CONVERSION OF 2D TO 3D IMAGE AND VIDEO

1MISS. MUJAWAR A. R., 2MR. NANAWARE J.D

Electronics Department, KBPCEPS, Satara
Shivaji University

ABSTRACT

There is a significant growth in the various technologies in the last few years. Innovation of 3D is the technology which brought thinking of human being to the next level. 3D has tremendous advantages which are specially seen in movies. Despite of this, the availability of 3D content is very less as compared to 2D. Methods involving human operators have been most successful but these are time consuming and costly. The methods which do not involve human operator increase the speed of the conversions. The method which do not require human operator is called automatic methods and the method which require human operator is called as semi automatic method. Both of these methods use many algorithms to convert 2D data into 3D. Some algorithm are based on local point mapping which uses low level attributes such as color, location, motion whereas some algorithms are based on global point mapping which uses repository of 3D images available online.

Keywords:- 3D images, stereoscopic images, image conversion, nearest neighbor classification.

1. INTRODUCTION

Two dimensional is a concept that describes anything that composes of length and width. In two dimensional everything in the image is presented at the same distance from the viewer. But users continuously demand richer, more immersive and closer to reality viewing experiences. After the introduction of color displays and high definition images, 3D video promises to be the next revolution in visual technology. The recent momentum in the production of 3D content for cinema applications is a good example that the revolution has started. Nevertheless, although many stereoscopic 3D movies have been produced recently, there is a lack of 3D video content, especially for the 3DTV industry. Moreover, 3D reproduction of conventional well-know 2D movies or TV programs is appealing for both users and content producers [1], [2]. Similarly, there are large no of 2D video available that exist in different compressed format. Though 3D content provide more realistic sense to viewer, the 3D broadcasting is very less as compared to 2D broadcasting. However, 3D can be regarded as the next revolution for many applications such as television, movies, and video games. The smash hit movie “Avatar” has demonstrated great success in the use of 3D and announced the approach of the 3D era [3]. Therefore, there is an urgent need for efficient and robust 2D-to-3D conversion algorithms.

The requirement of 3D over 2D:

• Faster product design (roughly 45% faster on average)
• Beat your competition to market
• Automatic flattening of sheet metal parts (with bend allowance)
• More effective communication with suppliers /customers.
• Visualize more ‘what-if’ scenarios during the design process.
• The ability to create renderings and animations for design proposals or reviews.
• More effective internal design reviews.
• Generation of virtual prototypes allows non-CAD people to participate in the process.
• Allow non-technical personnel such as sales department (and even customers) to quote, specify and configure product whilst maintaining your design & engineering integrity [4].

Generally, current 2D-to-3D conversion algorithms can be divided into two methods:

1) Semi-automatic methods with human-computer interactive operations.

2) Automatic methods which directly output 3D video from 2D input without any user interactions involved. Semi-automatic 2D-to-3D conversion unsurprisingly has better performance than fully-automatic methods because of the high-level knowledge provided by users [3], [4]. In a semi automatic method a skilled operator assigns depth to various parts of an image or video. Based on this sparse depth assignment, a computer algorithm estimates dense depth over the entire image or video sequence [1]. In automatic method, computer algorithm itself computes the depth of video or image and makes the conversion of given 2D input query to 3D output. Semi automatic method involves human operator whereas automatic method do not involve any human operator. Both of these methods use some algorithms to convert 2D content to 3D. Some algorithm are based on local point mapping which uses low level...
attributes such as color, location, motion whereas some algorithms are based on global point mapping which uses repository of 3D images available online. The paper is organized as follows. In Section II, The methods of 2D-to-3D video and image conversion are discussed. In Section III, the algorithms used in semiautomatic method are described. In Section IV, the algorithms used in automatic method are described. Issues and Challenges are described in section V. And paper is concluded in section VI.

2. METHODS OF 2D TO 3D VIDEO AND IMAGE CONVERSION

There are two types of conversion methods to convert 2D query input image/video to the 3D image for the more realistic view. They are:
1. Semi automatic method.

2.1 Semi automatic method

In this method human operator involvement is essential for the conversion process. The operator in this method does the work of delineating objects in the query image then placing them at suitable depth and finally correcting the errors (if any) after the final rendering. This method of conversion has been successfully used commercially by such companies as IMAX Corp., Digital Domain Productions Inc. (formerly In-Three Inc.), etc. Many films have been converted to 3D using this approach [1]. In this method due to the involvement of human operator, the speed of the conversion is less and thus the cost is more. Also, in this method human operator involvement is mandatory.

2.2 Automatic method

In this method no human operator is needed for the conversion of 2D to 3D. Thus in this method the speed of the conversion is more as compared to semi automatic method. Also the cost is less. The main step in the 2D to 3D conversion is to calculate the depth from a single 2D image. To perform this task, many methods are developed. Some methods use the camera to capture the photos of the same image from different angles, at different conditions, but not at the same time. On the other hand, there are some methods which uses readily available repository [1] of images to calculate the depth. And then the best match image is taken for the depth recovery.

3. ALGORITHMS USED BY SEMI AUTOMATIC METHOD

Algorithms for semi automatic methods are as follows-

3.1 2D To 3D Conversion Using Disparity Propagation.

In the disparity propagation algorithm there is 2D query input is available. Of that query image, some /many photos using camera are taken. All the photos are taken from different angles. Then first, a set of corresponding points in the image pair are found. Then, by means of the triangulation method, the depth information can be retrieved with a high degree of accuracy when all the parameters of the stereo system are known. Some steps which are to be done in the disparity propagation are: Multiple objects are first segmented by the input user-scribbles. Then, the initial disparity map is assigned to each key frame with the aid of various preset disparity models for each object. After the disparity assignment step, disparity maps for other frames of the video are obtained through a disparity propagation strategy taking into account both color similarity and motion information. Finally, the 3D video is synthesized according to the type of 3D display device [3]. The algorithm is demonstrated on seven sequences. Most of the sequences are taken from PhillipsWowVx web site [5]. Algorithm shows better result on six sequences than other methods in terms of average mean square error. Except the third sequence dice. In this sequence, because of sharp edges and color ambiguity, the algorithm result is very close to expected value but not the same.

3.2 Stereo Extraction from Video Footage

The basic units of broadcast video are frames, shots and scenes. When the input shot is given to this algorithm, the following steps are performed. User scribbles are marked on some of the frames to indicate desired disparity values. The marked disparities are propagated on the frames on which they were drawn. A classifier is trained for every disparity value marked by the user. The classifier is applied to the entire shot, and high confidence predictions are recorded. The disparity map of the entire shot is recovered in an optimization process which is constrained by the original scribbles and the high confidence predictions [6]. This method works well with videos with long shots. This method has limitation on short shots because the perception of depth in the middle of shot is lost.

3.3 Video Stereolization

This algorithm converts conventional videos into stereoscopic videos. The input image sequence is first passed through three individual automatic modules: structure-from-motion is applied to the input image sequence with dominant rigidly moving objects to recover a sparse set of 3D points. The MOS module is used to automatically segment the foreground, it is particularly effective in a follow shot in which the foreground is relatively static and the background is rapidly changing. Finally, the PDC module inspects the size change of an object’s image to estimate relative depth
changes between frames. If there are still undefined regions, the users need to label them in some key frames by simple scribbling [7]. With these algorithm users labeling task is simplified. Also this algorithm is user friendly. But this algorithm is only applied to stereoscopic videos is the major disadvantage.

4. ALGORITHMS USED BY AUTOMATIC METHOD

Algorithms for automatic method are as follows-

4.1 2D To 3D Conversion by Learning A Local Point Transformation

In this technique low level image or video attributes at each pixel are to be considered. So, in this method depth is assigned to a pixel based on its attributes [1]. The low level image/video attributes to be considered are: Color, Location and Motion. A pivotal element in this approach is a point transformation used to compute depth from image attributes. Then by referring the original image, this algorithm considers each pixel in the other photos that by how much location it is moved from one place to another then how much color get changed and finally what about its motion at that pixel. From the above discussion it is clear that for the 2D-to-3D conversion by learning a local point transformation which uses the low level attributes such as color, location, motion, the function of color, location, motion for the depth estimation and thus for the conversion. This method of conversion, considers the location of each pixel also. The location transformation is nothing but simply the average depth at the same location computed from the all depth map in the repository I. And thus by using this color, location and motion transformation techniques we can compute depth of the given 2D query image and get resultant 3D output. This algorithm has the main advantage of the computational efficiency. And the point transformation can be learned off-line and applied basically in real time. But this method uses purely local, low level image or video attributes is its major disadvantage. To overcome this disadvantage, another algorithm is developed which uses online available repository of 3D images.

4.2 2D To 3D Conversion Based On Global Nearest-Neighbor Depth Learning

This algorithm uses the online available 3D image repository using a nearest-neighbor regression type idea. This algorithm uses one assumption and a key observation. The key observation is that among millions of 3D images available on-line, there likely exist many whose 3D content matches that of a 2D input (query) which wants to convert to 3D. And an assumption that two images that are photo metrically similar also have similar 3D structure (depth) [1].

For the conversion using this method, there are some steps and they are:

I. Search for representative depth fields from online available 3D images repository like on Make3D#[1][4] and finding out the best match for query image.
II. Depth fusion in which after getting ‘k’ number of images, and then by using the median filtering [1] this algorithm is combining the k representative depth field.
III. Depth smoothing process in which the variations are removed but depth discontinuities are preserved.
IV. Stereo rendering provides the right image of fictitious stereo-pair [1] and also the smooth depth field is also obtained. The main advantage of this algorithm is that it uses online available repository of 3D images and this global method with CBF takes about 1 second to process.

4.3 2D To 3D Conversion System Using Edge Information

This work describes a novel 2D-to-3D conversion method based on use of the edge information. Importantly, the edge of an image has a high probability of being the edge of the depth Map. After the pixels are grouped together, a relative depth value can be assigned to each region. The depth of each segment is then assigned by using an initial depth hypothesis. Next, the blocky artifact is removed using cross bilateral filtering. Finally, multi-view images are rendered by depth image-based rendering (DIBR) and display on a 3D display [8]. The proposed method uses only a single image with only slight side effects. If the images taken from low quality camera, due to unsharp edges depth estimation can go wrong.

5. ISSUES AND CHALLENGES

The main challenges introduced in the 2D to 3D conversion using disparity propagation are sharp edges, color ambiguity, large displacement and zoom in-zoom out. After performing experiments on various seven sequences this algorithm provides good result on challenging factor such as zoom in-zoom out but provide poor results on color ambiguity and sharp edges but results are near the expected values. The algorithm namely Semi-automatic Stereo Extracation from Video Footage uses variety of broadcast videos. This algorithm works well on sport videos with long shots, but has limitation where the scene changes within a shot. The depth perception in the middle of the shot is lost. Video Stereolization algorithm has limitation that it is applied to only stereoscopic videos and lack of parallax is also found. This algorithm needs to further develop to overcome these problems. Algorithm which converts 2D to 3D using local point transformation has disadvantage that it uses only low level image and video attributes. And algorithm which uses global nearest-neighbor depth learning has limitation that it needs to perform ‘k’ number of repetitions to get the
exact match image which reduces its efficiency. 2D to 3D conversion using sharp edges can work poor if the camera quality is poor due to which efficiency is poor is its limitation.

6. CONCLUSION

The amount of 3D content has greatly increased in recent years and there is more demand for 3D video in many scenarios. Recovering 3D information from conventional 2D video is an inherently ill-posed problem. All these reasons make 2D-to-3D conversion not only valuable in practical applications, but also meaningful for research [2]. A single solution to convert the entire class of 2D images to 3D models does not exist. Combining depth cues enhances the accuracy of the results. It has been observed that machine learning is a new and promising research direction in 2D to 3D conversion. And it is also helpful to explore the alternatives than to confine ourselves only in the conventional methods based on depth maps [3]. 2D to 3D conversion method which involves human operator in many algorithms are time consuming but having high efficiency whereas some algorithm which is not involving operator are less time consuming but having less efficiency. Many algorithms have been developed in the last few years to convert 2D to 3D image and video which are less complex. For having the more realistic visualization of any image/video 3D is must and thus 2D to 3D conversion technique with simplified algorithm with simple technique of conversion is essential. The simpler conversion techniques with simplified algorithms are also in under research process.

REFERENCES

[5] Dr. Ir. E. A. Hendricks, Dr. Ir. P. A. Redert “Research Assignment Converting 2D to 3D: A Survey” Information and Communication Theory Group (ICT), December 2005

AUTHOR

Anjum Mujawar completed diploma in Electronics and Telecommunication and received B.E. in Electronics degree in 2008-2009 and 2011-2012, respectively from Karmaveer Bhaurao Patil College of Engineering And Polytechnique, Satara and presently doing M.E. second year in Electronics from the same college.