Dynamic Shadow Removal and Quality Assessment of High-Resolution Orthophotos for Pavement Inspection

Yueh-Che Li [1], Jyun-Ping Jhan [2]

[1] Graduate Institute of A.I. Cross-disciplinary Technology, NTUST, Taiwan – M11452017@mail.ntust.edu.tw

[2] Graduate Institute of A.I. Cross-disciplinary Technology, NTUST, Taiwan – jpjhan@mail.ntust.edu.tw

Keywords: Shadow Removal, DSM and orthophoto, Pavement Inspection, Mobile Mapping, Quality Assessment

ABSTRACT

Traditional pavement inspection and data collection are often constrained by traffic conditions, operational safety, and equipment costs, making it difficult to achieve both efficiency and large-scale coverage. To address these limitations, this study employs a Pavement Roughness Index and Distress Extraction System (PRIDEs), which integrates high-resolution industrial cameras, high-precision global navigation satellite system (GNSS), wheel pulse sensors, and an onboard computer to acquire high-quality images under high-speed driving conditions. Using photogrammetry and computer vision techniques, camera poses are reconstructed to generate dense point clouds, digital surface models (DSMs), and orthophotos for detailed pavement distress analysis. However, the acquired imagery is affected by dynamic shadows and lens-focusing induced blur, resulting in ghosting artifacts and inconsistent orthophoto quality. To mitigate these issues, this study proposes a masking strategy during orthophoto generation, where U-Net is employed to detect shadow regions and Laplacian variance is used to identify blurred areas. By integrating these masks, more uniform and higher-quality orthophotos can be produced. Experimental results demonstrate that the proposed approach effectively reduces false positives and false negatives of crack detection caused by shadows and blur, thereby improving the reliability of orthophotos for automated pavement condition assessment.

1. Introduction

Pavement distress detection is closely associated with vehicle ride comfort and safety. Typical distress types include potholes, depressions, and cracks, and require the measurement of continuous road elevation profiles for ride quality assessment. To improve the efficiency and safety of pavement data acquisition, this study develops a system termed the Pavement Roughness Index and Distress Extraction System (PRIDEs). The system enables reliable acquisition of high-resolution pavement imagery under high-speed operating conditions and supports subsequent digital surface model (DSM) and orthophoto generation, spatial localization of pavement defects, and International Roughness Index (IRI) estimation. Compared with conventional approaches that rely on expensive LiDAR systems, PRIDEs provides a more cost-effective and flexible solution for large-scale pavement inspection.

PRIDEs employs photogrammetric techniques to generate DSMs and orthophotos of pavement surfaces, enabling subsequent pavement deterioration detection. However, under strong illumination conditions, the tail fin and supporting structures of the fixed camera module cast dynamic shadows onto the pavement imagery, leading to ghosting artifacts in the resulting orthophotos. In addition, due to lens focusing characteristics, image quality gradually degrades from the center to ward the edges, resulting in inconsistencies along mosaicing seams. Therefore, this study focuses on addressing dynamic shadowing and lens-induced blurring effects in pavement imagery and develops a workflow to improve orthophoto quality and enhance the reliability of subsequent pavement defect detection.

1.1 Related Work

Pavement distress detection has gradually shifted from traditional manual measurements to automated approaches. (Chang et al., 2026) developed an image-based pavement roughness detection system and demonstrated its capability to achieve an elevation accuracy of 1 mm, as well as its feasibility in estimating the IRI, indicating that image-based systems can serve as effective tools for pavement condition assessment. Building upon this foundation, PRIDEs further improves data acquisition efficiency and system integration by incorporating high-resolution imaging, high-precision positioning, and distance-triggering mechanisms, enabling reliable pavement monitoring under high-speed operating conditions.

However, image-based pavement distress inspection methods are prone to interference from shadows caused by changes in lighting conditions. Shadows not only alter the brightness distribution of images but may also obscure texture information on the pavement , thereby affecting subsequent defect identification and image analysis results. To address the shadow detection issue, (Zhu et al., 2018) proposed the Bidirectional Feature Pyramid Network with Recurrent Attention Residual Modules (BDRAR). By combining a bidirectional feature pyramid architecture with recurrent attention residual modules, this approach effectively integrates multi-scale features and shadow boundary information to improve the accuracy of shadow region identification. (Hu et al., 2020) introduced Direction-aware Spatial Context (DSC) features, which further account for spatial context information from different directions to improve shadow detection and removal performance in complex scenes. The aforementioned studies demonstrate that deep learning methods combining multi-scale feature learning, attention mechanisms, and spatial context modeling have become a key research direction in shadow segmentation and removal.

1.2 Objective

Imagery has been demonstrated to be effective for high-precision pavement distress inspection, and shadows can also be detected using computer vision and deep learning techniques. However, current research remains insufficient in addressing challenges such as dynamic shadows caused by vehicle-mounted structures, as well as orthophoto ghosting and quality inconsistencies induced by blurred regions near image boundaries.

To address these issues, this study employs the PRIDEs system to acquire pavement imagery and integrates image processing with deep learning techniques to mitigate the effects of dynamic shadows and image blur, thereby improving orthophoto quality and enhancing the reliability of subsequent pavement distress detection.

Specifically, masking techniques are applied to exclude shadowed and blurred regions during orthophoto mosaicking, reducing ghosting artifacts and improving image consistency. A shadow detection approach combining image processing and deep learning is developed to identify dynamic shadow regions, while a sharpness mask based on texture complexity is introduced to evaluate and filter low-quality image areas.

2. PRIDEs and Study Area

This section introduces the PRIDEs system and the data acquisition process.

2.1 PRIDEs System

As shown in **Figure 1**, the PRIDEs system integrates an industrial camera system, a real-time kinematic (RTK) global navigation satellite system (GNSS), a wheel pulse transducer (WPT), and an industrial computer, all mounted on the rear structure of a Tesla Model Y. The system utilizes two FLIR high-resolution monochrome industrial cameras installed on either side of the rear mounting frame, with a spacing of 1.5 m and a mounting height of approximately 1.6 m. As listed in **Table 1**, each camera has a resolution of 4096 × 3000 pixels and is equipped with a 4.5 mm fisheye lens, providing a field of view (FOV) of approximately 120° × 85°. Under this configuration, the system achieves a ground sampling distance of 1 mm and a coverage width of up to 8 m.

To ensure equidistant image sampling, the system records travel distance using the WPT and generates trigger signals accordingly. Compared with time-based triggering methods, the WPT is independent of vehicle speed and maintains a fixed sampling interval, thereby improving the consistency of ortho mosaicking. Dual cameras are synchronously triggered at intervals of 1.25 m, corresponding to an acquisition rate of approximately 24 frames per second (FPS), enabling stable image capture at speeds of up to 110 km/h.

For data recording and system control, the PRIDEs system is equipped with an industrial computer with a high-speed SSD to support continuous high-speed data writing. Dedicated control software is developed to integrate image acquisition, triggering, and GNSS positioning. Since the mounting positions of the cameras and the GNSS antenna are not co-located, system calibration is required to compensate for spatial offsets. The system calibration and positioning accuracy assessment are described in section 3.1.

Figure 1. PRIDEs and its control interface and components

Image Resolution	4096 × 3000
Pixel Size (um)	2.74
Lens Type	Fisheye
Focal Length (mm)	4.5
FOV (H×V)	120°× 85°
Frame Rate	24.5 FPS
Pavement Coverage	Approx. 8 meters
GSD @ 1.6 m	1mm

Table 1. Camera Specifications

2.2 Study Area

The study area is located between the Tucheng and Sanxia interchanges on National Freeway No. 3. The PRIDEs system was used to collect pavement images covering all lanes in both directions, with four lanes per direction and approximately 7 km per lane, resulting in a total coverage of about 56 lane-kilometers and approximately 90,000 images. The spatial distribution of the collected data is shown in **Figure 2**.

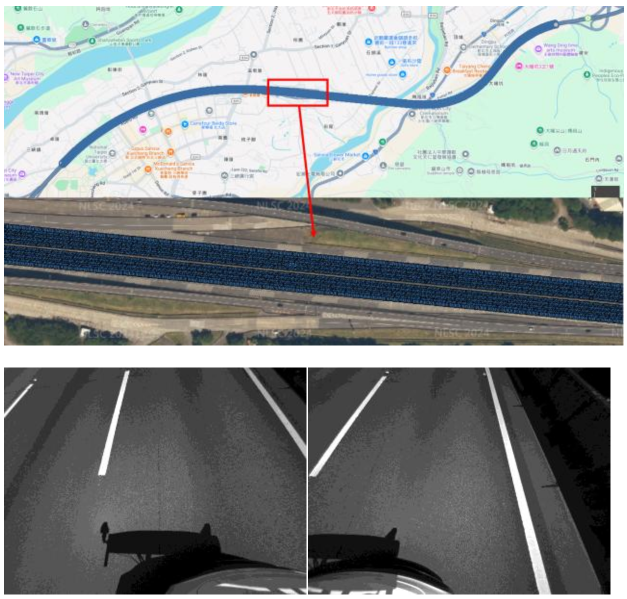


Figure 2. Image acquisition trajectory and sample images.

Due to the large dataset size, a representative section of approximately 500 m in one direction, covering eight lanes in total and including a weigh station segment, was selected for orthophoto quality analysis with respect to shadow and blur masking. **Figure 2**. also illustrates shadows cast by rear vehicle-mounted structures, which are expected to introduce significant ghosting artifacts during orthophoto mosaicking.

3. Methodology

As shown in **Figure 3**, this study first performs systematic calibration of the PRIDEs system. Subsequently, 3D reconstruction techniques are applied to generate high-resolution DSMs and orthophotos from pavement images collected along the National Freeway No. 3. To improve the quality and consistency of orthophoto mosaicking, shadowed and blurred regions in the original images are detected and masked, thereby reducing ghosting artifacts and non-uniform blurring effects in the resulting orthophotos. Finally, the positioning accuracy is evaluated, and the quality differences between the original and enhanced orthophotos are compared.

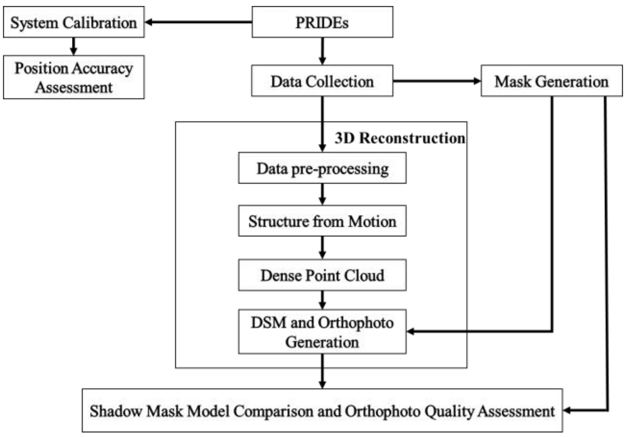


Figure 3. Workflow

3.1 System Calibration

To correct the spatial offset between the camera and the GNSS antenna and to assess the positioning accuracy of the PRIDEs system, a calibration site was established at the Huaxia Campus of National Taiwan University of Science and Technology. As shown in **Figure 4**, the site provides open-sky conditions, ensuring reliable RTK positioning during image acquisition. A total of 13 ground control points (GCPs) were distributed across the campus parking lot and roadways, with their coordinates measured using RTK GNSS.

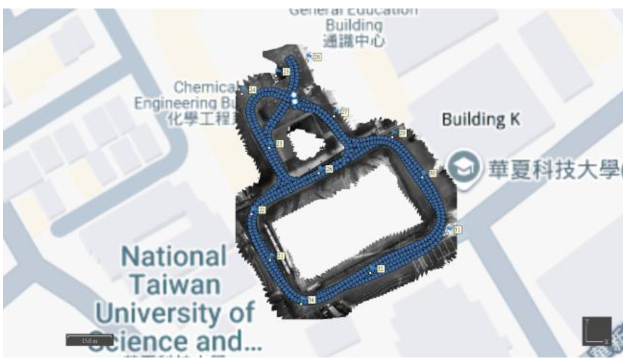


Figure 4. Captured images in the calibration field and distribution of GCPs.

This study adopts an in-situ calibration approach, in which field-acquired images are used to simultaneously estimate camera interior orientation parameters (IOPs) and the spatial offset between the camera and the GNSS antenna during bundle adjustment. As shown in **Figure 4** , the PRIDEs system captures images around the calibration site in both clockwise and counterclockwise directions to reduce systematic bias.

All images are processed in Agisoft Metashape (Agisoft, 2024), where Structure-from-Motion (SfM) is applied to recover camera poses. With onboard RTK positioning solutions and GCP measurements incorporated, the spatial offset between the camera and the GNSS antenna can be effectively calibrated.

3.2 3D Reconstruction

The 3D reconstruction workflow consists of data pre-processing, SfM reconstruction, dense point cloud generation and filtering, DSM interpolation, and orthophoto mosaicking.

3.2.1 Data Pre-processing: During image acquisition, vehicles in adjacent lanes are inevitably captured. These moving objects, when observed across multiple frames, may be incorrectly treated as static features during SfM image matching, thereby degrading the stability of pose estimation and 3D reconstruction. To address this issue, vehicle masking is performed prior to reconstruction to remove dynamic objects. In this study, a YOLO-11 (Jocher, G. and Qiu, J., 2024) instance segmentation model pre-trained on the COCO dataset is employed to detect vehicles and generate corresponding masks. **Figure 5** shows an example of the masking results.

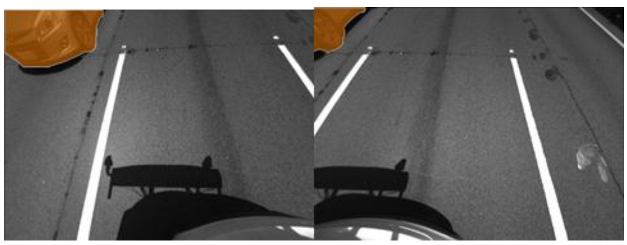


Figure 5. Vehicle masks

3.2.2 SfM Reconstruction: During SfM reconstruction, the pre-calibrated spatial offset between the camera and the GNSS antenna is incorporated as a fixed geometric constraint, enabling the PRIDEs system to achieve centimeter-level positioning accuracy without relying on ground control points (GCPs) when RTK positioning solutions are available. However, given the large volume of image data, exhaustive pairwise matching of all images would result in significant computational cost. To address this issue, a spatial search strategy based on image reference coordinates is adopted, in which planar distances between images are computed, and image pairs are selected for matching only when the distance is below a threshold of 7 m.

3.2.3 Point Cloud Generation and Filtering: Once accurate camera positions and orientations are obtained, a dense 3D point cloud is generated through dense image matching. Due to the PRIDEs imagery having more than 60% overlap both along the driving direction and across adjacent lanes, a high-confidence point cloud can be produced. However, as matching inevitably introduces errors, points with confidence values below 2 are removed to reduce noise during DSM interpolation.

3.2.4 DSM Interpolation and Orthophoto Mosaicking: Interpolating the dense point cloud produces a regularly gridded DSM, which reveals pavement surface variations such as potholes, rutting, and depressions, and enables the generation of continuous elevation profiles for IRI estimation (Chang et al., 2026). To balance computational efficiency and reconstruction quality, dense matching is performed at one-quarter image resolution (Medium quality in Metashape), resulting in a DSM with a resolution of approximately 5 mm.

Subsequently, through orthorectification and image mosaicking, all input images are projected onto a common spatial reference to generate a complete orthophoto with a spatial resolution of up to 1 mm. As a result, the orthophoto can be utilized not only for pavement distress detection, such as cracks, potholes, and patches, but also for spatial localization, measurement, and visualization of pavement deterioration.

However, conventional orthophoto mosaicking methods primarily rely on the geometric relationships between the DSM and camera poses, selecting images based on proximity without considering their quality. Consequently, the resulting orthophotos may suffer from ghosting artifacts caused by vehicle-induced shadows, as well as inconsistent blurring due to variations in lens focus, thereby degrading the overall image quality.

3.3 Mask Generation

To improve the quality of orthophoto mosaicking, this study applies shadow and blur masks to exclude low-quality regions during the mosaicking process.

3.3.1 Shadow Mask: This study employs U-Net as the semantic segmentation model for shadow detection. U-Net adopts an encoder–decoder architecture and preserves multi-scale spatial information through skip connections, making it well suited for pixel-level segmentation tasks. A total of 210 pavement images acquired by the PRIDEs system were manually annotated to train the shadow segmentation model, including 163 images for training, 22 for validation, and 25 for testing. Model training was conducted on a workstation equipped with an NVIDIA RTX 5080 GPU (16 GB).

As shown in **Figure 6**, the U-Net model demonstrates strong performance in detecting shadow regions, particularly those cast by the fixed camera components of the PRIDEs system. The resulting shadow masks are then used to exclude shadowed regions during orthophoto mosaicking.

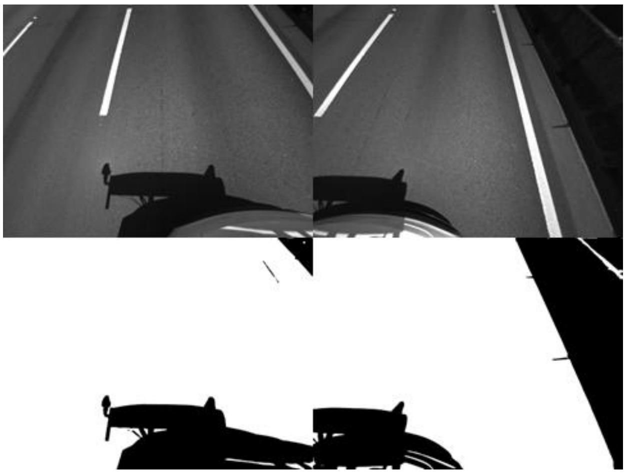


Figure 6. Shadow masks of sample images shown in Figure 2.

3.3.2 Blur Mask: Images captured with wide-angle lenses typically exhibit higher texture complexity in the central regions, while peripheral areas tend to be more blurred. As the degree of blur is closely related to local image sharpness, each image is partitioned into 10 × 10 pixel blocks, and the Laplacian variance of each block is computed according to Equations (1)–(3) to quantify local sharpness..

$$L_{(x,y)} = I_{(x,y-1)} + I_{(x,y+1)} + I_{(x-1,y)} + I_{(x+1,y)} - 4I_{(x,y)} \quad (1)$$

$$L_{avg} = \frac{1}{N}\sum_{i=1}^{N} L_i \quad (2)$$

$$L_{var} = \frac{1}{N}\sum_{i=1}^{N} (L_i - L_{avg})^2 \quad (3)$$

Specifically, the Laplacian operator $L_{(x,y)}$ is used to measure local intensity variation, where $I_{(x,y)}$denotes the pixel intensity at location (x,y). The average Laplacian value L_{avg} is computed within each block, and the variance L_{var} is then calculated to represent the degree of sharpness. Higher variance values indicate stronger high-frequency content and sharper regions, whereas lower values correspond to smoother or blurred areas.

Figure 7. presents a visualization of image blur estimation, where variance values below 20 are shown in red, 40 in yellow, 80 in green, and values above 180 in blue. As observed, higher variance values indicate sharper regions with richer high-frequency information and greater texture variation. In contrast, lower variance values correspond to smoother regions, which may be affected by defocus blur or shadow occlusion, leading to loss of detail.

Since the selected blocks are discrete and do not form continuous regions, high-resolution areas are first retained, followed by circular or elliptical fitting to construct the final effective image mask. To evaluate the impact of different sharpness thresholds on orthophoto mosaicking quality, two thresholds (variance > 80 and variance > 180) are applied to generate corresponding masks for comparative analysis.

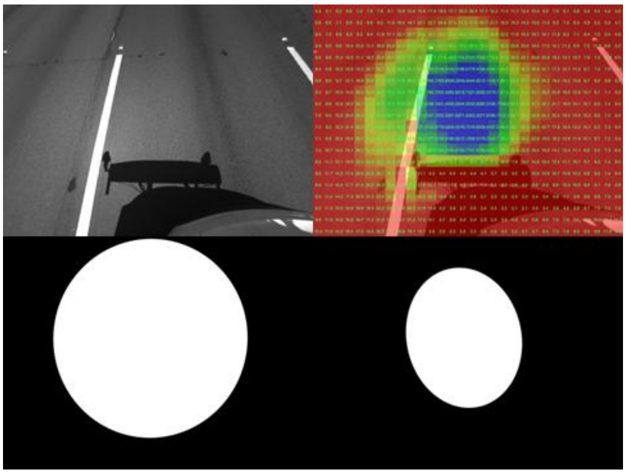


Figure 7. Calculation of Laplace Variance and Blur Masks The bottom left shows the mask for scores above 80, and the bottom right shows the mask for scores above 180.

3.4 Shadow Mask Model Comparison and Orthophoto Quality Assessment

Although this study employs a self-trained U-Net model for shadow masking, comparative experiments with DSC, BDRAR, and OpenCV-based methods are conducted to evaluate the performance and reliability of the proposed approach. Model accuracy is assessed using Intersection over Union (IoU), as defined in Equation (4), which is calculated as the ratio of the area of intersection to the area of union:

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union} \tag{4}$$

Furthermore, to evaluate the quality of orthophotos after applying shadow and blur masks, the method described in Section 3.3.2 is adopted for quantitative analysis and comparison.

4. Results and Analysis

This section presents the positioning accuracy of the PRIDEs system, the 3D reconstruction results, the comparative evaluation of shadow masking methods, and the quality analysis of orthophotos.

4.1 PRIDEs Positioning Accuracy Assessment

During SfM processing, images with RTK positioning solutions were combined with GCPs to estimate the spatial offset between the cameras and the GNSS antenna. The calibration parameters listed in **Table 2** describe the relative geometric configuration between the cameras and the GNSS antenna. The estimated offsets of the two cameras in the X direction are approximately 75 cm and −75 cm, respectively, resulting in a total baseline of about 1.5 m, which is consistent with the physical mounting distance of the camera rig and confirms the validity of the calibration results.

Furthermore, to validate the positioning accuracy of the PRIDEs system under dynamic conditions, an additional image dataset was acquired, and reconstruction was performed using the fixed calibration parameters. Positioning accuracy was evaluated based on the root mean square error (RMSE) of the GCPs. As shown in Table 1, the positioning errors are 1.6 cm in the east (E) direction, 2.9 cm in the north (N) direction, and 5.0 cm in height (H), demonstrating centimeter-level positioning accuracy consistent with RTK performance.

These results indicate that the PRIDEs system is well suited for freeway pavement image acquisition and provides sufficient accuracy to support the localization and spatial analysis of pavement distress.

GNSS offset	X(m)	Y(m)	Z(m)
Cam1	-76.8	-158.0	96.5
Cam2	74.8	-1.625	92.0
GCP(RMSE)	1.6	2.9	5.0

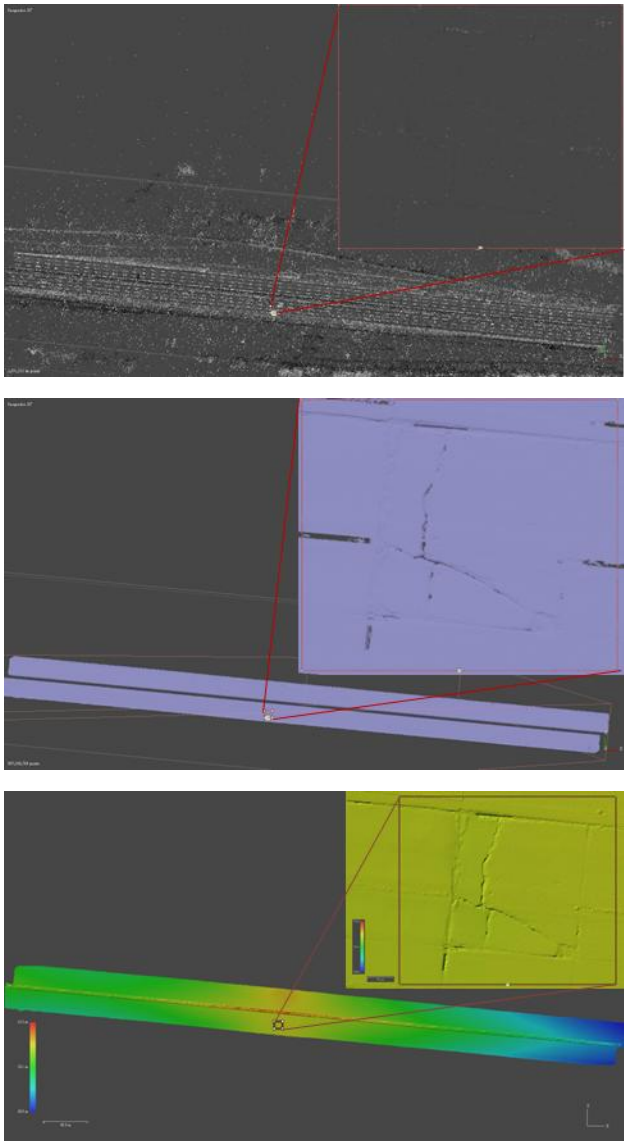
Table 2. Camera-to-GNSS Spatial Offset and RMSE of GCPs

4.2 3D Reconstruction Results

As shown in **Figure 8**, the 3D reconstruction results include the image distribution and sparse point cloud obtained from SfM processing, as well as the generated dense 3D point cloud, a DSM with a spatial resolution of 5 mm, and an orthophoto with a spatial resolution of 1 mm. The results demonstrate that the PRIDEs system is capable of efficiently processing large volumes of image data and generating mosaicked results across multiple lanes.

Furthermore, zoomed-in views of the DSM and orthophoto demonstrate that the high-resolution surface provides clear interpretability, capturing both overall elevation variations along the roadway and localized pavement distresses such as major cracks and potholes. Correspondingly, the 1 mm resolution orthophoto provides a detailed representation of the road surface condition, where cracks consistent with those observed in the DSM can also be identified.

However, as the orthophoto has not yet undergone shadow masking and blur removal, noticeable ghosting artifacts remain, indicating the need for further refinement.



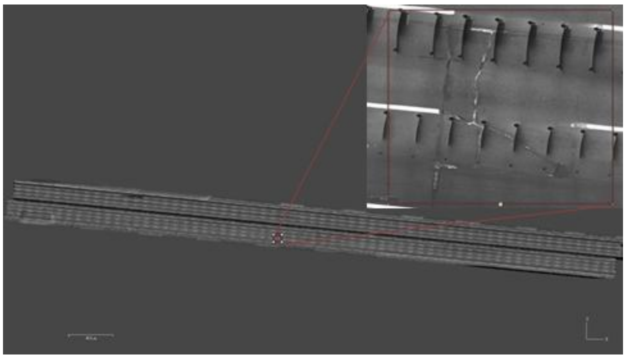


Figure 8. 3D reconstruction results: (top to bottom) SfM image distribution and sparse point cloud, dense point cloud, DSM, and orthophoto.

4.3 Shadow Mask Comparisons

Figure 9. Comparison of masks generated by different methods presents a visual comparison of shadow detection results, including the input image, ground truth, and the outputs of the proposed method, BDRAR, DSC, and OpenCV-based approaches, while the corresponding IoU values are summarized in **Table 3**. The results show that the proposed method achieves more accurate delineation of shadow boundaries and attains the highest IoU of 0.9870 among all compared methods.

Compared with the traditional OpenCV-based approach, the proposed method demonstrates greater robustness in handling pavement texture variations and complex shadow boundaries, while effectively avoiding segmentation fragmentation caused by threshold selection and local brightness variations. Furthermore, compared with deep learning-based methods such as BDRAR and DSC, the proposed method achieves higher segmentation accuracy, indicating its superior capability in capturing the spatial distribution characteristics of shadow regions under the conditions of this study.

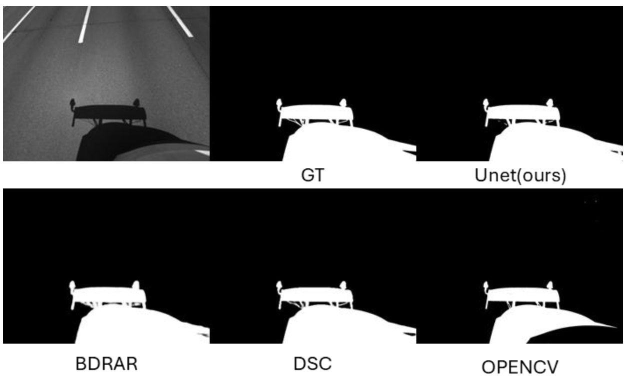


Figure 9. Comparison of masks generated by different methods

Model	IoU
Unet (ours)	**0.987**
DSC	**0.981**
BDRAR	**0.941**
OpenCV	**0.721**

Table 3. IoU comparison of different methods.

4.4 Orthophoto Quality Assessment

This section evaluates the quality of orthophotos generated by the PRIDEs system, with a focus on the effects of shadow and blur artifacts. Specifically, the performance of the shadow mask and blur mask is analyzed, followed by a comparative assessment of orthophoto results under different masking conditions. In addition, the impact of different masking strategies on pavement distress detection is further evaluated..

4.4.1 Orthophoto with Shadow Mask: As shown in **Figure 10**, without masking, the original orthophoto exhibits ghosting artifacts caused by shadows, leading to inconsistent texture variations. After applying the shadow mask, the shadow regions are effectively removed, resulting in improved brightness consistency across the image..

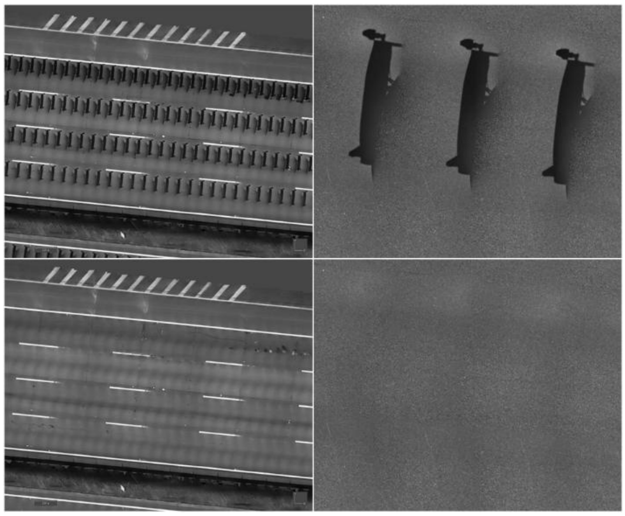


Figure 10. Before and after applying a shadow mask

4.4.2 Blur Mask Analysis: As shown in **Figure 11**, a region of approximately 6×3 m^2, including four continous images and adjacent lanes, was selected for analysis. Three cases were compared: applying only the shadow mask, and applying combined shadow and blur masks with thresholds of 80 and 180. The results indicate that when only the shadow mask is applied, inconsistencies in sharpness remain between continous images and adjacent lanes. With a threshold of 80, the blur mask is insufficient to fully address these inconsistencies, and discontinuities are still observed. In contrast, the threshold of 180 achieves the best performance, effectively improving image consistency. In terms of quantitative evaluation, the mean Laplacian variances for the three cases are 70.1, 74.4, and 80.9, respectively, demonstrating that image sharpness improves with increasing threshold.

However, despite the superior performance of the threshold of 180, local inconsistencies still exist between adjacent lanes. This issue is mainly attributed to the system geometry: with a lane width of approximately 3.5 m and a camera baseline of about 1.5 m, the captured distance increases to approximately 2 m when observing across lanes. This leads to variations in image resolution and viewing geometry, thereby affecting mosaicking consistency. Therefore, future improvements may include increasing the camera baseline or optimizing the camera configuration to reduce cross-lane distance differences and enhance overall multi-lane mosaicking consistency.

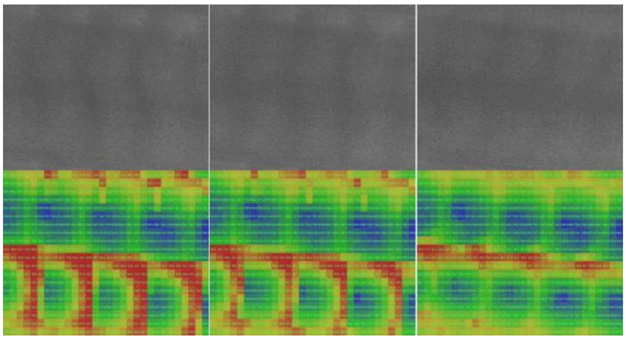


Figure 11. Comparison of orthophotos under different blur mask thresholds.

4.4.3 Comparison of Orthophoto with different Masks: As shown in **Figure 12**, the quality differences of orthophotos before and after masking using the PRIDEs system are presented. The unprocessed orthophoto contains prominent shadow bands and regions with locally inconsistent sharpness. After applying the shadow mask, high-contrast shadow regions are effectively removed, resulting in improved brightness consistency across the retained pavement areas; however, local inconsistencies in image quality still remain. By further integrating both shadow and blur masks, the resulting orthophoto exhibits more consistent brightness and sharpness, thereby effectively reducing noise and interference in subsequent deep learning-based pavement distress detection. Comparison of Masking Effects on Distress Detection.

To evaluate the impact of orthophoto quality on pavement distress detection, this study employs DeepLabv3+ for distress detection. As shown in **Figure 12**, shadow regions are often misclassified as pavement defects, while actual defects may be obscured by shadows and remain undetected. In addition, sharpness discontinuities in the image may lead to false detections along seam regions. In contrast, when the image exhibits more homogeneous sharpness and is free from shadow interference, the detection results become more stable and the number of misclassifications is significantly reduced.

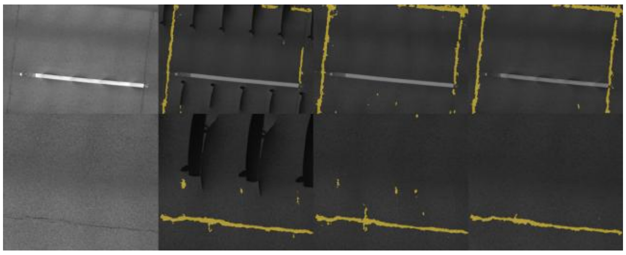


Figure 12. Comparison of Crack Detection

5. Conclusions

This study demonstrates the feasibility and practical potential of applying the PRIDEs system to freeway pavement inspection. The system enables high-resolution image acquisition under normal traffic conditions without requiring road closures, providing both operational efficiency and large-area coverage. The generated high-resolution DSM (5 mm) and orthophotos (1 mm) allow detailed observation of pavement conditions, meeting the requirements of high-precision pavement inspection. However, the acquired imagery is still affected by shadows cast by vehicle structures and local blur caused by lens defocus, which may interfere with pavement distress detection.

To mitigate these effects, shadow and blur masks are incorporated into the orthophoto mosaicking process to identify and exclude low-quality regions. Experimental results show that the proposed method improves the consistency of orthophoto quality and effectively reduces false detections in crack identification.

References

Agisoft. 2024. “Discover intelligent photogrammetry with Metashape.” https://www.agisoft.com/.

Chang, J.-R., Jhan, J.-P., Yang, P.-S., and Lin, K.-H.: Development and Accuracy Validation of an Image-Based Pavement Roughness Inspection System, Journal of Transportation Engineering, Part B: Pavements, 152, 04026005, 10.1061/jpeodx.Pveng-1762, 2026.

Hu, X., Fu, C., Zhu, L., Qin, J., and Heng, P.: Direction-Aware Spatial Context Features for Shadow Detection and Removal, IEEE Transactions on Pattern Analysis & Machine Intelligence, 42, 2795–2808, 10.1109/TPAMI.2019.2919616, 2020.

Jocher, G. and Qiu, J., 2024. Ultralytics YOLO11. https://github.com/ultralytics/ultralytics. Software, version 11.0.0.

Zhu, L., Deng, Z., Hu, X., Fu, C.-W., Xu, X., Qin, J., and Heng, P.-A.: Bidirectional Feature Pyramid Network with Recurrent Attention Residual Modules for Shadow Detection, Computer Vision – ECCV 2018, Cham, 2018//, 122–137,