iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 2017 Issue: onferendeMonth of publication: September 15, 2017

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue IX, September 2017- Available at www.ijraset.com

Depth Estimation from a Single Image in a Self-Driving Car Using Neural Networks

Rashmila Mahajan[1]
[1]*Veermata Jijabai Technological Institute, University of Mumbai,India2017*

Abstract: [3] the project is based on a solution to tackle the problem of determining depth from a 2 dimensional image. The motive behind extracting depth from an image is to calculate the distance between a self- driving car and obstacles on its path. The algorithm proposes disaneconomica alternative to the existing methods. [1] It is applied individually on the frames extracted from the video recorded by the camera mounted on the bonnet of self- driving car. It involves implementation of neural network which are trained using manually collected samples, where appearing width is the input and actualdepth is the output. [2] Once trained, the neural network is used to estimate the actual depth of the obstacle. Further, the task has been performed using Polynomial curve fitting. A comparison has been done between the results obtained using the neural networks and Polynomial curve fitting. It has been observed that neural network gives more

accurate results. [1] Moreover, the efficiency of the proposed algorithm has been verified by computing the error value between the estimated and the actual depth. The error value lies in a very low range which proves the successful working of the algorithm.

I. INTRODUCTION

[7]In case of a two dimensional image, it is not possible to estimate the real depth of an object. The depth is interpreted as the distance between the camera and object. In a self-driving car, a similar problem has been observed where the car wasn't able to estimate the real depth of an obstacle. Current methods involve usage of sensors like LIDAR and ultrasonic, which are not economic. [5]As the number of sensors needed to be installed increase, the cost and the overall weight increases. This leads to more drainage in power of battery. To tackle these problems, a cheaper and efficient alternative is required. The depth is extracted from a 2D image by using relative size method. His involves a camera mounted on the bonnet which records the video of path taken by the car indicating the obstacles. Frames extracted from this video are considered as individual images, required to estimate the depth of an obstacle using the proposed algorithm. Database for training samples is created using images of varying obstacles clicked from different distances. While collecting training samples, a single object is moved back and forth along a straight line, perpendicular to the camera. For every single image, appearing width is calculated by finding difference between the two extreme lower edge coordinates of the obstacle and its corresponding depth is included in database. Using these samples, a neural network is trained considering a single feature as input which is the appearing width, for which actual depth is expected as output. The neural network consists of a single hidden layer. A curve is plotted with abscissa as appearing width, calculated using sample images and ordinate as actual depth. An equivalent quadratic equation is obtained which describes the relation between output and input.

II. WORKING LOGICANDFINAL TESTING

Relationship between appearing width and depth is exploited. The main observation is the reduction in appearing width as actual depth increases and the increment in appearing width as actual depth increases, which forms the basis of the algorithm. A test sample is given to neural network to obtain the actual depth of the desired obstacle. This test sample is also given as input to the Polynomial curve fitting equation in an attempt to get the actual depth. Finally the results are compared to find the more efficient algorithm.

III. METHODOLOGY

The project uses input from a collection of 130images. The input data in this project are single images, i.e. any particular scene has been captured with a single image. All images are having objects at different distances and so their sizes are varying and all images are unique.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue IX, September 2017- Available at www.ijraset.com

Different sized objects(ranging from width 6cmto46cm)were used and a set of pictures for every object kept at different distances from camera clicked (range 20cm to350cm).These ranges are at the scale of small object models, and have to be scaled by 10tobesuitablefor real large-sized vehicles on road.

Depth of objects refers to the distance of objects (for example, vehicles on road, or obstructions like road blocks, humans and animal stress passing on road) from the camera.

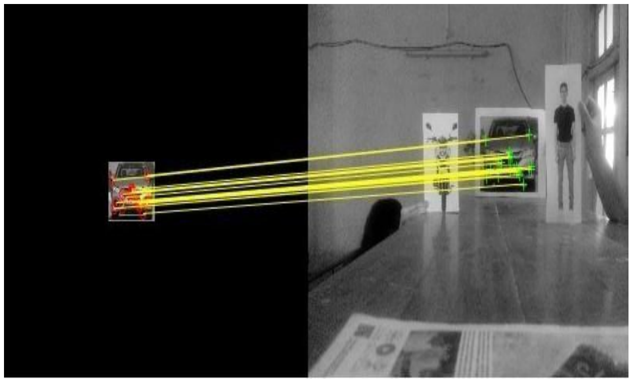

Figure1Measurement of appearing width and depth was done by considering length of the marked line on above figure.

Object detection- [5]Surf features are used to detect the object.

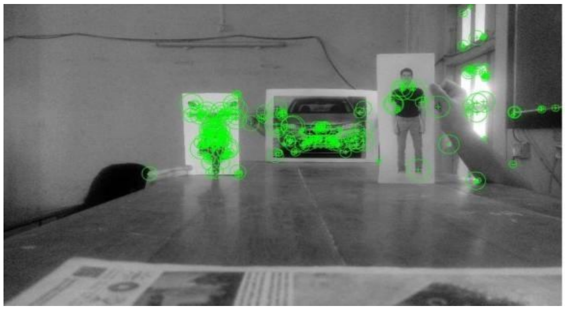

Figure2StrongestSURFfeaturesdetected from scene image

Figure 3[15] SURF features of box image of car had higher match with scene image features, as compared to the box image features of two wheeler or truck. Hence car was detected at the region with the highest matching of SURF features.

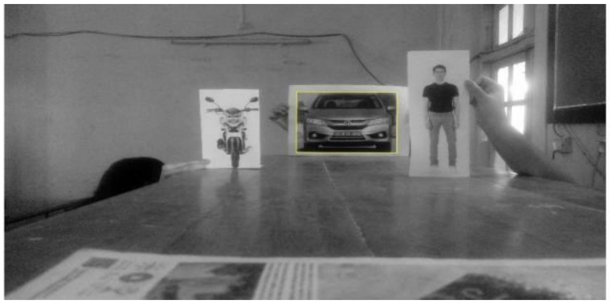

Figure4Detectedobjectis shown encircle dbya box

After detection of objects, appearing size was calculated and given as an input to the artificial neural network (to be explained in next section) and the targeted actual distance of the obstacle was output variable of the network.

A. Training Nn[14]

Artificial neural networks had to be trained using these images to create models for predicting depth of test image. As one can visualize, shifting objects away will decrease the appearing width or size of objects, so distance or depth is inversely related to appearing width. First network had input as appearing width and output as appearing depth. The second network had inputs as appearing width and the depth obtained from the first network and output as actual depth. But the inverse relation between width of object and depth from camera is not linear; it depends on the distance from camera itself. So, the same trained model could not produce acceptable results for different objects. Hence, separate network were trained for separate objects so that an identified object from a test image can be passed through the appropriate network for that category of object. Networks were trained with iterations until the results came nearly accurate and every network had different

IV. RESULTS

The appearing width for the detected object was fed to its respective neural network.

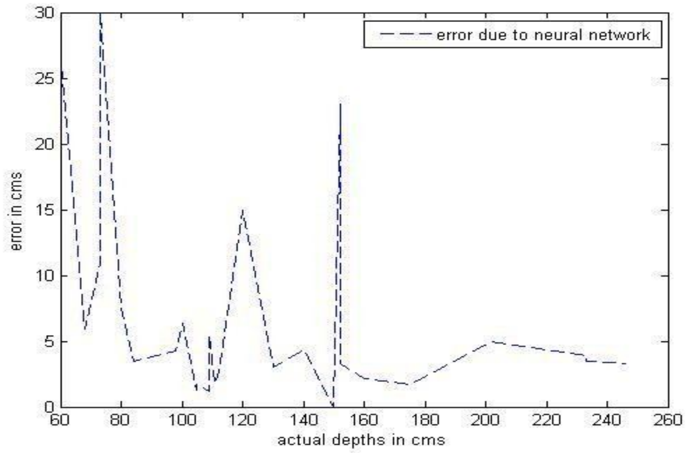

Figure5

REFERENCES

[1] Saxena, A., Sun, M. and Ng, A.Y.,2008."Make 3D:Depth Perception from a single still Image ".IEEE Transactions on Pattern Analysis and Machine Intelligence. parameters, such as number of hidden layers, or nodes, or training error.

[2] Shelhamer, E., Barron, J. and Darrell, T.,2015." Scene Intrinsic and Depth from a single image". The IEEE International Conference on Computer Vision(ICCV)Workshops ,pp. 37-44.

[3] Poyade, M., Reyes-Lecuona, A., Viciana-Abad, L., "Influence of Binocular Disparity in depth perception

[4] Mechanisms in virtual environments"., Book-New Trends on Human-Computer [4]DavidG.Lowe,2004,"Distinctive Image Features from Scale-Invariant Key points". International Journal of ComputerVision,Vol60, Issue2,pp. 91b110.

[5] Jauregi, E., Lazkano and Sierra, B., University of Basque Country, Donostia,2010,"Objectrecognition using region detection and feature extraction". Proceedings of Towards Autonomous Robotic Systems(TAROS)(ISSN:2041-6407).

[6] Cole, L. Austin, D. And Cole, L.,2004, "Visual object recognition using Template matching". In Proceedings o Australian Conference on Robotics and Automation

[7] Gavin, H., Duke University, 2017, "The Levenberg- Marquart method for nonlinear least squares curve-fitting problems". Journal of Computation and Applied Mathematics. Additional References

[8] Mortensen,E.,Deng,H.,Shapiro,L.,Interaction2009,pp13-22."RegistrationofDepthImageandColorImageBasedonHarris-SIFT"

[9] RetrievedonDecember,2016 https://en.wikipedia.org/wiki/Speeded_up_robust_features

[10] RetrievedonFebruary,2016 https://www.mathworks.in

[11] RetrievedonNovember,2016from https://www.boundless.in

[12] Harris,C. andMikeStephens,1988, "A Combined Corner and Edge Detector", Proceedings of The Fourth Alvey Vision Conference (Manchester, UK),pp. 147–151.\

[13] Retrieved onNonember,2016 https://en.wikipedia.org/wiki/Depth_perception

[14] Azad, P., As four,T.,DillamannR.,2009,"Combining Harris Interest Point and SIFT Descriptor for Fast Scale- Invariant Object Recognition", In IEEE/RST International Conference on Intelligent Robots and Systems.,StLoius,USA

[15] Simin Zhao, Xiangming pXu, WeilongZheng,JianwenLing,2012,

[16] https://ieeexplore.ieee.org/document/542