Advanced Robot Perception

Fortgeschrittene Konzepte der Wahrnehmung für Robotersysteme

Georg von Wichert, Siemens Corporate Technology

Nochmal 2. Hausaufgabe

- Berechnung der Kovarianzmatrix

$$\operatorname{Cov}(X) = \operatorname{E}\left((X-\mu)(X-\mu)^T\right) = \operatorname{E}(XX^T) - \mu\mu^T .$$

$$E(x) = \int xp(x)dx$$

$$E(X) = \sum XP(X)$$

Nochmal 2. Hausaufgabe

- Berechnung der Kovarianzmatrix

$$\mu = E(\mathbf{x}) = \int \mathbf{x} \cdot p(\mathbf{x}) d\mathbf{x}$$

$$\Sigma = E(\mathbf{x} - \mu) = \int (\mathbf{x} - \mu)^T (\mathbf{x} - \mu) \cdot p(\mathbf{x}) d\mathbf{x}$$

- In unserem Fall ist eine geschlossene Lösung schwierig

$$p(\mathbf{x}) = p(\mathbf{x}|z) = \ldots$$

Probabilistisches Sensormodell

Physikalisches Modell des Messvorgangs

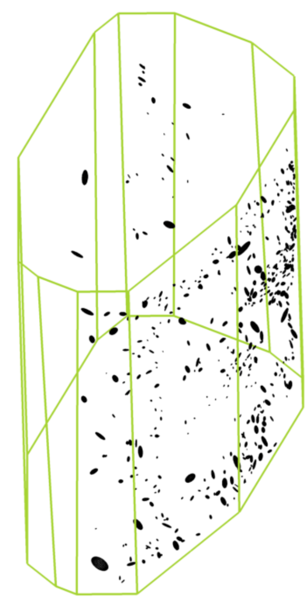
- Punktweise Projektion der SIFT-Merkmalspunkte in die Bilder der Beiden Kameras

$$\tilde{\mathbf{x}} = \begin{pmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R & \mathbf{t} \end{pmatrix} \tilde{\mathbf{p}}_w$$

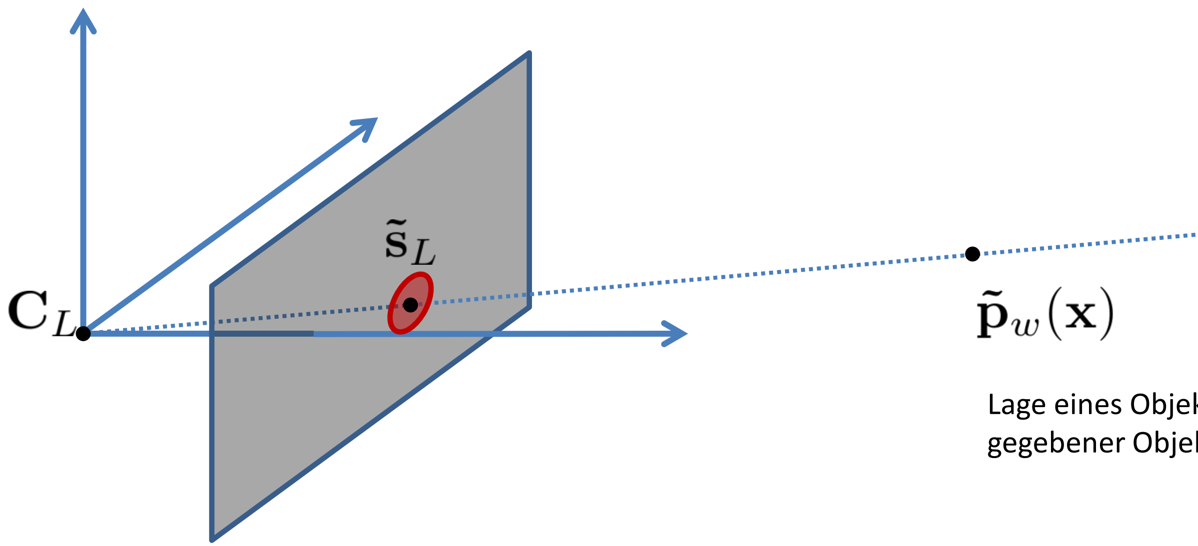
Wesentliche Fehlerquellen

- Fehler bei der Detektion der SIFT-Merkmalspunkte
- (Modellfehler im 3D Objektmodell)

→ müssen im Sensormodell berücksichtigt werden



Zur Herleitung des Messmodells



Lage eines Objektpunktes bei gegebener Objektpose **x**

$$\tilde{\mathbf{s}} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R & \mathbf{t} \end{pmatrix} \tilde{\mathbf{p}}_w + \nu \qquad \nu \sim N(0, Q)$$

Normalverteiltes Rauschen mit Kovarianzmatrix Q

Probabilistisches Sensormodell

Observation

Sample from object pose distribution

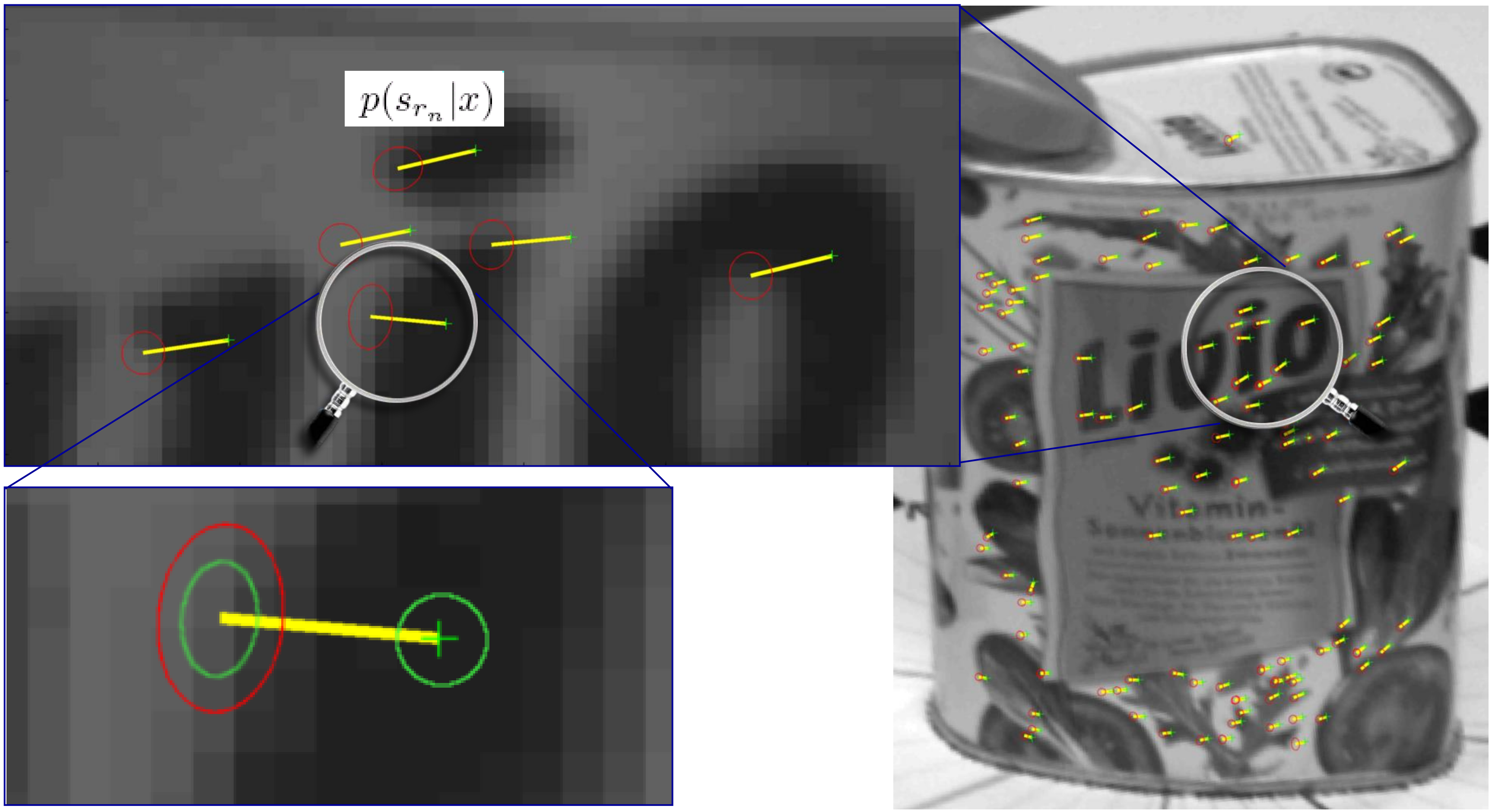
$$p(z|x) = \prod_{1}^{N} p(s_{l_n}|x)$$

Interest points in image

- Projektion der Modellfehler in die jeweilige Bildebene
- Gauß’sche Modellierung des Detektionsfehlers
- Alle Features werden als statistisch unabhängig angenommen

Probabilistisches Sensormodell

- **Beispiel für eine Teilmessung**



Probabilistisches Sensormodell

- **Vergleich zweier Lagehypothesen**

Nochmal 2. Hausaufgabe

- Berechnung der Kovarianzmatrix

$$\mu = E(\mathbf{x}) = \int \mathbf{x} \cdot p(\mathbf{x}) d\mathbf{x}$$

$$\Sigma = E(\mathbf{x} - \mu) = \int (\mathbf{x} - \mu)^T (\mathbf{x} - \mu) \cdot p(\mathbf{x}) d\mathbf{x}$$

- In unserem Fall ist eine geschlossene Lösung schwierig

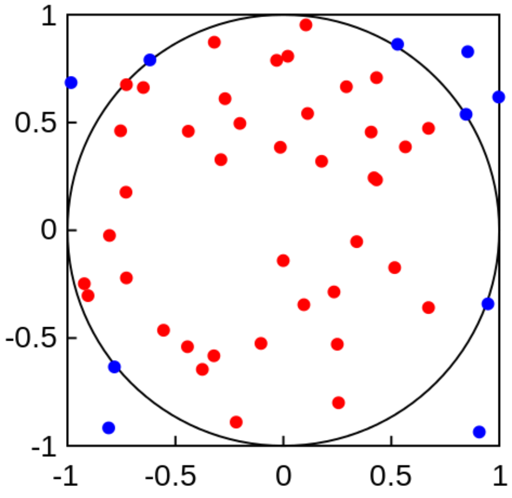
$$p(\mathbf{x}) = p(\mathbf{x}|z) = \ldots$$

Monte-Carlo-Integration

- Ersatz des Integrals durch eine Summe über eine gleichverteilte Stichprobe

$$I = \int_\Omega f(\overline{\mathbf{x}})\, d\overline{\mathbf{x}} \quad \Longrightarrow \quad I \approx Q_N \equiv V \frac{1}{N} \sum_{i=1}^{N} f(\overline{\mathbf{x}}_i) = V \langle f \rangle$$

$$V = \int_\Omega d\overline{\mathbf{x}} \qquad \overline{\mathbf{x}}_1, \cdots, \overline{\mathbf{x}}_N \in \Omega,$$



Vorletzes Mal: Aktive Wahrnehmung

- Wahrnehmungsprozess immer fehlerbehaftet
  - „Signal-Rausch-Verhältnis“ in der Robotik immer sehr niedrig, d.h. massive Störungen
  - Expliziter Umgang mit den Fehlern!
  - Bayes-Filter als Modell für Wahrnehmungsvorgang unter Unsicherheit

$$\boxed{Bel(x_t) = \underbrace{\eta\ P(z_t \mid x_t)}_{\text{Korrektur/ Mess-Update}} \underbrace{\int P(x_t \mid u_t, x_{t-1})\ Bel(x_{t-1})\ dx_{t-1}}_{\text{Prädiktion / Zeit-Update}}}$$

Aktive Wahrnehmung

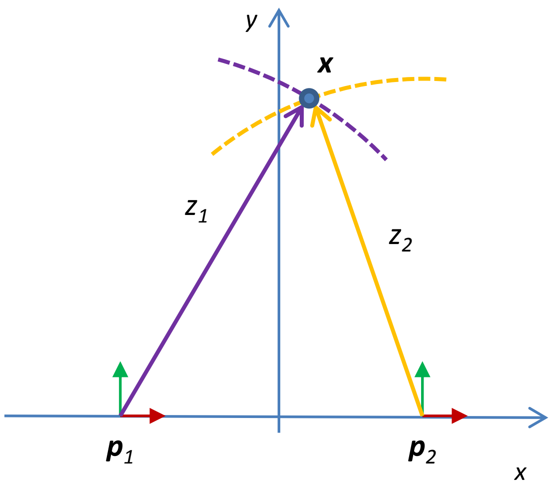
- Was kann ich als Maschine tun, um meine Wahrnehmungsergebnisse zu verbessern?

- Speziell in der Robotik: Roboter sind beweglich
  - Aktive Steuerung der Wahrnehmung möglich
  - Beispiel: Einnehmen eines „guten“ Standpunktes
  - Beispiel: „Gute“ Konfiguration der Sensoren

- Was heißt hier „gut“?

Beispiel: Datenfusion mit dem Extended Kalman Filter

- Beobachtung eines beweglichen Objekts in der Ebene mit zwei ungerichteten Abstandsensoren



$$\mathbf{p}_i = \begin{bmatrix} p_{i1} \\ p_{i2} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \quad \mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$z_i = \|\mathbf{x} - \mathbf{p}_i\| = \sqrt{(x_1 - p_{i1})^2 + (x_2 - p_{i2})^2}$$

Sensormodell:

$$\mathbf{z}_t = \mathbf{h}(\mathbf{x}_t) + \nu_t = \begin{bmatrix} z_1(\mathbf{x}_t) \\ z_2(\mathbf{x}_t) \end{bmatrix} + \nu_t$$

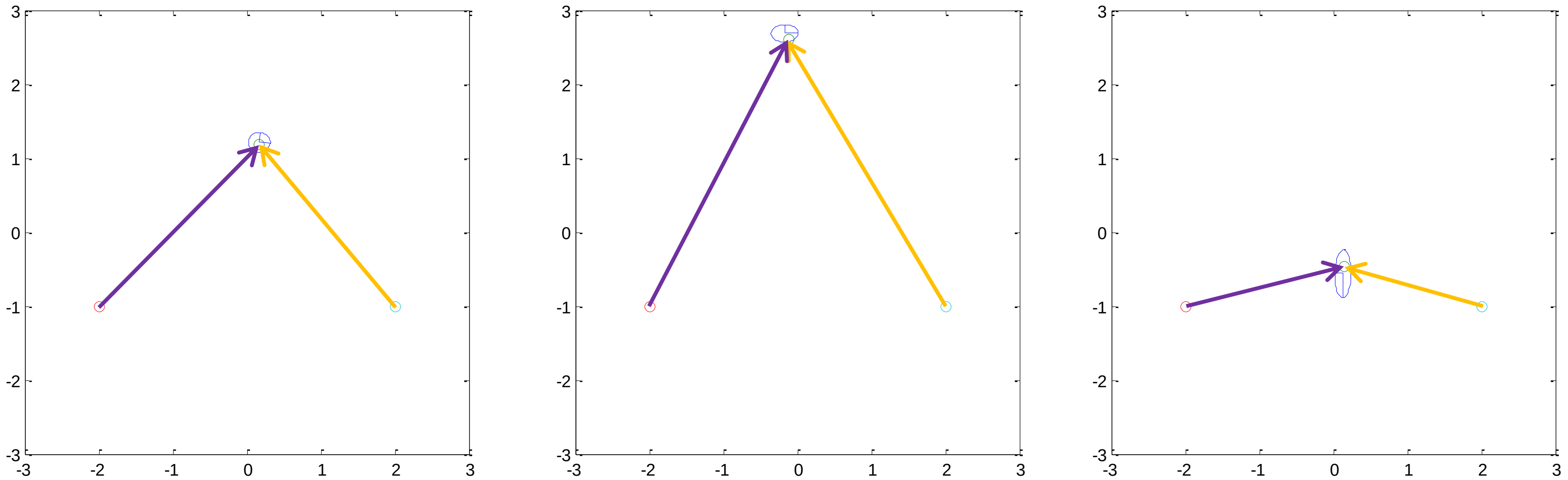
Messrauschen: $\nu_t \sim \mathcal{N}(0, \mathrm{Q})$

Prozessmodell:

$$\mathbf{x}_t = \mathbf{g}(\mathbf{x}_{t-1}, \mathbf{u}_t) + \xi_t = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{t-1} + \xi_t$$

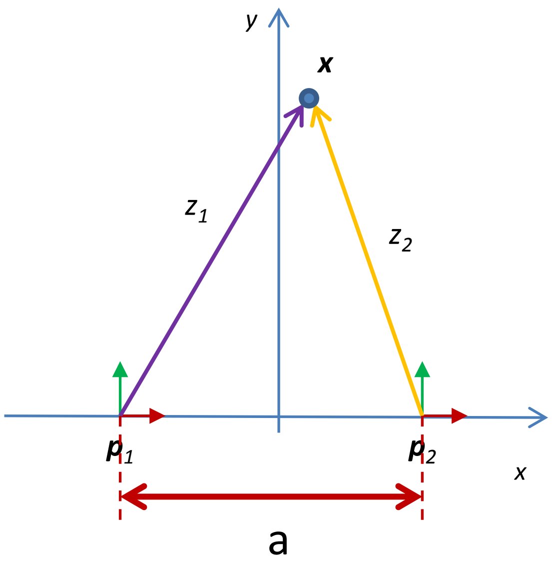
Prozessrauschen: $\xi_t \sim \mathcal{N}(0, \mathrm{R})$

Beispiel: Datenfusion mit dem Extended Kalman Filter



- Beobachtung eines beweglichen Objekts in der Ebene mit zwei ungerichteten Abstandsensoren
- Wie sieht die ideale Konfiguration aus?
  - Rechtwinkliges Dreieck

Aktive Wahrnehmung



- Steuerung von a zur Laufzeit
- Denkbare Stellstrategie:
  - Mögl. rechtwinkliges Dreieck herstellen!
  - Natürlich nur bei diesem Aufbau sinnvoll!
- Wir suchen ein grundlegendes Prinzip für die Steuerung der Wahrnehmungsparameter bzw. Wahrnehmungsaktionen!
- Frage: Können wir quantifizieren, wie viel wir die **Unsicherheit** über den unsicheren System/Umweltzustand durch die jeweilige Messung verringern?
  - Maximieren!

Informationstheorie

- Entwickelt für Kommunikationstechnik
  - "A Mathematical Theory of Communication“, 1948
    - Zentrales Thema: Übertragung von Information über einen gestörten Kanal
    - Applikationen in Kodierung und Kryptographie
- Wir werden das hier ohne Herleitung einführen
  - Details: Spezialvorlesung, **Lehrstuhl für Nachrichtentechnik,** Prof. Kramer
  - Wir wollen hier nur sehen, dass/wie wir die Informationstheorie für unsere Zwecke nutzen können!

Claude Elwood Shannon
30.4.1916 – 24.2.2001

Entropie

- Entropie ist der **mittlere Informationsgehalt** eines Ereignisses

$$H(X) = E_x[I(x)] = \sum_k P(X = x_k) I(x_k) = -\sum_k p_k \log p_k$$

- Das Äquivalent für kontinuierliche Zufallsvariablen ist die differentielle Entropie

$$H(X) = -\int p(x) \log(p(x)) \; dx$$

- Die Entropie ist ein Maß für die Unsicherheit der Verteilung p(x).
- Sie ist maximal, wenn $p(x)$ eine Gleichverteilung ist.

Bedingte (differentielle) Entropie

- Nehmen wir an, wir haben zwei Zufallsvariablen *x* und *z* mit einer Verbundverteilung *p*(*x*, *z*)
- Wie groß ist die Entropie von *x*, wenn wir *z* kennen?

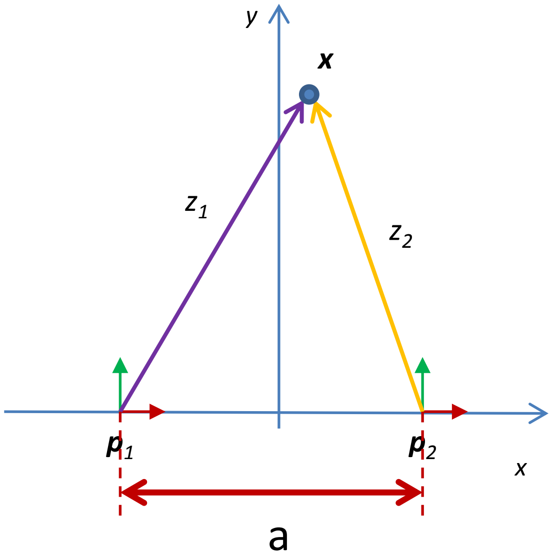
$$H(X|Z) = -\iint p(x,z) \log(p(x|z)) \; dx \; dz$$

$$= -\iint p(x|z)p(z) \log(p(x|z)) \; dx \; dz$$

$$= -\int p(z) \underbrace{\int p(x|z) \log(p(x|z)) \; dx}_{H(X|z)} \; dz = \int p(z) H(X|z) \; dz = E_z(H(X|z))$$

z.B. Entropie von *x*, nachdem wir eine bestimmte Messung *z* gemacht haben

Aktive Wahrnehmung



- Steuerung von a zur Laufzeit
- Denkbare Stellstrategie:
  - Mögl. rechtwinkliges Dreieck herstellen!
  - Natürlich nur bei diesem Aufbau sinnvoll!
- Wir suchen ein grundlegendes Prinzip für die Steuerung der Wahrnehmungsparameter bzw. Wahrnehmungsaktionen!
- Minimieren der Entropie nach der Messung

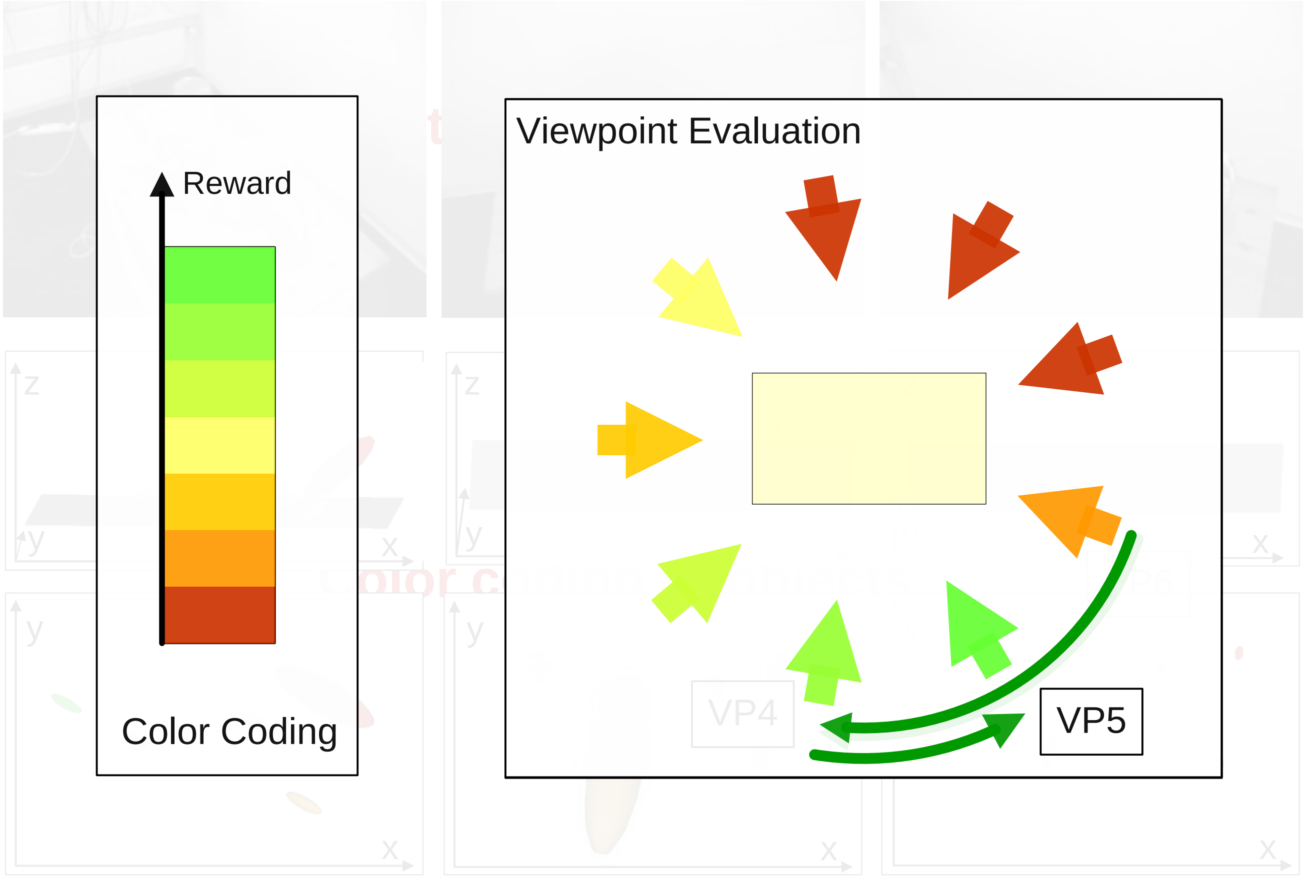
$$H_a(X_t \mid Z_t) \quad \text{wobei} \quad bel(x_t) = p(x_t \mid z_t(a))$$

durch **geeignete Wahl** der Aktion ***a***

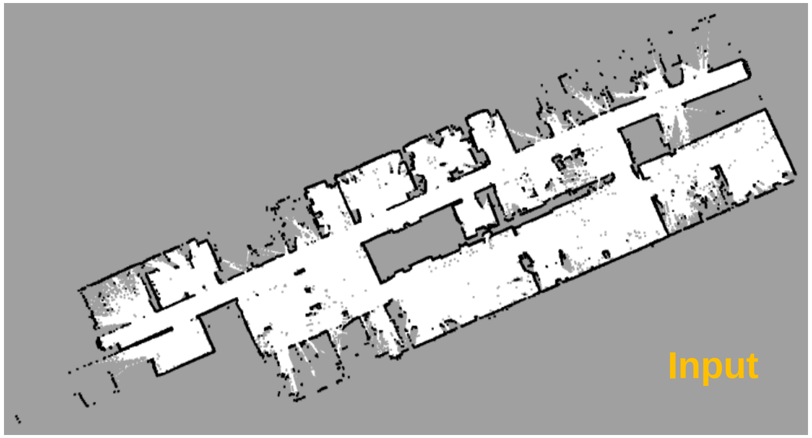
Active Perception for Robots: Example

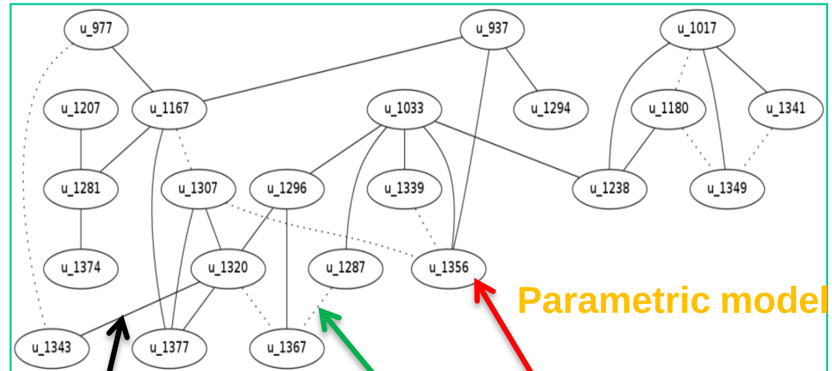
Initial scenario

Observations



Results using weak context knowledge as prior

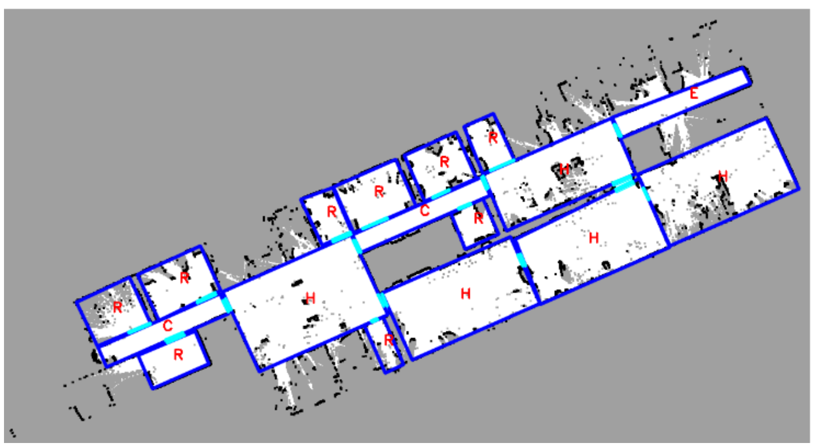


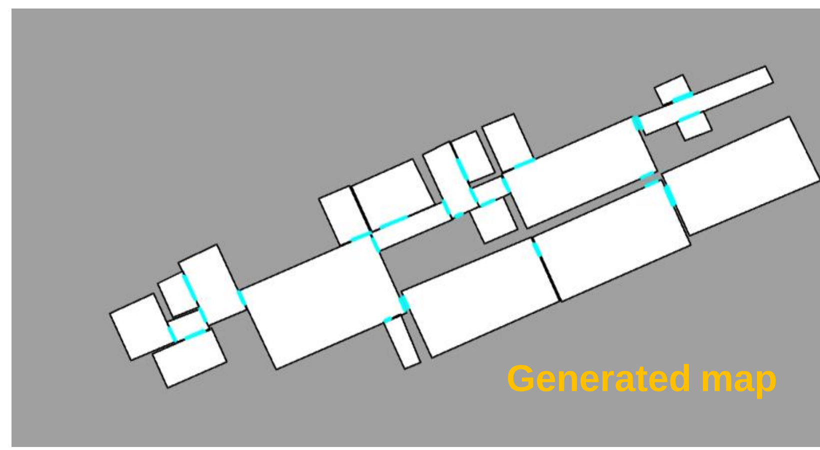


Solid edge: connected by a door

Dashed edge: connected by a wall

Node: unit represented by parameters



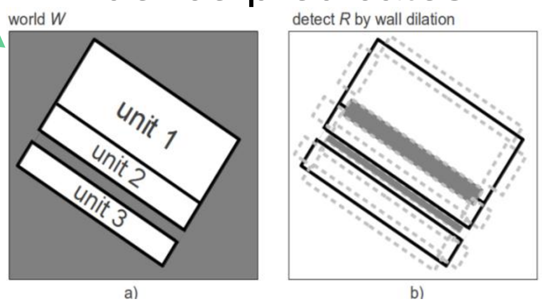


Knowledge reasoning in MLNs

size \ ratio	small	big
small	room-like	corridor-like
big	hall-like	hall-like

predicate	explanation
$RoLi(u_p)$	Unit u_p has a room-like geometry.
$CoLi(u_p)$	Unit u_p has a corridor-like geometry.
$HaLi(u_p)$	Unit u_p has a hall-like geometry.
$MulDoor(u_p)$	Unit u_p has multiple doors.
$Adj(u_p, u_q)$	Unit u_p and u_q are adjacent.

Evidence predicates



predicate	explanation
$Room(u_p)$	Unit u_p has the type of room.
$Corr(u_p)$	Unit u_p has the type of corridor.
$Hall(u_p)$	Unit u_p has the type of hall.
$Other(u_p)$	Unit u_p has the type of other.
$SaLe(u_p, u_q)$	Unit u_p and u_q have each a wall with the same length.

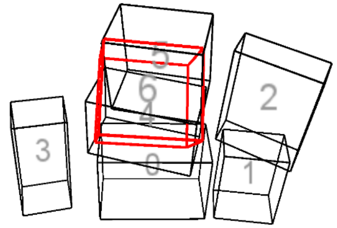
Query predicates

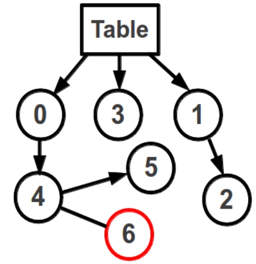
basic features:
$Adj(u_p, u_q) \rightarrow Adj(u_q, u_p)$
$SaLe(u_p, u_q) \rightarrow SaLe(u_q, u_p)$
reasoning on type:
$HaLi(u_p) \rightarrow Hall(u_p)$
$HaLi(u_p) \rightarrow \neg Room(u_p)$
$HaLi(u_p) \rightarrow \neg Corr(u_p)$
$HaLi(u_p) \rightarrow \neg Other(u_p)$
$RoLi(u_p) \rightarrow \neg Hall(u_p)$
$CoLi(u_p) \rightarrow \neg Hall(u_p)$
$RoLi(u_p) \wedge \neg MulDoor(u_p) \rightarrow Room(u_p)$
$RoLi(u_p) \wedge \neg MulDoor(u_p) \rightarrow \neg Corr(u_p)$
$RoLi(u_p) \wedge MulDoor(u_p) \rightarrow Other(u_p)$
$CoLi(u_p) \wedge \neg MulDoor(u_p) \rightarrow Other(u_p)$
$CoLi(u_p) \wedge MulDoor(u_p) \rightarrow Corr(u_p)$
$CoLi(u_p) \wedge MulDoor(u_p) \rightarrow \neg Room(u_p)$
reasoning on $SaLe$:
$\neg Adj(u_q, u_p) \rightarrow \neg SaLe(u_p, u_q)$
$Room(u_p) \wedge Room(u_q) \wedge Adj(u_p, u_q) \rightarrow SaLe(u_p, u_q)$
$Room(u_p) \wedge Hall(u_q) \wedge Adj(u_p, u_q) \rightarrow \neg SaLe(u_p, u_q)$
$Room(u_p) \wedge Corr(u_q) \wedge Adj(u_p, u_q) \rightarrow \neg SaLe(u_p, u_q)$
$Hall(u_p) \wedge Corr(u_q) \wedge Adj(u_p, u_q) \rightarrow \neg SaLe(u_p, u_q)$
$Other(u_p) \wedge Hall(u_q) \wedge Adj(u_p, u_q) \rightarrow \neg SaLe(u_p, u_q)$
$Other(u_p) \wedge Corr(u_q) \wedge Adj(u_p, u_q) \rightarrow \neg SaLe(u_p, u_q)$
$Other(u_p) \wedge Room(u_q) \wedge Adj(u_p, u_q) \rightarrow \neg SaLe(u_p, u_q)$

Logic rules

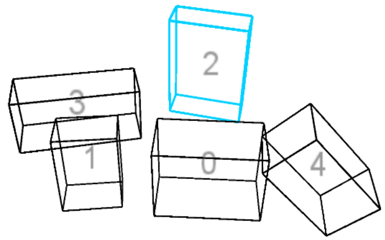
[Liu2013-IROS]

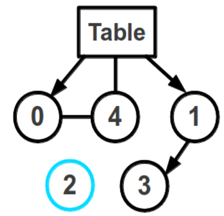
Experimental results



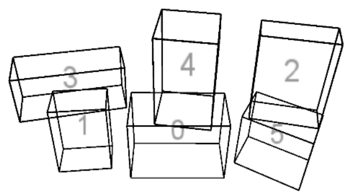


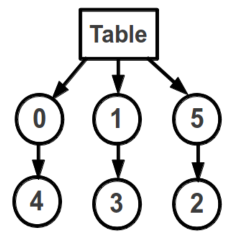
false(0)=0.074, hidden(0)=0.406
false(1)=0.036, hidden(1)=0.424
false(2)=0.092, hidden(2)=0.384
false(3)=0.102, hidden(3)=0.386
false(4)=0.094, hidden(4)=0.400
false(5)=0.130, hidden(5)=0.348
false(6)=0.912, hidden(6)=0.466



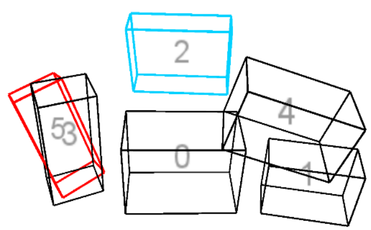


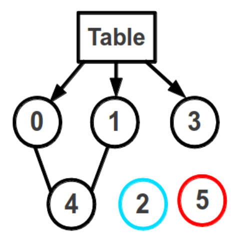
false(0)=0.082, hidden(0)=0.428
false(1)=0.044, hidden(1)=0.432
false(2)=0.150, hidden(2)=0.816
false(3)=0.110, hidden(3)=0.382
false(4)=0.496, hidden(4)=0.478





false(0)=0.070, hidden(0)=0.416
false(1)=0.036, hidden(1)=0.454
false(2)=0.158, hidden(2)=0.384
false(3)=0.110, hidden(3)=0.400
false(4)=0.152, hidden(4)=0.374
false(5)=0.074, hidden(5)=0.358





false(0)=0.096, hidden(0)=0.404
false(1)=0.088, hidden(1)=0.406
false(2)=0.136, hidden(2)=0.808
false(3)=0.138, hidden(3)=0.386
false(4)=0.492, hidden(4)=0.506
false(5)=0.914, hidden(5)=0.460

Wahrnehmung als probabilisticher

- Wahrnehmungsprozess immer fehlerbehaftet
  - Expliziter Umgang mit den Fehlern!
  - Bayes-Filter als Modell für Wahrnehmungsvorgang unter Unsicherheit
  - Aktive Informationsbeschaffung!

- Wichtig für autonome Systeme in unsicheren Umgebungen:
  Fähigkeit „Erwartungen“ zu formulieren
  - Grundvoraussetzung für stetiges Aktualisieren des Weltzustands auf Basis fehlerbehafteter Wahrnehmungen
  - Vermeidung von Haluzinationen

$$\boxed{Bel(x_t) = \underbrace{\eta\, P(z_t \mid x_t)}_{\text{Korrektur/ Mess-Update}} \underbrace{\int P(x_t \mid u_t, x_{t-1})\, Bel(x_{t-1})\, dx_{t-1}}_{\text{Prädiktion / Zeit-Update}}}$$

Fortgeschrittene Konzepte der Wahrnehmung für Robotersysteme

Basisthemen

Algorithmen-klassiker

- Sensortypen
- Lineare Algebra+Geometrie
- Segmentierung in Pointclouds
- Registrierung
- Grundlagen Bildverarbeitung
- Muster- und Objekterkennung
- Local Features

Wahrnehmung als Prozess

Umgang mit Unsicherheit

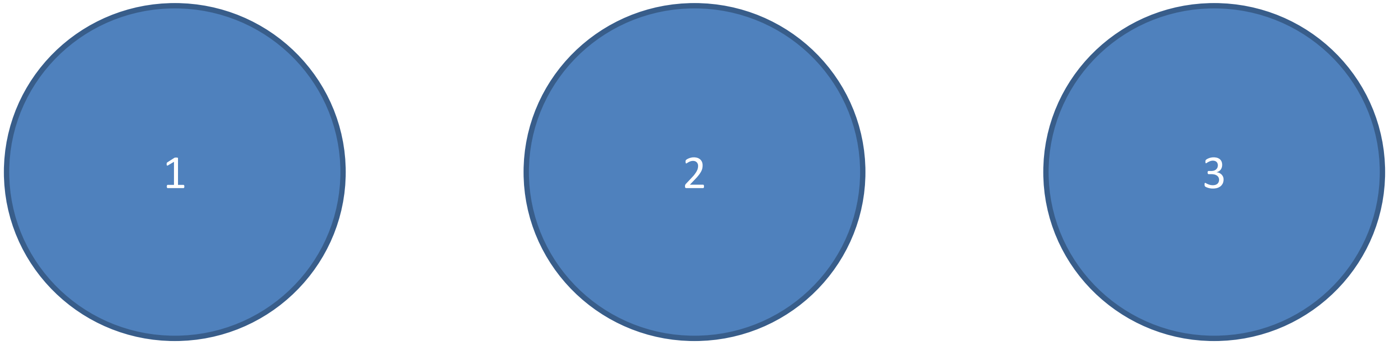
- Umgang mit Unsicherheiten / Probabilistik
- Zustandschätzung und Datenfusion
- Sequential Bayes und der Kalmanfilter
- Probabilistische Sensormodelle
- Grundlagen der Informationstheorie und aktive Wahrnehmung
- Offene Fragen

Fortgeschrittene Konzepte der Wahrnehmung für Robotersysteme

Was sollten Sie (mindestens) mitnehmen?

- Ein paar wichtige Basistechniken und grundlegende Zusammen-hänge für den Fall, dass Sie selber ein Wahrnehmungsproblem lösen (w/s)ollen
  - Betrachten Sie das als Pointer in die Literatur. Sie werden dort noch viele relevante Ideen und Konzepte finden
- Wahrnehmung ist ein kognitiver Vorgang / Denkprozess
  - Geeignete Modelle, die die wahrzunehmende Welt beschreiben sind erforderlich
  - Der explizite Umgang mit Unsicherheit extrem wichtig für die Robustheit
  - Probabilistische Konzepte wie das Bayes-Filter liefern deshalb gute Denkmodelle für die Beschäftigung mit Wahrnehmungsthemen
    - vgl. Mining im „Big Data“

EIN PAAR DENKANSTÖSSE ZUM ABSCHLUSS



$$p(\mathbf{x}|\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{x})\ p(\mathbf{x})}{p(\mathbf{z})} = \frac{p(\mathbf{z}|\mathbf{x})\ p(\mathbf{x})}{\int p(\mathbf{z}|\mathbf{x})\ p(\mathbf{x}) d\mathbf{x}}$$

$\mathbf{x}$ $\longrightarrow$ unser **mentales „Bild“** von der Welt

$\mathbf{z}$ $\longrightarrow$ unser **Eingansdatenstrom**