

Betrachtung von Verfahren zur Posenbestimmung und Merkmalsexktraktion

Thorsten Jost
INF-MA – SR – Wintersemester 2008/2009
1. Dezember 2008

Agenda

- ▶ Motivation
- ▶ „Feature Detection“
- ▶ SIFT
- ▶ MOPS
- ▶ SURF
- ▶ SLAM
- ▶ Monte Carlo

Motivation

- ▶ Bildverarbeitung ist weit verbreitet
 - ▶ Industrie
 - ▶ Forschung
 - ▶ Security
 - ▶ Vermessung
- ▶ Extraktion von Information
- ▶ Teilgebiet: Objekterkennung

Motivation

- ▶ Aufgaben der Objekterkennung
 - ▶ Wiederfinden (Segmentieren)
 - ▶ Identifizieren
 - ▶ Lage, Orientierung
 - ▶ Positionsbestimmung

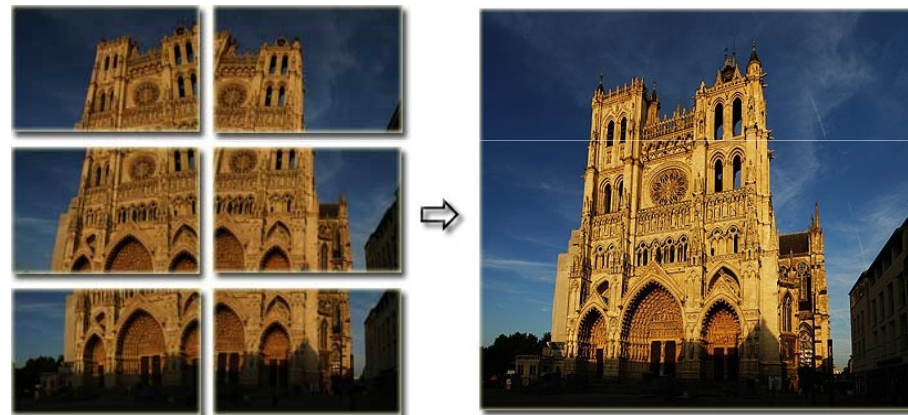
Motivation

- ▶ Ansprüche an Objekterkennung
 - ▶ Invarianz gegenüber
 - ▶ Position, Drehung, Größe
 - ▶ Affine und 3D Transformation
 - ▶ Verdeckung
 - ▶ Rauschen, Helligkeit
 - ▶ Geschwindigkeit
 - ▶ Echtzeit?

Motivation

- ▶ Mögliche Einsatzgebiete

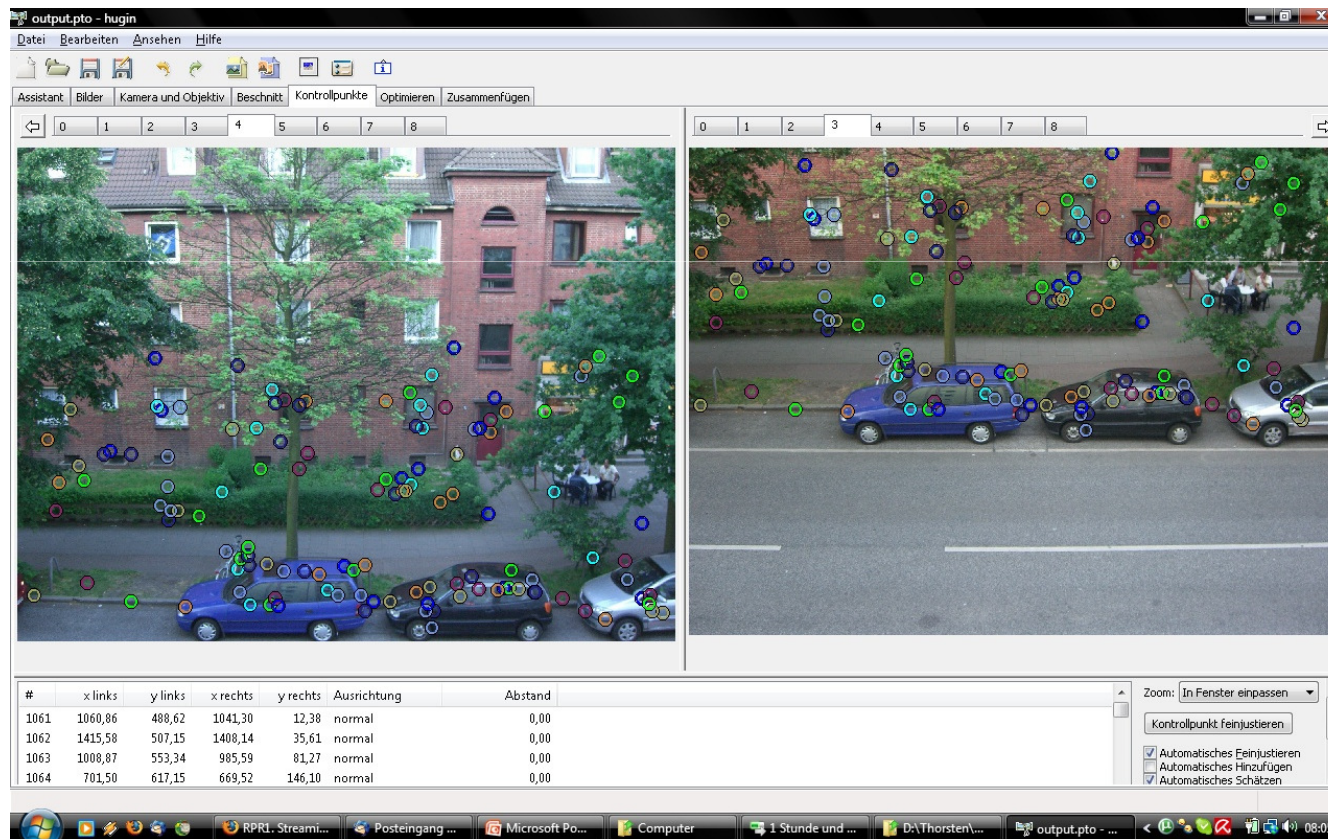
- ▶ Objekterkennung
- ▶ Stitching



- ▶ Rescue
- ▶ Infotainment
- ▶ Autonome Navigation!
- ▶ ...

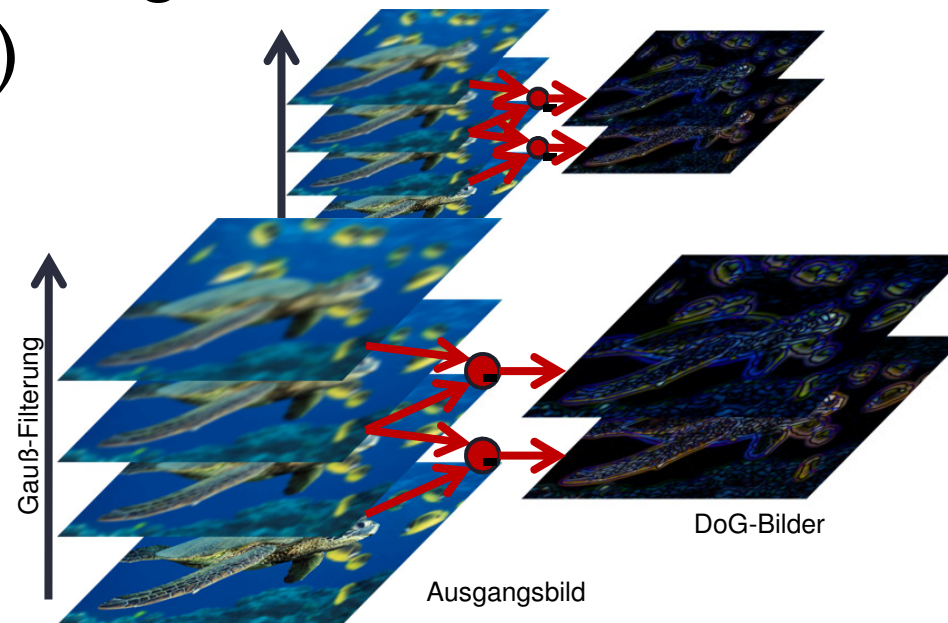
„Feature Detection“

► Verknüpfung von Bildpunkten in Bildpaaren



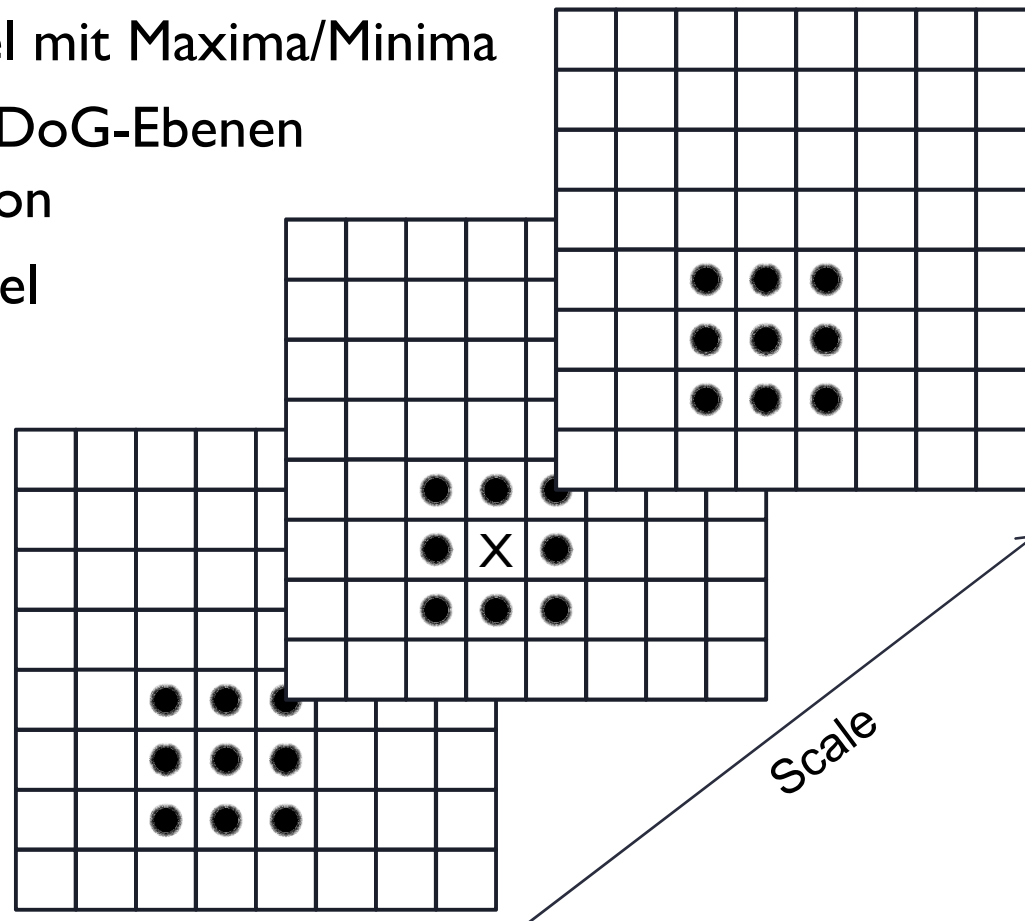
„Feature Detection“ mit SIFT

- ▶ Laplace-Pyramide
- ▶ Suche nach *markanten* Bildpunkten (Keypoints)
- ▶ Berechnung von ‚Difference of Gaussian‘-Bildern (DoG)



„Feature Detection“ mit SIFT

- ▶ Keypoints sind Pixel mit Maxima/Minima
- ▶ Innerhalb von drei DoG-Ebenen und einer 3x3 Region
- ▶ 26 benachbarte Pixel
- ▶ ca. 1000 Keypoints (512x512 Image)

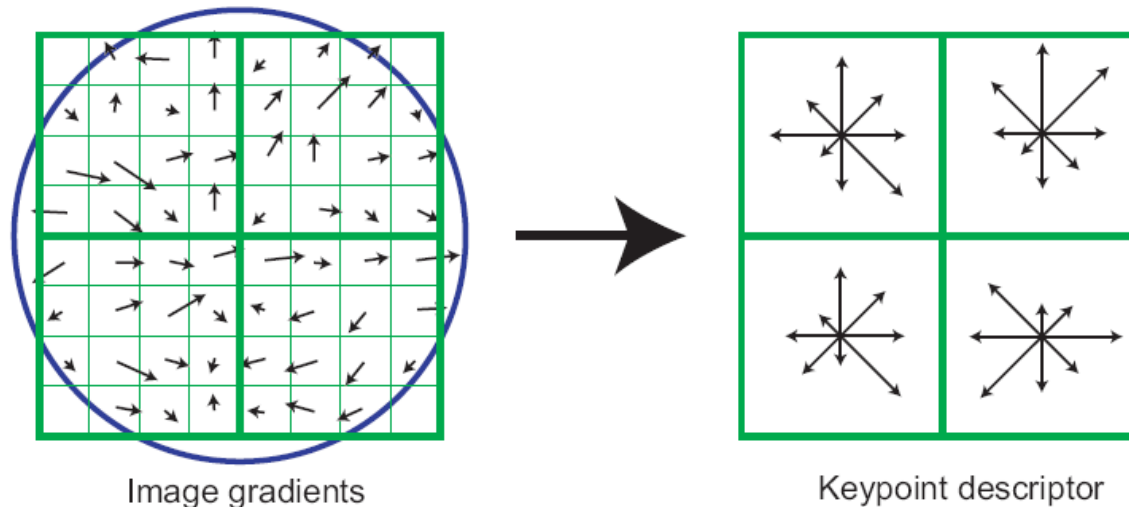


„Feature Detection“ mit SIFT

- ▶ Charakterisierung der Keypoints erforderlich
- ▶ Durch Gradientenbetrag und –richtung
- ▶ Werte werden in Histogramm akkumuliert
 - ▶ 36 „Behälter“ → ein Behälter 10% des Vollkreises
- ▶ Normierter Gradientenbetrag
 - ▶ → Belichtungs-Invarianz
- ▶ Der höchste Wert bestimmt die Orientierung des Keypoint
 - ▶ → Rotations-Invarianz

Lokale Abbildungs-Beschreibung

- ▶ Funktionsweise ähnlich der von Säugetieren
 - ▶ Visueller Cortex
- ▶ Gradienten-Richtungshistogramm
- ▶ SIFT-Key Vektor hat 160 Dimensionen!
 - ▶ $8 \times 4 \times 4 + 8 \times 2 \times 2$ (2 Oktaven)



Beispiele



„Feature Detection“ mit MOPS

- ▶ Multi-Image Matching using Multi-Scale Oriented Patches
 - ▶ M. Brown, R. Szeleiski, S. Winder, Microsoft Research

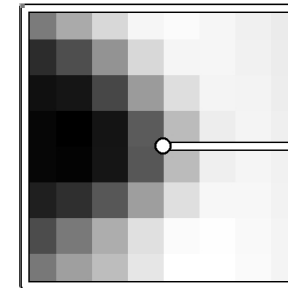


„Feature Detection“ mit MOPS

- ▶ Interest Points
 - ▶ Gauß-Pyramide wie bei SIFT
 - ▶ Jedoch „Multi-Scale Harris Corners“, statt Maxima/Minima
 - ▶ Partielle Ableitung der „Bildfunktion“
 - ▶ Sub-Pixel Genauigkeit
- ▶ Non-Maxima Rejection
 - ▶ Räumliche Verteilung der „Interest Points“
 - ▶ „Stärke“ IP → sortierte Liste
 - ▶ Minimaler Abstands-Radius
 - ▶ Bis eine gewünschte Anzahl von IPs gefunden wurde

„Feature Detection“ mit MOPS

- ▶ Direct Patch Based Sampling
 - ▶ „Oriented Patch“
 - ▶ 8x8 – Patch
 - ▶ 5 Pixel Zwischenräume
 - ▶ Berechnung des Features eine Ebene oberhalb der IP-Ebene
 - ▶ Zusätzliche Anwendung der Gauß-Funktion
 - ▶ → wenig „Aliasing“



„Feature Detection“ mit MOPS

- ▶ Feature Space Outlier Rejection
 - ▶ Beste Übereinstimmung von zwei Patches
 - ▶ (Quellbild $\leftrightarrow$ Suchbild)
 - ▶ Vermutlich ein Treffer
 - ▶ Zweitbeste Übereinstimmung von zwei Patches
 - ▶ Vermutlich KEIN Treffer
 - ▶ Durchschnitt aller Fehler der zweitbesten Übereinstimmung
 - ▶ $\rightarrow$ Outlier Distance
 - ▶ Verwerfung aller Patches, deren Fehler $>$ Outlier Distance
 - ▶ Reduzierung von 80% der falschen Patches
 - ▶ Verlust von 10% der korrekten Patches

„Feature Detection“ mit SURF

- ▶ Speeded Up Robust Features
 - ▶ H. Bay, T. Tuytelaars, L. Van Gool, ETH Zürich und Katholieke Universiteit Leuven
 - ▶ „Fast Hessian Detector“ für Interest Points
 - ▶ $\det(H)$ für Lage und Skalierung



$$\det(\mathcal{H}(x, \sigma)) = \det \begin{pmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{pmatrix} \approx D_{xx}D_{yy} - (0.9 D_{xy})^2$$

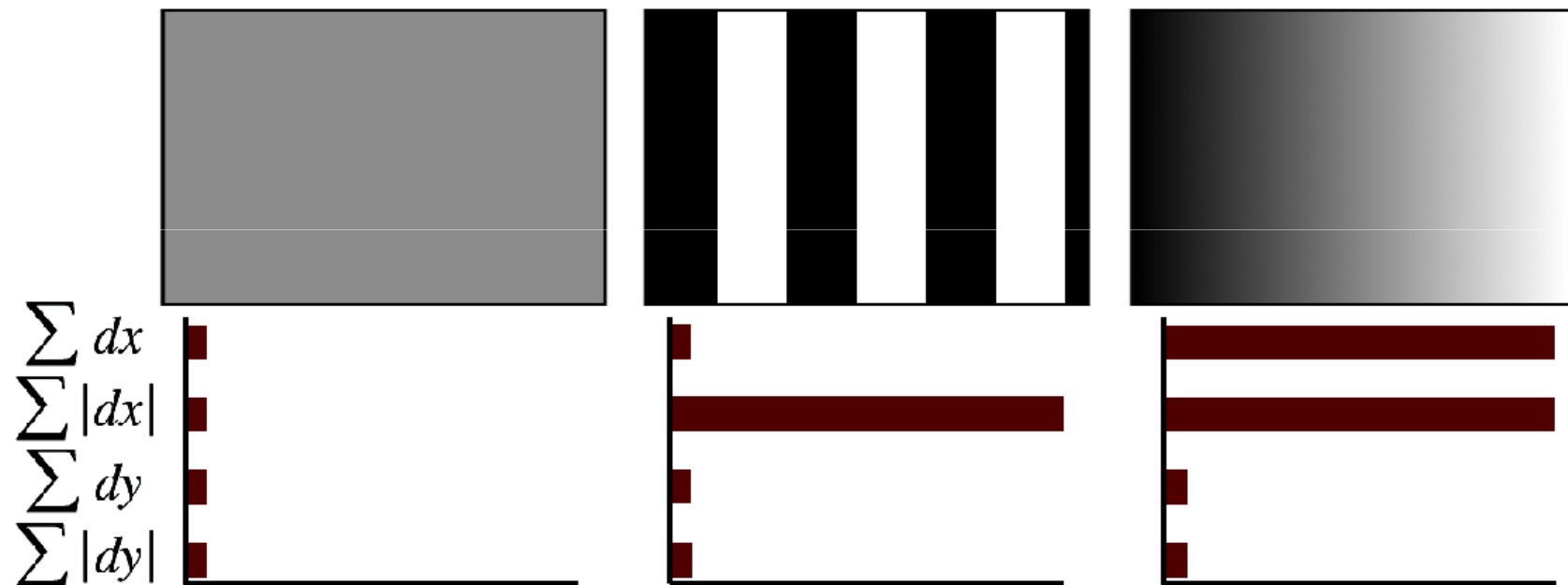
„Feature Detection“ mit SURF

- ▶ „Haar-Wavelet Response“ für IP-Descriptor
 - ▶ Basiert auf den bekannten SIFT-Features
 - ▶ Vermindert aber dessen Komplexität
- ▶ Orientierung wird durch die „Wavelet-Response“ berechnet
- ▶ Feature-Vektor enthält Informationen aus 4x4 Sub-Regionen

$$v = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \right)$$

- ▶ U-SURF → keine Orientierung → schneller, aber weniger genau

„Feature Detection“ mit SURF



SLAM

- ▶ Simultaneous Localization and Map-Building Using Active Vision
 - ▶ A. J. Davision und D.W. Murray, IEEE
 - ▶ Verwendet den „Kalman-Filter“
 - ▶ Orientierung anhand von Messdaten → Bewegungsmodell
 - Geschwindigkeit, Lenkwinkel, ...
 - ▶ Verknüpfung verschiedener Messverfahren
 - ▶ Bewegungsmodelle
 - ▶ Vision
 - ▶ ...

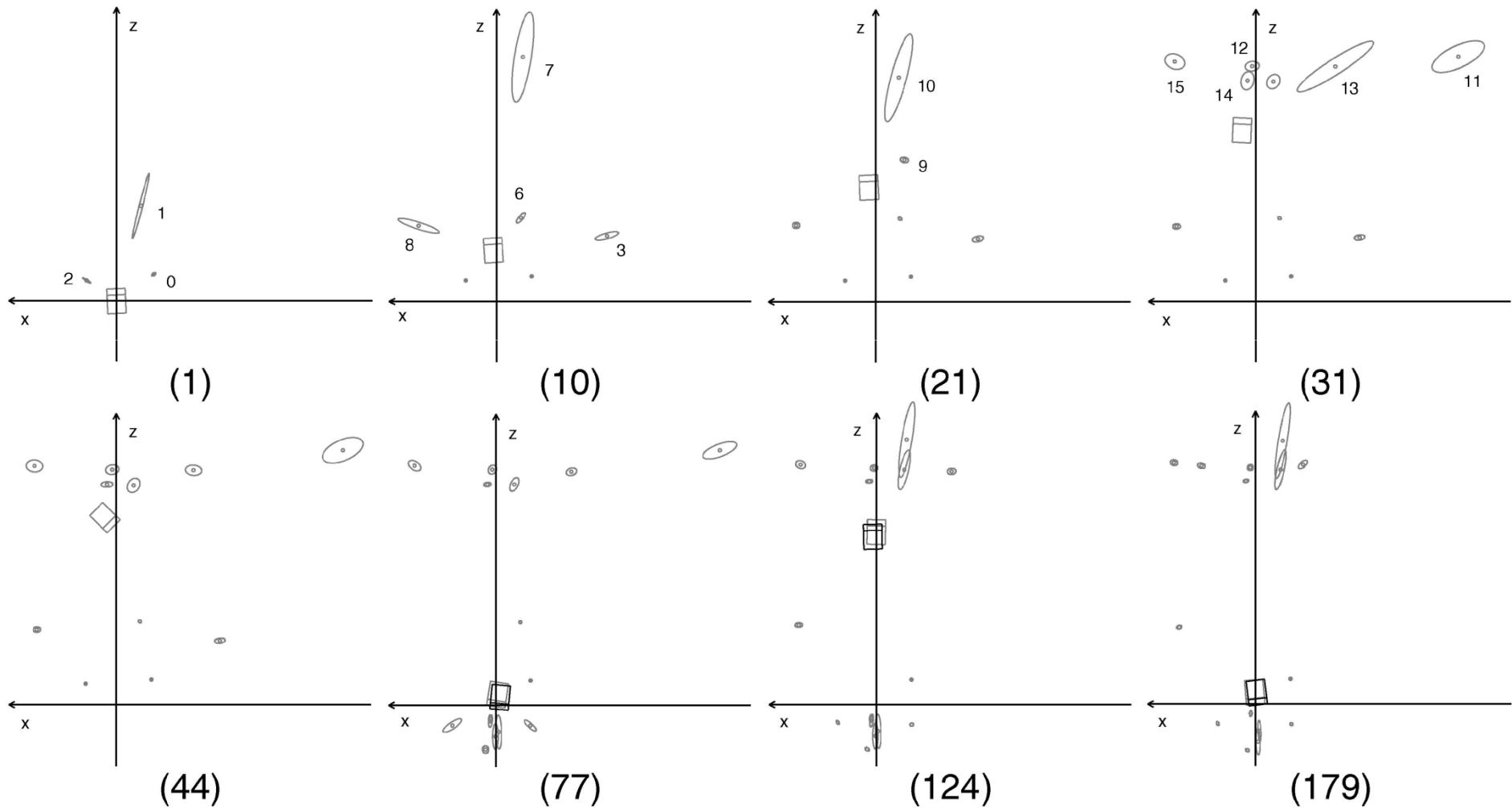
SLAM

- ▶ Zustandsvektor enthält alle Informationen
 - ▶ Roboter $\mathbf{x}_v = (\hat{z}, \hat{x}, \hat{\phi})^T$
 - ▶ Features $\mathbf{\hat{y}}_i = (\hat{X}_i, \hat{Y}_i, \hat{Z}_i)^T$
 - ▶ Und weitere Komponenten möglich
 - ▶ Geschwindigkeit, ...
 - ▶ Verknüpfung verschiedener Messverfahren zu einem Zustandsvektor

SLAM

- ▶ Harris Corner Detection
 - ▶ 15x15 Pixel Patches
- ▶ Feature – Stereo – Matching
 - ▶ Position wird mit „Ungenauigkeitsangabe“ im Zustandsvektor gespeichert
 - ▶ Qualität des IP wird mit der „Verfolgbarkeit“ bewertet
- ▶ Matching
 - ▶ Innerhalb einer Bildregion (Ellipse)
 - ▶ Wird anhand der Bewegung abgeschätzt

SLAM



Monte Carlo Localization

- ▶ Robust Vision-based Localization for Mobile Robots Using an Image Retrieval System Based on Invariant Features
 - ▶ J. Wolf, W. Burgard, H. Burkhardt, Universität Hamburg und Universität Freiburg



Monte Carlo Localization

- ▶ Beruht auf der Wahrscheinlichkeit der eigenen Position
 - ▶ Wird bei Bewegung neu berechnet
 - ▶ Angabe der Position und der Wahrscheinlichkeit, dass diese die „echte“ Position ist

- ▶ Zwei Schritte
 1. Prognose
 2. Korrektur

Monte Carlo Localization

- ▶ Bilddatenbank beinhaltet Vergleichsbilder
- ▶ „Ähnliche“ Bilder werden gesucht, wobei

$$M_i = gM(g|M)$$

- ▶ Position, in der M aufgenommen wurde ist bekannt
- ▶ Berechnet man g, kennt man die Position in der M_i aufgenommen wurde

Monte Carlo Localization

- ▶ Ähnlichkeit von Suchbild und Datenbankbild
 - ▶ Schnittmenge über die „globalen Features“



Literatur

Object Recognition from Local Scale-Invariant Features

David G. Lowe (1999)

Local Feature View Clustering for 3D Object Recognition

David G. Lowe (2001)

Distinctive Image Features from Scale-Invariant Keypoints

David G. Lowe (2004)

Vision-based Mobile Robot Localization And Mapping using Scale-Invariant Features

Stephen Se, David Lowe, Jim Little (2001)

http://en.wikipedia.org/wiki/Scale-Invariant_Feature_Transform

Mai 2008

<http://user.cs.tu-berlin.de/~nowozin/autopano-sift>

Mai 2008

<http://www.cs.ubc.ca/~mbrown/autostitch/autostitch.html>

Mai 2008
