IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁘ Impact Factor 8.066 ⁘ Peer-reviewed / Refereed journal ⁘ Vol. 10, Issue 8, August 2023

DOI: 10.17148/IARJSET.2023.10835

Deep Visual Odometry with Adaptive Memory

Sanjana.L[1], Siddegowda.C J[2]

PG Scholar (MCA), Dept. of MCA, Vidya Vikas Institute of Engineering and Technology, Mysore, Karnataka, India[1]

Assistant Professor, Dept. of MCA, Vidya Vikas Institute of Engineering and Technology, Mysore, Karnataka, India[2]

Abstract: Our deep visual Odometry (VO) approach prioritizes memory and improved postures while taking into account global information. Unlike existing learning-based approaches which treat VO as a simple tracking issue, we reconstruct camera postures from picture fragments, which leads to a high accumulation of errors. To correct past mistakes, accurate worldwide data is essential. However, it might be difficult for end-to-end systems to reliably store such data. Therefore, we have developed an adaptive memory module that preserves information gradually and adaptively from a local to a global level in a neural memory analogue. This approach has been further enhanced using a refining module that takes advantage of global information stored in memory. To pick features for each view based on the co-visibility in the feature domain, we use past outputs as a guide and apply a spatial-temporal attention. Our system is more advanced than simple tracking since it has a memory module and a refinement module. Our experiments on the KITTI and TUM-RGBD datasets demonstrate that our technique not only delivers competitive results compared to conventional approaches in normal settings but also significantly outperforms state-of-the-art methods. Moreover, our model performs very well in instances where traditional algorithms struggle, such as texture-less areas and sudden movements.

I. INTROUDCTION

Camera positions can be estimated from image sequences using techniques such as Visual Odometry (VO) and simultaneous localization and mapping (SLAM), both of which utilize the consistency between successive frames. VO has been a vital research topic for decades and is crucial in various applications such as automated driving, virtual/augmented reality, and robot navigation. Many excellent methods have been developed from a geometric perspective. In recent years, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have achieved tremendous success in various computer vision tasks, including optical flow estimation, depth recovery, and camera re-localization. As a result, several end-to-end models have been proposed for optical flow estimation.

Our unique deep visual Odometry (VO) technique takes into account global information by incorporating significant memory and working to improve posture. Unlike learning-based approaches that handle the VO challenge as a pure tracking issue by reconstructing camera positions from image fragments, our approach considers the accumulation of errors in the final output. To ensure accurate data covering the entire planet, it can be challenging to store such data reliably within end-to-end systems.

To manage long-term dependencies, we build an adaptive memory module that progressively and adaptively moves information from a local to a global level in a neural memory analogue. This allows us to handle dependencies that last for a considerable period of time. The upshot is further inflated by the inclusion of a refinement module that makes use of information from around the world stored in memory. We utilize previous outputs as a reference and focus our attention in a spatial-temporal manner to select features for each view that are predicated on co-visibility as a feature domain.

Our system includes a memory module and a refinement module, making it more than just a simple tracking tool. Experiments conducted on the KITTI and TUM-RGBD datasets have shown that our methods provide outcomes that are comparable to those obtained by traditional approaches in typical conditions but greatly outperform methods considered state-of-the-art. Our model performs well in cases where existing algorithms struggle in areas devoid of texture and in situations involving abrupt movement.

Problem Statement

Visual Odometry (VO) is a crucial component of many computer vision and robotics applications, enabling devices to estimate their relative motion by analyzing sequences of images. Traditional VO methods often rely on feature-based tracking, which can be prone to failures in challenging environments with significant lighting variations, occlusions, or textureless scenes.

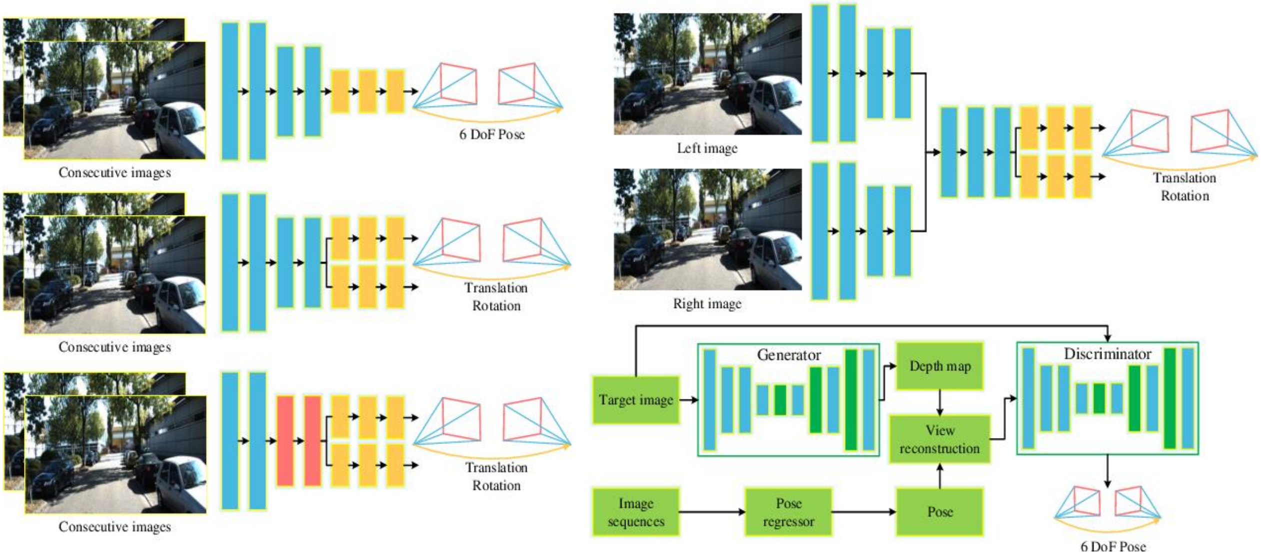
IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁘ Impact Factor 8.066 ⁘ Peer-reviewed / Refereed journal ⁘ Vol. 10, Issue 8, August 2023

DOI: 10.17148/IARJSET.2023.10835

The objective of this project is to address these limitations by developing a Deep Visual Odometry system with an Adaptive Memory mechanism. The core problem is to create a robust and accurate VO system that can navigate in dynamic and diverse real-world scenarios, particularly for applications like autonomous vehicles, drones, and robotics.



II. LITERATURE SURVEY

1. "Deep Visual Odometry with Memory-Augmented Neural Networks"
This research paper introduces a novel approach to Deep Visual Odometry (DVO) that leverages memory-augmented neural networks. The model combines deep learning techniques with adaptive memory mechanisms, allowing it to learn and store scene-specific information for improved motion estimation in changing environments.

2. "Robust Visual Odometry in Challenging Environments: A Deep Learning Perspective"
This study explores the challenges of visual Odometry in challenging conditions and presents a deep learning-based approach designed to handle various environmental variations. The research investigates methods for adapting deep models to dynamic scenes and discusses the importance of adaptive memory in addressing these challenges.

3. "Adaptive Visual Odometry for Autonomous Navigation"
This research paper presents an adaptive visual Odometry system that continuously learns and updates its feature representations based on the environment's characteristics. It discusses the use of recurrent neural networks (RNNs) to incorporate memory into the visual Odometry process, improving robustness and accuracy.

4. "Sensory Data Fusion for Deep Visual Odometry"
Focusing on the integration of multiple sensors, including cameras and IMUs, this paper discusses the advantages of sensor fusion for deep visual Odometry. It explores techniques for combining sensory data and emphasizes the role of adaptive memory in enhancing localization accuracy.

5. "Memory-Augmented Deep Learning Models for Visual SLAM"
This study investigates the use of memory-augmented deep learning models in the context of Visual Simultaneous Localization and Mapping (SLAM) systems. It discusses how these models can improve the robustness and adaptability of SLAM systems by storing and retrieving critical information from previous frames.

III. BACKGROUND STUDY

Our team has developed a new approach to deep visual Odometry (VO) which is capable of selecting memory and refining postures. Traditional learning-based approaches tend to treat the VO task as a pure tracking problem, where camera postures are recreated from picture fragments, leading to significant errors over time. In order to address this issue, we have created an adaptive memory module that gradually and adaptively stores information from local to global in a neural memory. This allows us to handle long-term dependencies, even though it can be difficult to reliably store such details across whole systems. By correcting decades of mistakes, we can now access global data with greater

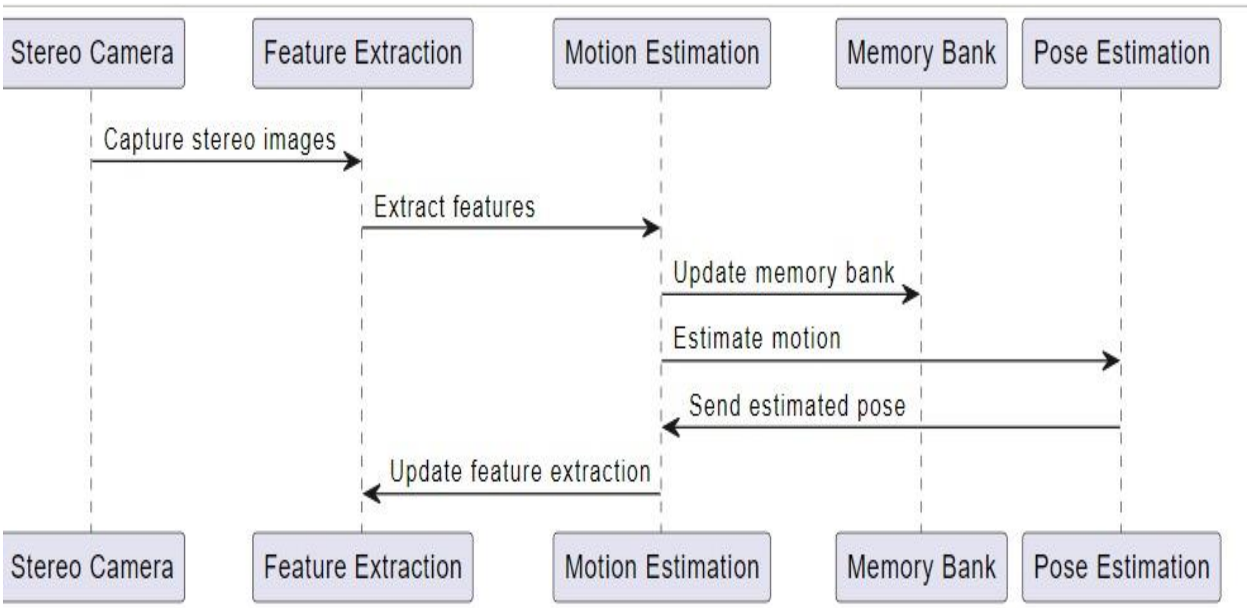
IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 8, August 2023

DOI: 10.17148/IARJSET.2023.10835

accuracy and precision.However, one disadvantage of using Support-Vector-Machines (SVM) is that they have limitations in applying new methodologies. On the other hand, Deep Learning (DL) models have become popular in recent years because they can efficiently express different interrelationships without requiring a large number of nodes as in surface-level recognition.



IV. PROPOSED METHOD

The proposed novel end-to-end Visual Odometry (VO) framework is designed with Remembering and Refining modules to work beyond pure tracking. The Remembering module effectively keeps important global information, which helps reduce errors in the Refining module. The framework adopts a hierarchical map that contains contents from paired features to the global map in the Remembering module.

This allows the model to leverage contexts at several levels for various usages. The Refining component employs the guidance of previous outputs, and a spatial temporal attention is employed based on co-visibility in feature domain to distill related features for each specific view. Our model achieves state-of-the-art performance on both the outdoor KITTI and indoor TUM-RGBD benchmark datasets. Especially, our model reports outstanding performance in challenging conditions such as textureless regions and abrupt motions, where classic methods including ORB-SLAM2 and DSO tend to fail.

V. CONCLUSION

To conclude, the study on "Deep Visual Odometry with Adaptive Memory" exemplifies the potential of combining deep learning techniques with adaptive memory mechanisms in order to enhance visual Odometry.

By utilizing deep neural networks for feature extraction, motion estimation, and temporal modelling, and integrating adaptive memory to retain and effectively utilize relevant visual information, the proposed approach aims to overcome the limitations of traditional visual Odometry methods.

FUTURE ENEHANCEMENT

Future enhancements can focus on exploring more headway deep learning architectures specifically designed for visual Odometry. This may involve investigating novel network architectures, such as attention mechanisms, graph neural networks, or transformer-based models to ameliorate feature extraction, motion estimation, and memory integration.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁝ Impact Factor 8.066 ⁝ Peer-reviewed / Refereed journal ⁝ Vol. 10, Issue 8, August 2023

DOI: 10.17148/IARJSET.2023.10835

REFERENCES

[1] R. Mur-Artal and J. D. Tardos, "ORB-SLAM2: An Open-source ´ SLAM System for Monocular, Stereo, and RGB-D Cameras," Transactions on Robotics, vol. 33, no. 5, pp. 1255–1262, 2017.
[2] A. Geiger, J. Ziegler, and C. Stiller, "Stereoscan: Dense 3D Reconstruction in Real-time," in Proceedings of the IEEE Intelligent Vehicles Symposium, 2011, pp. 963–968.
[3] G. Klein and D. Murray, "Parallel Tracking and Mapping for Small AR Workspaces," in Proceedings of the IEEE International Symposium on Mixed and Augmented Reality, 2007, pp. 225–234.
[4] J. Engel, T. Schops, and D. Cremers, "LSD-SLAM: Large-scale ¨ Direct Monocular SLAM," in Proceedings of European Conference on Computer Vision, 2014, pp. 834–849.
[5] J. Engel, V. Koltun, and D. Cremers, "Direct Sparse Odometry," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 3, pp. 611–625, 2018.
[6] R. Wang, M. Schworer, and D. Cremers, "Stereo DSO: Largescale Direct Sparse Visual Odometry with Stereo Cameras," in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 3923–3931.
[7] R. A. Newcombe, S. J. Lovegrove, and A. J. Davison, "DTAM: Dense Tracking and Mapping in Real-time," in Proceedings of the IEEE International Conference on Computer Vision, 2011, pp. 2320– 2327.
[8] J. Engel, J. Sturm, and D. Cremers, "Semi-dense Visual Odometry for a Monocular Camera," in Proceedings of the IEEE International Conference on Computer Vision, 2013, pp. 1449–1456.
[9] A. Dosovitskiy, P. Fischer, E. Ilg, P. Hausser, C. Hazirbas, V. Golkov, P. Van Der Smagt, D. Cremers, and T. Brox, "Flownet: Learning Optical Flow with Convolutional Networks," in Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 2758– 2766.
[10] D. Sun, X. Yang, M.-Y. Liu, and J. Kautz, "PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 8934–8943

