OBSTACLE DETECTION FOR PEOPLE MOVERS USING VISION AND RADAR

Johan C. van den Heuvel, Senior Project Manager
Jan C.M. Kleijweg, Project Manager
Wannes van der Mark, Research Scientist
Christiaan M. Lievers, Research Scientist
Leon J.H.M. Kester, Research Scientist
TNO Physics and Electronics Laboratory
Oude Waalsdorperweg 63, 2597 AK The Hague, The Netherlands
Tel: +31 70 374 0453 - Fax: +31 70 374 0654
E-mail: vandenHeuvel@fel.tno.nl

SUMMARY

People movers (autonomous road vehicles) are a new means of public transport with promising capabilities. A roadmap is presented that shows the anticipated development of crucial technologies for people movers. A critical technology is obstacle detection. An obstacle detection system is presented based on stereo vision and radar that uses sensor fusion and tracking for obstacle detection. The obstacle detection system was tested on an autonomous test vehicle and showed promising results.

INTRODUCTION

People movers represent a new class in public transport. They are autonomous systems, i.e. without a human driver, that use the normal road, since they are fitted with tires. At present people movers use a reserved part of the road, however, in the future they will mix with normal traffic. They support traditional public transport by offering frequent transport to and from stations of large public transport systems. In addition, people movers are a promising concept for transportation inside cities in order to realise better access to the inner city and reduction of car traffic.

In the Netherlands various systems have been realised: Parkhopper at Amsterdam Airport, ParkShuttle at business park Rivium, and the Cybercaps demonstration at the theme park Floriade. Due to the people mover activities in the Netherlands, a roadmap has been developed by the various stakeholders: industry, municipalities, public transport authorities, research institutes, and the ministry of transport. The roadmap shows that obstacle detection systems are a critical technology.

By using the technological developments in the field of intelligent transportation systems, people movers can now be realised in a cost-effective and environment-friendly way. Since people movers are autonomous systems, obstacle detection systems are an essential component of the system. In this paper an obstacle detection system is presented that is based on stereo vision and radar. Sensor fusion with tracking is used to obtain a good detection probability of obstacles with a low false alarm rate and to predict situations that require action.

TECHNOLOGY ROADMAP

In view of the potential of people movers for public transport, a roadmapping study was performed by all the relevant organisations: industry, municipalities, public transport authorities, research institutes, and the ministry of transport. Connekt, an organisation that was founded by several Dutch ministries, companies, and organisations, initiated the roadmapping study. The mission of Connekt is to manage the innovation cycle in traffic and transport. The developed roadmap aims to clarify the perspective of people movers in the future. For this objective, the demand-side, the supply-side and the regulations were addressed as shown in Figure 1.

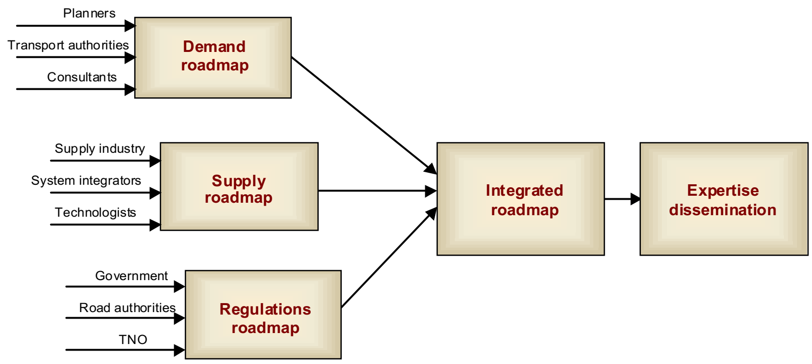


Figure 1: Roadmapping structure.

The conclusion from the demand-side is that a distinction has to be made between long distance transport using large vehicles and short distance transport using smaller vehicles. Short distance transport with people movers will be mainly in theme parks and residential areas. With long-distance people movers, a high frequency transport between a station and a residential area can be provided. It is expected that both systems will integrate in the future to provide an improved service to the passengers.

From the viewpoint of regulations, system safety, insurance and certification were addressed. Most important was the issue of people movers on the public road. It was investigated which steps should be taken to make this possible.

Developments from the supply side were identified in nine functions: energy supply, traction, obstacle detection, localisation, vehicle control, communication, man-machine interface, safety, and reliability. It was clear that there is a strong dependency on technology developments in the automotive industry. Technologies for obstacle detection are a clear example of this. Two critical technologies were identified: energy supply and obstacle detection.

Figure 2 shows the roadmap for these technologies. In this paper we present the development of an obstacle detection system based on radar and (stereo) vision for future people movers.

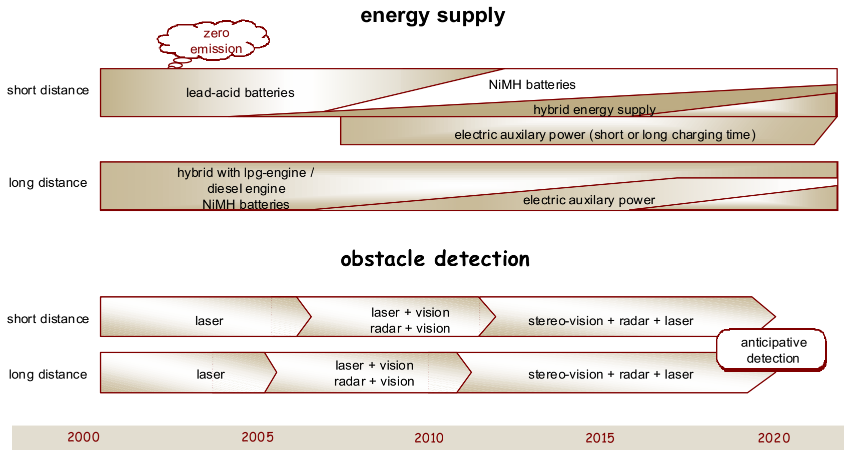


Figure 2: Road map of energy supply and obstacle detection.

AUTONOMOUS TEST VEHICLE

For the development and testing of obstacle detection systems, a dedicated and advanced test vehicle is used: the RoboJeep (Figure 3).

Figure 3: The autonomous test vehicle: RoboJeep.

The test vehicle is based on a modified Jeep Wrangler 4x4 platform. Sensors include radar, sonar, laser radar, stereovision cameras, odometers and GPS navigation hardware. Actuators control the automatic gearbox, steering wheel, throttle and brakes. On-board computation uses two ruggedised 400 MHz Pentium II PC's (Beckhoff) connected via a 100 Mbps Ethernet Hub. Both PC's run Windows NT. Sensor and Actuator interfacing is provided by industrial grade PLC hardware (Beckhoff Modules) connected to the PC via field buses (Profi-bus and/or CAN-bus). The system operates on an independent 24V battery, which is recharged by the vehicle's engine. Primary safety during testing is provided by a 'Stop' button that de-couples all computer controlled actuators and returns vehicle control to the driver. The hardware is very modular both in architectural and in physical sense. The platform control architecture is based on the Real-time Control System (RCS), a reference model architecture

for intelligent systems developed at the National Institute of Standards and Technology (NIST) (1)(2).
The RoboJeep is capable of driving autonomously for the testing of obstacle detection in autonomous vehicles (3). In addition, it can drive on the public road for testing of obstacle detection systems in order to enhance road safety. At present the RoboJeep is an excellent test vehicle for data acquisition using a multitude of state-of-the-art sensors in structured and unstructured terrain. Data acquisition in a number of scenarios has been performed. In this paper, we will present obstacle detection developments based on the acquired data with the RoboJeep and real-time obstacle detection results from radar and vision sensors mounted on the RoboJeep.
Future work will focus on the further development and testing of sensor-fusion algorithms for obstacle detection. Here, the RoboJeep is ideally suited for the real-time testing of obstacle detection algorithms and systems. The testing of the obstacle detection algorithms will include driving on the public road in real traffic situations.

VISION-BASED OBSTACLE DETECTION

Stereo vision is a well-known technique for obtaining geometrical information of a scene (4). Range information (depth) can be extracted from stereo images by matching corresponding image points. In addition, elevation and azimuth information is given by combining the distance with the image points locations. Due to the presence of range information, data fusion with (short-range) radar is established.
Figure 4 shows a diagram of the steps, which are performed to supply the sensor fusion with visually detected obstacles. These steps consist of image rectification, depth estimation and obstacle detection & segmentation.

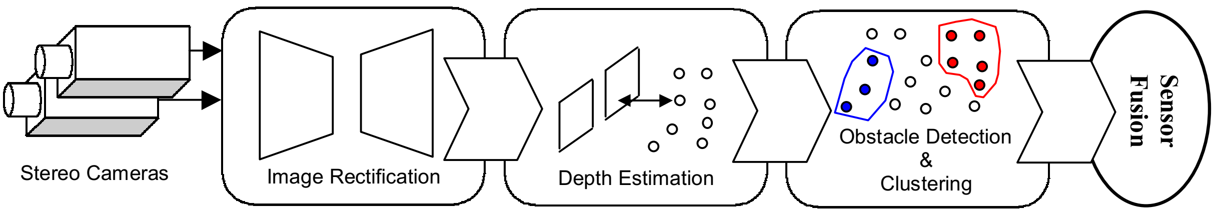


Figure 4: Overview of the vision based obstacle detection

Image rectification

The most commonly used model for camera projection and stereo reconstruction is the pinhole model. This camera model assumes that light rays enter the camera through a small hole. However, real cameras use a lens to capture an image of the world. The spherical lens causes radial distortions in the image. It is possible to correct a normal camera image in such a way that it becomes equal to an image taken by a pinhole camera if this distortion is known. With a calibration procedure (5) it is possible to measure this and other properties of cameras.
Points in the world and their projections in both cameras form one plane, know as the epipolair plane. This plane intersects both camera images along the epipolair lines. Finding corresponding stereo points is therefore reduced to searching along the epipolair lines. To exploit this fact it is desirable that the epipolair lines of a stereo pair are parallel with the image lines. Unfortunately, this is not the case because of small differences in the camera's orientations. Camera calibration can also be used to measure the orientation and position

difference between cameras, which in turn can be used to compute the image conversion, needed to rectify the stereo images.
In the first step the lens distortion is removed from each newly captured image. Then the images are rectified to make them suitable for stereo matching.

Depth estimation

Together with the stereo camera parameters from calibration and the disparity between corresponding stereo points, real-world distances can be retrieved. In order to find corresponding pairs of stereo points, they first have to be compared for different disparities, after which the best matching pairs can be determined.
The maximum range at which the stereo vision can be used for detecting obstacles depends on the image and depth resolution. These should be sufficiently large, in order to allow enough time for taking evasive action or braking. However, using a high depth and image resolution results in high computation costs for matching all the stereo points. For real-time applications only very efficient stereo algorithms can be used.
We have developed our own efficient implementation for stereo disparity estimation. Absolute differences of pixel intensities are used in the algorithm to compute stereo similarities between points. By computing the sum of the absolute differences (SAD) for pixels in a window surrounding the points, distinction between similarity scores for stereo points can be increased. The disparity associated with the smallest SAD value is selected as best match. By fitting a second-degree polynomial through the minimum and the two closest values, the disparity estimate can be improved to sub-pixel accuracy.
Without any optimisation the complexity of a dense algorithm is equal to $O(xywhd)$, with x and y being the width and height of the image, w and h the dimensions of the rectangular local windows around the pixels and d the interval of possible disparities. Because the windows overlap for equal disparities, a new SAD value can be computed from the value of a neighbouring window. Only subtraction and addition is needed to remove and add the absolute difference values of the pixels, which do not belong to both windows. With this technique the overall complexity can be reduced to $O(xyd)$.
The computation is further optimised by using Single Instruction Multiple Data (SIMD) instruction sets, such as Intel's MMX, SSE and SSE2 or AltiVec on PowerPC type processors. Similar to the stereo algorithms described in (6) and (7) the instructions are used to compute the SAD values and search minima for multiple disparities simultaneously.

Obstacle Detection & Segmentation

As indicated previously the depth information obtained from stereo can be used to reconstruct the 3D location of points.
In the obstacle detection step, the points belonging to obstacles must be found. For data fusion with other sensors it is important to cluster obstacle points in such a way that individual obstacles can be identified.
For both the obstacle detection and clustering we use the approach of Talukder et al. (8). They have developed an efficient method, which uses two simple criterions to decide if a pair of reconstructed points belongs to the same obstacle. The first criterion states that the distance between to points of an obstacle must not exceed a pre-set maximum and must be larger then a certain minimum. The second criterion demands that the line connecting both points must have a sufficiently high angle with the ground plane, because obstacles have predominantly vertically oriented features.

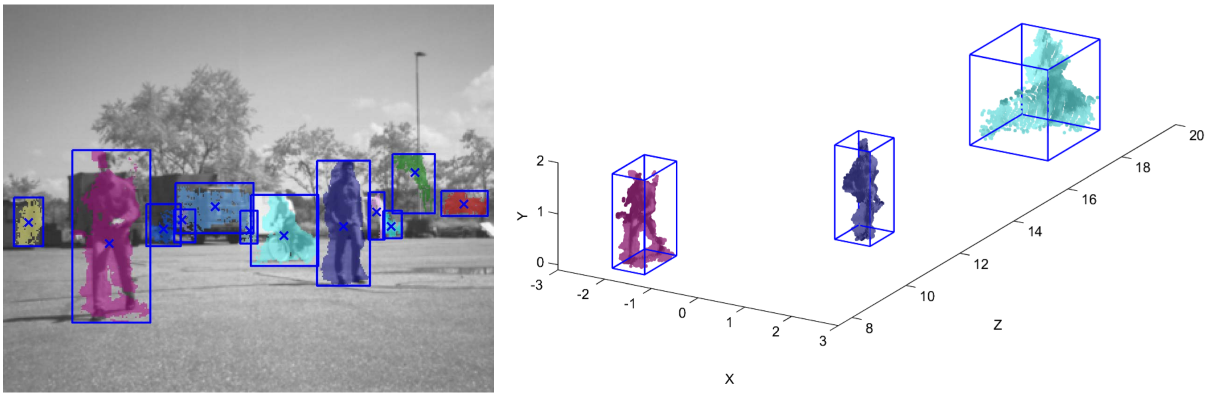
By keeping track of which point pairs up with which other point it is possible to link them into single obstacles. A tree-based labelling scheme can be used for this. Each detected obstacle point refers to a single label. Every obstacle consists of a set of labels organised in a tree structure. The label, which is the base node of the tree, identifies the obstacle. Only linear search is needed to find out to which obstacle a point belongs because each point or label only refers to a single other label. Pruning can be used to optimise identification of points. Setting the retrieved label at the base of the tree directly as the label of the point reduces needed search steps in the future.

By adding compatible points to the trees, points can be clustered into separate obstacles. However, it is possible that two separate obstacles need to be merged into one. If two clusters contain a pair with points, which satisfy the criterions, they should be merged. Because tree structures are used it is also easy to merge obstacle clusters. Only the base node of one of the clusters has to be included into the tree of the other to merge them.

The advantage of the Talukder approach becomes clear after the clustering of points into single obstacles. It opens up the possibility of geometric reasoning with the found obstacles. Simple properties of the clusters such as number of points, volume of the enclosing box or height can be used.

In the current implementation obstacles are removed which only contain a few points. These types of obstacles contain mostly points, which are due to remaining errors in the stereo matching. Also obstacles which have a small volume are removed. Because of their small dimensions they pose no threat to the vehicle.

The result of vision obstacle detection for a traffic situation is shown in Figure 5. The left image shows the detected and segmented obstacles in different colours encapsulated by their bounding box. The right image shows the three nearest obstacles reconstructed in 3D space.



Detected obstacles in the image — Detail of the nearest obstacles in 3D

Figure 5: Vision based obstacle detection

RADAR-BASED OBSTACLE DETECTION

Radar-based obstacle detection is done with two FMCW radars. A wide beam low frequency (2.4 GHz) radar for short distances up to about 60 meters, and a narrow beam high frequency (76 GHz) radar for longer range.

An FMCW radar is a radar that transmit a continuous wave (CW) signal with a changing frequency, in other words frequency modulated (FM). The period of the modulation is about 1 ms, during that period the frequency that is transmitted is increased linearly with 250 MHz.

One period is called a sweep. With the 2.4 GHz radar, for example, the frequency starts at 2.3 GHz and increases linearly to 2.55 GHz. Objects in the radar beam reflect the transmitted signal, this reflected signal is measured by the radar. In Figure 6 a time-frequency plot is shown of both the transmitted and reflected signal.

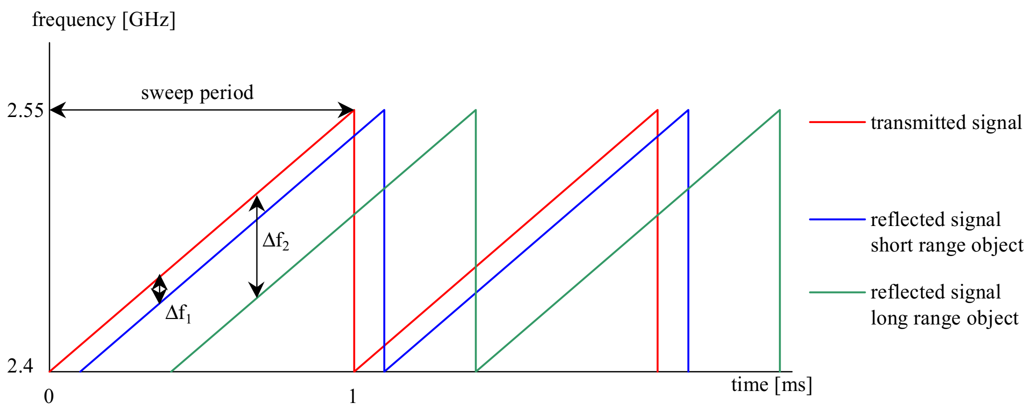


Figure 6: FMCW signals

The frequency difference between a currently transmitted signal with a reflected signal is Δf. The farther an object is from the radar, the larger the frequency difference. This is caused by the time of flight of a signal reflected by an object; the reflections from an object far away need more time to return to the radar than reflections from an object close to the radar.
The frequency difference between currently transmitted signal and the reflected signals is measured. This frequency difference is a measure for the distance of an object. The strength of a reflected signal is a measure for the size of an object.
From the change in range with time, the velocity of an object can be calculated. The reflected signal of an object at constant distance to the radar remains the same, so does the frequency difference. The frequency difference and therefore also the phase of a reflected signal from a moving object will change, the larger the rate of change, the larger the velocity of that object. Using multiple sweeps a range-velocity diagram can be calculated, see Figure 7.

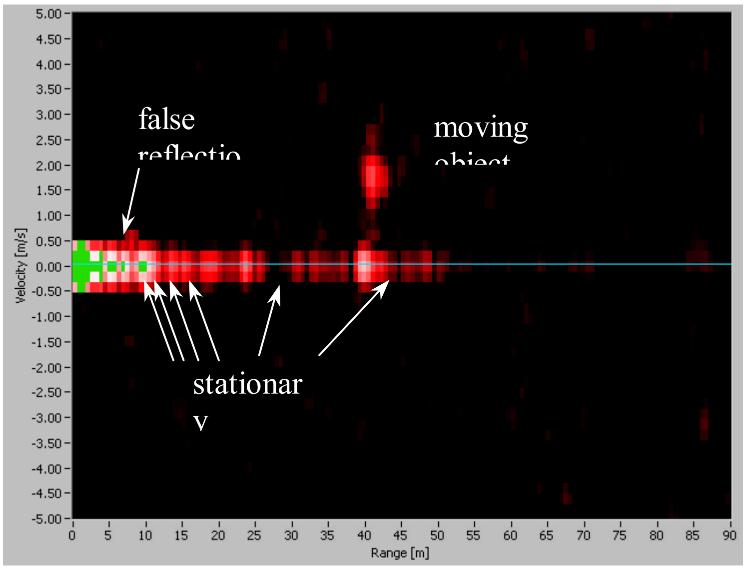


Figure 7: Range-velocity diagram

On the line with velocity 0 m/s the reflected signals from stationary objects can be seen. The colour indicate the signal strength; black indicates a signal below a lower threshold, green a signal above an upper threshold, from the lower to the upper threshold the colour ranges from red to white. Several stationary objects can be seen in Figure 7: at 4, 6, 7, 10, 23 and 40 meters. At 42 meters there is a reflection from a moving object.

Just for indication of maximum ranges and resolutions, the following values are given: maximum range is about 200 m, range resolution is about 60 cm, maximum detectable velocity is about 110 km/hr and the velocity resolution is about 1.8 km/hr.

With a CFAR (Continuous False Alarm Rate) algorithm objects are detected in this range-velocity diagram. The CFAR algorithm generates a list with detections, both false alarms and detections from real targets. The list with detections is clustered, that means that detections that are close by in range and velocity are grouped. Clusters of real targets likely contain several detections, so clusters with only a few detections are considered false alarms and are filtered out.

SENSOR FUSION AND TRACKING

Both the vision-based system and the radar-based system give a list with detections. Since the sensors are different, the list contains different parameters of the obstacles; the vision-based system delivers the range and angle of an obstacle whereas the radar-based system delivers range and velocity information. These detections are fed into a tracking algorithm in which the following hypotheses are considered:

- Does this new detection belong to a detection in the past from the same obstacle?
- Is this detection the first detection from a new obstacle?
- Is this detection a false alarm?

A list of detections from one obstacle in time is called a track from that obstacle. The added value of a tracker is that previous detections (history) are used to follow an obstacle. With use of the history the location of the object can be predicted. Even when an obstacle is not detected in a single measurement, the obstacle is still tracked since it can't disappear at once.

Both sensors measure range; this is used to add a new detection from one sensor to a track started with detections from the other sensor. The list with current tracks is searched for a track with the same range as the new detection. If such a track is found, the new detection can be added to the current track.

After adding a detection from the other sensor than the sensor that started the track, not only the parameters of an obstacle measured by the first sensor are known, but also the parameters measured by the other sensor. For example a track that is started with radar detections only contains range and velocity information. When a detection from the camera system is added to the track, the angle of the obstacle is known also and the track contains range, velocity and angle information.

First results with sensor fusion and tracking show that the track continuity improves and the number of false tracks decreases.

CONCLUSIONS

We have presented an obstacle detection system for people movers based on vision and radar. The system gives separate detection of obstacles for stereo vision and for radar. These separate detections are fused in the sensor-fusion module that also performs the tracking of the obstacles. The system has been tested with good results in an autonomous test vehicle (RoboJeep).

REFERENCES

(1) Albus, J.S., “A Reference Model Architecture for Intelligent Systems Design”, Intelligent Systems Division, Manufacturing Engineering Laboratory, National Institute of Standards and Technology, September 1994

(2) T.W. van den Berg, W. Huiskamp, J.C. van den Heuvel, “Unmanned Vehicle Control using Simulation and Virtual Reality Techniques”, IEEE Intelligent Transportation Systems Conference 2001, pp. 895-900, Oakland, USA, August 25-29, 2001.

(3) W. van der Mark, F.C.A. Groen, J.C. van den Heuvel, “Stereo based navigation in unstructured environments”, IEEE Instrumentation and Measurement Technology Conference 2001, pp. 2038-2043, Budapest, Hungary, May 21-23, 2001.

(4) O. Faugeras, “Three-Dimensional Computer Vision: A Geometric Viewpoint”, MIT Press, 1993.

(5) Z. Zhang, “Flexible Camera Calibration by Viewing a Plane from Unknown Orientations”, International Conference on Computer Vision, 1999.

(6) Luigi Di Stefano, Stefano Mattoccia , “Fast Stereo Matching for the VIDET System using a General Purpose Processor with Multimedia Extensions”, Fifth IEEE International Workshop on Computer Architectures for Machine Perception International (CAMP'00) Padova, Italy, September 11 - 13, 2000.

(7) K. Mühlmann and D. Maier and J. Hesser and R. Männer, “Calculating Dense Disparity Maps from Color Stereo Images, an Efficient Implementation” In: CVPR 2001. S., Kauai Marriott, Hawaii, 2001.

(8) A. Talukder, R. Manduchi, A. Rankin, L. Matthies, “Fast and Reliable Obstacle Detection and Segmentation for Cross-country Navigation”, IEEE Intelligent Vehicle Symposium, Versailles, France, June 18-21, 2002.