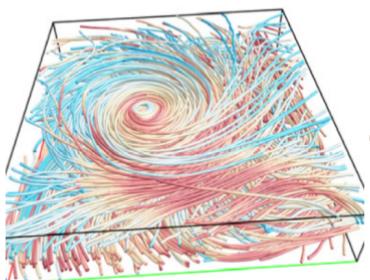




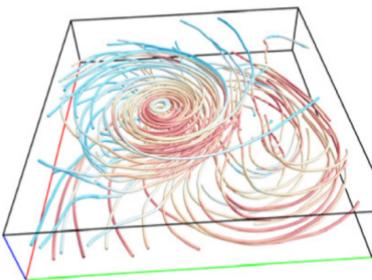
ML Based Large-Scale Scientific Data Compression for Interactive VR Game Creation

21074937, Ph.D. Student, CSE, EMIA 6500K, HKUST



Original Scientific Data

Data
Compressio
n



Compressed Scientific Data

Real-time
Interaction



Interactive VR Game

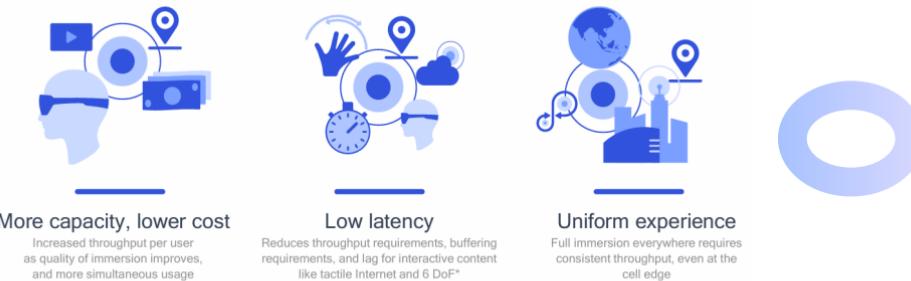
Wenshuo ZHANG

2025-04-28

1. Introduction / Background
2. Contribution / Methods
3. Demo Showcase
4. Discussion for possible next proj.

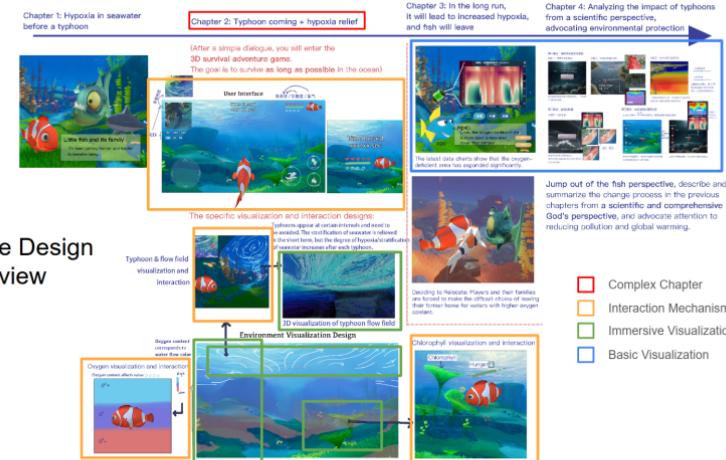
VR Need Fast Response

Interactive and immersive VR games require low-latency content loading and processing



Requirement of VR

Game Design Overview



Requirement of VR Game in scientific visualization and interaction scenario

- Immersive plot design and user experience
- Accurate scientific data visualization
- Low-latency scientific data switching when interacting
- Very small data capacity for edge computing (VR computing)

SciVis is Time Consuming

Scientific data files are very large, and processing them can cause significant delays

nc-demo-dataset

Data Card Code (0) Discussion (0) Suggestions (0) Settings

PRE_case241_3D_fields_050.nc (237.09 MB)

About this file

This file does not have a description yet.

Unable to show preview

Previews for binary data are not supported

Data Explorer

Version 1 (237.09 MB)

PRE_case241_3D_fields_050.nc

PRE_case241_3D_fields_050_f

Summary

2 files

+ New Version

Example data of one day

The volume of scientific data is quite huge.

(1) Over 200 MB per day. Left Fig. is an example data showing recorded Indicator data on 2019.01.01 in Guangdong-Hong Kong-Macao Greater Bay Area.

(2) Data from more than one day is involved. VR game involves long time scientific data visualization and interaction, i.e., over 30 days or the whole year, for better storytelling and data showing.

(3) The delay caused by the increase in data volume will be seriously aggravated. Data volume per day multiplied by day amount involved will be an astronomical figure

Data Governance and Compression Methods are in Great Need in Scientific Visualization Scene

Key Question

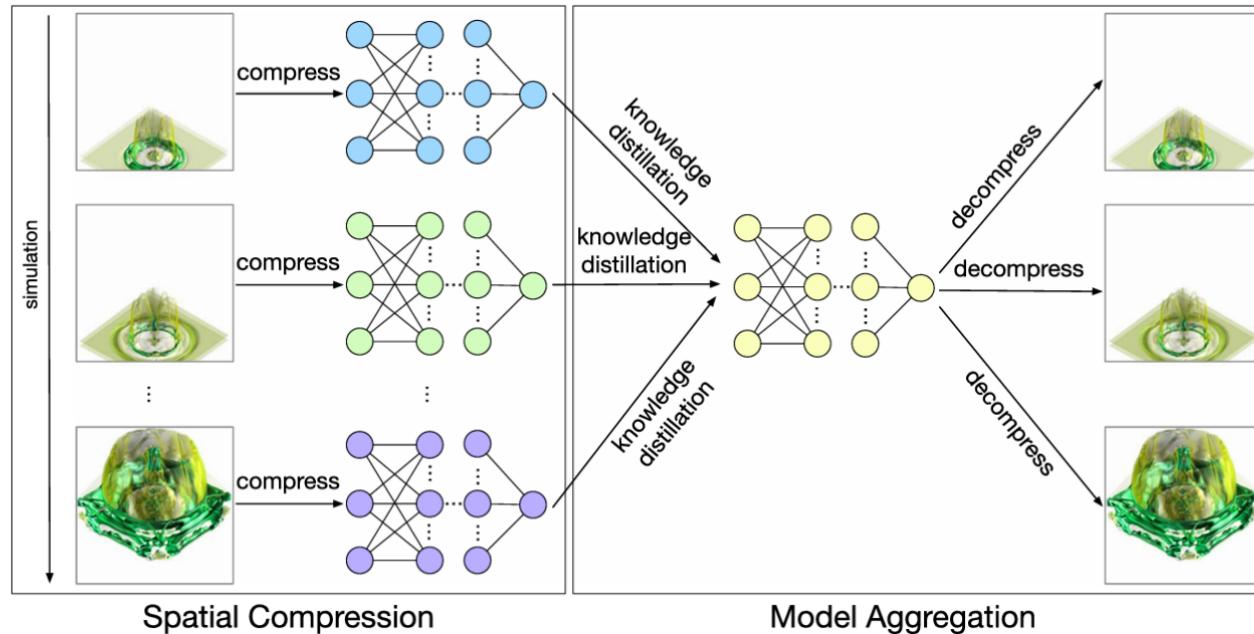
Question: How to speed up data processing process?

- Sub Q.1: How to fully use graphics computing ability especially when the parallel processing and rendering capabilities of graphics processing unit are growing so fast?
- Sub Q.2: How to maximize data compression and merging to adapt to edge computing scenarios (VR or local computing scenarios)
- Sub Q.3: How to ensure that the final result of the compressed data after post processing is consistent with the uncompressed original scientific data?

Using neural networks to compress and learn scientific data in a NeRF-like manner is a very good solution

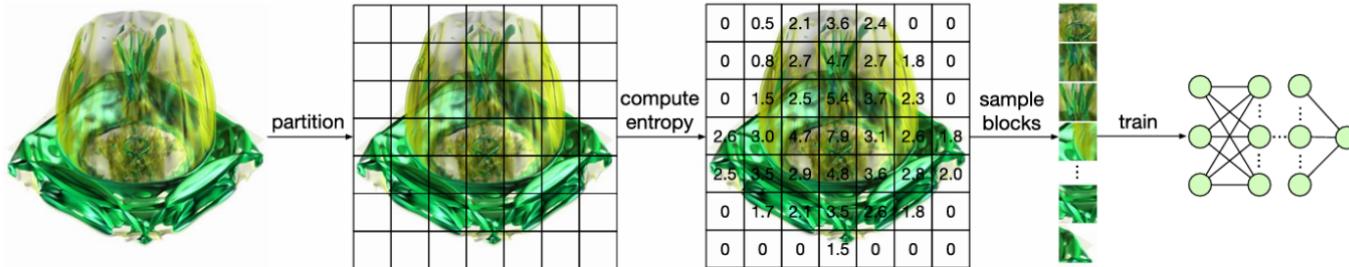
1. Introduction / Background
2. Contribution / Methods
3. Demo Showcase
4. Discussion for possible next proj.

Scientific Data Compression Framework



It consists of spatial compression and model aggregation. In spatial compression, each time step is independently compressed using a tiny network. Then, in model aggregation, a KD algorithm is utilized to merge these into single model.

Each Step Compression



The data is partitioned into a set of non-overlapping blocks and sample blocks based on the entropy values for network optimization. When infer, model accept time step index to generate learnt data.

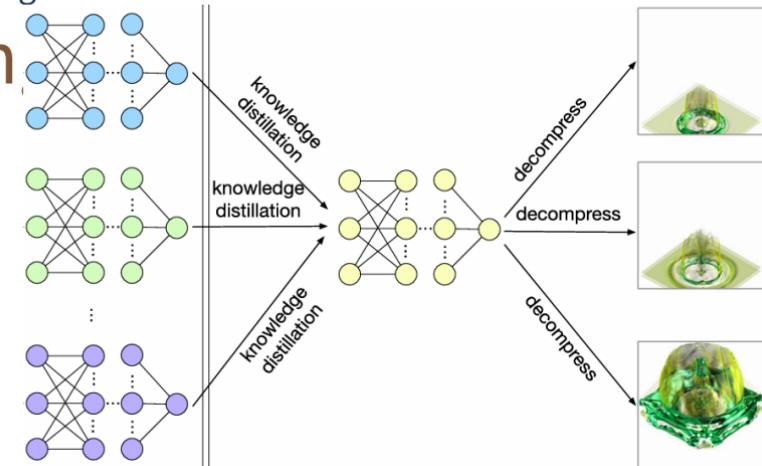
KD Algorithm for Model Mergin

To achieve extracting knowledge from these teacher models and inject it into a tiny model to reduce the model storage cost, we use a KD algorithm to aggregate the teacher models. Given a set of models $M = \{M_1, M_2, \dots, M_T\}$ and adistilled model M_{distill} , the

KDloss

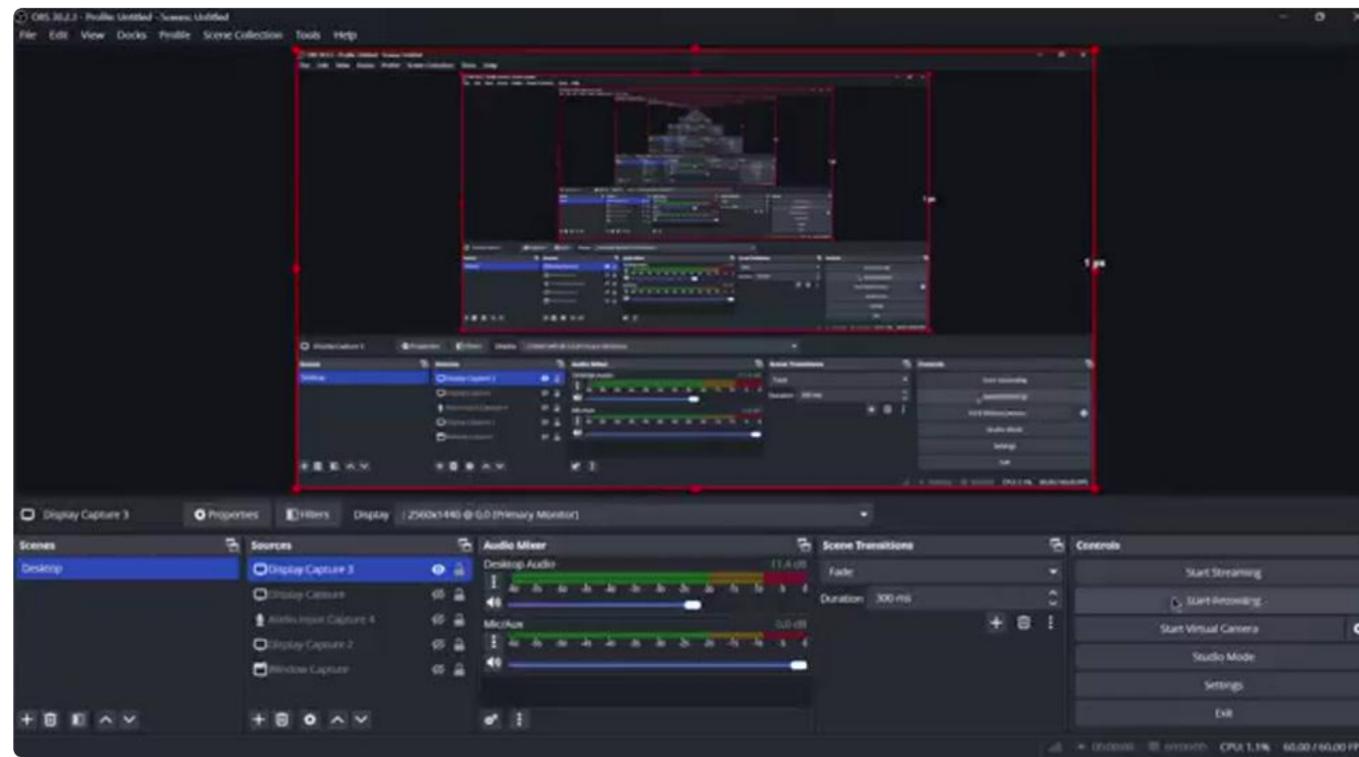
$$\mathcal{L}_{\text{KD}} = \sum_{t=1}^T \sum_{c \in C} \|M_{\text{distill}}(c, t) - M_t(c)\|^2$$

After this loss is applied, models set M is distilled into $M_{\text{distilled}}$



1. Introduction / Background
2. Contribution / Methods
3. Demo Showcase
4. Discussion for possible next proj.

Demo Video



Displaying and processing one day's scientific data (over 200 MB) only needs 0.75s on Nvidia 3070 and model is only

1. Introduction / Background
2. Contribution / Methods
3. Demo Showcase
4. Discussion for possible next proj.

Traditional Practice for VR Environment Creation

Example of general VR environment and a well polished rendered scenario in architectural design



Example of VR environment



Example of architectural design

Time-consuming and
laborious

What you really perceive is what I expect you to see



Real VR
environment

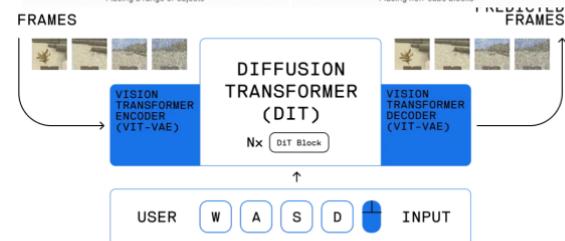


Rerceived
VR
environment



User with VR glass

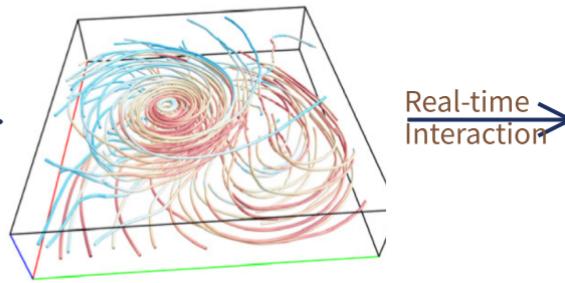
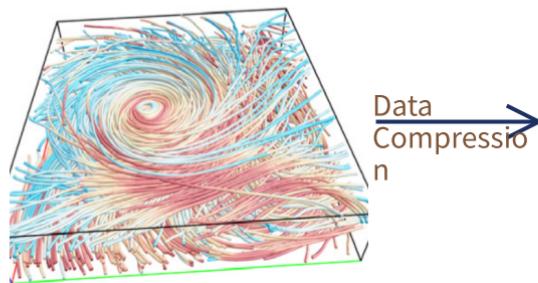
Controlled Oasis-like





ML Based Large-Scale Scientific Data Compression for Interactive VR Game Creation

21074937, Ph.D. Student, CSE, EMIA 6500K, HKUST



Interactive VR Game

Thanks for your attention!