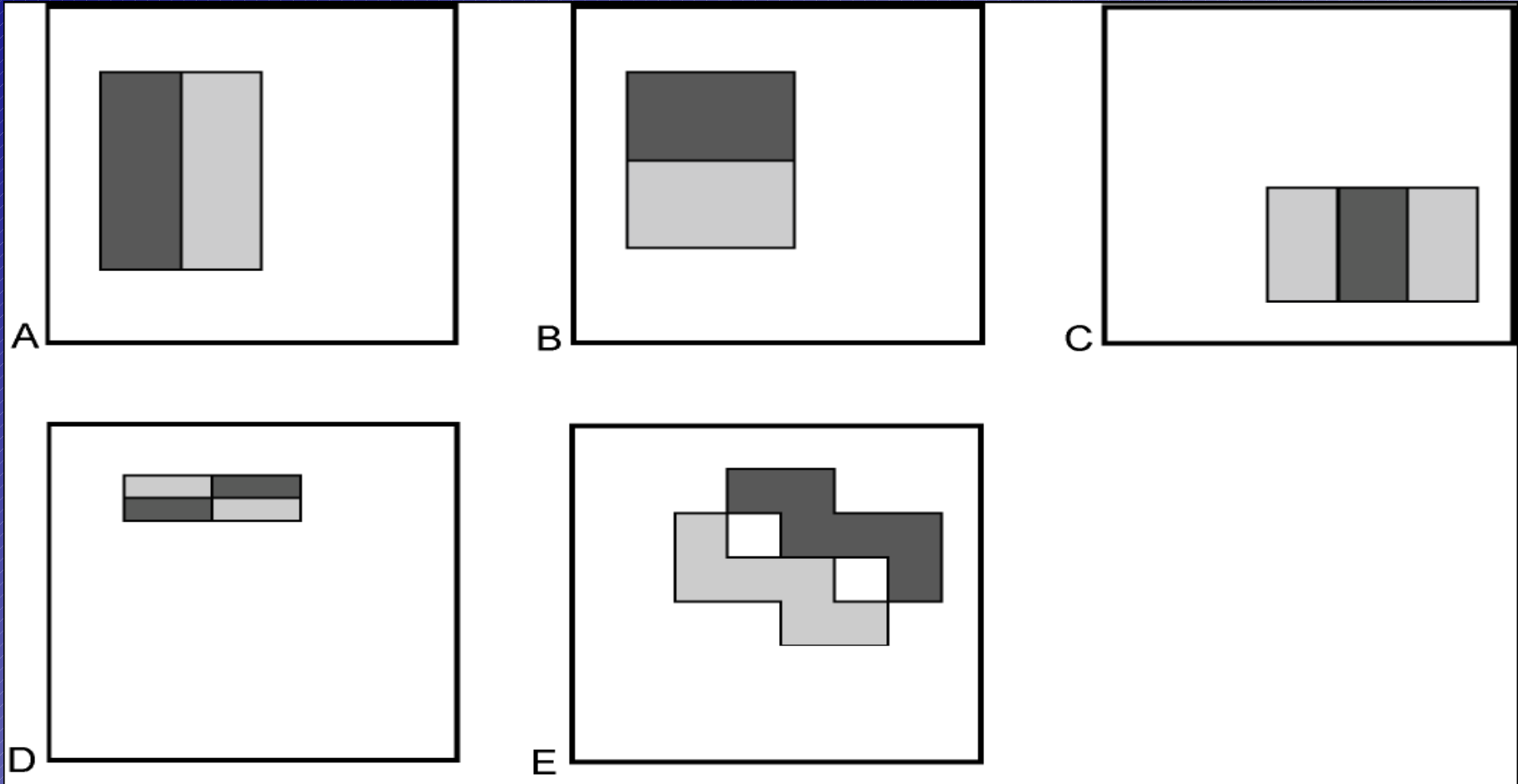
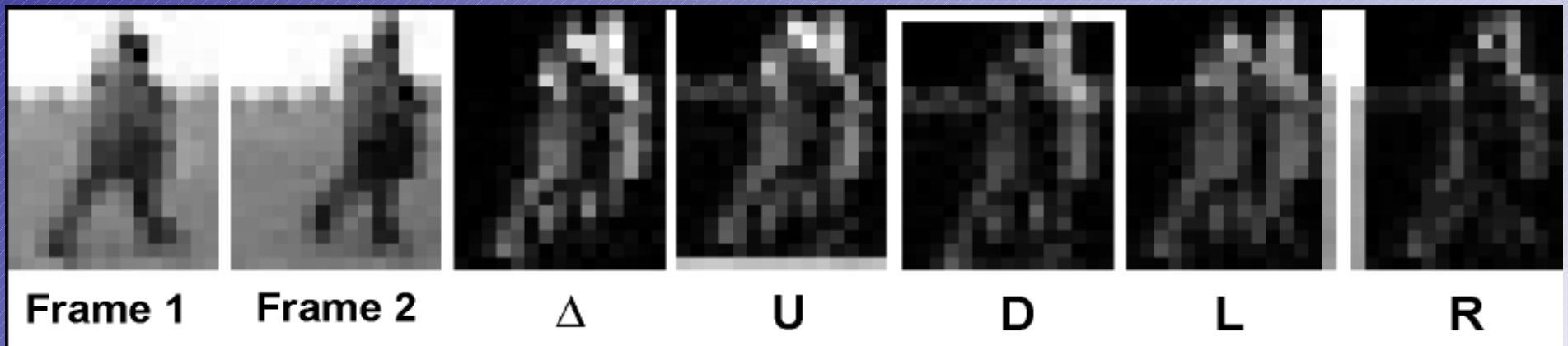


Abbildung 2. Rechteckfilter von P. Viola und Michael J. Jones

Erweiterung der Rechteckfilter

- Informationen über Bewegung werden aus Bildpaaren gewonnen
 - lokale Intensitätsunterschiede zwischen Bildregionen
 - Bewegungsrichtung kann durch Unterschiede zwischen ersten und verschobenem zweiten Bild ermittelt werden.
- 5 Arten von Bildern

- $\Delta = \text{abs}(I_t - I_{t+1})$
- $U = \text{abs}(I_t - I_{t+1} \uparrow)$
- $D = \text{abs}(I_t - I_{t+1} \downarrow)$
- $L = \text{abs}(I_t - I_{t+1} \leftarrow)$
- $R = \text{abs}(I_t - I_{t+1} \rightarrow)$
- I_t und I_{t+1} Bilder zu den Zeiten t und $t+1$
- $\leftarrow \uparrow \rightarrow \downarrow$ sind Verschiebungsoperatoren



Experimentelle Evaluation

- Evaluation der Detektoren
 - 8 Videosequenzen von Straßenszenen mit Passanten
 - 6 davon nutzte man um einen Trainingsatz einzurichten um den statischen und dynamischen Detektor zu evaluieren
 - 2 wurden für Testzwecke der Detektoren verwendet
 - Passanten wurden manuell markiert



Abbildung 3: Beispielbilder aus den 8 aufgenommenen Videosequenzen

- Dynamische Detektor wurde mit aufeinander folgenden Bildpaaren evaluiert
- Statische Detektor wurde mit statischen Mustern evaluiert
- Evaluation der Klassifikatoren jeder Stufe der Kaskade
 - 2500 positive und 2500 negative Beispiele
 - Jedes der Beispiele ist ein Paar von aufeinander folgenden Bildern



Abbildung der positiven Trainingsbilder der Klassifikatoren

Resultate

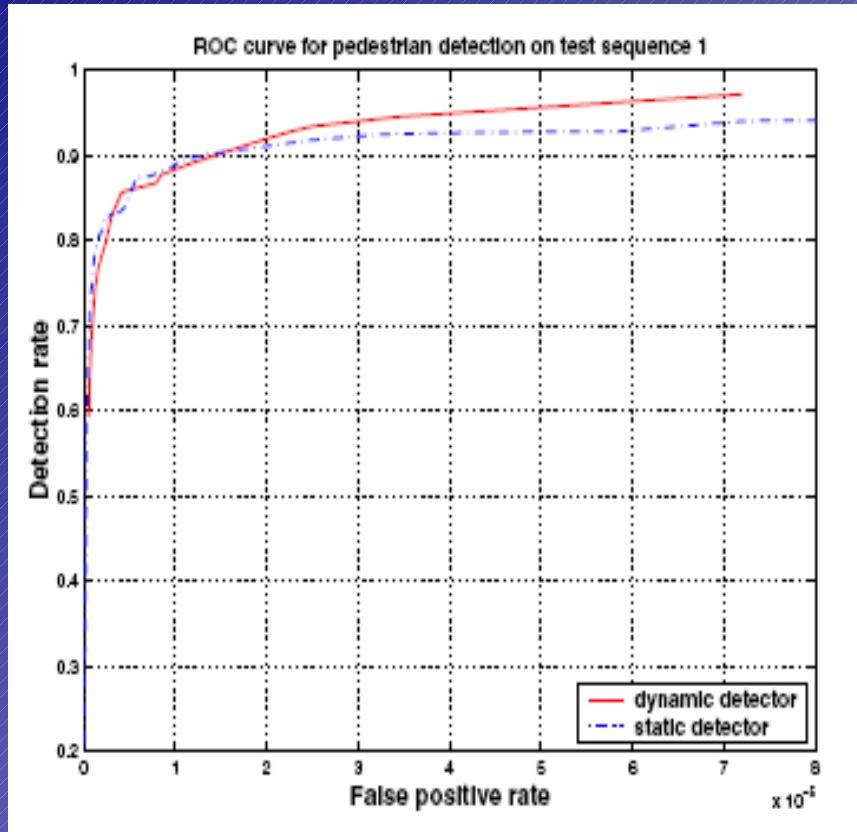


Abbildung 4: ROC Kurve zur Passantenerkennung. Testsequenz 1.

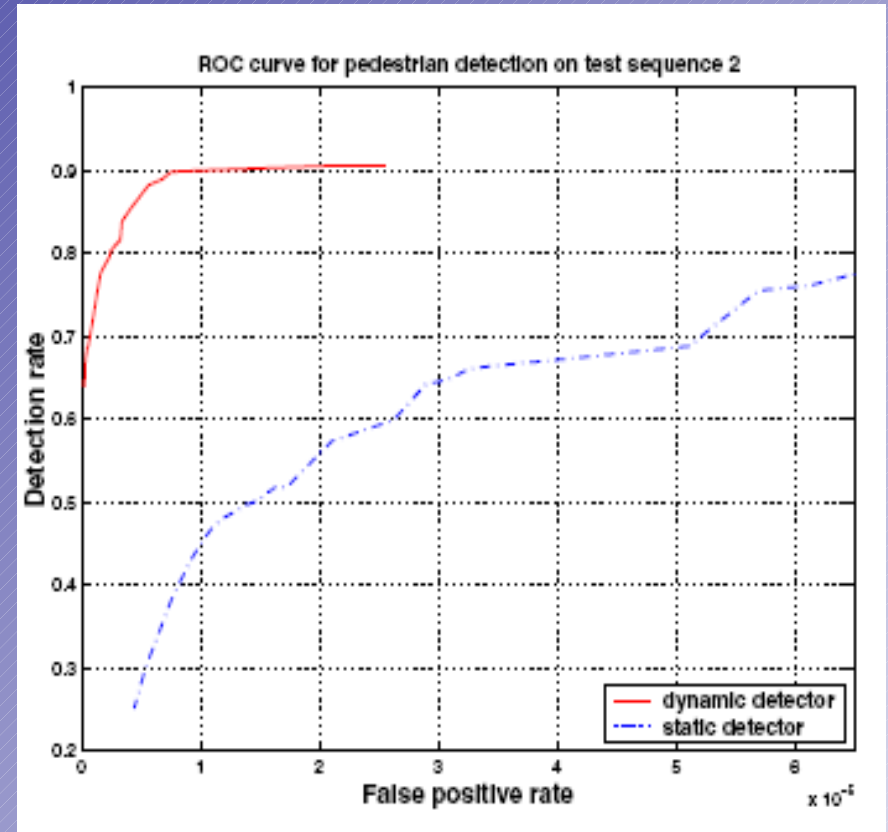
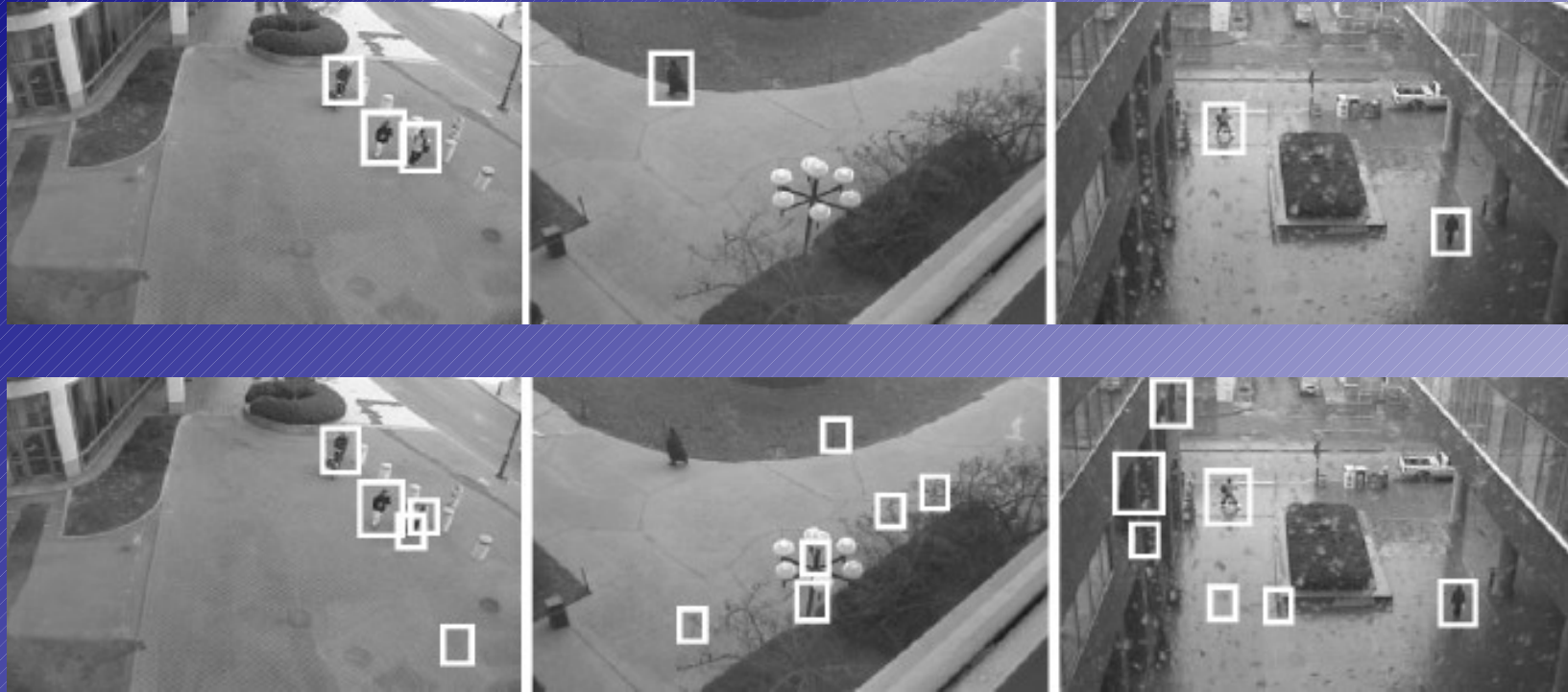


Abbildung 5: ROC Kurve zur Passantenerkennung. Testsequenz 2.

- Testsequenz 1. Der dynamische Detektor ist akkurater als der statische. Bei einer Detektionsrate von ca. 80%, haben beide Detektoren eine Fehlerrate von ca. 1/400000. (siehe Bild 5)
- Testsequenz 2. Der dynamische Detektor ist akkurater als der statische. Bei einer Detektionsrate von ca. 80%, hat der dynamische Detektor eine Fehlerrate von ca. 1/400000 wobei der statische eine von ca. 1/15000 hat. (siehe Bild 4)
- Es benötigte 0,25 Sekunden um alle Fußgänger in einem 360x240 Pixel Bild zu erfassen
- Der verwendete Prozessor war ein 2,8 GHz P4 Prozessor



Abbildungen 6 und 7: Beispiele von Abtastungen mit einem dynamischen Detektor (oben) und einem statischen Detektor (unten).

Vorstellung alternativer Systeme

- System von Gavrilu und Philomen
 - Einsatz auf statischen Bilder mit Passanten
 - Es werden Rechtecke auf den Bildern extrahiert und danach mit vorhandenen Exemplaren abgestimmt.
 - Fein abgestimmtes und praktisches System
 - Hat eine Erkennungsrate von 75%
- System von Papageorgiou
 - Erfasst Fußgänger mithilfe der Elementarwelle
 - Testbericht oder Ergebnisse nicht vorhanden
 - Kein direkter Vergleich möglich

Literaturangabe

- 1. C. Papageorgiou, M. Oren, and T. Poggio. A general framework for object detection. In *International Conference on Computer Vision*, 1998.
- 2. P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2001.
- 3. P. Viola, Michael J. Jones and Daniel Snow. Detecting Pedestrians Using Patterns of Motion and Appearance. In *IEEE Conference on Computer Vision*, 2003.
- 4. V. Philomen D. Gavrila. Real-time object detection for "smart" vehicles. In *IEEE International Conference on Computer Vision*, pages 87–93, 1999.
- 5. PETS (Performance Evaluation of Tracking and Surveillance)
<http://www.cvg.rdg.ac.uk/slides/pets.html>
- 6. Yoan Freund, Robert E. Shapire: A short introduction to Boosting.
- 7. A Real-Time Scene Understanding System for Airport Apron Monitoring, David Thirde, Mark Borg and James Ferryman
- 8. W4S: A Real-Time System for Detecting and Tracking People in 2 1/2D
Ismail Haritaoglu, David Harwood and Larry S. Davis ,Computer Vision Laboratory
University of Maryland, College Park, MD 20742, USA

? Fragen ?