Analysis of Patch Shapes in Dense Matching for Reducing Disparity Errors in UAV Stereo Images

Kim H.[1], Lee C.[2], Kim T.[2*]

[1]*Program in Smart City Engineering, Inha University, Incheon, Republic of Korea.*
twojin514@inha.edu
[2]*Department of Geoinformatic Engineering, Inha University, Incheon, Republic of Korea.*
ichaeyeon269@inha.edu
[2]*Department of Geoinformatic Engineering, Inha University, Incheon, Republic of Korea.*
tezid@inha.edu

Abstract: Traditional dense stereo matching for 3D reconstruction from UAV images have limitations in complex boundary and low-texture areas, leading to degraded accuracy. To address this, we propose a method that integrates multi-shape patches, each specialized for different directional characteristics, and estimates the optimal disparity using a voting-based support aggregation. Experiment results show that the proposed method reduced the end-point-error by up to 38.7% compared to using a single square patch and enabled successful disparity estimation in complex structures and low-texture areas. This demonstrates that the proposed method was a highly effective strategy for generating high fidelity 3D spatial information.

Keywords: Dense stereo matching, UAV image, Disparity estimation, 3D reconstruction

1. Introduction

Digital twin and smart city require 3D modeling technologies capable of accurately replicating the real world in virtual space. Unmanned aerial vehicles (UAVs) have emerged as a key data acquisition platform for this purpose, as they can rapidly capture ultra-high-resolution stereo images. Such UAV images offer significant advantages for constructing detailed three-dimensional information of complex urban environments and natural terrains. Dense matching, which generates dense point clouds from UAV stereo image, is a core technology for 3D reconstruction (Lee et al., 2024).

However, traditional local dense matching methods that are based on square patch face fundamental limitations in fully handling the structural complexity. These approaches fail to account for the directional characteristics of object boundaries and tend to average disparities in areas with strong depth discontinuities, leading to distorted boundaries (Stentoumis et al., 2014). In addition, in low-texture areas, reliable feature correspondences are often difficult to obtain, resulting in frequent disparity estimation failures. Consequently, the overall precision of the 3D reconstruction degrades, and essential disparity information of objects is lost.

To overcome these limitations, this study proposes a dense matching method based on multi-shape patches integrated using voting-based support aggregation. To assess the contribution of each shape, the study first investigates the tendencies of disparity estimation when applying each patch shape individually. Our approach enables reliable, high-precision 3D reconstruction with UAVs, even in complex environments where methods using a single patch fail.

2. Method

This study proposes a dense matching method for accurate disparity estimation from UAV stereo images. The proposed method combines a multi-resolution image pyramid with multi-shape patch-based support aggregation. The proposed method follows a coarse-to-fine strategy that progressively refines disparity from low to high resolutions. Each step consists of four main steps: multi-resolution image pyramid and cost calculation, multi-shape patch-based support aggregation, disparity determination, and post-processing.

First, a multi-resolution environment is established by constructing an image pyramid from the input stereo images. This is intended to efficiently address the wide search range problem and to robustly handle large disparity displacements. At each pyramid level, the initial matching cost is calculated using the Census transform, which is robust to illumination changes. The Census transform utilizes the relative intensity order of neighboring pixels rather than their absolute brightness values. As a result, it can minimize the effects of exposure differences or shadows that may occur in UAV images.

Second, this study applies a voting-based support aggregation that utilizes multi-shape patches. This approach is to solve the problem of traditional square patches averaging information across object boundaries. In addition to squares, we utilize patches with various shapes that can account for directionality, such as horizontal/vertical rectangles, a cross, a diagonal cross, and quadrant. For each patch, if a disparity candidate shows consistency between the left and right images, the support for that disparity is defined as being incremented. The accumulation of these support values quantitatively represents the reliability of each disparity candidate at every pixel.

Once the support aggregation for all patches is complete, the disparity with the highest support count at each pixel is determined as the final disparity for that level. This disparity from the low-resolution level is then propagated to the next higher-resolution level, where it serves two roles. It functions as an anchor that centers the search range, effectively reducing the search space. It works as a seed that influences the disparity determination of surrounding areas. This coarse-to-fine strategy plays a key role in maximizing computational efficiency and reducing uncertainty in ambiguous regions.

Finally, a post-processing is applied to the disparity map determined at each level, which includes a median filter and a left-right consistency check. This process removes noise generated during the matching process and effectively corrects inconsistent disparities in areas such as occlusions. As a result, the overall quality of the final disparity map is improved.

3. Results

To validate the performance of the proposed method, we used the UAVStereo dataset, which consists of scenes with characteristics: residential, forest, and mining area (Zhang, X et al., 2024). To analyze the effectiveness of the multi-shape patches, we established a baseline using a square patch. We then compared this baseline against an approach that integrates all shape patches except for the quadrant patch. The visual results of the disparity estimation for this comparison are presented in Figure 1.

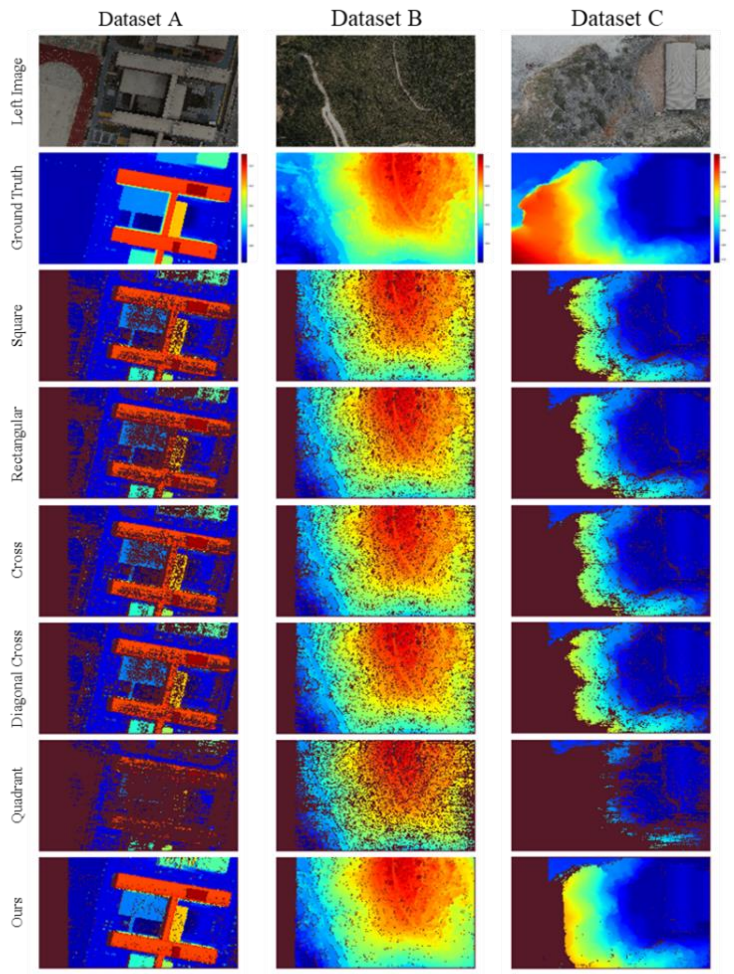

Figure 1: Comparison of disparity maps by patch shape.

Although each patch shape exhibited distinct advantages, it became clear that no single shape was sufficient to handle all conditions. While the square patch performed reliably in uniformly textured areas, it caused a blurring effect at object boundaries, where disparity would bleed across the edges. The horizontal rectangular patch, being aligned with the epipolar line, processed horizontal boundaries sharply, but it caused distortion on the vertical walls of buildings. The vertical rectangular patch exhibited the opposite characteristic, showing strength in vertical structures while distorting horizontal boundaries. The cross and diagonal patches minimized pixel sampling along their respective axial and diagonal boundaries, yielding sharp results in those areas. However, due to the limited amount of information they process, they tended to produce noise in larger areas like wide walls. The quadrant patch, which represents the most extreme case of this tendency, showed potential for minimizing blur by including only one side of a boundary, but it produced the most unstable results due to insufficient information.

The proposed method significantly reduced the error compared to the baseline across all datasets. Specifically, the end-point-error (EPE) for the residential area of Dataset A decreased from 118.75 pixels to 77.69, for the forest area of Dataset B from 53.88 pixels to 33.03, and for the mining area of Dataset C from 189.95 pixels to 152.22. These results demonstrate that the proposed multi-shape patch-based support aggregation was an effective method for improving disparity accuracy.

4. Conclusion

To overcome the limitations of single patch in complex boundary and low-texture areas of UAV image, this study proposed a method that combines patches of various shapes. The proposed method employs a voting-based approach to aggregate support from patches that have strengths in different directional and structural characteristics. The experiment results demonstrate that this approach improved stability in low-texture areas while simultaneously increasing overall disparity accuracy. These findings demonstrate that aggregating support from a variety of patch shapes effectively overcome inherent methodological limitations. This enables the robust generation of high fidelity 3D spatial information, in complex environments observed by UAV.

Acknowledgements

This work was carried out with the support of “Cooperative Research Program for Agriculture Science and Technology Development (Project No. RS-2021-RD009566)” Rural Development Administration, Republic of Korea.

References

Lee, E., Park, S., Jang, H., Choi, W. and Sohn, H. G., 2024. Enhancement of low-cost UAV-based photogrammetric point cloud using MMS point cloud and oblique images for 3D urban reconstruction, Measurement, 226, 114158.

Stentoumis, C., Grammatikopoulos, L., Kalisperakis, I. and Karras, G., 2014. On accurate dense stereo-matching using a local adaptive multi-cost approach. ISPRS Journal of Photogrammetry and Remote Sensing, 91, 29-49.

Zhang, X., Cao, X., Yu, A., Yu, W., Li, Z. and Quan, Y., 2023. UAVStereo: A multiple resolution dataset for stereo matching in UAV scenarios, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16, 2942-2953.