

A Neural Network Enhanced Stereo Vision Obstacle Detection and Avoidance System for Unmanned Ground Vehicle

FanJun Liu

School of Mechanical Engineering, Xi'an Jiaotong
University, Xi'an China
e-mail: liufj515@163.com

Binggang Cao

School of Mechanical Engineering, Xi'an Jiaotong
University, Xi'an China
e-mail: liufj@yeah.net

Abstract—This paper presents a neural network enhanced stereo vision obstacle detection and avoidance system for unmanned ground vehicle. We build a neural network to learn the mapping for the left image to the right image under the assumption of a flat road. Using the trained neural network we map the left image to the right directly and we get the left remapped image. So obstacles can be detected using correlation values between the right image and the left remapped image. With detection result the system tells the unmanned vehicle how to avoid obstacles. Our system does not require intrinsic calibration of stereo cameras and it does not perform the two IPM transforms. With neural network's parallel processing our system reduces the computation expense and increases the real-time performance. Experimental results show that our system is practicable.

Keywords- neural network; stereo vision; obstacle detection; unmanned ground vehicle.

I. INTRODUCTION

Obstacle detection and avoidance play the important role in an autonomous driving vehicle. Obstacle detection means to extract areas that the intelligent vehicle cannot or should not be traversed by. obstacle areas such as rocks, fences, trees, and steep slopes are examples. Commonly used obstacle detection methods include vision-based methods, radar-based methods and multi-sensor fusion methods. Vision based methods can be mainly divided into two classes: one is based on monocular vision, such as optical flow method [1,2], the other is based on stereo vision which generates depth estimations from two or more stereo images [3~6].

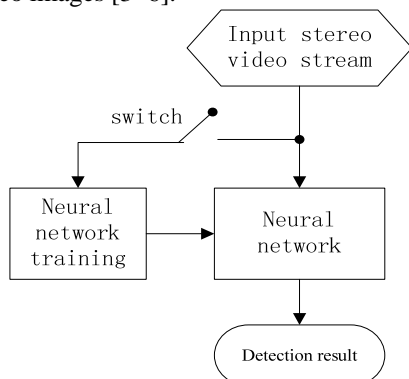


Figure 1. The diagram of our system

This paper presents an artificial neural network enhanced stereo vision obstacle detection and avoidance system for unmanned ground vehicle. The diagram of the main components is shown in Figure 1. We can decompose the system into two parts: the neural network training and the obstacles detection and avoidance. Accordingly those will be described in Section 2 and Section 3 of this paper. In Section 2, we describe the difference of our method and IPM obstacle detection method firstly, then we built a neural network to learn the mapping from the left image to the right image secondly, and we give the mapping result and the different image lastly. In Section 3, we use the correlation coefficient between the left remapped image and the right image to determine obstacle and give the experimental results.

II. IPM AND NEURAL NETWORK

In this section, we describe the difference of our method and IPM obstacle detection method firstly, then we built a neural network to learn the mapping from the left image to the right image secondly, and we give the mapping result and the different image lastly.

A. The application of IPM in obstacle detection

Because a complete three-dimensional reconstruction is very time consuming, the application of IPM (inverse perspective projection)[7] to stereo images plays a key role especially in obstacle detection. Much of the research described in this thesis depends on projective geometry [8]. In the existing literature [9~12], the road is usually assumed to be flat. Given the calibration parameters of the camera, both left and right images can be changed using the inverse perspective transformation. In [9], the authors mapped left and right images onto the same coordinates using the inverse perspective projection. Image points on the road produced no disparities, i.e., they had the same coordinates in both transformed images, whereas points on obstacles produced large disparities. Thus, obstacles can be detected by subtracting one image from the other. The IPM detection method in [9] is shown in Figure 2.

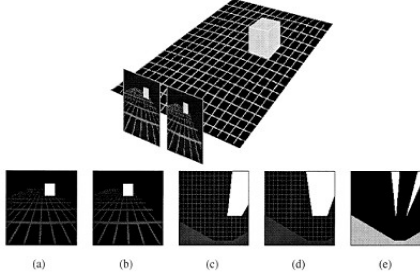


Figure 2. Diagram of the obstacle :(a) left image,(b)right image,(c) remapped image to (a) ,(d) remapped image to (b), (e)difference image between (c) and (d)

Although these IPM obstacle detection methods are simple and have been applied in the ARGO autonomous vehicle, they are not very robust :(1)If we use IPM transform, we must know the specific acquisition conditions (optics, orientation, camera position, etc.)[13] and the camera must be mounted in a fixed position, so other kind of sensors can sense the calibration of the system and the surrounding environment[14]. (2) the calibration parameters must remain unchanged while the vehicle is moving and distortions of the two cameras cannot be too large. Small sensor position changes can result in false indications of obstacles. (3)They are no suitable for high speed applications because of the computation expense with the IPM transforms.

In our system, obstacle detection method is much similar to the IPM method. Under the assumption of a flat road, we map the left image to the right using a neural network directly and we get the left remapped image. If there is no obstacle, the left remapped image should be same with the right image. Obstacles can be detected using correlation values between the right image and the left remapped image. So our method does not require the camera calibration parameters and it does not perform the two IPM transforms. It is faster and more robust then IPM detection method.

B. Construction and training of the neural network

In this section we will build a neural network and let it learn the mapping from the left image to the right image. The main work of this section is shown in Figure 3.

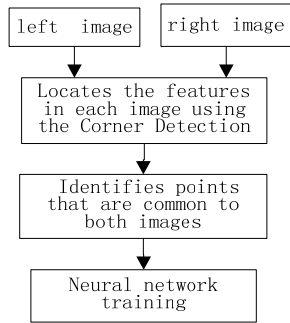


Figure 3. The neural network training

I) The network structure

In our application, we build a neural network with two-layer structure: the hidden layer and the output layer. The network structure is illustrated in Figure 4. The coordinates of pixels in the left image are used as the neural network's input, and the coordinates of pixels in the right image are used as the output. Thus, the network is a feed-forward neural network (FFNN) with two inputs and two outputs.

The transfer function for hidden layer is the symmetrically tangent hyperbolic sigmoid-function. The output layer transfer function is pure linear. In particular, more number of hidden neurons does not correlate with more correct classification rate. Experiments prove that one hidden layer with four neurons is enough to our system.

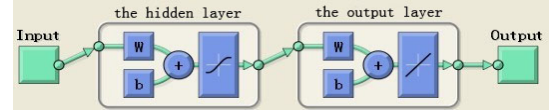


Figure 4. The network structure for our system

The graphs and symbols of linear and hyperbolic tangent sigmoid transfer functions are shown as Figure 5.

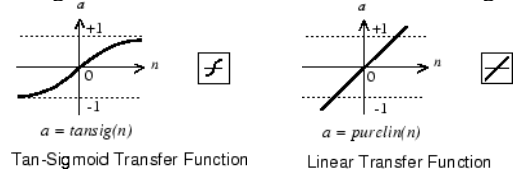


Figure 5. The input function and the output function

II) Getting the training-dataset and training

We get a pair of stereo images from the video stream: the left image and the right image. Both images must be unlikely to contain obstacle in the region of interest (ROI) Then we can get the training dataset by two steps:(1)Detect the Corners in each image. (2) Identify points which are in both images using Sum of Squared Differences (SSD) algorithm [15][16]. For more details of getting matching features points we can refer to "Matlab(7.8.0 R2009a) demos/Blocksets/Video and Image Processing/Registration/Disparity Estimation for Stereo Vision"

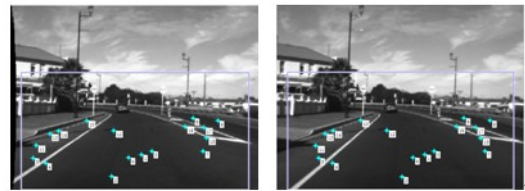


Figure 6. The matched method points in ROI

Figure 6 shows the matched feature points in ROI in the left and right images. We use the coordinates of matched feature points as the neural network training-dataset. All testing images were obtained from the Multimedia Imaging Technology portal hosted by the University of Auckland.

As we know, there are many neural network training algorithms, such as Back Propagation, Resilient Propagation and Quick Propagation. To our system these algorithms have little effect on the quality of the results of the training,. but there is much difference in the time

required for training. Experiments indicate the Back Propagation algorithm is better to our system. So we selected back propagation the neural network training algorithm.. The training-dataset is shown in Table 1.

TABLE 1: THE NEURAL NETWORK TRAINING-DATASET

Input vector: the coordinates of feature points in the left image																	
Horizontal coordinates (u)	383	324	408	143	108	353	539	553	485	216	101	285	125	146	528	465	515
Vertical coordinates (v)	388	446	380	410	398	397	380	301	294	300	365	326	335	327	347	316	319
Target vector: the coordinates of feature points in the right image																	
Horizontal coordinates (u)	337	260	364	90	58	304	496	534	470	196	66	266	101	124	505	450	498
Vertical coordinates (v)	389	446	380	412	398	397	380	300	294	300	365	326	335	328	346	316	319

In fact, the neural network training process is the process of estimating geometry and intrinsic calibration of stereo cameras. But all are done by the neural network.

C. The Mapping from the left image to the right image

We get a pair of images for the stereo videos. Figure 7 shows the left image and the right images. We use the trained neural network to map the left image to the right, and we get the left remapped image. We get a pixel-wise difference between the left remapped image and the right image Figures 8 shows the left remapped image and the difference images. Figure 8 also shows that the vehicles and people above the ground are apparent, while their shadows, which we do not consider to be obstacles, have been removed.

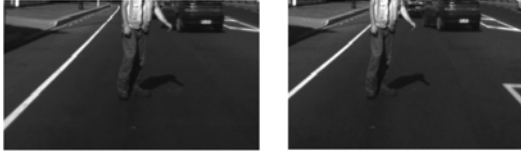


Figure 7. The left and right images



Figure 8. The left remapped image and the difference image

III. OBSTACLE DETECTION AND AVOIDANCE

From the difference image, we can see the obstacles. But this is not the terminal goal. Our goal is to tell the vehicle whether it can go through. In this section, we use the correlation coefficient between the left remapped image and the right image to determine obstacle and give the experimental results.

A. Determination of obstacles

We can determine whether there are obstacles using normalized product correlation (NProd) between the left remapped image and the actual right image. Because

NProd is insensitive to differences in lighting conditions and weather conditions[17].

The normalized product correlation (NProd) is computed as follows.

Suppose that images X , Y , of size $m \times n$ ($X(i, j)$, $Y(i, j)$, $1 \leq i \leq m$, $1 \leq j \leq n$ are points in X , Y), then the normalized product correlation (NProd) algorithm can be expressed by Formula (1):

$$R = \frac{\sum_{i=1}^m \sum_{j=1}^n X(i, j)Y(i, j)}{\left[\sum_{i=1}^m \sum_{j=1}^n X^2(i, j) \right]^{1/2} \left[\sum_{i=1}^m \sum_{j=1}^n Y^2(i, j) \right]^{1/2}} \quad (1)$$

Supposing that X is the left remapped image and Y is the actual right image, then as the value of R increases, the assumption of the road being flat is correct to a greater degree, and the possibility of obstacles being present therefore decreases. So we can set a threshold T . If $R > T$ we can say there is no obstacle in ROI. Experimental results show that the threshold can be, for example, $T=0.95$. If $R < T$, then obstacles are present.

B. Image sub-regions in ROI

To determine the orientations of obstacles, we can separate ROI into 3 sub-regions: Left-front Front and Right-front. This is shown in Figure 9. If the obstacle is located in the Left-front sub-region, then the unmanned vehicle turns right; if the obstacle is located in the Right-front sub-region, it turns left; if the obstacle the obstacle is located on the Front, then it can stop or turn left or right. By this rule, the unmanned vehicle will be able to avoid obstacles.



Figure 9. Image sub-regions in ROI

C. Experimental results

The simulation is implemented by applying this detection method to 80 pairs of images of stereo sequences that were recorded on the same road. The stereo sequences are obtained from the Multimedia Imaging Technology portal hosted by The University of Auckland.

In Figure 10, the horizontal axis represents the image sequence number, and the vertical axis represents the value of R . Figure 10 shows the results in each of the 3 image regions. The blue, green, red lines track the R values of Left-front, Front and Right-front sub-regions in ROI respectively. Here the $T=0.95$, we can see from Figure 10 that, obstacles are detected in Left-front sub-region. in the 34~54th and 70~74th frames, Front sub-region obstacles are in the 35~40th frames and Right-front obstacles are in the 32~37th frames. Figure 7 also demonstrates that obstacles appear and disappear in our ROI.

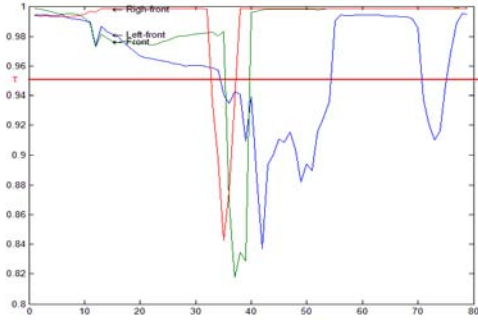


Figure 10. The correlation result of 3 regions in ROI

IV. CONCLUSION

Obstacle detection and avoidance using stereo vision is a promising approach to unmanned ground vehicle navigation. In this paper, we propose a neural network enhanced stereo vision obstacle detection and avoidance system for unmanned ground vehicle. We use an artificial neural network to simulate the perspective mapping geometrical transform of left image to the right image. Our system is easily trained by simply driving the vehicle through the environment without obstacles. It can also keep learning the appearance of the ground during operation. Once the network trained well the system can work intelligently to detect the obstacle. Our system does not require intrinsic calibration of stereo cameras. With neural network's parallel processing and without IPMs our system reduces the computation expense and increases the real-time performance. Our algorithm has been tested successfully in Xi'an Jiaotong University's SpringRobot system[18].

ACKNOWLEDGEMENTS

The research is supported by a grant from National Natural Science Foundation of China (No. 11072183, No.

21106113) and the research fund of State Key Laboratory of Automobile Safety and Energy Conservation (No. KF10092) and a grant from Jiangsu Provincial Natural Science Foundation of China (No. BK2009144) and a grant from Suzhou Research Program of Application Foundation(SYG201138). The authors would like to express their thanks for these funds.

REFERENCES

- [1] J. L. Barron, D. J. Fleet and S. S. Beauchemin, Performance of optical flow techniques, International Journal of Computer Vision, pp. 34-77, 1994.
- [2] Yunfang Zhu, Yishu Wang, Optical flow based obstacle detection in static environment Journal of Zhejiang University(Engineering Science) 2008-06
- [3] G.-W. Zhao and S. Yuta, Obstacle detection by vision system for an autonomous vehicle, In Intelligent Vehicles '99 Symposium, Tokyo, Japan, July 14-16, 1993, pp. 31-36, (1993).
- [4] A. Broggi, C. Caraffi, R. Fedriga, and P. Grisleri, Obstacle detection with stereo vision for off-road vehicle navigation, Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on, vol. 3, no. 65- 65, pp. 20-26, June 2005.
- [5] Zhu Zhigang, et al., 1999. A Real-Time Visual Obstacle Detection System Based on Reprojection Transformation, J. Computer Research and Development, 1, 36, 77-84.
- [6] M. Bertozzi and A. Broggi, "GOLD: A Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection," IEEE Transactions on Image Processing, Vol. 7, No. 1, January 1998, pp. 62-81.
- [7] H.A. Mallot, H.H. Bu'lhoff, J.J. Little, S. Bohrer, Inverse perspective mapping simplifies optical flow computation and obstacle detection.
- [8] O. Faugeras. Three-Dimension Computer Vision. MIT Press, 1993.
- [9] Massimo Bertozzi, Alberto Broggi, A. F., 1998. Stereo inverse perspective mapping: theory and applications. Image and Vision Computing 16(1998)585-590.
- [10] Hebert, M. and Ponce, J., A New Method for Segmenting 3-D Scenes into Primitives, 6th International Conference on Pattern Recognition, pp836-838, Oct. 1982.
- [11] Yi-peng Zhou, Fast and Robust Stereo Vision Algorithm for Obstacle Detection, Journal of Bionic Engineering 5(2008)247-252
- [12] Yang, H.S. and Boyer, K.L. and Kak, A.C., Range Data Extraction and Interpretation by Structural Light, pp 199-205, Conference on Artificial Intelligence Applications, 1984.)
- [13] D.A. Pomerleau. RALPH: rapidly adapting lateral position handler, in: I. Masaky (Ed.), Proceedings of IEEE Intelligent Vehicles '95, Detroit, IEEE Computer Society, 1995, pp. 506-511.
- [14] K. Storjohann, T. Zielke, H.A. Mallot, W. von Seelen. Visual obstacle detection for automatically guided vehicle, in: Proceedings of the IEEE International Conference on Robotics and Automation, vol.11, 1990, pp. 761-766.
- [15] Trucco, E; Verri, A; Introductory Techniques for 3-D Computer Vision (1998) Prentice Hall.
- [16] Hannah M J; Computer Matching of Areas on Stereo Images, Ph.D Thesis, Stanford University (1974)
- [17] Liu Fanjun, Cao Binggang, 2009. Color Image Similarity Measurement and Its Application in Matching. IEEE. ICICTA 2009, 566-568.
- [18] <http://www.aiar.xjtu.edu.cn/>