International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459 (Online), An ISO 9001:2008 Certified Journal, Volume 3, Special Issue 1, January 2013)

International Conference on Information Systems and Computing (ICISC-2013), INDIA.

2D TO 3D CONVERSION FOR IMAGES USING HOUGH TRANSFORM

Sneha. K[1], T. Sheela[2]

[1] *PG Student, ME Computer and Communication, Sri Sai Ram Engineering College, Chennai 44;*
[2] *Professor, Department of IT, Sri Sai Ram Engineering College, Chennai 44.*

Snehanair88@gmail.com.
Sheela_saiit@yahoo.com.

Abstract

Three dimensional imaging is a common technique used in computer vision and radar imaging. The main challenges in 3D imaging is the noise content of the input image and degraded quality of the reconstructed image. In this paper a novel technique is introduced to get efficient conversion of 2D images to 3D. In the this technique after the standardization, the image it is filtered to reduce the noise content. Our method works on edge images obtained by applying the canny edge algorithm to the source image. In the transformation phase Standard Hough Transform (SHT) & Improved Hough Transform (IHT) transforms has been applied for the correctness and the precision of detecting lines. Image alignment and reconstruction consists of a synthesis of translations, rotation, and isotropic scaling. The proposed system is efficient in reconstructing the image and it ensures a clear boundary distinction and proper three dimensional view of the object.

***Index Terms*-- Gaussian Smoothing, Canny edge algorithm, IHT, SHT**

I. INTRODUCTION

Producing three dimensional images with proper brightness and contrast is a strong requirement in several areas like computer vision, radar imaging and biomedical image analysis. The world of 3D adds the third dimension of depth, which can be perceived by human vision. Human eyes are located at slightly different positions, and these can perceive different views of the real world scene. The brain can now reconstruct the depth information from these different views. A 3D display takes advantage of this phenomenon, creating two slightly different images of scenes and then resenting them to the individual eyes. By an appropriate disparity and calibration of parameters, a correct 3D perception can be realized.

The process of acquiring 3D information from the surrounding environment can be classified into three groups of methods :

- Active or Passive method.
- Image Based or Direct method.
- Monocular or Multiple Views.

An active 3D acquisition group includes methods that introduce a source of energy such as light or ultrasonic waves. In passive methods the process of recording an image does not alter the surrounding environment in any way. A camera flash unit does not emit structured light, but the direction of illumination or the location of shadows changes with the location of a camera.

Therefore, camera flash can be used as a factor that alters the environment. The second group divides the acquisition methods into ones that collect data in the form of images and direct methods that do not require such data representation like range sensors. The monocular 3D acquisition method defines a number of views for the data acquisition process.

3D information can be extracted from a single view. But this does not mean that one image is enough to recover the depth. For example, in the case of a range from focus approach, several images are taken from a single view. For each image, the focus will be different. By identifying the sharpest areas in each image, one can get the information about depth of an underlying scene. Other monocular methods like the range of brightness, attenuation or texture use some extra information that must be known in advance.

II. BACKGROUND

Among the several approaches capable of generating 3D content include active depth sensor, triangular stereo vision, and 3D graphics rendering. Active methods use active sensors, such as structured light and time-of-flight sensor, to retrieve depth maps. The triangular stereo vision requires multiple cameras to record the content. Above methods require specific devices and are only effective in generating new content, making them invisible for 2D contents.

2D contents that require time-consuming manual editing of the depth information have become a barrier to the mass market, necessitating the development of an efficient 2D-to-3D conversion system.

Most of the 2D-3D conversion algorithms work on depth information generated from a single view image and convert to a multi view image using the estimated depth maps. A depth map generation method is classified mainly into a single frame and multi frame methods. Single frame methods include depth assignment using image classification machine learning, depth from focus or defocus depth from geometric perspective and depth from relative height. The depth is assigned based on the object properties of different regions.

2D to 3D depth generation algorithms generally face two challenges. One is the depth uniformity in the same object. The other challenge involves retrieving an appropriate depth relationship of all objects. To overcome these two challenges, this work presents a novel algorithm that uses the depth cues pattern texture and linear perspective. This algorithm extracts the pattern texture cue by edge detection and linear perspective by vanishing point detection and gradient plane assignment. Gaussian smoothing is applied to enhance the visual quality.

The rest of this paper is organized where Section 3 describes the existing 2d to 3d conversion systems. Section 4 deals with the proposed technique, 2D to 3D conversion system. Section 5 includes conclusion along with recommendations for future research.

III. Related Work

In early stages estimated depth maps are used to extract the information from a 2D image. Battiato [5] derived depth maps by using multiple cues from a single image. Hoiem [6] created depth maps using machine learning algorithms, in this the depth of objects are assigned based on the trained classes.

Sanyo's 2D to 3D Conversion Adaptive Algorithm uses a Computed Image Depth Method (CID) which combines image segmentation with depth extraction. In [2] the edge information is used to segment the 2D image into objects. A depth map is then generated by hypothesize depth gradient mode and a bilateral filter is used for efficient imaging.

The asymmetric edge adaptive filter can generate depth maps with comparatively accurate object edges, by avoiding the high computation after the introduction of depth generation method.

It can reduce the area of artifacts in rendering views by using asymmetric smoothing of depth maps, ensuring an improvement in the image quality with reduced artifacts and distortions.

The light scattering model uses atmosphere scattering for the depth map generation. This algorithm is used for estimating the depth map of outdoor images having a portion of the sky. When a light ray gets reflected because of a scattering medium, scattering of the ray may take place at different points of the light ray, and it will change the light rays original path. When light reaches the viewer from all points of the light ray, the viewer actually gets the effect of a new light source.

The energy minimization model uses shading for the depth map generation. It is based on finite elements. The image is divided into small triangular areas. The reflectance map R (p, q) is then approximated using a linear function. The depth map is generated by converting the energy minimization model into a problem of the form of solving a linear equation until a specified error threshold is reached.

The frontal Texel method uses the depth cue patterned texture for a depth map generation. Patterned texture gives a good 3D impression because of the two properties: the distortion of each Texas and the rate of change of Texel distortion through the image region. Loh and Hartley [9] proposed a method suitable for perspective views. The algorithm proceeds by making the frontal Texel as the reference point. The frontal texel will be unique. Any incorrect hypothesis of the frontal Texel leads to incorrect estimates of the surface orientation, it cannot be realized by a reconstructed surface. So a search through all possible frontal texels with the surface consistency constraint will help in generating a unique fontal texel estimate.

Statistical estimators mode uses the depth cue statistical patterns in the depth map generation. The mean depth of an image can be estimated by machine learning, then it is possible to estimate the depth value per each pixel. Several sets of global features and local features, such as texture variations, texture gradients, haze, image features are collected. The parameters are estimated using the maximum likelihood estimator and linear regression based on the training data and then the depth map is generated.

EXPLORING RESEARCH AND INNOVATIONS

International Journal of Emerging Technology and Advanced Engineering

Website: www.ijetae.com (ISSN 2250-2459 (Online), An ISO 9001:2008 Certified Journal, Volume 3, Special Issue 1, January 2013)

International Conference on Information Systems and Computing (ICISC-2013), INDIA.

Table 1
Comparative Study

DEPTH CUES	IMAGE CONTENT	ACCURACY	DENSE OR SPARSE DEPTH MAP
ATMOSPHERE SCATTERING	Scene include sky	Relative error rate: 10% of outdoor images	Dense
SHADING	The image must not be that much dark. The depth of Shadow area cannot be recovered	The mean error rate is 4.6%. Maximum error 40%	Dense on the surface
PATTERN TEXTURE	Requires segmented image.	The average error , ie,the angle between estimated and actual surface is 8 degree.	Sparse or dense
STATISTICAL PATTERN	All images.	Average error: 0.132 in log scale with base 10.	Dense

IV. PROPOSED WORK

This work describes a novel 2D-to-3D conversion method based on the use of the edge information.

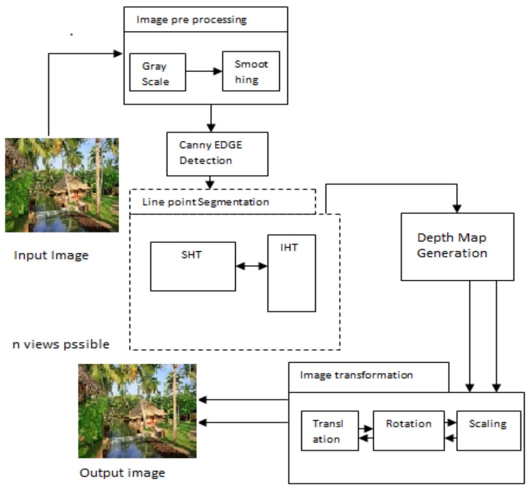

Fig 1: Proposed system

5.1 Gray Scale Conversion

The Gray scale images are distinct from black and white images and also called as binary images. Gray scale images have shades of gray in between.

Colors in an image may be converted to a shade of gray by calculating the effective brightness or luminance of the color image and uses this value to create a shade of gray that matches the desired brightness. The effective gray scale value of a pixel is calculated with the following formula:

Y= (Red + Green + Blue) / 3 (1)

Based on the equation (1) effective value of an arbitrary pixel is calculated.

5.2 Gaussian Smoothing

The image after gray scale conversion may contain noise. To get a quality image the presence of noise is not desirable. So Gaussian smoothing is applied to the image. In this the raw image is convolved with a Gaussian value. The Gaussian function is:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (2)$$

Using the Gaussian function a matrix of Gaussian values will be generated as per the above equation (2) and the image will be convolved with the Gaussian matrix. This method will remove the noise and the image will get a blurred effect.

5.3 Canny Edge Detection

The algorithm runs in 4 separate steps:

1. **Finding gradients**: The edges are marked where the gradients of the image have large magnitudes.
2. **Finding the edge direction.**
3. **Non-maximum suppression**: edge pixel values less than a fixed threshold is discarded.
4. **Hysteresis:** Eliminate breaking up of edges**.**

The Canny edge algorithm finds edges where the grayscale intensity of the image changes. The edges are found by determining gradients of the image. Gradients of each pixel are determined by applying the Sobel operator in both x and y direction. Then the direction of the image is found using matrix created using a Sobel operator. Non maximum suppression is done by preserving the edges having gradient greater than the fixed threshold value, and deleting everything else. In hysteresis two thresholds are taken the edge pixels stronger than the high thresholds are marked as strong and edge pixels weaker than the lower threshold are suppressed. The edge pixels between the two thresholds and connected to earlier found edges are marked as strong. The strong edges are taken to create the edge image.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459 (Online), An ISO 9001:2008 Certified Journal, Volume 3, Special Issue 1, January 2013)
International Conference on Information Systems and Computing (ICISC-2013), INDIA.

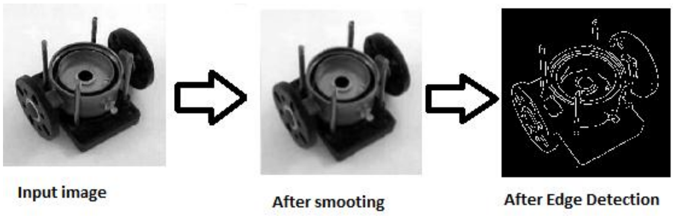

Fig 2: An image after smoothing and edge detection.

5.4 Line And Point Segmentation

The major operation of line and point segmentation is done next. Edges are characterized by object boundaries and are used for segmentation and identification of objects in an image. The Hough transform is used to detect separate straight lines of the image and thereby identify the true geometric structure of the object. Standard Hough transform (SHT) is used to trace objects like walls, doors etc. Improved Hough transform (IHT) is used to find point values and corners using the point segmentation.

If we use these edge or boundary points as input to the Hough transform, a curve is generated in polar space (r, θ) for each edge point in the Cartesian space. Curves generated by collinear points in the image intersect in peaks (r, θ) in the Hough transform space. These intersection points characterize the straight lines of the image. Mapping back from Hough transform space into Cartesian space yields a set of line descriptions of the image.

5.5 Depth Map Generation

For the depth map generation vanishing line detection and gradient plane assignment is used. The main steps in this are:

1. *Vanishing line detection:* Initially some image features, like Vanishing Point (VP) and vanishing lines are detected.
2. *Gradient plane generation*: During this processing step, the position of vanishing point (relative to the original image) and the slopes of vanishing lines is analyzed.
3. *Depth gradient assignment***:** A gray level corresponding to the depth level is assigned to every pixel corresponding to depth gradient planes.
4. *Depth map generation by fusion:* In this step qualitative depth map and geometric depth map are combined to generate the final depth map.

5.6 Image Transformation

Image transformation defining three types of pattern transformation.

1. Translation moves the generating function across the image.
2. Rotation rotates the image generating function by the angle.
3. Anisotropic scaling scales the generating function anisotropically.

VI. Results

For the implementation of the system Java is used. The simulation results of Gaussian smoothing and canny edge detection is shown below.

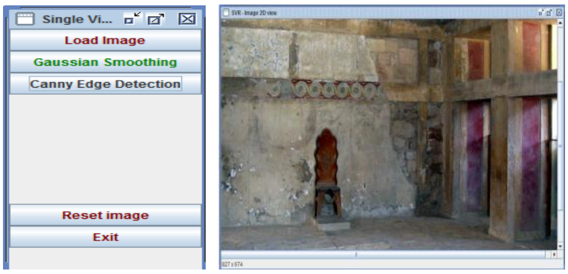

Fig 3: (a) Front end, (b) Loaded image

Java swing is used to develop the display area.

Fig 4: (a) Gaussian smoothing, (b) Canny edge detection

Using this edge image the lines, points, vanishing points etc. have to be calculated and transformations have to be done

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459 (Online), An ISO 9001:2008 Certified Journal, Volume 3, Special Issue 1, January 2013)
International Conference on Information Systems and Computing (ICISC-2013), INDIA.

VII. Conclusion

My project plan consists of building a novel 2D to 3D conversion system with highest possible accuracy. The geometric features of the image will be extracted using edge detection and line and point segmentation. The proposed system is expected to give a clear 3D vision of a 2D image.

REFERENCES

[1] Chao-Chung Cheng, Chung-Te Li and Liang-Gee Chen, “A 2D-To-3D Conversion System Using Edge Information” in Proceedings of IEEE International Conference on Consumer Electronics, 2010.

[2] Thuy Tuong Nguyen, Xuan Dai Pham and Jae Jeon, “Rectangular Object Tracking Based on Standard Hough Transform,” IEEE Electron Device Lett., vol. 20, pp. 569–571, Nov. 2009.

[3] W. Tam and L. Zhang, “3D-TV Content Generation: 2D-To-3D Conversion,” Proceedings of IEEE ICME, (2006).

[4] J. Ko, M. Kim and C. Kim, “2D-To-3D Stereoscopic Conversion: Depth-Map Estimation in a 2D Single-View Image,” Proceedings of SPIE, Vol. 6696, (2007)

[5] S. Battiato, S. Curti, E. Scordato, M. Tortora, and M. La Cascia, “Depth map generation by image classification,” Three-Dimensional Image Capture and Applications VI, vol. 5302, pp. 95-104, 2004.

[6] D. Hoiem, A. A. Efros, and M. Hebert, “Automatic photo pop-up,” in Proc. ACM SIGGRAPH, 2005.

[7] Q. Wei, “Converting 2D to 3D: A Survey”, Information and Communication Theory Group, Delft University of Technology, the Netherlands, (2005).

[8] J. F. Canny, “A Computational Approach To Edge Detection,” IEEE Trans. Pattern Analysis and Machine Intelligence, 8, 679-714, (1986) .

[9] Loh, A.M.; Hartley, R. (2005) “Shape from Non-Homogeneous,Non-Stationary,Anisotropic, perspective texture”, Proceedings, the British Machine Vision Conference 2005