Event-based HDR Structured Light

Jiacheng Fu Yue Li Xin Dong Wenming Weng Yueyi Zhang Zhiwei Xiong*
University of Science and Technology of China
jc_fu@mail.ustc.edu.cn zwxiong@ustc.edu.cn

Abstract

Event-based structured light (SL) systems have attracted increasing attention for their potential in high-performance 3D measurement. Despite the inherent HDR capability of event cameras, reflective and absorptive surfaces still cause event clutter and absence, which produce overexposed and underexposed regions that degrade the reconstruction quality. In this work, we present the first HDR 3D measurement framework specifically designed for event-based SL systems. First, we introduce a multi-contrast HDR coding strategy that facilitates imaging of areas with different reflectance. Second, to alleviate inter-frame interference caused by overexposed and underexposed areas, we propose a universal confidence-driven stereo matching strategy. Specifically, we estimate a confidence map as the fusion weight for features via an energy-guided confidence estimation. Further, we propose the confidence propagation volume, an innovative cost volume that offers both effective suppression of inter-frame interference and strong representation capability. Third, we contribute an event-based SL simulator and propose the first event-based HDR SL dataset. We also collect a real-world benchmarking dataset with ground truth. We validate the effectiveness of our method with the proposed confidence-driven strategy on both synthetic and real-world datasets. Experimental results demonstrate that our proposed HDR framework enables accurate 3D measurement even under extreme conditions. The code and data are available at https://github.com/Quma233/Event-based-HDR-SL.

1 Introduction

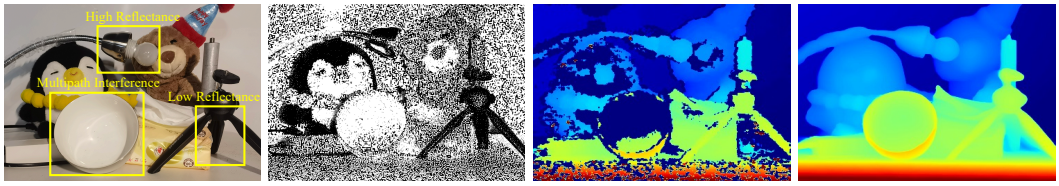
Structured light (SL) systems perform high-quality depth estimation by projecting predefined patterns onto a scene and analyzing their deformation upon interaction with object surfaces. Their high precision and dense reconstruction capabilities make them well-suited for applications in both consumer electronics and industrial applications [1, 2, 3]. However, reconstructing scenes under high dynamic range (HDR) conditions remains a significant challenge, as conventional cameras only capture the brightness levels with limited range. The target surfaces, such as rust, glossy paint, or carbon fiber, always lead to overexposure or underexposure of the projected patterns in captured images [4, 5], resulting in degraded or failed reconstruction. Existing solutions typically adopt multi-exposure fusion and multi-projection fusion techniques, but these approaches require extensive manual tuning of exposure settings and projection intensity to achieve satisfactory results [1].

Recently, event cameras [6], high-speed neuromorphic sensors with HDR capability, have been integrated into SL systems, demonstrating advantages in capturing high-speed dynamic scenes and maintaining imaging quality under complex illumination conditions. Previous work has shown that event-based SL systems offer improved HDR scene reconstruction compared to frame-based cameras, however, they still suffer from limitations of low reconstruction quality and insufficient robustness, especially in overexposed and underexposed areas [7, 8, 9]. Practical tests further indicate that,

*Corresponding author

39th Conference on Neural Information Processing Systems (NeurIPS 2025).

https://doi.org/10.52202/085713-1126



(a) HDR Scene (b) Event Frame with Speckle Coding (c) Disparity from (b) (d) Ours

Figure 1: Despite the inherent HDR capabilities of event cameras, they remain inadequate to handle HDR scenes with high reflectance variations (a), which cause event clutter and absence in the captured event frame (b). (c) shows the disparity estimated from (b) using SGBM, indicating the compromised performance of common event-based structured light systems in HDR environments. To address this, our HDR 3D measurement framework introduces a multi-contrast HDR coding strategy that employs range-partitioned sensing to image regions with different reflectance. As shown in (d), combined with our confidence-driven stereo matching strategy, our method demonstrates the ability to reconstruct high-reflectance surfaces (lamp), low-reflectance surfaces (black tripod), and regions affected by multipath interference (bowl), achieving superior reconstruction results even under extreme HDR conditions.

despite the inherent dynamic range advantages of event cameras, existing event-based SL systems remain inadequate for real-world HDR scene reconstruction, as illustrated in Fig. 1.

In this work, we pioneer the first HDR 3D measurement framework for event-based SL. First, we propose a multi-contrast HDR coding strategy that uses three pairs of predefined speckle patterns with different contrast and content for event triggering and depth encoding. Leveraging the HDR capability of event cameras, different contrast enables imaging across different reflection regions without tedious projection parameter tuning, while content-wise different speckles provide coding and imaging gains, maximizing the depth information encoded in three frames.

Second, to alleviate inter-frame interference caused by overexposed and underexposed regions in individual frames, we propose a universal confidence-driven stereo matching strategy. Specifically, it includes an energy-guided confidence estimation (ECE) module, which leverages the energy distribution of binary speckle patterns as a prior to estimate a confidence map for each of the three input frames as the fusion weight. Building upon this, we innovate the confidence propagation volume (CPV) — a confidence-driven cost volume construction method. During volume construction, confidence maps are propagated along the channel and disparity dimensions, and subsequently used to weight the left and right features. This enables accurate extraction of valid coding information across the three frames, leading to a cost volume with suppressed inter-frame interference and enhanced representation capability.

Third, for model training, we contribute a SL simulator based on Blender [10] and propose the first event-based HDR SL dataset. We also build a real-world benchmarking dataset with ground truth (GT) based on our built event-based SL system and a high-precision HDR SL scanner.

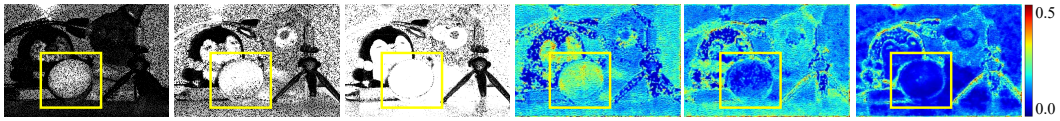
We integrate the confidence-driven stereo matching strategy into IGEV, a competing stereo network [11], to form our HDR 3D reconstruction method. Experiments on synthetic and real-world datasets demonstrate the superior performance of our method, which achieves robust 3D reconstruction even under extreme HDR conditions (Fig. 1)(d). Besides, we also demonstrate the universality of our confidence-driven strategy, which significantly enhances the accuracy of networks across diverse paradigms.

2 Related Work

2.1 Event-based Structured Light

Existing event-based SL studies primarily focus on exploring event-triggering mechanisms and corresponding algorithms to improve acquisition speed and reconstruction accuracy [7, 8, 9, 12, 13, 14, 15]. Leveraging the inherent HDR capability of event cameras, existing event-based methods offer certain advantages in HDR scene perception compared to frame-based cameras, yet they still struggle to handle complex HDR scenes. Matsuda *et al.* explored a line-scanning-based reconstruction method

https://doi.org/10.52202/085713-1126



(a) Event Frames (b) Corresponding Confidence Maps

Figure 2: Visualization of event frames derived by multi-contrast HDR coding and corresponding estimated confidence maps.

and tested it on a shiny steel sphere [7]. However, the reconstruction quality was limited by the constraints of early-generation event sensors. Muglikar *et al.* managed to reconstruct a plaster of David under varying lighting conditions [8], however, the target surface was diffuse and did not represent a typical HDR scenario, which limited the persuasiveness of the results. In [9], the proposed sinusoidal fringe-based encoding demonstrated ability in handling high-reflectance and high-contrast scenes, but it still suffered from insufficient robustness, limited resolution and unsatisfactory detail recovery. To address these challenges, we propose the first HDR 3D measurement framework tailored for event-based SL systems, enabling accurate and robust reconstruction even in extreme scenarios.

2.2 Cost Volume-based Stereo Matching

Current cost volume–based deep stereo methods fall into two categories. One is cost-filtering approaches, which build a dense 4D cost volume and aggregate matching costs via 3D CNN [16, 17, 18, 19]. These methods proposed various cost-volume paradigms: GC-Net introduces a simple yet effective concatenation volume with strong representational power [16]; GwcNet presents a group-wise correlation volume that combines the benefits of concatenation and correlation volume [18]. Recent attention-based variants ACVNet and FastACV [19] employ attention-weighted concatenation volumes to suppress redundancy while preserving discriminative matching cues. The other category is iterative-optimization approaches [20, 21]. This category is exemplified by RAFT-Stereo [20], which introduces a deep iterative architecture using a GRU-based updater to refine disparity by sampling cost values from an all-pairs correlation volume. More recently, IGEV-Stereo [11] combines these two paradigms and achieves impressive results. However, when naively applied in our HDR setting, existing approaches would suffer from inter-frame interference that undermines cost-volume expressiveness and degrades the reconstruction quality.

3 Event-based HDR Structured Light

3.1 Multi-Contrast HDR Coding

To perceive HDR scenes using sensors with limited dynamic range, it is necessary to adopt a range-partitioned sensing strategy. Considering that event cameras only respond to changes in illuminance, we propose a multi-contrast encoding strategy to enable the imaging of surfaces with different reflectance.

Specifically, we employ N pairs of binary speckle patterns [3] to trigger events, where each pair consists of two sequentially projected frames with distinct contrast and content. We define the white foreground speckles to constitute the foreground area and the rest to be the background area. In each pair, the first frame F_1 assigns a uniform intensity I_b to the background area and 0 to the foreground, while the second frame F_2 assigns I_f to the foreground and 0 to the background. Let $I(x, y, t - \Delta t)$ and $I(x, y, t)$ denote the observed intensities at pixel (x, y) when projecting F_1 and F_2, respectively, where $t - \Delta t$ and t are the timestamps corresponding to the two projections. An event is triggered when the intensity change between two projections exceeds triggering thresholds:

$$e(x,y,t) = \begin{cases} 1 & \text{if } \log\left(\dfrac{I(x,y,t)}{I(x,y,t-\Delta t)}\right) \geq C_1 \\ -1 & \text{if } \log\left(\dfrac{I(x,y,t)}{I(x,y,t-\Delta t)}\right) \leq -C_2 \end{cases}, \tag{1}$$

where C_1 and C_2 are triggering thresholds for positive and negative events, respectively.

This configuration ensures that the foreground speckle area exhibits a significant positive intensity change, thereby triggering events. In contrast, non-speckle background regions undergo a negative

https://doi.org/10.52202/085713-1126

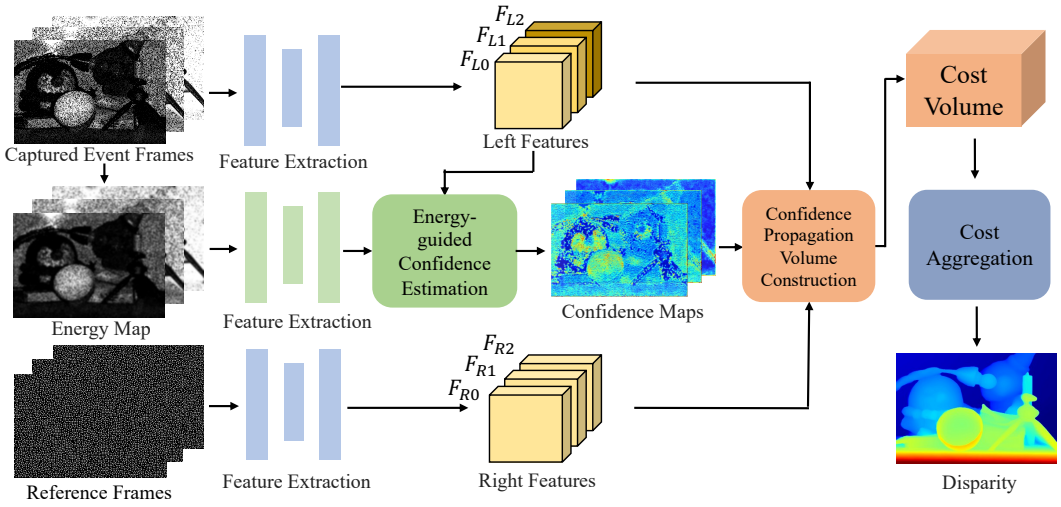


Figure 3: Pipeline of the Confidence-Driven Stereo Matching strategy.

change. This design aims to suppress positive intensity changes in non-speckle regions, thereby reducing unwanted event triggers that may arise from high reflectance. Note that, for bandwidth efficiency, we do not use negative events in our setting, which also ensures that no interfering negative events will be triggered. By adjusting I_b and I_f in each pair, we construct multiple contrast levels to enable imaging across regions with varying reflectance as shown in Fig. 2(a). Compared to frame-based methods that require extensive per-scene adjustment of projection settings, a key advantage of our strategy is that, leveraging the inherent HDR capability of event cameras, it can cover the dynamic range of most complex scenes using only three preset $[I_b, I_f]$. This significantly simplifies the cost of system deployment.

As for the content, we adopt N mutually independent speckle frames in each pair. As shown in [22], this strategy offers greater coding distinguishability than using one single pattern, which in turn improves reconstruction accuracy. Besides, projecting several mutually independent random-speckle frames is equivalent to taking multiple independent samples of the surface micro-facet normal distribution, which reduces the likelihood that all frames at a given pixel suffer from event clutter or absence. In practice, it is found that selecting $N = 3$ strikes an optimal balance between accuracy and efficiency (see the full supporting experiments in the supplement). This setting is used in all subsequent method designs and experiments.

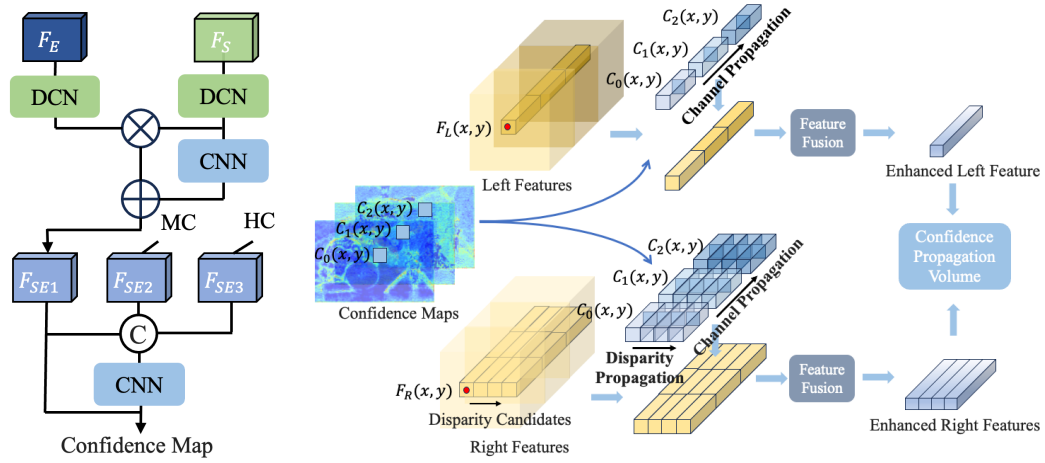
3.2 Confidence-Driven Stereo Matching

Our proposed confidence-driven stereo matching strategy is illustrated in Fig. 3. The inputs of the network are three rectified stereo pairs: the left frames are three event frames produced by the multi-contrast coding, and the right frames are corresponding projected reference frames. The Energy-Guided Confidence Estimation (ECE) processes each left image together with its energy map to predict a per-frame confidence map. These confidence maps, combined with the features from both views, are subsequently fused to construct the Confidence Propagation Volume (CPV) that underpins the final disparity-estimation stage.

3.2.1 Energy-Guided Confidence Estimation

Our design of the Energy-Guided Confidence Estimation (ECE) module is motivated by two key observations. On the one hand, as shown in Fig. 2(a), overexposed and underexposed areas carry no reliable coding information. When the three pairs are directly used to compute the cost as in [22], overexposed and underexposed regions not only fail to provide valid matching cues, but also contaminate the correct cost computed from well-exposed regions during multi-frame cost fusion, ultimately leading to reconstruction failure. We refer to this phenomenon as inter-frame interference. To alleviate this, low-quality areas should contribute less to cost fusion or to the construction of the matching feature, which calls for a frame-wise confidence estimation mechanism. On the other hand, low- and high-reflectance regions tend to exhibit event absence and event clutter, respectively,

https://doi.org/10.52202/085713-1126



(a) Energy-Guided Confidence Estimation

(b) Confidence Propagation Volume Construction

Figure 4: The overview of the proposed ECE (a) and CPV (b). (a) illustrates how the ECE module processes the low-contrast input. MC and HC stand for medium contrast and high contrast, respectively. F_{SE1}, F_{SE2}, and F_{SE3} are the fused features estimated from the low-, medium-, and high-contrast branches.

leading to overly low or high local energy responses. Moreover, our binary event frames contain only the projected speckles without background interference, making the local energy distribution a meaningful prior for speckle quality. This motivates us to estimate the confidence maps using both the speckle pattern and its corresponding energy map.

Based on these insights, we propose ECE, as shown in Fig. 4(a). For each left frame, features of the energy map F_E and the raw speckle data F_S are processed using deformable convolution (DCNv4), which adaptively adjusts sampling positions to accommodate the diverse and spatially irregular random speckles caused by extreme reflectance variations in HDR scenes. This flexibility improves the network's ability to assess speckle quality, leading to more accurate confidence estimation compared to conventional CNNs with fixed sampling grids. Next, we modulate the energy map features through element-wise multiplication to enhance the raw speckle features, resulting in the fused feature F_{SE1}. This operation leverages the complementary nature of the two inputs, where the raw speckle provides fine-grained pixel-level details and the energy map offers noise-robust confidence priors. Finally, since the confidence reflects the relative quality among the three frames, the confidence estimation for each frame is made jointly based on all three fused features through a residual structure. We visualize the learned confidence maps, as shown in Fig. 2(b). The visualizations show a strong correspondence with the raw input data, demonstrating clear interpretability: regions suffering from overexposure or underexposure exhibit low confidence, while well-exposed areas yield high confidence values.

3.2.2 Confidence Propagation Volume Construction

Leveraging the estimated confidence maps, we propose a cost-volume construction strategy as shown in Fig. 4(b). For each spatial location (x, y), the scalar confidences $\{C_0(x, y), C_1(x, y), C_2(x, y)\}$ of the three input frames are individually expanded along the channel dimension to match the channel number of each frame's feature. These expanded vectors are then concatenated to form the fusion weight vector $\mathbf{C}(x, y)$. The confidence-weighted feature is computed by element-wise multiplication with the corresponding frame features:

$$\tilde{\mathbf{F}}_L(x, y) = \mathbf{F}_L(x, y) \odot \mathbf{C}(x, y).$$

This process weights the features of each frame according to the quality of the projected speckle. In this way, the resulting feature can minimize interference caused by suboptimal imaging regions.

Although the right features are extracted from reference frames that are free from under- or over-exposure, it is still necessary to impose fusion weights on them in order to maintain consistency

https://doi.org/10.52202/085713-1126

with the confidence-weighted left features during disparity searching. Therefore, during the disparity search of $\mathbf{F}_L(x, y)$, we propagate the fusion weight $C(x, y)$ along the disparity dimension to weight the right features within the disparity searching range:

$$\tilde{\mathbf{F}}_R(d; x, y) = C(x, y) \odot \mathbf{F}_R(x{+}d, y),$$

where $\mathbf{F}_R(x{+}d, y)$ denotes the right feature sampled at $(x{+}d, y)$, and $\mathbf{F}_R(d;\, x, y)$ denotes the corresponding confidence-weighted right feature. This ensures that both left and right features are modulated by the same fusion weights, enabling consistent and confidence-aware stereo matching.

The confidence-weighted features from both views are further refined by two independent feature fusion modules, each consisting of 1×1 convolutions, to reorganize information along the channel dimension, enhancing feature representation ability while reducing redundancy. Finally, the enhanced left feature is matched with the right features in disparity searching range to assemble the CPV. By integrating the proposed confidence-driven strategy into IGEV [11], we derive our final method for HDR 3D reconstruction.

4 Experiment

4.1 Dataset and Implementation Details

Synthetic Dataset. To support supervised training, we design an event-based SL simulator in Blender [10] and propose the first event-based HDR SL dataset. Using the calibrated intrinsic and extrinsic parameters from our real SL system, we construct its digital twin in Blender. Since Blender lacks a native event-camera sensor model, we render two RGB frames, one with the projected speckle pattern and one without, and take their pixel-wise difference to approximate the imaging process of the event camera. In addition, a physically based projector is implemented using Blender's node-tree system, allowing its optical properties to be precisely configured by the intrinsics. HDR scenes are built from high-fidelity, textured meshes sourced from the OmniObject3D dataset [23]. To simulate the reflectance behavior of real-world HDR surfaces, we vary each object's metallicity, micro-facet roughness and self-emission coefficient. The resulting scenes contain mirror-like specular peaks alongside low-reflectance regions, offering realistic dynamic range diversity. High-precision scene depth is also rendered by Blender for supervision. Samples are shown in the supplement. We render a total of 8,550 image pairs at a resolution of 1280×704. The dataset is split into 7,500 training samples and 1,050 test samples.

Real-World Dataset. We build a monocular event-based SL system using the Prophesee EVK4 event camera, which offers a resolution of 1280×720 and a dynamic range exceeding 120 dB [24], and the DLP6500 projector from Texas Instruments [25]. In addition, we use an HDR structured light scanner [26] to capture GT point cloud. A total of 15 complex HDR scenes, including highly reflective, absorptive, transparent, and fine-grained objects, are captured. Each sample contains three event frames generated by our multi-contrast coding and corresponding GT point cloud.

Implementation Details. For multi-contrast HDR coding, the projection intensities of the three pairs are set to [32,55], [32,200], and [0,255], respectively. The depth sensing rate of our system is 200 Hz. Our HDR 3D reconstruction method is implemented in PyTorch, trained on a synthetic dataset, and evaluated on both synthetic and real-world data. The disparity search range is set to 256. All experiments are conducted on the NVIDIA 3090 GPUs.

4.2 Baselines and Evaluation Metrics

For comparison, we evaluate our method against several baselines. They include two traditional stereo algorithms, Block Matching (BM) [27] and Semi-Global Block Matching (SGBM) [28], two monocular structured-light depth estimation methods, CTD [29] and GigaDepth [30], as well as three representative cost volume-based stereo networks, RAFT-Stereo [20], FastACV [19], and IGEV [11]. In our setting with three input frames, BM and SGBM are applied independently to each frame, and the final disparity map is obtained by averaging the disparities of high-confidence regions. For the deep learning-based methods, the three frames are concatenated along the channel dimension and jointly fed into the network. We evaluate disparity quality with three metrics: end-point error EPE (mean absolute pixel error), Bad τ rates for $\tau \in \{0.5, 1, 2, 3, 5\}$ px (fraction of pixels whose error exceeds τ), and D1 (pixels whose error is > 3 px and $> 5\%$ of GT). Together, they summarize average accuracy, thresholded error rates, and gross outliers.

Methods	EPE ↓	Bad 0.5 ↓	Bad 1.0 ↓	Bad 2.0 ↓	Bad 3.0 ↓	Bad 5.0 ↓	D1 ↓	Time(s)
BM	31.4518	0.3919	0.3213	0.3156	0.3120	0.3053	0.3112	**0.0030**
BM-3	18.2216	0.3141	0.2453	0.2401	0.2349	0.2274	0.2302	0.0120
SGBM	13.3384	0.2697	0.1563	0.1421	0.1390	0.1359	0.1363	0.0060
SGBM-3	6.3327	0.2243	0.1082	0.0918	0.0874	0.0830	0.0831	0.0210
CTD	26.8375	0.3341	0.2755	0.2638	0.2598	0.2465	0.2489	0.0190
GigaDepth	8.6688	0.1425	0.1210	0.1135	0.1111	0.1082	0.1070	0.0210
FastACV	0.7112	0.1068	0.0600	0.0342	0.0243	0.0159	0.0154	0.0670
RAFT-Stereo	0.4136	0.1286	0.0507	0.0219	0.0136	0.0077	0.0077	0.8490
IGEV	0.3863	0.0548	0.0311	0.0177	0.0127	0.0084	0.0076	0.4200
Ours	**0.2937**	**0.0359**	**0.0203**	**0.0122**	**0.0090**	**0.0063**	**0.0062**	0.4080

Table 1: Quantitative results on the synthetic dataset. All metrics are error rates (↓ lower is better). Only BM and SGBM are single-frame methods, others are 3-frame methods.

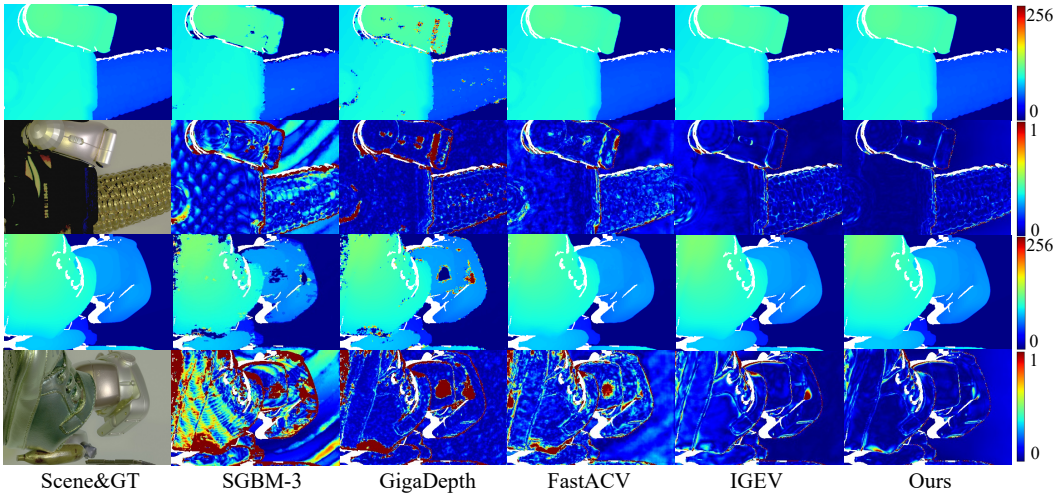


Figure 5: Qualitative results on the synthetic dataset. Two samples are shown, with the first row showing the disparities and the second row showing the scene and the error maps.

4.3 Synthetic Results

The results in Table 1 clearly demonstrate the superiority of our proposed 3D measurement framework, significantly outperforming existing event-based monocular and stereo methods. The performance of BM and SGBM reflects the typical limitations of conventional event-based structured light systems in HDR scenes, where large reconstruction errors are observed. Their three-frame variants (BM-3, SGBM-3) exhibit noticeable improvements, which validates the effectiveness of our proposed multi-contrast HDR coding. Among all baselines, deep learning-based stereo networks (RAFT-Stereo, FastACV, IGEV) achieve markedly better performance. In comparison, our method achieves further improvements over the strongest baseline, IGEV, demonstrating the effectiveness of the proposed confidence-driven stereo matching strategy.

Qualitative results are shown in Fig. 5. As observed, traditional methods and monocular estimation networks produce large areas of mismatches and perform poorly near object boundaries. While deep learning-based stereo methods partially solve these issues, however, as indicated in the error map, they still exhibit noticeable reconstruction errors in regions with specular reflections or fine structures. In contrast, our method achieves superior reconstruction quality in both general and HDR-challenging regions, delivering more accurate disparity estimates. See more qualitative results in the supplement.

4.4 Ablation Study

To validate the effectiveness of each component in our framework, we conduct an ablation study based on our method and present quantitative results in Table 2.

Content Variation in Multi-Contrast HDR Coding. The first two variants (#A and #B) compare the impact of different 3-frame inputs, with and without speckle content variation. As can be seen,

https://doi.org/10.52202/085713-1126

Table 2: Ablation study on the content variation in multi-contrast HDR coding (Content), confidence propagation volume (CPV), energy prior (EnP), and the energy-guided confidence estimation module (ECE).

Variants	Content	CPV	EnP	ECE	**EPE ↓**	**Bad 0.5 ↓**	**Bad 1.0 ↓**	**Bad 2.0 ↓**	**Bad 3.0 ↓**	**Bad 5.0 ↓**	**D1 ↓**
#A	✗	✗	✗	✗	0.5915	0.0776	0.0507	0.0324	0.0221	0.0141	0.0145
#B	✓	✗	✗	✗	0.3863	0.0548	0.0311	0.0177	0.0127	0.0084	0.0076
#C	✓	✓	✗	✗	0.3433	0.0415	0.0242	0.0145	0.0111	0.0077	0.0073
#D	✓	✓	✓	✗	0.3215	0.0376	0.0218	0.0131	0.0097	0.0067	0.0065
#E	✓	✓	✓	✓	**0.2937**	**0.0359**	**0.0203**	**0.0122**	**0.0090**	**0.0063**	**0.0062**

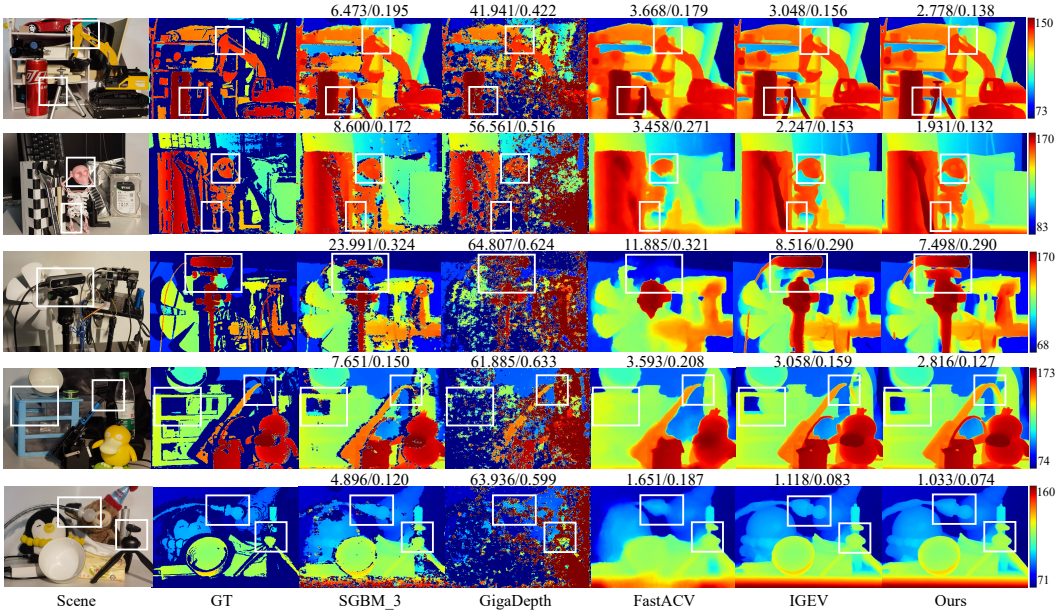


Figure 6: Qualitative and quantitative results of real-world HDR scenes. The numbers above each result denote the EPE and Bad 2.0 error for the corresponding scene.

applying our proposed content-varying coding reduces Bad 1.0 by 38.7%, indicating that introducing content diversity enhances the distinguishability of speckle coding and enables more accurate local matching. Meanwhile, Bad 5.0 decreases by 47.6%, suggesting that our method not only brings imaging gains in challenging HDR regions but also provides richer cues for disparity completion.

Confidence Propagation Volume. For variant #C, we substitute the ECE with convolutions and cancel the use of the energy prior, resulting in a compromised confidence estimation. Nevertheless, comparing the results of variants #B and #C, it can be found that applying CPV can still effectively suppress the inter-frame interference and provide better results.

Energy-guided Confidence Estimation. We add the energy prior (EnP) to variant #C to produce variant #D. The results are improved, demonstrating the effectiveness of energy prior in confidence estimation. Finally, we substitute convolutions with the proposed ECE and achieve the best result in variant #E. It can be deduced that the high-quality confidence map estimation provides better fusion weights for multi-frame information, which maximizes the suppression of the interference and enhances the representation capability.

4.5 Real-World Results

To evaluate the generalizability of our method, we test it on real-world HDR scenes captured by our custom-built SL system. For qualitative and quantitative comparison, we first convert the estimated disparity maps into depth and reconstruct corresponding 3D point clouds. The GT point cloud captured by the Photoneo MotionCam-3D M+ is then aligned to each predicted point cloud using

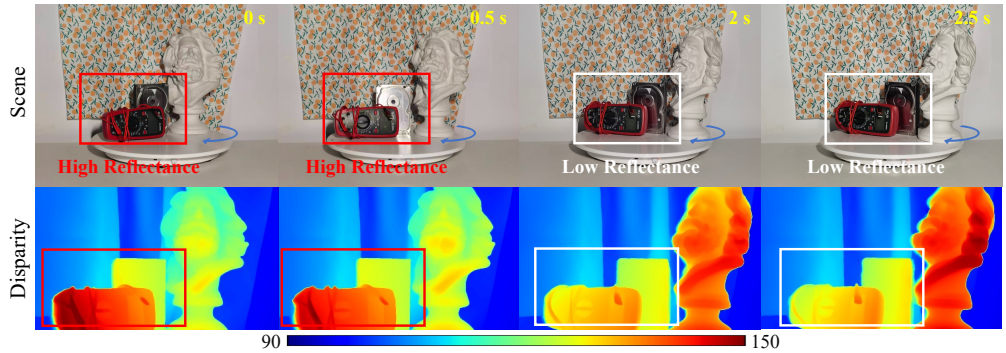


Figure 7: Reconstruction results of the proposed method in a challenging dynamic HDR scene.

Methods	EPE ↓	Bad 0.5 ↓	Bad 1.0 ↓	Bad 2.0 ↓	Bad 3.0 ↓	Bad 5.0 ↓	D1 ↓
FastACV	0.7112	0.1068	0.0600	0.0342	0.0243	0.0159	0.0154
Ours(FastACV)	0.4983	0.0753	0.0475	0.0261	0.0177	0.0113	0.0112
RAFT-Stereo	0.4136	0.1286	0.0507	0.0219	0.0136	0.0077	0.0077
Ours(RAFT-Stereo)	0.3248	0.0717	0.0324	0.0170	0.0117	0.0070	0.0069

Table 3: Quantitative evaluation of the universality of our confidence-driven stereo matching strategy on the synthetic dataset. Ours(*) indicates our strategy with * as the baseline.

iterative closest point (ICP) registration. Finally, the aligned GT point cloud is projected onto the image plane to generate the disparity maps used for evaluation. The qualitative and quantitative results are presented in Fig. 6. Among traditional methods, SGBM-3 can recover coarse disparity but suffers from large mismatched regions. The monocular method GigaDepth performs even worse, yielding noisy and fragmented outputs, indicating inferior robustness compared to stereo-based methods. Deep learning-based stereo networks generate more complete disparity maps. However, both FastACV and IGEV fail to reconstruct fine-scale structures. For instance, in the second row, neither method accurately recovers the legs and the left eye of the small skeleton figure, resulting in missing or overly smoothed geometry. In contrast, our method successfully reconstructs these fine-grained details, yielding a complete and structurally consistent skeleton. Our method also demonstrates remarkable disparity inference and completion capabilities. As shown in the third row, our method manages to reconstruct a black metallic object with extremely weak reflectance, where all other methods fail due to poor speckle quality. Moreover, our method shows the capability to reconstruct transparent objects, as shown in the fourth row, highlighting its strong generalization ability and robustness to non-Lambertian surfaces.

The quantitative results are consistent with the qualitative observations, confirming the superiority of our methods. However, the improvements are less pronounced than those on the synthetic dataset. This is primarily due to the sparsity and incompleteness of the real-world ground-truth data, particularly in extreme HDR regions where even the GT fails to make an accurate reconstruction.

We further evaluate our method under dynamic HDR conditions. To this end, we construct a dynamic HDR scene by placing three geometrically complex objects with diverse surface reflectance properties on a rotating platform with a controllable angular velocity. The captured dynamic scenes and their corresponding reconstructions are shown in Fig. 7. The results demonstrate that our method achieves high-quality reconstruction with high inter-frame consistency. Notably, for metallic surfaces whose reflected intensity varies drastically—from highly saturated to extremely weak—as the platform rotates, our approach consistently produces accurate and stable reconstructions, clearly showcasing its robustness in dynamic HDR environments.

4.6 Universality of Confidence-Driven Stereo Matching

We further validate the universality of our confidence-driven stereo matching strategy. Quantitative results are shown in Table 3. Specifically, we observe a reduction in EPE from 0.7112 to 0.4983 on FastACV, and from 0.4136 to 0.3248 on RAFT-Stereo, corresponding to performance gains of 29.9% and 21.5%, respectively. Similar improvements are consistently observed across all Bad τ rates and

https://doi.org/10.52202/085713-1126

D1 metric. It has been proved that incorporating our strategy can significantly improve the accuracy of stereo matching across models with different paradigms. See also the qualitative results in the supplement.

5 Conclusion and Discussion

Conclusion. In this work, we innovate the first framework for event-based HDR structured light. The proposed event-based HDR coding scheme encodes depth information using three speckle projections with variations in both intensity and content, enabling robust representation across challenging HDR scenes. Furthermore, the proposed confidence-driven stereo matching strategy fully leverages the multi-frame encoding by employing ECE and CPV to suppress inter-frame interference and enhance the representation ability of the cost volume, leading to highly accurate disparity estimation. We validate the effectiveness of our method on both synthetic and real-world datasets. Our method significantly outperforms existing baselines and demonstrates robust reconstruction performance even under extreme HDR conditions.

Limitation. Our method does not explicitly address occlusions. There exist no speckles in occluded regions, leaving the disparity estimation of these regions ill-posed. Due to the network's inherent disparity completion behavior, this often results in over-smoothed or expanded boundaries near the left edges in large-disparity scenes. A potential solution is to estimate a confidence score for the predicted disparity and filter out unreliable estimates based on this confidence, which we leave as future work.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China under Grants 62131003 and 62021001, and the Fundamental Research Funds for the Central Universities under Grant WK2100000059.

References

[1] S. Zhang. High-speed 3d shape measurement with structured light methods: A review. *Optics and Lasers in Engineering*, 106:119–131, 2018.

[2] Zhiwei Xiong, Yueyi Zhang, Feng Wu, and Wenjun Zeng. Computational depth sensing: Toward high-performance commodity depth cameras. *IEEE Signal Processing Magazine*, 34(3):55–68, 2017.

[3] Y. Zhang, Z. Xiong, Z. Yang, and F. Wu. Real-time scalable depth sensing with hybrid structured light illumination. *IEEE Transactions on Image Processing*, 23(1):97–109, 2013.

[4] Pei Zhou, Yue Cheng, Jiangping Zhu, and Jialing Hu. High-dynamic-range 3-d shape measurement with adaptive speckle projection through segmentation-based mapping. *IEEE Transactions on Instrumentation and Measurement*, 72:1–12, 2022.

[5] Ke Wu, Jie Tan, Hai Lun Xia, and Cheng Bao Liu. An exposure fusion-based structured light approach for the 3d measurement of a specular surface. *IEEE Sensors Journal*, 21(5): 6314–6324, 2020.

[6] Guillermo Gallego, Tobi Delbrück, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew J Davison, Jörg Conradt, Kostas Daniilidis, et al. Event-based vision: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 44(1):154–180, 2020.

[7] N. Matsuda, O. Cossairt, and M. Gupta. Mc3d: Motion contrast 3d scanning. In *2015 IEEE International Conference on Computational Photography (ICCP)*, pages 1–10. IEEE, 2015.

[8] M. Muglikar, G. Gallego, and D. Scaramuzza. Esl: Event-based structured light. In *2021 International Conference on 3D Vision (3DV)*, pages 1165–1174. IEEE, 2021.

[9] Yuhui Li, Heng Jiang, Chen Xu, and Lilin Liu. Event-driven fringe projection structured light 3-d reconstruction based on time–frequency analysis. *IEEE Sensors Journal*, 24(4):5097–5106, 2024.

[10] Blender Online Community. Blender - a 3d modelling and rendering package. `https://www.blender.org/`, 2023. Version 3.x.

[11] Gangwei Xu, Xianqi Wang, Xiaohuan Ding, and Xin Yang. Iterative geometry encoding volume for stereo matching. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 21919–21928, 2023.

[12] X. Huang, Y. Zhang, and Z. Xiong. High-speed structured light based 3d scanning using an event camera. *Optics Express*, 29(22):35864–35876, 2021.

[13] Jiacheng Fu, Yueyi Zhang, Yue Li, Jiacheng Li, and Zhiwei Xiong. Fast 3d reconstruction via event-based structured light with spatio-temporal coding. *Optics Express*, 31(26):44588–44602, 2023.

[14] Xingyu Lu, Lei Sun, Diyang Gu, and Kaiwei Wang. Sge: structured light system based on gray code with an event camera. *Optics Express*, 32(26):46044–46061, 2024.

[15] Jiacheng Fu, Yue Li, Wenming Weng, Xin Dong, Yueyi Zhang, and Zhiwei Xiong. Boosting event-based structured light via cyclic spatial coding. *Optics Letters*, 50(20):6365–6368, 2025.

[16] Alex Kendall, Hayk Martirosyan, Saumitro Dasgupta, Peter Henry, Ryan Kennedy, Abraham Bachrach, and Adam Bry. End-to-end learning of geometry and context for deep stereo regression. In *Proceedings of the IEEE international conference on computer vision*, pages 66–75, 2017.

[17] Jia-Ren Chang and Yong-Sheng Chen. Pyramid stereo matching network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5410–5418, 2018.

[18] Xiaoyang Guo, Kai Yang, Wukui Yang, Xiaogang Wang, and Hongsheng Li. Group-wise correlation stereo network. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3273–3282, 2019.

[19] Gangwei Xu, Yun Wang, Junda Cheng, Jinhui Tang, and Xin Yang. Accurate and efficient stereo matching via attention concatenation volume. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(4):2461–2474, 2023.

[20] Lahav Lipson, Zachary Teed, and Jia Deng. Raft-stereo: Multilevel recurrent field transforms for stereo matching. In *2021 International Conference on 3D Vision (3DV)*, pages 218–227. IEEE, 2021.

[21] Jiankun Li, Peisen Wang, Pengfei Xiong, Tao Cai, Ziwei Yan, Lei Yang, Jiangyu Liu, Haoqiang Fan, and Shuaicheng Liu. Practical stereo matching via cascaded recurrent network with adaptive correlation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 16263–16272, 2022.

[22] Q. Tang, C. Liu, Z. Cai, H. Zhao, X. Liu, and X. Peng. An improved spatiotemporal correlation method for high-accuracy random speckle 3d reconstruction. *Optics and Lasers in Engineering*, 110:54–62, 2018.

[23] Tong Wu, Jiarui Zhang, Xiao Fu, Yuxin Wang, Jiawei Ren, Liang Pan, Wayne Wu, Lei Yang, Jiaqi Wang, Chen Qian, et al. Omniobject3d: Large-vocabulary 3d object dataset for realistic perception, reconstruction and generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 803–814, 2023.

[24] Prophesee. *Event Camera Evaluation Kit 4*, 2024. URL `https://www.prophesee.ai/event-camera-evk4/`. Accessed: Sep. 12, 2024.

[25] Texas Instruments. *DLP6500 Digital Micromirror Device (DMD)*, 2023. URL `https://www.mouser.com/new/texas-instruments/ti-dlp6500-dmd/`. Accessed: Sep. 12, 2024.

 https://doi.org/10.52202/085713-1126

[26] Photoneo. Motioncam-3d m+. https://www.photoneo.com/products/motioncam-3d-m-plus/, 2024. Accessed: 2025-05-13.

[27] Gary Bradski. Opencv: Open source computer vision library. https://opencv.org/, 2000. Version 4.x.

[28] H. Hirschmuller. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):328–341, 2007.

[29] Gernot Riegler, Yiyi Liao, Simon Donne, Vladlen Koltun, and Andreas Geiger. Connecting the dots: Learning representations for active monocular depth estimation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 7624–7633, 2019.

[30] Simon Schreiberhuber, Jean-Baptiste Weibel, Timothy Patten, and Markus Vincze. Gigadepth: Learning depth from structured light with branching neural networks. In *European Conference on Computer Vision*, pages 214–229. Springer, 2022.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: See the abstract and introduction.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: See the limitation section.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [NA]

https://doi.org/10.52202/085713-1126

Justification: No theoretical result.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: See implementation details.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [No]

Justification: We will public them if the manuscript is accepted.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [Yes]

   Justification: See implementation details.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: See tables.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

https://doi.org/10.52202/085713-1126

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: See implementation details.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: We follow the ethics.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [NA]

    Justification: There is no societal impact of the work.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to

generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.

- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

    Justification: The paper poses no such risks.

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [NA]

    Justification: The paper does not use existing assets.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

 https://doi.org/10.52202/085713-1126

Answer: [NA]

Justification: Will be released once accepted.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: The paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: The paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: The core method development in this research does not involve LLMs as any important components.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.

https://doi.org/10.52202/085713-1126