A Survey on Automated 2D Image to 3D Model Generation

Prof. A.R.Kamble, Yash Devgirikar, Rohan Choudhar, Harshwardhan Barmade, Aditya Khetre

ABSTRACT

With the rise of digital world, Traditional learning methods, while time-tested, often fall short in meeting the dynamic educational needs of today's learners. They are constrained by several drawbacks that impede effective comprehension and engagement. This abstract sheds light on the limitations of traditional learning approaches and introduces how the integration of 3D model reconstruction can mitigate these challenges, thereby enhancing the learning experience. We proposed deep Conventional techniques, which create a virtual view from a reference picture, use two distinct steps: depth image-based rendering (DIBR) with estimated depth and depth (or disparity) estimation for the reference image and based on these steps we can create 3D model. Which allows educators to create immersive and interactive learning environments that engage students on a profound level. By converting abstract concepts into tangible and visually stimulating representations, 3D models bridge the gap between theory and practice, making complex subjects more accessible.

Keywords: 2D Image, 3D Model, CNN, DIBR.

Introduction:

In the realm of education, the age-old paradigm of traditional learning, characterized by lectures, textbooks, and two-dimensional visual aids, has long been the standard approach. However, the rapid advancement of technology has ushered in a new era of learning, where immersive and interactive experiences take center stage. This paper delves into the compelling just a position between traditional learning methods and the emerging trend of 3D model-based learning. It seeks to explore the key differences, advantages, and potential drawbacks of these two contrasting approaches, as well as their impact on the educational landscape.

In stark contrast, 3D model-based learning harnesses the power of cutting-edge technology to transform education into a dynamic and interactive endeavor. By creating three-dimensional representations of concepts, objects, and environments, this approach engages learners on a deeper level, making it easier to comprehend complex subjects and fostering experiential learning. This paper will analyze the potential benefits of 3D model based learning, including its ability to accommodate diverse learning styles, stimulate critical thinking, and enhance knowledge retention.

Literature Survey:

The paper[1] explore applications in diverse fields, such as precise measurement registration, creation of 3D human models, image-to-shape translation, 2D-to-3D image conversion, and detailed 3D modeling for indoor spaces. These applications span areas like medical imaging, gaming, navigation, and object reconstruction. The papers often discuss the evaluation of their proposed methods, comparing them with existing techniques or traditional approaches. This involves showcasing the potential and effectiveness of the proposed methods through experiments and comparisons with 3D scanning models or other benchmarks. The research papers in the field of automatic 3D reconstruction explore a wide range of applications, including precise measurement registration, creation of 3D human models, image-to-shape translation, 2D-to-3D image conversion, and detailed 3D modeling for indoor spaces. These

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

applications find relevance in diverse fields such as medical imaging, gaming, navigation, and object reconstruction.

A common thread in these paper[3] and [5] is the emphasis on evaluating and comparing proposed methods with existing techniques or traditional approaches. This involves conducting experiments and benchmarking against 3D scanning models or other established standards to showcase the potential and effectiveness of the new methods. These papers highlight the challenges associated with aligning 2D images with 3D models, especially when dealing with different modalities. The complexity arises due to the differences in geometric or visual features between the two types of data. Various approaches and methods are discussed for converting 2D images into 3D representations. These include self-supervised frameworks (SIST), systematic methodologies involving human silhouettes and feature points, genetic algorithms (GAs), and convolutional neural networks (CNNs) and some papers emphasize the use of deep learning techniques for 3D reconstruction from 2D images. This includes the application of deep learning models for decoding ShapeNet rendering images and the use of multi-scale deep CNNs for automatic 2D-to-3D conversion. Some papers address the practical aspects of their proposed methods, considering factors such as time efficiency, cost-effectiveness, and reduction of expenses compared to alternative technologies. This includes discussions on the use of basic cameras and LiDAR systems for achieving accurate results. So we had studied these all paper and compare the capabilities and limitations of the old technique with the new features and technology and identify the gaps in terms of functionality, accuracy, efficiency, and quality. Hence, we are using CNN and DIBR algorithm for automatic 3d reconstruction.

Proposed Methodology:

The system architecture for a ”VisioLearn: Mastering Visual Concepts with DepthCraft” project involves acquiring 2D data, preprocessing it, extracting features and depth information, generating a 3D model, refining the model, visualizing it, providing a user interface, managing data, and potentially integrating machine learning and optimizing performance. The architecture's specifics depend on project goals and requirements, with options for monolithic or distributed systems, cloud-based processing, and integration with other components.

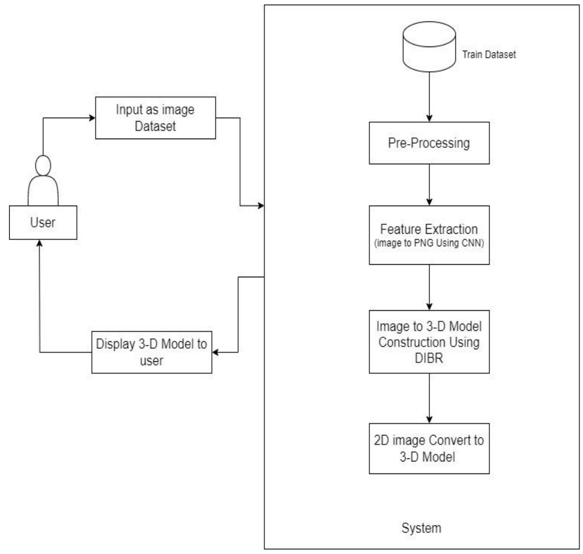

Figure 1: System Architecture

The system consists of several modules: User Authentication ensures secure access with valid credentials. View and Authorize Users empowers administrators to manage and authorize user accounts. The Image Scanning module enables users to select categories and scan images, leveraging a camera or uploaded images. The View Result module presents the 3D model generated from the scanned image, accompanied by relevant information. Post-Result Actions offer users options to navigate, such as returning to the home page or initiating another image scan for additional information. Throughout, error handling, a user-friendly interface, security measures, logging, and scalability considerations are integrated for a robust and efficient user experience.

Observation :

The research papers in the field of automatic 3D reconstruction collectively explore a wide range of applications that have implications across diverse industries. These applications include precise measurement registration, the creation of detailed 3D human models, image-to-shape translation, 2D-to-3D image conversion, and the development of intricate 3D models for indoor spaces. The significance of these applications extends to fields such as medical imaging, gaming, navigation, and object reconstruction.

A significant aspect of these papers is the meticulous evaluation and comparison of the proposed methods with existing techniques or traditional approaches. Rigorous experiments and benchmarking against established standards, such as 3D scanning models, are conducted to demonstrate the efficacy and potential superiority of the newly proposed methods.

One of the prominent challenges addressed in these papers revolves around aligning 2D images with 3D models, particularly when dealing with different modalities. This challenge arises due to variations in geometric or visual features between the two types of data. To tackle this, the papers discuss various approaches, including self-supervised frameworks, systematic methodologies involving human silhouettes and feature points, genetic algorithms, and the application of convolutional neural networks (CNNs). The use of CNNs, with their ability to learn hierarchical features, is especially highlighted for its effectiveness in handling complex mappings and relationships between 2D and 3D data.

Practical considerations are also emphasized in these papers. Factors such as time efficiency, cost-effectiveness, and reduction of expenses compared to alternative technologies are taken into account. Some papers stress the use of basic cameras and LiDAR systems, making the implementation of these 3D reconstruction methods more accessible while still achieving accurate results.

The role of CNNs is pivotal throughout the 3D reconstruction process. These neural networks are employed for decoding ShapeNet rendering images, implementing multi-scale deep CNNs for automatic 2D-to-3D conversion, and creating detailed 3D human models. The deep learning capabilities of CNNs are leveraged to understand spatial hierarchies and intricate features in the input data, contributing to the accuracy and quality of the reconstructed 3D models.

Furthermore, the integration of Depth Image-Based Rendering (DIBR) algorithms is a critical component of these studies. DIBR algorithms are employed to synthesize novel views and manage depth information, significantly enhancing the overall accuracy and quality of the reconstructed 3D models.

In summary, these research papers collectively identify gaps and challenges in the realm of 3D reconstruction, presenting innovative solutions that leverage advanced technologies like CNNs and DIBR algorithms. The focus on practical considerations and the potential for wide-ranging applications underscores the transformative nature of these advancements in the field.

Conclusion:

This paper studied 8 papers. The highlights and observations are papered in chapter 2. The gap has been analyzed, based on which problem statement is designed along with its objectives. The detail plan of all the activities is mentioned in section 1.5.

The process of 2D image to 3D model construction involves transition from images to dynamic three-dimensional representations. Beginning with the acquisition and enhancement of 2D images, the methodology incorporates sophisticated techniques such as image recognition, depth estimation, and 3D reconstruction. Further refinements, validation processes, and user interaction elements contribute to the creation of realistic 3D models for better engagement and rich experience.

Future Scope:

In the ever-evolving landscape of education, the traditional methods of learning, characterized by lectures and two-dimensional visual aids, have been the longstanding norm. However, the rapid progress of technology has ushered in a new era, one where immersive and interactive experiences take precedence. This paper delves into the intriguing juxtaposition between conventional learning approaches and the emergent trend of 3D model-based learning. It aims to explore the fundamental differences, advantages, and potential drawbacks of these two contrasting methodologies, shedding light on their respective impacts on the educational landscape.

In stark contrast to traditional methods, 3D model-based learning harnesses cutting-edge technology to transform education into a dynamic and interactive endeavor. By creating three-dimensional representations of concepts and environments, this approach engages learners on a deeper level, making it easier to comprehend complex subjects and fostering experiential learning. This paper analyzes the potential benefits of 3D model-based learning, including its ability to accommodate diverse learning styles, stimulate critical thinking, and enhance knowledge retention. Looking forward, the future scope of 3D model-based learning holds significant promise for transformative educational experiences and improved outcomes, with possibilities ranging from extended subject coverage and customized learning paths to the integration of virtual and augmented reality technologies. The evolution of 3D model-based learning stands poised to shape a more engaging and personalized educational landscape.

References:

[1] P. Ndjiki-Nya, M. Koppel, D. Doshkov, H. Lakshman, P. Merkle, K. Muller, T. Wiegand, “Depth image-based rendering with advanced texture synthesis for 3-D video” in IEEE Transactions on Multimedia, vol. 13, no. 3, pp. 453-465, 2019.

[2] J. Xie, R. Girshick, and A. Farhadi, “Deep3d: Fully automatic 2D-to-3D video conversion with deep convolutional neural networks,” in Proc. Eur. Conf. Com- put. Vis., pp.842-857,2016.

[3] H. Yoo, J. Son, B. Ham, and K. Sohn, “Real-time rear obstacle detection using reliable disparity for driver assistance,” in Expert Systems with Application, vol. 56, no. 1, 2016.

[4] J. Kopf, M. F. Cohen, and R. Szeliski, “First-person hyperlapse videos, ”in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, 2019.

[5] A. Saxena, M. Sun, and A.Y. Ng, “Make3D: Learning 3D-scene structure from a single still image” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 5, 2009

[6] Zuzana Cernekov ́a, Ioannis Pitas ˇ , Senior Member, IEEE, and Christophoros Nikou, Member, IEEE “] Information Theory-Based Shot Cut/Fade Detection and Video Summarization”