PAN ROBOTS
Plug&Navigate robots for smart factories

FP7-ICT-314193, grant 314193/2012

3D Omnidirectional Stereo Vision Based Perception

Sergiu Nedevschi
Technical University of Cluj-Napoca, Romania

SICK Sensor Intelligence. | ELETTRIC80 | UNIMORE Università degli Studi di Modena e Reggio Emilia | VTT | Universitatea Tehnică din Cluj-Napoca | IberianPartners Coca-Cola Iberian Partners, S.A.

TUCN Team

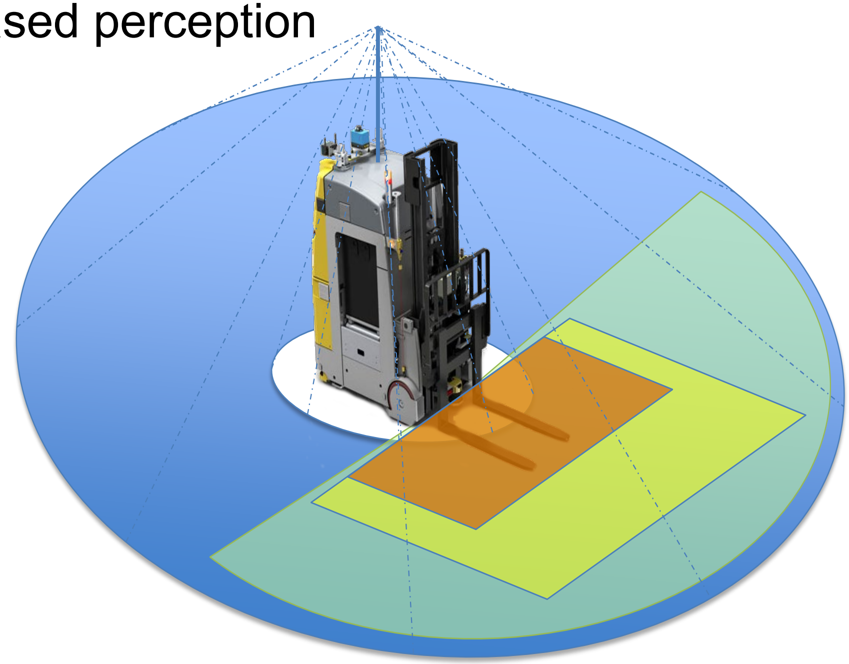
Coordinator: Sergiu Nedevschi

Team:
Marita Tiberiu
Andrei Vatavu
Arthur Costea
Robert Varga
Istvan Szakats
Aliz-Eva Nagy

On-board Perception
PAN-Robots Solution

Objectives

- On-board multi-sensor environment perception:
  - 2D safety and perception, enhanced by
  - 3D omnidirectional stereo vision based perception



3D Omnidirectional Stereo Vision

Specific Objectives

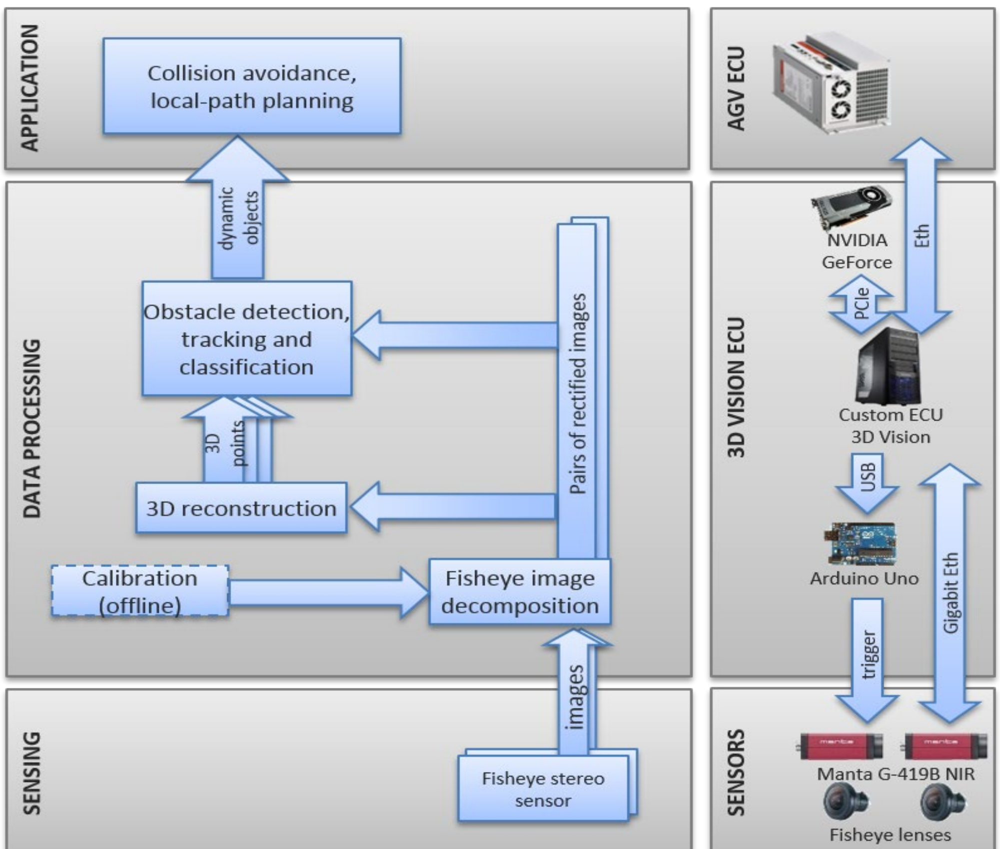
- To develop and implement a 3D sensing mechanism for AGV surrounding perception
- 360^{o} horizontal FOV, 150^{o} vertical FOV
- Detect, track and classify objects
- Enhance the 2D safety around the AGV

Technical Requirements

- Object detection, tracking and classification
  - Position/Dimension error: < 3 % of object distance
  - Velocity accuracy: ±0.5 m/s at 1 σ
  - Object classification rate: > 90%
- Area to be processed
  - 11 m in front and 4 m in lateral direction (starting from the longitudinal axis of the AGV), and a height of 4.5 m
- Data processing rate
  - >=10 Hz

3D Omnidirectional Stereo Vision

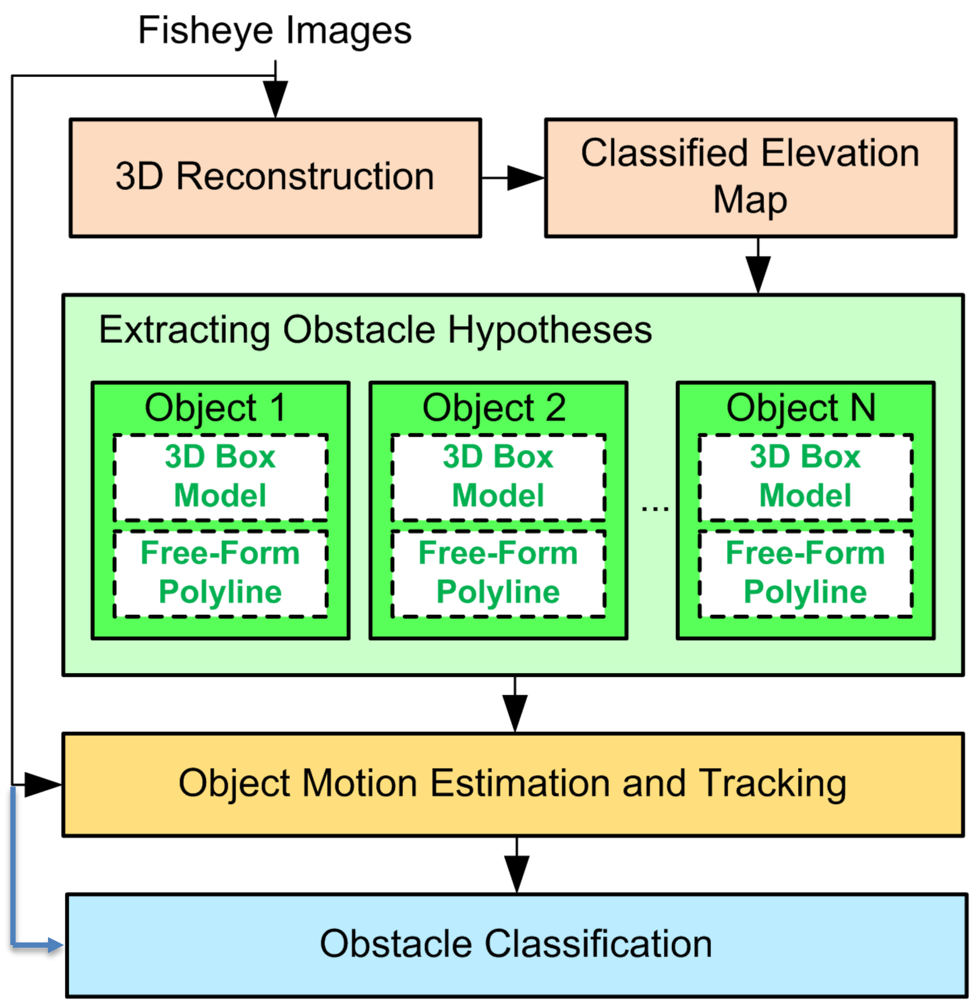
System Overview



3D Omnidirectional Stereo Vision

Main tasks

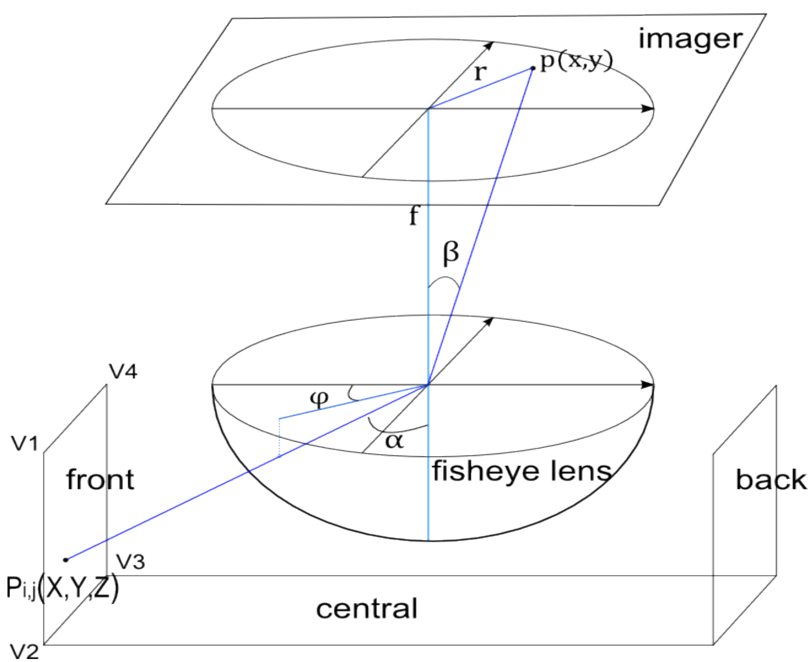
- 3D Reconstruction
- DEM
- Extracting Object Hypotheses
- Object Classification
- Object Tracking

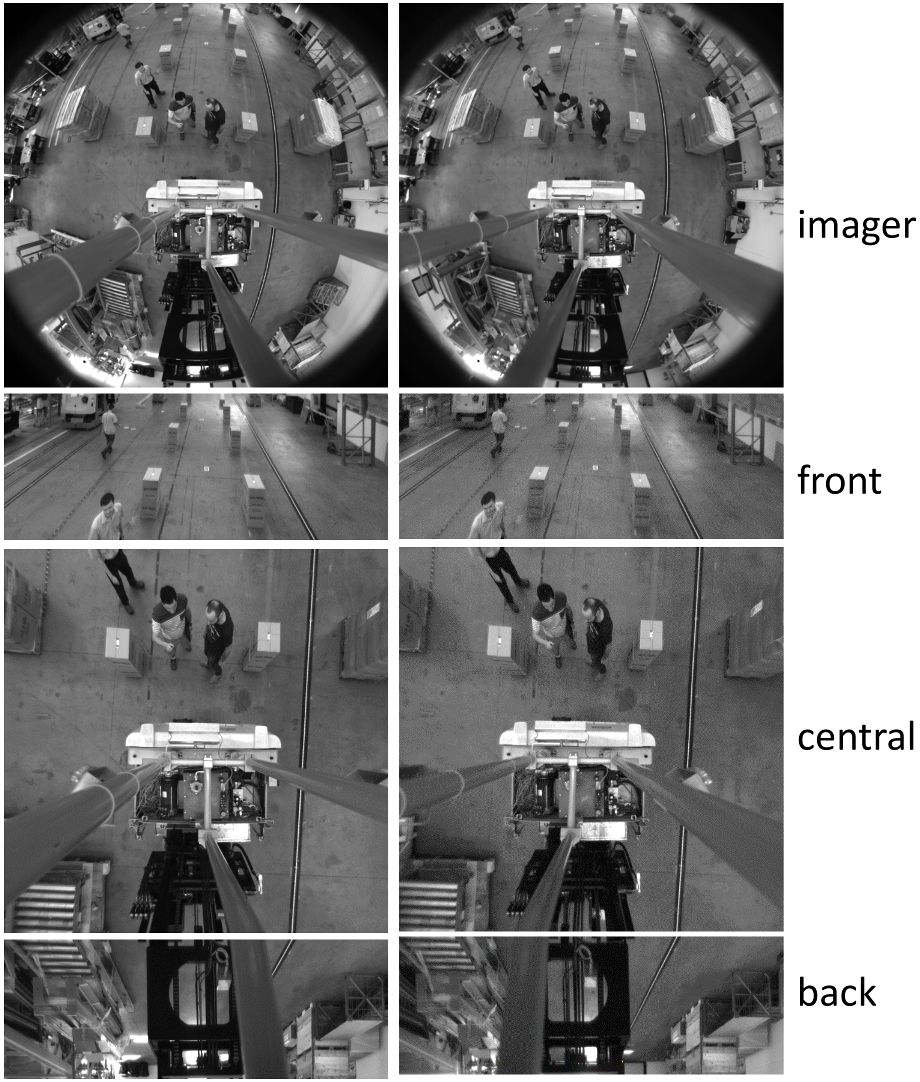


Processing Flow

3D Omnidirectional Stereo Vision 3D Reconstruction

Multi-Channel Rectification

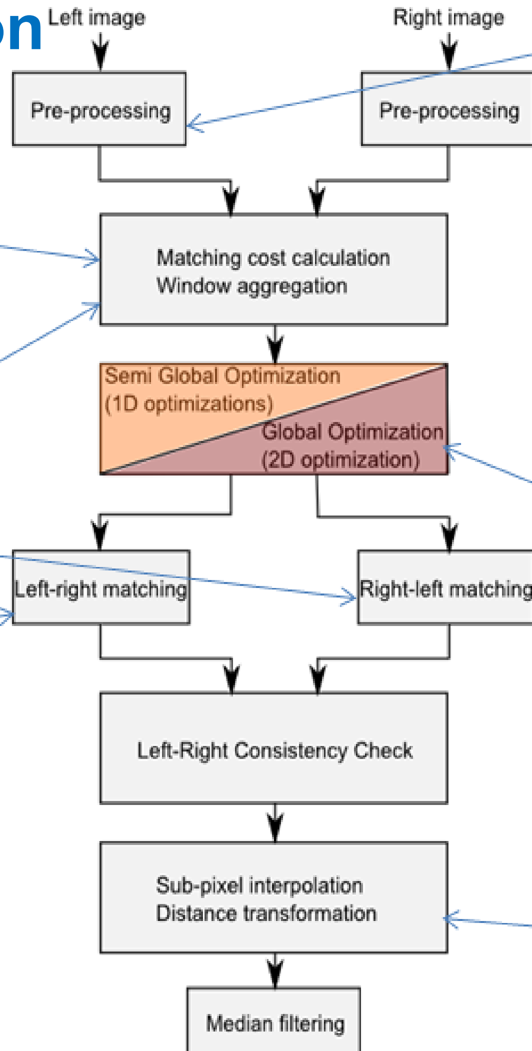




Plug&Navigate robots for smart factories

3D Omnidirectional Stereo Vision 3D Reconstruction

Stereo Reconstruction



Rectification, Parametric filtering (Gaussian, Laplace, LOG), **Non-parametric transforms** (Rank, Census)

SAD, SSD, NCC, Hamming Distance

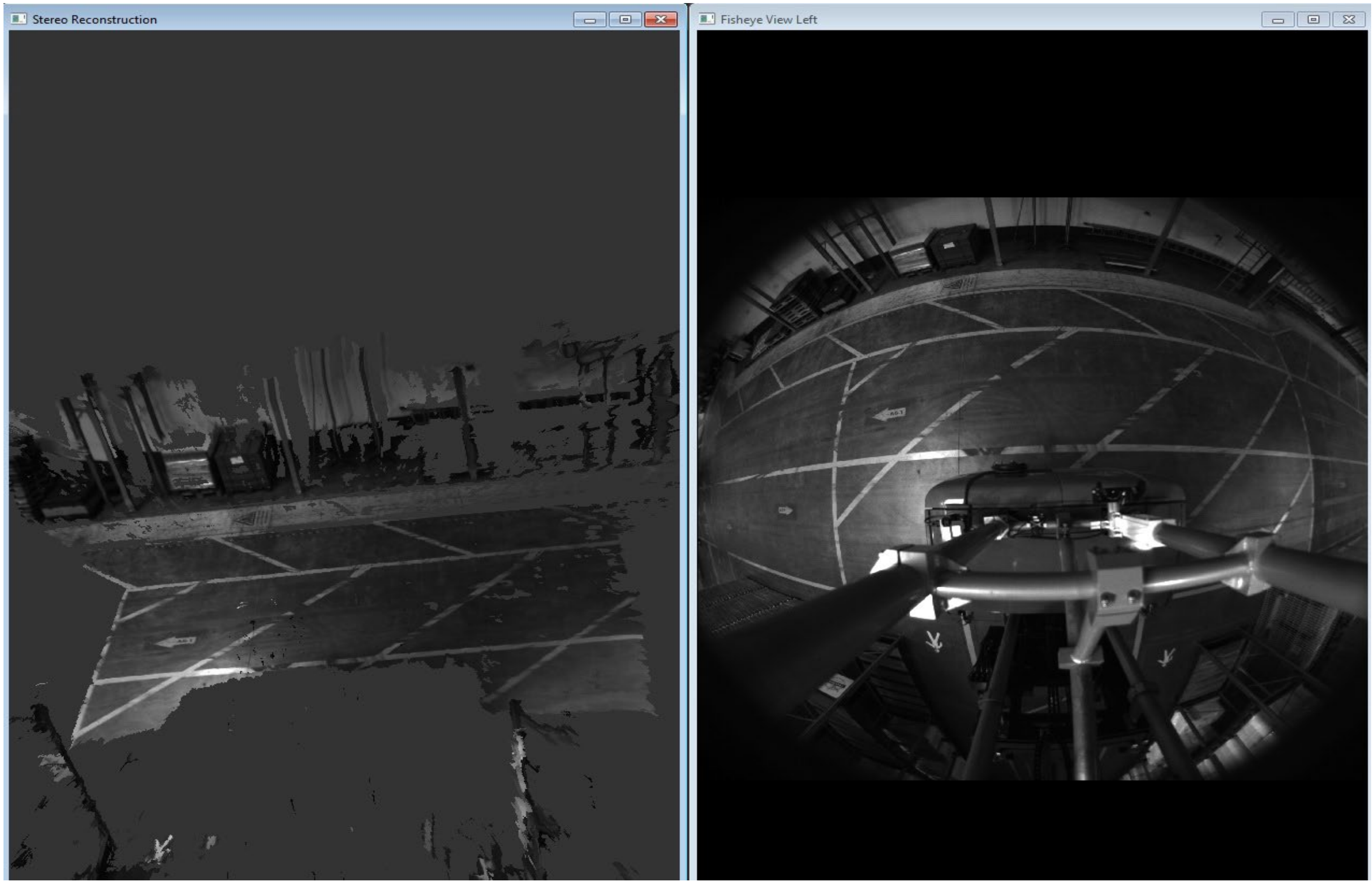
Single window, Multiple windows

Locale Methods: no optimization
Global Methods: 2D optimizations (simulated annealing, graph-cut, dynamic programming, etc.)
Semi-global methods: multiple 1D optimizations (4, 8, 16)

Winner-takes-all, Confidence factor based filtering

Sub-pixel level interpolation: parabolic, symmetric V, histogram based, etc.

3D Omnidirectional Stereo Vision 3D Reconstruction



Plug&Navigate robots for smart factories

Digital Elevation Map

- An accurate probabilistic inverse sensor model of the stereovision sensor based DEM creation algorithm was proposed

$$P(c_{i,j} \mid z) = \alpha \cdot P(z \mid c_{i,j}) \cdot P(c_{i,j})$$

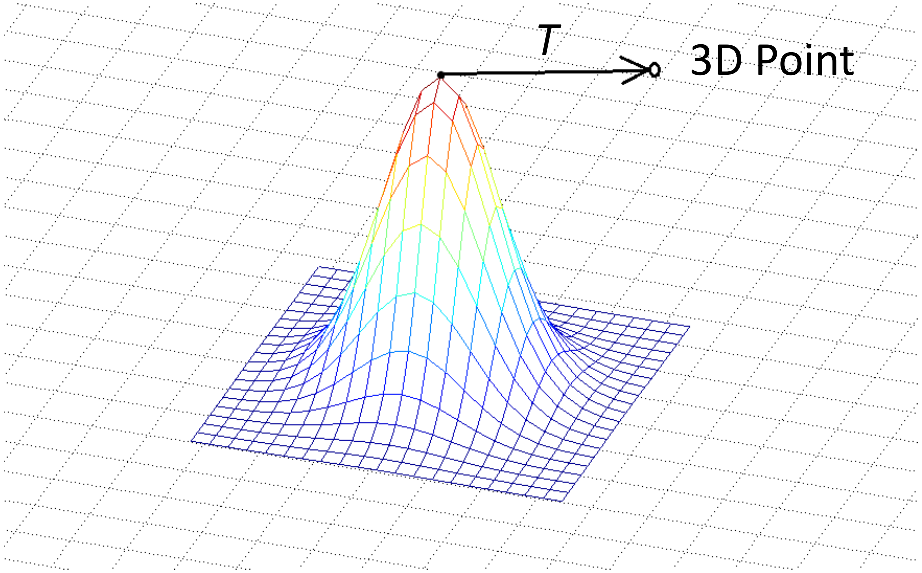
- More generally the direct sensor model is proposed as:

$$P(z_t \mid x_t, m) = \begin{cases} N(z_t^*, z_t, \sigma); z_t \in ROI \\ 0 \end{cases}$$

z_t^* - true position
z_t - the measured position
x_t - position of the robot (sensor)
σ - parameters of the normal distribution
m – known position of objects in the environment

Digital Elevation Map

- 2 components of the direct sensor model are used:
  - Systematic errors: the translation vector *T* from the measurement to the true position
  - Gaussian noise around the true measurements

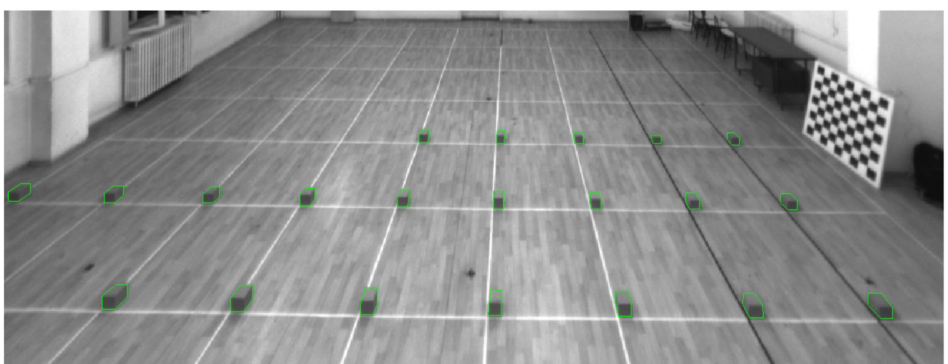


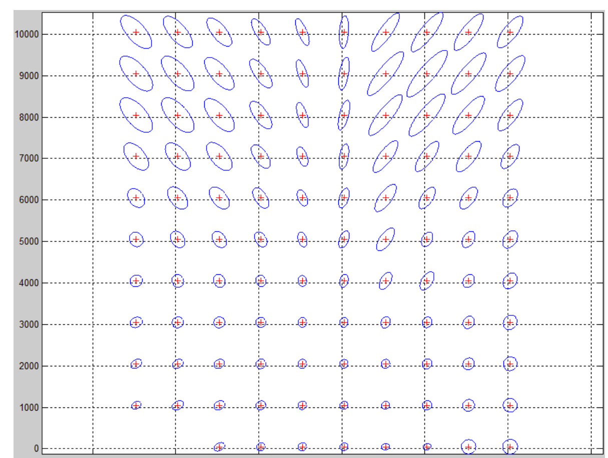
3D Omnidirectional Stereo Vision

Digital Elevation Map

Sensor Models

- Algorithm for finding the direct sensor model:
  - Mark the objects with a polygon in the rectified images
  - Collect all the 3D points reconstructed inside the polygon for several frames
  - Calculate the mean, standard deviation, covariance and the vector from the calculated mean to the ground truth position corresponding to each template object
  - Filter the results
  - Perform a trilinear interpolation in order to generalize the model to the whole ROI





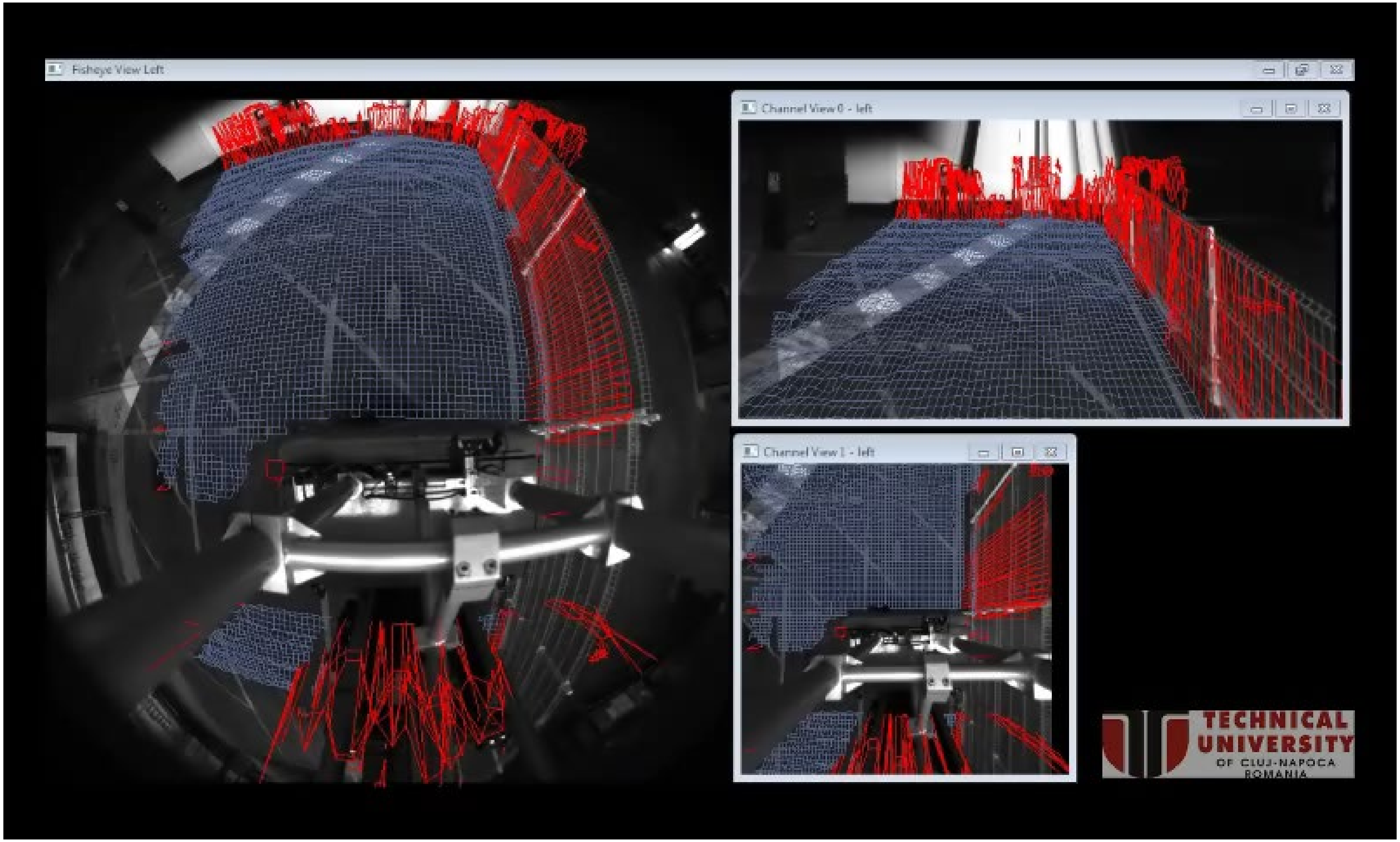
Uncertainty Based Digital Elevation Map Creation

- Create a histogram of heights for each cell with a probabilistic contribution of points based on the invers sensor model
- Filter the histogram (Gaussian filtering)
- Take the consistent height which is closest to the biggest elevation.

Digital Elevation Map Classification

- Plane fitting using the elevation map
- The 3D points lying in a range of +/-30 cm w.r.t. the expected height of the ground floor are fitted to planar surface using a RANSAC estimator
- The 3D points and by consequence the DEM cells are discriminated and labeled in road and obstacle points and cells based on the estimated surface.

3D Omnidirectional Stereo Vision
Uncertainty based DEM

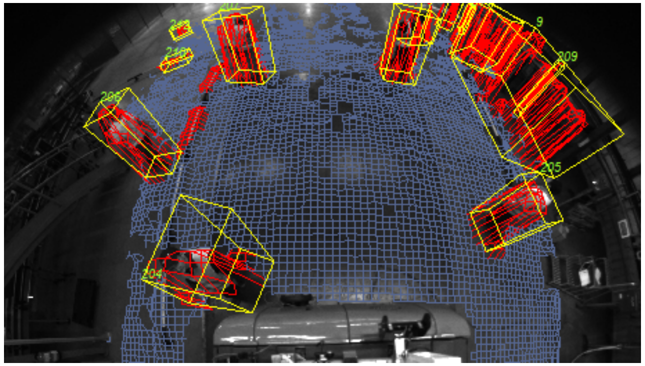


3D Omnidirectional Stereo Vision
Object Hypotheses

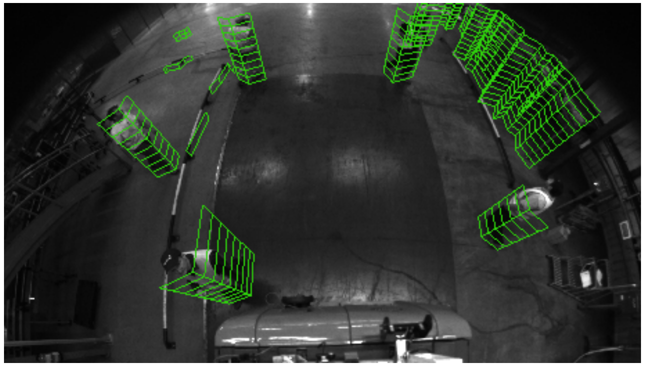
Extracting Object Hypotheses

- The object hypotheses are extracted by clustering the DEM cells into connected entities – blobs.

- Each object hypotheses is defined by two separate models:
  - **-A 3D cuboid** – used to select the ROI for **obstacle classification**.
  - **-A free-form polygonal representation** – used to extract the **object motion** by applying a fast ICP-based alignment solution.



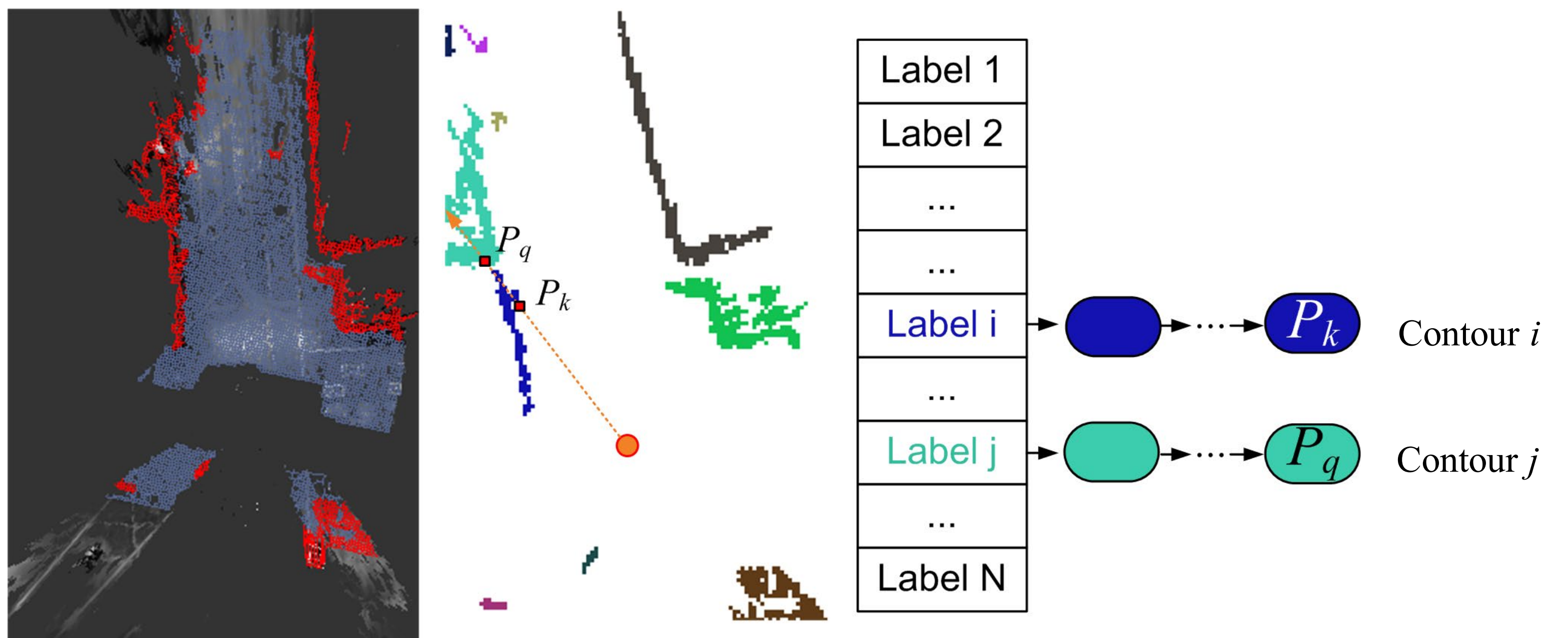
3D Cuboids



Free-Form Polylines

Multiple Depths Border Scanner Algorithm

- Current solution:
  - Is able to collect obstacle contour points situated at different depths along the same ray (one point per obstacle, multiple points per ray).



The classified DEM (top view)

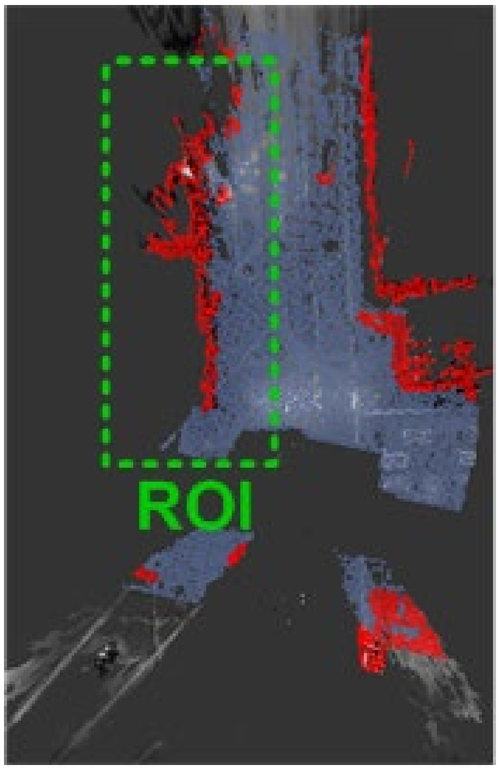
DEM Cells grouped into individual clusters (blobs)

The extracted object contours

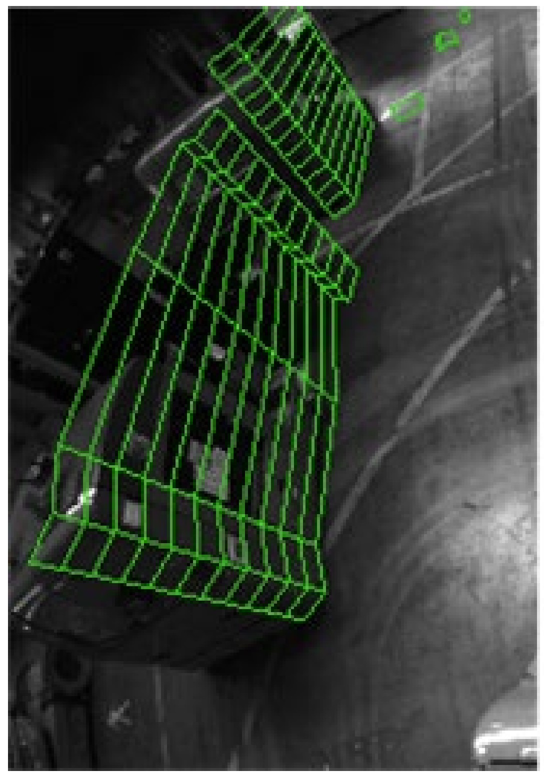
3D Omnidirectional Stereo Vision

Object Hypotheses

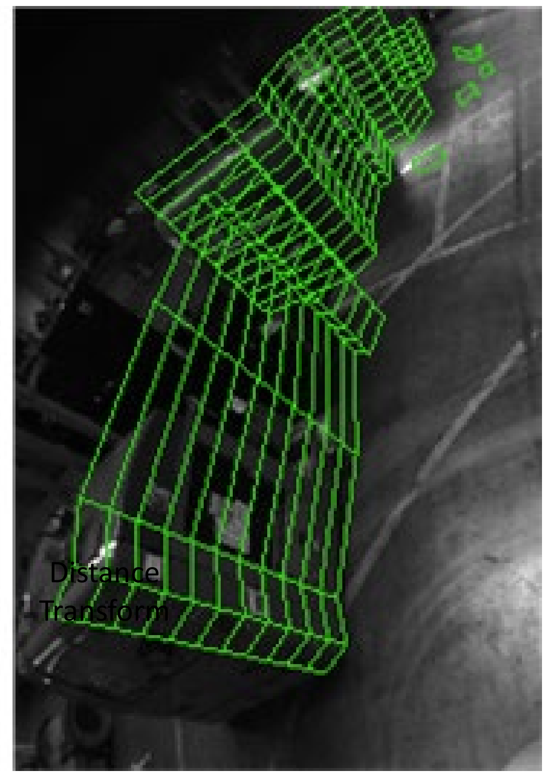
Border Scanner Algorithm



The classified DEM
(top view)



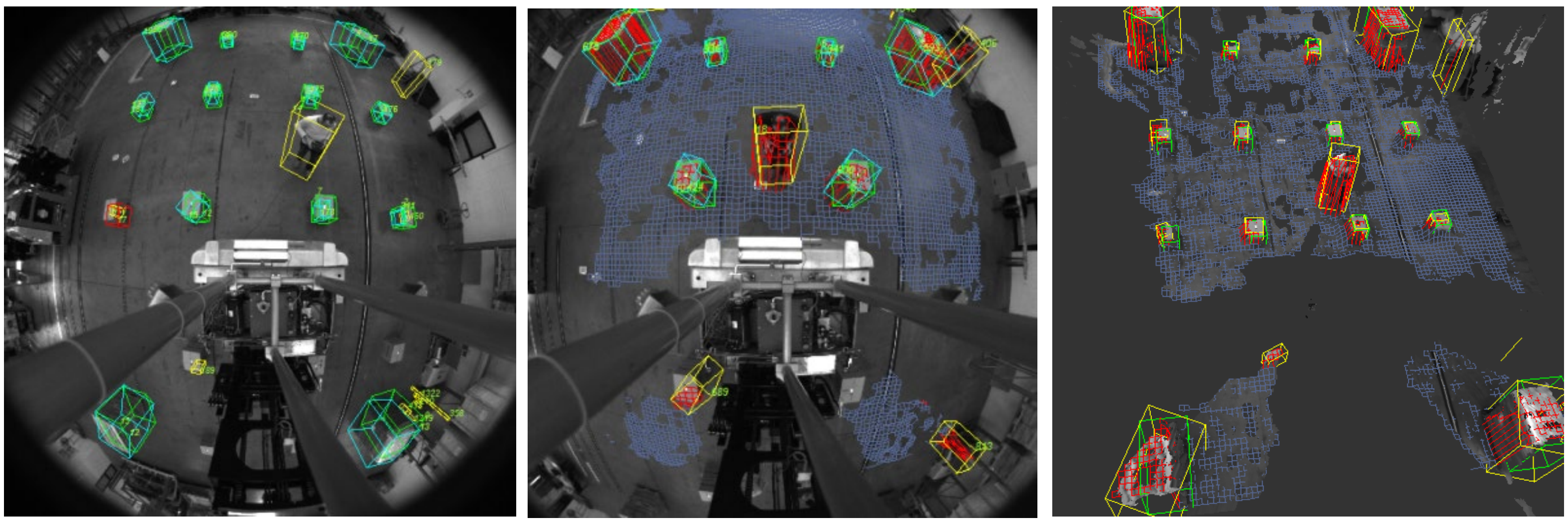
Previous Approach



Multiple Depth Border
Scanner
(Current Solution)

Object Detection Evaluation

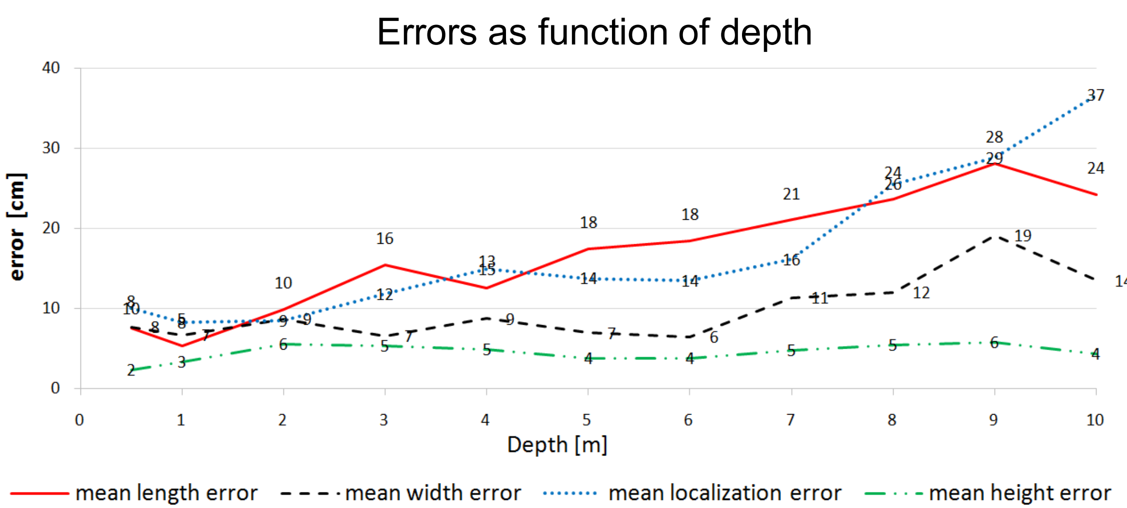
Evaluation based on objects with known position and size



Evaluation Metrics	Result
Detection Rate	92.8 %
Frames Number	456
Total Objects	4595
Missing Objects	331

Object Detection Evaluation

Evaluation based on objects with known position and size



Evaluation Metrics	Result
Mean Localization Error	0.156 m
Mean Width Error	0.10 m
Mean Height Error	0.054 m
Mean Length Error	0.167 m
Dist Err [%] at 1 σ	2.53 %
Width Err at 1 σ	0.09 m
Length Err at 1 σ	0.11 m
Height Err at 1 σ	0.04 m

Kalman Filter Based Solution

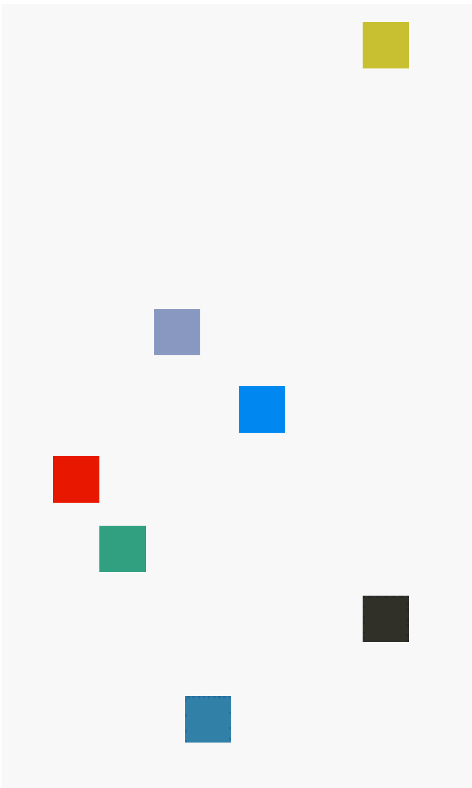
- Estimate, recursively in time, the state of the detected obstacles around the AGV given all measurements up to the current time *t.*
- The obstacle state, at a time *t*:

$$X_t = [\underbrace{x_t, z_t}_{\text{Position}}, \underbrace{vx_t, vz_t}_{\text{Speed}}, \underbrace{W_t, L_t, H_t}_{\text{Size}}]$$

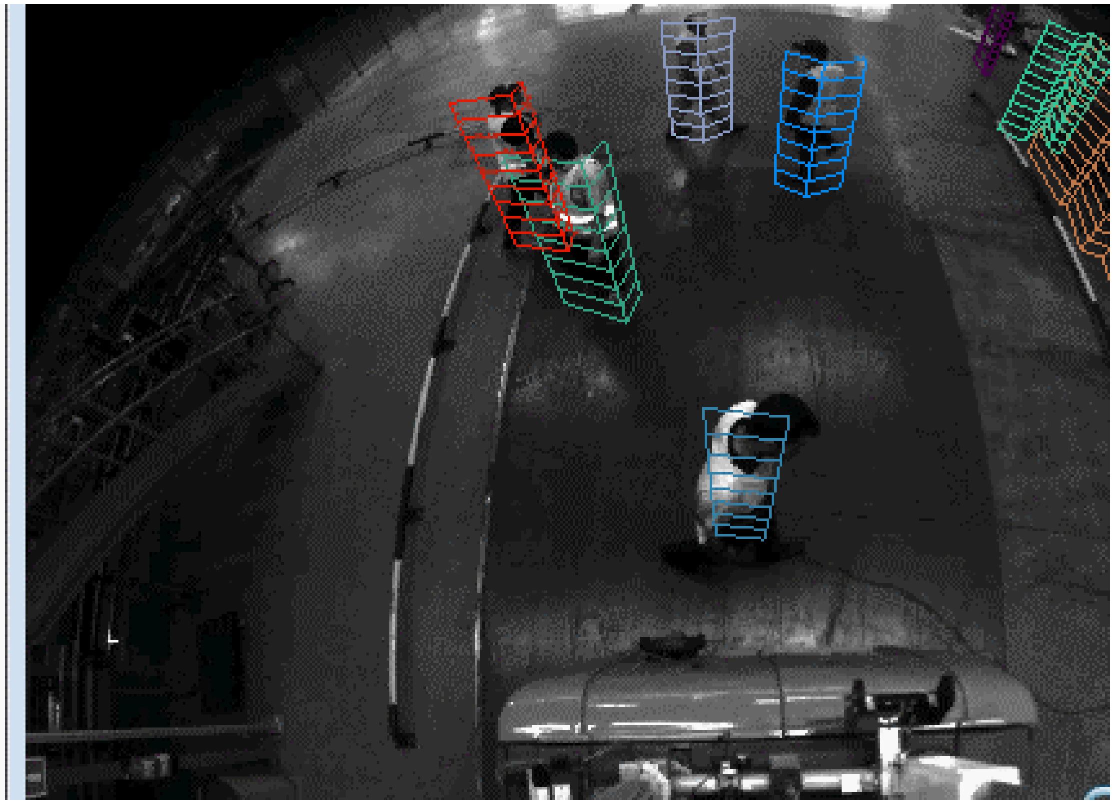
- The overall obstacle tracking solution is decomposed into several steps:
  1. **AGV motion compensation**
  2. **State Prediction**
  3. **Position and Size measurement**
  4. **Data Association**
  5. **ICP-based Speed measurement**
  6. **Update**

3D Omnidirectional Stereo Vision
Object Tracking

3D Omnidirectional Stereo Vision
Object Tracking



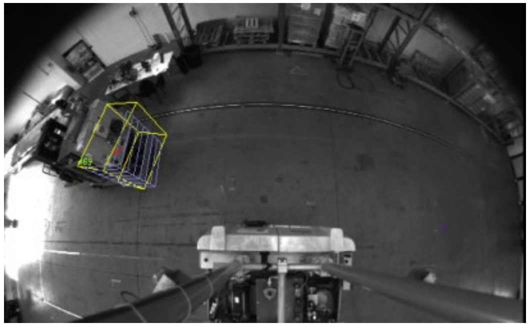
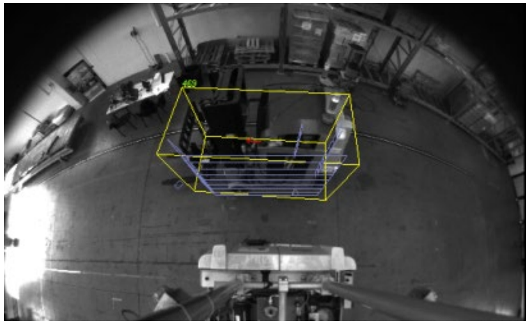
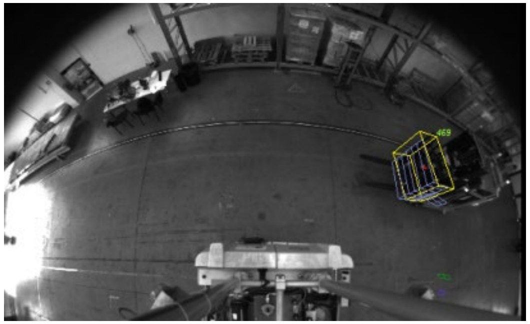
Obstacle Trajectories in Time
(top view)



Obstacle Tracking

Tracking Evaluation

2. Evaluation using AGVs with different known speeds and orientations

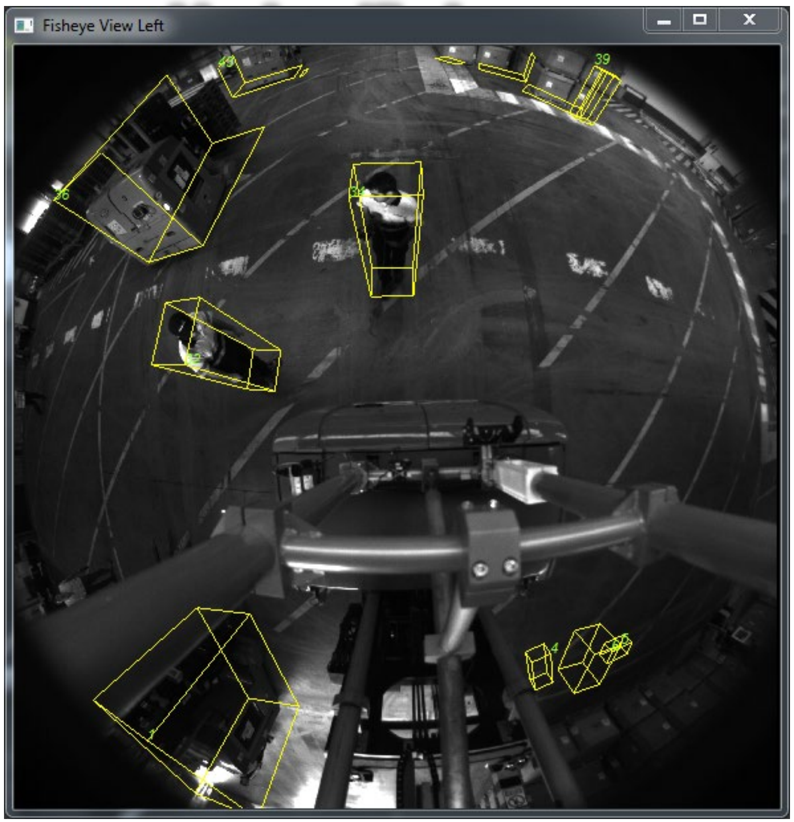
Tracked speed and orientation evaluation

			Orientation			
		Accuracy metric	**0°**	**90°**	**135°**	**Spec at 1σ**
Ground Truth Speed	**0.75m/s**	Speed err. STDEV (m/s)	0.25	0.31	0.14	< 0.5 m/s
		Orientation err. STDEV (deg.)	2.90	7.14	11.44	< 15 °
	1.2m/s	Speed err. STDEV (m/s)	0.21	0.18	0.15	< 0.5 m/s
		Orientation err. STDEV (deg.)	1.77	10.86	10.42	< 15 °
	1.8m/s	Speed err. STDEV (m/s)	0.33	0.19	0.22	< 0.5 m/s
		Orientation err. STDEV (deg.)	1.87	4.87	3.73	< 15 °

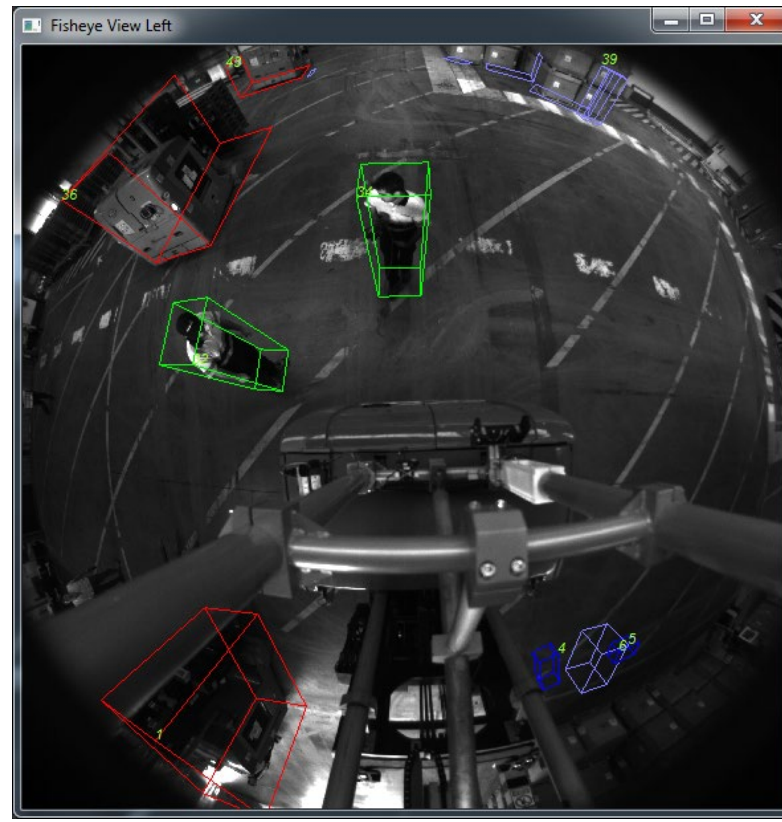
3D Omnidirectional Stereo Vision
Object Classification

Classes

Pedestrian, **AGV**, **Large Obstacle**, **Small Obstacle** (height < 50 cm),



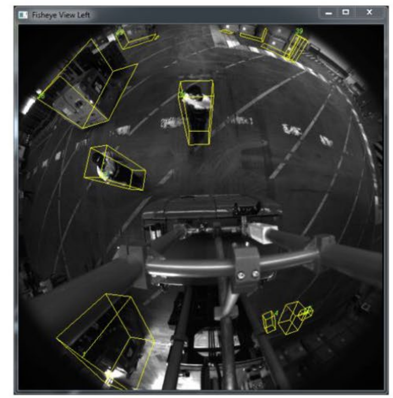
Initial obstacle hypotheses



Classified Obstacles

3D Omnidirectional Stereo Vision
Object Classification

Classification Process
- 3D box -> 2D image
- Resize image
- Compute descriptor
- Classify descriptor


Visual Codebook Based Image Description
- Use HOG as local descriptor
  24 dimensional vector
  for a 8 x 8 pixel region

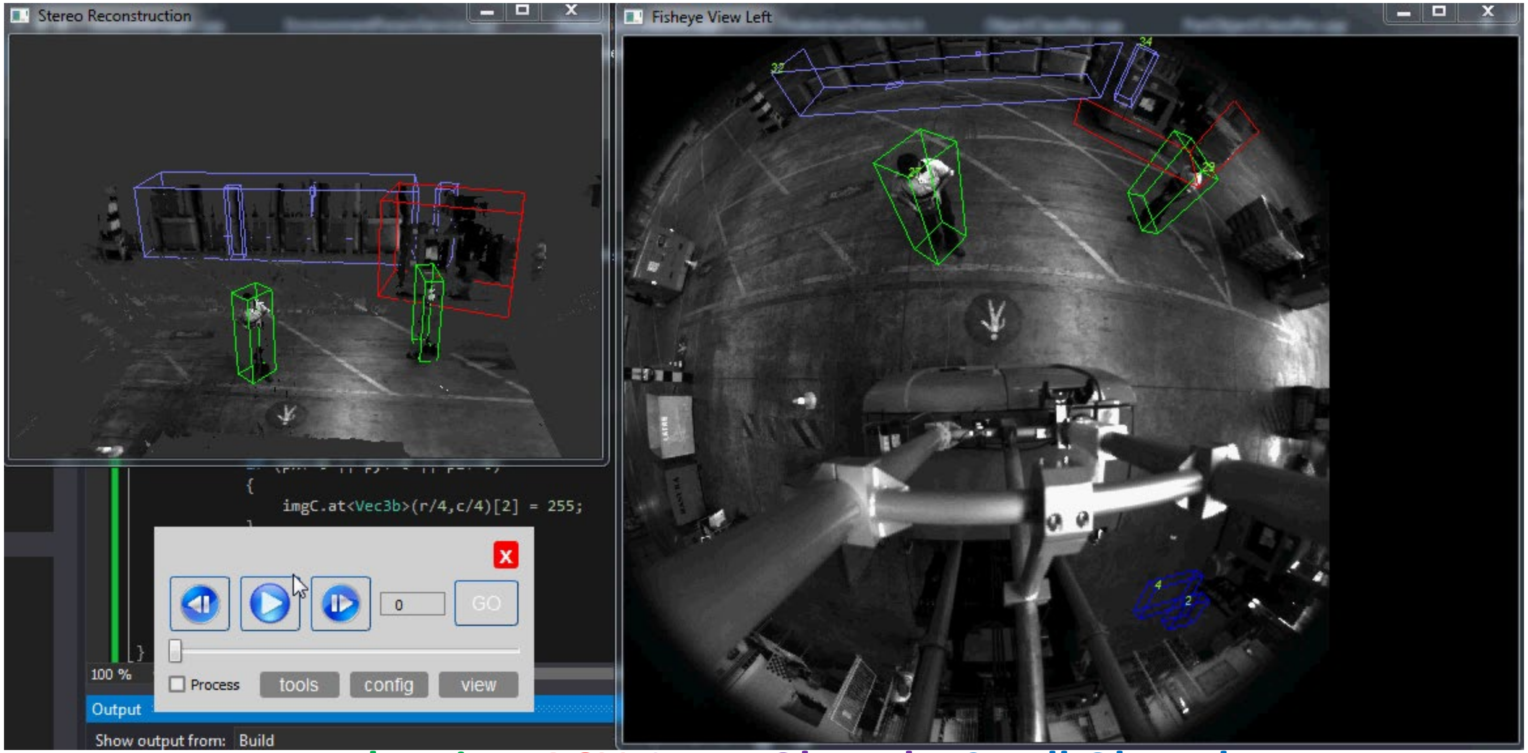
- Dense local descriptor computation
- Create a visual codebook for HOG descriptors - 100 words.

3D Omnidirectional Stereo Vision
Object Classification

Classification Features: word count in 25 regions

Classification

- Boosting classifier with 2048 decision stumps
- Use classification history for tracked objects



Pedestrian, AGV, Large Obstacle, Small Obstacle

Classification Evaluation

- Obstacle database from 32 different scenarios, annotated manually:
  - 8709 pedestrians
  - 6753 AGVs
  - over 130000 other obstacles
- 75% used for training, 25% used for evaluation

Class	Recall	Precision
Pedestrian	91%	92%
AGV	93%	94%
Other	96%	97%

Recall – how many pedestrians were detected from all the pedestrians?
Precision – how many pedestrians were really pedestrians? (penalize false positives)

PAN ROBOTS
Plug&Navigate robots for smart factories

3D Omnidirectional Stereo Vision
Push Beyond State of the Art

- **Original Omnidirectional Stereo Solution**
  - Multi-channel virtual pinhole imagers
  - Multi-channel based rectification algorithms
  - Multi-channel based SGM stereo reconstruction
- **Uncertainty Based DEM**
  - Direct and invers stereo sensor model estimation method
  - Invers sensor based DEM creation algorithm
- **Original Object Tracking and Classification** methods in the context of omnidirectional stereo vision
- **The Concept of 2D Safety Enhanced by 3D Perception**

PAN ROBOTS
Plug&Navigate robots for smart factories

Project Coordinator:
KAY FUERSTENBERG
SICK AG
Merkurring 20
22143 Hamburg, Germany
kay.fuerstenberg@sick.de

Project Dissemination Manager:
CESARE FANTUZZI
UNIVERSITY OF MODENA AND REGGIO EMILIA
via Amendola 2, (pad. Morselli)
42122 Reggio Emilia, Italy
cesare.fantuzzi@unimore.it

Presentation author:
Sergiu Nedevschi
TECHNICAL UNIVERSITY OF CLUJ-NAPOCA
Str. Memorandumului
400114 Cluj-Napoca, Romania
Sergiu.nedevschi@cs.utcluj.ro

PAN-Robots is funded by the European Commission, under the 7th Framework Programme Grant Agreement n. 314193.