SRGCD: Stability-Driven Region Growth Framework for 3D Change Detection

Supplementary Material

A. Additional Implementation Details

To ensure consistency with existing 3D change detection frameworks, our method adopts the same preprocessing and sampling protocol as SiameseKPConv and PGN3DCD, including voxel downsampling, block cropping, and feature construction.

This section provides the implementation hyperparameters of the Mutual Geometric Consistency Prior (MGCP) and Stability-Guided Controlled Attention (SGCA), as well as the complete feature dimension flow of the network and the training configurations used in our experiments.

A.1. Mutual Geometric Consistency Prior (MGCP)

MGCP is applied at the deepest layer to identify geometrically consistent unchanged regions. Tab. 5 summarizes the dataset-specific geometric thresholds used for Urb3DCD and HKCD.

Table 5. MGCP parameters for Urb3DCD and HKCD.

Parameter	Urb3DCD	HKCD
Normal-estimation neighbors k_{normal}	16	20
Distance threshold τ_d	5.0 m	3.0 m
Normal-angle threshold τ_θ	15°	15°
Mutual nearest neighbor k	1	1
Input sampling resolution	1.0 m grid	1.5 m grid

A.2. Stability-Guided Controlled Attention (SGCA)

SGCA consists of three components: Stability-Guided Attention (SGA), Stability-Adaptive Propagation (SAP), and Stability-Proportional Temperature (SPT). Among them, SGA introduces no additional hyperparameters. In SAP, we set the propagation coefficient to $\alpha = 0.25$, so that stable regions receive only a small amount of update while unstable regions can absorb more information from stable ones. In SPT, the temperature range is set to $[T_{\min}, T_{\max}] = [0.4, 1.0]$, and the per-layer temperature is dynamically determined according to the stability ratio $r^{(\ell)}$ via linear interpolation within this range. In practice, the deepest layer often yields an effective temperature close to 0.6, indicating that the attention is not overly sharp even at the coarsest stage.

A.3. Feature Dimension Flow

Tab. 6 summarizes the feature dimensions across all stages of our network. The encoder contains five stages (E0–E4), each consisting of two KPConv blocks that progressively expand the feature dimensionality from the input (1 for Urb3DCD, 4 for HKCD) to 2048. During the coarse-to-fine refinement process, the refined feature from the deeper level is first upsampled and then directly processed by SGCA to produce the next-level refined representation, without concatenation with encoder features. At each propagation stage, a lightweight segmentation head is applied for supervision. Specifically, the refined feature is concatenated with the corresponding encoder difference feature and the upsampled mask from the previous stage. The fused feature is then projected to 64 channels, followed by a final mapping to a 1-dimensional binary change score ($64 \to 1$). The four segmentation heads take fused features of dimensions 1024+1024+1, 512+512+1, 256+256+1, and 128+128+1, respectively.

Table 6. Feature dimension flow of the encoder and the coarse-to-fine stability propagation stages.

Stage	Input Dim	Output Dim
Encoder E0	1 (4)	128
Encoder E1	128	256
Encoder E2	256	512
Encoder E3	512	1024
Encoder E4 (deepest)	1024	2048
Propagation L0 (deepest)	2048	2048
Propagation L0 $\to$ L1	2048	1024
Propagation L1 $\to$ L2	1024	512
Propagation L2 $\to$ L3	512	256
Propagation L3 $\to$ L4	256	128

A.4. Training Configuration

All experiments on Urb3DCD and HKCD follow the same optimization setting. We use stochastic gradient descent (SGD) with an initial learning rate of 0.01, momentum of 0.98, and weight decay of 1×10^{-3}. An exponential decay scheduler is applied with a decay factor $\gamma = 0.9885$, and the learning rate is updated at every epoch. Batch normalization momentum is controlled by a BNMomentumScheduler with a base momentum of 0.02. Both datasets are trained with a batch size of 4 for 200 epochs. No data augmentation is used, and the training pipeline directly operates on the grid-sampled inputs produced by the dataset-specific pre-transform. Urb3DCD uses a 1.0 m grid size, while HKCD uses a 1.5 m grid size. Dropout with a rate of

0.5 is applied in all segmentation heads. The boundary loss weight is fixed to $\lambda = 2.0$ in all experiments. All models are trained on a single NVIDIA GeForce RTX 4090 GPU.

B. Additional Ablation Study

B.1. Ablation on Supervision Strategy

We evaluate the influence of the supervision strategy and the sensitivity of the boundary loss weight λ. All experiments are conducted on Urb3DCD-V1. As shown in Tab. 7, we compare single-level supervision (final change map only) with multi-level supervision (all refinement levels), and further vary λ in the multi-level setting.

Table 7. Ablation on supervision strategy and boundary loss weight λ on Urb3DCD-V1.

Setting	**Supervision**	λ	**mIoU (%)**
Single-level	L4	2.0	92.37
Multi-level	L1–L4	1.0	93.52
		2.0	**94.11**
		3.0	93.85

Multi-level supervision significantly improves stability in the region-growth process compared with supervising only the final change map. The best mIoU is obtained with $\lambda = 2.0$, which provides a good balance between interior completeness and boundary compactness. Smaller λ weakens the influence of the boundary-aware term and slightly blurs the predicted edges, while larger λ tends to overemphasize boundary sharpness with limited gains.

B.2. Ablation on Geometric Seed Definitions

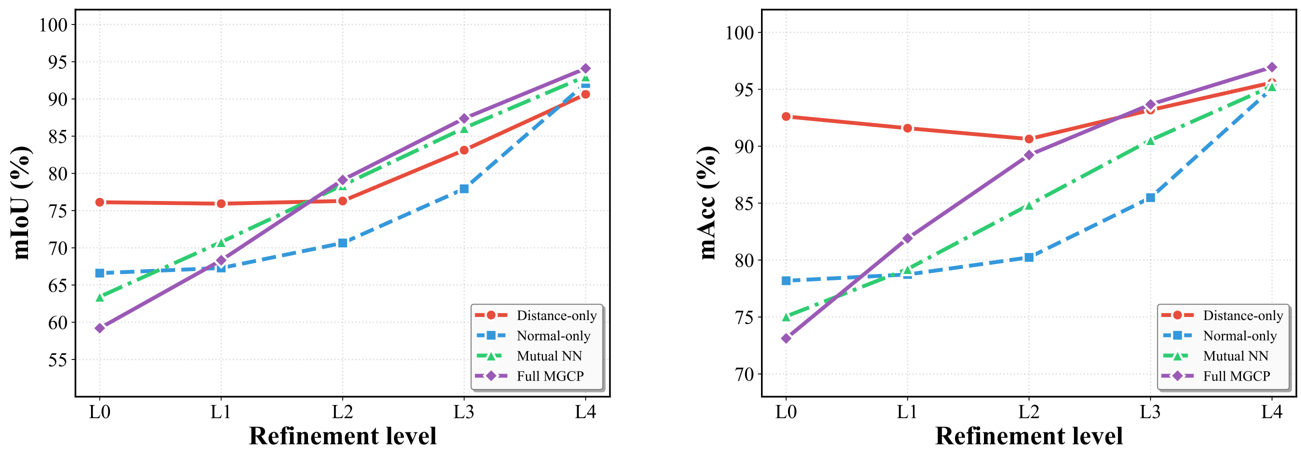
To analyze the influence of geometric initialization on the stability-driven evolution process, we compare four seed definitions at the deepest layer L0: (i) distance-only, (ii) normal-only, (iii) mutual nearest-neighbor, and (iv) full MGCP (distance + normal + mutual). Each seed type is propagated through the same network, and the change-detection performance is evaluated across refinement levels L0–L4. L0 corresponds to the raw geometric decision, whereas L1–L4 reflect the progressively refined estimates produced by SGCA-based stability propagation.

As shown in Tab. 8 and visualized in Fig. 6, we attempt to interpret the behavior of different seed definitions through the lens of the stability-driven region-growth mechanism. In our framework, stability propagation expands outward from the initial stable seeds; therefore, the stability purity at L0 plays a more critical role than the magnitude of the initial IoU. Although Distance-only and Normal-only seeds exhibit high IoU at L0, their stability purity is noticeably lower, meaning the "stable region" contains points

Table 8. Comparison of geometric seeds across refinement levels (L0–L4) on Urb3DCD-V1. All numbers are percentages.

Seed Type	Layer	IoU (unch/chg)	Acc (unch/chg)
Distance-only	L0	[90.54, 61.71]	[91.47, 93.76]
	L1	[91.63, 60.23]	[90.85, 92.32]
	L2	[92.74, 59.84]	[90.23, 91.04]
	L3	[93.83, 72.41]	[92.15, 94.21]
	L4	[94.92, 86.34]	[94.73, 96.45]
Normal-only	L0	[83.50, 49.70]	[91.65, 64.72]
	L1	[86.27, 48.31]	[91.42, 66.05]
	L2	[89.38, 51.92]	[91.10, 69.37]
	L3	[92.41, 63.45]	[92.05, 78.92]
	L4	[96.76, 87.39]	[96.21, 93.95]
Mutual NN	L0	[76.45, 50.41]	[98.01, 52.10]
	L1	[82.63, 58.92]	[98.10, 60.33]
	L2	[88.47, 68.35]	[98.24, 71.47]
	L3	[93.12, 79.06]	[98.45, 82.63]
	L4	[97.26, 88.79]	[98.73, 91.82]
Full MGCP	L0	[71.73, 46.68]	[98.76, 47.48]
	L1	[81.04, 55.62]	[98.90, 64.92]
	L2	[88.90, 69.34]	[99.05, 79.41]
	L3	[94.21, 80.55]	[99.14, 88.21]
	L4	[99.16, 89.05]	[99.21, 94.69]

whose stability is uncertain. When stability propagation starts from such seeds, the region-growth direction becomes less reliable, and unstable areas may be incorrectly incorporated into the expanding stable region. This tends to reduce the stability of boundary transitions, reflected in a substantial drop in IoU of the changed class and consequently the mIoU at L1–L2, despite strong initial values. Moreover, the SGCA module does not immediately override these early deviations. Due to the Stability-Adaptive Propagation (SAP) mechanism, updates to stable regions are intentionally moderated to avoid disrupting reliably stable structures. This design choice makes stability propagation conservative, meaning that early-stage instability cannot be corrected in one step and instead requires multi-layer progressive refinement. This explains why Distance-only and Normal-only seeds show a typical "initially high $\rightarrow$ decline $\rightarrow$ recovery" pattern across levels. Mutual NN seeds, despite lower L0 IoU, provide higher stability purity. Stability growth therefore begins from a more trustworthy foundation, allowing SGCA to refine residual instability early and producing a smoother, almost monotonic improvement. Full MGCP yields the most stable seeds due to its joint distance, normal, and mutual-consistency constraints. Stability propagation thus proceeds along reliable structures from



(a) Layer-wise mIoU of different seeds.

(b) Layer-wise mAcc of different seeds.

Figure 6. Evolution of performance from L0 to L4 under different geometric seed definitions.

the beginning, producing consistently stronger refinement across all levels.

C. Additional Visualizations

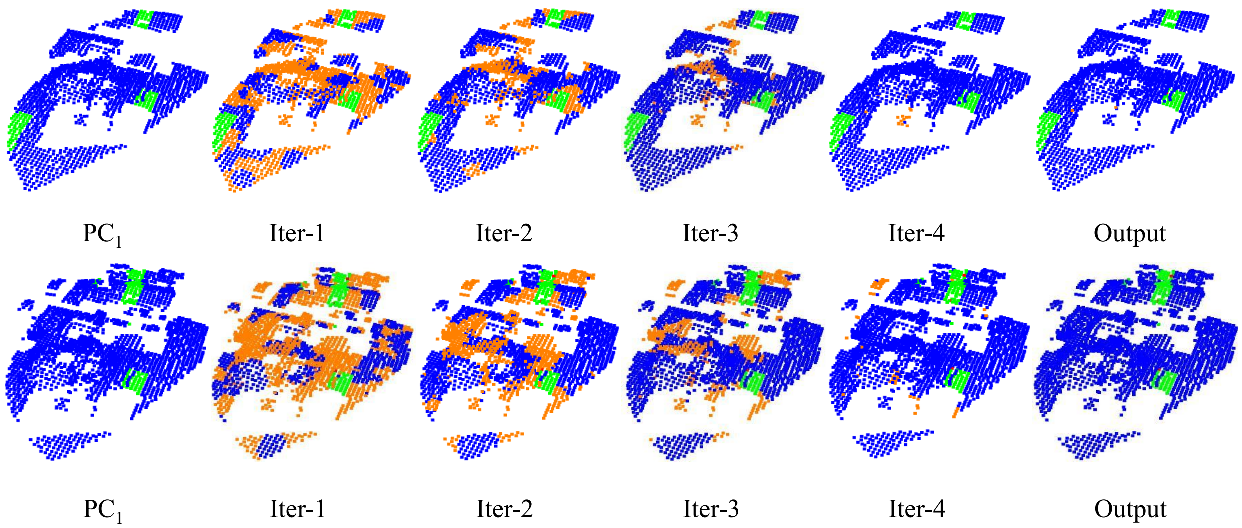
C.1. Multi-level Region-Growth Visualization

To further illustrate the internal behavior of SRGCD across different scene types, we provide additional visualizations of the layer-wise refinement process on both Urb3DCD-V1 and HKCD datasets. In the visualizations, Iter-k corresponds to the output of propagation layer L(k-1), i.e., Iter-1 shows the MGCP initialization at L0, and Iter-2 through Iter-4 show the SGCA refinements at L1-L3, with the final Output representing L4.

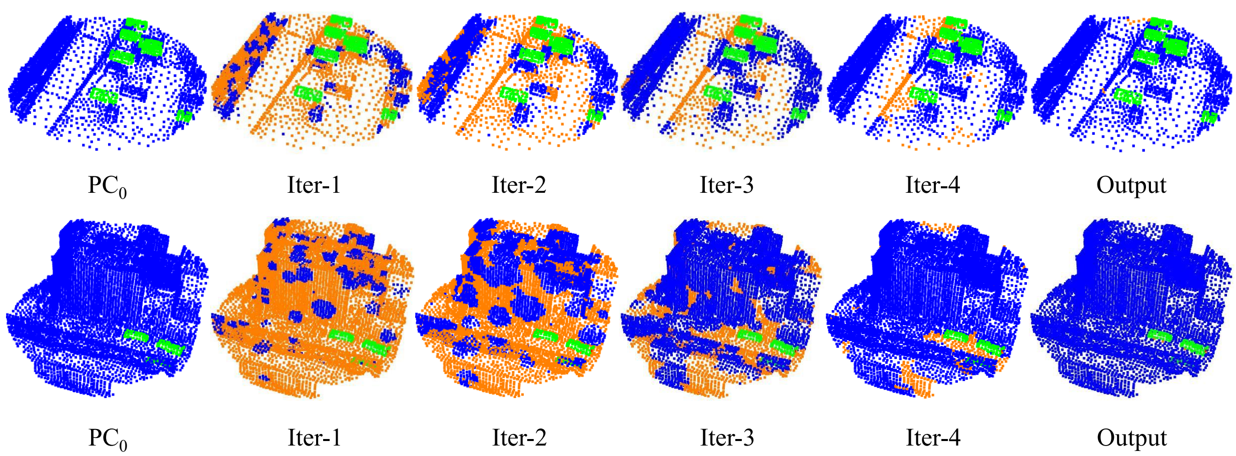
As shown in Fig. 7a, the Urb3DCD-V1 scenes contain complex residential structures with multiple buildings and intricate geometric details. At Iter-1, MGCP identifies only sparse unchanged seeds (blue points) due to strict geometric constraints, leaving most unchanged regions undetected and thus misclassified as changed (orange points). Through subsequent iterations, SGCA progressively propagates stability from these reliable seeds to neighboring uncertain regions. Notably, the unchanged regions expand in a spatially coherent manner—stability grows from building interiors toward boundaries rather than through scattered point-wise corrections. By Iter-4, nearly all unchanged structures are correctly recovered, with only minor residual errors concentrated at geometric discontinuities (e.g., occlusion boundaries).

Fig. 7b presents HKCD scenes with sparse changes, including moving vehicles and subtle building facade modifications. These cases are particularly challenging due to extreme class imbalance, where unchanged points vastly outnumber changed ones. Conventional segmentation methods often misclassify small changed regions as unchanged under such conditions. In contrast, SRGCD maintains the integrity of changed regions throughout all iterations: the vehicle region (green points) remains isolated at Iter-1 and is never invaded by expanding unchanged regions. This demonstrates that our controlled propagation mechanism does not aggressively expand unchanged seeds into truly changed areas, effectively mitigating the class imbalance issue.

Across all visualized scenes, a consistent pattern emerges: (i) MGCP initializes with sparse but highly reliable seeds, (ii) SGCA expands these seeds layer-by-layer in a spatially coherent manner, and (iii) residual errors in the final outputs are primarily located at geometrically ambiguous boundaries rather than randomly distributed. While some unchanged regions remain unrecovered (orange points), the dominance of correctly classified points ensures excellent overall performance. These visualizations validate the effectiveness of our stability-driven region growth mechanism across diverse scene complexities and data characteristics.



(a) Urb3DCD-V1 dataset



(b) HKCD dataset

Figure 7. Multi-level region-growth visualization on Urb3DCD-V1 and HKCD datasets. (a) Progressive stability propagation in complex residential scenes. (b) Controlled propagation preserving small changed regions under extreme class imbalance. Blue: correctly unchanged, green: correctly changed, red: missed changes, orange: false alarms.