Novelty Assessment Report

Paper: Compositional Diffusion with Guided search for Long-Horizon Planning
PDF URL: https://openreview.net/pdf?id=b8avf4F2hn
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-01

Abstract

Generative models have emerged as powerful tools for planning, with compositional approaches offering particular promise for modeling long-horizon task distributions by composing together local, modular generative models. This compositional paradigm spans diverse domains, from multi-step manipulation planning to panoramic image synthesis to long video generation. However, compositional generative models face a critical challenge: when local distributions are multimodal, existing composition methods average incompatible modes, producing plans that are neither locally feasible nor globally coherent. We propose Compositional Diffusion with Guided Search (CDGS), which addresses this \emph{mode averaging} problem by embedding search directly within the diffusion denoising process. Our method explores diverse combinations of local modes through population-based sampling, prunes infeasible candidates using likelihood-based filtering, and enforces global consistency through iterative resampling between overlapping segments. CDGS matches oracle performance on seven robot manipulation tasks, outperforming baselines that lack compositionality or require long-horizon training data. The approach generalizes across domains, enabling coherent text-guided panoramic images and long videos through effective local-to-global message passing. More details: https://cdgsearch.github.io/

Core Task Landscape

This paper addresses: **Compositional Generation of Long-Horizon Sequences from Short-Horizon Models**

A total of **50 papers** were analyzed and organized into a taxonomy with **23 categories**.

Taxonomy Overview

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

- **Robotic Manipulation and Task Planning**
- **Spatiotemporal Forecasting and Extrapolation**
- **Video Generation and Synthesis**
- **Motion and Trajectory Synthesis**
- **Sequential Recommendation and User Modeling**
- **Domain-Specific Temporal Modeling**
- **Reasoning and Memory Architectures**
- **Multi-Step Time Series Forecasting**
- **Long-Horizon Vision Tasks**
- **Bio-Inspired Neuromorphic Memory Systems**

Complete Taxonomy Tree

- Compositional Generation of Long-Horizon Sequences from Short-Horizon Models Survey Taxonomy
- Robotic Manipulation and Task Planning
  - Diffusion-Based Trajectory Composition ★ (4 papers)
  - [0] Compositional Diffusion with Guided search for Long-Horizon Planning (Anon et al., 2026) View paper
  - [5] Generative trajectory stitching through diffusion composition (Luo, 2025) View paper
  - [13] What Do You Need for Diverse Trajectory Stitching in Diffusion Planning? (Q Clark, 2025) View paper
  - [50] Compositional Visual Planning via Inference-Time Diffusion Scaling (SCALING, n.d.) View paper
  - Hierarchical Skill Chaining and Subgoal Decomposition (5 papers)
  - [2] Structured motion generation with predictive learning: Proposing subgoal for long-horizon manipulation (Namiko Saito, 2023) View paper
  - [7] Generative skill chaining: Long-horizon skill planning with diffusion models (Mishra, 2023) View paper
  - [8] Extendable long-horizon planning via hierarchical multiscale diffusion (Chen Chang, 2025) View paper
  - [20] Planning to practice: Efficient online fine-tuning by composing goals in latent space (Kuan Fang, 2022) View paper
  - [41] Hierarchies, search, and generative models in sequential decision-making (KujanpÃ¤Ã¤, 2025) View paper
  - Imitation Learning for Long-Horizon Tasks (3 papers)
  - [14] Learning to generalize across long-horizon tasks from human demonstrations (Ajay Mandlekar, 2020) View paper
  - [23] ManiLong-Shot: Interaction-Aware One-Shot Imitation Learning for Long-Horizon Manipulation (Zixuan Chen, 2025) View paper
  - [37] One-Shot Transfer of Long-Horizon Extrinsic Manipulation Through Contact Retargeting (Albert Wu, 2024) View paper
  - Model-Based Reinforcement Learning and World Models (2 papers)
  - [25] Advances in Long-Horizon Planning via Transformer and Diffusion World Models (Chen, 2025) View paper
  - [48] Example-Driven Model-Based Reinforcement Learning for Solving Long-Horizon Visuomotor Tasks (Bo-Han Wu, 2021) View paper
  - Vision-Language-Action Systems (3 papers)

- [30] Towards Long-Horizon Vision-Language-Action System: Reasoning, Acting and Memory (D Li, 2025) View paper
- [42] NOD-TAMP: Generalizable Long-Horizon Planning with Neural Object Descriptors (Cheng Shuo, 2023) View paper
- [43] Towards Long-Horizon Vision-Language Navigation: Platform, Benchmark and Method (Xin-Shuai Song, 2024) View paper

- Spatiotemporal Forecasting and Extrapolation
  - Radar Echo and Precipitation Nowcasting (8 papers)
  - [3] An Long Short-Term Memory Model with Multi-Scale Context Fusion and Attention for Radar Echo Extrapolation (Guangxin He, 2024) View paper
  - [4] SepConv-ens: An ensemble of separable convolution-based deep learning models for weather radar echo temporal extrapolation (Gabriela Czibula, 2024) View paper
  - [15] Short-Term Precipitation Radar Echo Extrapolation Method Based on the MS-DD3D-RSTN Network and STLoss Function (Wulin Yang, 2024) View paper
  - [18] 3D-UNet-LSTM: A Deep Learning-Based Radar Echo Extrapolation Model for Convective Nowcasting (Shiqing Guo, 2023) View paper
  - [21] Hourly rolling correction of precipitation forecast via convolutional and long shortâ□□term memory networks (Ruyi Yang, 2022) View paper
  - [22] An Improved LSTM-based Method Capturing Temporal Correlations and Using Attention Mechanism for Radar Echo Extrapolation (Zhiyun Yang, 2022) View paper
  - [44] Data-Driven Prediction of Severe Convection at Deutscher Wetterdienst (DWD): A Brief Overview of Recent Developments (Richard Muller, 2024) View paper
  - [46] Data Driven Prediction of Severe Convection at DWD. An Overview of Recent Developments (Richard Muller, 2024) View paper
  - Traffic Flow Prediction (1 papers)
  - [33] LSTTN: A Long-Short Term Transformer-based Spatio-temporal Neural Network for Traffic Flow Forecasting (Qinyao Luo, 2024) View paper
  - General Spatiotemporal Inference (1 papers)
  - [45] Decoupling Long- and Short-Term Patterns in Spatiotemporal Inference (Junfeng Hu, 2021) View paper
- Video Generation and Synthesis
  - Long-Horizon Video Diffusion Models (1 papers)
  - [12] VideoSSM: Autoregressive Long Video Generation with Hybrid State-Space Memory (Yifei Yu, 2025) View paper
  - Robotic Video Generation (1 papers)
  - [28] RoboEnvision: A Long-Horizon Video Generation Model for Multi-Task Robot Manipulation (Yang Liu-di, 2025) View paper
  - Temporal Consistency in Video Translation (1 papers)
  - [19] Long-term temporally consistent unpaired video translation from simulated surgical 3d data (Rivoir, 2021) View paper
- Motion and Trajectory Synthesis
  - Human Motion Synthesis with Recurrent Networks (2 papers)
  - [16] Recurrent network models for human dynamics (Katerina Fragkiadaki, 2015) View paper
  - [17] Combining recurrent neural networks and adversarial training for human motion synthesis and control (Zhiyong Wang, 2019) View paper
  - Keyframe-Driven Motion Inbetweening (1 papers)
  - [10] Longâ□□term Motion Inâ□□betweening via Keyframe Prediction (Seokhyeon Hong, 2024) View paper
  - Trajectory Prediction and Distillation (2 papers)
  - [31] Modular and Multimodal Generative Adversarial Imitation Learning for Modeling Flight Trajectories (Christos Spatharis, 2025) View paper
  - [39] Distilling Knowledge for Short-to-Long Term Trajectory Prediction (Sourav Das, 2024) View paper
- Sequential Recommendation and User Modeling (3 papers)
  - [1] GNN-based long and short term preference modeling for next-location prediction (Jinbo Liu, 2023) View paper
  - [27] Temporal User Profiling with LLMs: Balancing Short-Term and Long-Term Preferences for Recommendations (Mansoury, 2025) View paper
  - [38] Long- and Short- Term Sequential Recommendation with Temporal Interval (Kun He, 2022) View paper
- Domain-Specific Temporal Modeling
  - Medical and Biological Temporal Synthesis (2 papers)
  - [6] Color Spike Camera Reconstruction via Long Short-Term Temporal Aggregation of Spike Signals (Yanchen Dong, 2025) View paper
  - [9] 4D VQ-GAN: Synthesising Medical Scans at Any Time Point for Personalised Disease Progression Modelling of Idiopathic Pulmonary Fibrosis (Zhao An, 2025) View paper
  - Materials Science and Physical Process Prediction (1 papers)
  - [29] Predicting Grain Growth in Polycrystalline Materials Using Deep Learning Time Series Models (Eliane Younes, 2025) View paper
  - Face Aging and Long-Term Visual Evolution (1 papers)
  - [24] A concatenational graph evolution aging model (Jin-Li Suo, 2012) View paper
  - Security and Intrusion Detection Temporal Modeling (2 papers)
  - [11] Multi-temporal dependency handling in video smoke recognition: A holistic approach spanning spatial, short-term, and long-term perspectives (Feng Yang, 2024) View paper
  - [26] Learning long- and short-term dependencies for network intrusion detection (Qiong-lan Na, 2024) View paper
- Reasoning and Memory Architectures (2 papers)
  - [34] h1: Bootstrapping LLMs to Reason over Longer Horizons via Reinforcement Learning (S. Motwani, 2025) View paper
  - [40] An Integrated Model of Contex, Short-Term, and Long-Term Memory (Gosmann, 2018) View paper
- Multi-Step Time Series Forecasting (3 papers)
  - [36] Multi-step time series forecasting with an ensemble of varied length mixture models (Yicun Ouyang, 2018) View paper
  - [47] Statistical Methods (Watson, 2023) View paper
  - [49] Comment on wes-2021-33 (Nicola Bodini, 2021) View paper
- Long-Horizon Vision Tasks (1 papers)
  - [35] Long-RVOS: A Comprehensive Benchmark for Long-term Referring Video Object Segmentation (Liang Tianming, 2025) View paper

- Bio-Inspired Neuromorphic Memory Systems (1 papers)
  - [32] Bioâ□□Inspired Sb2S3 Memristors Supporting Dualâ□□Mode Operation via Selfâ□□Regulating Transport Pathways for Allâ□□inâ□□one Neuromorphic Perception and â□¦ (Y Yang, 2025) View paper

Narrative

Core task: Compositional generation of long-horizon sequences from short-horizon models. This field addresses the challenge of extending predictions or plans far into the future by composing outputs from models trained on shorter temporal windows. The taxonomy reveals a diverse landscape spanning robotic manipulation and task planning, spatiotemporal forecasting (weather radar, precipitation), video generation and synthesis, motion and trajectory synthesis, sequential recommendation, domain-specific temporal modeling (medical imaging, materials science), reasoning and memory architectures, multi-step time series forecasting, long-horizon vision tasks, and bio-inspired neuromorphic memory systems. Within robotic manipulation, diffusion-based trajectory composition has emerged as a particularly active direction, leveraging generative models to stitch together short-horizon skills or subgoals into coherent long-horizon behaviors. Meanwhile, spatiotemporal forecasting branches explore recurrent and attention-based architectures for extrapolating radar echoes or precipitation patterns, and video synthesis methods tackle the challenge of maintaining temporal consistency over extended sequences.

A central tension across these branches involves balancing computational efficiency with the ability to capture long-range dependencies and avoid compounding errors. In robotic planning, works like Trajectory Stitching Diffusion[5] and Diverse Trajectory Stitching[13] explore how to compose pre-trained diffusion models for different skills, while Compositional Diffusion Planning[0] sits within this cluster by emphasizing modular composition of trajectory segments to achieve extended task horizons. Compared to approaches that rely on hierarchical abstractions (e.g., Subgoal Manipulation[2]) or skill chaining (Skill Chaining Diffusion[7]), the diffusion-based composition methods offer flexible probabilistic blending of short-horizon priors. In contrast, spatiotemporal forecasting branches such as LSTM Radar Extrapolation[3] and SepConv Radar Ensemble[4] focus on recurrent or convolutional architectures for weather prediction, highlighting a different set of trade-offs around spatial resolution and ensemble uncertainty. The original paper's emphasis on diffusion-based trajectory composition places it squarely within the robotic manipulation branch, where it contributes to ongoing efforts to scale planning horizons without retraining monolithic models.

Related Works in Same Category

The following **3 sibling papers** share the same taxonomy leaf node with the original paper:

1. Generative trajectory stitching through diffusion composition

Authors: Luo, Yunhao, Mishra, Utkarsh A., Yunhao Luo, et al. (12 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Effective trajectory stitching for long-horizon planning is a significant challenge in robotic decision-making. While diffusion models have shown promise in planning, they are limited to solving tasks similar to those seen in their training data. We propose CompDiffuser, a novel generative approach that can solve new tasks by learning to compositionally stitch together shorter trajectory chunks from previously seen tasks. Our key insight is modeling the trajectory distribution by subdividing it ...

Relationship Analysis

Both papers belong to the Diffusion-Based Trajectory Composition category, using diffusion models to compose short trajectory segments into long-horizon plans for robotic manipulation. They share the core approach of compositional diffusion with overlapping segments and iterative information propagation between neighboring chunks. The key difference is that the original paper (CDGS) introduces a guided search mechanism with population-based sampling and likelihood-based pruning to address mode-averaging issues, while the candidate paper (CompDiffuser) focuses on bidirectional conditioning through noisy-sample conditioning and autoregressive sampling to ensure physical consistency at connection points.

---

2. What Do You Need for Diverse Trajectory Stitching in Diffusion Planning?

Authors: Q Clark, F Shkurti | **Year/Venue**: 2025 | **URL**: View paper

Abstract

â□¦ of the different whole-sequence composition strategies. The â□¦ of long-horizon planning and short-horizon consistency. â□¦ of both long-horizon planning and short-horizon composition, so it â□¦

Relationship Analysis

Both papers belong to the Diffusion-Based Trajectory Composition category, using diffusion models to compose short trajectory segments into long-horizon plans. They overlap in addressing the challenge of composing multi-modal local distributions to generate coherent global plans, with both identifying mode-averaging as a key problem. The original paper (CDGS) addresses this through guided search with iterative resampling and likelihood-based pruning during inference, while the candidate paper focuses on architectural properties (shift equivariance and local receptiveness) and training augmentation strategies to enable composition, analyzing when and why diffusion planners can stitch trajectories.

---

3. Compositional Visual Planning via Inference-Time Diffusion Scaling

Authors: TD SCALING | **URL**: View paper

Abstract

â□¦ a compositional generation perspective on long-horizon planning, we compose plans from overlapping, short-horizon â□¦ model on randomly sampled short chunks from long-horizon â□¦

Relationship Analysis

Both papers belong to the Diffusion-Based Trajectory Composition category, using diffusion models to compose short trajectory segments into long-horizon plans for robotic manipulation. They share the core challenge of addressing mode-averaging issues when composing multimodal local distributions, with both employing message-passing mechanisms during diffusion denoising to enforce global consistency. The key difference is that the original paper (CDGS) uses population-based search with likelihood-based pruning and iterative resampling to filter infeasible candidates, while the candidate paper operates on Tweedie estimates with synchronous/asynchronous message passing on a factor graph to enforce boundary agreement, representing a training-free guidance approach rather than a search-based filtering strategy.

Contributions Analysis

Overall novelty summary. The paper proposes Compositional Diffusion with Guided Search (CDGS), a method for composing short-horizon diffusion models into long-horizon robot manipulation plans. It resides in the 'Diffusion-Based Trajectory Composition' leaf, which contains only four papers total, including this work and three siblings. This is a relatively sparse research direction within the broader taxonomy of 50 papers across 36 topics, suggesting the specific approach of embedding search within diffusion denoising for compositional planning is not yet heavily explored.

The taxonomy reveals that robotic manipulation encompasses multiple alternative paradigms: hierarchical skill chaining and subgoal decomposition (five papers), imitation learning (three papers), model-based RL (two papers), and vision-language-action systems (three papers). The diffusion-based trajectory composition leaf sits adjacent to these approaches, sharing the goal of long-horizon planning but diverging in its use of probabilistic generative models rather than hierarchical abstractions or reinforcement learning. The scope note explicitly excludes non-diffusion methods, positioning this work within a narrower methodological niche focused on generative composition.

Among 18 candidates examined across three contributions, the iterative resampling mechanism shows one refutable candidate from 10 examined, while the core CDGS framework and likelihood-based pruning appear more novel (zero refutable candidates from eight and zero examined, respectively). The limited search scope means these statistics reflect top-K semantic matches and citation expansion, not exhaustive coverage. The iterative resampling mechanism's overlap with prior work suggests this component may have precedent, while the integration of search-based mode exploration within diffusion denoising appears less directly anticipated by the examined literature.

Given the sparse four-paper leaf and limited 18-candidate search, the work appears to occupy a relatively unexplored intersection of diffusion models and search-based planning for compositional generation. The taxonomy context suggests the field is fragmented across diverse methodological branches, with diffusion-based composition representing a minority approach. However, the analysis cannot rule out relevant work outside the top-K semantic neighborhood or in adjacent communities not captured by the search strategy.

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

Contribution 1: Compositional Diffusion with Guided Search (CDGS)

Description: CDGS is a novel inference-time algorithm that integrates guided search into the diffusion denoising process to compose short-horizon local generative models into coherent long-horizon plans. The method addresses the mode-averaging problem in compositional generative models through population-based sampling, iterative resampling for global consistency, and likelihood-based pruning of infeasible candidates.

This contribution was assessed against **8 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Compositional Monte Carlo Tree Diffusion for Extendable Planning

URL: View paper

Brief Assessment

Monte Carlo Tree Diffusion[52] focuses on tree search within individual diffusion-generated trajectories for planning, not on compositional diffusion methods that address mode-averaging when composing local generative models. The candidate's approach operates at the trajectory level rather than composing short-horizon local distributions.

---

2. Hybrid Diffusion for Simultaneous Symbolic and Continuous Planning

URL: View paper

Brief Assessment

Hybrid Diffusion Planning[53] focuses on simultaneous symbolic and continuous planning through hybrid diffusion of discrete and continuous variables, not on compositional diffusion with guided search for avoiding mode-averaging in long-horizon planning.

---

3. Prismatic World Model: Learning Compositional Dynamics for Planning in Hybrid Systems

URL: View paper

Brief Assessment

Prismatic World Model[54] focuses on decomposing hybrid dynamics into composable primitives using mixture-of-experts for model-based planning in continuous control, not on compositional diffusion models with guided search for long-horizon planning or addressing mode-averaging in generative models.

---

4. Unifying Modern AI with Robotics: Survey on MDPs with Diffusion and Foundation Models

URL: View paper

Brief Assessment

Diffusion Foundation Robotics[51] appears to be a survey paper that mentions compositional approaches in passing. The provided context is too fragmentary to assess whether it describes methods that refute CDGS's novelty claims about guided search integration, population-based sampling, iterative resampling, or likelihood-based pruning.

---

5. Inference-time Scaling of Diffusion Models through Classical Search

URL: View paper

Brief Assessment

Inference Time Scaling[55] focuses on inference-time control through classical search algorithms (BFS/DFS) with verifier-guided optimization, not on compositional diffusion for long-horizon planning with mode-averaging solutions.

---

6. ComposableNav: Instruction-Following Navigation in Dynamic Environments via Composable Diffusion

URL: View paper

Brief Assessment

ComposableNav[56] focuses on instruction-following navigation in dynamic environments by composing motion primitives for robot navigation tasks, not on addressing mode-averaging in long-horizon planning through guided search mechanisms.

---

7. Compositional Visual Planning via Inference-Time Diffusion Scaling

URL: View paper

Brief Assessment

Compositional Visual Planning[50] addresses compositional diffusion for visual planning through message passing on tweedie estimates with synchronous/asynchronous schemes, whereas CDGS focuses on population-based sampling with iterative resampling and likelihood-based pruning for mode-averaging in robotic planning. The technical approaches and problem formulations differ substantially.

---

8. Controllable Graph Generation with Diffusion Models via Inference-Time Tree Search Guidance

URL: View paper

Brief Assessment

Graph Generation Guidance[57] focuses on controllable graph generation using MCTS-guided diffusion for molecular structures, not on compositional planning for long-horizon robotic tasks or multi-step manipulation. The technical domains and problem formulations are fundamentally different.

Contribution 2: Iterative resampling mechanism for local-to-global message passing

Description: The method introduces an iterative resampling procedure that alternates between forward noising and denoising steps to propagate information across distant segments through overlapping variables. This enables effective local-to-global message passing, ensuring that compositional sampling produces globally coherent candidate plans.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. ShapeShift: Towards Text-to-Shape Arrangement Synthesis with Content-Aware Geometric Constraints

URL: View paper

Brief Assessment

ShapeShift[60] addresses shape arrangement synthesis with geometric constraints, not compositional diffusion for long-horizon planning. The candidate does not discuss iterative resampling between overlapping segments for message passing in compositional generative models.

---

2. Seismic Data Interpolation via Denoising Diffusion Implicit Models With Coherence-Corrected Resampling

URL: View paper

Brief Assessment

Seismic Data Interpolation[67] applies resampling in seismic data reconstruction, not compositional planning. The resampling alternates between forward noising and denoising to improve continuity between revealed and missing seismic traces, which is a different technical domain and objective than propagating information across distant segments in compositional generative models for planning tasks.

---

3. Multi-view people tracking via hierarchical trajectory composition

URL: View paper

Brief Assessment

Hierarchical Trajectory Composition[66] focuses on multi-view people tracking through hierarchical composition of trajectory fragments, not on iterative resampling for compositional generative models with diffusion-based denoising.

---

4. AccDiffusion: An Accurate Method for Higher-Resolution Image Generation

URL: View paper

Brief Assessment

AccDiffusion[59] focuses on patch-wise higher-resolution image generation with patch-content-aware prompts and dilated sampling. It does not address iterative resampling for message passing across overlapping segments in compositional generative models.

---

5. Compositional foundation models for hierarchical planning

URL: View paper

Prior Art Analysis

Compositional Hierarchical Planning[62] demonstrates prior work on iterative refinement mechanisms that enable information propagation across hierarchical planning levels through repeated sampling and refinement steps. The paper explicitly describes an iterative refinement procedure that alternates between different planning stages to ensure consistency, which serves a similar purpose to the original paper's iterative resampling for local-to-global message passing. Both approaches address the challenge of maintaining coherence across compositional models through iterative procedures that propagate information between segments.

Evidence

Evidence 1 - **Rationale**: Both papers describe iterative refinement procedures that enable information propagation across model components. The candidate paper's iterative refinement mechanism serves the same fundamental purpose as the original's iterative resampling: ensuring global consistency through repeated refinement steps that incorporate feedback from different parts of the system. - **Original**: we apply iterative resampling [39]: repeatedly alternating between forward noising τ(t) ∼p(τ (t)|τ(t-1))and denoising steps. this procedure enables the score network's predictions for each segment to incorporate information from distant neighbors via overlapping variables, encouraging global consist... - **Candidate**: to create consistent plans across our disparate models, we propose an iterative refinement mechanism to ensure consistency using feedback from the downstream models [28]. at each step of the language model's generative process, intermediate feedback from a likelihood estimator conditioned on an ima...

Evidence 2 - **Rationale**: Both algorithms demonstrate iterative procedures with multiple refinement steps (u iterations in original, m iterations in candidate) that enable information propagation. The candidate's iterative refinement from visual planning to task planning mirrors the original's resampling mechanism for propagating information across segments. - **Original**: algorithm 2composedscore require:noisy sampleτ (t), denoising timestepτ, pretrained local plan score functionεθ require:start and goal:xs,xg require:number of resampling stepsu 1:foru=1,...,udo 2:calculateε(τ (t),t)using eq. 3 3:ifu<uthen 4:calculateτ (t-1) using eq. 2 5:add noise tox s/xg: 6:x (t-1... - **Candidate**: algorithm 1 decision making with hip 1: models: large language model pllm, subgoal classifier fϕ, noise model of diffusion eϕ, observation trajectory classifier gψ, inverse dynamics pψ 2: hyperparameters: guidance scales ω, ω′, denoising diffusion steps k 3: input: current observation xt, language g...

---

6. Hydra: A hyper agent for dynamic compositional visual reasoning

URL: View paper

Brief Assessment

Hydra Visual Reasoning[63] focuses on visual reasoning tasks using a reinforcement learning agent to dynamically select instruction samples for compositional reasoning. It does not address iterative resampling for global consistency in compositional generative models with overlapping segments for planning tasks.

---

7. Streamlining Robust Constrained Production Optimization: An Integrated Framework Utilizing Automatically Differentiated Gradient from Deep-Learning-Based â□¦

URL: View paper

Brief Assessment

Production Optimization Framework[65] focuses on production optimization using least-squares support vector regression with iterative sampling refinements for surrogate modeling in oil-water and compositional reservoir models. This is fundamentally different from the original paper's iterative resampling for message passing in compositional diffusion models for planning tasks.

---

8. Artificial intelligence for catalyst design and synthesis

URL: View paper

Brief Assessment

AI Catalyst Design[58] focuses on catalyst design using ML models for compound stability prediction, not on compositional generative models or diffusion-based planning with iterative resampling for global coherence.

---

9. Roomtex: Texturing compositional indoor scenes via iterative inpainting

URL: View paper

Brief Assessment

Roomtex[64] focuses on 3D scene texturing through iterative inpainting of visual content, not on compositional generative models for planning. The iterative refinement in Roomtex operates on RGB texture generation for 3D meshes, fundamentally different from the original paper's iterative resampling for propagating information across plan segments in diffusion-based compositional sampling.

---

10. Constructing a 3D Town from a Single Image

URL: View paper

Brief Assessment

The candidate paper (3D Town Construction[61]) focuses on constructing 3D scenes from single images, which is a computer vision task fundamentally different from the original paper's compositional diffusion framework for sequential planning and generation.

Contribution 3: Likelihood-based pruning using DDIM inversion

Description: The approach employs a novel pruning mechanism based on DDIM inversion to approximate local plan likelihoods and filter out incoherent global plans. The method defines a smoothness measure based on diffusion trajectory curvature to identify and eliminate plans with locally inconsistent segments that result from mode-averaging.

This contribution was assessed against **0 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Compositional Diffusion with Guided search for Long-Horizon Planning View paper
- [1] GNN-based long and short term preference modeling for next-location prediction View paper
- [2] Structured motion generation with predictive learning: Proposing subgoal for long-horizon manipulation View paper
- [3] An Long Short-Term Memory Model with Multi-Scale Context Fusion and Attention for Radar Echo Extrapolation View paper
- [4] SepConv-ens: An ensemble of separable convolution-based deep learning models for weather radar echo temporal extrapolation View paper
- [5] Generative trajectory stitching through diffusion composition View paper
- [6] Color Spike Camera Reconstruction via Long Short-Term Temporal Aggregation of Spike Signals View paper
- [7] Generative skill chaining: Long-horizon skill planning with diffusion models View paper
- [8] Extendable long-horizon planning via hierarchical multiscale diffusion View paper
- [9] 4D VQ-GAN: Synthesising Medical Scans at Any Time Point for Personalised Disease Progression Modelling of Idiopathic Pulmonary Fibrosis View paper
- [10] Longâ□□term Motion Inâ□□betweening via Keyframe Prediction View paper
- [11] Multi-temporal dependency handling in video smoke recognition: A holistic approach spanning spatial, short-term, and long-term perspectives View paper
- [12] VideoSSM: Autoregressive Long Video Generation with Hybrid State-Space Memory View paper
- [13] What Do You Need for Diverse Trajectory Stitching in Diffusion Planning? View paper
- [14] Learning to generalize across long-horizon tasks from human demonstrations View paper
- [15] Short-Term Precipitation Radar Echo Extrapolation Method Based on the MS-DD3D-RSTN Network and STLoss Function View paper
- [16] Recurrent network models for human dynamics View paper
- [17] Combining recurrent neural networks and adversarial training for human motion synthesis and control View paper
- [18] 3D-UNet-LSTM: A Deep Learning-Based Radar Echo Extrapolation Model for Convective Nowcasting View paper
- [19] Long-term temporally consistent unpaired video translation from simulated surgical 3d data View paper
- [20] Planning to practice: Efficient online fine-tuning by composing goals in latent space View paper
- [21] Hourly rolling correction of precipitation forecast via convolutional and long shortâ□□term memory networks View paper
- [22] An Improved LSTM-based Method Capturing Temporal Correlations and Using Attention Mechanism for Radar Echo Extrapolation View paper
- [23] ManiLong-Shot: Interaction-Aware One-Shot Imitation Learning for Long-Horizon Manipulation View paper
- [24] A concatenational graph evolution aging model View paper
- [25] Advances in Long-Horizon Planning via Transformer and Diffusion World Models View paper
- [26] Learning long- and short-term dependencies for network intrusion detection View paper
- [27] Temporal User Profiling with LLMs: Balancing Short-Term and Long-Term Preferences for Recommendations View paper
- [28] RoboEnvision: A Long-Horizon Video Generation Model for Multi-Task Robot Manipulation View paper
- [29] Predicting Grain Growth in Polycrystalline Materials Using Deep Learning Time Series Models View paper
- [30] Towards Long-Horizon Vision-Language-Action System: Reasoning, Acting and Memory View paper
- [31] Modular and Multimodal Generative Adversarial Imitation Learning for Modeling Flight Trajectories View paper
- [32] Bioâ□□Inspired Sb2S3 Memristors Supporting Dualâ□□Mode Operation via Selfâ□□Regulating Transport Pathways for Allâ□□inâ□□one Neuromorphic Perception and â□¦ View paper
- [33] LSTTN: A Long-Short Term Transformer-based Spatio-temporal Neural Network for Traffic Flow Forecasting View paper
- [34] h1: Bootstrapping LLMs to Reason over Longer Horizons via Reinforcement Learning View paper
- [35] Long-RVOS: A Comprehensive Benchmark for Long-term Referring Video Object Segmentation View paper
- [36] Multi-step time series forecasting with an ensemble of varied length mixture models View paper
- [37] One-Shot Transfer of Long-Horizon Extrinsic Manipulation Through Contact Retargeting View paper
- [38] Long- and Short- Term Sequential Recommendation with Temporal Interval View paper
- [39] Distilling Knowledge for Short-to-Long Term Trajectory Prediction View paper
- [40] An Integrated Model of Contex, Short-Term, and Long-Term Memory View paper

- [41] Hierarchies, search, and generative models in sequential decision-making View paper
- [42] NOD-TAMP: Generalizable Long-Horizon Planning with Neural Object Descriptors View paper
- [43] Towards Long-Horizon Vision-Language Navigation: Platform, Benchmark and Method View paper
- [44] Data-Driven Prediction of Severe Convection at Deutscher Wetterdienst (DWD): A Brief Overview of Recent Developments View paper
- [45] Decoupling Long- and Short-Term Patterns in Spatiotemporal Inference View paper
- [46] Data Driven Prediction of Severe Convection at DWD. An Overview of Recent Developments View paper
- [47] Statistical Methods View paper
- [48] Example-Driven Model-Based Reinforcement Learning for Solving Long-Horizon Visuomotor Tasks View paper
- [49] Comment on wes-2021-33 View paper
- [50] Compositional Visual Planning via Inference-Time Diffusion Scaling View paper
- [51] Unifying Modern AI with Robotics: Survey on MDPs with Diffusion and Foundation Models View paper
- [52] Compositional Monte Carlo Tree Diffusion for Extendable Planning View paper
- [53] Hybrid Diffusion for Simultaneous Symbolic and Continuous Planning View paper
- [54] Prismatic World Model: Learning Compositional Dynamics for Planning in Hybrid Systems View paper
- [55] Inference-time Scaling of Diffusion Models through Classical Search View paper
- [56] ComposableNav: Instruction-Following Navigation in Dynamic Environments via Composable Diffusion View paper
- [57] Controllable Graph Generation with Diffusion Models via Inference-Time Tree Search Guidance View paper
- [58] Artificial intelligence for catalyst design and synthesis View paper
- [59] AccDiffusion: An Accurate Method for Higher-Resolution Image Generation View paper
- [60] ShapeShift: Towards Text-to-Shape Arrangement Synthesis with Content-Aware Geometric Constraints View paper
- [61] Constructing a 3D Town from a Single Image View paper
- [62] Compositional foundation models for hierarchical planning View paper
- [63] Hydra: A hyper agent for dynamic compositional visual reasoning View paper
- [64] Roomtex: Texturing compositional indoor scenes via iterative inpainting View paper
- [65] Streamlining Robust Constrained Production Optimization: An Integrated Framework Utilizing Automatically Differentiated Gradient from Deep-Learning-Based â¦ View paper
- [66] Multi-view people tracking via hierarchical trajectory composition View paper
- [67] Seismic Data Interpolation via Denoising Diffusion Implicit Models With Coherence-Corrected Resampling View paper