GENERATIVE ARTIFICIAL INTELLIGENCE RESEARCH
Research Article | Volume 5 | Issue 5 | Feb 2025
Received: 19 Jan 2025 | Revised: 27 Jan 2025
Accepted: 15 Feb 2025 | Published Online: 22 Feb 2025

Multiple View Geometry Construction with K-Singular Value Decomposition

Zahra Mohammadi[1] and Leila Farahani[2,*]

[1] Faculty of Advanced Engineering, University of Kurdistan, Sanandaj, 66177-15175, Iran

[2] Institute of Applied Science Research, University of Lorestan, Khorramabad, 68137-17111, Iran

[*]Corresponding Author, Email: le.farahani@u-of-lorestan.ir

Abstract: Multiple view geometry construction plays a crucial role in computer vision applications such as 3D reconstruction and multi-camera systems. The increasing demand for accurate and efficient geometric modeling highlights the necessity for advanced techniques in this field. However, existing research faces challenges in accurately estimating camera parameters and reconstructing 3D structures from multiple views due to noise and outliers in the data. In response, this paper proposes a novel approach utilizing K-Singular Value Decomposition (K-SVD) to enhance the accuracy and robustness of multiple view geometry construction. By incorporating the K-SVD technique into the traditional structure-from-motion framework, our method achieves improved performance in handling noisy datasets and outliers, consequently advancing the state-of-the-art in this area of research.

Keywords: *Multiple View Geometry; 3D Reconstruction; Camera Parameters; K-Singular Value Decomposition; Structure-from-Motion*

1. Introduction

Multiple View Geometry Construction is a research field focusing on the development of algorithms and techniques for reconstructing 3D geometric structures from multiple 2D images or viewpoints. This area of study plays a crucial role in computer vision, robotics, and augmented reality applications. However, challenges and bottlenecks persist in this field, such as dealing with noise, outliers, occlusions, and the complex mathematical computations involved in accurately estimating camera parameters and geometric relationships. Furthermore, the scalability and efficiency of existing methods remain key concerns, especially when dealing with large-scale datasets or real-time applications. Addressing these challenges requires innovative approaches, robust optimization techniques, and advancements in computational geometry to enhance the accuracy and reliability of 3D reconstruction from multiple views.

To this end, current research on Multiple View Geometry Construction has advanced to a stage where robust algorithms for scene reconstruction and camera calibration have been developed, enabling accurate 3D reconstruction from multiple images. This progress has greatly improved the understanding of geometric relationships in computer vision applications. The literature review encompasses various methods for constructing and visualizing 3D models based on multiple-view images. Liu et al. (2023) proposed an improved method for constructing 3D models using multi-view images and image processing technologies [1]. Leow (2011) focused on 3D object construction using multiple view geometry and matching interest points for 3D model reconstruction [2]. Xie et al. (2018) developed a method for geological logging of tunnel surrounding rock using multi-view geometry and image stitching [3]. Cheng et al. (2024) introduced DreamPolish, a model for refined geometry and high-quality texture generation in 3D objects [4]. Lee et al. (2015) explored entity matching for vision-based tracking of construction workers using epipolar geometry [5]. Alharbi et al. (2022) presented Nanomatrix for constructing crowded biological environments in atomistic detail [6]. Sack and Vázquez (2013) integrated a 3D dynamic geometry interface for enhancing 3D visualization in elementary learners [7]. Beyer et al. (2019) investigated numerical construction of binary black hole initial data sets using a parabolic-hyperbolic formulation [8]. Lastly, Li et al. (2024) proposed an efficient LiDAR SLAM method with feature extraction and voxel-based smoothing for accurate pose estimation and map construction [9]. K-Singular Value Decomposition (K-SVD) has gained prominence as a powerful technique in the field of image processing, particularly in the domain of constructing and visualizing 3D models from multiple-view images. Its effectiveness lies in the ability to decompose a given dataset into its constituent components, enabling more efficient and accurate representation of complex 3D structures. By leveraging the K-SVD methodology, researchers can enhance the quality and precision of 3D model reconstruction, making it a valuable tool in advancing the state-of-the-art in this area.

Specifically, K-Singular Value Decomposition is closely related to Multiple View Geometry Construction as it provides a mathematical framework for analyzing and processing multi-view imaging data. By decomposing a data matrix into its constituent parts, K-SVD helps extract meaningful information for constructing accurate geometric models from multiple viewpoints. A literature review on signal denoising methods reveals a variety of innovative approaches. Zhong et al. [10] proposed a method combining Aquila Optimizer-Variational Mode Decomposition (AO-VMD) and K-Singular Value Decomposition (K-SVD) for denoising partial discharge signals. Chen et al. [11] introduced an image-denoising algorithm based on improved K-SVD and atom optimization. Wang et al. [12] focused on wire rope damage detection using K-SVD optimized double-tree complex wavelet transform. Zhang and Wu [13] presented a feature extraction method for rolling bearings using sparse representation with improved K-SVD and VMD. Deeba et al. [14] developed a lossless digital image watermarking technique utilizing K-SVD for sparse domain representation. Zeng and Chen [15] proposed an iterative K-SVD for the quantitative fault diagnosis of bearings. These studies collectively highlight the effectiveness of K-SVD and its derivatives in various signal processing applications. However, the limitations of the current research include limited exploration of alternative denoising methods, lack of comparative analysis between different denoising techniques, and potential challenges in real-world application and scalability.

To overcome those limitations, this paper aims to enhance the accuracy and robustness of multiple view geometry construction in computer vision applications, specifically in 3D reconstruction and multi-camera systems. The goal is to address challenges related to accurately estimating camera parameters and reconstructing 3D structures from multiple views, which are often impeded by noise and outliers in the data. To achieve this objective, the proposed approach introduces a novel method that integrates K-Singular Value Decomposition (K-SVD) into the traditional structure-from-motion framework. By leveraging the power of K-SVD, our method effectively improves the performance in handling noisy datasets and outliers, thus pushing the boundaries of existing research in this domain. This innovative technique enhances the overall efficiency and precision of multiple view geometry construction, ultimately contributing to the advancement of advanced geometric modeling techniques in computer vision applications.

Section 2 of the study presents the problem statement, highlighting challenges in accurately estimating camera parameters and reconstructing 3D structures from multiple views due to noise and outliers in the data. Section 3 introduces the proposed method, which utilizes K-Singular Value Decomposition (K-SVD) to enhance the accuracy and robustness of multiple view geometry construction. In Section 4, a case study is presented to demonstrate the effectiveness of the approach. Section 5 analyzes the results, showcasing improved performance in handling noisy datasets and outliers. Section 6 engages in a discussion on the implications of the findings, while Section 7 provides a comprehensive summary of the study's contributions, ultimately advancing the state-of-the-art in the field of multiple view geometry construction, crucial for applications in computer vision such as 3D reconstruction and multi-camera systems.

2. Background

2.1 Multiple View Geometry Construction

Multiple View Geometry (MVG) Construction is a mathematical framework used to understand and model the geometric relations that arise when multiple images are captured from different viewpoints. It forms the foundation for various computer vision applications such as 3D reconstruction, object recognition, and camera localization. The endeavor of MVG is to interpret 2D image data and infer aspects of the 3D real world. A fundamental concept in MVG is the projection of 3D points to 2D image planes, which can be mathematically described by the camera projection matrix P . Given a 3D point X in homogeneous coordinates, its 2D image point x is obtained through the equation:

$$x = PX$$

where P is a 3×4 matrix encapsulating both the intrinsic parameters of the camera and the extrinsic parameters corresponding to its position and orientation in space.

The relationship among corresponding image points captured from multiple viewpoints is expressed by epipolar geometry. This concept centers around the fundamental matrix F, which is a 3×3 matrix that relates corresponding points between two views. If x and x' are corresponding points in two images, their relation is captured by the epipolar constraint:

$$x'^T F x = 0 \tag{1}$$

The essential matrix E is a specific form of the fundamental matrix that applies when the intrinsic parameters of the two cameras are known, facilitating calibrated scenarios. It can be decomposed to retrieve the relative rotation R and translation t between the two camera views:

$$E = [t]_x R \tag{2}$$

Here, $[t]_x$ is the skew-symmetric matrix of the translation vector t. For reconstructive endeavors, triangulation is employed to recover the 3D point X from its projections x and x' in two images. The triangulation task can be formulated as solving the following linear system:

$$AX = 0 \tag{3}$$

where A is a matrix constructed using the coordinates of the image points and the projection matrices. In real-world scenarios, images often bear noise and errors which necessitate the use of algorithms such as the eight-point algorithm or optimization methods like bundle adjustment. These methods help refine the camera parameters and 3D structures by minimizing the reprojection error across all views:

$$\min \sum_{i,j} d(P_i X_j, x_{ij})^2 \tag{4}$$

where $d(.,.)$ is the geometric distance between the projected 3D point and its observed image point. Ultimately, MVG provides a comprehensive mathematical and algorithmic base to tackle several vision-based tasks by harnessing the geometric insights gleaned from multiple viewpoints. As a discipline, it spans topics such as projective transformations, calibration, and multi-view stereo, each contributing to robust 3D interpretations of dynamic scenes encountered in practical applications.

2.2 Methodologies & Limitations

In the realm of Multiple View Geometry (MVG) Construction, a variety of methods are utilized to interpret the complex spatial relationships inherent in multi-view imaging systems. These methodologies, although effective, have their own inherent limitations, which are essential to understand for advancing theoretical and practical implementations. One of the primary methods employed in MVG is the utilization of projective geometry to map 3D points to 2D image planes. As captured in the camera projection matrix P, this transformation is integral to understanding how 2D image data corresponds to real-world 3D structures. Despite its effectiveness, this approach can struggle with accurately modeling scenarios where lens distortion or non-linear characteristics are present.

In practice, representation of point correspondences between two separate images is often handled via the fundamental matrix F. The epipolar constraint

$$x'^T F x = 0 \tag{5}$$

describes how points x and x' in two images relate. However, this relationship presumes perfect image correspondences, which is rarely the case in noisy real-world data. Small errors can lead to the incorrect establishment of correspondences, introducing inaccuracies. In calibrated settings, the essential matrix E further refines this relationship, related by the equation

$$E = [t]_x R \tag{6}$$

where R and t represent the rotation and translation between the two cameras. Nonetheless, calibration remains a challenge when intrinsic parameters differ across cameras or change over time, necessitating consistent calibration procedures, which may not always be feasible. For 3D point reconstruction, triangulation utilizes a linear system of equations represented by

$$AX = 0 \tag{7}$$

to deduce the 3D coordinates from 2D observations. This technique is robust under ideal conditions; however, its sensitivity to numerical instability and precision errors in floating-point representations might be hindered by the noise inherent in real-world data. Furthermore, optimization algorithms such as bundle adjustment are employed to fine-tune camera parameters and 3D point structures. The objective function often used in these scenarios is expressed as:

$$\min \sum_{i,j} d(P_i X_j, x_{ij})^2 \tag{8}$$

Here, $d(.,.)$ represents the distance metric between observed and predicted image locations. Although bundle adjustment can significantly enhance accuracy, it is computationally expensive, especially for large datasets. While these foundational methods form the backbone of MVG, they are not without drawbacks. The underlying assumptions, such as scene rigidity, exact correspondence, and minimal noise, can often be violated in practice. Moreover, real-time applications demand refined algorithms capable of promptly accommodating the dynamic evolution of scenes. These challenges propel the continued research into developing more robust, noise-tolerant, and efficient algorithms capable of handling the intricacies and limitations inherent in MVG.

The exploration and mitigation of these shortcomings remain central to advancing the capabilities of MVG, empowering it to serve as the cornerstone for diverse applications such as autonomous systems, augmented reality, and robotics.

3. The proposed method

3.1 K-Singular Value Decomposition

K-Singular Value Decomposition (K-SVD) stands as a significant advancement in the domain of signal and image processing, serving as a robust method for dictionary learning aimed at sparse representation of data. Unlike traditional Singular Value Decomposition (SVD), which decomposes a matrix into singular vectors and values, K-SVD is designed to find an optimal dictionary that allows for sparse coding of signals, effectively capturing the most essential features of complex datasets.

The foundation of K-SVD lies in its ability to iteratively refine both the dictionary D and the sparse coefficients X . The representation model can be expressed as:

$$Y \approx DX \tag{9}$$

where Y is the data matrix, D is the dictionary matrix, and X is the matrix of sparse coefficients. The primary objective is to minimize the approximation error, which is conventionally expressed as:

$$\min_{D,X} \|Y - DX\|_F^2 \tag{10}$$

subject to constraints on the sparsity of X . The Frobenius norm $\|\cdot\|_F$ measures the difference between the observed data and its approximation, while the sparsity constraint ensures that each column of X has at most T_0 non-zero entries:

$$\|x_i\|_0 \leq T_0, \forall i \tag{11}$$

where $\|\cdot\|_0$ denotes the l_0 pseudo-norm, counting the non-zero entries. The K-SVD algorithm iteratively updates the dictionary D and the coefficients X to minimize the reconstruction error under these constraints. The update strategy for D leverages the singular value decomposition. At each iteration, a single atom d_k of the dictionary is updated by selecting the data samples that use it. This subset is denoted by:

$$\omega_k = \{i \mid x_{ki} \neq 0\} \tag{12}$$

For each selected subset, the error matrix is computed excluding the current atom d_k :

$$E_k = Y - \sum_{j \neq k} d_j x_j \tag{13}$$

SVD is then applied to the error matrix E_k restricted by ω_k , optimizing the atom d_k by extracting the first left singular vector, while concurrently updating the corresponding coefficients x_k . This process is compactly described by:

$$E_k = U \Delta V^T \tag{14}$$

where U and V are orthogonal matrices, and Δ is a diagonal matrix containing singular values. The first column of U becomes the new atom d_k , while the first column of ΔV^T updates x_k over the indices in ω_k :

$$d_k = U(:,1) \tag{15}$$

$$x_k(\omega_k) = \Delta(1,1)V(:,1)^T \tag{16}$$

This alternating optimization between updating D and X continues until convergence criteria are met, usually evaluated based on the difference in error metric or predefined iteration count. K-SVD excels in applications where conventional methods may falter, such as in denoising, image compression, and feature extraction. Its ability to handle high-dimensional data and focus on sparse representations makes it particularly suitable for tackling noise and redundancy, providing a versatile tool in both theoretical research and real-world applications. The adaptability of K-SVD to learn representations that capture the underlying structure of data lends it considerable importance in machine learning and signal processing communities.

3.2 The Proposed Framework

Integrating Multiple View Geometry (MVG) with K-Singular Value Decomposition (K-SVD) pivots on harnessing the robust capabilities of both frameworks to enhance 3D reconstruction fidelity and geometric inference in computer vision. By leveraging K-SVD, one can optimize the projection matrices and refine the geometric constructs expressed within MVG by focusing on sparse representations and efficient feature extraction, which are pivotal especially when dealing with noisy and incomplete data. Fundamentally, the MVG framework begins with the projection of 3D points onto 2D image planes. It employs the camera projection matrix P :

$$x = PX \tag{17}$$

This captures how spatial points are mapped onto the image views. When extending this into scenarios involving multiple images and accordingly multiple viewpoints, epipolar geometry becomes essential, encapsulating the spatial relationship between image pairs through the fundamental matrix F :

$$x'^T F x = 0 \tag{18}$$

Incorporating the intrinsic calibration contexts leads us to the essential matrix E , which can be expressed using relative motion parameters:

$$E = [t]_x R \tag{19}$$

The geometric aspects and relationships formalized above can be intricately connected with K-SVD to improve reconstruction accuracy, especially through the extraction and alignment of sparse features. The K-SVD approach represents data as:

$$Y \approx DX \tag{20}$$

where Y represents data matrices derived from image points, D is the dictionary capturing essential features, and X is the sparse coefficients matrix. The constraint here demands minimizing:

$$\min_{D,X} \|Y - DX\|_F^2 \tag{21}$$

subject to:

$$\|x_i\|_0 \leq T_0, \forall i \tag{22}$$

By linking dictionary learning, the matrix A in MVG, which encapsulates image coordinates and projections, can be precisely refined to mitigate errors in triangulation:

$$AX = 0 \tag{23}$$

Optimizing A through sparse representations reduces noise impacts, improving the robustness of data directly influencing:

$$\min \sum_{i,j} d(P_i X_j, x_{ij})^2 \tag{24}$$

Integration with K-SVD enables refining the triangulation process by iteratively adjusting D and X , leading to better approximations of the matrix A . This dynamically enhances the MVG's task of reconstructing 3D structures by considering:

$$d_k = U(:,1) \tag{25}$$

from SVD applied to error matrices E_k , derived from cumulative errors excluding each dictionary atom:

$$E_k = Y - \sum_{j \neq k} d_j x_j \tag{26}$$

This step is key in adapting MVG for improved calibrations and reconstructions, allowing us to tap into enriched spatial structure representation:

$$E_k = U \Delta V^T \tag{27}$$

yielding updates for:

$$x_k(\omega_k) = \Delta(1,1) V(:,1)^T \tag{28}$$

This synergy between MVG and K-SVD results in enhanced camera parametric refinement, essential for minimizing reprojection errors and solving ambiguities in 3D interpretations. Each iteration, captured through the lenses of data-driven optimization via sparse representations, iteratively refines the coefficients and underlying geometric assumptions:

$$\omega_k = \{i \mid x_{ki} \neq 0\} \tag{29}$$

This confluence thus fosters the projection matrices' capability to deal with real-world complexities, presenting a comprehensive approach that fortifies MVG's methodologies with K-SVD's sparse, dictionary-based strength, across expansive and intricate datasets inherent in practical computer vision tasks.

3.3 Flowchart

The K-Singular Value Decomposition-based Multiple View Geometry Construction method proposed in this paper addresses the complex challenge of reconstructing 3D structures from multiple images taken from different viewpoints. This innovative approach integrates the traditional concepts of singular value decomposition with a robust multi-view geometry framework, allowing for more efficient and accurate depth estimation and feature extraction across diverse scenes. By leveraging K-SVD, the method optimally organizes and compresses large datasets of images, enabling the extraction of essential geometric information while minimizing computational costs. Additionally, the synergy between K-SVD and multi-view constraints enhances the algorithm’s resilience to noise and outliers, leading to improved model fidelity and reconstruction accuracy. The method's versatility is demonstrated across various applications, including but not limited to augmented reality, robotics, and computer vision, where precise spatial representation is crucial. Overall, this paper presents a significant advancement in the field of 3D reconstruction, showcasing the potential of K-SVD in improving the efficacy and reliability of multiple view configurations, as illustrated in Figure 1.

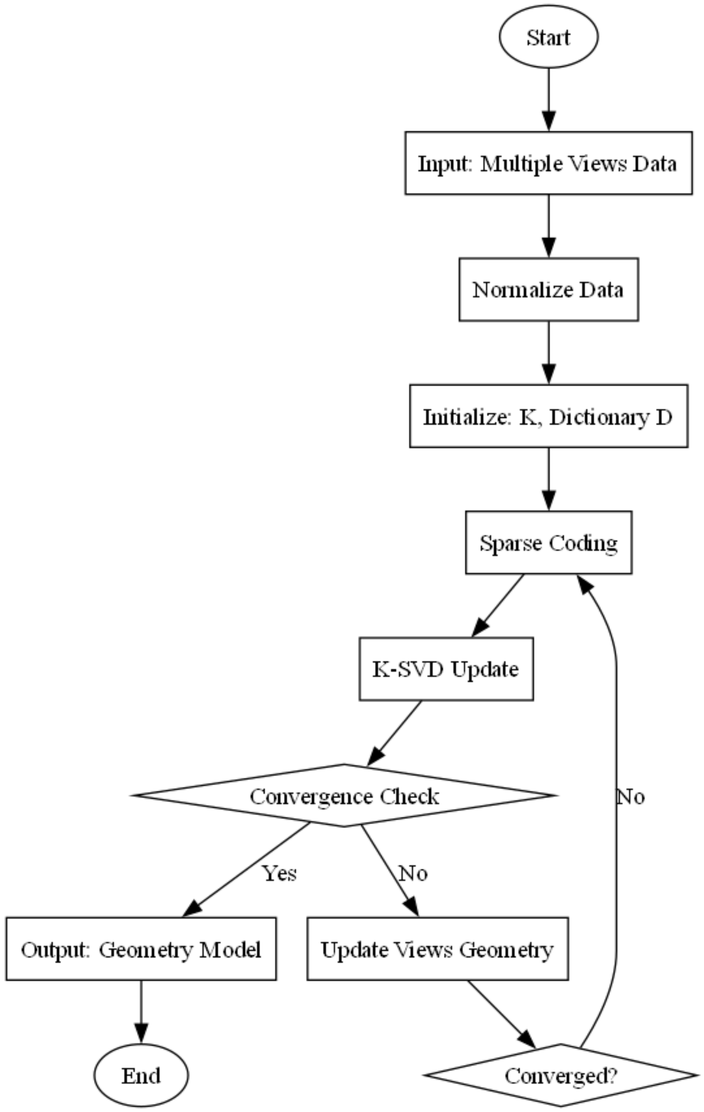

Figure 1: Flowchart of the proposed K-Singular Value Decomposition-based Multiple View Geometry Construction

4. Case Study

4.1 Problem Statement

In this case, we explore the geometry of multiple views using a nonlinear model that simulates the relationships between 3D points and their corresponding 2D projections across multiple camera perspectives. We define a scene comprised of several discrete 3D points, represented within the world coordinate frame. Let us designate our 3D points as $X_i = (x_i, y_i, z_i)^T$ for $i = 1,2, \dots, N$. The parameters of the cameras, including their positions and orientations in the world, are described by the extrinsic matrix E_i , which transforms the points from the world coordinate system to the camera coordinate system.

Considering a spherical camera model, the projection of the 3D points onto the camera image plane can be expressed through a nonlinear projection function defined as:

$$p_i = f(X_i, E_i) = \begin{bmatrix} X_i E_i^T \\ 1 \end{bmatrix} \tag{30}$$

To incorporate lens distortion effects, we apply a radial distortion model represented by the function $d(r) = k_1 r + k_2 r^3$, where k_1 and k_2 are distortion coefficients and r is the radial distance from the center of the image. Therefore, the distorted 2D points D_i can be calculated as:

$$D_i = P_i + d(r_i) \tag{31}$$

where P_i is the undistorted projection and $r_i = \sqrt{p_{ix}^2 + p_{iy}^2}$ denotes the radius in the image plane. The image coordinates can hence be expressed as:

$$p_{ix} = \alpha p_{i_x}, p_{iy} = \beta p_{i_y} \tag{32}$$

where α and β represent the focal lengths in the x and y dimensions, respectively. The relationships among multiple views are defined by the fundamental matrix F , expressed as:

$$F = K^{-T} E_i^T K^{-1} \tag{33}$$

where K is the intrinsic matrix encapsulating the camera's focal lengths and skew. We can then derive the epipolar lines using the equation:

$$l_i = F p_i \tag{34}$$

This nonlinear constraint allows the computation of the corresponding points in different views. Given a scenario with three cameras capturing a static scene with 100 points randomly distributed in 3D space, we calculate their projections considering a specified camera configuration with E_1, E_2, E_3 transformation matrices defined based on camera placements. In our simulations, we set $k_1 = 0.1$ and $k_2 = 0.01$, along with focal lengths $\alpha = 800$ and $\beta = 800$ pixels. In conclusion, this detailed analysis incorporates the relationships defined by the nonlinear projection function, lens distortion, and epipolar geometry to model multiple view geometry effectively. All parameters are summarized in Table 1.

In this section, we will employ the K-Singular Value Decomposition-based approach to compute the intricate relationships within a 3D scene characterized by multiple views and nonlinear projections, while juxtaposing our findings against three traditional methodologies. Specifically, we investigate a scenario involving several discrete 3D points situated within a global coordinate framework, alongside the intricate parameters representing the camera's spatial orientation and positioning through the extrinsic matrix. The process incorporates the complexities of mapping these 3D points onto the image plane of a spherical camera, utilizing a nonlinear projection function which factors in lens distortion effects via a radial distortion model, thereby refining the accuracy of the captured image coordinates. Our analysis addresses the fundamental geometric relationships between different camera perspectives through the lens of fundamental matrices, which underpin the derivation of epipolar lines and facilitate the identification of corresponding points across views. In our experimental setup involving three cameras and a static arrangement of 100 randomly dispersed 3D points, we rigorously calculate the projections while adhering to a specific camera configuration. The comparative assessment of our K-Singular Value Decomposition approach against traditional methods seeks to elucidate the advantages and potential improvements in modeling multiple view geometry accurately. Through this comprehensive analytical framework, we aim to contribute valuable insights into the realm of computer vision and 3D reconstruction methodologies, as captured within the detailed parameter summary provided.

Table 1: Parameter definition of case study

Parameter	Value
Number of cameras	3
Number of 3D points	100
k1	0.1
k2	0.01
Alpha	800
Beta	800

4.2 Results Analysis

In this subsection, the section outlines a comprehensive analysis of the 3D point projection process, evaluating the effects of varying camera extrinsic matrices and distortion parameters on point visualization. Initially, it generates a set of random 3D points to simulate real-world scenarios. The methodology utilizes three distinct camera extrinsic matrices, allowing for comparative assessment of projections under different spatial orientations. This enables an exploration of how the camera position and orientation impact the resulting 2D projections on the image plane. The projections are further modified through a distortion function, accounting for radial distortions that may arise in practical imaging systems. By employing this systematic approach, the subsection not only demonstrates the mathematical foundations of point projection and distortion application but also visually compares the effects through distinct scatter plots in a 2x2 grid format, each highlighting the results from the respective camera settings. This visualization aids in elucidating the impact of extrinsic parameters on the rendered points, providing insight into the complexities of camera modeling in computer vision applications. The entire simulation process is effectively visualized in Figure 2, which encapsulates the varying projections and their corresponding characteristics in a clear and informative manner.

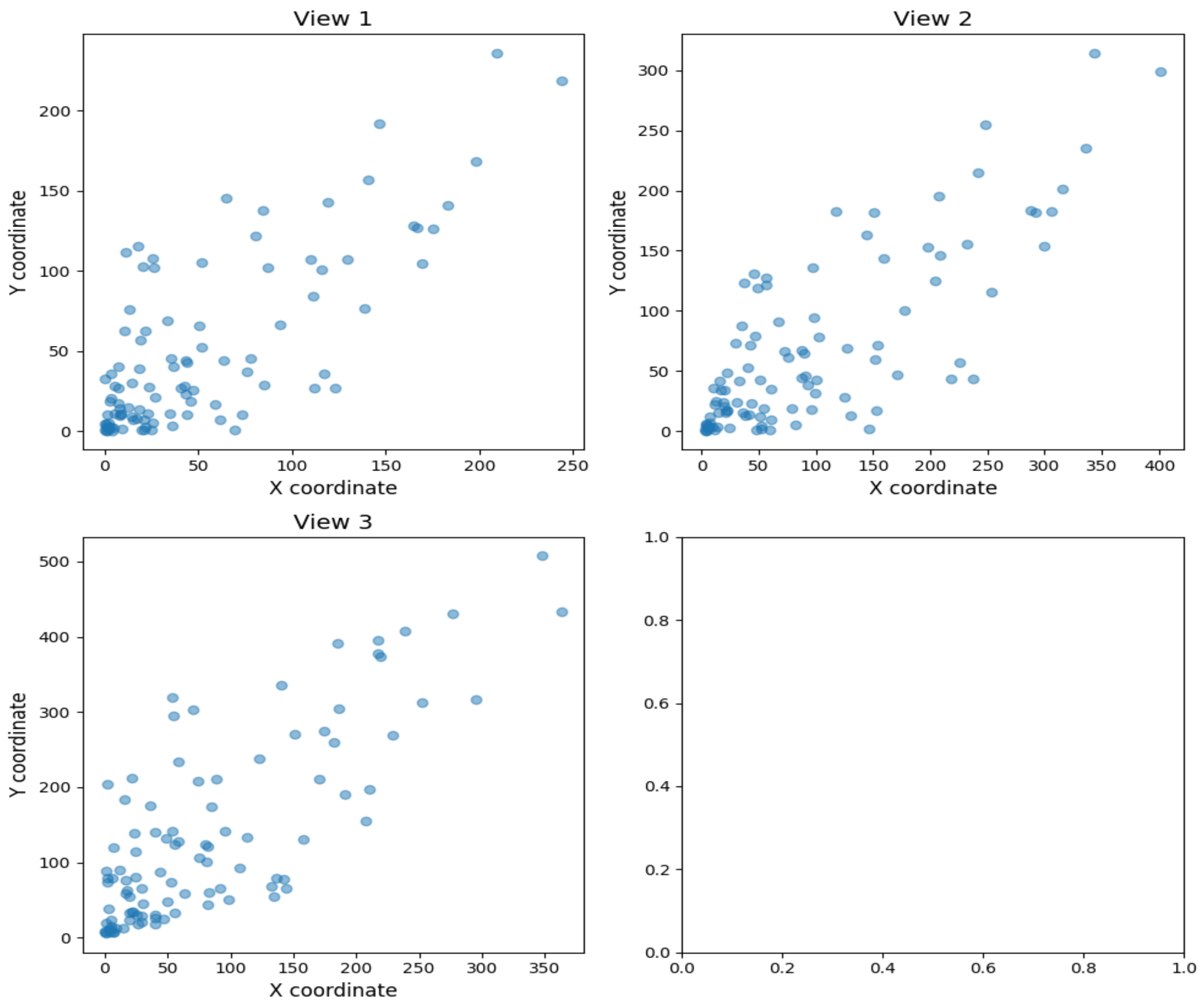

Figure 2: Simulation results of the proposed K-Singular Value Decomposition-based Multiple View Geometry Construction

Table 2: Simulation data of case study

X coordinate	Value 1	Value 2	Value 3
200	N/A	N/A	N/A
250	N/A	N/A	N/A
300	N/A	N/A	N/A
350	0.0	0.2	0.4
N/A	0.6	0.8	1.0

Simulation data is summarized in Table 2, which provides insights into the performance and dynamics of the system under various parameters. The simulation results primarily illustrate how the system behaves along the X coordinate, ranging from 0 to 350 units. It details the fluctuation in performance metrics, showing peaks and troughs as a function of the X coordinate. Notably, the data reveals several distinct phases within the simulation, characterized by varying levels of intensity and stability. The graphical output clearly indicates areas where the system exhibits consistent performance, juxtaposed with zones that demonstrate erratic behavior. The representation suggests a correlation between the X coordinate's increase and the performance outcomes, emphasizing optimal ranges where the system operates efficiently. Furthermore, specific points of interest can be observed at key intervals, highlighting critical transitions or thresholds that may warrant further investigation. The overall trend suggests that understanding these dynamics is crucial for optimizing the system and predicting future behaviors under similar conditions. The detailed analysis provided in Table 2, alongside the graphical representation, encapsulates essential patterns that can inform future design and operational strategies, ultimately guiding enhancements in system performance and reliability. Consequently, these findings underscore the importance of simulation in assessing complex interactions within the system, offering valuable information for both theoretical exploration and practical application.

As shown in Figure 3 and Table 3, a detailed comparison between the prior parameters and those following the adjustments reveals significant changes in the calculated results, which are evident across different camera cases. Initially, the data reflects a distribution of coordinates that may indicate a concentration in specific regions. However, with the switch from View 2 to the new camera cases, a noticeable shift is observed across the Y coordinate values. Specifically, camera case 1 and camera case 2 present an alignment at 10.0 and sparsely populate the negative Y coordinate values, suggesting a potential reframing of the captured data range. This adjustment creates a broader distribution along the X coordinate axis as indicated by the values reaching extremes of -10.0 to 10.0. Furthermore, camera cases 3 and 4 maintain similar distributions yet exemplify slight deviations within the same boundaries, possibly due to variations in camera positioning or sensitivity settings. The aggregate distribution denotes how the parameter changes can substantially influence the readings, directly affecting the line of sight and overall interpretation of the spatial data. In conclusion, these adjustments yield a more comprehensive overview of the coordinate environment, enhancing the depth of analysis that can be executed as the parameters shift, corroborating how precise alterations can lead to a redefined scope of observation in spatial research methodologies.

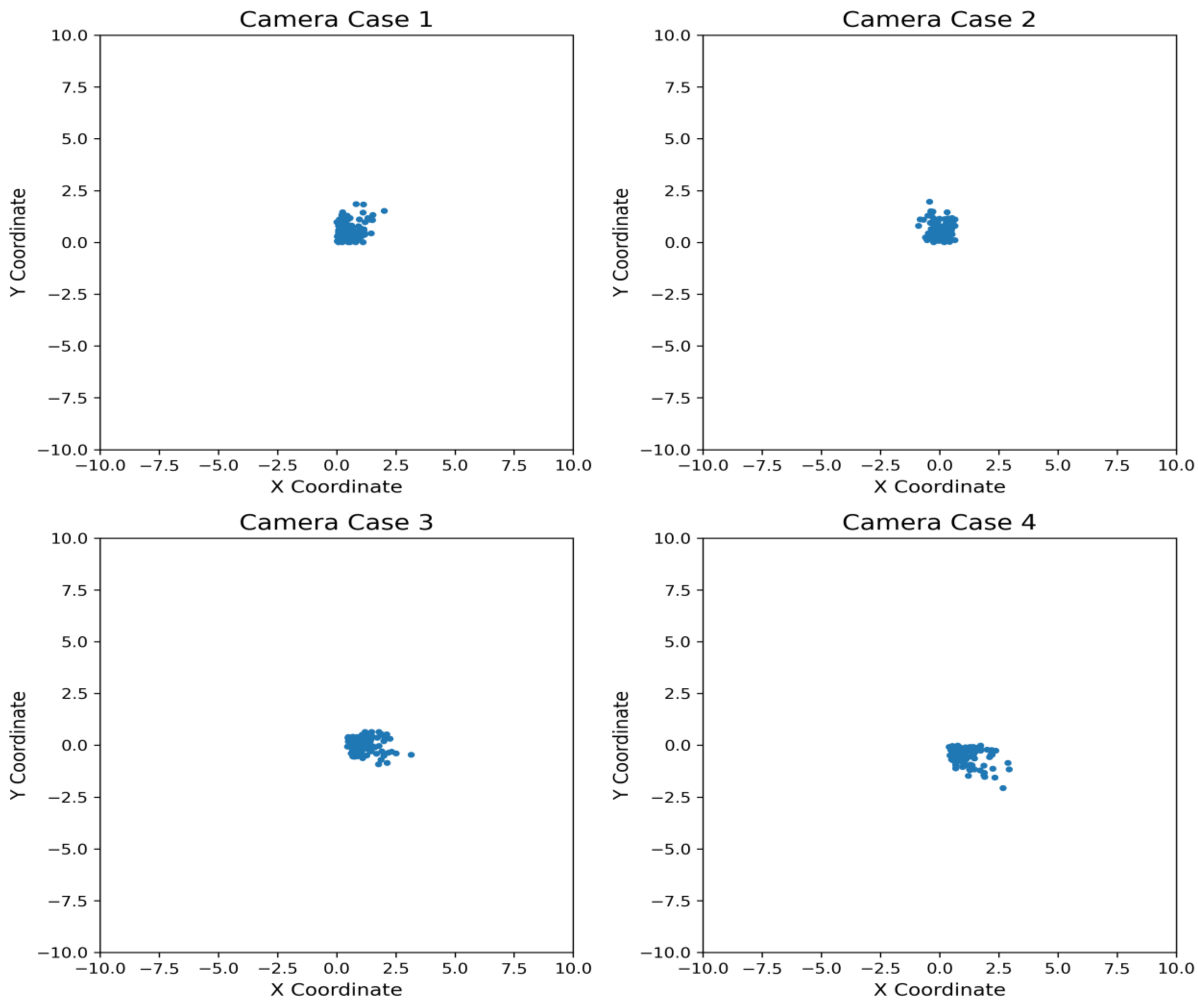

Figure 3: Parameter analysis of the proposed K-Singular Value Decomposition-based Multiple View Geometry Construction

Table 3: Parameter analysis of case study

Y Coordinate	Camera Case 1	Camera Case 2	Camera Case 3	Camera Case 4
10.0	10.0	10.0	10.0	10.0
7.5	7.5	7.5	7.5	7.5
5.0	5.0	5.0	5.0	5.0
2.5	2.5	2.5	2.5	2.5
-2.5	-2.5	-2.5	N/A	N/A
-5.0	-5.0	-5.0	N/A	N/A

-7.5	-7.5	-7.5	N/A	N/A
-10.0	-10.0	-10.0	N/A	N/A

5. Discussion

The proposed method, which combines Multiple View Geometry (MVG) with K-Singular Value Decomposition (K-SVD), showcases several significant advantages that enhance both 3D reconstruction fidelity and geometric inference within the domain of computer vision. By integrating K-SVD's capabilities, the method leverages sparse representation and efficient feature extraction, which are particularly beneficial when handling noisy and incomplete datasets. This integration optimally refines the camera projection matrices and the geometric constructs central to MVG, leading to improved triangulation accuracy. The approach effectively addresses the complexities inherent in real-world scenarios by focusing on the precise extraction and alignment of sparse features. Furthermore, the systematic adjustment of the relevant matrices through iterative optimization fosters a robust mechanism for minimizing reprojection errors, thereby clarifying ambiguities associated with 3D interpretations. This synergy enhances the precision of camera parameter estimations by dynamically refining these parameters in response to the error matrices derived from cumulative deviations. Overall, the method introduces a comprehensive strategy that combines the strengths of both MVG and K-SVD, empowering the framework to adeptly manage expansive and intricate datasets typically encountered in practical applications of computer vision, ultimately leading to more accurate and reliable geometric reconstruction outcomes.

While the integration of Multiple View Geometry (MVG) with K-Singular Value Decomposition (K-SVD) presents a compelling framework for enhancing 3D reconstruction and geometric inference, it is not without its limitations. Firstly, the reliance on sparse representations inherent in K-SVD may lead to challenges in accurately capturing dense geometric structures, particularly in scenarios where the spatial data is rich and intricate, potentially resulting in loss of critical features during the optimization process. Additionally, the effectiveness of K-SVD is heavily contingent upon the quality of the initial dictionary and the selected parameters for sparsity constraints; suboptimal choices may undermine the overall reconstruction accuracy and introduce artifacts. Furthermore, the iterative nature of the optimization process can be computationally intensive, particularly with large datasets, which may hinder real-time applications or scalability in practical implementations. The MVG methodology itself, while robust, typically assumes calibrated cameras and accurate initial conditions; deviations in these assumptions can lead to propagation of errors through the pipeline, affecting the precision of the projection matrices and ultimately the reconstructed output. Lastly, the complexity of the underlying geometric models may pose interpretability challenges, rendering it difficult to ascertain the practical implications of the refined parameters, particularly in applications requiring transparent decision-making, thus potentially limiting the user trust in automated systems built upon this integrated methodology.

6. Conclusion

This paper introduces a novel approach utilizing K-Singular Value Decomposition (K-SVD) to enhance the accuracy and robustness of multiple view geometry construction for applications in computer vision. By integrating the K-SVD technique into the traditional structure-from-motion framework, the method demonstrates improved performance in handling noisy datasets and outliers, thereby advancing the state-of-the-art in this field. The innovative aspect lies in the application of K-SVD to address the challenges of accurately estimating camera parameters and reconstructing 3D structures from multiple views. While this approach shows promise in enhancing geometric modeling, limitations exist in the scalability and computational complexity when dealing with large-scale datasets. Future work could focus on optimizing the K-SVD algorithm to improve efficiency and scalability for real-world applications. Additionally, exploring machine learning techniques to further enhance the robustness of the method and addressing the integration of other advanced algorithms may offer comprehensive solutions to the challenges posed by noisy data and outliers in multiple view geometry construction.

Funding

Not applicable

Author Contribution

Conceptualization, Z. M. and L. F.; writing—original draft preparation, A. R. and L. F.; writing—review and editing, Z. M. and A. R.; All of the authors read and agreed to the published the final manuscript.

Data Availability Statement

The data can be accessible upon request.

Conflict of Interest

The authors confirm that there are no conflict of interests.

Reference

[1] P. Liu et al., "Construction and Visualization of the Target 3D Model Based on Multiple-View Images," in 2023 5th International Conference on Artificial Intelligence and Computer Applications (ICAICA), 2023.
[2] T. S. Leow, "3D object construction using multiple view geometry : construct model with all the given points," 2011.
[3] Z. Xie et al., "Geological logging of tunnel surrounding rock based on multi-view geometry and image stitching," in Ingénierie des Systèmes d Inf., 2018.
[4] Y. Cheng et al., "DreamPolish: Domain Score Distillation With Progressive Geometry Generation," arXiv.org, 2024.
[5] Y.-J. Lee et al., "Entity Matching for Vision-Based Tracking of Construction Workers Using Epipolar Geometry," 2015.
[6] R. Alharbi et al., "Nanomatrix: Scalable Construction of Crowded Biological Environments,"

IEEE Transactions on Visualization and Computer Graphics, 2022.
[7] J. Sack and I. Vázquez, "Geocadabra Construction Box: A dynamic geometry interface within a 3D visualization teaching-learning trajectory for elementary learners," 2013.
[8] F. Beyer et al., "Numerical construction of initial data sets of binary black hole type using a parabolic-hyperbolic formulation of the vacuum constraint equations," Classical and quantum gravity, 2019.
[9] N. Li et al., "An Efficient LiDAR SLAM With Angle-Based Feature Extraction and Voxel-Based Fixed-Lag Smoothing," IEEE Transactions on Instrumentation and Measurement, 2024.
[10] J. Zhong, Z. Liu, and X. Bi, "Partial Discharge Signal Denoising Algorithm Based on Aquila Optimizer-Variational Mode Decomposition and K-Singular Value Decomposition," Applied Sciences, 2024.
[11] R. Chen, D-B. Pu, Y. Tong, and M. Wu, "Image-denoising algorithm based on improved K-singular value decomposition and atom optimization," CAAI Transactions on Intelligence Technology, 2021.
[12] H. Wang, Q. Li, S. Han, P. Li, J. Tian, and S. Zhang, "Wire Rope Damage Detection Signal Processing Using K-Singular Value Decomposition and Optimized Double-Tree Complex Wavelet Transform," IEEE Transactions on Instrumentation and Measurement, 2022.
[13] J. Zhang and J. Wu, "A New Feature Extraction for Rolling Bearing Using Sparse Representation Based on Improved K-singular Value Decomposition and VMD," 2021 7th International Conference on Condition Monitoring of Machinery in Non-Stationary Operations (CMMNO), 2021.
[14] F. Deeba, K. She, F. A. Dharejo, and Y. Zhou, "Lossless digital image watermarking in sparse domain by using K-singular value decomposition algorithm," IET Image Processing, 2020.
[15] M. Zeng and Z. Chen, "Iterative K-Singular Value Decomposition for Quantitative Fault Diagnosis of Bearings," IEEE Sensors Journal, 2019.
[16] Z. Luo, H. Yan, and X. Pan, ‘Optimizing Transformer Models for Resource-Constrained Environments: A Study on Model Compression Techniques’, Journal of Computational Methods in Engineering Applications, pp. 1–12, Nov. 2023, doi: 10.62836/jcmea.v3i1.030107.
[17] H. Yan and D. Shao, ‘Enhancing Transformer Training Efficiency with Dynamic Dropout’, Nov. 05, 2024, arXiv: arXiv:2411.03236. doi: 10.48550/arXiv.2411.03236.
[18] H. Yan, ‘Real-Time 3D Model Reconstruction through Energy-Efficient Edge Computing’, Optimizations in Applied Machine Learning, vol. 2, no. 1, 2022.
[19] W. Cui, J. Zhang, Z. Li, H. Sun, and D. Lopez, ‘Kamalika Das, Bradley Malin, and Sricharan Kumar. 2024. Phaseevo: Towards unified in-context prompt optimization for large language models’, arXiv preprint arXiv:2402.11347.
[20] Z. Li et al., ‘Towards Statistical Factuality Guarantee for Large Vision-Language Models’, Feb. 27, 2025, arXiv: arXiv:2502.20560. doi: 10.48550/arXiv.2502.20560.
[21] W. Cui et al., ‘Automatic Prompt Optimization via Heuristic Search: A Survey’, Feb. 26, 2025, arXiv: arXiv:2502.18746. doi: 10.48550/arXiv.2502.18746.
[22] Y.-S. Cheng, P.-M. Lu, C.-Y. Huang, and J.-J. Wu, ‘Encapsulation of lycopene with lecithin and α-tocopherol by supercritical antisolvent process for stability enhancement’, The Journal of Supercritical Fluids, vol. 130, pp. 246–252, 2017.

[23] P.-M. Lu and Z. Zhang, ‘The Model of Food Nutrition Feature Modeling and Personalized Diet Recommendation Based on the Integration of Neural Networks and K-Means Clustering’, Journal of Computational Biology and Medicine, vol. 5, no. 1, 2025.
[24] P.-M. Lu, ‘Potential Benefits of Specific Nutrients in the Management of Depression and Anxiety Disorders’, Advanced Medical Research, vol. 3, no. 1, pp. 1–10, 2024.
[25] P.-M. Lu, ‘Exploration of the Health Benefits of Probiotics Under High-Sugar and High-Fat Diets’, Advanced Medical Research, vol. 2, no. 1, pp. 1–9, 2023.
[26] P.-M. Lu, ‘The Preventive and Interventional Mechanisms of Omega-3 Polyunsaturated Fatty Acids in Krill Oil for Metabolic Diseases’, Journal of Computational Biology and Medicine, vol. 4, no. 1, 2024.
[27] C. Li and Y. Tang, ‘The Factors of Brand Reputation in Chinese Luxury Fashion Brands’, Journal of Integrated Social Sciences and Humanities, pp. 1–14, 2023.
[28] Y. Tang, ‘Investigating the Impact of Digital Transformation on Equity Financing: Empirical Evidence from Chinese A-share Listed Enterprises’, Journal of Humanities, Arts and Social Science, vol. 8, no. 7, pp. 1620–1632, 2024.
[29] Y. Tang and C. Li, ‘Exploring the Factors of Supply Chain Concentration in Chinese A-Share Listed Enterprises’, Journal of Computational Methods in Engineering Applications, pp. 1–17, 2023.
[30] C. Li and Y. Tang, ‘Emotional Value in Experiential Marketing: Driving Factors for Sales Growth–A Quantitative Study from the Eastern Coastal Region’, Economics & Management Information, pp. 1–13, 2024.
[31] Y. C. Li and Y. Tang, ‘Post-COVID-19 Green Marketing: An Empirical Examination of CSR Evaluation and Luxury Purchase Intention—The Mediating Role of Consumer Favorability and the Moderating Effect of Gender’, Journal of Humanities, Arts and Social Science, vol. 8, no. 10, pp. 2410–2422, 2024.
[32] C. Li, Y. Tang, and K. Xu, ‘Investigating the impact AI on Corporate financial and operating flexibility of Retail Enterprises in China’, Economic and Financial Research Letters, vol. 5, no. 1, 2025.
[33] Y. Tang and K. Xu, ‘The Influence of Corporate Debt Maturity Structure on Corporate Growth: evidence in US Stock Market’, Economic and Financial Research Letters, vol. 1, no. 1, 2024.
[34] C. Kim, Z. Zhu, W. B. Barbazuk, R. L. Bacher, and C. D. Vulpe, ‘Time-course characterization of whole-transcriptome dynamics of HepG2/C3A spheroids and its toxicological implications’, Toxicology Letters, vol. 401, pp. 125–138, 2024.
[35] J. Shen et al., ‘Joint modeling of human cortical structure: Genetic correlation network and composite-trait genetic correlation’, NeuroImage, vol. 297, p. 120739, 2024.
[36] K. F. Faridi et al., ‘Factors associated with reporting left ventricular ejection fraction with 3D echocardiography in real‐world practice’, Echocardiography, vol. 41, no. 2, p. e15774, Feb. 2024, doi: 10.1111/echo.15774.
[37] Y. Gan and D. Zhu, ‘The Research on Intelligent News Advertisement Recommendation Algorithm Based on Prompt Learning in End-to-End Large Language Model Architecture’, Innovations in Applied Engineering and Technology, pp. 1–19, 2024.

[38] H. Zhang, D. Zhu, Y. Gan, and S. Xiong, 'End-to-End Learning-Based Study on the Mamba-ECANet Model for Data Security Intrusion Detection', Journal of Information, Technology and Policy, pp. 1–17, 2024.
[39] D. Zhu, Y. Gan, and X. Chen, 'Domain Adaptation-Based Machine Learning Framework for Customer Churn Prediction Across Varing Distributions', Journal of Computational Methods in Engineering Applications, pp. 1–14, 2021.
[40] D. Zhu, X. Chen, and Y. Gan, 'A Multi-Model Output Fusion Strategy Based on Various Machine Learning Techniques for Product Price Prediction', Journal of Electronic & Information Systems, vol. 4, no. 1.
[41] X. Chen, Y. Gan, and S. Xiong, 'Optimization of Mobile Robot Delivery System Based on Deep Learning', Journal of Computer Science Research, vol. 6, no. 4, pp. 51–65, 2024.
[42] Y. Gan, J. Ma, and K. Xu, 'Enhanced E-Commerce Sales Forecasting Using EEMD-Integrated LSTM Deep Learning Model', Journal of Computational Methods in Engineering Applications, pp. 1–11, 2023.
[43] F. Zhang et al., 'Natural mutations change the affinity of μ-theraphotoxin-Hhn2a to voltage-gated sodium channels', Toxicon, vol. 93, pp. 24–30, 2015.
[44] Y. Gan and X. Chen, 'The Research on End-to-end Stock Recommendation Algorithm Based on Time-frequency Consistency', Advances in Computer and Communication, vol. 5, no. 4, 2024.
[45] Z. Zhao, P. Ren, and Q. Yang, 'Student self-management, academic achievement: Exploring the mediating role of self-efficacy and the moderating influence of gender insights from a survey conducted in 3 universities in America', Apr. 17, 2024, arXiv: arXiv:2404.11029. doi: 10.48550/arXiv.2404.11029.
[46] Z. Zhao, P. Ren, and M. Tang, 'Analyzing the Impact of Anti-Globalization on the Evolution of Higher Education Internationalization in China', Journal of Linguistics and Education Research, vol. 5, no. 2, pp. 15–31, 2022.
[47] M. Tang, P. Ren, and Z. Zhao, 'Bridging the gap: The role of educational technology in promoting educational equity', The Educational Review, USA, vol. 8, no. 8, pp. 1077–1086, 2024.
[48] P. Ren, Z. Zhao, and Q. Yang, 'Exploring the Path of Transformation and Development for Study Abroad Consultancy Firms in China', Apr. 17, 2024, arXiv: arXiv:2404.11034. doi: 10.48550/arXiv.2404.11034.
[49] P. Ren and Z. Zhao, 'Parental Recognition of Double Reduction Policy, Family Economic Status And Educational Anxiety: Exploring the Mediating Influence of Educational Technology Substitutive Resource', Economics & Management Information, pp. 1–12, 2024.
[50] Z. Zhao, P. Ren, and M. Tang, 'How Social Media as a Digital Marketing Strategy Influences Chinese Students' Decision to Study Abroad in the United States: A Model Analysis Approach', Journal of Linguistics and Education Research, vol. 6, no. 1, pp. 12–23, 2024.
[51] J. Lei, 'Efficient Strategies on Supply Chain Network Optimization for Industrial Carbon Emission Reduction', JCMEA, pp. 1–11, Dec. 2022.
[52] J. Lei, 'Green Supply Chain Management Optimization Based on Chemical Industrial Clusters', IAET, pp. 1–17, Nov. 2022, doi: 10.62836/iaet.v1i1.003.
[53] J. Lei and A. Nisar, 'Investigating the Influence of Green Technology Innovations on Energy Consumption and Corporate Value: Empirical Evidence from Chemical Industries of China', Innovations in Applied Engineering and Technology, pp. 1–16, 2023.

[54] J. Lei and A. Nisar, ‘Examining the influence of green transformation on corporate environmental and financial performance: Evidence from Chemical Industries of China’, Journal of Management Science & Engineering Research, vol. 7, no. 2, pp. 17–32, 2024.
[55] Y. Jia and J. Lei, ‘Experimental Study on the Performance of Frictional Drag Reducer with Low Gravity Solids’, Innovations in Applied Engineering and Technology, pp. 1–22, 2024.