Enhanced 2-Dimensional to 3-Dimensional Conversion for Medical Images

C. Saranya, G. Padmavathi

***Abstract*--The advent of modern 3D technological devices and the desire to create 3D images from the numerous 2D images have increased tremendously. Image-based 3D-modeling techniques for creating a 3D representation of a scene from one or more 2D images have received great attention and methods to improve the conversion process have been probed by both academicians and researchers.**

This research study focuses on particular conversion method recommended that facilitates various image processing algorithms to improve the conversion of 2D images to 3D images. The primary steps involved are Motion, Edge detection and image segmentation, Depth estimation and Shift algorithm. A weighted motion detection registration method is used to calculate the difference between the current image frame and the previous image frame. During edge detection, Sobel edge detector is used to detect the edges. The result of motion detection and edge detection are combined together and then a gray level closing is performed to make the edges connected and smooth.

An edge registration module is used to store the motion and edge information in the memory. The segmentation process uses two algorithms, namely, K-Means and Mean Shift. An enhanced connected component algorithm which improves the traditional algorithm to use Max-Tree is used to create refined components. The final step of the proposed algorithm is the shift algorithm, which reconstructs the 3D image.

To prove the efficiency of the proposed algorithm, several experiments are conducted. Various parameters like Root Mean Square Error (RMSE), Peak Signal to Noise Ratio, Standard Deviation and Speed of conversion are used to analyze the performance and efficiency of the proposed conversion algorithm. The experimental results proved that the depth map generated by mean-shift algorithm and enhanced connected component produce efficient and improved results.

Index Terms: RMSE, Dept Estimation, Edge detection, Shift Algorithm, Sobel Edge Detector

I. INTRODUCTION

In medical domain, 3D imaging has become one of the highest growth segments in the field of medical imaging. They have the potential to change the way of interaction between a patient's human body and the doctor. In recent years, medical imaging systems have envisaged improved computing power, better display technology, improved resolution and improved speed of processing.

This in turn has made 3D images popular in various medical applications like computer guided & robotic surgery, cardiology, sports medicine, pediatrics and urology. In order to capture a 3D scene, special equipments such as stereo or multi-view cameras and a depth camera are normally needed [10].

Manuscript Received 25 October, 2011.

C. Saranya, M.Phil Research Scholar, Department of Computer Science, Avinashilingam Deemed University for Women, Coimbatore-43., saranya86kcs@gmail.com

Dr. G. Padmavathi, Professor & Head, Department of Computer Science, Avinashilingam Deemed University for Women, Coimbatore-43., ganapathi.padmavathi@gmail.com

Even if 3D image contents have been produced and become available, the amount of 3D contents is not enough to satisfy the user demand yet. On the other hand, there are abundant 2D image contents captured by conventional single-view cameras. Moreover, the rapid development of image and video technology along with its three-dimensional (3D) counterparts, people are increasingly not satisfied with the ordinary two-dimensional (2D) image or video. Hence, generation of 3D scenes from 2D contents is an alternative and attractive solution that overcomes the current discrepancy and fills up the lack of 3D image contents. The most common method is to obtain the corresponding depth information from the existing 2D image using which the 3D image is generated based on human visual principles. The key point of the conversion algorithm is how to get more accurate depth information of a 2D image in a fast manner and how to optimize the algorithm that generates the depth map.

This paper focus on proposing an algorithm that uses various techniques to enhance the conversion process. The goal is to design an automatic system that require less human intervention, provide better visualization of 3D scenes in a fast and accurate manner, so that they can operate with more accuracy in unknown environments. To achieve this aim, this work develops depth map estimation method for 2D to 3D conversion of medical images using short-term motion assisted color segmentation, which combines the pictorial, monocular and binocular depth cues of human vision [2]. The proposed method utilizes a motion/edge registration technique and connected component algorithm for depth map generation. The work performs clustering before connected component analysis to increase accuracy and speed. Two clustering algorithms, namely, k-means clustering algorithm and Mean-shift clustering algorithm are used to create blocks or regions. Further, the connected component algorithm is enhanced to use a MaxTree data structure [8] to save space.

The rest of the paper is organized as below. Section 2 presents a previous works related to the topic of discussion. Section 3 presents the proposed methodology. The proposed converter was tested with various images and the results were evaluated and compared with existing methods. The results obtained are presented in Section 4. Section 5 summarizes the work with future research directions.

II. LITERATURE REVIEW

Although the main focus of the study is on depth estimation from a single or multiple still images, there are many studies that have proposed techniques for constructing 3D images from 2D counterparts.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

This section presents a short review of these techniques. The 3D reconstruction techniques can be categorized into three groups, namely, explicit measurements with laser or radar sensors 3D reconstruction [14], 3D reconstruction using two (or more than two) images [16] and 3D reconstruction using video sequences [3]. Among the vision-based approaches, most work has focused on stereovision and on algorithms that require multiple images, such as optical flow [1], structure from motion [7] and depth from defocus [5]. A large class of algorithms reconstructs the 3D shape of known objects, such as human bodies, from images and laser data [19]. Structured lighting [17] offers another method for depth reconstruction. There are some algorithms that can perform depth reconstruction from single images in very specific settings. Nagai *et al.* [13] performed surface reconstruction from single images for known, fixed, objects such as hands and faces. Torresani and Hertzmann [20] worked on reconstructing non-rigid surface shapes from video sequences. Michels *et al.* [12] used supervised learning to estimate 1D distance to obstacles, for the application of autonomously driving a small car. Delage *et al.* [6] generated 3D models of indoor environments containing only walls and floor, from single monocular images. Single view metrology [4] assumes that vanishing lines and points are known in a scene, and calculates angles between parallel lines to infer 3D structure from Manhattan images. Even though several studies have been performed in the field of 3D image reconstruction, it is clear that the field has not yet reached its maturity.

III. PROPOSED METHODOLOGY

The proposed algorithm consists of three key parts, (i) Motion/edge detection and image segmentation (ii) Depth estimation and (iii) Shift algorithm. The procedure of the first two steps is illustrated in Fig 1. The input consists of multiple frames indicating the left and right eye movement. A frame buffer is used to store the previous frame information, and the information is sent into the motion and edge detection part for motion extraction. Motion detection methods calculate the difference between the current image frame and the previous image frame and are generally termed as ‘short-term motion detection’. These methods even though efficient are sometimes affected by motion jitters, at which time, the performance degrades. To solve this problem, a weighted registration is used on the difference calculation (Equation 1).

$$D_{current}=Y_{current}-Y_{previous}+D_{previous}w \quad (1)$$

where D denotes the motion difference, Y denotes the luminance value and w denotes the weighting factor. Through the above difference equation, the motion difference is memorized for a short time with respect to ‘w’ and then the motion information would not suffer from the motion jitter.

In the edge detection part, Sobel edge detector is used to extract the edges. The result of motion detection and edge detection are combined together, and then a gray level closing is performed to make the edges connected and smooth. An edge registration module is used to store the motion and edge information in the memory. At the same time, the segmentation process also begins. The segmentation step uses the traditional K-Means and mean-shift algorithms. After image segmentation, the original image is divided into a number of regions, each with individual label number. The problem in this step is that the after segmentation, different objects are associated to the same region. This circumstance represents the lack of spatial information. To solve this, a connected component searching method is used to determine the similar colour components in the segmented region. The traditional connected component algorithm has the following disadvantages.

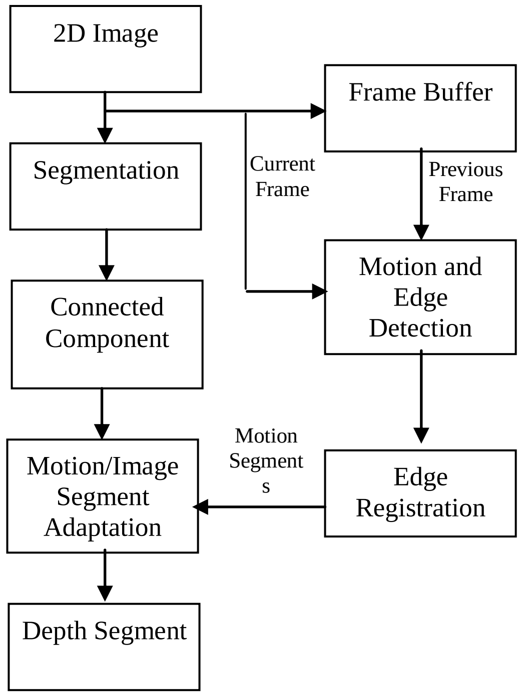


Fig 1 : Motion/edge detection, image segmentation, Depth estimation

1. Creates glitches (unusual patterns) in the resultant image that makes it difficult to use directly for depth map creation.
2. Large Computation overhead because of the distance calculation and comparison has to be performed for each and every pixel in all the segments has to be computation. This overhead in turn decreases the algorithm speed.

The first problem is solved by using Range Estimation and Component Separation (RECS) algorithm (Section 3.1) and the second problem is solved by using a MAX-Tree data structure (Section 3.2.). The proposed method then combines the refined components and motion segments to obtain the depth map. The last step of the proposed algorithm is the shift algorithm, which reconstructs the 3D image. After obtaining depth estimation, the left-view and right-view image can be obtained by stereovision characteristic. The left and right views are obtained using image depth information and by employing the characteristics of human vision. In order to obtain the right-view image, the pixel’s position with respect to the plane is considered.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

If the pixel is close to the zero plane, then it will have a leftwards shift, else a rightward shift is performed. If it is on the zero plane, then its position is left unchanged. A median smoothening filter is then applied to fill in the holes obtained through the shift procedure. The resultant image is the converted 3D image.

A. Range Estimation and Component Separation (RECS) algorithm

The RECS algorithm consists of two steps. The first step uses connected component refinement procedure to refine components so as to reduce the glitches (Fig 2a) and the second step enhances the components by using range estimator (Fig 2b). After the motion/edge detection and connected component, the position and range of the moving objects are estimated by the range estimation. It performs raster-scan and inverse raster-scan on the image with the following procedures. If there is an edge registered in the neighbor of the current scanning pixel, the pixel is recognized as a pixel in the estimated range, which is shown in Fig 2b.

A component separation algorithm separates one component into two if there are two connected components occluded with each other. If a component is inside the estimated range and is surrounded by both image and motion edge segments, the component is set as the foreground component and its depth value is assigned to the nearest depth. Furthermore, if combining with a depth from geometry perspective method such as the vanishing point detection [18], detailed depth can be assigned on the objects to get a fine depth map.

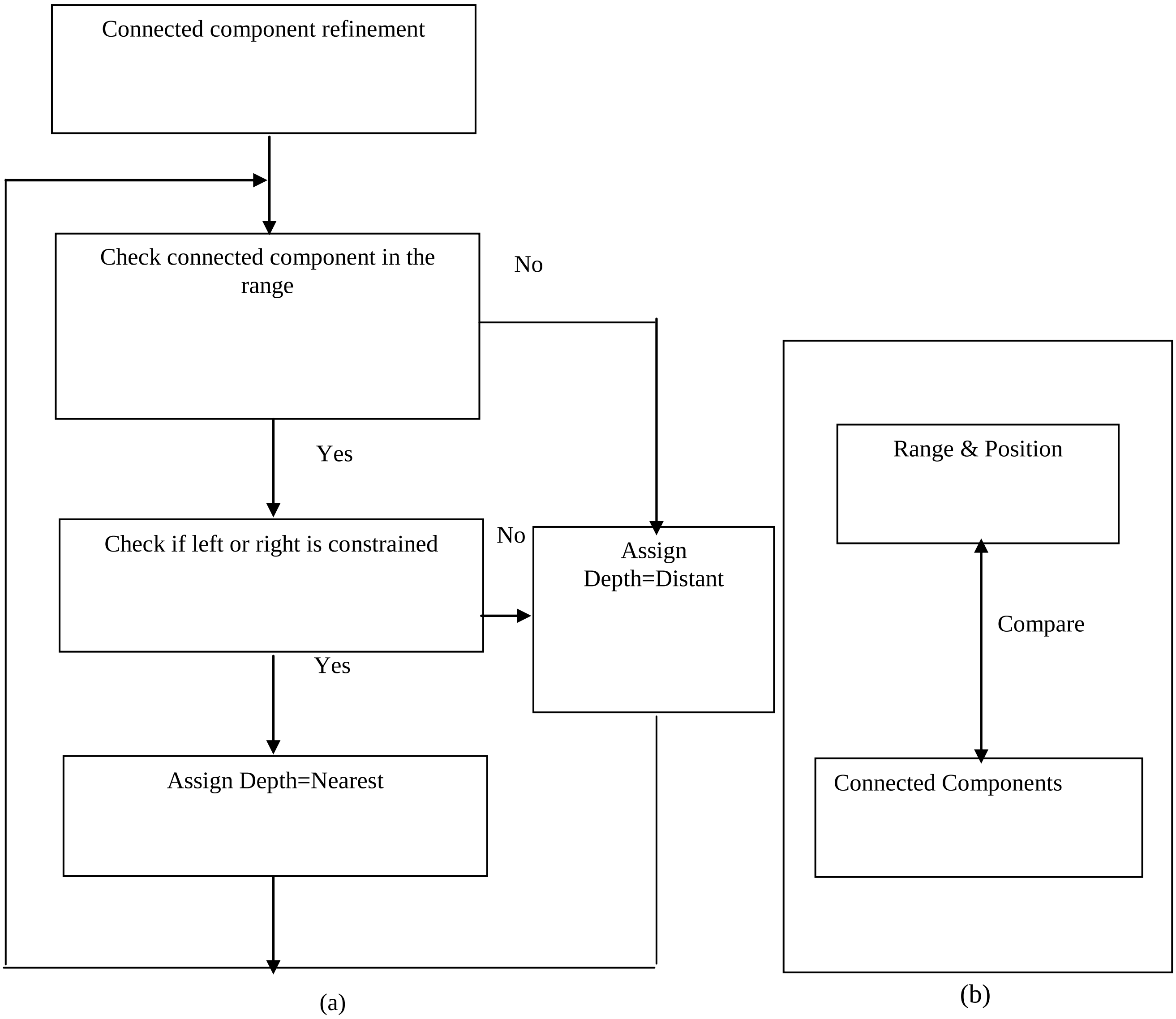


Fig 2: RECS Algorithm (a) Procedure (b) Range Estimation

B. X-Tree based Connected Components

The Max-Tree based connected component algorithm takes every pixel in an image as a node and the nodes that belong to the same connected area are grouped together to build a tree, in such a way, that the different connected areas are represented by the different trees. Here, a node structure is defined as below.

```
Node{
        Coordinates(x,y);
        Bool color;
        Int size;
    }
```

An image is considered as a two-dimensional array with (x, y) denoting a pixel's row and column position, color is used to represent whether or not the pixel is foreground. Size is always set to 1, except for root node, which will have the connected area's size. The procedure scans the image from left to right and then top to bottom. The first pixel is treated as root node with size one. The four neighbors of the pixel as shown in Fig 3 are considered. Let P be the pixel under consideration (Centre area of Fig 3) and P1 be the neighboring node (Shaded area of Fig 3). If both P and P1 have the same colour property, then the same root is shared by both P and P1 and the root node size is incremented by 1.

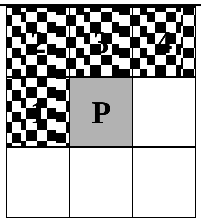


Fig 3 : Scanning Process

The main advantage of using Max-Tree representation for connected components is to avoid the repeated scanning of all neighbors during distance calculation while maintaining the integrity. A Max-Tree is an inclusion tree that considers connected components of upper-level sets of an image. A node of the Max-Tree corresponds to a connected component C_0 of an upper-level set. Its parent is defined as the connected component C_1 of the higher upper-level set such that $C_0 \neq C_1$ and $C_0 \subset C_1$. The leaves of a Max-Tree correspond to the regional maxima of the image. As reported in [15], connected attribute filters for grayscale images can be defined as a pruning of its Max-Tree, in such a way that the threshold decomposition property [9] holds. This enables the straightforward reconstruction of the filtered image from the pruned tree. Given a non-increasing criterion, many different pruning rules can be considered, each leading to a different connected-set filter [21]. The rules used in this study are described below.

All the neighbors of a pixel are considered to be scanned only if the current pixel and its neighbor are in the same tree. The current pixel can either belong to foreground or background. Using this information, the rules are formulated. When a current pixel p(x, y) has been scanned, the number of neighbouring pixels scanned depends on whether it is a foreground or background. If it is foreground then four neighborhood pixels are considered, if it is a background, then only two neighborhood pixels are considered. The rules are

1. If the current pixel, first and second neighbouring pixels are foreground, then combine current pixel, first and second neighborhood pixels in the same tree and check only the fourth pixel
2. If the current pixel, third and fourth neighboring pixels are foreground, then all the pixels (current and four neighbours) can be combined into the same tree and a scan need not be performed.
3. If the current pixel and its first neighbor is a background, a scan need to be performed only on the second neighbor and the rest can be grouped under the same tree.
4. If the current pixel and its second neighbor is a background, both can be grouped together and a scan need not be performed.

These rules are framed based on the equivalence relationship that exists between the current pixel and its neighbors. An equivalence relation is a relation that partitions a set so that every element of the set is a member of one and only one cell of the partition. Two elements of the set are considered equivalent (with respect to the equivalence relation) if and only if they are elements of the same cell. The intersection of any two cells is empty; the union of all the cells equals the original set. Two pixels are said to be equivalent, if they belong to the same region of an image.

IV. EXPERIMENTAL RESULTS

This section presents the image dataset used and the various performance metrics used to analyze the performance of the proposed algorithm. The experiments were planned in two stages. The first stage is used to evaluate the two segmentation algorithms, namely, K-Means and Mean-Shift. The second stage evaluates the 3D-2D converter and compares the result with the existing method. The existing method use K-means segmentation algorithm with traditional connected component algorithm. The proposed method use Mean-Shift algorithm with enhanced connected component algorithm with Max-Tree.

In order to verify the performance of the proposed algorithm, experimental evaluation is performed using a series of different benchmark images. Six MRI images in varied sizes were randomly selected (Fig 4) to analyze the scalability and flexibility of the algorithms.

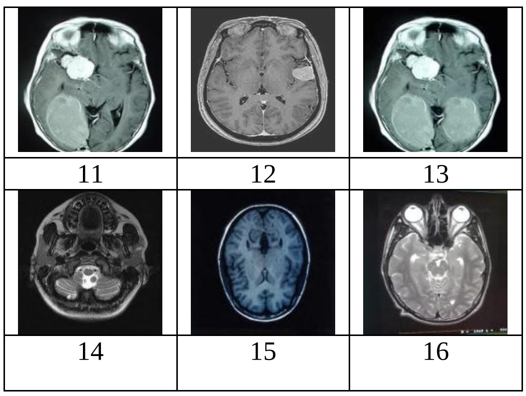


Fig 4 : Test Images

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Performance evaluation of the methods developed is the most important step in any research. Different researchers use different parameters for analysis. In general, all the works with a common aim of finding whether the proposed method works efficiently and shows improvement over the existing methods. In stage 1, two parameters, namely, visual comparison and segmentation time were selected to evaluate the two selected segmentation algorithm. In Stage 2, four metrics were used to analyze the proposed conversion algorithm. They are Root Mean Square Error (RMSE), Peak Signal to Noise Ratio (PSNR), and Speed of Conversion.

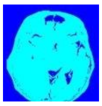
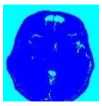
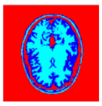
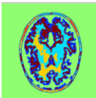
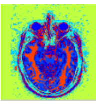
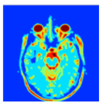
Image	K-Means	Mean-Shift
I1		
12		
I3		
14		
I5		
16		

Fig 5. Segmentation Results

A. Stage 1 Results

Two segmentation algorithms, namely K-means and Mean-Shift algorithms were considered to be used in the proposed algorithm. The main objective of this work is to identify an algorithm for segmenting medical images. As each medical image have different characteristics, using the same segmentation technique for all types of images is impossible. The present method is developed for MRI image, but the same can be tested for other images also. Fig 5 shows the result of segmentation on the test images.

For segmentation algorithm to be perfect, there should be clear distinction between the various regions of the image and the edges of these regions has to be identified unmistakably [11]. Supporting this theory, the proposed model segments the image in a more accurate fashion, by dividing it into more regions separating it using random colours. The edges of the image are more evident in the proposed model while the same cannot be held good for the base model, which clearly supports the argument that the segmentation process is more reliable and accurate in the proposed system.

Table I : Segmentation Time

Image	Mean-Shift	K-Means
I1	0.88	1.43
I2	1.12	1.94
I3	1.41	1.61
I4	1.59	1.89
I5	0.94	1.67
I6	0.99	1.81

Table I shows the segmentation time of the selected algorithms. The results prove that the mean-shift algorithm is efficient in terms of segmentation speed also. The mean-shift algorithm shows 37.19% time efficiency when compared with K-means algorithm. All these results prove that the performance of mean-shift algorithm shows significant improvement in terms of segmentation accuracy, speed when compared to K-Means algorithm.

B. Stage 2 Results

Table II shows the Root Mean Square Error (RMSE) obtained when comparing the existing algorithm (with K-means and traditional connected component algorithm) and the proposed algorithm (with mean shift and enhanced connected component algorithm). From the results, it could be seen the performance of the proposed algorithm shows significant improvement in terms of RMSE, indicating that the proposed version produces accurate 3D conversion. The results again shows that the proposed algorithm as an improved version to the existing algorithm. While considering RMSE, the proposed algorithm showed 10.81% efficiency gain over the existing algorithm.

Table II : RMSE

Image	Existing	Proposed
I1	0.316	0.274
I2	0.397	0.311
I3	0.374	0.299
I4	0.518	0.482
I5	0.623	0.597
I6	0.339	0.301

Peak Signal to Noise Ratio is a measure of the peak error and is a benchmark method used to evaluate the quality difference between two similar images. Table III shows the Peak Signal to Noise Ratio (PSNR) of the proposed and existing 2D to 3D conversion algorithms.

Table III : PSNR

Image	Existing	Proposed
I1	28.42	34.40
I2	29.63	35.25
I3	30.12	36.09
I4	29.44	35.58
I5	29.94	36.16
I6	30.43	36.69

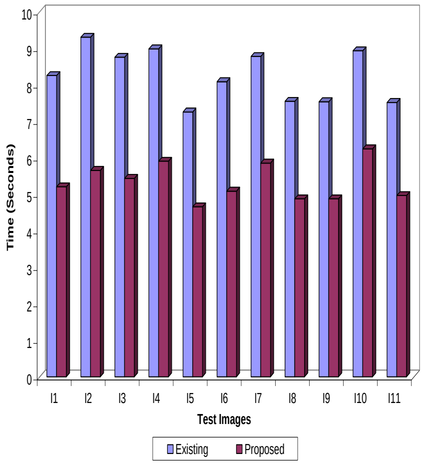


Fig 6: Time Taken for existing and proposed method

The high PSNR values obtained (34.40-36.69dB) by the proposed system when compared to existing algorithm (28.42-30.88dB) indicate that it is an optimal choice for converting the 2D image to 3D counterparts. The proposed algorithm is efficient by 16.49% than the existing algorithm while considering PSNR performance metric. The final parameter considered is the time parameter, which is used to analyze the speed of the proposed and existing algorithms. The results obtained are shown in Fig 6. Execution time results again project the fact that the proposed converter is faster than the existing algorithm. The average speed is 8.27 seconds and 5.35 seconds by the existing and proposed converter, thus showing an efficiency time gain of 35.38%. The converted 3D images for the selected test images are shown in Fig 7.

V. CONCLUSION

3D imaging has become one of the highest growth segments in the field of medical imaging and has the potential to change the way of interaction between a patient's human body and the doctor. Due to the overwhelming amount of 2D images available, a common method used is to convert 2D images into a 3D form. The primary steps involved are Motion/edge detection and image segmentation, Depth estimation and Shift algorithm. A weighted motion detection registration method was used to calculate the difference between the current image frame and the previous image frame. The input image is segmented using two algorithms, namely, K-Means and Mean-Shift.

Segmentation results with an image divided into regions.

To further refine the segmentation process an enhanced connected component algorithm that uses Max-Tree structure was used. The last step is the shift algorithm,which reconstructs the 3D image. The experimental results prove that the depth map generated by mean-shift algorithm and enhanced connected component is an improved conversion method and produce efficient and accurate results. In future, the edge detection algorithm, methods for parallelism were to be considered.

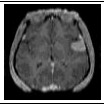
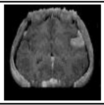
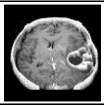
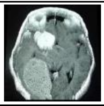
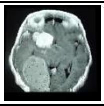
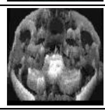
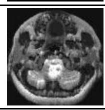
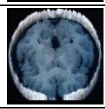
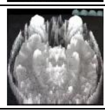
Image	K-Means	Mean-Shift
I1		
I2		
I3		
14		
15		
16		

Fig 7. Converted Results

REFERENCES

1. Barron, J.L., Fleet, D.J. and Beauchemin, S.S. (1994) Performance of optical flow techniques. International Journal of Computer Vision, Vol.12, Pp.43–77.
2. Chang, Y.L., Fang, C.Y., Ding, L.F., Chen, S.Y. and Chen, L.G. (2007) Depth map generation for 2D-to-3D conversion by short-term motion assisted color segmentation, ICME 2007, Pp.1958-1961.
3. Cornelis, N., Leibe, B., Cornelis, K. and Van Gool, L. (2006) 3d city modeling using cognitive loops, Video proceedings of CVPR (VPCVPR), Pp. 1-8.
4. Criminisi, A., Reid, I. and Zisserman, A. (2000) Single view metrology, International Journal of Computer Vision, Vol. 40, Pp.123–148.
5. Das, S. and Ahuja, N. (1995) Performance analysis of stereo, vergence, and focus as depth cues for active vision, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No.12, Pp.1213–1219.
6. Delage, E., Lee, H. and Ng, A. Y. (2005) Automatic single-image 3d reconstructions of indoor Manhattan world scenes, 12th International Symposium of Robotics Research (ISRR), Pp.305-321
7. Forsyth, D. A. and Ponce, J. (2003) Computer vision: a modern approach, New York: Prentice Hall.
8. Fu, Y., Chen, X. and Guo, H. (2009) A new Connected Component Analysis Algorithm Based on Max-Tree, Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing, IEEE Comp. Society, Pp.843-845.
9. Guichard, F. and Morel, J.M. (2002) Mathematical morphology "almost everywhere", Proceedings of ISMM, Csiro Publishing, Pp.293–303.
10. Jung, J.I. and Ho, Y.S. (2010) Depth map estimation from single-view image using object classification based on Bayesian learning, 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), Tampere, Pp.1-4.
11. Meer, P. and Georgescu. B. (2001) Edge detection with embedded confidence. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No.12, Pp. 1351–1365.
12. Michels, J., Saxena, A. and Ng, A. Y. (2005) High speed obstacle avoidance using monocular vision and reinforcement learning, 22nd international conference on machine learning (ICML), Pp. 593-600.

13. Nagai, T., Naruse, T., Ikehara, M. and Kurematsu, A. (2002) HMM-based surface reconstruction from single images, IEEE International Conference on Image Processing (ICIP), Vol. 2, Pp. II-561 - II-564.
14. Quartulli, M. and Datcu, M. (2001) Bayesian model based city reconstruction from high resolution ISAR data, IEEE/ISPRS joint workshop remote sensing and data fusion over urban areas, Pp. 58 – 63.
15. Salembier, P., Oliveras, A. and Garrido, L. (1998) Antiextensive connected operators for image and sequence processing, IEEE Trans. on Image Processing, Vol.7, No.4, Pp.555–570.
16. Scharstein, D. and Szeliski, R. (2002) A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, Internationa Journal of Computer Vision, Vol.47, No.1, Pp. 7–42.
17. Scharstein, D. and Szeliski, R. (2003) High-accuracy stereo depth maps using structured light, Computer vision and pattern recognition (CVPR), Vol.1, Pp.195-202.
18. Steinbach, E., Hanjalic, A. and Girod, B. (1996) 3D Motion and Scene Structure Estimation with Motion Dependent Distortion of Measurement Windows, Proceedings [EEE International Conference Image Processing, ' 96, Vol. 1, Pp. 61-64.
19. Thrun, S. and Wegbreit, B. (2005) Shape from symmetry, International conference on computer vision (ICCV), Vol. 2, Pp. 824.1831.
20. Torresani, L. and Hertzmann, A. (2004) Automatic non-rigid 3D modeling from video, European conference on computer vision, 299-312.
21. Wilkinson, M.H.F. and Westenberg, M.A. (2001) Shape preserving filament enhancement filtering, Proc. of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Pp. 770–777.

AUTHORS PROFILE

C.SARANYA, M.Phil Research Scholar, Avinashilingam Deemed University for Women. Area of specialization: Image Processing. Currently I am working on Image Conversion.

Dr. (Mrs).G. PADMAVATHI is the Professor and Head of the Department of Computer Science of Avinashilingam Deemed University for Women. Coimbatore. She has 23 years of teaching experience and one year of industrial experience. Her areas of interest include: Real time communication, Cryptography in network security. She has 140 Publications in her Research area. In presently she is guiding PhD's and M.Phil scholars. She has been profiled a various organizations her academic contributions. She is currently the principle investigator of four projects funded by UGC and DRDO. She is a life member of many Preferred Organizations CSI, ISTE, WSEAS, AACE, and ACRS.