



# Inhaltsanalyse unter Zuhilfenahme von Metadaten

**Seminar - Bridging the Semantic Gap**

**Präsentation von Alexander Bellm**



## Objektgrenzenerkennung für ontologiegestützte Bildklassifizierung

- **System**
- **Bildsegmentierung**
  - Randerkennung
  - Bereichentwicklung
  - Zusammenlegen benachbarter Bereiche durch die MOT (Matrix Oriented Technique)
- **Vorläufige Ergebnisse**

## Wissensunterstützte semantische Videoobjekterkennung

- **System**
- **multimedia analysis ontology**
- **Formula One domain ontology**
- **Zusammenfügung**
- **Beispiele**
- **Ausblick**



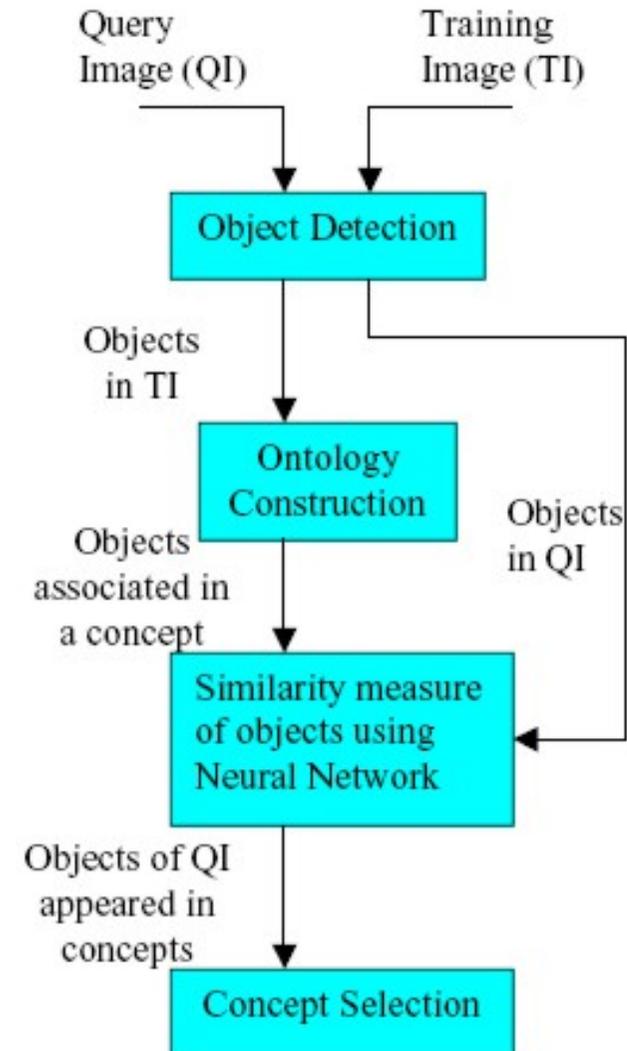
## 2 Ziele:

1. Gewinnung von Schlagwörtern die das Bild beschreiben und identifizieren
2. Das Konstruieren einer Ontologie

## Ablauf von 1:

- Objekterkennung
- Ein neuronales Netz erzeugt Schlagwörter für die Objekte
- Die Schlagwörter werden durch einen ontologiebasierten Algorithmus selektiert

Das automatische konstruieren einer Ontologie ist Teil der zukünftigen Forschungsarbeit





# Randerkennung

## Vorgehensweise:

- Grundlage: Intensitätswert
- 4 mögliche Szenarien
- Positionsbezeichnungen
- Kalkulation
- Zuordnung der Pixel:
  - EPS ( edge pixel set) wenn MOE(x,y) größer als wählbarer Grenzwert
  - RPS (region pixel set) sonst

1	2	1
0	0	0
-1	-2	-1
HOE		
2	1	0
1	0	-1
0	-1	-2
NOE		
1	0	-1
2	0	-2
1	0	-1
VOE		
0	1	2
-1	0	1
-2	-1	0
SOE		

(x-1, y-1)	(x-1, y)	(x-1,y+1)
(x, y-1)	(x, y)	(x, y+1)
(x+1,y-1)	(x+1,y)	(x+1,y+1)

$$\begin{aligned}
 \text{HOE}(x, y)_i &= | I(x-1,y-1) + 2I(x,y-1) + I(x+1,y-1) \\
 &\quad - I(x-1,y+1) - 2I(x,y+1) - I(x+1,y+1) | \\
 \text{VOE}(x, y)_i &= | I(x-1,y-1) + 2I(x-1,y) + I(x-1,y+1) \\
 &\quad - I(x+1,y-1) - 2I(x+1,y) - I(x+1,y+1) | \\
 \text{NOE}(x, y)_i &= | I(x,y-1) + 2I(x-1,y-1) + I(x-1,y) \\
 &\quad - I(x+1,y) - 2I(x+1,y+1) - I(x,y+1) | \\
 \text{SOE}(x, y)_i &= | I(x,y-1) + 2I(x+1,y-1) + I(x+1,y) \\
 &\quad - I(x-1,y) - 2I(x-1,y+1) - I(x,y+1) | \\
 \text{MOE}(x, y)_i &= \max \{ \text{HOE}(x, y)_i, \text{VOE}(x, y)_i, \\
 &\quad \text{NOE}(x, y)_i, \text{SOE}(x, y)_i \}
 \end{aligned}$$

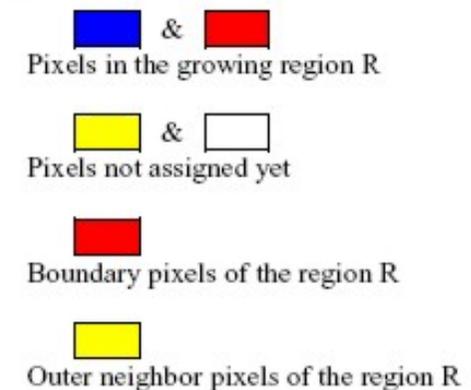
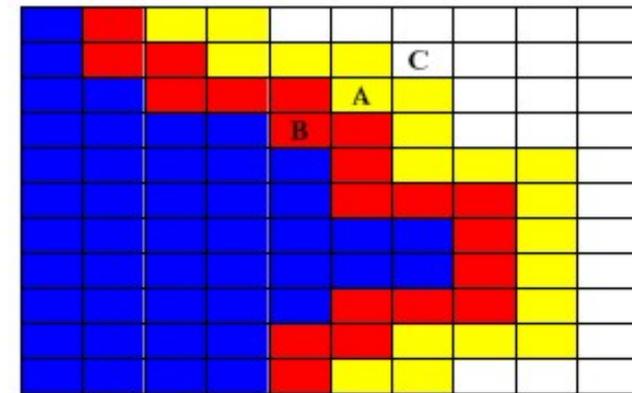


## Bereichentwicklung

1 2 3 4 5 6 7 8 9 10 11 12 13 14

### Ablauf:

- zufällige Wahl eines Pixels aus RPS
- Einteilung in „boundary pixel“ und „inner pixel“
- Ein Pixel ist „boundary pixel“ wenn eines seiner 8 Nachbarpixel nicht zur selben Region gehört
- Das erste Pixel einer Region ist also immer ein „boundary pixel“
- Prüfung aller Nachbarpixel eines „boundary pixel“ in RPS, ob sie die Bedingungen erfüllen:
  - Die Farbabweichung zwischen dem Pixel und seinen Nachbarpixel in der Region sind vernachlässigbar klein
  - Die Farbabweichung zwischen dem Pixel und dem Durchschnittswert der Nachbarpixel in der Region sind vernachlässigbar klein
  - Die Farbabweichung zwischen dem Pixel und dem Durchschnittswert der Region sind vernachlässigbar klein
- Sind die Bedingungen erfüllt wird die Einteilung erneut durchgeführt und alles wiederholt bis die Region nicht mehr wächst





## Zusammenlegen benachbarter Bereiche

1 2 3 4 5 6 7 8 9 10 11 12 13 14

### Problem:

1. Auftreten von „noise regions“
2. Ein Objekt wird in verschiedene Regionen aufgeteilt

-1	-1	5	5	5	5	-1	3
2	-1	5	5	5	5	-1	3
2	-1	-1	5	5	-1	-1	3
2	2	-1	5	5	-1	3	3
2	2	-1	5	5	-1	3	3
2	2	2	-1	-1	3	3	3
2	2	2	2	2	-1	-1	3
2	2	2	2	2	2	-1	3
2	2	2	2	2	2	2	-1

### Lösung:

Einsatz der Matrix Oriented Technique (MOT)

- Matrix (jede Zelle entspricht einem Pixel)
- Inhalt der Zellen sind die Regionsnummern
- Randpixel haben den Wert -1

2	-1	5	5	5	5	5	-1
2	-1	5	5	5	5	-1	-1
2	-1	5	5	5	5	-1	3
2	-1	5	5	5	5	-1	3
2	-1	-1	-1	-1	-1	-1	3
2	-1	4	4	4	4	-1	3
2	-1	4	4	4	4	-1	3
2	-1	4	4	4	4	4	-1
2	-1	4	4	4	4	4	-1

Für 1: Wenn eine Region nur aus sehr wenigen Pixel besteht, wird sie der benachbarten Region mit dem kleinsten Farbunterschied zugeordnet.

Für 2: Jede Grenze erhält den Wert des Farbunterschiedes ihrer Regionen, so dass über einen definierten Grenzwert bestimmt werden kann, welche Regionen zusammengelegt werden sollen.



# Vorläufige Ergebnisse

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14

Original Image		Detected Edge	
Original Image		Detected Edge	
Detected Objects		Detected Objects	
Original Image		Detected Edge	
Detected Objects		Detected Objects	



# System

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14

Unterstützt die Erkennung von themenspezifischen Objekten

**multimedia analysis ontology**

**domain specific ontology**

Integration durch bereichsunabhängige Grundklassen (Object, Feature, Feature Parameter)

Repräsentiert das Wissen über den entsprechenden Themenbereich

**video stream**



Wahl der Abfolge der Algorithmen durch die vorhandenen optischen low-level Eigenschaften und das räumliche Verhalten zur Objekterkennung



**Algorithm Repository**



**Output**



# multimedia analysis ontology

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14

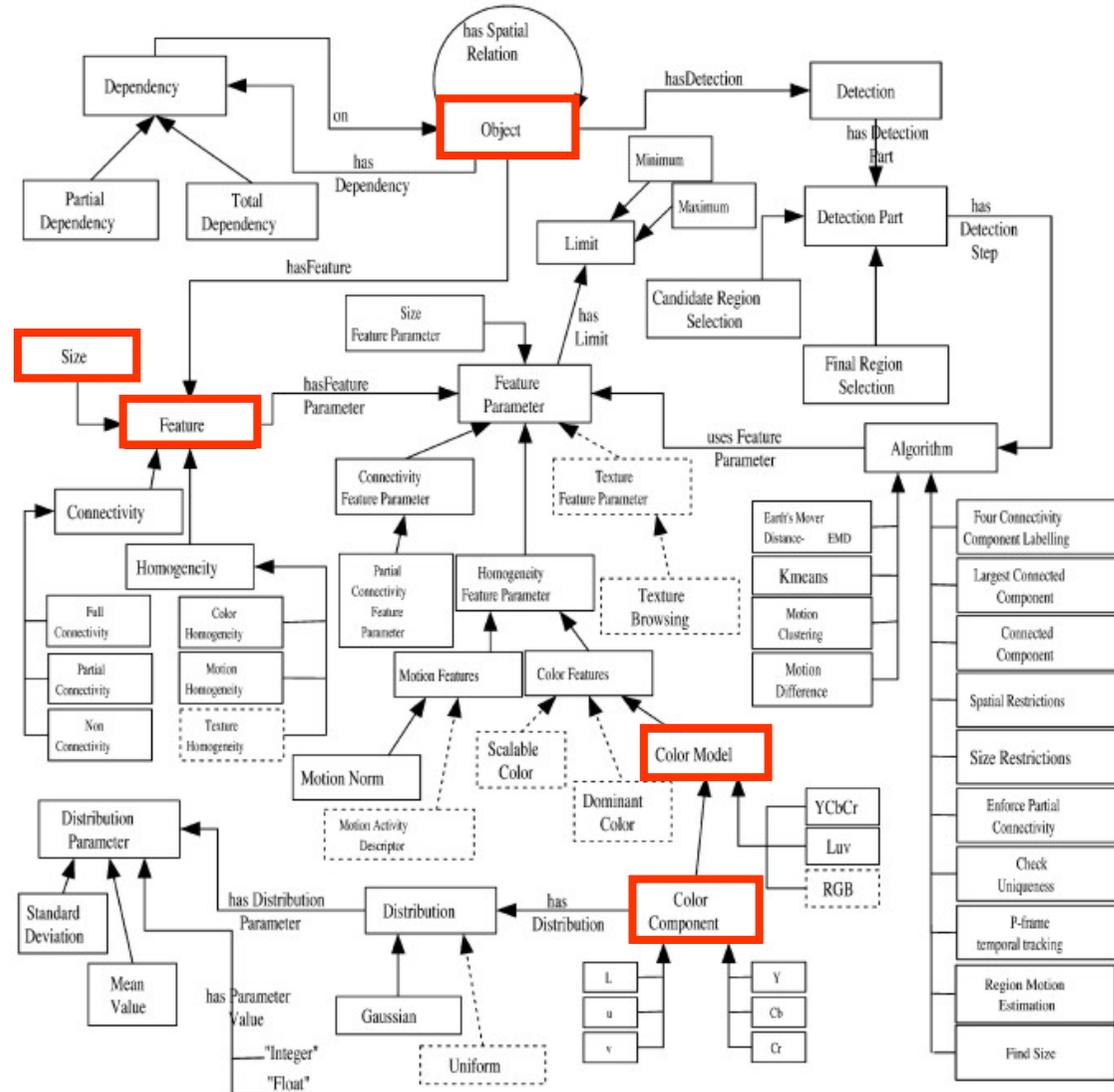
## Object:

Superklasse aller Videoobjekte die während der Analyse erfasst werden

## Feature:

Superklasse aller low-level-Eigenschaften die den Objekten zugeordnet werden

## Size + Color





# Formula One domain ontology

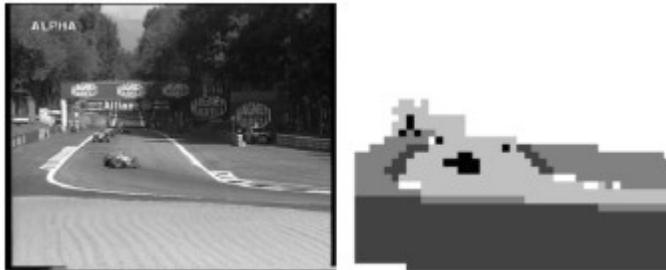
1 2 3 4 5 6 7 8 9 10 11 12 13 14

## Auto:

- zusammenhängende Region
- homogene Bewegung
- innerhalb des Objekts Straße
- Bewegungswert > Minimaler Wert
- Größe hat keinen Maximalwert

## Straße:

- zusammenhängende Region
- homogene Farbe
- Größe > vordefinierter Minimalwert
- größte Region des Videos



## Gras:

- homogene Farbe
- teilweise zusammenhängende Region
- jede Teilregion > minimale Größe
- adjacent-to-Beziehung zu Straße
- adjacent-to-Beziehung zu Sand

## Sand:

- homogene Farbe
- teilweise zusammenhängende Region
- jede Teilregion > minimale Größe
- adjacent-to-Beziehung zu Straße
- adjacent-to-Beziehung zu Gras



# Zusammenfügung

1 2 3 4 5 6 7 8 9 10 11 12 13 14

## Zusammenfügen der „multimedia analysis ontology“ und der „Formula One domain ontology“

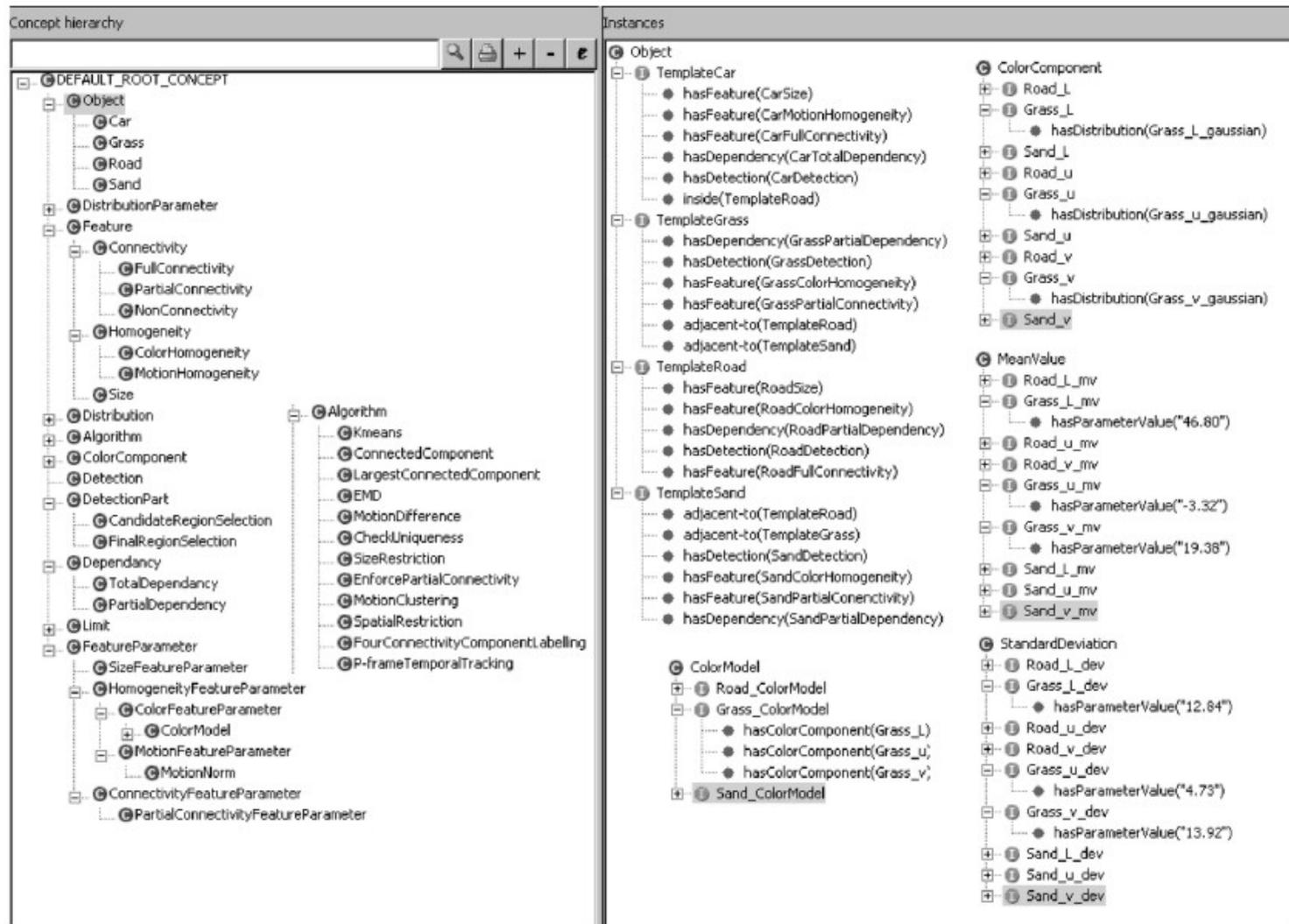
Grundlage dafür bilden die in beiden Ontologien vertretenen Klassen „Object“, „Feature“ und „Feature Parameter“

Hauptaugenmerk dieses Ansatzes basiert auf visuellen Objekteigenschaften wie Homogenität und Zusammenhängigkeit sowie auf den räumlichen Beziehungen der einzelnen Objekte zueinander



# Zusammenfügung

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14





# Beispiele

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14



Input Frame



anfängliche Segmentierung



Endresultat des betrachteten Systems



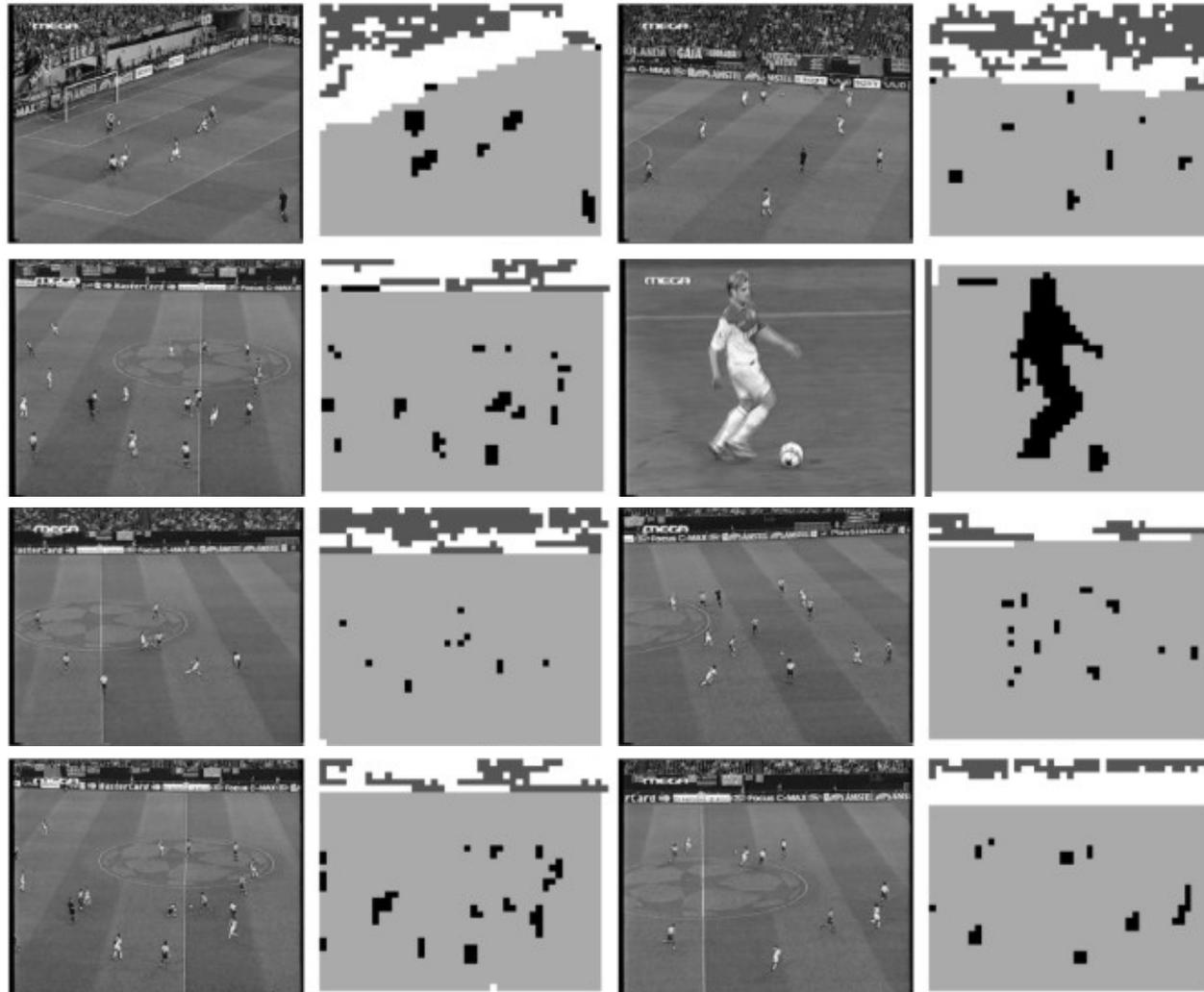
Numerische Auswertung für die Formel Eins:

Object	correct detections	false detections	missed
Road	97%	2%	1%
Grass	90%	7%	3%
Sand	88%	8%	4%
Car	74%	22%	7%



# Beispiele

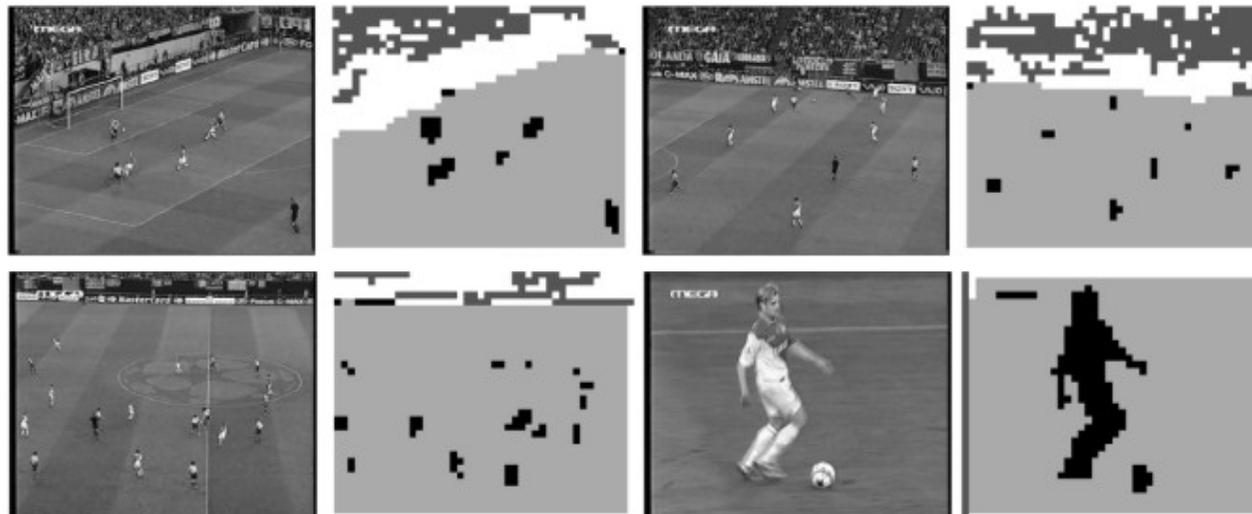
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14





# Beispiele

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14



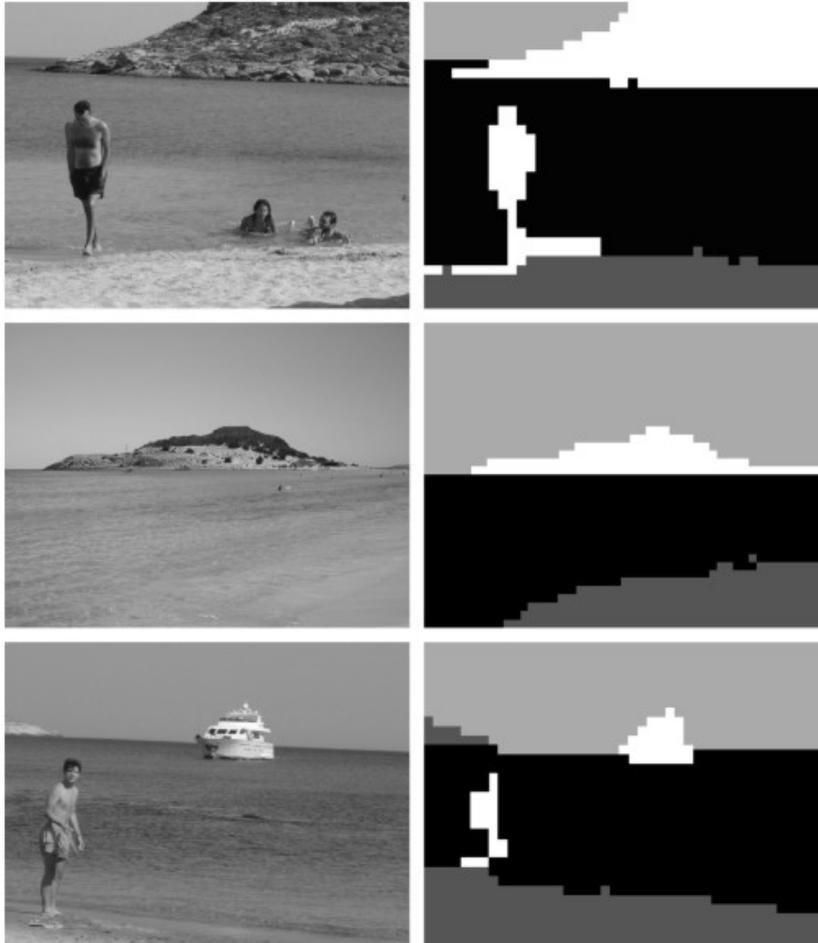
Numerische Auswertung  
für den Fußball:

Object	correct detections	false detections	missed
Field	100%	0%	0%
Player	82%	5%	13%
Spectators	70%	2%	28%



## Beispiele

1 2 3 4 5 6 7 8 9 10 11 12 13 14



### Interessant hier:

In diesem Beispiel sind die räumlichen Beziehungen zwischen den Objekten die dominanten Eigenschaften.

Z.B.: Sand is adjacent-to Sea  
Sea is below-of Sky  
Sky is above-of Sea  
Sky is above-of Sand



## Geplante Systemerweiterungen:

- Erweiterung der domain ontology durch komplexere Objektdarstellung
- Raum-Zeit Beziehungen zur Definition bzw. Erkennung von Events