IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Volume-3, Issue-9 E-ISSN: 2347-2693

A Survey on: Conversion of 2 Dimensional Images to 3 Dimensional Model

Pranoti Vaidya[1*], Jyoti Deore[2], Rakhi Thakare[3] and Pooja Pangavhane[4]

[1,2,3,4] Department, of Computer Engineering,*
Late G.N.Sapkal College of Engineering, Savitribai Phule Pune University, India

www.ijcseonline.org

Received: Sep /05/2015 Revised: Sep/12/2015 Accepted: Sep/28/2015 Published: sep/30/ 2014

Abstract— In this paper we proposed that a system to present an android application for 3Dimentional model construction from 2Dimentional images. The system enables user to have more informative, interactive and user specific experiences with augmented information by recognizing/tracking the content of offline 2D images. For android device which required Low computational power and an image matching technique based on the combination of two binary feature descriptors are applied Our application can be applicable to many areas such as education and entertainment industry.

Keywords— Augmented reality,BRISK,FREAK,Image processing,Image Matching and detection

I. INTRODUCTION

Three Dimensional views are most significant view for detailing. To get more information and detail view of any object 3D model is more frequently used. This paper proposed an application to construct 3D model from related 2D image i.e. using augmented reality. With the development of 3D applications, the conversion of existing 2D images to three Dimensional images which becomes an important component of Three Dimensional content production. The process of conversion of existing Two Dimensional images into Three Dimensional is commercially feasible and is fulfilling the growth of high quality stereoscopic images. The leading technique for such content conversion is to develop a depth map for each and every frame of Two Dimensional material. After observing the world, the human brain usually combines that the heuristic depth cues for that the generation of the depth perception. However For android device it's challenging to augment number of target images hence to overcome this we already stored Data on cloud. Two state-of-the-art of binary descriptors, binary robust invariant Scalable key points (BRISK)[3] and fast retina key points freak Are applied for recognizing/tracking target images with low Computational power and high level of accuracy.

Augmented reality (AR) [1,2] is a term for the live direct or indirect view of a physical, real-world environment whose elements is multiply or augmented by computer-generated sensory input. The applications and related technologies for AR are attracting increasing attention from both the scientific community and companies originally involved in different research areas. In particular, the progress achieved in the fields of computer vision and mobile computing are mainly shifting the focus towards the development of systems for AR for mobile devices. AR is thus creating newer and newer opportunities for exploring mechanisms of interaction between humans, and virtual and physical environments. However, even if in strong expansion, the current state of the art of AR technologies and applications is still below market expectations, especially when considering the quality of the interaction offered. While some aspects that are closely linked to the AR technology (i.e. marker-tracking, rendering, etc.), are gradually evolving, on the other end, there are still several aspects, technical and social, requiring further investigation. One of these aspects is the creation and analysis of appropriate interaction techniques for AR applications, which allow the user to interact with virtual content in an intuitive manner. It is possible to explore the development of new interaction techniques in different directions including: ubiquitous computing, tangible computing, and social computing. In ubiquitous computing [5], we can analyze the interactions with the user and its activities, within a dynamic environment. In tangible computing the user interacts with interfaces, modelled as physical objects and associated with digital information.

II. LITRATURE SURVEY

Among two Dimensional-to-three Dimensional image conversion methods, those involving human operators have been most successful but also time-consuming and not cheap. In this modern era popularity of three Dimensional hardware is increased but, three Dimensional contents are still dominated by its two Dimensional counterpart. Until now days many of researchers have proposed different methods to close this gap. Mainly, these conversion methods are categorized in an automatic method and semi-automatic method. In automatic method human intervention is not involved, where as in semi-automatic method human operator is involved. Automatic methods, that typically make use of a deterministic three Dimensional scene model, have not yet achieved the same level of accuracy as they often depend on assumptions that are easily violated in practice. Adopt the radically several approach of "learning" the three Dimensional scene structure. Simplified and computationally-efficient version of recent two Dimensional-to-three Dimensional image conversion algorithm was developed. Given a repository of three Dimensional images, found K pairs photometric content most closely matches that of a two Dimensional query to be converted. Then, fused the k corresponding depth fields and align the fused depth with the two Dimensional query. While far from the perfect, obtain the result which are presented and then results demonstrate that online repositories of three Dimensional contents can be used for effective two Dimensional-to- three Dimensional image conversions. There are varied attributes that can be considered during the different conversion, like for video conversion motion and optical flow are mostly considered parameter, while for image conversion local attributes of images are considered. Computational time and design cost are the main design metrics which are considered while designing algorithm. Each method is having its own pros and cons, depending on specification which method will be suitable for conversion is decided. Three Dimensional signal processing has become a vital trend in the related visual processing field. But at this stage, the lack of three-Dimensional content is now becoming more useful and important factor that limits its development. It will be great help to solve this problem if extracted the depth information from the monocular image and automatically convert it to three Dimensional video. On the basis of some research findings in recent years, introduced the main two kinds of methods to do the conversion: off-line conversion and real-time conversion. Taking off-line conversion for example, systematically presents all aspects of the conversion process, including the various methods that each aspect can use, and then we describe the principles and basic methods which are use for real-time conversion, and finally after that discuss the challenge and development that two Dimensional image to three Dimensional model conversion technologies will be in the future.

III PROPOSED SYSTEM

The overall framework of is shown in Fig.1.After capturing an image from the by a mobile camera, the system first identify the query image by equivalent and matching with reference images in the given database. This recognition result is then sent to the context management system. The context management system links to the corresponding site and contents information of each database. When users capture live video sequences, tracking procedure is performed. At that time, Three Dimensional virtual models recovered from the database are rendered on the images in the off-line booklet or natural images based on the estimated camera pose. The key step in 2D to 3D conversion process is the generation of a dense depth map [7]. In these years, several type of depth map generation algorithms have been proposed according to the principle of human visual system. Each of the algorithms has its own strengths and weakness. Most depth extraction algorithms usually use certain depth cue but some of them combine two or more depth cues to generate depth map. 2D-to-3D depth generation algorithms typically face two challenges. First is the depth uniformity inside the same object. A more flexible and suitable grouping of pixels results in a better outcome for the depth uniformity inside the object. The other challenge includes retrieving an appropriate depth relationship between all the objects. These methods result in the false depth information when the object is with contrasting self-motion vectors. The pixels can belong to the same object that may be assigned with changed and different types of depth values. In this manner the depth map generation from single 2D images is an ill-posed problem.

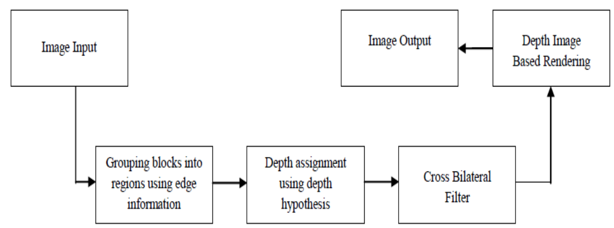


Figure 1: 2D-to-3D Conversion System

IV SYSTYEM ARCHITECTURE

(A)Recognition: symmetric and combinatorial matching

For object recognition and finding visual correspondences between images, many types of feature descriptors were proposed to represent key points robustly in scale, noise level and rotation changes. (SURF) Speeded-Up Robust Feature and (SIFT) Scale Invariant Feature Transform are good examples. Though, for a stand-alone mobile system, all these type of algorithms are not very suitable due to their high computational cost. Therefore, in the system, two state-of-the-art binary feature descriptors are applied to achieve comparable matching performance at less computation time. The numerous same types of key points can be extracted and it often leads to many false matches. To overcome these challenges, applied two strategies: symmetric and combinatorial matching. The step by step overall procedure for the image recognition is provide and summarized as follows:

1. Calculate the key points in both a query (*Iq*) and a reference image (*Ir*) in database by multi-scale AGAST [3].

2. Nearest neighbor matching should be perform, and then remain matches if the ratio of the both shortest and the second-shortest descriptor distance is lesser than a predefined threshold □ (□=0.9). This step is denoted by *M* (*Ir* | *Iq*) .

3. Compute the matches from *I r* to *I q* . Then calculate the symmetric matches; *Ms* (*Ir* | *Iq*) □□*M* (*Ir* | *Iq*) ∩□*M* (*Iq* | *Ir*) .

4. Follow step 2 and 3 in both BRISK and FREAK [4] descriptor matching. Candidate matches are obtained for each of two cases, and then common matches between them are always used to finalize the matches between the images.

5. Consider the total number of matched key points as a match score and recognize the image by retrieving the related image with higher match score.

(B) Tracking: tracking-by-detection

For live image tracking, select a tracking-by- detection approach, which is suitable for fast motion and occlusion of the object in uncontrollable mobile environments. The basic idea of this approach is that feature points are derived from incoming frames at matching and run-time is performed against a feature point's database for which the three-Dimensional locations are known [5]. Then, three –Dimensional pose can be estimated by finding the correspondences. At first find the correspondences discussed in previous section. After this, some of the outliers are eliminated by RANSAC algorithm [6].

ACKNOWLEDGMENT

It gives us great pleasure to acknowledge our topic titled: Two Dimensional Image to Three Dimensional model conversion. This topic is presented as we are interested in the field of artificial intelligence. We would like to thanks all people who have lit our path with their kind guidance.

REFERENCES

[1] V. Vlahakis, N. Loannidis, J. Karigiannis, M. Tsotros, and M. Gounaris, "Archeoguide: An Augmented Reality Guide for Archaeological Sites," *IEEE Computer Graphics and Applications,* vol. 22, pp.52-60, **2002**.

[2] iTacitus, http://www.itacitus.org (accessed on June, 29, **2013**).

[3] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," *IEEE International Conference on Computer Vision (ICCV)*, pp. 2548-2555, **2011**.

[4] A. Alahi, R. Ortiz, and P. Vandergheynst, "FREAK: Fast retina keypoints," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 510-517, **2012**.

[5]V. Lepetit, "On Computer Vision for Augmented Reality," *International Symposium on Ubiquitous Virtual Reality (ISUVR)*, pp. 13-16, **2008**.

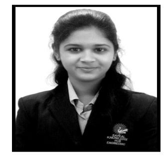
[6] M. A. Fischler, R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Communications of ACM (CACM)*, vol. 24, no. 6, pp.381-395, **1981**.

[7] Y-L. Chang, et al, "Depth Map Generation For 2D-To-3D Conversion By Short-Term Motion Assisted Color Segmentation" in Proceedings of ICME, **2007**

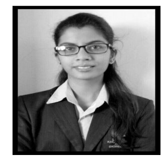
[10] W. J. Tam, and L. Zhang, "3D-TV content generation: 2D-to-3D conversion," in Proc. ICME, pp. 1869-1872, **2010**

AUTHORS PROFILE

Name: Pranoti R. Vaidya
Department of Computer Engineering
Lt. G.N.Sapkal College of Engineering, Nashik
Savitribai Phule Pune University, India



Name: Jyoti K. Deore
Department of Computer Engineering
Lt. G.N.Sapkal College of Engineering, Nashik
Savitribai Phule Pune University, India



Name: Rakhi S. Thakare
Department of Computer Engineering
Lt. G.N.Sapkal College of Engineering, Nashik
Savitribai Phule Pune University, India

Name: Pooja pangavhane
Department of Computer Engineering
Lt. G.N.Sapkal College of Engineering, Nashik
Savitribai Phule Pune University, India