Image-Based Situation Awareness Audit 8.5.2018

Sakari Lampola

Previous Audit 28.2.2018

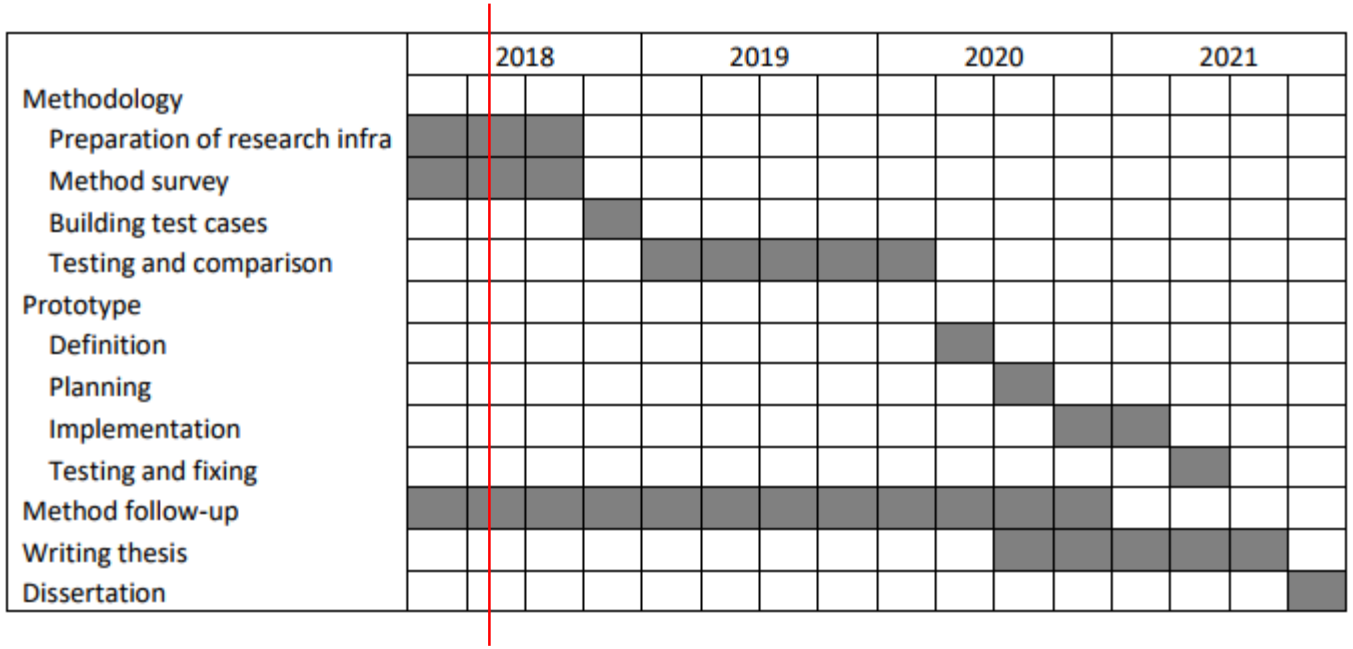
Next steps

- Kalman filter parameter adjustments (Q1) 🟢
- Dataset selection (Q1) 🟢
- Stereo vision (Q2) 🟢
- Camera yaw, pitch, roll estimation (Q2) 🟢
- Speech recognition (Q2) 🟡
- Semantic segmentation (Q2) 🟡
- ~~Experiments in the wild (Q2)~~
- Paper (Q3) 🔴
- Speech analysis (Q3) 🔴
- Speech generation (Q3) 🔴
- Use cases (Q4) 🟡

Other

- Body forecast
  - kinetic
  - based on class history
  - based on swarm history
- R matrix estimation
- Monograph or papers?

Project Plan



1. Methodology / Preparation of research infra
    a. Software platforms are constructed and tested
    b. Off-the-shelf models are acquired and tested
    c. Necessary skills on platforms are learned
2. Methodology / Method survey
    a. Current state-of-art methods are studied
    b. Methods are constructed and tested on the software platforms
3. Method follow-up
    a. Screening of conference papers related to the subject
    b. Possibly integrating new methods to the project

Work Done

- Dataset selection
  - KITTI
  - COCO
  - VIZVIZ
- Stereo
- Kalman filter parameter optimization
- 3D bounding box
- Grid world
- Moving camera and world coordinates
- Optical flow
- Use cases
- Semantic segmentation implementation search
- Speech recognition test

Hands-on experience with tools and methods achieved.

The Big Picture

Thing Detection

Detections

Tracking

Things

State Estimation

Stereo Depth Estimation

Depth

Structure From Depth

Thing Locations And Velocities

Stuff Segmentation

Stuff

Stuff Location

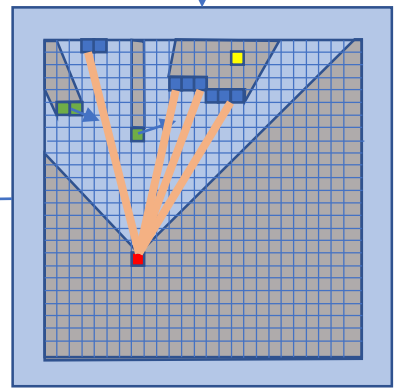
World Generation

Optical Flow

ROI Flow

Egomotion Estimation

Observer Location



Planning

Actions

Functionality

Data

Contribution

World Representation

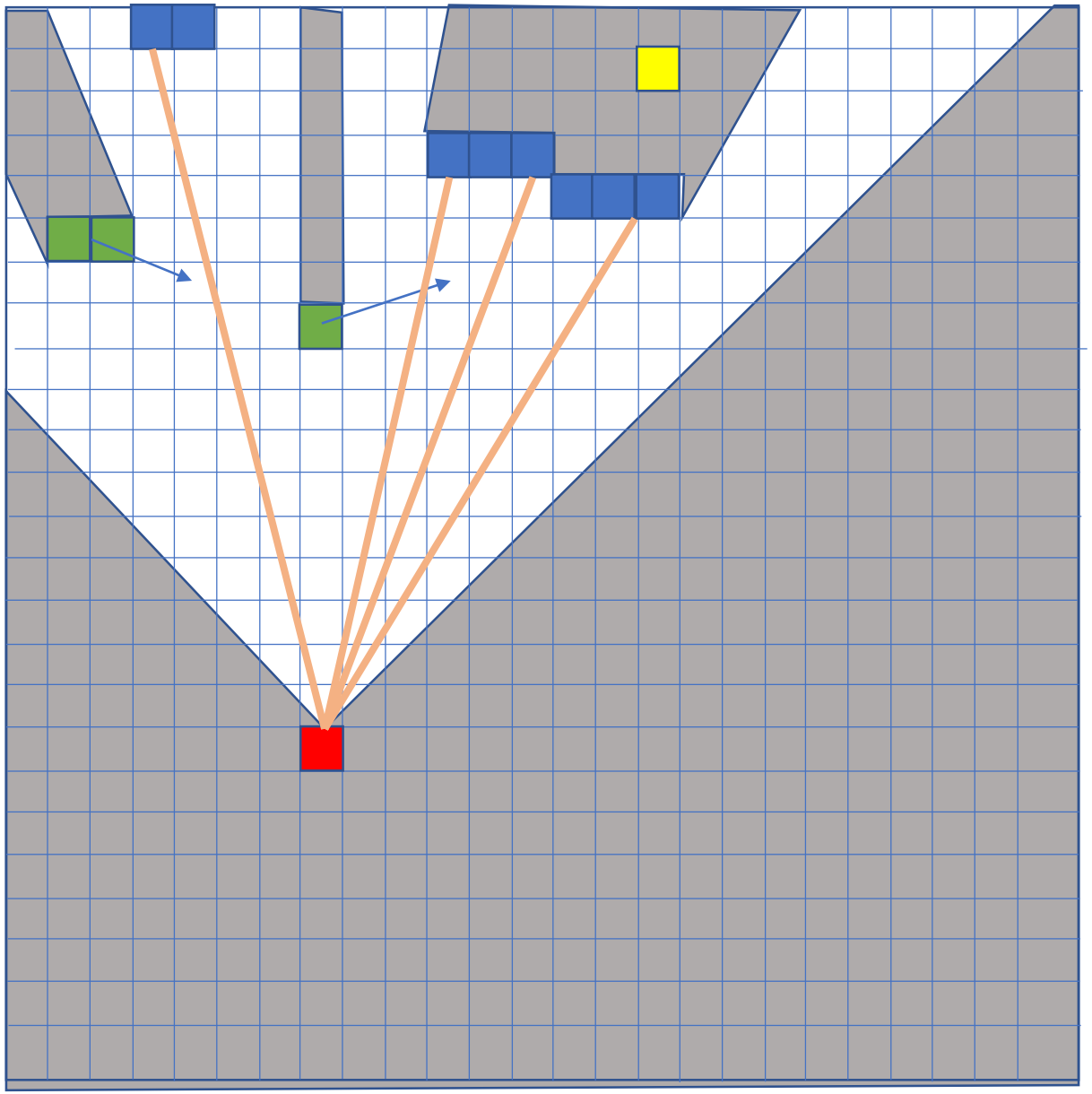
World representation

- thing
- stuff
- unknown
- observer
- goal

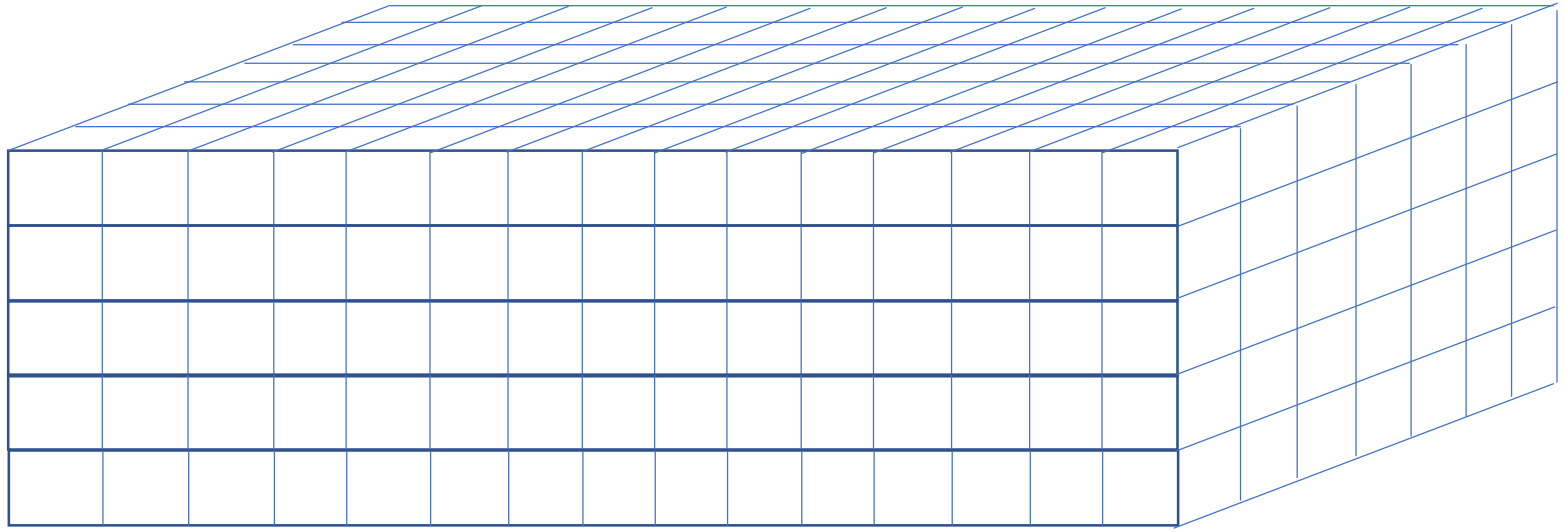
Assumptions:

- Stuff does not move
- Things move with constant velocity
- State estimation provides thing location and velocity + uncertainties
- State consists of
  - Stuff locations
  - Thing locations and velocities
  - Observer location
  - Confidences (known/unkown)
- Things move independently
- State is fully observable
- Observer motion can be estimated from relative stuff locations

Top view



Voxels



Examples:

- Outdoor environment: area of interest 100m*100m*10m, 1m3 voxel => 100 000 voxels
- Indoor environment: area of interest 10m*10m*3 m, 10cm*10cm*10cm voxel => 300 000 voxels

?

Things, stuff, observer and confidence are represented as occupancy maps.

Next Steps

- For each functionality:
  - Paper and literature references are collected
  - Document (paper) is written
- For each data:
  - Representation is described
- Prototype is constructed
  - Based on COCO things and stuff
  - Selected methods are integrated and tested
- Summary document is written:
  - Overall description
  - Test results

Initial list of papers:

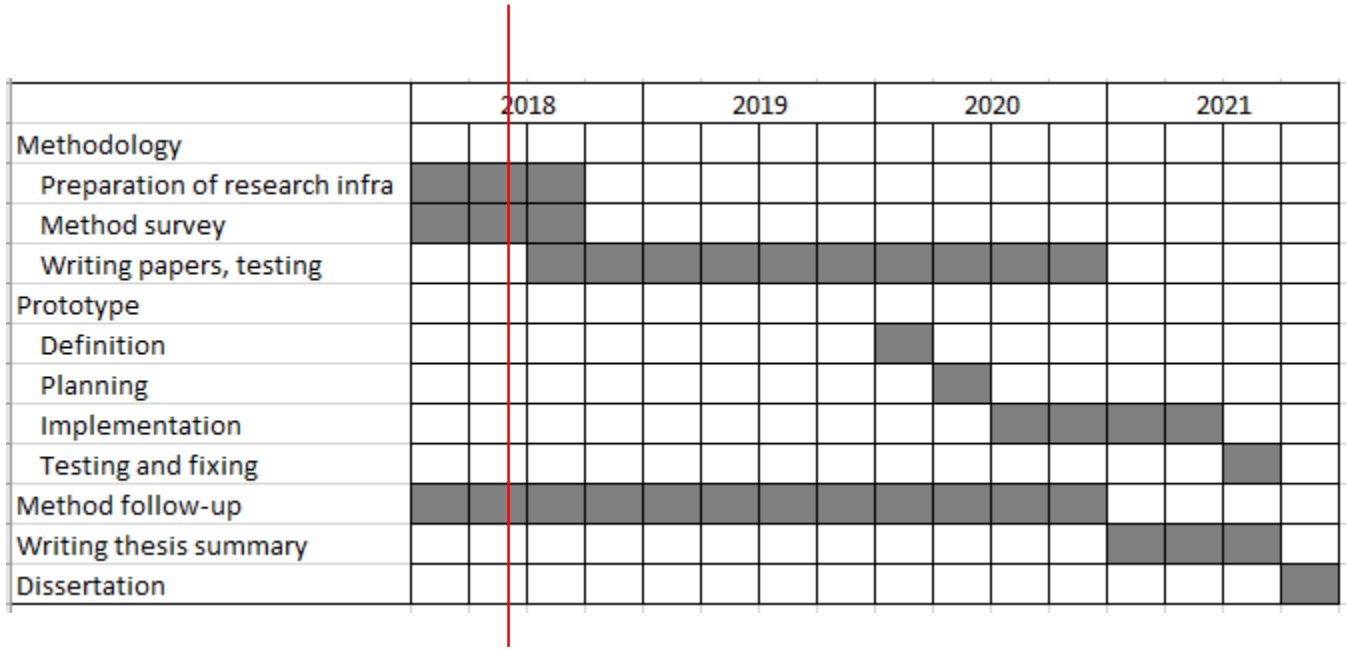
- ”Thing Detection for Image-Based Situation Awareness”
- ”Stereo Depth Estimation for Image-Based Situation Awareness”
- ”Stuff Segmentation for Image-Based Situation Awareness”
- ”Optical Flow for Image-Based Situation Awareness”
- ”Tracking for Image-Based Situation Awareness”
- ”Structure from Depth for Image-Based Situation Awareness”
- ”Egomotion Estimation for Image-Based Situation Awareness”
- ”State Estimation for Image-Based Situation Awareness”
- ”World Generation for Image-Based Situation Awareness”
- ”Planning for Image-Based Situation Awareness”

Names will probably change….

Summary includes:

- World representation
- Overall system structure
- Prorotype description
- Test results

Revised Project Plan



Thank you!

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