Wed 2024.04.24

Low-ordered Orthogonal Voxel Finite Element with INT8 Tensor Cores for GPU-based Explicit Elastic Wave Propagation Analysis

Tsuyoshi Ichimura, Kohei Fujita, Muneo Hori, Maddegedara Lalith

Faster explicit elastic wavefield simulations are required for large and complex three-dimensional media using a structured finite element method. Such wavefield simulations are suitable for GPUs, which have exhibited improved computational performance in recent years, and the use of GPUs is expected to speed up such simulations. However, available computational performance on GPUs is typically not fully exploited, and the conventional method involves some numerical dispersion. Thus, in this paper, we propose an explicit structured-mesh wavefield simulation method that uses INT8 Tensor Cores and reduces numerical dispersion to speed up computation on GPUs. The proposed method was implemented for GPUs, and its performance was evaluated in a simulation experiment of a real-world problem. The results demonstrate that the proposed method is 17.0 times faster than the conventional method.

link: http://arxiv.org/abs/2404.13683v1

Hyper-SD: Trajectory Segmented Consistency Model for Efficient Image Synthesis

Yuxi Ren, Xin Xia, Yanzuo Lu, Jiacheng Zhang, Jie Wu, Pan Xie, Xing Wang, Xuefeng Xiao

Recently, a series of diffusion-aware distillation algorithms have emerged to alleviate the computational overhead associated with the multi-step inference process of Diffusion Models (DMs). Current distillation techniques often dichotomize into two distinct aspects: i) ODE Trajectory Preservation; and ii) ODE Trajectory Reformulation. However, these approaches suffer from severe performance degradation or domain shifts. To address these limitations, we propose Hyper-SD, a novel framework that synergistically amalgamates the advantages of ODE Trajectory Preservation and Reformulation, while maintaining near-lossless performance during step compression. Firstly, we introduce Trajectory Segmented Consistency Distillation to progressively perform consistent distillation within pre-defined time-step segments, which facilitates the preservation of the original ODE trajectory from a higher-order perspective. Secondly, we incorporate human feedback learning to boost the performance of the model in a low-step regime and mitigate the performance loss incurred by the distillation process. Thirdly, we integrate score distillation to further improve the low-step generation capability of the model and offer the first attempt to leverage a unified LoRA to support the inference process at all steps. Extensive experiments and user studies demonstrate that Hyper-SD achieves SOTA performance from 1 to 8 inference steps for both SDXL and SD1.5. For example, Hyper-SDXL surpasses SDXL-Lightning by +0.68 in CLIP Score and +0.51 in Aes Score in the 1-step inference.

link: http://arxiv.org/abs/2404.13686v1

Detecting Compromised IoT Devices Using Autoencoders with Sequential Hypothesis Testing

Md Mainuddin, Zhenhai Duan, Yingfei Dong

IoT devices fundamentally lack built-in security mechanisms to protect themselves from security attacks. Existing works on improving IoT security mostly focus on detecting anomalous behaviors of IoT devices. However, these existing anomaly detection schemes may trigger an overwhelmingly large number of false alerts, rendering them unusable in detecting compromised IoT devices. In this paper we develop an effective and efficient framework, named CUMAD, to detect compromised IoT devices. Instead of directly relying on individual anomalous events, CUMAD aims to accumulate sufficient evidence in detecting compromised IoT devices, by integrating an autoencoder-based anomaly detection subsystem with a sequential probability ratio test (SPRT)-based sequential hypothesis testing subsystem. CUMAD can effectively reduce the number of false alerts in detecting compromised IoT devices, and moreover, it can detect compromised IoT devices quickly. Our evaluation studies based on the public-domain N-BaIoT dataset show that CUMAD can on

average reduce the false positive rate from about 3.57% using only the autoencoder-based anomaly detection scheme to about 0.5%; in addition, CUMAD can detect compromised IoT devices quickly, with less than 5 observations on average.

link: http://dx.doi.org/10.1109/BigData59044.2023.10386399

A Complete System for Automated 3D Semantic-Geometric Mapping of Corrosion in Industrial Environments

Rui Pimentel de Figueiredo, Stefan Nordborg Eriksen, Ignacio Rodriguez, Simon Bøgh

Corrosion, a naturally occurring process leading to the deterioration of metallic materials, demands diligent detection for quality control and the preservation of metal-based objects, especially within industrial contexts. Traditional techniques for corrosion identification, including ultrasonic testing, radio-graphic testing, and magnetic flux leakage, necessitate the deployment of expensive and bulky equipment on-site for effective data acquisition. An unexplored alternative involves employing lightweight, conventional camera systems, and state-of-the-art computer vision methods for its identification. In this work, we propose a complete system for semi-automated corrosion identification and mapping in industrial environments. We leverage recent advances in LiDAR-based methods for localization and mapping, with vision-based semantic segmentation deep learning techniques, in order to build semantic-geometric maps of industrial environments. Unlike previous corrosion identification systems available in the literature, our designed multi-modal system is low-cost, portable, semi-autonomous and allows collecting large datasets by untrained personnel. A set of experiments in an indoor laboratory environment, demonstrate quantitatively the high accuracy of the employed LiDAR based 3D mapping and localization system, with less then $0.05m$ and 0.02m average absolute and relative pose errors. Also, our data-driven semantic segmentation model, achieves around 70\% precision when trained with our pixel-wise manually annotated dataset.

link: http://arxiv.org/abs/2404.13691v1

A sustainable development perspective on urban-scale roof greening priorities and benefits

Jie Shao, Wei Yao, Lei Luo, Linzhou Zeng, Zhiyi He, Puzuo Wang, Huadong Guo

Greenspaces are tightly linked to human well-being. Yet, rapid urbanization has exacerbated greenspace exposure inequality and declining human life quality. Roof greening has been recognized as an effective strategy to mitigate these negative impacts. Understanding priorities and benefits is crucial to promoting green roofs. Here, using geospatial big data, we conduct an urban-scale assessment of roof greening at a single building level in Hong Kong from a sustainable development perspective. We identify that 85.3\% of buildings reveal potential and urgent demand for roof greening. We further find green roofs could increase greenspace exposure by \textasciitilde61\% and produce hundreds of millions (HK\$) in economic benefits annually but play a small role in urban heat mitigation (\textasciitilde0.15\degree{C}) and annual carbon emission offsets (\textasciitilde0.8\%). Our study offers a comprehensive assessment of roof greening, which could provide reference for sustainable development in cities worldwide, from data utilization to solutions and findings.

link: http://arxiv.org/abs/2404.13692v1

PV-S3: Advancing Automatic Photovoltaic Defect Detection using Semi-Supervised Semantic Segmentation of Electroluminescence Images

Abhishek Jha, Yogesh Rawat, Shruti Vyas

Photovoltaic (PV) systems allow us to tap into all abundant solar energy, however they require regular maintenance for high efficiency and to prevent degradation. Traditional manual health check, using Electroluminescence (EL) imaging, is expensive and logistically challenging making automated defect detection essential. Current automation approaches require extensive manual expert labeling, which is time-consuming, expensive, and prone to errors. We propose PV-S3 (Photovoltaic-Semi Supervised Segmentation), a Semi-Supervised Learning approach for semantic

segmentation of defects in EL images that reduces reliance on extensive labeling. PV-S3 is a Deep learning model trained using a few labeled images along with numerous unlabeled images. We introduce a novel Semi Cross-Entropy loss function to train PV-S3 which addresses the challenges specific to automated PV defect detection, such as diverse defect types and class imbalance. We evaluate PV-S3 on multiple datasets and demonstrate its effectiveness and adaptability. With merely 20% labeled samples, we achieve an absolute improvement of 9.7% in IoU, 29.9% in Precision, 12.75% in Recall, and 20.42% in F1-Score over prior state-of-the-art supervised method (which uses 100% labeled samples) on UCF-EL dataset (largest dataset available for semantic segmentation of EL images) showing improvement in performance while reducing the annotation costs by 80%.

link: http://arxiv.org/abs/2404.13693v1

Resampling-free Particle Filters in High-dimensions

Akhilan Boopathy, Aneesh Muppidi, Peggy Yang, Abhiram Iyer, William Yue, Ila Fiete

State estimation is crucial for the performance and safety of numerous robotic applications. Among the suite of estimation techniques, particle filters have been identified as a powerful solution due to their non-parametric nature. Yet, in high-dimensional state spaces, these filters face challenges such as 'particle deprivation' which hinders accurate representation of the true posterior distribution. This paper introduces a novel resampling-free particle filter designed to mitigate particle deprivation by forgoing the traditional resampling step. This ensures a broader and more diverse particle set, especially vital in high-dimensional scenarios. Theoretically, our proposed filter is shown to offer a near-accurate representation of the desired posterior distribution in high-dimensional contexts. Empirically, the effectiveness of our approach is underscored through a high-dimensional synthetic state estimation task and a 6D pose estimation derived from videos. We posit that as robotic systems evolve with greater degrees of freedom, particle filters tailored for high-dimensional state spaces will be indispensable.

link: http://arxiv.org/abs/2404.13698v1

Semantic-Rearrangement-Based Multi-Level Alignment for Domain Generalized Segmentation

Guanlong Jiao, Chenyangguang Zhang, Haonan Yin, Yu Mo, Biqing Huang, Hui Pan, Yi Luo, Jingxian Liu

Domain generalized semantic segmentation is an essential computer vision task, for which models only leverage source data to learn the capability of generalized semantic segmentation towards the unseen target domains. Previous works typically address this challenge by global style randomization or feature regularization. In this paper, we argue that given the observation that different local semantic regions perform different visual characteristics from the source domain to the target domain, methods focusing on global operations are hard to capture such regional discrepancies, thus failing to construct domain-invariant representations with the consistency from local to global level. Therefore, we propose the Semantic-Rearrangement-based Multi-Level Alignment (SRMA) to overcome this problem. SRMA first incorporates a Semantic Rearrangement Module (SRM), which conducts semantic region randomization to enhance the diversity of the source domain sufficiently. A Multi-Level Alignment module (MLA) is subsequently proposed with the help of such diversity to establish the global-regional-local consistent domain-invariant representations. By aligning features across randomized samples with domain-neutral knowledge at multiple levels, SRMA provides a more robust way to handle the source-target domain gap. Extensive experiments demonstrate the superiority of SRMA over the current state-of-the-art works on various benchmarks.

link: http://arxiv.org/abs/2404.13701v1

PEMMA: Parameter-Efficient Multi-Modal Adaptation for Medical Image Segmentation

Nada Saadi, Numan Saeed, Mohammad Yaqub, Karthik Nandakumar

Imaging modalities such as Computed Tomography (CT) and Positron Emission Tomography (PET) are key in cancer detection, inspiring Deep Neural Networks (DNN) models that merge these scans for tumor segmentation. When both CT and PET scans are available, it is common to combine them as two channels of the input to the segmentation model. However, this method requires both scan types during training and inference, posing a challenge due to the limited availability of PET scans, thereby sometimes limiting the process to CT scans only. Hence, there is a need to develop a flexible DNN architecture that can be trained/updated using only CT scans but can effectively utilize PET scans when they become available. In this work, we propose a parameter-efficient multi-modal adaptation (PEMMA) framework for lightweight upgrading of a transformer-based segmentation model trained only on CT scans to also incorporate PET scans. The benefits of the proposed approach are two-fold. Firstly, we leverage the inherent modularity of the transformer architecture and perform low-rank adaptation (LoRA) of the attention weights to achieve parameter-efficient adaptation. Secondly, since the PEMMA framework attempts to minimize cross modal entanglement, it is possible to subsequently update the combined model using only one modality, without causing catastrophic forgetting of the other modality. Our proposed method achieves comparable results with the performance of early fusion techniques with just 8% of the trainable parameters, especially with a remarkable +28% improvement on the average dice score on PET scans when trained on a single modality.

link: http://arxiv.org/abs/2404.13704v1

Concept Arithmetics for Circumventing Concept Inhibition in Diffusion Models

Vitali Petsiuk, Kate Saenko

Motivated by ethical and legal concerns, the scientific community is actively developing methods to limit the misuse of Text-to-Image diffusion models for reproducing copyrighted, violent, explicit, or personal information in the generated images. Simultaneously, researchers put these newly developed safety measures to the test by assuming the role of an adversary to find vulnerabilities and backdoors in them. We use compositional property of diffusion models, which allows to leverage multiple prompts in a single image generation. This property allows us to combine other concepts, that should not have been affected by the inhibition, to reconstruct the vector, responsible for target concept generation, even though the direct computation of this vector is no longer accessible. We provide theoretical and empirical evidence why the proposed attacks are possible and discuss the implications of these findings for safe model deployment. We argue that it is essential to consider all possible approaches to image generation with diffusion models that can be employed by an adversary. Our work opens up the discussion about the implications of concept arithmetics and compositional inference for safety mechanisms in diffusion models. Content Advisory: This paper contains discussions and model-generated content that may be considered offensive. Reader discretion is advised. Project page: https://cs-people.bu.edu/vpetsiuk/arc

link: http://arxiv.org/abs/2404.13706v1

SVGEditBench: A Benchmark Dataset for Quantitative Assessment of LLM's SVG Editing Capabilities

Kunato Nishina, Yusuke Matsui

Text-to-image models have shown progress in recent years. Along with this progress, generating vector graphics from text has also advanced. SVG is a popular format for vector graphics, and SVG represents a scene with XML text. Therefore, Large Language Models can directly process SVG code. Taking this into account, we focused on editing SVG with LLMs. For quantitative evaluation of LLMs' ability to edit SVG, we propose SVGEditBench. SVGEditBench is a benchmark for assessing the LLMs' ability to edit SVG code. We also show the GPT-4 and GPT-3.5 results when evaluated on the proposed benchmark. In the experiments, GPT-4 showed superior performance to GPT-3.5 both quantitatively and qualitatively. The dataset is available at https://github.com/mti-lab/SVGEditBench.

link: http://arxiv.org/abs/2404.13710v1

ArtNeRF: A Stylized Neural Field for 3D-Aware Cartoonized Face Synthesis

Zichen Tang, Hongyu Yang

Recent advances in generative visual models and neural radiance fields have greatly boosted 3D-aware image synthesis and stylization tasks. However, previous NeRF-based work is limited to single scene stylization, training a model to generate 3D-aware cartoon faces with arbitrary styles remains unsolved. We propose ArtNeRF, a novel face stylization framework derived from 3D-aware GAN to tackle this problem. In this framework, we utilize an expressive generator to synthesize stylized faces and a triple-branch discriminator module to improve the visual quality and style consistency of the generated faces. Specifically, a style encoder based on contrastive learning is leveraged to extract robust low-dimensional embeddings of style images, empowering the generator with the knowledge of various styles. To smooth the training process of cross-domain transfer learning, we propose an adaptive style blending module which helps inject style information and allows users to freely tune the level of stylization. We further introduce a neural rendering module to achieve efficient real-time rendering of images with higher resolutions. Extensive experiments demonstrate that ArtNeRF is versatile in generating high-quality 3D-aware cartoon faces with arbitrary styles.

link: http://arxiv.org/abs/2404.13711v1

TF2AIF: Facilitating development and deployment of accelerated AI models on the cloud-edge continuum

Aimilios Leftheriotis, Achilleas Tzenetopoulos, George Lentaris, Dimitrios Soudris, Georgios Theodoridis

The B5G/6G evolution relies on connect-compute technologies and highly heterogeneous clusters with HW accelerators, which require specialized coding to be efficiently utilized. The current paper proposes a custom tool for generating multiple SW versions of a certain AI function input in high-level language, e.g., Python TensorFlow, while targeting multiple diverse HW+SW platforms. TF2AIF builds upon disparate tool-flows to create a plethora of relative containers and enable the system orchestrator to deploy the requested function on any peculiar node in the cloud-edge continuum, i.e., to leverage the performance/energy benefits of the underlying HW upon any circumstances. TF2AIF fills an identified gap in today's ecosystem and facilitates research on resource management or automated operations, by demanding minimal time or expertise from users.

link: http://arxiv.org/abs/2404.13715v1

A Practical Multilevel Governance Framework for Autonomous and Intelligent Systems

Lukas D. Pöhler, Klaus Diepold, Wendell Wallach

Autonomous and intelligent systems (AIS) facilitate a wide range of beneficial applications across a variety of different domains. However, technical characteristics such as unpredictability and lack of transparency, as well as potential unintended consequences, pose considerable challenges to the current governance infrastructure. Furthermore, the speed of development and deployment of applications outpaces the ability of existing governance institutions to put in place effective ethical-legal oversight. New approaches for agile, distributed and multilevel governance are needed. This work presents a practical framework for multilevel governance of AIS. The framework enables mapping actors onto six levels of decision-making including the international, national and organizational levels. Furthermore, it offers the ability to identify and evolve existing tools or create new tools for guiding the behavior of actors within the levels. Governance mechanisms enable actors to shape and enforce regulations and other tools, which when complemented with good practices contribute to effective and comprehensive governance.

link: http://arxiv.org/abs/2404.13719v1

The Framework of a Design Process Language

Arnulf Hagen

The thesis develops a view of design in a concept formation framework and outlines a language to describe both the object of the design and the process of designing. The unknown object at the outset of the design work may be seen as an unknown concept that the designer is to define. Throughout the process, she develops a description of this object by relating it to known concepts. The search stops when the designer is satisfied that the design specification is complete enough to satisfy the requirements from it once built. It is then a collection of propositions that all contribute towards defining the design object - a collection of sentences describing relationships between the object and known concepts. Also, the design process itself may be described by relating known concepts - by organizing known abilities into particular patterns of activation, or mobilization. In view of the demands posed to a language to use in this concept formation process, the framework of a Design Process Language (DPL) is developed. The basis for the language are linguistic categories that act as classes of relations used to combine concepts, containing relations used for describing process and object within the same general system, with some relations being process specific, others being object specific, and with the bulk being used both for process and object description. Another outcome is the distinction of modal relations, or relations describing futurity, possibility, willingness, hypothetical events, and the like. The design process almost always includes aspects such as these, and it is thus necessary for a language facilitating design process description to support such relationships to be constructed. The DPL is argued to be a foundation whereupon to build a language that can be used for enabling computers to be more useful - act more intelligently - in the design process.

link: http://arxiv.org/abs/2404.13721v1