

State-of-the-art in Large-Scale Volume Visualization Beyond Structured Data

J. Sarton¹, S. Zellmann², S. Demirci³, U. Güdükbay³,
W. Alexandre-Barff⁴, L. Lucas⁴, J.M. Dischler¹, S. Wesner², I.
Wald⁵

¹ University of Strasbourg, France

² University of Cologne, Germany

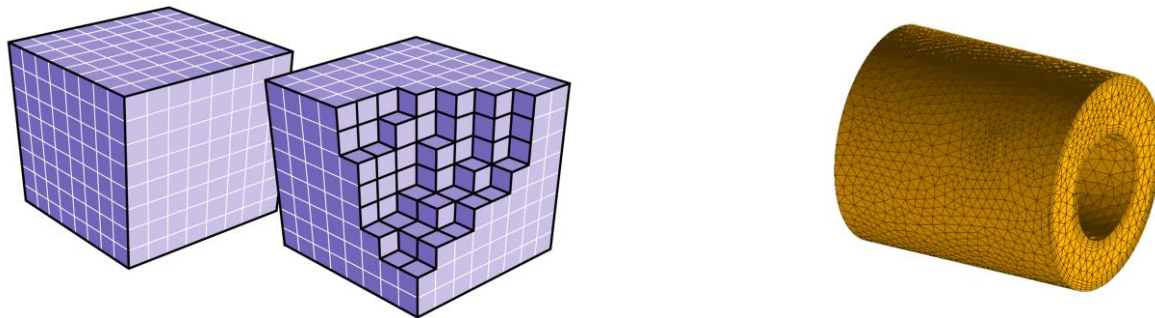
³ Bilkent University, Ankara, Turkey

⁴ University of Reims Champagne-Ardenne, France



Context and motivation

Volume data produced by acquisition or simulation



Scientific visualization to: analyze datasets, extract information, guide phenomenon modeling, validate or invalidate models, evaluate experimental results, ...

Volume rendering techniques: used to produce 2D images from this data

Context and motivation

Volume data produced by acquisition or simulation

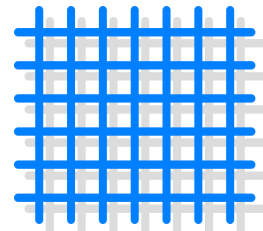
-> Usually massive: Large-scale volume data

Scientific visualization Memory subsystem are the primary bottleneck

Large-scale volume rendering techniques to cope with the input data and the auxiliary data structures needed to render them

Context and motivation

Direct volume rendering: historically associated with **regular grids** for medical imaging, microscopy, ...



Simple representation (implicit topology)

Regular representation **adapted to GPU** architecture and texture units

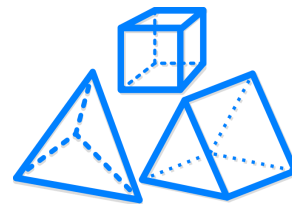
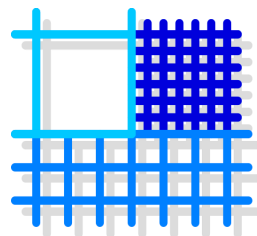
High resolution → **huge datasets** (up to 10^{12} voxels)

Context and motivation

Better adapt to the domain of large-scale simulations:

- hierarchical representations
- unstructured volume data

→ reduce memory consumption



But

- these representations need **large and complex structures** for efficient interactive visualization
- large-scale simulations produced: time series, multi-variate volume data...

Large-scale direct volume rendering techniques

- ray tracing techniques
- sampling and traversal methods of visualization algorithms
- in-core or out-of-core methods
- single or multi CPU/GPU approaches

And hardware-accelerated ray tracing

Not discuss: in-situ visualization solutions

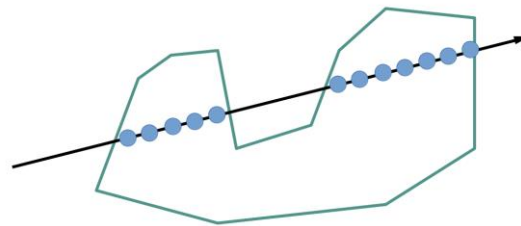
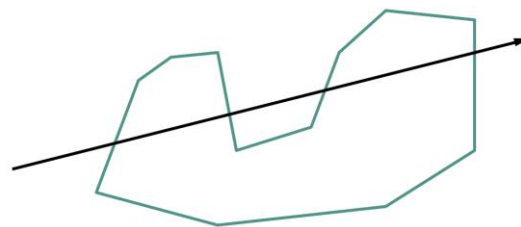
Volume rendering with ray-traversal

- **Traversal**

obtaining the domains $[t_{min}, t_{max}]$

- **Sampling**

volume lookup and classification
(with a Transfer Function)



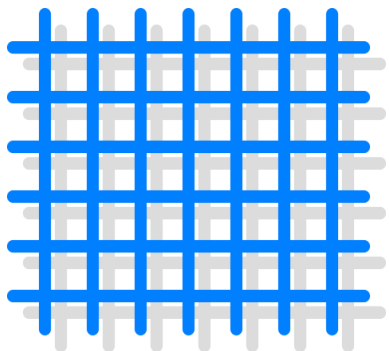
Hardware acceleration

Specific hardware operations for ray-tracing

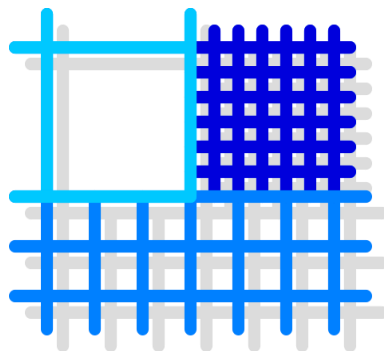
- BVH traversal
- Ray/triangle intersections

Emerging technology relevant to many of the papers discussed

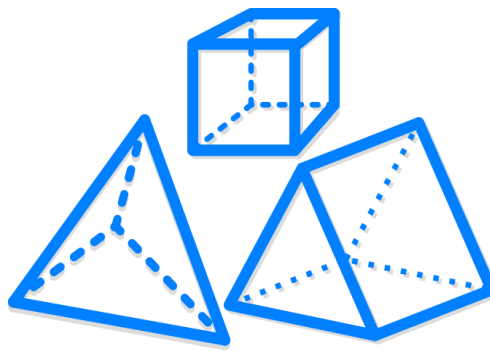
Data driven structure



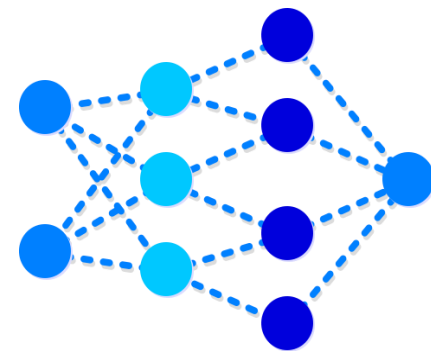
**Structured
volume data**



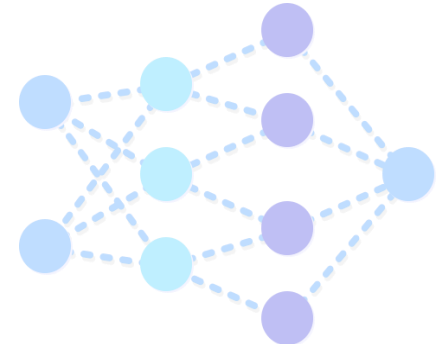
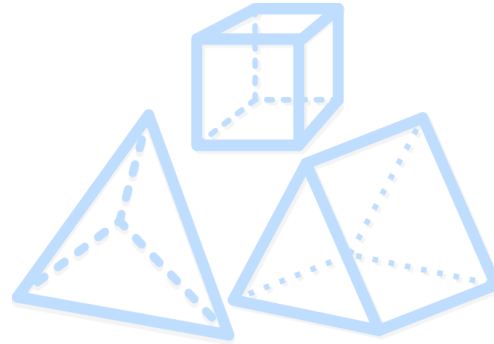
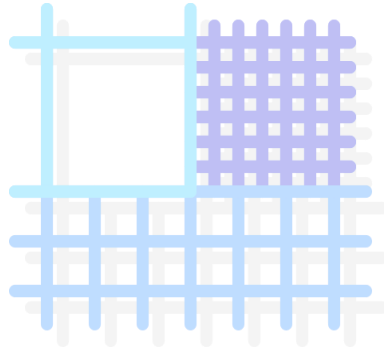
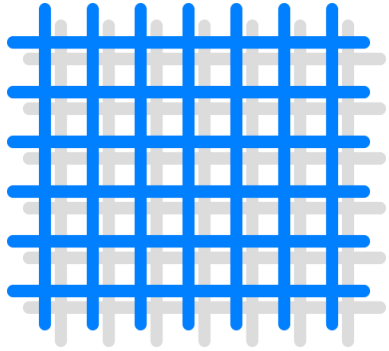
**Adaptive Mesh
Refinement
volume data**



**Unstructured
volume data**

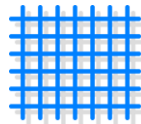


**Compressed
and neural
representations**



Structured Volume Data

Background



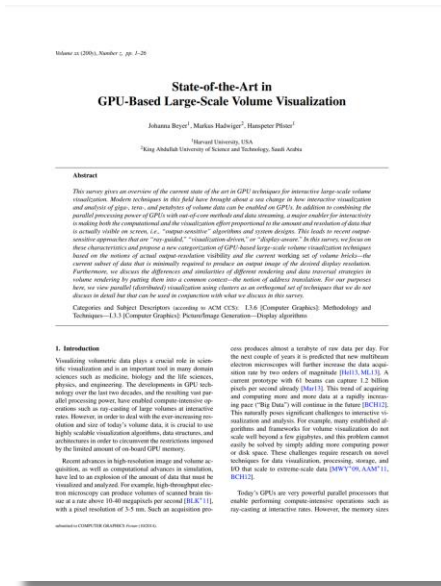
Structured Volume
Data

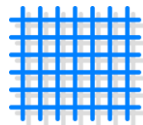
Cartesian grids

GPU techniques for interactive large-scale (structured) volume visualization

Beyer et al. State-of-the-art in GPU-based large-scale volume visualization, Computer Graphics Forum (2015)

Volume rendering of regular grids and mostly review methods for scalar data and a single time step





Structured Volume
Data

Recent works focus on:

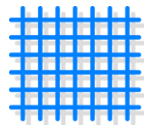
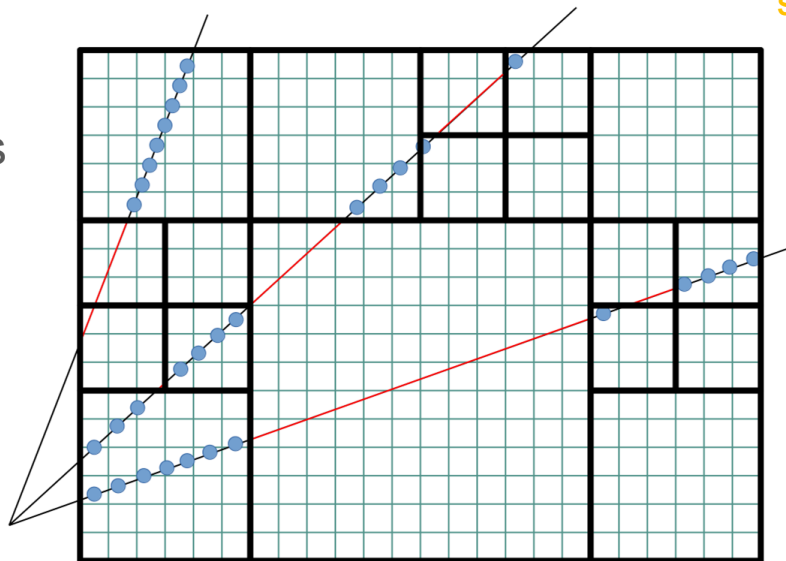
- Reduce pressure on the memory subsystem by reducing the number of samples: **Empty space skipping data structures**
- Stream volume data that does not fully fit into the node's main memory: **Paging and out-of-core approaches**

Empty space-skipping

Grid and octree-based methods

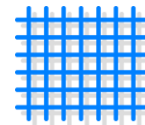
min/max trees from coarser regions
of space : octree leaves

Hadwiger et al. *SparseLeap: Efficient empty
space skipping for large-scale volume
rendering*, IEEE TVCG (2018)



Structured Volume
Data

Empty space skipping



Structured Volume
Data

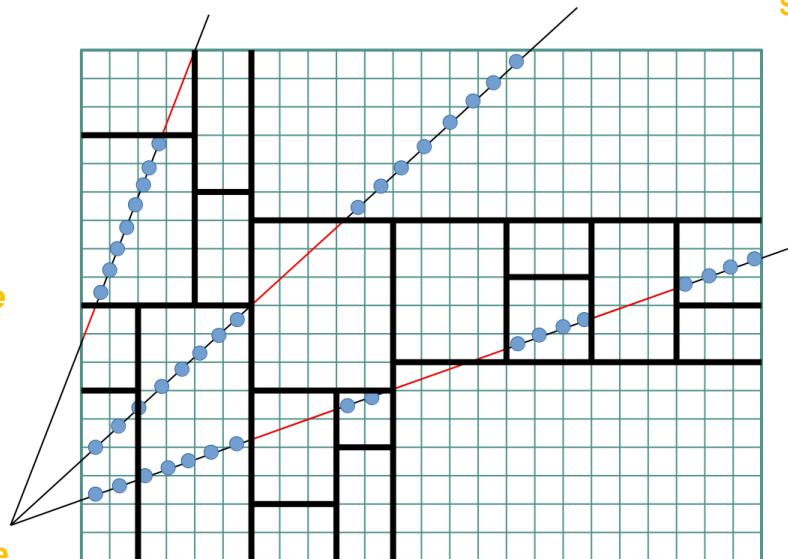
High-quality k-d trees

build an exact data structure
according to TF updates

Vidal et al. *Simple empty-space removal for interactive volume rendering*, Journal of Graphics Tools (2008)

Zellmann et al. *Rapid k-d tree construction for sparse volume data*, EGPGV (2018)

Zellmann et al. *Binned k-d tree construction for sparse volume data on multi-core and GPU systems*, IEEE TVCG (2021)



Empty space skipping

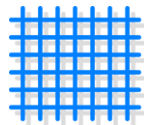
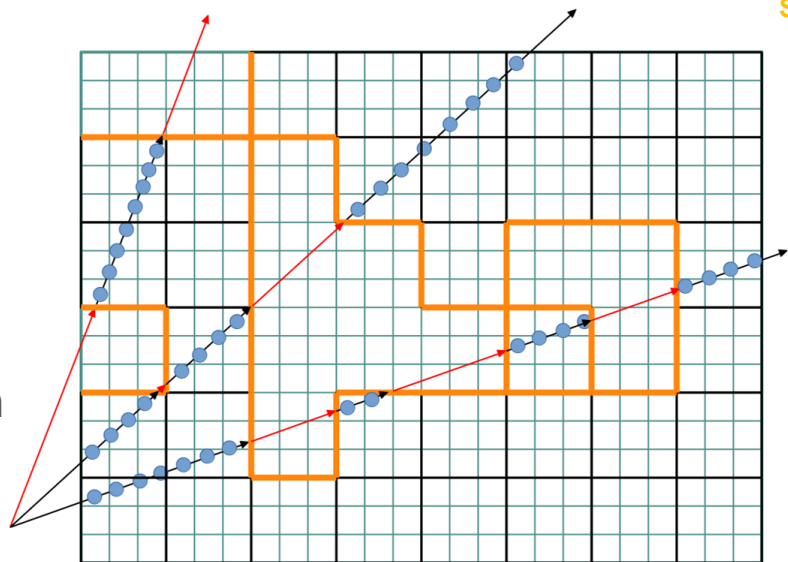
Hardware accelerated ray-tracing based space skipping

BVH traversal

Ganter and Manzke, *An analysis of region clustered BVH volume rendering on GPU*, Computer Graphics Forum (2019)

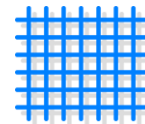
BVH traversal + ray/triangle intersection

Wald et al. *Faster RTX-accelerated empty space skipping using triangulated active region boundary geometry*, EGPGV (2021)



Structured Volume
Data

Paging and out-of-core



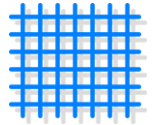
Structured Volume
Data

When volume data does not fully fit into the node's main memory

Pageable memory to stream data to the GPU or asynchronously fetch the data from disk

- Out-of-core volume rendering on multi-core CPU architectures
- GPU out-of-core on-demand rendering and processing

Paging and out-of-core



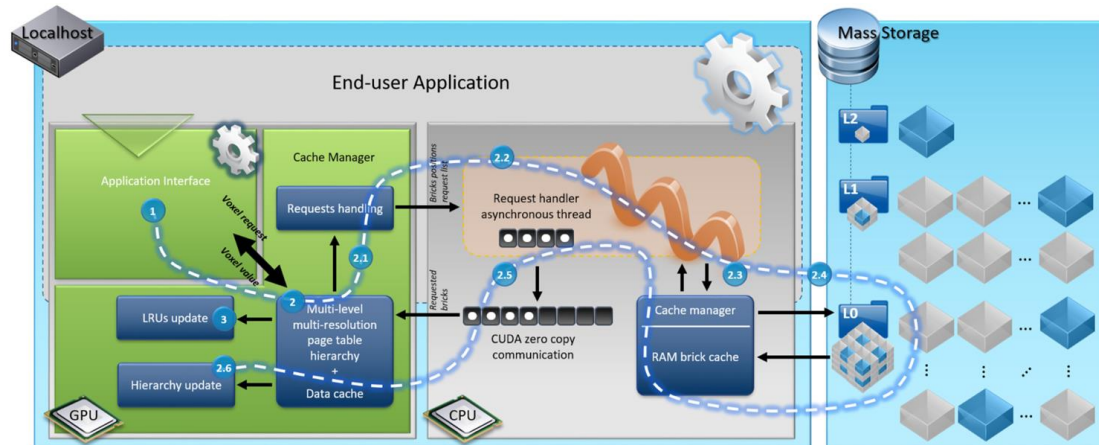
Structured Volume
Data

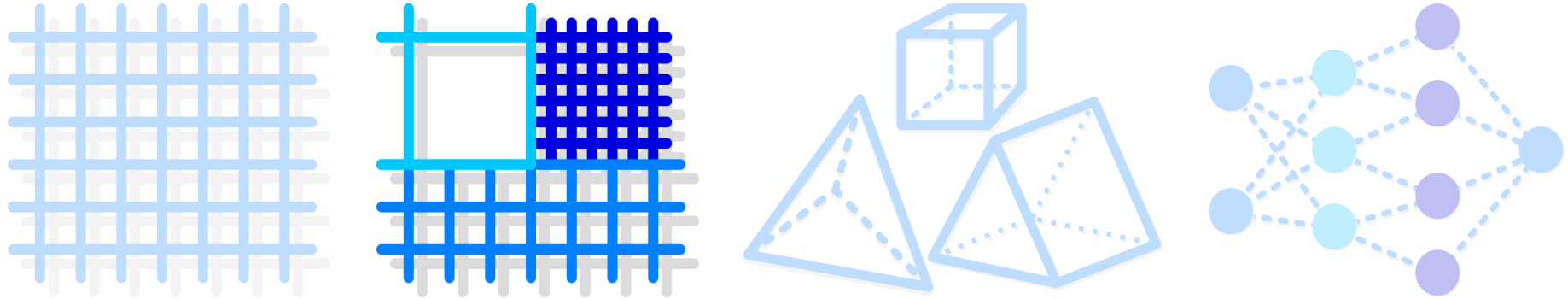
GPU out-of-core on-demand rendering and processing

Multi-resolution virtual addressing structure fully managed on GPU

Asynchronous and demand-driven system can scale well for both visualization and image processing applications

Sarton et al. *Interactive visualization and on-demand processing of large volume data: A fully GPU-based out-of-core approach*, IEEE TVCG (2020)

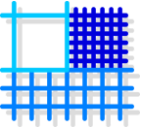




Adaptive Mesh Refinement Volume Data

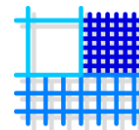
AMR data: terminology & definitions

- Different flavors of AMR. State-art focuses on type where subgrids are Cartesian
- Each cell stores its refinement level $l=\{0,1,\dots\}$
- Assume that cells are voxels (unit length)
- Refinement level determines voxel size: $s=2^l$
- Blockstructured AMR, Octree AMR, wide Octrees, etc.

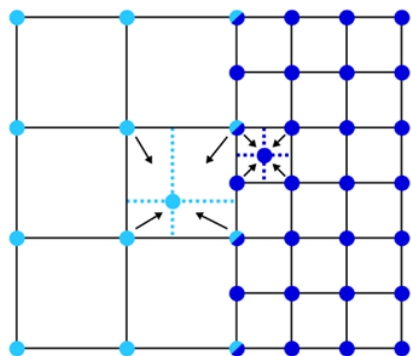


Adaptive Mesh
Refinement
Volume Data

Motivating problem

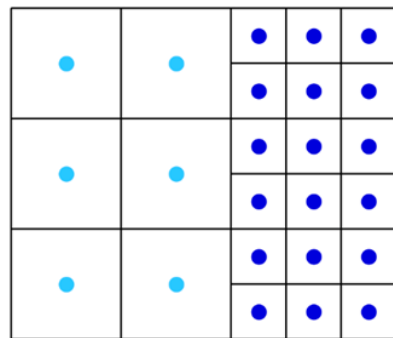


Adaptive Mesh
Refinement
Volume Data



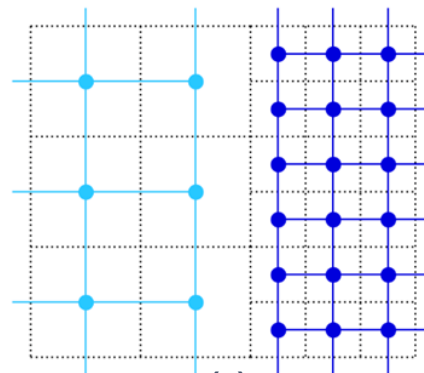
(a)

Vertex-centered
AMR:
Simple
reconstruction with
trilern



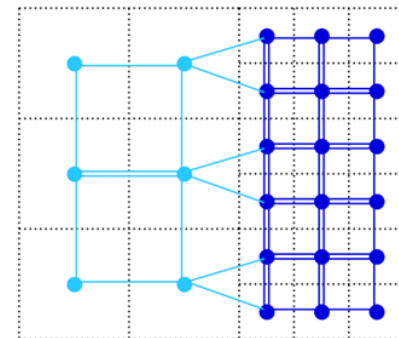
(b)

Cell-centered
AMR



(c)

T-Junction
Boundaries



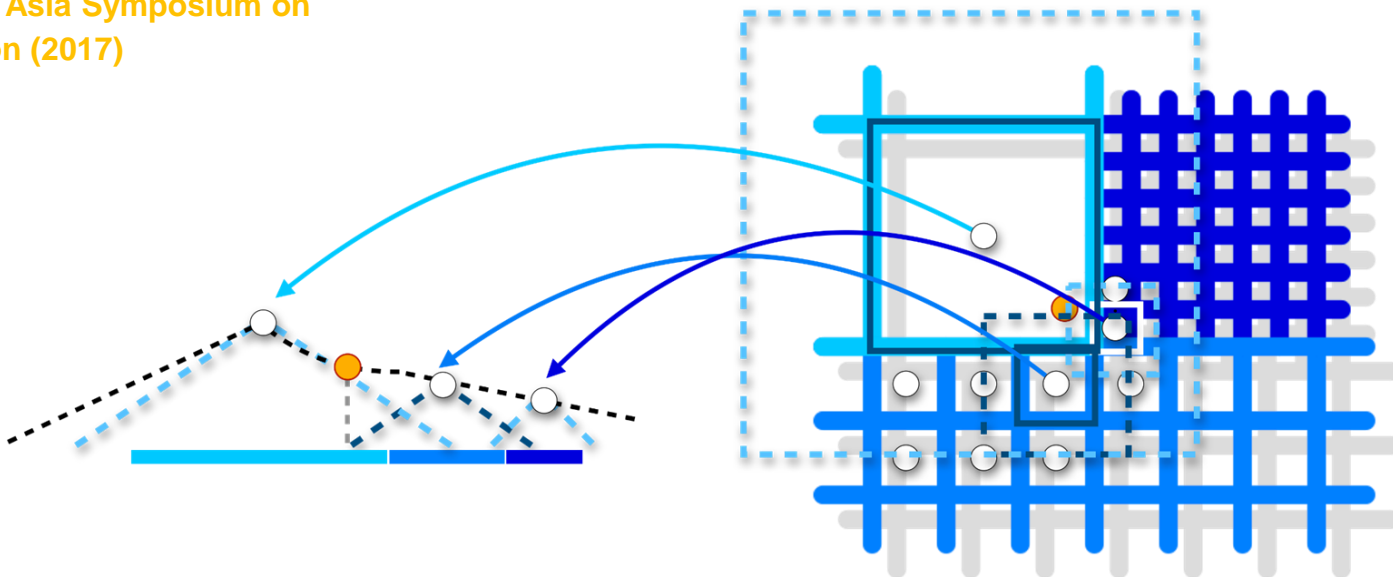
(d)

Stitching

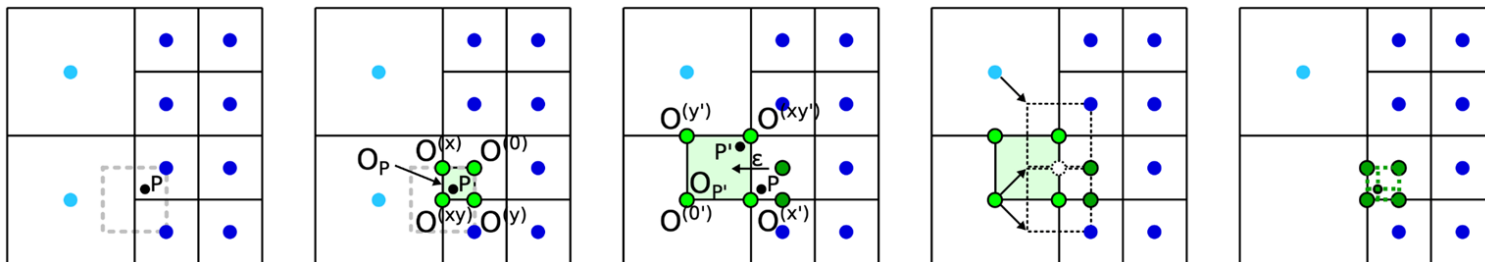
Hat-Shaped Basis Functions

Wald et al. *CPU volume rendering
of adaptive mesh refinement data*,
SIGGRAPH Asia Symposium on
Visualization (2017)


Adaptive Mesh
Refinement
Volume Data



Octant Method



Octant method not an interpolator on its own

Provides tessellation of base domain into dual cells—w/o explicitly storing dual cells

Assignment of dual corner values requires interpolant (e.g., hat basis functions)

Wang et al. *CPU isosurface ray tracing of adaptive mesh refinement data*, **IEEE TVCG** (2019)

ExaBricks Data Structure



Problem: high-quality interpolators require to locate multiple cells per single sample

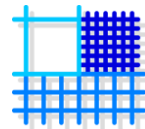
No adjacency information nor refinement level of neighboring cells available

Each located cell requires traversal of hierarchy (kd-tree over cells, etc.)

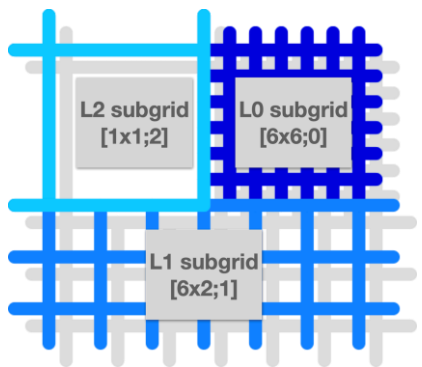
Prohibitive, especially on GPUs

Wald et al. *Ray Tracing Structured
AMR Data Using ExaBricks*, **IEEE
TVCG (2021)**

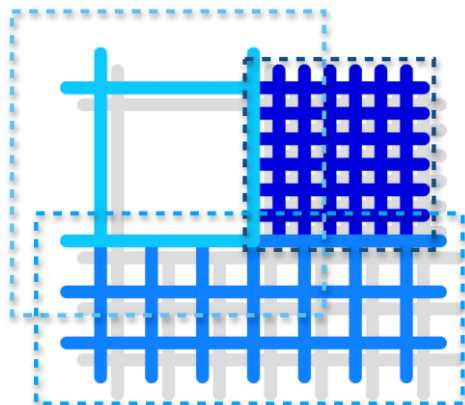
ExaBricks Data Structure



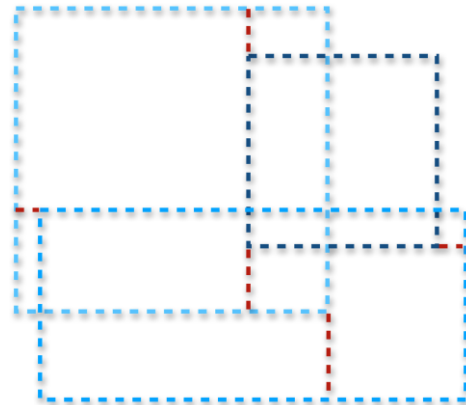
Adaptive Mesh
Refinement
Volume Data



Step 1: Build coarse bricks



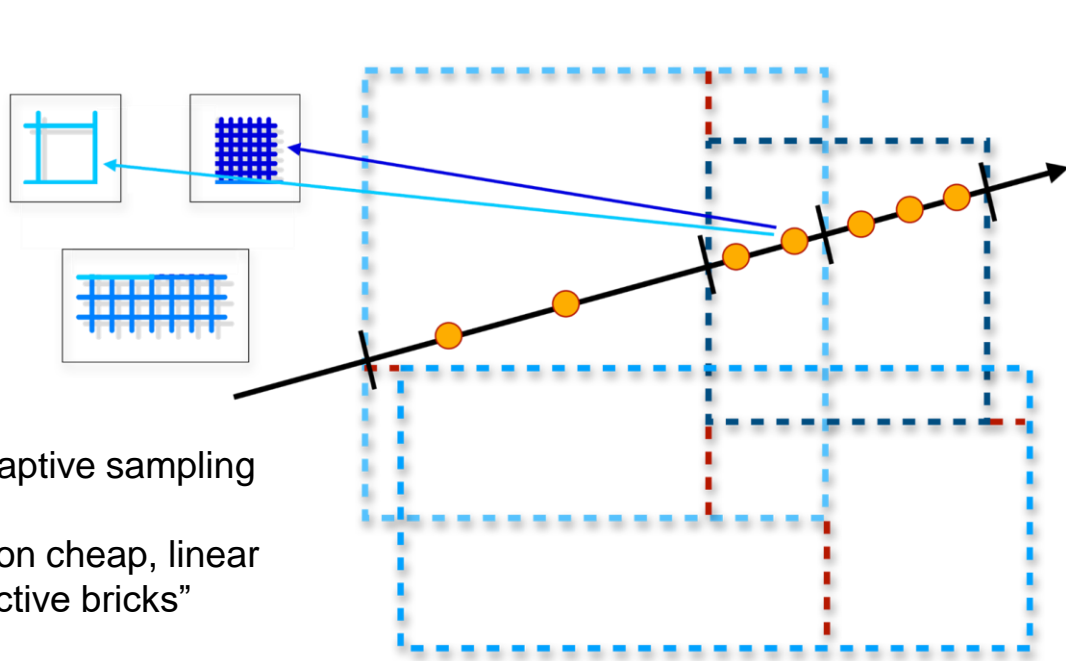
Step 2: Brick overlap regions



Step 3: OptiX overlap BVH

Wald et al. *Ray Tracing Structured
AMR Data Using ExaBricks*, IEEE
TVCG (2021)

ExaBricks Data Structure



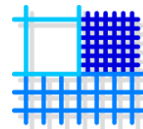
Adaptive Mesh
Refinement
Volume Data

Overlap regions for adaptive sampling

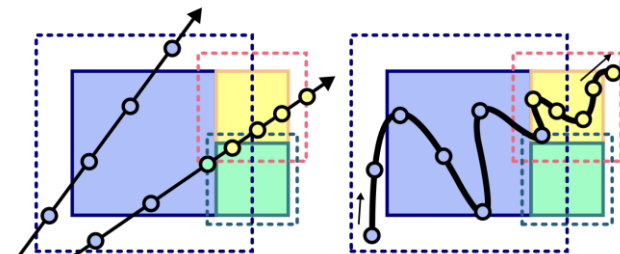
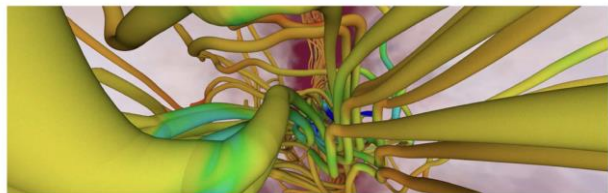
Sampling inside a region cheap, linear traversal over list of “active bricks”

Wald et al. *Ray Tracing Structured AMR Data Using ExaBricks*, IEEE TVCG (2021)

Future Research Challenges



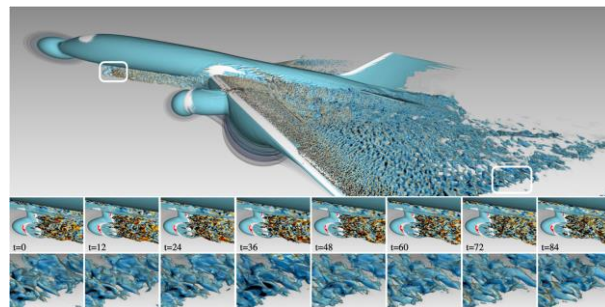
Adaptive Mesh
Refinement
Volume Data



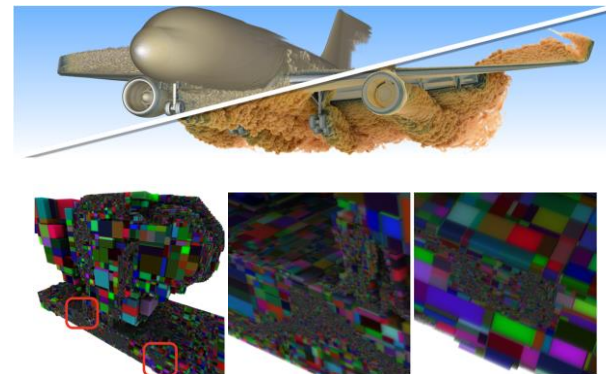
(a)

(b)

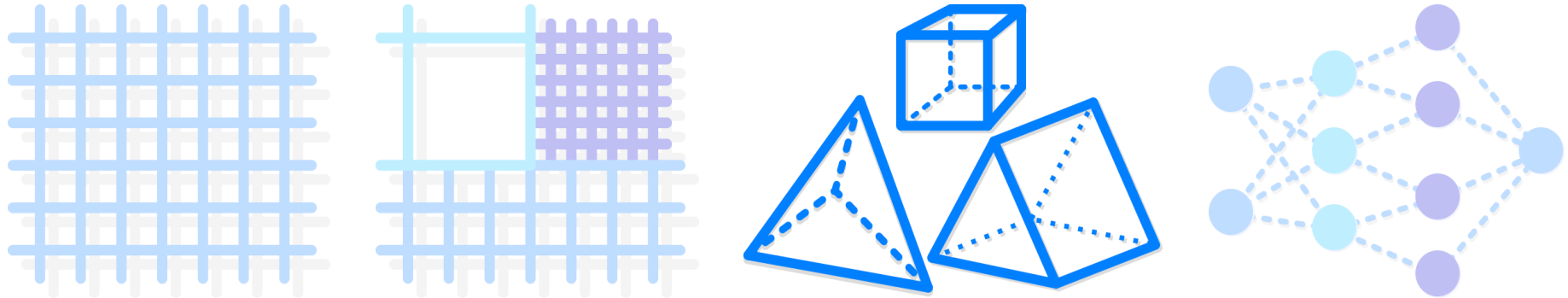
Zellmann et al. *Point Containment Queries on Ray Tracing Cores for AMR Flow Visualization*, **CiSE (2022)**



Zellmann et al. *Design and Evaluation of a GPU Streaming Framework for Visualizing Time-Varying AMR Data*, **EGPGV (2022)**



Zellmann et al. *Beyond ExaBricks: GPU Volume Path Tracing of AMR Data*, **ArXiv preprint**



Unstructured Volume Data

Ultimate way of refining the data to adapt to regions of interest



Unstructured
Volume Data

Tetrahedra only / different types of polyhedra (tetrahedra, hexahedra, pyramids, wedges) / high order, twisted and bent elements

Requires managing the topology, geometry, and scalar field to locate and render these polyhedral cells

Recent trends in traversing and sampling unstructured volumes



Unstructured
Volume Data

- element marching approaches

Muigg et al. *Scalable hybrid unstructured and structured grid raycasting*, IEEE TVCG (2007)

Muigg et al. *Interactive volume visualization of general polyhedral grids*, IEEE TVCG (2011)

Sahistan et al. *Ray-traced shell traversal of tetrahedral meshes for direct volume visualization*, IEEE Vis (2021)

- fixed sampling pattern

Binyahib et al. *A scalable hybrid scheme for ray-casting of unstructured volume data*, IEEE TVCG (2019)

Sahistan et al. *GPU-based data-parallel rendering of large, unstructured, and non-convexly partitioned data*, arXiv (2022)

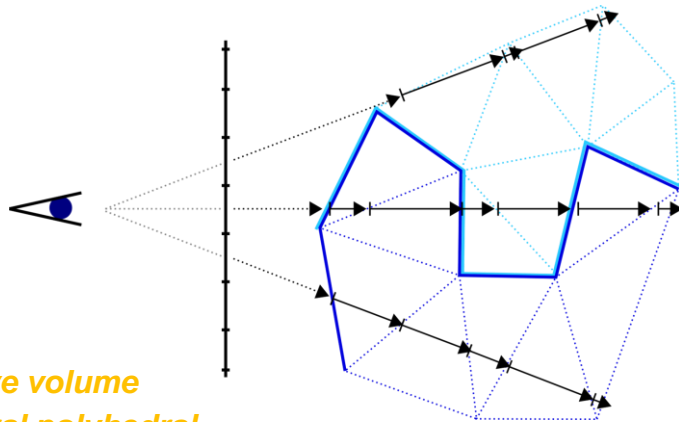
- point ray sampling with hardware-accelerated ray-tracing

Morriscal et al. *Accelerating unstructured mesh point location with RT cores*, IEEE TVCG (2022)

Wald et al. *A memory efficient encoding for ray tracing large unstructured data*, IEEE TVCG (2021)

Element marching

Traversal: marching from cell to cell via connectivity information

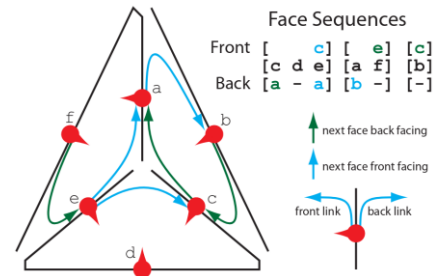


Unstructured
Volume Data

Muigg et al. *Scalable hybrid unstructured and structured grid raycasting*, IEEE TVCG (2007)

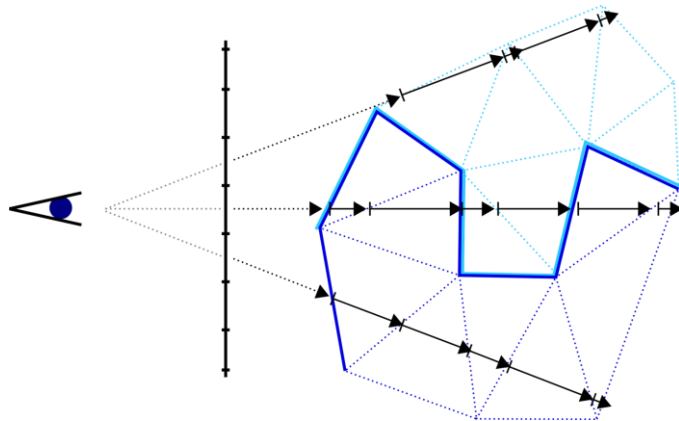
Muigg et al. *Interactive volume visualization of general polyhedral grids*, IEEE TVCG (2011)

- Decomposes the unstructured grid into bricks using a k-d tree
- Find entry/exit points: using rasterization and depth peeling
- Compact data structure to store the connectivity information: list of faces



Element marching

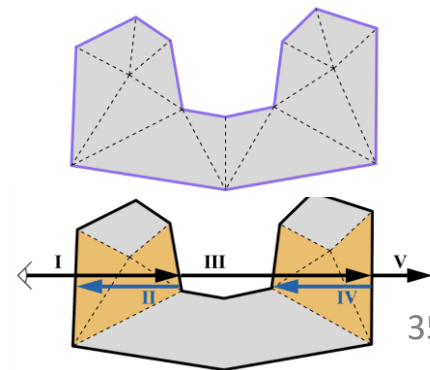
Traversal: marching from cell to cell via connectivity information



Unstructured
Volume Data

Sahistan et al. *Ray-traced shell traversal of tetrahedral meshes for direct volume visualization*, IEEE Vis (2021)

- Hardware-accelerated ray tracing for element marching
- OptiX triangle BVH to find entry and exit points (over tetrahedral meshes)
- Data compression with XOR compaction



Parallel and distributed rendering



Unstructured
Volume Data

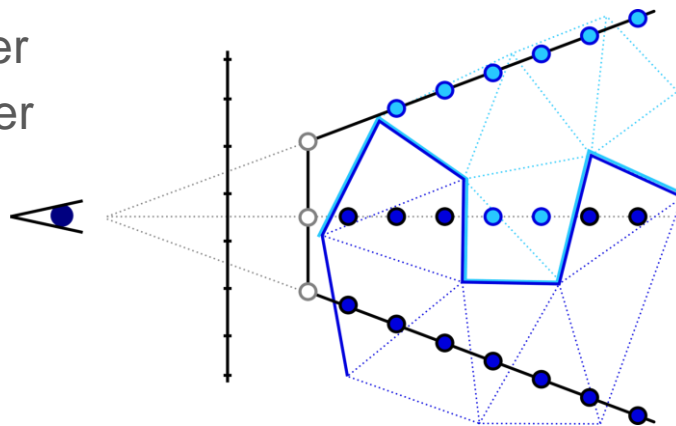
Binyahib et al. *A scalable hybrid scheme for ray-casting of unstructured volume data*, IEEE TVCG (2019)

A scalable unstructured volume rendering algorithm

- Small cells -> object-order
- Large cells -> image-order

Fixed sampling pattern

Compositing



Obtain the samples via **point queries**

Rathke et al. *SIMD parallel ray tracing of homogeneous polyhedral grids, EGPGV (2015)*

Use a hierarchical acceleration structure

Decouples traversal and sampling

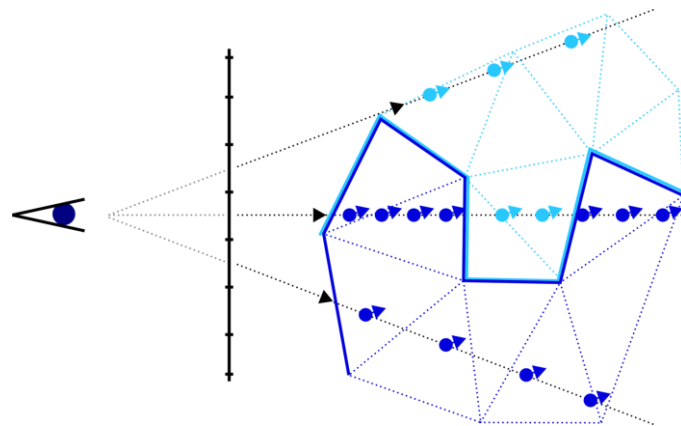
Hardware accelerated cell location

Wald et al. *A memory efficient encoding for ray tracing large unstructured data, IEEE TVCG (2021)*

Morriscal et al. *Accelerating unstructured mesh point location with RT cores, IEEE TVCG (2022)*



Unstructured
Volume Data



Hardware accelerated cell location

Using ray-tracing cores -> cast zero length rays ($t_{min} = t_{max} = 0$) for each sample

Hardware BVH traversal + ray/triangle intersections

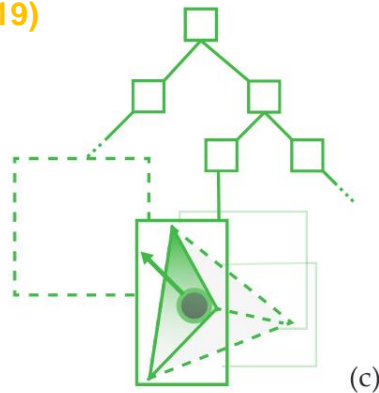
Good performances but memory intensive

Morriscal et al. *Accelerating unstructured mesh point location with RT cores, IEEE TVCG (2022)*

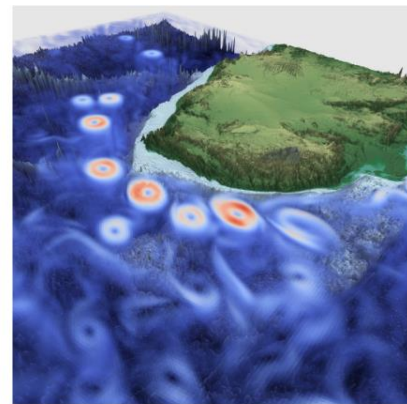


Unstructured
Volume Data

Wald et al. *RTX beyond ray tracing: Exploring the use of hardware ray tracing cores for tet-mesh point location, HPG (2019)*



(c)



Empty space skipping



Unstructured
Volume Data

Unstructured meshes:

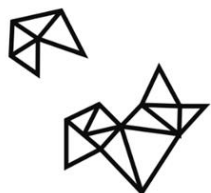
- often non-convex boundaries
 - > empty space = boundary mesh and RGB α TF
- individual elements can vary significantly in size



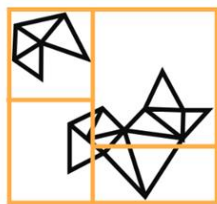
Hardware accelerated cell location with

- **Space skipping**: data structure (k-d tree partitioning) traversed using OptiX and RTX + min-max tree-based empty space skipping
- **Adaptive sampling**: adapt sampling rate according to TF variance in partitions

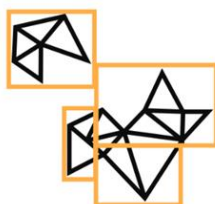
Morriscal et al. *Efficient space skipping and adaptive sampling of unstructured volumes using hardware accelerated ray tracing*, IEEE Vis (2019)



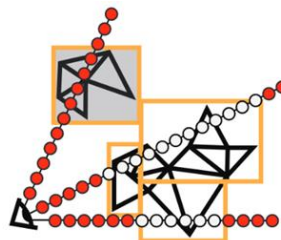
(a) Input unstructured mesh



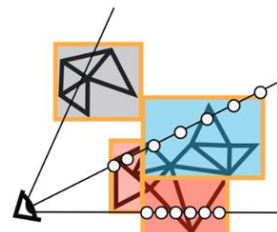
(b) Build KD-tree



(c) Shrink partitions



(d) Empty space skipping



(e) Adaptive sampling

Element marcher: only requiring face connectivity

-> can be compressed significantly



Unstructured
Volume Data

Vs.

Hardware-accelerated point queries: require BVH and complete vertex index lists

-> very high memory costs

Mesh data compression

Meshlet decomposition - compact indices representation with element regrouping

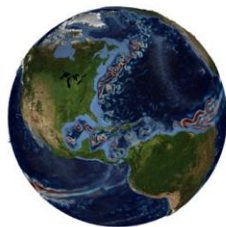


Acceleration structure compression

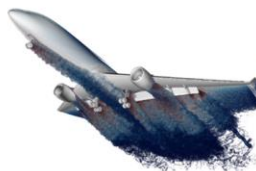
Eight-wide BVH + quantization - rearranging nodes and primitives

Collapse leaves

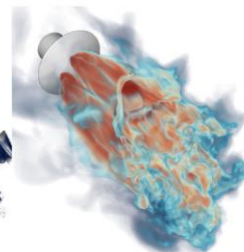
Wald et al. *A memory efficient encoding for ray tracing large unstructured data*, IEEE TVCG (2021)



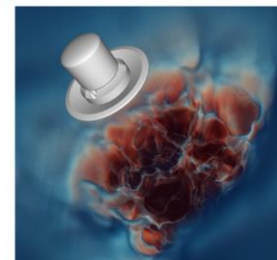
DKRZ full ocean
485 M verts, 749 M elements
compression rate 5.9 : 1



NASA Exa-Jet
656 M verts, 652 M elements
compression rate 4.9 : 1



NASA Mars Lander (small)
143 M verts, 789 M elements
compression rate 14.0 : 1



NASA Mars Lander (large)
576 M verts, 2.9 B elements
compression rate 12.3 : 1

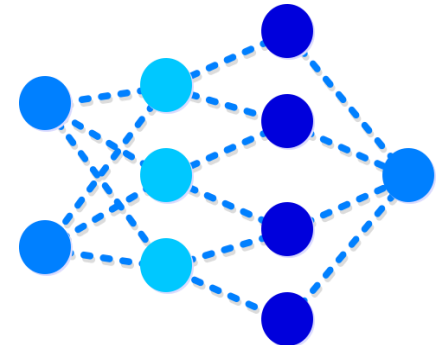
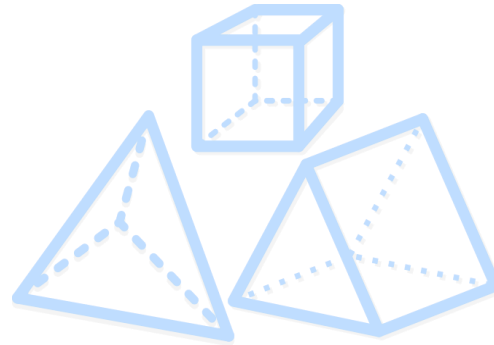
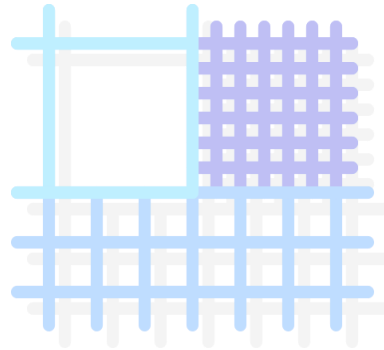
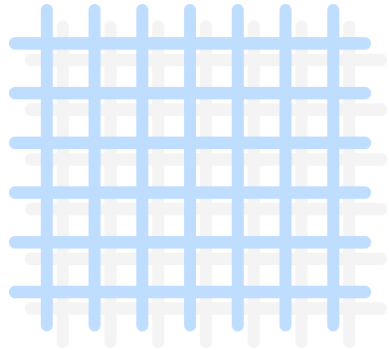
Extreme compression but bottleneck for acceleration structure
building time



Morriscal et al. *Quick clusters: A GPU-parallel partitioning for efficient path tracing of unstructured volumetric grids*, IEEE TVCG (2022)

Lazy sorting and clustering scheme -> much lower pre-processing times

-> trades sampling performance for pre-processing times



Compressed and Neural Representations

Hierarchical Compression

Unified resolution-precision compression

- Reducing grid resolution
- Reducing scalar precision



Compressed &
Neural
Representations

Fine control over resolution vs. precision

Hoang et al. *Efficient and flexible hierarchical data layouts for a unified encoding of scalar field precision and resolution*, IEEE TVCG (2021)

Bhatia et al. *AMM: Adaptive Multilinear Meshes*, IEEE TVCG (2022)

Neural Representations

Coordinate-based scene representation networks

- $F(x, y, z) \rightarrow \tau, \sigma, \delta, \text{rgb}$

Lu et al. \rightarrow Compressive neural representations

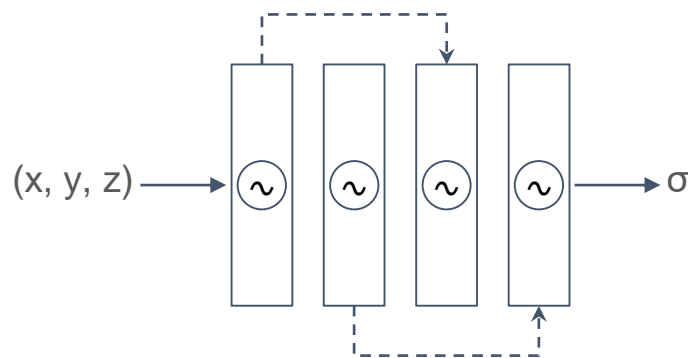
Weiss et al. \rightarrow Faster inference

Lu et al., *Compressive neural representations of volumetric scalar fields*, Computer Graphics Forum 40, 3 (2021)

Weiss et al., *Fast neural representations for direct volume rendering*. Computer Graphics Forum 41, 6 (2022)



Compressed & Neural Representations



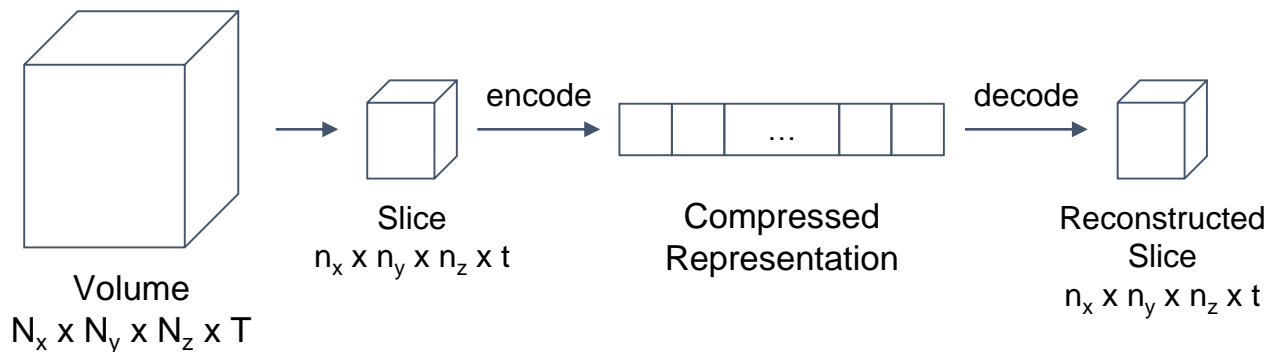
Neural Representations



Autoencoder based methods for time-varying data,

Pan et al., *Adaptive deep learning based time-varying volume compression*. IEEE International Conference on Big Data (2019)

Jain et al., *Compressed volume rendering using deep learning*. In Proceedings of LDAV (2017)



Findings

Classification

	Platform		Architecture			Traversal		Sampling	
	CPU	GPU	in-core	out-of-core	distributed	software	hardware	hierarchical	direct/linear
● [HAAB*18] [BMA*19] [SCRL20]		✓		✓		✓		✓	
● [WWJ19]	✓			✓		✓		✓	
● [GM19] [WZM21]		✓	✓				✓		✓
● [ZSL21]	✓	✓	✓			✓			✓
● [WBUK17] [WWW*19] [WMU*20]	✓		✓			✓		✓	
● [WZU*21] [ZWS*22a] [ZSM*22]		✓	✓				✓		✓
● [ZWS*22b]		✓	✓			✓	✓	✓	✓
● [MHDH07] [MHDG11]		✓	✓			✓			✓
● [SDM*21]		✓	✓				✓		✓
● [BPL*19]		✓			✓				✓
● [SDW*22]		✓			✓		✓		✓
● [WUM*19] [MUWP19] [WMZ21] [MWUP22] [MSG*22]		✓	✓				✓	✓	
● [HSB*21] [BHM*22]				✓				✓	
● [LJLB21]		✓	✓						

Classification

	Platform		Architecture			Traversal		Sampling	
	CPU	GPU	in-core	out-of-core	distributed	software	hardware	hierarchical	direct/linear
● [HAAB*18] [BMA*19] [SCRL20]		✓		✓		✓		✓	
● [WWJ19]	✓			✓		✓		✓	
● [GM19] [WZM21]		✓	✓				✓		✓
● [ZSL21]	✓	✓	✓			✓			✓
● [WBUK17] [WWW*19] [WMU*20]	✓		✓			✓		✓	
● [WZU*21] [ZWS*22a] [ZSM*22]		✓	✓				✓		✓
● [ZWS*22b]		✓	✓			✓	✓	✓	✓
● [MHDH07] [MHDG11]		✓	✓			✓			✓
● [SDM*21]		✓	✓				✓		✓
● [BPL*19]		✓			✓				✓
● [SDW*22]		✓			✓		✓		✓
● [WUM*19] [MUWP19] [WMZ21] [MWUP22] [MSG*22]		✓	✓				✓	✓	
● [HSB*21] [BHM*22]				✓				✓	
● [LJLB21]		✓	✓						

Classification

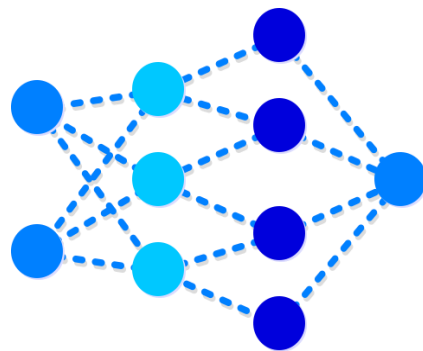
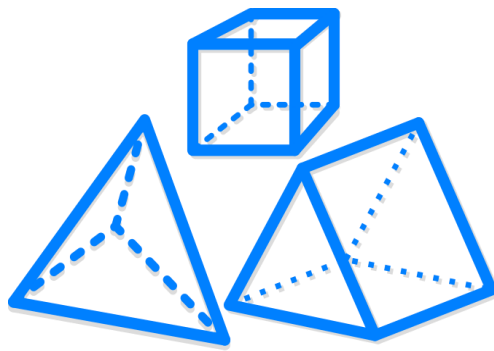
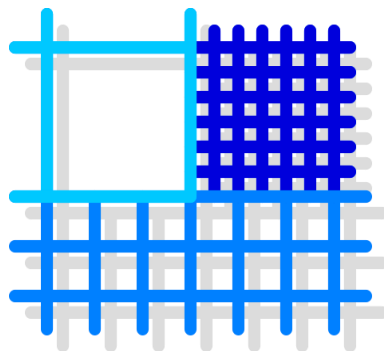
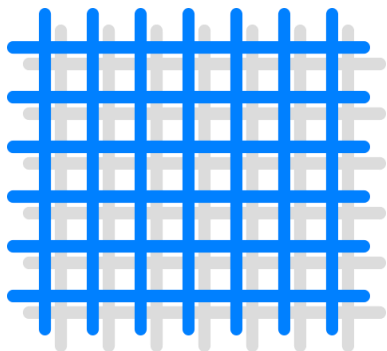
	Platform		Architecture			Traversal		Sampling	
	CPU	GPU	in-core	out-of-core	distributed	software	hardware	hierarchical	direct/linear
● [HAAB*18] [BMA*19] [SCRL20]		✓		✓		✓		✓	
● [WWJ19]	✓			✓		✓		✓	
● [GM19] [WZM21]		✓	✓				✓		✓
● [ZSL21]	✓	✓	✓			✓			✓
● [WBUK17] [WWW*19] [WMU*20]	✓		✓			✓		✓	
● [WZU*21] [ZWS*22a] [ZSM*22]		✓	✓				✓		✓
● [ZWS*22b]		✓	✓			✓	✓	✓	✓
● [MHDH07] [MHDG11]		✓	✓			✓			✓
● [SDM*21]		✓	✓				✓		✓
● [BPL*19]		✓			✓		✓		✓
● [SDW*22]		✓			✓		✓		✓
● [WUM*19] [MUWP19] [WMZ21] [MWUP22] [MSG*22]		✓	✓				✓	✓	✓
● [HSB*21] [BHM*22]				✓				✓	
● [LJLB21]		✓	✓						

Classification

	Platform		Architecture			Traversal		Sampling	
	CPU	GPU	in-core	out-of-core	distributed	software	hardware	hierarchical	direct/linear
● [HAAB*18] [BMA*19] [SCRL20]		✓		✓		✓		✓	
● [WWJ19]	✓			✓		✓		✓	
● [GM19] [WZM21]		✓	✓				✓		✓
● [ZSL21]	✓	✓	✓			✓			✓
● [WBUK17] [WWW*19] [WMU*20]	✓		✓			✓		✓	
● [WZU*21] [ZWS*22a] [ZSM*22]		✓	✓				✓		✓
● [ZWS*22b]		✓	✓			✓	✓	✓	✓
● [MHDH07] [MHDG11]		✓	✓			✓			✓
● [SDM*21]		✓	✓				✓		✓
● [BPL*19]		✓			✓				✓
● [SDW*22]		✓			✓		✓		✓
● [WUM*19] [MUWP19] [WMZ21] [MWUP22] [MSG*22]		✓	✓				✓	✓	✓
● [HSB*21] [BHM*22]				✓				✓	
● [LJLB21]		✓	✓						

- Going more toward unstructured format than regular format
- Sampling workloads become increasingly incoherent:
 - ❖ secondary shadow or scattering rays into the volume
 - ❖ arbitrary sampling e.g., to compute flow visualizations
 - ❖ path tracing is adopted by the sci-vis community
- Compressing given data representations; extremely lossy compression by encoding the volume as a neural network (NeRF)

Conclusion



Our contribution:

Review of recent research in interactive volume visualization for large-scale data, divided into 4 main data representations.

Mostly focus on general-purpose ray tracing techniques.

Conclusion

Sampling is critical for reconstructing values, but becomes costly the less regular the data.

Introduction of hierarchical and less structured volume data to reduce memory overhead, but at the cost of complex visualization algorithms.

Popularization of GPUs with accelerated ray tracing capability, which helps with visualization algorithm flexibility but is still heavy in memory consumption and interactive changes.

Thank you for your attention!

