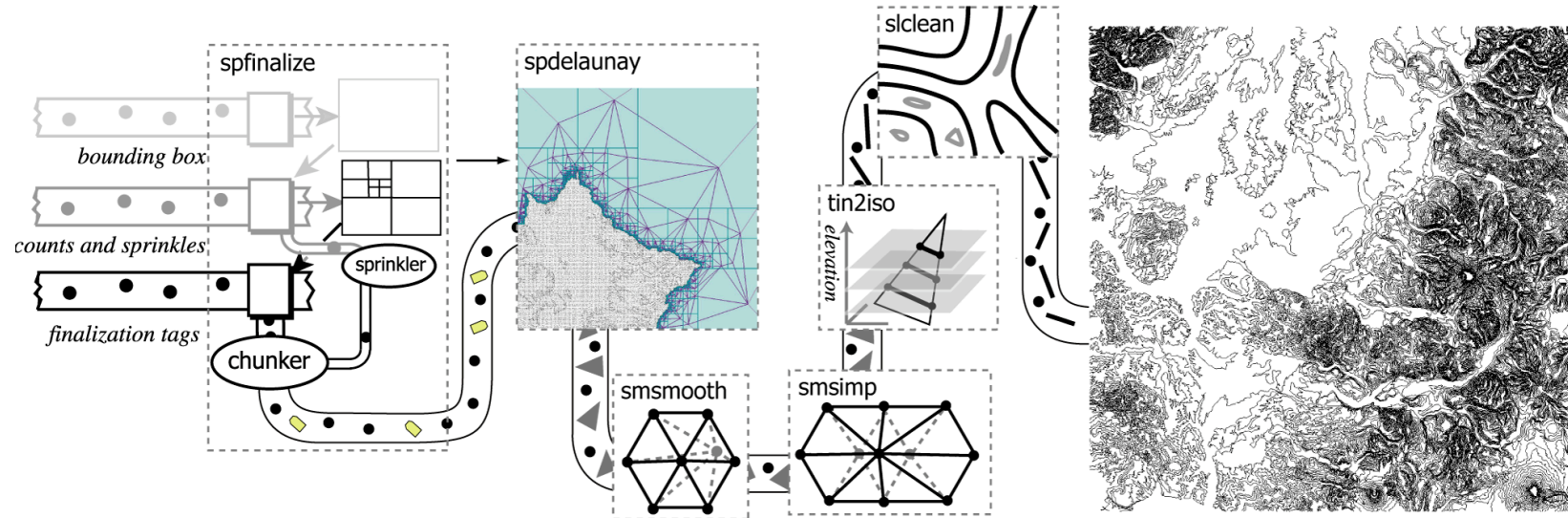


Streaming processing of spatial data



David Millman

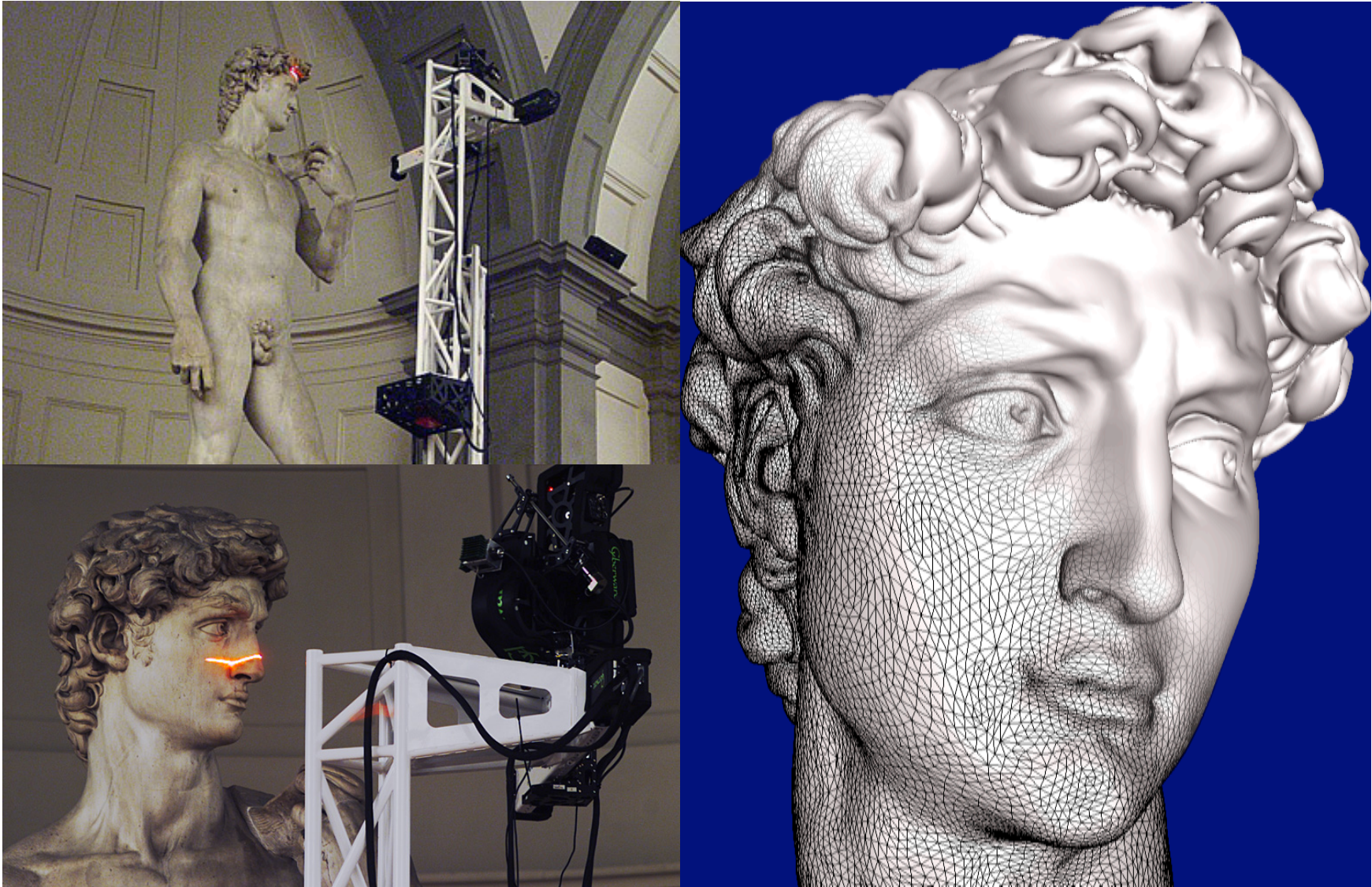
joint work with

Martin Isenburg,

Shawn Brown and

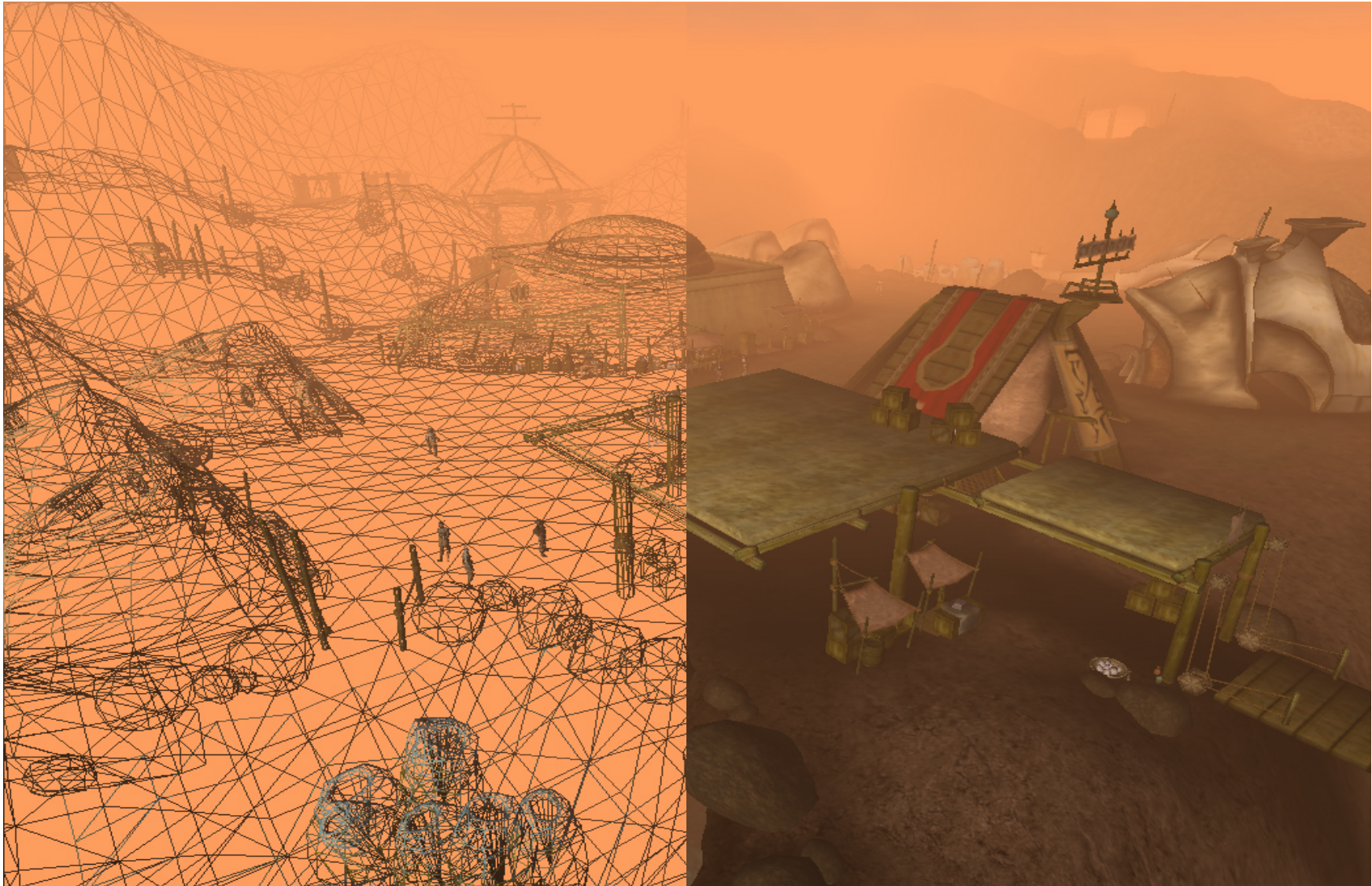
Jack Snoeyink

Scanning Artifacts



scanning of “Michelangelo’s David” courtesy of Marc Levoy

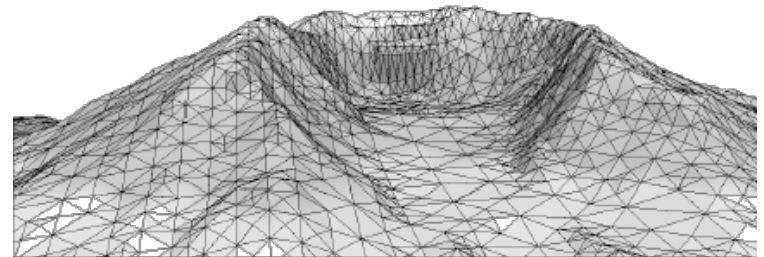
Modeling & Games



screen shot of “The village of Gnisis”, The Elder Scrolls III

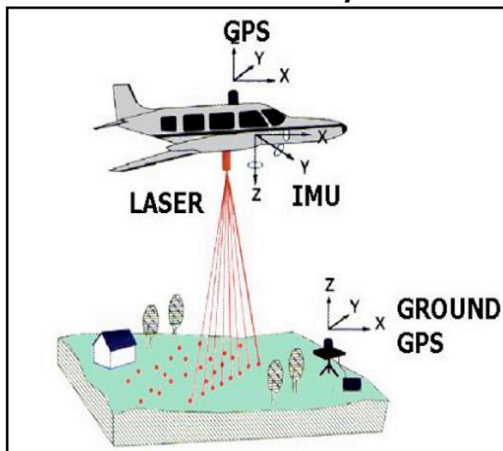
Geometry processing applications

- Computer-aided design (CAD): modeling
- Graphics/Games: level of detail rendering
- Finite element (FEM): analysis
- Robotics: path planning
- Geographic Information Systems (GIS): Terrain maps [M01,vK97]
 - TIN triangulation [PFLM76]
 - Contour lines
 - Raster DEM [T90]

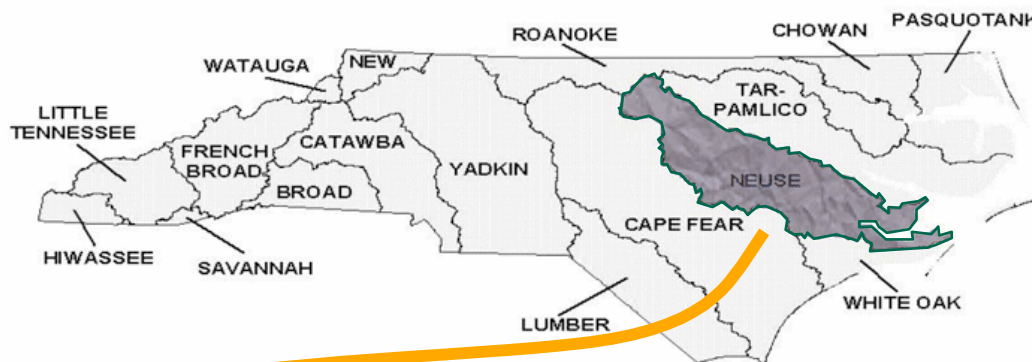


LIDAR to TIN to contour or raster

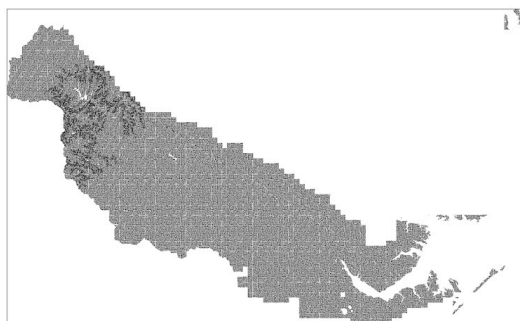
Airborne LIDAR System



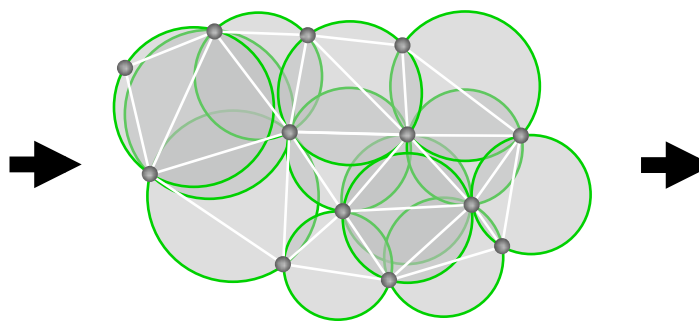
FLOODPLAIN MAPPING PROGRAM
FINAL BASIN PLAN — NEUSE RIVER BASIN



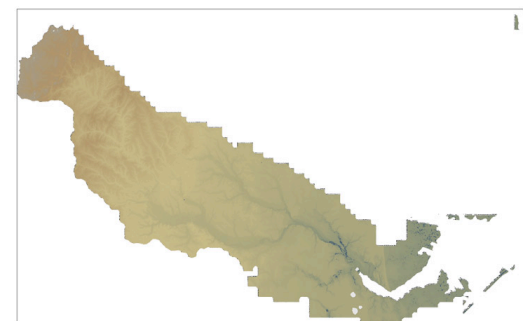
0.5 Billion raw
points (11 GB)



via Delaunay
Triangulation (40 GB)



rasterized onto
Digital Elevation Grid



Pipes

data in → **filter** → **results out**

```
ps aux | grep oeyi > myprocs.txt
```

- Easy to combine simple tools
- Avoids writes to temporary files
- OS handles buffers & time, some apps have their own buffers
- Data must stream in a format that the application expects

Advantages of pipes

- ***Composable*** - string modules together to solve complex tasks with high throughput.
- ***Low Latency*** - Allow output from the last module in the string while the first is still reading its input
- ***Memory Coherence*** - As data is finalized its memory may be released
- ***Parallelism*** - The operating system does load balancing for additional processors or multi-core chips
- ***Modularity*** - Prototype modules can be tested and later replaced with "better" techniques

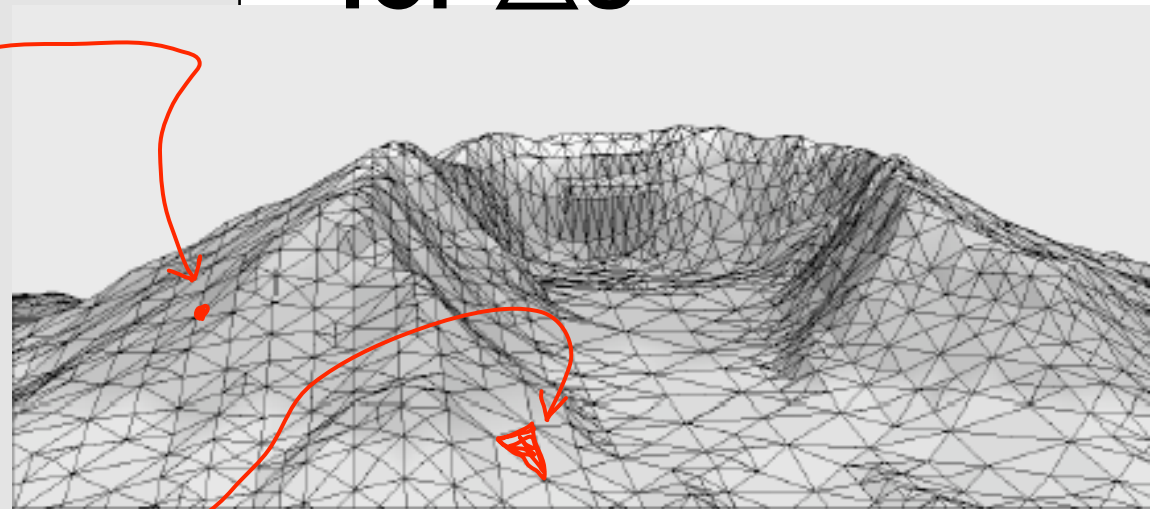
Spatial locality

- Waldo Tobler's 1st Law of Geography:
 - *“Everything is related to everything else, but near things are more related than distant things.”*
 - This law affects how data is used and how it is collected.
- Finalization: document in the data format when an item is last used.
 - For meshes
 - For point sets


```
TIN
BEGT
TNAM helens
VERT 2832
134 449 76.12
134 450 76.76
134 451 77.12
134 453 77.28
  :   :   :
TRI 5488
2810 2815 2821
2801 2811 2805
2793 2802 2796
2783 2794 2788
  :   :   :
```

TIN format

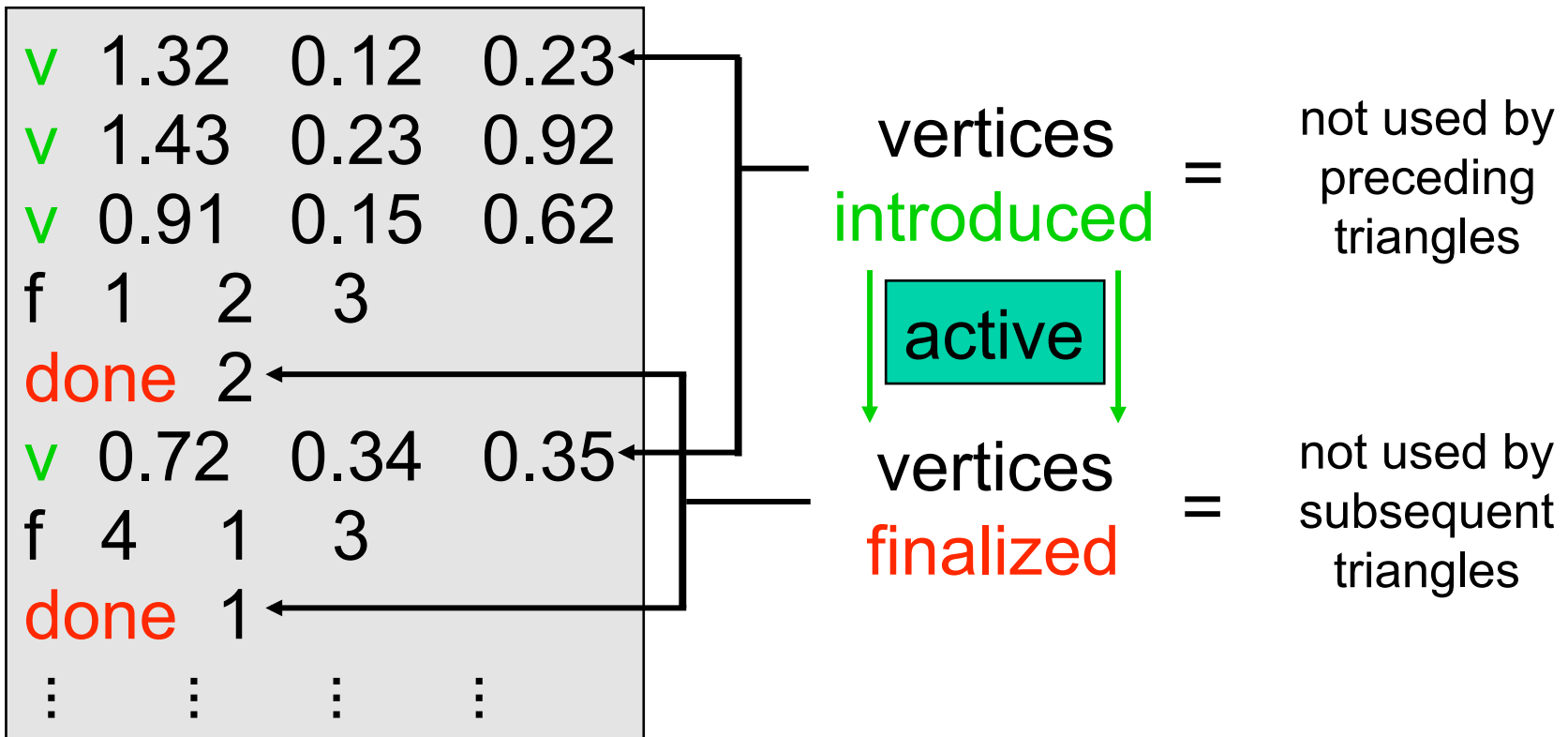
List coords for
verts & indices
for Δ s



**Batch: Must read all
coords before
processing any Δ**

Streaming Mesh [105]

- interleave vertices & triangles
- **introduce** and **finalize** vertices



St. Matthew Statue

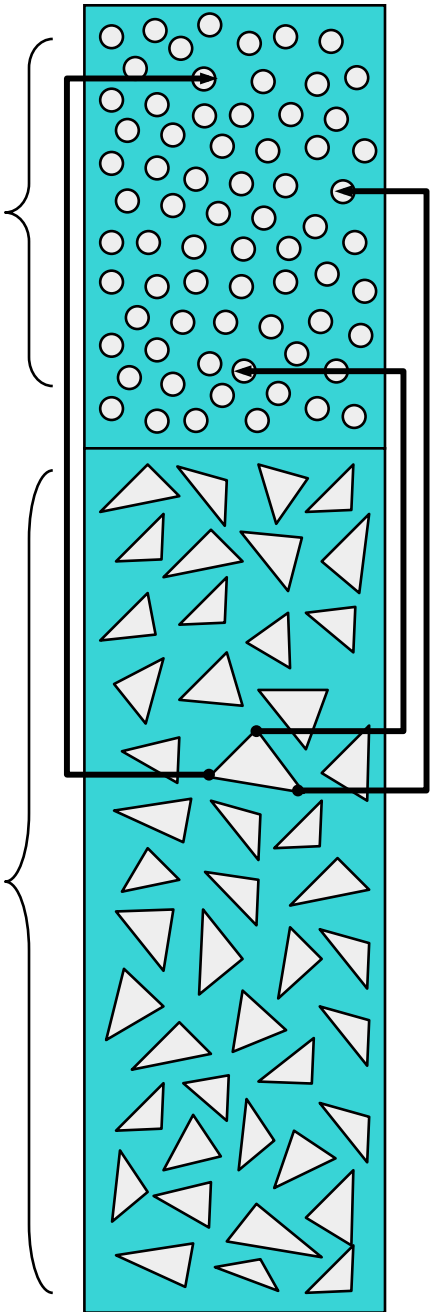


- standard indexed format

2 GB

4 GB

186 million vertices
372 million triangles



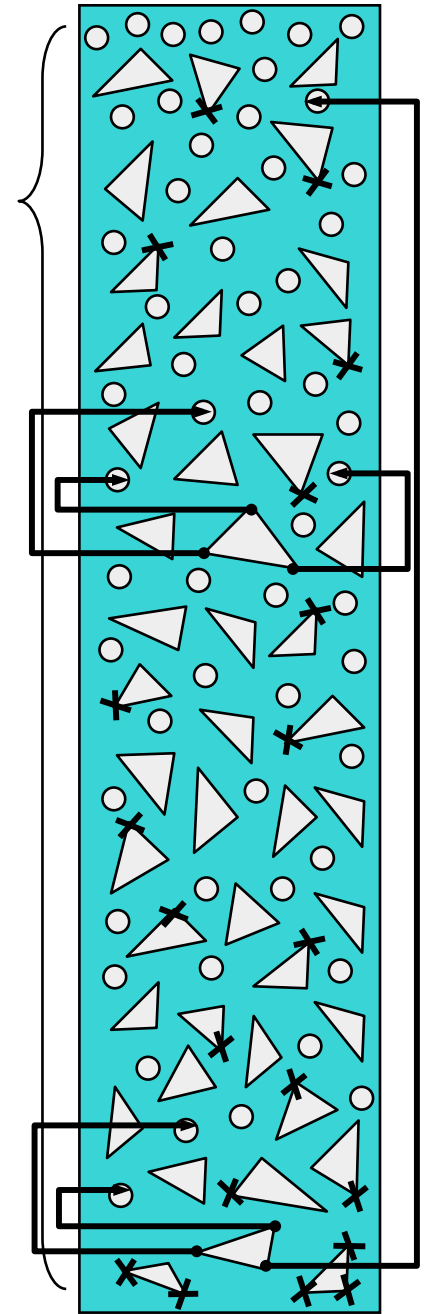
St. Matthew Statue



- standard indexed format
- **streaming format**

6 GB

186 million vertices
372 million triangles

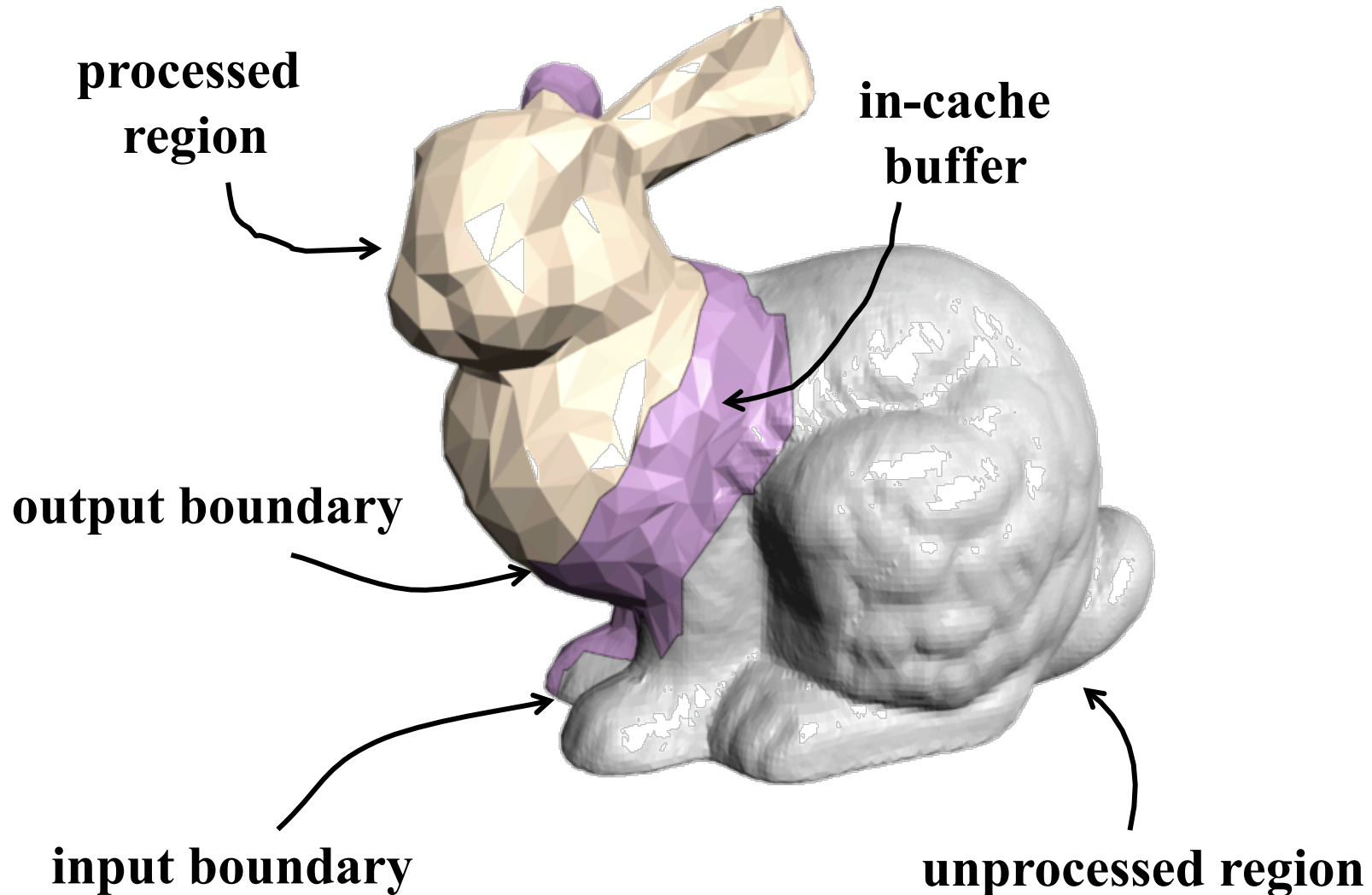


Advantages of finalization

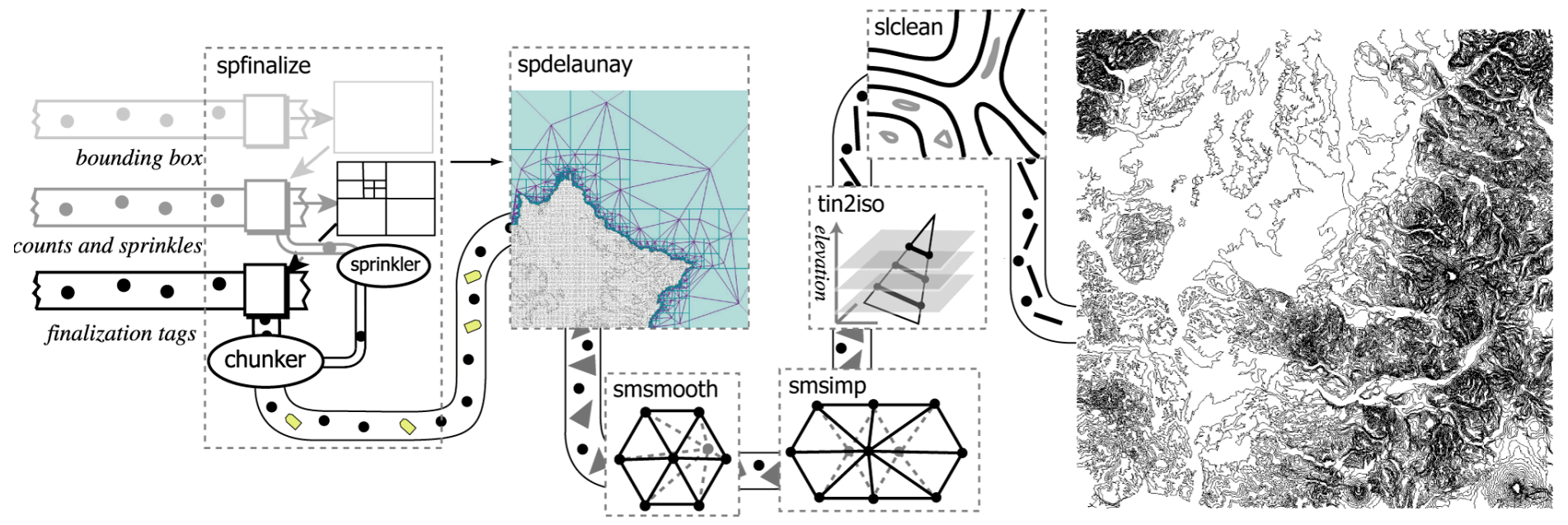
- Can flush data no longer needed
- Pipeline geometry processing
- Simple API

- Disadvantage: can't do everything...
fixed traversal; not progressive

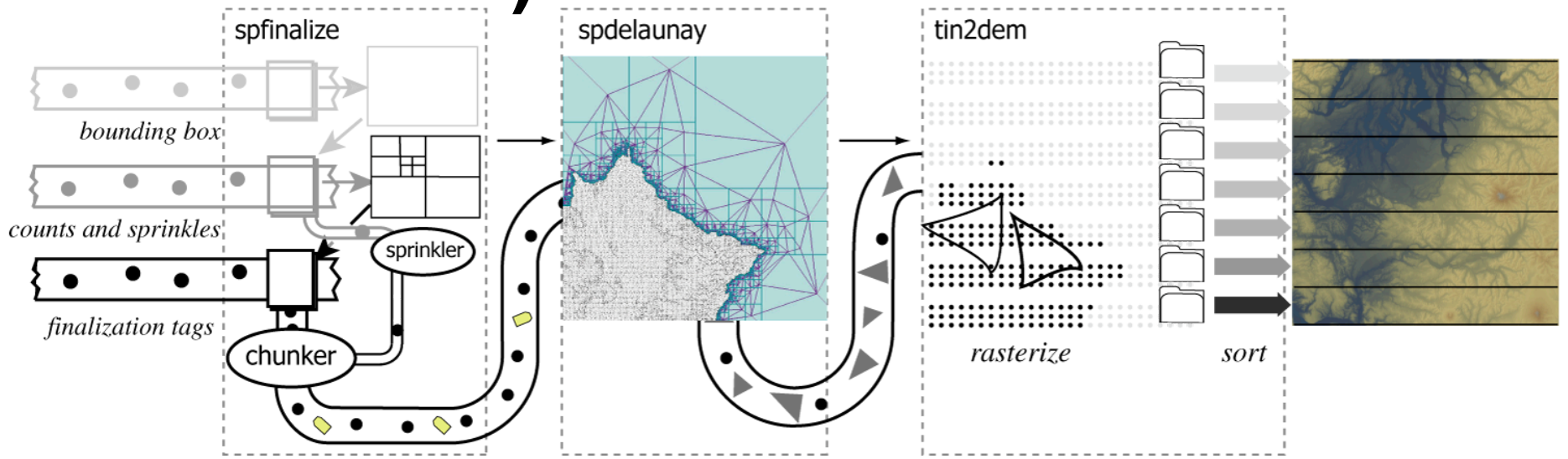
Stream Processing with a small memory footprint



Points to contours in 7 modules



Points to raster DEM (TO BE DELETED)



- **Need the sort phase, but can at least produce raster in strips...**
- **Effectively produces rasters that are larger than memory size.**

Thank you!