Image-Based Situation Awareness Audit 1.3.2018

Sakari Lampola

Previous Audit 11.1.2018

Previous Audit

Open questions:

- Role of classical object tracking alrorithms?
- What to do with multiple bounding boxes around one object?
- Appropriate minimum confidence level?
- What to do with false detections inside other objects?
- What to do with false detections from the background?
- How to set Kalman filter parameters for image object filtering?
- Hungarian algorithms special case for hidden objects

To do:

- Close open questions
- Image object status
- Image object velocity estimation
- Probabilistic approach for matching detected and image objects
- 2d -> 3d transformation
- World object state estimation

Other:

- Semantic segmentation
- Organisations to follow: ICCV, ICRA, NIPS, IROS, arXiv
- Camera motion (yaw, pitch)
- Grid or continuos presentation?
- Class specific attributes
- Object history

Work Done

Image Object Velocity Estimation

Image object velocity is necessary for:

- predicting image object locations when matching new measurements
- identifying image objects
- predicting image object locations for hidden objects

Image object

- id
- status
- x_min
- x_max
- y_min
- y_max
- vx_min
- vx_max
- vy_min
- vy_max
- class
- confidence
- appearance

Estimation algorithm

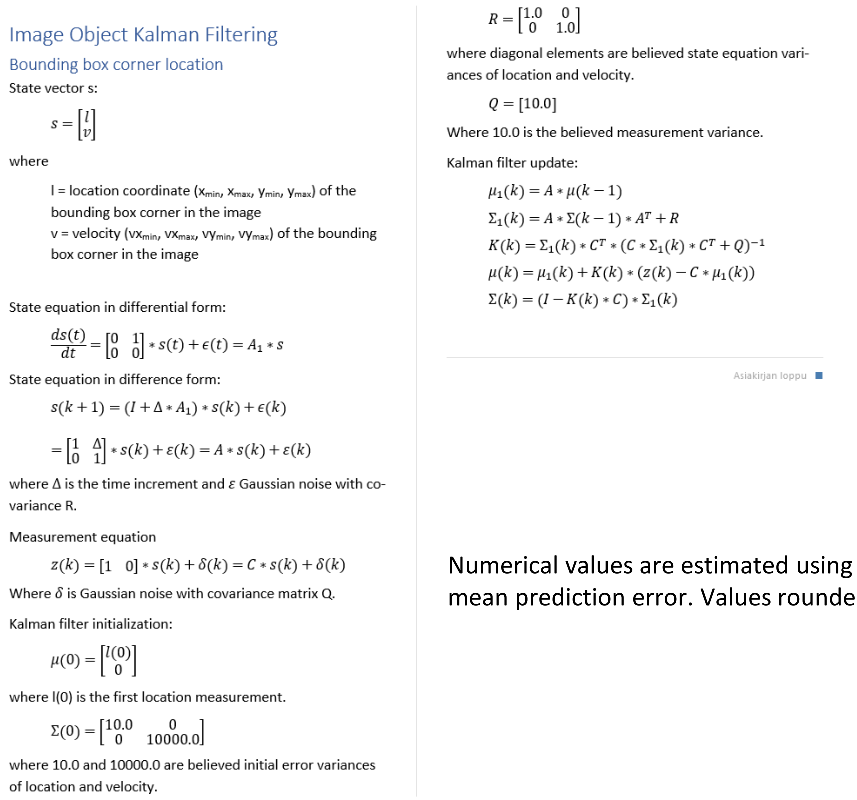


Image Object Kalman Filtering

Bounding box corner location

State vector s:

$$s = \begin{bmatrix} l \\ v \end{bmatrix}$$

where

l = location coordinate (x_{min}, x_{max}, y_{min}, y_{max}) of the bounding box corner in the image
v = velocity (vx_{min}, vx_{max}, vy_{min}, vy_{max}) of the bounding box corner in the image

State equation in differential form:

$$\frac{ds(t)}{dt} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} * s(t) + \epsilon(t) = A_1 * s$$

State equation in difference form:

$$s(k+1) = (I + \Delta * A_1) * s(k) + \epsilon(k)$$

$$= \begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix} * s(k) + \varepsilon(k) = A * s(k) + \varepsilon(k)$$

where Δ is the time increment and ε Gaussian noise with covariance R.

Measurement equation

$$z(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} * s(k) + \delta(k) = C * s(k) + \delta(k)$$

Where δ is Gaussian noise with covariance matrix Q.

Kalman filter initialization:

$$\mu(0) = \begin{bmatrix} l(0) \\ 0 \end{bmatrix}$$

where l(0) is the first location measurement.

$$\Sigma(0) = \begin{bmatrix} 10.0 & 0 \\ 0 & 10000.0 \end{bmatrix}$$

where 10.0 and 10000.0 are believed initial error variances of location and velocity.

$$R = \begin{bmatrix} 1.0 & 0 \\ 0 & 1.0 \end{bmatrix}$$

where diagonal elements are believed state equation variances of location and velocity.

$$Q = [10.0]$$

Where 10.0 is the believed measurement variance.

Kalman filter update:

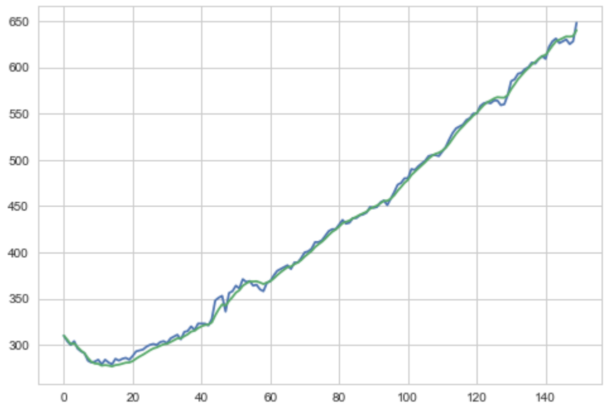
$$\mu_1(k) = A * \mu(k-1)$$
$$\Sigma_1(k) = A * \Sigma(k-1) * A^T + R$$
$$K(k) = \Sigma_1(k) * C^T * (C * \Sigma_1(k) * C^T + Q)^{-1}$$
$$\mu(k) = \mu_1(k) + K(k) * (z(k) - C * \mu_1(k))$$
$$\Sigma(k) = (I - K(k) * C) * \Sigma_1(k)$$

Asiakirjan loppu ■

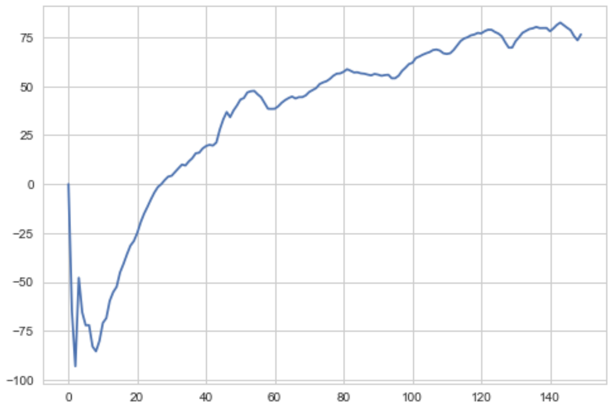
Numerical values are estimated using grid search and 10 step ahead mean prediction error. Values rounded.

Image Object Velocity Estimation

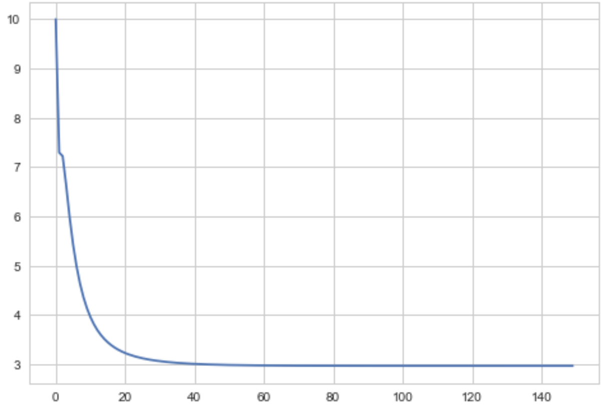
Moving object (car)



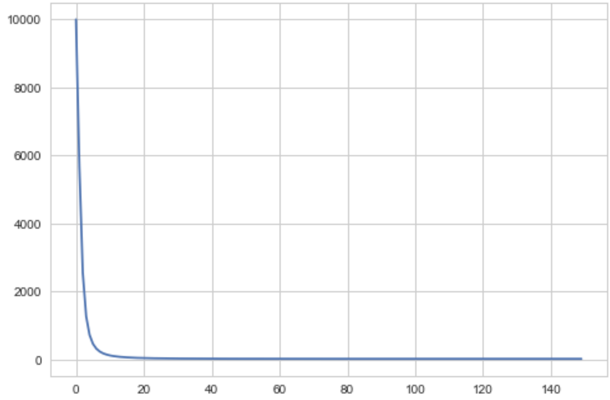
Measured and filtered location (upper left corner)



Estimated velocity



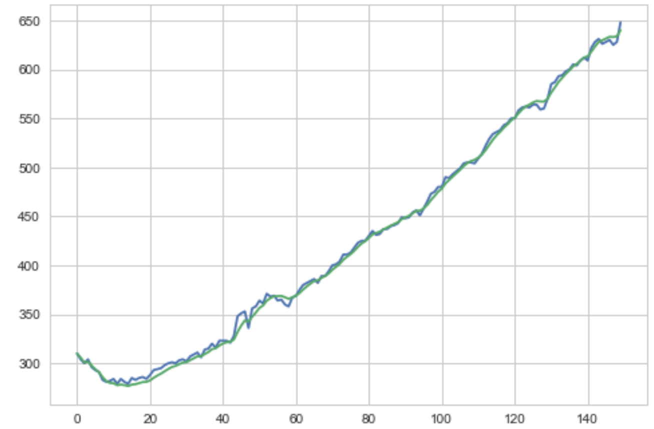
Location variance

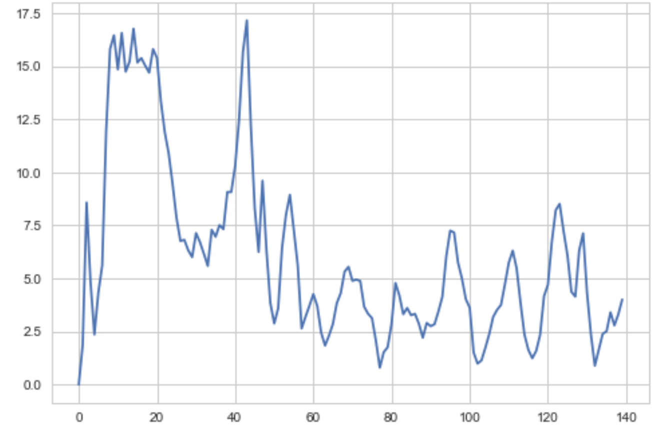


Velocity variance

Image Object Velocity Estimation

Moving object (car)

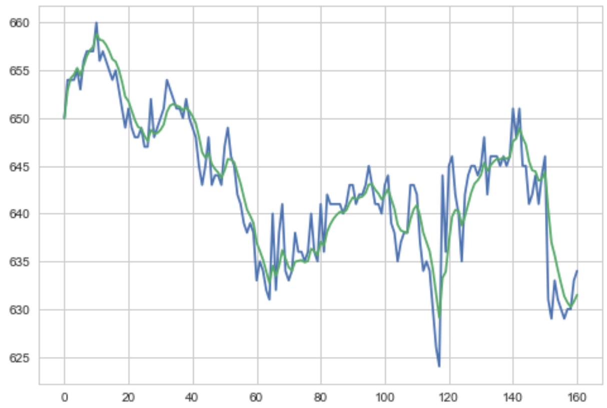




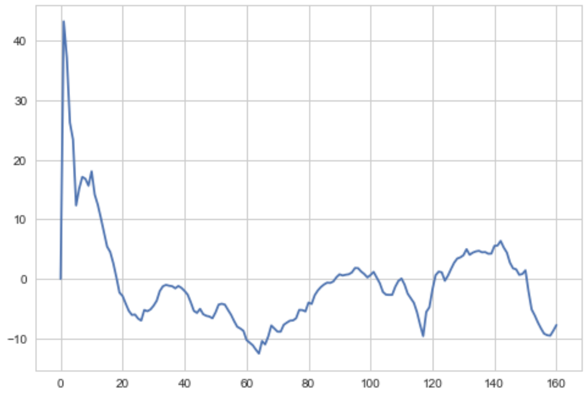
10 step ahead mean prediction error

Image Object Velocity Estimation

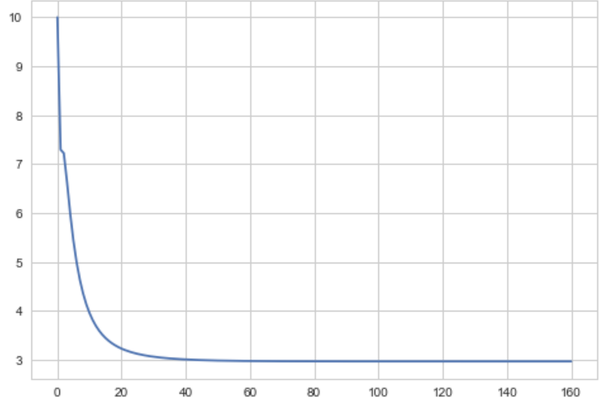
Static object (calf)



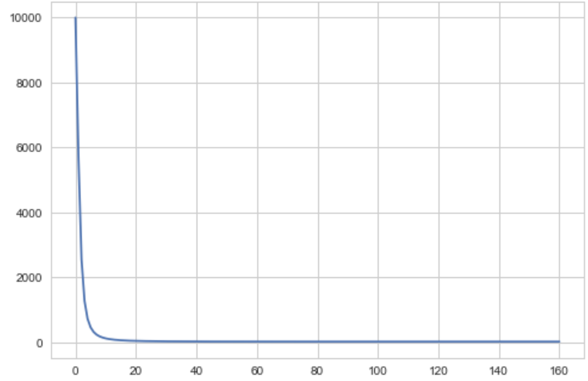
Measured and filtered location (upper left corner)



Estimated velocity



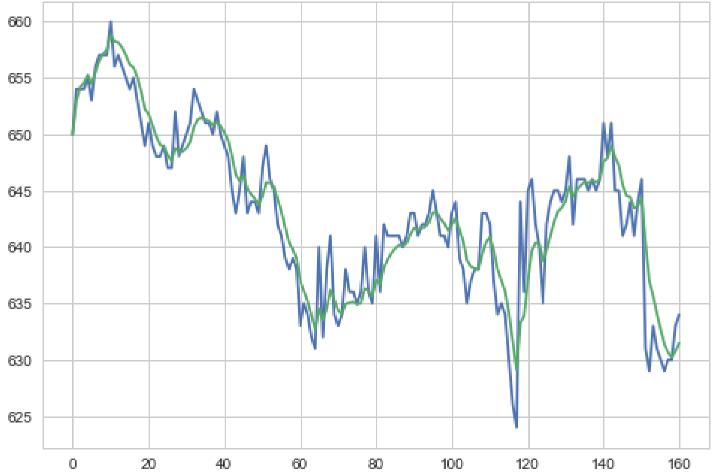
Location variance

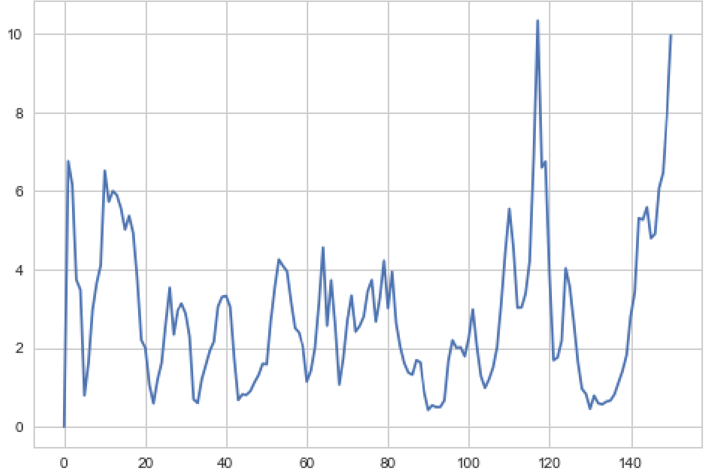


Velocity variance

Image Object Velocity Estimation

Static object (calf)

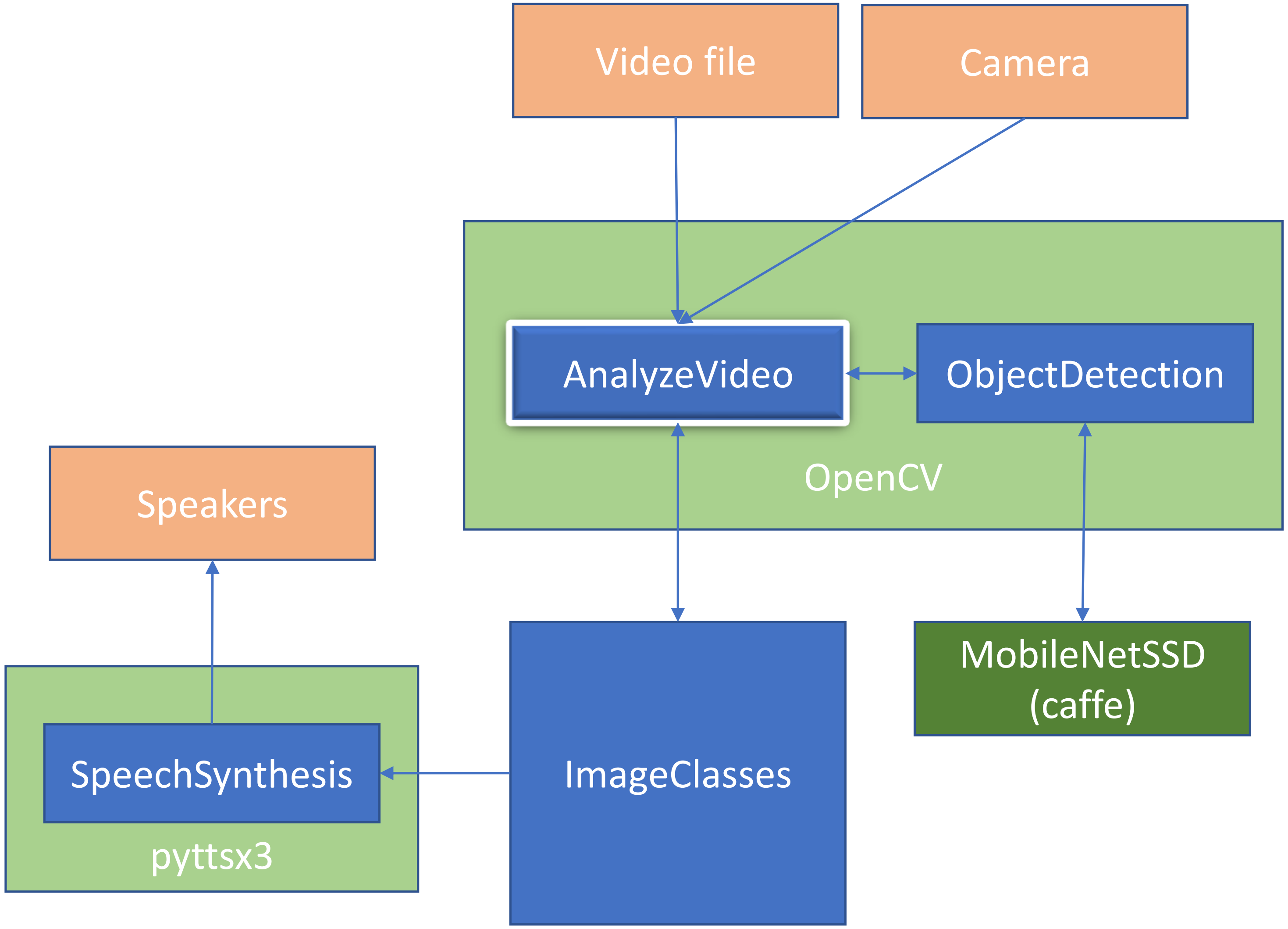




10 step ahead mean prediction error

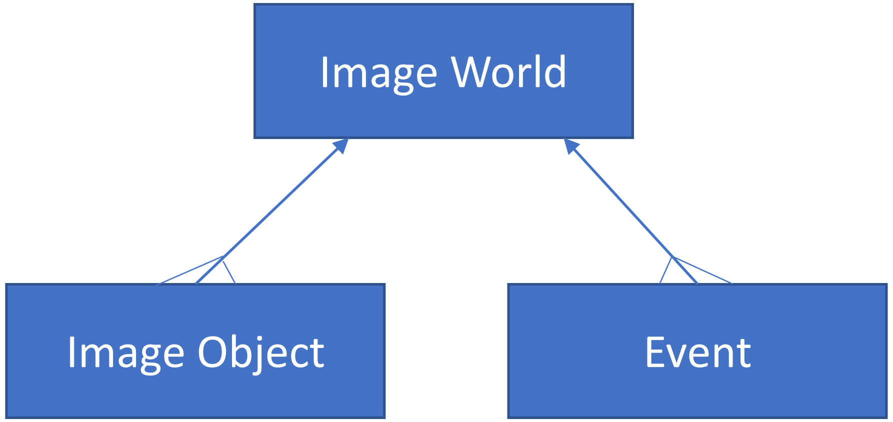
Speech Synthesis

Software Architecture



Speech Synthesis

Entities

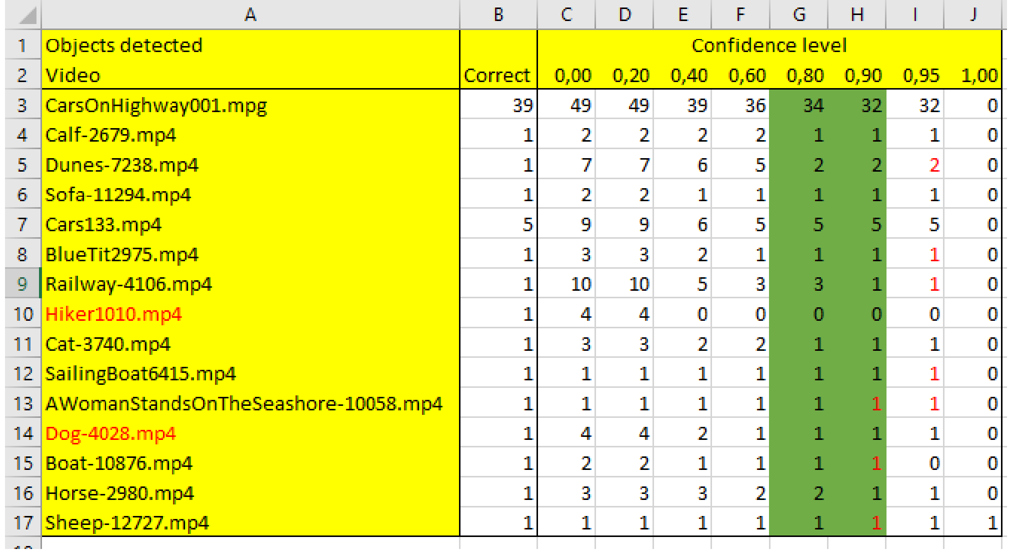


Event
- time
- text
- priority

- Event is generated when
  - new image object is created
  - image object status is changed
- Event will pause the video for the duration of speech (not in the final version)
- Events are collected (history)

Confidence Level

SSD Mobilenet implementation:
# extract the confidence (i.e., probability) associated with the prediction

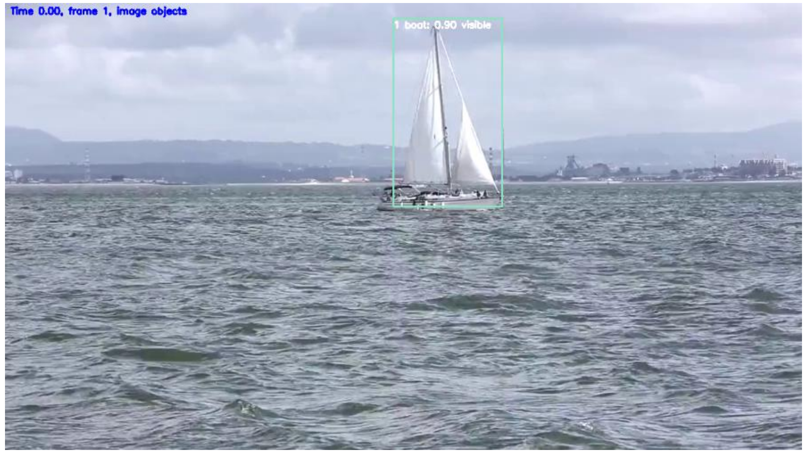


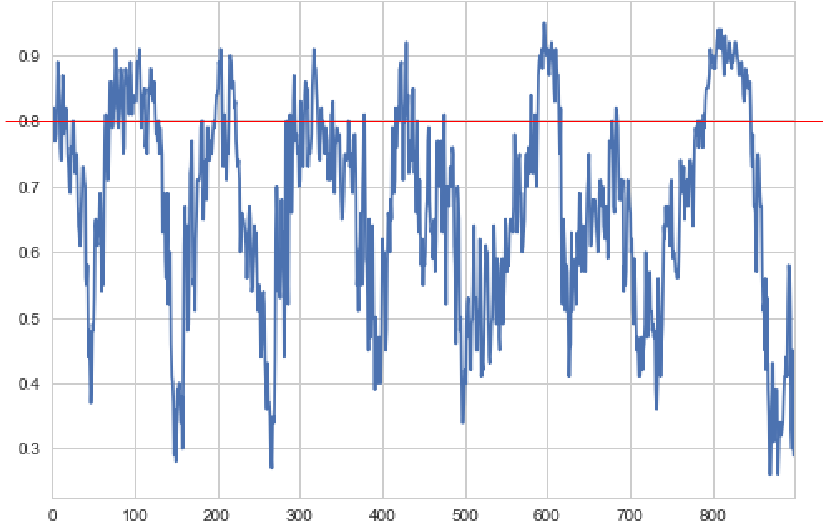
	A	B	C	D	E	F	G	H	I	J
1	Objects detected		Confidence level							
2	Video	Correct	0,00	0,20	0,40	0,60	0,80	0,90	0,95	1,00
3	CarsOnHighway001.mpg	39	49	49	39	36	34	32	32	0
4	Calf-2679.mp4	1	2	2	2	2	1	1	1	0
5	Dunes-7238.mp4	1	7	7	6	5	2	2	2	0
6	Sofa-11294.mp4	1	2	2	1	1	1	1	1	0
7	Cars133.mp4	5	9	9	6	5	5	5	5	0
8	BlueTit2975.mp4	1	3	3	2	1	1	1	1	0
9	Railway-4106.mp4	1	10	10	5	3	3	1	1	0
10	Hiker1010.mp4	1	4	4	0	0	0	0	0	0
11	Cat-3740.mp4	1	3	3	2	2	1	1	1	0
12	SailingBoat6415.mp4	1	1	1	1	1	1	1	1	0
13	AWomanStandsOnTheSeashore-10058.mp4	1	1	1	1	1	1	1	1	0
14	Dog-4028.mp4	1	4	4	2	1	1	1	1	0
15	Boat-10876.mp4	1	2	2	1	1	1	1	0	0
16	Horse-2980.mp4	1	3	3	3	2	2	1	1	0
17	Sheep-12727.mp4	1	1	1	1	1	1	1	1	1

Good value is between 0.8 and 0.9. 0.8 selected for now.

The ’good’ value also depends on other hyperparameters.

Confidence Level

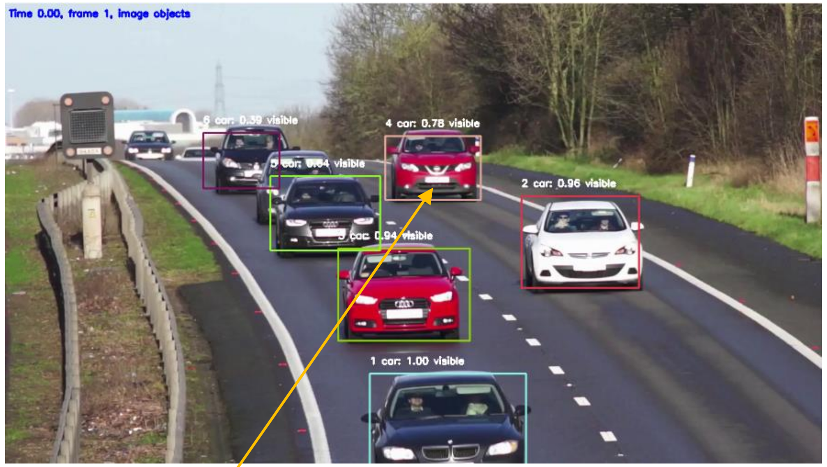


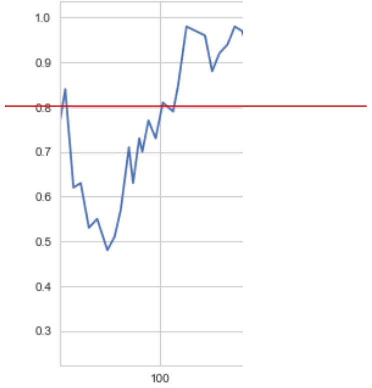


status = visible

status = hidden

Confidence Level

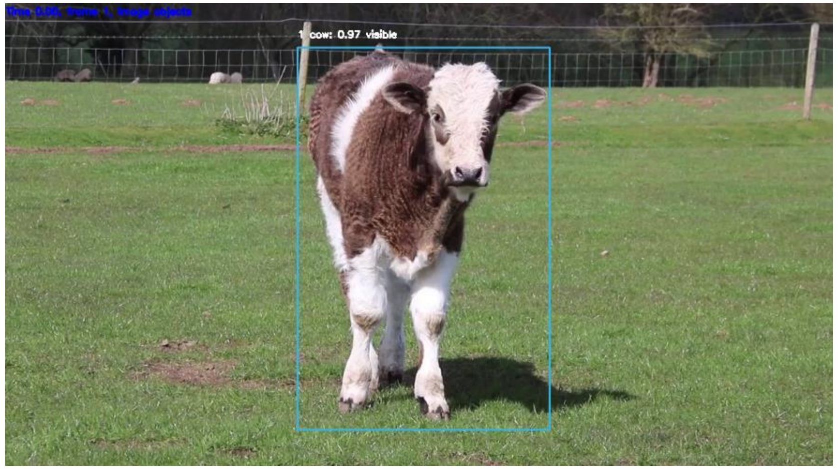


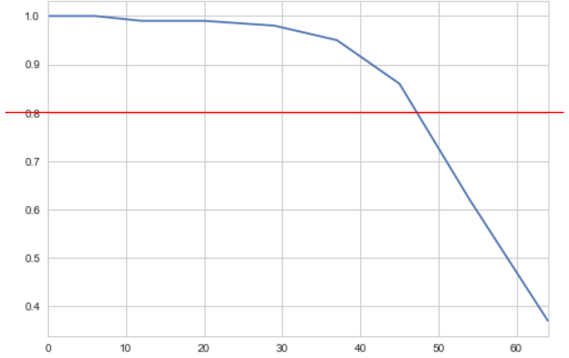


status = visible

status = hidden

Confidence Level





status = visible

status = hidden

Hyper-parameters

- Minimum confidence level
- Retention time
- Margin width
- Minimum size
- Similarity distance

Work in Progress

Perception

“The first step in achieving SA is to perceive the status, attributes, and dynamics of relevant elements in the environment. Thus, Level 1 SA, the most basic level of SA, involves the processes of monitoring, cue detection, and simple recognition, which lead to an awareness of multiple situational elements (objects, events, people, systems, environmental factors) and their current states (locations, conditions, modes, actions).”

Next Steps

Next steps

Comprehension:

1. Closing the open questions
2. 2d -> 3d transformation
3. World object state estimation

To Be Discussed

Method follow-up

- Google search enough?
- Good way of following new papers?

Thank you!

lampola@student.tut.fi
https://github.com/SakariLampola/Thesis