RESEARCH ARTICLE OPEN ACCESS

3d Image Reconstruction Using Automatic Depth Extraction from 2d Images Using Machine Learning

S. Sathiya Priya*, S. Subheetha** P.Ezhilarasi***
**Department of xxxxxxxxx, St. Joseph's College of Engineering, Chennai.*
***Department of xxxxxxxxx, St. Joseph's College of Engineering, Chennai.*
****Department of xxxxxxxxx, St. Joseph's College of Engineering, Chennai.*
Corresponding Author: S. Sathiya Priya

ABSTRACT
The availability of 3D players and displays has significantly risen in the last years. Nonetheless, the amount of 3D content has not experienced an increment of such magnitude. To overcome this problem, many algorithms for converting images and video from 2D to 3D have been proposed. Here, we have proposed an automatic learning-based 2D-to-3D image conversion approach, based on the key hypothesis that color images with similar structure likely present a similar depth structure. This project presents simple and efficient modified line growing stereo matching algorithm using region based segmentation which focuses on removal of Ambiguous pixels in segmented region. MODIFIED RELIABLE DISPARITY PROPAGATION FOR DEPTH EXTRACTION algorithm is developed for extraction depth information from two colored stereo image set. The adaptive bilateral filter is used for eliminating unreliable disparity estimation along with removal of impulse noise while preserving the edges. The efficiency of the proposed method has been verified by using test bed image sets. For the regions corresponding to background, depth values are achieved merging information from linear perspective and texture characteristics. We presented a method which combined the foreground segmentation. According to the separated foreground and background results from foreground segmentation algorithm, the viewer can use their acquired visual experience to initiate this operation with some depth information of background.
Keywords - Histogram, Depth extraction, Disparity map, Adaptive Histogram Equalization

---

Date of Submission: 30-03-2019 Date of acceptance: 13-04-2019

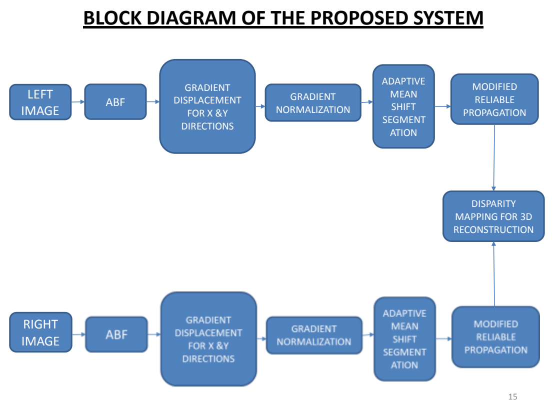
I. INTRODUCTION

The algorithm extracts the depth structure of the scene in an automatic way, representing an important saving in time and cost. For this purpose, many advanced depth extraction algorithms have been proposed that estimate the depth structure from defocus, motion, or shading. The main problem of these methods is that they cannot be universally applied to all the images that compose a movie (or generic video content). Stereo matching continues to be an active survey area as is proven by a large number of recent publications dedicated to this topic.The 2D-to-3D conversion process is typically performed in two main steps .The first one is the depth estimation from a given monocular image, and the second one is the DepthImage-Based Rendering (DIBR) of a new image or images to form a stereo pair, or a multi-view set of images. For this rendering step, there exists many algorithm that generate good quality results, while the depth estimation from single images is still a more challenging process. For this reason, this paper focuses on inferring the depth information from 2D images. Recently, a new family of methods based on a machine-learning approach have appeared, which automatically estimate the depth of an image or video sequence by processing a repository of color + depth images (RGBD) that is used as prior knowledge.In general matching algorithms can be characterized into local and global methods. Local approaches are utilizing the color or intensity values within a finite window to determine the disparity for each pixel. Global approaches are incorporating explicit smoothness assumptions and determine all disparities simultaneously by applying energy minimization techniques such as graph cuts, edge propagation, dynamic programming, scan line optimization or simulated annealing. Our objective is to remove the mismatches caused by both occlusions and false targets. In our approach, the stereo matching problem is examined as an optimization problem. [1] Shweta Patil et al proposed state of the art methods of 2D to 3D image and/or video conversion. In this modern era popularity of 3D hardware is increased but, 3D contents are still dominated by its 2D equivalent. Till now many researchers have proposed different methods to close this gap. Mainly, these conversion methods are classified into automatic method and semi-automatic method. In automatic method human intervention is not involved, where as in semi-automatic method human operator is involved. There are distinct characteristics that can be considered during conversion, like for video conversion motion and optical flow are mostly considered parameter, and for image conversion local attributes of images are considered. Computational time and design cost are the main design parameters that should be considered while designing algorithm. Each method is having its own advantages and disadvantages depending on specification which method will be suitable for conversion is decided. [2] Janusz Konrad; et al proposed Despite a significant growth in the last few years, the availability of 3D content is still diminished by that of its 2D equivalent. To close this gap, multiple 2D-to-3D image and video conversion methods have been proposed. Methods involving human operators have been most successful but additionally time-consuming and costly. [3] Moshe Guttmann; et al proposed a semi-automatic system that converts conventional video shots to stereoscopic video pairs. The system uses just a few user-scribbles in a sparse set of frames and also combines a diffusion scheme, which takes into account the local saliency and the local motion at each video location, connected with a classification scheme that assigns depth to image patches. The system tolerates both scene motion and camera motion. In typical shots, contained hundreds of frames, even in the face of significant motion, it is enough to mark scribbles on the first and last frames of the shot. Once if marked, plausible stereo results are obtained in a matter of seconds, leading to a scalable video conversion system. Finally, we ratify our results with ground truth stereo video. [4] Chao-Chung Cheng; et al proposed even though three-dimensional (3D) displays that enhances the visual quality more than two-dimensional (2D) displays do, the depth information required for 3D displays is unavailable in the conventional 2D content. Therefore, converting 2D videos into 3D ones has become an important issue in rising 3D applications. This work gives a novel algorithm that automatically converts 2D videos into 3D ones. The proposed algorithm uses the edge information to segment the image into object groups. A depth map is then assigned based on the assumption of depth gradient model. Subsequently the depth map is block-based assigned by cooperating with a cross bilateral filter to generate visually comfortable depth maps efficiently and also diminish the block artifacts. A multiview video can be readily obtained by using a depth image-based rendering method.[5] Raymond Phan et al proposed a semi-automated method for converting conventional 2D images into stereoscopic3D.User-defined strokes corresponding to a rough approximation of the depth values in the scene are defined for the image of interest. With these, our system certify the depth values for the rest of the image, producing a depth map which can be used to create stereoscopic 3D image pairs. Our work is build on a similar scheme, using the Random Walks segmentation paradigm. Though, the related work is quite complex, with many processing steps required to produce the final stereoscopic image pair. Combined with its evident shortcomings, but noting the merits, we have proposed a system employing Random Walks, while incorporating information from the popular Graph Cuts segmentation paradigm. Thus, a final integrated depth map is produced, combining the merits of both. The results show that we can produce good quality stereoscopic image pairs, while using a much more simplified method in comparison to the related word. [6] Miao Liao; et al proposed a semiautomatic system that converts conventional videos into stereoscopic videos by combining motion analysis with user interaction, aiming to shift as much as possible tag work from the user to the computer. In addition to the widely used structure from motion (SFM) techniques, we have developed two new methods that examine the optical flow to provide additional qualitative depth constraints. They removed the camera movement restriction imposed by SFM therefore the general motions can be used in scene depth estimation-the central problem in mono-to-stereo conversion. With these algorithms, the user's labelling task is remarkably simplified. We further developed a quadratic programming approach to consolidate both quantitative depth and qualitative depth (such as these from user scribbling) to recover dense depth maps for all frames, from which stereoscopic view can be synthesized. In addition to visual results, we present user study results showing that our approach is more intuitive and less labor intensive, while producing 3D effect comparable to that from current state-of-the-art interactive algorithms. [7] Ruo Zhang; et al proposed being the first shape-from-shading (SFS) technique was developed by Horn in the early 1970s, many different approaches have emerged. In this paper, six well-known SFS algorithms are achieved and compared. The accomplishment of the algorithms was examined on synthetic images using mean and standard deviation of depth (Z) error, mean of surface gradient (p, q) error, and CPU timing. Each algorithm works well for certain images, yet performs poorly for others. In general,

minimization approaches are more sturdy, while the other approaches are faster.
[8] Pan Ji; et al proposed the 2D to 3D conversion technique which plays a crucial role in the development and promotion of three-dimensional television (3DTV) for it can provide appropriate supply of high-quality 3D program material. In this paper, a novel automatic 2D to 3D conversion method using multi-depth cues is conferred. The depth cues used in our system that will be integrated into one depth map based on the types of the 2D scenes, which include perspective geometry, defocus, visual saliency and adaptive depth models. After the depth maps are extracted, the original 2D image or video is changed to stereoscopic one for showing on 3D display devices. Our method is proved on various sequences and the experimental results show that the resulting image or video is both realistic and visual pleasing. [9] Murali Subbarao; et al proposed new method named STM which is used for determining distance of objects and rapid auto focusing of camera systems. STM uses image defocus data which is based on a new Spatial-Domain Convolution/Deconvolution Transform. The method needs only two images taken with different camera parameters such as lens position, focal length, and aperture diameter. Both images can be randomly blurred and neither of them needs to be a focused image. [10] Fayao Liu; et al proposed, we handle the problem of depth estimation from single monocular images. Compared with depth estimation using multiple images such as stereo depth perception, depth from monocular images is much more difficult.

II. PROPOSED METHOD

Fig 1. The block diagram of proposed work



1. **Adaptive Bilateral Filter(ABF)**
2. **Image enhancement using Mean adjustment**
3. **Modified EM based on Modified Canny Edge Detection**
4. **Modified Reliable Disparity Propagation Algorithm For Depth Extraction**

1. Adaptive Bilateral Filter

Introduction of different types of noise in the image causes variation in brightness or colour information. The various noises such as Gaussian, Impulse, Speckle, Periodic and Poisson noise deteriorates the picture quality. Gaussian noise is the most popular one that arises during acquisition process caused by poor illumination and/or high temperature. This paper proposes an improved modified adaptive Bilateral Filter to remove Gaussian noise from color images. This technique has been implemented in MATLAB-9 and various performance metrics taken into consideration are: Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE), Mean Absolute Error (MAE), and Normalized Colour Difference (NCD). ABF is used for digital images whereas Bilateral filter is used for Square matrices.

2. Image Enhancement Using Mean Adjustment

Image enhancement is a computer image processing technique used to improve contrast in images. It differs from ordinary histogram equalization in the respect that the adaptive method computes several histograms, each corresponding to a different section of the image, and uses them to redistribute the lightness values of the image. It is therefore suitable for improving the local contrast of an image and bringing out more detail.However, AHE has a tendency to over-amplify noise in relatively homogeneous regions of an image. A variant of adaptive histogram equalization called image enhancement using mean adjustment prevents this by limiting the amplification.

3. Modified Expectation Maximization Segmentation Based On Modified Canny Edge Detection

Modified Expectation Maximization is used to estimate the gradient of the image. An image gradient (High frequency region) is the directional change in the intensity or color in an image. Image gradients shall be used to extract information from images.
The algorithm runs in 5 separate steps:
1. Smoothing: Blurring of the image inorder to remove noise.
2. Finding gradients: The edges can be marked where the gradients of the image has large magnitudes.
3. Non-maximum suppression: Only local maxima should be marked as the edges.
4. Double thresholding: Potential edges are determined by means of thresholding.

5. Edge tracking by hysteresis: Final edges are determined by suppressing all edges which are not connected to a very clear (strong) edge.

4. Modified Reliable Disparity Propagation Algorithm For Depth Extraction

The intensity based area correlation techniques have been investigated extensively for commercial applications in stereo vision but are also one of the newest methods used in computer vision. The algorithm that we are about to demonstrate here falls in the category of the correlation based stereo algorithms.The tokens which are used in the correspondence process are the image pixels themselves with one feature the intensity at the pixel The algorithm uses two geometric constraints the epipolar constraint to reduce the search for correspondences and the constraint that the disparity or depth is locally constant in the vicinity of a pixel The second limitation is clearly only an approximation. The algorithms also has the limitation that the intensities at two corresponding pixels are almost the same .This is in fact a physical constraint related to the hypothesis that the observed objects are approximately.

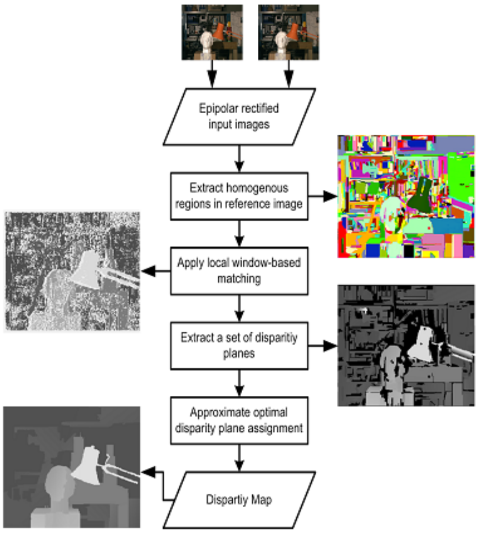


Fig 2 Flowchart of segment based Stereomatching

Algorithms augmented input data,intermediate and final results of the Proposed method

III. RESULTS

Adaptive bilateral filter

Fig 3 Input image

Fig 4 Output image

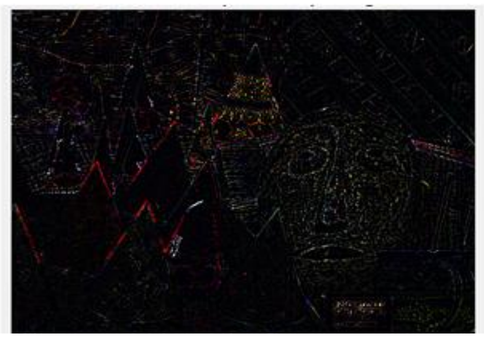


Fig 5 Difference between the input and Output image3. It removes the noises present in the input image

Fig 6 Image Enhancement Using Adaptive Mean Adjustment

Fig 7 shows the total gradient of image 3 **Fig 8** Modified Expectation Maximization Segmentation

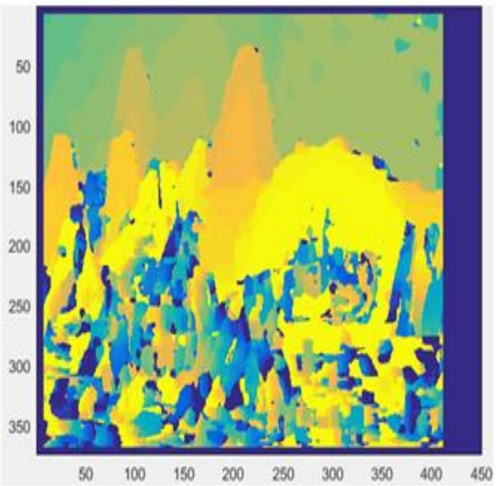
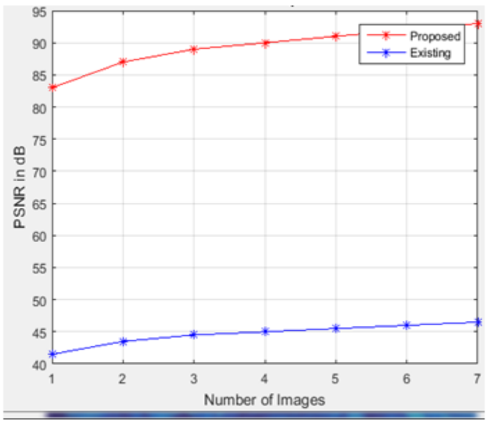


Fig 9 Modified relief propagation algorithm **Fig 10** 3D image reconstruction



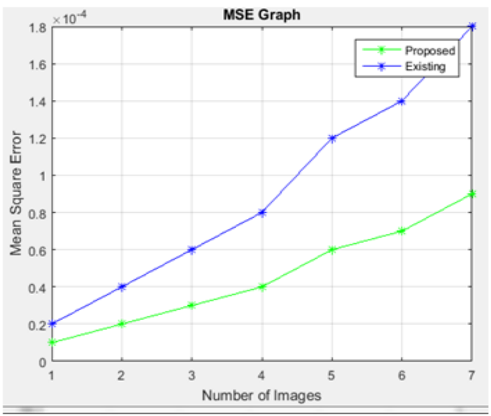


Figure 11 shows the PSNR comparison graph between the proposed and the existing system. The proposed method shows a greater value of PSNR compared to the existing system which indicates better image quality
Figure 12 shows the Mean Square Error Comparison graph between the proposed and the existing system. The proposed system shows error which is closely equal to zero

IV. CONCLUSION

Stereo Vision is becoming increasingly important for emerging embedded system applications that require real time disparity map computation. This paper presented hardware architecture for real-time disparity map computation from stereo images, targeting embedded stereo vision applications. The architecture integrates an edge detector that speeds up the system performance, and was designed with emphasis on a number of configuration parameters. A prototype implementation indicates real-time performance (50 fps for 1,280 1,024 input image frames) under a variety of configurations, while maintaining acceptable quality on the disparity maps. Our immediate plans involve the integration of a rectification and interpolation step, both as custom hardware units.

REFERENCES

[1]. S. Patil and P. Charles, "Review on 2D-to-3D image and video conversionmethods," in International Conference on Computing CommunicationControl and Automation, Feb 2015, pp. 728–732.

[2]. J. Konrad, M. Wang, P. Ishwar, C. Wu, and D. Mukherjee, "Learningbased,automatic 2D-to-3D image and video conversion," IEEE Transactionon Image Processing, vol. 22, no. 9, pp. 3485–3496, 2013.

[3]. M. Guttmann, L. Wolf, and D. Cohen-Or, "Semi-automatic stereo extractionfrom video footage," in International Conference on ComputerVision, Sept 2009, pp. 136–142.

[4]. C.-C. Cheng, C.-T. Li, and L.-G. Chen, "A novel 2D-to-3D conversionsystem using edge information," IEEE Transactions on Consumer Electronics,vol. 56, no. 3, pp. 1739–1745, 2010.

[5]. R. Phan, R. Rzeszutek, and D. Androutsos, "Semi-automatic 2D to3D image conversion using scale-space random walks and a graphcuts based depth prior," in IEEE International Conference on ImageProcessing, Sept 2011, pp. 865–868.

[6]. M. Liao, J. Gao, R. Yang, and M. Gong, "Video stereolization: Combiningmotion analysis with user interaction," IEEE Transactions onVisualization and Computer Graphics, vol. 18, no. 7, pp. 1079–1088,2012.

[7]. R. Zhang, P.-S. Tsai, J. Cryer, and M. Shah, "Shape-from-shading: a survey,"IEEE Transactions on Pattern Analysis and Machine Intelligence,vol. 21, no. 8, pp. 690–706, 1999.

[8]. P. Ji, L. Wang, D.-X. Li, and M. Zhang, "An automatic 2D to 3D conversionalgorithm using multi-depth cues," in International Conferenceon Audio, Language and Image Processing, July 2012, pp. 546–550. 13, no. 3, pp.271–294, 1994.

[9]. M.Subbarao and G.Surya, "Depth from defocus:Aspatialdomainapproach," International Journal of Computer Vision, vol. 13, no. 3, pp. 271–294, 1994. Fayao Liu, Chunhua Shen,

[10]. "Learning Depth from Single Monocular Images Using Deep Convolutional Neural Fields" IEEE T. Pattern Analysis and Machine Intelligence.,feb 20l5.

S. Sathiya Priya" 3d Image Reconstruction Using Automatic Depth Extraction from 2d Images Using Machine Learning " International Journal of Engineering Research and Applications (IJERA), Vol. 09, No.03, 2019, pp. 38-43