

# Implementation of Stereo Matching Technique in Android Application

P.Karckey, PG Scholar

Department of Computer Science and Engineering  
Indra Ganesan College of Engineering  
Tiruchirappalli, Tamil Nadu, India 620012  
marcxiya@gmail.com

Dr. G. Balakrishnan

Department of Computer Science Engineering  
IndraGanesan College of engineering  
Tiruchirappalli, Tamil Nadu, India 620012  
balakrishnan.g@gmail.com

**Abstract** – Blind people faces number of visual challenges everyday and visually impaired also has various problems while moving around spaces which are not familiar for them. There are so many computer vision techniques available but implementing them in mobiles is difficult. Hence an application that helps the blind or visually impaired people, can be developed using android through stereo matching techniques in mobile phones. The proposed methodology deals with the fast stereo matching method of high resolution images for finding the Disparity scenes. The visually impaired can identify the object based on distance using this mobile application after the sound conversion. This application is used by all kind of android mobile phones which is having the stereo camera.

**Keywords** - *Android; OpenCV; Stereo Matching; Disparity Map.*

## I. INTRODUCTION

Visually impaired has lots of problems to explore unknown spaces independently, safely and efficiently. They find their navigation difficult as they often lack the needed information for bypassing obstacles and hazards. With the recent advances in inclusive technology it is possible to extend the support given to people with visual impairment during their mobility. Computer vision systems are undergoing a great development, which may allow visually impaired people to have access to better information of the surrounding environment, such as to detect the presence of obstacles or points of interest and calculate how far the user is from them. Even though these kinds of products are available for blind people, they are much harder to carry while travelling for visually impaired people. So in order to avoid difficulties, mobile is a single device that can be used for many purposes for all kind of peoples in the world.

Mobile Phones have become an inevitable part of our daily life. Stereo matching method is capable of estimating scene depth information in the real time. The main challenge of stereo vision, also called stereopsis, is the reconstruction of 3D information of a scene captured from two different points of view. This can be done by finding pixel correspondences between both (left and right) images. The horizontal displacement of corresponding pixels is called disparity

This approach builds a prior over the disparity space by forming a triangulation on a set of robustly matched correspondences, named 'support points. As demonstrated this experiments, this method is able to achieve state-of-the-art

performance with significant speedups of up to three orders of magnitude when compared to prevalent approaches and obtain 300 MDE/s (million disparity evaluations per second) on a single CPU core. The following are some of the literature review which is based on this paper. The drawbacks of all the following reviews are overcome by this research based paper.

## II. RELATED WORK

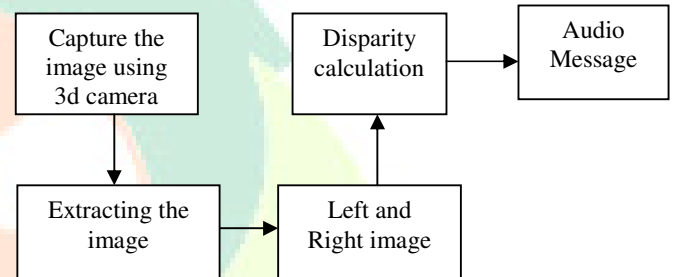
A novel disparity map refinement method is introduced in vision based surveillance framework for the task of detecting objects of interest in dynamic outdoor environments from two stereo video sequences taken at different times and from different viewing angles by a mobile camera platform [2]. Here the depth evaluation using block matching algorithm is performed [3]. A novel map based rendering (DMBR) method to provide the consumers with depth-preserving 3D services [4]. A novel based acquiring stereo images using the Unreal Development Kit method permits the creation and capture of virtual stereo scenes in a carefully controlled environment. In this method, Images are captured with physical stereo cameras in stereo matching tests using OpenCV algorithms [5]. The computer stereo vision is an important technique for robotic navigation and other mobile scenarios where depth perception is needed, but it usually requires two cameras with a known horizontal displacement, a prototype using a state-of-the-art mobile phone, which has to be manually displaced in order to record images from different lines of sight. Since the displacement between the 2 images is not known in advance, it is measured using the phone's inertial sensors. The evaluation of the accuracy of single-camera approach by performing distance calculations to everyday objects in different indoor and outdoor scenarios, and compared the results with a stereo camera phone [6]. To obtain better result [16] used a Patch Match method, and slanted plane on the Middlebury bench mark [11] proposed a novel aggregation method inspired by the anisotropic diffusion technique used in image filtering. Moreover, near realtime performance is demonstrated with a GPU implementation of the algorithm [12] calculated the likelihood matching error based on rank transform, and then it is parallelized and implemented in GPU programming [13]. The classical optical flow method performs well when integrated with modern optimization and implementation techniques. Although the local optimization approach provides good results, its computational requirements are significantly

lesser than that of the global. In global optimization [14] generated high quality results by integrating color-weighted correlation and improved hierarchical belief propagation. The development of promising stereo vision algorithms that efficiently tradeoff accuracy for large reductions in required computational resources [15] although methods such as Rank transform, mutual information, and NCC have achieved robustness to some extent, they are unable to achieve high accuracy. As the solution to this problem [16] proposed a new stereo matching measure that is robust to various radiometric variations, such as local and global radiometric variations is introduced and at the same time it is accurate. Dense and accurate matching can be obtained by global methods, which enforce smoothness explicitly by minimizing an MRF-based energy function which can be decomposed as the sum of a data fitting term and a regularization term. Since for most energies of practical use such an optimization is NP-hard, approximate algorithms have been proposed, e.g. graph-cuts [17, 18], belief propagation. A parallel VLSI hardware design for belief propagation that achieves real time performance on VGA imagery was proposed. The application of global methods to high-resolution images is, however, limited by their high computational and memory requirements, especially in the presence of large disparity ranges. Furthermore, models based on binary potentials between pixels favour front-parallel surfaces. 'Ground control points' are used in [19] to improve the occlusion cost sensitivity of dynamic programming algorithms. In [20,21] disparities are 'grown' from a small set of initial correspondence seeds. Though these methods produce accurate results and can be faster than global approaches, they do not provide dense matching and struggle with texture less and distorted image areas. Approaches to reduce the search space have been investigated for global stereo methods [23, 22]. However, they mainly focus on memory requirements and start with a full search using local methods. Furthermore, the use of graph-cuts imposes high computational costs particularly for largescale imagery. In contrast, in this paper a proposed Bayesian approach to stereo matching that is able to compute accurate disparity maps of high resolution images at frame rates close to real time without the need for global optimization. Even though several methods are available those methods have following drawbacks, Single camera mobile is used for capturing images in two different locations but this will not give accurate result for that a Stereo Camera Mobile is used in proposed methodology for capturing the image, which will give accurate result for Image Separation and disparity calculation. There are so many problems arises like disturbance, acquisition problem and calibration problem if taken using Single Camera. There are so many techniques available for finding disparity using Mat lab, etc but the proposed methodology uses android platform. The 3d image extraction and finding disparity in mobile using stereo Camera is not available in Android. Some applications are cost effective and it always needs a web. These are the some reviews based to this project work. The following is the methodology used for this paper

and it has the architecture diagram to develop the application and to implement the stereo matching technique in android.

### III. METHODOLOGY

Recovery of 3D shape is a critical problem in many vision application domains such as object modeling, scene understanding and high level visual activity recognition or robotics applications. Obtaining a precise and accurate depth map is the ultimate goal for 3D shape recovery and 3D image reconstruction. The topic of the paper is focused on the process of the depth map computation from the images that are captured by the stereo camera. The Ultimate finding of project is to Design and development of android application which assist blind users to navigate on their own.



**Fig 3.1 Architecture of the Proposed System**

The Stereo Matching method is capable for estimating scene depth information in the real time. The main challenge of stereo vision, also called as stereopsis, is the reconstruction of 3D information of a scene captured from two different points of view. This can be done by finding pixel correspondences between both (left and right) the images. Here, this Object Extraction will show the difference of foreground and background object in the Captured Image that is foreground or background image will be darkened by the gray color.

Next the disparity will be found. The horizontal displacement of corresponding pixel is called disparity. Using the disparity concept the distance of an object will be calculated by using some techniques. Finally, Sound Conversion techniques will be applied. And the Sound will be heard according to the distance of an object. There will be variations in the sound when the object is close to us; this will be happened while the camera runs lively and the user can also use head set while navigating for example, if the object is closer to left side of the user then the sound will be heard through left side head set and if the object is placed in right side of the user then sound will be heard by right side head set and if the object is placed in front of the user both side the sound will be heard.

Therefore, these techniques are used to develop the android application and this android application will be useful for Visually Impaired or Blind people to Navigate own by holding the mobile device and when the object is close to the device some sound will be produced. Using this sound the Visually Impaired or Blind people get useful for Navigating by own without any ones help.

#### IV. STEREO IMAGE EXTRACTION

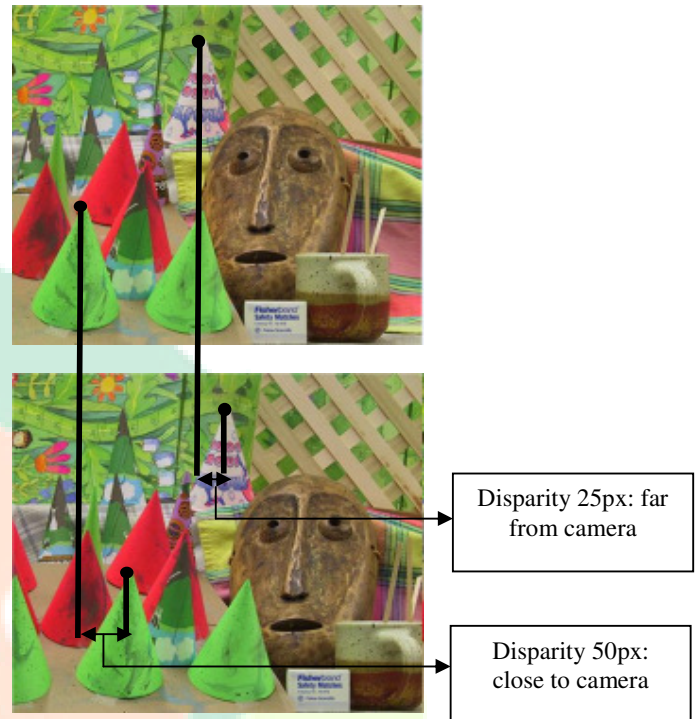
In Stereo Image Extraction the process of splitting 3d image into left and right image after capturing a 3D image using Stereo Mobile which will be useful for finding disparity using Stereo Matching Methods. In the Proposed system a new camera application is created for an android device. Creating an android project using ADT Bundle Tool and including necessary Libraries (OPENCV lib) along with new application and also the predefined Opencv functions are checked whether loaded or not.

Then Opencv Manager App should be installed in android device. Suppose if it is not installed before installing a newly created android apk it will automatically redirect to play store to download it. The following explains about the creation of camera application in android device.

##### A. Stereo Matching in Passive mode

For each pixel in the left image, we can now search for the pixel on the same scan line in the right image, which captured the same object point. Because a single pixel value is typically not discriminative enough to reliably find the corresponding pixel, one usually tries to match small windows (e.g. 7x7 pixels) around each pixel against all possible windows in the right image on the same row. As further restriction, we don't need to search the entire row but only a limited number of pixels to the left of the left image pixel's x-coordinate, corresponding to the slightly cross-eyes gaze necessary to focus near objects. This accelerates the matching and restricts the depth range where points can be triangulated. If a sufficiently good and unique match was found, we associate the left image pixel with the corresponding right image pixel. The association is stored in the disparity map in form of an offset between the pixels x-positions.

This matching technique is called local stereo matching, as it only uses local information around each pixel. Obviously, it can only match a region between left and right image when it is sufficiently distinct from other image parts on the same scan line. Thus, local stereo matching will fail in regions with poor or repetitive texture. Other methods, known as global stereo matching, can also exploit neighboring information. They don't just consider each pixel (or image patch) individually to search for a matching partner; instead they try to find an assignment for all left and right image pixels at once. This global assignment also takes into account that surfaces are mostly smooth and thus neighboring pixels will often have similar depths. Global methods are more complex and need more processing power than the local approach, but they require less texture on the surfaces and deliver more accurate results, especially at object boundaries.



**Fig 4.1 Result of stereo matching**

The above figure shows that the 3d image which is extracted as left and right respectively. In that extracted image it produces the disparity value and output disparity image. The above methodology is used to implement the stereo matching technique in android application. The following section is discussed deeply about the stereo matching.

#### V. STEREO MATCHING

The following are some of the supporting points which are used to the stereo matching technique.

- Capture the image using 3d camera
- Extracting the Image
- Left and Right Image
- Disparity Calculation
- Audio Message

As support points, it denotes pixels which can be robustly matched due to their texture and uniqueness. While a variety of methods for obtaining stable correspondences are available[17, 21, 22], the matching support points on a regular grid using the  $\ell_1$  distance between vectors formed by concatenating the horizontal and vertical Sobel filter responses of 9 x 9 pixel windows to be both efficient and effective. In all experiments the Sobel is used with masks of size 3 x 3 and a grid with fixed step-size of 5 pixels. A large disparity search range of half the input image width was employed to impose no restrictions on the disparities.

The Following describes about probabilistic generative model which, given a reference image and the support points, can be used to draw samples from the other image.



This algorithm aims to minimize the following global energy function,  $E$ , for disparity image,  $D$ .

$$(D) = ((p,)+ \sum_{p \in NP1} [I(p, Dp - Dq) = 1] + \sum_{q \in NP2} [I(p, Dp - Dq) = 1]) \quad \text{----- (1)}$$

with  $P2 \geq P1$  where  $E(D)$  is the energy for disparity image,  $D$ ,  $p, q$  represent indices for pixels in the image  $Np$  is the neighbourhood of the pixel  $p$   $C(p, Dp)$  is the cost of pixel matching with disparity in  $Dp$   $P1$  is the penalty passed by the user for a change in Disparity values of 1 between neighbouring pixels  $P2$  is the penalty passed by the user for a change in disparity values greater than 1 between neighbouring pixels  $I[.]$  is the function which returns 1 if the argument is true and 0 otherwise the minimized function produces a perfect disparity map with smoothing governed by parameters  $P1$  and  $P2$ ; however, minimizing the function for a 2D image space is an NP-complete problem. The semi-global matching function approximates the 2D minimization by performing multiple 1D, or linear, minimizations. The matching function aggregates costs on multiple paths which converge on the pixel under examination. Cost is computed for the disparity range specified by the minimum disparity and number of disparities parameters. By default, the matching algorithm aggregates costs for 5 directions. You can set the full dynamic programming parameter to true to force the algorithm to aggregate costs for 8 directions. Let,  $S(p, d)$  be the aggregate cost for pixel  $p$  and disparity  $d$ . Then

$$(p, d) = (p, r) \quad \text{----- (2)}$$

Where,  $r$  is a direction used for converging to the pixel  $p$ .  $Lr(p, d)$  is the minimum cost of the path taken in direction  $r$  from pixel  $(p)$  for disparity  $d$ . The cost  $Lr(p, d)$  is given in the following equation:

$$Lr(p, d) = C(p, d) + \min (Lr(p-r1, d-1) + P1, Lr(p-r, d+1) + P1, \min_i (Lr(p-r1, d-1) + P2) - \min_k \quad \text{----- (3)}$$

The equation uses the following costs to find the disparity by adding current cost,  $C(p, d)$ , to previous pixel in direction  $r$ : The minimum of the cost at previous pixel with disparity  $d$ . The cost at previous pixel with disparity  $d - 1$  and  $d + 1$  with added penalty  $P1$ . The cost at previous pixel with disparities less than  $d - 1$  and greater than  $d + 1$  with added penalty  $P2$ . In order to limit the ever increasing value of  $Lr(p, d)$  on the path, minimum value of the previous pixel is subtracted. The upper value of  $Lr(p, d)$  is bounded by  $C_{max} + P2$ , where  $C_{max}$  is the maximum value of cost  $C$ . The cost function  $C(p, d)$  is computed in the following manner:

$$C(p, d) = \min (d, p, p-d, IL, IR, d, p-d, p, I, L, R) \quad \text{----- (4)}$$

Where,  $IL$  and  $IR$  are left and right rectified images, respectively.

$$(p, p-d, I, L, I, R) = \min (p-d-0.5, p-d+0.5, I, L, I, R) - I, L, R \quad \text{----(5)}$$

The value of  $C$  is aggregated over a window of a user-defined size  $l$ . After computing  $S(p, d)$  for each pixel  $p$  for each disparity  $d$ , the algorithm chooses the disparity which provides the minimum cost for that pixel. The complexity of this algorithm is given in the following equation:

$$O(w.h.n) \quad \text{----- (6)}$$

where,  $w$  equals the width of the rectified image  $h$  equals the height of the rectified image  $n$  equals the number of disparities.

## DISPARITY ESTIMATION

The Stereo Matching process is equal to the human stereo vision process, also known as binocular vision system. Estimating the disparity map from the left and right images, maximum posterior (MAP) is estimated to compute the disparities. This approach can also be used for images with all kinds of color or gray-scale depths. As from here, a mapping from the left image to the right image will be assumed.

**Definition 1 (Disparity range).** The disparity range  $r = [0; \_max]$  is the shift range for which the calculations are done. As only stereoscopic images are regarded, pixels from the left image can only find its correspondence by shifting to the left.  $\_max$  limits the disparity range and is determined by the designer. For objects close to the camera larger shifts are expected.

**Definition 2 (Starting pixel).** A starting pixel has its coordinates at  $(x_L; y)$  in the left image, whose corresponding pixel in the right image has to be found.

**Definition 3 (Matching set).** The matching set consists of those pixels from the right image with which the starting pixel is compared to. This set has  $\delta_{max} + 1$  elements with coordinates  $((x_L; y); (x_L + 1; y); \dots; (x_L + \delta_{max}; y))$ .

**Definition 4 (Matching pixel).** A matching pixel has its coordinates at  $(x_R; y)$  from the matching set which is compared with the starting pixel. The resulting disparity is  $d \in r$ .

These are the some of the definitions and formulas to calculate the disparity between two images (left and right) which are extracted from the 3d image captured by the stereo camera.

## CONVERSION INTO ANDROID

The Following are the steps for integrating C++ and Java coding for Android Application. To Run C++ coding in Android NDK (Native Development Kit) is installed along ADT Bundle Tool and other S/w is also used to compile and run in c++ file for developing android application. Configuring the path by adding all software's and extensions are added using eclipse. In the Project folder the JNI folder is created and inside that two new files are created for linking the C++ coding into java.

1. Android.mk file is used to call the c++ file into Android Java file.
2. Application.mk file is used to configure the emulator.

The CDT GCC Built-in compiler setting is included from preprocessor built path in c++ general then the NDK path is added in GNU C++ and the Cygwin GCC is used to compile the C++ coding. Additional coding is added in main project for using above steps. The Library is added along with this application. Htc Extensions are added in this project to run this application in htc Mobile. Emulator is configured by installing required sdk and also the application is created on Android operating system ( 4.5 Lollipop ) is run successfully then apk will be automatically generated. After installing the newly created A mobile then it shows find disparity button. If the button is pressed then it will show the disparity map on that itself. This Emulator shows in laptop / Desktop when the code runs without error in eclipse. The full application cannot be run in Emulator but the generated apk runs in all versions of android devices. The emulator shows only the icons of the created android applications. (Disparity) is the icon for the proposed methodology that is android application.

## CONCLUSION AND FUTURE WORK

In this paper, the implementation system using a Stereo Camera Mobile Phone which is able to measure the distance of proposed methodology overcomes with existing methods that is two images should be taken by the user using single camera which gives some false. The proposed methodology gives accurate result for finding Existing Method, Proposed method Disparity Map. Therefore the existing method is **2 Time** of stereo vision everyday objects. The proposed measurement but in the disparity, the Disparity Map shows the distance of each object. Thus the Disparity Map for a Stereo Image is obtained successfully using the Stereo Camera Mobile. The following are the benefits of proposed System; this application will run on all versions of Android Mobiles. The visually impaired can highly get benefits by using this mobile application after sound conversion. Blind or Visually impaired can make use of sound conversion technology and navigate on their own. All types of people can also use this android application, to calculate distance of an object.

## REFERENCES

- [1] Liyuan Li, et.al. "Stereo-based human detection for mobile service robots" Control, Automation, Robotics and Vision Conference, ICARCV 2004.
- [2] Haberdar, H. , et.al. "Disparity Map Refinement for Video Based Scene Change Detection Using a Mobile Stereo Camera Platform", Pattern Recognition (ICPR), 2010 20th International Conference.
- [3] Hamzah, R.A., et.al. "Depth evaluation in selected region of disparity mapping for navigation of stereo vision mobile robot", et.al. Industrial Electronics & Applications (ISIEA), 2010 IEEE Symposium.
- [4] Cheolkon Jung ,et.al. "Disparity-map-based rendering for mobile 3D TVs" Consumer Electronics, IEEE Transactions on 2011.
- [5] Miknis, M. , et.al. "Virtual environment stereo image capture using the Unreal Development Kit", Computer & Information Technology (GSCIT), 2014.
- [6] Clemens Holzmann and Matthias Hochgatterer "Measuring Distance with Mobile Phones Using Single- Camera Stereo Vision". 2012, 32<sup>nd</sup> International Conference. Visual Object Classes Challenge 2012 (VOC2012) Results. 2012.
- [7] Sandhini.G, et.al. "Robotics Bees", IROS'93, International Conference, DIST, Genoa Univ., Italy.
- [8] Brown M. Z., et.al. "Advances in computational stereo," IEEE Trans. Pattern Anal. Mach. Intell., Aug. 2003.
- [9] Zucheul Lee, Jason Juang, and Truong Q. Nguyen, "Local Disparity Estimation with Three-Mode Cross Census and Advanced Support Weight", IEEE Trans. on Electrical and Computer Engineering, 2013.
- [10] Xiaoyan Hu, and Philippos Mordohai, "A Quantitative Evaluation of Confidence Measures for Stereo Vision", IEEE Trans. on Pattern Analysis And Machine Intelligence, Vol. 34, No. 11, Nov 2012.
- [11] Leonardo De-Maeztu, Arantxa Villanueva, and Rafael Cabeza, "Near Real-Time Stereo Matching Using Geodesic Diffusion", IEEE Trans. on Pattern Analysis And Machine Intelligence, Feb 2012.
- [12] X.Xiang, M.Zhang, G.Li, Y.He, "Real-time stereo matching based on fast belief propagation", Springer on Machine Vision and Applications, Jan 2012.
- [13] J Beau J.Tippets, Dath-Jye Lee, James K.Archibald, and Kirt D.Lillywhite, "Dense Disparity Real-Time Stereo Vision Algorithm for Resource-Limited Systems", IEEE Transactions on Circuits and Systems for Video Technology, Oct 2011.
- [14] Sang Hwa Lee and Siddharth sharma, "Real-Time Disparity Estimation Algorithm for Stereo Camera Systems", IEEE Trans. Consumer Electronics, vol. 57, no 3, Aug 2011.
- [15] Yong Seok Heo, Kyong Mu Lee, Sang Uk Lee, "Robust Stereo Matching Using Adaptive Normalized Cross Correlation," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 32, Apr 2011.
- [16] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 3017-3024, Jun. 2011.
- [17] Boykov, Y., Veksler, O., Zabih, R.: Markov "Random fields with efficient approximations" in, CVPR, 1998.
- [18] Kolmogorov, V., Zabih, "Computing visual correspondence with occlusions using graph cuts", in International Conference on Computer Vision, 2001.
- [19] Liang, C.K., Cheng, C.C., Lai, Y.C., Chen, L.G., Chen, H.H. "Hardware-efficient belief propagation in Computer Vision and Pattern Recognition", 2009.
- [20] Bobick, A.F., Intille, "Large occlusion stereo", in International Journal of Computer Vision 33 1999.
- [21] Cech, J., S\_ara, "Efficient sampling of disparity space for fast and accurate matching", in Computer Vision and Pattern Recognition, 2007.
- [22] Kostkova, J., Sara, R. "Stratified dense matching for stereopsis", in complex scenes, in BMVC, 2003.