

03-05-H
-709.53

Echtzeitbildverarbeitung (14)

Prof. Dr. Udo Frese

Anwendungen von Partikel Filtern
Kognitive Bildverarbeitung
Rekapitulation des Prüfungsstoffes

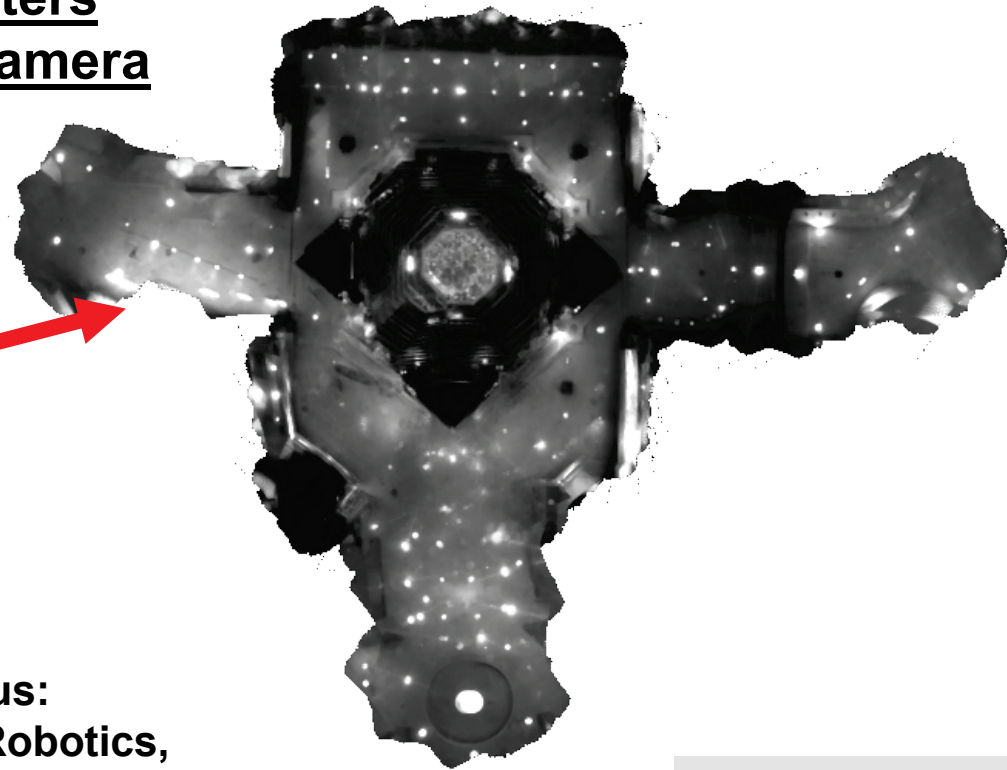
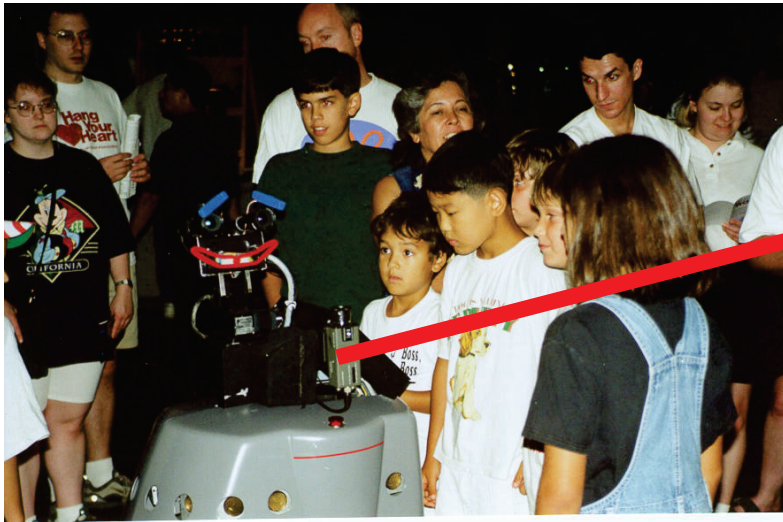
Was bisher geschah

- ▶ **Partikel Filter sind Bayesfilter**
- ▶ **repräsentieren a-posteriori Verteilung durch gewichtete Stichproben**
- ▶ **Dynamik: ziehen aus Zustandsübergangsverteilung**
 - ▶ praktisch: ideale Bewegung + gezogenes Rauschen
- ▶ **Messung: gewichten**
 - ▶ praktisch oft: Differenz zu idealer Messung in Fehlerverteilung
- ▶ **re-sampling: ziehen aus Partikeln mit Wahrscheinlichkeit proportional zu Gewicht**
 - ▶ systematisch, nicht unabhängig ziehen
- ▶ **initialisieren auf Grund der ersten (paar) Messungen**
- ▶ **jede Unsicherheit *einmal* einbringen, d.h. *entweder* zufällig ziehen oder gewichten.**

Anwendungen von Partikel Filtern

Anwendungen von Partikel Filtern

Lokalisation eines Museumsroboters mit einer zur Decke blickenden Kamera



Quelle: Dellaert et. al, 1999, entnommen aus:
Wolfram Burgard, Introduction to Mobile Robotics,
Lecture 9, Universität Freiburg, 2005

<http://ais.informatik.uni-freiburg.de/lehre/ss05/robotics/>

Video: Sebastian Thrun, Stanford, <http://robots.stanford.edu/>

Video 1

Anwendungen von Partikel Filtern

Frage an das Auditorium: Welche Vorzüge oder auch Nachteile hat eine zur Decke gerichtete Kamera im Vergleich zu einer nach vorne gerichteten Kamera oder einem Laserscanner?

Anwendungen von Partikel Filtern

Frage an das Auditorium: Welche Vorzüge oder auch Nachteile hat eine zur Decke gerichtete Kamera im Vergleich zu einer nach vorne gerichteten Kamera oder einem Laserscanner?

▶ **Vorteile:**

- ▶ keine Besucher im Blickfeld
- ▶ Decke verändert sich selten
- ▶ direkte Korrespondenz zwischen 2D Bild und 2D Roboterbewegung (bei bekannter Deckenhöhe).

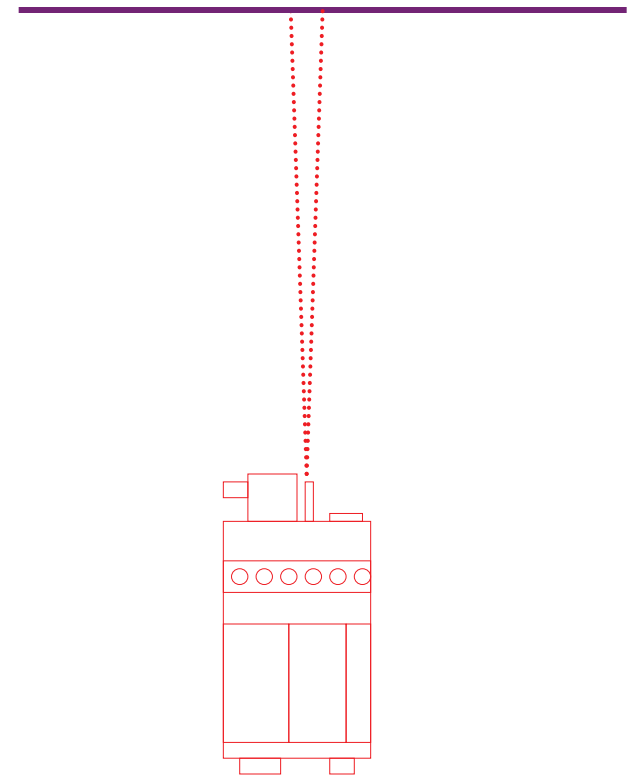
▶ **Nachteile:**

- ▶ sehr mehrdeutig
- ▶ Deckenlichter blenden Kamera.
- ▶ starker Unterschied zwischen Licht an/aus
- ▶ keine Hindernisvermeidung

Anwendungen von Partikel Filtern

Lokalisation eines Museumsroboters mit einer zur Decke blickenden Kamera

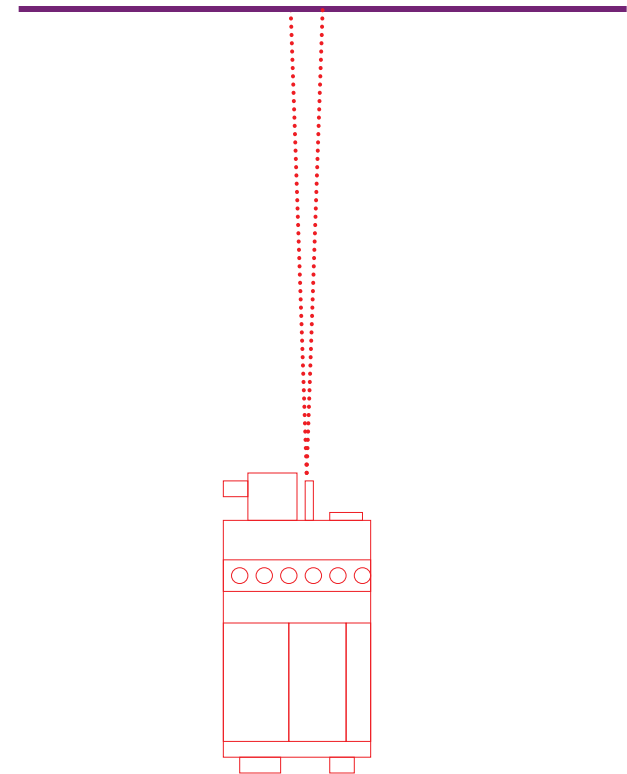
- ▶ Partikel Filters wegen Mehrdeutigkeiten
- ▶ ⇒ nicht Positionsbestimmung aus einem einzelnen Bild, wie Übungszettel 5
- ▶ ⇒ sondern aus Kontext der Bewegung verbunden durch Odometrie (Raddrehungen)



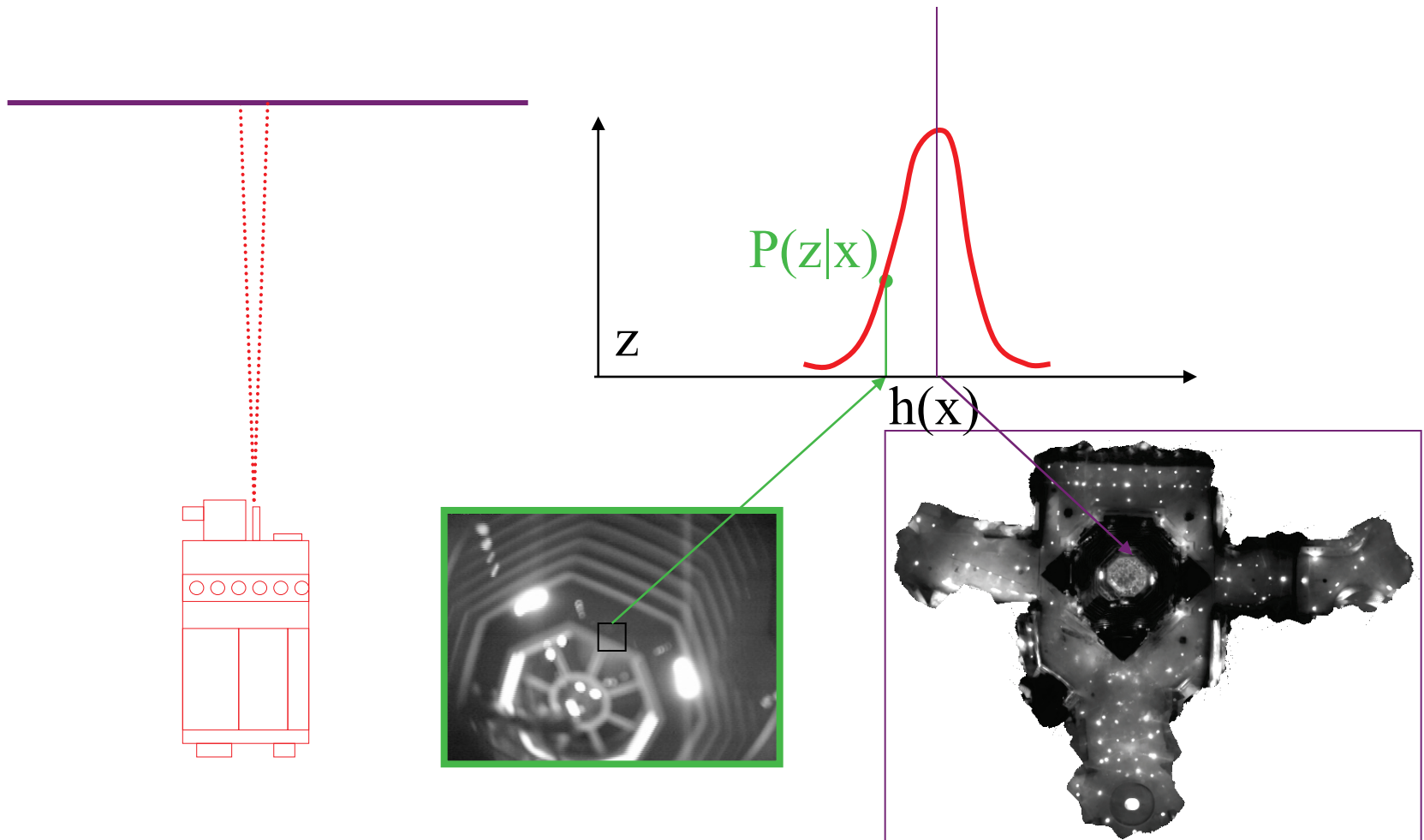
Anwendungen von Partikel Filtern

Partikelfilter für Lokalisation

- ▶ **Zustand $\mathbf{x} = (p_x, p_y, p_\Theta)$:**
 - ▶ Position (p_x, p_y)
 - ▶ Orientierung (p_Θ)
- ▶ **Dynamikmodell $p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u})$:**
 - ▶ Odometrie mit Rauschen
- ▶ **Idee:**
 - ▶ keine Merkmale aus Bild extrahieren, sondern direkt Bild vergleichen
 - ▶ kleiner Ausschnitt vermeidet perspektivische Effekte bei unterschiedlicher Deckenhöhe
- ▶ **Messmodell $p(\mathbf{z}_t | \mathbf{x}_t)$:**
 - ▶ Bildausschnitt von der Kamera
 - ▶ mittlerer quadratischer Abstand zum x entsprechenden Abschnitt einer Deckenkarte

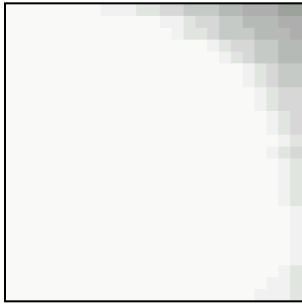


Anwendungen von Partikel Filtern

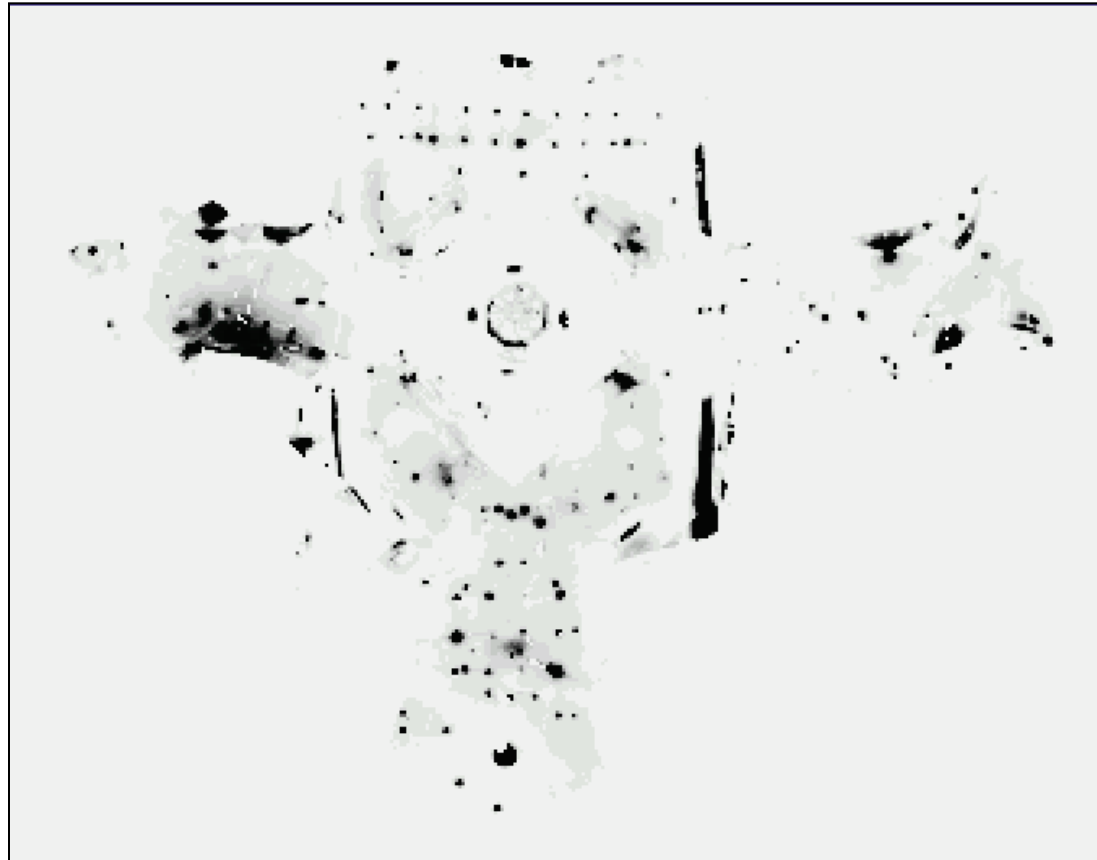


Anwendungen von Partikel Filtern

Messung z :

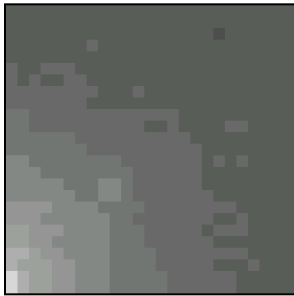


$P(z|x)$:

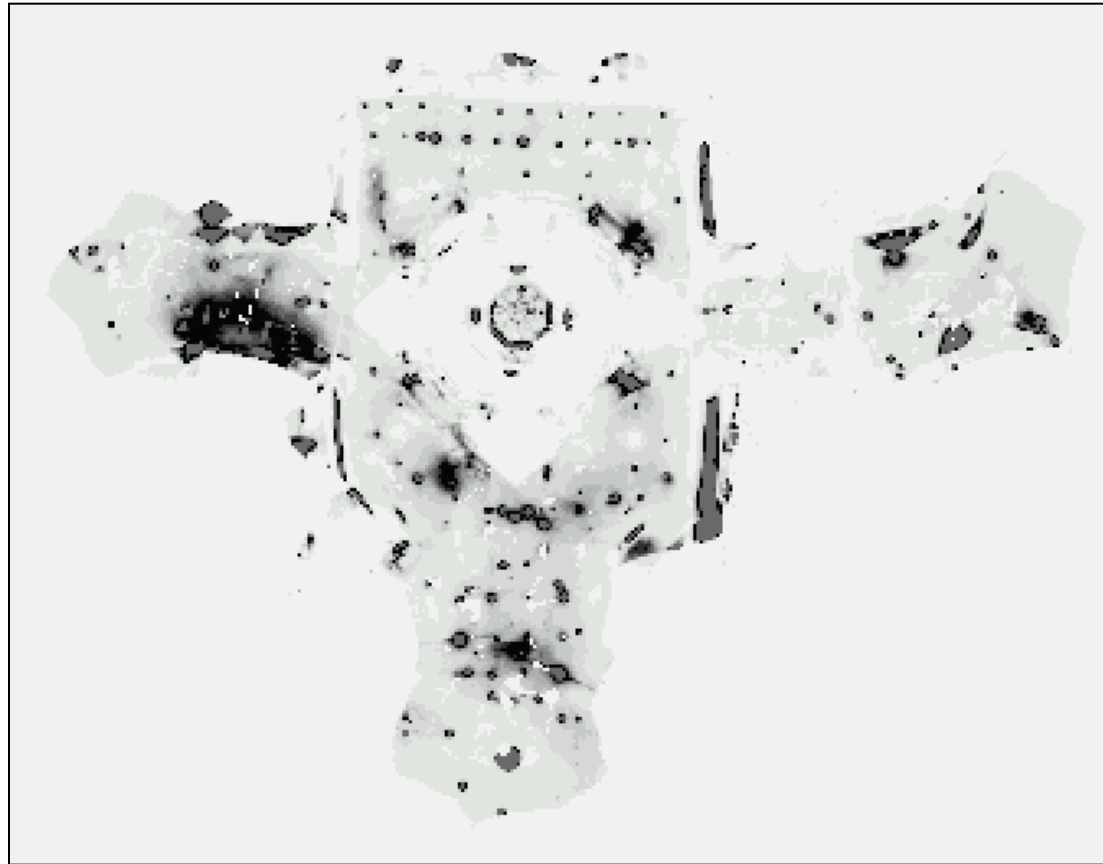


Anwendungen von Partikel Filtern

Messung z :



$P(z|x)$:



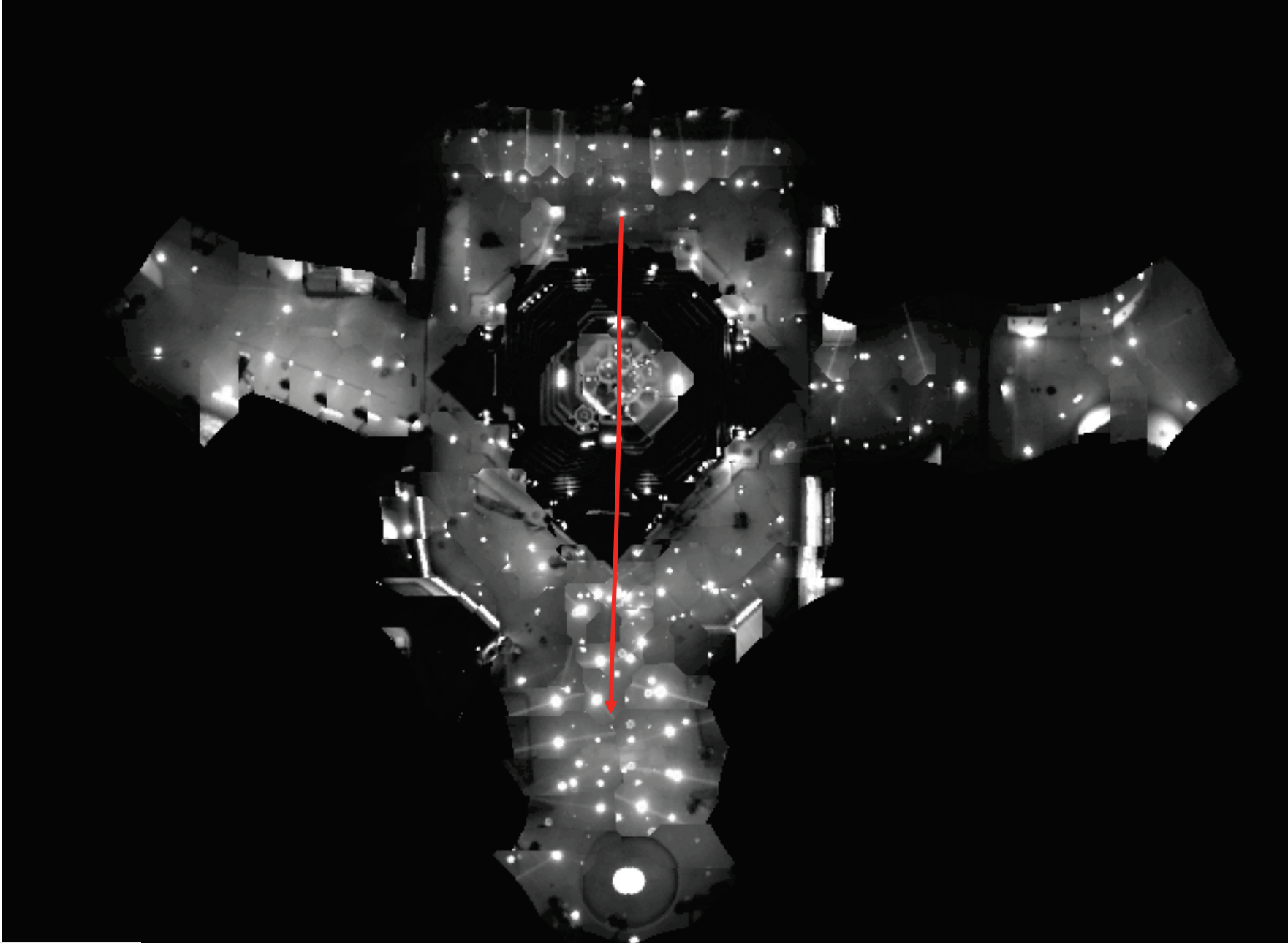
Anwendungen von Partikel Filtern

Messung z :



$P(z|x)$:



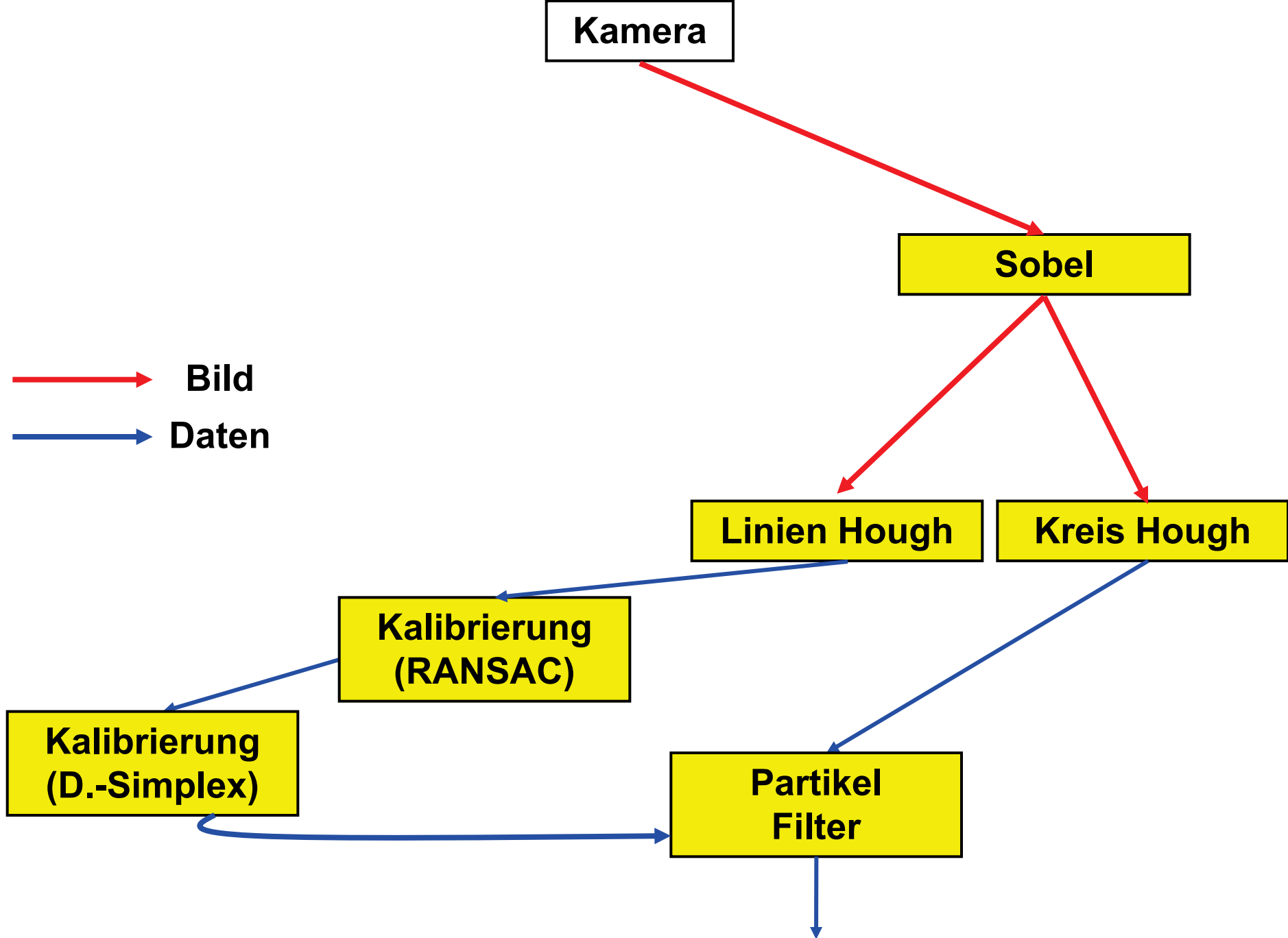


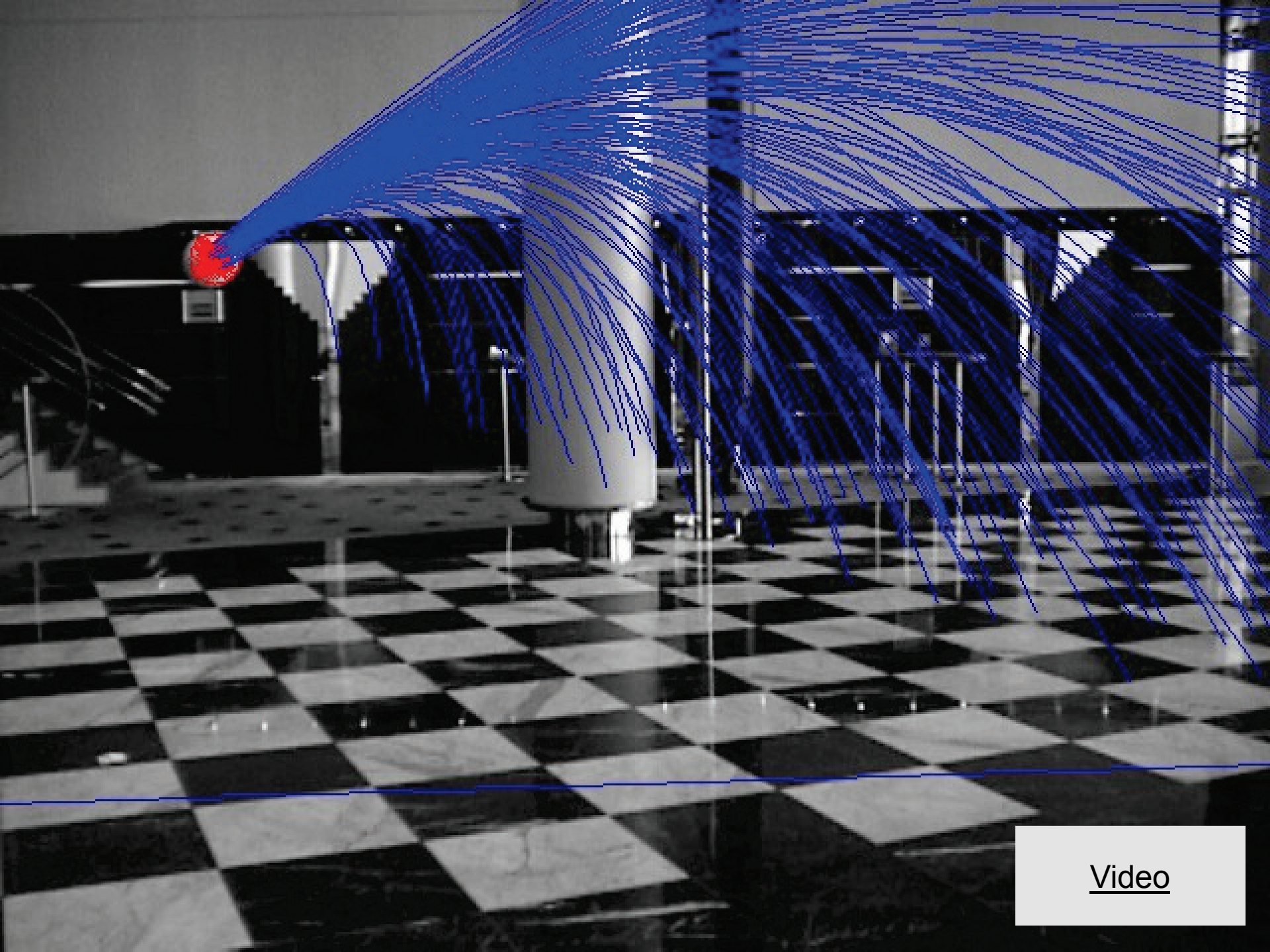
Video

▶ **Ballvorhersage Beispiel aus den Übungen**



Video

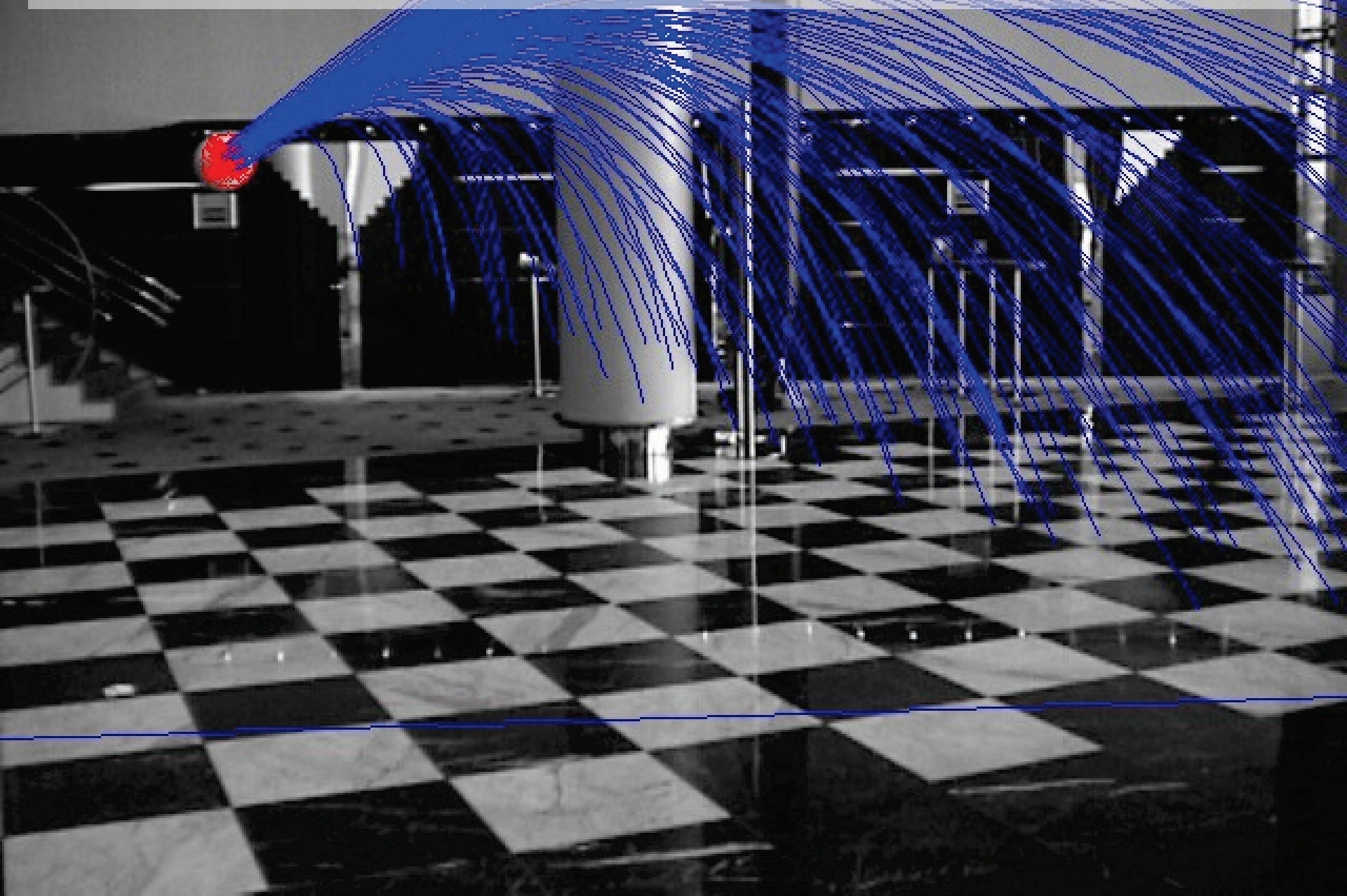




[Video](#)

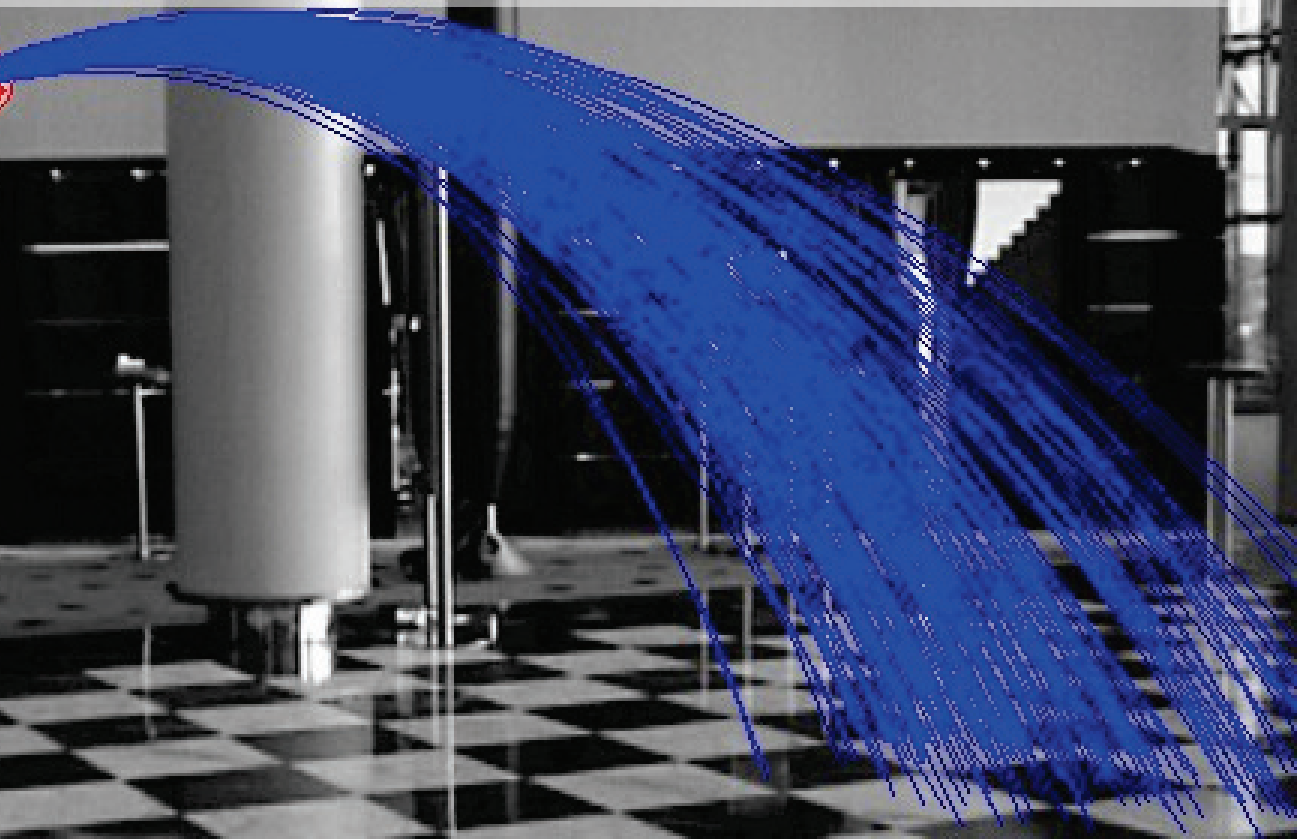


Frage an das Auditorium: Was könnt Ihr über die Struktur der Unsicherheit sagen? Warum ist sie so?

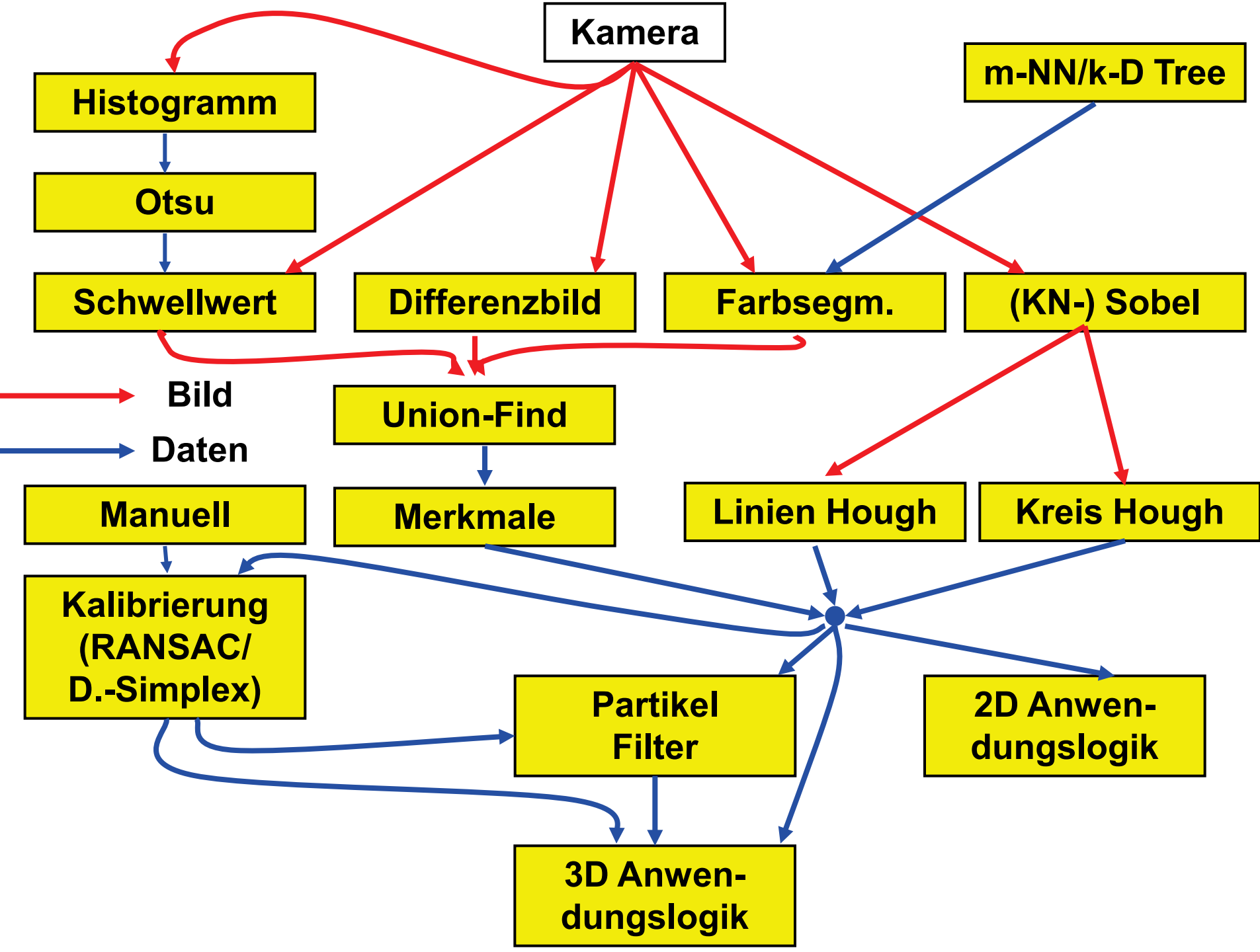


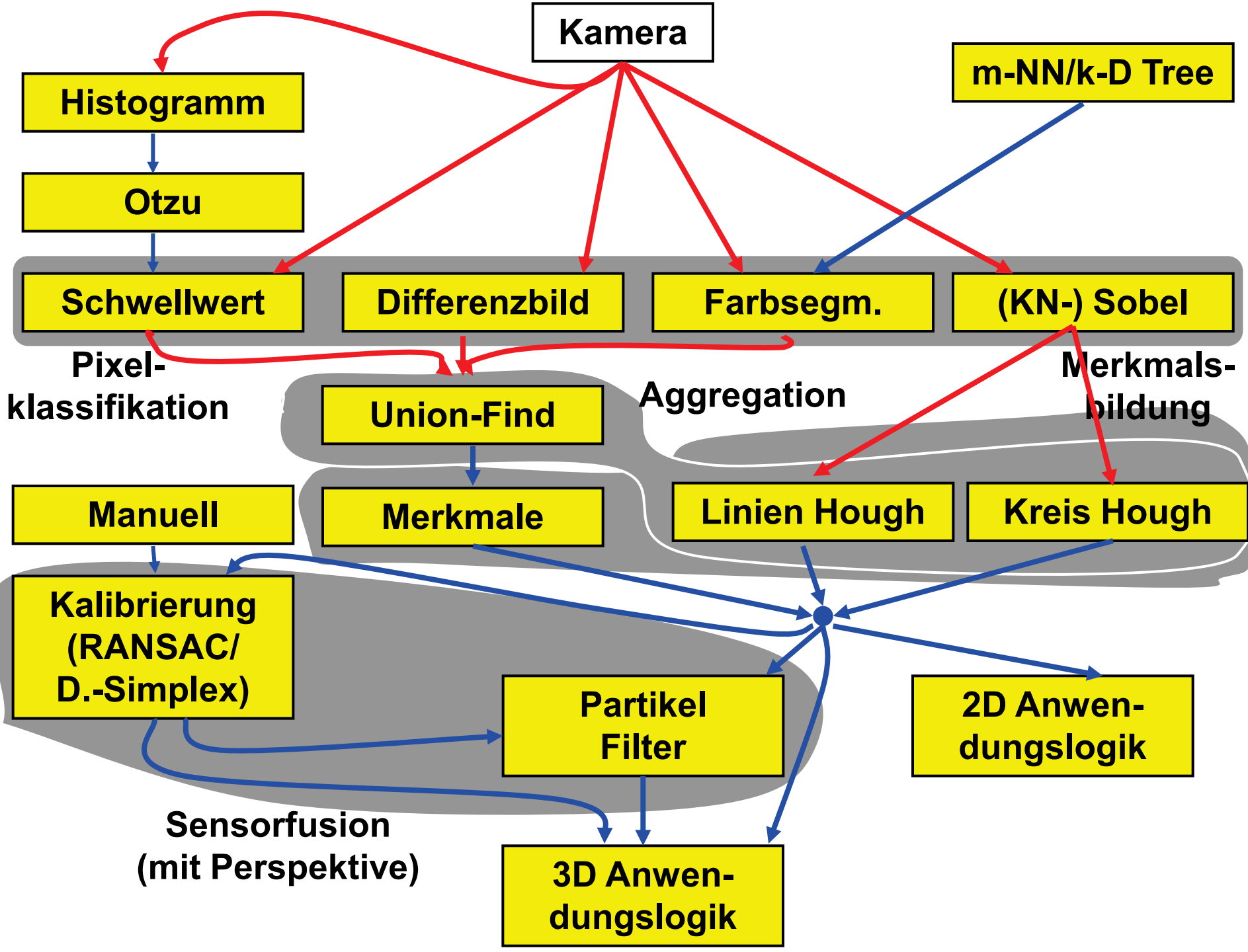
Frage an das Auditorium:

Warum ist die Unsicherheit jetzt so viel kleiner?



Kognitive Bildverarbeitung





Kognitive Bildverarbeitung

Frage an das Auditorium: Was fällt bei den Pfeilen auf?

Kognitive Bildverarbeitung

Frage an das Auditorium: Was fällt bei den Pfeilen auf?

- ▶ **alle nach unten!**
- ▶ **Informationsfluss verläuft**
 - ▶ von einfachen, bildnahen Merkmalen
 - ▶ durch verschiedene Erkennungsalgorithmen
 - ▶ zu höheren, objektnahen Merkmalen.
- ▶ **keine Rückkopplung.**
 - ▶ keinen Informationsfluss von objektnahen Merkmalen zurück zu bildnahen Merkmalen.
- ▶ **rein „datengetriebener“ Ansatz**

Kognitive Bildverarbeitung

Frage an das Auditorium:
Was ist das?

(Quelle: G. Binnig,
Definiens AG,
Vortrag 2004)


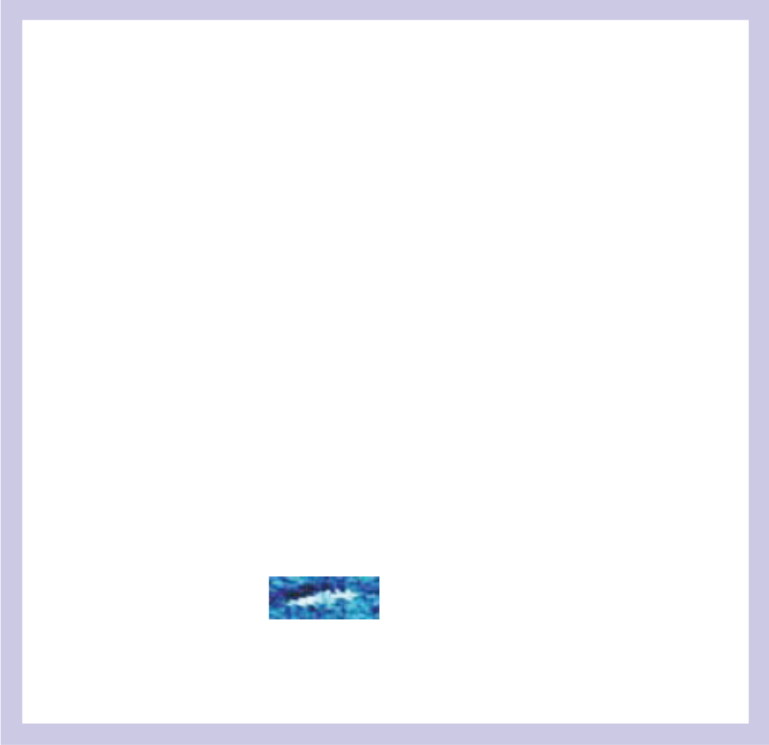
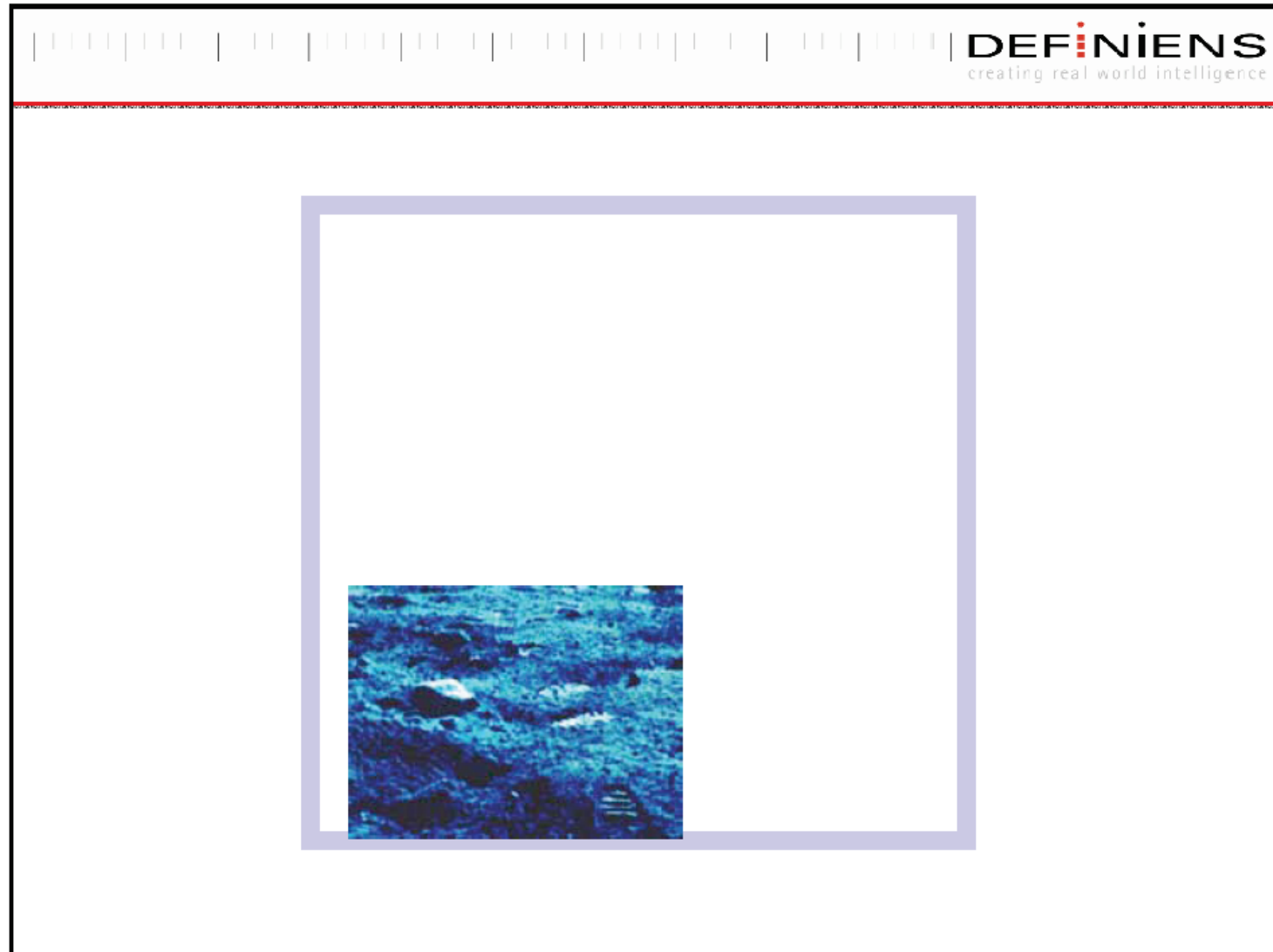


Image Understanding with Cognition Networks



Kognitive Bildverarbeitung

Frage an das Auditorium:
Was ist das?



Kognitive Bildverarbeitung

Frage an das Auditorium:
Was ist das?



Kognitive Bildverarbeitung

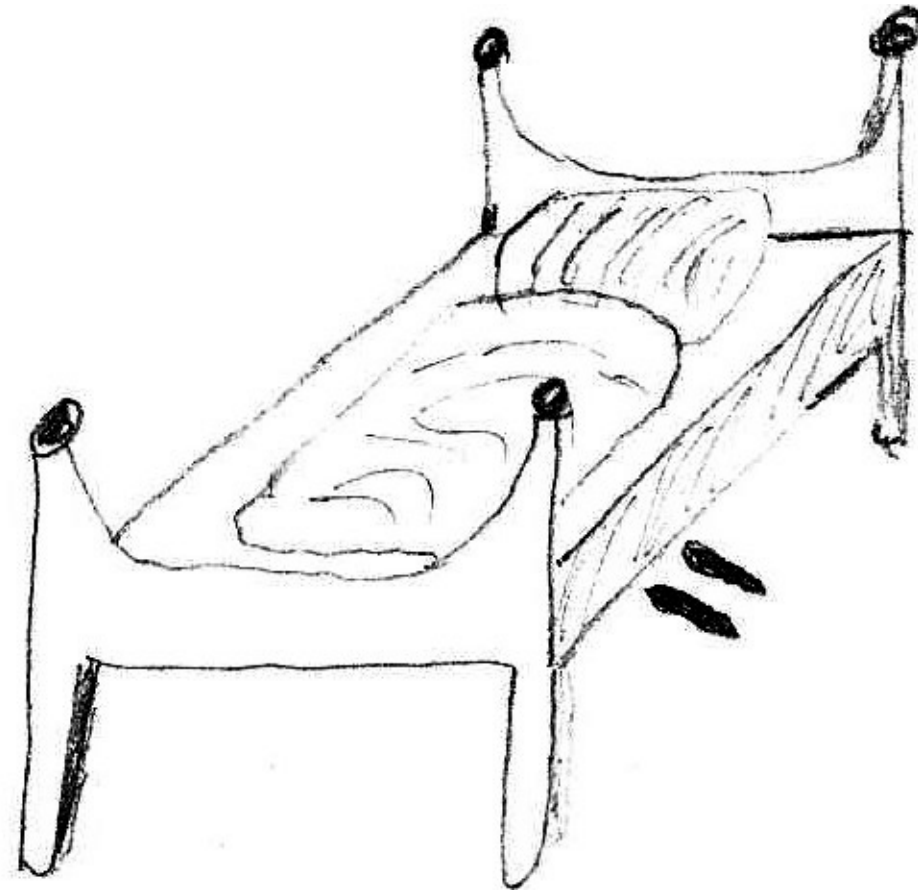
Frage an das Auditorium: Was ist das?

(Quelle (Idee): S. Ullman, Sequence-seeking and counter streams: A computational model for bi-directional information flow in the visual cortex. *Cerebral Cortex*, 5(1) 1-11)



Kognitive Bildverarbeitung

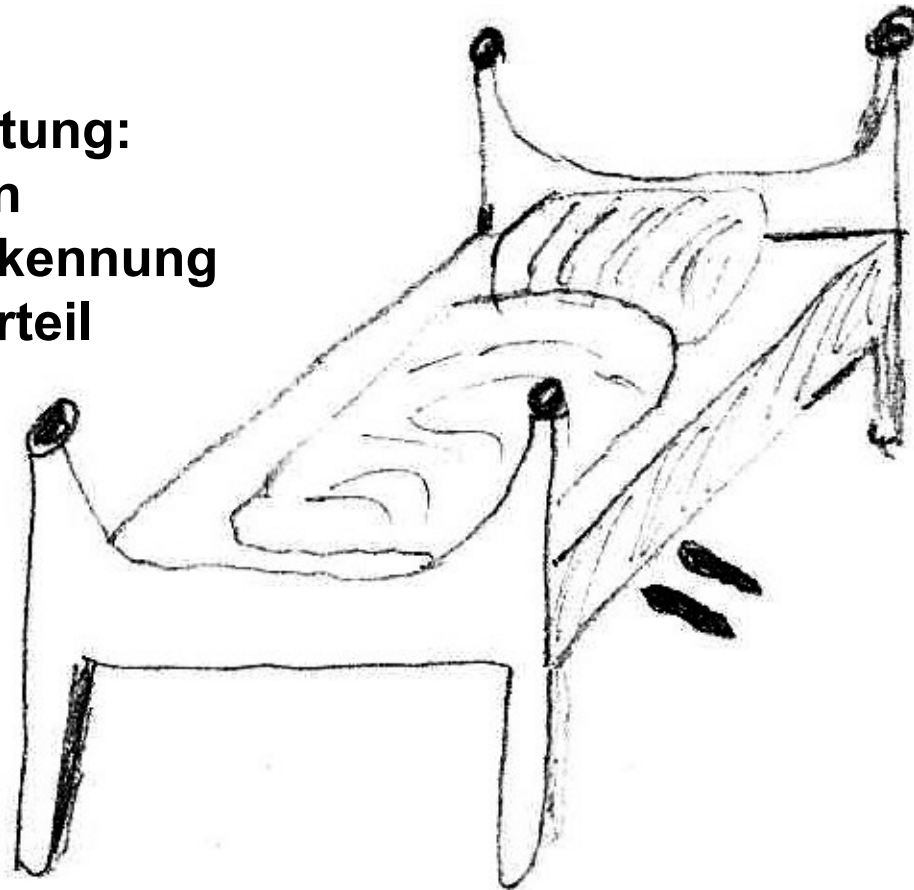
Frage an das Auditorium: Was ist das?



Kognitive Bildverarbeitung


Frage an das Auditorium: Was ist das?

- ▶ **Pantoffeln!**
- ▶ **These der Kognitiven Bildverarbeitung:**
“Die Leichtigkeit mit der Menschen auf allen Ebenen Kontext in die Erkennung einbringen, ist der wesentliche Vorteil gegenüber technischer Bildverarbeitung.“



Kognitive Bildverarbeitung

- ▶ Quelle:
G. Binnig,
Definiens AG,
Vortrag 2004

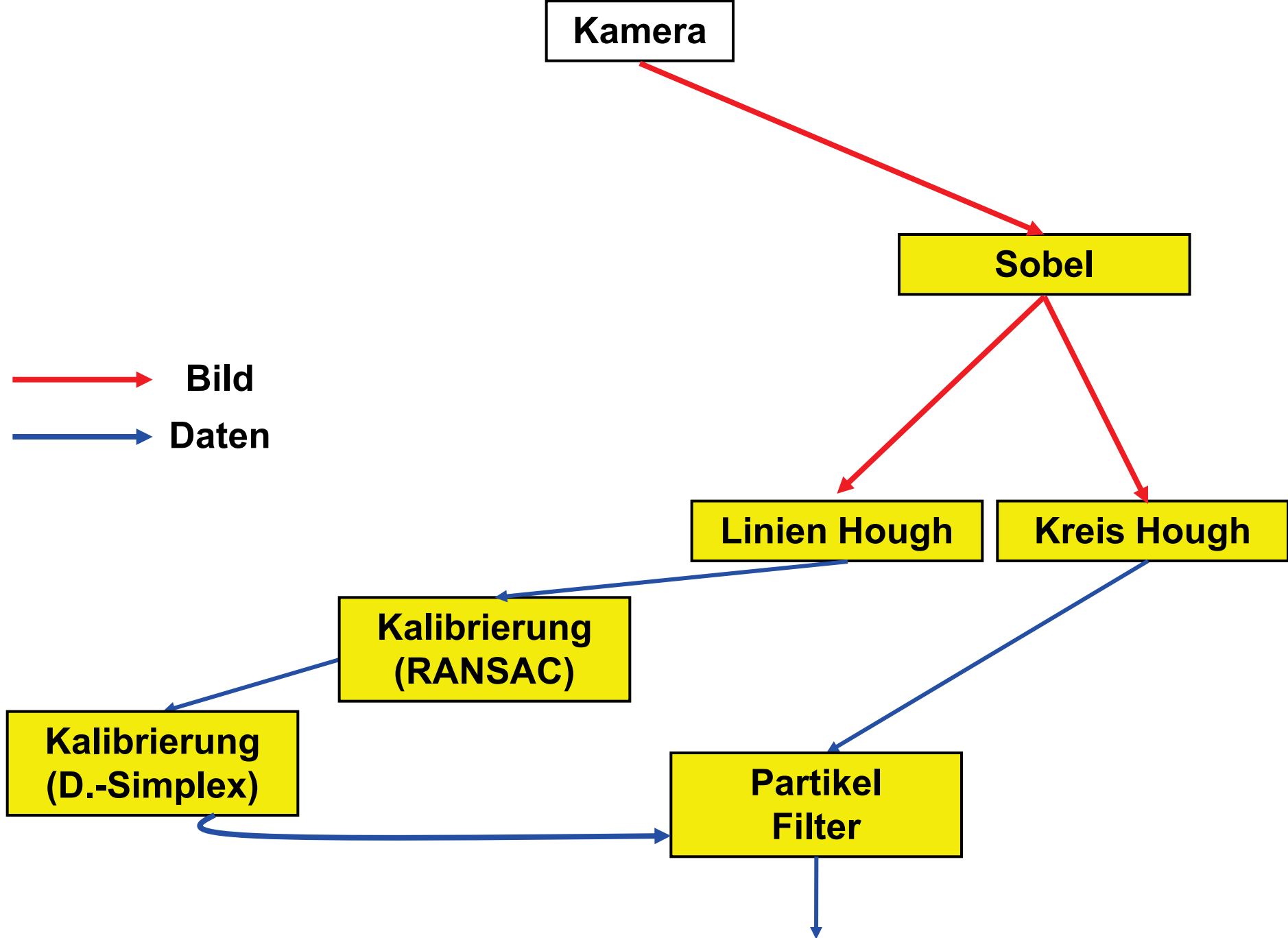


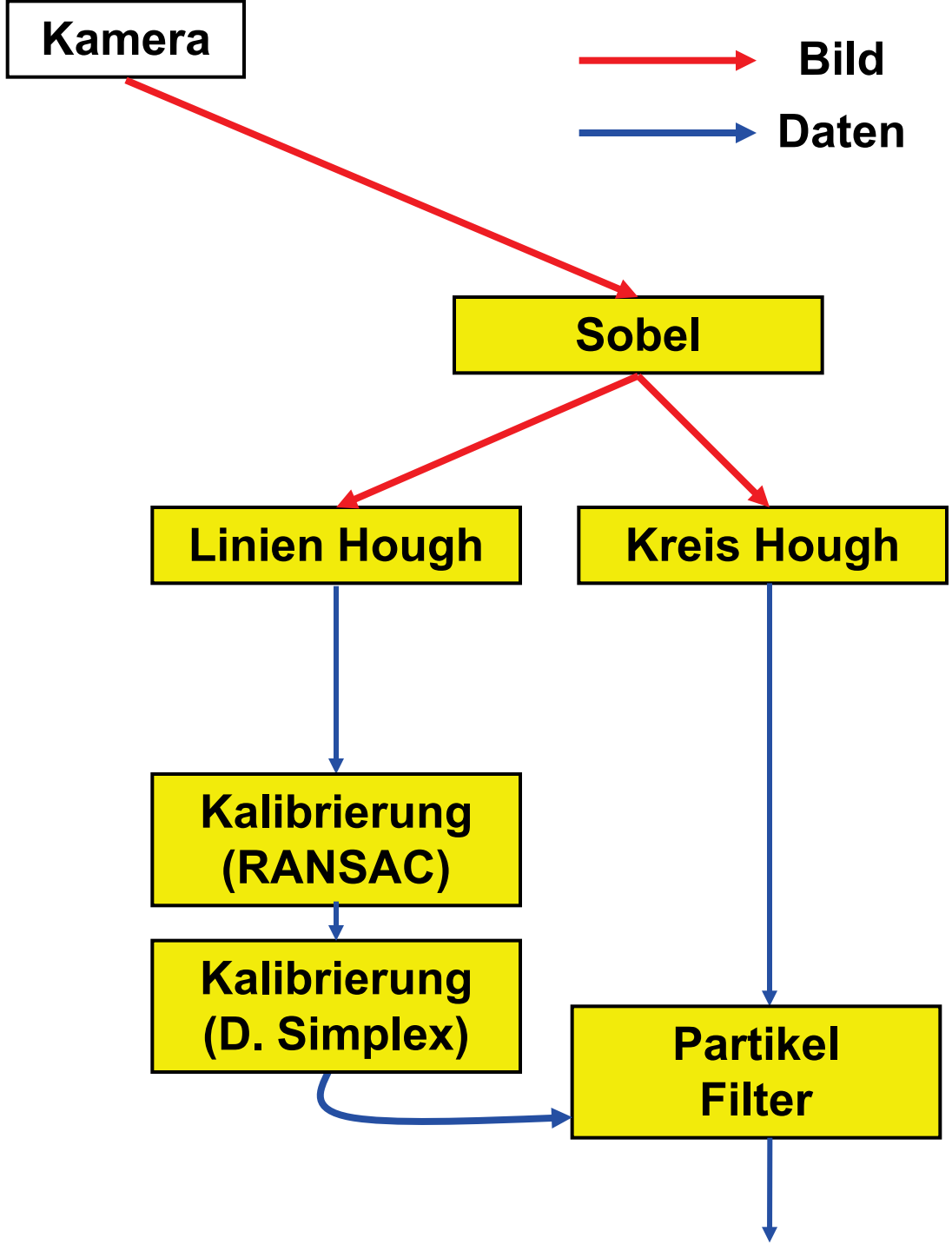
The **challenge**
of **understanding** images
is not just to analyze a piece of information
locally but also
to bring the context into play
(G. Binnig)

Kognitive Bildverarbeitung

Kontext Einbringen

- ▶ **Wie?**
- ▶ **Z.B. Globaler Probabilistischer Ansatz**
 - ▶ $\arg \max_{x_t} p(X_t=x_t | I_{0..t}=i_{0..t})$,
 - ▶ X_t uns interessierender Zustand
 - ▶ I_t das Bild zum Zeitpunkt t .
- ▶ **aus den Modulen werden Faktoren $p(\dots)$ in einem Produkt von bedingten Wahrscheinlichkeiten.**
- ▶ **\Rightarrow harte Entscheidungen möglichst spät treffen**
- ▶ **\Rightarrow niedrigerer Wert in einem $p(\dots)$ wird akzeptiert, wenn dadurch höherer Wert in anderen $p(\dots)$ entsteht.**

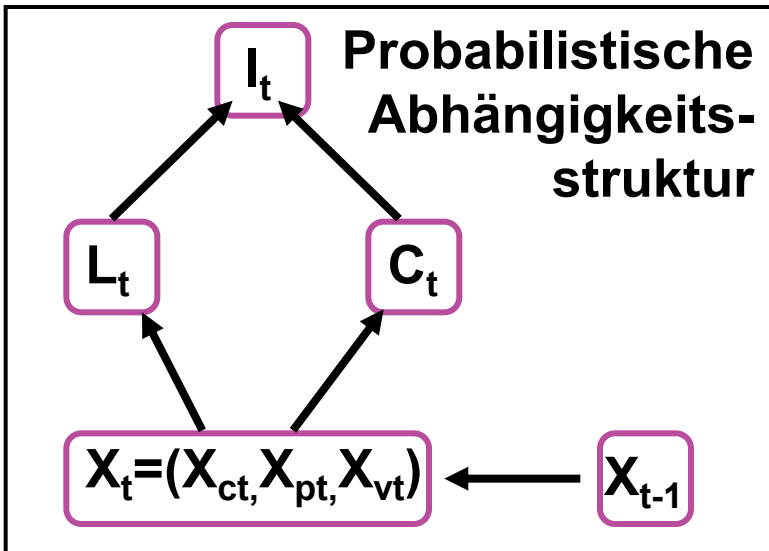




- $X_t = (X_{ct}, X_{pt}, X_{vt})$
- $X_{ct} = \text{Kamerapose}$
- $X_{pt} = \text{Ballposition}$
- $X_{vt} = \text{Ballgeschwindigkeit}$
- $C_t = \text{Kreis(-e) im Bild}$
- $L_t = \text{Linien im Bild}$
- $I_t = \text{Bild}$

$$p(x_t | i_{0..t}) \propto$$

$$\int p(i_t | c_t, l_t) p(l_t | x_{ct}) p(c_t | x_{pt}, x_{ct}) \\ p(x_t | x_{t-1}) p(x_{t-1} | i_{0..t-1}) dx_{t-1} dl_t dc_t$$



Kamera

- **Bild**
- **Daten**
- **Wahrs.**

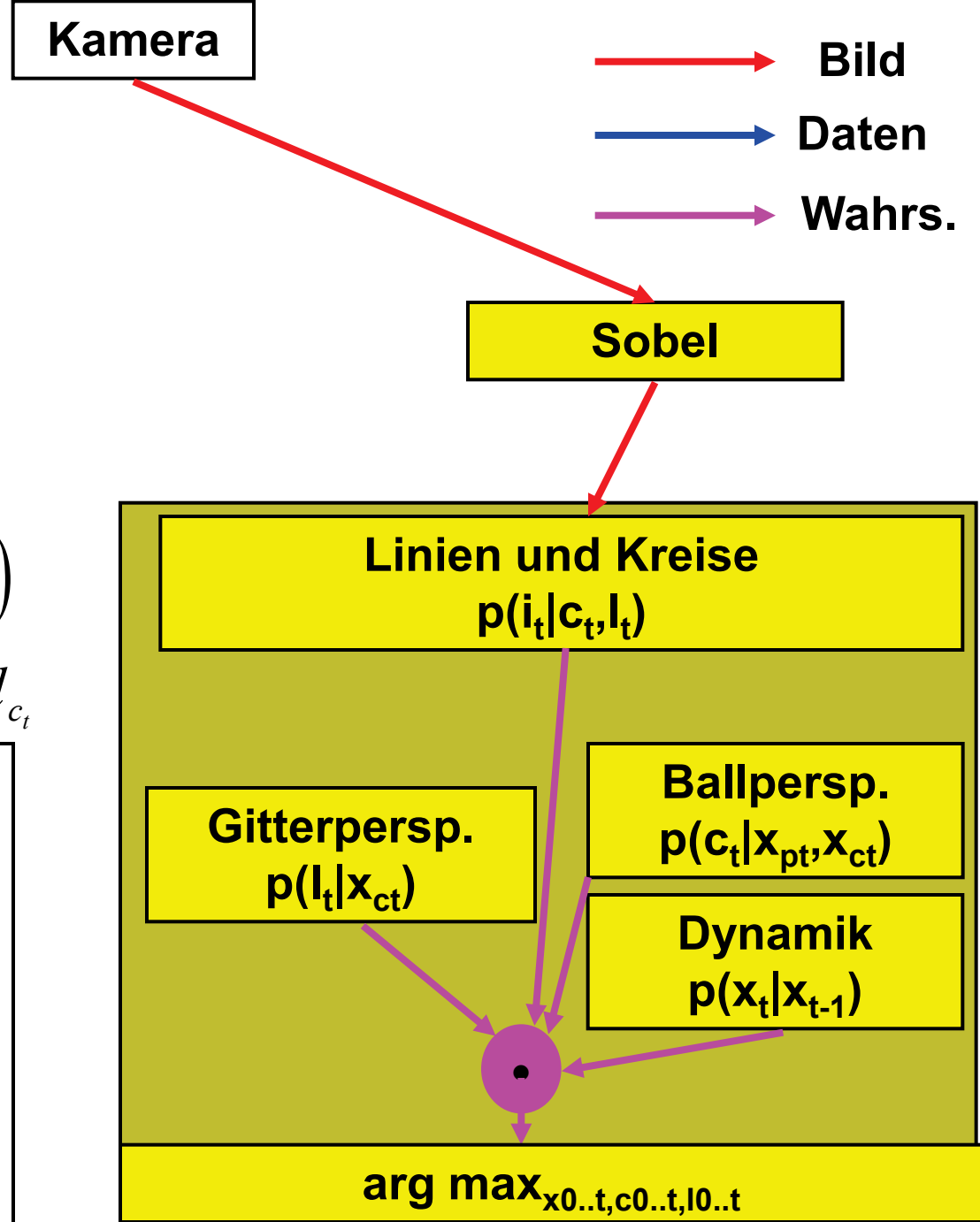
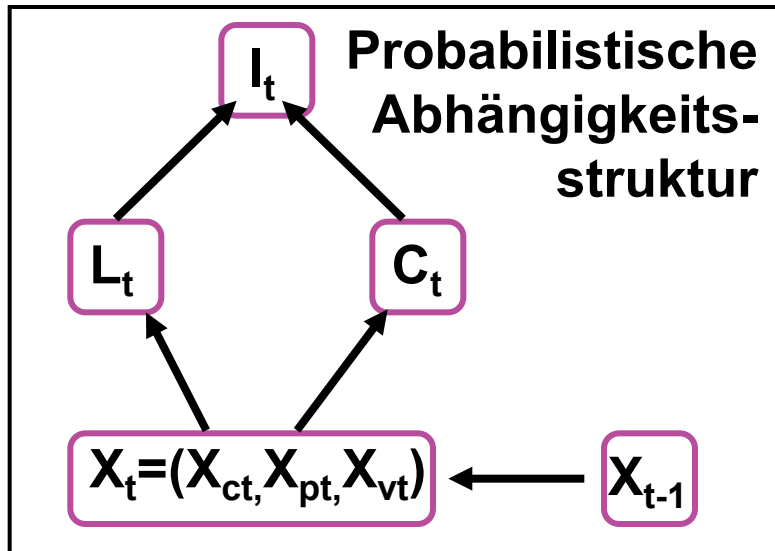
Sobel

Frage an das Auditorium: Was bedeuten anschaulich die verschiedenen p()?

- $X_t = (X_{ct}, X_{pt}, X_{vt})$
- $X_c = \text{Kamerapose}$
- $X_p = \text{Ballposition}$
- $X_v = \text{Ballgeschwindigkeit}$
- $C_t = \text{Kreis(-e) im Bild}$
- $L_t = \text{Linien im Bild}$
- $I_t = \text{Bild}$

$$p(x_t | i_{0..t}) \propto$$

$$\int p(i_t | c_t, l_t) p(l_t | x_{ct}) p(c_t | x_{pt}, x_{ct}) p(x_t | x_{t-1}) p(x_{t-1} | i_{0..t-1}) dx_{t-1} dl_t dc_t$$



Kognitive Bildverarbeitung

Vorteile: Globaler Kontext

- ▶ schwach erkennbarer Kreis wird eher als Ball interpretiert, wenn Größe und Position zum Flughafenmodell passen.
- ▶ Kalibrierungsunsicherheit wird berücksichtigt
- ▶ Flughäfen verbessern Kalibrierung.
- ▶ z.B. Köpfe von Zuschauer führen zu Ballhypothesen, die aber nicht plausibel sind, weil die Flugbahn nicht passt.

Kognitive Bildverarbeitung

Nachteile: Extrem schwierige Optimierung

- ▶ Einzugsbereich des Minimums ist oft nur wenige Pixel groß.
- ▶ Probleme, die im klassischen Ansatz anwendungsbezogen und heuristisch gelöst werden, muss hier der Optimierer implizit lösen.

Anwendungsbeispiel 3D Modellierung einer Stadt

N. Cornelis, B. Leibe, K. Cornelis, L. Van Gool, 3D City Modeling using Cognitive Loops, 3rd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06), Chapel Hill



Kognitive Bildverarbeitung

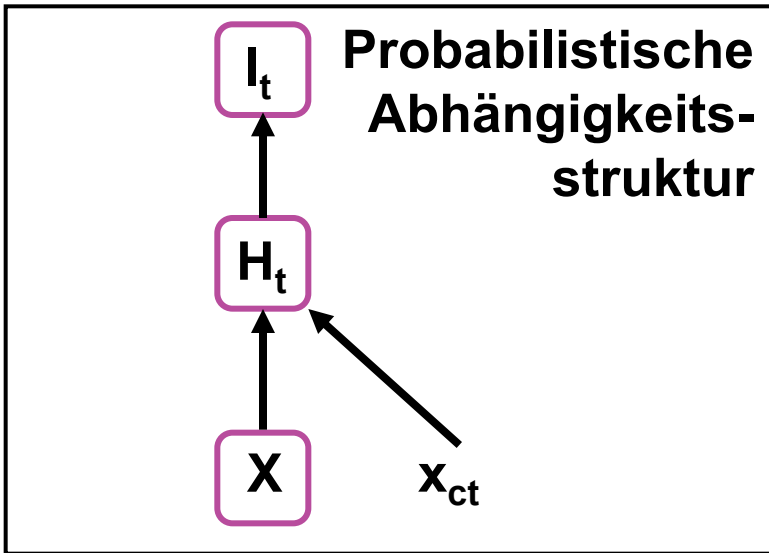
Anwendungsbeispiel 3D Modellierung einer Stadt

- ▶ an sich schwierige aber rein geometrische Aufgabe...
- ▶ ... aber semantischer Kontext kann helfen.
- ▶ **Idee: Erkenne Autos um besser 3D Modelle zu erstellen.**
- ▶ **Erkennung \Rightarrow Geometrie**
 - ▶ Autos sind schlechte Merkmale (beweglich, reflektierend, schwer modellierbar)
 - ▶ ignoriere Autos für geometrische Rechnungen
 - ▶ ersetze erkannte Autos im 3D Modell durch ein Prototypauto
- ▶ **Geometrie \Rightarrow Erkennung**
 - ▶ Erwartungshaltung für Position und Größe für Erkennung
 - ▶ macht Erkennung robuster

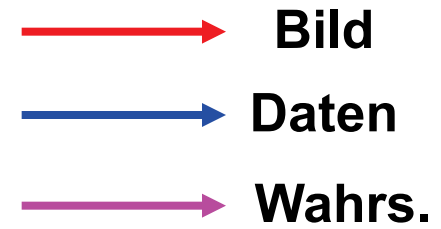
- ▶ X = Autoposen
- ▶ x_{ct} = Kamerapose
- ▶ H_t = Bildposition eines "Autobildes"
- ▶ I_t = Bild

$$p(x | i_{0..t})$$

$$\propto \int p(i_t | h_t) p(h_t | x) p(x | i_{0..t-1}) dh_t$$



Stereokamera



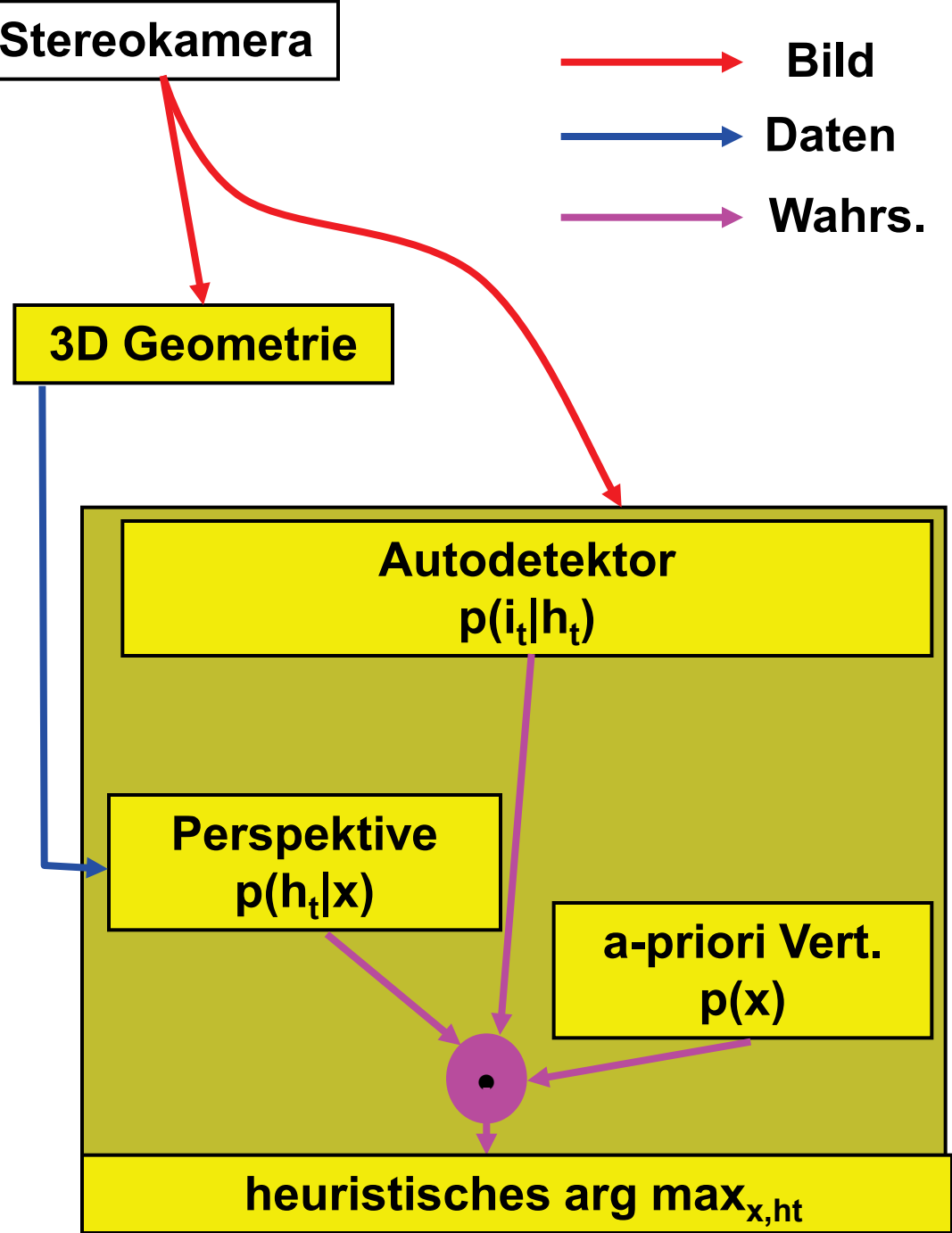
3D Geometrie

Autodetektor
 $p(i_t | h_t)$

Perspektive
 $p(h_t | x)$

a-priori Vert.
 $p(x)$

heuristisches arg max_{x,ht}



Kognitive Bildverarbeitung

Kognitive Bildverarbeitung

- ▶ **Kognitive Bildverarbeitung übernimmt Elemente der menschlichen visuellen Wahrnehmung.**
- ▶ **u.a. einbringen von Kontext um Erkennung robuster zu machen.**
- ▶ **nicht Module für verschiedene Erkennungsebenen, die aneinander Daten weitergeben,...**
- ▶ **..., sondern probabilistische Modellierung der verschiedenen Ebenen mit Suche nach der plausibelsten Erklärung durch gemeinsame Optimierung über alle Ebenen.**
- ▶ **sehr aktiver Forschungsbereich mit großem Potential aber noch nicht vielen etablierten Methoden**

Rekapitulation des Prüfungsstoffes

Rekapitulation des Prüfungsstoffes

Fachgespräch:

- ▶ **Übungszettel 6: 6.7.10 abgeben (hammer@)**
- ▶ **2.8.10, Cartesium 0.041**
- ▶ **in Gruppen**
- ▶ **Anmeldung per E-Mail**
- ▶ **Note nach Punkten**
- ▶ **Fachgespräch am 2.8.10**
 - ▶ Überprüfung, dass alle an den Übungszetteln beteiligt waren
 - ▶ kurzes (4 min) Erklären einer Aufgabe
 - ▶ nur Aufgaben mit $\geq 50\%$ der Punkte
 - ▶ Übungszettel mitbringen!

Note	\geq Punkte
1.0	76
1.3	72
1.7	68
2.0	64
2.3	60
2.7	56
3.0	52
3.3	48
3.7	44
4.0	40

Rekapitulation des Prüfungsstoffes

Mündliche Prüfung:

- ▶ **einzel**
- ▶ **Anmeldung per E-Mail**
- ▶ **2.8.10 oder 23.9.10, Cartesium 0.041**
- ▶ **Thema: Der Inhalt der Vorlesung:**
 - ▶ Algorithmen
 - ▶ Anwendungen
 - ▶ Konzepte
 - ▶ Echtzeitimplementierung
- ▶ **Prüfung startet mit kleinen Anwendungsaufgabe.**
 - ▶ 3 min Bedenkzeit
 - ▶ erläutern, wie man die Anwendung angehen könnte.
- ▶ **Lernmaterialien auf der Webseite („Prüfungsstoff“)**
 - ▶ Liste an Beispielanwendungen

Rekapitulation des Prüfungsstoffes

Was sollte man aus der Vorlesung (für die Prüfung) mitnehmen?

▶ **Algorithmen:**

- ▶ Was wird berechnet?
- ▶ Woraus?
- ▶ Wie funktioniert die Berechnung („Idee in 2min“)?
- ▶ Siehe Diagramm und Kurzzusammenfassung auf Lernmaterialfolien

▶ **Anwendungen:**

- ▶ Was würde passieren, wenn man ein bestimmtes Verfahren anwendet?
- ▶ Welche Probleme treten auf?
- ▶ Wie kann man sie angehen?
- ▶ Welche Voraussetzungen macht man bei einem Verfahren?
- ▶ Wie kann man die Umgebung / Beleuchtung / Kamera passend gestalten?
- ▶ Prüfung startet mit kleinem Anwendungsproblem.

Rekapitulation des Prüfungsstoffes

Was sollte man aus der Vorlesung (für die Prüfung) mitnehmen?

▶ **Konzepte:**

- ▶ Wie ist es formell definiert (Gleichung)?
- ▶ Was bedeutet das real?
- ▶ Wo (in welchem Algorithmus) verwendet man es?
- ▶ Siehe Liste in den Lernmaterialfolien

▶ **Echtzeitimplementierung:**

- ▶ Wie optimiert man im Detail am Beispiel **Linien – Hough Transformation?**
- ▶ LUT, Laufzeiger, Festkommaarithmetik, Ausnutzung des Gradienten

Das Ende

Dank an:

- ▶ **Das Auditorium für die Aufmerksamkeit, und interessante Fragen und Antworten**
- ▶ **Tobias Hammer für die Korrektur der Übungszettel**
- ▶ **Thomas Röfer für die Folien „RoboCup“**
- ▶ **Wolfram Burgard für die Folien „Partikel Filter“**
- ▶ **Berthold Bäuml für Hinweise zur Kognitiven Bildverarbeitung**