Tutorial: Tensor Approximation in Scientific Visualization

Scientific Visualization Applications

Susanne K. Suter, Rafael Ballester-Ripoll, Renato Pajarola

University of Zurich[UZH]

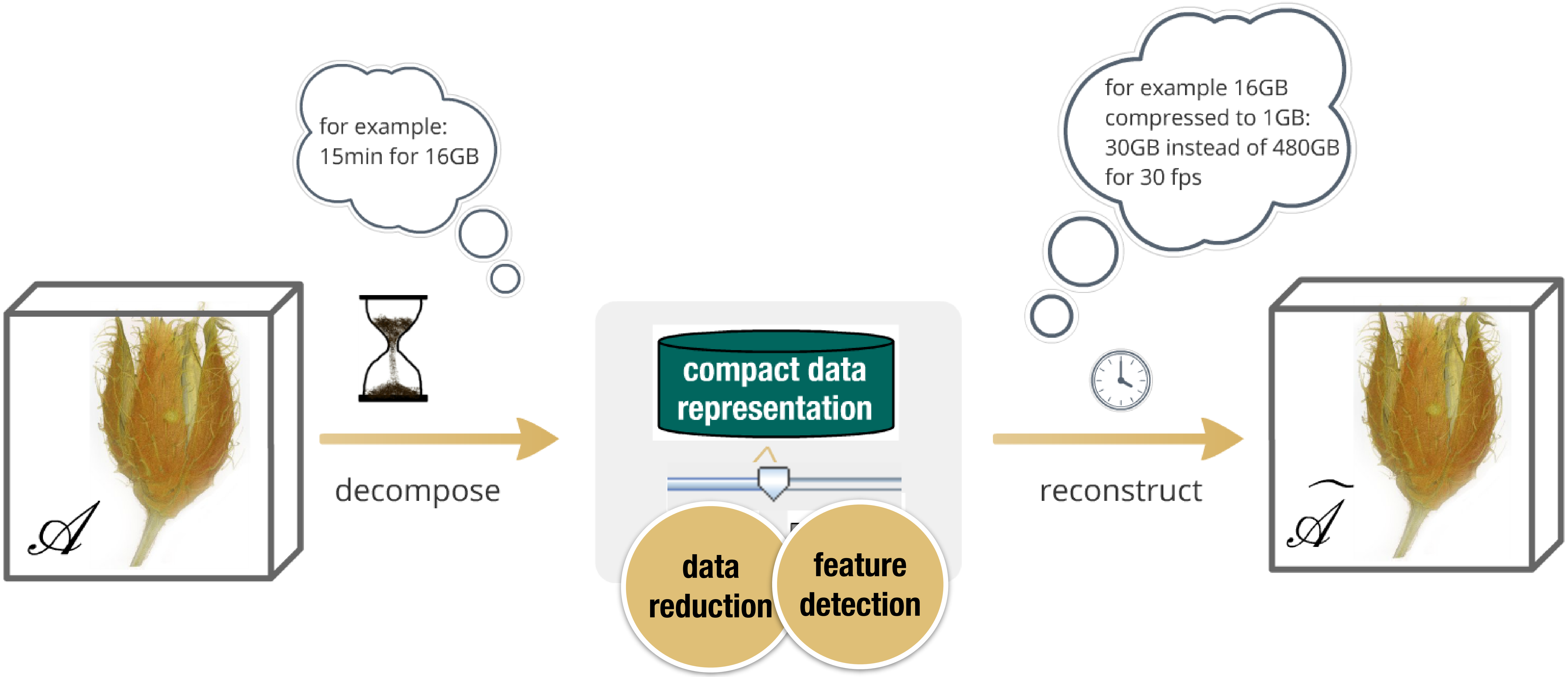
VISUALIZATIONAND MULTIMEDIALAB

Outline

- Part 1: Compact data representations compared
  - wavelets
  - tensor approximation (Tucker model)
  - compression and multiscale features
- Part 2: Multiresolution TA Hierarchies

Part 1: Compact Data Representations Compared

Compact Data Representation



key challenges:
- reduce storage costs and data transfer costs
- real-time reconstruction (higher-order linear algebra operations)
- compact representation with fast data access on GPU

Compact Data Representation in DVR

- Discrete cosine transform
  [Yeo & Liu, 1995]
- Fourier transform
  [Chiueh et al., 1997]
- Wavelet transform
  [Rodler, 1999; Guthe et al, 2002]
- Vector quantization
  [Schneider & Westermann, 2003; Fout & Ma, 2007; Parys & Knittel, 2009]
- Tensor approximation
  [Wu et al., 2008; Suter et al., 2010+2011+2013]
- Sparse coding
  [Gobbetti et al., 2012]

Balsa et al.. A Survey of Compressed GPU Direct Volume Rendering.
Eurographics state-of-the-art report, May 2013.

VISUALIZATIONAND MULTIMEDIALAB

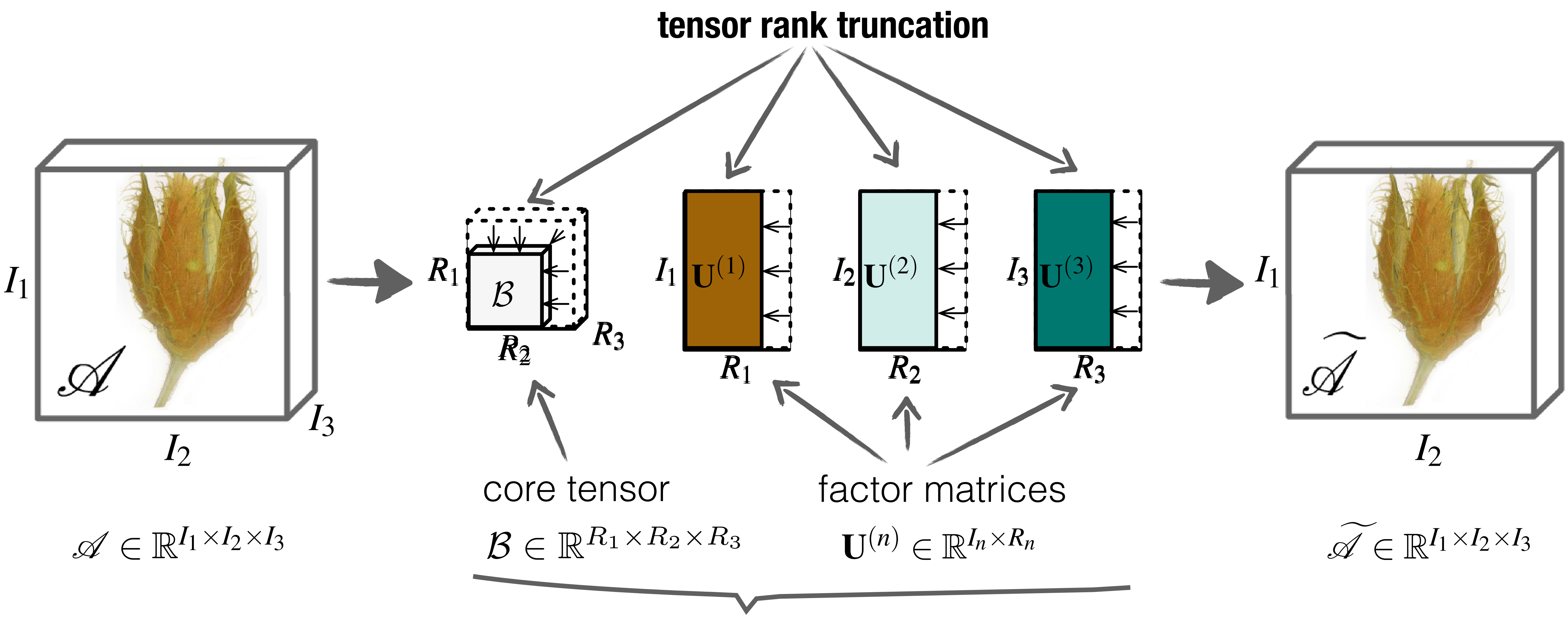
Feature Extraction

- Typically done with multiresolution analysis
  - significant components at low frequencies
    - details at high frequencies
- Visualize features at multiple spatial scales
  - achieved through tensor rank truncation

Suter et al., VMV 2010. Application of tensor approximation to multiscale volume feature representations.

VISUALIZATIONAND MULTIMEDIALAB

Tucker Tensor Rank Truncation



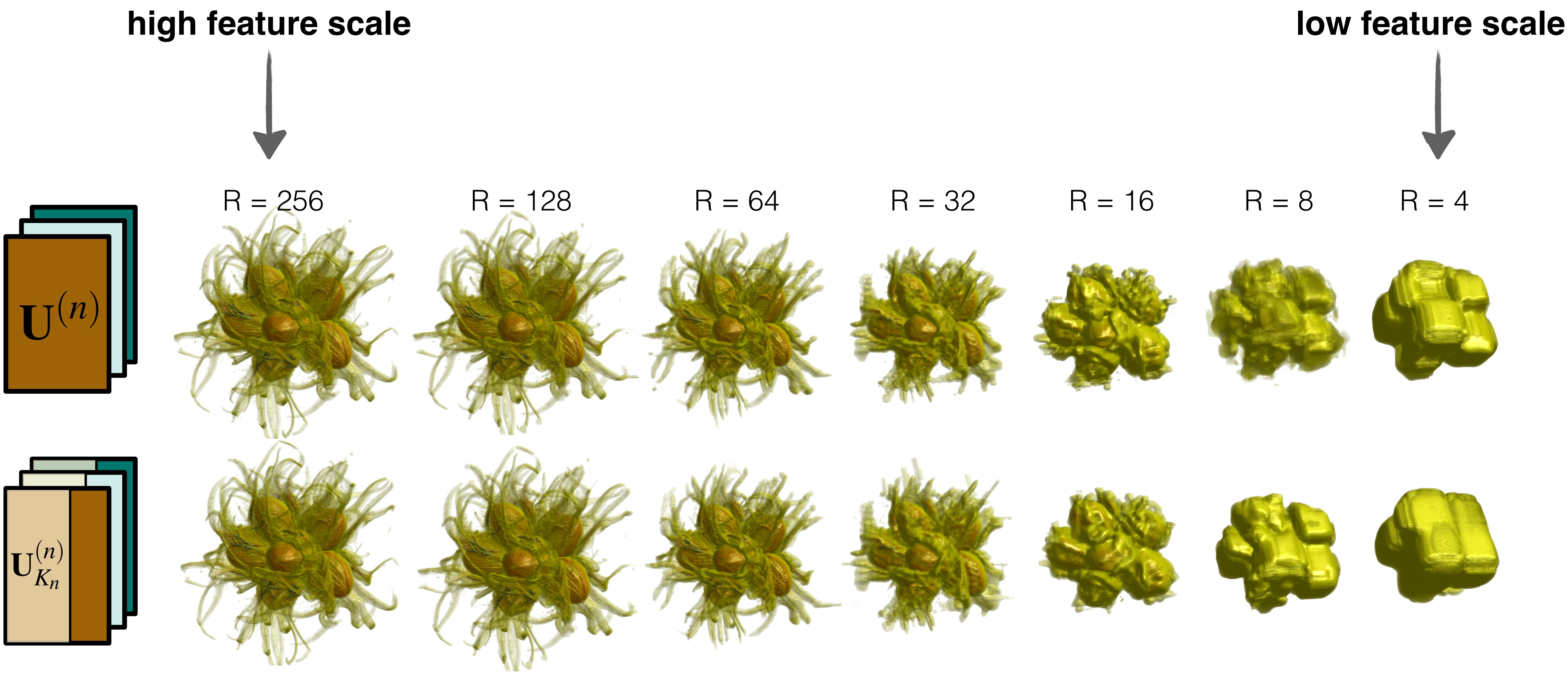
Tucker tensor decomposition

$$\widetilde{\mathcal{A}} = \mathcal{B} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)}$$

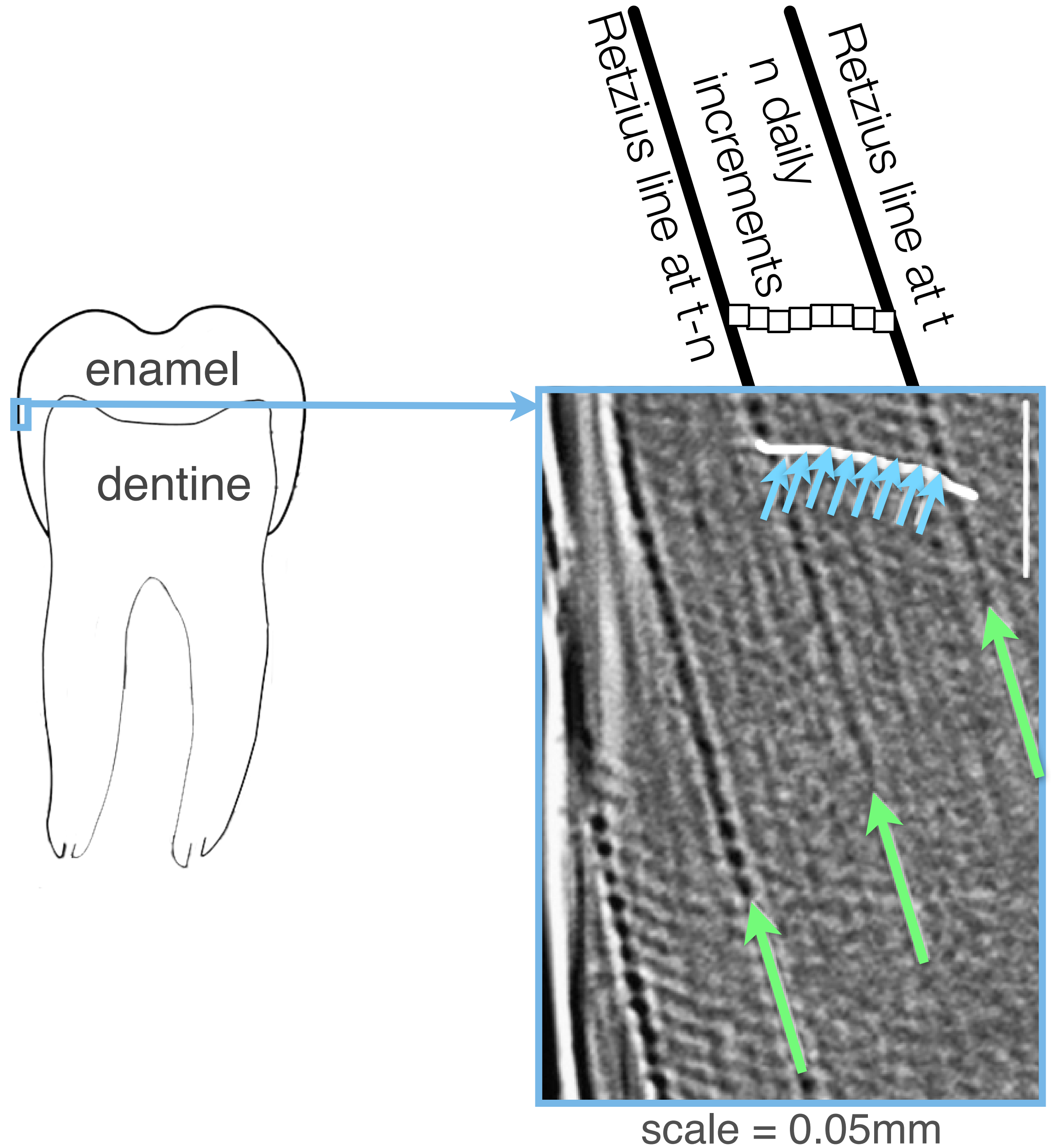
I_n input tensor length
R_n multilinear rank
where $R_n < I_n$

VISUALIZATIONAND MULTIMEDIALAB

Example: Rank Truncation



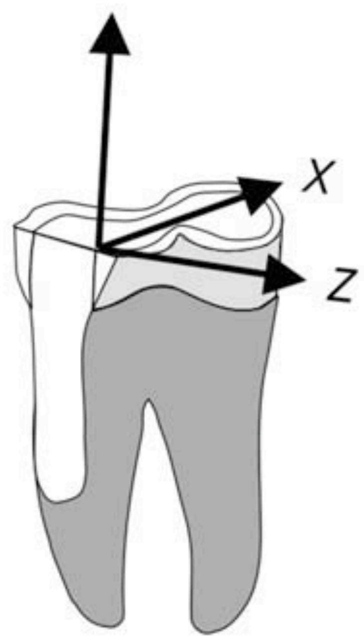
Application: Multiscale Dental Growth Pattern

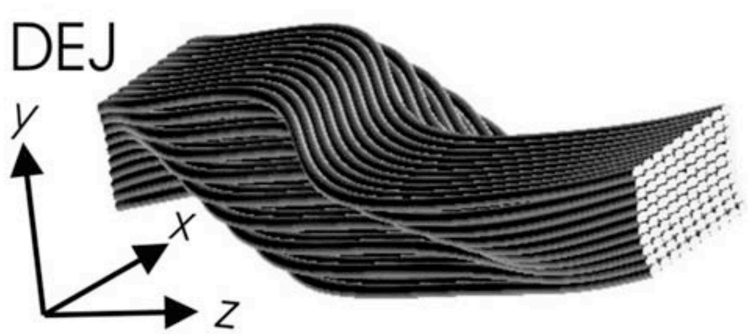


Suter, Zollikofer and Pajarola. Application of tensor approximation to multiscale volume feature representations. In *Proceedings Vision, Modeling and Visualization*, pages 203–210, 2010.

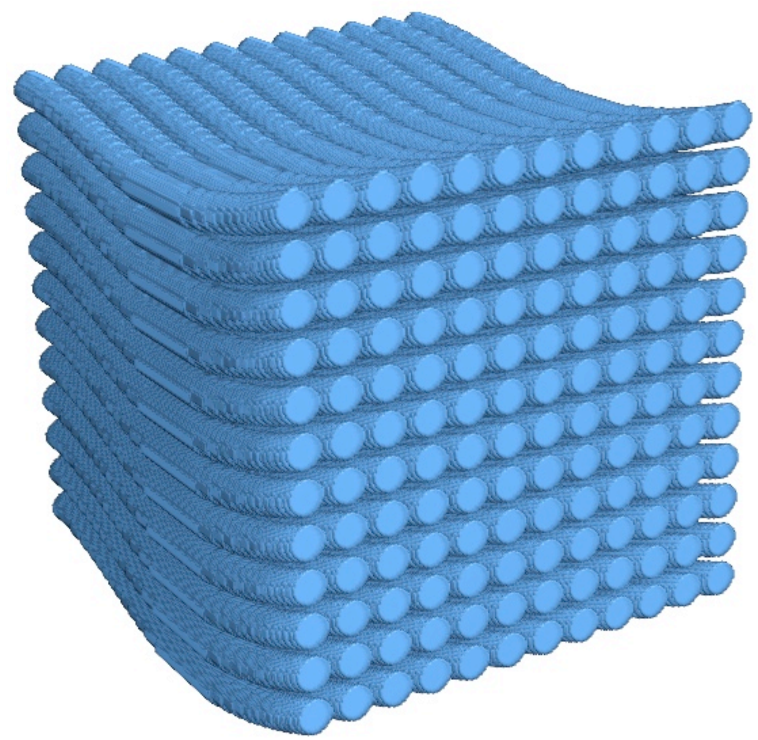
Synthetic 3D Dental Growth Structures

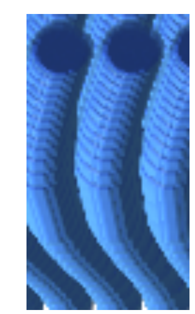
[Suter et al., 2010]





Source: Macho et al., 2003

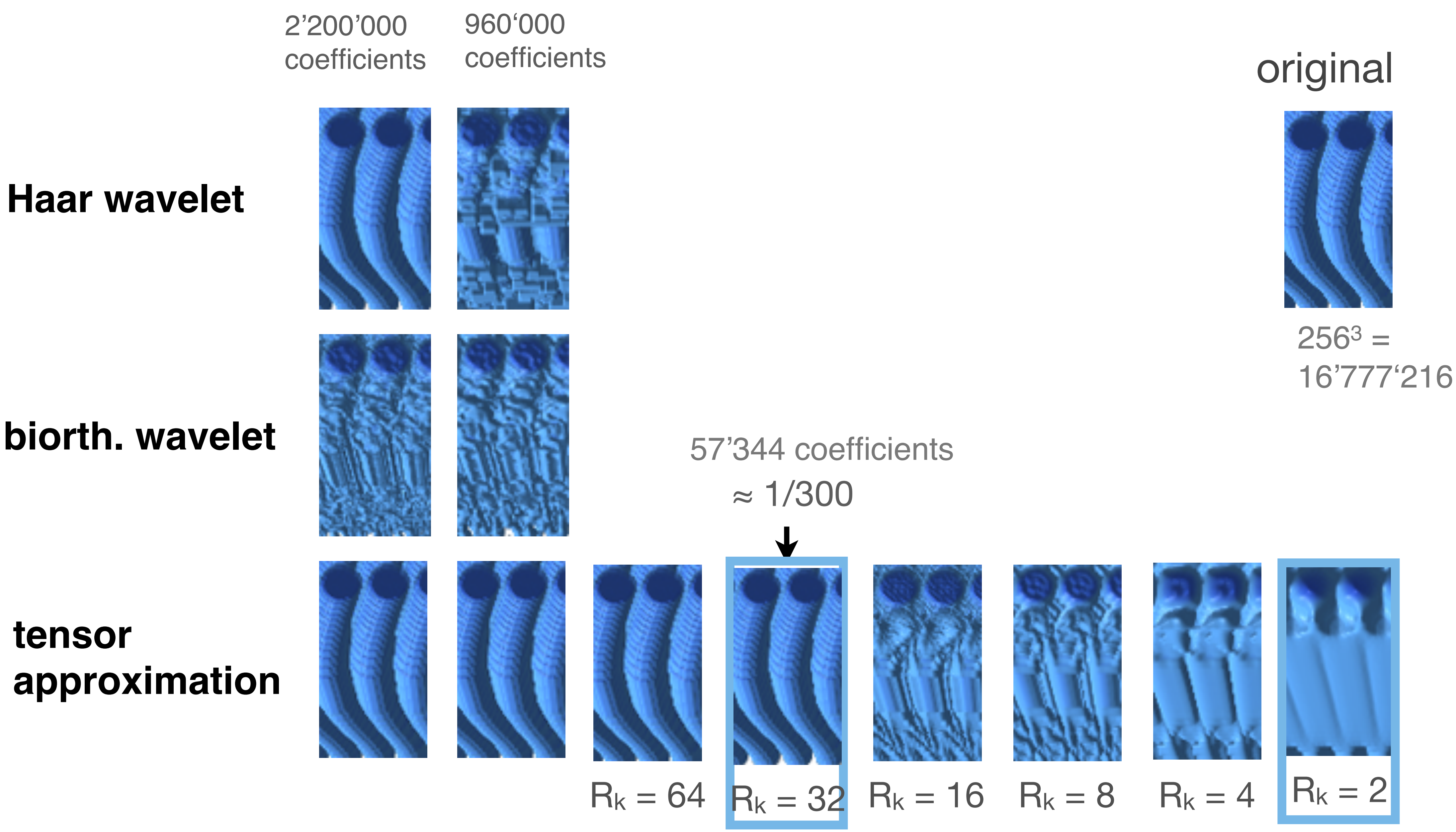




VISUALIZATIONAND MULTIMEDIALAB

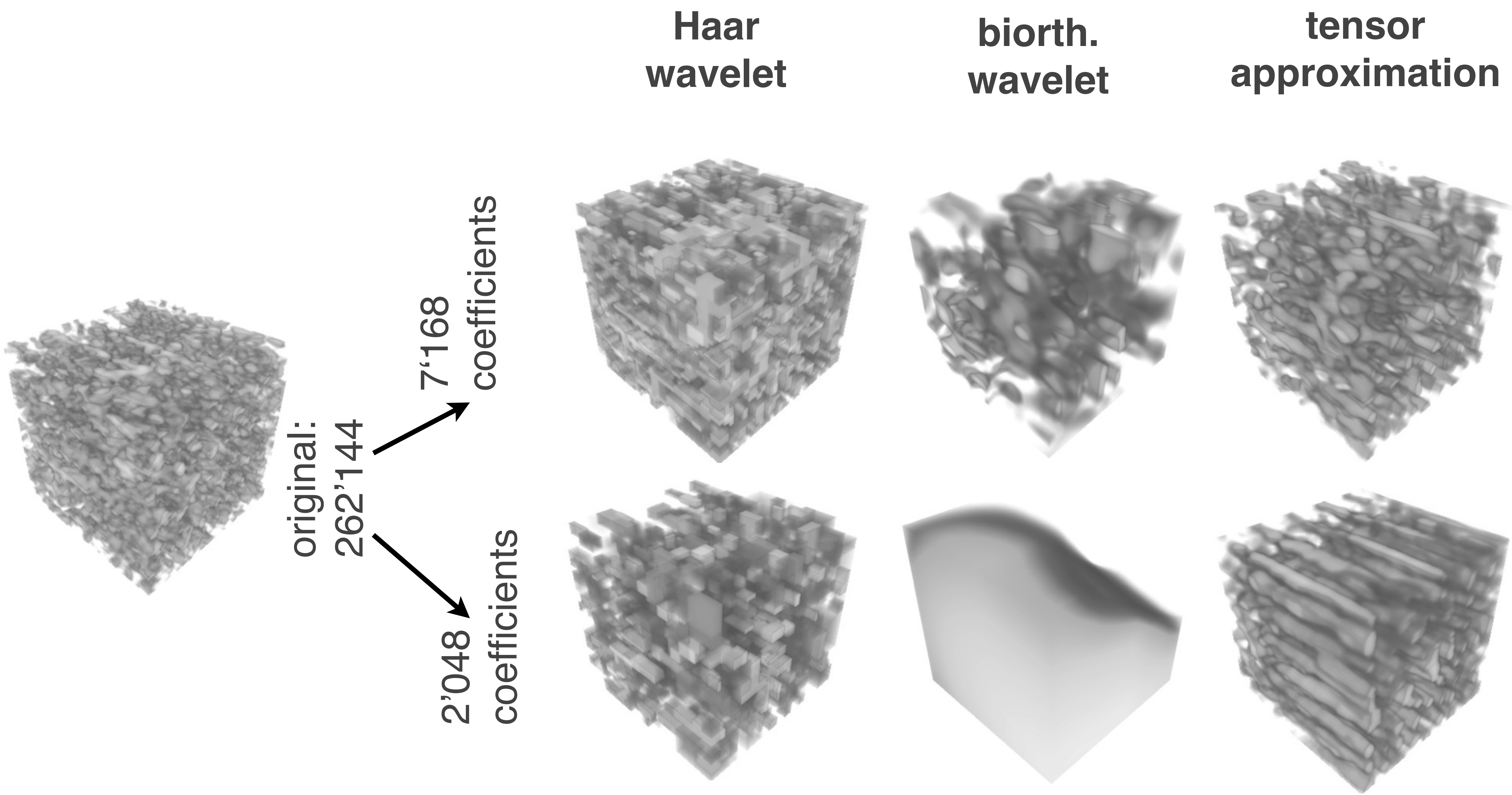
Non-axis-aligned Synthetic Features

[Suter et al., 2010]



Real Dental Growth Structures

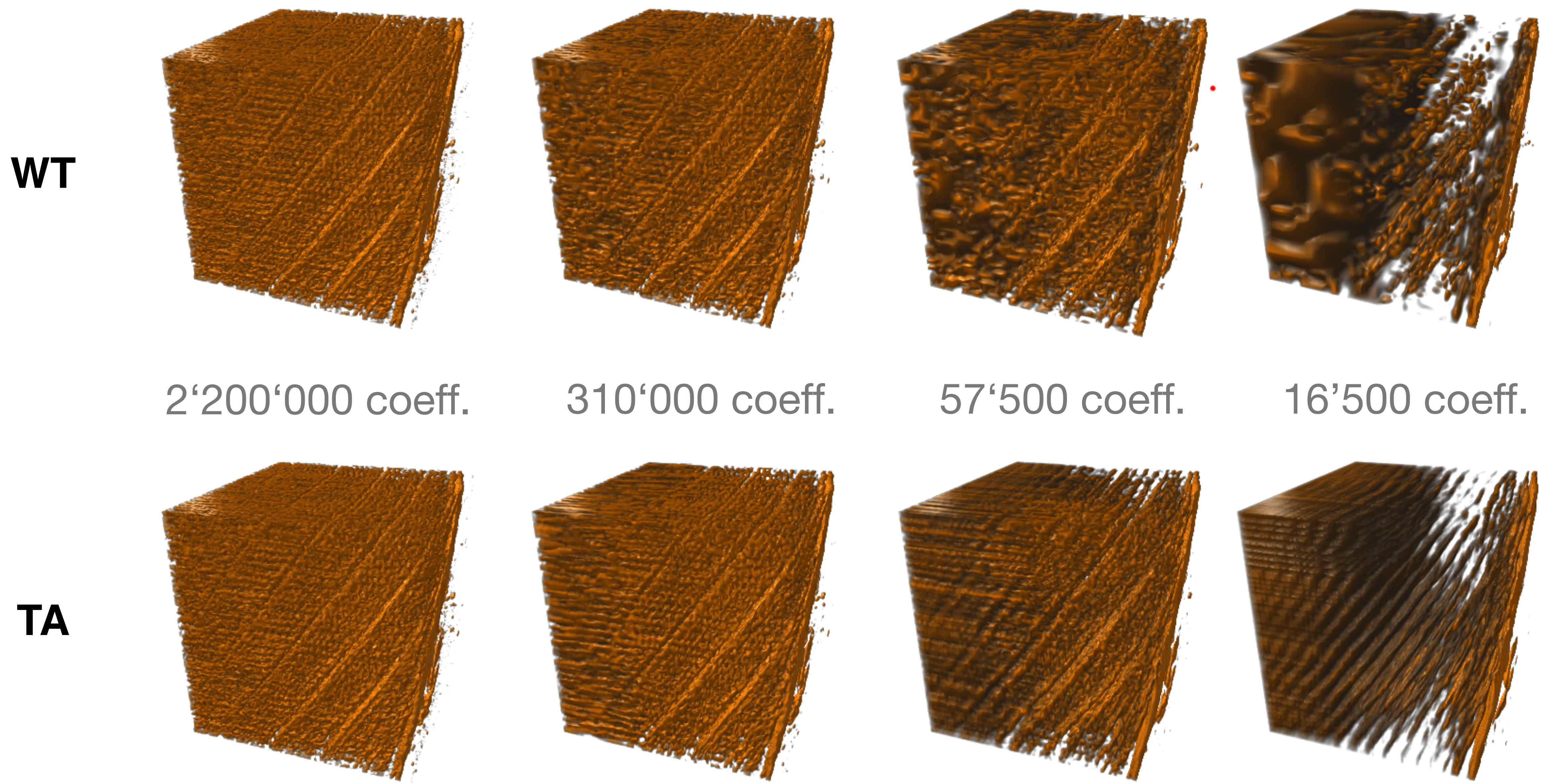
[Suter et al., 2010]



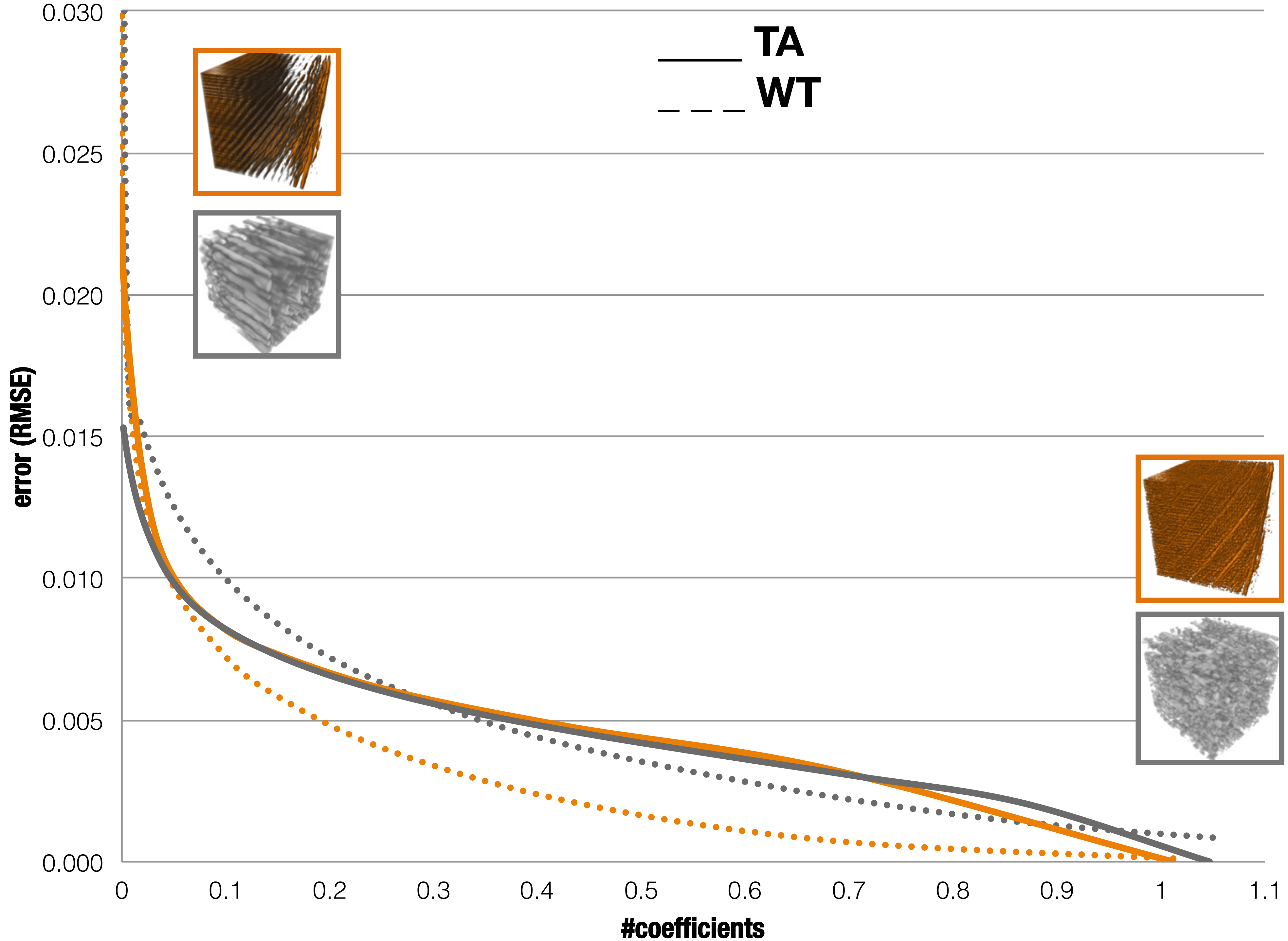
Real Multiscale Dental Growth Structures

[Suter et al., 2010]

original size:
$256^3 = 16'777'216$



Reconstruction “Error” vs. Compression



WT vs. TA

Wavelet Transform (WT)

- Recursive decomposition at each scale into coarser resolutions
- Traditional multiresolution:
  - projects signal at different resolutions onto a prescribed bases without knowledge on data
- Axis-aligned data reduction

Tensor Approximation (TA)

- Bases are adopted for a given dataset
  - search for major direction/ variability within dataset
- Higher quality images at high data reduction ratios
- Goal: lossy, but keep features
  - analyze and count

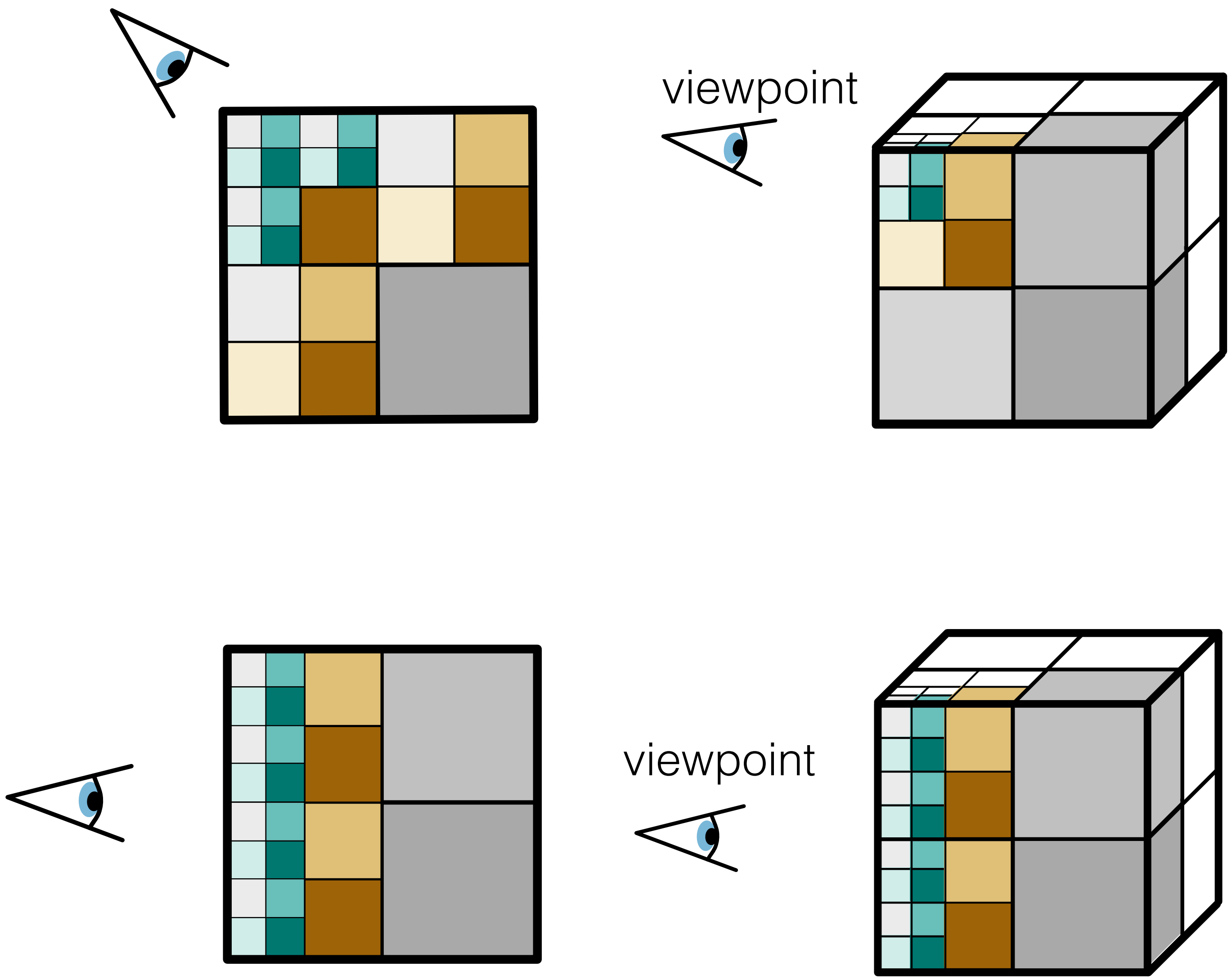
Part 2: Multiresolution TA Hierarchies

University of Zurich[UZH]

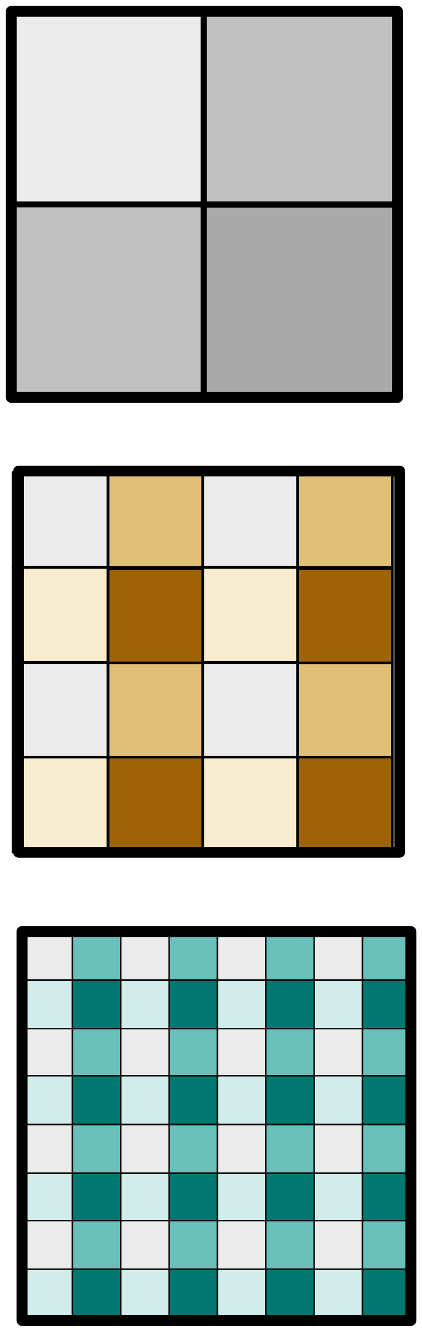
VISUALIZATIONANDMULTIMEDIALAB

Multiresolution Analysis

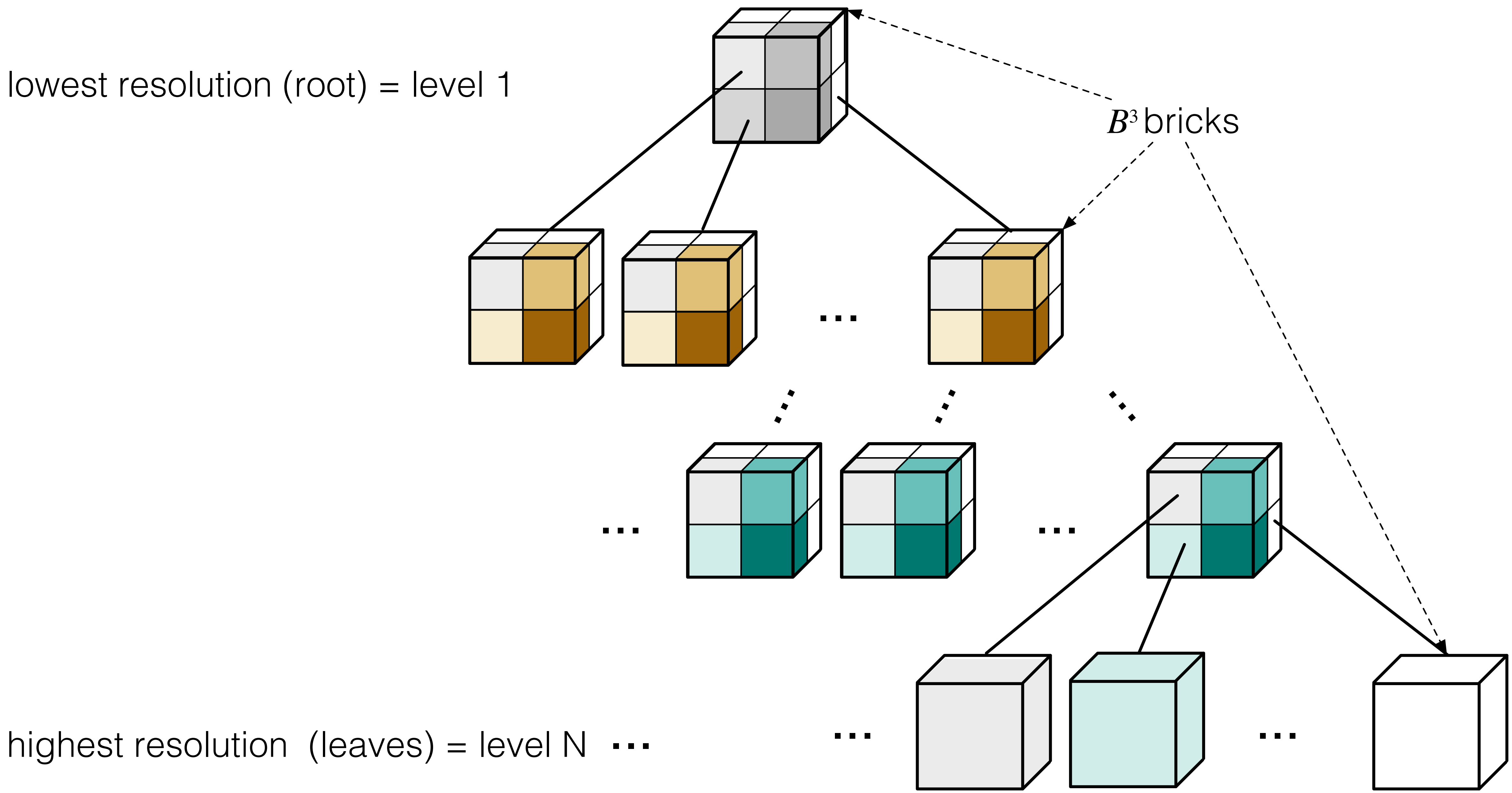
multiresolution visualization



progressive image refinement



Multiresolution Tree Data Structure



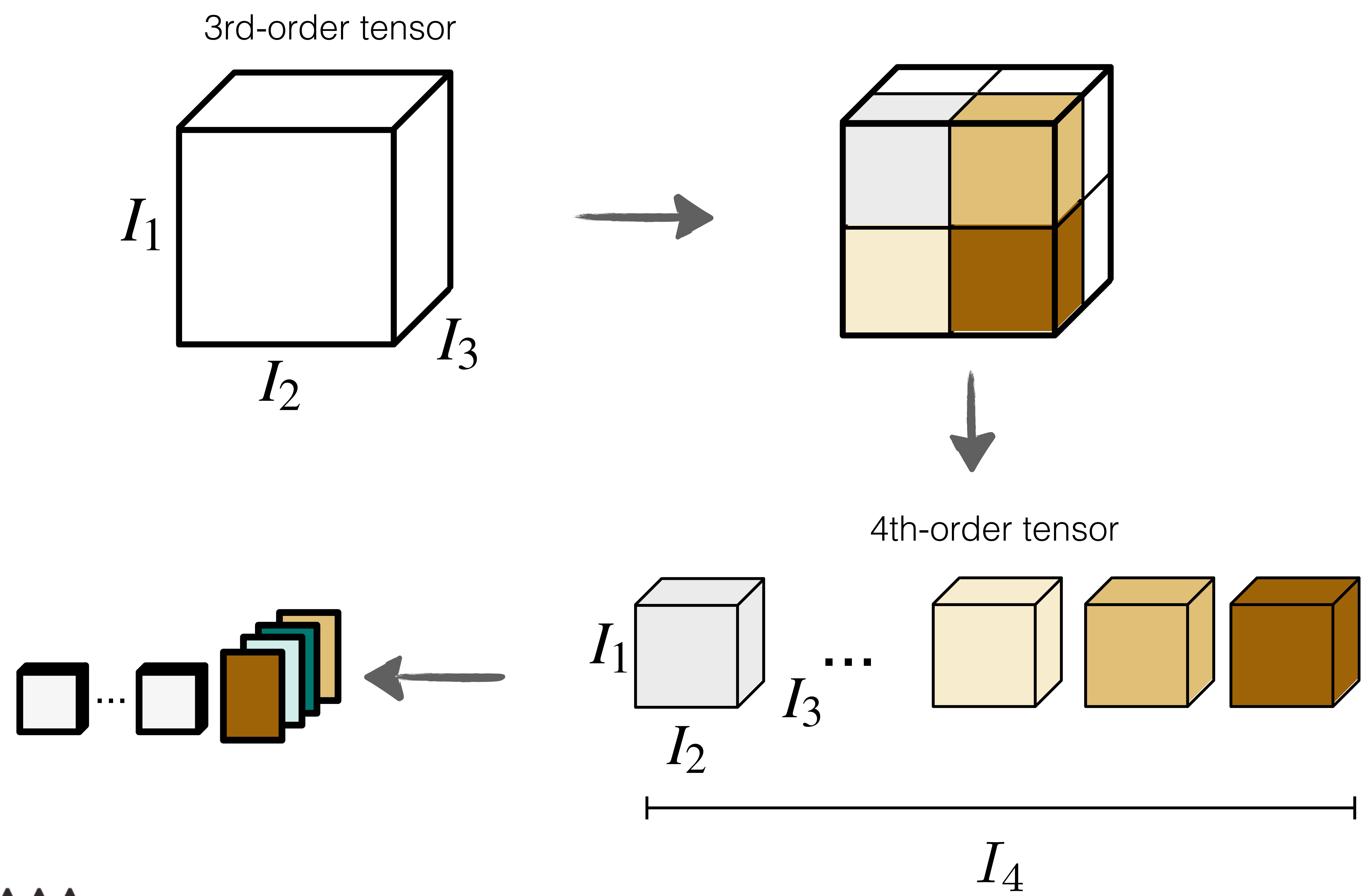
Hierarchical Tucker Model

- Multiresolution analysis
  - significant components at low frequencies
  - less important components at high frequencies
  - high-frequency components have smaller spatial support
  - thus, high-frequency components receive shorter basis vectors
- Why?
  - exploit more redundancy
  - receive smoother borders

Wu et al.. Hierarchical tensor approximation of multidimensional visual data. *IEEE Transactions on Visualization and Computer Graphics* 14(1):186-199, January/February 2008.

Tensor Ensembles

[Wu et al., 2008]



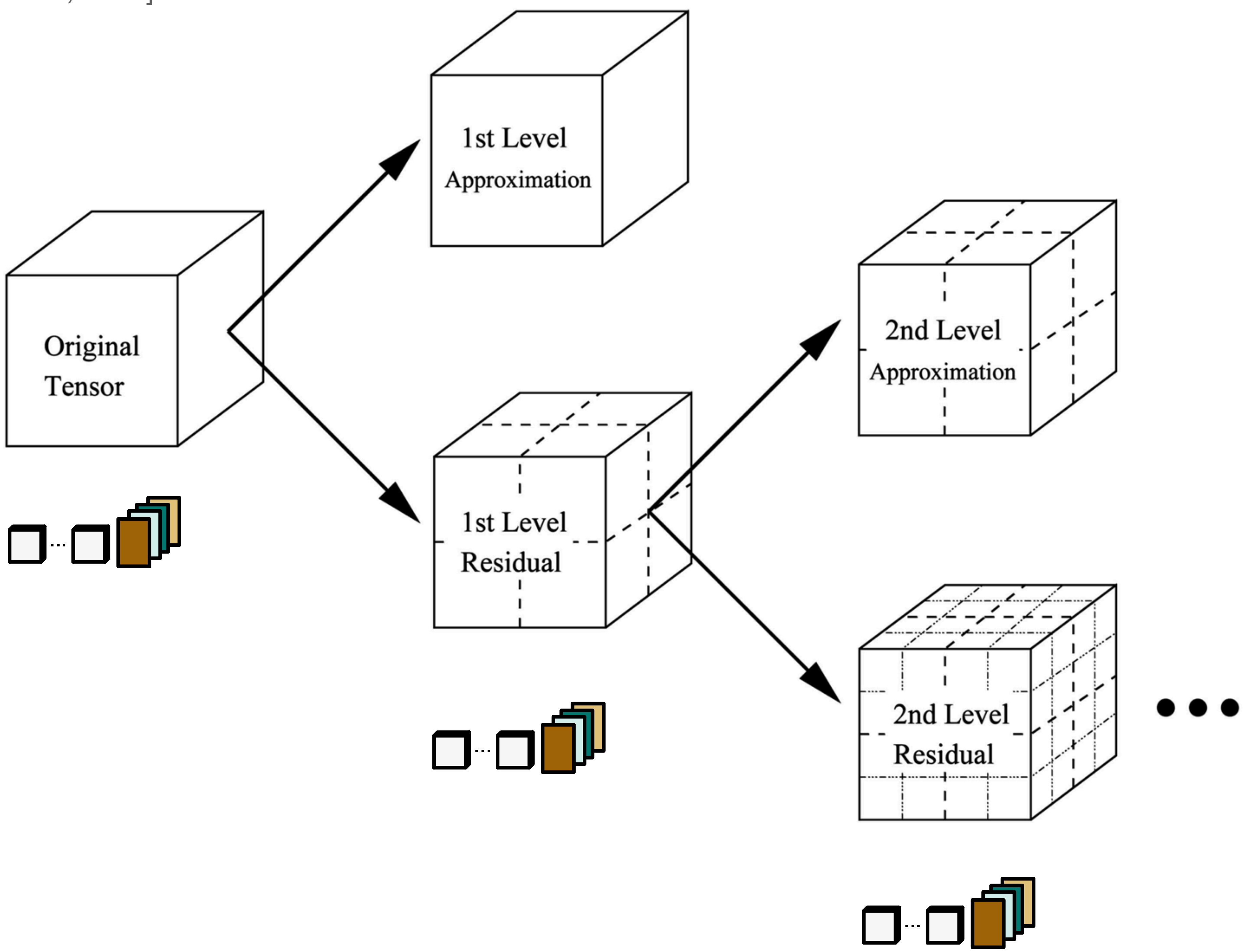
Tensor Ensemble Ranks

[Wu et al., 2008]

- Multilinear rank $(R_1, R_2, \ldots R_N)$ defined per hierarchy:
  - start with $R_n = I_n / 8$ or $I_n / 16$
  - each next hierarchy rank R_n is half of the rank of the previous hierarchy level
  - for example: 32, 16, 8, 4, 2

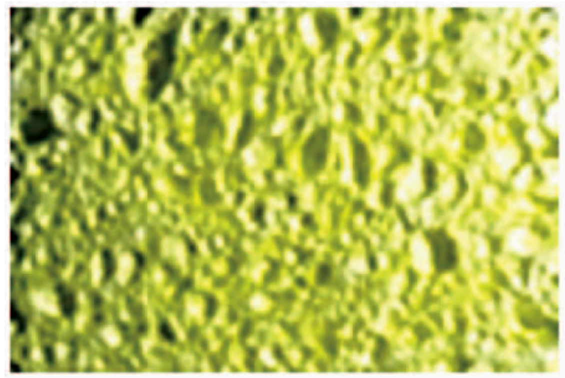
Residual-based Hierarchy

[Wu et al., 2008]



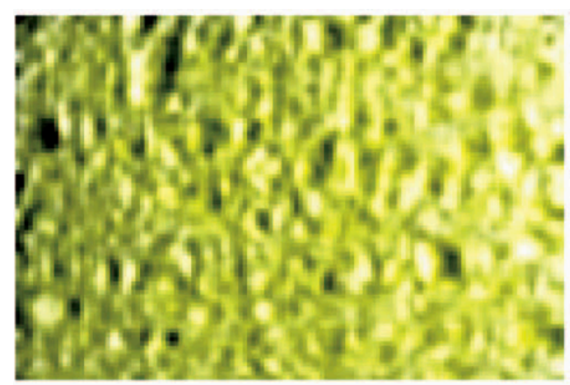
Hierarchical TA and WT on a BTF

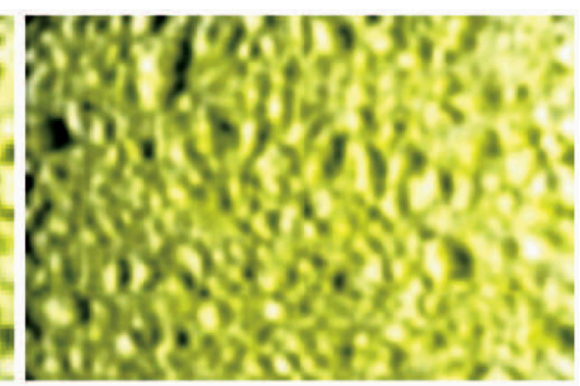
[Wu et al., 2008]



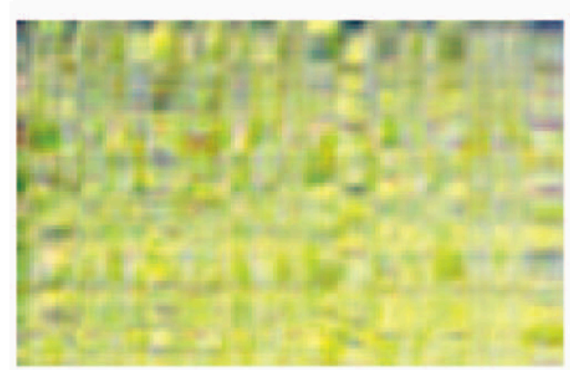
original: sponge BTF
- 45 views
- 60 illuminations
- image: 128x128

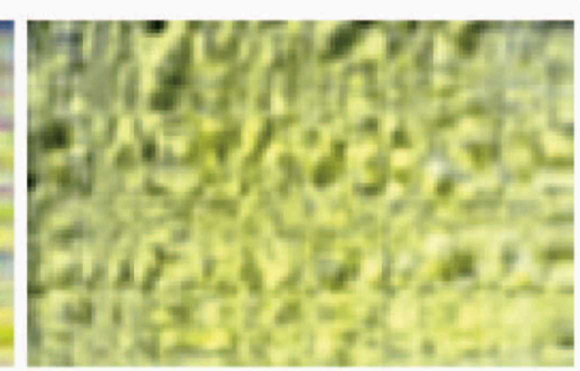
biorth. wavelet | **wavelet packet** | **single level TA** [Wang et al., 2005] | **multi level TA**





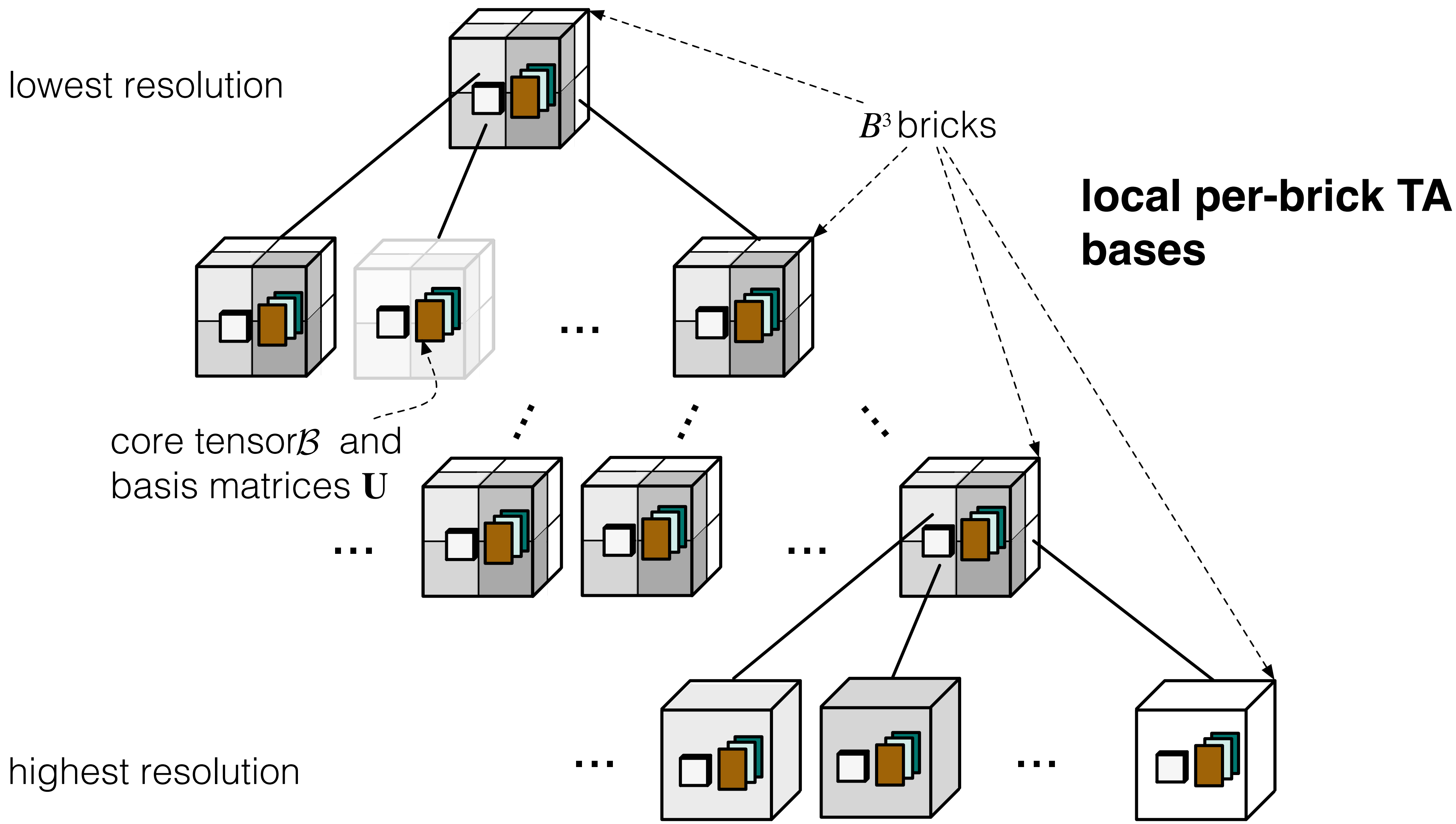
compression ratio: 55





compression ratio: 3,922

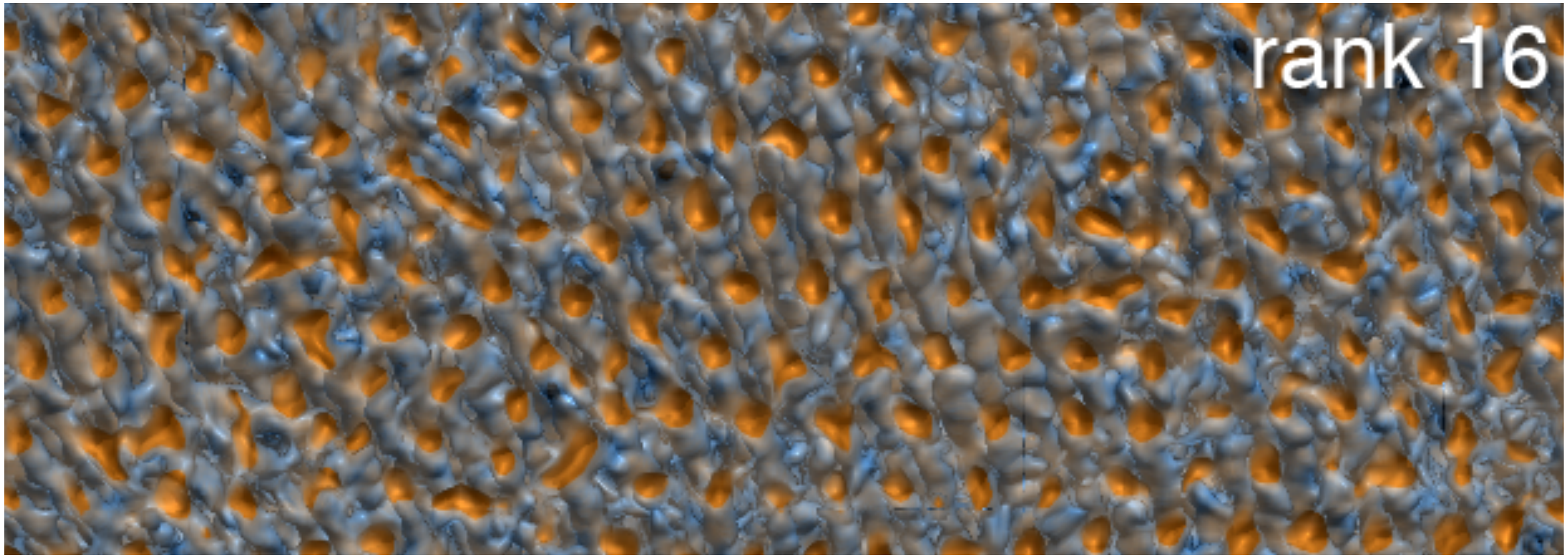
Multiresolution Direct Volume Rendering

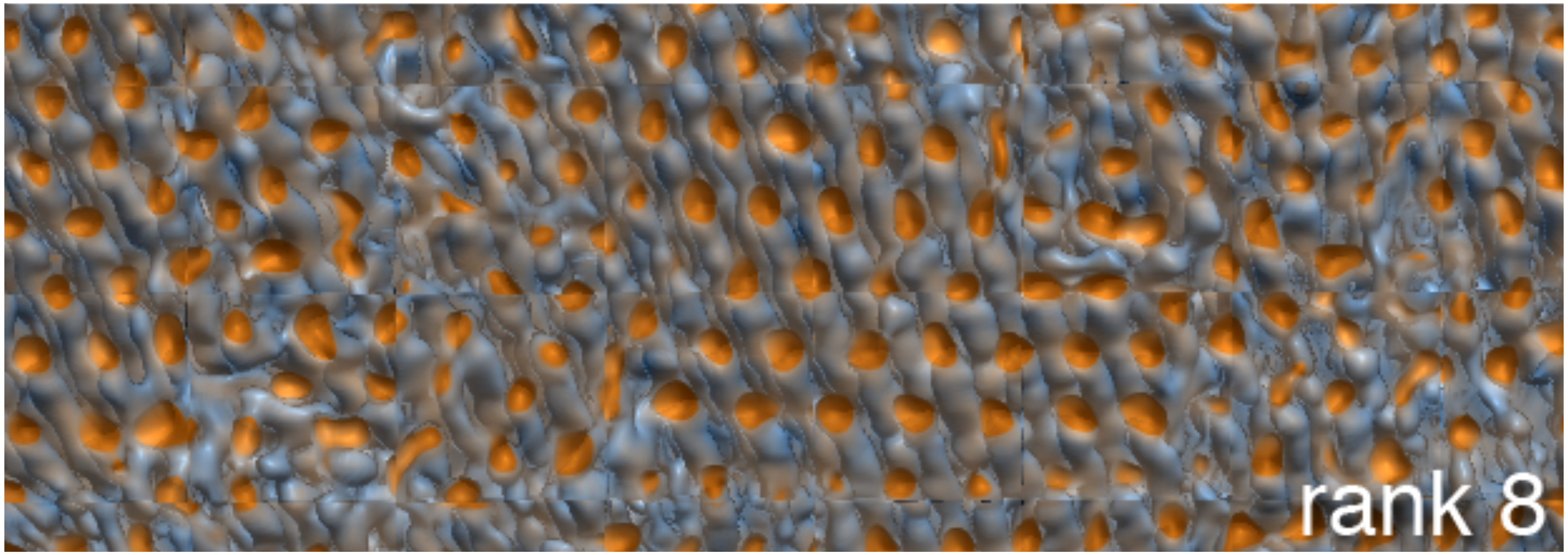


Suter et al.. Interactive multiscale tensor reconstruction for multiresolution volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2135–2143, December 2011.

Rank-reducibility and Feature Extraction

[Suter et al., 2011]

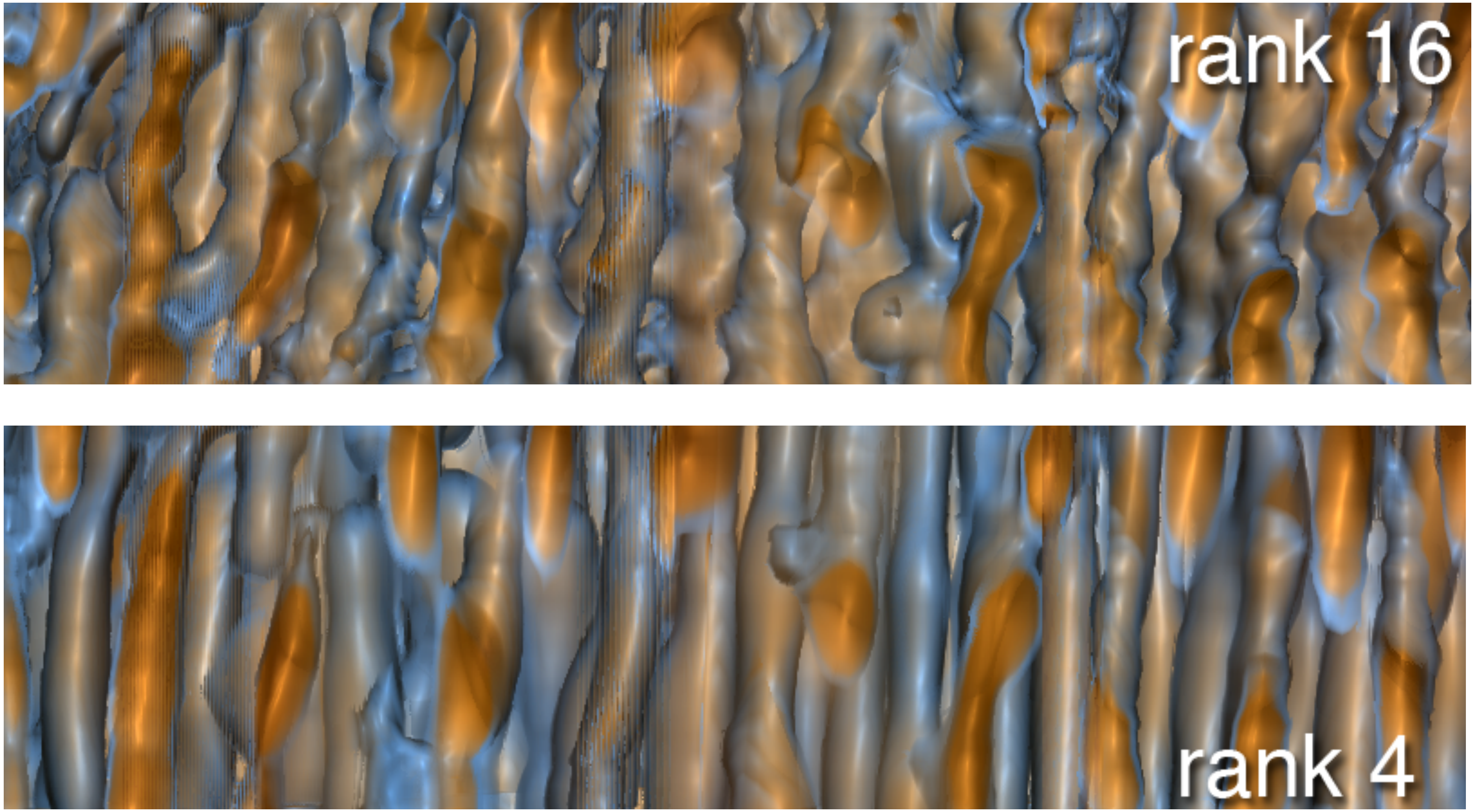




VISUALIZATIONAND
MULTIMEDIALAB

Rank-reducibility and Feature Extraction

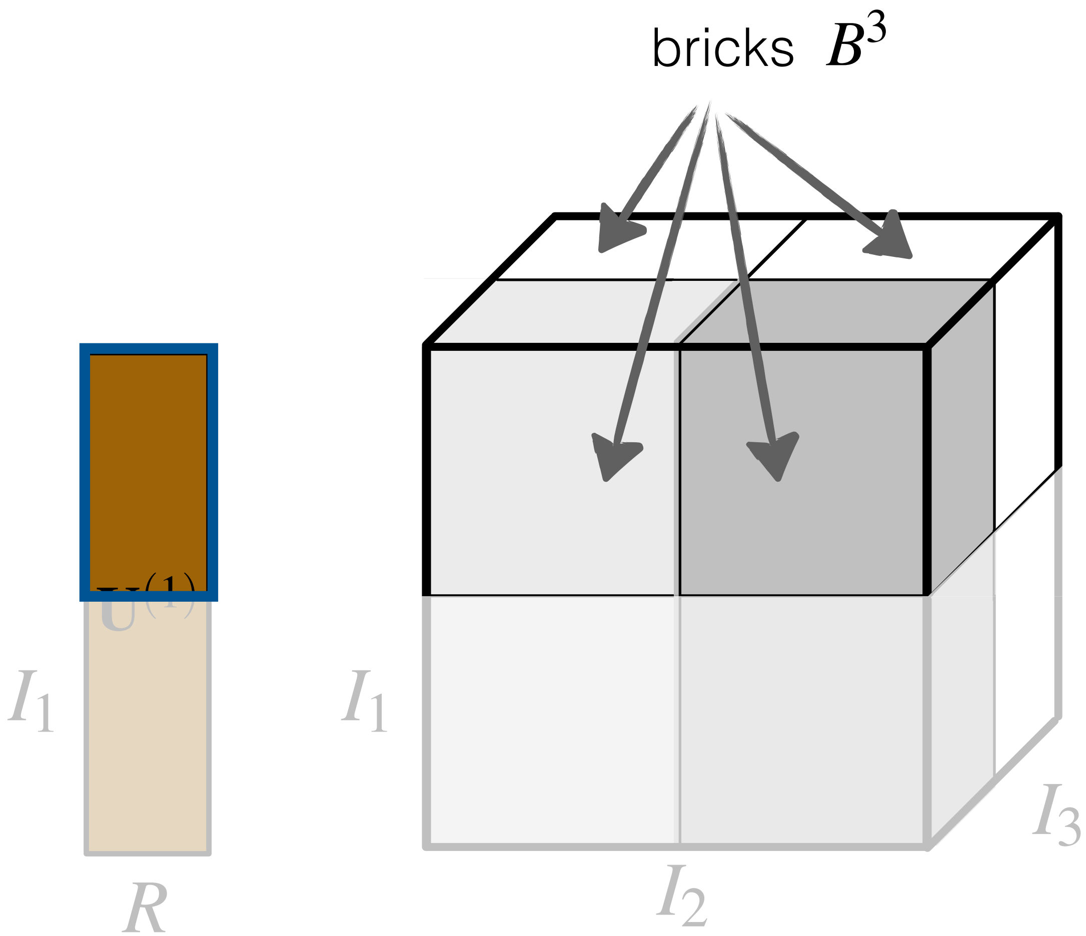
[Suter et al., 2011]



Why Improve Existing Model?

- Flexible rendering based on user choice of a feature scale
- Improve brick-border artifacts
- Exploit redundancies
- Reduce storage costs

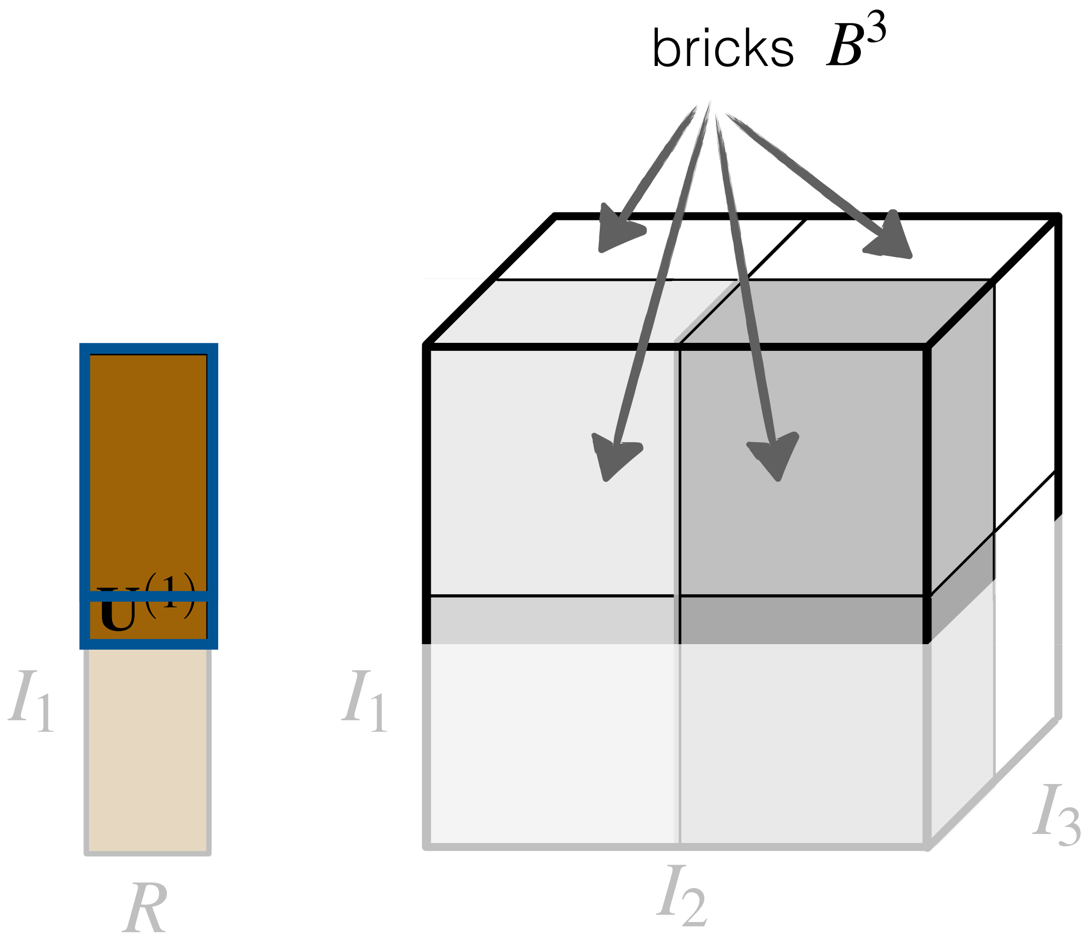
Motivation for Global Bases



$$I_1 = I_2 = I_3 = 2 \cdot B$$
$$B = 64$$
$$R = 32$$

- Exploit redundancies among bricks

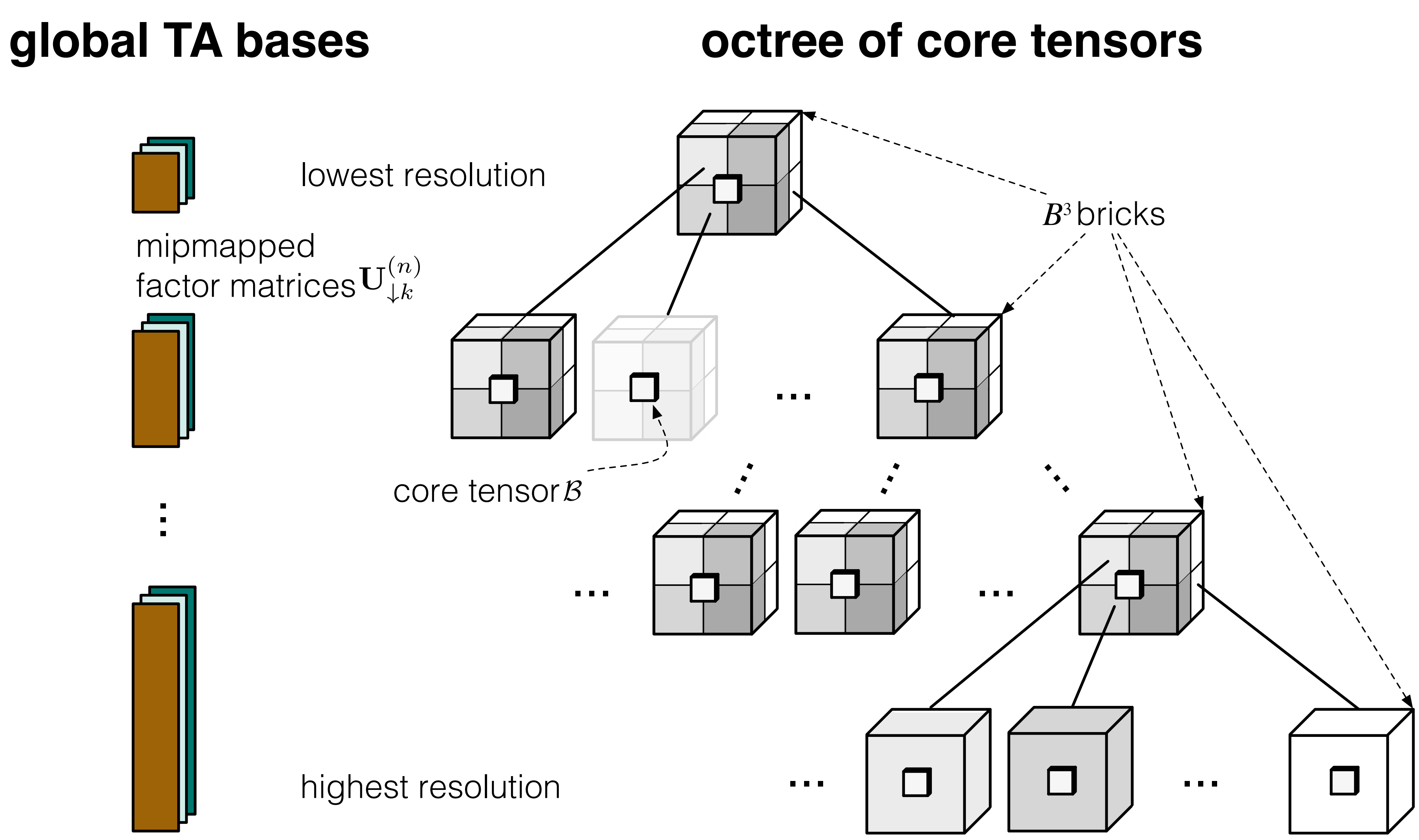
Motivation for Global Bases



$$I_1 = I_2 = I_3 = 2 \cdot B$$
$$B = 64$$
$$R = 32$$

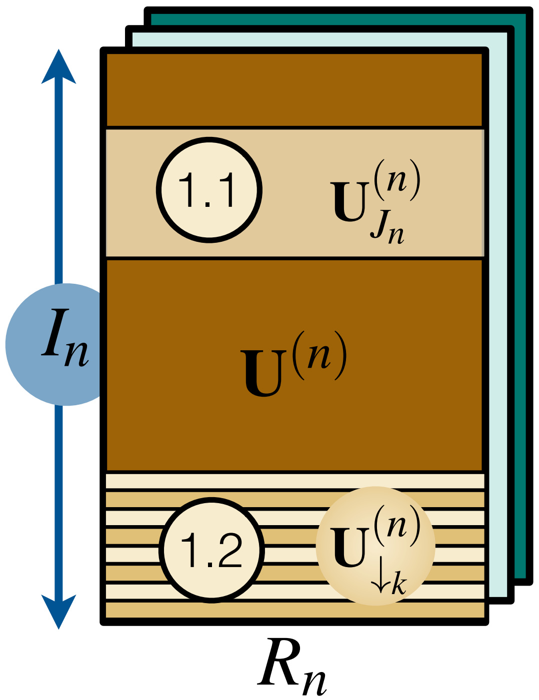
- Use knowledge about neighboring bricks (brick-border artifacts)

TA Hierarchy with Global Bases

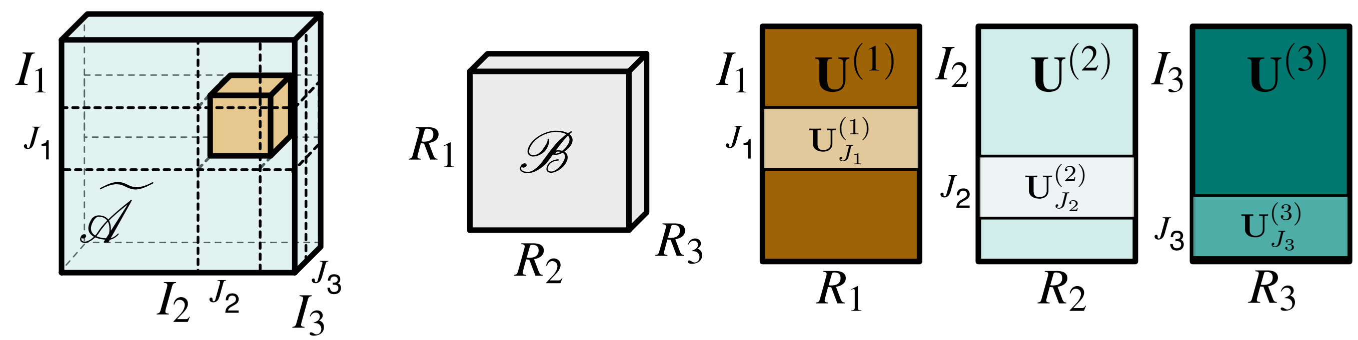


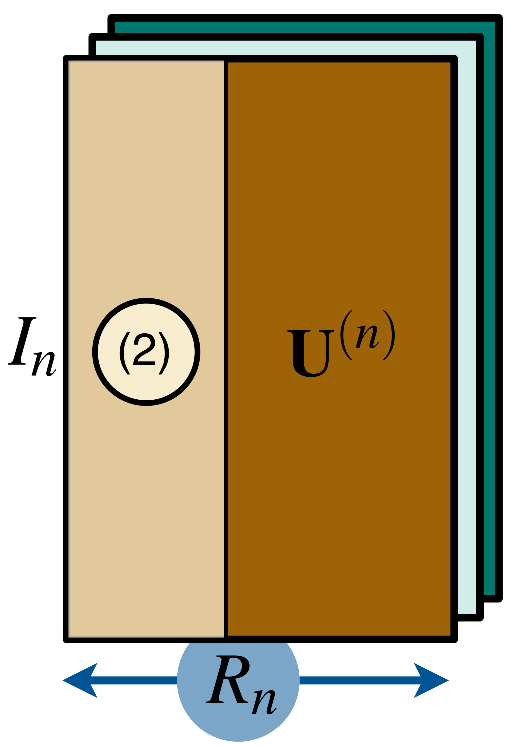
Suter, Makinya, Pajarola, EuroVis 2013. TAMRESH - Tensor Approximation Multiresolution Hierarchy for Interactive Volume Visualization.

TA Bases Properties

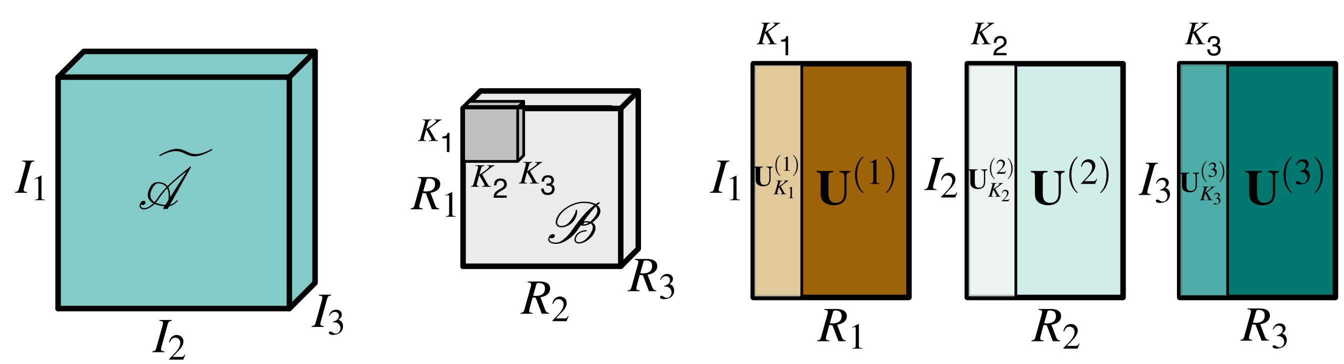


1.1 spatial selection
1.2 spatial subsampling



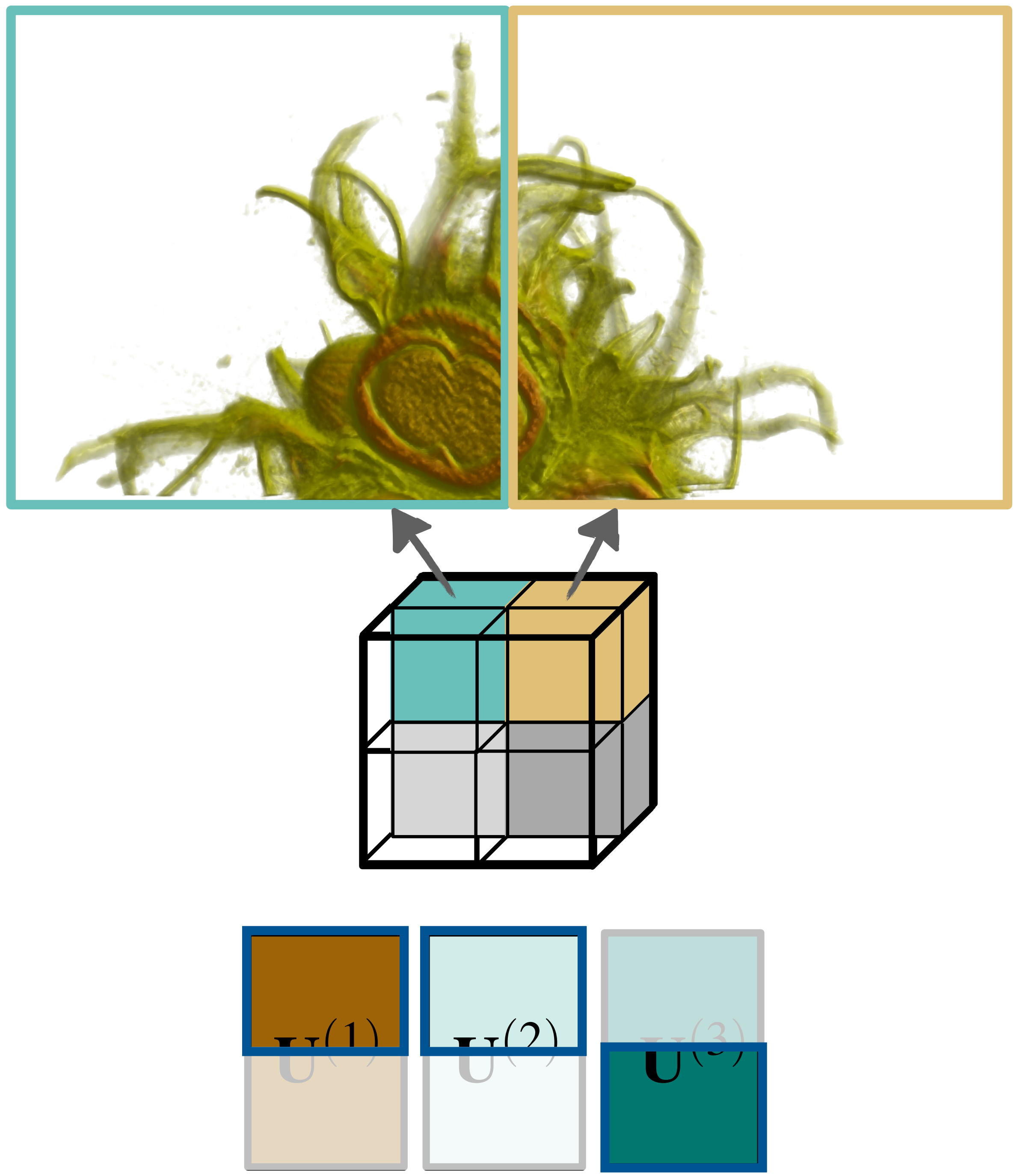


2 tensor rank truncation



VISUALIZATIONAND MULTIMEDIALAB

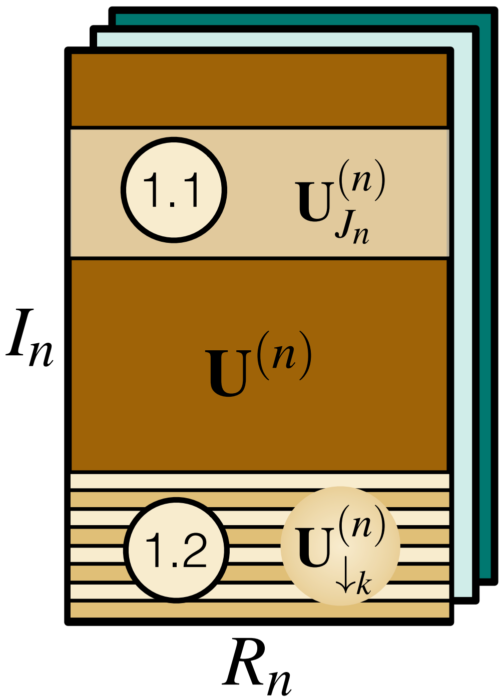
Spatial Selection



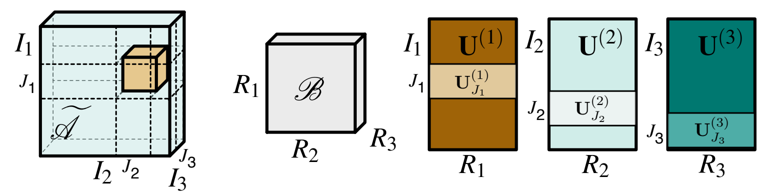
Spatial Subsampling

VISUALIZATIONAND MULTIMEDIALAB

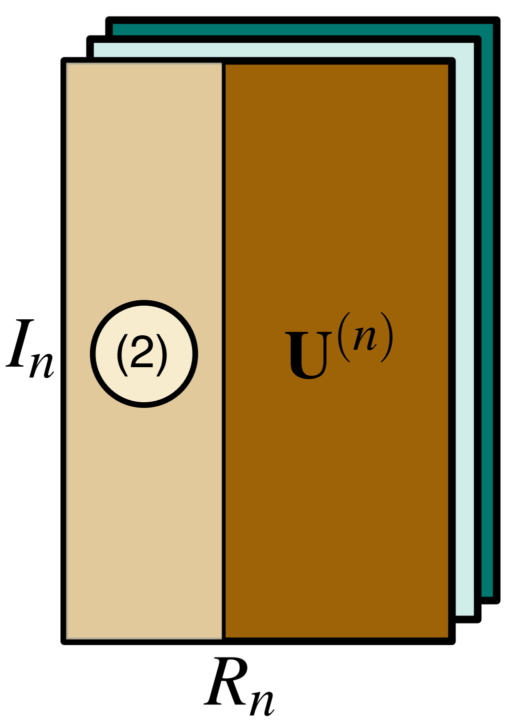
TA Bases Properties



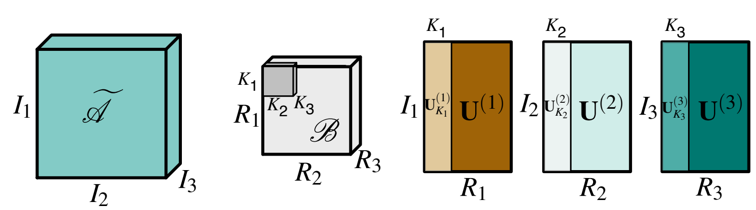
1.1 spatial selection
1.2 spatial subsampling



➡ use for state-of-the-art **multiresolution** volume rendering



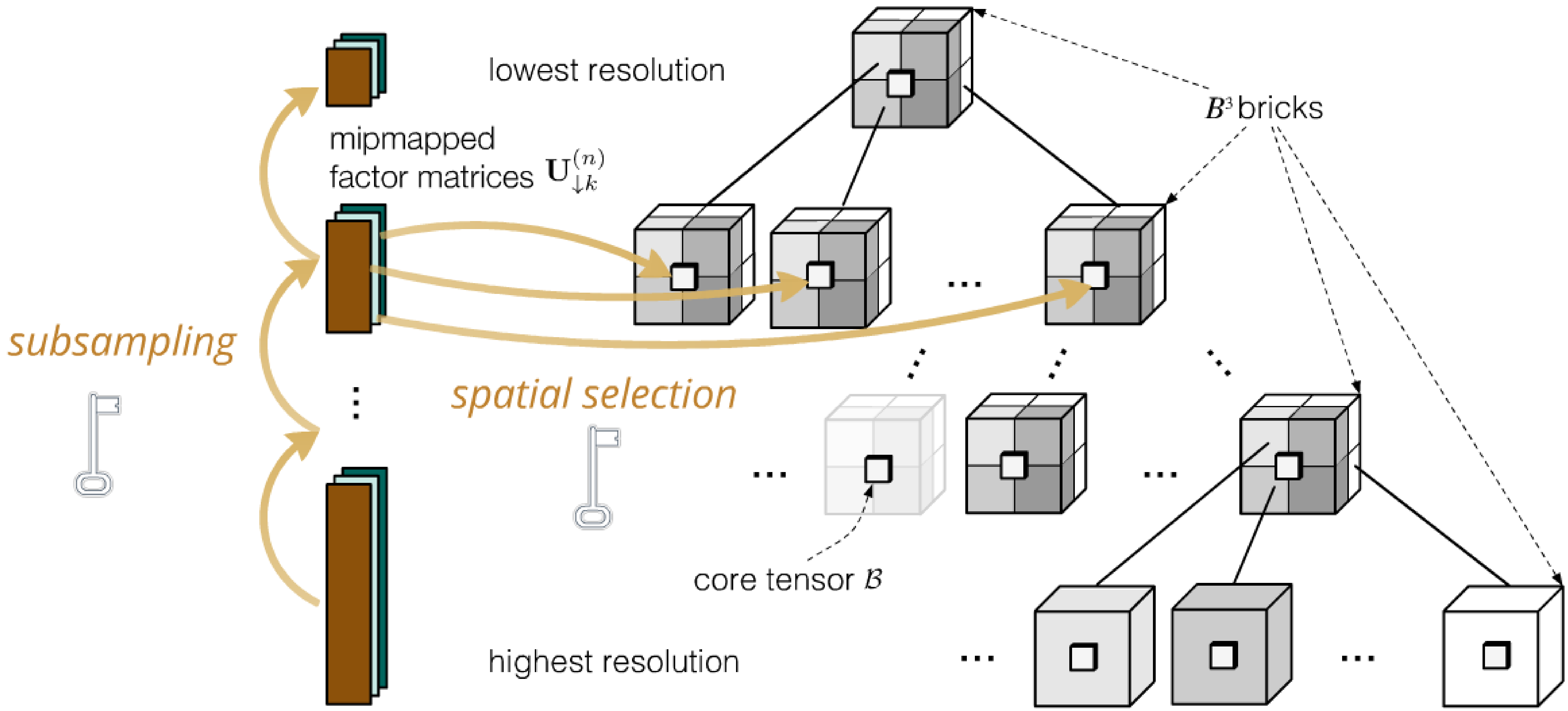
2 tensor rank truncation



➡ use for **multiscale** volume feature visualization

VISUALIZATIONAND MULTIMEDIALAB

Multiresolution and Multiscale DVR



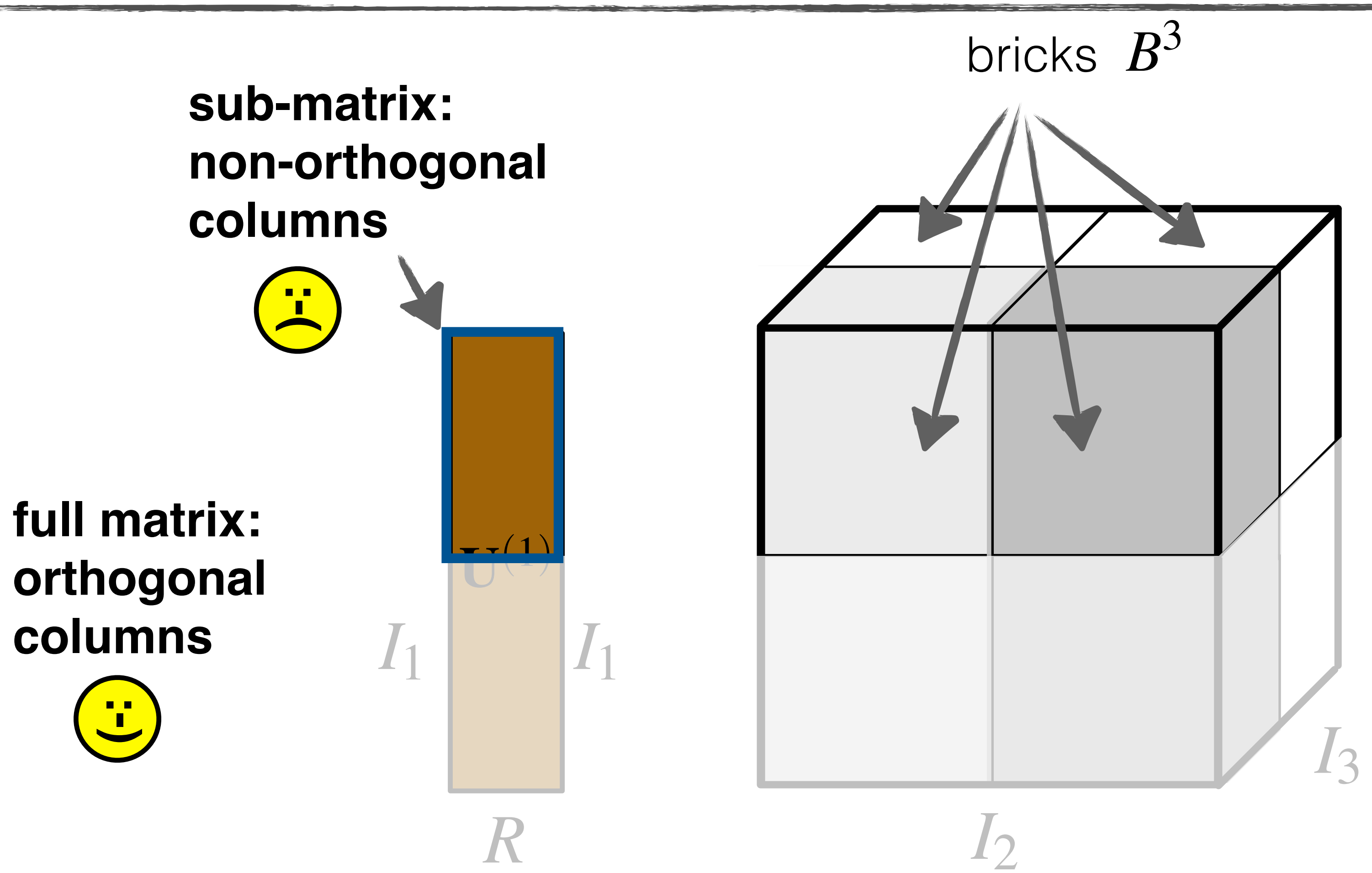
Suter, Makhinya and Pajarola. TAMRESH: Tensor approximation multiresolution hierarchy for interactive volume visualization. *Computer Graphics Forum*, 2013.

Coupling Global Bases with Rank Truncation

- Multiscalability through tensor rank truncation
  - **necessary condition**:
    produce core tensors from orthogonal TA factor matrices

VISUALIZATIONAND
MULTIMEDIALAB

Integrate Rank Truncation



$$I_1 = I_2 = I_3 = 2 \cdot B$$
$$B = 64$$
$$R = 32$$

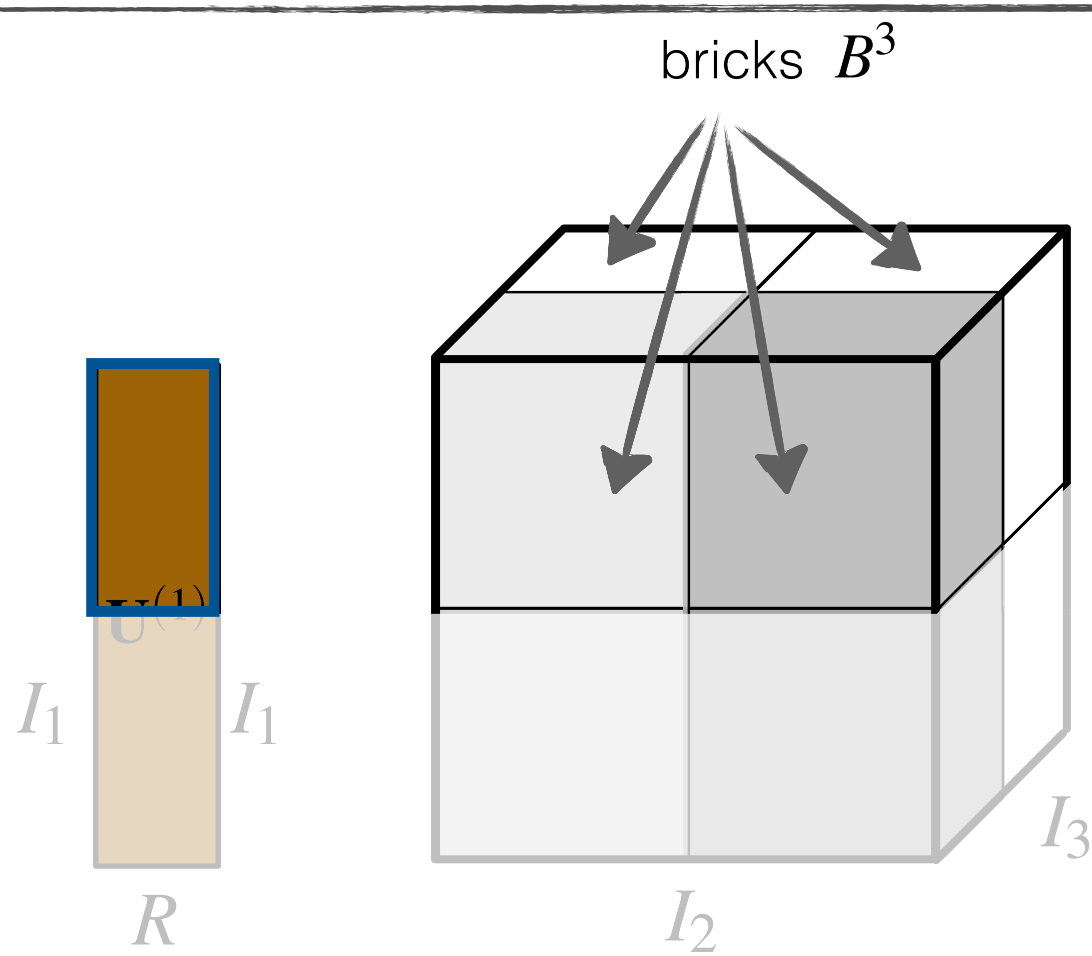
- TA-based multiresolution model uses spatial selectivity

➡ Orthogonality property in core tensors lost

➡ Rank truncation not possible

VISUALIZATIONAND MULTIMEDIALAB

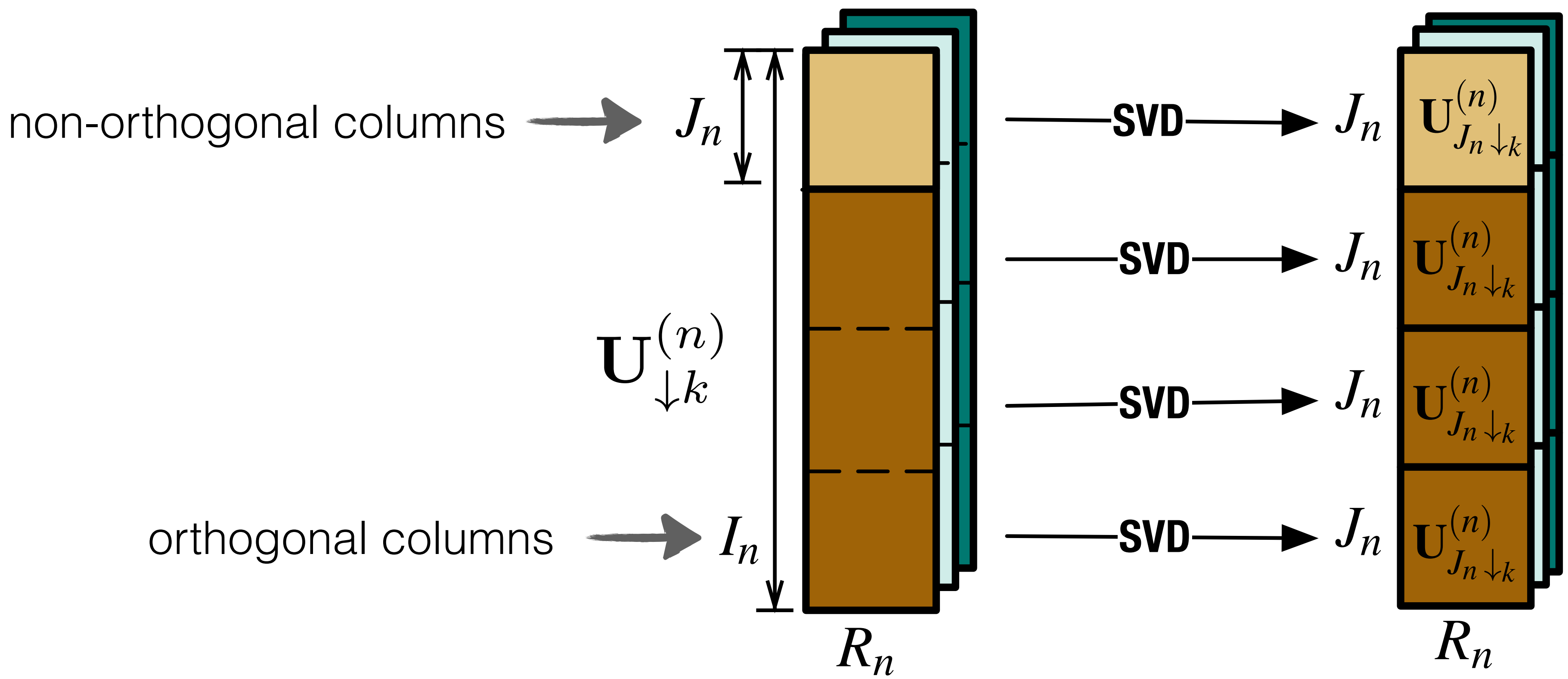
Brick-row-block Correspondence



$$I_1 = I_2 = I_3 = 2 \cdot B$$
$$B = 64$$
$$R = 32$$

- One row-block corresponds to bricks of same spatial position

Maintaining All-orthogonal Core Tensors



- Row-block SVDs on TA factor matrices
- Adapted from Tsai & Shih. K-clustered tensor approximation: A sparse multilinear model for real-time rendering. ACM Transactions on Graphics, 2012.

Feature Scale Parameter

- Feature scale parameter modeled through tensor **rank truncation**
- Different approximation errors
  - truncated reconstructions vs. original
- Error in terms of RMSE
  - per brick
- Preprocessing routine

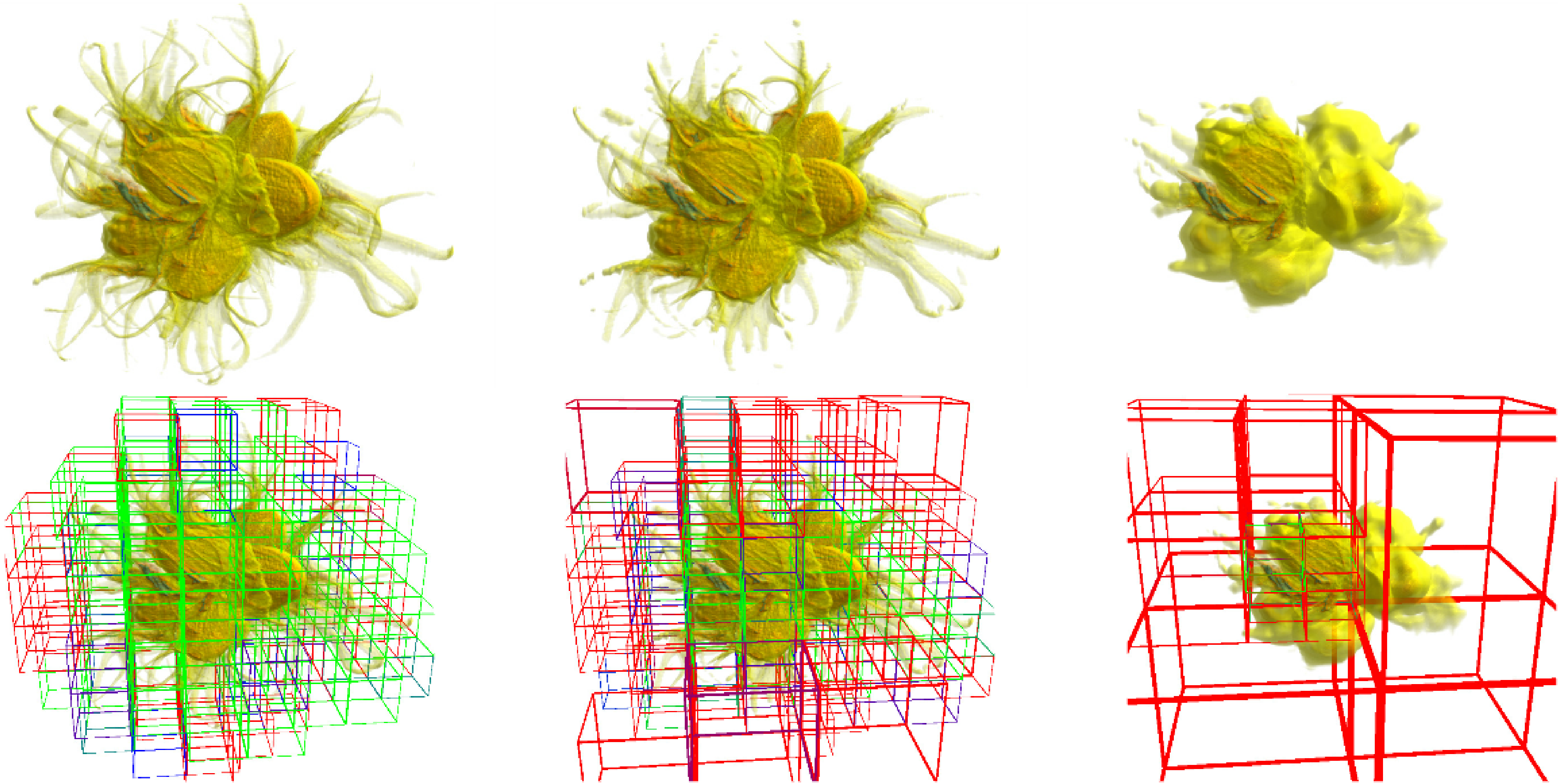
VISUALIZATIONAND MULTIMEDIALAB

Multiscale and Multiresolution

[Suter et al., 2013]

high feature scale *many details*

low feature scale *basic shape*



resolution
(subcube size)

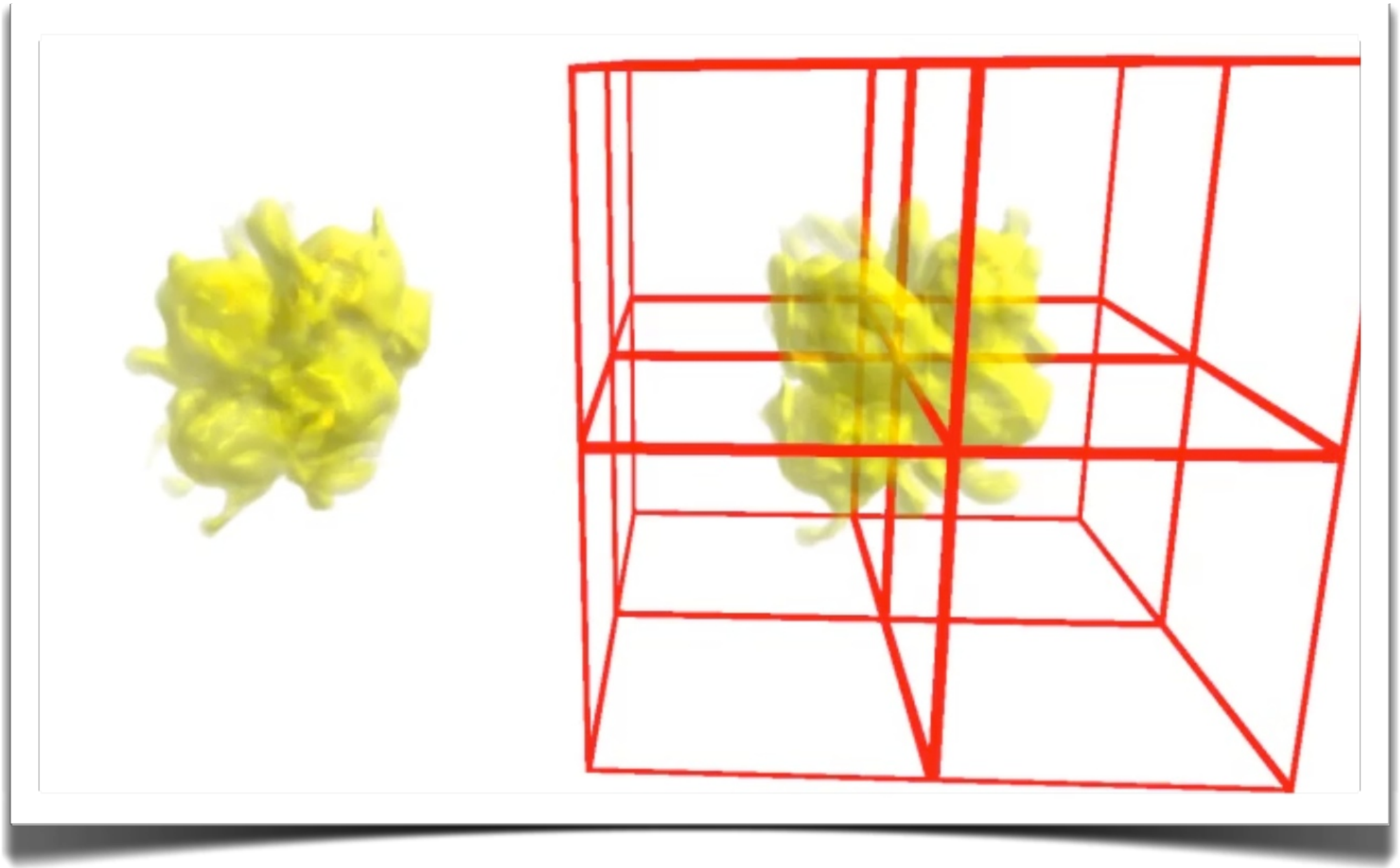
scale
(rank)

8 32

Link Feature Scale to Rendering

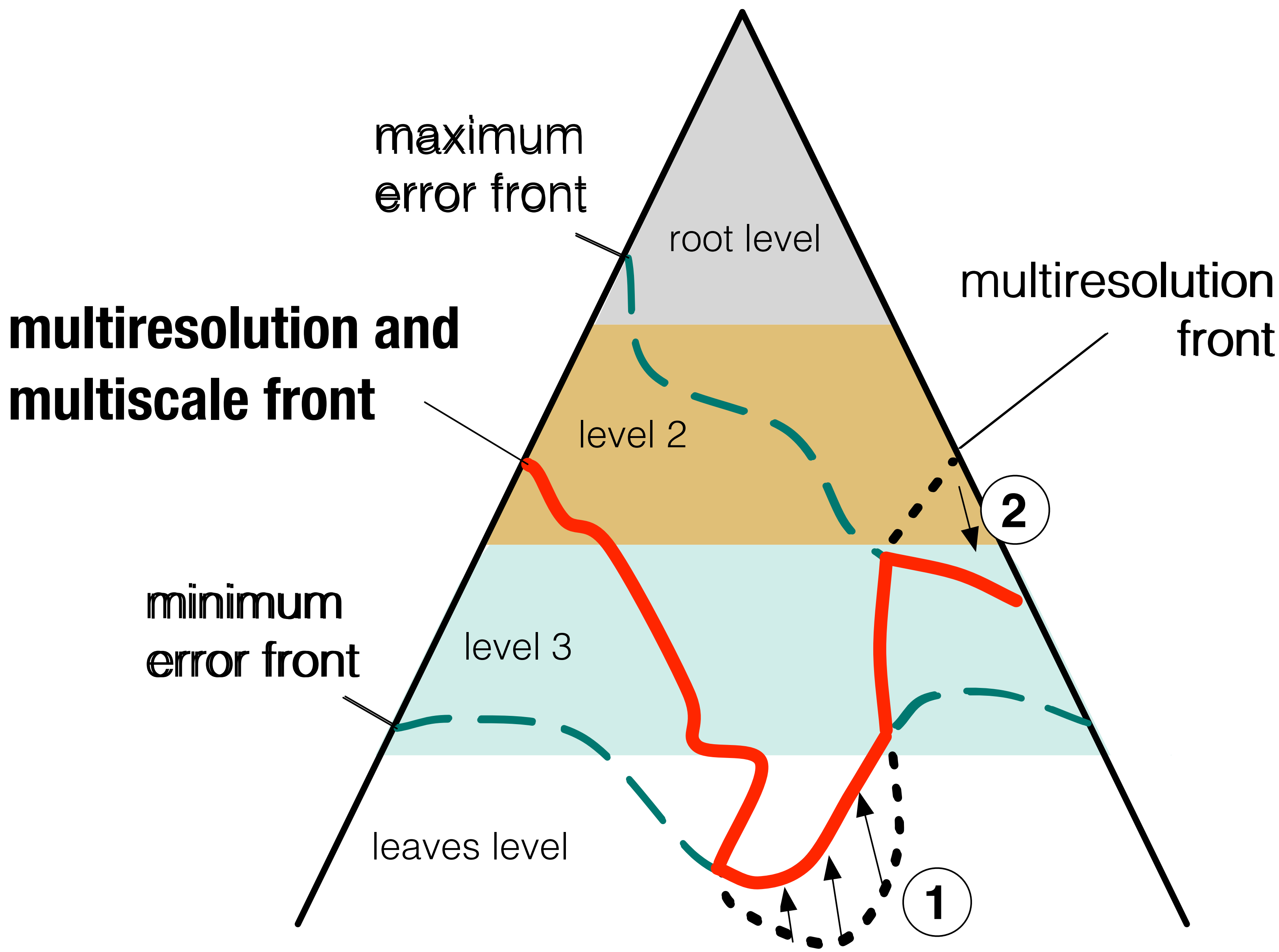
- Multiscale rendering does not map to typical multiresolution hierarchy rendering
- Feature scale per brick
- One feature scale matches to different resolutions of a brick
- Swap between resolution levels
- Link feature scale parameter to the renderer

Increase Feature Scale Parameter



swap resolution

Adaptive Multiresolution and Multiscale Visualization



1 prevent block refinement

2 force block refinement

Multiscale Feature Visualization Demo (1)

Flower model

Model size: 1024^3

Screen resolution: 1024^2

Storage Costs of TA Hierarchy Models

- Theoretical costs
  - without empty space skipping
  - without pruning/thresholding of coefficients
- Assumptions
  - brick size $B = 64$
  - initial rank $R_{init} = 32$
- **Suter et al., 2011**: $\approx 0.17 \cdot I^3$
- **Suter et al., 2013**: $\approx 192 \cdot I + 0.14 \cdot I^3$

VISUALIZATIONAND MULTIMEDIALAB

Storage Costs of TA Hierarchy Models

- Core tensors
  - Wu et al., 2008: $O(\log(I) \cdot R^3)$
  - Suter et al., 2013: $O(R^3)$
- Factor matrices
  - Wu et al., 2008: $O(4 \cdot I \cdot R) + O(\frac{I^6}{B^6})$
  - Suter et al., 2013: $O(6 \cdot I \cdot R)$
- Rank
  - Wu et al., 2008: $R = \frac{I}{16}$
  - Suter et al., 2013: $R = \frac{B}{2} = 32$
- Pruning is an important factor for Wu et al., 2008

VISUALIZATIONAND MULTIMEDIALAB

Quantization of TA Hierarchy Models

- Compact representation coefficients usually floating point numbers
- Quantize coefficients

	factor matrices	core tensors
Wu et al., 2008	8-bit	8…20-bit
Suter et al., 2011	16-bit	8-bit
Suter et al., 2013	32-bit*	8-bit

*no quantization

VISUALIZATIONAND MULTIMEDIALAB

Storage Costs of TA Hierarchy Models

- Theoretical costs
  - without empty space skipping
  - without pruning/thresholding of coefficients
- Assumptions
  - brick size $B = 64$
  - initial rank $R_{init} = 32$
- **Wu et al., 2008**:
- **Suter et al., 2011**: $0.20 \cdot I^3$
- **Suter et al., 2013**: $768 \cdot I + 0.14 \cdot I^3$

VISUALIZATIONAND MULTIMEDIALAB

Summary
Scientific Visualization Applications

- Compact data representations in scientific visualization
  - Tucker models
- Multiscale feature extraction
  - tensor rank truncation
- Hierarchical (multiresolution) Tucker models
  - residual-based approach (pruning important) [Wu et al., 2008]
  - simple brick-based multiresolution model [Suter et al., 2011]
  - global bases; multiresolution and multiscalability [Suter et al., 2013]
- Compression via TA

Tutorial: Tensor Approximation in Scientific Visualization

Implementation Examples

Susanne K. Suter, Rafael Ballester-Ripoll, Renato Pajarola

University of Zurich UZH

VISUALIZATION AND MULTIMEDIA LAB

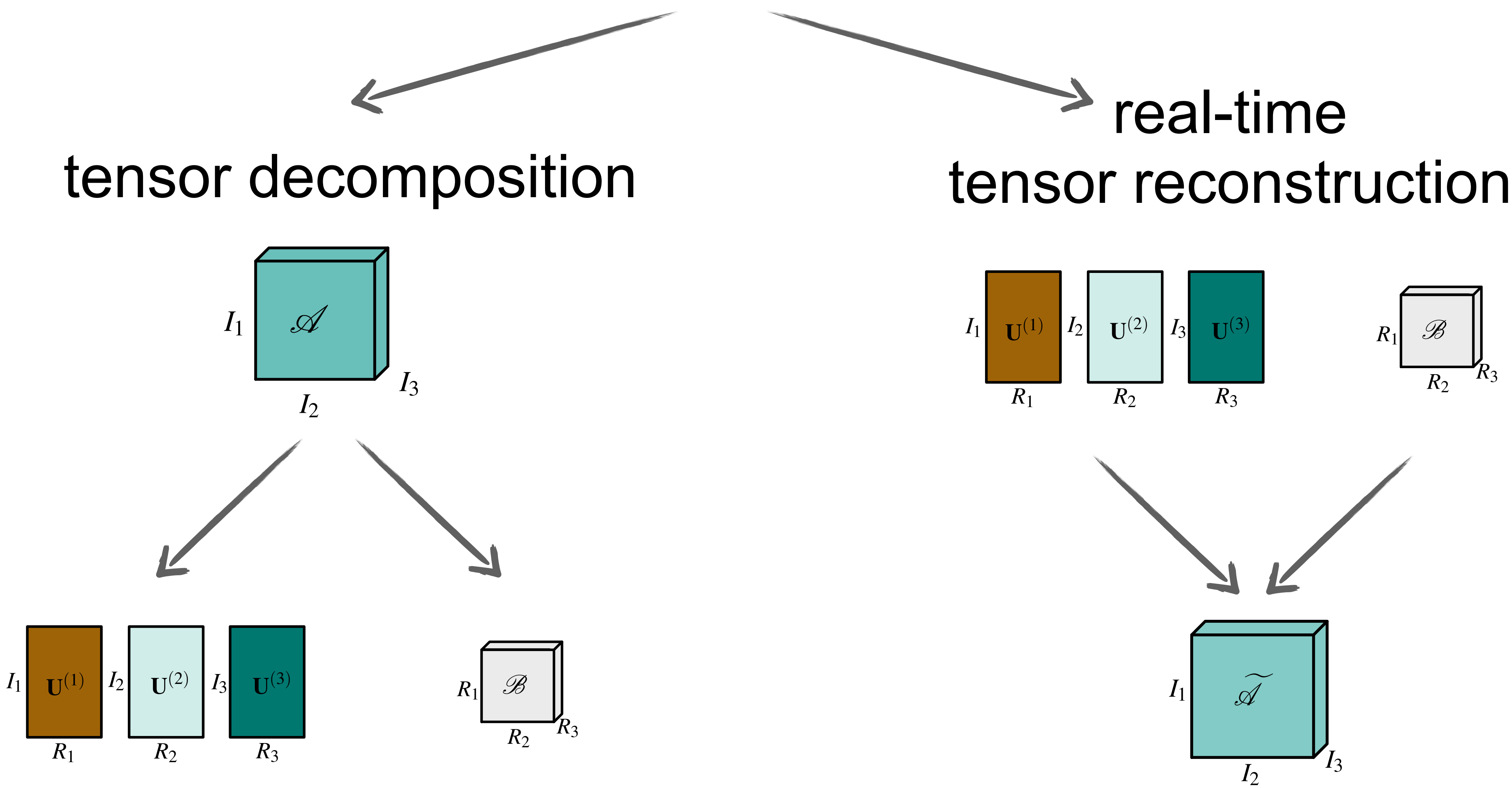
Outline

- Part 1: Typical decomposition algorithms/operations
- Part 2: GPU-based tensor reconstruction

VISUALIZATIONAND MULTIMEDIALAB

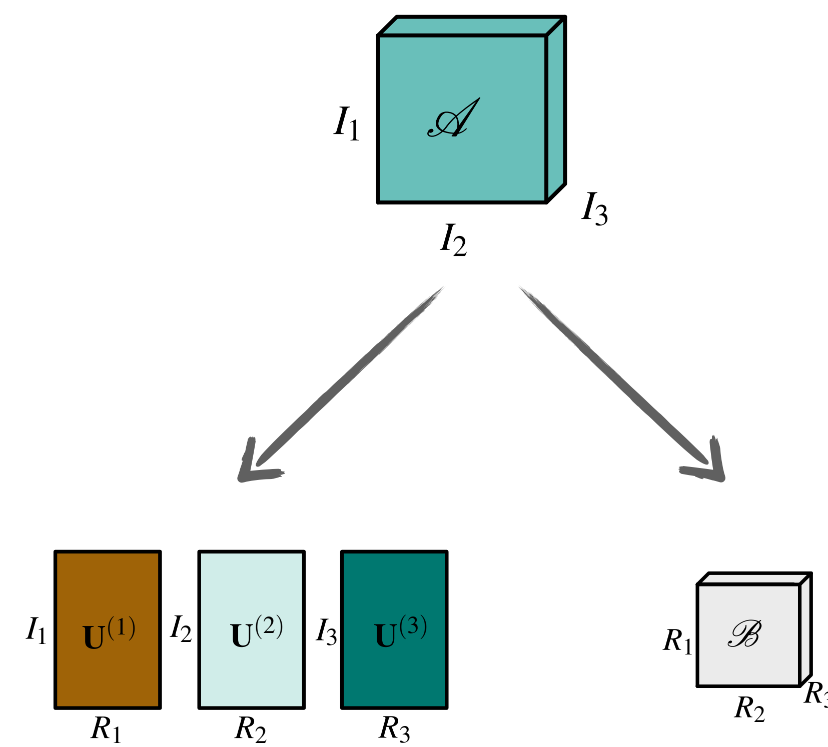
Typical TA Operations

tensor approximation

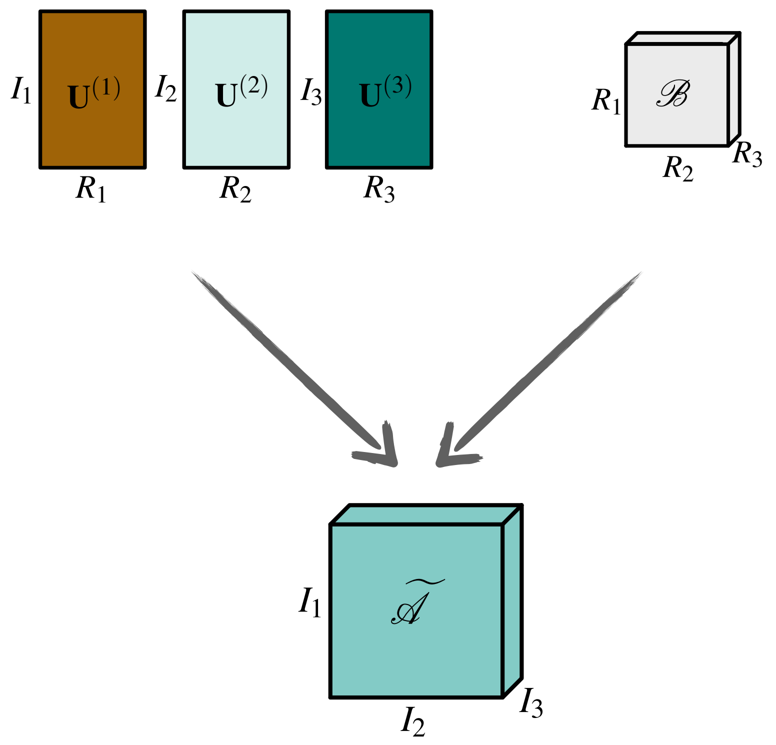


Tensor Decomposition

- Create factor matrices
  - higher-order SVD (HOSVD)
    - tensor unfolding
  - alternating least-squares (ALS) algorithms
    - higher-order orthogonal iteration (HOOI)
    - higher-order power method (HOPM)
- Generate core tensor
  - tensor times matrix (TTM) multiplications

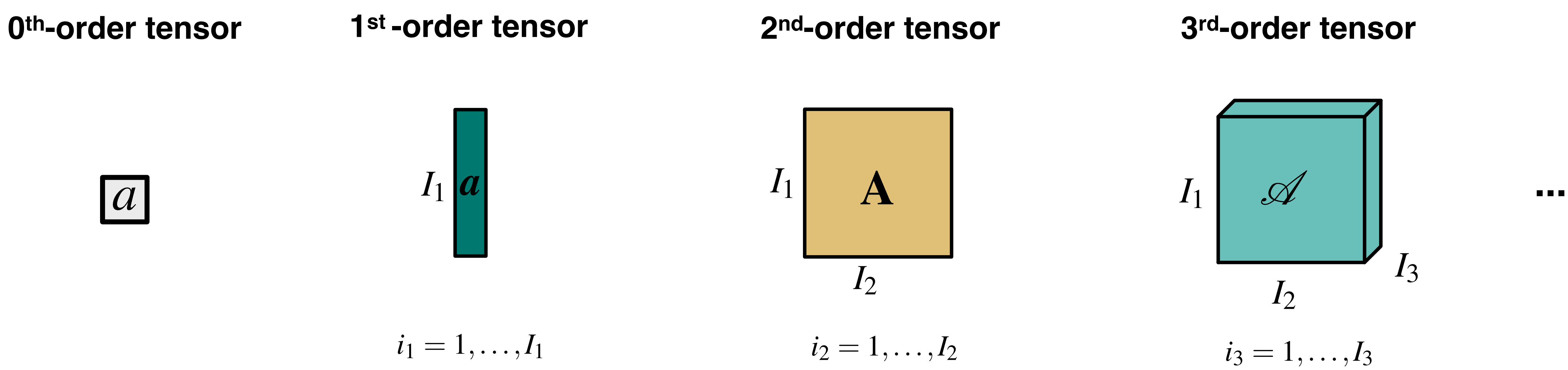


Tensor Reconstruction

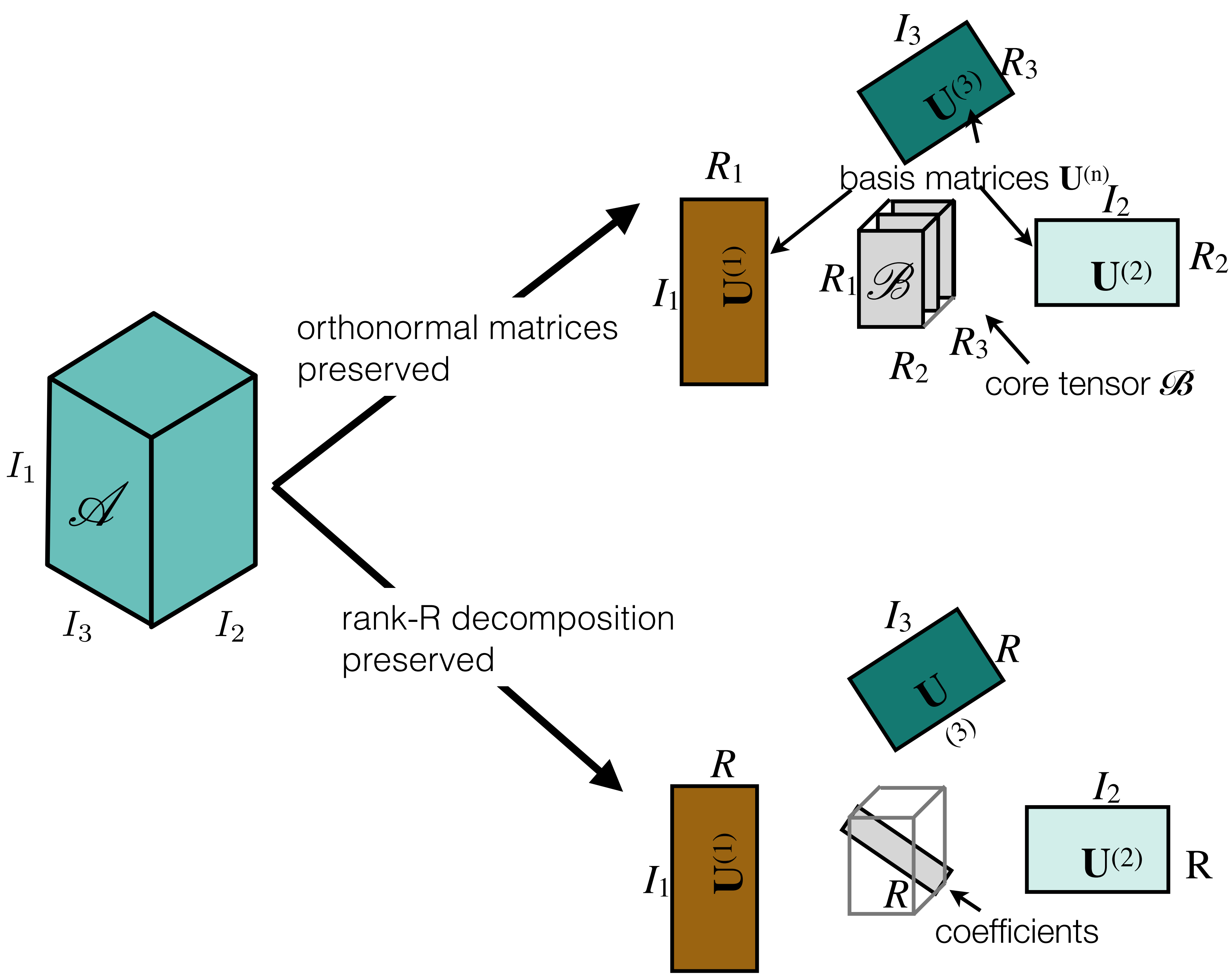


- Realtime (!) reconstruction
  - tensor times matrix (TTM) multiplications

Tensor: A Multidimensional Array



SVD Extension to Higher Orders



Tucker

- Three-mode factor analysis (**3MFA/Tucker3**) [Tucker, 1964+1966]
- Higher-order SVD (**HOSVD**) [De Lathauwer et al., 2000a]

CP

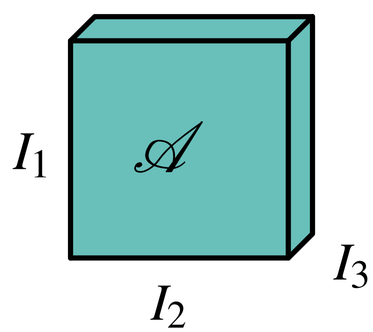
- **PARAFAC** (parallel factors) [Harshman, 1970]
- **CANDECOMP** (CAND) (canonical decomposition) [Caroll & Chang, 1970]

Part 1: Typical Decomposition Algorithms and Operations

Downloads

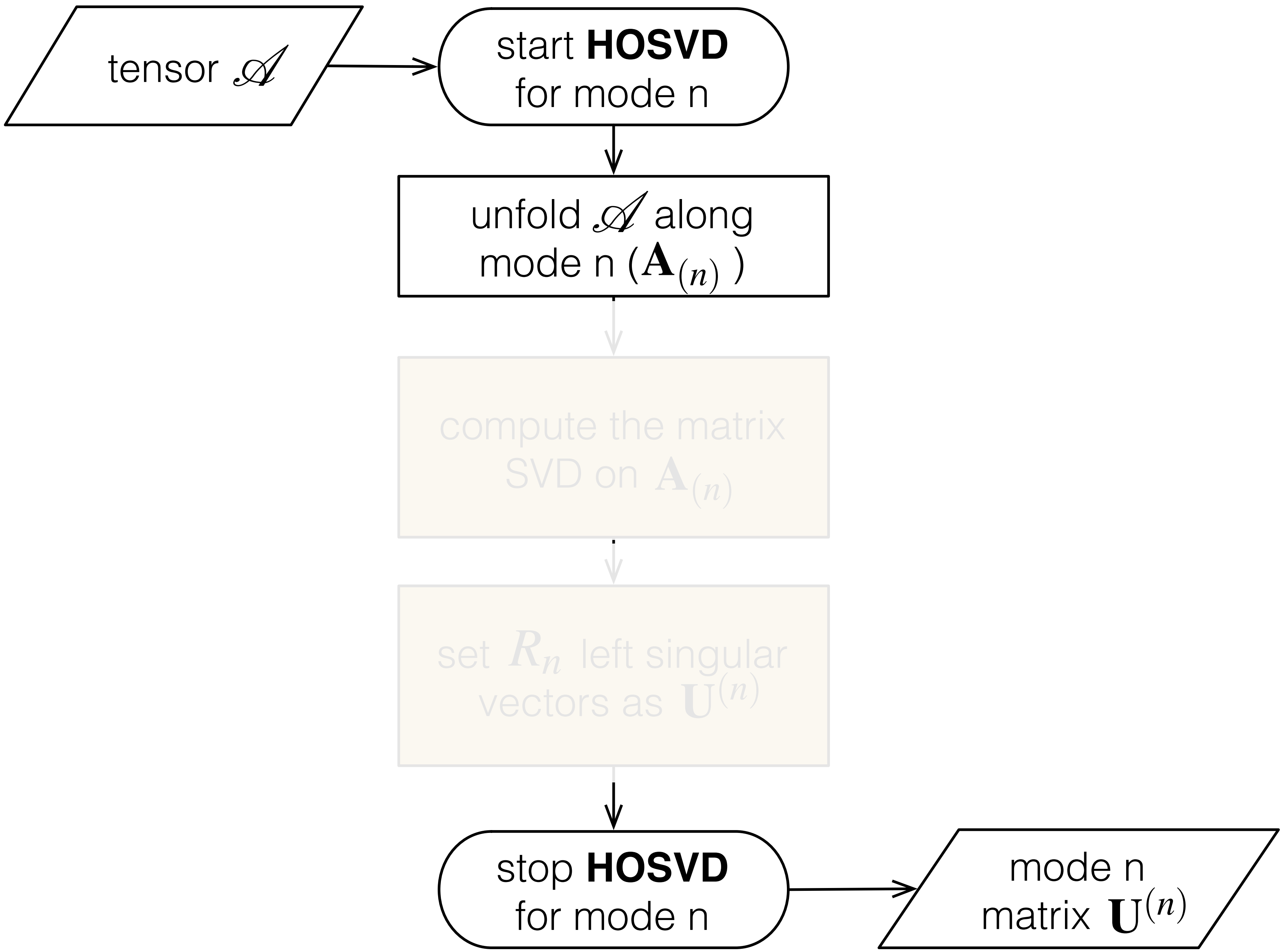
- MATLAB tensor toolbox
  - http://www.sandia.gov/~tgkolda/TensorToolbox
- vmmlib: C++ library for vectors, matrices, and tensor approximation
  - http://vmml.github.io/vmmlib/
- Tensor tutorial notes
  - http://vmml.ifi.uzh.ch/links/TutorTensorAprox.html

Test Dataset: Hazelnut



- A microCT scan of dried hazelnuts
- $I_1 = I_2 = I_3 = 512$
- Values: unsigned char (8bit)
- http://vmml.ifi.uzh.ch/research/datasets.html

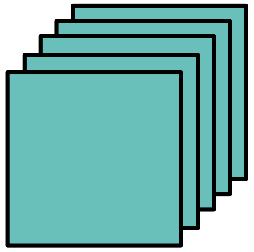
Higher-order SVD (HOSVD)

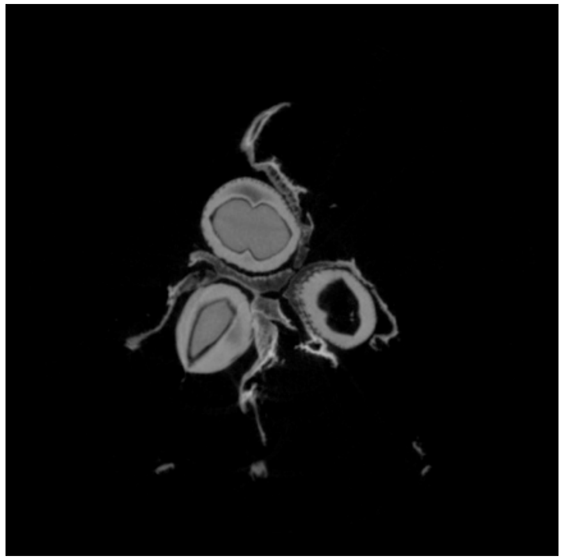


De Lathauwer, de Moor, Vandewalle. A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications*, 21(4):1253–1278, 2000.

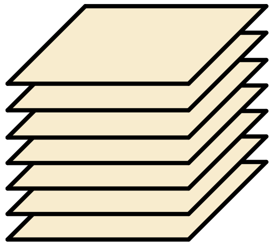
Slices of a Tensor3

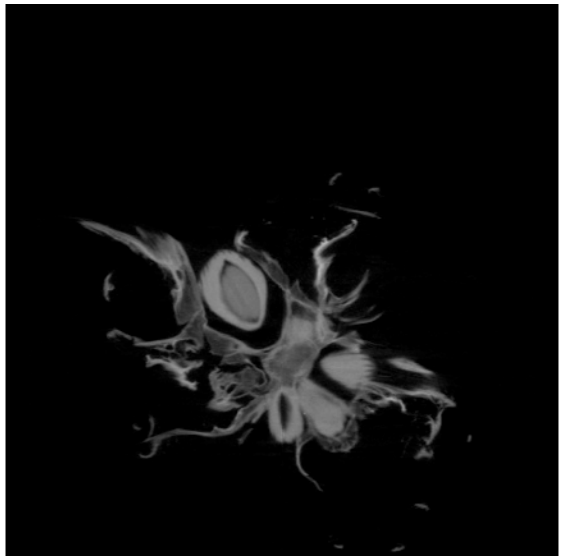
frontal slices



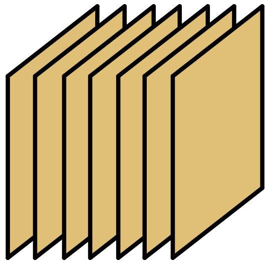


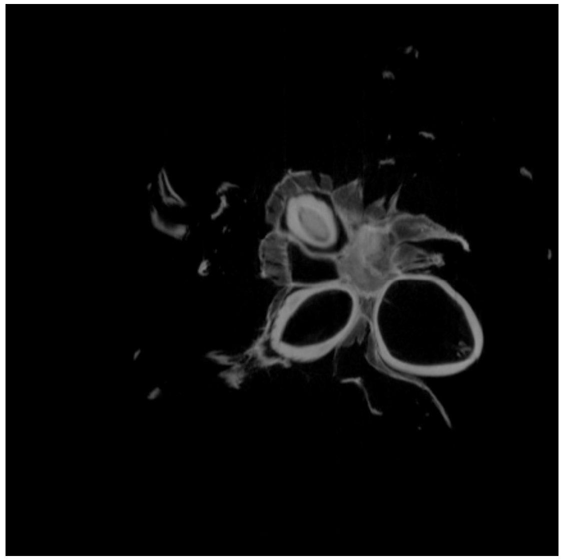
horizontal slices





lateral slices





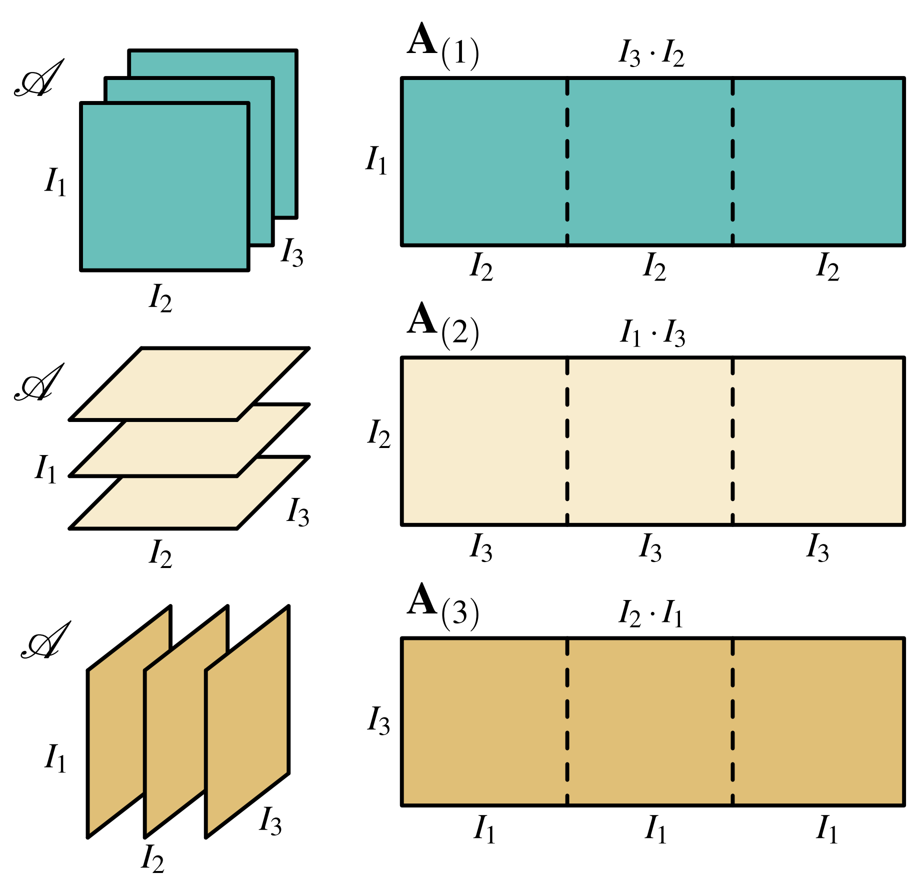
[vmmlib]

```
matrix< 512, 512, values_t > slice;

t3.get_frontal_slice_fwd( 256, slice );

t3.get_horizontal_slice_fwd( 256, slice );

t3.get_lateral_slice_fwd( 256, slice );
```

VISUALIZATIONAND MULTIMEDIALAB

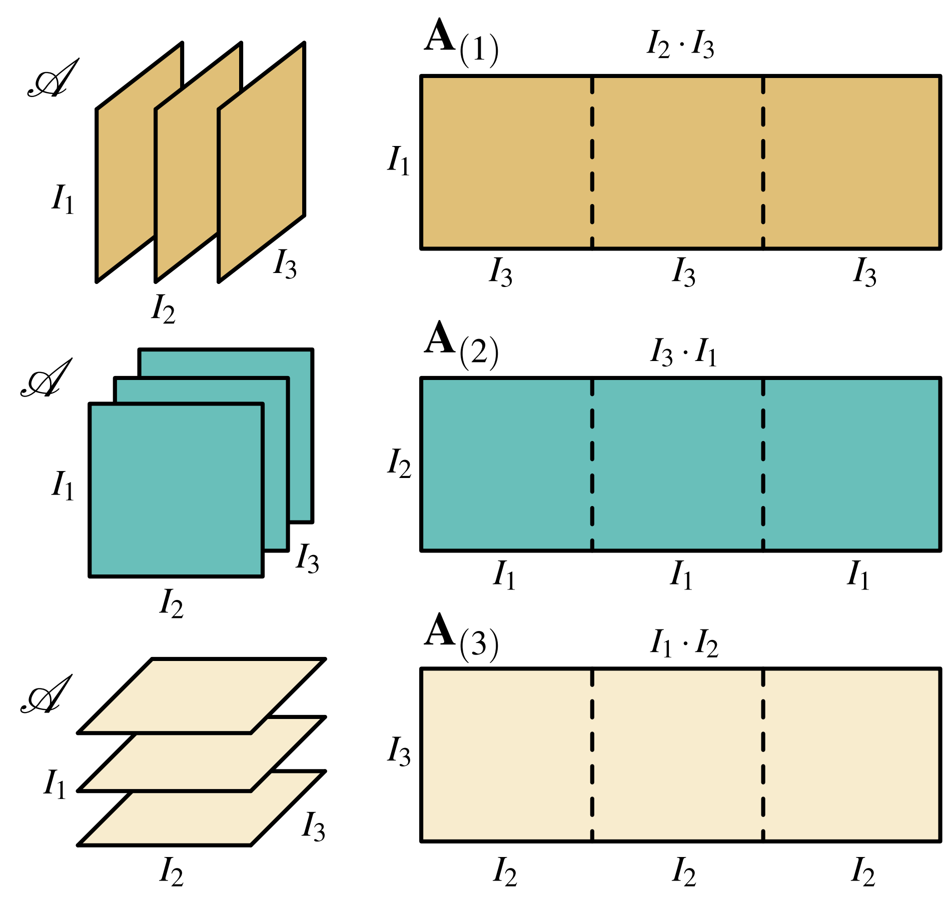
Tensor Unfolding (Matricization)

forward cyclic unfolding



Kiers. Towards a standardized notation and terminology in multiway analysis. *Journal of Chemometrics*, 14(3):105–122, 2000.

backward cyclic unfolding



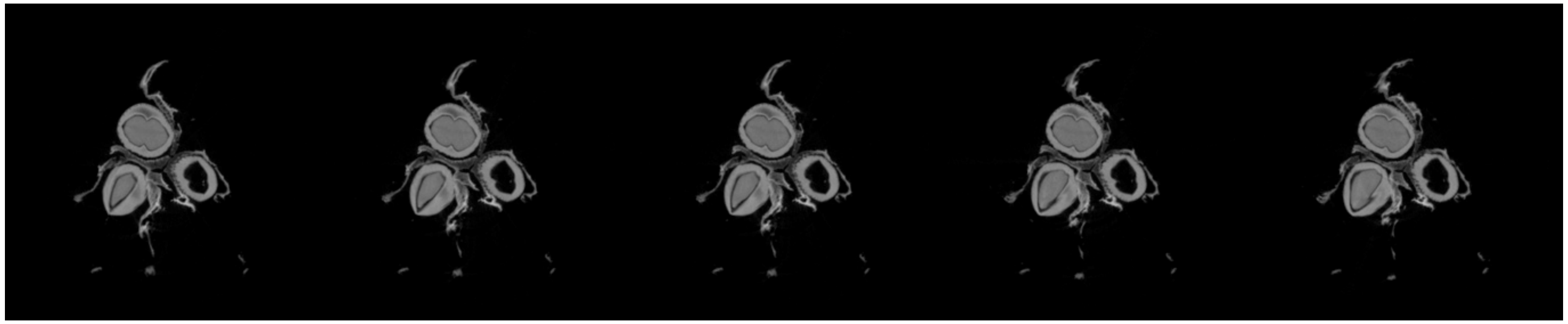
De Lathauwer, de Moor, Vandewalle. A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications*, 21(4):1253–1278, 2000.

Tensor Unfolding Example

262’144

512

mode-1 unfolding …

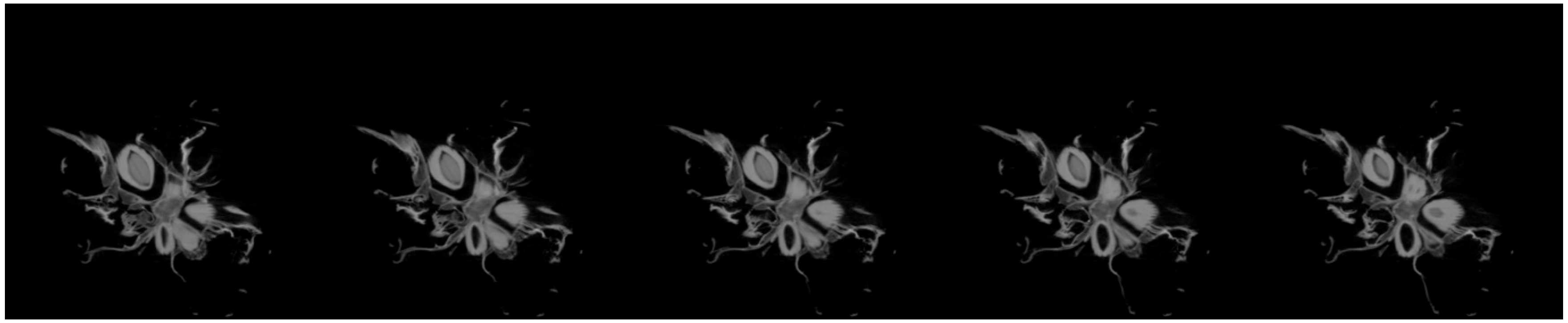


…

262’144

512

mode-2 unfolding …

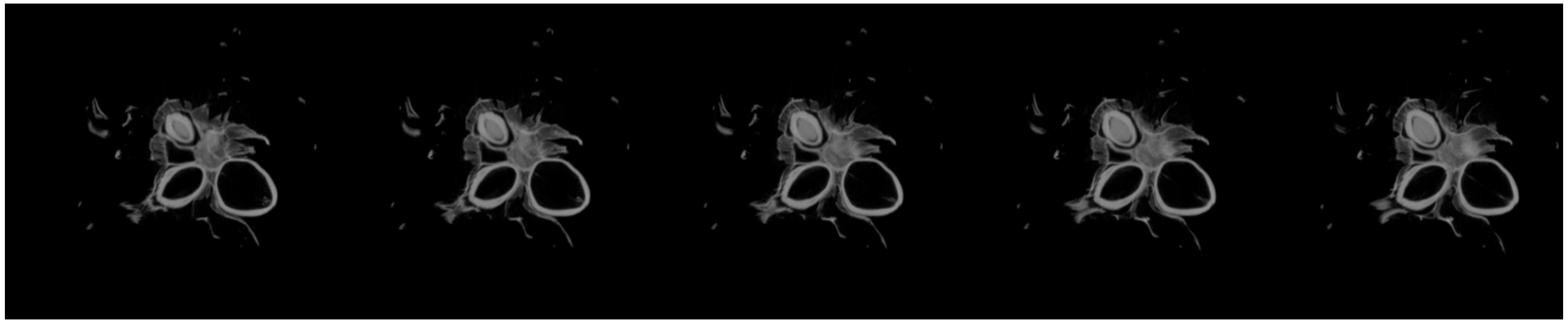


…

262’144

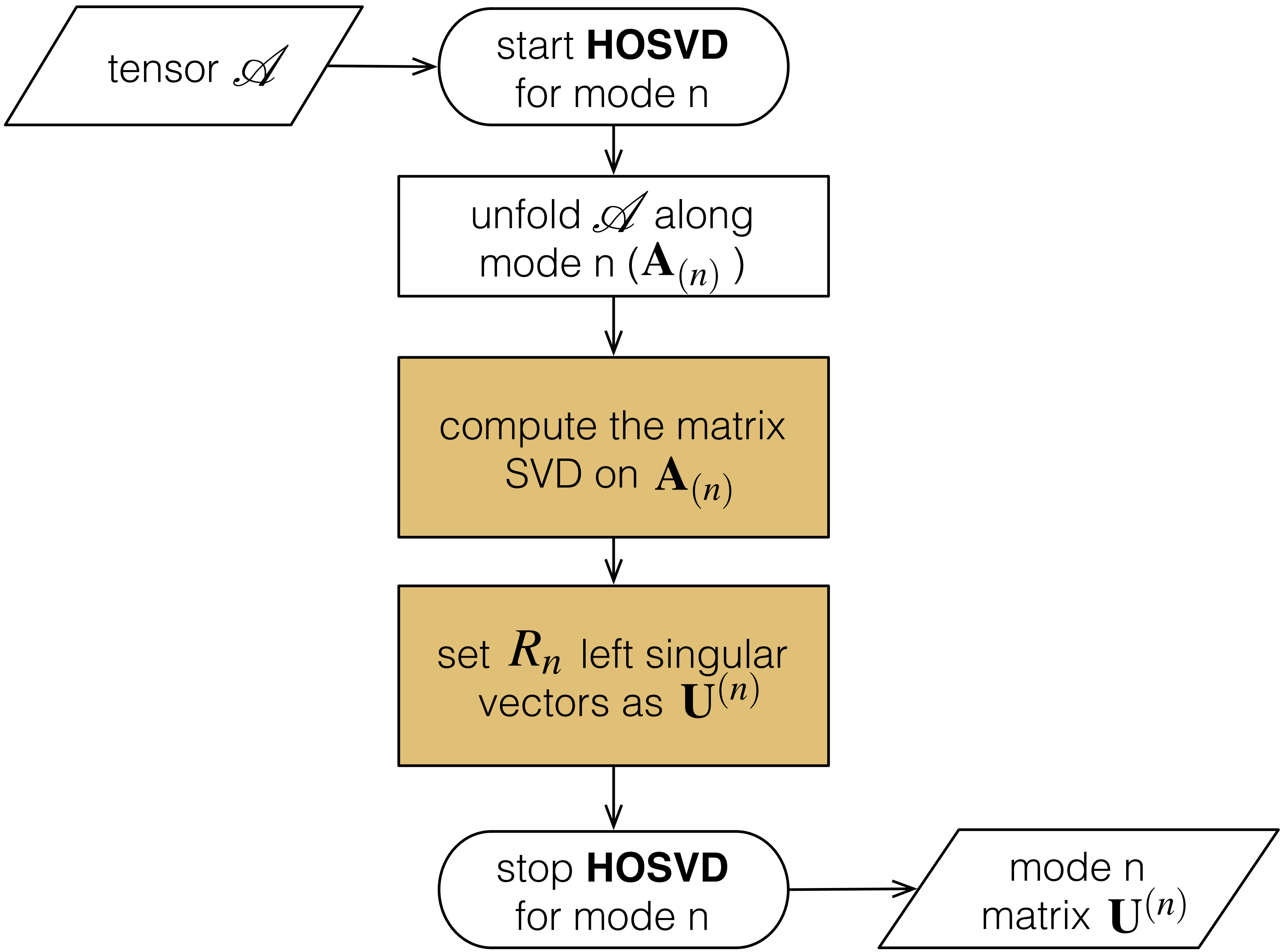
512

mode-3 unfolding …



…

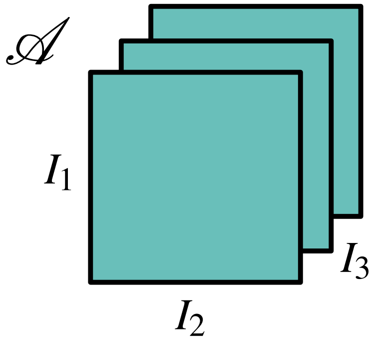
Higher-order SVD (HOSVD)



De Lathauwer, de Moor, Vandewalle. A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications*, 21(4):1253–1278, 2000.

VISUALIZATIONAND MULTIMEDIALAB

Large Data Tensors (in vmmlib)



[vmmlib]

```
const size_t d = 512;
typedef tensor3< d,d,d, unsigned char > t3_512u_t;
typedef t3_converter< d,d,d, unsigned char > t3_conv_t;
typedef tensor_mmapper< t3_512u_t, t3_conv_t > t3map_t;


std::string in_dir = "./dataset";
std::string file_name = "hnut512_uint.raw";
t3_512u_t t3_hazelnut;
t3_conv_t t3_conv;

t3map_t t3_mmap( in_dir, file_name, true, t3_conv ); //true -> read-only
t3_mmap.get_tensor( t3_hazelnut );
```

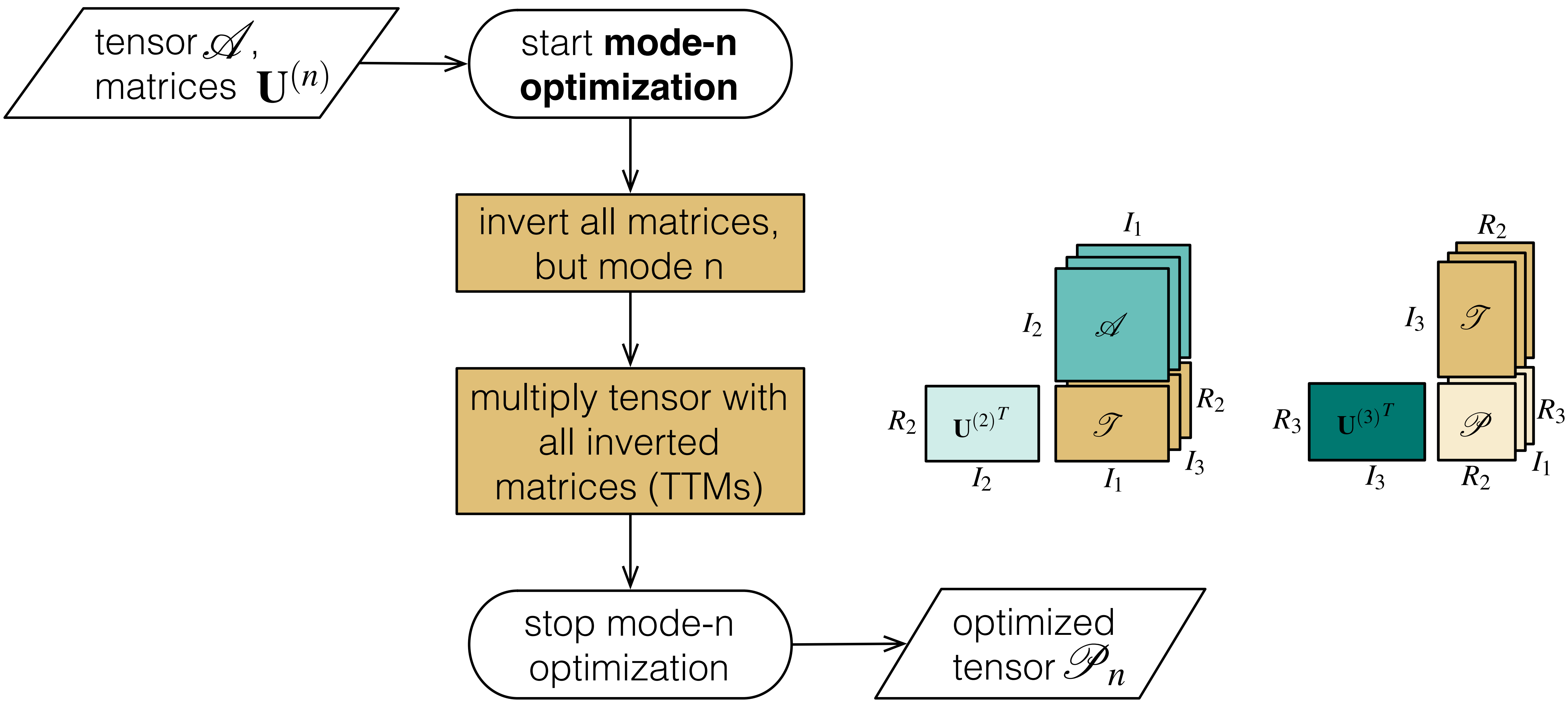
VISUALIZATIONAND MULTIMEDIALAB

Optimize Factor Matrices

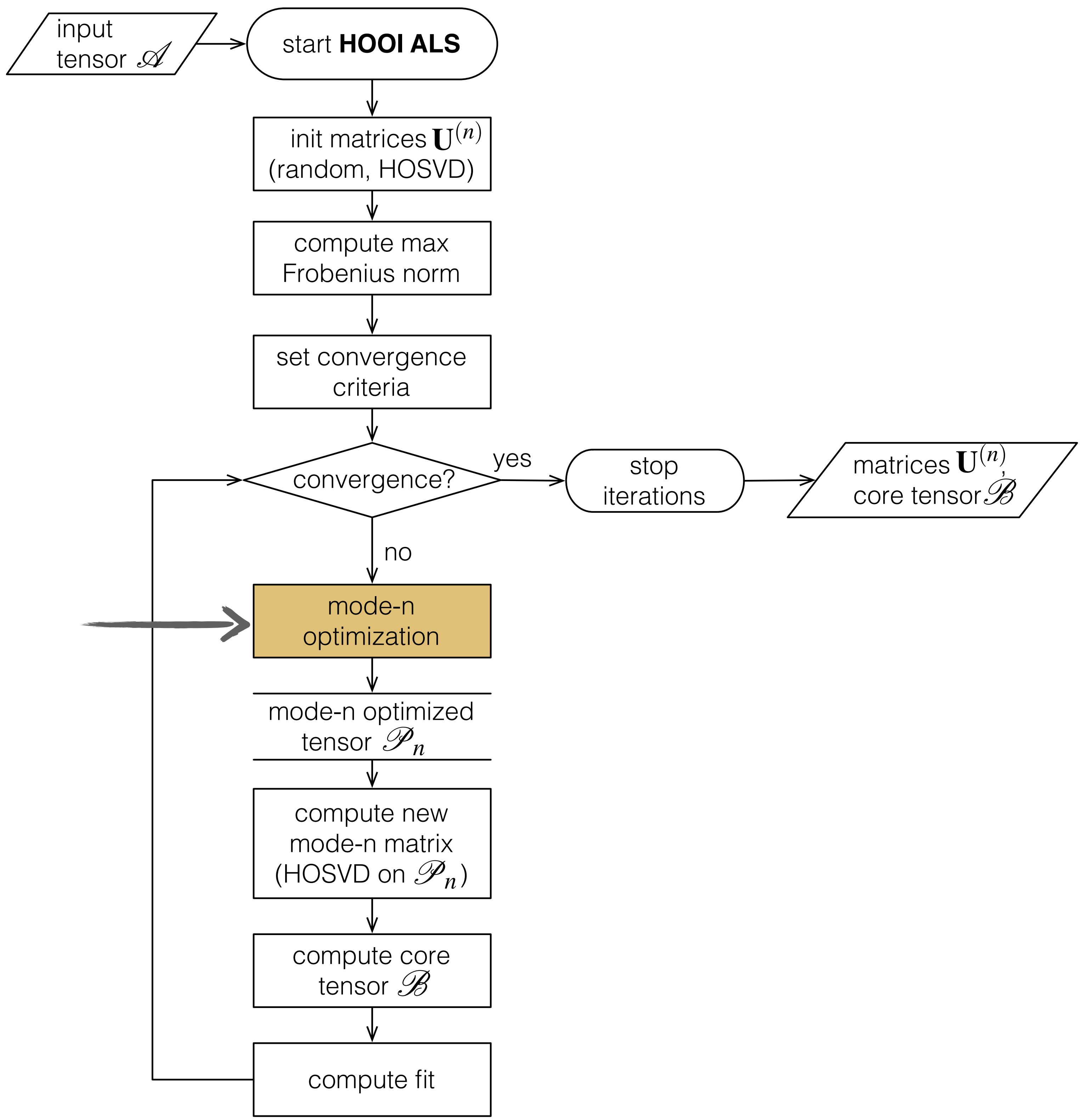
- Higher-order orthogonal iteration
  - optimize factor matrix of mode n
  - keep factor matrices of all other modes fixed
  - generate optimized data tensor
    - project original data tensor on the inverted factor matrices of all other modes
  - receive optimized mode-n factor matrix
    - apply HOSVD to the optimized tensor

De Lathauwer, de Moor, Vandewalle. On the best rank-1 and rank-(R1,R2,...,RN) approximation of higher-order tensors. *SIAM Journal on Matrix Analysis and Applications*, 21(4):1324–1342, 2000.

Optimize Mode-n Factor Matrix

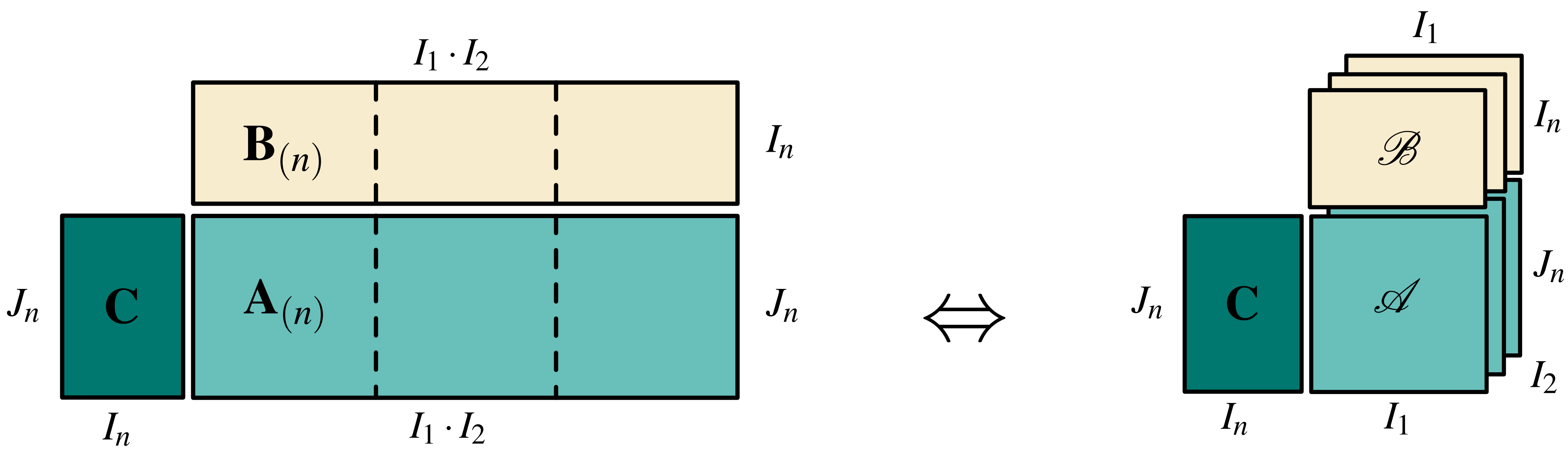


Higher-order Orthogonal Iteration (HOOI)



VISUALIZATIONAND MULTIMEDIALAB

Tensor Times Matrix Multiplication



$\mathbf{A}_{(n)} = \mathbf{C}\mathbf{B}_{(n)}$

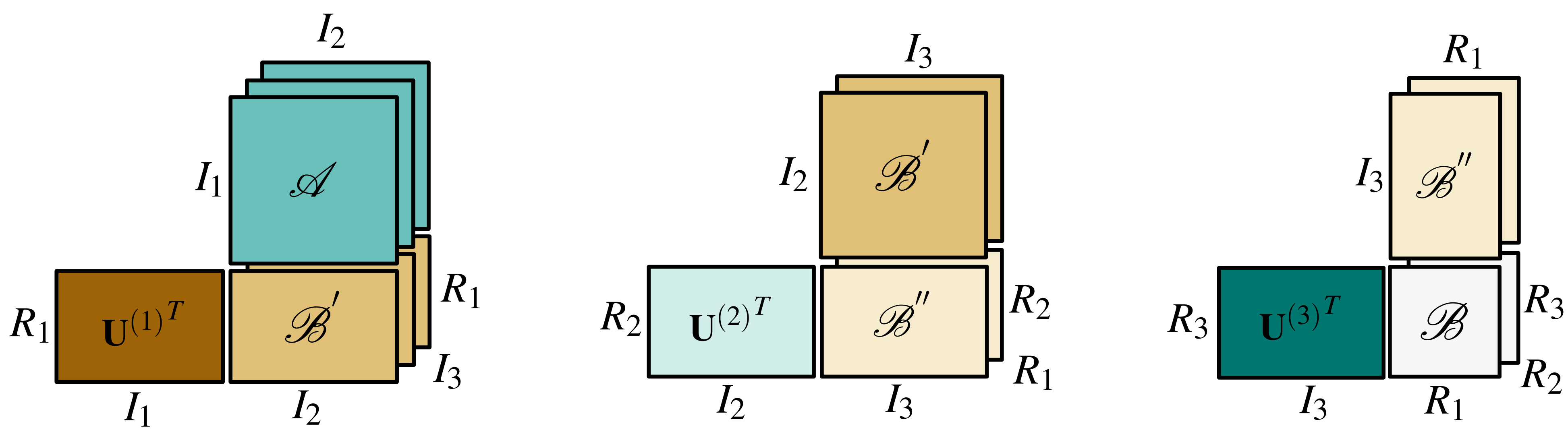
$\mathscr{A} = \mathscr{B} \times_n \mathbf{C}$

n-mode product

[De Lathauwer et al., 2000a]

$$(\mathscr{B} \times_n \mathbf{C})_{i_1 \ldots 1_{n-1} j_n i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_n} b_{i_1 i_2 \ldots i_N} \cdot c_{j_n i_n}$$

Example TTMs: Core Computation



$$\mathscr{B} = \mathscr{A} \times_1 \mathbf{U}^{(1)^{(-1)}} \times_2 \mathbf{U}^{(2)^{(-1)}} \times_3 \cdots \times_N \mathbf{U}^{(N)^{(-1)}} \xrightarrow[\text{factor matrices}]{\text{orthogonal}} \mathscr{B} = \mathscr{A} \times_1 \mathbf{U}^{(1)^T} \times_2 \mathbf{U}^{(2)^T} \times_3 \cdots \times_N \mathbf{U}^{(N)^T}$$

- Three consecutive TTM multiplications
- For orthogonal matrices, use the transposes of the three factor matrices (otherwise the (pseudo)-inverses)

VISUALIZATIONAND MULTIMEDIALAB

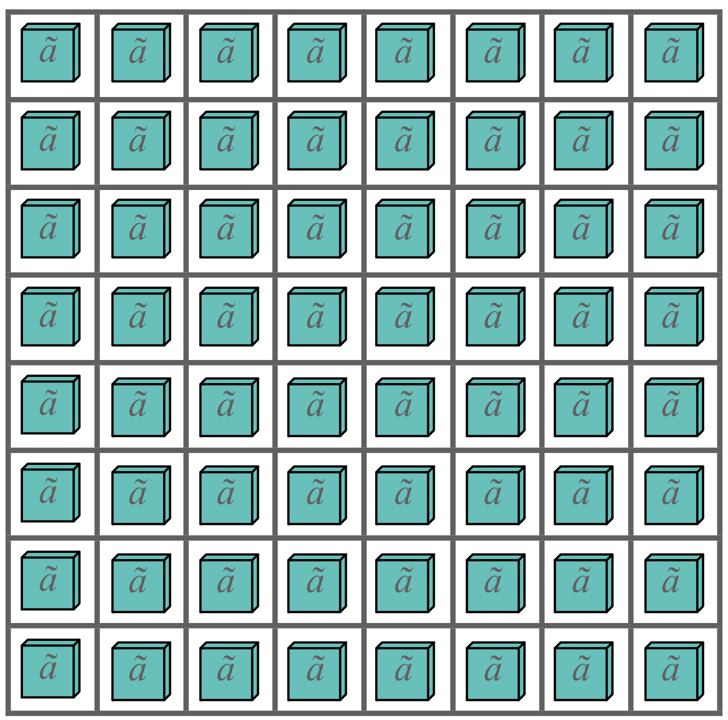
Part 2: GPU Reconstruction

Suter et al.. Interactive multiscale tensor reconstruction for multiresolution volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2135–2143, December 2011.

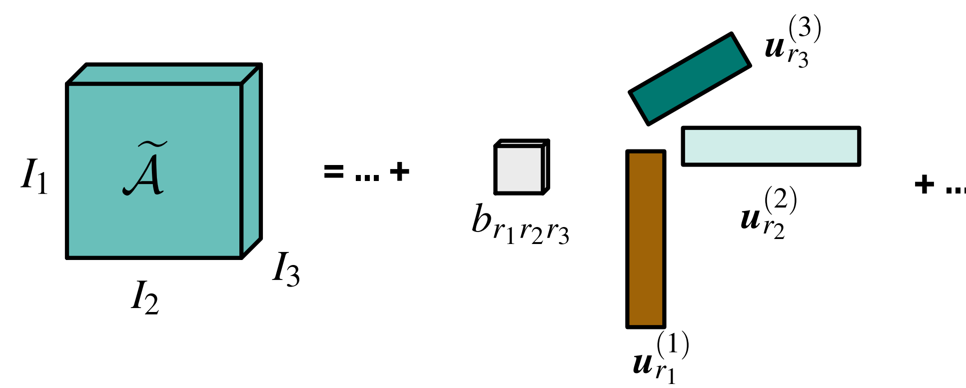
University of Zurich[UZH]

VISUALIZATIONAND MULTIMEDIALAB

Parallel Tensor Reconstruction

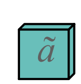


parallel computing grid per brick



$$\widetilde{a}_{i_1 i_2 i_3} = \sum_{r_1} \sum_{r_2} \sum_{r_3} b_{r_1 r_2 r_3} \cdot u^{(1)}_{i_1 r_1} \cdot u^{(2)}_{i_2 r_2} \cdot u^{(3)}_{i_3 r_3}$$

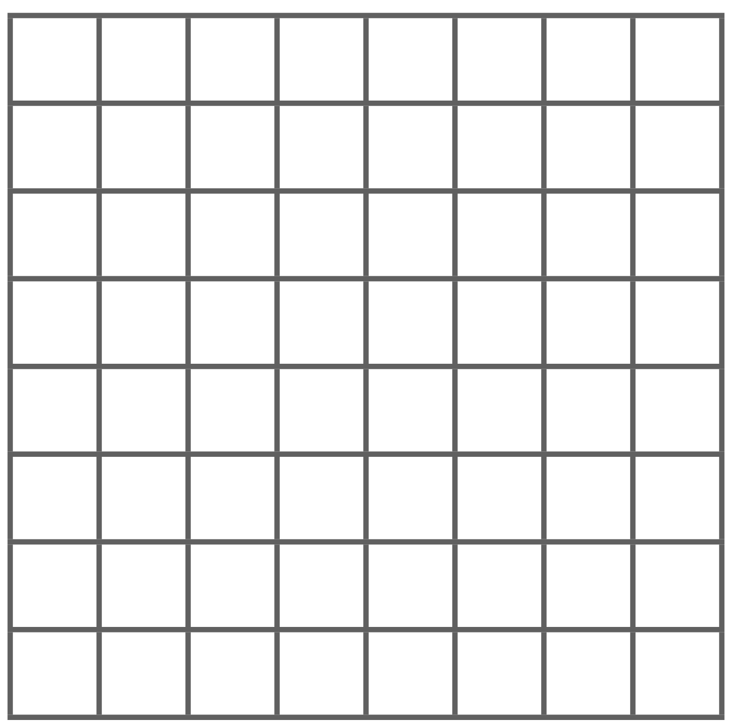
triple-for-loop



computational cost per voxel is <u>cubic</u>: O(R^3)

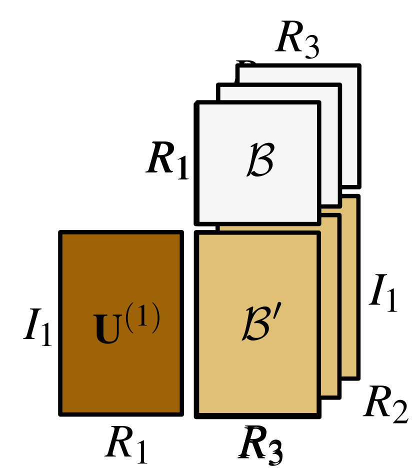
VISUALIZATIONAND MULTIMEDIALAB

Faster Parallel Tensor Reconstruction

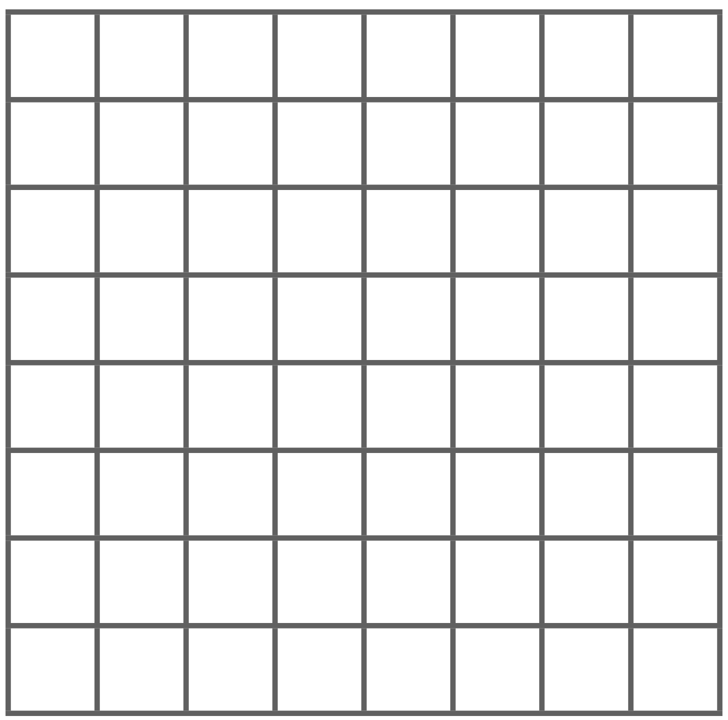


parallel computing grid per brick

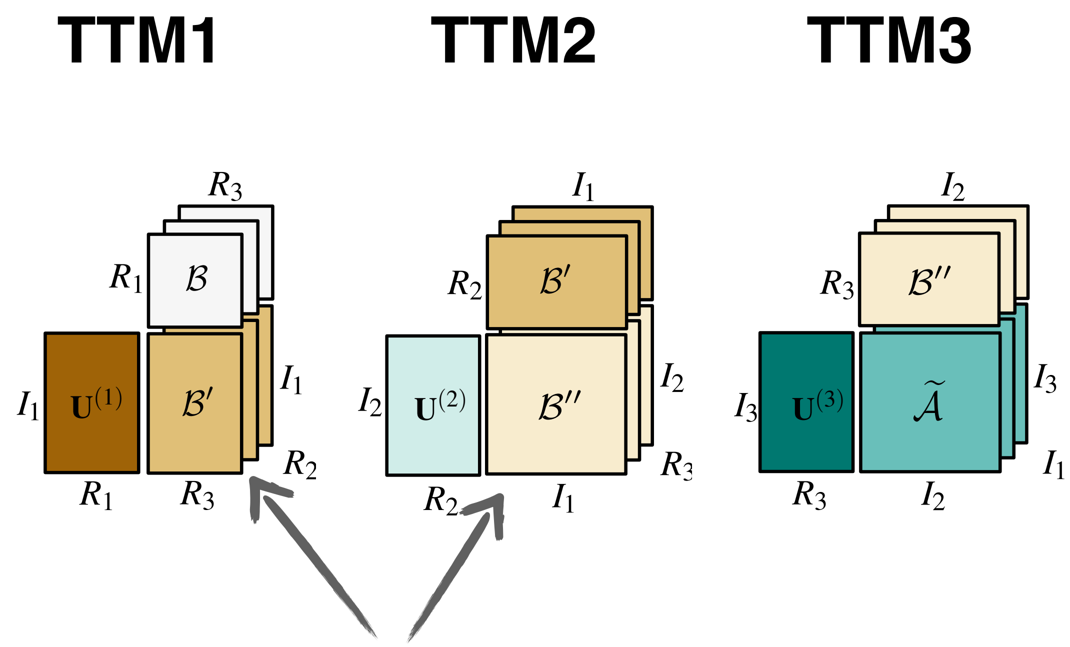
tensor times matrix (TTM) multiplication or n-mode product



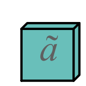
Faster Parallel Tensor Reconstruction



parallel computing grid per brick

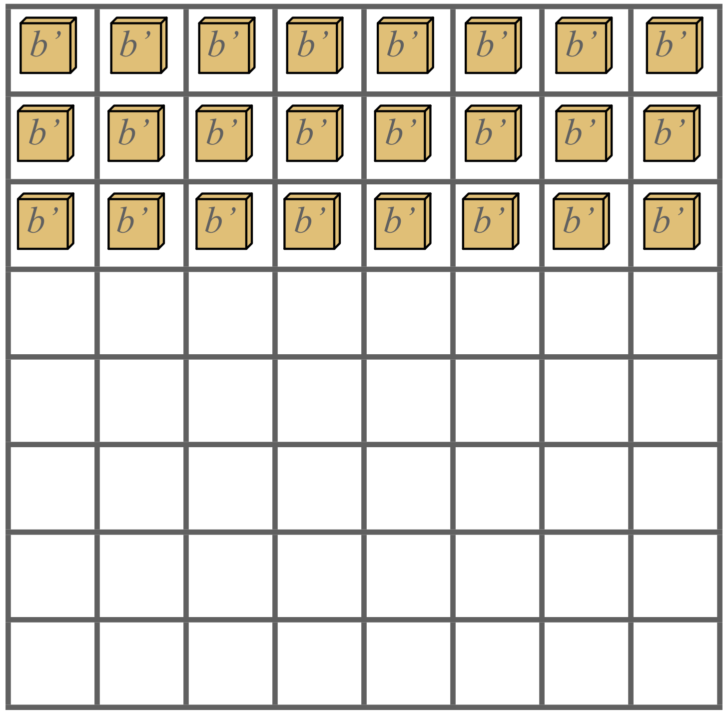


store intermediate results (B' and B'')



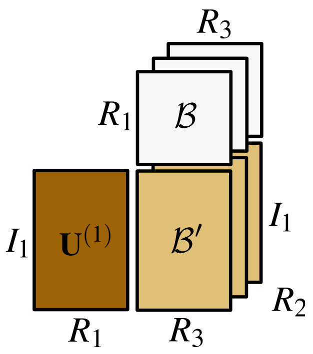
computational cost per voxel is <u>linear</u>: O(R)

Compute Intermediate Tensor B'

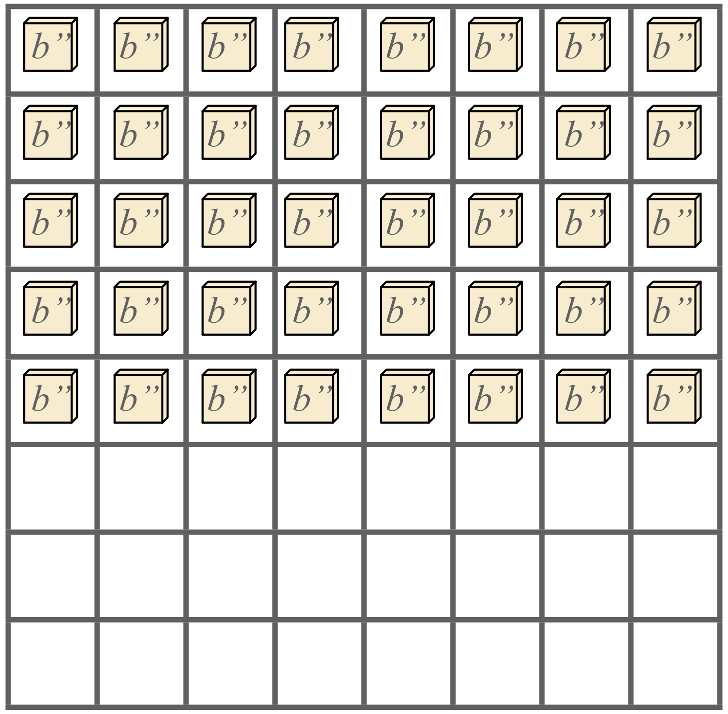


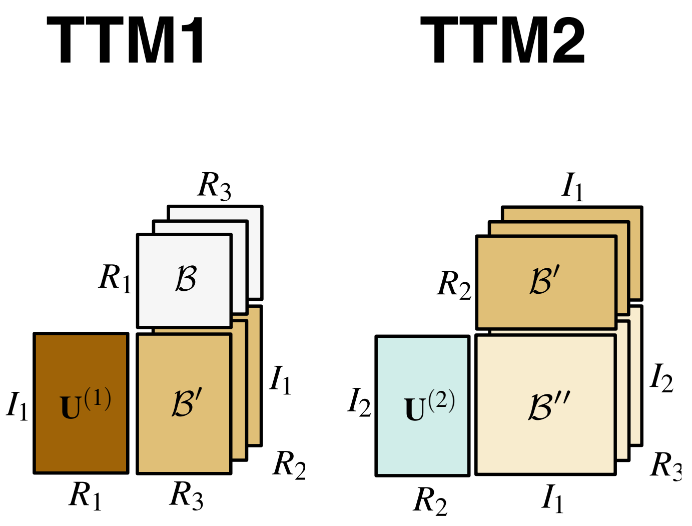
parallel computing grid per brick

TTM1



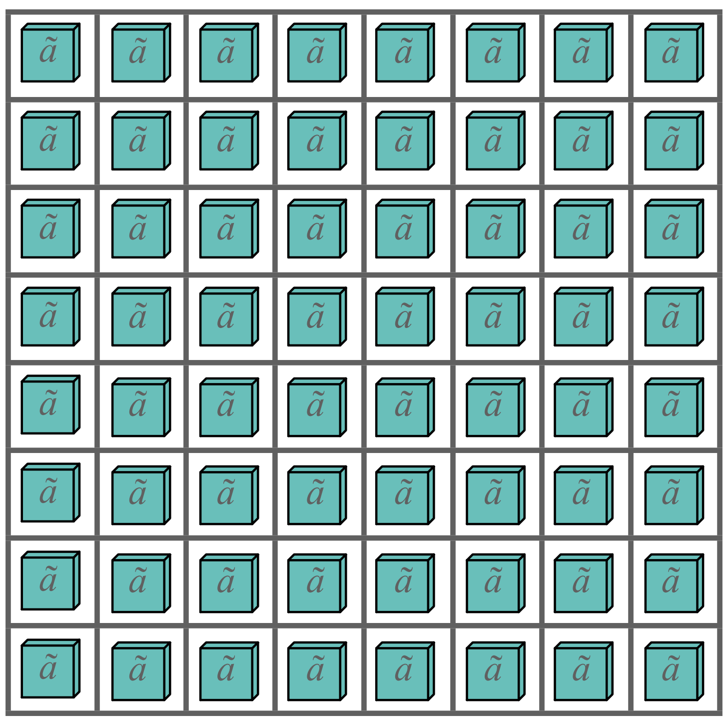
Compute Intermediate Tensor B”

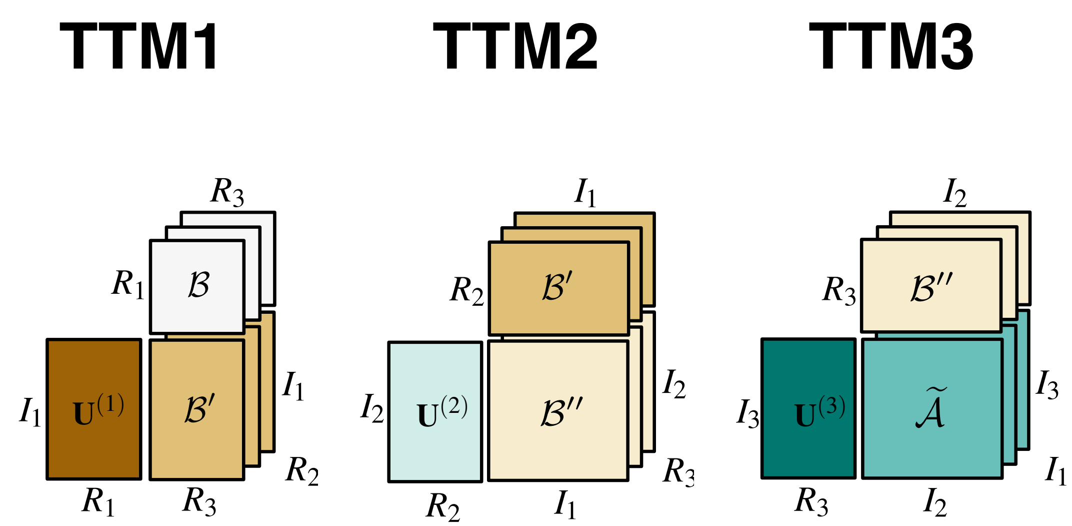




parallel computing grid per brick

Compute Approximated Tensor Ã

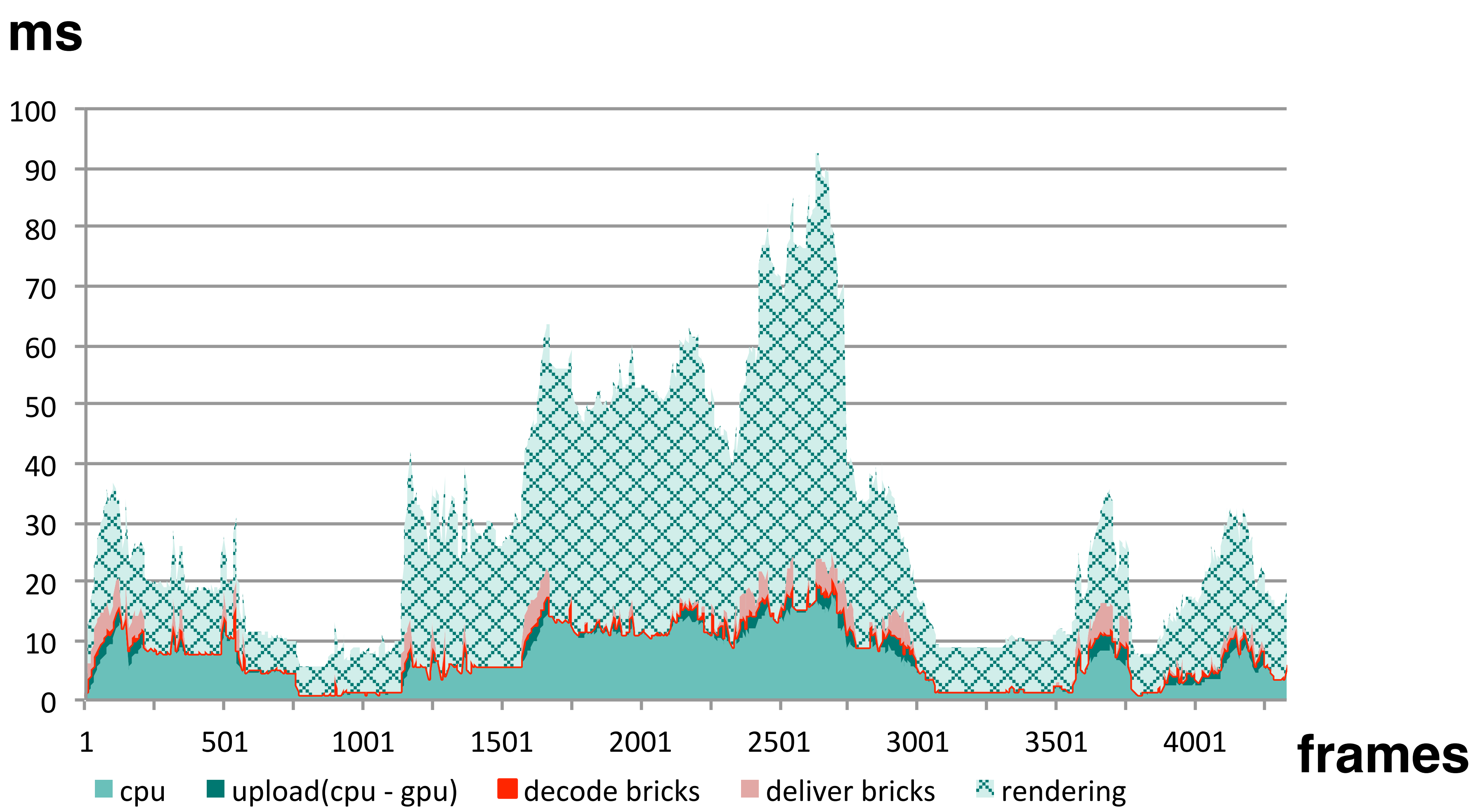


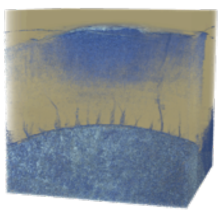


parallel computing grid per brick

Reconstruction Performance

2048^3

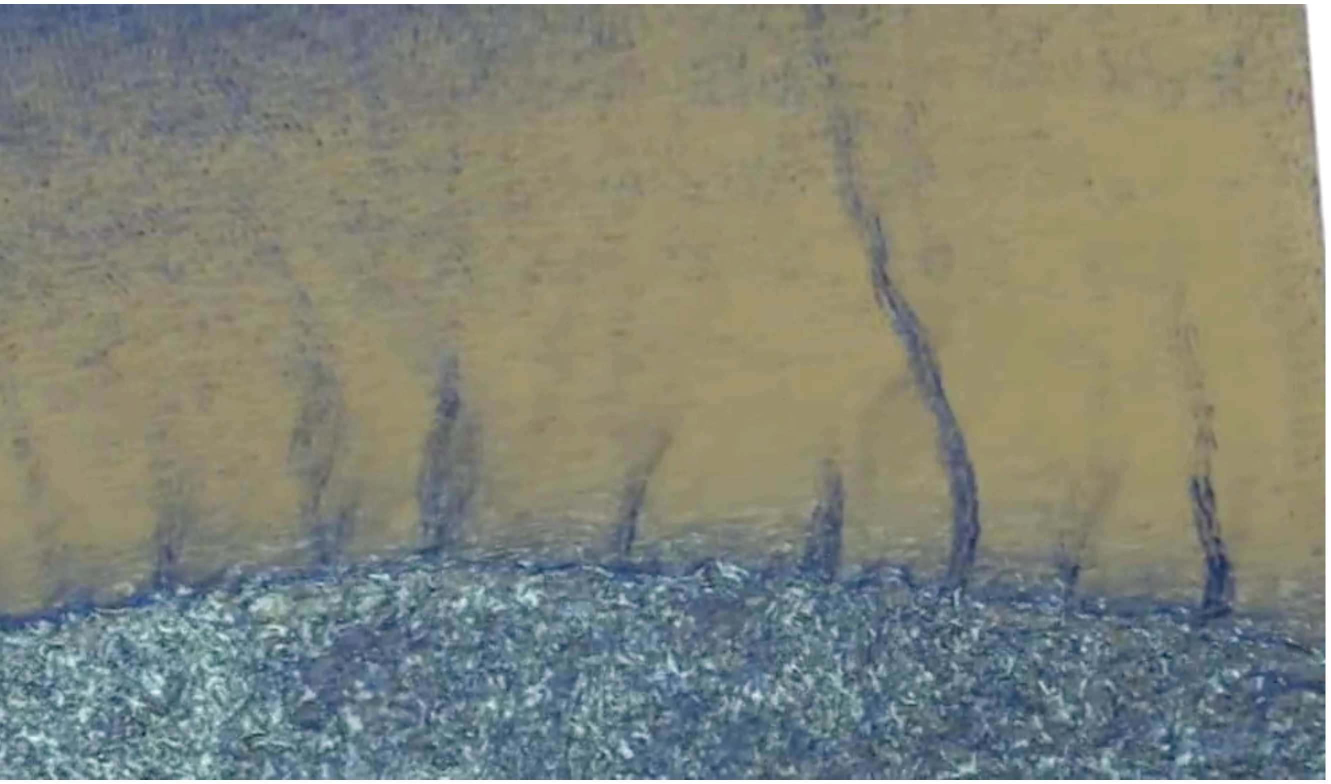


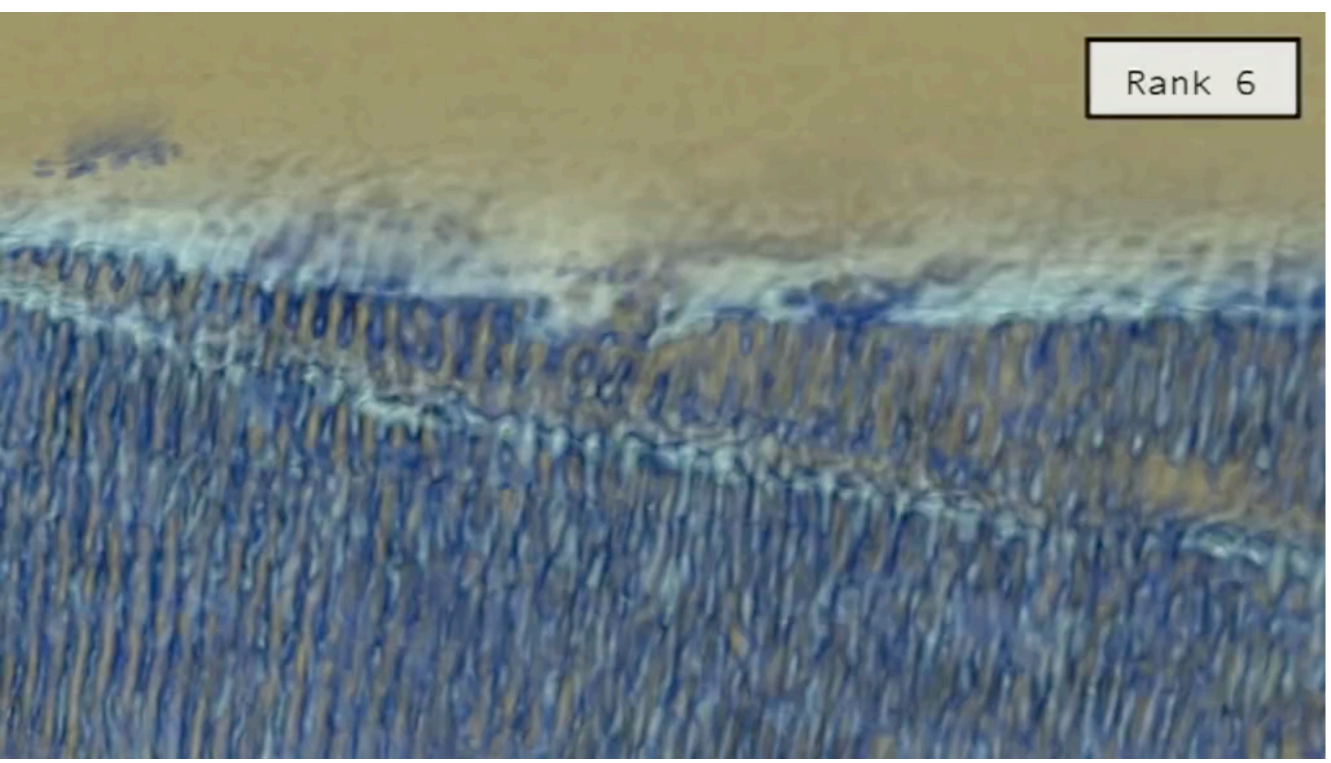


- Intel Core 2 E8500 3.2GHz Linux PC, 4GB memory
- NVIDIA GeForce GTX 480, 1.5GB memory

VISUALIZATIONAND MULTIMEDIALAB

great ape molar (17GB -> 5.5 GB)



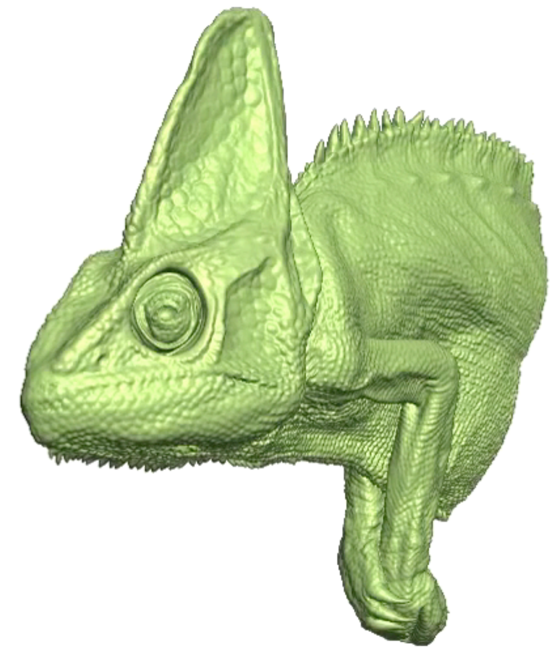


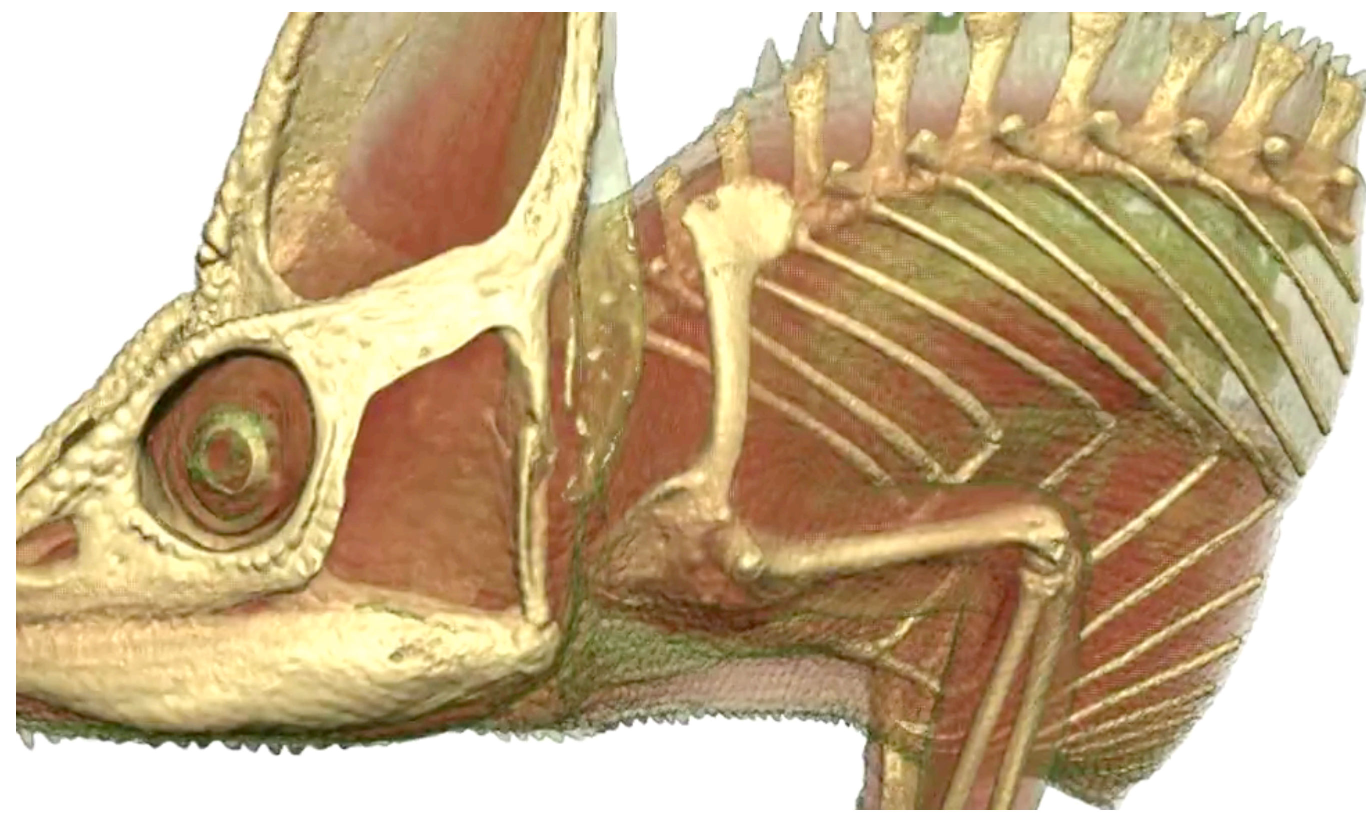
demo videos

http://www.youtube.com/user/VMMLuzh

[Suter et al., 2011]

chameleon (2 GB -> 230 MB





Conclusion

- Critical implementation steps
- Tensor decomposition
  - initial decomposition or a large input tensor
    - memory mapping
  - tensor times matrix (TTM) multiplications
    - parallel matrix matrix multiplications
  - higher-order SVD
- Tensor reconstruction
  - GPU implementation of TTM

Acknowledgments

- This work was supported by:
  - Forschungskredit of the University of Zürich
  - Swiss National Science Foundation (SNSF) grant project n °200021_132521
  - EU FP7 People Programme (Marie Curie Actions) REA Grant Agreement n°290227
  - All vmmlib collaborators, contributors and users
  - Computer-Assisted Paleoanthropology group at University of Zürich (acquisition of the test datasets)

VISUALIZATIONAND MULTIMEDIALAB