INTERSAFE2

On Board 6D Visual Sensors for Intersection Driving Assistance Systems

S. Nedevschi, T. Marita, R. Danescu, F. Oniga, S. Bota, I. Haller, C. Pantilie, M. Drulea, C. Golban

Sergiu.Nedevschi@cs.utcluj.ro

Technical University of Cluj-Napoca, ROMANIA

Volkswagen AG

European Commission

SEVENTH FRAMEWORK PROGRAMME

Outline

INTER SAFE2

- Requirements for the stereo sensor
- Proposed approach
- Stereo sensor architecture
  - Low level architecture
    - Stereo reconstruction
    - Optical flow
    - Ego motion estimation
  - High level architecture
    - Structured environment description
      - Lane perception
      - Painted road object perception
      - Obstacle perception
    - Unstructured environment description
      - Classified occupancy grid
      - Attributed poly-line map
- Conclusions/Discussions

Requirements for the stereo sensor

INTER SAFE2

- Main roles of the stereovision sensor in an intersection driving assistance system are related to *sensing* and *perception* of the environment in front of the ego vehicle
- Perception functions are applied to both
  - Static road and intersection elements
  - Dynamic road and intersection elements
- Perception means: detect, measure the relative position and speed, track, classify and localize the objects on the road and in the intersection

Requirements for the stereo sensor (2)

INTER SAFE2

- Intersection specific constraints:
  - Physical setup of the stereo sensor - wider field of view, increased spatial and pixel resolution cameras
  - Multi-modal information required from the stereo system, improved resolution, improved accuracy
  - Multiple processing paradigms

Proposed approach

- A two level stereo architecture
- The low level architecture - 6D visual sensor
  - Dense stereo reconstruction with improved accuracy - 3D position information
  - Accurate dense optical flow allowing for motion field estimation -3D motion vector information
  - Visual odometry
- The high level architecture implementing two perception paradigms:
  - Structured environment perception
  - Unstructured environment perception
- Fusion

Proposed approach (2)

- **Structured environment perception**
  - road geometry - estimated from visible lane/road delimiters
  - obstacles – modeled as cuboids having position, size, orientation and speed
  - corresponds to un-crowded environments with visible lane delimiters

Proposed approach (3)

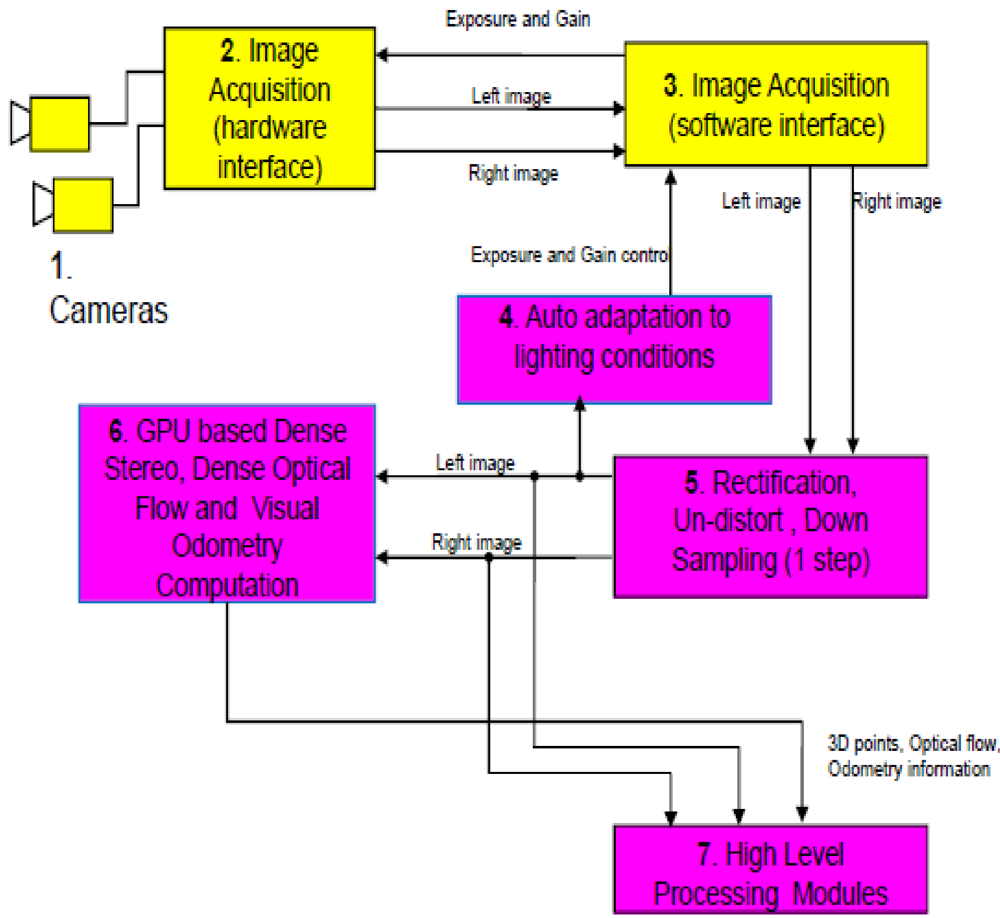
INTER SAFE2

- **Unstructured environment perception**
  - Road geometry and obstacles – detected and classified from digital elevation maps – classified occupancy grid
    - Road and intersection drivable area, 3D curbs, side walks and traffic isles
    - Obstacles are characterized by attributed poly lines
  - Appropriate for crowded environments like intersections in which lane delimiters are frequently not visible
- **Fusion** of the results provided by the two paradigms
  - Refine the results in both qualitative and quantitative parameters

Stereo sensor:
Low level architecture

INTER SAFE2

- Low level 6D stereo architecture and implementation
- Provides the primary information needed by the high level processing modules:
  - 6D point information (3D position and 3D motion)
  - Ego motion estimation
  - Intensity images



Stereo sensor: Low level architecture (2)

INTER SAFE2

- Improved stereo reconstruction
  - Solution focused on both accuracy and processing speed using an improved variant of the Semi-Global Matching (SGM) method
  - Research was focused to improve performances of the stereo reconstruction algorithm related to:
    - Running time
    - Integer disparity accuracy
    - Sub-pixel disparity accuracy

Stereo sensor: Low level architecture (3)

INTER SAFE2

- Proposed Semi-Global method
  - Improvements over existing SGM algorithms
    - Reduced computational complexity (by a factor of 2) without affecting quality
    - Improved depth accuracy and point spread using an innovative adaptive sub-pixel estimation
  - Improvements over the local method
    - Handles more efficiently featureless regions and repetitive patterns
    - Increased density and reduced number of errors
    - Improved accuracy and reduced point spread
  - Features
    - GPU based implementation, running time 19ms for full image (13ms ROI)

Stereo sensor: Low level architectures (4)

INTER SAFE2

- Tyzx DeepSea (local method, as reference) reconstruction (JAI A10 - 768 x 576)

Stereo sensor: Low level architecture (5)

INTER SAFE2

- Our proposed SGM method (JAI A10 – 736 x 576)

Dense optical flow

INTER SAFE2

- Increased accuracy differential algorithm based on a cost function using L1 norm which is less sensitive to noise
- Sequential implementations cannot reach real-time performance
- High degree of parallelism favored a GPU-based implementation
- Performed simultaneously with stereo reconstruction, in one step, at 20ms per frame
- Improvement are required in overlapped and left areas

Ego motion estimation

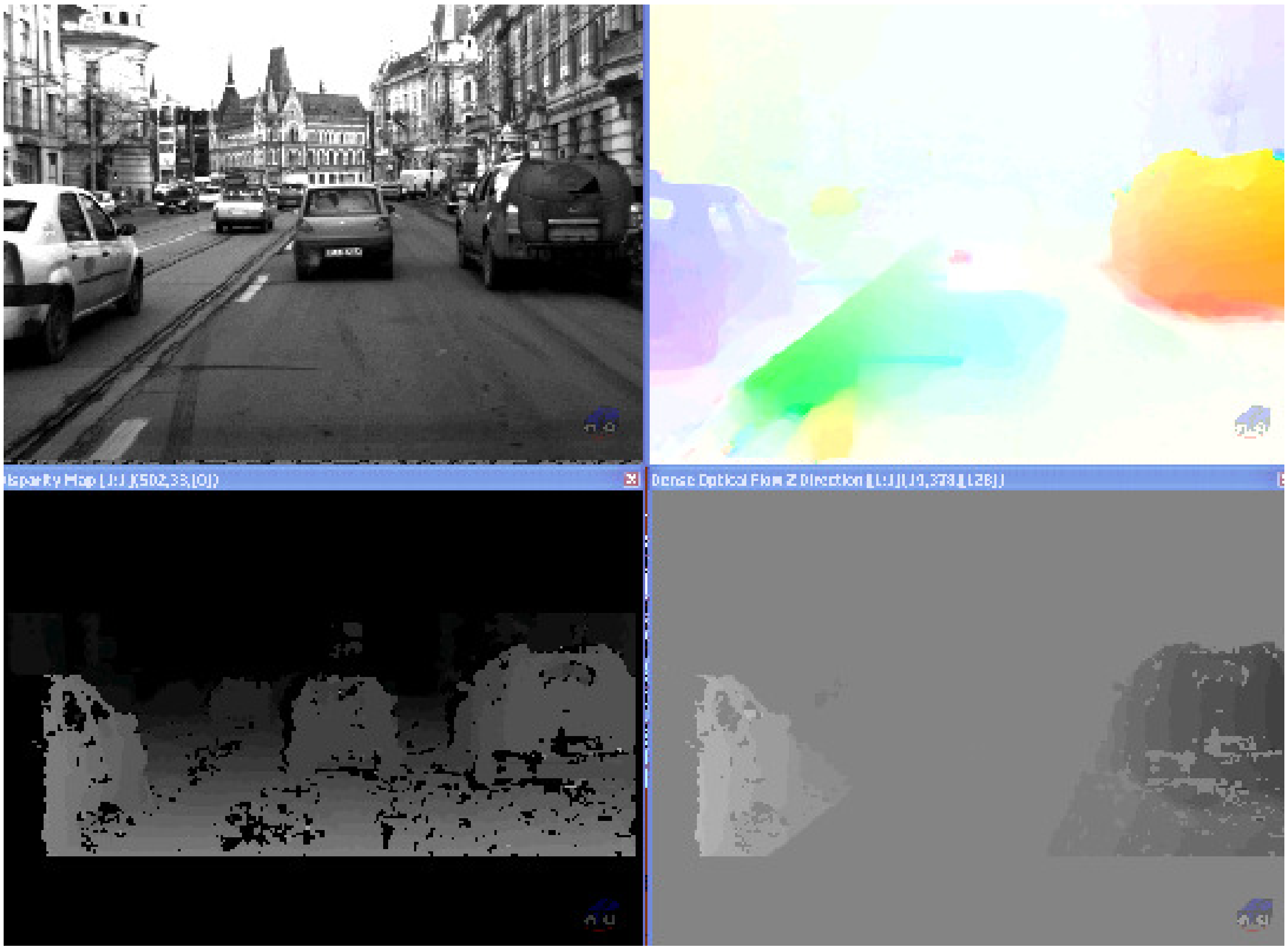
- Vehicle motion can be estimated using information provided by the car sensors:
  - Yaw rate, speed, global time

- Maximum accuracy can be reached only if the 6 degrees of freedom of the ego motion are determined precisely

- The developed solution is using consecutive stereo pairs and is carrying out a direct evaluation based on Levenberg Marquard optimization and RANSAC

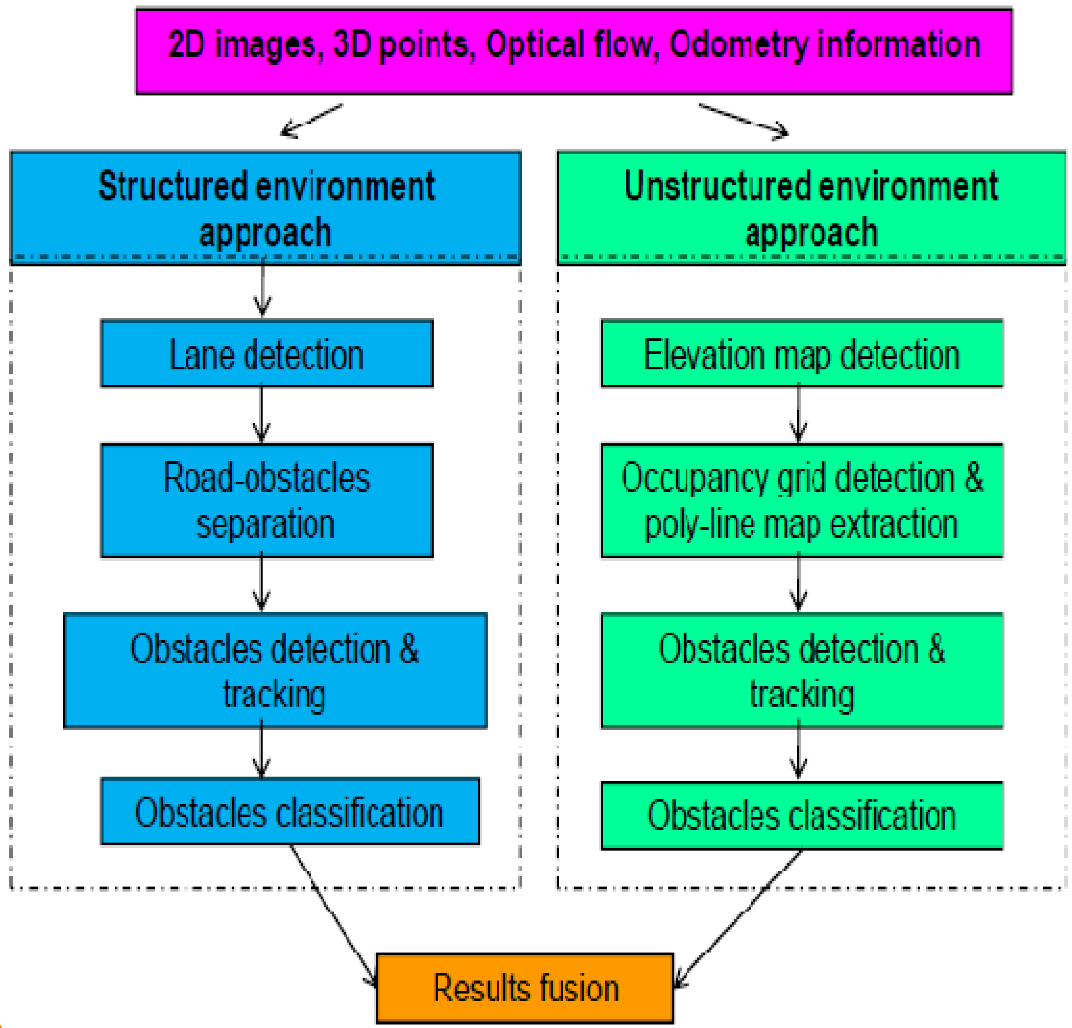
Dense optical flow and Ego motion

INTER SAFE2



Stereo sensor: High level architecture

- Two environment description paradigms:
  - Structured description
  - Unstructured description



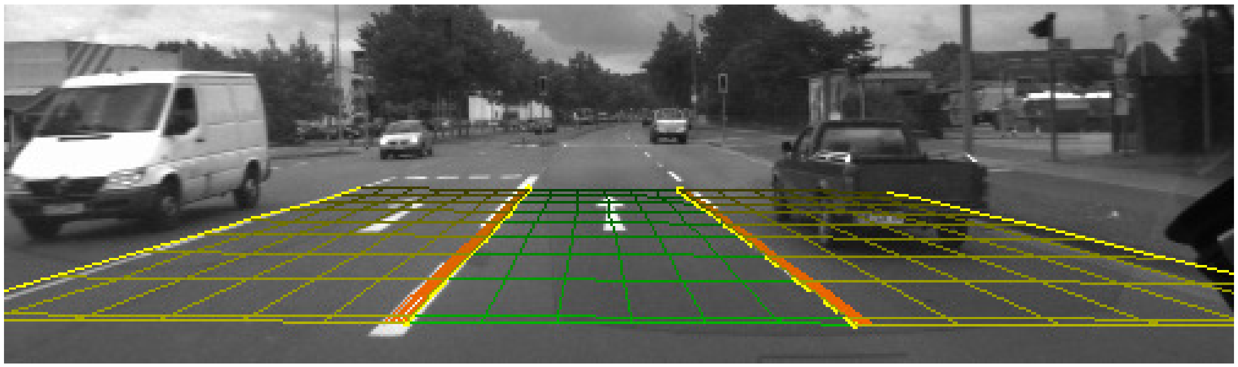
Structured environment description

- A complete description of the environment can be constructed using:
  - Lane perception
  - Painted road object perception
  - Obstacle perception

Structured environment description (2)

Lane perception

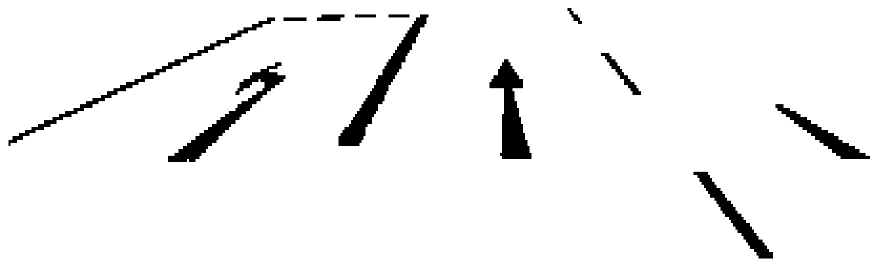
- Input: 3D points, left grayscale image
- Road vertical profile detection and 3D points discrimination: road, obstacles
- 3D lane markings and lane delimiters extraction
- Current and side lanes 3D model estimation: particle filter based approach

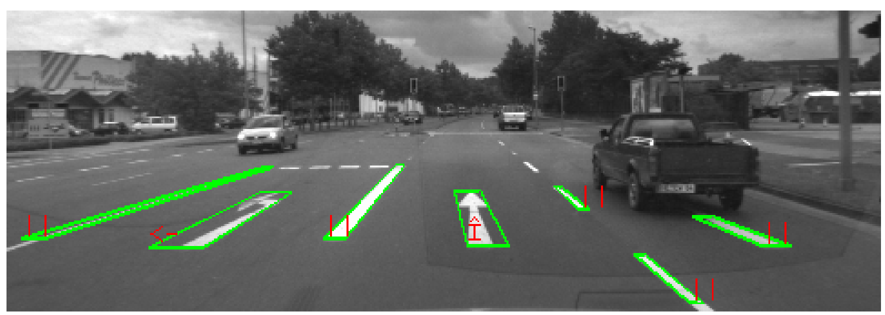


Structured environment description (3)

Painted road object perception: lane markings, arrows

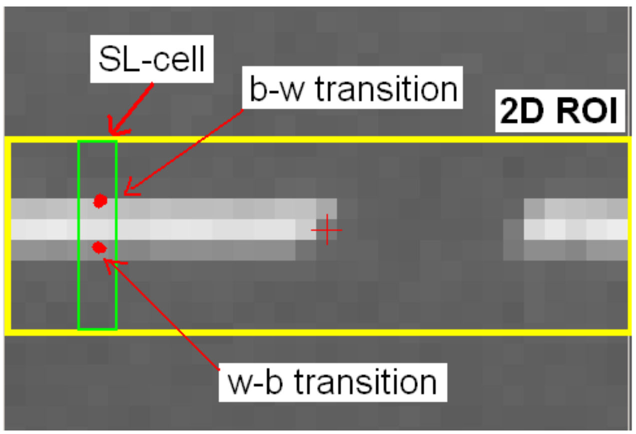
- Steps: 2D segmentation, 3D localization, classification





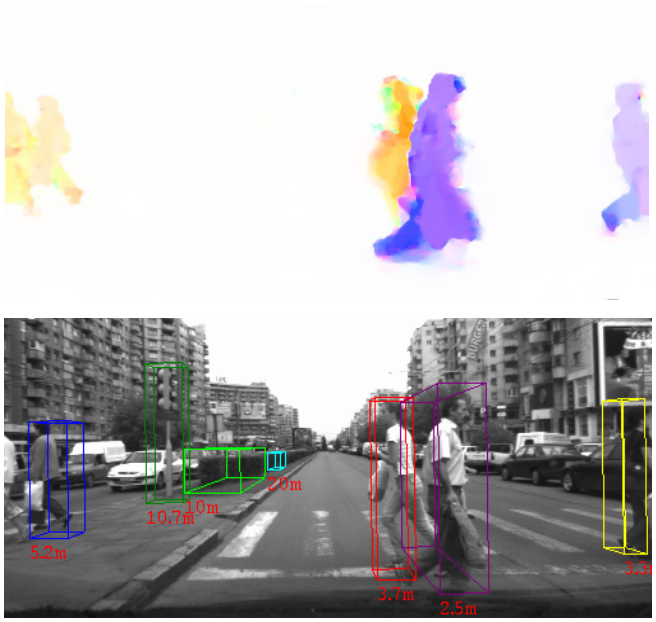
Painted road object perception: stop line, crossing lines

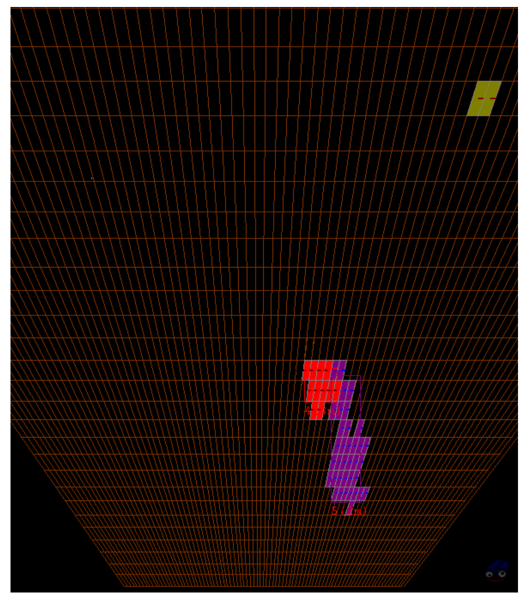
- Steps: Selection of 2D and 3D ROI, detection of the horizontal line structures in 2D ROI, 3D validation and grouping, 2D hypothesis validation and model based classification, refinement of the 3D position, 3D objects tracking



Improved obstacle detection

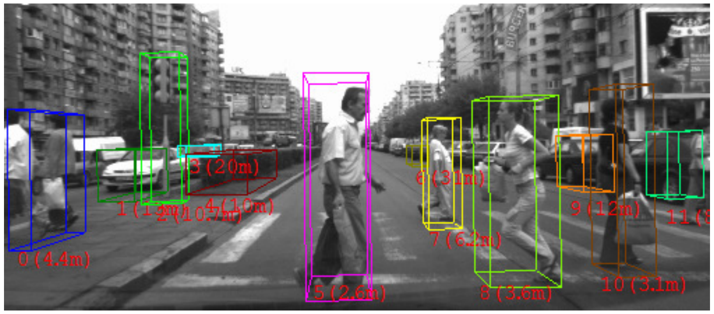
- Based on dense depth, dense/sparse motion and intensity information
- Output: obstacle position and size

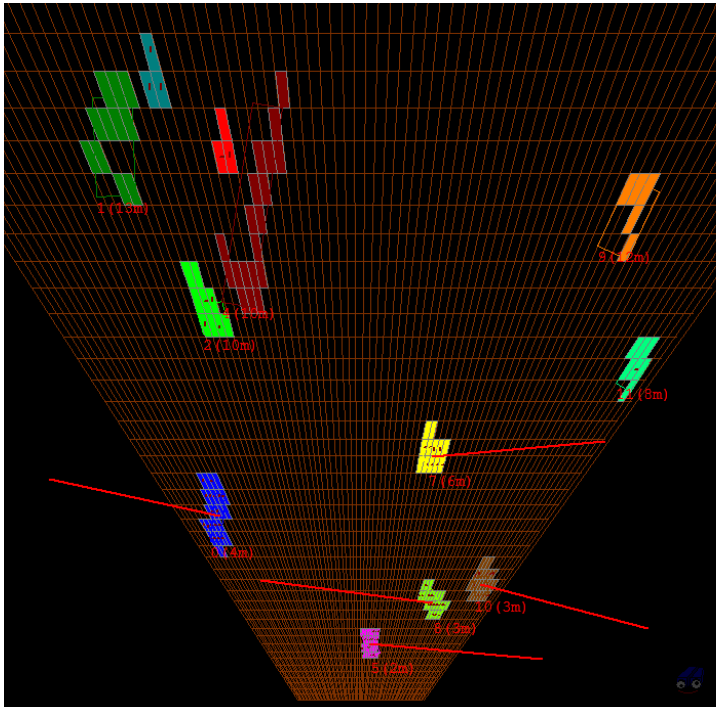




Structured environment description (6)

- Output: speed, and orientation



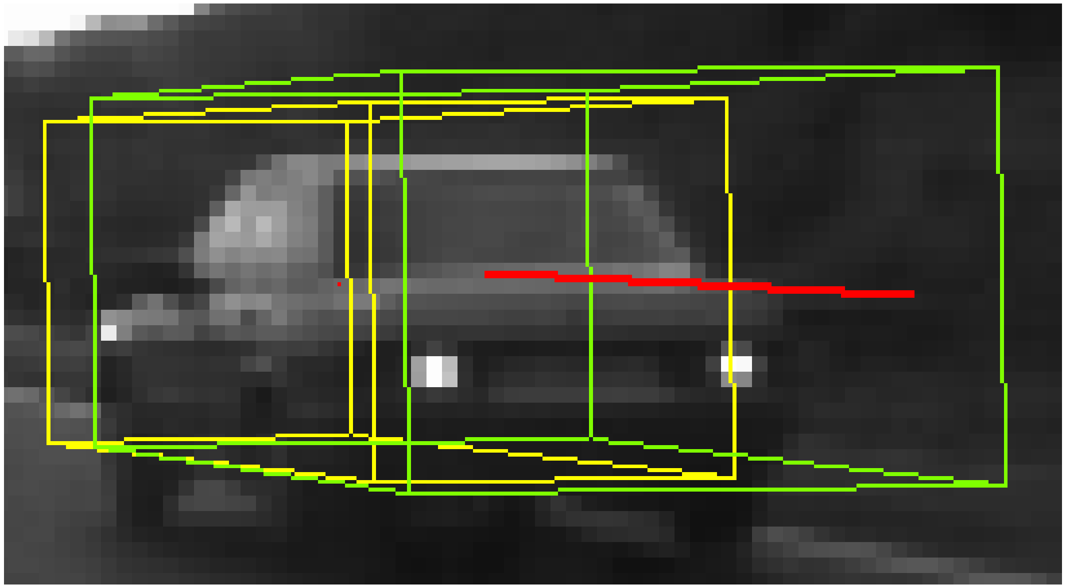


Structured environment description (7)

INTER SAFE2

Optical flow based obstacle tracking

- Steps: prediction, association, object merging/splitting and Kalman filtering.
- The innovation consists in the use of robust motion parameters of the ego vehicle and of the detected objects from the previous frame for generating the predictions.



Elevation map based environment representation

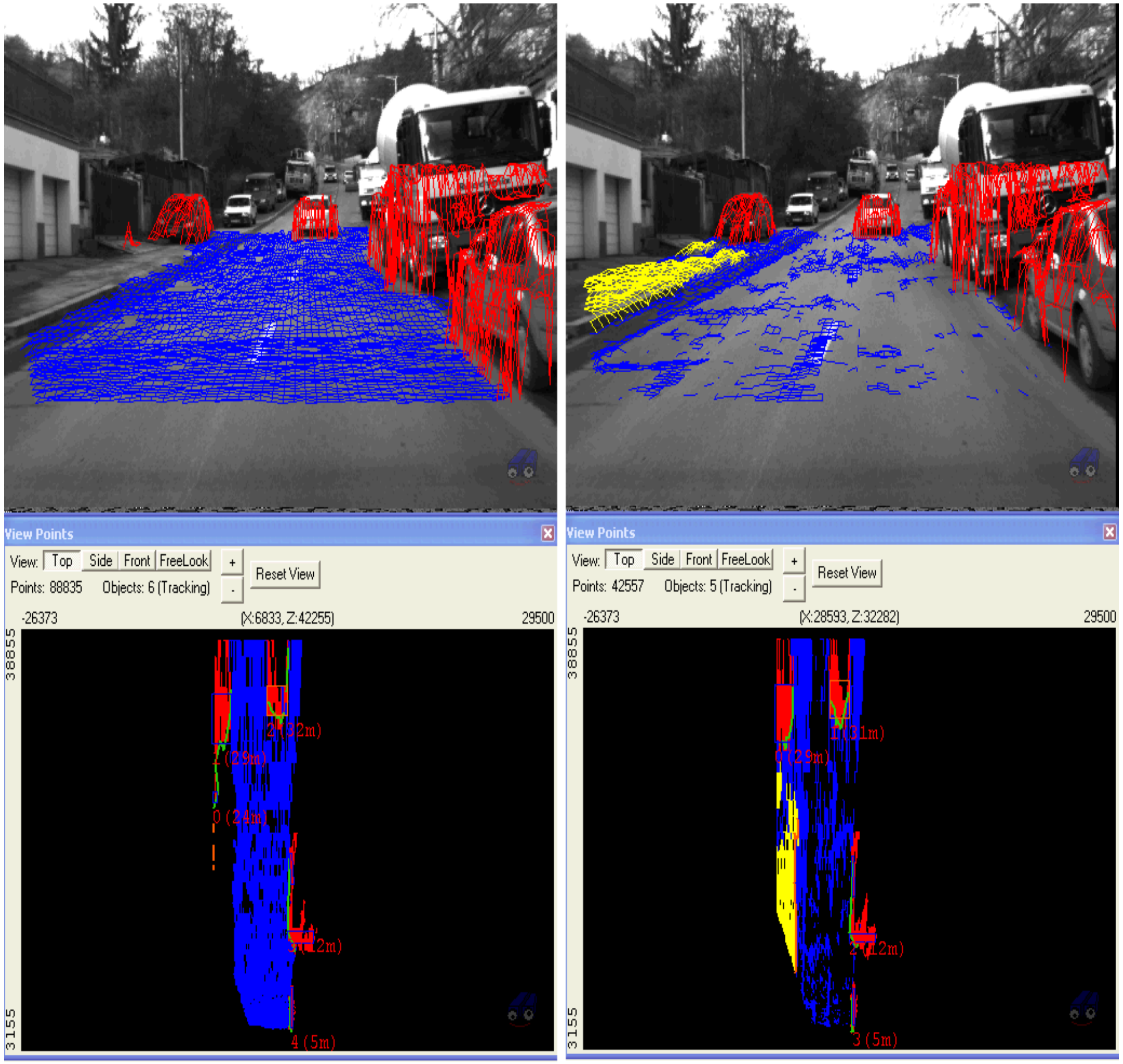
- Classified Occupancy Grid obtained by elevation map cells classification
  - Quadratic surface fitting to the region in front of the ego vehicle followed by a region growing process
  - The road surface is used for discrimination between road, side walks and traffic isle, and obstacles cells

- Curb detection algorithm based on a cubic polynomial model: curb points detection, RANSAC based extraction of the best polynomial

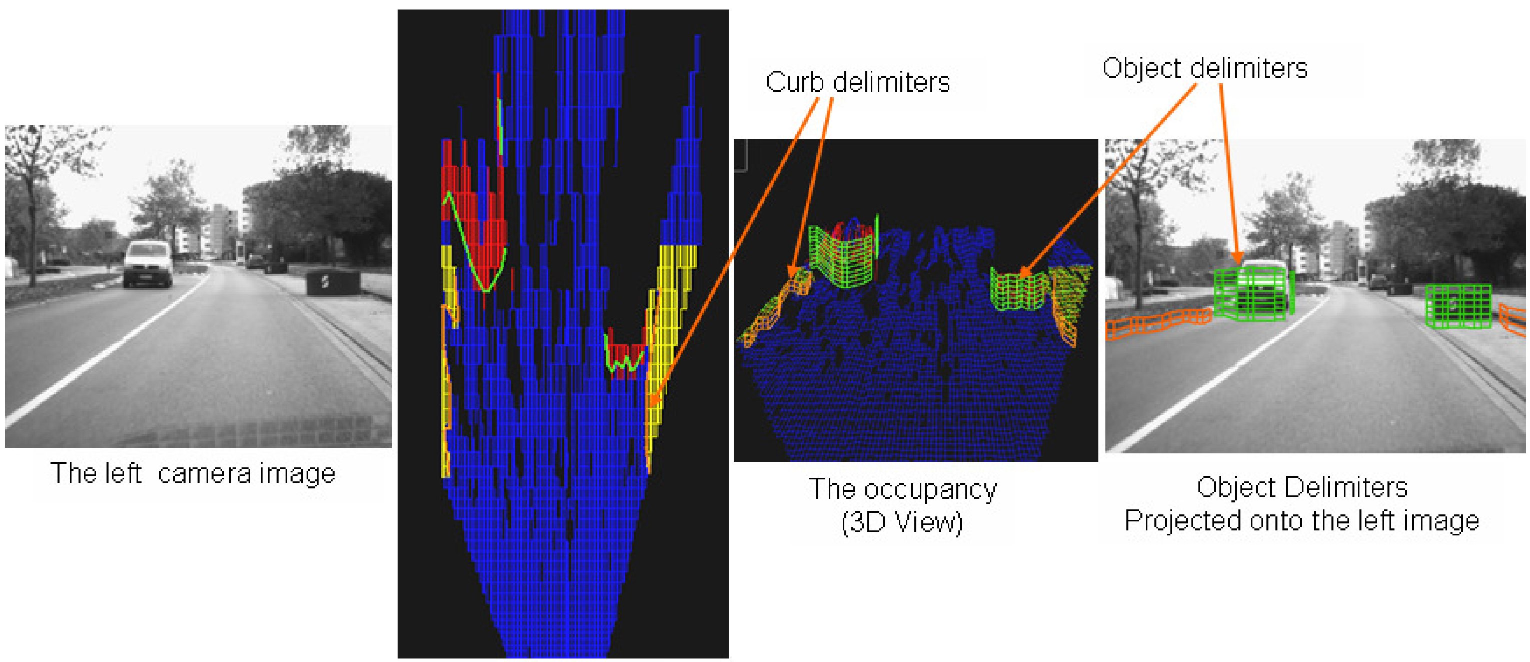
- Attributed poly-line based representation of the driving area delimiters and object delimiters

Unstructured environment description (2)

INTER SAFE2



Attributed poly-line based representation of the environment



The left camera image

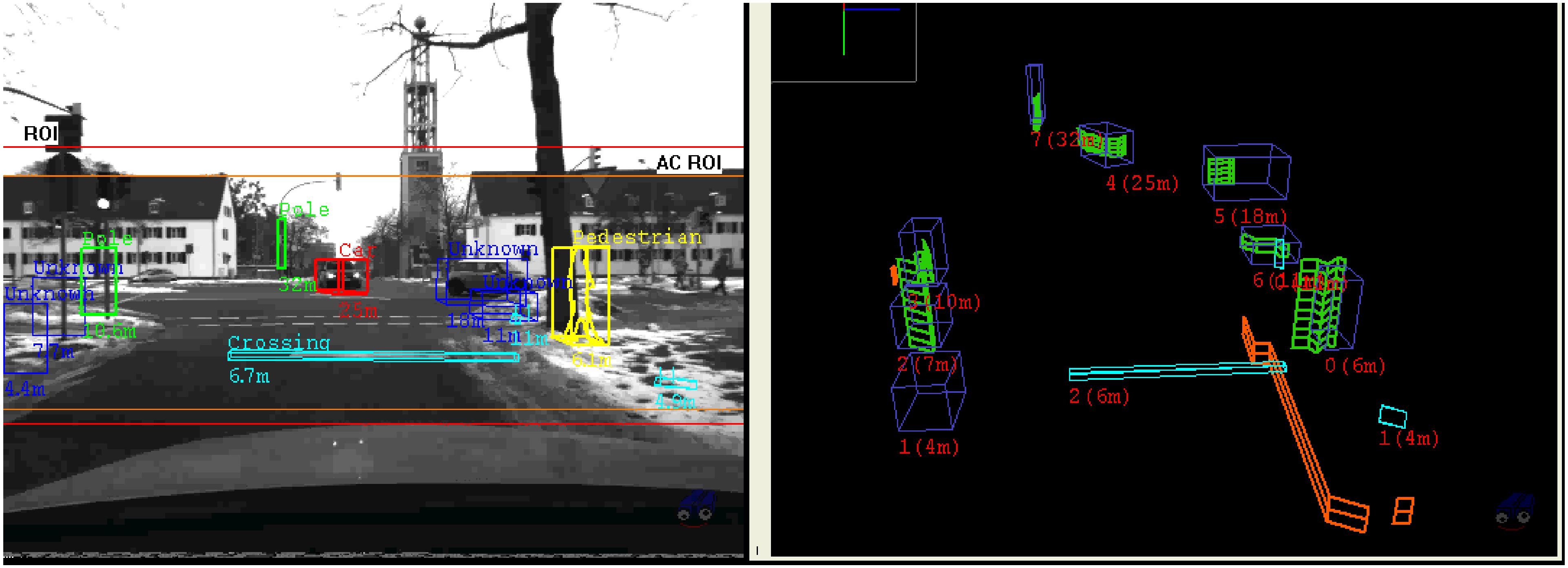
The occupancy (Top View)

The occupancy (3D View)

Object Delimiters Projected onto the left image

Fusion

INTER SAFE2



Conclusions

INTER SAFE2

- Two levels stereo sensor architecture
- Dense stereo, optical flow and visual odometry is a must
- Motion vector and ego motion based grouping
- Motion vector and ego motion based tracking
- Structured environment and unstructured environment paradigms
- Compressed representation based on attributed poly-line representation

Thank you

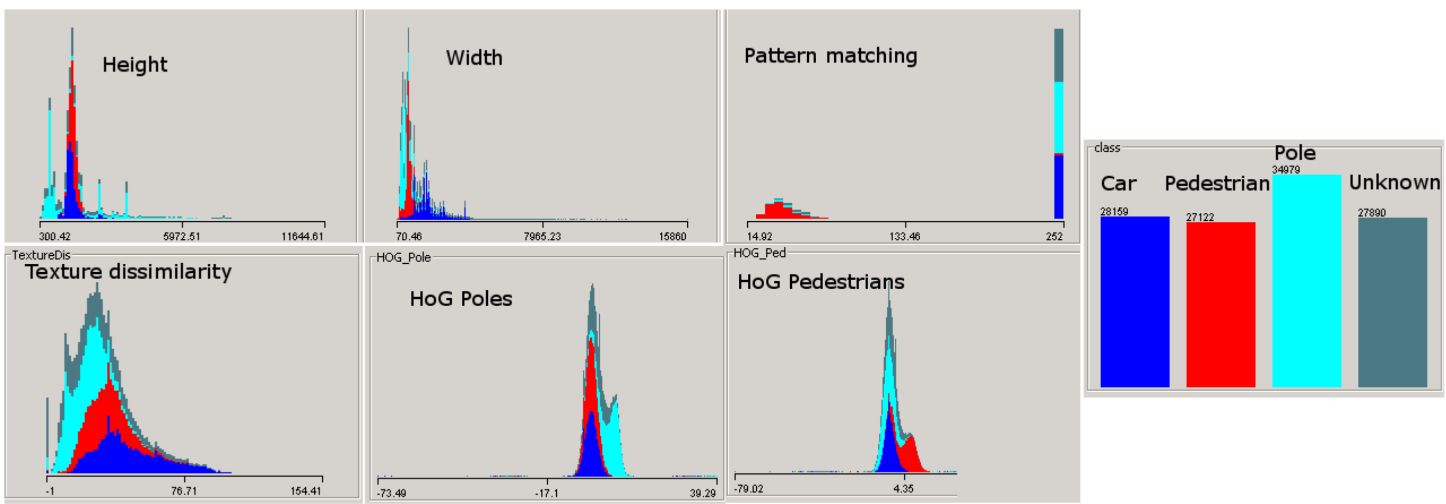
http://cv.utcluj.ro

Structured environment description (8)

INTER SAFE2

Relevant objects classification: car, pedestrian, pole

- Input: Object hypothesis generated by the object detection module, together with the orientation and tracking information.
- Coarse classification based on size, shape and speed information
- Refining using a hierarchy of statistical classifiers



Structured environment description (9)

INTER SAFE2