Novelty Assessment Report

Paper: StreamSplat: Towards Online Dynamic 3D Reconstruction from Uncalibrated Video Streams
PDF URL: https://openreview.net/pdf?id=SaiDRQU7Ez
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2025-12-29

Abstract

Real-time reconstruction of dynamic 3D scenes from uncalibrated video streams demands robust online methods that recover scene dynamics from sparse observations under strict latency and memory constraints. Yet most dynamic reconstruction methods rely on hours of per-scene optimization under full-sequence access, limiting practical deployment. In this work, we introduce **StreamSplat**, a fully feed-forward framework that instantly transforms uncalibrated video streams of arbitrary length into dynamic 3D Gaussian Splatting (3DGS) representations in an online manner. It is achieved via three key technical innovations: 1) a probabilistic sampling mechanism that robustly predicts 3D Gaussians from uncalibrated inputs; 2) a bidirectional deformation field that yields reliable associations across frames and mitigates long-term error accumulation; 3) an adaptive Gaussian fusion operation that propagates persistent Gaussians while handling emerging and vanishing ones. Extensive experiments on standard dynamic and static benchmarks demonstrate that StreamSplat achieves state-of-the-art reconstruction quality and dynamic scene modeling. Uniquely, our method supports the online reconstruction of arbitrarily long video streams with a $1200\times$ speedup over optimization-based methods. Our code and models will be made publicly available.

Core Task Landscape

This paper addresses: **Online Dynamic 3D Reconstruction from Uncalibrated Video Streams**

A total of **50 papers** were analyzed and organized into a taxonomy with **17 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Feed-Forward Dynamic Scene Reconstruction**
- **Optimization-Based Dynamic Reconstruction**
- **Incremental and Online Reconstruction**
- **Static Human Reconstruction from Monocular Video**
- **Sparse Feature-Based Reconstruction**
- **Depth-Guided Neural Reconstruction**
- **Specialized Application Reconstruction**
- **Motion and Dynamics Estimation**
- **Uncalibrated Reconstruction Techniques**

Complete Taxonomy Tree

- Online Dynamic 3D Reconstruction from Uncalibrated Video Streams Survey Taxonomy
- Feed-Forward Dynamic Scene Reconstruction
  - Real-Time Dynamic Gaussian Splatting ★ (5 papers)
  - [0] StreamSplat: Towards Online Dynamic 3D Reconstruction from Uncalibrated Video Streams (Anon et al., 2026) View paper
  - [4] DGS-LRM: Real-Time Deformable 3D Gaussian Reconstruction From Monocular Videos (Lin, 2025) View paper
  - [5] Feed-Forward Bullet-Time Reconstruction of Dynamic Scenes from Monocular Videos (Liang, 2024) View paper
  - [22] SplineGS: Robust Motion-Adaptive Spline for Real-Time Dynamic 3D Gaussians from Monocular Video (Jongmin Park, 2024) View paper
  - [24] QUEEN: QUantized Efficient ENcoding of Dynamic Gaussians for Streaming Free-viewpoint Videos (Girish, 2024) View paper
  - Generative Model-Based 3D Reconstruction (3 papers)
  - [1] Lyra: Generative 3D Scene Reconstruction via Video Diffusion Model Self-Distillation (Bahmani, 2025) View paper
  - [10] Consistent4D: Consistent 360Â° Dynamic Object Generation from Monocular Video (Jiang Yanqin, 2023) View paper
  - [50] PDF-MPI: From Monocular Video to Volumetric Video (Ling-Dong Wang, 2025) View paper
  - Multi-Human 4D Reconstruction (2 papers)
  - [15] Forge4D: Feed-Forward 4D Human Reconstruction and Interpolation from Uncalibrated Sparse-view Videos (Hu Ying-dong, 2025) View paper
  - [20] Humans as a Calibration Pattern: Dynamic 3D Scene Reconstruction from Unsynchronized and Uncalibrated Videos (Choi, 2024) View paper
- Optimization-Based Dynamic Reconstruction
  - Deformable Object Reconstruction (2 papers)
  - [13] PAD3R: Pose-Aware Dynamic 3D Reconstruction from Casual Videos (Liao, 2025) View paper
  - [48] REC-MV: REconstructing 3D Dynamic Cloth from Monocular Videos (Lingteng Qiu, 2023) View paper
  - Unsynchronized Multi-View Reconstruction (2 papers)
  - [2] Sparse dynamic 3d reconstruction from unsynchronized videos (Enliang Zheng, 2015) View paper
  - [11] Visual-hull reconstruction from uncalibrated and unsynchronized video streams (Sudipta N. Sinha, 2004) View paper

- Incremental and Online Reconstruction
  - Dense Volumetric Reconstruction (4 papers)
  - [12] Incremental Dense Reconstruction From Monocular Video With Guided Sparse Feature Volume Fusion (Xingxing Zuo, 2023) View paper
  - [21] Flora: Dual-Frequency LOss-Compensated ReAl-Time Monocular 3D Video Reconstruction (Yue Gong, 2023) View paper
  - [23] MonoFusion: Real-time 3D reconstruction of small scenes with a single web camera (Vivek Pradeep, 2013) View paper
  - [49] SST: Real-time End-to-end Monocular 3D Reconstruction via Sparse Spatial-Temporal Guidance (Chenyangguang Zhang, 2023) View paper
  - Online Human-Scene Reconstruction (2 papers)
  - [6] MegaSaM: Accurate, fast and robust structure and motion from casual dynamic videos (Zhengqi Li, 2025) View paper
  - [25] ODHSR: Online Dense 3D Reconstruction of Humans and Scenes from Monocular Videos (Zetong Zhang, 2025) View paper
  - Dynamic Pose Reconstruction (2 papers)
  - [32] Cross-Dimensional Refined Learning for Real-Time 3D Visual Perception from Monocular Video (Ziyang Hong, 2023) View paper
  - [47] LivePose: Online 3D Reconstruction from Monocular Video with Dynamic Camera Poses (Noah Stier, 2023) View paper
- Static Human Reconstruction from Monocular Video (9 papers)
  - [3] Real-time 3D pose reconstruction of human body from monocular video sequences (Liang-Jia Zhu, 2009) View paper
  - [8] Real-Time Reconstruction of Multi-Body Pedestrian Pre-Impact Posture in Collision Accidents From Monocular Images (Meijun Wang, 2025) View paper
  - [9] Link to the Past: Temporal Propagation for Fast 3D Human Reconstruction from Monocular Video (Matthew Marchellus, 2025) View paper
  - [14] Realtime dynamic 3D facial reconstruction for monocular video in-the-wild (Shuang Liu, 2017) View paper
  - [29] MUC: Mixture of Uncalibrated Cameras for Robust 3D Human Body Reconstruction (Zhu, 2024) View paper
  - [30] Inferring 3D body pose from uncalibrated video (Xianjie Qiu, 2005) View paper
  - [36] Model-based human gait tracking, 3D reconstruction and recognition in uncalibrated monocular video (Ehsan Adeli, 2012) View paper
  - [37] Real-time 3D morphable shape model fitting to monocular in-the-wild videos (Patrik Huber, 2017) View paper
  - [45] 3d reconstruction of human motion and skeleton from uncalibrated monocular video (Yen-Lin Chen, 2009) View paper
- Sparse Feature-Based Reconstruction (5 papers)
  - [7] Scene coordinate reconstruction: Posing of image collections via incremental learning of a relocalizer (Brachmann, 2024) View paper
  - [18] A fast 3D scene reconstructing method using continuous video (Bo-Yi Sung, 2017) View paper
  - [28] Projective reconstruction and metric models from uncalibrated video sequences (Oram, 2001) View paper
  - [33] 3d reconstruction of architectural scenes from uncalibrated video sequences (Frahm Jan-Michael, 2009) View paper
  - [46] 3D reconstruction system based on incremental structure from motion using a camera with varying parameters (Soulaiman El hazzat, 2018) View paper
- Depth-Guided Neural Reconstruction (2 papers)
  - [27] DG-Recon: Depth-Guided Neural 3D Scene Reconstruction (Jihong Ju, 2023) View paper
  - [38] Recollection from Pensieve: Novel View Synthesis via Learning from Uncalibrated Videos (Wang Ruoyu, 2025) View paper
- Specialized Application Reconstruction
  - Face and Facial Reconstruction (2 papers)
  - [31] 3D face pose tracking from an uncalibrated monocular camera (Zhiwei Zhu, 2004) View paper
  - [40] BabyNet: Reconstructing 3D faces of babies from uncalibrated photographs (Araceli Morales, 2023) View paper
  - Medical and Laparoscopic Reconstruction (2 papers)
  - [26] A multi-stage neural network approach for coronary 3D reconstruction from uncalibrated X-ray angiography images. (Kritika Iyer, 2023) View paper
  - [35] Real-time deformable SLAM with geometrically adapted template for dynamic monocular laparoscopic scenes. (Xuanshuang Tang, 2024) View paper
  - Aerial and Trajectory Reconstruction (2 papers)
  - [41] Reconstruction of 3D flight trajectories from ad-hoc camera networks (Jingtong Li, 2020) View paper
  - [42] Real-time dense 3D Reconstruction from monocular video data captured by low-cost UAVs (Hermann Max, 2021) View paper
  - Industrial and Simulation Applications (2 papers)
  - [39] Integrated Pipeline for Monocular 3D Reconstruction and Finite Element Simulation in Industrial Applications (Zheng Bowen, 2025) View paper
  - [43] EPRecon: An Efficient Framework for Real-Time Panoptic 3D Reconstruction from Monocular Video (Zhen Zhou, 2025) View paper
- Motion and Dynamics Estimation (2 papers)
  - [19] Immediate vehicle movement estimation and 3D reconstruction for Mono cameras by utilizing epipolar geometry and direction prior (Zoltan Rozsa, 2022) View paper
  - [44] Volumetric change detection using uncalibrated 3D reconstruction models (Yakov Diskin, 2015) View paper
- Uncalibrated Reconstruction Techniques (3 papers)
  - [16] Real-time 3D features reconstruction through monocular vision (Alfredo Liverani, 2010) View paper
  - [17] Robust 3D reconstruction from uncalibrated small motion clips (Zhaoxin Li, 2021) View paper
  - [34] Augmented reality using uncalibrated video sequences (K. Cornelis, 2000) View paper

Narrative

Core task: online dynamic 3D reconstruction from uncalibrated video streams. The field addresses the challenge of recovering time-varying geometry from video without prior camera calibration, spanning a diverse set of methodological branches. Feed-forward dynamic scene reconstruction methods emphasize rapid, single-pass inference using learned priors, often leveraging neural representations or Gaussian splatting for real-time performance. Optimization-based approaches iteratively refine geometry and motion through energy minimization, trading speed for accuracy. Incremental and online reconstruction techniques process frames sequentially, maintaining consistency across time, while static human reconstruction from monocular video focuses on capturing detailed body shape and pose from single-camera setups. Sparse feature-based methods rely on keypoint tracking and structure-from-motion pipelines, whereas depth-guided neural reconstruction integrates depth sensors or learned depth cues to regularize geometry. Specialized application branches target domains such as medical imaging, architectural scenes, or aerial footage, and motion and dynamics estimation explicitly models temporal deformations. Uncalibrated reconstruction techniques handle unknown or varying intrinsics, a critical capability for casual video capture.

Recent work has concentrated on real-time dynamic Gaussian splatting, where methods like StreamSplat[0], DGS-LRM[4], and SplineGS[22] push the frontier of feed-forward reconstruction by representing scenes as collections of evolving 3D Gaussians. StreamSplat[0] emphasizes streaming video input and online adaptation, positioning itself within the real-time dynamic Gaussian splatting cluster alongside neighbors such as Bullet-Time Reconstruction[5] and QUEEN[24], which explore multi-view synchronization and quality-efficiency trade-offs. In contrast, optimization-based pipelines like Lyra[1] and incremental methods such as MegaSaM[6] prioritize geometric fidelity over speed, iteratively refining reconstructions as new frames arrive. The tension between feed-forward speed and optimization-based accuracy remains a central theme, with StreamSplat[0] leaning toward the former by exploiting learned scene priors and efficient splatting, while closely related works like SplineGS[22] explore temporal smoothness through spline-based motion modeling.

Related Works in Same Category

The following **4 sibling papers** share the same taxonomy leaf node with the original paper:

1. DGS-LRM: Real-Time Deformable 3D Gaussian Reconstruction From Monocular Videos

Authors: Lin, Chieh Hubert, LV Zhaoyang, C. Lin, Wu, et al. (37 authors total) | **Year/Venue**: 2025 • arXiv.org | **URL**: View paper

Abstract

We introduce the Deformable Gaussian Splats Large Reconstruction Model (DGS-LRM), the first feed-forward method predicting deformable 3D Gaussian splats from a monocular posed video of any dynamic scene. Feed-forward scene reconstruction has gained significant attention for its ability to rapidly create digital replicas of real-world environments. However, most existing models are limited to static scenes and fail to reconstruct the motion of moving objects. Developing a feed-forward model for d...

Relationship Analysis

Both papers belong to the Real-Time Dynamic Gaussian Splatting category, using feed-forward approaches with 3D Gaussian representations for dynamic scene reconstruction. They overlap in their goal of achieving real-time dynamic 3D reconstruction from video without per-scene optimization, both employing deformable Gaussian representations and transformer-based architectures. However, StreamSplat focuses on online reconstruction from uncalibrated video streams with bidirectional deformation fields and adaptive Gaussian fusion for streaming data, while DGS-LRM emphasizes feed-forward reconstruction from monocular posed videos with per-pixel deformable Gaussians trained on synthetic data with ground-truth scene flow supervision.

---

2. Feed-Forward Bullet-Time Reconstruction of Dynamic Scenes from Monocular Videos

Authors: Liang, Hanxue, Ren Jiawei, Hanxue Liang, Mirzaei, et al. (30 authors total) | **Year/Venue**: 2024 • arXiv.org | **URL**: View paper

Abstract

Recent advancements in static feed-forward scene reconstruction have demonstrated significant progress in high-quality novel view synthesis. However, these models often struggle with generalizability across diverse environments and fail to effectively handle dynamic content. We present BTimer (short for BulletTimer), the first motion-aware feed-forward model for real-time reconstruction and novel view synthesis of dynamic scenes. Our approach reconstructs the full scene in a 3D Gaussian Splattin...

Relationship Analysis

Both papers belong to the Real-Time Dynamic Gaussian Splatting category, employing feed-forward approaches for dynamic scene reconstruction using 3D Gaussian Splatting representations. They overlap in addressing online/real-time reconstruction from monocular videos without requiring per-scene optimization, both achieving significant speedups over optimization-based methods. The key difference is that StreamSplat focuses on uncalibrated video streams with bidirectional deformation fields and adaptive Gaussian fusion for online streaming scenarios, while BTimer emphasizes bullet-time reconstruction at a target timestamp by aggregating information from all context frames, achieving 150ms reconstruction time but not explicitly designed for continuous online streaming.

---

3. SplineGS: Robust Motion-Adaptive Spline for Real-Time Dynamic 3D Gaussians from Monocular Video

Authors: Jongmin Park, Minh-Quan Viet Bui, Juan Luis Gonzalez Bello, Jaeho Moon, J. Bello, et al. (7 authors total) | **Year/Venue**: 2024 • Computer Vision and Pattern Recognition | **URL**: View paper

Abstract

Synthesizing novel views from in-the-wild monocular videos is challenging due to scene dynamics and the lack of multi-view cues. To address this, we propose SplineGS, a COLMAP-free dynamic 3D Gaussian Splatting (3DGS) framework for high-quality reconstruction and fast rendering from monocular videos. At its core is a novel Motion-Adaptive Spline (MAS) method, which represents continuous dynamic 3D Gaussian trajectories using cubic Her-mite splines with a small number of control points. For MAS, ...

Relationship Analysis

Both papers belong to the Real-Time Dynamic Gaussian Splatting category, focusing on feed-forward approaches for dynamic scene reconstruction using 3D Gaussian representations. They overlap in addressing dynamic 3D reconstruction from monocular video without requiring COLMAP preprocessing, both employing deformation fields to model temporal dynamics. However, StreamSplat emphasizes online streaming reconstruction from uncalibrated videos with bidirectional deformation and adaptive Gaussian fusion, while SplineGS focuses on Motion-Adaptive Splines with cubic Hermite splines and control point pruning for robust trajectory modeling.

---

4. QUEEN: QUantized Efficient ENcoding of Dynamic Gaussians for Streaming Free-viewpoint Videos

Authors: Girish, Sharath, Li Tianye, Sharath Girish, Mazumdar, et al. (15 authors total) | **Year/Venue**: 2024 • Neural Information Processing Systems | **URL**: View paper

Abstract

Online free-viewpoint video (FVV) streaming is a challenging problem, which is relatively under-explored. It requires incremental on-the-fly updates to a volumetric representation, fast training and rendering to satisfy real-time constraints and a small memory footprint for efficient transmission. If achieved, it can enhance user experience by enabling novel applications, e.g., 3D video conferencing and live volumetric video broadcast, among others. In this work, we propose a novel framework for...

Relationship Analysis

Both papers belong to the Real-Time Dynamic Gaussian Splatting category, focusing on feed-forward approaches for dynamic scene reconstruction using 3D Gaussian representations. While StreamSplat addresses online dynamic 3D reconstruction from uncalibrated video streams with bidirectional deformation fields and adaptive Gaussian fusion for arbitrary-length streaming, QUEEN focuses on efficient encoding and streaming of free-viewpoint videos through quantization-sparsity frameworks and residual learning between consecutive frames. The key distinction is that StreamSplat emphasizes uncalibrated online reconstruction with temporal coherence mechanisms, whereas QUEEN prioritizes compression efficiency and bandwidth optimization for streaming pre-captured dynamic content.

Contributions Analysis

Overall novelty summary. StreamSplat introduces a feed-forward framework for online dynamic 3D reconstruction using Gaussian Splatting representations, processing uncalibrated video streams without per-scene optimization. The paper resides in the 'Real-Time Dynamic Gaussian Splatting' leaf, which contains five papers total, indicating a moderately populated but emerging research direction. This leaf sits within the broader 'Feed-Forward Dynamic Scene Reconstruction' branch, distinguishing itself from optimization-based methods that require iterative refinement. The focus on streaming input and online adaptation positions StreamSplat at the intersection of real-time performance and dynamic scene modeling.

The taxonomy reveals neighboring research directions that contextualize StreamSplat's contributions. Adjacent leaves include 'Generative Model-Based 3D Reconstruction' (leveraging diffusion priors) and 'Multi-Human 4D Reconstruction' (specialized for human subjects), both under the same feed-forward parent branch. The 'Incremental and Online Reconstruction' branch contains methods like dense volumetric reconstruction and online human-scene reconstruction, which share the streaming constraint but differ in representation choice (volumetric vs. Gaussian). StreamSplat's uncalibrated input handling also connects to the 'Uncalibrated Reconstruction Techniques' branch, though that category emphasizes augmented reality applications rather than dynamic scene modeling.

Among fifteen candidates examined across three contributions, none were identified as clearly refuting StreamSplat's novelty. The core framework (Contribution 1) examined nine candidates with zero refutable overlaps, suggesting limited prior work on fully feed-forward, online Gaussian splatting for dynamic scenes. The probabilistic sampling mechanism (Contribution 2) and bidirectional deformation field with adaptive fusion (Contribution 3) examined four and two candidates respectively, also without refutation. This limited search scope—fifteen papers from semantic retrieval—indicates that while no direct precedents emerged, the analysis does not exhaustively cover all related work in real-time reconstruction or deformation modeling.

Given the constrained literature search and the moderately populated taxonomy leaf, StreamSplat appears to occupy a distinct niche within real-time dynamic Gaussian splatting. The absence of refutable candidates among fifteen examined papers suggests technical differentiation from sibling works, though the small sample size precludes definitive claims about field-wide novelty. The combination of online processing, uncalibrated input, and adaptive Gaussian fusion distinguishes StreamSplat from optimization-heavy or batch-processing alternatives, but broader validation against the full corpus of dynamic reconstruction methods remains necessary.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: StreamSplat framework for online dynamic 3D reconstruction

Description: The authors present StreamSplat, a fully feed-forward system that instantly transforms uncalibrated video streams of arbitrary length into dynamic 3D Gaussian Splatting representations in an online manner, achieving real-time performance with a 1200× speedup over optimization-based methods.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Anysplat: Feed-forward 3d gaussian splatting from unconstrained views

URL: View paper

Brief Assessment

Anysplat[55] focuses on feed-forward novel view synthesis from uncalibrated static image collections, not online dynamic 3D reconstruction from video streams. The candidate does not address temporal dynamics, streaming processing, or deformation fields for moving scenes.

---

2. SLAM3R: Real-Time Dense Scene Reconstruction from Monocular RGB Videos

URL: View paper

Brief Assessment

SLAM3R[51] focuses on dense 3D reconstruction from RGB videos using pointmaps and global coordinate registration, not dynamic 3D Gaussian Splatting representations. The technical approaches differ fundamentally: SLAM3R regresses 3D pointmaps and aligns them globally, while StreamSplat predicts dynamic 3DGS with deformation fields and adaptive fusion for temporal coherence in dynamic scenes.

---

3. Forge4D: Feed-Forward 4D Human Reconstruction and Interpolation from Uncalibrated Sparse-view Videos

URL: View paper

Brief Assessment

Forge4D[15] focuses on 4D human reconstruction from sparse-view videos with emphasis on novel-time interpolation via dense motion prediction, whereas StreamSplat targets general dynamic scene reconstruction from arbitrary-length uncalibrated video streams with different technical approaches (probabilistic sampling, bidirectional deformation).

---

4. A-nerf: Articulated neural radiance fields for learning human shape, appearance, and pose

URL: View paper

Brief Assessment

A-nerf[57] focuses on learning articulated human body models from monocular videos using neural radiance fields with skeleton-based deformation, not general online dynamic 3D reconstruction from uncalibrated video streams using feed-forward Gaussian splatting methods.

---

5. MapAnything: Universal Feed-Forward Metric 3D Reconstruction

URL: View paper

Brief Assessment

MapAnything[56] focuses on static metric 3D reconstruction from images with optional geometric inputs (depth, poses, intrinsics), not online dynamic scene reconstruction from uncalibrated video streams. The candidate does not address temporal dynamics, deformation fields, or streaming video processing that are central to StreamSplat.

---

6. Flare: Feed-forward geometry, appearance and camera estimation from uncalibrated sparse views

URL: View paper

Brief Assessment

Flare[52] focuses on static 3D reconstruction from sparse uncalibrated views using feed-forward methods, not online dynamic scene reconstruction from video streams. The candidate addresses pose estimation and static geometry, while the original paper tackles dynamic 3D Gaussian Splatting with temporal deformation fields.

7. PanoRecon: Real-Time Panoptic 3D Reconstruction from Monocular Video

URL: View paper

Brief Assessment

PanoRecon[53] focuses on panoptic 3D reconstruction (geometry + semantic/instance segmentation) from monocular video, not dynamic 3D Gaussian Splatting representations. The technical approaches differ fundamentally: PanoRecon uses volumetric features and voxel clustering, while StreamSplat uses Gaussian primitives with deformation fields.

8. Feed-Forward Bullet-Time Reconstruction of Dynamic Scenes from Monocular Videos

URL: View paper

Brief Assessment

Bullet-Time Reconstruction[5] focuses on bullet-time reconstruction at a single target timestamp from casual monocular videos, not continuous online streaming reconstruction from uncalibrated video streams of arbitrary length.

9. Large spatial model: End-to-end unposed images to semantic 3d

URL: View paper

Brief Assessment

Large Spatial Model[58] focuses on semantic 3D reconstruction from unposed images with language-based segmentation capabilities, not on online dynamic 3D reconstruction from video streams. The candidate addresses static scene reconstruction with semantic understanding, while the original paper specifically targets dynamic scene modeling with temporal coherence across video streams.

Contribution 2: Probabilistic sampling mechanism for 3D Gaussian position prediction

Description: The authors propose a probabilistic position sampling strategy that predicts a truncated normal distribution for each 3D offset rather than direct regression. This approach captures geometric uncertainty and avoids local minima common in feed-forward models.

This contribution was assessed against **4 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. LiftPose3D, a deep learning-based approach for transforming two-dimensional to three-dimensional poses in laboratory animals

URL: View paper

Brief Assessment

LiftPose3D[61] focuses on transforming 2D poses to 3D poses in laboratory animals using deep learning, not on probabilistic sampling for 3D Gaussian position prediction from uncalibrated video streams. The candidate's sparse context mentions hand-annotated joint locations and uncalibrated images but provides no evidence of probabilistic sampling mechanisms for Gaussian splatting or dynamic scene reconstruction.

2. 3D GAUSSIAN SPLATTING FOR REAL TIME RADIANCE FIELD RENDERING USING INSTA360 CAMERA

URL: View paper

Brief Assessment

Gaussian Splatting Insta360[64] focuses on applying 3D Gaussian Splatting to large-scale scene reconstruction using Insta360 cameras, without proposing probabilistic sampling mechanisms for position prediction. The candidate implements the original Gaussian Splatting method rather than introducing novel sampling strategies.

3. A causal bayesian network and probabilistic programming based reasoning framework for robot manipulation under uncertainty

URL: View paper

Brief Assessment

Causal Bayesian Manipulation[62] focuses on probabilistic programming for robot manipulation using Bayesian networks and physics simulation, not on predicting 3D Gaussian positions from uncalibrated video inputs for scene reconstruction.

4. Unsupervised learning of platform motion in synthetic aperture sonar

URL: View paper

Brief Assessment

Unsupervised Sonar Motion[63] addresses platform motion estimation in synthetic aperture sonar using variational autoencoders for disentangling displacement from coherence measurements. This is fundamentally different from predicting 3D Gaussian positions with truncated normal distributions for scene reconstruction.

Contribution 3: Bidirectional deformation field with adaptive Gaussian fusion

Description: The authors introduce a bidirectional deformation field that models both forward and backward motion between consecutive frames, combined with an adaptive fusion mechanism based on time-dependent opacity. This enables robust cross-frame associations and maintains temporal coherence while naturally handling emerging and vanishing scene content.

This contribution was assessed against **2 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. FutureGS: Structured Gaussian Fields for Future-Aware Dynamic Scene Modeling

URL: View paper

Brief Assessment

FutureGS[60] focuses on future scene prediction using bidirectional LSTM for temporal encoding, while the original paper addresses online reconstruction from streaming video with bidirectional deformation between consecutive frames and opacity-based fusion. The technical approaches and application domains differ substantially.

2. LSTM-Kalman Filter-Based Multi-Sensor Signal Fusion for UAV Altitude Prediction in Non-Gaussian Environments.

URL: View paper

Brief Assessment

LSTM-Kalman UAV[59] focuses on UAV altitude prediction using LSTM-Kalman filtering for sensor fusion under non-Gaussian noise. This is fundamentally different from the original paper's bidirectional deformation field for 3D Gaussian splatting in dynamic scene reconstruction, with no overlap in methodology or application domain.

Appendix: Text Similarity Detection

Textual similarity detection checked 19 papers and found 2 similarity segment(s) across 1 paper(s).

The following **1 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. Forge4D: Feed-Forward 4D Human Reconstruction and Interpolation from Uncalibrated Sparse-view Videos

Detected in: Contribution: contribution_1

⚠ **Note**: This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] StreamSplat: Towards Online Dynamic 3D Reconstruction from Uncalibrated Video Streams View paper
- [1] Lyra: Generative 3D Scene Reconstruction via Video Diffusion Model Self-Distillation View paper
- [2] Sparse dynamic 3d reconstruction from unsynchronized videos View paper
- [3] Real-time 3D pose reconstruction of human body from monocular video sequences View paper
- [4] DGS-LRM: Real-Time Deformable 3D Gaussian Reconstruction From Monocular Videos View paper
- [5] Feed-Forward Bullet-Time Reconstruction of Dynamic Scenes from Monocular Videos View paper
- [6] MegaSaM: Accurate, fast and robust structure and motion from casual dynamic videos View paper
- [7] Scene coordinate reconstruction: Posing of image collections via incremental learning of a relocalizer View paper
- [8] Real-Time Reconstruction of Multi-Body Pedestrian Pre-Impact Posture in Collision Accidents From Monocular Images View paper
- [9] Link to the Past: Temporal Propagation for Fast 3D Human Reconstruction from Monocular Video View paper
- [10] Consistent4D: Consistent 360Â° Dynamic Object Generation from Monocular Video View paper
- [11] Visual-hull reconstruction from uncalibrated and unsynchronized video streams View paper
- [12] Incremental Dense Reconstruction From Monocular Video With Guided Sparse Feature Volume Fusion View paper
- [13] PAD3R: Pose-Aware Dynamic 3D Reconstruction from Casual Videos View paper
- [14] Realtime dynamic 3D facial reconstruction for monocular video in-the-wild View paper
- [15] Forge4D: Feed-Forward 4D Human Reconstruction and Interpolation from Uncalibrated Sparse-view Videos View paper
- [16] Real-time 3D features reconstruction through monocular vision View paper
- [17] Robust 3D reconstruction from uncalibrated small motion clips View paper
- [18] A fast 3D scene reconstructing method using continuous video View paper
- [19] Immediate vehicle movement estimation and 3D reconstruction for Mono cameras by utilizing epipolar geometry and direction prior View paper
- [20] Humans as a Calibration Pattern: Dynamic 3D Scene Reconstruction from Unsynchronized and Uncalibrated Videos View paper
- [21] Flora: Dual-Frequency LOss-Compensated ReAl-Time Monocular 3D Video Reconstruction View paper
- [22] SplineGS: Robust Motion-Adaptive Spline for Real-Time Dynamic 3D Gaussians from Monocular Video View paper
- [23] MonoFusion: Real-time 3D reconstruction of small scenes with a single web camera View paper
- [24] QUEEN: QUantized Efficient ENcoding of Dynamic Gaussians for Streaming Free-viewpoint Videos View paper
- [25] ODHSR: Online Dense 3D Reconstruction of Humans and Scenes from Monocular Videos View paper
- [26] A multi-stage neural network approach for coronary 3D reconstruction from uncalibrated X-ray angiography images. View paper
- [27] DG-Recon: Depth-Guided Neural 3D Scene Reconstruction View paper
- [28] Projective reconstruction and metric models from uncalibrated video sequences View paper
- [29] MUC: Mixture of Uncalibrated Cameras for Robust 3D Human Body Reconstruction View paper
- [30] Inferring 3D body pose from uncalibrated video View paper
- [31] 3D face pose tracking from an uncalibrated monocular camera View paper
- [32] Cross-Dimensional Refined Learning for Real-Time 3D Visual Perception from Monocular Video View paper
- [33] 3d reconstruction of architectural scenes from uncalibrated video sequences View paper
- [34] Augmented reality using uncalibrated video sequences View paper
- [35] Real-time deformable SLAM with geometrically adapted template for dynamic monocular laparoscopic scenes. View paper
- [36] Model-based human gait tracking, 3D reconstruction and recognition in uncalibrated monocular video View paper
- [37] Real-time 3D morphable shape model fitting to monocular in-the-wild videos View paper
- [38] Recollection from Pensieve: Novel View Synthesis via Learning from Uncalibrated Videos View paper
- [39] Integrated Pipeline for Monocular 3D Reconstruction and Finite Element Simulation in Industrial Applications View paper
- [40] BabyNet: Reconstructing 3D faces of babies from uncalibrated photographs View paper
- [41] Reconstruction of 3D flight trajectories from ad-hoc camera networks View paper
- [42] Real-time dense 3D Reconstruction from monocular video data captured by low-cost UAVs View paper
- [43] EPRecon: An Efficient Framework for Real-Time Panoptic 3D Reconstruction from Monocular Video View paper
- [44] Volumetric change detection using uncalibrated 3D reconstruction models View paper
- [45] 3d reconstruction of human motion and skeleton from uncalibrated monocular video View paper
- [46] 3D reconstruction system based on incremental structure from motion using a camera with varying parameters View paper
- [47] LivePose: Online 3D Reconstruction from Monocular Video with Dynamic Camera Poses View paper
- [48] REC-MV: REconstructing 3D Dynamic Cloth from Monocular Videos View paper
- [49] SST: Real-time End-to-end Monocular 3D Reconstruction via Sparse Spatial-Temporal Guidance View paper
- [50] PDF-MPI: From Monocular Video to Volumetric Video View paper
- [51] SLAM3R: Real-Time Dense Scene Reconstruction from Monocular RGB Videos View paper
- [52] Flare: Feed-forward geometry, appearance and camera estimation from uncalibrated sparse views View paper
- [53] PanoRecon: Real-Time Panoptic 3D Reconstruction from Monocular Video View paper

- [54] Splatt3r: Zero-shot gaussian splatting from uncalibrated image pairs View paper
- [55] Anysplat: Feed-forward 3d gaussian splatting from unconstrained views View paper
- [56] MapAnything: Universal Feed-Forward Metric 3D Reconstruction View paper
- [57] A-nerf: Articulated neural radiance fields for learning human shape, appearance, and pose View paper
- [58] Large spatial model: End-to-end unposed images to semantic 3d View paper
- [59] LSTM-Kalman Filter-Based Multi-Sensor Signal Fusion for UAV Altitude Prediction in Non-Gaussian Environments. View paper
- [60] FutureGS: Structured Gaussian Fields for Future-Aware Dynamic Scene Modeling View paper
- [61] LiftPose3D, a deep learning-based approach for transforming two-dimensional to three-dimensional poses in laboratory animals View paper
- [62] A causal bayesian network and probabilistic programming based reasoning framework for robot manipulation under uncertainty View paper
- [63] Unsupervised learning of platform motion in synthetic aperture sonar View paper
- [64] 3D GAUSSIAN SPLATTING FOR REAL TIME RADIANCE FIELD RENDERING USING INSTA360 CAMERA View paper