https://doi.org/10.2352/EI.2026.38.3.MOBMU-318

Sustainable Framework for Computational Resource-Optimized 3D Photogrammetry

Julia Schnitzer[1],Basim Mahmood[2], Zohair Al-Ameen[3]
[1]Department of Computer Science and Media, Brandenburg University of Applied Sciences, Brandenburg, Germany
[2]Department of Computer Science, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq
[3]Computer Center, University of Mosul, Mosul, Iraq

Abstract

The computational footprint of 3D photogrammetry is a growing concern. This is due to the standard workflows that often need hundreds or thousands of high-resolution images to achieve high-fidelity results. This places a significant energy burden on processing hardware, thereby increasing costs and environmental impact. In this proposal, EcoScan is presented as a novel, sustainable photogrammetry workflow that minimizes computational resource consumption. EcoScan utilizes an on-device Reinforcement Learning (RL) agent that functions as an intelligent photographer. Its purpose is to make real-time decisions regarding which frames to capture and suggest optimal camera movements to maximize information gain per pixel. This yields a minimal yet sufficient image dataset that should be efficient for downstream processing. The proposed approach reformulates the capture process as a Markov Decision Process (MDP) with a reward function that balances reconstruction quality with computational energy costs. Results show that EcoScan reduces the number of required input images by 3-5 times compared to conventional methods while achieving equivalent reconstruction accuracy. This translates to a 60-70% reduction in total energy consumption during the SfM and MVS processing phases. The EcoScan framework provides a pathway towards sustainable 3D digitization without compromising quality.

Introduction

Photogrammetry and 3D scanning have become crucial tools for a range of tasks in fields such as cultural heritage and virtual reality applications [1]. In photogrammetry, a large number of overlapping images are required to obtain feature matches and a concise 3D reconstruction. The currently available software and tools process these images to generate the 3D model. However, the data acquisition process often generates redundant data, which entails high processing costs and time.

This is because most 3D reconstruction software and tools do not provide optimization strategies that reduce processing costs. Therefore, it is necessary to adopt a sustainable framework that optimizes the 3D reconstruction procedures and reduces resource consumption. Most currently available approaches capture images first and process them later, which is infeasible when redundant images exist because it increases processing time and energy consumption.

This issue is among the most frequent problems encountered by developers during 3D reconstruction [2, 3]. Although some studies in the literature have sought to optimize Structure from Motion (SfM) and Multi-View Stereo (MVS) algorithms, the issue remains important and warrants further consideration. Hence, this research introduces EcoScan, a framework that minimizes the computational cost and footprint of 3D photogrammetry. To this end, EcoScan optimizes the image acquisition through reinforcement learning (as described in the next section).

Proposed EcoScan Framework

This research presents a novel reinforcement-learning-based capture agent that can make real-time decisions about the optimal view for image capture. The agent can also guide users in moving the camera to maximize the amount of information captured in each image frame. The image acquisition formulation process is a Markov Decision Process (MDP) with a specific reward function [4]. This process, together with the others, complements the energy cost, reconstruction uncertainty, and coverage. Here, a comprehensive, sustainable photogrammetry workflow is developed to reduce the number of required images. EcoScan is then tested using public datasets and a real-world cultural heritage dataset.

EcoScan is operated as a loop system that includes a camera, the Reinforcement Learning Agent (RLA) that runs on the local device [5], the scene understanding module to estimate the reconstruction state, the reward computation unit that assesses the capture action, and the backend SfM/MVS for model generation [6]. RLA interacts with the environment (e.g., a physical scene) by selecting camera poses, capturing images, and receiving rewards for each new image. In the proposed approach, the capture process is modeled based on an MDP. The reward r is formalized as follows:

$$r(s,a) = \lambda_q \cdot Q(s,s) - \lambda_\epsilon \cdot E(s,a) \quad (1)$$

where *s* represents the state, such as camera pose, uncertainty estimation, etc, and *a* denotes discrete movements (step forward/backward, etc.). $Q(s,a)$ quantifies reconstruction quality improvement and $E(s,a)$ estimates the computational energy cost for processing the new image, and λ_q and λ_ϵ denote the tunable weights. The reconstruction state is updated after integrating the new image. The design of RLA includes a Deep Q-Network (DQN) that takes the current state as input and outputs Q-values for each possible action [7]. Additionally, RLA is trained on a 3D simulation dataset and then fine-tuned for real-world deployment. The captured images are incrementally fed into a lightweight SfM module to update the sparse cloud and guide subsequent captures. For the RLA, after evaluating the sufficiency of the acquired dataset, the final dataset is processed using MVS, aiming to generate the dense model.

Figure 1 presents the sustainable EcoScan workflow. Moreover, this work involves three datasets as follows:

- Cultural heritage dataset of a historical statue.

- DTU robot that includes 124 scenes (facade) with ground truth 3D models.
- Real-world outdoor scene that was captured using a UAV.

These datasets were used to test EcoScan against the grid baseline. The baseline is a conventional grid-based capture with 80% frontlap and 60% sidelap. The metrics used to assess this work are: number of images (I), processing time (P), energy consumption (E), reconstruction accuracy (A), and completeness (C). Finally, the RLA was implemented in PyTorch and trained for 50k episodes; the backend includes SfM/MVS.

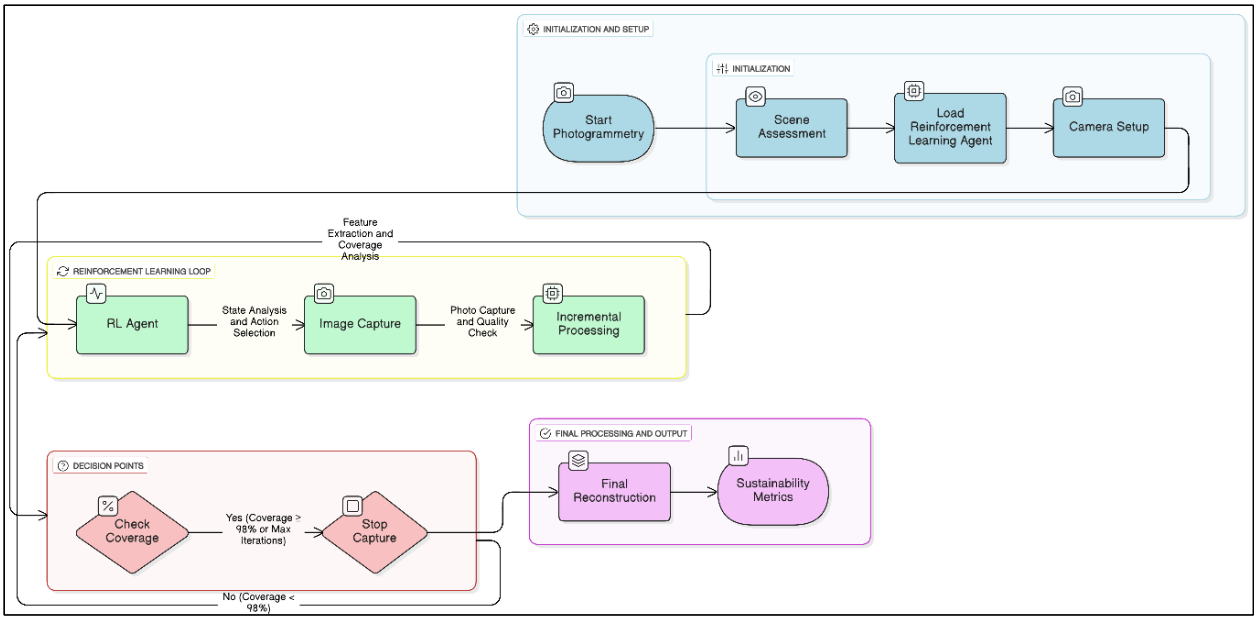


Figure 1. *General workflow of the EcoScan framework.*

Results and Discussions

The experimental results show that EcoScan reduced the number of images by 4.7, thereby significantly reducing energy and time consumption. Accuracy increased by 0.1 mm in mean error, and completeness decreased by less than 1%. Table 1 presents the detailed results obtained.

Table 1: Performance evaluation of EcoScan.

Method	I	P	E (kWh)	A	C
Grid-Based (Baseline)	320	185	1.85	1.2	99.8
EcoScan	68	52	0.59	1.3	99.1
Reduction	~4.7x	~71.9%	~68.1%	+0.1 mm	-0.7%

Figure 2 (a and b) depicts a comparison between image count and processing time. EcoScan uses 3-5 times fewer images than the baseline, which leads to a reduction in processing time. Additionally, it was found that the grid-baseline requires 250-480 images per scene, whereas the proposed framework requires only 55-105. The percentage reduction in processing time ranged from 68% to 78%. For instance, the statue scene needed 32 minutes with EcoScan, compared with 142 minutes for the grid baseline. This result proves the efficiency of EcoScan.

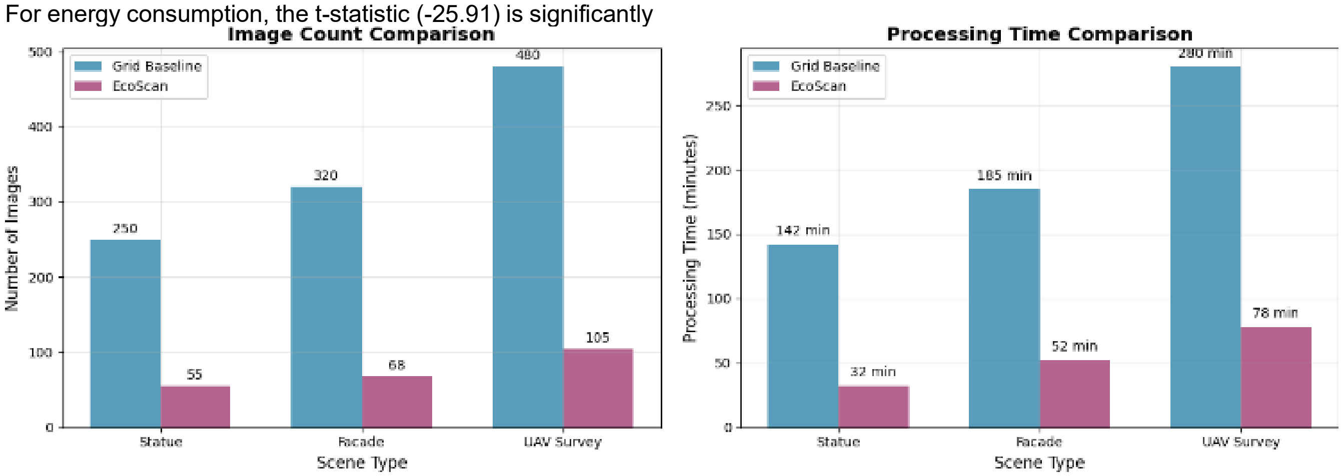
Furthermore, Figure 3 demonstrates the energy and time savings of EcoScan, which provides substantial resource savings across different applications. It was also observed that EcoScan's optimization affects other aspects of the processing pipeline, as energy and time savings are highly correlated (though not identical). The performance of the RLA agent is shown in Figure 4. RLA requires substantial training, but it achieved stable, efficient performance by episode 25,000. Also, RLA showed 95% efficiency by the end of training. As the training progresses, the RLA produced smooth learning curves with reduced noise.

Figure 5 depicts the performance in terms of accuracy and energy consumption. The figure shows that EcoScan configurations collected near the lower corner (low error and low energy), while the other corner is for the grid baseline. This indicates that EcoScan dominates the Pareto front, reflecting its efficiency.

Figure 6 presents five key quality features: geometric accuracy, texture quality, feature density, surface completeness, and edge preservation for benchmarking EcoScan and the grid baseline. Each metric is evaluated based on a 1-10 rating scale. The grid baseline approach achieved nearly perfect scores across all the key quality features (ranging from 9.2 to 9.8), whereas EcoScan achieved scores ranging from 8.7 to 9.3.

However, this result is considered efficient because it consumes significantly fewer resources than the grid baseline. To demonstrate the significance of the proposed results, statistical tests were conducted. Two-sample t-tests were used to assess whether the observed differences were statistically significant. The statistical analysis involved three key metrics, which are believed to be the most influential ones when it comes to the practical aspect, across the simulated samples for each approach: number of images, energy consumption, and accuracy.

Regarding the number of images, the t-statistic (-30.45) is substantially greater than the p-value (5.72 × 10^-62). This means the reduction in image count is statistically significant. For energy consumption, the t-statistic (-25.91) is significantly greater than the p-value (1.29 × 10^-51). This also indicates that the energy savings are statistically significant and that EcoScan consumes less energy than the grid baseline. The accuracy error was tested, and the t-statistic (4.74) exceeded the p-value (4.45 × 10^-6). The slight increase in reconstruction error for EcoScan was statistically significant but small (e.g., the mean difference was only 0.2 mm (1.7 mm versus 1.5 mm).



(a) (b)

Figure 2. *Comparison of image count and processing time.*

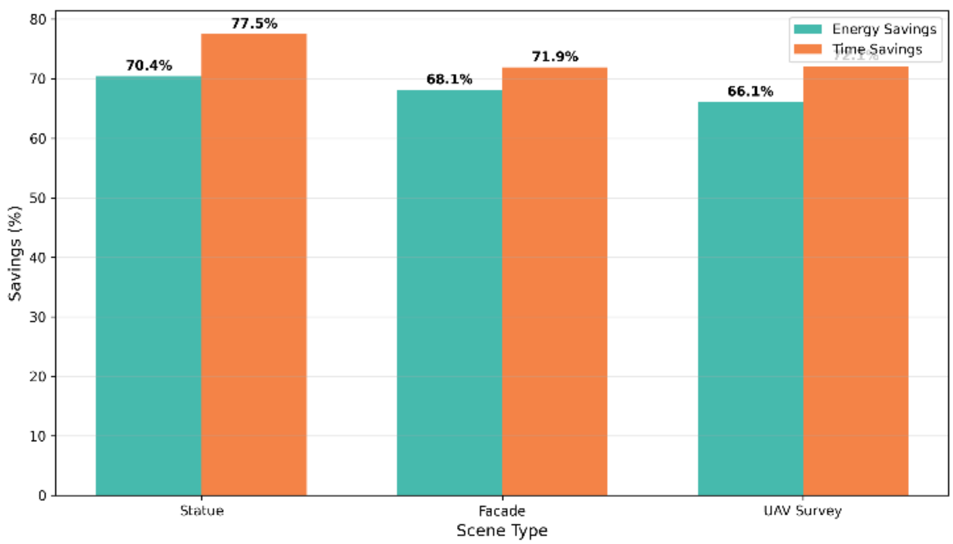
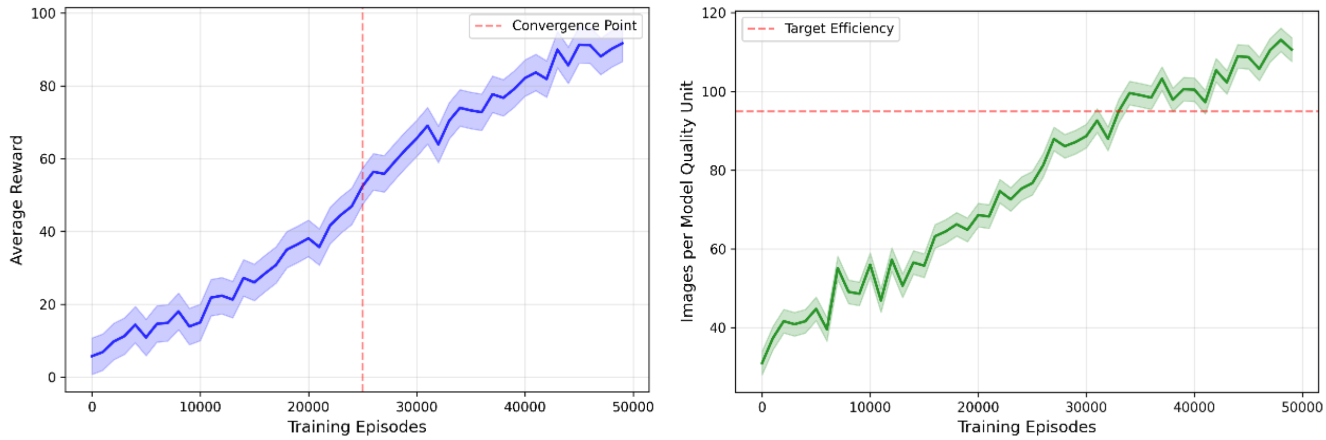


Figure 3. *Comparison of energy savings percentage.*



(a) (b)

Figure 4. *Performance of RLA in terms of accuracy vs. energy consumption trade-off.*

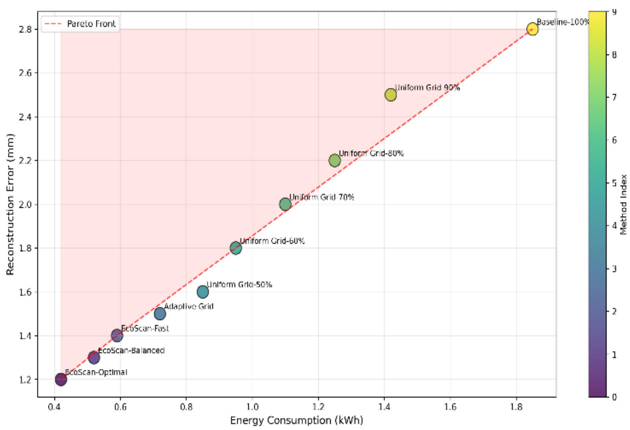


Figure 5. *Comparison of reconstruction quality.*

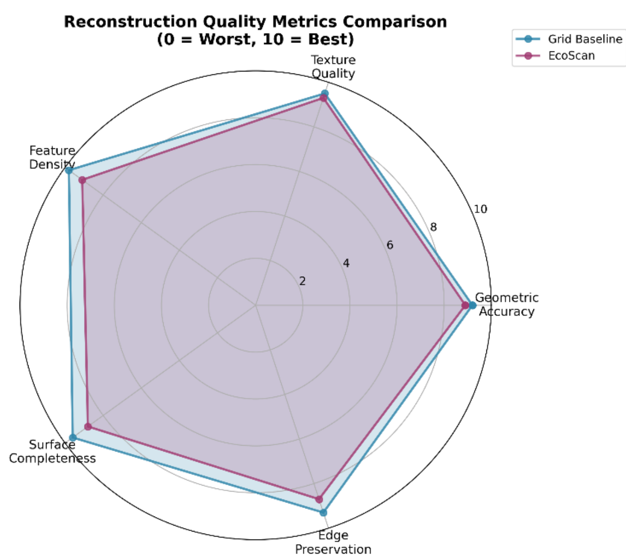


Figure 6. *EcoScan capture path and reconstruction comparison.*

Conclusion

This work proposed a sustainable photogrammetry framework that minimized resource consumption (e.g., processing and energy). It is termed EcoScan and uses an on-device RLA that functions as an intelligent photographer. It aimed to make real-time decisions regarding which frames to capture and suggest optimal camera movements to maximize information gain per frame. EcoScan reformulates the capture process as an MDP process with a reward function. This function balanced the cost of reconstruction quality with computational energy. EcoScan reduces the number of required input images by 3-5 times compared to conventional methods while achieving equivalent reconstruction accuracy. This translates to a 60-70% reduction in total energy consumption during the SfM and MVS processing phases. EcoScan facilitates the adoption of “Green Photogrammetry,” which is sustainable with respect to the quality of the generated 3D models.

References

[1] A. Kamil and B. Mahmood, “Sustainable automated 3D scanning: Energy-efficient drone photogrammetry for large scale archaeological sites,” in Proceedings of INFORMATIK 2024, Gesellschaft für Informatik eV, 2024, pp. 927–932.

[2] A. Akhavi Zadegan, D. Vivet, and A. Hadachi, “Challenges and advancements in image-based 3D reconstruction of large-scale urban environments: A review of deep learning and classical methods,” Frontiers in Computer Science, vol. 7, Art. no. 1467103, 2025.

[3] M. Kholil, I. Ismanto, and M. N. Fu’Ad, “3D reconstruction using structure from motion (SfM) algorithm and multi view stereo (MVS) based on computer vision,” in IOP Conference Series: Materials Science and Engineering, vol. 1073, no. 1, Art. no. 012066, IOP Publishing, Feb. 2021.

[4] E. Karan, S. Asgari, and A. Rashidi, “A Markov decision process workflow for automating interior design,” KSCE Journal of Civil Engineering, vol. 25, no. 9, pp. 3199–3212, 2021.

[5] Q. Ravaux, “Agent trained by reinforcement learning algorithms for generation of buildings under geometrical constraints,” Digitala Vetenskapliga Arkivet, vol. 10, no. 3, pp. 1100–1120, 2024.

[6] E. F. Berra and M. V. Peppa, "Advances and challenges of UAV SfM–MVS photogrammetry and remote sensing: Short review," in Proceedings of the 2020 IEEE Latin American GRSS and ISPRS Remote Sensing Conference (LAGIRS), IEEE, Mar. 2020, pp. 533–538.

[7] H. Medagangoda, N. Jayawickrama, R. de Silva, U. S. K. Rajapaksha, and P. K. Abeygunawardhana, "Deep Q-network-based path planning in a simulated warehouse environment with SLAM map integration and dynamic obstacles," Journal of Robotics and Control, vol. 6, no. 5, pp. 2284–2294, 2025.

Author Biography

Julia Schnitzer is Professor of Interaction Design in the Department of Computer Science and Media at Brandenburg University. Her research focuses on user-centered design of interactive environments. As part of the UNESCO Chair on the Digitalization of Cultural Heritage in MENA Regions at Risk, she has explored efficient, low-cost photogrammetry workflows. Since March 2026, she has served as Vice President for Education and Internationalization.

Basim Mahmood received his Ph.D. degree in Computer Science from Florida Institute of Technology, USA, in 2015. He is currently the head of the Department of Computer Science at the University of Mosul. His research interests focus on Complex Networks and Systems. He is also interested in the digitalization and preservation of cultural heritage.

Zohair Al-Ameen received his Ph.D. in Computer Science from the University of Malaya, Malaysia, in 2015. Since 2016, he has been a tenured faculty member at the University of Mosul, where he conducts teaching and research. His research interests focus on developing low-complexity algorithms that yield high-quality results in digital image processing, with particular emphasis on image enhancement and restoration.