Avoiding Forward Car Collision using Stereo Vision System

T. Surgailis, A. Valinevicius, V. Markevicius, D. Navikas, D. Andriukaitis
Department of Electronics Engineering, Kaunas University of Technology,
Studentu St. 50, LT-51368 Kaunas, Lithuania
algimantas.valinevicius@ktu.lt

***Abstract*—This paper presents a model of stereo vision system for intelligent vehicles to avoid forward car collision. Approached vehicle extraction and distance estimation algorithm based on dense edge map calculation and ROI filter. Primary experiments have been executed modeling vision system in a virtual computer reality. Special developed software has been used to simulate virtual environment and vision system. Results shows that approached algorithm successfully extracted not only vehicles but and other potential obstacles, such as, pedestrians.**

***Index Terms*—automotive electronics, vehicle detection, vehicle driving, object recognition.**

I. INTRODUCTION

Car accidents are the key problem of today's traffic world. Particularly this strongly felt in urban areas where cars density is largest. According to Lithuanian 2011 year car accidents statistic most common type of car accidents are collisions. It represents 44% of total car accidents. Accidents happened with pedestrians and obstacles also could be assigning to similar car accidents type. By summing of everything, we are getting 82% of total accidents [1]. Most of car accidents happened due to human factor: somnolence, inattention, absence of mind and slow reaction. All these factors are critical for a driving safety. Using stereo vision system is a way to reduce number of car accidence. Stereo vision based traffic observation systems are in active development period. Scientist and many famous car manufacturers are developing durable, accurate and real time car driver assistance systems. Many of works done by researching and developing vision system for lane tracking [2], pedestrian or bicyclists detection [3], car detection [4] or driver monitoring [5].

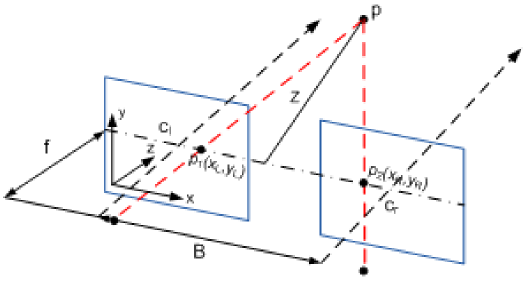

Fig. 1. Typical structure of stereo vision [6].

Manuscript received February 18, 2012; accepted May 25, 2012.

The main purpose of using stereo vision system is capability to transform 2D view into 3D information usually called depth map. Typical visual system consists of two digital video cameras which are displaced horizontally from one another to obtain different views of a scene, in a manner similar are used to human binocular vision (Fig. 1). 3D information derived by finding the corresponding points across multiple cameras images. Correspondences lie to the same epipolar line. To simplify the matching process, camera images are often rectified. The result is epipolar line corresponds on image scan line. Before image rectification system has to meet these requirements: cameras image planes are parallel, focal points height and lengths are the same. To get a depth map a disparity value must be calculated for every image pixel. A disparity value for a target point *p* calculated using following expression

$$d(x,y)=p_1(x,y)-p_2(x,y). \quad (1)$$

Disparity value estimation directly related to main stereo vision problem – correspondences. Local and global methods are used to find correspondences. Local methods usually use less computation resources and are much faster while global opposite.

A depth map estimation based on a triangulation. Given camera geometry with a focal length *f* and the distance *B* between two cameras are main system parameters. Depth map can be computed using

$$Z(x,y)=fB/d(x,y). \quad (2)$$

Depth map is a main 3D data source for car extraction and distance estimation.

II. STEREO VISION SYSTEM ARCHITECTURE

Stereo vision system architecture consists of three data processing levels (Fig. 2):

- Low level procedures covers image capturing from digital cameras, data buffering and image rectification according to system calibration data.
- Medium level is the most critical for all vision system accuracy. Selected stereo matching algorithms and estimated 3D information strongly influence higher level results. Today one of the most advanced stereo matching algorithms is based on segmentation, adaptive belief or reliable propagation [6]–[8].

– High level is the last processing stage where depth and image data are used for object extraction and recognition. Further object list is used for front vehicle tracking, distance or velocity estimation. Calculated data are used for car indication or car control.

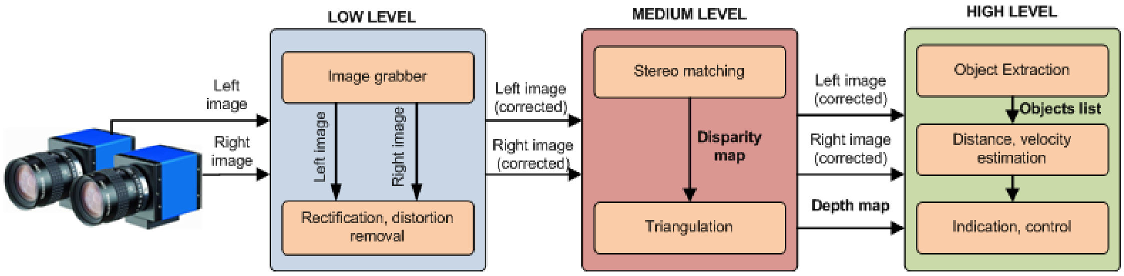

Fig. 2. Stereo vision system architecture consist of three data processing levels.

III. CAR EXTRACTION APPROACH

Suggested approach based on dense edge map calculation and ROI filter. Fig. 3 illustrates a block diagram of vision system data processing. Totally eight steps sequence to extract and mark car from 2D image set. Car extraction process starts from capturing of view images. Next, captured 2D images processed in the middle data processing level. Here 2D data transformed into 3D information, well known as depth map. ROI filter eliminates redundant 3D information. The rest of data is processed by dense edge map estimation algorithm.

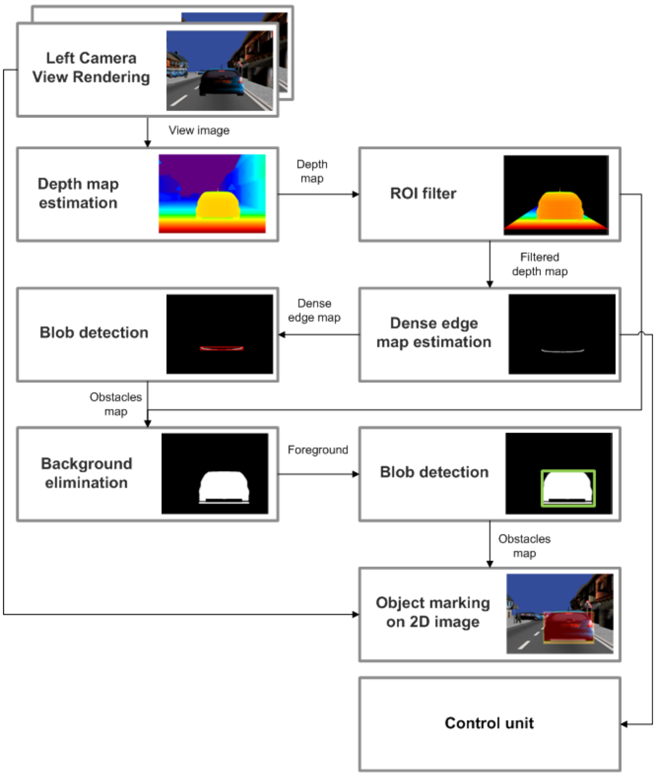

Fig. 3. Vision system's data processing block diagram.

The result of algorithm is top view of depth map with highlighted object edges. Edges detected using binary large object method. After this process we have all useful information about object position, height, width and distance to it. These results directly can be use for vehicle control or representation. To represent detected object on 2D source image, three more steps sequence. Source view background eliminated using depth map and dense edge map data. Blob bounds detected using same binary large object search method. These bounds further used for object marking on 2D view image.

IV. ROI FILTER

Region of Interest filter is used to eliminate redundant 3D information. ROI can be represented as a frustum in a 3D environment (Fig. 4). According to fact that system designed to observe front vehicles a frustum is projected on a roadway and covers one or more drive lines. All data outside this space are eliminated leaving only significant information. The main benefit of applying ROI filter is saved extra computation resources for further object extraction processes.

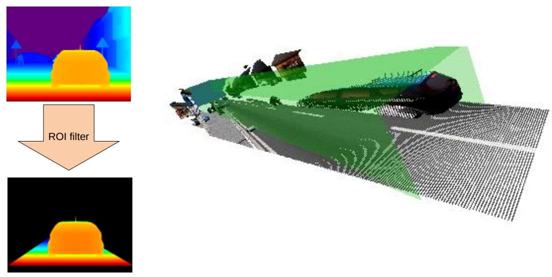

Fig. 4. Illustration of region of interest filtering.

V. EDGE DENSE MAP ESTIMATION

Edge dense map estimation basically is a depth maps transformation from the front view to the top view perspective. A value of every new map element calculated using following expression:

$$B(i,j)=\sum_{n=0}^{h-1} p(i,j,n), \tag{3}$$

where

$$p(i,j,n)=\begin{cases}1, & j=\dfrac{A(i,n)-\min(A)}{c},\\ 0, & otherwise,\end{cases} \tag{4}$$

where A is 3D data array; B is dense edge data array, $i\in\{0,1..h\text{-}1\}$, h is image height, $j\in\{0,1..w\text{-}1\}$, w is image width, c is a factor calculated

$$c=\frac{\max(\mathrm{A})-\min(\mathrm{A})}{\mathrm{h}}. \tag{5}$$

This transformation is modified following the idea to analyze not whole object but only their edges. The result of this transformation is an edge's map in a perspective of top view. Fig. 5 depicts the transformation principle in a 3D environment. The most density of points is located on object's edges. Edges are highlighted by summing points in a vertical direction. The more object's edge is higher and sharper the more points will be counted in a vertical direction. Flat and low altitude objects have a small amount of points; therefore, objects like a road can be easily eliminated using simple threshold.

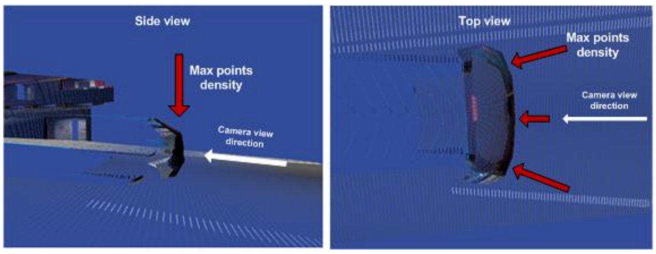

Fig. 5. Illustration of 3D object dense edge estimation principle.

VI. System Testing Platform

A virtual 3D computer reality was chosen as very first vision system and approached car extraction algorithm test environment. Specially designed software allows simulating stereo vision system behavior in a virtual 3D reality. Software based on OpenGL and OpenCV programming libraries and is written in a C++ language. Software runs on Core 2 Duo 2.53GHz CPU and 4 GB memory of RAM. Similar kind of test method has been used by other authors [9] and [10] to test stereo vision systems for relative applications. In order that tested situations would be more realistic, a segment of street with all following infrastructure: cars, road signs, buildings, pedestrians were designed and rendered in a virtual 3D reality. In this test stage a depth map was generated using OpenGL graphic API methods. This allows getting the best depth map quality. Depth map can be estimated and using any kind of stereo matching algorithm. Simulation software allows imitating stereo camera view and capture left and right camera images. Estimated depth map can be represented as a point cloud converting 3D data to real word coordinates. Different traffic situations, vision systems configurations can be emulated and repeated many times. This is the main advantage of described simulation software.

VII. Experiment Results

As was mention before, approached car extraction algorithm was tested simulating different traffic situations in a virtual 3D reality. Some situations examples presented in Fig. 8. Left image depicts a camera view and marked obstacle on the road. Right image is a representation of a dense edge map. Situation no. 1 represents a traffic circumstance then forward car is far away from driving car. Vision system detected potential obstacle and marked with a green rectangle. Green color indicates that vehicle is in safe distance. Objects away from a camera represented as small blobs in a dense edge map. Blob position in a vertical direction determined by distance from camera to object. Greater object distance leads greater blob distance from map bottom. The opposite situation presented in the situation no. 2. Detected car marked with a red colored rectangle. Red color indicates that car is in the unsafe distance and driver must pay attention. Close objects represent as big blobs located near to dense edge map bottom. The rest two situations (no. 3, 4) shows system ability detect not only forward cars but and different kind of obstacles, for instance pedestrians crossing the street. Another aspect of vision system is ability to detect multi numbers of obstacles.

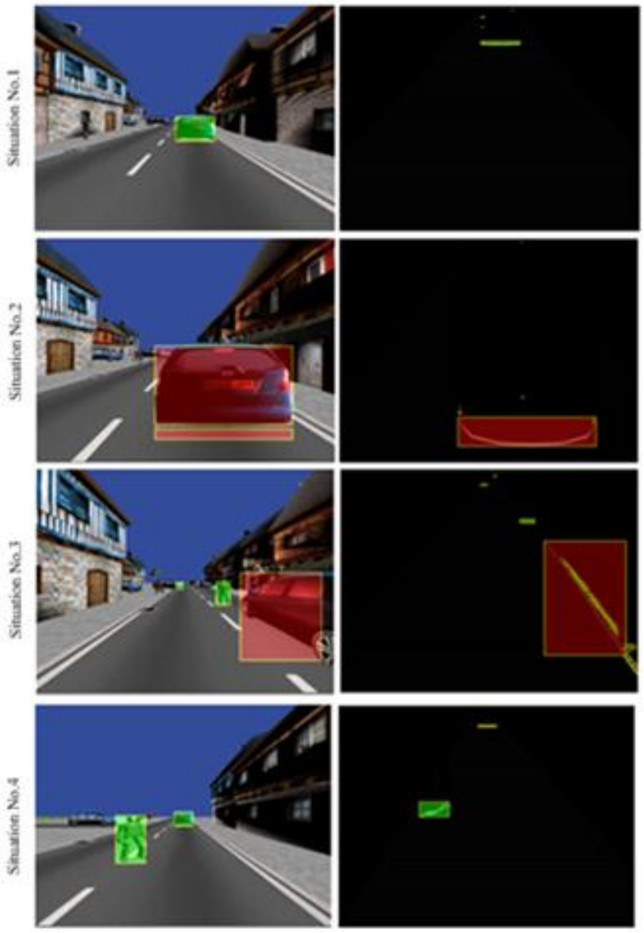

Fig. 6. Four traffic situation examples simulated using virtual 3D reality. Left image – camera view, right image – edge dense map.

VIII. Conclusions

After range of generated different traffic situations, testing results shows a good approached car extraction algorithm ability to detect objects and measure distance to it on the road. Vision system was able to detect not only massive object such as vehicle, but and less massive, such as pedestrians.

To reduce computational recourses ROI filter was used to eliminate redundant 3D data. This filter actually determines system analysis space. To simplify object extraction and safe calculation time, edge dense map is used. Here object examined only as dense edge curve. The average time of object extraction and marking took ~125 ms. This was achieved using Core 2 Duo 2.54 GHz processor and 4GB memory of RAM. To apply vision system for cars need to reach at least 30 ms for all data processing. This can be done using optimized code, multi threading, and new generation or GPU processors.

To use approached vision system model for avoidance of front car collision need to execute more researches, testing model with a real stereo vision prototype.

REFERENCES

[1] *The Lithuanian Road Administration under the Ministry of Transport and Communications,* [Online]. Available: http://www.lra.lt/lt.php/eismo_saugumas/eismo_ivykiu_statistika/27

[2] R. Danescu, S. Nedevschi, "New Results in Stereovision Based Lane Tracking", IEEE, pp. 230–235, 2011.

[3] S. Bota, S. Nedesvchi, "Multi-Feature Walking Pedestrian Detection Using", *Intelligent Transport Systems,* IET, vol. 2, no. 2, pp. 92–104, 2008. [Online]. Available: http://dx.doi.org/10.1049/iet-its:20070039

[4] T. Kowsari, S. S. Beauchemin, J. Cho, Real-time vehicle detection and tracking using stereo vision and multi-view AdaBoost // in *Proc. of 14th International IEEE Conference on Intelligent Transportation Systems (ITSC),* 2011, pp. 1255–1260.

[5] L. M. Bergasa, J. Nuevo, M. A. Sotelo. "Real-Time System for Monitoring Driver Vigilance", *Intelligent Transportation Systems,* IEEE, vol. 7, no. 1, pp. 63–77, 2006. [Online]. Available: http://dx.doi.org/10.1109/TITS.2006.869598

[6] Z. Wang, Z. Zheng, "A Region Based Stereo Matching Algorithm Using Cooperative Optimization", in *Proc. of CVPR,* IEEE, 2008, pp. 1–8.

[7] A. Klaus, M. Sormann, K. Karner, "Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure", *Pattern Recognition,* IEEE, - no. 3, pp. 15–18, 2006.

[8] X. Sun, X. Mei, S. Jiao, "Stereo Matching with Reliable Disparity Propagation", in *Proc. of International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission*, IEEE, 2011 pp. 132–139.

[9] C. H. Messom, "Stereo Vision Controlled Humanoid Robot Tool-kit", *CiteSeerX,* 2009.

[10] M. Chessa, F. Solari, S. P. Sabatini, "A virtual reality simulator for active stereo vision", *CiteSeerX,* 2009.