Proc. of the 9th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Zhongqiang Liu, Jian Dai and Kate Robinson
Published by Research Publishing, Singapore.
doi: 10.3850/GRF-25280825_isgsr-169-P204-cd

ISGSR2025

A STUDY ON REAL-TIME 3D RECONSTRUCTION BASED ON NeWCRFs: A CASE STUDY OF EXCAVATION ENGINEERING

Chenxi Han
Dept. of Geotechnical Engineering, Tongji Univ, Shanghai 200092, China. E-mail: hcx@tongji.edu.cn
Hongwei Huang
Dept. of Geotechnical Engineering, Tongji Univ, Shanghai 200092, China. E-mail: huanghw@tongji.edu.cn
Siyi Guo
Dept. of Geotechnical Engineering, Tongji Univ, Shanghai 200092, China. E-mail: gsiyi@tongji.edu.cn
Linghan Ouyang
School of Engineering, Cardiff University, Cardiff CF24 3AA, UK. E-mail: ouyangl@cardiff.ac.uk

Abstract: Surface surcharge loads around excavation engineering are critical to project stability. However, despite strict regulations, existing monitoring methods lack both precision and real-time capabilities. This study introduces a UAV-based monitoring system that integrates 3D reconstruction and load space extraction to enhance anomaly detection accuracy. The system was tested in ultra-large and ultra-deep excavations in Nanjing's Jiangbei New Area, and a comprehensive surcharge load categories database was established. The NeWCRFs model enables depth estimation from a single RGB image, point cloud generation, and load distribution visualization through spatial projection. Additionally, the system incorporates image enhancement, real-time depth prediction, point cloud modeling, and bird's-eye visualization within a user-friendly interface, facilitating practical deployment. The findings of this study significantly enhance the accuracy of abnormal surcharge load monitoring and provide strong support for excavation safety management.

Keywords: NeWCRFs, 3D reconstruction, point cloud, image enhancement, UAV, excavation engineering.

1. Introduction

With advancements in construction technology and engineering demands, excavations are trending toward large-scale, ultra-deep pit clusters. However, complex geological and environmental conditions challenge stability control, particularly in monitoring surcharge loads, which still relies on subjective observation and lacks timely execution (Huang 2017). High-precision automated inspections are costly and difficult to scale, leading to inadequate supervision of abnormal surcharge loads.

A potential solution is a real-time inspection algorithm deployable on portable drones, offering a low-cost, mobile, and rapid approach. Effective control requires acquiring 2D and 3D load information, including spatial position, volume, and type. In computer vision, 3D reconstruction provides a promising method for drone-based real-time inspections.

Research in 3D vision has evolved through four stages: stereo vision using depth/stereo cameras, high-precision LiDAR scanning, Structure from Motion (SfM) for survey-grade tasks, and deep learning-based monocular 3D reconstruction (e.g., NeRF, 3D Gaussian Splatting). While computationally demanding, deep learning methods are now mainstream, with real-time optimization techniques emerging. Notably, Alibaba's NeWCRFs enables single-image 3D reconstruction with improved precision.

This study conducts a 3D visualization experiment in Nanjing's Jiangbei New District using NeWCRFs and develops a drone-based algorithm for real-time surcharge load inspections, paving the way for remote, intelligent monitoring of excavations (Yuan 2022)

2. Methodology

The core of the algorithm is to establish a real-time, lightweight 3D reconstruction algorithm based on monocular images, which can be deployed on portable civil drones. NeWCRFs (New Conditional Random Fields) is a probabilistic graphical model used for sequence labeling and structural prediction. It represents an innovation and extension of traditional Conditional Random Fields (CRFs). Compared to traditional CRFs, NeWCRFs have stronger capabilities in modeling complex dependencies, capturing long-range contextual information, and handling multimodal data. As a result, it is well-suited for the rapid inspection tasks of surcharge loads around excavations.

2.1. *Window Fully-connected CRFs*

In traditional monocular depth estimation, Conditional Random Fields (CRFs) are often used to construct an energy function by incorporating observational cues such as texture and location information along with the final predictions. This energy function is then optimized to obtain depth predictions.

$$E(\mathbf{x}) = \sum_{i} \psi_u(x_i) + \sum_{ij} \psi_p(x_i, x_j) \tag{1}$$

$$\psi_p = \mu(x_i, x_j) f(x_i, x_j) g(I_i, I_j) h(p_i, p_j) \tag{2}$$

Here, x_i is the predicted value of node i, and j represents all other nodes. The unary potential ψ_u is computed from image features, while the pairwise potential ψ_p models node connects pair of nodes with $\mu(x_i, x_j) = 1$, if $i \neq j$ and $\mu(x_i, x_j) = 0$. I_i and p_i denote the color and position of node i, respectively.

However, due to the high computational cost, this method is typically limited to neighborhood CRFs and cannot achieve fully connected CRFs (FC-CRFs). NeWCRFs introduces a window-based fully connected CRF algorithm as shown in Fig 1 (b) compared with Fig 1 (a), which segments the image into multiple patch-based windows. Each window contains N×N image patches, where each patch consists of N×N pixels. In the algorithm, each block is treated as a node, and all patches within a window are fully connected, while patches between different windows are not connected. This approach reduces computational complexity by considering only the patches within a single window when calculating the energy function. To address the isolation issue between windows, the algorithm adopts the idea from swim-transformers by shifting the patches within each window by (N/2, N/2) and then recalculating depth energy function, thereby connecting isolated pixels across windows.

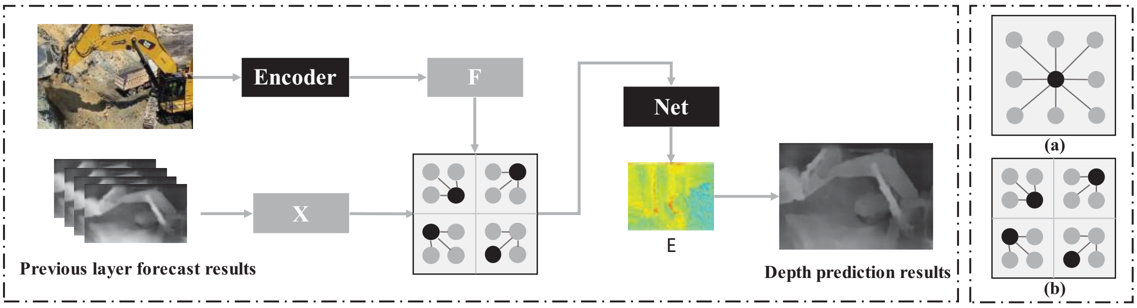

Fig. 1. Algorithm flow and sliding window method full connection principle.

2.2. NeWCRFs neural network structure

By replacing manually designed potential functions that represent predicted probability distributions with neural networks, NeWCRFs can capture high-dimensional information and describe complex connections that are difficult to represent through handcrafted designs. NeWCRFs constructs a bottom-up structure that performs CRF optimization across four levels. We integrate the neural window FC-CRFs module into the network as a decoder, which predicts the next layer of depth based on coarse depth and image features. Meanwhile, a Swin Transformer is used as the encoder to extract features. The patches in the four CRF levels are sized 4×4, 8×8, 16×16, and 32×32 pixels, respectively, from top to bottom. At the top of the algorithm, a Pyramid Pooling Module (PPM) is used to aggregate information from the entire image. Each window contains 7×7 patches to facilitate the computation of unary and pairwise potentials using the Transformer. The energy function is then fed into an optimization network, composed of two fully connected layers, and the network outputs the optimized depth map. (Yuan 2022) And The entire depth prediction process is shown in Fig 1.

3. Case Study of Nanjing Jiangbei

The underground complex project in the Jiangbei New District of Nanjing is a key infrastructure project for the construction of a national-level new area. It involves an ultra-large and ultra-deep excavation group, covering more than 100,000 square meters with a depth exceeding 40 meters. The construction conditions in the Jiangbei New District are highly complex, as the project is located just 600 meters from the Yangtze River, requiring careful consideration of the impacts of pressurized water and geological voids on the engineering work. Additionally, the project is situated near an existing structure and a super-high-rise building currently under construction in the first construction zone. Therefore, it is essential to establish an automated, rapid inspection method to strictly monitor the surcharge loads around the excavations, preventing potential safety hazards in critical excavations.

3.1. Image acquisition and enhancement process

The image data is acquired using a cost-effective, small rotary-wing civilian drone. For this case study, the DJI-AVATA portable drone is employed. This drone can be remotely controlled by an operator on the ground to capture 4K high-definition monocular images and videos. During each image data collection, the drone must fly carefully around the edge of the excavation and the surrounding environment to record video. Subsequently, frames are extracted from the recorded video using a frame sampling method to obtain the image data. To ensure that the image data acquired by the drone meets the requirements for subsequent analysis, image enhancement is necessary. This section provides a detailed explanation of two aspects: adaptive Gamma transformation and internal parameter calibration.

3.1.1. *Adaptive Gamma Transform*

Due to the complex on-site conditions of the project, some images may suffer from overexposure or underexposure, which can negatively impact subsequent image processing and analysis. To address this, adaptive Gamma transformation is applied to correct the image brightness. Based on the overall brightness distribution, the Gamma value is adaptively adjusted to optimize the image's resolution and detail representation. (Babakhani 2015)

$$X^{\gamma} = Y = \frac{1}{2} \tag{3}$$

The above is the formula for adaptive gamma transformation. In this context, the size of X can represent the current brightness, while Y represents the target brightness after adjustment. The target brightness is set to the average value, which is used to calculate the gamma value.

Fig. 2.Comparison of the original image with the enhanced image.

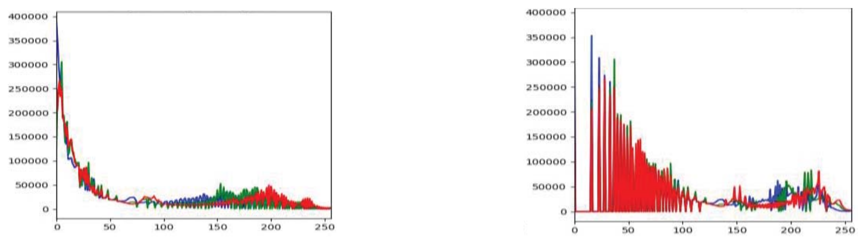

Fig. 3. Comparison of RGB channels before and after transform.

After processing, the image's balance and recovery are significantly improved as shown in Fig 2, providing a clearer data foundation for subsequent analysis. The brightness adjustment of the RGB channels is shown in Fig 3.

3.1.2. *Internal parameter calibration*

Optical lenses inevitably introduce geometric distortions, which depend on the camera's internal parameters. To ensure the geometric accuracy of the images captured by the drone, it is necessary to calibrate the camera's internal parameters. The calibration process is based on the classical camera calibration method, where images of a calibration board with known dimensions are used to accurately compute the camera's focal length, principal point location, and distortion parameters. For the calibration of focal length and principal point position, calibration images of the board from multiple angles are used, and Zhang's Calibration Method is employed to solve for the intrinsic parameter matrix of the focal length, minimizing image distortion. (Zhang 1999)

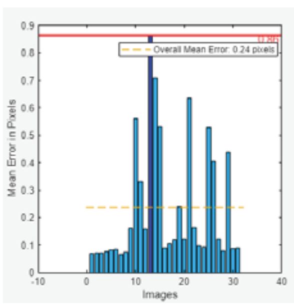

Fig. 4. Intrinsic calibration error histogram.

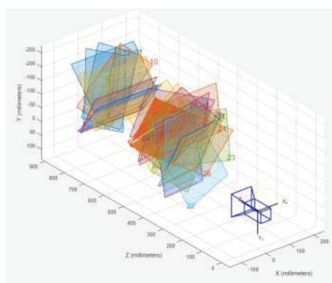

Fig. 5. Chessboard's pose during the calibration.

The important internal parameters for calibration include the camera internal control matrix and five distortion coefficients, and the specific results are as follows:

$$\begin{bmatrix} fx & 0 & cx \\ 0 & fy & cy \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1.4807e+03 & 0 & 1.9918E+03 \\ 0 & 1.4801e+03 & 1.1115e+03 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

$$[\mathrm{k1}, \mathrm{k2}, \mathrm{p1}, \mathrm{p2}, \mathrm{k3}] = [-0.0744, -0.0535, -2.95e-04, 5.93e-04, 0.0314] \tag{5}$$

The calibration error is 0.24 pixels, which meets the accuracy requirement (0.84 pixels), so the calibration is valid.

3.2. NeWCRFs 3D Reconstruction and Comparison

After performing image enhancement preprocessing, NeWCRFs is run to carry out 3D reconstruction of the excavation site, enabling the acquisition of the reconstructed point cloud along with textured details. After performing image enhancement preprocessing, NeWCRFs is run to carry out 3D reconstruction of the excavation site, enabling the acquisition of the reconstructed point cloud along with textured details. And the results of the reconstruction will be presented in the section 3.4. Due to the optimal design of the computing power of the method, the algorithm can be used to complete real-time reconstruction which is not possible with traditional 3D reconstruction algorithms, and its various aspects are shown in Table 1.

Table 1. Comparison between 3D reconstruction algorithms.

Method	SfM	NeRF	NeWCRFs
Reconstruction Efficiency	Medium	Low	real-time
RAM Usage	High	Medium	Low
Density of Point Clouds	Medium	Dense	Medium
Point Cloud Quality	Medium	High	Medium
Scope of Reconstruction	Small	Large	Large
Other Functions	Position Estimation	Composite New Views	Composite New Views

3.3. *Bird's Eye View Transform (BEV)*

In the current construction specifications for deep excavation engineering, abnormal surcharge loads are identified based on the horizontal spatial distance between the load and the foundation pit. To obtain the spatial information of the surcharge load, a Bird's Eye View (BEV) transformation is applied to the 3D reconstructed point cloud data, making it easier to calculate the spatial position of the load-bearing objects. The specific implementation involves writing Python code to move the viewpoint to a top-down view and project the point cloud onto a 2D plane while retaining color information. A coordinate system is then created based on the spatial positions of the point cloud. A grid is drawn on the BEV with meter units, allowing for the determination of the relative position of each load-bearing object to the excavation on the construction site. The BEV effect will be showcased in the 3.4 section.

3.4. *Graphic User Interface and risk control*

To facilitate user interaction and provide an intuitive view of the inspection results, this study developed a depth estimation and point cloud visualization application based on PyQt5 and OpenGL. The application is divided into two main parts: the graphical user interface (UI) and the 3D rendering section. The UI is structured into several areas: an image display area, a predicted depth map display area, an OpenGL 3D rendering area, a BEV viewing area, and a button section. The OpenGL 3D rendering area is interactive, allowing users to view the 3D reconstructed point cloud in real-time with 360° rotation, using mouse clicks, dragging, and scrolling. The BEV display area on the far right generates a real-time bird's eye view from the point cloud and includes a coordinate system for real-time assessment of the relative position of surcharge loads and the excavation.

Using this intuitive interaction and observation interface above, the generated point cloud data can be used to calculate the volume of irregular pile loads and extract the location information. This information can be used to make threshold judgments or input into numerical models for subsequent stability analysis, which can effectively and timely control the pit stability safety risk caused by surface surcharge loads on the construction site.

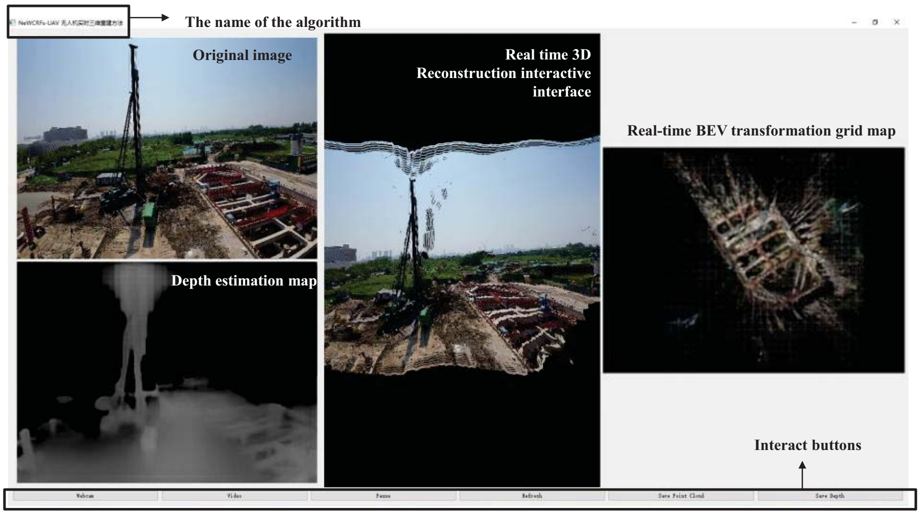

Fig. 6. Various functional areas of the user graphical interface.

Since the algorithm can be deployed on any civilian UAV with vision capabilities, the method can be used in the majority of excavation construction scenarios with pre-reporting (additional permissions may need to be obtained for projects near key areas such as airports).

4. Conclusion

Based on the basic principle of NeWCRFs, this paper obtains data through UAV and uses image enhancement algorithm for image preprocessing, summarizes and proposes a new UAV real-time 3D reconstruction method suitable for excavation, and converts the reconstructed 3D model into BEV image and integrates it into an interactive GUI interface. The feasibility of the above method is verified by field test at the construction site of excavations in Jiangbei New Area, Nanjing.

5. Acknowledgments

The authors wish to thank the support from National Key R&D Program of China (Grant No. 2023YFC3009303).

References

Babakhani, P., & Zarei, P. (2015). Automatic gamma correction based on average of brightness. Advances in Computer Science: an International Journal, 4(6), 156-159.

Huang, Z. (2017, December). Analysis and Research on the Impact of Surcharge Loading near the Foundation Pit and Over Excavation within the Pit on Deformation in Hangzhou Soft Soil Area. Master's thesis, Zhejiang University of Technology.

Yuan, W., Gu, X., Dai, Z., Zhu, S., & Tan, P. (2022). Neural window fully-connected CRFs for monocular depth estimation. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 3916-3925.

Zhang, Z. (1999, September). Flexible camera calibration by viewing a plane from unknown orientations. In Proceedings of the seventh IEEE international conference on computer vision, 666-673.