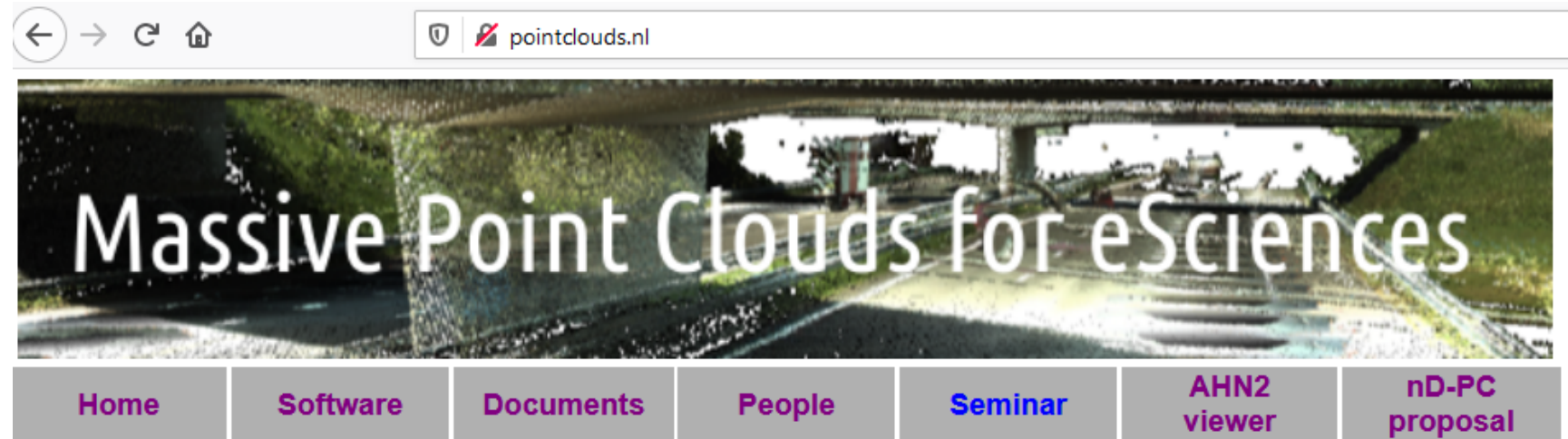# nD-PointCloud continuous level representation for spatio-temporal phenomena in Open Point Cloud Maps

Peter van Oosterom, Edward Verbree, Martijn Meijers, Haicheng Liu, Vitali Diaz Mercado
GIS technology, Section Digital Technologies,
Department Architectural Engineering and Technology
Faculty of Architecture and the Built Environment, TU Delft

Presentation at the project kick-off, 15 November 2021

Sequel of eScience project 2013-2015

Overview

- project details
- direct use of point clouds
- Open Point Cloud Map
- nD-PointClouds
- big data management
- conclusion

New project: 2021-2024

- nD-PointCloud continuous level representation for spatio-temporal phenomena in Open Point Cloud Maps

- 3 years, started 1 nov'21

- eScience Center Personnel: 2.5 PYR

- Deltares for proof-of-principe water management application

- Total budget in cash k€ 479

- software to be used/extended (a.o.): Oracle/PostgreSQL/MonetDB, Potree, PDAL, Entwine

Composition of the Research Team

Name	Affiliation	Period / % / PYR	Expertise and type of involvement
Postdoc Vitali Diaz Mercado	TUD	3 years / 100%	Big spatial data processing
Peter van Oosterom	TUD	3 years / 15%	Spatial DBMS, project leader
Edward Verbree	TUD	3 years / 5%	Point could data expert, advisor
Martijn Meijers	TUD	3 years / 5%	Map generalization expert, advisor
Haicheng Liu	TUD	1 year / 100%	Last year of PhD on Point Clouds
Fedor Baart/Maarten Pronk	Deltares	3 years / 5%	Water management domain advisor
Markus Schütz	TU Wien	2 visits	Point cloud visualization expert
eScience RSE 1	NLeSc	3 years / 50%	Cluster/cloud computing engineer
eScience RSE 2	NLeSc	2 years / 50%	3D Web graphics engineer

Project organization in WP's

Work Package	Year 1		Year 2		Year 3	
WP1: Organization, Communication	■	■	■	■	■	■
WP2: Conceptual Model	■	■				■
WP3: Data Management Technologies	■	■	■	■		
WP4: HPC/HTC Web-services		■	■	■	■	
WP5: nD-PointCloud Visualization			■	■	■	■
WP6: Proof-of-Principle water management	■			■	■	■

Work via sprints:
→ short period of intensive collaboration with clear goal
→ alternate with period of follow-up refinement

Relationship between the WPS

Overview

- project details
- direct use of point clouds
- Open Point Cloud Map
- nD-PointClouds
- big data management
- conclusion

Direct point cloud use (incl. analysis)

- benefits:
  - no conversion time
  - no data loss
  - analysis may be better
  - LoD continuous (raster pixels factor 2, vector hard)
  - very realistic representations (e.g tree with leaves)
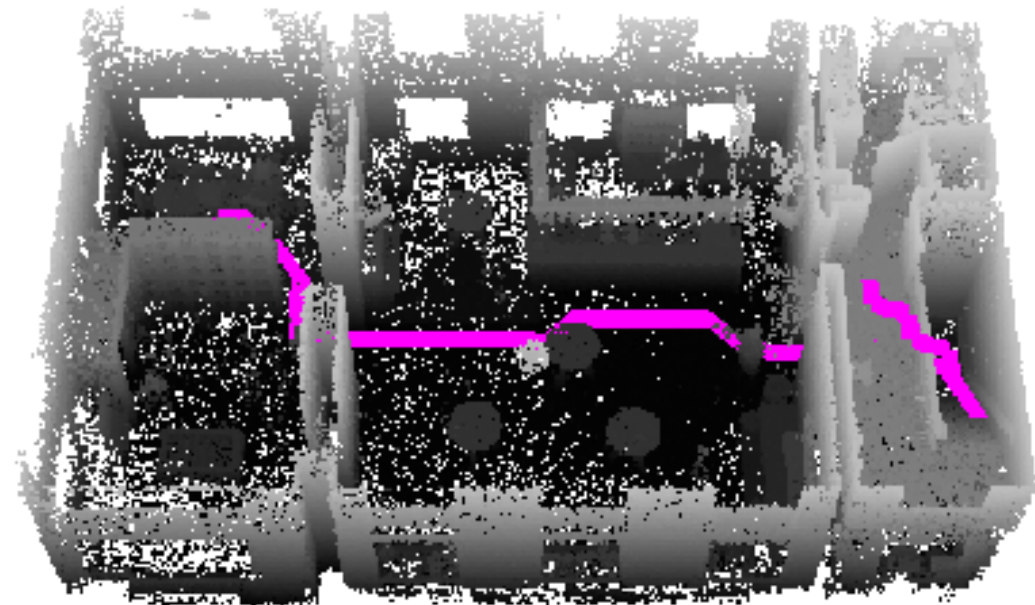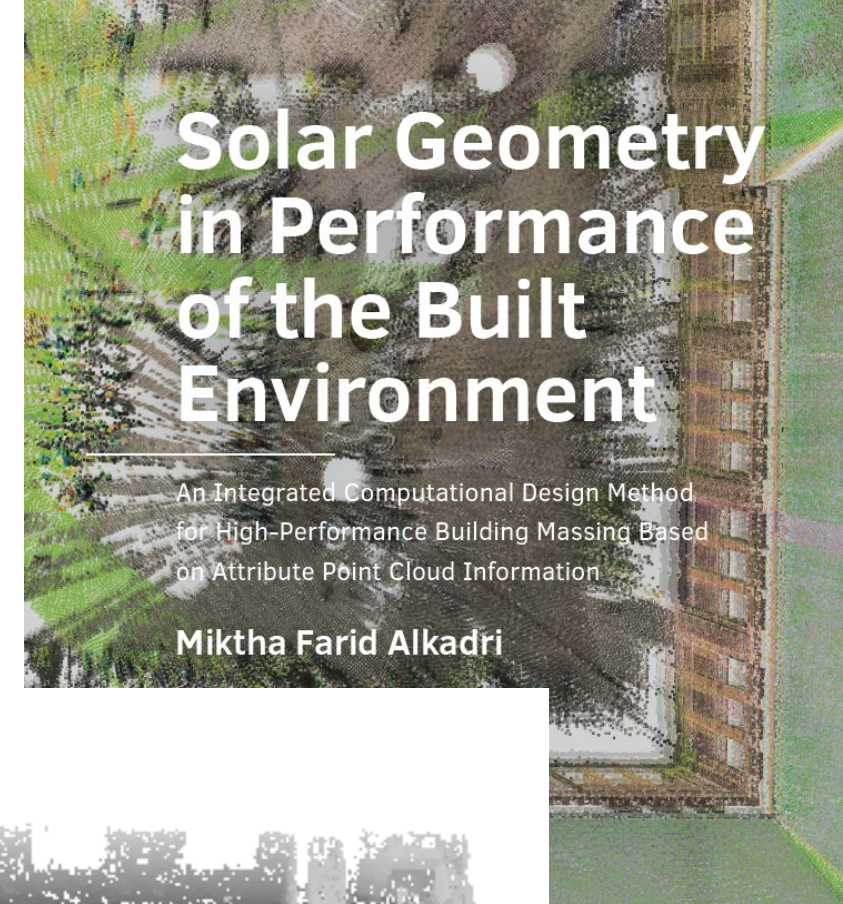- drawbacks:
  - lot of data
  - redevelop algorithms
- have it as option
  together with conversions PC ←→vector, PC←→ raster
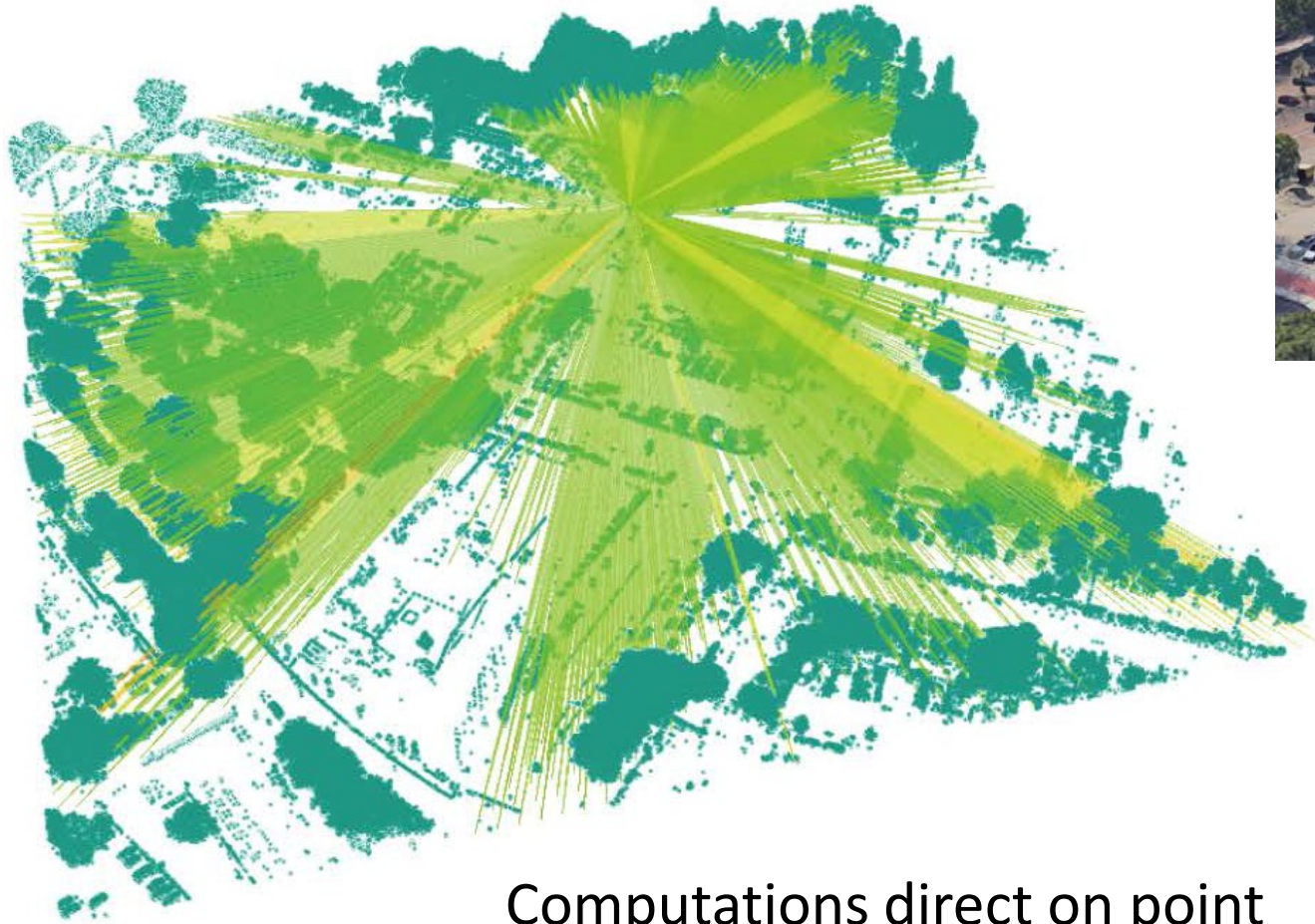
Types of analysis, direct point clouds

- solar energy potential
- solar envelopes
- 3D routing (e.g. drone, both indoor/outdoor)
- viewshed/ line-of-sight
- change detection (deformations)
- volume analysis computations
- vegetation analysis

- continuous LoD also for analysis not only visualization

Rodenberg MSc thesis 2017 →

Solar Geometry in Performance of the Built Environment

An Integrated Computational Design Method for High-Performance Building Massing Based on Attribute Point Cloud Information
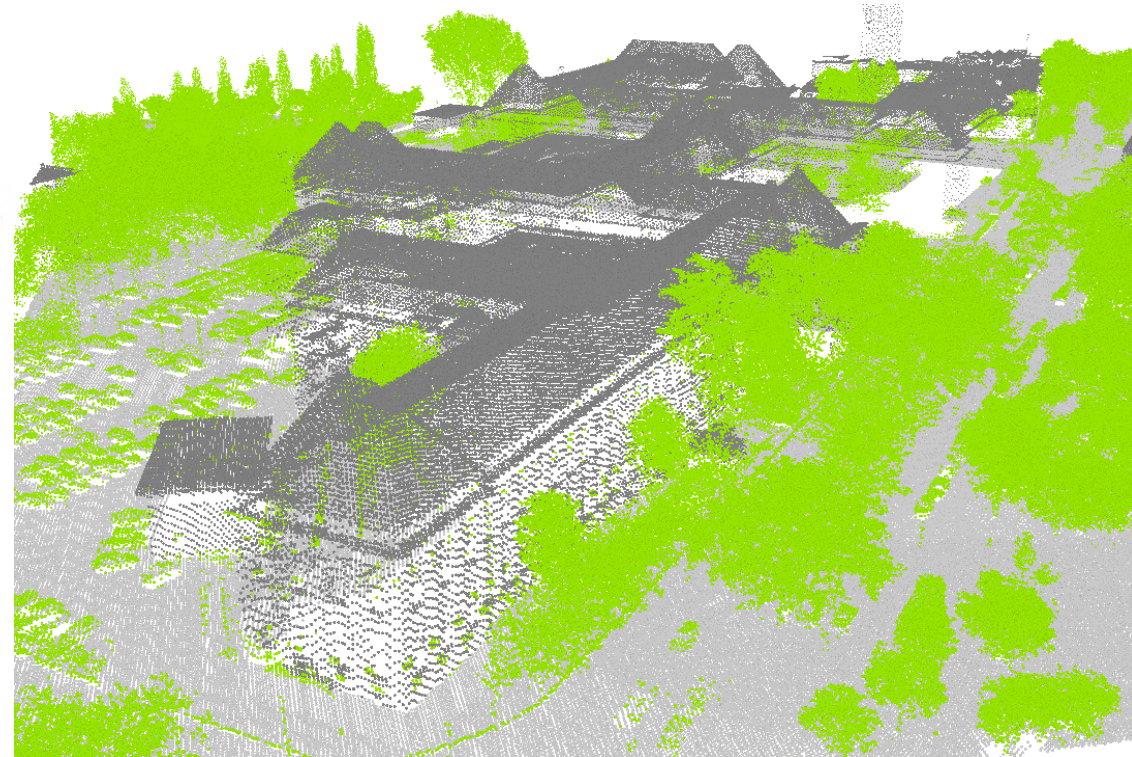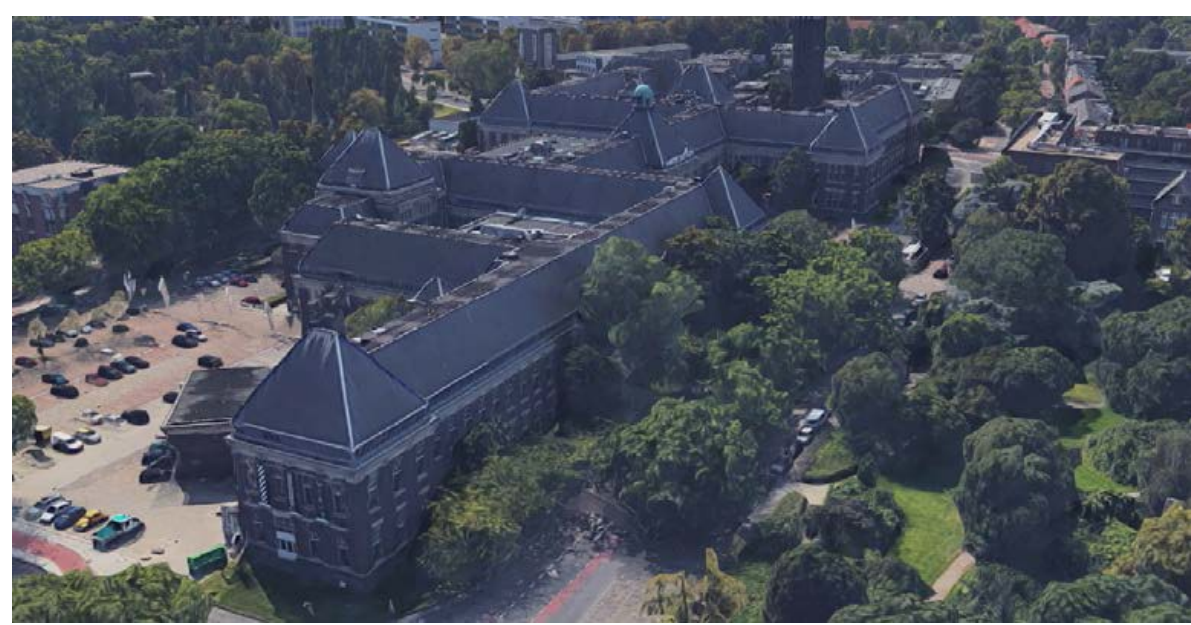
Miktha Farid Alkadri

Viewshed, line-of-sight

Computations direct on point
clouds (note vegetation),

Guanting PhD visit TUD 2017

Change detection

AHN2-AHN3

Green=added

Red=removed

ChronoCity (Synthesis project 2017)
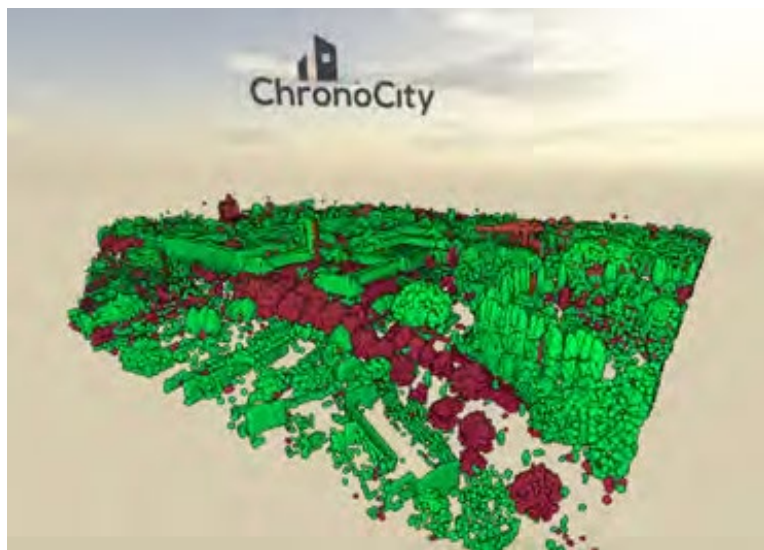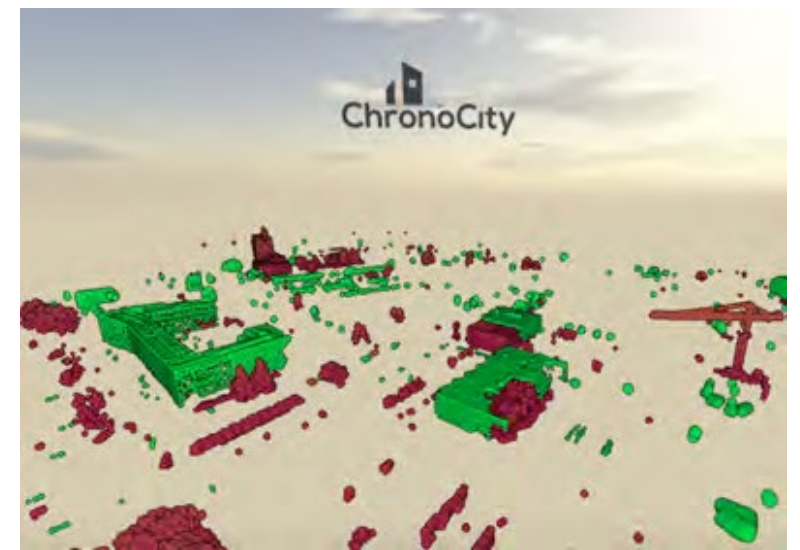

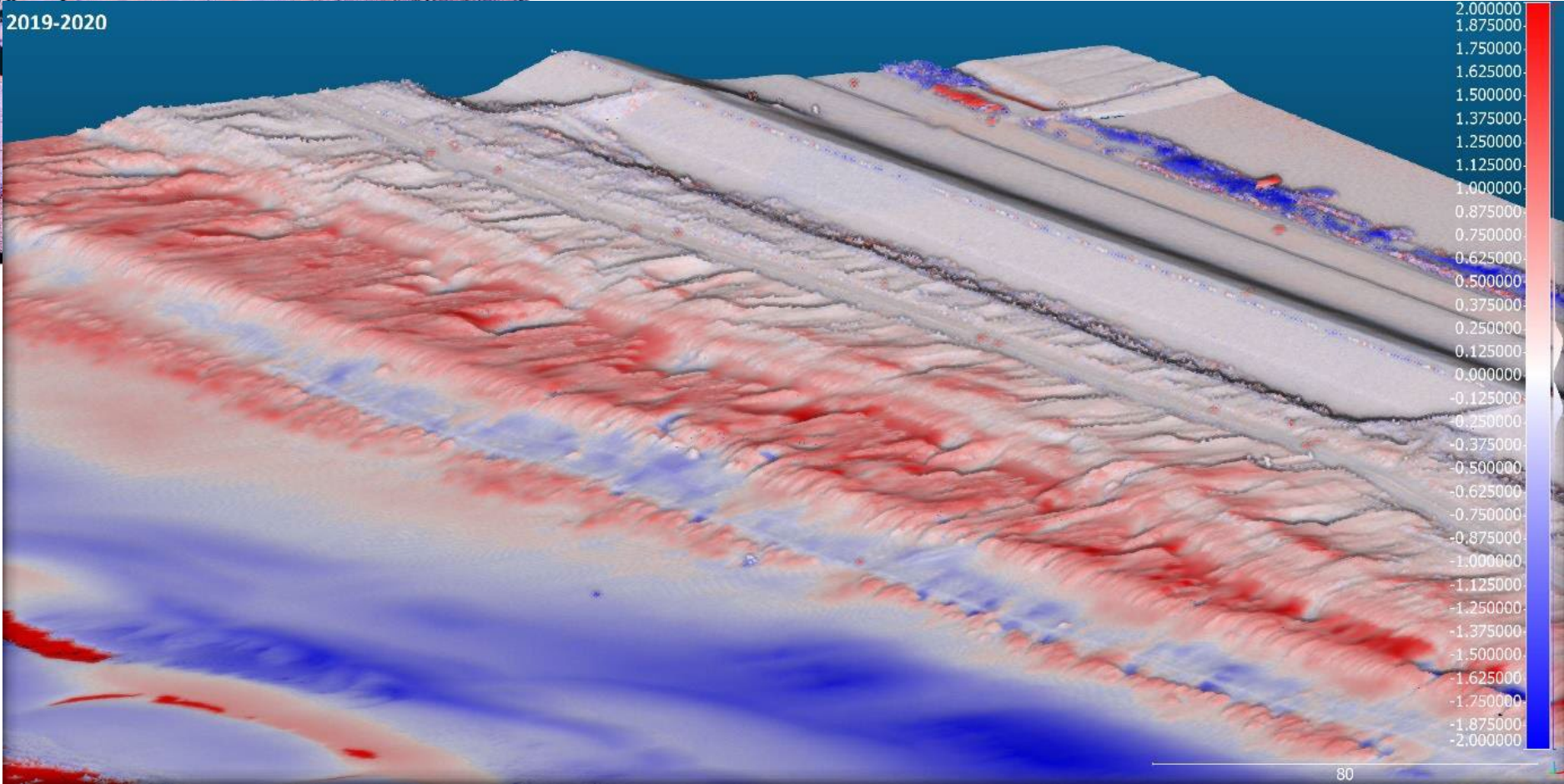Threshold

Left: too small many changes
Right: proper size

09-02-2020 between 12:00 and 13:00

Coastal monitoring

2019-2020

one Ciara storm hour

one year difference

MSc GIMA thesis 2021
Ynke Schreijer

Hydrology/ flow over surface

MSc GIMA thesis 2021
Stijn Ticheloven
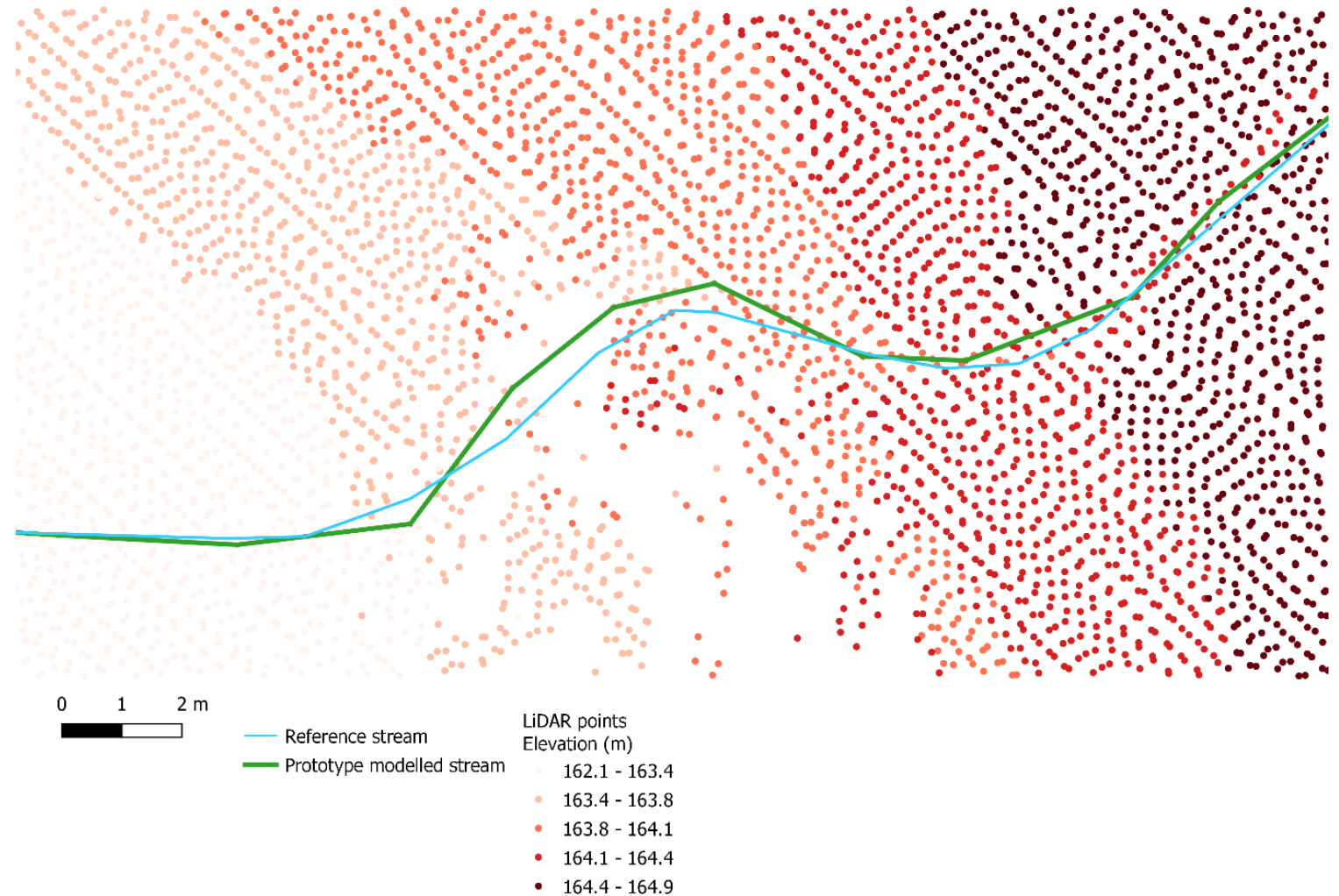
Representations of space, time, scale
…after grid/voxel or object/vector

- **new 3rd representation**: nD-PointCloud (nD-PC)

- many scientific domains (spatial): geography, medicine, physics, astronomy, hydrology, architecture, archaeology, arts, CAD, social media/ moving objects, gaming…

- nD-PC in whole processing chain: acquisition, DBMS, analysis, simulation, dissemination, visualization,…

- deep integration space/time/scale
  1. more efficient, store, exchange, compute
  2. more functionality (smooth zoom/ analysis)

- BIG spatial data: 35 trillion points (in astronomy, geo-info)

Overview

- project details
- direct use of point clouds
- Open Point Cloud Map
- nD-PointClouds
- big data management
- conclusion

Open Point Cloud Map (OPCM)

- 3D Webviewer/computation tool will be operational after project

- Open Point Cloud Map, pilot phase will select data cases from:
  - several countries (NL, UK, Denmark, USA, etc.)
  - include time series of point clouds
  - add different data sets, (local) higher density, different attributes
  - point cloud from different nature: wet and dry
  - dense matching stereo, PS-InSAR, MBES,..

- Target functionality:
  - play with point cloud layers (on/off, styling)
  - compare point cloud (differences), temporal/other animation
  - scalable architecture (distributed serving/standardized protocols)
  - store point cloud data blocks in DBMS (not files)

Open Point Cloud Map
distributed architecture: multiple servers
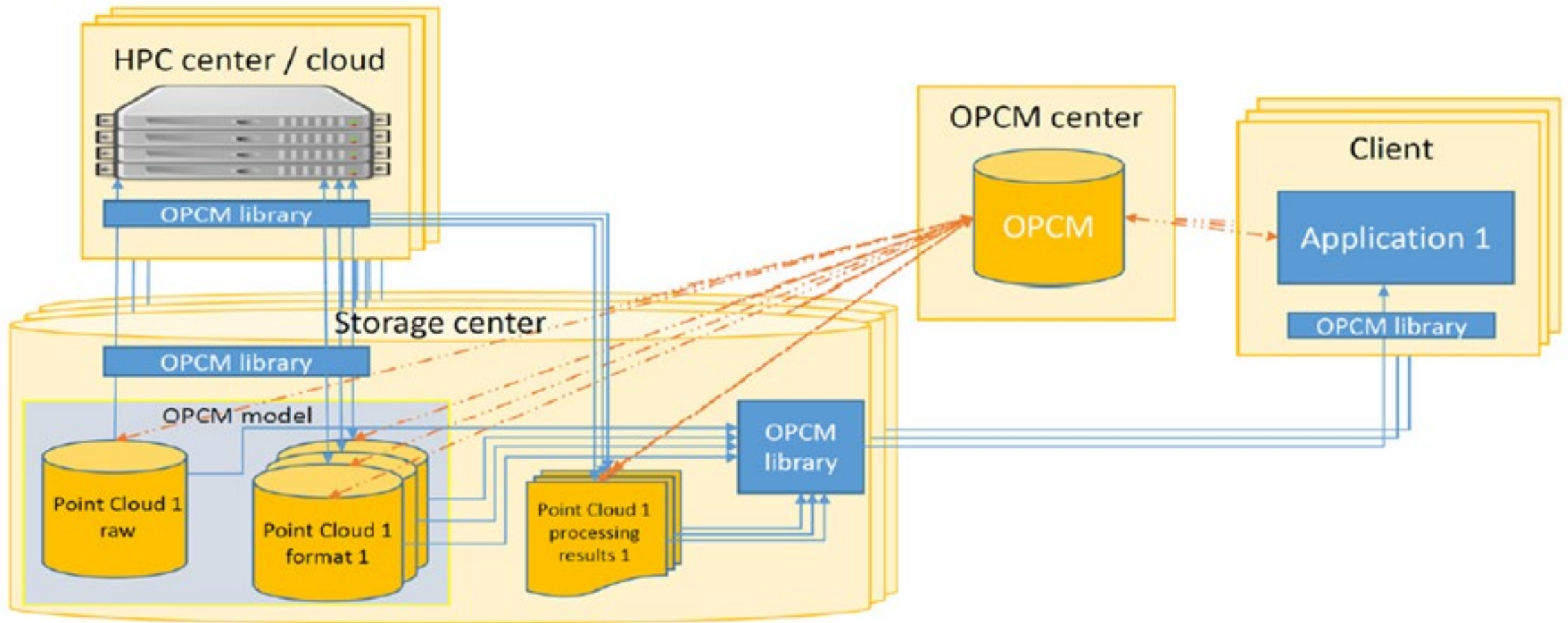
Overview GIS facilities

- gdmc.nl (=gist.bk.tudelft.nl) CentOS 7 (64 bits), 2 cpus, 394 Gb server for dissemination, events, papers, ..

- ahn2.pointclouds.nl (=ahn.bk.tudelft.nl) Ubuntu 16.04.4 LTS, 4 cpus, NAS 5 Tb on svm104.storage.tudelft.net server for 3D interactive PC viewer whole NL (potree)

- external domain names: isoladm.org (Wiki ISO 19152), pointclouds.nl (NLeSc massive PC), maps4society.nl (NWO programme with practice), nd-pc.org, geo5d.nl, maps4science.nl

- software: Oracle (Spatial and Graph), Enterprise Architect, FME, Esri ArcGIS Pro, LASTools, .., open source: PostgreSQL/PostGIS, QGIS, Potree, PDAL, Entwine,..

Laser scan data

- most data via partners:
  1. AHN (Het Waterschapshuis/RWS)
  2. Cyclomedia
  3. USGS 3D Elevation Program (3DEP)
     34,230,375,700,789 points

- Apple iPad Pro (2021) 12.9 inch 2TB

- equipment via our partners, e.g.:
  1. Leica Nederland – BLK2GO scanner
  2. Geology, CiTG – ZEB Revo RT

iPad Pro

LiDAR. The LiDAR Scanner measures how long it takes light to reflect back from objects, so it can create a depth map of any space you're in and unlock immersive AR experiences. And it works with the powerful ISP to more accurately focus images and videos in low-light conditions and reduce capture time.

HP DL380p Gen8 → pakhuis, physical server for benchmarks, DBMS, parallel computing

HP DL380p Gen8 server

1. 2 x 8-core Intel Xeon processors (32 threads), E5-2690 at 2.9 GHz
2. 128 GB main memory (DDR3)
3. Linux RHEL 6.5 operating system

- **Disk storage – direct attached**
  1. 2*41 TB SATA 7200 rpm in RAID-5 (external Yotta-III)
 2. 6 TB SAS 15K rpm in RAID 5 (internal)
  3. 400 GB SSD (internal)

- **NAS 26 TB on svm104.storage.tudelft.net**

# Exadata X4-2: Oracle SUN hardware for Oracle database software (previously 'contrib by Oracle')

- database Grid: multiple Intel cores, computations
  Eight, quarter, half, full rack with resp. 24, 48, 96, 192 cores

- storage Servers: multiple Intel cores, massive parallel smart scans
  (predicate filtering, less data transfer, better performance)

- hybrid columnar compression (HCC): query + archive modes

Exadata DB Machine X4-2 Quarter Rack
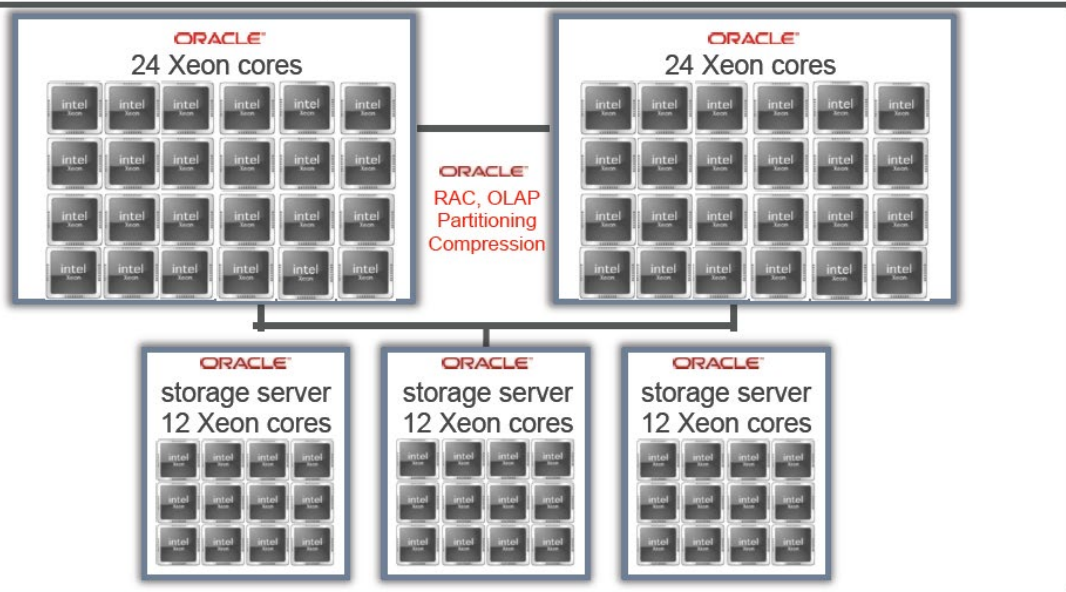
Delft High-Performance Computer: Delft Blue

- nr. 250 (in world top-500), 2 petaflops, available from January 2022

- Fujitsu's innovations, hardware and expertise are combined with the TU Delft expertise (esp. team of Kees Vuik, Prof Numerical Analysis)

- 400 nodes of 3 types: 1. x86 CPU compute intensive, 2. memory intensive max 1,5 TB of main memory, and 3. GPU Nvidia Tesla V100s

- high-speed parallel storage subsystem: BeeGFS / HDR100 Infiniband

- BK has 6% use right
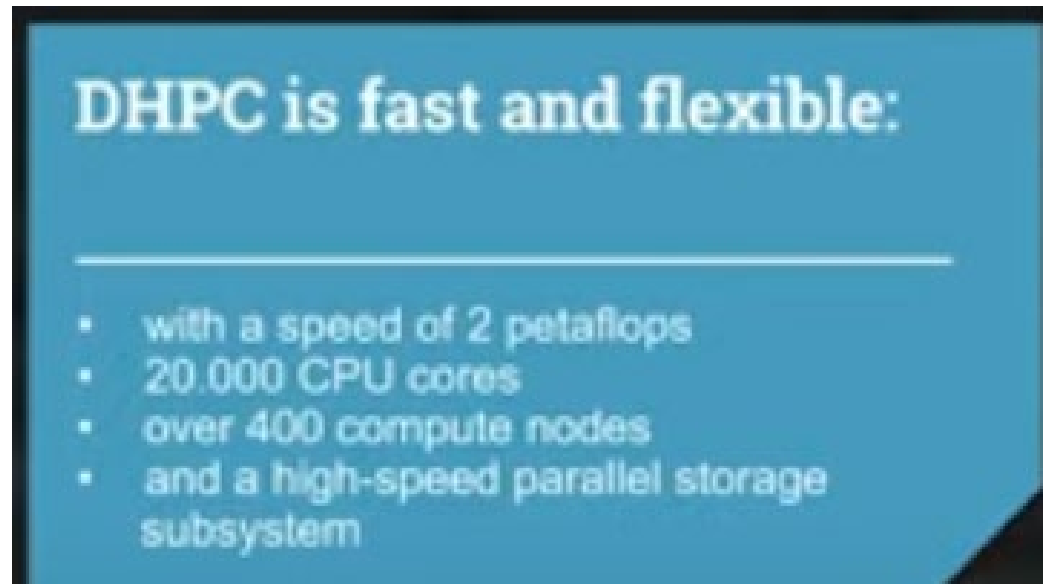
Overview

- project details
- direct use of point clouds
- Open Point Cloud Map
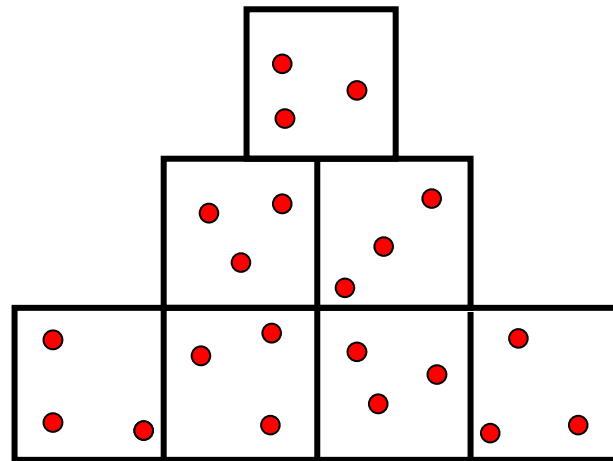- nD-PointClouds
- big data management
- conclusion

Motivation nD-PointCloud

- point cloud data sets are often used for monitoring
  → dynamic point clouds
  → time added as additional organizing dimension

- organizing point cloud data in LoD's/importance levels is an approach to manage large data sets
  → LoD: discrete (multi-scale) or continuous (vario-scale)
  → scale treated as additional organizing dimension

- how to manage higher dimensional point clouds (4D, 5D)

Large number of points (e.g. whole of NL…)

- overview queries just want subset
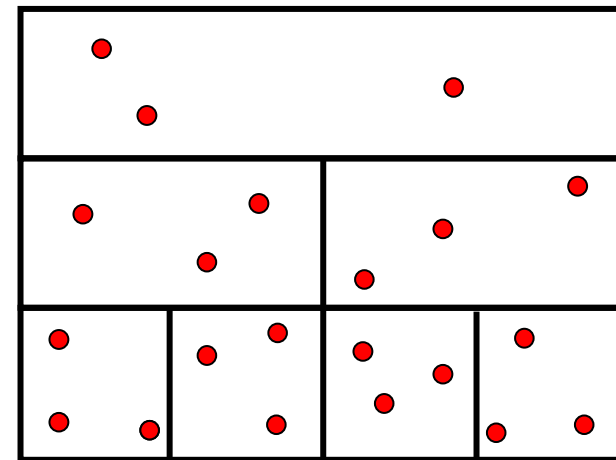- organize in data pyramid (LoD)

2D schematic view, data blocks….   stretched over domain   density low
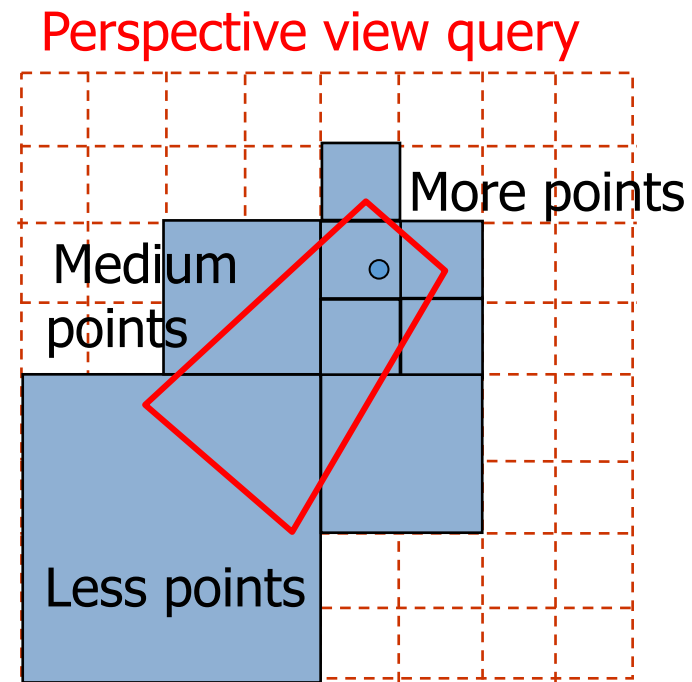
LoD 2

LoD 1

LoD 0

high

Data pyramid/multi-scale

- allows fast spatial searching including LoD selection

- the further away from viewer
  the lesser points selected (i.e.
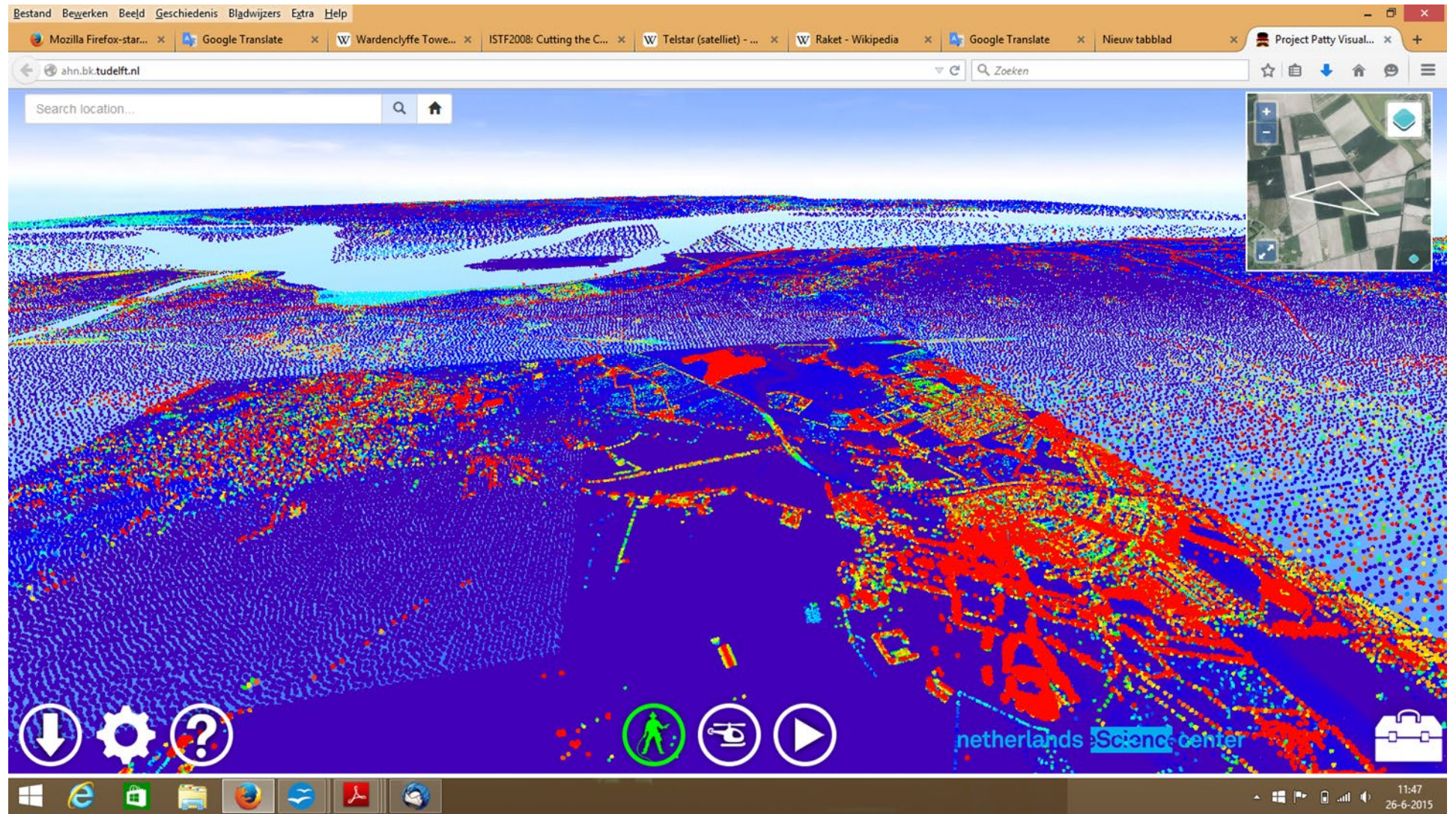 the higher level blocks/points)

- drawbacks:
  1. discrete number of levels
  2. filling/creating structure

Perspective view query

More points

Medium points

Less points

http://ahn2.pointclouds.nl (outdated from previous project)

640.000.000.000 points on-line 3D viewer
bit newer version but with less functionality at https://potree.entwine.io/data/ahn.html
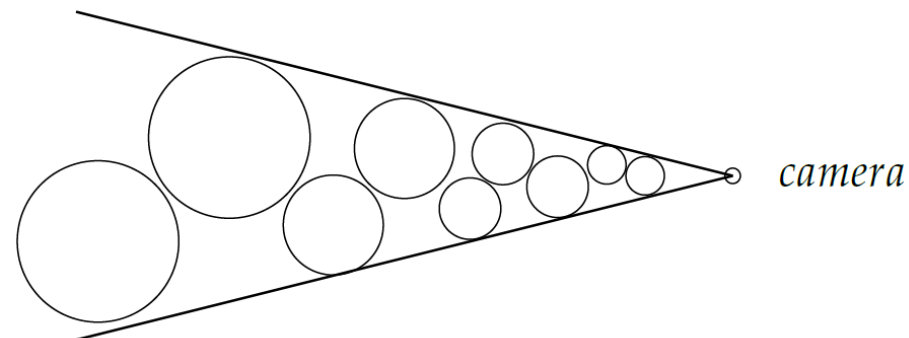
Getting rid of discrete level postprocessing (1/2)

- MSc Geomatics thesis 'Vario-scale visualization of the AHN2 point cloud' by Jippe van der Maaden (TU Delft, April 2019).
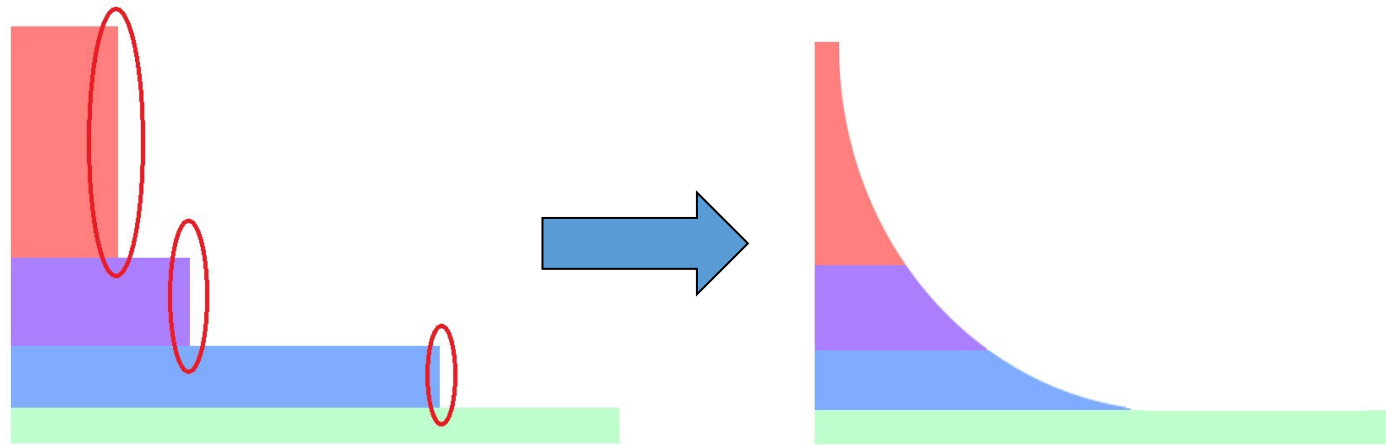
- assign radius to points based on distance to camera

- select points with empty circle/sphere

- remove 70-90% of points

Getting rid of discrete level postprocessing (2/2)

Real-Time Continuous Level of Detail Rendering of Point Clouds

Markus Schütz*
TU Wien

Katharina Krösl†
TU Wien,
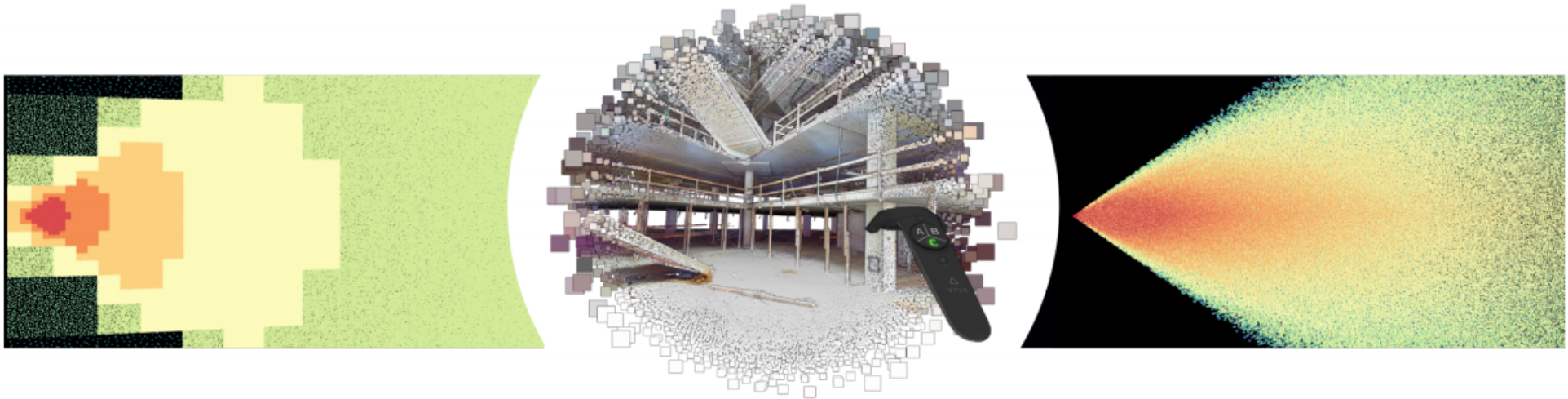VRVis Forschungs-GmbH

Michael Wimmer‡
TU Wien

Figure 1: Left: Discrete LOD structure with sudden jumps in density and popping artifacts under motion. Middle: Screenshot of our method in virtual reality. No visible seams across different levels of detail due to a continuous reduction in density, and also continuously reduced density towards the periphery to reduce geometric complexity where less is needed. Endeavor point cloud courtesy of NVIDIA [4]. Right: Continuous transition from one LOD to another.

paper at IEEE VR 2019 (postprocessing at GPU, from 86.000.000 to 5.000.000 points every 5 to 6 frames → 90 fps left+right = 180 fps)

Further improvements ... beyond discrete levels

- might result in artefacts when looking at perspective view image (possible 'see' blocks of different levels)

- also not optimal within block (near viewer perhaps not enough points, further from viewer perhaps too much points)

- would a true vario-scale option possible?

    → Vario-scale geo-info research at TU Delft

Vario-scale for point cloud data

- lesson from vario-scale research: add one continuous dimension to the geometry to represent scale -> continuous LoD
(2D data vario-scale represented by 3D geometry)

- apply this to point cloud data
  1. compute the imp/cLoD value
  2. add this as organizing dimension, either x,y,imp (z and others attributes) or x,y,z,imp (and others as attributes)
  3. Cluster/index the 3D or 4D point
  4. Define perspective view selections, view frustum with one more dimension: the further, the higher imp's
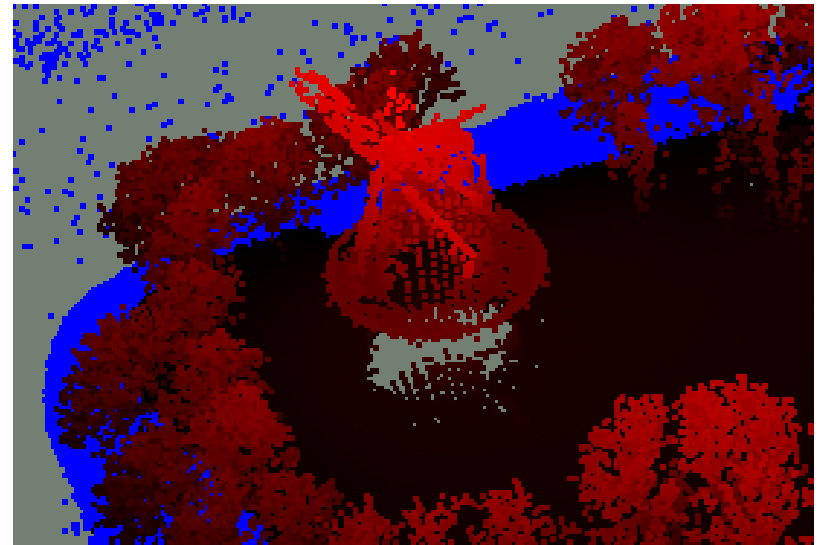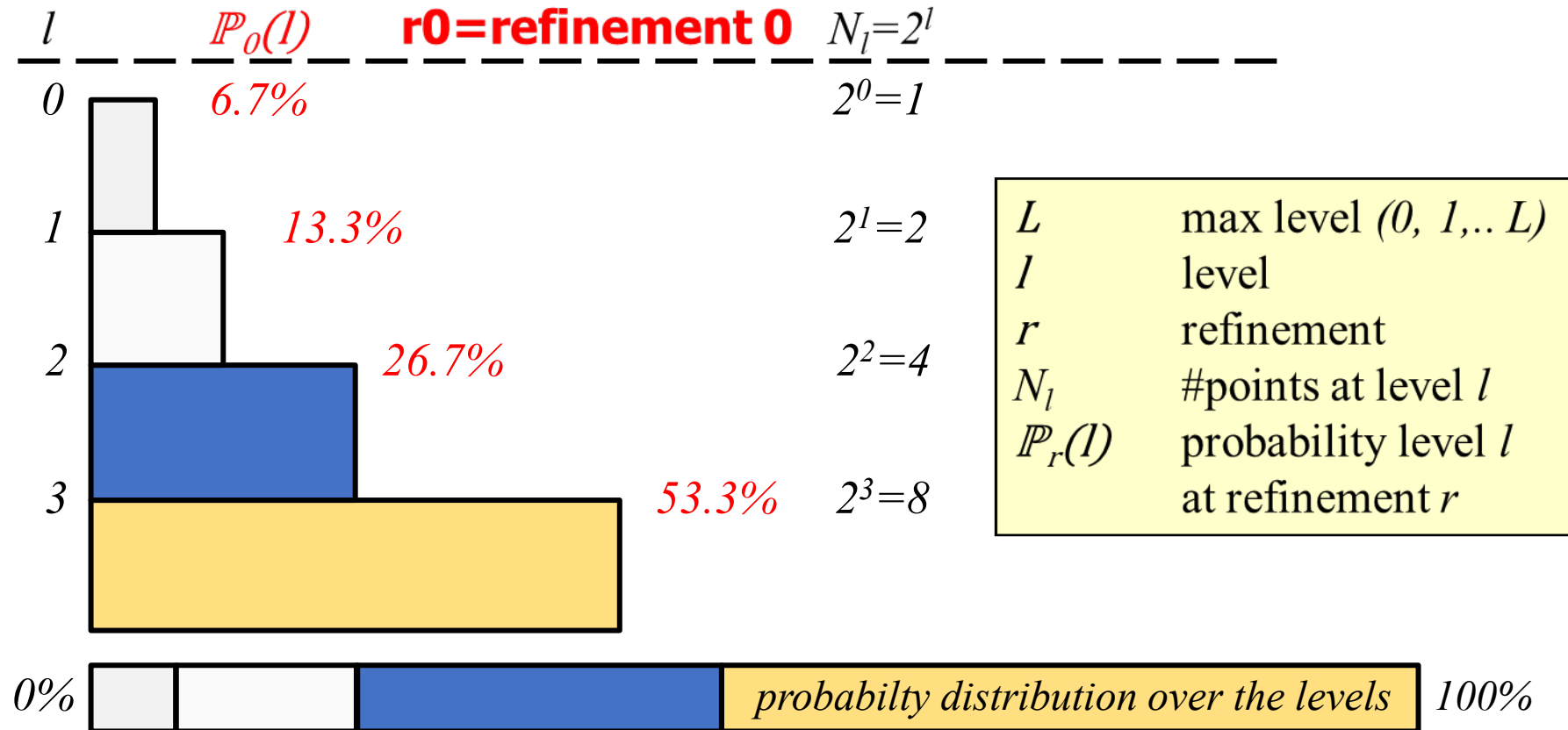
How to create cLoD distribution?

- same type of distribution as ideal data pyramid

- but without discrete levels

- many years no clue

- 2019: finally a solution, keynote at ISPRS Geospatial week…
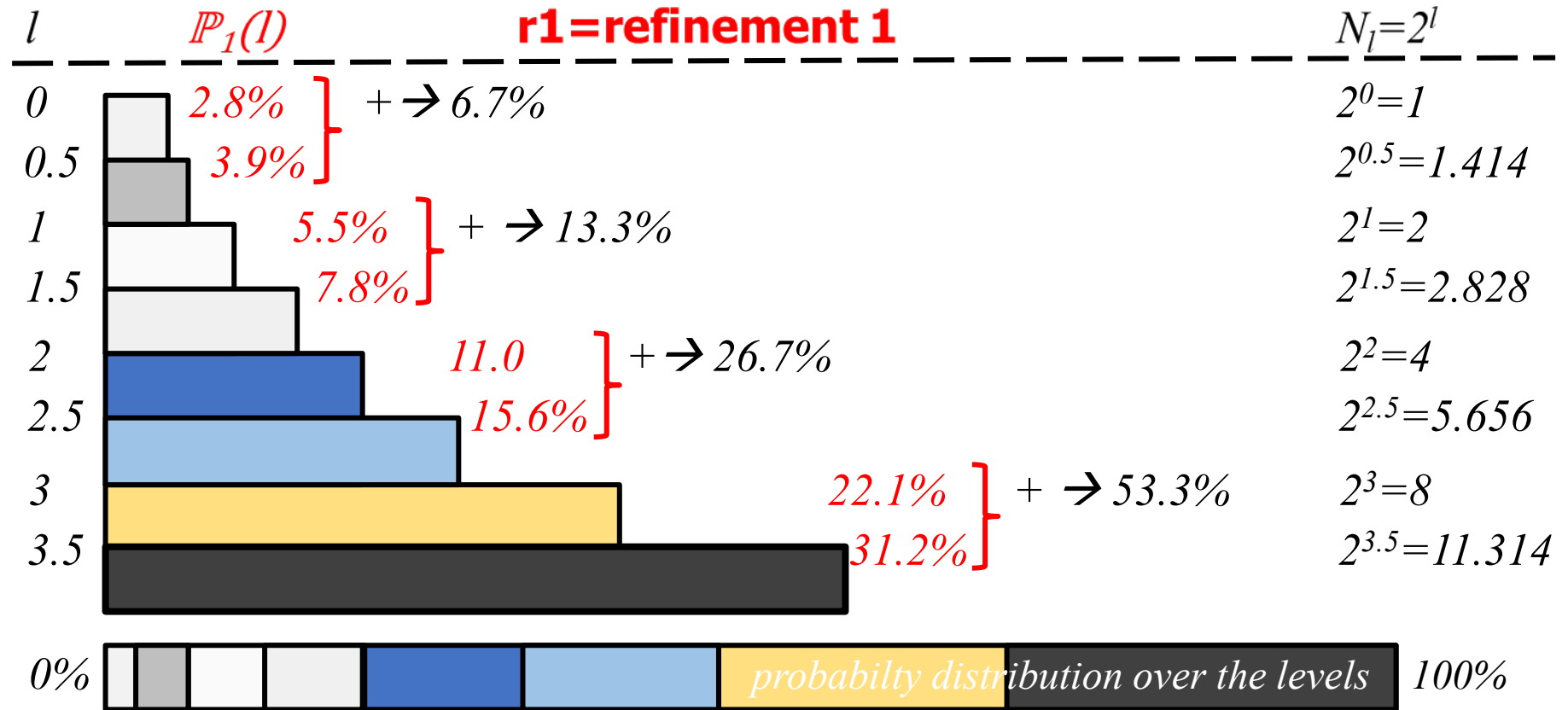
Ideal distribution refined LoD r0

- **1D example**, duplicate data (per dimension) for next level: $N_l=2^l$
- same principle as vector, raster LoD's (and quadtree/octree approach)

l	$\mathbb{P}_0(l)$	**r0=refinement 0**	$N_l=2^l$
0	6.7%		$2^0=1$
1	13.3%		$2^1=2$
2	26.7%		$2^2=4$
3	53.3%		$2^3=8$

L	max level $(0, 1,.. L)$
l	level
r	refinement
N_l	#points at level l
$\mathbb{P}_r(l)$	probability level l at refinement r

0%   *probabilty distribution over the levels*   100%

example with L=3 (4 levels)

Ideal distribution refined LoD r1
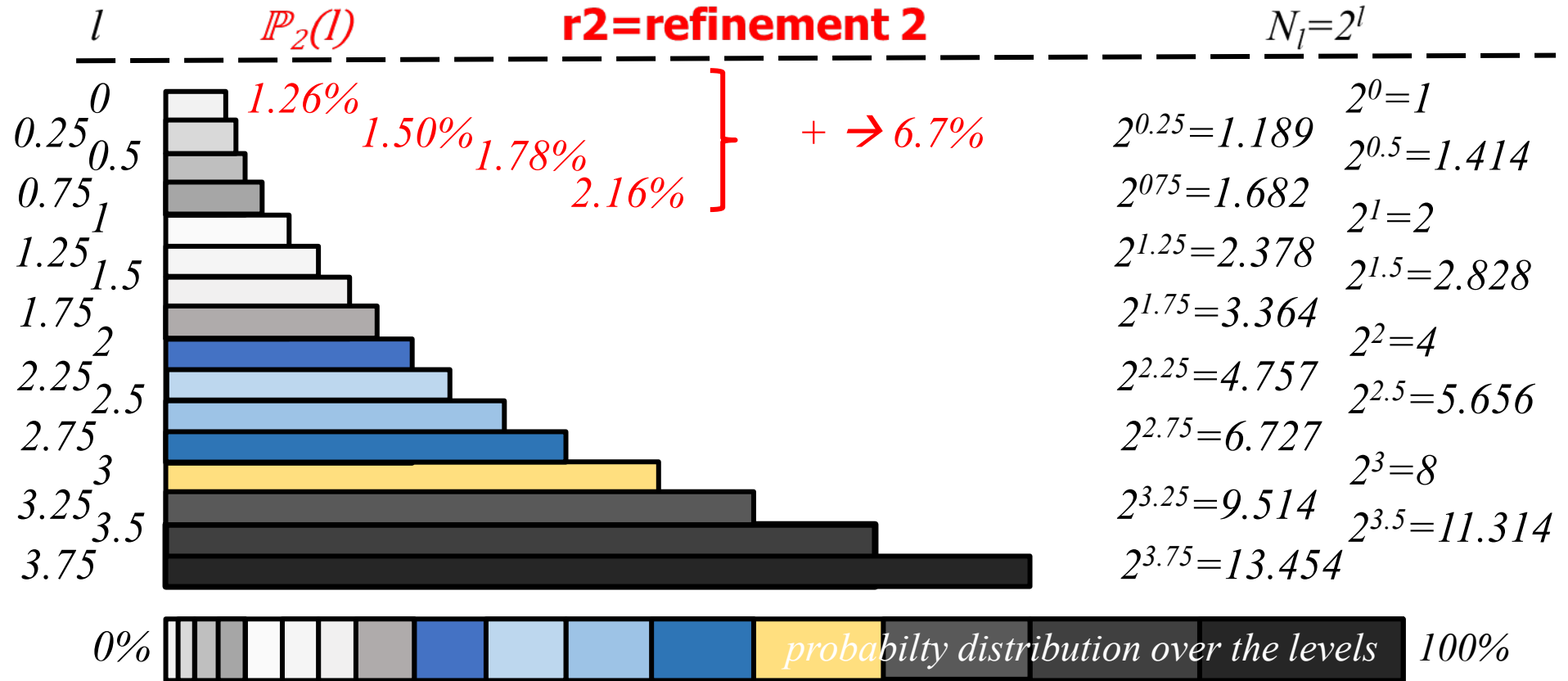
- next to integer levels also compute ideal data amount for half levels
- use exactly same for formula: $N_l = 2^l$

l	$\mathbb{P}_1(l)$	**r1=refinement 1**	$N_l = 2^l$
0	2.8%	$+ \to$ 6.7%	$2^0 = 1$
0.5	3.9%		$2^{0.5} = 1.414$
1	5.5%	$+ \to$ 13.3%	$2^1 = 2$
1.5	7.8%		$2^{1.5} = 2.828$
2	11.0	$+ \to$ 26.7%	$2^2 = 4$
2.5	15.6%		$2^{2.5} = 5.656$
3	22.1%	$+ \to$ 53.3%	$2^3 = 8$
3.5	31.2%		$2^{3.5} = 11.314$

0%   *probabilty distribution over the levels*   100%

summing *r1* probabilities *l=0* and *l=0.5* totals to *r0* probability for *l=0*

Ideal distribution refined LoD r2

- next refinement, also compute for quarter levels
- use exactly same for formula: $N_l=2^l$

l	$\mathbb{P}_2(l)$	**r2=refinement 2**	$N_l=2^l$

$$0 \qquad 1.26\%$$
$$0.25$$
$$0.5 \qquad 1.50\%\ 1.78\%$$
$$0.75$$
$$1 \qquad 2.16\%$$
$$1.25$$
$$1.5$$
$$1.75$$
$$2$$
$$2.25$$
$$2.5$$
$$2.75$$
$$3$$
$$3.25$$
$$3.5$$
$$3.75$$

$$+ \to 6.7\%$$

$2^0=1$
$2^{0.25}=1.189$
$2^{0.5}=1.414$
$2^{075}=1.682$
$2^1=2$
$2^{1.25}=2.378$
$2^{1.5}=2.828$
$2^{1.75}=3.364$
$2^2=4$
$2^{2.25}=4.757$
$2^{2.5}=5.656$
$2^{2.75}=6.727$
$2^3=8$
$2^{3.25}=9.514$
$2^{3.5}=11.314$
$2^{3.75}=13.454$

0% *probabilty distribution over the levels* 100%

note: Ideal distribution is maintained (check by summing first 4 levels)

Getting rid of (refined) discrete levels
→ real continuous levels, nD case

- for ideal continuous distribution function over levels (nD):

$$f(l, n) = \frac{2^{(n-1)l}(n-1)\ln 2}{2^{(n-1)(L+1)} - 1}$$
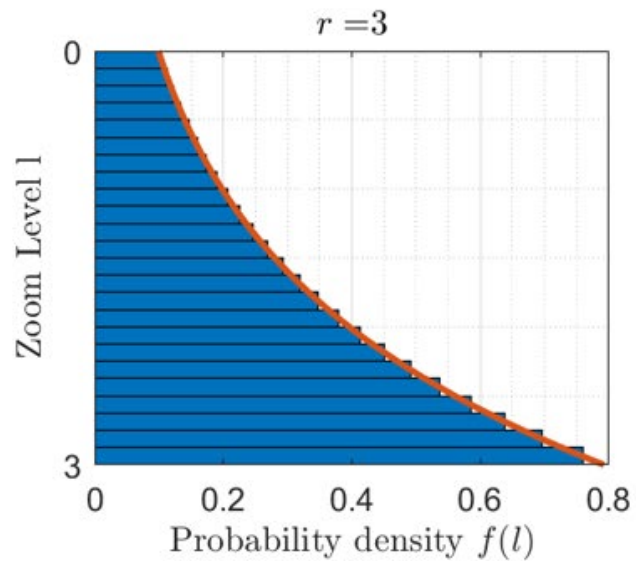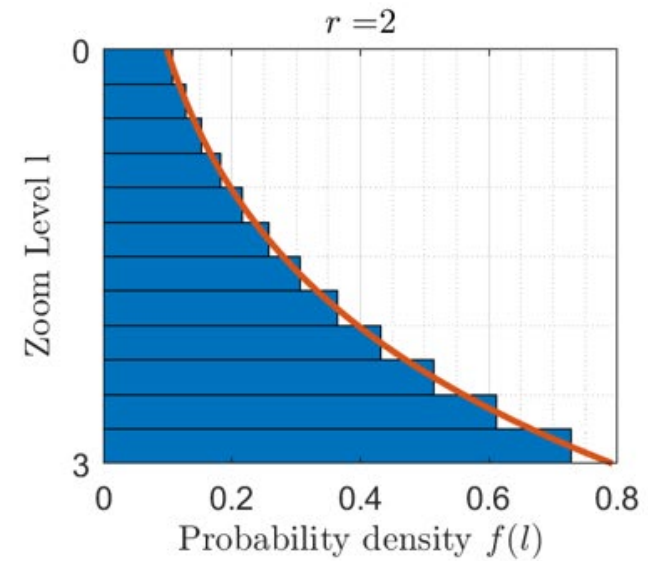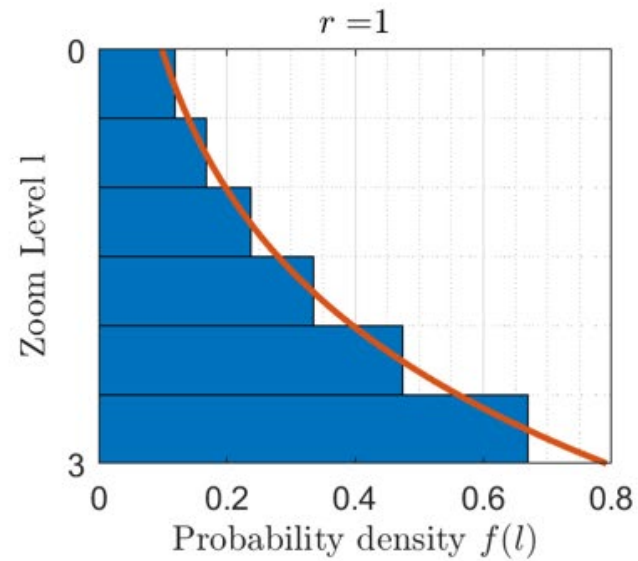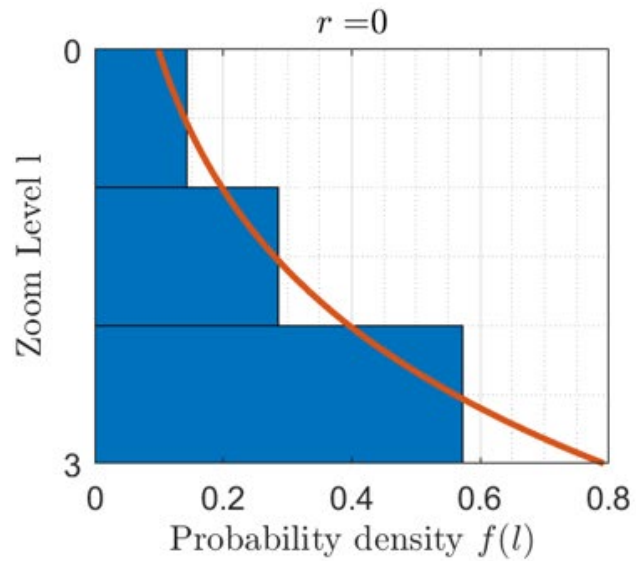
for l between 0 and $L+1$
and n number of dimensions

- this function has Cumulative Distribution Function (CDF):

$$F(l, n) = \frac{2^{(n-1)l} - 1}{2^{(n-1)(L+1)} - 1}$$

for l between 0 and $L+1$
and n number of dimensions

- using random generator U (uniform between 0 and 1) to generate level l (between 0 and $L+1$) for next point in nD space:

$$l = \frac{\ln((2^{(n-1)(L+1)} - 1)U + 1)}{(n-1)\ln 2}$$
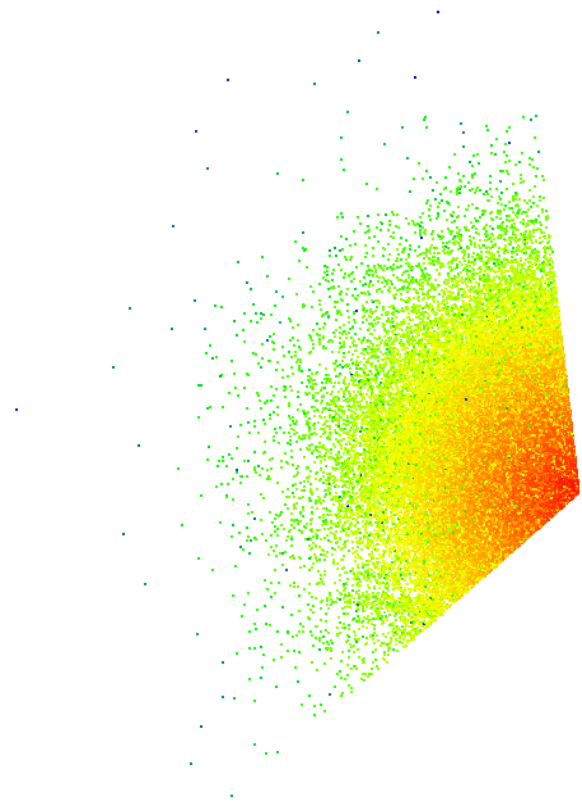
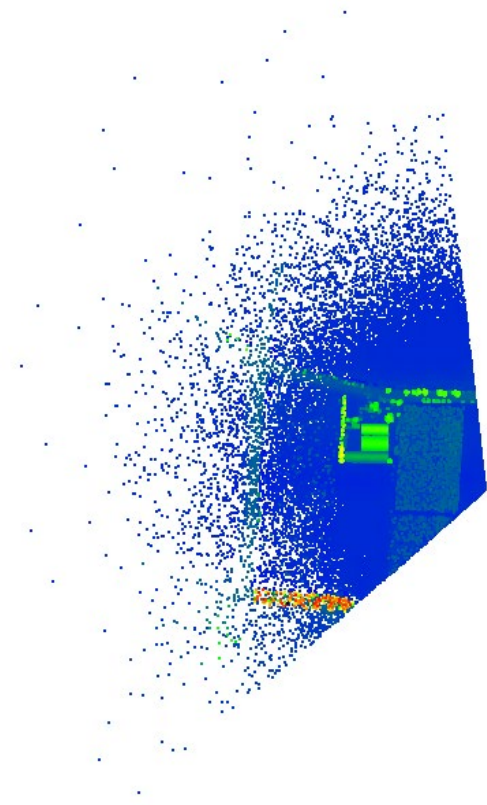# More refinements → continous function
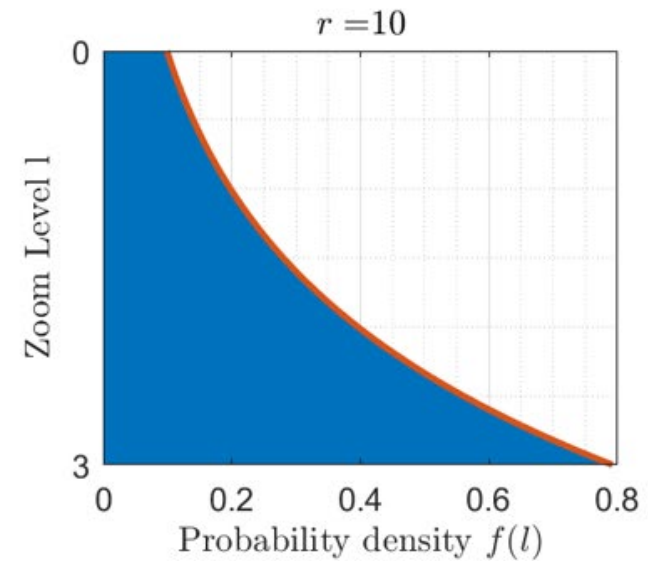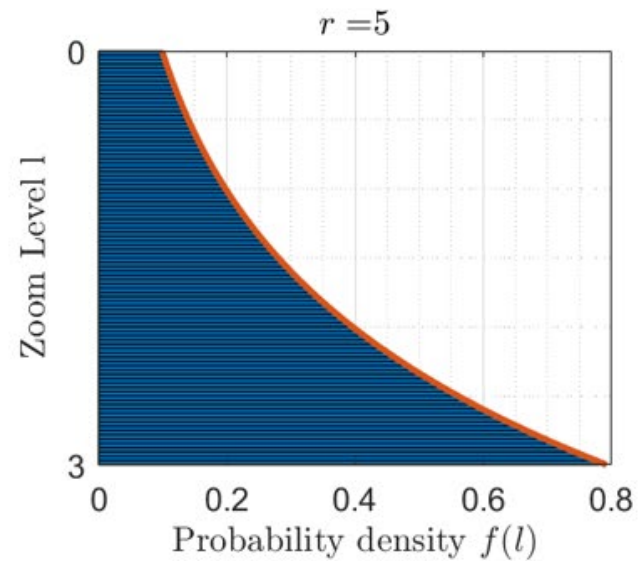
Perspective view
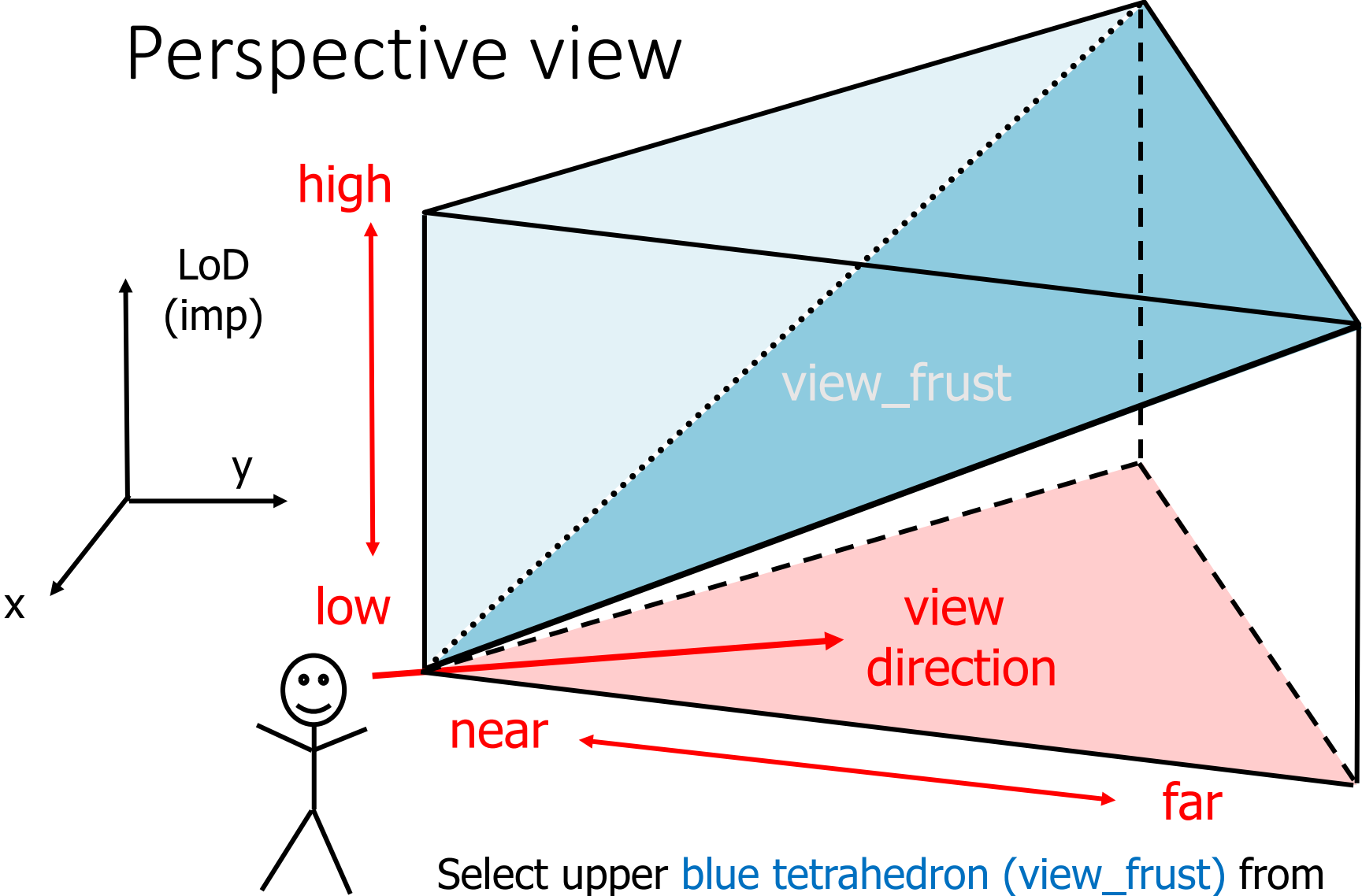
Select upper blue tetrahedron (view_frust) from prism-part of vario-scale x,y,imp point cloud data cube

cLoD perspectief view query (3rd person view)
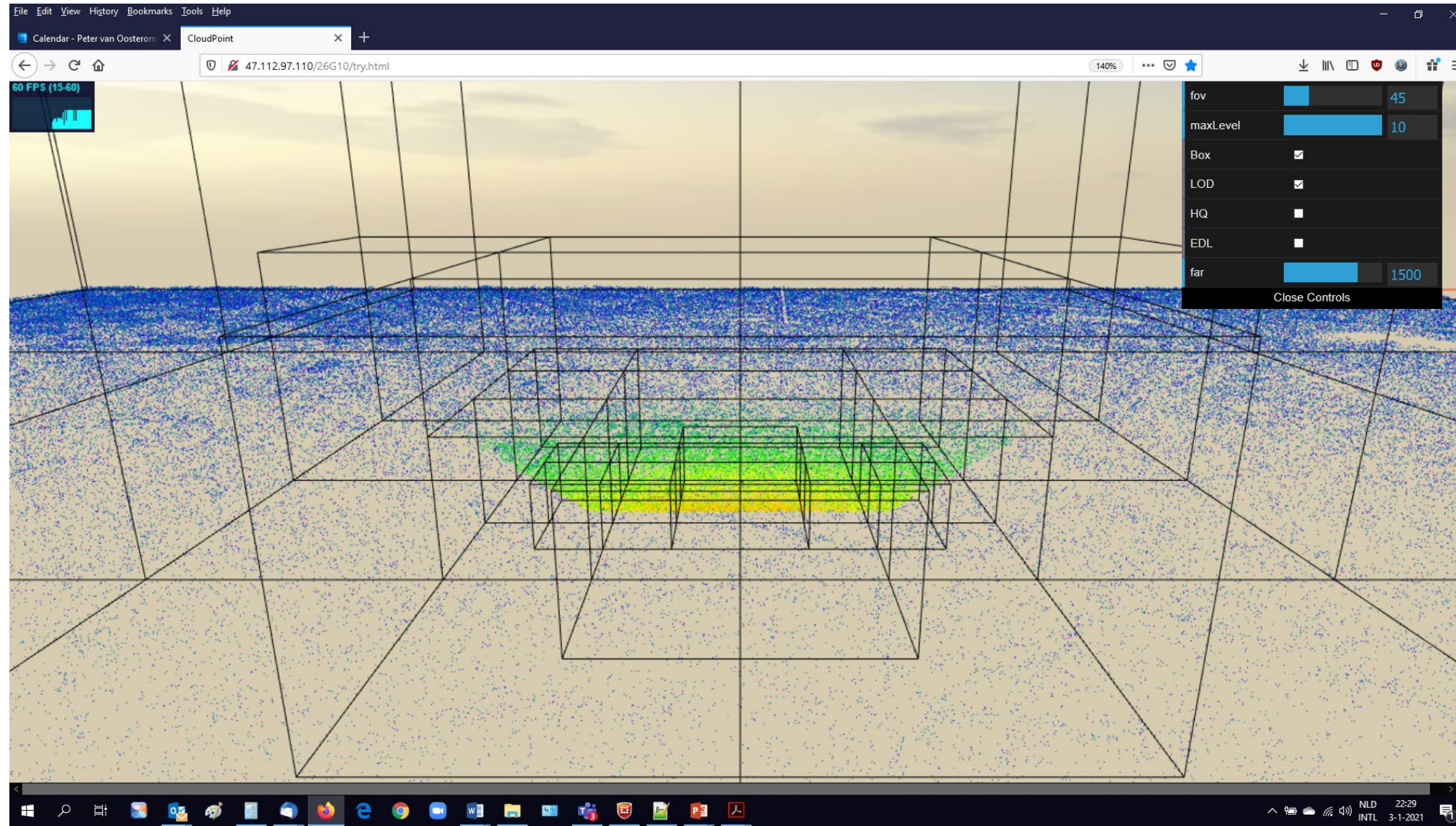

colored by cLoD   and Z

from PhD research Haicheng Liu

cLoD in 3D web application

developed by Xuefeng Guan (Wuhan Univ) after 1 year visit TUD

note: colour is density (3rd person view)

Augmented Reality (AR) & cLoD point clouds on smartphone

- **MSc thesis 2020 Liyao Zhang**

- challenges:

  1. dealing with huge-amount data of point cloud datasets with limited memory, CPU and GPU resources of mobile devices

  2. reaching relatively high visual quality and performance requirements

- using ARCore / Unity game engine: The point cloud model in the scene will be updated every X frames accords to new cLoD based selection result. X is a parameter called UpdateFrequency

Potential Applications

- outdoor: architecture, industrial design

- indoor: home renovation, estate sales, architectural design

- outdoor: change visualization (AHN3-AHN4), detection, measurement
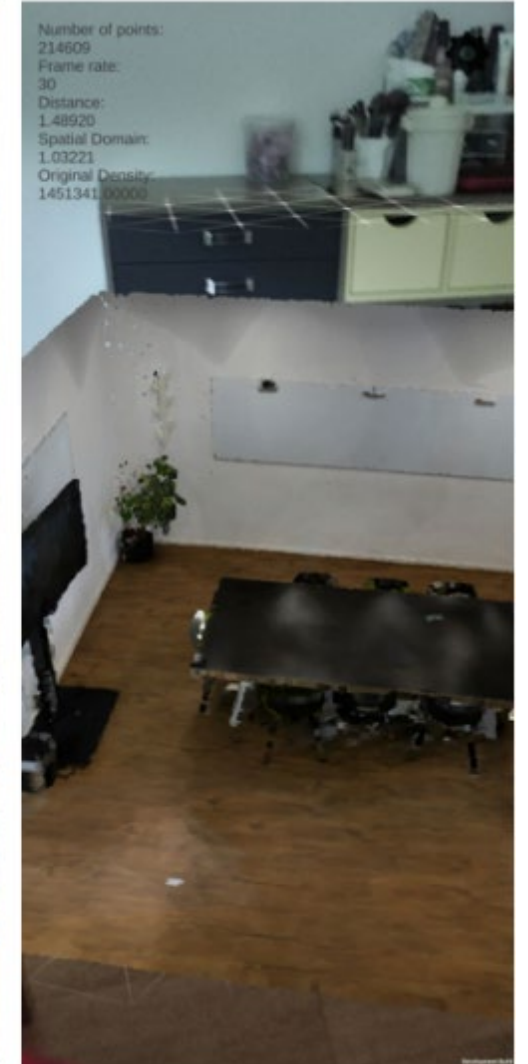
- …

Overview

- project details
- direct use of point clouds
- Open Point Cloud Map
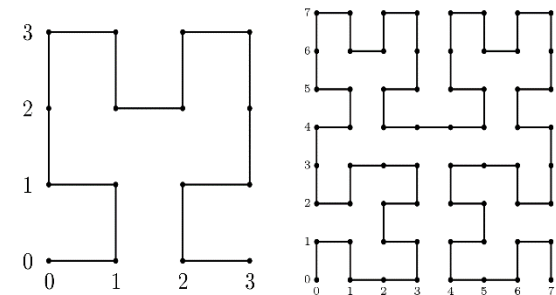- nD-PointClouds
- big data management
- conclusion

In detail: Space Filling Curves (SFCs)

- from nD to 1D (computer memory addresses) and back…

Morton (Peano)

- apply linear ordering to a multidimensional domain (spatial clustering)
- organize a flat table efficiently
- full resolution keys: avoid storing x,y[,z] + t/l
  → recovered from SFC key

- use Index Organized Table (data stored in the B-Tree index)
- queries need to be re-written to SFC-ranges, benefit from spatial clustering → efficient

Hilbert

- SFCs based on hyper-cubes
  - Morton/Hilbert both nD and quadrant recursive
  - Consider relative scaling of dimensions
 - Space reserved on the hypercube for future data
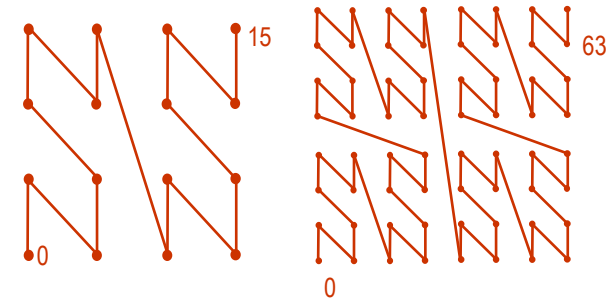
nD-PointClouds data management

- use DataBase Management System (DBMS), possibly distribued
- based on modelling theory for nD point cloud data
- tools to support modelers, developers and users in point cloud data organization design decisions for (given 1. data sets and 2. required functionalities in applications):
  - what are the organizing dimensions,
  - what are the associated attributes,
  - what type of organization: Morton-code/ kd-tree/ nD simplices-part,
  - what relative scale of various dimensions,
  - parameters such as clustering/ blocking size,
  - what compression,
  - what approach and level of parallelism (incl. hardware aspects),

Overview: load & query point clouds

loading (upper right of figure)
- covert nD points to SFC keys
- store in IOT table (organized on SFC key)
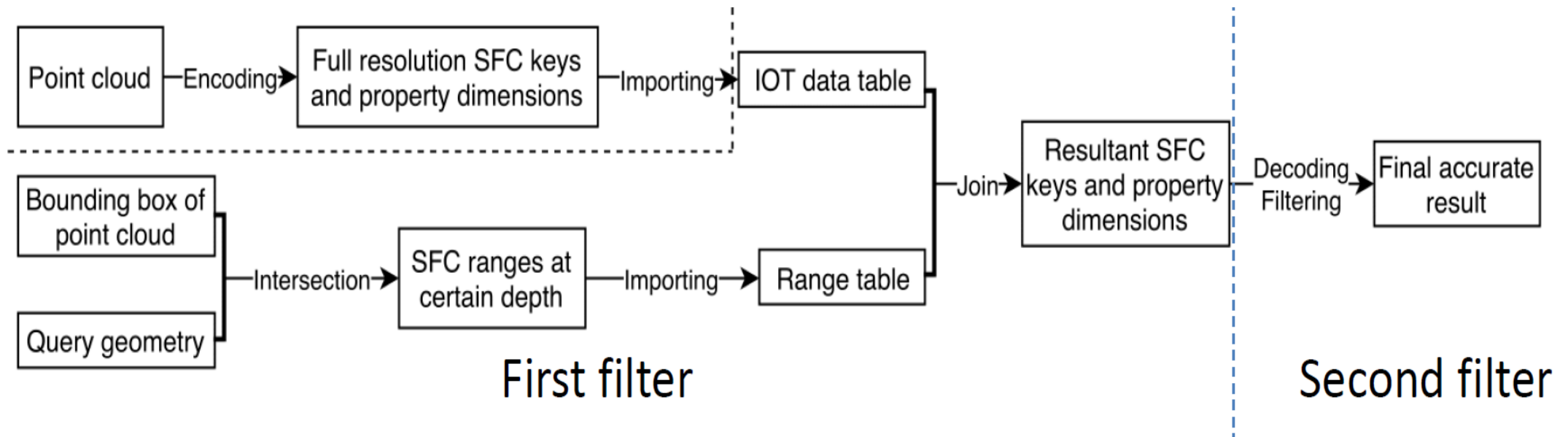
querying (remainder)
- convert query geometry to SFC ranges,
- first filter: join SFC range table with IOT table
- second filter (optional): test points inside query geometry

Overlap_Codes (query_geometry, domain) → ranges

- based on concepts of Region Quadtree & Quadcodes
- works for any type of query geometry (point, polyline, polygon)
- also works in 3D (Octree) and higher dimensions
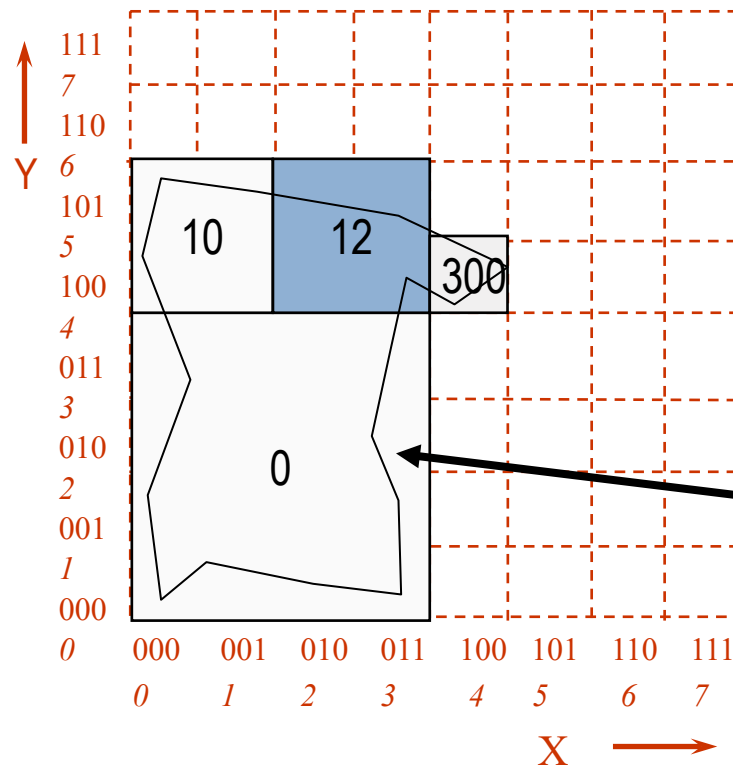
Quadcode 0:   Morton range 0-15

Quadcode 10:   Morton range 16-19

Quadcode 12:   Morton range 24-27

Quadcode 300:   Morton range 48-48

(Morton code gaps resp. 0, 4, 20)

query_geometry, polygon

Note : SW=0, NW=1, SE=2, NE=3

Quadcells / ranges and queries

```
CREATE TABLE query_results_1 AS (

SELECT * FROM

  (SELECT x,y,z FROM ahn_flat WHERE
  (hm_code between 1341720113446912 and 1341720117641215) OR
  (hm_code between 1341720126029824 and 1341720134418431) OR
  (hm_code between 1341720310579200 and 1341720314773503) OR
  (hm_code between 1341720474157056 and 1341720478351359) OR
  (hm_code between 1341720482545664 and 1341720503517183) OR
  (hm_code between 1341720671289344 and 1341720675483647) OR
  (hm_code between 1341720679677952 and 1341720683872255)) a

WHERE (x between 85670.0 and 85721.0)

  and (y between 446416.0 and 446469.0))
```
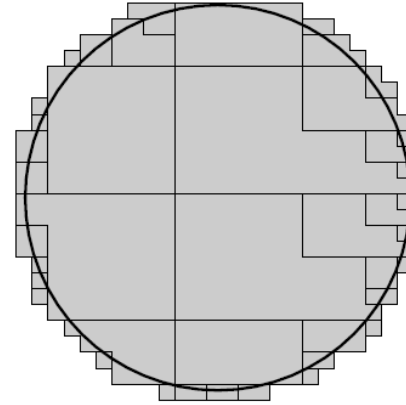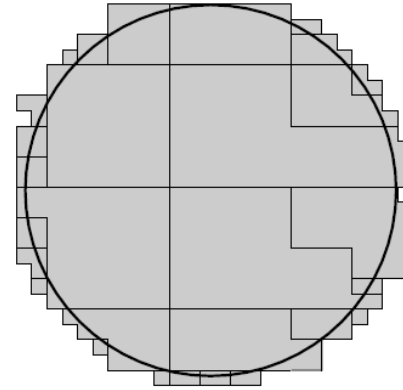
Query 1 (small rectangle)
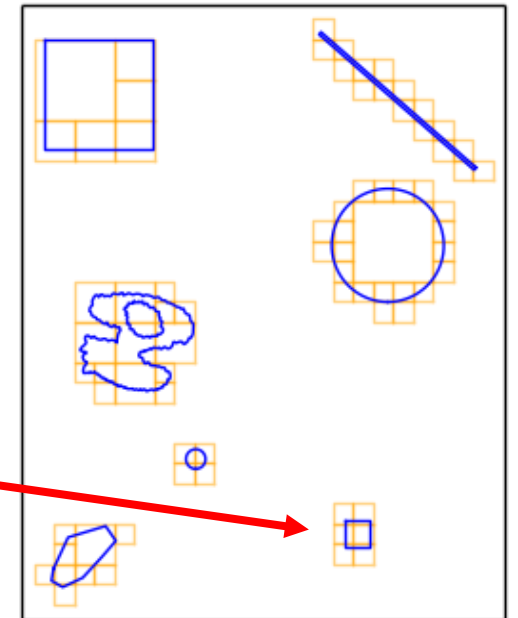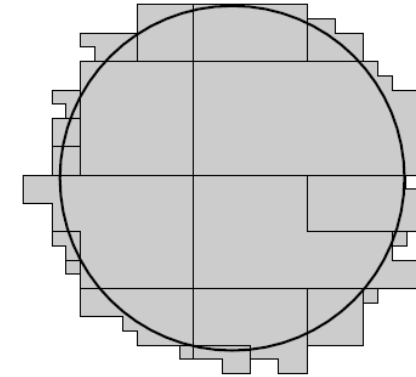
Merge ranges → add space
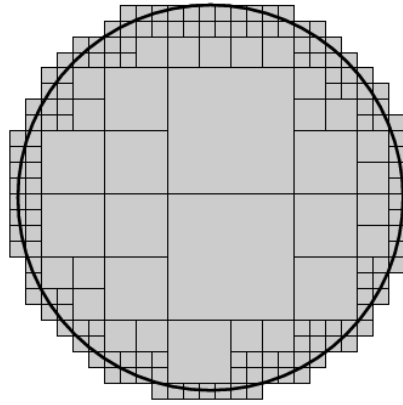
(a) Original 176 tree cells

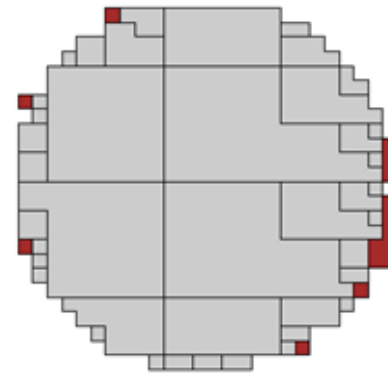(b) Merging of direct neighbours leading to 42 cells
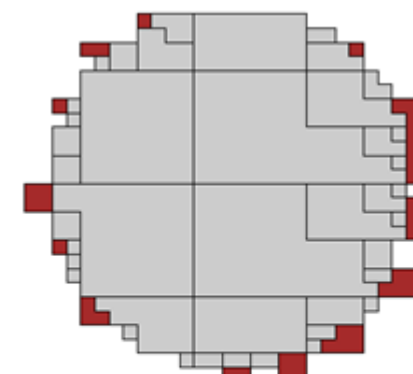
(c) Merged to maximum number of 30

(d) Merged to maximum number of 20

effect of range merging with gaps is adding space (with possible more unwanted point inside) → red area

(e) The expansion of the original geometry (case (b)) in red after merging to 30

(f) The expansion of the original geometry (case (b)) in red after merging to 20

source: MSc thesis Stella Psomadaki, 2016

Drawback of high dimensional SFC?

- nD SFC keys have benefits: space-time-scale (and perhaps even other attributes) in compact organization

- may select on multiple/trans dimensions at same time efficiently (perspective view)

- possible drawbacks of high dimensional point cloud:
  1. need big SFC code (128 bits number or other encoding, like varchar)
  2. if just limited number of dimensions are specified for selection → other dimensions than range form min-to-max: 'infinite tall prisms' many (empty?) cells, what are the query performance consequences

- needs further exploration
  (as the relative scaling of dimensions need attention → basis for defining trans-dimension distance → actual grouping/ clustering)

Overview

- project details
- direct use of point clouds
- Open Point Cloud Map
- nD-PointClouds
- big data management
- conclusion

nD-PointClouds/OPCM will become reality…

... with this new eScience project !

3. **Platforms for Big & Complex Spatial Data**

nD-PointCloud continuous level representation for spatio-temporal phenomena in Open Point Cloud Maps

Prof. dr. ir. Peter J.M. van Oosterom (Delft University of Technology)

This innovative eScience technology proposal aims to make point clouds the primary representation of spatio-temporal features throughout the entire processing chain. This includes data acquisition, storage, analysis, visualization and dissemination. Today, point clouds are mainly used in the data acquisition phase, while gridded (raster) or object (vector) models are used in the other phases. Handling the extract-transform-load actions becomes an increasing problem in using big data. Based on a novel use of high-resolution nD space filling curves, this proposal will realize a deep integration of space, time and scale as the basis for data organization, enabling High Performance/Throughput Computing for enormous point clouds. By enabling operations directly on the raw point cloud data, nD-PointCloud represents a major advance in domains requiring lossless spatio-temporal data of an extremely high accuracy. A distributed Open Point Cloud Map (OPCM) infrastructure will be developed that supports the sharing of big data nD-PointCloud, and enables interactive real-time visualizations using perspective views without data density shocks, continuous zoom-in/out and progressive data streaming between the server and client. Applications from the water management domain will be used as Proof-of-Principle. If successful, nD-PointCloud will become the preferred model enabling progress in (research) fields like cultural heritage, land administration, vegetation monitoring, building modelling, transportation and mobility.