Joint Depth/Texture Bit-Allocation For Multi-View Video Compression

Yannick Morvan[1], Dirk Farin[1] and Peter H. N. de With[1,2]

[1]Video Coding and Architectures research group
Signal Processing Systems
Eindhoven University of Technology, The Netherlands.

[2]LogicaCMG, Eindhoven.

Picture coding Symposium 2007

TU/e technische universiteit eindhoven

PHILIPS

Introduction to multi-view video

- Two possible applications for multi-view video:
  - 3D TV: present depth using a multi-view display.
  - free-viewpoint video: interactively change the viewpoint.

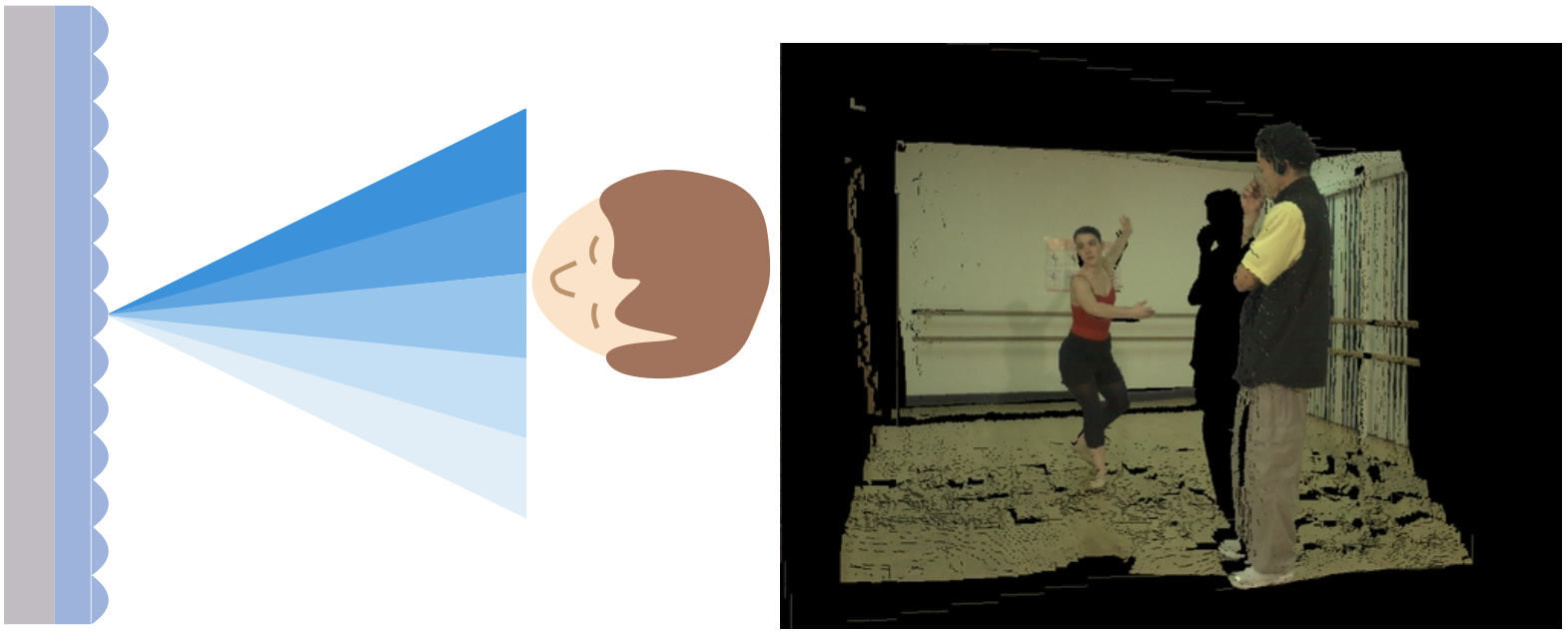
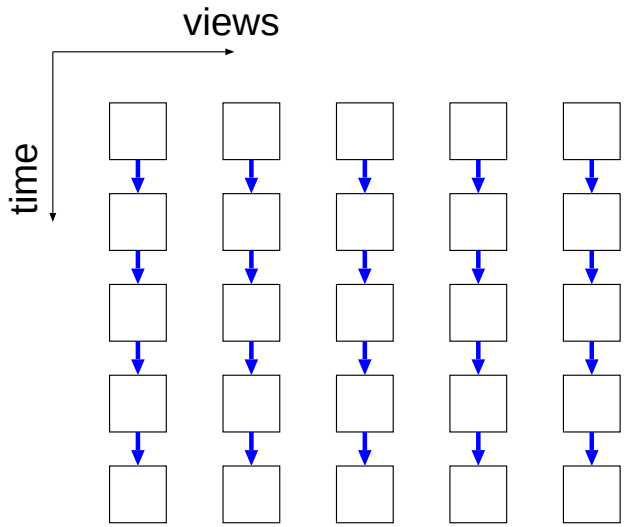


Figure: 3D TV and free-viewpoint video

TU/e

Compression of multi-view video 1/2

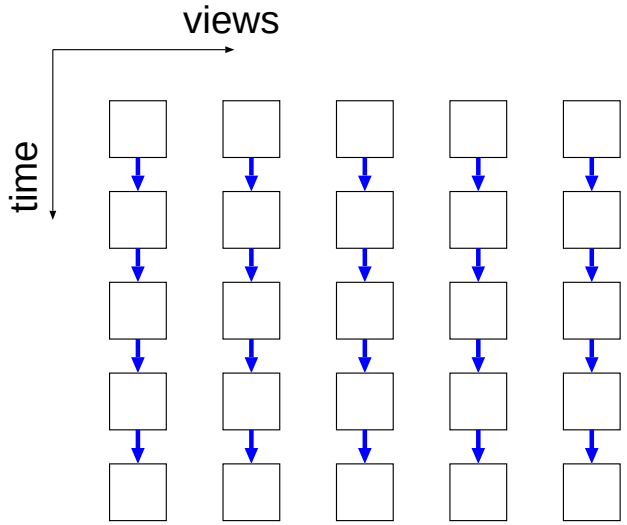
- Simple approach: encode each view independently (Simulcast coding).
- Exploits only the redundancy between consecutive frames (temporal prediction).



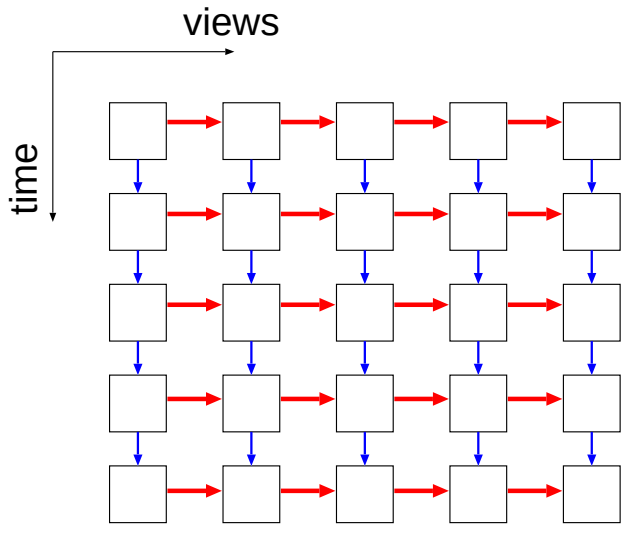
TU/e

Compression of multi-view video 1/2

- Simple approach: encode each view independently (Simulcast coding).
- Exploits only the redundancy between consecutive frames (temporal prediction).



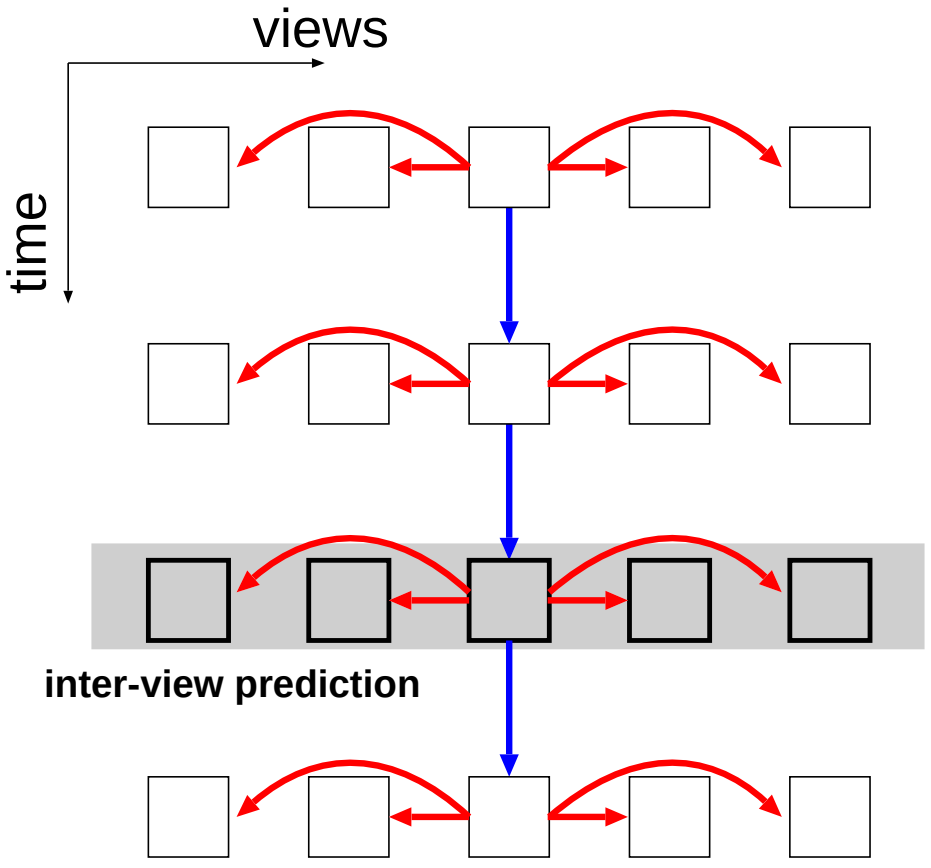
- More efficient compression: also exploit the inter-view correlation.



TU/e

Compression of multi-view video 2/2

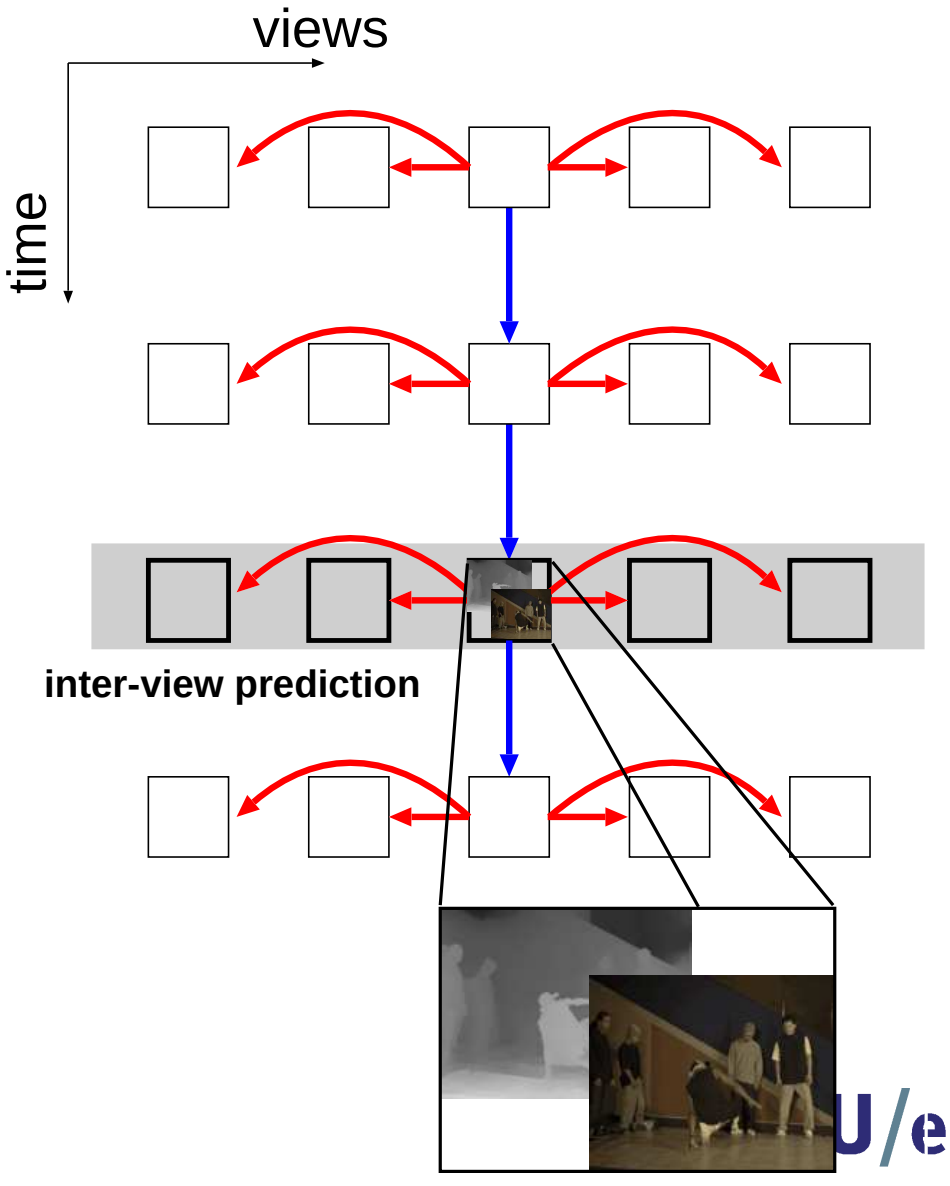
- Neighboring views can be predicted using
  1. disparity estimation
  2. view synthesis.
- View synthesis with Depth Image Based Rendering.



TU/e

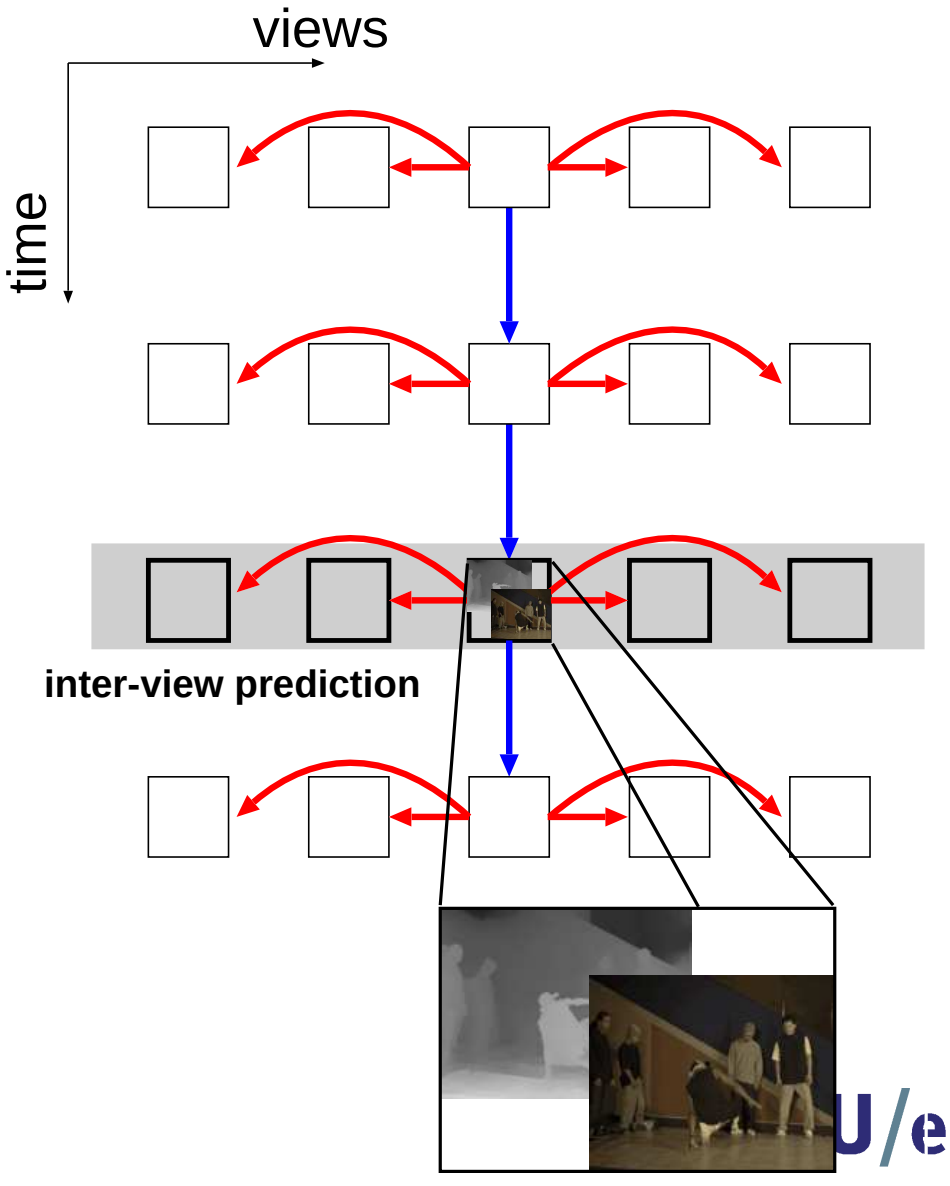
Compression of multi-view video 2/2

- Neighboring views can be predicted using
  1. disparity estimation
  2. view synthesis.
- View synthesis with Depth Image Based Rendering.
- $\Rightarrow$ coding algorithm relies on depth and texture.



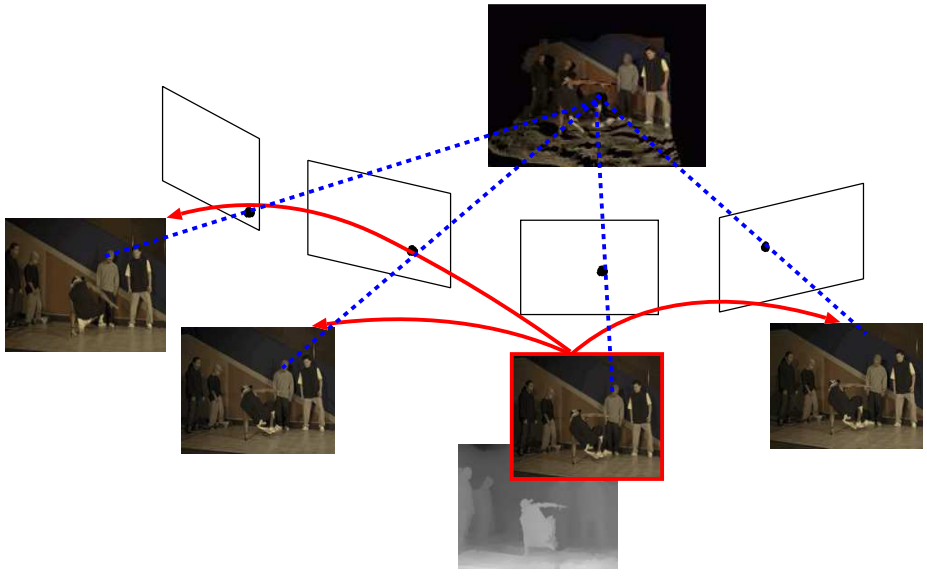
Compression of multi-view video 2/2

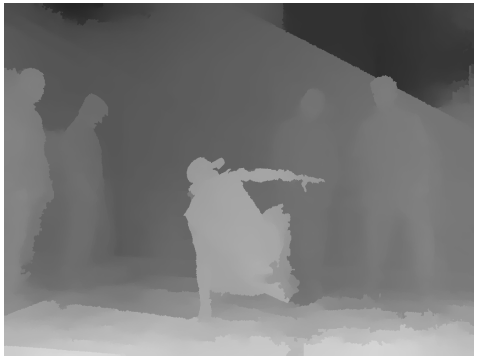
- Neighboring views can be predicted using
  1. disparity estimation
  2. view synthesis.
- View synthesis with Depth Image Based Rendering.
- $\Rightarrow$ coding algorithm relies on depth and texture.
- Coding experiments on view-synthesis prediction are currently performed within MPEG MVC.



View-synthesis using one depth+texture

- Coding improvement is not the only reason to use depth images.
- Further advantages of depth+texture representation:
  1. multiple views can be predicted from a single depth+texture image
  2. rendering of arbitrary views

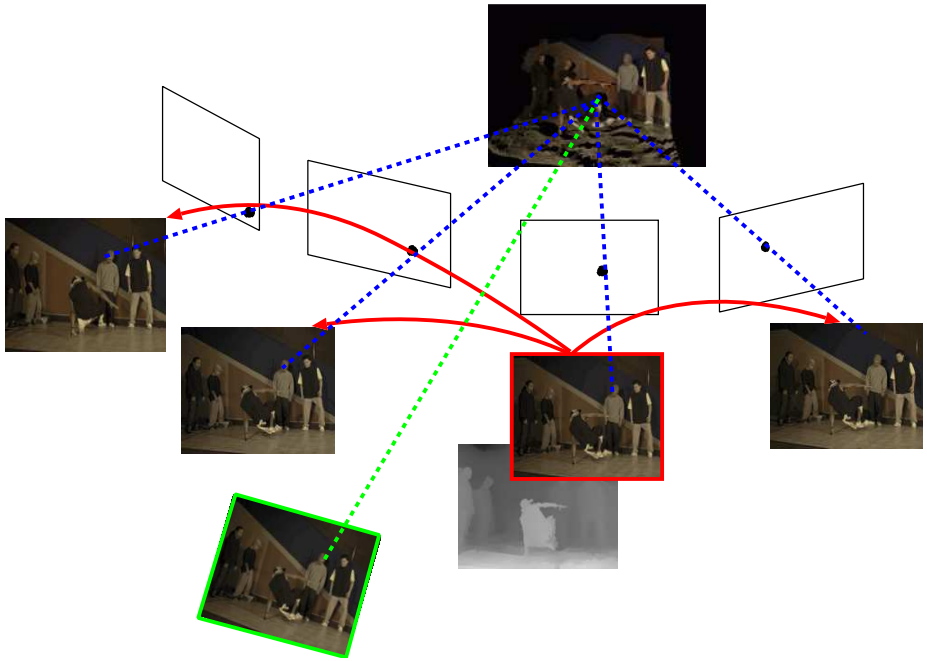


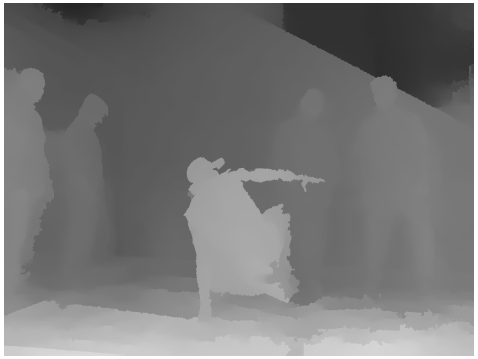


TU/e

View-synthesis using one depth+texture

- Coding improvement is not the only reason to use depth images.
- Further advantages of depth+texture representation:
  1. multiple views can be predicted from a single depth+texture image
  2. rendering of arbitrary views

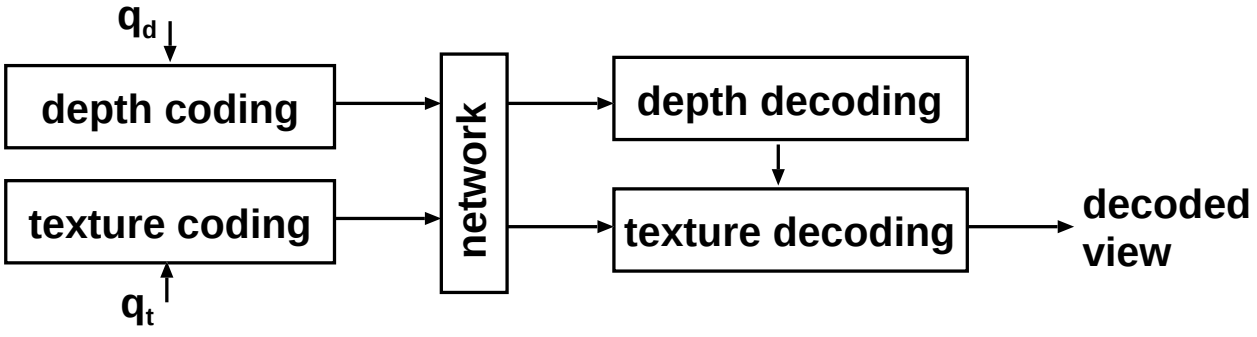




TU/e

Depth-based view prediction 1/2

- Depth and texture data are compressed and transmitted independently.
- MPEG MVC suggests that depth data can be coarsely quantized: the depth bit-rate corresponds to 10%-25% of the texture bit-rate.
- Currently, the impact of depth quantization on the quality of view synthesis is not understood. 

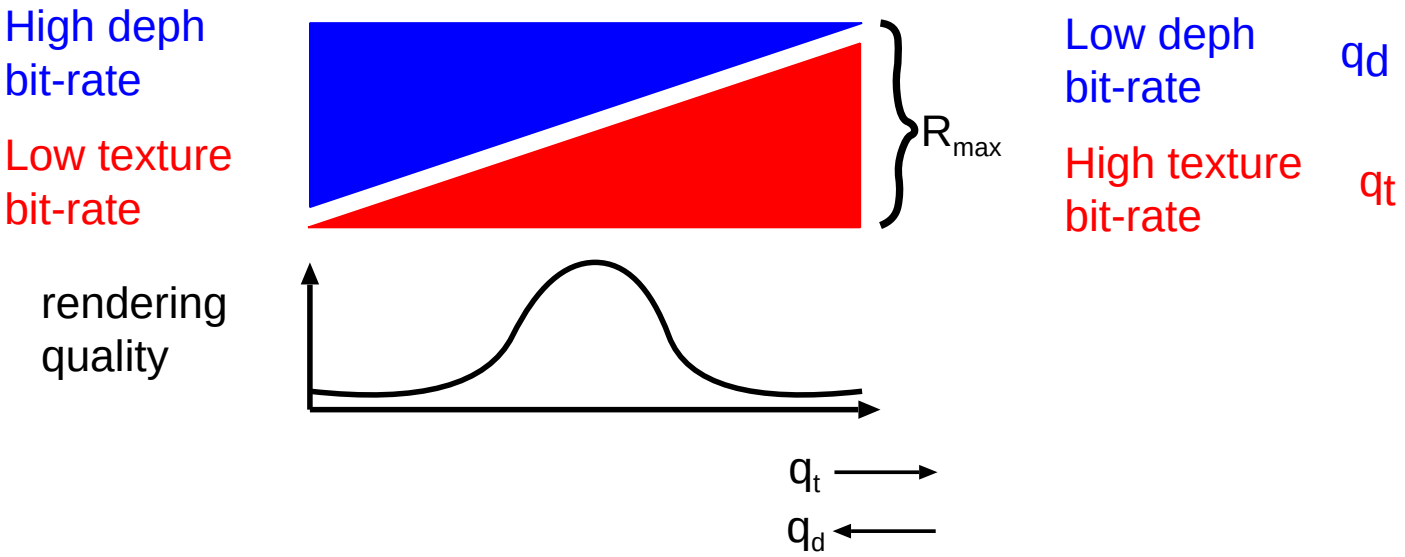


TU/e

Depth-based view prediction 2/2

Problem statement

Determine the quantization parameters q_t and q_d so that the view-prediction accuracy is maximized for a bit-rate budget R_{max}.



Joint bit-allocation of depth and texture

joint rate and rendering function

- Depth and texture data can be coded using two quantization parameters q_t and q_d.
- We define a joint rate function as:

$$R_{max}(q_t, q_d) = R_t(q_t) + R_d(q_d),$$

- and a joint rendering function as:

$$D_{render}(q_t, q_d)$$

TU/e

Optimization problem

joint bit-allocation problem

- The optimization can be formulated as

$$(q_t^{opt}, q_d^{opt}) = \underset{q_d, q_t \in Q}{\arg\min}\, D_{render}(q_t, q_d),$$

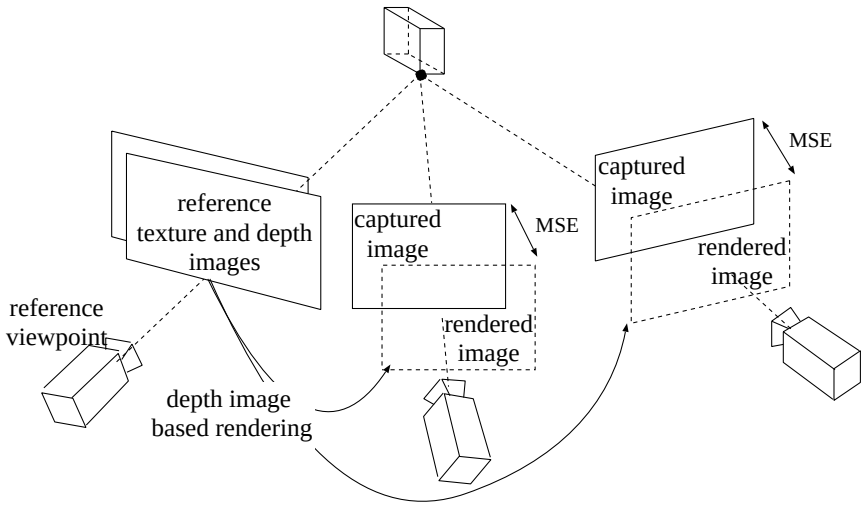
under the constraint that

$$R_t(q_t^{opt}) + R_d(q_d^{opt}) \leq R_{max}$$

TU/e

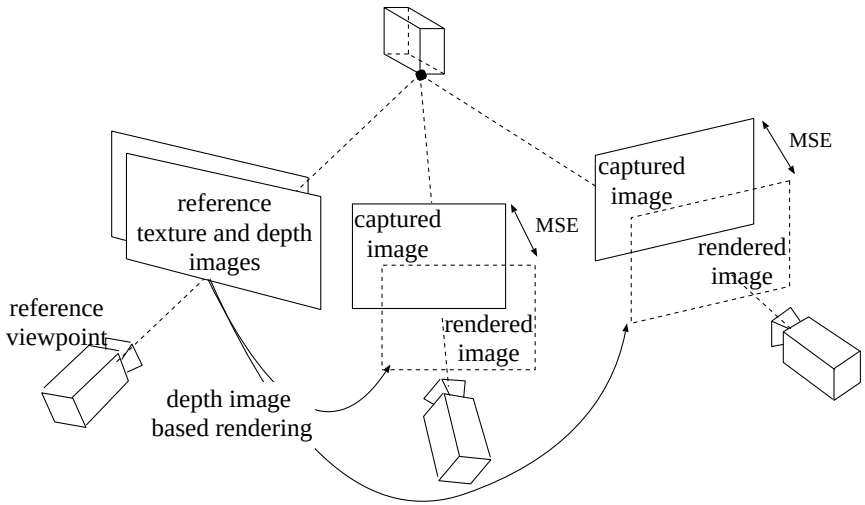
Definition of rendering distortion

- To measure the rendering distortion $D_{render}(q_t, q_d)$:
  1. code a texture and depth image using quantizers (q_t, q_d),



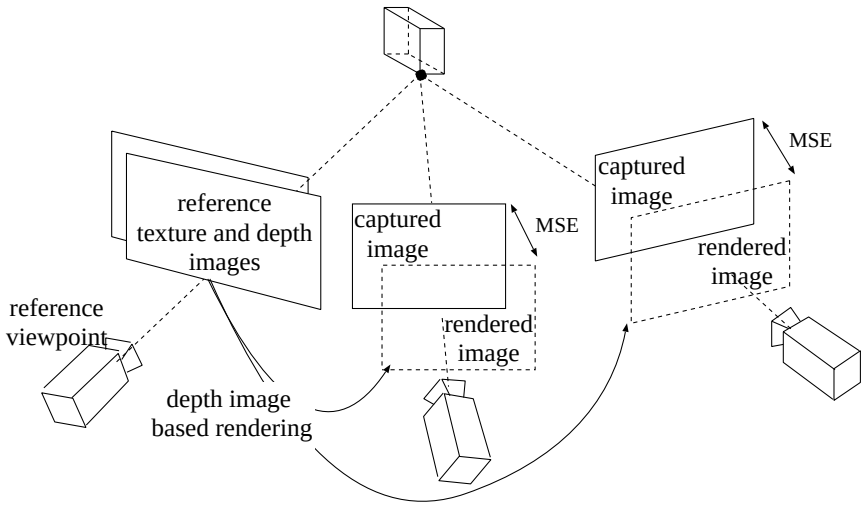
Definition of rendering distortion

- To measure the rendering distortion $D_{render}(q_t, q_d)$:
  1. code a texture and depth image using quantizers (q_t, q_d),
  2. render image at the position of the predicted view,



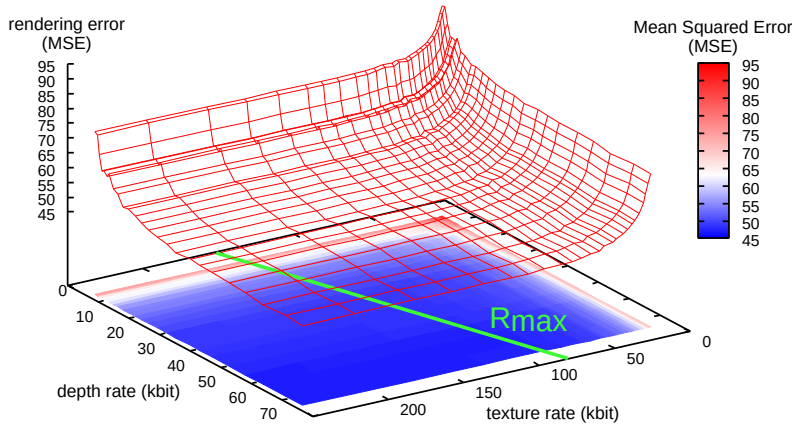
Definition of rendering distortion

- To measure the rendering distortion $D_{render}(q_t, q_d)$:
  1. code a texture and depth image using quantizers (q_t, q_d),
  2. render image at the position of the predicted view,
  3. calculate prediction error (MSE).

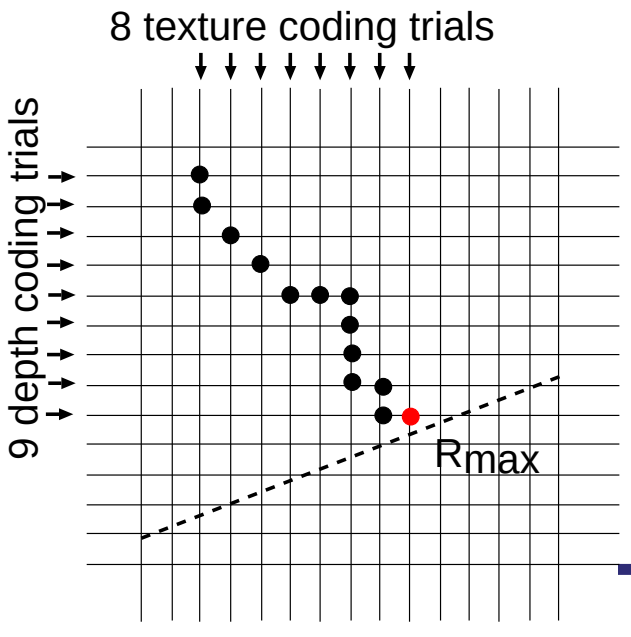


R-D surface analysis

- $D_{render}(q_t, q_d)$ is a function of 2 parameters (q_t, q_d): R-D surface.
- The R-D surface generation requires $2 \times k$ coding trials. (k q_t, k q_d).



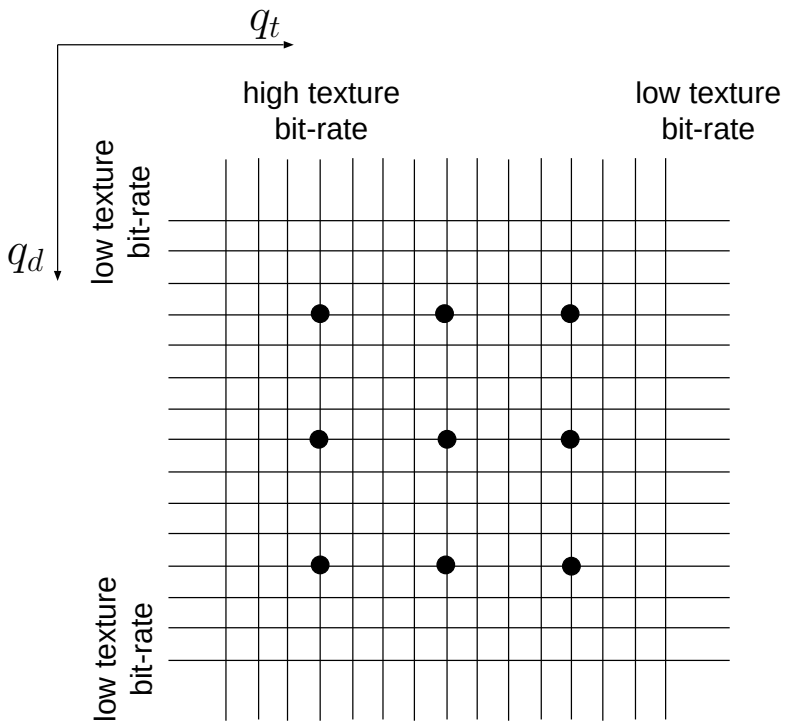
- Coding trial is computationally expensive.
- Using less measurement points does not necessarily reduce the number of coding trials.
- For example, gradient descent has comparable complexity.



TU/e

Hierarchical optimization 1/2

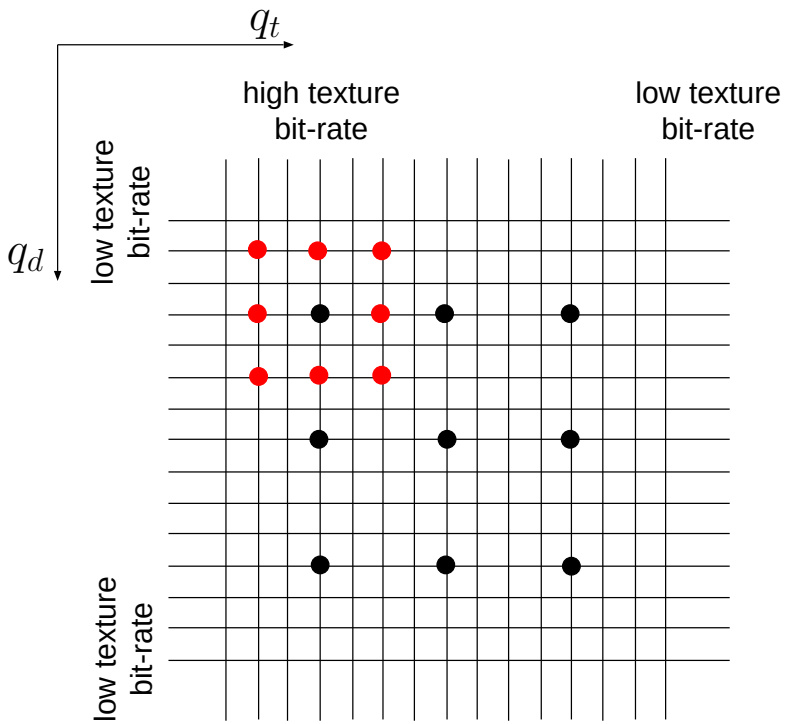
- We propose optimization with less coding trials.
- Perform a coarse-to-fine search of appropriate quantizers:
  1. compute coarsely spaced candidates (q_t, q_d),
  2. select the candidate with lowest rendering distortion below bit-rate budget R_{max},
  3. reduce the search range and perform the search recursively.
- The approach is similar to a *three-step* search in motion-estimation.



TU/e

Hierarchical optimization 1/2

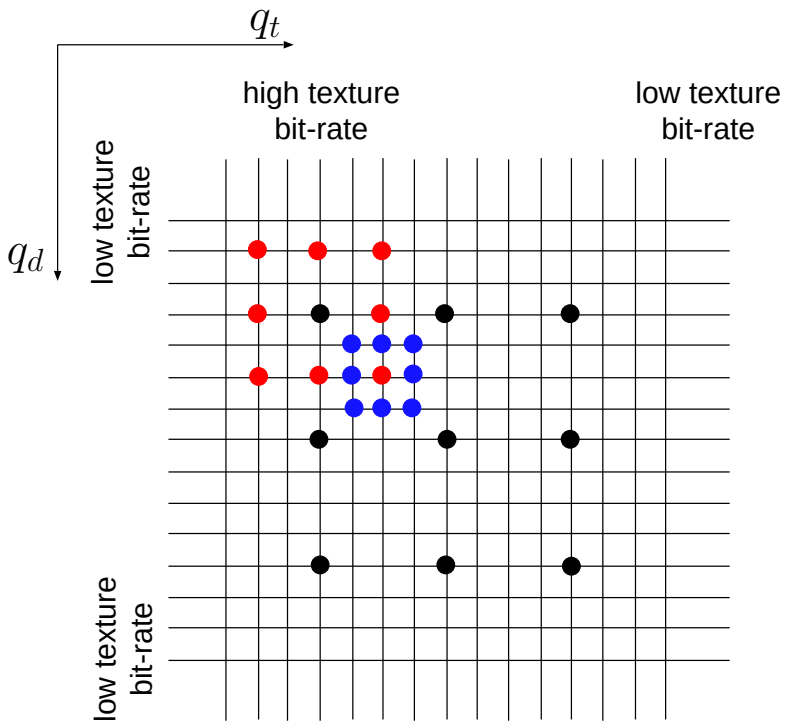
- We propose optimization with less coding trials.
- Perform a coarse-to-fine search of appropriate quantizers:
  1. compute coarsely spaced candidates (q_t, q_d),
  2. select the candidate with lowest rendering distortion below bit-rate budget R_{max},
  3. reduce the search range and perform the search recursively.
- The approach is similar to a *three-step* search in motion-estimation.



TU/e

Hierarchical optimization 1/2

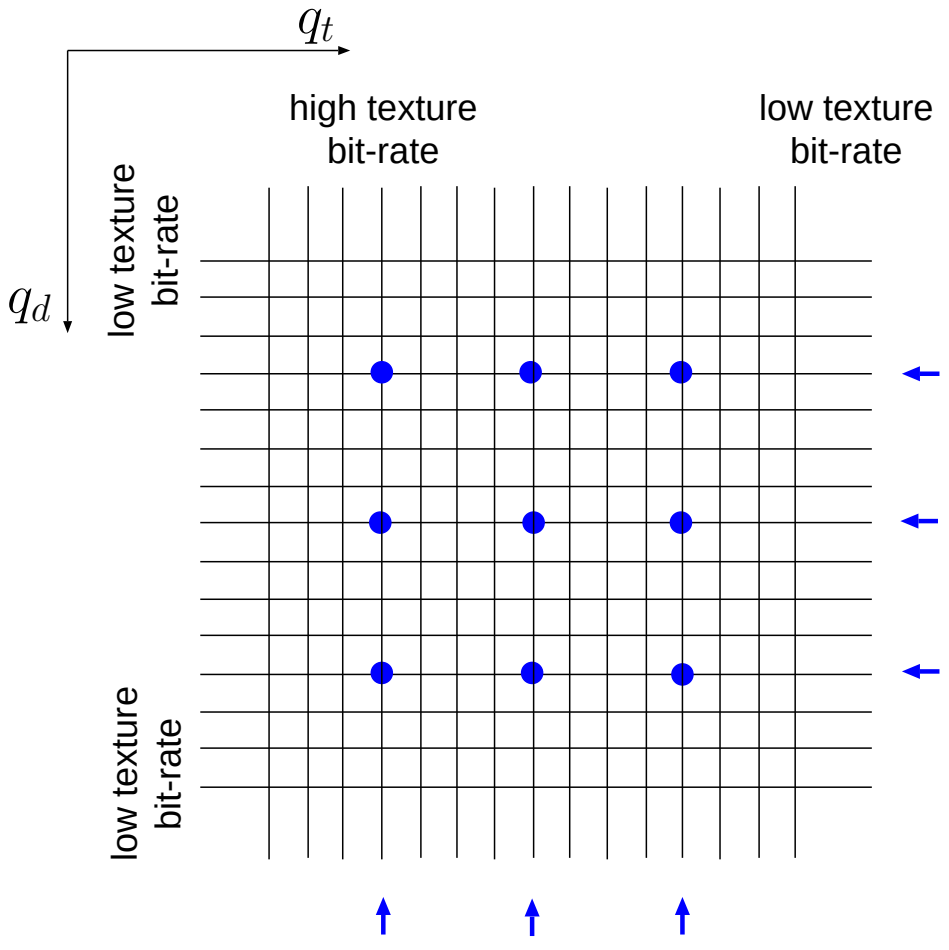
- We propose optimization with less coding trials.
- Perform a coarse-to-fine search of appropriate quantizers:
  1. compute coarsely spaced candidates (q_t, q_d),
  2. select the candidate with lowest rendering distortion below bit-rate budget R_{max},
  3. reduce the search range and perform the search recursively.
- The approach is similar to a *three-step* search in motion-estimation.



TU/e

Hierarchical optimization 2/2

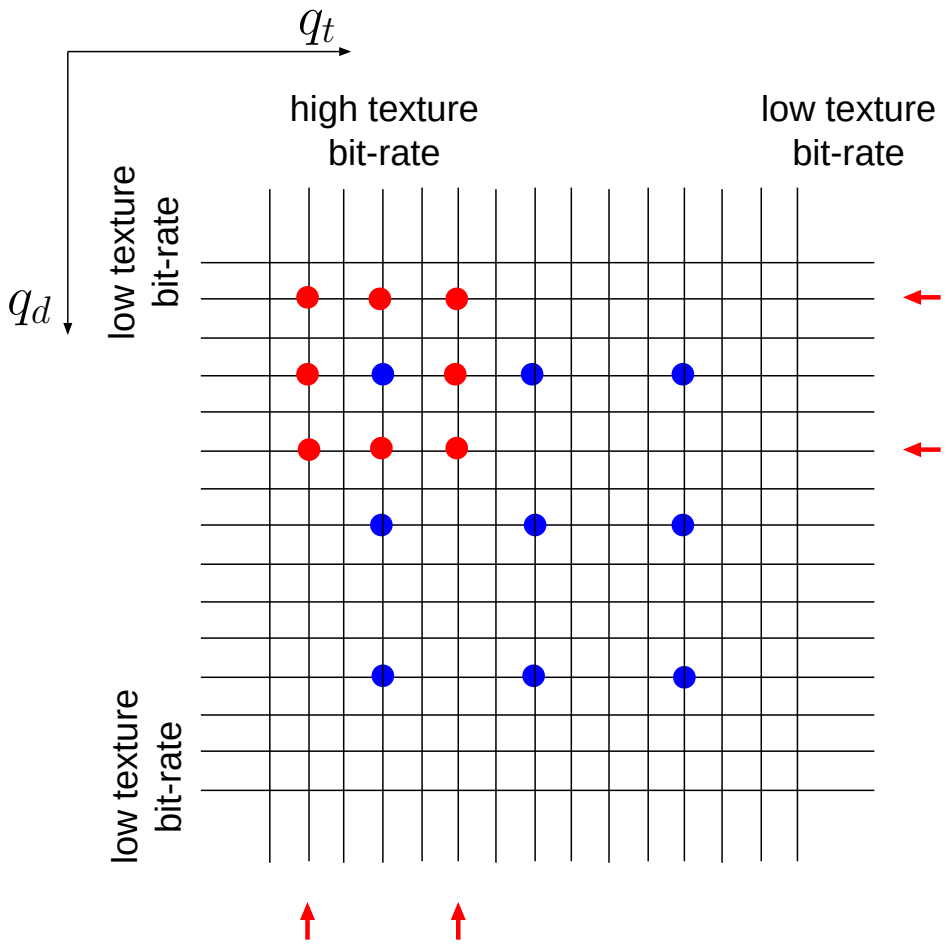
- Hierarchical search reduces the computational complexity.
- Reduced number of coding trials:
  - $3 + 3$ trials $\Rightarrow$ 9 R-D points



TU/e

Hierarchical optimization 2/2

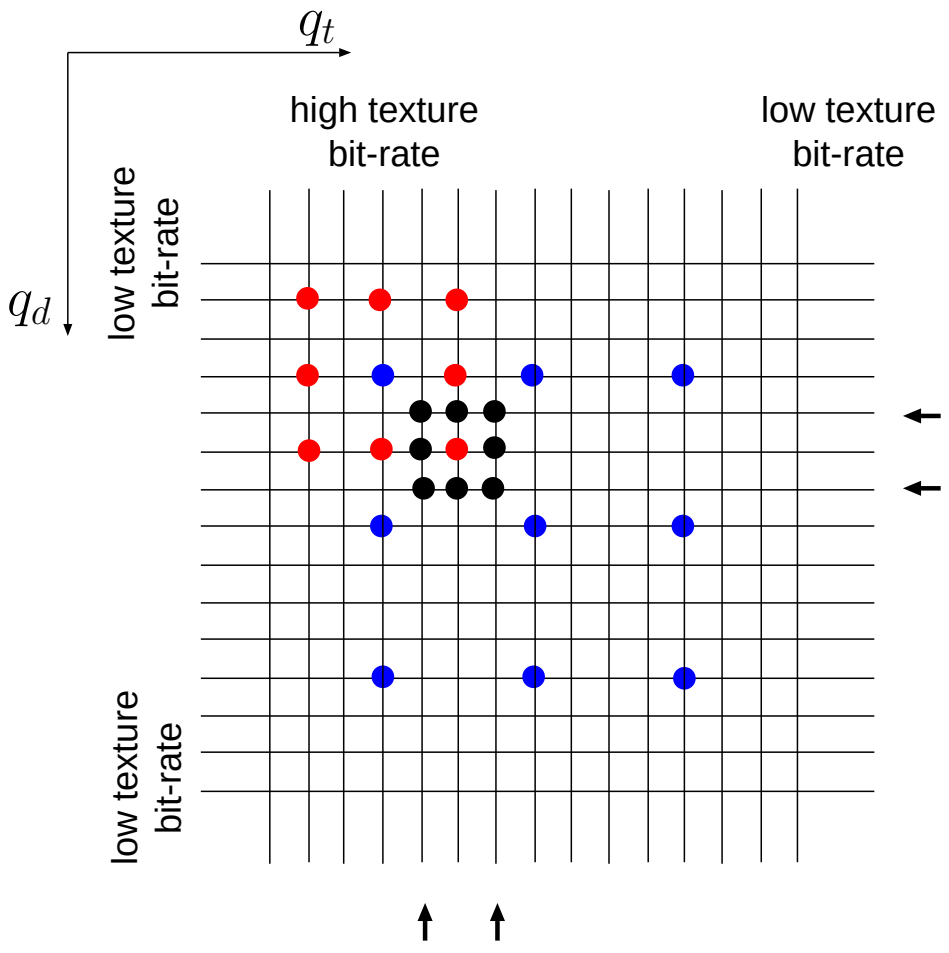
- Hierarchical search reduces the computational complexity.
- Reduced number of coding trials:
  - $3 + 3$ trials $\Rightarrow$ 9 R-D points
  - $2 + 2 \Rightarrow$ 9+8 R-D points



TU/e

Hierarchical optimization 2/2

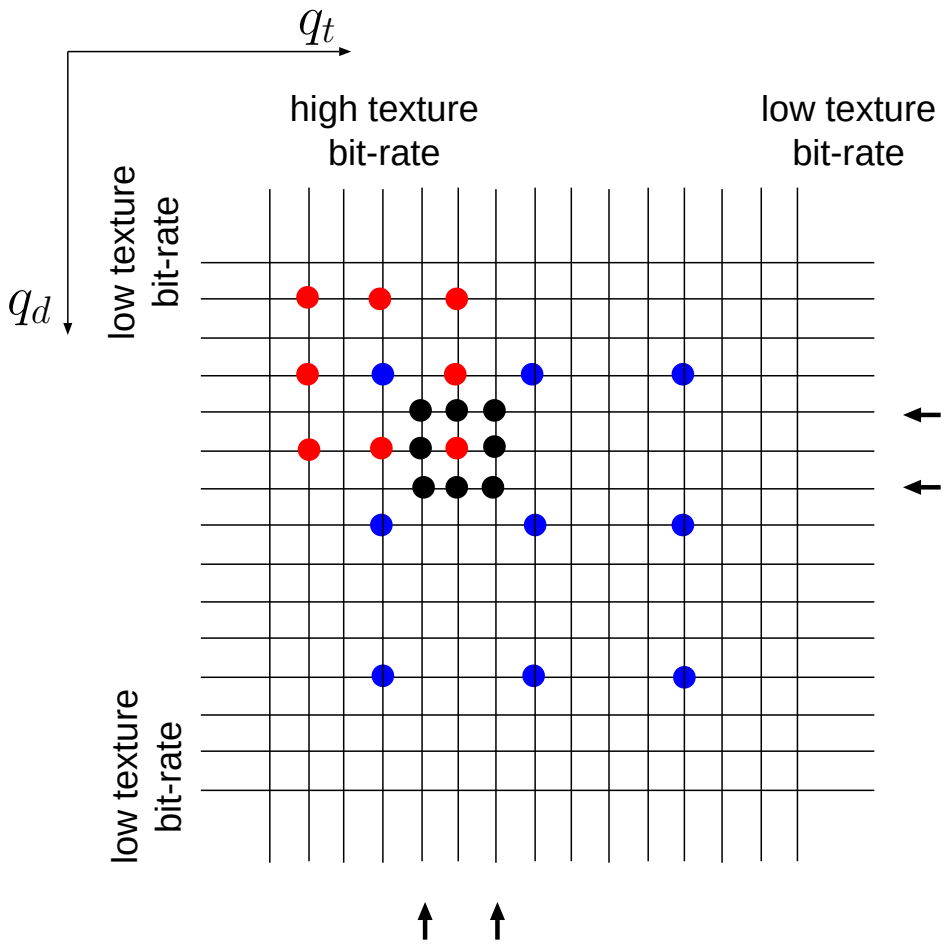
- Hierarchical search reduces the computational complexity.
- Reduced number of coding trials:
  - $3 + 3$ trials $\Rightarrow$ 9 R-D points
  - $2 + 2 \Rightarrow$ 9+8 R-D points
  - $2 + 2 \Rightarrow$ 9+8+8 R-D points



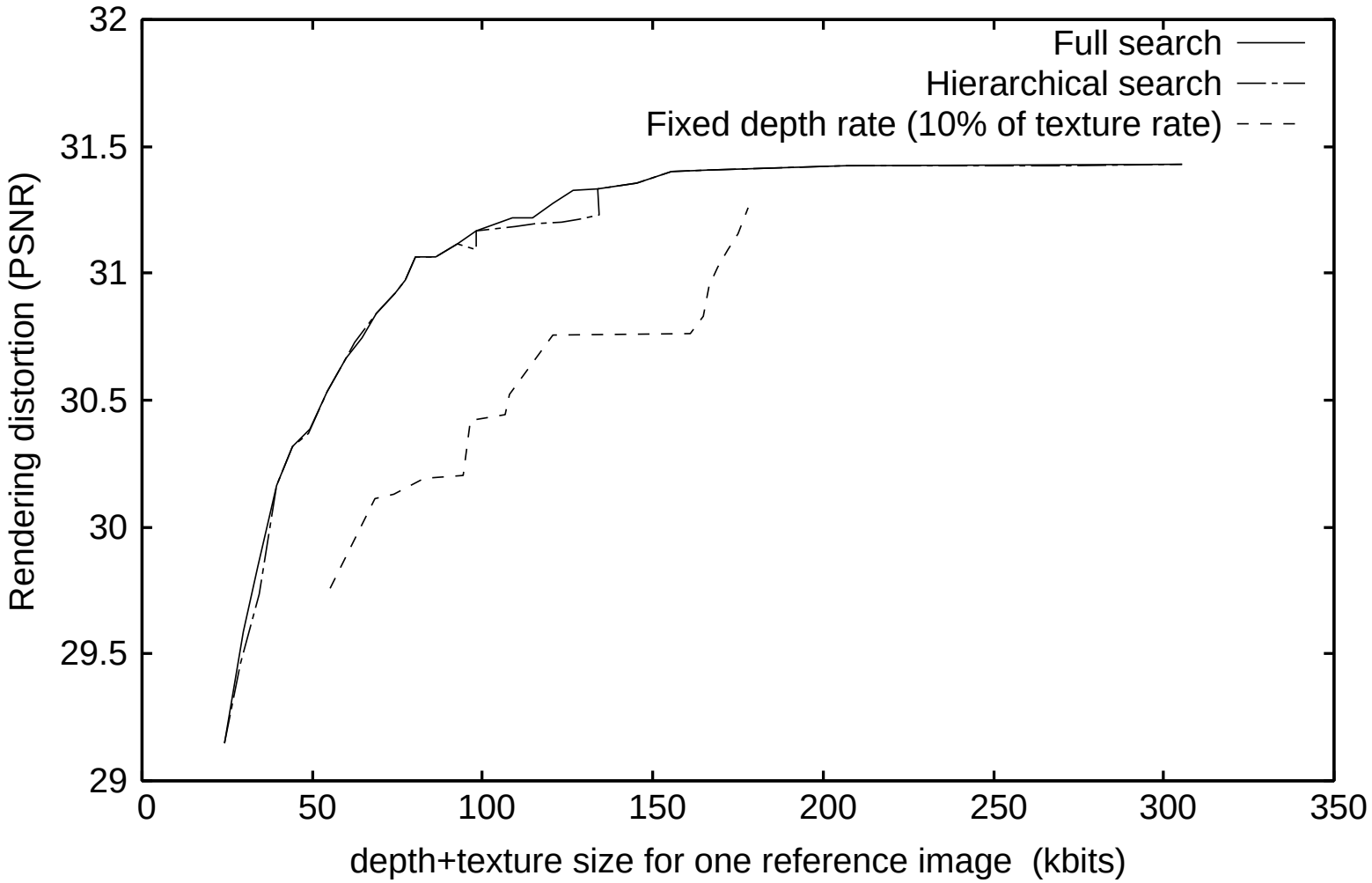
Hierarchical optimization 2/2

- Hierarchical search reduces the computational complexity.
- Reduced number of coding trials:
  - $3 + 3$ trials $\Rightarrow$ 9 R-D points
  - $2 + 2 \Rightarrow$ 9+8 R-D points
  - $2 + 2 \Rightarrow$ 9+8+8 R-D points

  - 14 coding trial $\Rightarrow$ 25 R-D points.



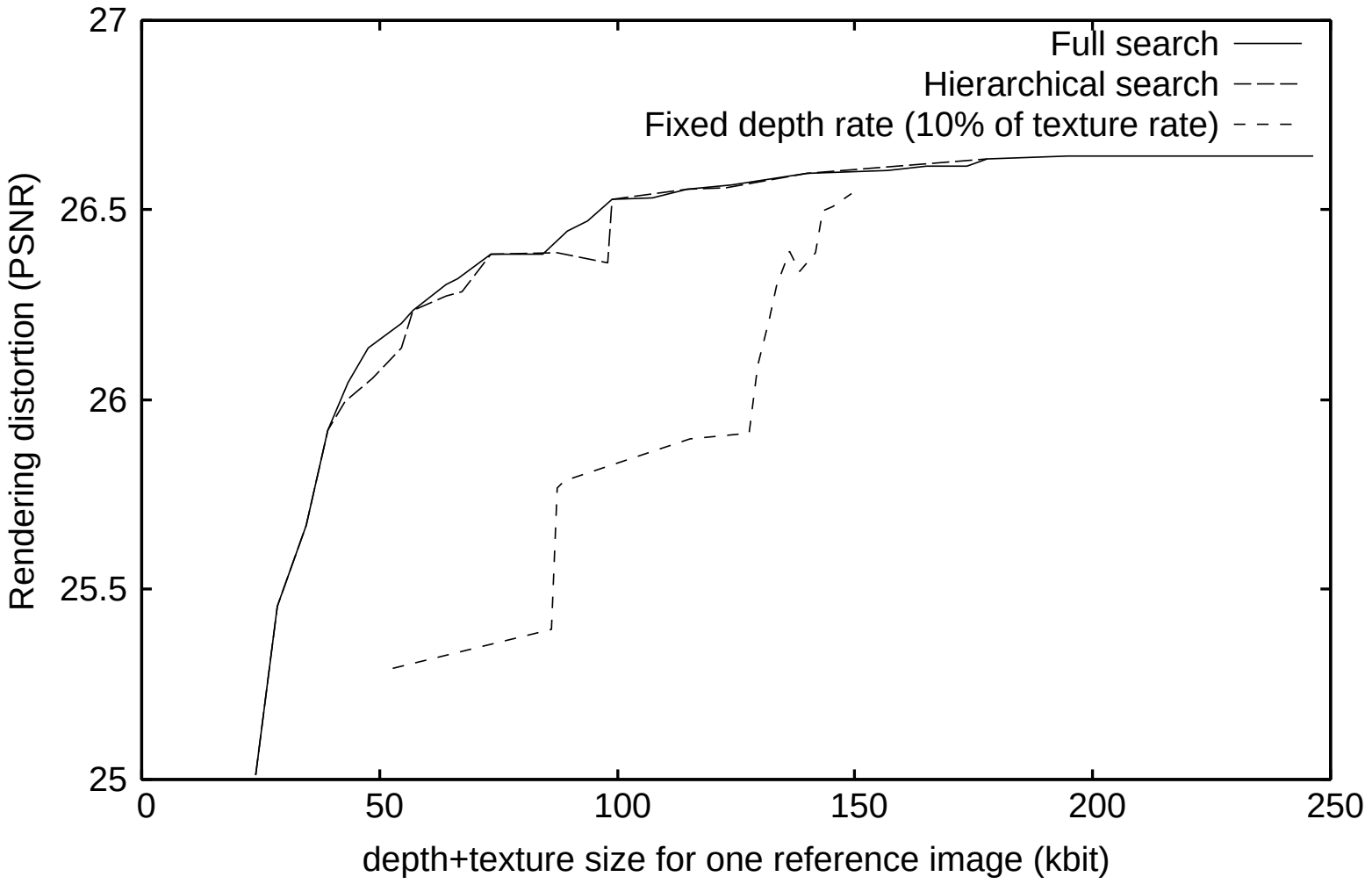
Experimental results 1/2



Breakdancers sequence

TU/e

Experimental results 2/2



Ballet sequence

Conclusions

- We have presented an algorithm that addresses the problem of joint bit allocation for depth and texture.

- The algorithm estimates the quantization parameters so that the view-rendering quality is maximized.
- The parameters are optimized using a hierarchical search requiring only a low number of coding trials.

TU/e

Conclusions

- We have presented an algorithm that addresses the problem of joint bit allocation for depth and texture.
- The algorithm estimates the quantization parameters so that the view-rendering quality is maximized.
- The parameters are optimized using a hierarchical search requiring only a low number of coding trials.
- Up to 1 *dB* rendering quality improvement.

TU/e

Conclusions

- We have presented an algorithm that addresses the problem of joint bit allocation for depth and texture.

- The algorithm estimates the quantization parameters so that the view-rendering quality is maximized.
- The parameters are optimized using a hierarchical search requiring only a low number of coding trials.

- Up to 1 *dB* rendering quality improvement.
- Low and constant complexity: 14 coding trials.

TU/e

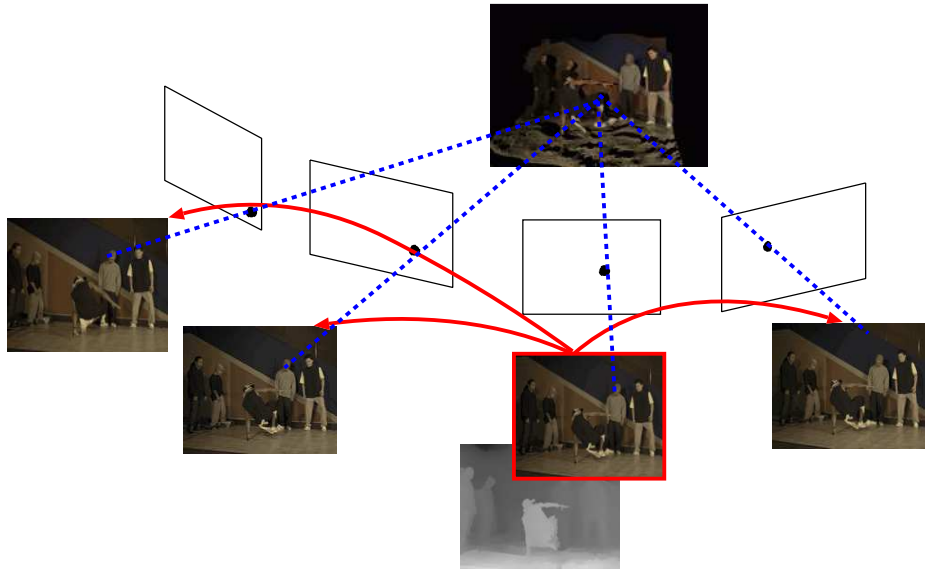
Conclusions

- We have presented an algorithm that addresses the problem of joint bit allocation for depth and texture.

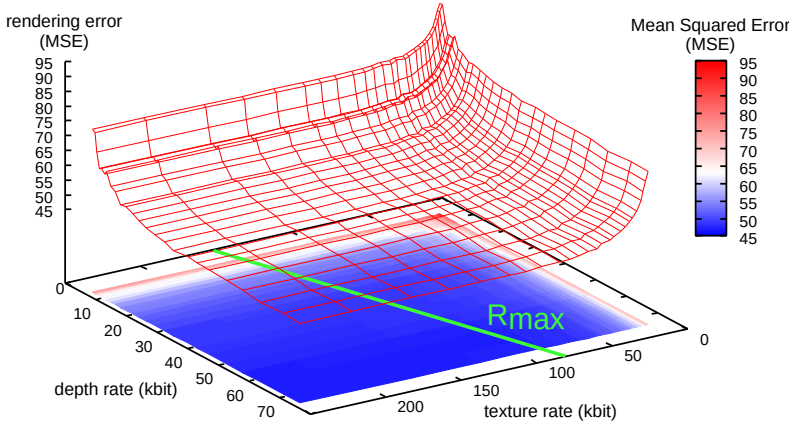
- The algorithm estimates the quantization parameters so that the view-rendering quality is maximized.
- The parameters are optimized using a hierarchical search requiring only a low number of coding trials.

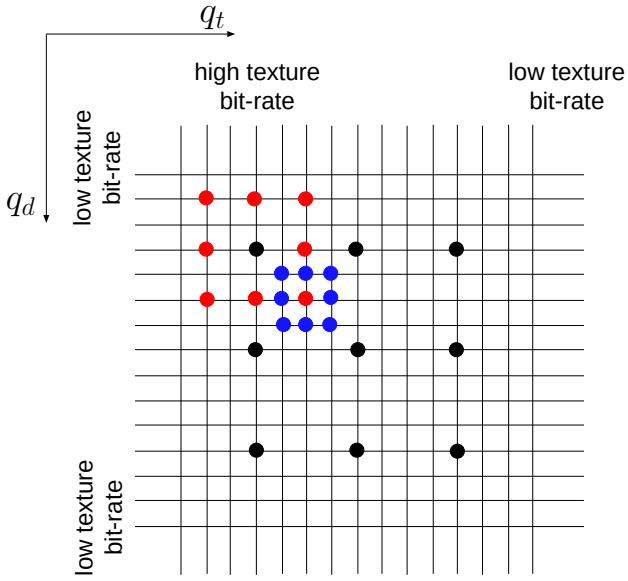
- Up to 1 *dB* rendering quality improvement.
- Low and constant complexity: 14 coding trials.
- Can be readily integrated into an MVC coder.

TU/e



Questions ?


rendering error
(MSE)
95
90
85
80
75
70
65
60
55
50
45
0
10
20
30
40
50
60
70
depth rate (kbit)
200
150
100
50
0
texture rate (kbit)
Rmax
Mean Squared Error
(MSE)
95
90
85
80
75
70
65
60
55
50
45


q_t
high texture
bit-rate
low texture
bit-rate
q_d
low texture
bit-rate
low texture
bit-rate