

# Reliable Multi-view 3D Reconstruction for ‘Just-in-time’ Edge Environments

Md. Nurul Absur\*, Abhinav Kumar† Swastik Brahma† Saptarshi Debroy\*

\*City University of New York, USA; †University of Cincinnati, USA

Emails: \*mbsur@gradcenter.cuny.edu; †{kumar3a3, brahmask}@ucmail.uc.edu; \*saptarshi.debroy@hunter.cuny.edu

**Abstract**—Multi-view 3D reconstruction applications are revolutionizing critical use cases that require rapid situational-awareness, such as emergency response, tactical scenarios, and public safety. In many cases, their near-real-time latency requirements and ad-hoc needs for compute resources necessitate adoption of ‘Just-in-time’ edge environments where the system is set up on the fly to support the applications during the mission lifetime. However, reliability issues can arise from the inherent dynamism and operational adversities of such edge environments, resulting in spatiotemporally correlated disruptions that impact the camera operations, which can lead to sustained degradation of reconstruction quality. In this paper, we propose a novel portfolio theory inspired edge resource management strategy for reliable multi-view 3D reconstruction against possible system disruptions. Our proposed methodology can guarantee reconstruction quality satisfaction even when the cameras are prone to spatiotemporally correlated disruptions. The portfolio theoretic optimization problem is solved using a genetic algorithm that converges quickly for realistic system settings. Using publicly available and customized 3D datasets, we demonstrate the proposed camera selection strategy’s benefits in guaranteeing reliable 3D reconstruction against traditional baseline strategies, under spatiotemporal disruptions.

**Index Terms**—Reliability, Edge Computing, 3D Reconstruction, Portfolio Theory, Resource Management.

## I. INTRODUCTION

Multi-view 3D reconstruction [1] applications are increasingly becoming fundamental for complex 3D video processing application pipelines that create immersive 3D environments, such as, virtual, augmented, and mixed reality [2]. For multi-view 3D reconstruction, 2D images of a scene are simultaneously captured by cameras from multiple viewpoints and then fused together to get a 3D visualization. In many cases, such applications are hosted by critical use cases, such as emergency response, tactical scenarios, and public safety, that rely on rapid and high quality reconstruction results for mission success. In order to provide such rapid situational awareness, these use cases often rely on ‘just-in-time’ (JIT) edge environments where: i) raw video data is captured by end-devices, such as drones and robots; ii) specialty on-premise edge servers equipped with CPU/GPU and 3D video processing capabilities (which are sometimes hosted on multi-utility vehicles), process the raw video data to produce reconstructed scenes in 3D; and iii) ground consumers (e.g., tactical or first responder units) visualize the reconstructed scene on their hand-held devices.

The already tricky proposition of resource management in JIT edge environments becomes even more challenging as such ad-hoc system deployments with weakly-coupled

components are often prone to ‘disruptions’ that can have potentially catastrophic impact the reconstruction quality. In general, such disruptions may occur due to various unintended factors, such as, abrupt changes in the operating environment [3], and/or malicious factors, such as jamming attacks on wireless channels [4]. Although such disruptions can affect any edge system component (e.g., servers, network, cameras), their impact (i.e., in terms of reconstructed outcome) are most profound on the cameras and/or camera-enabled devices due to lack of redundancy (in terms of their numbers) within JIT edge environments and the relative importance of each and every camera (and its viewpoint) in generating high quality reconstruction.

These disruptions, triggered by the aforementioned causes can lead to: i) ‘spatially correlated’ failures in many cameras, impacting multiple or all at the same time and/or ii) a series of ‘temporally correlated’ failures across many cameras over a period of time. An example of spatially correlated disruption can be a sudden obstruction of view of multiple cameras due to smoke in an indoor fire emergency event, leading to lack of useful data (from such cameras) towards 3D reconstruction. Such lack of useful data even from a single camera can often result in significant reconstruction quality degradation, leading to useless reconstruction. *Thus, there is a need for reliable camera management in the presence of such spatiotemporally correlated disruptions to ensure minimum reconstruction quality satisfaction. Although there exist strategies for mission critical edge resource management [5], little work is done to ensure reliability in such environments in the presence of disruptions.*

In this paper, we propose an intelligent JIT edge resource management technique that makes multi-view 3D reconstruction in such environments reliable under system-wide disruptions. In particular, our methodology involves an innovative camera selection strategy inspired by *portfolio theory* that is aimed at mitigating the potential catastrophic effects of spatial and temporal disruptions that are commonplace in JIT edge environments [6]. First through benchmarking experiments with 3D reconstruction datasets, we demonstrate the importance and non-triviality of camera selection towards useful 3D reconstruction under disruptions. Then we formulate the camera selection problem for minimum reconstruction quality satisfaction as a portfolio theory-inspired investment fluctuation minimization problem. Next, we propose a genetic algorithm-based solution to the optimization problem that demonstrates rapid convergence in realistic scenarios, ensuring practical utility.

We evaluate the proposed edge resource management technique through a detailed simulation that mimics spatiotem-

This material is based upon work supported by the National Science Foundation (NSF) under Award Numbers: CNS-1943338 and CNS-2401928

porally correlated disruptions on JIT edge environments. Using both our own and publicly available multi-view 3D reconstruction datasets and state-of-the-art 3D reconstruction pipeline, we run extensive experiments to evaluate and compare the quality of 3D reconstruction in 3 ways: i) without disruptions, ii) under disruptions but addressed using traditional baseline approaches towards ensuring reliability, and iii) under disruption but resolved using our proposed portfolio theoretic approach. The results demonstrate that under different degrees of system-wide disruptions and camera selection stipulations, our proposed method can ensure more reliable reconstruction in terms of higher mean value and lower standard deviation of reconstruction quality than baseline strategies. The results also overwhelmingly manifest our proposed approach's superiority in guaranteeing useful reconstructed results than traditional baseline approaches, under disruptions. Overall, the results validate our theoretical constructs and emphasize the practical enhancements that our proposed approach offers to mission-critical 3D reconstruction applications hosted by JIT edge environments.

The rest of the paper is organized as follows. Section II discusses the related work in this area. Section III introduce background information on multi-view 3D reconstruction and experimentally analyses problem evidence. Section IV describes the system model and formulates the portfolio theoretic optimization problem. Section V discusses performance evaluation. Section VI concludes the paper.

## II. RELATED WORK

Disruption management in distributed and networked systems to improve system reliability and security has been an active area of research for sometime. Jeong et al. [7] discussed the importance of reducing service disruption by introducing the ReSeT system, which reduces the handover time of mobile nodes. In [8], the authors reflect on major critical analyses of edge computing regarding quality, reliability, and latency. Li et al. [9] reflect on the importance of latency management for real-life applications and propose a two-phase scheduling strategy to mitigate disruptions in edge environments. An auction theory inspired strategy is proposed in [10] to tackle disruptions caused by improper resource management and latency constraints.

There have some significant contributions in edge resource management for applications' latency satisfaction, however very few with the focus of disruption management. Among the notable works in this space, in [11], the authors propose a novel two-stage adaptive robust system to ensure resilience against any unpredictable failures. Aral et al. [12] explore possible failures while achieving quality of service (QoS) properties, such as geographical prevalence and low latency. In [13], the authors explore all possible threats that can degrade system efficiency, and propose a resilient security solution for IoT systems. The authors in [14] propose an adaptive distributed inference algorithm that can outperform traditional algorithms in terms of resilience against failures.

Although rare, in recent times, portfolio theory has been applied to computing science and engineering problems. In [6], the authors propose PortFawn, an open-source Python library that is capable of managing summery-variance portfolio optimization approaches to reduce the investment risk

of the assets. There are some work that combine portfolio and optimization theories for computing problems. In [15], a combination of three genetic algorithm variants is proposed to show better performance while solving portfolio problems with constraints. The authors in [16] demonstrate the benefits of portfolio theory by designing a multi-objective and robust probabilistic model for portfolio optimization. Another significant application of portfolio theory in computing is [17] where portfolio theory is used for traffic allocation in 5G heterogeneous cloud-RAN to achieve better quality of experience. *In most of the aforementioned work, portfolio theory is applied to optimization problems where the underlying application is mostly finance and economics. However, there is a lack of literature where the principles of balanced portfolio management are applied to explore the possibilities of reliability management in cyber infrastructures, which this work seeks to propose.*

## III. PROBLEM EVIDENCE ANALYSIS

Traditional multi-view 3D reconstruction techniques for dynamic scenes, such as SfM+MVS pipelines [18], [19], use multiple cameras to estimate the 3D information of the captured scene. Given a set of high quality (i.e., high resolution) images observing a scene from different viewpoints with mutually overlapping regions, feature points, defined by a feature descriptor data structure such as SIFT [20], are detected for each image and matched across images to signify the same 3D points. Thus, with more cameras covering the 3D scene from different viewpoints, the chances of matching more of such 3D points increases, which in turn increases reconstruction quality.

The objective of problem evidence analysis is to inquire whether choosing different sets of cameras for reconstruction would yield reconstruction of different quality. To achieve such, we use the most widely used SfM+MVS pipeline based algorithm, viz., openMVG/openMVS [18], [19] and publicly available Danc1 [21] dataset (7 synchronous video sequences from static cameras). It goes without saying that the characteristics of the following results are not predicated on the choice of reconstruction algorithm; reconstruction methods other than SfM+MVS pipelines, such as deep neural network (DNN) based algorithms [22], tend to show similar characteristics of reconstruction quality versus selected cameras for reconstruction.

Fig. 1 demonstrate the qualitative results from visual perspective, while the quantitative aspect is underpinned by the number of 3D point cloud vertices generated after the reconstruction process under different camera configurations. Fig. 1(a) showcases the optimal scenario where all cameras contribute to the reconstruction (i.e., scenarios where no disruption impacts the cameras), amassing 166,877 vertices, indicative of the most detailed and comprehensive model. However, when under disruption and Camera#2 is omitted, i.e., is unable send useful data for reconstruction, as depicted in Fig. 1(b), there is a substantial reconstruction quality reduction in terms of both visual effect and the vertices count of 82,404. This illustrates Camera#2's significant role in the 3D scene's coverage. The removal of Camera#5, however, as shown in Fig. 1(c), results in a vertices tally of 118,914, demonstrating a notable but less dramatic decline in visual

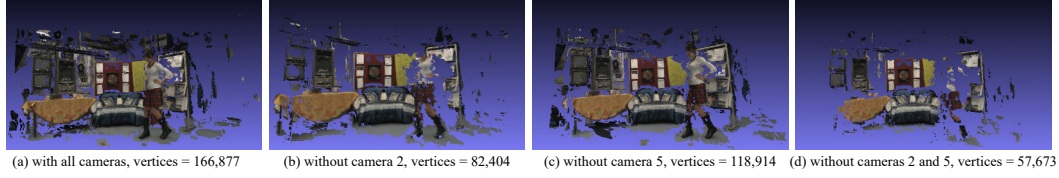


Fig. 1: Effect of camera selection on multi-view 3D reconstruction quality

details. Thus, one can argue that the role of Camera#2 is more significant than Camera#5 for generating high quality reconstruction with Dance1 [21] dataset. The combined exclusion of Camera#2 and Camera#5 in Fig. 1(d) yields the lowest vertices count of 57,673, further underscoring the criticality of these cameras in the 3D reconstruction process. *These results highlight the necessity for reliable methods to assess the trade-offs between the number of cameras and the reconstruction quality. Thus, one can argue that reconstruction dependability in situations where the integrity of camera operation as part of a JIT edge environment might be compromised, like disaster areas or tactical scenarios, depends on the careful selection of cameras. Creating methods for the best camera choice is a fascinating intellectual pursuit and a practical necessity because it directly impacts the ability to generate dependable, high-fidelity 3D models in challenging real-world situations.*

#### IV. SYSTEM MODEL AND PROBLEM FORMULATION

Our JIT edge system model, as depicted in Fig. 2, consists of a collection of  $N$  camera-enabled devices, labeled as  $\{1, \dots, N\}$ . In real-world implementations, such devices can be mobile. However, for this work, we assume static camera setting for simplicity of analysis. Each of these cameras capture an indoor or outdoor 3D scene (an outdoor park in our case) from different viewpoints in the form of a continuous stream of video frames. The video stream are transmitted to one or multiple edge servers ( $\{1, \dots, K\}$ ) in real-time where the video frames from different cameras captured at the same time-stamp are fused together using any state-of-the-art 3D reconstruction algorithm, to generate a 3D visualisation of the scene for that particular time-stamp. This process is repeated for all time-stamps and for the entire duration of the underlying application.

For most JIT edge implementations, cameras and their supporting devices can be heterogeneous in nature, in terms of the device hardware, camera specifications, image resolutions, among other properties. This makes the problem of device/camera selection for reasonable reconstruction quality satisfaction under disruption, a non-trivial undertaking. In this work, we assume heterogeneity only in terms of camera image resolutions for simplicity of analysis, i.e., we consider that each camera  $i$  is capable of sending video frames/images to edge servers at resolution  $R_i$ . Further, defining  $p_{c_i}$  to be the probability with which camera  $i$  is *not* disrupted, we consider that a camera  $i$ , transmits images with specified image resolution  $R_i$  with a probability  $p_{c_i}$ , and gets disrupted, producing no usable data, with the probability  $1 - p_{c_i}$ .

To model *fluctuations* of the disruption probabilities of the cameras, we consider that  $p_{c_i}$  is a *random variable*. Further, to model possible correlations between the disruptions experienced by cameras  $i$  and  $j$ , we consider that the correlation

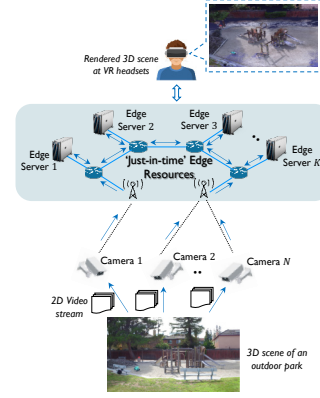


Fig. 2: Multi-view 3D reconstruction of an outdoor scene hosted at JIT edge environment

coefficient between  $p_{c_i}$  and  $p_{c_j}$  is  $\rho - i, j$ . This correlation significantly influences the likelihood of simultaneous failures, thereby impacting the system's resilience.

#### A. Portfolio Theoretic Formulation

In such a scenario, it can be noted that the resolution of the image obtained from camera  $i$ , say denoted as  $\mathcal{R}_i$ , is itself a random variable, with  $\mathcal{R}_i = 0$  (when camera  $i$  gets disrupted, which occurs with probability  $p_{c_i}$ ) and  $\mathcal{R}_i = R_i$  (when camera  $i$  is not disrupted), with  $\mathbb{E}[\mathcal{R}_i] = R_i \mathbb{E}[p_{c_i}]$  (where  $\mathbb{E}[\cdot]$  denotes expectation), which occurs with probability  $p_{c_i}$ ). Further, for cameras  $i$  and  $j$ , with  $p_{c_i}$  and  $p_{c_j}$  being correlated random variables, with  $\rho_{i,j}$  being their correlation coefficient,  $\mathcal{R}_i$  and  $\mathcal{R}_j$  also become correlated in nature whose covariance is  $\text{cov}(\mathcal{R}_i, \mathcal{R}_j) = R_i R_j \sigma_{p_{c_i}} \sigma_{p_{c_j}} \rho_{i,j}$ , where  $\sigma_{p_{c_i}}$  is the standard deviation of  $p_{c_i}$ . Thus, the more correlated the disruptions of two cameras, the higher the covariance among the resolutions of the images obtained from them, and subsequently the higher the chances of the two cameras *simultaneously* experiencing disruptions (and providing 0 resolution images *together*), which negatively impacts performance. To perform camera selection in such a scenario, suppose that camera  $i$  is selected with a probability  $\alpha_i$ , with  $\alpha = (\alpha_1, \dots, \alpha_N)$  denoting the camera selection vector, leading to the *expected* total resolution of the images obtained from the cameras to be  $\sum_{i=1}^N \alpha_i \mathbb{E}[\mathcal{R}_i]$ . For simplicity of analysis, in this work, we consider  $\alpha_i$  to be a binary variable where  $\alpha_i = 1$  signifies selection of camera  $i$  for reconstruction.

Portfolio theoretic resource management in the above problem scenario would enable selection of cameras such that a certain expected image quality (resolution) obtained from the cameras can be ensured (to sustain a desired quality of the reconstructed scene) while minimizing the

covariances among the selected cameras' image qualities to ensure reliability. Such a strategy would permit us to achieve optimal tradeoffs between the quality and reliability of a 3D reconstruction task in order to sustain its desired performance using disruption-prone limited system resources. Specifically, the portfolio theoretic optimization problem for selecting cameras can be formulated as (1).

$$\underset{\alpha=(\alpha_1, \dots, \alpha_N)}{\text{minimize}} \quad \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \text{cov}(\mathcal{R}_i, \mathcal{R}_j) \quad (1a)$$

$$\text{subject to} \quad \sum_{i=1}^N \alpha_i \mathbb{E}[\mathcal{R}_i] \geq \Theta \quad (\text{quality constraint}) \quad (1b)$$

$$\sum_{i=1}^N \alpha_i \leq \Psi \quad (\text{resource constraint}) \quad (1c)$$

$$\alpha_i \in [0, 1], \quad \forall i \quad (\text{selection constraint}) \quad (1d)$$

Note that (1) seeks to determine the camera selection strategy  $\alpha = (\alpha_1, \dots, \alpha_N)$  that would, to provide *reliability*, minimize the quantity  $\sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \text{cov}(\mathcal{R}_i, \mathcal{R}_j)$ , which determines the *variability* of the quality of the reconstructed scene. The constraints of (1) are the:

- *quality constraint* (1b), which ensures that the expected total resolution of images obtained from the cameras satisfy a given threshold  $\Theta$  (to make the reconstructed scene meet quality requirements). Note that  $\Theta$  controls the tradeoff between quality and reliability, with higher  $\Theta$  values emphasizing selection of cameras that make the reconstructed scene to expectedly have a good quality (while sacrificing reliability) and lower  $\Theta$  values reversing the trend;
- *resource constraint* (1c), which ensures that the total number of selected cameras do not exceed  $\Psi$  (to meet factors like bandwidth and energy limitations); and
- *selection parameter constraint* (1d), which ensures camera selection parameter definition.

## B. Genetic Algorithmic Solution

To solve the optimization problem in (1), we employ a genetic algorithm (GA) that efficiently explores the large solution space to find optimal camera configurations, as described below:

*Initial population:* The initial population is generated by randomly selecting camera configurations. Each individual represents a potential solution, encoded as a binary vector indicating camera inclusion.

*Fitness function:* The fitness function evaluates each camera configuration by balancing two key factors: achieving a minimum quality threshold ( $\Theta$ ) and minimizing the risk of correlated failures. The total quality of the selected cameras is computed as the weighted sum of their resolutions, where each camera's importance is captured by its respective weight ( $\alpha$ ). If the total quality falls below the threshold  $\Theta$ , a large penalty is assigned, ensuring only configurations meeting the minimum quality requirement are viable. If the quality exceeds  $\Theta$ , the risk of correlated failures is calculated based on the covariance of image resolutions between selected cameras, adjusted by their respective weights ( $\alpha$ ). This risk minimization encourages robust configurations that reduce the likelihood of simultaneous failures. The above described fitness function is defined as:

$$\mathcal{F} = \begin{cases} \infty, & \text{if } \sum_{i=1}^N \alpha_i > \Psi, \\ \Theta - \sum_{i=1}^N \alpha_i \mathbb{E}[\mathcal{R}_i], & \text{if } \sum_{i=1}^N \alpha_i \mathbb{E}[\mathcal{R}_i] < \Theta, \\ \sum_{i=1}^N \sum_{j=i}^N \alpha_i \alpha_j R_i R_j \rho_{ij} \sigma_i \sigma_j, & \text{otherwise.} \end{cases}$$

*GA parameters:* We define key parameters for the GA, including:

- **Population size** ( $P_s$ ): Number of individuals in the population.
- **Maximum generations** ( $G_m$ ): The maximum number of iterations or generations the algorithm will run.
- **Crossover rate** ( $C_r$ ): Probability of combining two parents to generate offspring.
- **Mutation rate** ( $M_r$ ): Probability of altering an individual's configuration to maintain genetic diversity.
- **Elitism count** ( $E_c$ ): The number of top-performing individuals retained in each generation.
- **Updated population** ( $P_{gen+1}$ ): The new population for the next generation, formed by the union of elite individuals and offspring:

$$P_{gen+1} \leftarrow E_c \cup \text{offspring}$$

where  $E_c$  represents the elite individuals and offspring represents the children generated through crossover and mutation.

## V. PERFORMANCE EVALUATION

Here, we discuss the performance evaluation of our proposed portfolio theoretic edge resource management technique.

### A. Simulation Environment

For the evaluation, we create a detailed simulation with a realistic JIT edge environment and video dataset that is customized for multi-view 3D reconstruction. Most of the simulation parameters are chosen based on the dataset and typical camera specifications and operational constraints within a JIT edge environment to ensure a realistic and practical evaluation. For one of the datasets consisting of 7 simultaneous video streams, we set the number of cameras ( $N$ ) for reconstruction to 7 with 3 cameras having a high resolution (1080p, i.e.,  $1920 \times 1080$  pixels) and the remaining 4 with a lower resolution (720p, i.e.,  $1280 \times 720$  pixels). For another dataset with 5 simultaneous video streams,  $N$  is set to 5 with 2 cameras having high resolution.

In order to simulate the disruptions, each camera's failure probability ( $p_{c_i}$ ) is considered to be a random variable that follows the Beta distribution. The Beta distribution is ideal for modeling probabilities due to its bounded range between 0 and 1. For the Beta distribution, parameter  $x$  values are set to  $[6, 6, 6, 2, 2.5, 3.5, 5]$  and parameter  $y$  values are set to  $[3, 3, 3, 3, 3.5, 2.5, 2]$ . Additionally, a  $7 \times 7$  correlation matrix  $p_{ij}$  is used to model the correlation of camera failures, with values ranging from 0.1 to 1, capturing how disruptions in cameras statistically depend on each other. This combination of Beta-distributed probabilities and the correlation matrix

ensures that the simulation realistically reflects disruption patterns and system performance under varying constraints.

### B. Datasets used

We use popular publicly available multi-view 3D reconstruction dataset, viz., Dance1 and Odzemok [21], for performance analysis of dynamic scenes. Both are indoor datasets with cluttered backgrounds with a moving person and static background. Both datasets have 7 static cameras with resolutions as explained before. We also use our own multi-view dataset [23] [24] that represent different degrees of indoor dynamic scenes, e.g., Walk and Handshake. The walk dataset is moderately dynamic, where a person walks across the scene from one end to another. On the other hand, the Handshake dataset represents a high dynamism, where two people walk from two different ends of the scene and shake hands in the middle. These datasets demonstrate the proposed method's utility in providing reliable reconstruction where the 3D scene is highly dynamic.

### C. Performance metrics

For the evaluation, we measure the performance of the proposed portfolio theoretic approach by estimating how reliably it can provide useful reconstruction. Consequently, we measure the reliability of 3D reconstruction (i.e., with and without our proposed portfolio theoretic approach) by ascertaining how consistently our approach generate reconstruction with a certain minimum quality (i.e., Threshold  $\Theta$ ). Such quality is measured by the number of vertices in the reconstructed 3D point cloud output. A higher number of vertices count signifies higher quality reconstruction, capturing greater scene details and enhancing the overall fidelity of the reconstruction. The consistency of generating useful reconstruction is measured in terms of mean and standard deviation (S.D.) of number of vertices in the reconstructed 3D scene over different disruption structures.

For simplicity, the quality threshold ( $\Theta$ ) is set at half of the highest resolution  $(1920 \times 1080)/2 = 1,036,800$ , a common practice in computer vision for image optimization across various sets of resolutions[25]. Such threshold approximation provides a balanced estimation that ensures high-quality reconstruction while accommodating a mix of camera resolutions. For all the following results, we use the upper bound of resource constraint ( $\Psi$ ), i.e.,  $\Psi'$  (from Eq. (1c)) as a system constant. Further, for all experiments, the mean and S.D. of reconstruction quality computed over 20 runs of Beta distributed camera failure probabilities.

### D. Baseline Strategies

In this paper, we define two baseline strategies to evaluate two separate aspects of our solution.

*Optimal (Brute force) vs. GA based solution:* In order to ascertain the accuracy of GA based camera selection, we choose a baseline strategy of optimal camera selection which is generated through exhaustive search of all possible camera selection combination and choosing the one with highest combined resolution. Of course, performing such exhaustive search with larger number of cameras to find the optimal camera selection is computationally challenging and thus impractical. However, the purpose of such baseline

comparison is to observe how camera selection with the proposed GA based solution performs against such optimal solution.

*Traditional vs. Portfolio-theoretic camera selection:* In another set of comparison, we use a traditional method of edge resource management (specifically camera selection) where cameras with highest resolution are always selected in comparison to portfolio theoretic approach that chooses cameras based on the solution of Eq. (1a).

### E. Optimality of genetic algorithmic solution

In Table I, we present the comparison between the camera selection strategy obtained by solving Eq. (1a) using our GA-based approach and the one obtained by solving it using a brute force method. While the brute force method evaluates all possible combinations exhaustively, the GA employs evolutionary principles, such as selection, crossover, and mutation, to efficiently explore the solution space. In Table I, we observe that for different camera selection scenarios (i.e., different  $\Psi$ ), the solutions provided by the two are predominantly the same. Such consistency validates GA's effectiveness in solving the optimization problem. One also has to note that although the GA seeks to find the near-optimal solution, it is after all an approximation strategy which can result in a solution that is slightly different from the optimal solution. However, what the proposed GA sacrifices in accuracy, it more than makes it up by employing an iterative approach that can handle larger problem sizes (i.e., number of cameras in our case) more feasibly, providing near-optimal solutions with significantly reduced computational requirements. Such consistency also offers a practical justification for using the GA in real-world applications, as it reconciles the exhaustive search with the iterative approach, ensuring reliable and relevant outcomes.

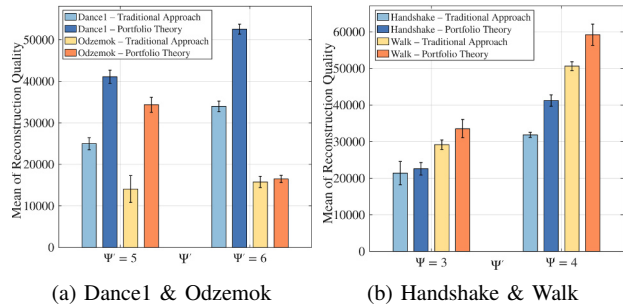


Fig. 3: Mean of Reconstruction Quality for Traditional vs. Portfolio.

### F. Reliability Comparison Against Traditional Approach

Here we present comparative results between two approaches —Portfolio Theoretic camera selection and Traditional Approach towards camera selection—in terms of their performance in ensuring reliable reconstruction under disruptions. As explained earlier, such reliability is measured in terms of the approaches' ability to ensure consistently high quality reconstruction over threshold ( $\Theta$ ). Further, in the following results, mean and standard deviation (S.D.) of reconstruction quality (in terms of number of vertices)



Camera	6 out of 7 ( $\Psi' = 6$ )		5 out of 7 ( $\Psi' = 5$ )		4 out of 7 ( $\Psi' = 4$ )	
	Genetic Algorithm	Optimal	Genetic Algorithm	Optimal	Genetic Algorithm	Optimal
1	✗	✗	✗	✗	✓	✗
2	✓	✓	✗	✗	✓	✗
3	✓	✓	✓	✓	✗	✗
4	✓	✓	✓	✓	✗	✓
5	✓	✓	✓	✓	✓	✓
6	✓	✓	✓	✓	✗	✓
7	✓	✓	✓	✓	✓	✓

TABLE I: Camera Selection Comparison (✓ => Selected, ✗ => Not selected)

are measured over multiple images sets (each image set is a set of images from different cameras belonging to the same timestamp) belonging to different datasets. Choosing multiple image sets from the same dataset examines how the proposed portfolio theoretic camera selection strategy preforms (in terms of generating reliable reconstruction) over a longer period as each dataset contains synchronized images from multiple cameras over a long term period in terms of streams of images. While choice of multiple datasets investigates the proposed method's versatility in dealing with 3D scenes that have a varied degree of dynamism as explained earlier.

1) *Mean of reconstruction quality for different disruption structures*: Figure 3 displays the mean reconstruction quality (in terms of number of vertices) over multiple image sets across datasets as a function of the upper bound on the number of cameras to be used for reconstruction ( $\Psi'$ ). Error bars represent the variance across 20 different runs, providing insight into the consistency of each method.

In Figure 3a (Dance1 & Odzemok), *Dance1* shows a clear gain: at  $\Psi' = 5$  Portfolio Theory is  $\sim 40,000$  vs.  $\sim 25,000$  for the Traditional Approach; at  $\Psi' = 6$  Portfolio Theory reaches  $\sim 52,000$  vs.  $\sim 33,000$ . For *Odzemok*, at  $\Psi' = 5$  Portfolio Theory is  $\sim 34,000$  vs.  $\sim 15,000$  for Traditional, whereas at  $\Psi' = 6$  both methods are much lower (Traditional  $\sim 16,000$ , Portfolio  $\sim 17,000$ ), with only a small gap. Across both datasets, Portfolio Theory generally exhibits tighter error bars.

In Figure 3b (Handshake & Walk), *Handshake* shows a modest advantage at  $\Psi' = 3$  (Portfolio  $\sim 24,000$  vs. Traditional  $\sim 22,000$ ) that widens at  $\Psi' = 4$  (Portfolio  $\sim 41,000$  vs.  $\sim 32,000$ ). *Walk* shows a consistent and larger gain: at  $\Psi' = 3$  Portfolio is  $\sim 34,000$  vs.  $\sim 29,000$ , and at  $\Psi' = 4$  Portfolio is nearly  $\sim 60,000$  vs.  $\sim 51,000$  for Traditional, again with narrower error bars for Portfolio Theory.

2) *S.D. of reconstruction quality for different disruption structures*: This section presents a quantitative comparison between Portfolio Theory and the Traditional Approach in terms of reconstruction reliability, measured by the standard deviation (S.D.) of reconstruction quality across different image sets in four datasets. Figure 4 shows the S.D. values as a function of  $\Psi'$ , with error bars representing variability over 20 different runs.

In Figure 4a (Dance1 & Odzemok), *Dance1* shows a clear reduction in variability: at  $\Psi' = 5$  the Traditional Approach records about 47,000, compared to  $\sim 35,000$  for Portfolio Theory; at  $\Psi' = 6$  the gap widens, with Traditional near 49,000 and Portfolio below 30,000. For *Odzemok*, at  $\Psi' = 5$  the S.D. is  $\sim 28,000$  for Traditional vs.  $\sim 28,000$  for Portfolio Theory (nearly equal), while at  $\Psi' = 6$  Traditional stays near 29,000 and Portfolio drops to about 20,000. Across both datasets, Portfolio Theory tends to produce smaller error bars,

indicating greater consistency.

In Figure 4b (Handshake & Walk), *Handshake* shows a large difference at  $\Psi' = 3$ , with Traditional at  $\sim 63,000$  and Portfolio near 31,000; at  $\Psi' = 4$ , Traditional is  $\sim 32,000$  and Portfolio  $\sim 28,000$ . For *Walk*, at  $\Psi' = 3$  the Traditional Approach records  $\sim 66,000$  vs.  $\sim 42,000$  for Portfolio Theory; at  $\Psi' = 4$  both are closer (Traditional  $\sim 39,000$ , Portfolio  $\sim 37,000$ ) but Portfolio still shows slightly less variability.

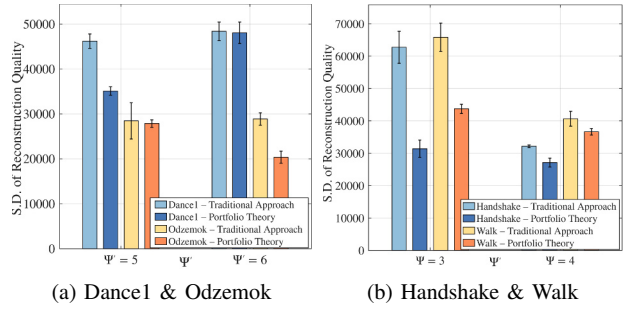


Fig. 4: Reconstruction S.D. of Reliability for Traditional vs. Portfolio.

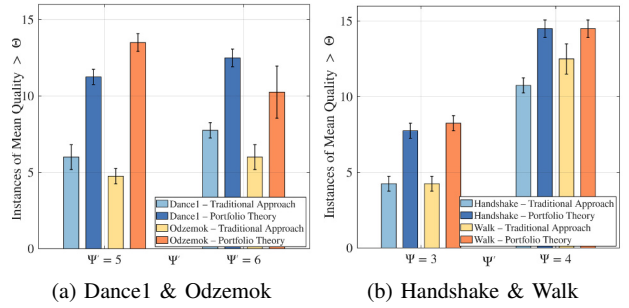


Fig. 5: Reconstruction Quality Over Threshold for Traditional vs. Portfolio.

3) *Mean reconstruction quality above  $\Theta$  for different disruption structures*: In this section, we evaluate the number of instances where the mean reconstruction quality surpasses a threshold  $\Theta$  across four datasets. Figure 5 shows these instances as a function of  $\Psi'$ , with error bars indicating variability over 20 runs and highlighting consistency.

In Figure 5a (Dance1 & Odzemok), *Dance1* at  $\Psi' = 5$  records about 6 instances for the Traditional Approach and 12 for Portfolio Theory; at  $\Psi' = 6$  these rise to roughly 8 and 13, respectively. For *Odzemok*, at  $\Psi' = 5$  the Traditional Approach has around 5 instances versus 14 for Portfolio

Theory; at  $\Psi' = 6$  the values are about 6 and 10. In both datasets, Portfolio Theory generally shows smaller error bars, indicating greater consistency.

In Figure 5b (Handshake & Walk), *Handshake* at  $\Psi' = 3$  achieves about 4 instances for the Traditional Approach compared to 7 for Portfolio Theory; at  $\Psi' = 4$  the counts are roughly 11 and 14. For *Walk*, at  $\Psi' = 3$  the Traditional Approach records about 5 instances versus 8 for Portfolio Theory; at  $\Psi' = 4$  these increase to about 13 and 15, respectively. Across both datasets, Portfolio Theory maintains narrower error bars, reflecting its stronger reliability.

### G. Impact of Disruption Correlations on Performance

Here, in Figures 6 and 7, we study how the degree of correlations ( $\rho_{i,j}$ ) among disruptions experienced by cameras impact 3D reconstruction performance using the *Dance1* and *Odzemok* datasets. Reconstruction reliability, measured through the mean vertices metric, demonstrates that a portfolio theory-based approach significantly outperforms traditional methods. By optimizing the trade-off between resolution and dependability, the portfolio theory approach exhibits consistent reliability across diverse correlation levels, underscoring its potential to advance 3D image reconstruction techniques for real-time, high-resolution applications.

1) *Reconstruction Reliability*: The analysis of mean reconstruction reliability across the *Dance1* and *Odzemok* datasets highlights the consistent superiority of the portfolio theory compared to the traditional approach, particularly as the correlation among higher-resolution cameras ( $\rho_{i,j}$ ) increases. For the *Dance1* dataset (Figure 6a), the traditional approach exhibits a sharp decline in reconstruction quality as  $\rho_{i,j}$  increases from 0.2 to 0.8. At  $\rho_{i,j} = 0.2$ , the traditional approach achieves a mean reconstruction quality of approximately 30,000, while the portfolio theory significantly outperforms it with a quality exceeding 50,000. As  $\rho_{i,j}$  increases, the traditional approach's quality plummets to below 15,000 at  $\rho_{i,j} = 0.8$ , whereas the portfolio theory demonstrates remarkable stability, retaining a quality of nearly 25,000 even under high correlation conditions.

A similar pattern is observed in the *Odzemok* dataset (Figure 6b). The traditional approach begins with a quality of around 25,000 at  $\rho_{i,j} = 0.2$  but declines sharply, reaching below 10,000 at  $\rho_{i,j} = 0.8$ . In contrast, the portfolio theory starts with a quality exceeding 40,000 at  $\rho_{i,j} = 0.2$  and maintains a significant advantage across all correlation levels, achieving approximately 20,000 at  $\rho_{i,j} = 0.8$ . These results underscore the robustness of the portfolio theory in managing high-resolution camera data with varying correlation levels. Its ability to maintain superior reconstruction quality and reliability, even under challenging conditions, highlights its potential for real-time 3D reconstruction tasks where consistency and performance are critical.

2) *Reconstruction S.D. of Reliability*: Analyzing the standard deviation (SD) of reconstruction quality across the *Dance1* and *Odzemok* datasets provides valuable insights into each method's robustness. A higher SD indicates greater variability and, consequently, reduced reliability. For the *Dance1* dataset (Figure 6c), as the correlation  $\rho$  increases from 0.2 to 0.8, the traditional approach exhibits a significant rise in SD, signaling diminishing robustness. In contrast,

the portfolio theory approach maintains a consistently lower SD, highlighting its ability to ensure stable and reliable reconstruction across varying correlation levels.

Based on the *Odzemok* dataset (Figure 6c), the findings confirm the *Dance1* patterns. The traditional approach shows that as correlation increases, robustness declines, indicating an increase in variability in dependability with a more significant standard deviation. In contrast, the portfolio theory demonstrates a significant standard deviation consistency at different correlation values. The technique's inherent robustness is highlighted by the slight increase in standard deviation, even at high correlation levels.

3) *Reconstruction Quality Over Threshold*: The analysis of the *Dance1* dataset (Figure 7a) demonstrates the portfolio theory's consistent superiority over the traditional approach across varying correlation levels ( $\rho$ ). At lower correlation levels ( $\rho = 0.2$ ), the portfolio theory achieves substantially higher reconstruction quality, indicating its exceptional performance when camera outputs exhibit minimal similarity. As  $\rho$  increases, both methods experience a decline in quality, yet the portfolio theory consistently maintains an edge over the traditional approach, confirming its robustness in ensuring higher-quality reconstructions.

Similarly, in the *Odzemok* dataset (Figure 7b), portfolio theory consistently outperforms the traditional approach, particularly at lower correlation levels ( $\rho = 0.2$ ). Even at higher correlations ( $\rho = 0.6$  and  $\rho = 0.8$ ), portfolio theory sustains a higher quality threshold, underscoring its capability to deliver superior reconstruction quality under challenging conditions. These results make portfolio theory the preferred choice for high-resolution, real-time 3D reconstruction applications requiring resilience to fluctuating camera correlations.

Our analysis reveals a clear pattern where the portfolio theory consistently outperforms the traditional approach to provide superior reconstruction quality and reliability in the presence of varying degrees of correlated disruptions. This consistently superior performance highlights the potential of the portfolio theory to deliver high-quality outputs for real-time applications, making it the preferred choice for systems that demand reliable quality in dynamic environments.

## VI. CONCLUSIONS

In this work, we introduced a novel portfolio-theoretic approach to optimize 3D reconstruction scenes, effectively balancing trade-offs between quality and reliability. Experimental results using real datasets demonstrated significant performance advantages over traditional methods, and the diverse datasets highlighted the scalability of our approach. We also proposed a genetic algorithm for efficiently finding camera selection strategies from our portfolio-theoretic perspective. Looking ahead, we will develop a comprehensive system-wide model that will integrate advanced technologies such as deep learning, system monitoring, and constrained optimization, aiming to understand and model spatiotemporally correlated disruptions while formulating a management strategy that will adaptively respond to these disruptions among cameras. Additionally, we will analytically quantify the trade-offs between quality, latency, and reliability under disruptions, ultimately striving to measure these trade-offs using advanced deep neural networks.

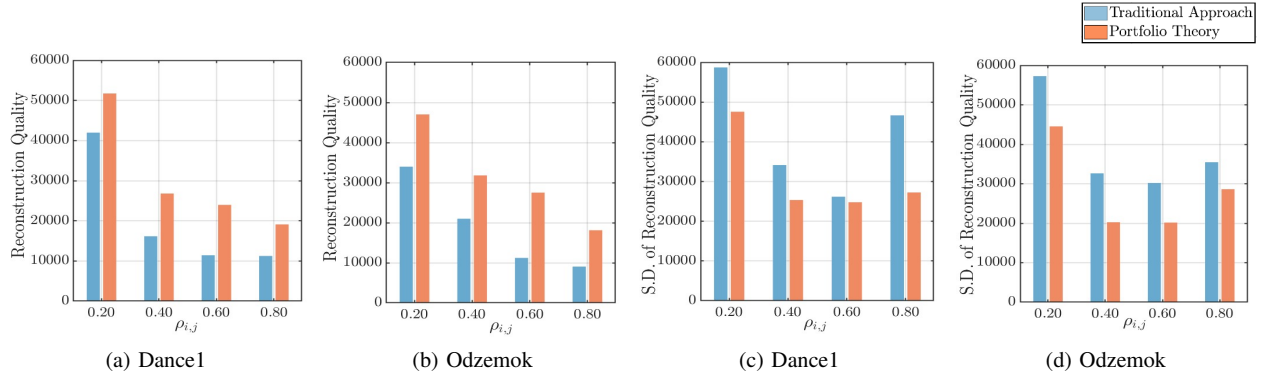


Fig. 6: Impact of Varying  $\rho_{i,j}$  on Reconstruction Quality

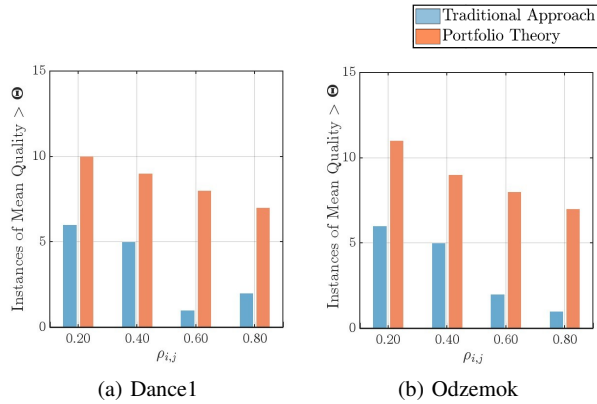


Fig. 7: Impact of Varying  $\rho_{i,j}$  on Reconstruction Reliability.

#### REFERENCES

- [1] X. Zhang, H. Gan, A. Pal, S. Dey, and S. Debroy, "On balancing latency and quality of edge-native multi-view 3d reconstruction," in *2023 IEEE/ACM Symposium on Edge Computing (SEC)*, 2023, pp. 1–13.
- [2] L. Han, T. Zheng, Y. Zhu, L. Xu, and L. Fang, "Live semantic 3d perception for immersive augmented reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 5, pp. 2012–2022, 2020.
- [3] S. Xiang and B. Tang, "Kernel-based edge-preserving methods for abrupt change detection," *IEEE Signal Processing Letters*, vol. 27, pp. 86–90, 2020.
- [4] B. Yu and L.-Y. Zhang, "An improved detection method for different types of jamming attacks in wireless networks," in *The 2014 2nd International Conference on Systems and Informatics (ICSAI 2014)*, 2014, pp. 553–558.
- [5] X. Zhang and S. Debroy, "Resource management in mobile edge computing: A comprehensive survey," *ACM Comput. Surv.*, vol. 55, no. 13s, Jul. 2023. [Online]. Available: <https://doi.org/10.1145/3589639>
- [6] M. Owahdi-Kareshk and P. Boulanger, "Portfolio optimization on classical and quantum computers using portfawn," 2021.
- [7] J. Jeong, J. Ha, and M. Kim, "Reset: Reducing the service disruption time of follow me edges over wide area networks," in *2019 22nd Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*, 2019, pp. 159–166.
- [8] H. Udeen, S. Abbas, A. Mumtaz, F. Hussain, and K. Kousar, "Critical analysis of edge computing," in *2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP)*, 2020, pp. 1092–1099.
- [9] X. Li, J. Wan, H.-N. Dai, M. Imran, M. Xia, and A. Celesti, "A hybrid computing solution and resource scheduling strategy for edge computing in smart manufacturing," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 7, pp. 4225–4234, 2019.
- [10] C. Avasalcay, C. Tsigkanos, and S. Dustdar, "Decentralized resource auctioning for latency-sensitive edge computing," in *2019 IEEE International Conference on Edge Computing (EDGE)*, 2019, pp. 72–76.
- [11] J. Cheng, D. T. Nguyen, and V. K. Bhargava, "Resilient edge service placement under demand and node failure uncertainties," 2023.
- [12] A. Aral and I. Brandic, "Dependency mining for service resilience at the edge," in *2018 IEEE/ACM Symposium on Edge Computing (SEC)*, 2018, pp. 228–242.
- [13] H. Kim, E. A. Lee, and S. Dustdar, "Creating a resilient iot with edge computing," *Computer*, vol. 52, no. 8, pp. 43–53, 2019.
- [14] P. Li, E. Koyuncu, and H. Seferoglu, "Adaptive and resilient model-distributed inference in edge computing systems," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 1263–1273, 2023.
- [15] I. Strumberger, N. Bacanin, and M. Tuba, "Constrained portfolio optimization by hybridized bat algorithm," in *2016 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, 2016, pp. 83–88.
- [16] M. Shaverdi, S. Yaghoubi, and H. Ensafian, "A multi-objective robust possibilistic model for technology portfolio optimization considering social impact and different types of financing," *Applied Soft Computing*, vol. 86, p. 105892, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1568494619306738>
- [17] V. Miliotis, N. Makris, V. Passas, and T. Korakis, "Portfolio theory application for 5g heterogeneous cloud-ran infrastructure," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.
- [18] P. Moulon, P. Monasse, R. Perrot, and R. Marlet, "OpenMVG: Open multiple view geometry," in *International Workshop on Reproducible Research in Pattern Recognition*. Springer, 2016, pp. 60–74.
- [19] Z. Li, S. Luo, W. Zeng, S. Guo, J. Zhuo, L. Zhou, Z. Ma, and Z. Zhang, "3d reconstruction system for foot arch detecting based on openmvg and openmvs," in *2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, 2022, pp. 1017–1022.
- [20] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004. [Online]. Available: <https://api.semanticscholar.org/CorpusID:174065>
- [21] A. Mustafa, H. Kim, J.-Y. Guillemaut, and A. Hilton, "General dynamic scene reconstruction from multiple view video," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, December 2015.
- [22] Y. Yao, Z. Luo, S. Li, T. Fang, and L. Quan, "Mvsnet: Depth inference for unstructured multi-view stereo," *European Conference on Computer Vision (ECCV)*, 2018.
- [23] X. Zhang, H. Gan, A. Pal, S. Dey, and S. Debroy, "On balancing latency and quality of edge-native multi-view 3d reconstruction," in *Proceedings of the Eighth ACM/IEEE Symposium on Edge Computing*, ser. SEC '23. New York, NY, USA: Association for Computing Machinery, 2024, p. 1–13. [Online]. Available: <https://doi.org/10.1145/3583740.3630267>
- [24] M. N. Absur, S. Brahma, and S. Debroy, "Poster: Reliable 3d reconstruction for ad-hoc edge implementations," in *2024 IEEE/ACM Symposium on Edge Computing (SEC)*, 2024, pp. 539–541.
- [25] D. Dumitrescu and C.-A. Boiangiu, "A study of image upsampling and downsampling filters," *Computers*, vol. 8, no. 2, 2019. [Online]. Available: <https://www.mdpi.com/2073-431X/8/2/30>