INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS…..

SURFACE POTHOLE DEPTH ESTIMATION USING STEREO MODE OF IMAGE PROCESSING

Vijaya Bashkar. A[1], Gowri Manohar. T[2]

[1]Department of Electrical and Electronics Engineering,
SVU College of Engineering, Tirupati-517501, Andhra Pradesh, India.
vbhaskar0590@gmail.com.
[2]Department of Electrical and Electronics Engineering,
SVU College of Engineering, Tirupati-517501, Andhra Pradesh, India.

1. INTRODUCTION

Surface inspection normally involves surface crack detection and properties retrieval for post-analysis and maintenance. Potholes could prove fatal especially for vehicles on-road due to their dissimilar and multi-form depths. Asphalt-surfaces are subjected to a broad spectrum of traffic levels, from two-lane rural routes to multi-lane interstate highways. They age and deteriorate, thus they require corrective measures to restore safety and ride-ability. The most common forms of distress on asphalt-surfaces are potholes - small, bowl-shaped depressions in the pavement surface. Pothole repair is necessary in those situations where potholes compromise safety and pavement ride ability. Pothole detection and estimation is one of the important tasks for the proper planning of reparation and rehabilitation of the asphalt-surfaces. Road maintaining companies need many technicians for manual collection of data, and many working hours for rough estimation of damage on the road. There are many factors which influence decisions for pothole patching, such as the level of traffic, the time until scheduled rehabilitation or overlay, the availability of personnel, equipment, and materials, and the tolerance of the travelling public. The cost-effectiveness of the overall patching operation is affected by material, labour, and equipment costs. The key of decision making for future reconstruction is estimation of damage from collected information. Stereo vision based image-processing techniques; on the other hand, provide three dimensional measurements, such that the geometric features of a pothole can be determined easily. Stereo vision provides information on the depth size of the pothole, without the need for using high cost specialized laser scanners. Stereo vision systems are used to determine the depth of the pothole from the images taken at the same time but from slightly at different viewpoint situation by using two cameras. Any two similar images at different viewpoints are related by the epipolar geometry, and corresponding points of the two images are constrained to lie on the pairs of conjugate epipolar lines. The stereo matching algorithm is to find out the points in the two similar images that represent the same scene point. In literature, there are two approaches to find the depth estimation of the pothole application. The first one is to try to get the depth estimation based on the disparity of the two similar images captured by the cameras placed in such positions, so that the view of the images look like two different views of the similar images at different viewpoints. Such systems are required by knowing the intrinsic and the extrinsic parameters of the used camera via camera calibration in order to make the rectification process in processing stage before matching the stereo matching images. The second approach explains that to rectify images directly from a set of matched points. We follow a similar approach and present a simple and easy algorithm to get depth estimation of the pothole application directly from the inliers and outliers of the fundamental matrix. The fundamental matrix is estimated by using the corresponding points of the images. The difficulties in estimating the fundamental matrix lies in fact that there are often a fair portion of the mismatches of a given set point of the corresponding images. Therefore it is important method which can be used for estimation of the matching points and the method used to estimate the fundamental matrix should ne navigated by robust control equipment. The known algorithms for stereo matching can be classified in two basic categories: feature based algorithm and area based algorithm. The algorithms of both categories often use special methods to improve the matching accuracy and reliability of the algorithms.

From the existing literature, Li et al. [1] presented a real-time 3D laser scanning system for pavement rutting, shoving and pothole detection. Using the combined approach their system can collect, measure and visualize 3D pavement surface data, which is used to identify possible distortions[2] Ezzatollah Salari, Eddie Chou, and James J Lynch, implemented Machine vision based methods for automatic pothole detection have also been proposed, which only require a camera as input. However, existing approaches rely on the texture of the road surface, resulting in low accuracy.[3] M. Mustaffar, T. C. Ling , O. C. Puan studied in developing a photogrammetric based pavement evaluation approach by utilizing ortho-images and image processing techniques. In [4], Karuppuswamy proposed new method based on integrating vision and motion system for detection of simulated potholes. However, its disadvantages are that it cannot detect potholes smaller than 2 feet in diameter, and a pothole has to be white in color. In [5], Koch et al. proposed a novel supervised approach for automated pothole detection based on asphalt pavement surface images. This approach is based on pothole texture extraction and comparison, where the surface texture inside a pothole candidate has to be described and compared with the texture of the surrounding region. [6]Ashraf Anwar Fahmy experimented SURF, and implemented Un-calibrated Rectification using inliers to calculate the depth without disparity map. This idea encouraged us to implement it in our paper.

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS…..

The framework of this paper involves depth estimation of potholes in pavement surfaces using stereo three dimensional vision processing consisting of disparity calculation through un-calibrated rectification. a simplest and efficient stereo vision algorithm known as un-calibrated image rectification has been used to estimate the depth of the potholes in MATLAB environment. Steps of stereo image processing involve pre-processing, image rectification, disparity calculation and depth map. For feature description, the well-known matching algorithm SURF has been proposed. Unlike previous methods, estimation of depth has been carried out without disparity map. By applying geometric and epi-polar constraints, the outliers have been removed; thereby the maximum and minimum disparity has been performed using inlier locations.

2. STEREO VISION

A. Basics

The geometric key problem in this stereo vision process is to find the corresponding points in stereo images. Corresponding points are the projections of a single 2D point in different image spaces. The difference in the position of the corresponding points in their respective images is called disparity and shown in Figure.1.From the above captured images from the cameras: left and right, optical centers: O1 and Or. Virtual image plane is projection of actual image plane through optical centre .Baseline; B is the separation between the optical centers. Scene point, P, image at P1 and Pr. Disparity, is the amount by which the two images of P are displaced relative to each other.

d=Xr-Xl   Disparity Depth, Z=Bf/d

In addition to provide the function for the mapping of the corresponding images points onto scene points, a camera model can be used to constraint the search for corresponding image point for the one dimension. The intersection of a plane with an image plane is called an epipolar plane. Every point of an epipolar line must be corresponding to a single point on the corresponding epipolar line. The search for a match for the points in the first image can be reduced to a one dimensional replica in the second image plane.

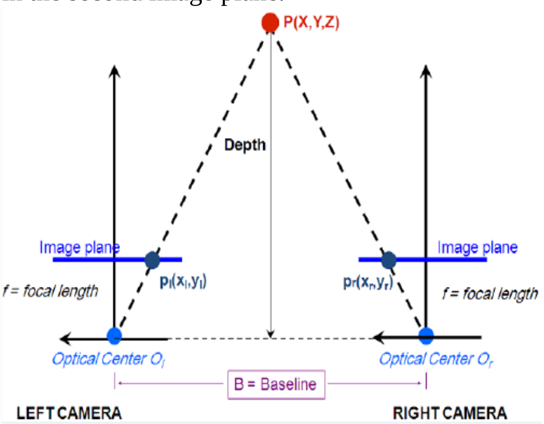

Figure 1: Stereo vision basics

B. Correspondence problem

There are two issues, how to select a candidate matches? And also how to determine the goodness of the perfect match? Two main classes of the correspondence (matching) algorithm: First is correlation based, attempt to establish a correspondence by matching image intensities, usually over a window of pixels in each image. Second Feature-based, attempt, attempt to establish a correspondence by matching sparse set of image features, usually edges. Disparity map is sparse, and number of points is related to the number of image features identified. Feature-based methods can be extracted from the scene, faster than the correlation-based methods, provide sparse disparity maps, suitable for applications like visual navigation, and relatively intensive to illumination changes.

The whole paper is divided in to sections: Section 1: Stereo processing for pothole depth estimation. Section 2: Results and discussions. Section 3: Conclusion. Section 4: References.

3. METHODOLOGY

Section 1:

1.1. Stereo processing for pothole depth estimation

Stereo vision processing involving Un-calibrated rectification i.e. finding the depth without the disparity map shown in Figure.2 has been implemented here as follows:-

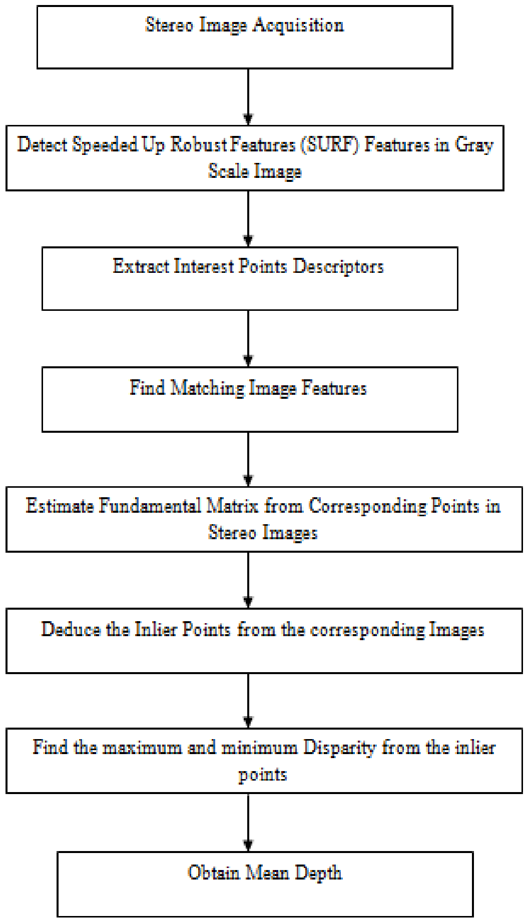

Figure 2: Stereo vision Un-calibrated rectification Process

1.2. Stereo Image Capture

A stereo image of pothole has been taken from the database shown in Figure.3 converted to gray scale images. The size of images has been reduced for easier analysis and to avoid software conflicts.

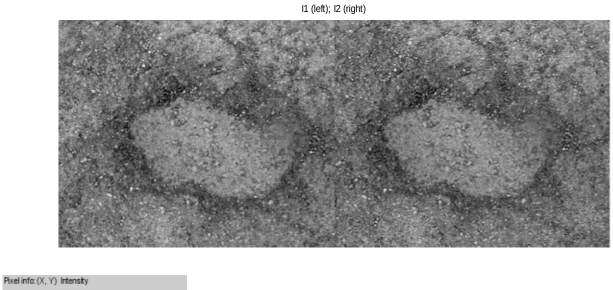

Figure 3: left and right image of a pothole

1.3.Colour-composite image

To know about the pixel-wise differences between the stereo images, a composite image of two different colour separations represented in Figure.4 has been introduced.

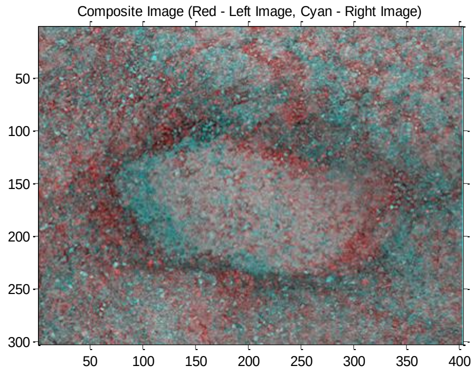

Figure 4: Composite image of left and right images

1.4. Collect interest feature points from each image

Since, there is certain amount of offset involved between the images a much necessary step of rectification has been performed. The rectification process includes finding a set of corresponding matching points, aligning them to calculate disparity. Here, SURF feature algorithm has been used which is shown in Figure.5 and Figure.6 respectively. SURF chooses region of interest as blobs from the rectilinear scale-space, and forms local features constructed on the gradient dissemination of the image. Speeded up Robust Features algorithm is based on summation of 2D Haar wavelet responses. It uses integral images. The steps of features detection as follows:

- Interest points are selected at distinctive locations in the image, such as corners, blobs and T-junctions. The most valuable property of an interest point detector is its reliability, i.e. whether it is reliable or not and finds the same interest points under different viewing conditions.
- Next, the neighbourhood of every interest point is represented by a feature vector. This descriptor has to be distinctive and, at the same time, robust to noise, detection errors, geometric and photometric deformations.
- Finally, the descriptor vectors are matched between different images. The matching is often based on a distance between the vectors, e.g. the Mahalanopbis or Euclidean distance. The dimension of the descriptor has a direct impact on the time this takes, and a lower number of dimensions is therefore desirable.

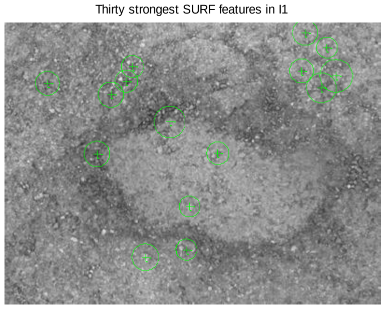

Figure 5: SURF feature extraction of left image

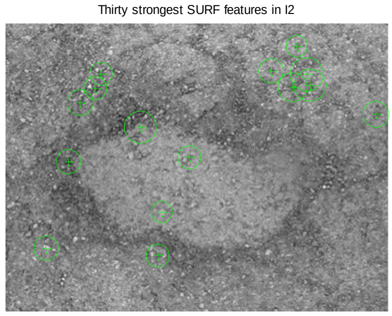

Figure 6: SURF feature extraction of Right image

Finding putative point correspondences

Extracting features and matching features for each blob has to be computed. Sum of absolute differences (SAD) has been used shown in figure 7. to find and calculate matching indices and matching points' location. The SAD algorithm is one of the modest algorithms measuring the dissimilarity of the left and right stereo pictures analogous with four-sided window. Intensity dissimilarities for each epi center pixel (x, y) in a window W (i, j) is computed as follows in (A):

$$\mathrm{SAD}(I, j, d) = \sum_{(x,y)\in W(i,j)}^{N} |\ \mathrm{Ileft} - \mathrm{Iright}(x - d, y)\ | \cdots (1)$$

Where I left and I right are functions of pixel intensities of the left and right stereo images individually. The disparity calculation has been repetitive through the x-direction encompass in the row section of the image of the investigated 3D image. The smallest distinction value over that frame specifies the best corresponding pixel, and its location.

Note that the image correspondences consist of some outliers.

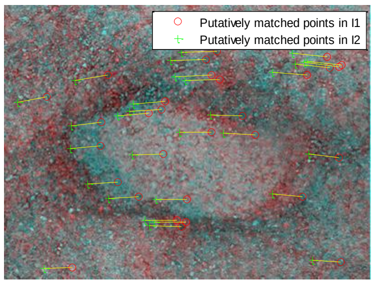

Figure 7: putatively matched correspondences

Removal of outliers

The appropriately matched points must fulfill epipolar constraints. The Fundamental matrix denoted as F is a 3×3 network which relates comparing focuses in stereo images. In epipolar theory, with identical picture directions, X and X′, of comparing focuses in a stereo image pair, FX depicts an epipolar line on which the relating point X′ on the other image must lie. That means, for all pairs of corresponding points holds

$$X' FX = 0 \quad ….(2)$$

RANSAC: Random Sample Consensus can be contemplated as a web crawler. It chooses again an indiscriminate sample of correspondences and calculates the inliers acumen the fundamental matrix.

Figure.8 denotes the formed inliers after RANSAC.

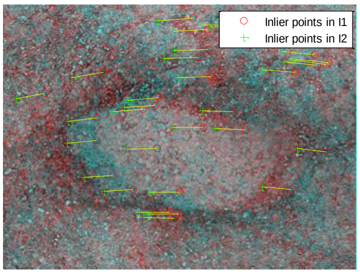

Figure 8: Inlier points of left and right images

Estimation of depth

Set of disparity values have been checked from the obtained inner points,

$$disparity = \sum_{i=1}^{N} | \text{Inleft}(x, y) - \text{Inright}(x, y) | \quad …. (3)$$

Where, In left, In right are the inliers points in the left and in the right stereo images respectively.

Depth values can be computed from disparity values as,

$$\text{Depth} = \frac{f+b}{disparity} \quad …. (4)$$

$$\text{Mean depth} = \frac{\text{depth (max)} + \text{depth (min)}}{2}$$

Where f and b are focal length of the camera in mm and baseline between two optical centers in mm respectively.

Disparity can be measured in pixels.

Rectified and cropped image in 3-D shown in Figure.9 as,

Rectified Stereo Images (Red - Left Image, Cyan - Right Image)

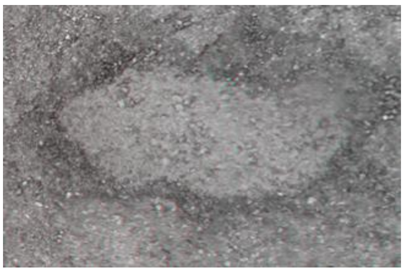

Figure 9: Rectified stereo image

Section 2: Results and Discussions

Depth estimation of Pothole image

Disparity check defined from the formula (3) gives the x coordinate location difference between the two images. From the values obtained from the inlier points, maximum and minimum disparities of pothole image shown in figures (10), (11), (12) are computed.

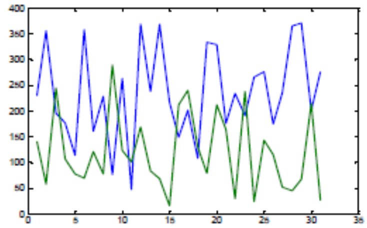

Figure 10: Plot of inlier SURF points of left image

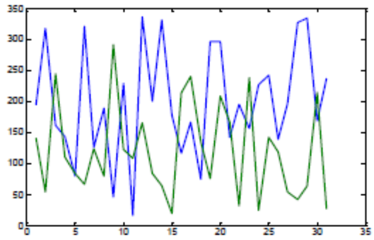

Figure 11: Plot of inlier SURF points of right image

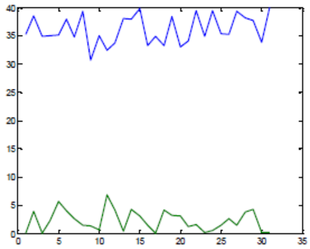

Figure 12: Disparity plot from inlier SURF points

Taking 0.02 mm/pixel as pixel size, Maximum depth for minimum disparity and minimum depth for maximum disparity is computed as shown in figure.13

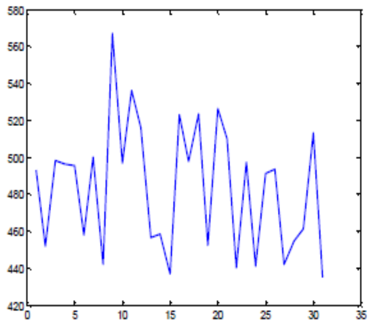

Figure 13: Estimated depth plot of pothole image

Table 1: Depth estimation of pothole image

Properties	Estimated Values(mm)
Maximum depth	567.1198
Minimum depth	435.3266
Mean depth	501.2232

Calculation of Errors

1. The min. depth error is calculated according to the following formula:

*E1%= abs((Minimum real depth- minimum depth estimated)/Minimum real depth)*100….(5)*

2. The max. depth error is calculated according to the following formula:

*E2%= abs((Maximum real depth- maximum depth estimated)/Maximum real depth) *100……(6)*

3. The depth error of the mean is calculated according to the following formula:

*E3%= abs((Mean real depth- mean depth estimated)/Mean real depth)*100….(7)*

Real depths have been measured and taken as a constant value.

Table 2: Error percent estimation

Error Properties	Percent error calculated in %
Min. Depth	3.26
Max. Depth	3.112
Mean Depth	0.264

4. CONCLUSION

This paper is aimed at pothole depth estimation in stereo mode of image processing via un-calibrated mode of rectification. The results obtained are shown to be calculated as error percent in tabular columns in comparison to the expected result. Therefore, reasonably accurate calculations of pothole mean depth values have been estimated through MATLAB environment. The future work can be carried out using calibrated rectification methodology for depth estimation to improve accuracy. Also, a mobile navigation system can be designed to interface with the vision system in dynamic environment.

REFERENCES

[1] Q. Li, M. Yao, X. Yao, B. Xu, "A real-time 3D scanning system for pavement distortion inspection", Meas. Sci. Technol. 21 (2010).

[2] Ezzatollah Salari, Eddie Chou, and James J Lynch, "Pavement distress evaluation using 3d depth information from stereo vision," Tech. Rep., 2012.

[3] Mustaffar M, Ling TC, Puan OC. "Automated Pavement Imaging Program (APIP) For Pavement Cracks Classification and Quantification – A Photogrammetric Approach", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B4. Beijing, 2008.

[4] J. Karuppuswamy, V. Selvaraj, M.M. Ganesh, E.L. Hall, Detection and avoidance of simulated potholes in autonomous vehicle navigation in an unstructured environment, Intelligent Robots and Computer Vision XIX: Algorithms, Techniques, and Active Vision, vol. 4197, pp. 70-80, 2000.

[5] C. Koch, I. Brilakis, Pothole detection in asphalt pavement image, Advanced Engineering Informatics, Vol. 25(3), pp. 507- 515, 2011.

[6] Ashraf Anwar Fahmy. "Stereo vision based Depth Estimation Algorithm in Un-calibrated rectification, International Journal of Video & Image Processing and Network Security IJVIPNS-IJENS, April 2013.