iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** V **Month of publication:** May 2018

DOI: http://doi.org/10.22214/ijraset.2018.5222

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 6 Issue V, May 2018- Available at www.ijraset.com

Depth Estimation using SVM for Image

G. Santosh[1], G. Gowri Pushpa[2]
[1,2] *Assistant Professor, Department of Computer Science and Engineering Anil Neerukonda Institute of Technology and Sciences, Sangivalasa, Visakahapatnam, India*

Abstract: ***The depth of an image is used to find out the distance from where it is taken. This is done by taking a set of images with different depths as the input and extracting Local and Global features from it. First, feature extraction based methods are utilized to evaluate image similarities. We propose a model that incorporates both of them to obtain significantly more accurate depth estimates than using either global or local properties alone. The features thus obtained are tested using support vector machines. The trained data is then used for testing to calculate the depth of any image approximately.***
Keywords: Depth estimation, local features, global features, PCA, SVM

I. INTRODUCTION

Image classification is the method of extracting the information from the image. Human being can see a scene and can easily classify the color, texture, height, distance, obstacles, whether the object in natural or manmade. By seeing the tree we can identify correctly, this is achieved by training the child from the childhood. Similarly the computer has to be trained to categories all the things. Many works has been proposed in these areas of classification. The classification are used in the area of feature extraction, data processing ,remote sensing etc. The classification can be categorized into two types, supervised learning and unsupervised learning. Supervised learning method should have a input and target output using which the data is trained .The algorithms proposed under supervised classifications[1] are probabilistic models, linear regression, Support vector machine[4], logistic regression etc. Unsupervised learning method only the input data is present using which the output data has to be obtained. Some of the unsupervised classification algorithms are k-mean and Apriori algorithm. The image classification uses these techniques for face detection, direction prediction, motion detection, Depth estimation and many more. To classify an image Internal processing are to be done in-order to classify. Few authors have used segmentation ,HOG, Transformation techniques as the preprocessing steps. Depth estimation of the image is a challenging job in today research era. Depth Estimation is the technique to identify the distance of the objects from the point of a camera which can be applied in varies applications such as for robot path tracking without obstacles if obstacle the to identify at what distance the obstacle is present. It is also necessary to identify whether the obstacle is manmade or natural. The images from the space are sent for classification in which the asteroids, meteoroids, satellite are at what distance from the earth can be predicted by the depth estimation. In recent days the Wireless sensors are used in different areas like agriculture for moisture predictions, rain forest for wild animals accountability or for border security purpose .the sensor are dropped from heights , where the human being cannot reach. To predict the distance at which the sensoris present can be predicted using depth estimation. This work is strict to area of the finding of the distance at which the sensor is present above the sea level.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue V, May 2018- Available at www.ijraset.com

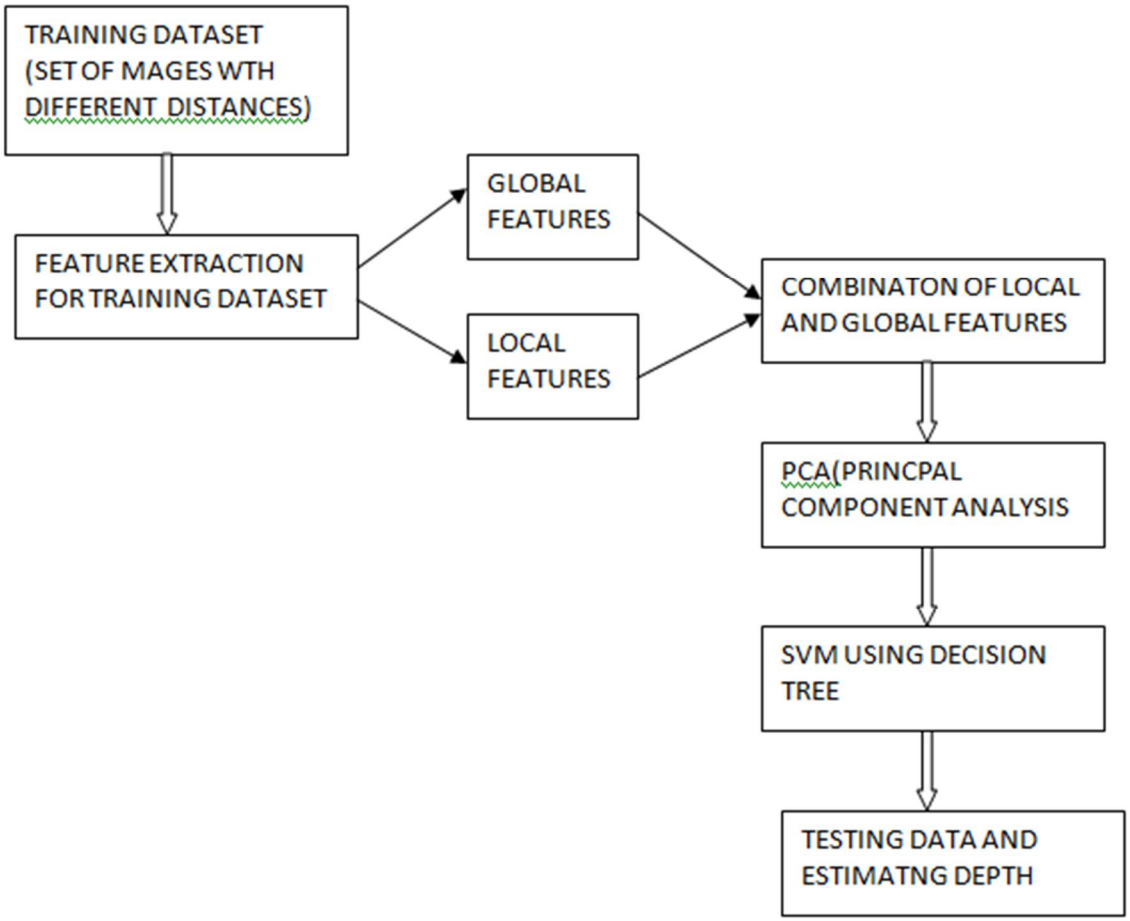


Figure 1: Flowchart of Proposed work

II. PROPOSED ALGORITHM

The work in figure 1.explains about application approach which can be modified depending on the application chosen .The current work illustrate to find the height of a building .which can be applied to find the height of the hill or the distance of car and a person . In this work a set of images with different distances are trained by supervise learning method. Varies pre-processing techniques are applied to convert the dataset to get the best training dataset. The pre-processing techniques are described in brief Section 2.1 and 2.2 describes how the global and local feature extraction help in processing the input image for training, and its necessity. Section 2.3 outlines the PCA need before the training. Section 2.4 specifies the SVM to decision tree which is applied for training and testing the dataset.

1) *Feature Extraction:* Feature are the information or list of numbers that are extracted from an image[6]. These are real valued numbers (integers, float or binary). There are a wider range of feature extraction algorithms in open computer vision. Features are the information extracted from the images, suppose we consider the image as data the information extracted from the data is known as features. Based on our application i.e., depth estimation of an image we have chosen two features first one is global feature extraction and second one is local feature extraction of an image.

a) *Global Features:* Global features are generally used in object detection and classification, detection is finding of the object (whether an object exists in image).Global feature means looking the whole image i.e., overall information present in an image. It is also known as global feature descriptors. In our project for extracting the global features we use fast Fourier Transform

b) *Fast Fourier Transform:* There are many transformation techniques like discrete Fourier, Cosine transform fast Fourier etc. Fast Fourier Transform[5] transforms a signal over a period and divides it into its frequency. These components are single sinusoidal oscillations at distinct frequencies each with their own amplitude and phase. Fast Fourier transform manages to reduce the

complexity of computing the Fourier transform from O(n^2) to O(nlogn) where n is the data size. Fast Fourier transform are widely used for many applications in engineering, science and mathematics. The fast Fourier transform is much faster when compared to ft but produces the same result as of discrete Fourier transform. The command or method for fft in python is fft.fft(image) with prerequisite NumPy. In our project the obtained matrix form after applying the command or method or function consists of the combination of amplitude and phase values, from this we take only amplitude values using the function in python abs (matrix form) (matrix form obtained from statement 4 mentioned command) and prerequisite is NumPy.

2) *Local Feature Extraction:* Local features are generally used for object recognition or identification. Recognition is finding of the object in other words focusing on something. Local features mean looking at particular part of the image and getting the details of that image at that particular area. These are also known as local descriptors. For finding out the local features for an input image, here we use windowed Fourier transform
3) *Windowed Fourier transform:* The windowed Fourier transform divides the image into window which is localized in time, used for the replacement of the Fourier transforms by the product of a sinusoidal windows. In our work we divide the entire image, here image taken as input is the resultant image obtained from the pre-processing technique. The image is then divided into 10*10 locations, each location is of size 25*25 i.e., width and height of the one location in that image. For each location we apply Fast Fourier Transform and get the absolute values of the resultant matrix for all 10*10 locations of that image. In general, the local features describe every part of the image, so when we apply Windowed Fourier Transform the details present in the image are more because the data obtained from that image is more, which indirectly states the large amount of data.

B. PCA (principal component analysis)

The PCA is a supervised learning method, analysis the data to identify patterns and reduces the dimensions of the dataset with minimal loss of information so it is also known as dimensionality reduction algorithm which reduces n dimensional to some d dimensional form for example if n=3 then d can be of any value either 2 or 1.Here in our work after feature extraction we reduce the output data which is in the form of 2 dimensional is converted into 1-Dimensional form by applying PCA algorithm.The PCA is mostly used so that we can reduce the computation part for our image. The PCA algorithm follows the following steps.

1) *Algorithm*

a) *Step 1:* Taking the whole dataset: Take the whole dataset i.e., feature values obtained from feature extraction part, consisting of n-dimensional (n=2) samples ignoring the class labels.
b) Note: Here the value of n is 2 then the resultant of PCA will be in the form of 1-D matrix.
c) *Step 2:* Compute n-Dimensional mean vectorCompute the n-Dimensional mean vector for the whole dataset taken in step1
d) The function or command used for calculating mean vector in python is mean (matrix, axis=0), the prerequisite is NumPy.

D. SVM (support vector machines)

A supervised machine learning algorithm which is used for classification is Support Vector Machine. Thefigure plot each data item of n feature as a point in n-dimensional space. Then, the hyper plane classifies the two classes. Hyper plane with minimum distance decides to which class the data belongs . he advantage of using SVM is it can easily classifies the data which can be separated by margin. The support vector which uses the subset of training points as decision function, which helps in increasing the memory efficiency. The drawback of SVM is cannot be applied for large dataset as training time is high and cannot be applied for overlapping dataset. Taking the advantage of SVM for classification into consideration. The hue dataset after applying Global and local feature to the images PCA is applied in order to reduce the dataset so that the drawback of SVM can overcome. After extracting features by applying PCA (principle component analysis) the dimension is reduced i.e., the result of global and local was 2-D and the resultant matrix is reduced to 1-D using PCA. After extracting the features and applying PCA all the features are combined int one dataset and finally gets a data for all folders and which is store in .csv file and also their classes are stored in separate folder. Using this training dataset storage we test the images present in testing folder using SVM support vector machines algorithm by fitting them using the fit function present in python except classes value and then predict the exact class representing the depth that it can fall in any of the four folders and calculate the result approximately.

III. EXPERIMENTAL RESULTS

The experimental results in this section shows the detail outputs of global and local features of trained images . a total of 1000 iages are categorized into 4 clusters, It produces a huge data size to which the PCA is applied to reduce the dimensions and then the

SVM is applied to the dataset . The comparative study shows the application of SVM for depth estimation is apt. The work used a dataset of 1000 images which is classified into different distances i.e. of 10metres, 100metres, 1000metres and greater than 1000m approximately. The SVM is applied to these groups as proposed [3] ,

over the global features the resultant confusion matrix is shown in the table 1. Global can predict the less distance images 99.6% where as for far images the prediction rate is very low i.e. ~50% -40%. The output shows having far >1000m is predicted as 1000m rest all are predicted approximately similar to the expected output. The local feature data is taken into consideration and the images are classified by SVM .Table 2 shows that the far images are ~95% appropriately predicted but the near images are ~80% classified which is better than the global features.

The local can predict far correctly but only 80% images of ~100M and 10M can be predicted. The line graph in figure 2 better shows that the global images can classify the near images depth easily where as the far distance is classified by the local features. For each section 250 images are tested to find the distance.

The table 4 shows 249 images of 10M predicted correctly whereas only 50% of <1000M is predicted correctly for global.200 images of 10M is correctly identified and similarly 238 images are correctly identified.

Table 4: Global and local feature tested results

		DISTANCE(IN METRES)			
FEATURE		10m	100M	1000M	<1000M
	GLOBAL	249	238	200	100
	LOCAL	200	210	230	238

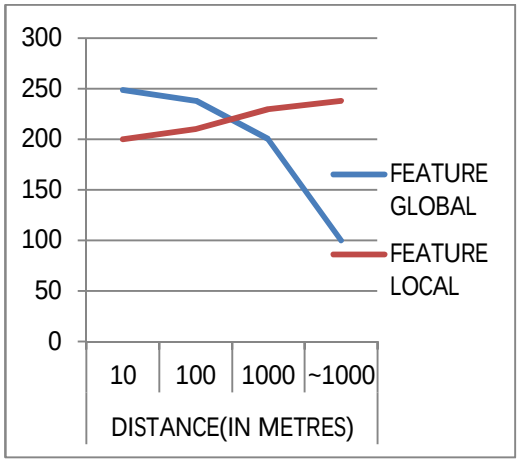


Figure 2:Number of images classified by local and global features

The advantages of both global and local is considered and by combining the feature results to a huge dataset. The dataset is reduce by applying PCA to which the SVM is applied results in a confusion matrix in Table 3 shows for far images the rate of approximation is ~93% and for near images ~98% which is better than applying individually.

Table 1: Confusion matrix for global feature

Global	10M	100M	1000M	<1000M
10M	0.996	0.02	0.01	0.01
100M	0.02	0.952	0.02	0.008
1000M	0.09	0.10	0.80	0.01
<1000M	0.05	0.2	0.35	0.4

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue V, May 2018- Available at www.ijraset.com

Table 2: Confusion matrix for local feature

Local	10M	100M	1000M	<1000M
10M	0.8	0.15	0.03	0.02
100M	0.06	0.84	0.06	0.04
1000M	0.01	0.04	0.92	0.03
<1000M	0.008	0.015	0.025	0.952

Table 3:Confusion matrix for local and global feature

Local and global	10M	100M	1000M	<1000M
10M	0.992	0.004	0.003	0.001
100M	0.01	0.98	0.01	0.0
1000M	0.02	0.014	0.912	0.064
<1000M	0.011	0.02	0.03	0.944

The graph in figure 3. shows that the global feature can predict near distance images whereas local feature can predict far distance. When both are combined on an average all the distance images are predicted correctly.

Table 5:The comparison of test images of local, global and combination of local and global

		DISTANCE(IN METRES)			
FEATURES		10M	100M	1000M	<1000M
	GLOBAL	249	238	200	100
	LOCAL	200	210	230	238
	LOCAL AND GLOBAL	248	245	228	236

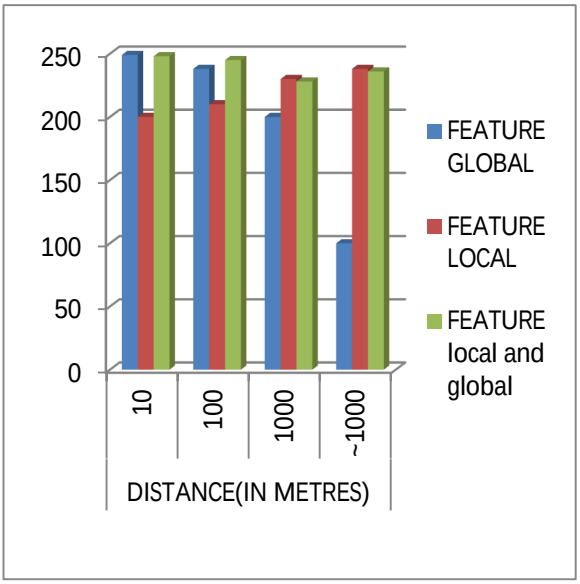


Figure 3: The difference in no. of images classified by global ,local and combination of global and local features

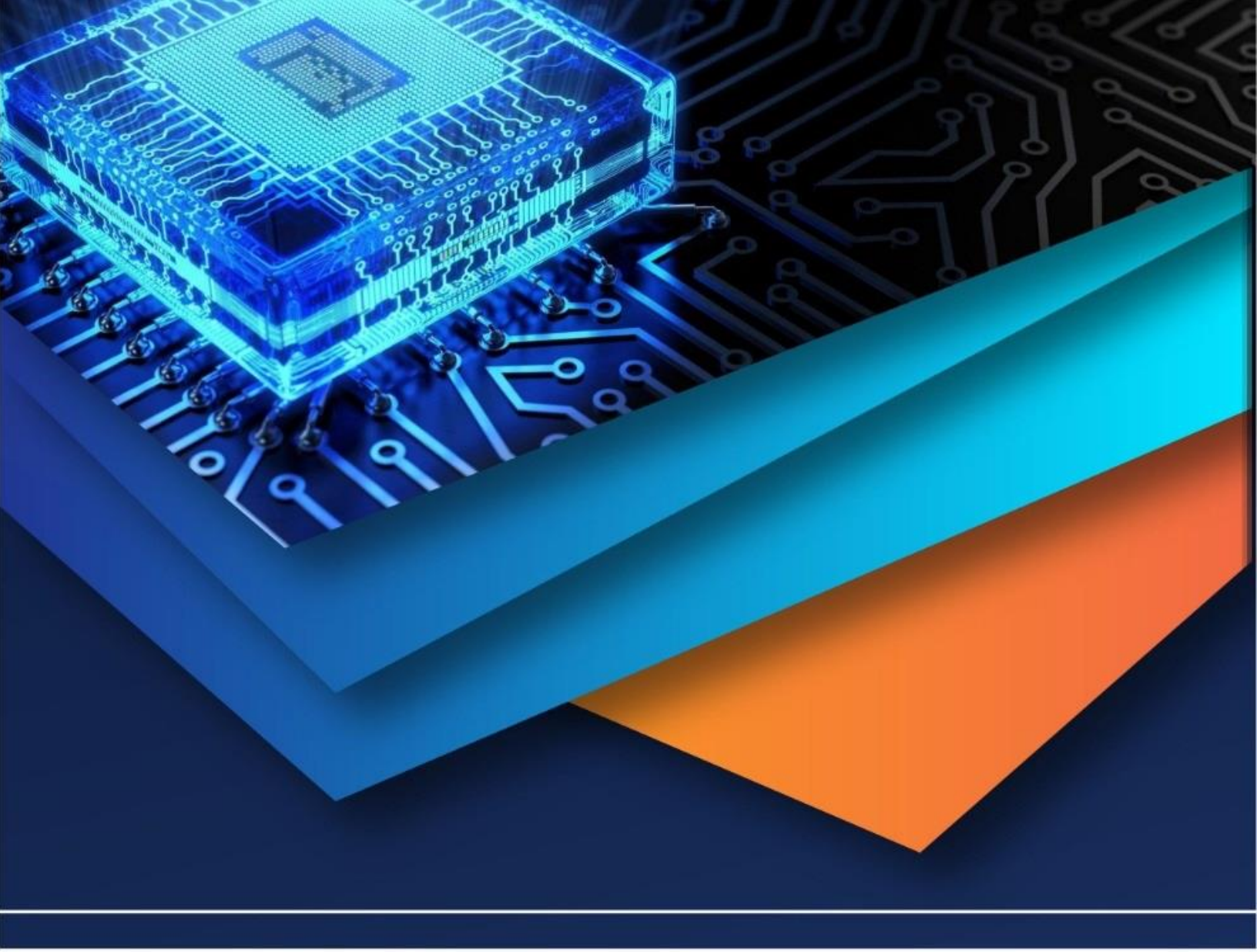
IV. CONCLUSION

In this, we have estimated depth of an image by performing training of dataset containing images and storing them in files, and the testing the new data and roughly predicting the depth using machine learning algorithm SVM classifier using python programming

language and found that SVM classifier gives more accuracy when both local and global features are combined together .they can classifies the problem for the static images at different distance. The work can be extended to other application.

REFERENCES

[1] Antonio Torralbo and Aude Oliva,"Depth Estimation for structured images",IEEE Transactions on pattern Analysis and Machine Intelligence,vol.24,no.9,pp.1-13 September 2002.
[2] Ashutosh Saxena, Sung H. Chung, Andrew Y. Ng,"3-D Depth Reconstruction from a Single Still Image"
[3] G. Santoshi, G Gowri Pushpa" Detect Pedestrian Orientation by Integrating Multiclass SVM Utilizing Binary Decision Tree", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 11, November 2015.
[4] G.Santoshi and S.R. Mishra" Pedestrian with Direction Detection Using the Combination of Decision Tree Learning and SVM", Springer Advances in Intelligent Systems and Computing 337,vo1. 1, pp. 249-255
[5] R. Cucchiara, C. Grana, A. Prati, and R. Vezzani, "Probabilistic posture classification for human-behavior analysis," IEEE Transactions on Systems, Man, And Cybernetics, Part A: Systems and Humans, vol. 35, no. 1, pp. 42–54, 2005.
[6] T. Gandhi and M. M. Trivedi, "Pedestrian protection systems:Issues, survey, and challenges," IEEE Transactions on Intelligence Transportation Systems, vol. 8, no. 3, pp. 413–430, September 2007.
[7] V. Aslantas,"A depth estimation algorithm with a single image", OPTICS EXPRESS, Vol. 15, No. 8, 16 Apr 2007



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY