

**Sensordatenverarbeitung**

# **SDV PARADIGMA (8B)**

**(2.-6.12.24)**



# Keine frühen harten Entscheidungen

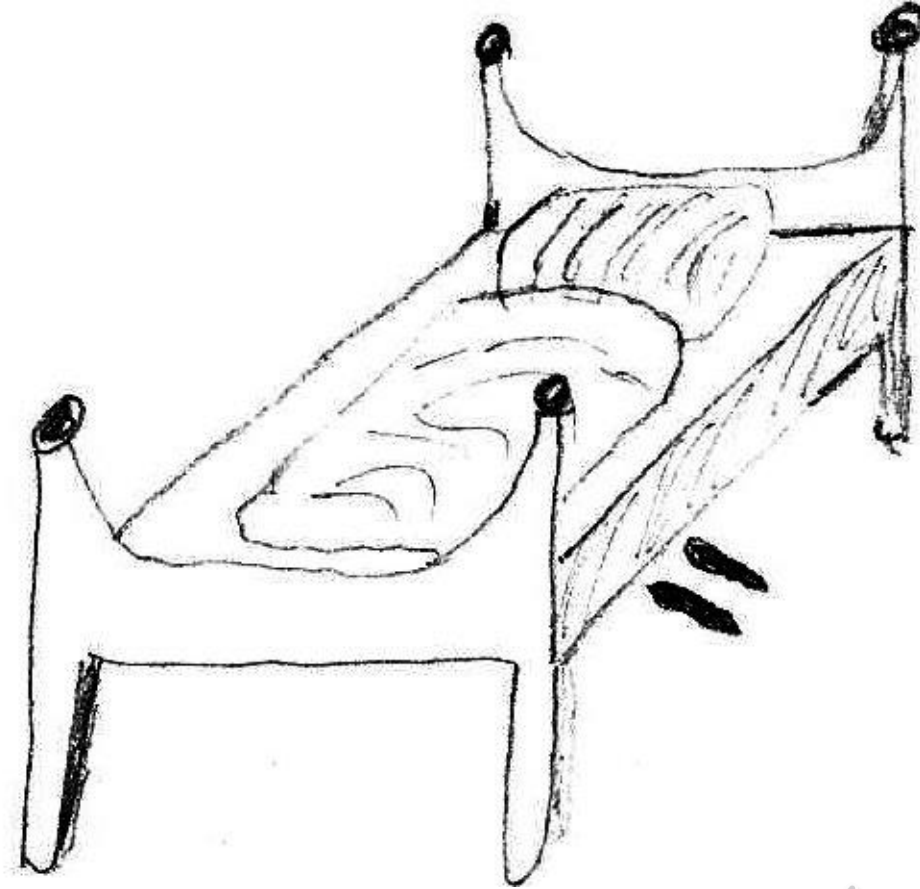


## Frage an das Auditorium: Was ist das?

(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)

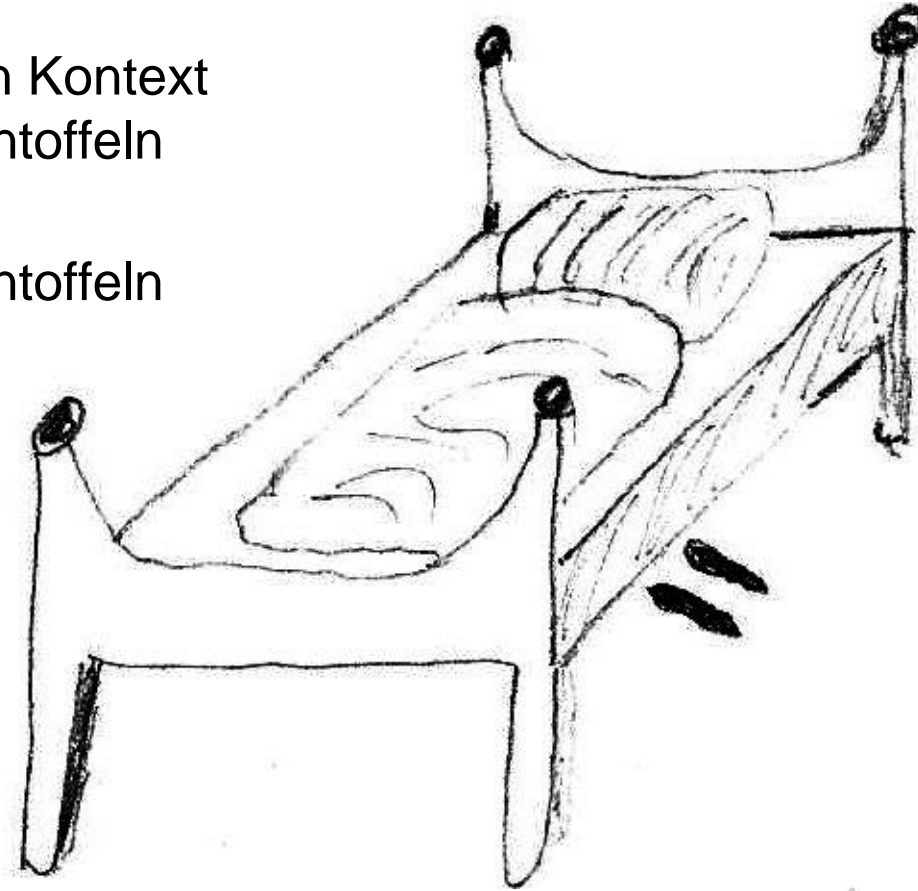


Frage an das Auditorium: Was ist das?



Frage an das Auditorium: Was ist das?

- Pantoffeln!
- Menschliche Wahrnehmung nutzt den Kontext des Bettes um die Erkennung der Pantoffeln zu stützen.
- Ohne Kontext des Bettes sind die Pantoffeln nicht visuell eindeutig genug.



- Außerdem: Ton hilft bei der Erkennung von Bildern
- „Finde das Tier“-Experiment
  - Gleich kommt ein Bild mit einem versteckten Tier
  - Finde das Tier!
  - Uhr läuft mit. Notiere die Zeit.
  - Es gibt einen akustischen Hinweis!
  - Wer ungerade viele Jahre alt ist stellt gleich den Ton aus  
→ kein Hinweis
  - Zeit und ob mit oder ohne Sound per Email an [ufrese@uni-bremen.de](mailto:ufrese@uni-bremen.de) schicken

**Jetzt Ton ausstellen, wenn Du ungerade viele Jahre alt bist!**





# Wo ist das verborgene Tier?



0

**Jetzt Ton wieder anstellen!**



- Ton hilft bei der Erkennung von Bildern
- Frage an das Auditorium: Welche Information liefert der Ton denn?

Bildquelle: [https://www.reddit.com/r/pics/comments/4vi9kq/find\\_a\\_cat/](https://www.reddit.com/r/pics/comments/4vi9kq/find_a_cat/)

Soundquellen:

Npeo, Kitty Meow, CC BY 3.0, freesound.org, #203121

Mattix, Retro Adventure Vibe 02, CC-BY-3.0, freesound.org, #479467

- Ton hilft bei der Erkennung von Bildern
- Frage an das Auditorium: Welche Information liefert der Ton denn?
- Hier ist eine Katze
- In Wirklichkeit außerdem grobe Richtung der Katze
- Menschliche Wahrnehmung berücksichtigt die aktuellen und früheren „Daten“ aller „Sensoren“ und Vorkenntnisse für die Erkennung
  - → Kontext
- Diese Fähigkeit ist eine große Stärke menschlicher Wahrnehmung, die Maschinen noch kaum können.

Bildquelle: [https://www.reddit.com/r/pics/comments/4vi9kq/find\\_a\\_cat/](https://www.reddit.com/r/pics/comments/4vi9kq/find_a_cat/)

Soundquellen:

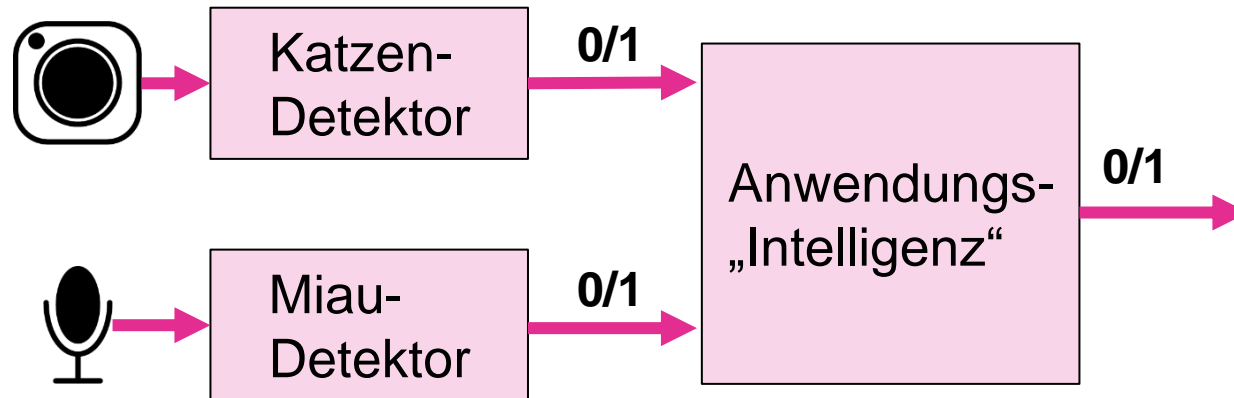
Npeo, Kitty Meow, CC BY 3.0, freesound.org, #203121

Mattix, Retro Adventure Vibe 02, CC-BY-3.0, freesound.org, #479467

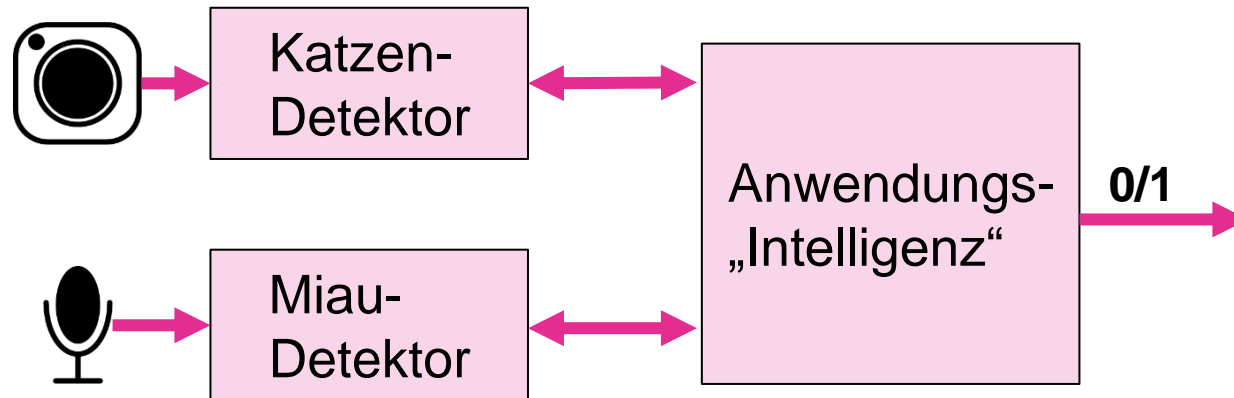


- Wie könnte eine Softwarearchitektur aussehen in der Bild und Ton sich gegenseitig als Kontext stützen?

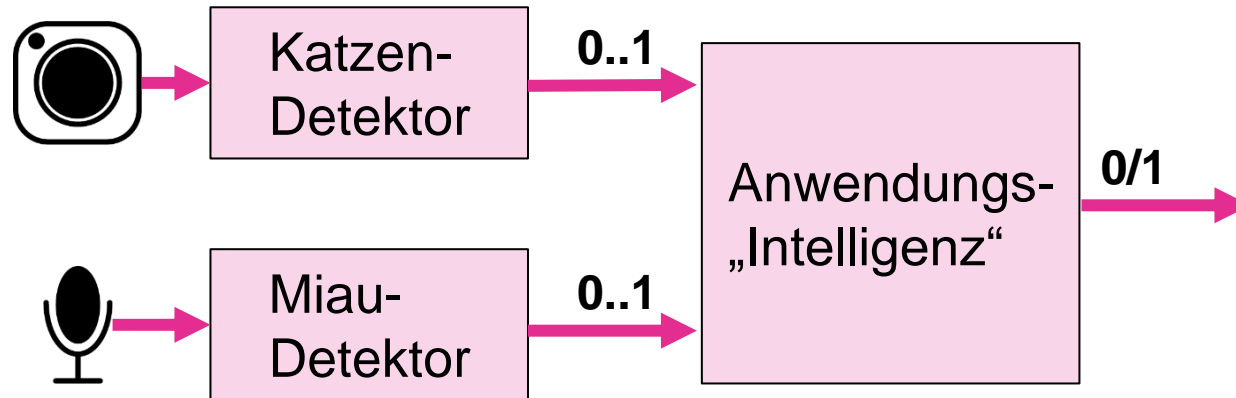




- Softwaremodul erkennt Teilaspekt und liefert Ergebnis
- Jedes Modul betrachtet nur seine Daten und **trifft harte Entscheidung** („Katze oder keine Katze“) → Kein Kontext
- Leichte Aufteilung in Module → Gut aus Sicht der Softwaretechnik
- Einfach aber nicht gut
- Extremfall segmentierungsbasierte BV: Harte Entscheidung pro Pixel



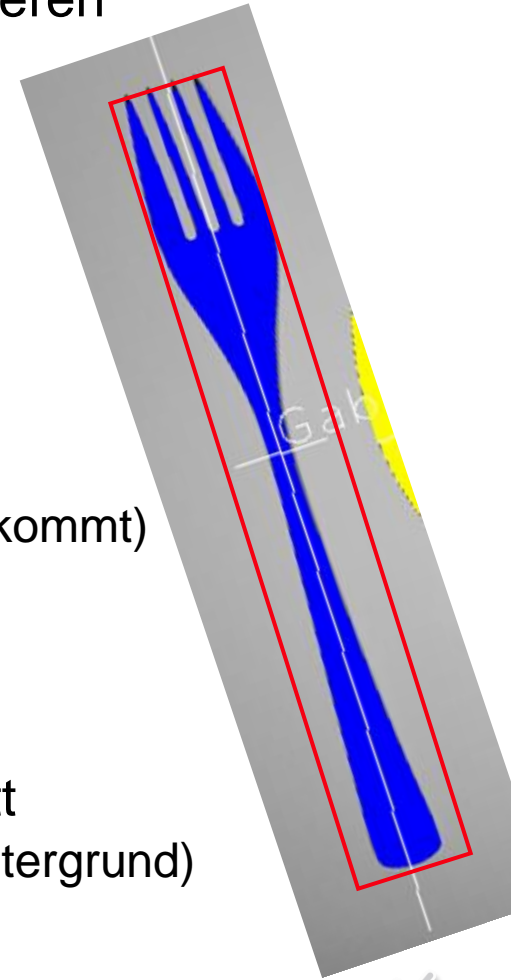
- Priming (Psychologie): *Beeinflussung der Verarbeitung (Kognition) eines Reizes* [durch andere Reize oder Vorerfahrung], Wikipedia
- Miau erkannt → Anwendungs-Intelligenz → Priming-Info an Katzen-Detektor → Detektor findet Katze → Anwendungs-Intelligenz
- Kognitiv plausibel, vermutlich mächtig
- Softwaretechnisch sehr verzahnt
- Unklar, wie genau zu realisieren



- Softwaremodul erkennt Teilaspekt und liefert Ergebnis
- Jedes Modul betrachtet nur seine Daten, trifft aber eine graduelle Entscheidung („Wie katzig sieht es hier aus?“)
- Anwendungsintelligenz kombiniert graduelle Entscheidungen
  - Durch das detektierte Miau reicht ein geringerer Wert vom Katzendetektor
  - Z.B. mit Wahrscheinlichkeiten und sogenannter Bayes-Formel
- **Harte Entscheidung so spät wie möglich** → Kontext
- Etablierte bessere Standardarchitektur in Sensordatenverarbeitung



- Ein Merkmal ist eine erkennbare Eigenschaft, die eine Person, eine Sache oder einen abstrakten Zusammenhang von anderen unterscheidet
- Indirektes Bildmerkmal: Merkmal das nach einem wesentlichen Vorverarbeitungsschritt berechnet wurde
- Eigenschaften einer segmentierten Region
  - Fläche, Umfang, Breite, Länge
  - Mittlere Farbe (einer Subregion)
  - Anzahl Pixel einer bestimmten Farbe
  - Farbhistogramm (Statistik, wie häufig welche Farbe vorkommt)
- Für Klassifizierung schon segmentierter Objekte
  - Durch anwendungsdefinierte Schwellwerte
  - Durch Machine Learning
- Segmentierung ist wesentlicher Vorverarbeitungsschritt
  - Fällt eine harte Entscheidung pro Pixel (Objekt oder Hintergrund)
  - „Softwaremodule“ Paradigma



- Menschliche Wahrnehmung berücksichtigt die aktuellen und früheren „Daten“ aller „Sensoren“ und Vorkenntnisse für die Erkennung
  - → Kontext
  - Fähigkeit ist große Stärke menschlicher Wahrnehmung
  - → Harte Entscheidungen so spät wie möglich, lieber graduelle Ausgaben