DANISH TECHNOLOGICAL INSTITUTE

Centre for Robot Technology

QUAD-AV project

Ambient sensing technologies of Automatically Guided Vehicles

Michael Nielsen

Ambient **A**wareness for **A**utonomous **A**gricultural **V**ehicles (QUAD-AV)

DANISH TECHNOLOGICAL INSTITUTE

SEVENTH FRAMEWORK PROGRAMME

Centre for Robot Technology

Autonomous vehicles are being increasingly adopted in agriculture to improve productivity and efficiency. For an autonomous agricultural vehicle to operate safely, environment perception and interpretation capabilities are fundamental requirements. QUAD-AV focuses on the development of sensors and sensor processing methods (stereo vision, thermal, radar, lidar) to provide an autonomous agricultural vehicle with such ambient awareness.

Positive obstacles

Negative obstacles

Moving obstacles / Live animals / People

Difficult Terrain

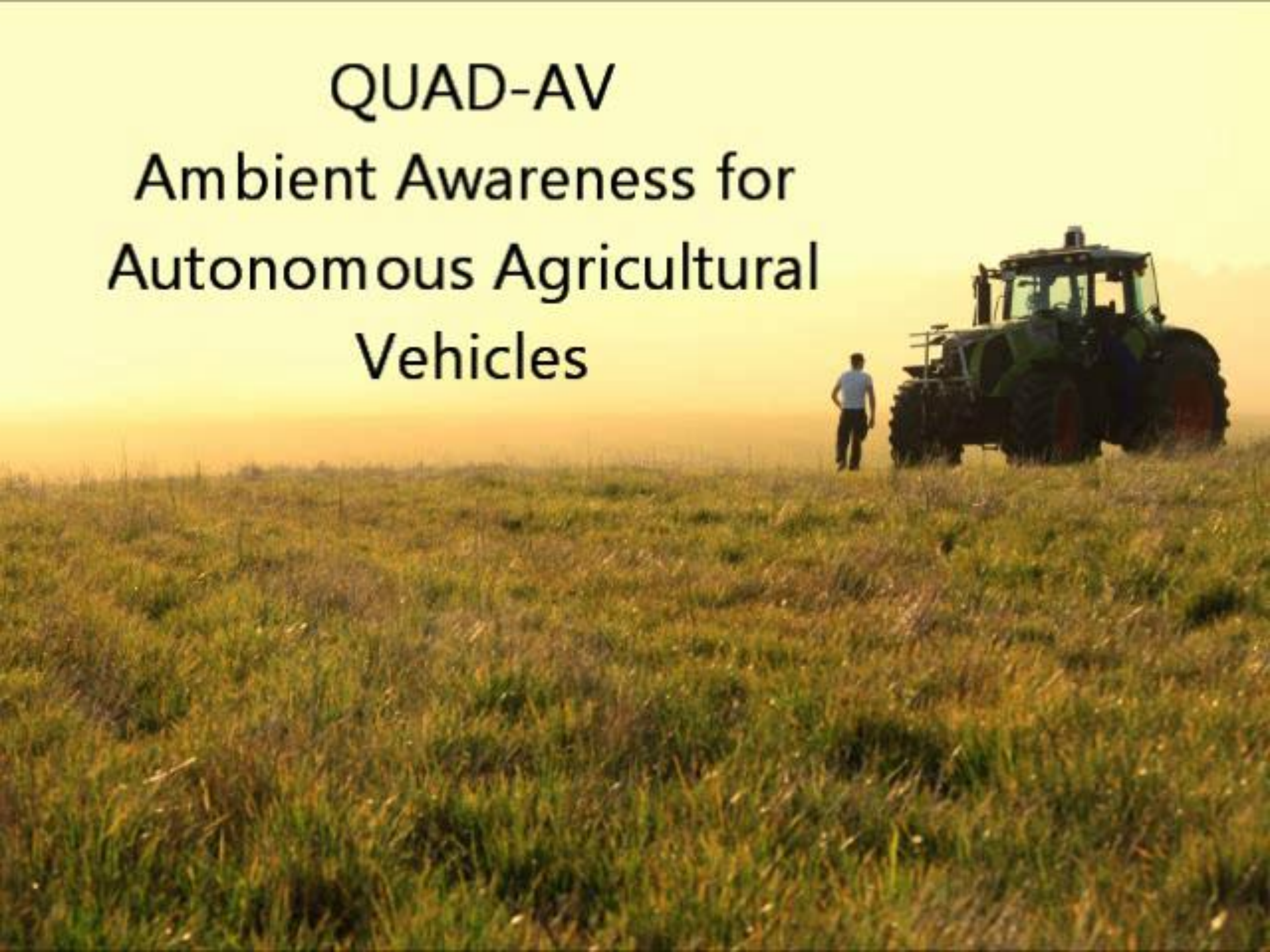
Fraunhofer

UNIVERSITÀ DEL SALENTO

CLAAS

irstea

TEKNOLOGISK INSTITUT


QUAD-AV
Ambient Awareness for
Autonomous Agricultural
Vehicles

Multi-sensor setup

Sensors Overview

Sensor	FOV (HxV) [deg]	FPS	Range	Visibility
Stereo with Multibaseline And HDR	70x52 (controllable by lens)	1-10Hz	1-50m in theory But in practice should be 1-20m	SNR and range is affected by visible particles in the air (fog, dust, smoke, rain, snow)
Thermography	30x24 or 40x30 (with stock lenses)	50 Hz (In theory, 20 Hz with robust processing)	Set it up like the stereo rig	Smoke does not affect it. Range is affected in cat 3 fog, but much less than visual cameras for cat1-2 fog and rain.
LIDAR	360x270	0.19 Hz	80m	Affected by rain, fog but not in smoke
RADAR	360x25	1 Hz	100m	Not affected

Sensors Overview

Sensor	Pros	Cons
Stereo	Colour pointclouds Texture visual classification Can model the terrain, see ditches, holes, positive obstacles incl living positive obstacles	High processing load Occlusion Difficult to see lakes and living things in vegetation. Needs visibility. Short narrow range (to solve other issues with visibility)
Thermography	Vehicles and living objects stand out Structures, water, poles also integrate a lot of heat.	No 3D representation Increasing noise that needs to be recalibrated every few minutes which takes a few seconds. Temperatures depend a lot on the weather so everything must be adaptive
LIDAR	Easy to setup, low noise point clouds 360 degrees full 3D	Only point clouds Slow rotation at the moment Affected by visibility issues
RADAR	Easy setup SLAM maps without GPS 360 degrees, 100 meters radius Can see lakes Unaffected by low visibility environments	Low resolution Only 2D maps

IRSTEA (old Cemagref): RADAR

- Carrier frequency 24.125 GHz
- Bandwidth 200 MHz
- Transmitted power (EIRP) 18 dBm
- Range 3-100 m
- Size (length-width-height) 27-24-30 cm
- Weight 10 kg

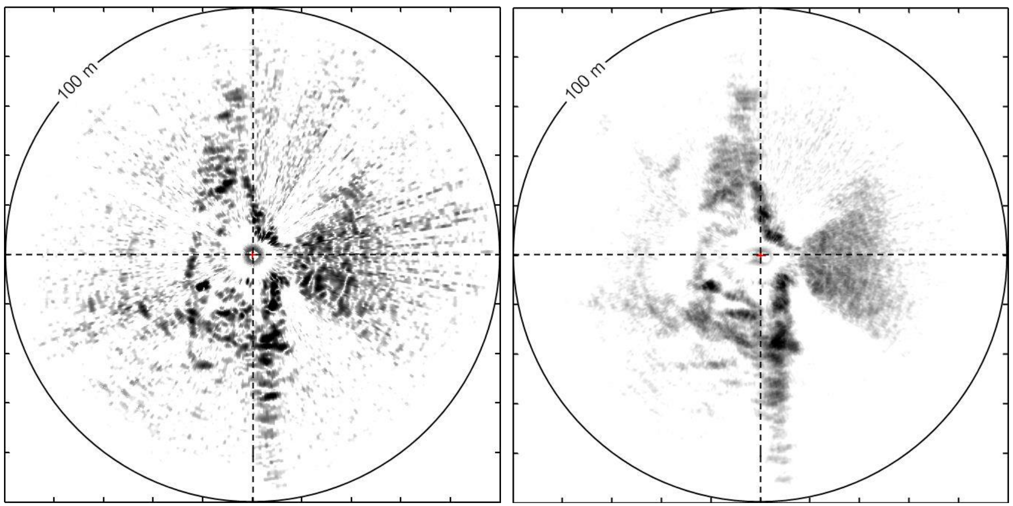
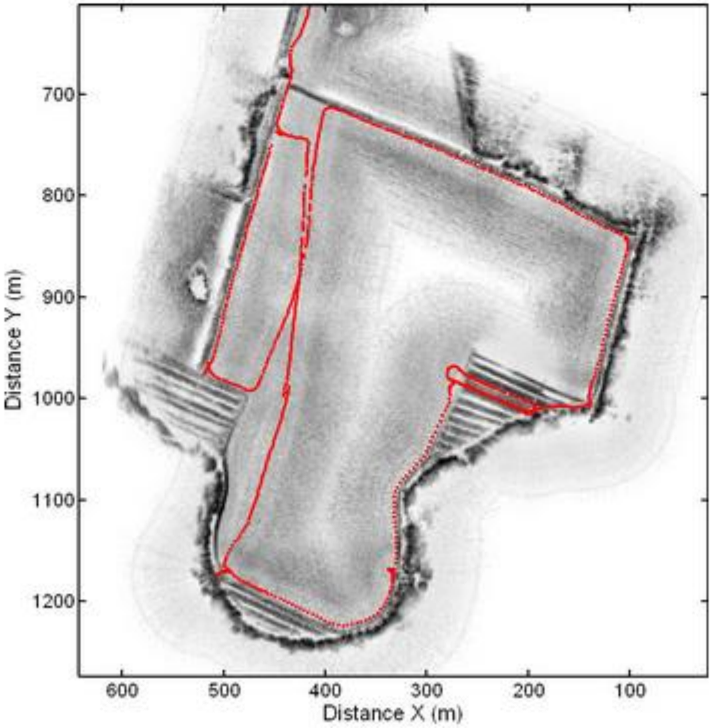


Figure 1. Example of panoramic radar images. The radar is positioned at the center of the image. (a) Initial radar image. (b) A multilook filter allows to reduce the speckle effect.

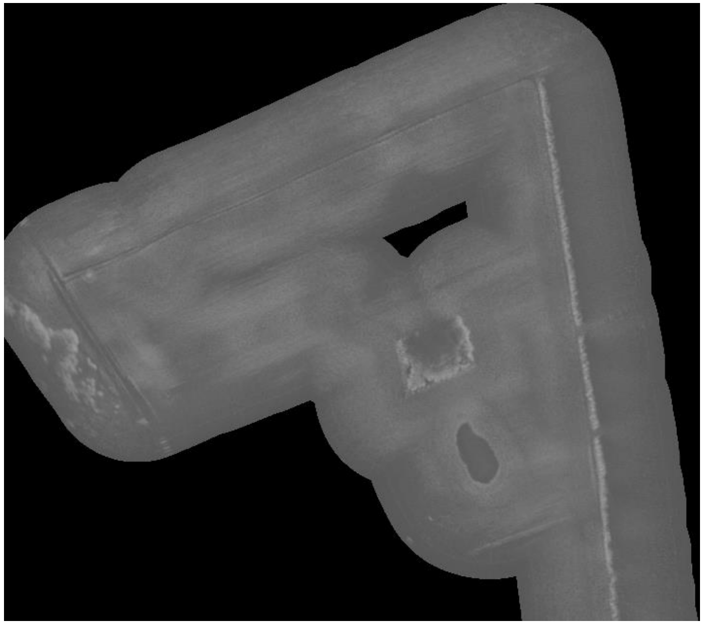
IRSTEA (old Cemagref): RADAR based SLAM (no gps used)

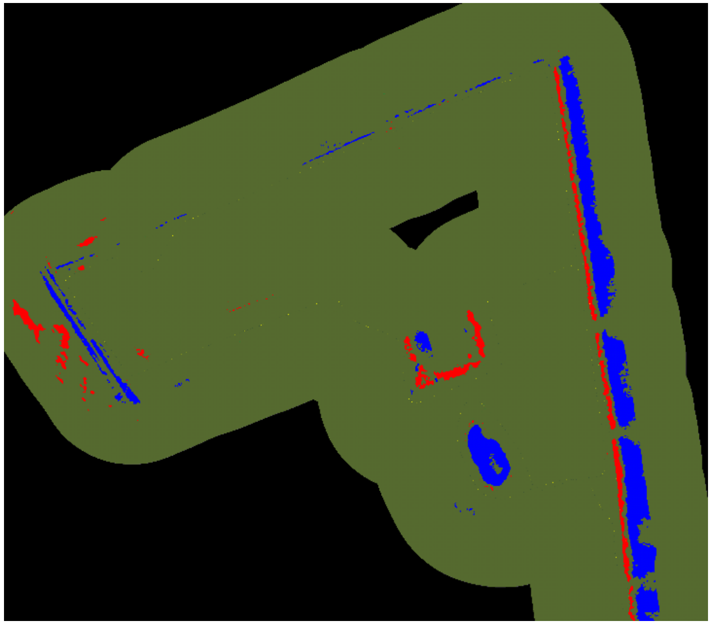


Radar map processing

- using computer vision techniques

atm. occluded regions and lakes are grouped together, but when the tractor moves closer to an occluded area, the map will be updated.



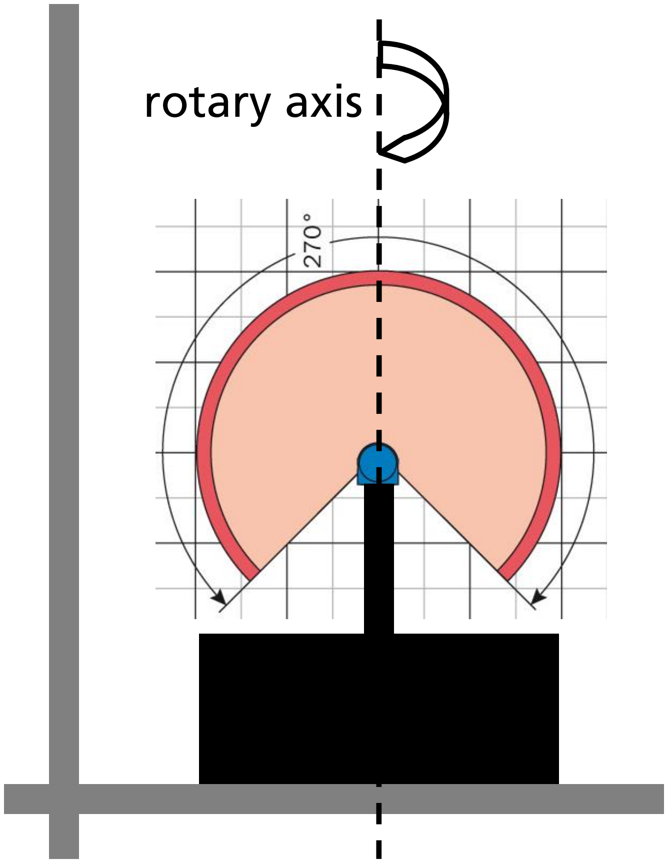


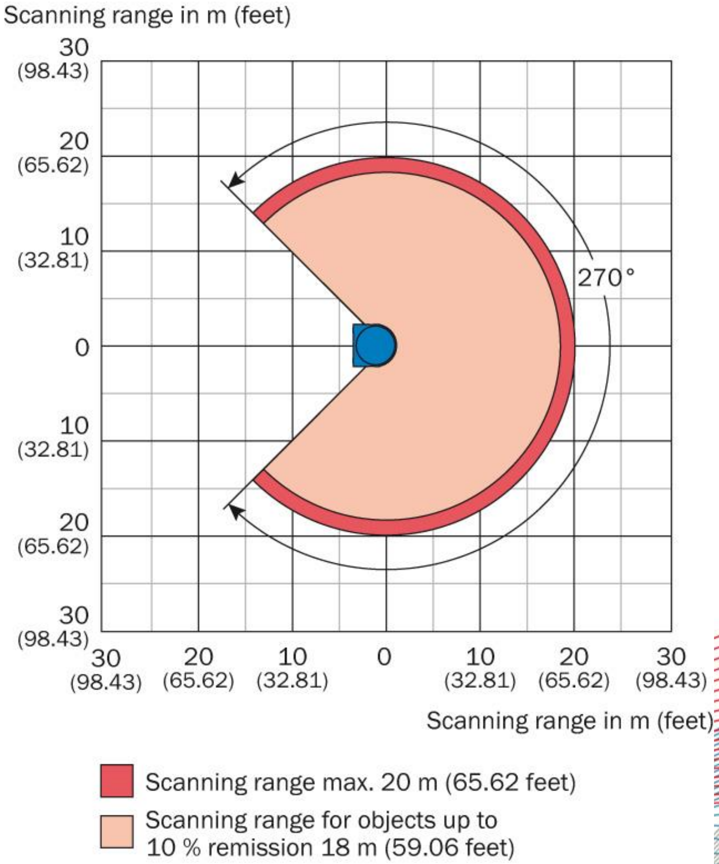
LIDAR

- based on off-the-shelf laser range finder (SICK LMS 111 or 211)
- mounted in different ways on a rotational unit to acquire 3D point clouds

LIDAR

- mounting and scanning range of LMS 111 as a 3D laser scanner





LIDAR

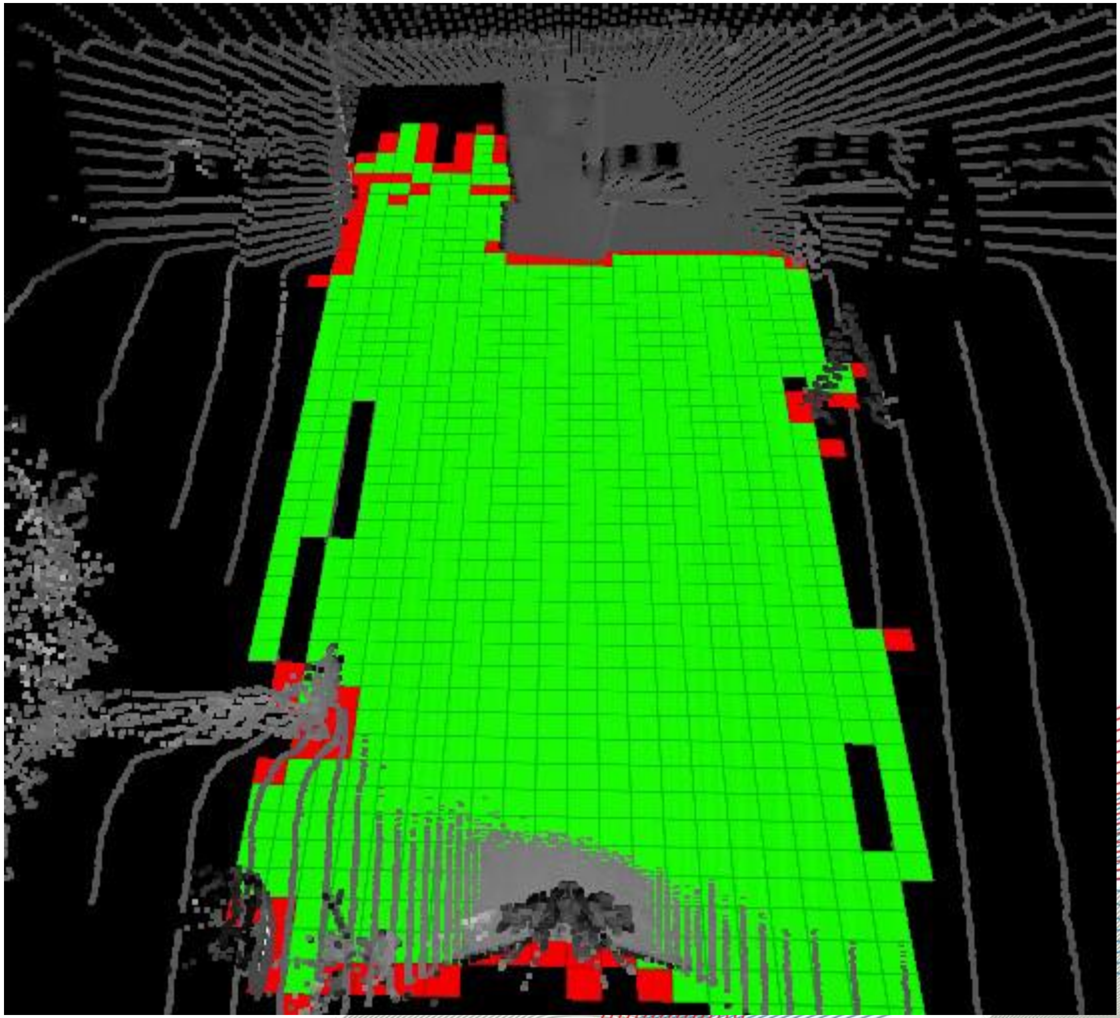
- traversability map for the vehicle created from a 3D point cloud

Idea:

- raster the sensed environment by a voxel grid
- use the slope between occupied voxels as a criterion for obstacles

Advantages:

- accessibility is independent of absolute height of objects
- downhill gradient is taken into account, too

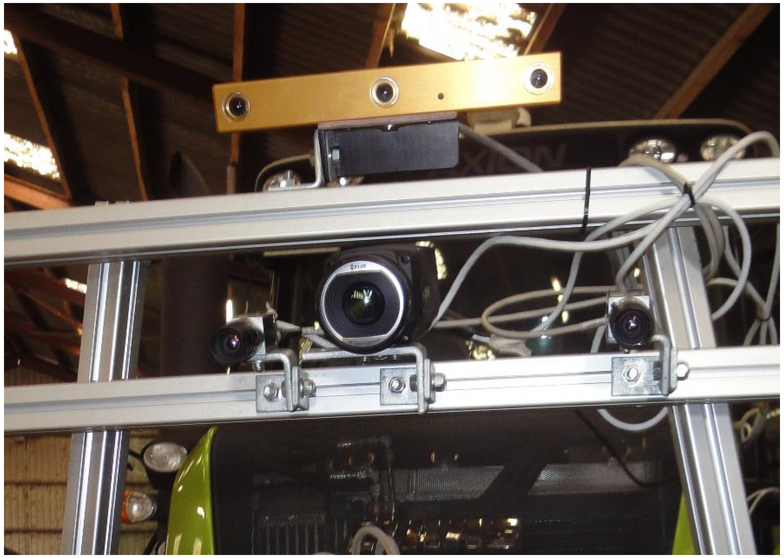


Stereo vision

DANISH TECHNOLOGICAL INSTITUTE

Centre for Robot Technology

Stereo system	Description
Multibaseline	Can fuse good results from short and long range
HDR	Can fuse good images from bright and dark areas



Stereo vision

Normal

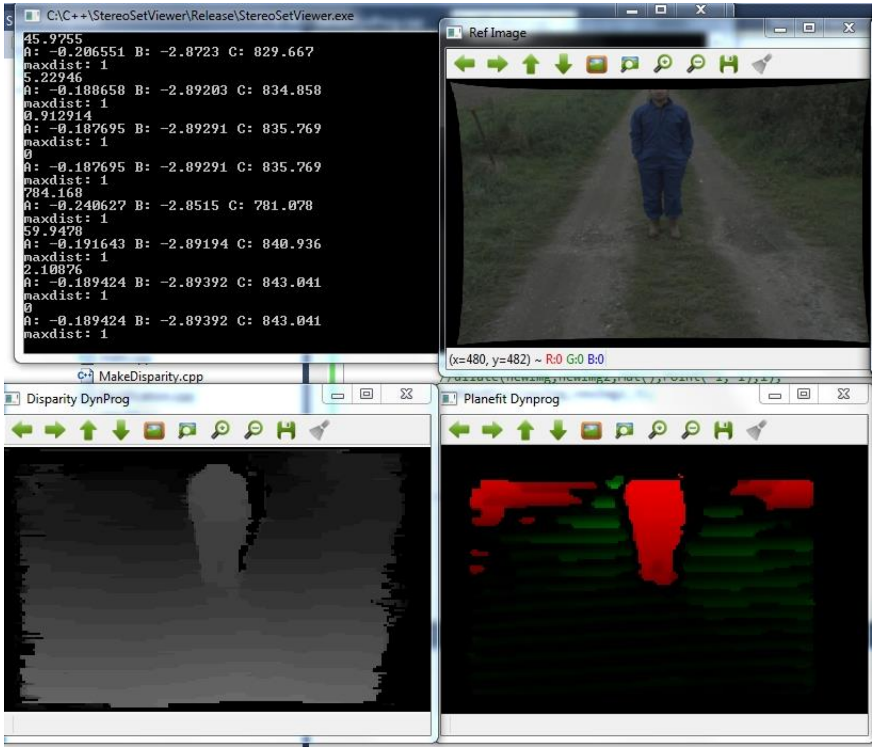
HDR

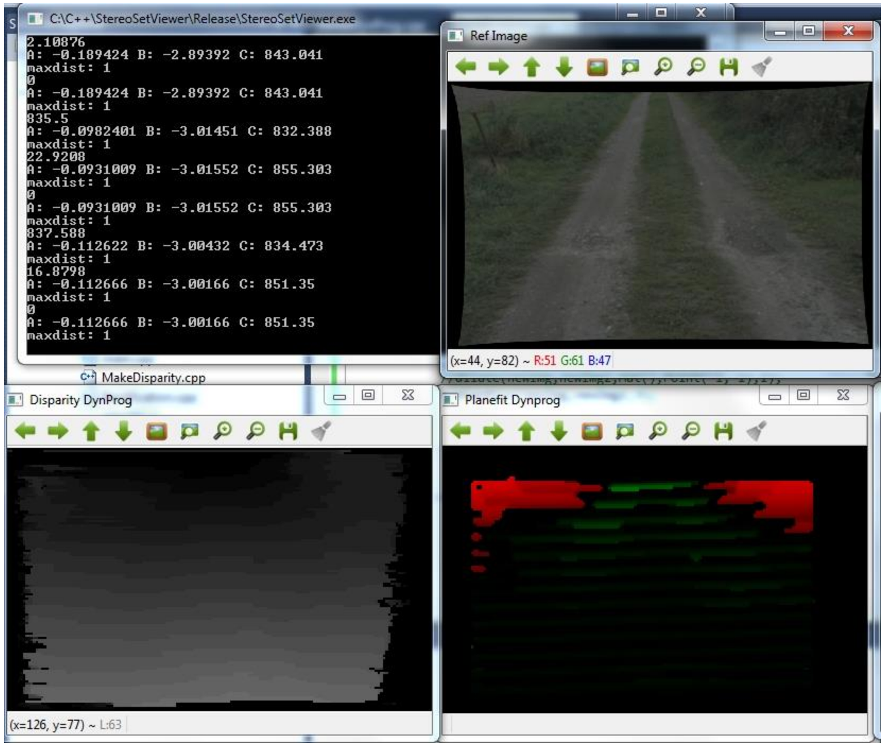
Outdoor vision HDR

Stereo Vision – Easy Task
Deviation from ground plane

DANISH TECHNOLOGICAL INSTITUTE

Centre for Robot Technology

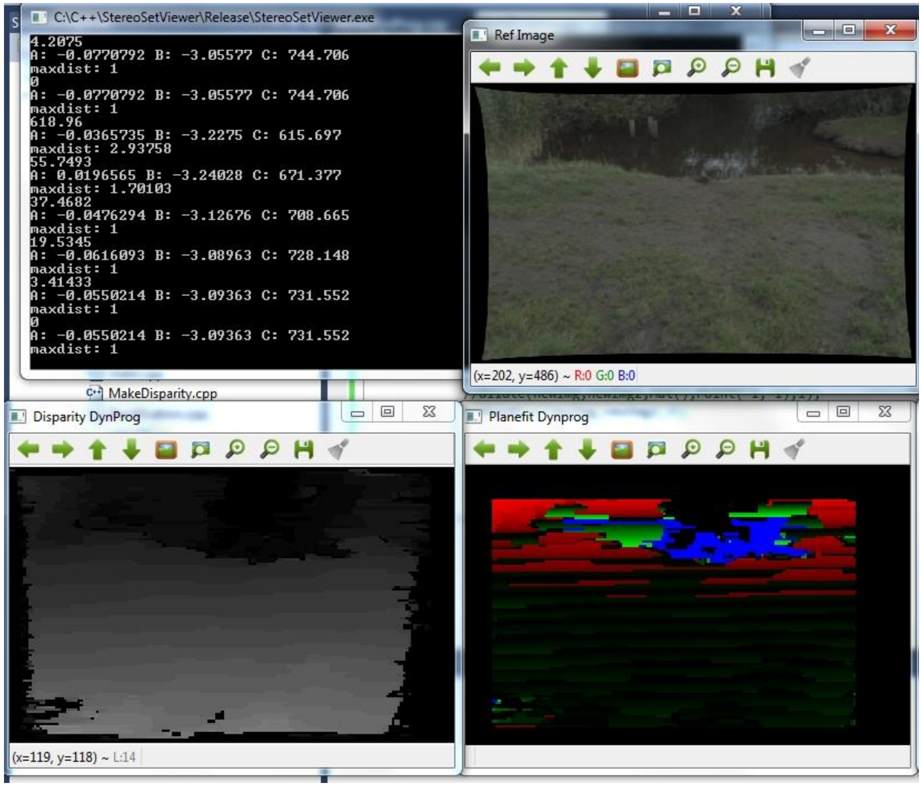


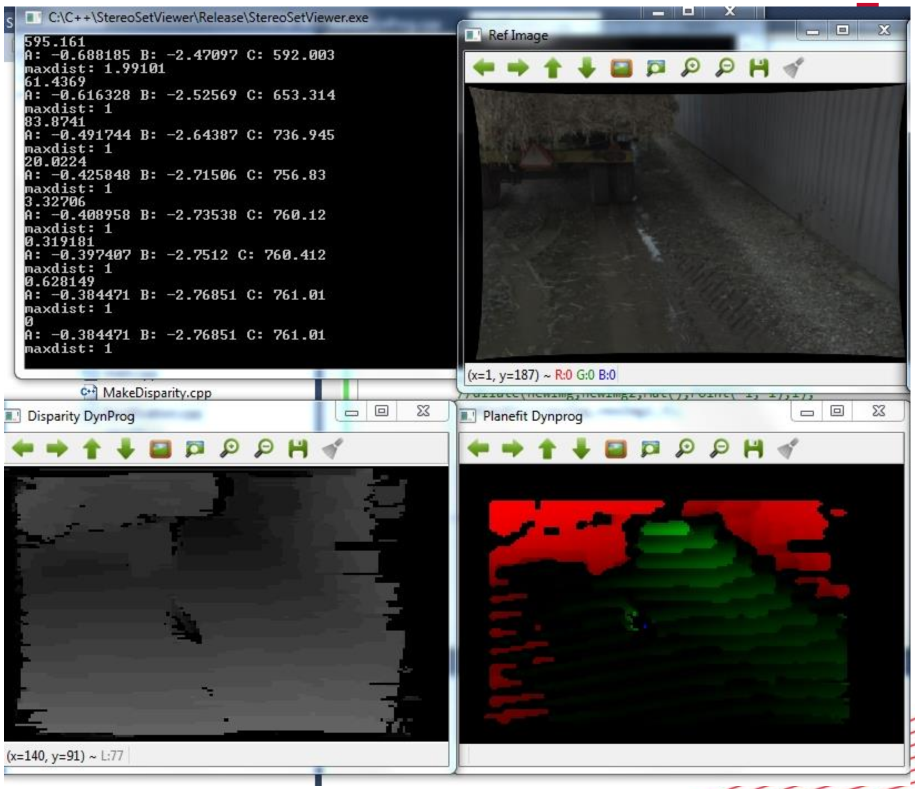


Stereo Vision
Negative obstacle, too

DANISH TECHNOLOGICAL INSTITUTE

Centre for

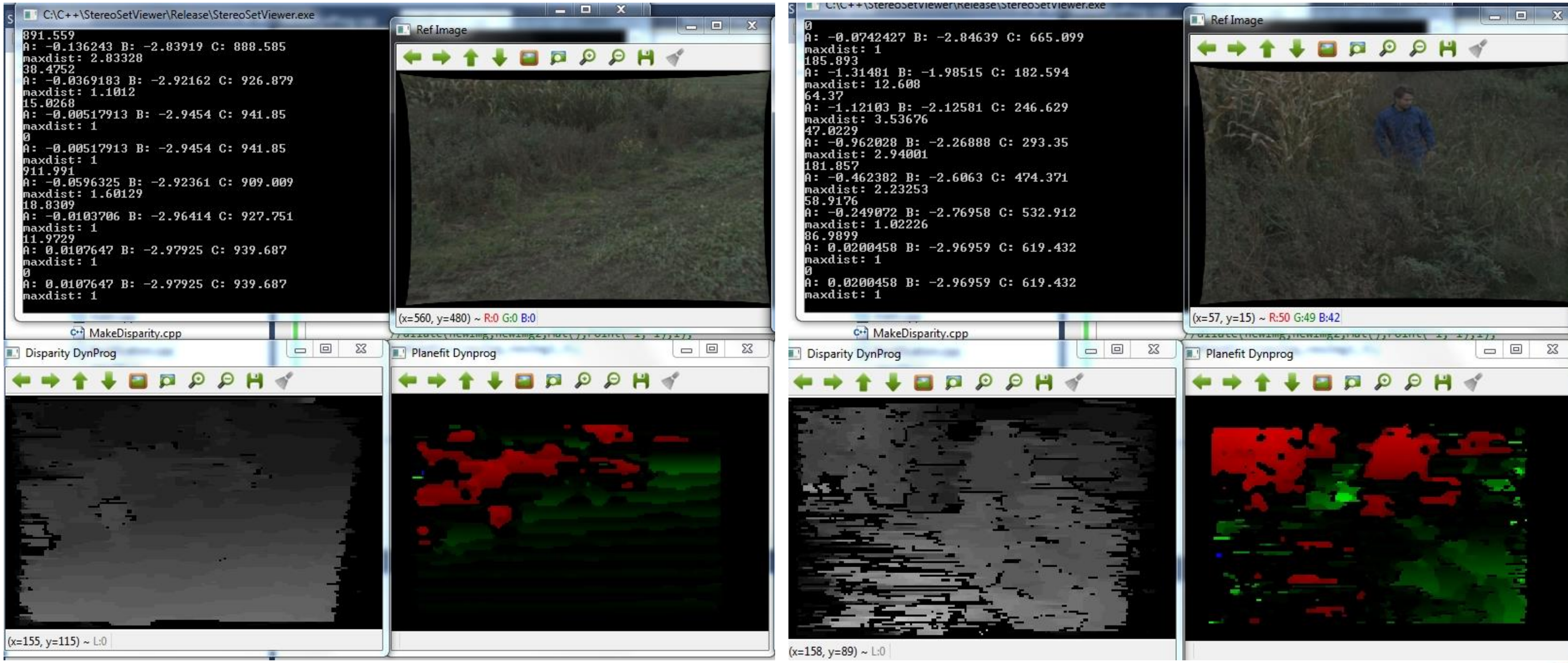




Stereo Vision
- but what about tall crops?

DANISH TECHNOLOGICAL INSTITUTE

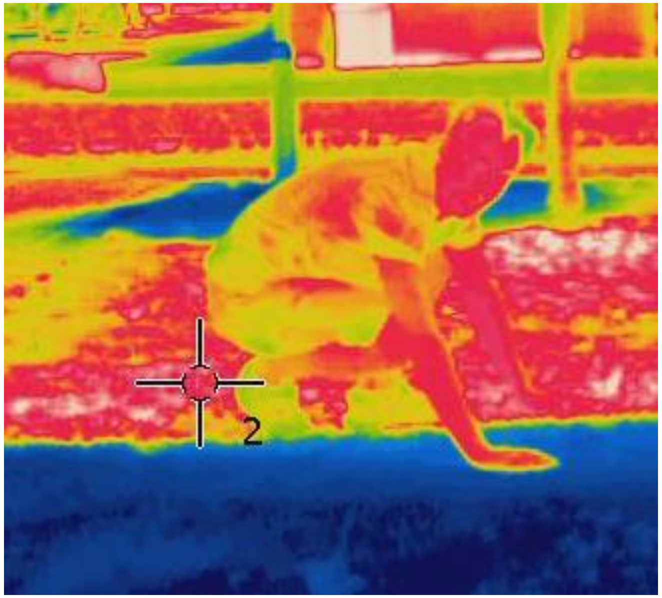
Centre for

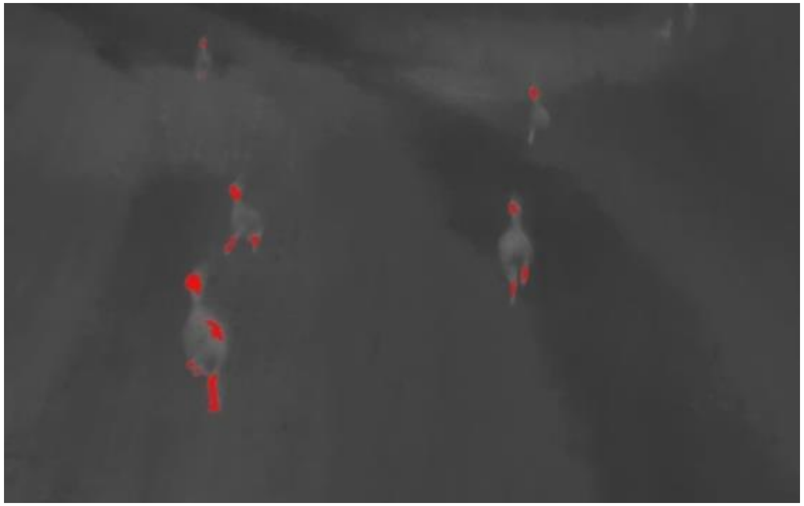


Thermal cameras

- Thermal camera challenge

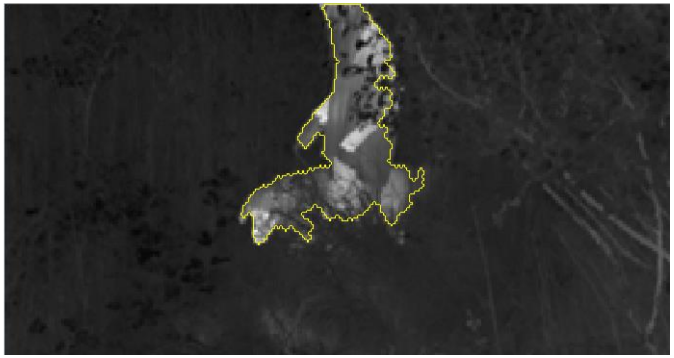
Light Background	Shadow	Sun
Warm (Crops)	Easy	***Hard***
Cold (Soil/rubble/stone/water)	Very easy	Easy





Thermography - segmentation

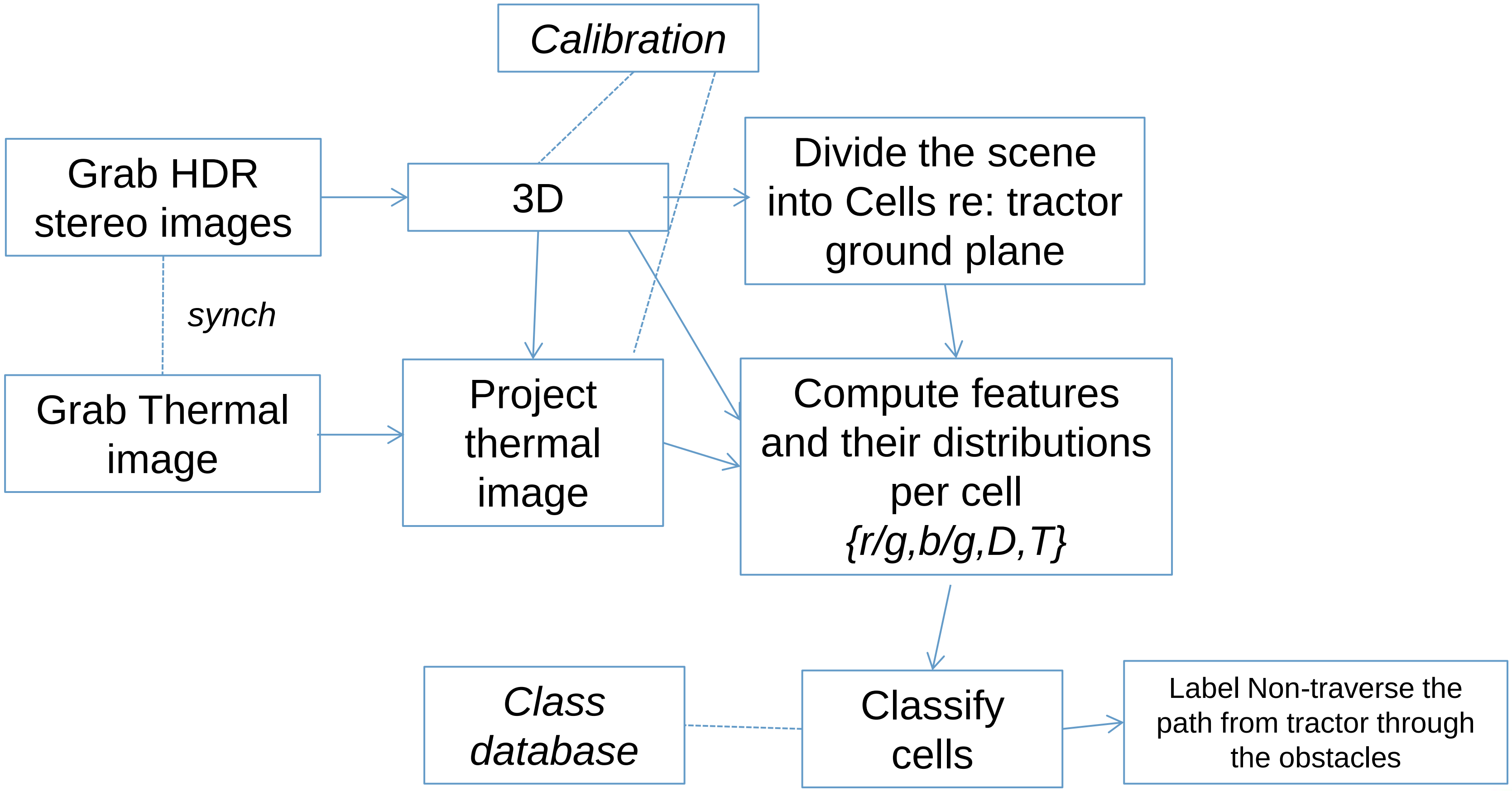
- Yellow outline marks obstacles



HDR Stereo + Thermography

- Instead of treating it as two separate classifiers we base the detection on a common feature vector
- A flexible trainable obstacle detector tested with two types of classifiers:
  - Accumulative histogram based classes
  - Gaussian Mixture models, where the classification is based on the scene GMM – i.e. the GMM for the screen is compared to a scene database
- The feature vector is based on the ***distributions*** of *normalized* colours, heights *from the ground*, and temperatures, where the colours are 2 features (C_1=r/g and C_2=b/g) and treated as one separate group and the height-temperature is another. Classes can then be defined using one or both groups with a weighting.
- This way to can teach the vehicle where it is OK to drive and where it is not, and most classes will only use height-temperature and special classes like crops that are tall can use the colour distributions as well.
- ***An efficient training scheme is to train traversable situations because these are predictable, whereas obstacles are not – those are the things that are not supposed to be there!***

HDR Stereo + Thermography



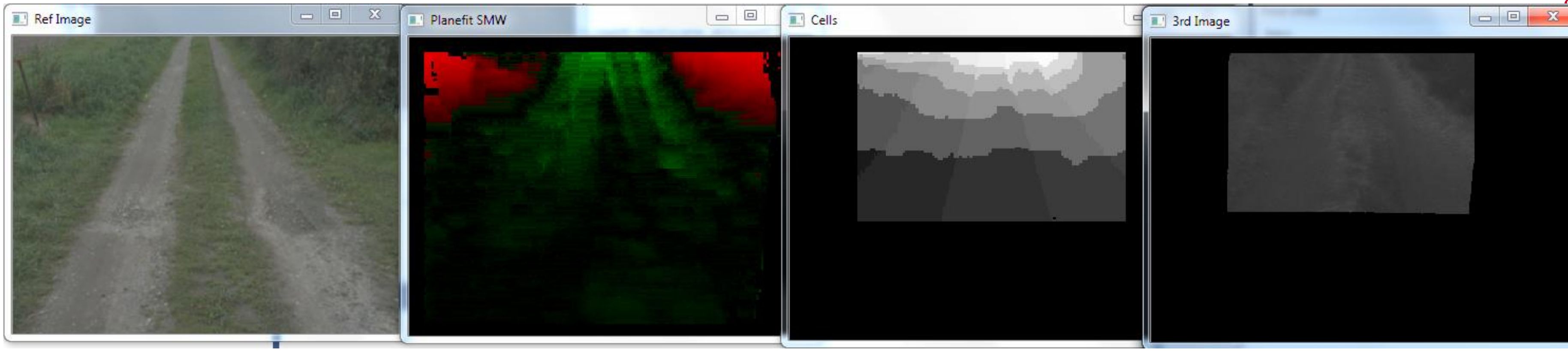
Concept

HDR images

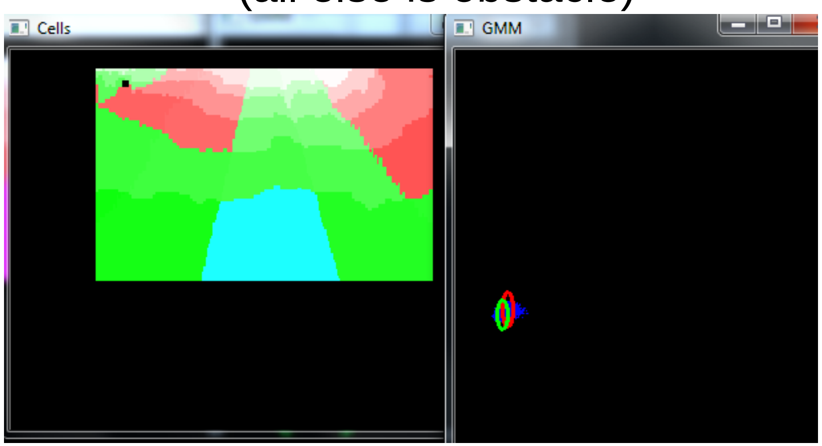
3D relation to the plane in front of the tractor

Divide into equally sized cells

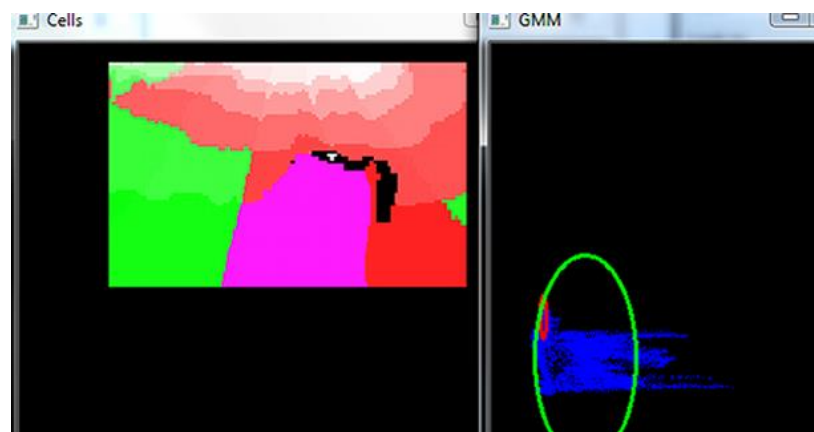
Project thermal image onto the point cloud



Teach – traversable situations (all else is obstacle)



Compare new situation and annotate traversable/non traversable

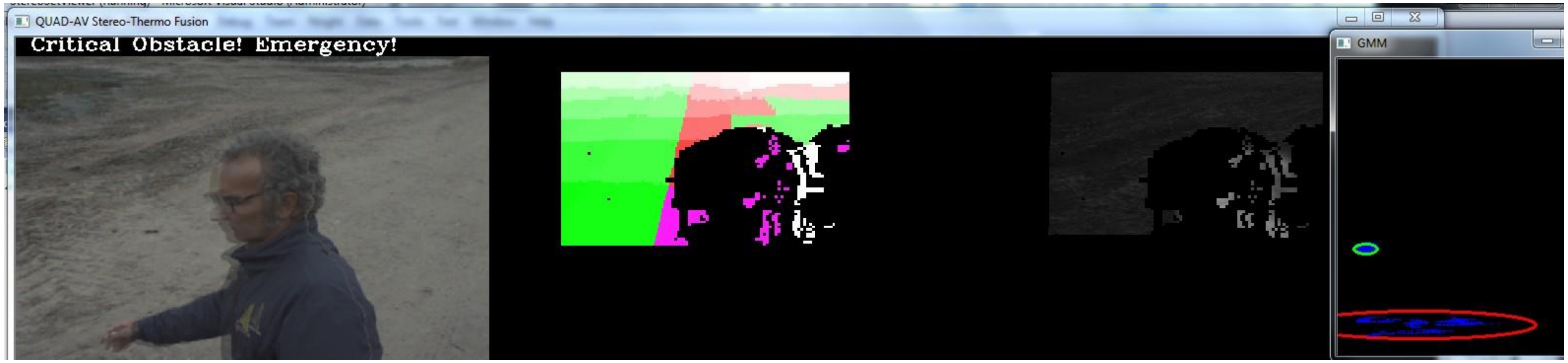


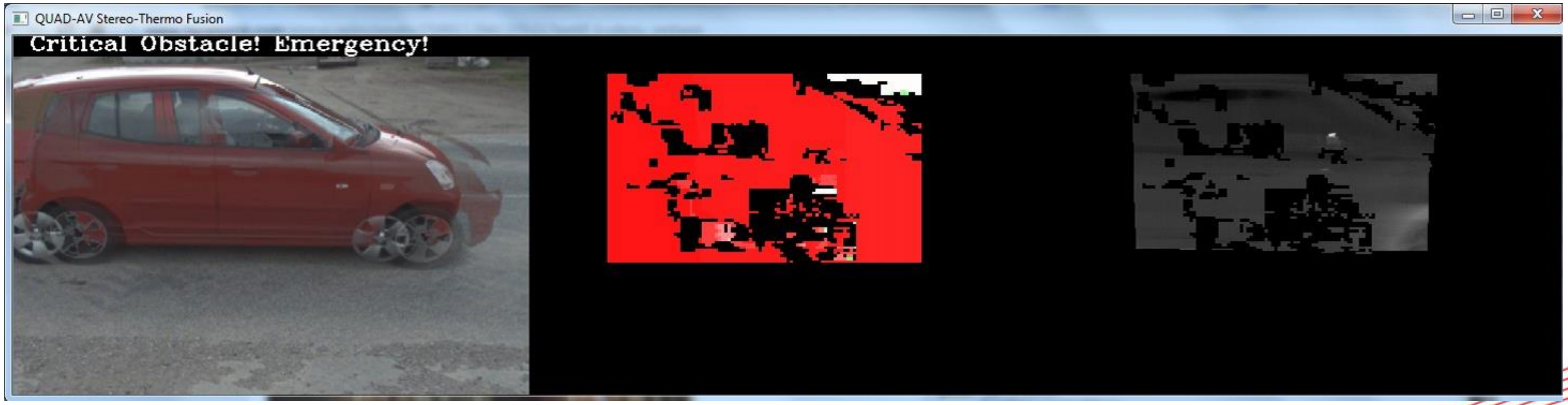
Results (RGB Image / Classified Cells / Thermal Image – smaller FOV)
Green: Traversable, Red: Not traversable, Purple/Cyan: Selected cell for showing the GMM)

DANISH TECHNOLOGICAL INSTITUTE

Centre for Robot Technology

Handles very close obstacles





Results (RGB Image / Classified Cells / Thermal Image – smaller FOV)
Green: Traversable, Red: Not traversable, Purple/Cyan: Selected cell for showing the GMM)

DANISH TECHNOLOGICAL INSTITUTE

Handles obstacles on flat ground as well as ramps and walls

Centre for Robot Technolog

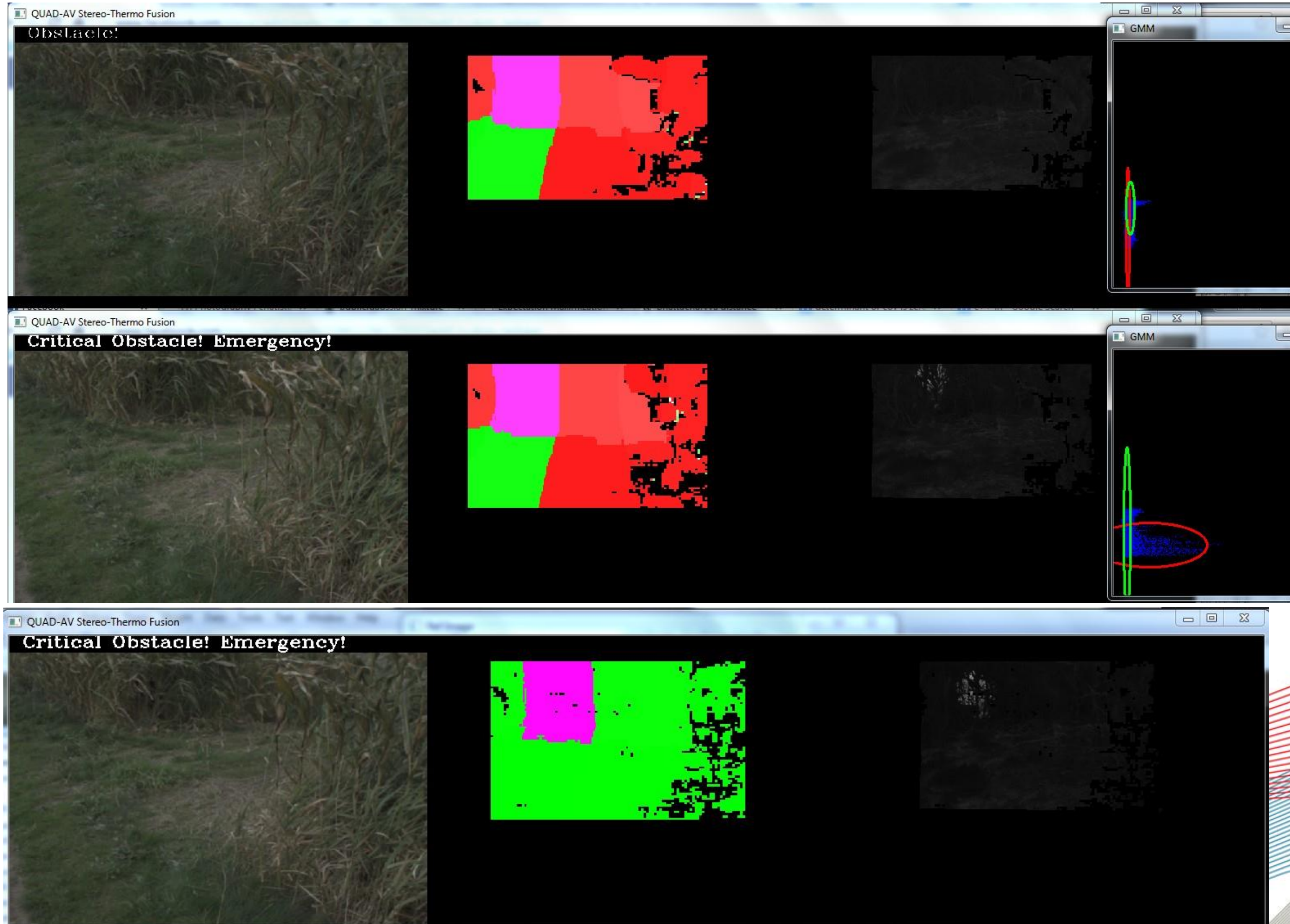
Results (RGB Image / Classified Cells / Thermal Image – smaller FOV)
Green: Traversable, Red: Not traversable, Purple/Cyan: Selected cell for showing the GMM)

DANISH TECHNOLOGICAL INSTITUTE

Centre for Robot Technology

Results (RGB Image / Classified Cells / Thermal Image – smaller FOV)
Green: Traversable, Red: Not traversable, Purple/Cyan: Selected cell for showing the GMM)

Can tell the difference between a wall of maize and a human hiding inside, so we can teach it that it is ok to drive through the corn but not the human

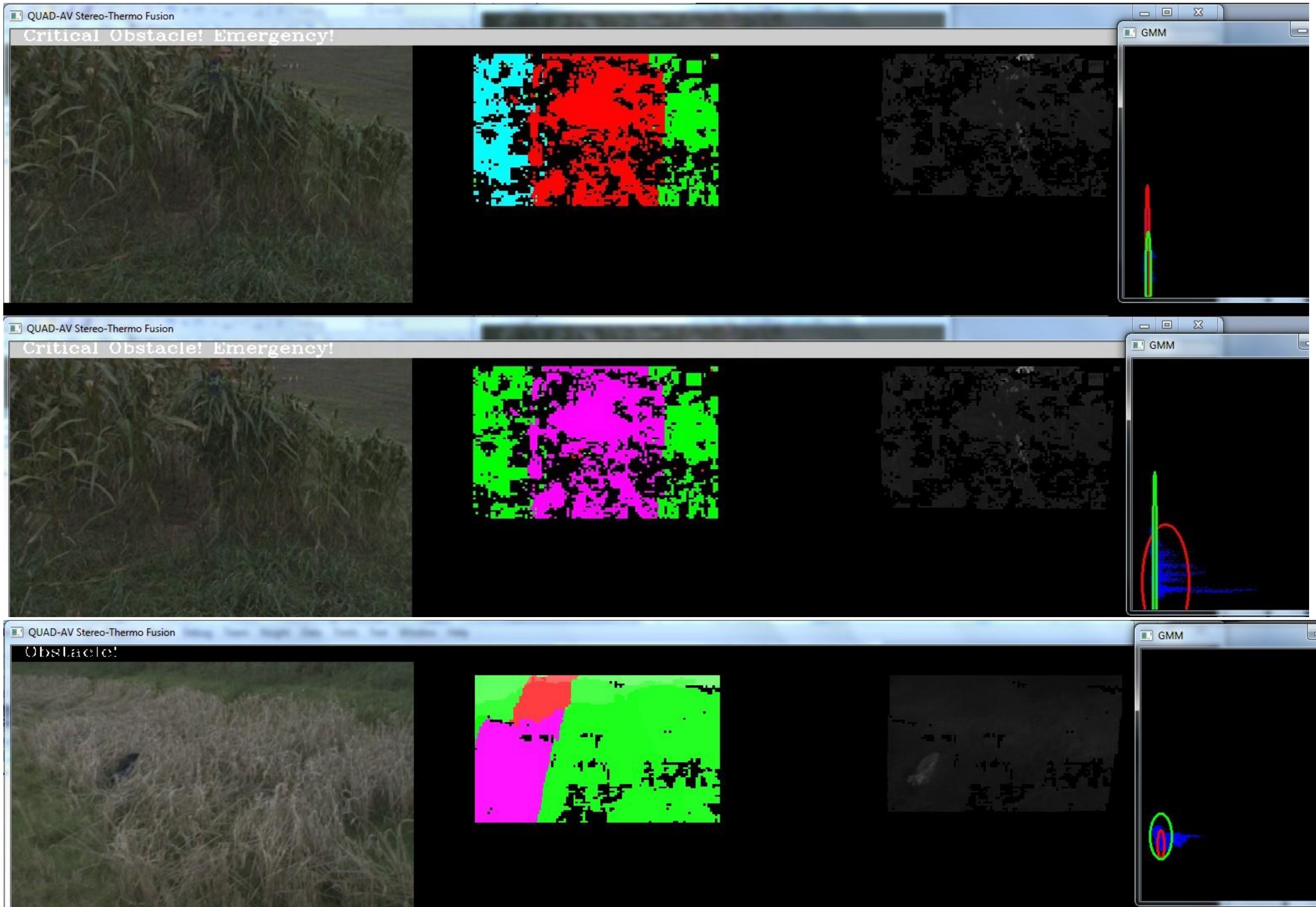


Results (RGB Image / Classified Cells / Thermal Image – smaller FOV)
Green: Traversable, Red: Not traversable, Purple/Cyan: Selected cell for showing the GMM)

DANISH TECHNOLOGICAL INSTITUTE

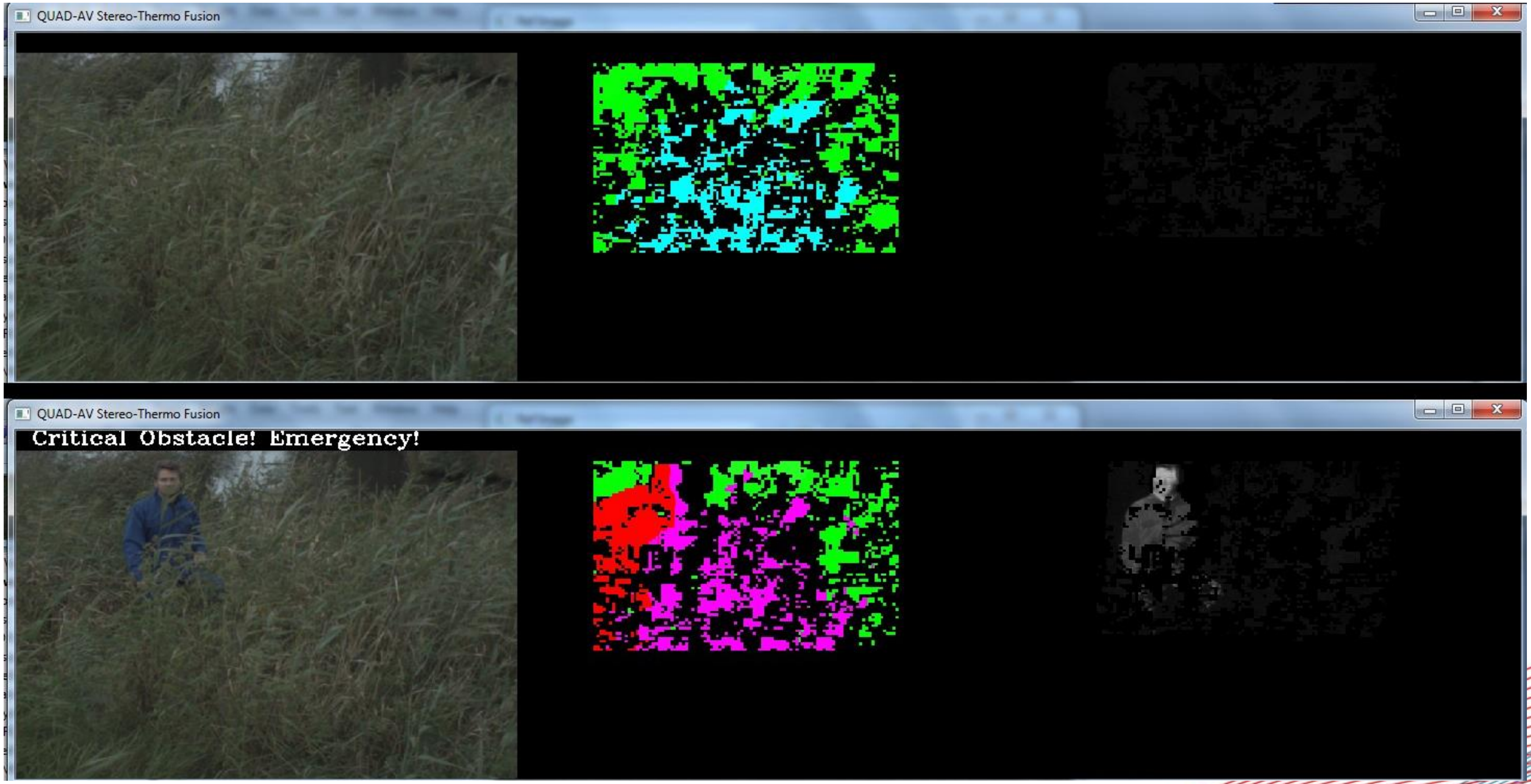
Centre for Robot Technology

Can learn that the tall crops are ok to drive through, except if a living obstacle is inside (cyan is traversable selection purple is obstacle selection)



Results (RGB Image / Classified Cells / Thermal Image – smaller FOV)
Green: Traversable, Red: Not traversable, Purple/Cyan: Selected cell for showing the GMM)

It can learn that the tall crops are ok to drive through, except if a living obstacle is inside (cyan is traversable selection purple is obstacle selection)

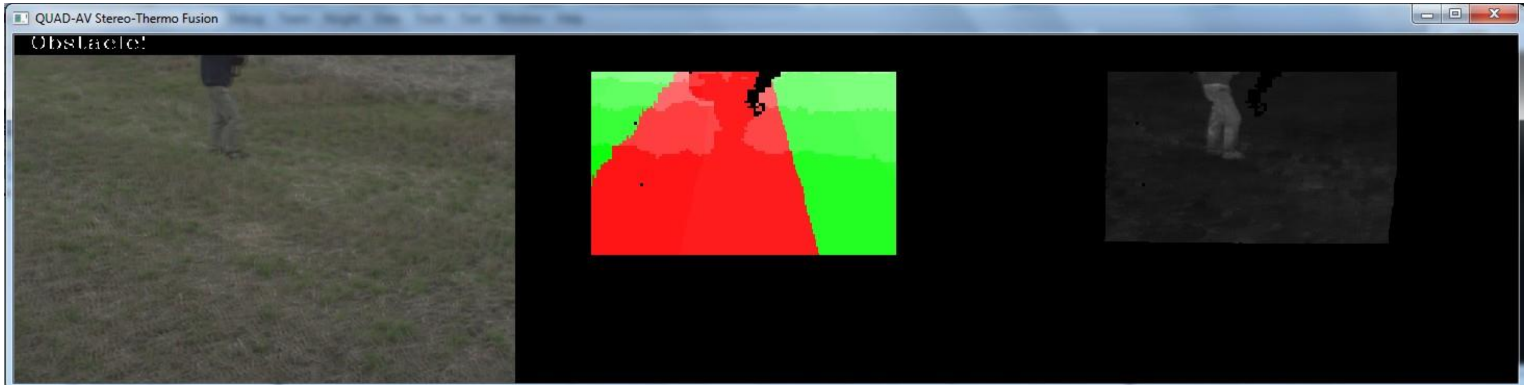


(never mind the lake behind it, haha!)

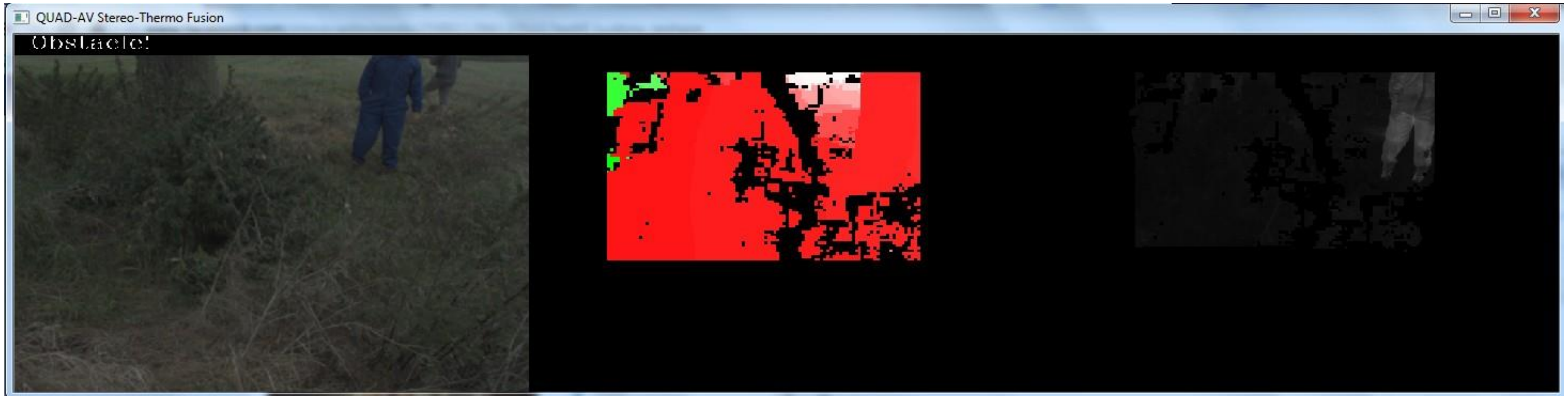
Results (RGB Image / Classified Cells / Thermal Image – smaller FOV)
Green: Traversible, Red: Not traversible, Purple/Cyan: Selected cell for showing the GMM)

DANISH TECHNOLOGICAL INSTITUTE

Centre for Robot Technology



Results (RGB Image / Classified Cells / Thermal Image – smaller FOV)
Green: Traversable, Red: Not traversable, Purple/Cyan: Selected cell for showing the GMM)



LIDAR and Stereo 3DImagery Integration

LIDAR and vision systems are complementary in many ways:

- Scanning LIDAR data are very accurate, yet sparse.
- LIDAR performs better in low-lighting conditions.
- However, scanning LIDARs typically provide low sampling rates (0.1-1Hz) resulting in difficulties to capture dynamic obstacles.
- In addition, scanning LIDARs sytems may feature limited sensing range according to their particular configuration, preventing the robot to perform reliable long-range navigation.
- Vision provides high-rate, dense, but less accurate measurements.

⇩

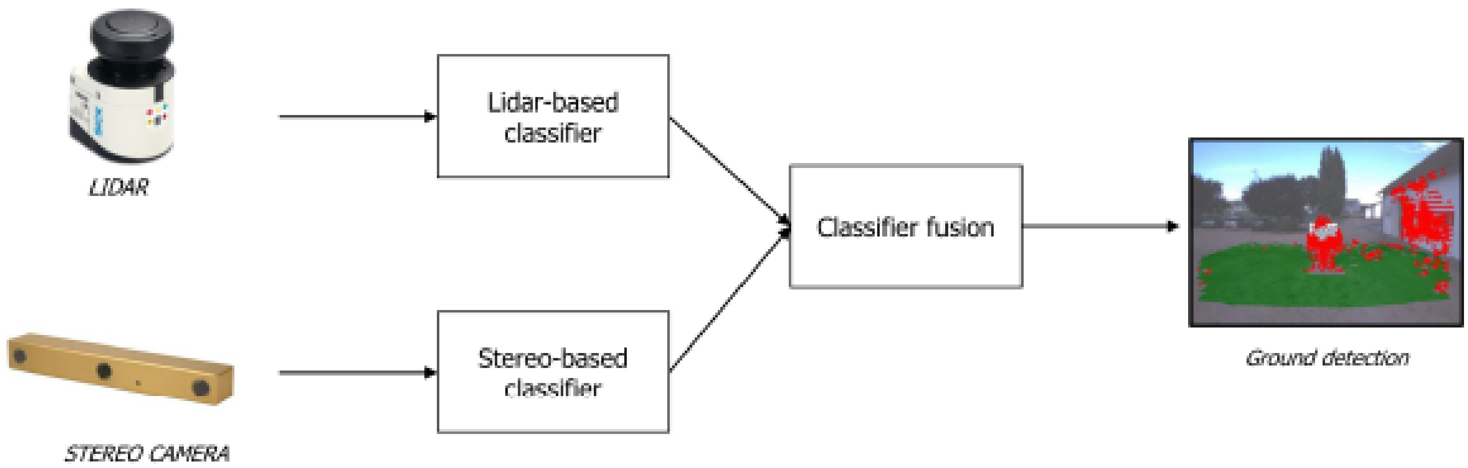
LIDAR and Stereo Combination

DANISH TECHNOLOGICAL INSTITUTE

Centre for Robot Technology

Statistical framework for multisensory perception

Input data from LIDAR and stereo camera are used by different classifiers,whose results are fused to produce a unique classification output.



Single sensor geometry-based Classifier

- A **multivariate Gaussian distribution** is assumed to model the **ground class** *M*(μ, Σ).
- In every newly-acquired scan, a given patch is classified as ground or non-ground based on its Mahalanobis distance, following an outlier detection approach

$$d^2 = (f-\mu)\Sigma^{-1}(f-\mu)^T$$

$d^2 > \chi^2_{m;\alpha} \rightarrow$ Non-Ground

$d^2 \leq \chi^2_{m;\alpha} \rightarrow$ Ground

CLASSIFIER

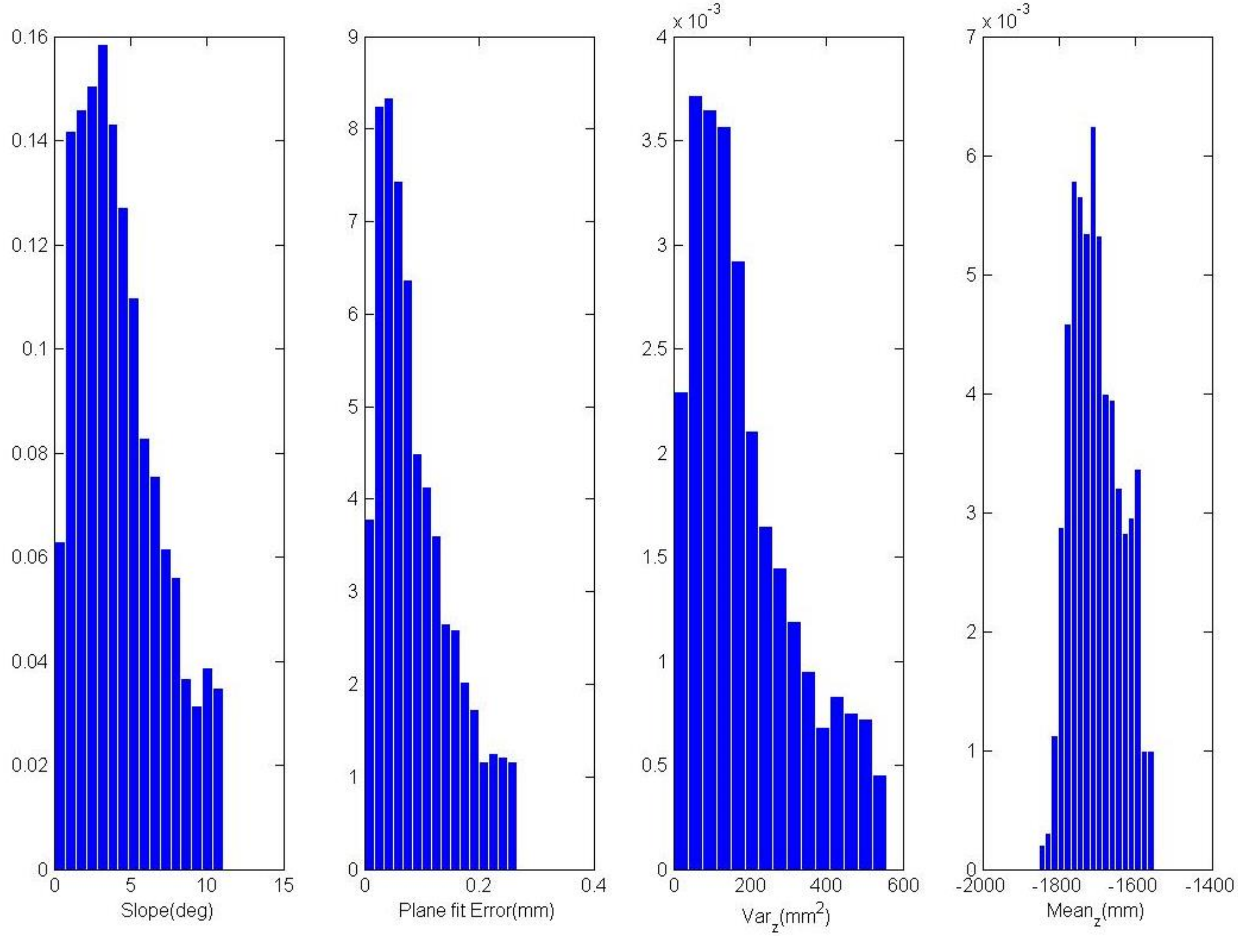
Geometric features

- Geometric features are statistics calculated from the elevation points associated with each terrain patch
- *Average slope of terrain*: the angle between the least-squares-fit plane and the horizontal plane.
- *Goodness of fit*: the mean-squared deviation of the points from the fitted plane along its normal.
- *Variance* in the heights (Z coordinate) of the range points.
- *Mean* of the heights of the range points.

Histograms of the distribution

DANISH TECHNOLOGICAL INSTITUTE

Centre for Robot Technology



All histograms exhibit an approximately **unimodal** distribution, which can be reasonably modelled with a Gaussian.

Single sensor geometry-based Classifier cont’ed

Objective:

- A *self-learning* classification system based on a rolling training set to:
  - construct the geometric model of the ground
    - ✓ automatically
    - ✓ online
  - segment scene into *ground* and *non ground* regions.

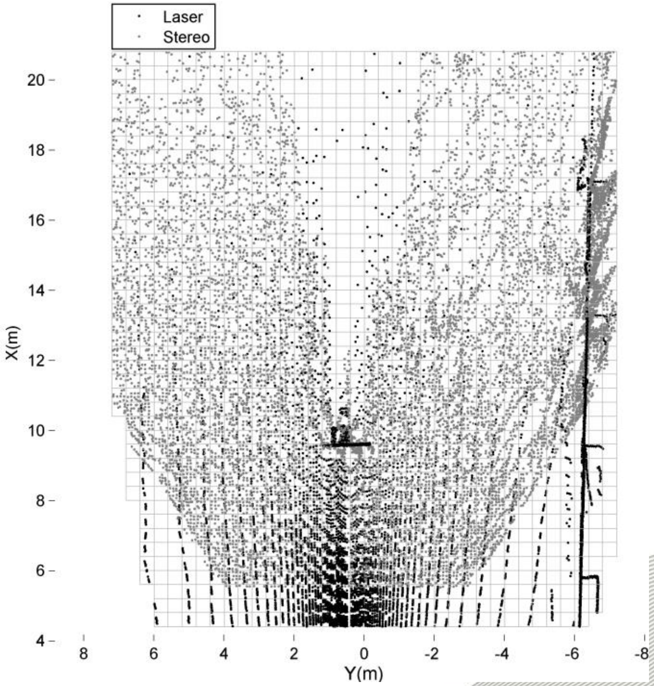
- The training set is automatically initialized at the beginning of the robot's operation via a bootstrapping approach and progressively updated.
- It is feasible to use the system for long range and long duration navigation, over changing environments.

Classifier Fusion

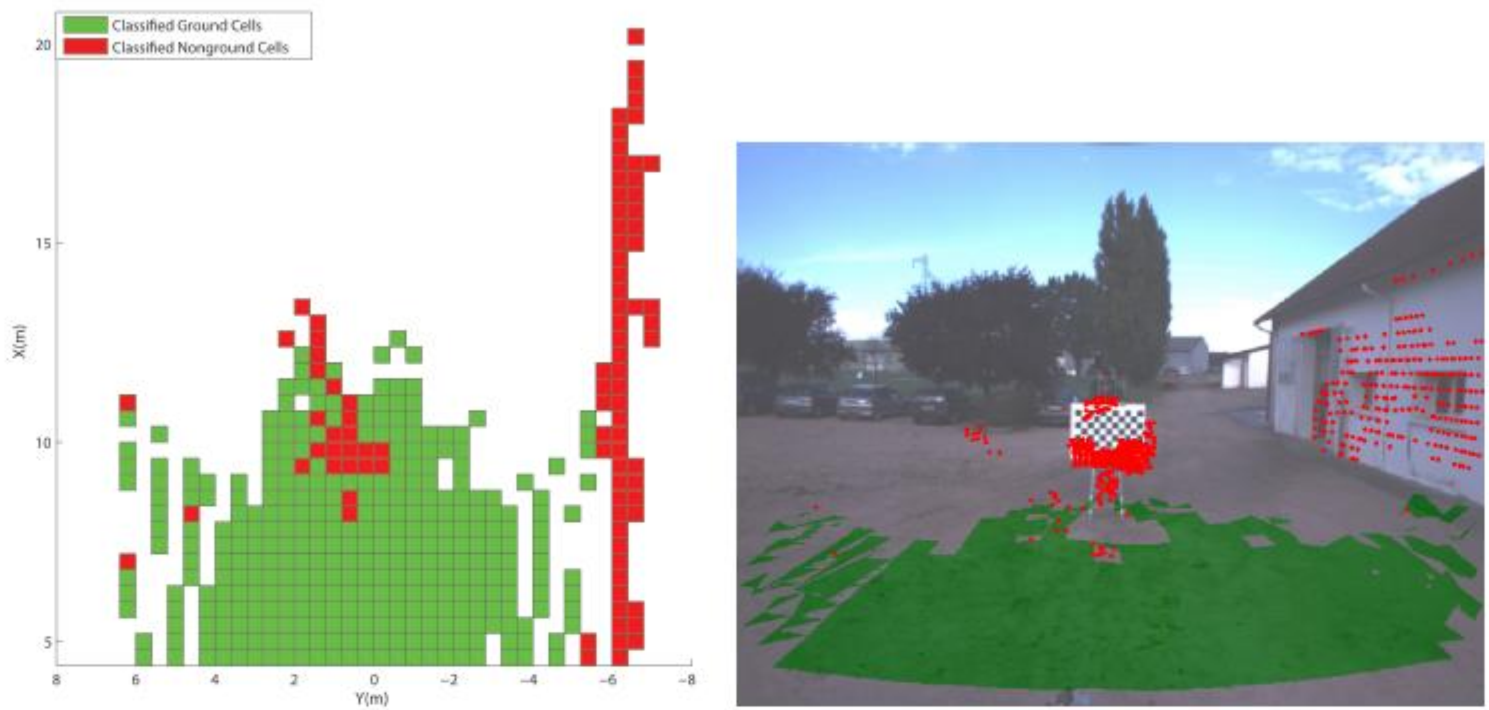
Objective:

- The single-sensor ground classifiers can be combined. Thus, one can exploit their individual advantages in order to reach an overall better performance than could be achieved by using each of them separately.
- One way is to weight the individual output obtained from the classifiers with their prior probabilities that can be statistically quantified using ground-truth data to obtain a unique classification results.

(a) Original visual image. (b) Reference grid divided into 0.4-m by 0.4-m cells. Points reconstructed by laser scanner are denoted in black. Points reconstructed by stereo vision are marked in grey.

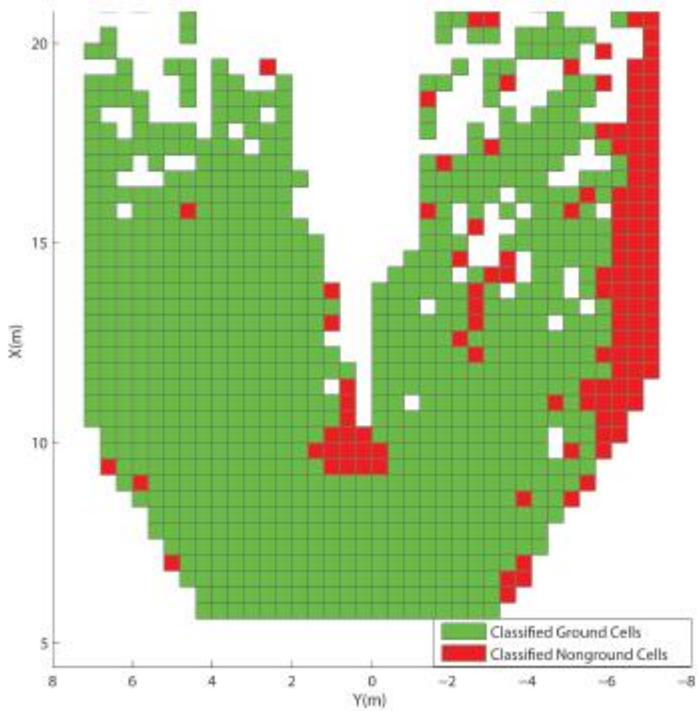


(a) Traversability map obtained by the LIDAR-based ground classifier. (b) Results projected over the original image. Pixels associated with ground- (non-ground-) labeled cells are marked using green (red).

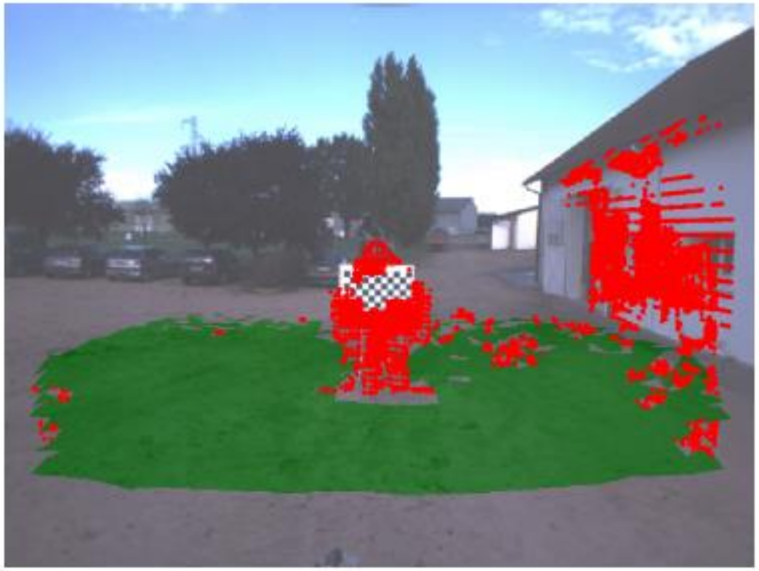


(a) (b)

(a) Traversability map obtained by the stereo-based ground classifier. (b) Results projected over the original image. Pixels associated with ground- (non-ground-) labeled cells are marked using green (red).



(a)



(b)

Traversability map obtained by the combined LIDAR-stereosystem

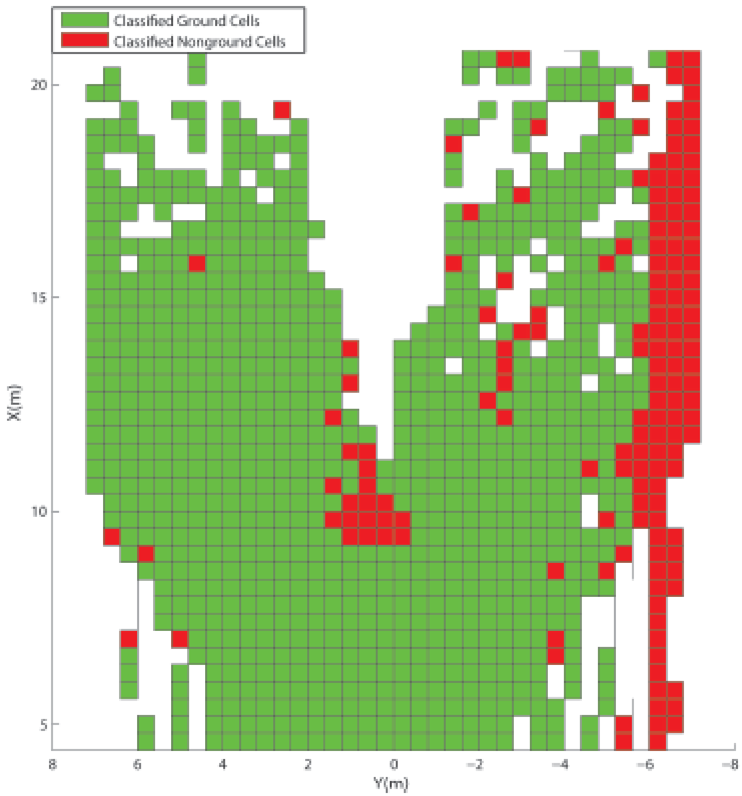


Table 3: Classification results obtained from the cells labeled by both single-sensor classifiers.

	LIDAR-based	Stereo-based	Combined
Precision	99.1%	96.5%	99.6%
Rejection Precision	87.5%	96.1%	98.3%

Conclusions

- Histogram training is very fast and can be implemented to be very fast on FPGA
  - Comparison is done using the intersection of histograms, which is very fast, even in a 4D data space.
  - But it requires much more training as it does not generalize well
  - Execution speed does not increase with training, because we are caring out distributions that are considered traversable in the same data space
- GMM classification was easier to train but are slower
  - The more classes you train, the slower it gets, but luckily fewer are needed with GMM.
  - A fast GMM E-M algorithm needs to be made in hardware. We used OpenMP to speed it up (using the GPU already for the 3D)
  - GMM clusters are compared using Bhattacharyya distance.
  - It tests different permutations of the clusters and finds the combination giving the minimum distance
- Training can be done by showing the vehicle where it can drive so it learns all the different environments on its path that are traversable, while a driver is driving it.
  - It can be born with certain knowledge about certain obstacles.
  - Certain classes can be assigned higher priorities

DANISH TECHNOLOGICAL INSTITUTE

Centre for Robot Technology

Contact Info

- Michael Nielsen
- Danish Technological Institute
- Phone: +45 72202984
- Email: mlnn@dti.dk


AXION