3D Image and Shape Compression summary

K. Makroleivaditis, Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr
Version 1.1.1

Artificial Intelligence &
Information Analysis Lab

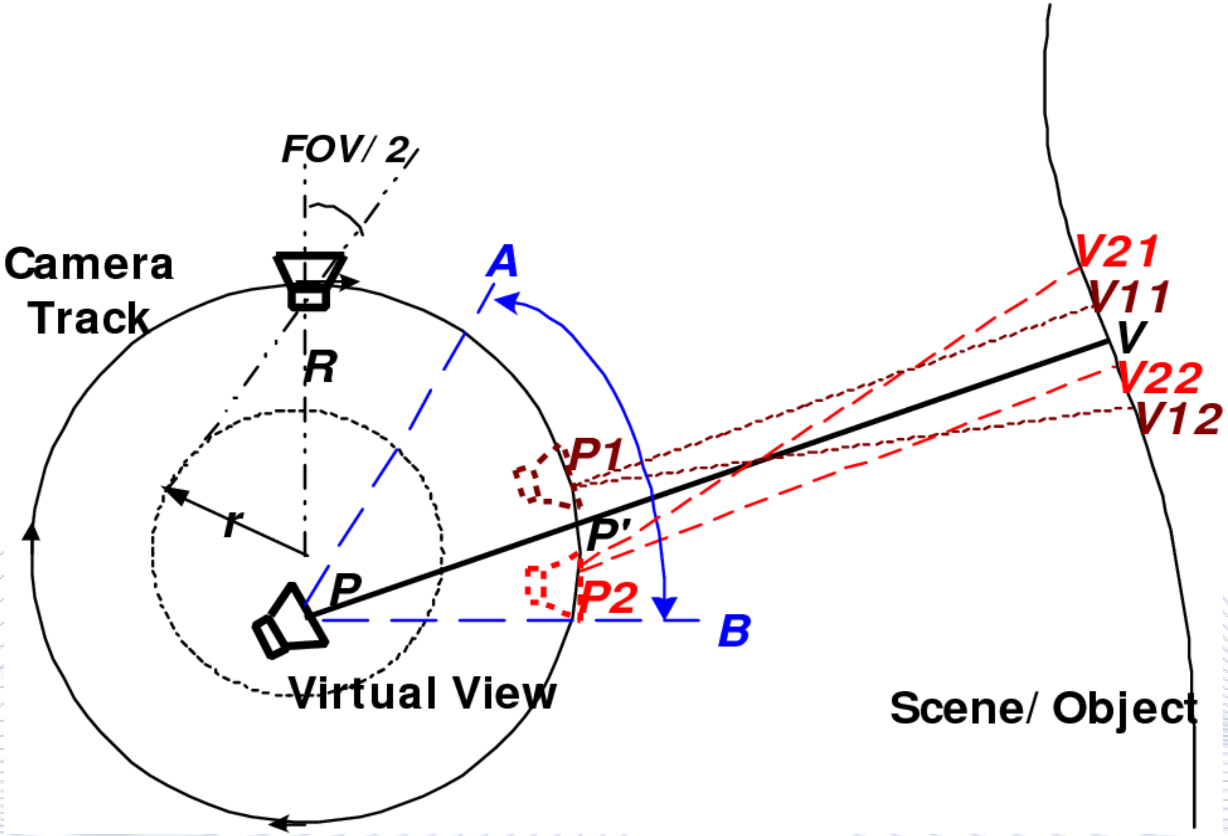
Introduction to 3D Image and Shape Compression

- **Image-based representations (IBR) compression**
- Depth map compression
- Layered depth image compression
- Model-based 3D shape compression
- Compression in MPEG-4
- Volumetric data compression

Image-based representations compression

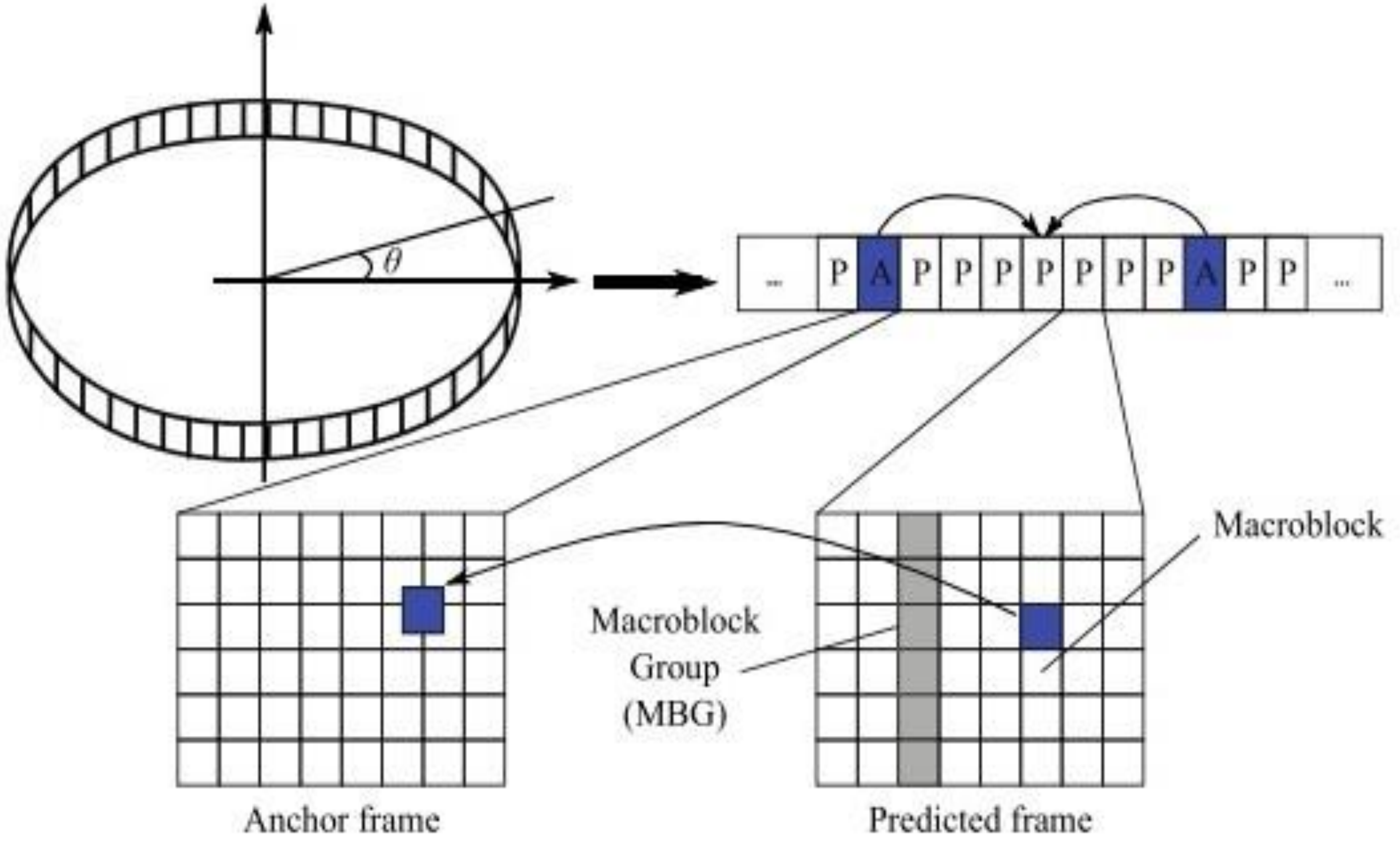
- Image-based representations (IBR) represent a three-dimensional object by using images rather than geometric shapes.
- In order to appropriately visualize a 3D object, IBR compression is mandatory due to the large amount of data that need to be stored.
- IBR data can be randomly accessed in order to generate the current view. Therefore, decoding only the specific bit-stream parts that are necessary for the current view, is the optimal adaption.
- Compression schemes for IBR prioritize fast random access.

Image-based representations compression



Concentric mosaic [SemanticScolar]

Image-based representations compression



Reference block coding structure

Artificial Intelligence & Information Analysis Lab

3D Image and Shape Compression

- Image-based representations (IBR) compression
- **Depth map compression**
- Layered depth image compression
- Model-based 3D shape compression
- Compression in MPEG-4
- Volumetric data compression
- DNA sequence compression

Depth Map Compression

- Depth images are composed by sharp edges and smooth regions.
- Depth map compression schemes function by subdividing the depth image into smooth or edge blocks.
- The algorithms issue rate distortion constraints, by utilizing the modeling function that diminishes the coding cost as much as possible.
- Furthermore, the addition of enhancement layers can produce an increased visual quality to the decoded depth map.

Depth Map Compression

(a) (b) (c) (d)

(a): Segmented depth map without depth reconstruction filter; (b): 3D object visualization without a reconstruction filter; (c): Segmented depth image with depth reconstruction filter; (d): 3D object visualization with reconstruction filter

3D Image and Shape Compression

- Image-based representations (IBR) compression
- Depth map compression
- **Layered depth image compression**
- Model-based 3D shape compression
- Compression in MPEG-4
- Volumetric data compression
- DNA sequence compression

Layered Depth Image Compression

- There is more than one depth layer for each pixel.
- A ***Layered Depth Image*** (LDI) is a two-dimensional array that keeps the layered depth pixels.
- LDI compression techniques focus on data redundancy in the various depth layers. Points of the same layer could correspond to various surface areas.
- Effective LDI compression schemes correlate pixels that use partial surfaces to correlate pixel data, rather than pixels which belong to the same layer.

3D Image and Shape Compression

- Image-based representations (IBR) compression
- Depth map compression
- Layered depth image compression
- **Model-based 3D shape compression**
- Compression in MPEG-4
- Volumetric data compression
- DNA sequence compression

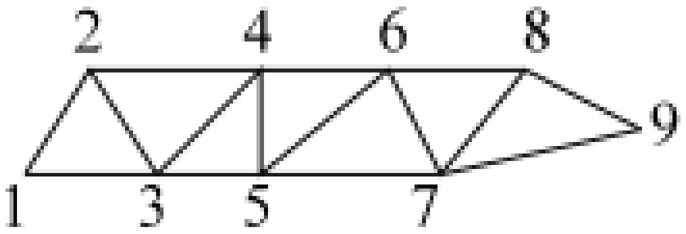
Model-based 3D shape compression

- *3D meshes* are the most well-known 3D shape representation models.
- Components required for 3D model representation:
  - ***Connectivity data***: represents how adjacent vertices are connected to each other.
  - ***Geometry data***: they define the position of the vertices.
  - ***Property data***: they define 3D shape attributes (colors, material reflectance etc.)
- In almost every algorithm, geometry data and property data are operated together.

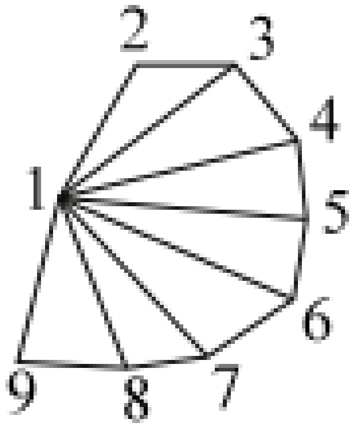
Model-based 3D shape compression

3D mesh of a human head

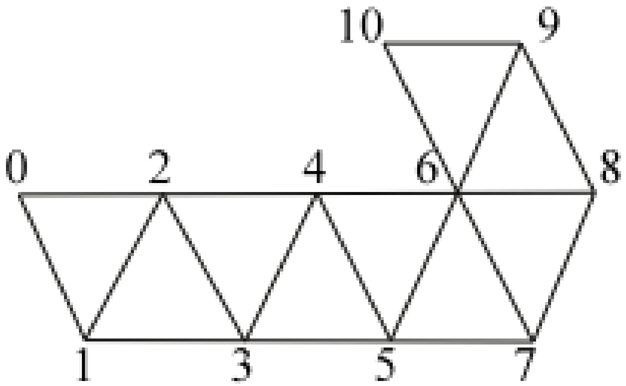
Model-based 3D shape compression



(a)



(b)



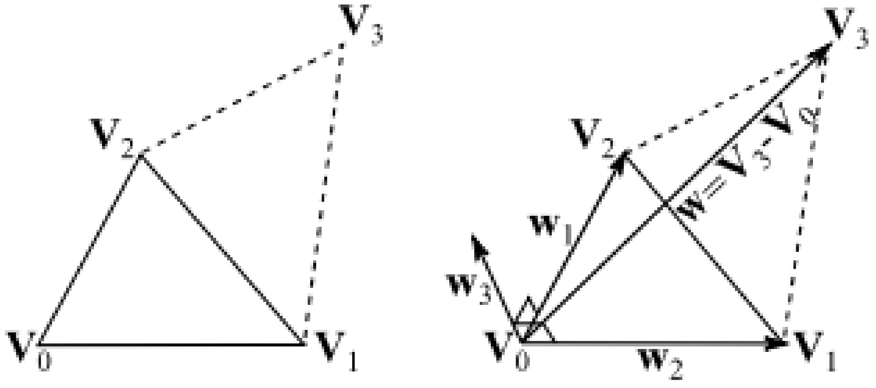
(c)

(a): Triangle strip ; (b): Triangle fan ; (c): Generalized triangle strip

Artificial Intelligence & Information Analysis Lab

Model-based 3D shape compression

- ***Vector quantization*** is a scheme applied in geometry information compression. It implements the *Cartesian-to-model-space* transformation.



Cartesian-to-model-space transformation

Model-based 3D shape compression

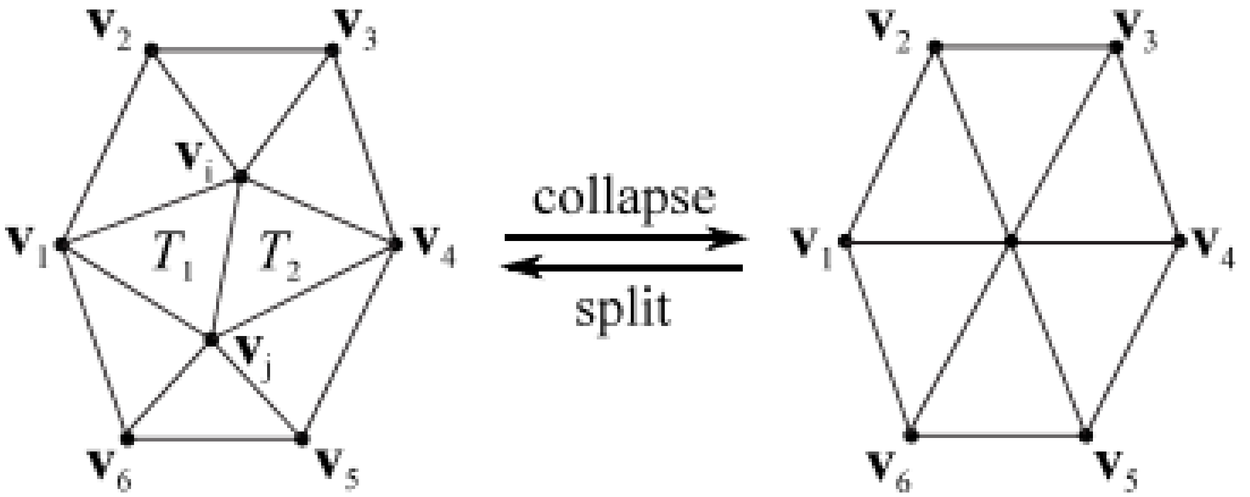
- To transform the vertex $\mathbf{V}_3$ into a model space vector, the following vectors need to be defined:
  - $\mathbf{w} = \mathbf{V}_3 - \mathbf{V}_0$
  - $\mathbf{w}_1 = \mathbf{V}_1 - \mathbf{V}_0$
  - $\mathbf{w}_2 = \mathbf{V}_2 - \mathbf{V}_0$
  - $\mathbf{w}_3 = (\mathbf{w}_1 \times \mathbf{w}_2)/sin\theta(\sqrt{(\mathbf{w}_1 \cdot \mathbf{w}_2)})$,
- θ is the angle between $\mathbf{w}_1$ and $\mathbf{w}_2$.

Model-based 3D shape compression

Progressive compression (neck bones)

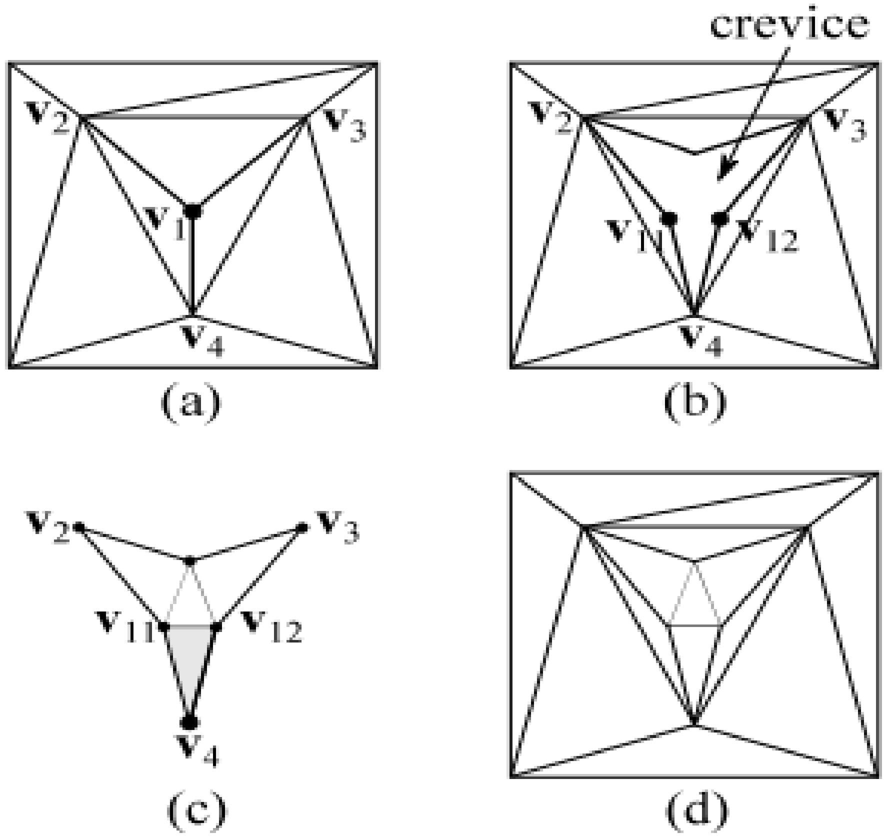
Model-based 3D shape compression

- ***Progressive meshes*** are built on the idea of simplifying a mesh by collapsing an edge.



Edge collapse operation

Model-based 3D shape compression

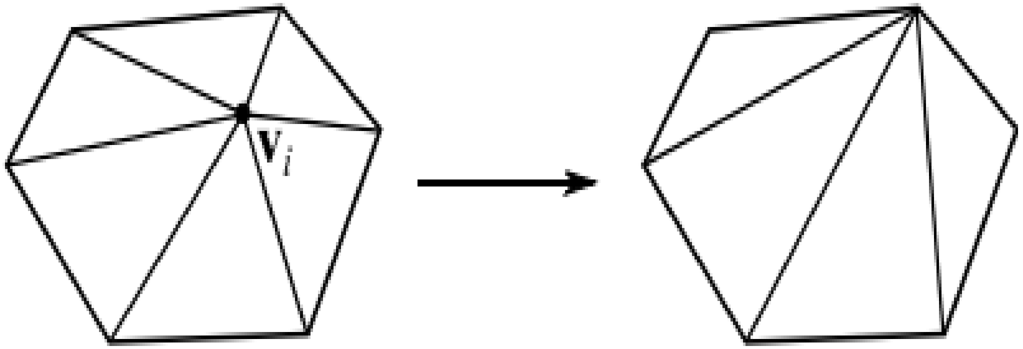


Progressive forest split

Artificial Intelligence & Information Analysis Lab

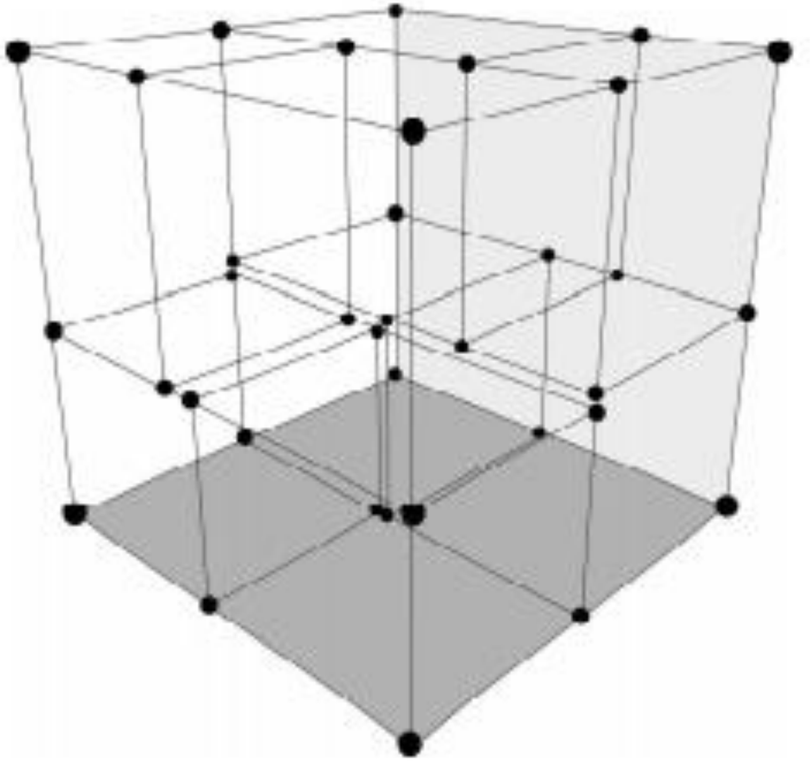
Model-based 3D shape compression

- The ***patch coloring*** scheme utilizes vertex decimation in order to simplify the original mesh. After a vertex and its edges are removed from the mesh, the blank space is filled with new triangles.



Vertex decimation and re-triangulation

Model-based 3D shape compression

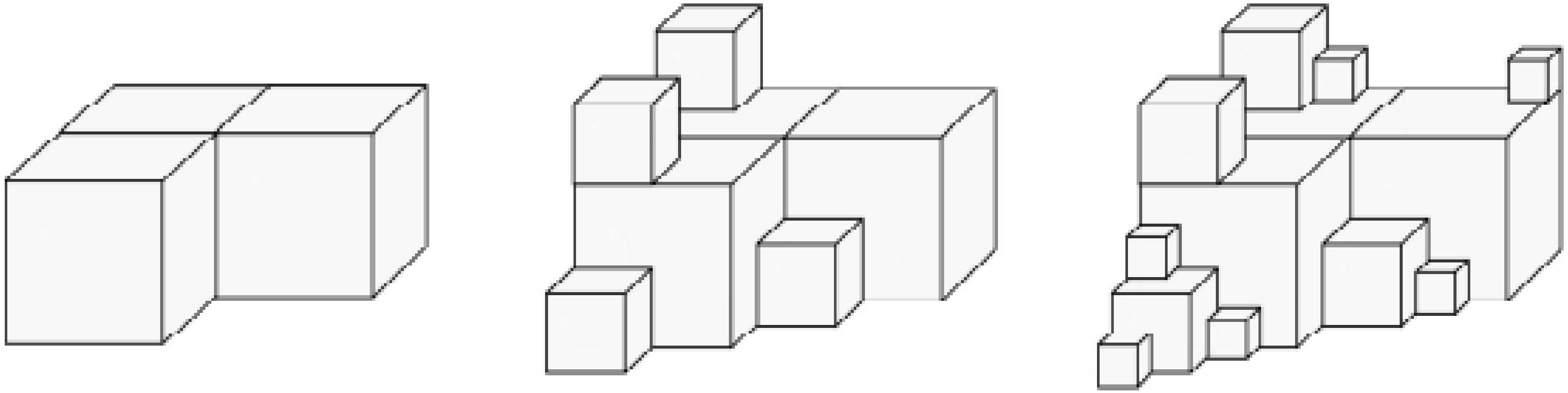


kd-tree division

Model-based 3D shape compression

- ***Octree decomposition*** can be implemented for the compression of arbitrary topology 3D shapes.
- In this technique, the model is divided into cubes of various sizes and an octree is obtained, which provides geometry and connectivity information.
- A great benefit that this scheme provides, is that the encoding process happens only if a cube is not empty, rather than considering the amount of vertices that each cube has.

Model-based 3D shape compression



Progressive compression by octree decomposition

Model-based 3D shape compression

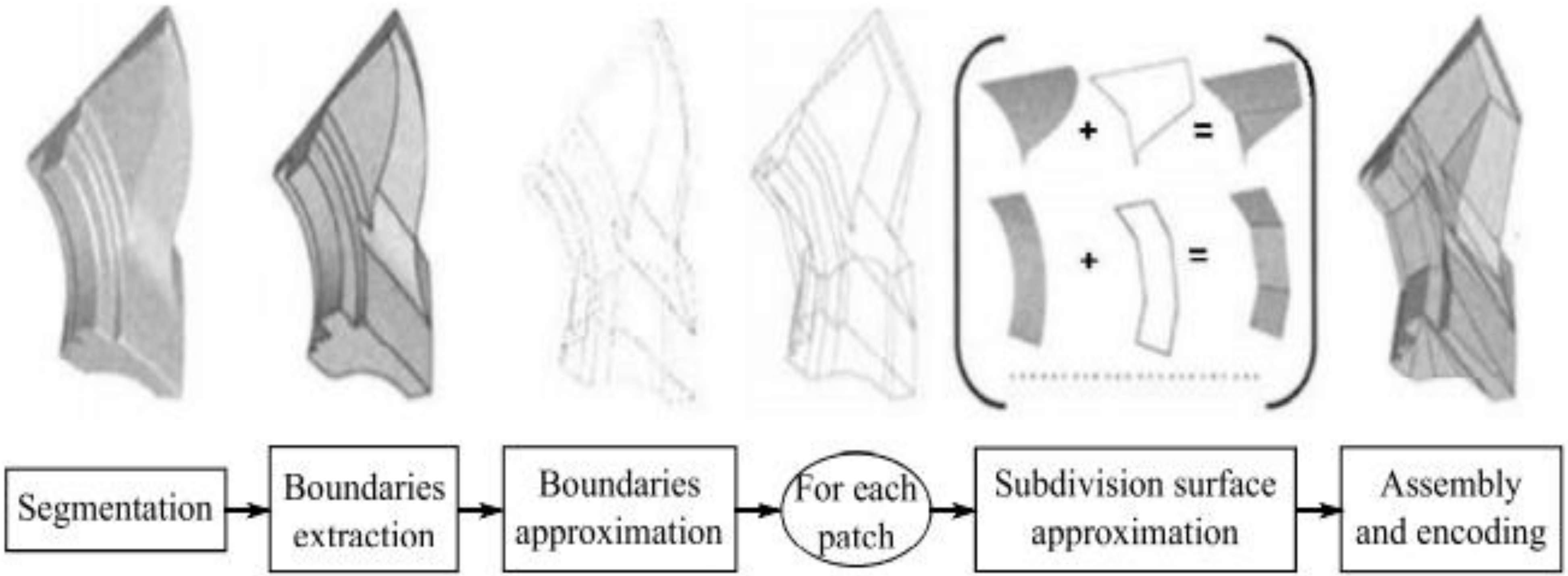
- ***Wavelet transform*** is another method for progressive mesh compression, that focuses on geometry data encoding.
- In this technique, the original arbitrary mesh is transformed to a *semiregular mesh* (every vertex has a valence value of 6).
- At the refinement state, each triangle in the coarse mesh is divided into four triangles. This scheme is called *quadrisection*.
- By implementing the wavelet transform, geometry data can be expressed with the coarse mesh and the wavelet coefficients.

Model-based 3D shape compression

- ***Polygonal mesh compression*** requires a different approach from the previous mentioned algorithms. Triangle mesh compression schemes can be implemented, but in this case that might lead to connectivity data loss.

- For the compression of manifold polygonal meshes, polygonal connectivity is represented by vertex valence and face degree. Low-valence vertices are surrounded by high-degree faces (and vice versa).

Model-based 3D shape compression



Overview of the subdivision surface compression framework

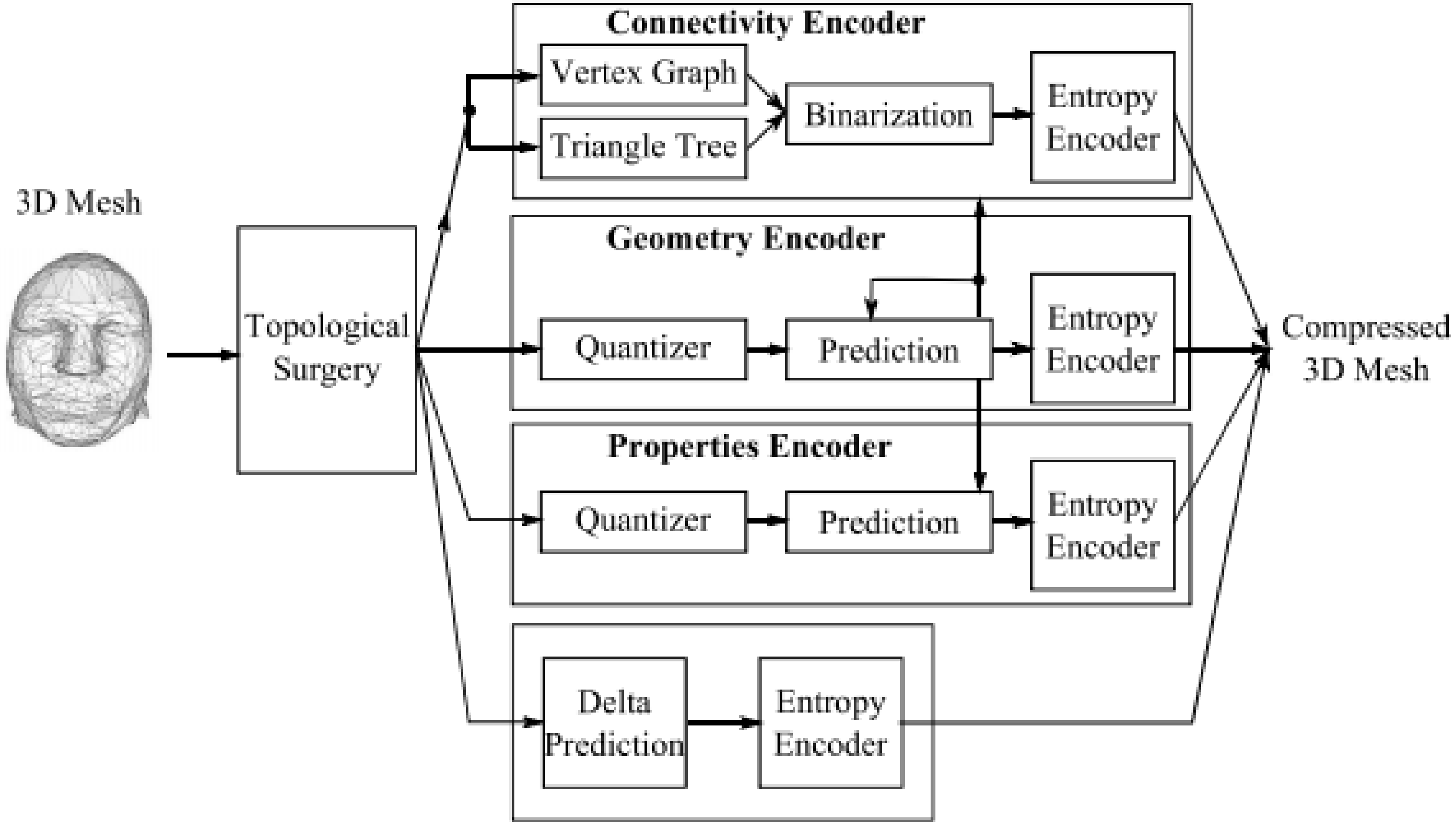
3D Image and Shape Compression

- Image-based representations (IBR) compression
- Depth map compression
- Layered depth image compression
- Model-based 3D shape compression
- **Compression in MPEG-4**
- Volumetric data compression
- DNA sequence compression

Compression in MPEG-4

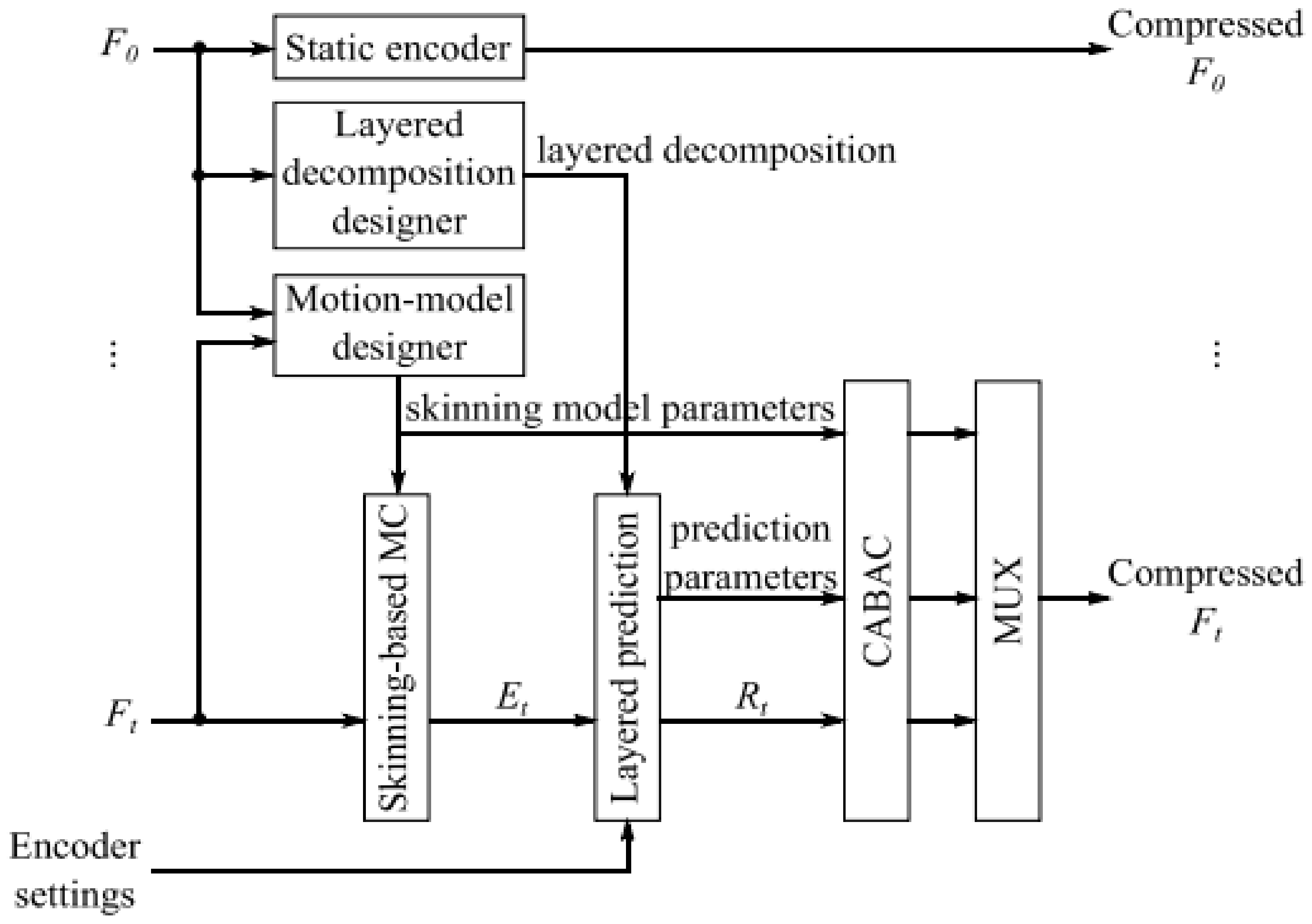
- ***3D mesh coding*** (3DMC) is an MPEG-4 technique that compresses 3D meshes without declining the output visual quality.
- Connectivity, geometry and property data are compressed.
- Visual quality can be upgraded by the *3DMC-X* extension, where the handling of texture coordinates is better structured.
- ***Differential 3DMC*** (D3DMC) is an extension that is suitable for dynamic 3D mesh compression.
- In D3DMC, the encoder uses the previously decoded mesh in order to predict the current positions of the vertex.

Compression in MPEG-4



Block diagram of 3DMC encoder.

Compression in MPEG-4

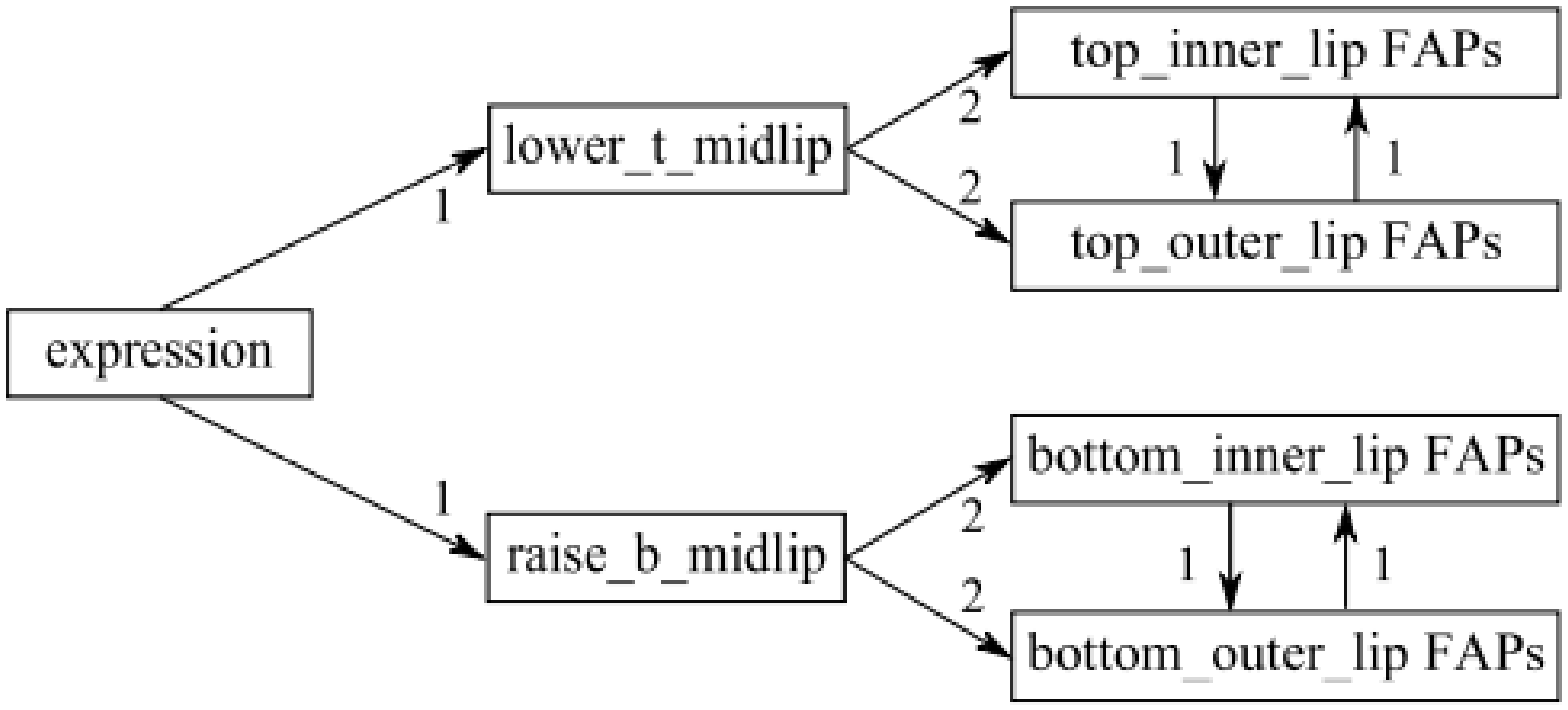


FAMC encoder

Compression in MPEG-4

- ***Facial Definition Parameters*** (FDP) describe facial geometry.
- FDP data are implemented in initializing the geometric model, precisely defining the facial shape and skin texture.
- ***Facial Animation Parameters*** (FAP) are transmitted in order to deform the facial model. They describe facial animation.
- Dispatching a part of the FAP data, and using an interpolation technique to compute the rest, is an efficient way to diminish the amount of FAP data.

Compression in MPEG-4



Facial Interpolation Graph for the interpolation of lip-FAPs used in FAP

Compression in MPEG-4

- ***Body Definition Parameters*** (BDP) determine body geometry.
- ***Body Animation Parameters*** (BAP) determine body motion.
- Usually, a mesh of polygons defines each body part, which can be compressed by utilizing aforementioned compression schemes.
- ***BAP sparsing*** is an MPEG-4 compression technique. It utilizes a *BAP sparsing layer*, an enhancement layer that suspends BAPs and releases them after less than one second. Note that this is unperceivable by the human eye.

3D Image and Shape Compression

- Image-based representations (IBR) compression
- Depth map compression
- Layered depth image compression
- Model-based 3D shape compression
- Compression in MPEG-4
- **Volumetric data compression**
- DNA sequence compression

Volumetric data compression

- Volumetric data are extensively used in environmental and medical imaging.

- Volumetric data compression results in a reduction to the number of bits per voxel.

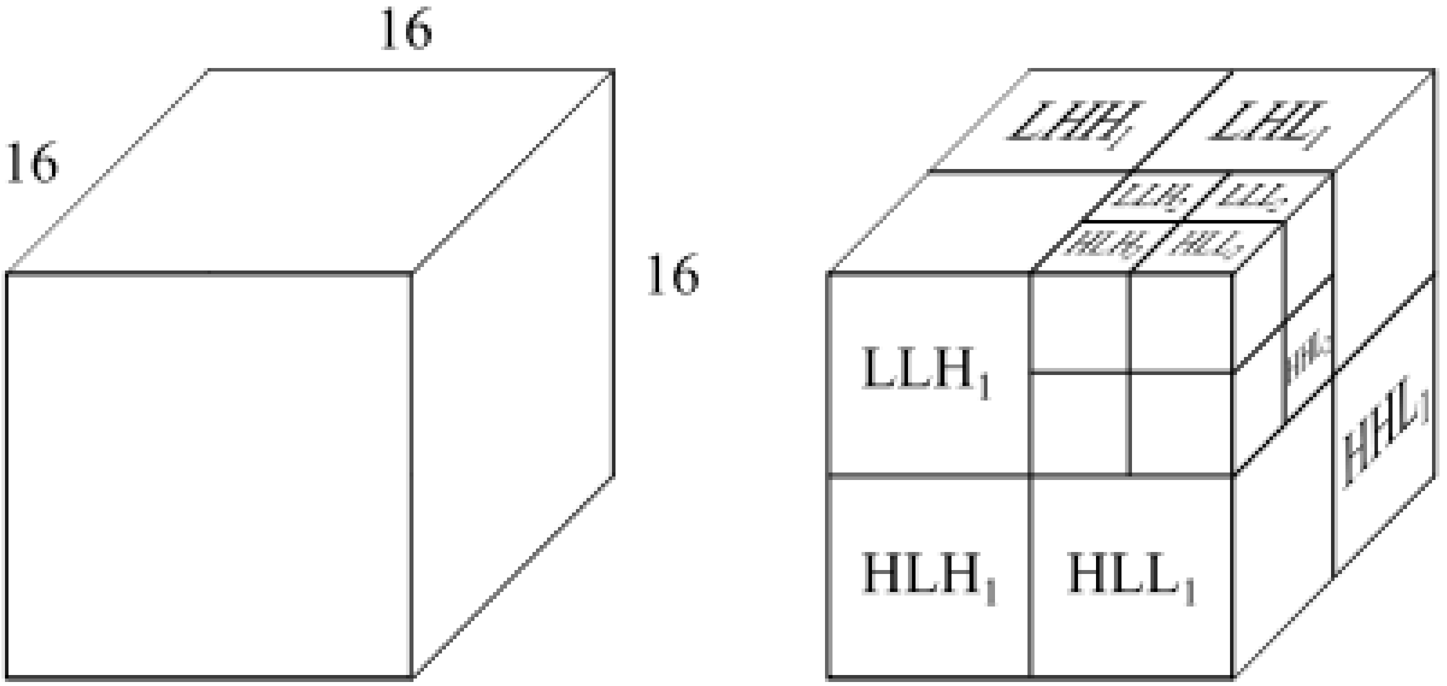
- The *Haar wavelet transform* converts a data sequence to a new one, whose length is half of the original one, by calculating the average value between two ensuing data samples.

- The 3D volume data are broken down into blocks of size $16 \times 16 \times 16$ voxels, the ***unit blocks***.

Volumetric data compression

- Afterwards, the unit block is recursively split into 8 subblocks of size $8 \times 8 \times 8$ voxels, each of which is subdivided in 8 *cells* of size $4 \times 4 \times 4$ voxels.
- ***H*** stands for ***High*** pass filtering, ***L*** stands for ***Low*** pass filtering. For instance, HHL applies two high-pass filters in succession, and then the low-pass one.
- The algorithm maintains only the wavelet coefficients that are greater than a specific threshold (the rest are zeroed).
- Zeroed cells are labeled by a zero value. On the contrary, non-zeroed cells are labeled by a positive integer in an increasing order.

Volumetric data compression



(a) (b)

(a): A $16 \times 16 \times 16$ unit block; (b):The same unit block after the wavelet transform

Bibliography

[NIK2000] N. Nikolaidis and I. Pitas, 3D Image Processing Algorithms, J. Wiley, 2000.

[PIT2021] I. Pitas, "Computer vision", Createspace/Amazon, in press.

[PIT2017] I. Pitas, "Digital video processing and analysis " , China Machine Press, 2017 (in Chinese).

[PIT2013] I. Pitas, "Digital Video and Television " , Createspace/Amazon, 2013.

[PIT2000] I. Pitas, Digital Image Processing Algorithms and Applications, J. Wiley, 2000.

[SZE2011] R.Szelinski, " Computer Vision " , Springer 2011

[DAV2017] Davies, E. Roy. "Computer vision: principles, algorithms, applications, learning ". Academic Press, 2017

[TRU1998] Trucco E, Verri A. "Introductory techniques for 3-D computer vision", Prentice Hall, 1998.

[GEO] http://geozone.ch/en/visionlidar-en

Q & A

Thank you very much for your attention!

More material in
http://icarus.csd.auth.gr/cvml-web-lecture-series/

Contact: Prof. I. Pitas
pitas@csd.auth.gr