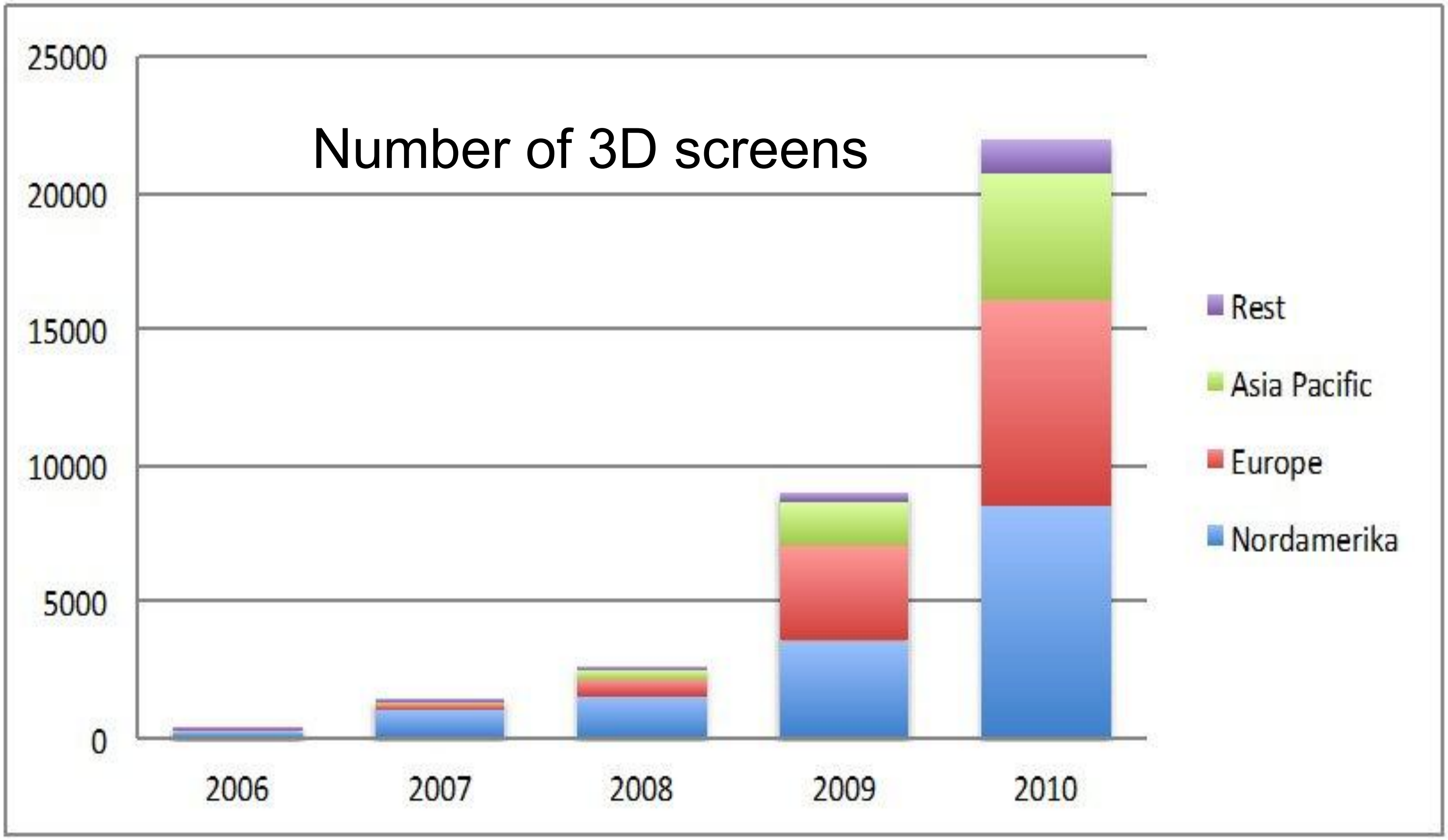
3D Video Coding

- 3D Cinema
- Stereo Video
- Disparity
- Autostereoscopic Displays
- 3D Video Coding for Stereo Displays
- 3D Video Coding for Autostereoscopic Displays

The Development of 3D Cinema



Source: Flying Eye, 2011

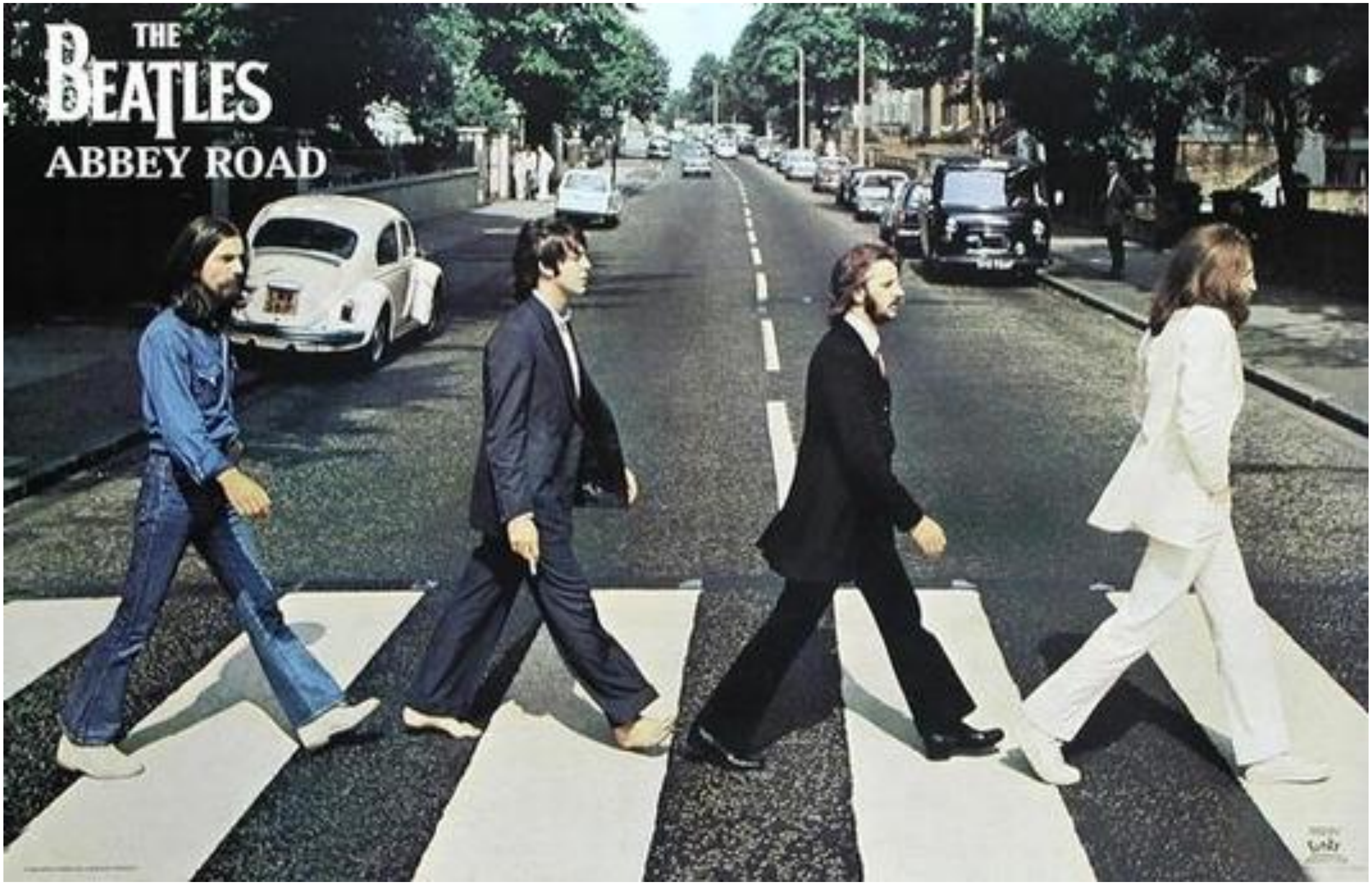
3D Cinema: Stereo Video

anaglyph

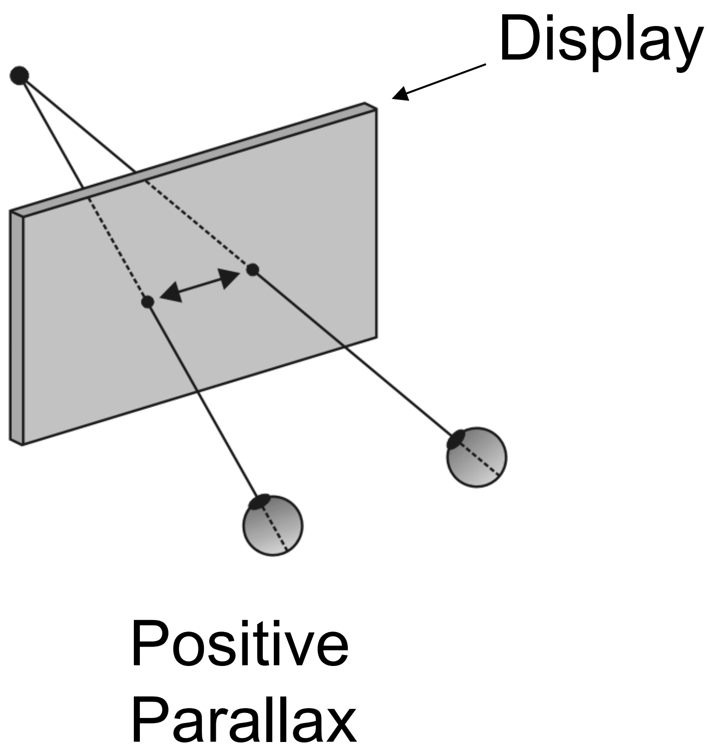
polarized   shutter

2 Views

3D on 2D Displays

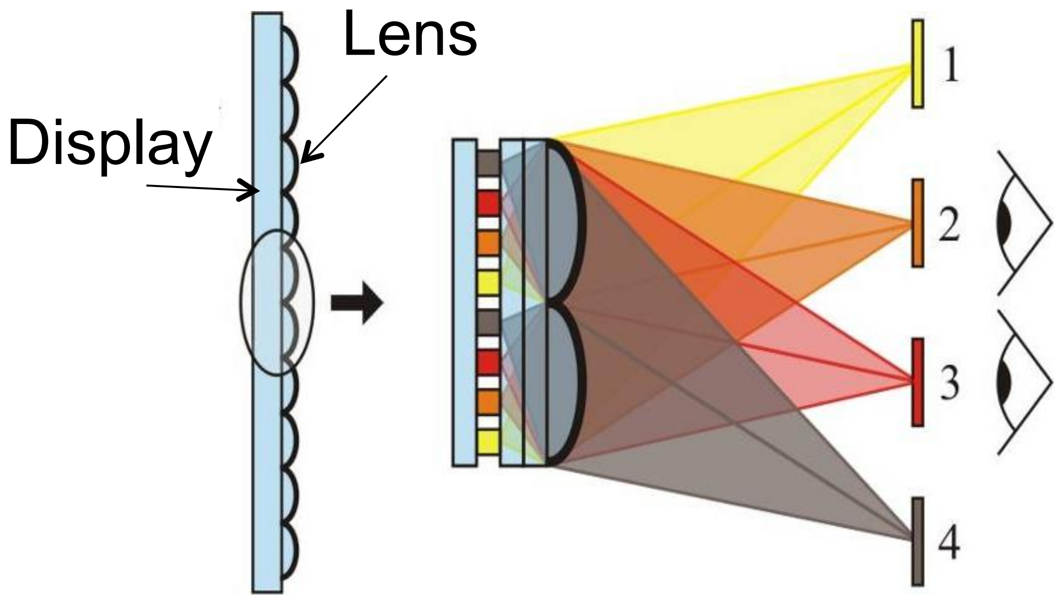


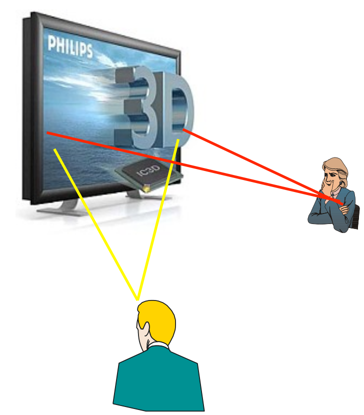
Why Glasses?



- Need to project a different image into each eye
- Glasses provide control over what each eye sees

Autostereoscopic Displays

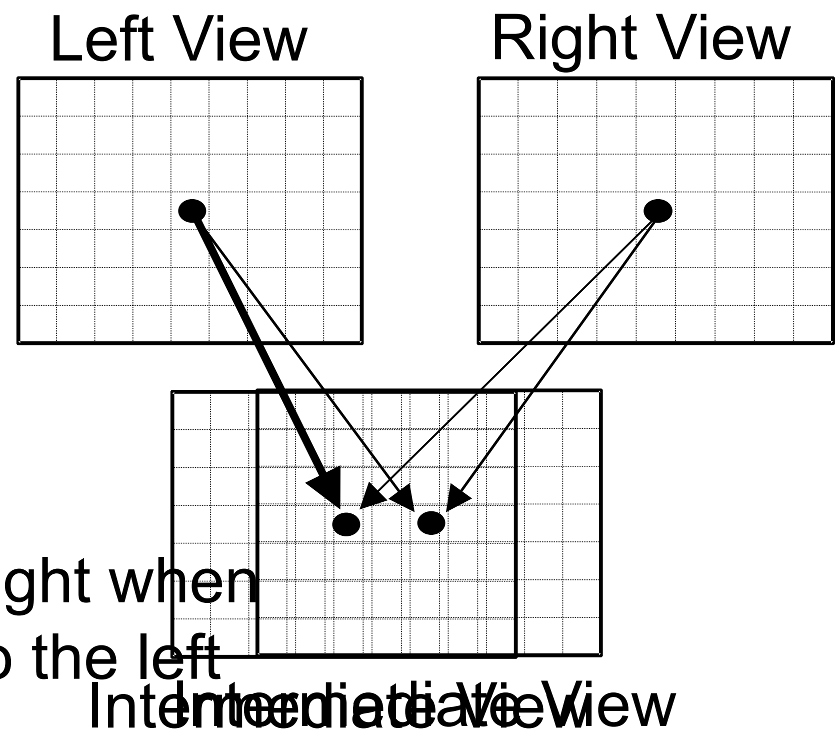




- Field of view is repeated
- Size of field of view is governed by number of views:
  Size: 1.50 m \ Number of Views: 50
- Problem for acquisition and transmission

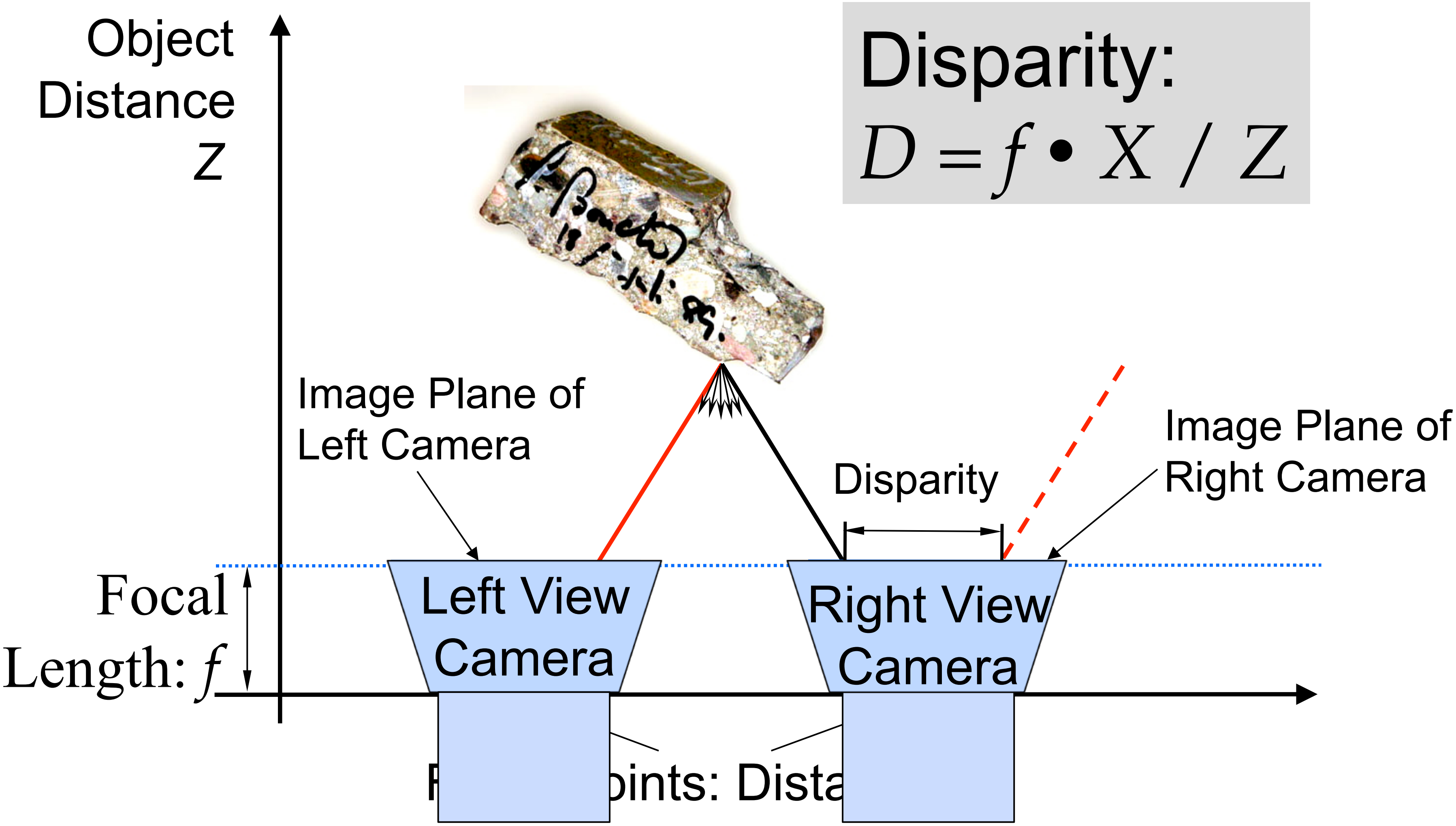
Two Cameras and Interpolation

- Continuous reproduction of all intermediate views
- Weighted averaging of left and right view



Increase left weight when view is moved to the left

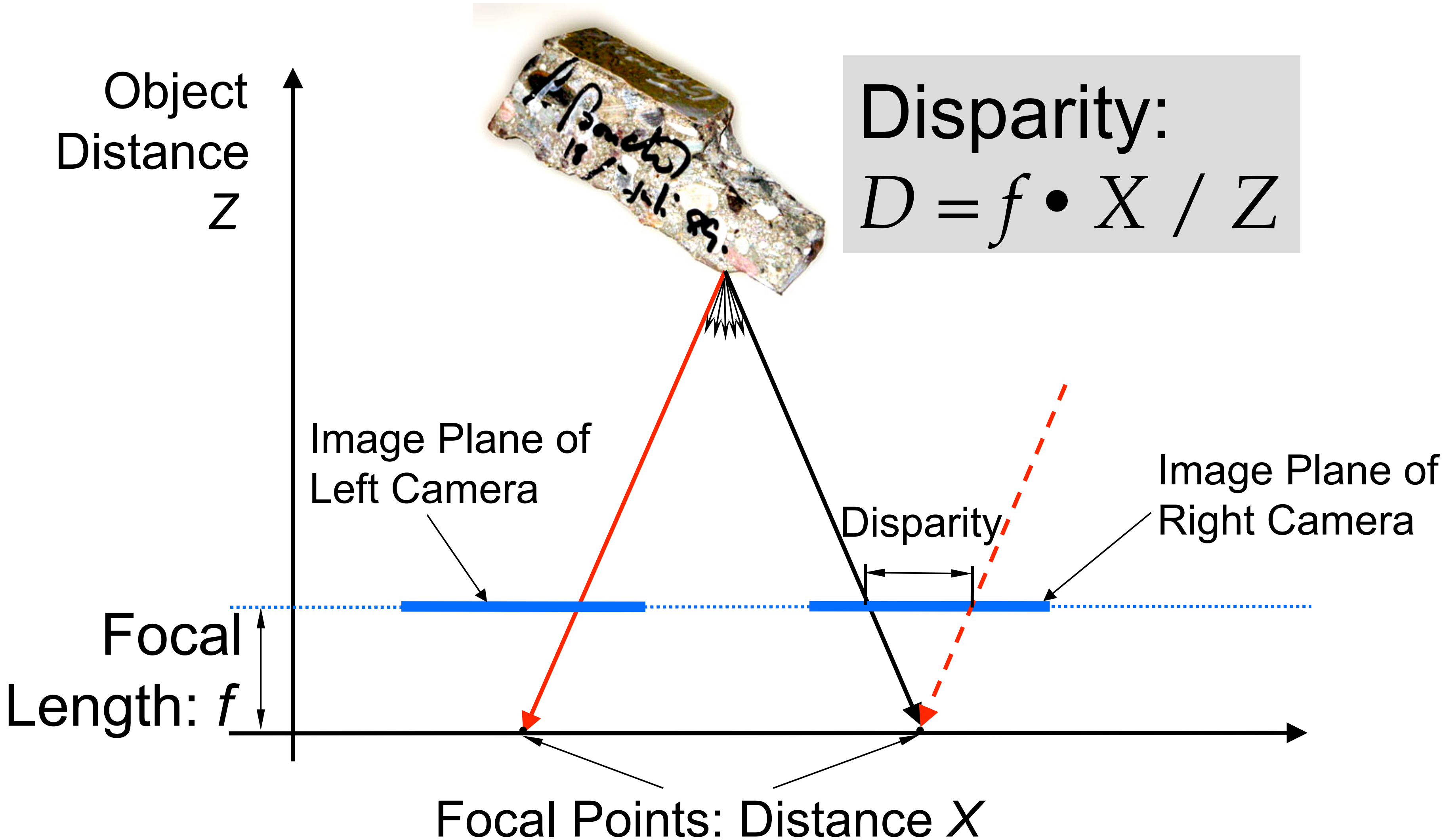
Acquisition: Plenoptic Sampling



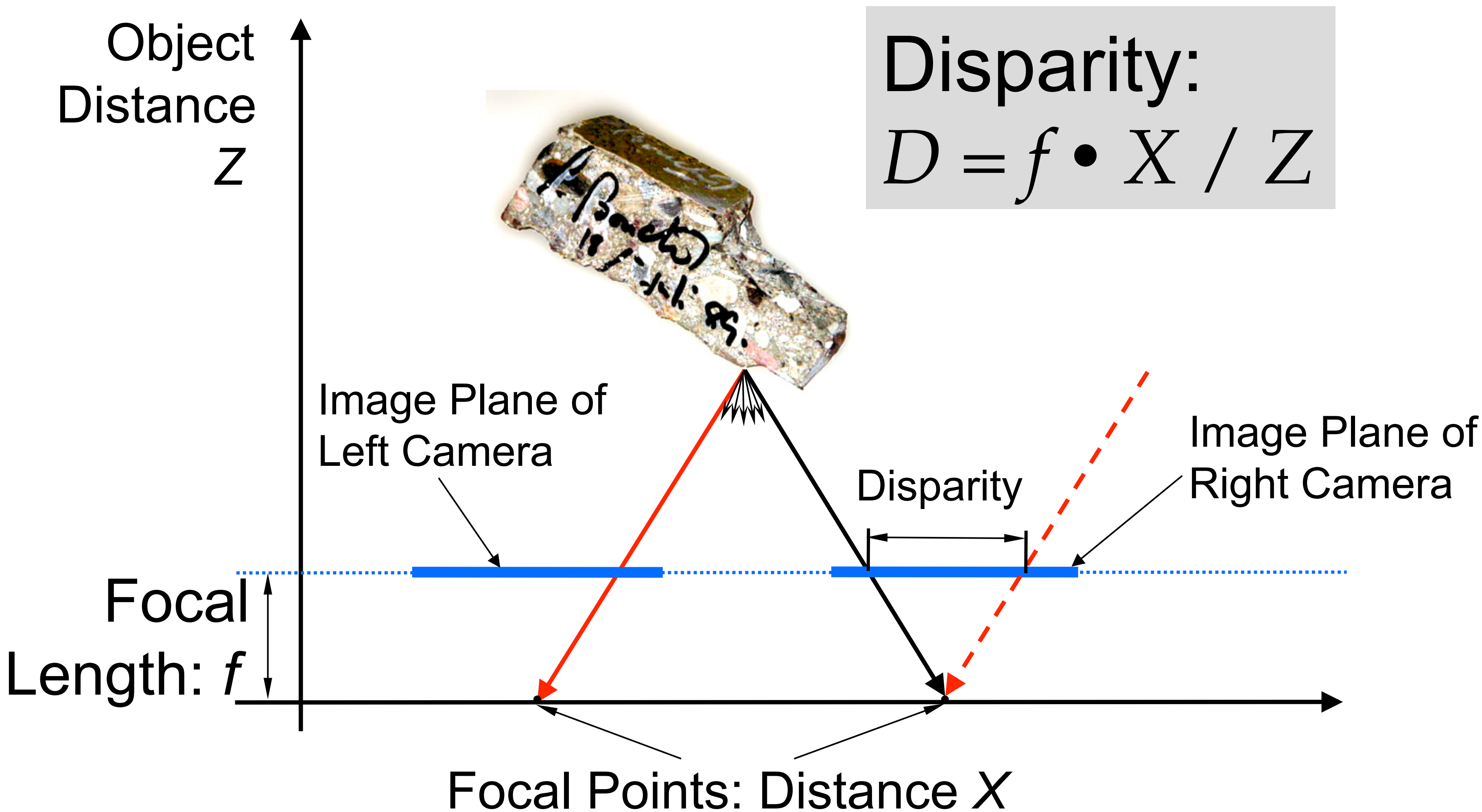
Acquisition: Plenoptic Sampling

Disparity:

$$D = f \bullet X \,/\, Z$$

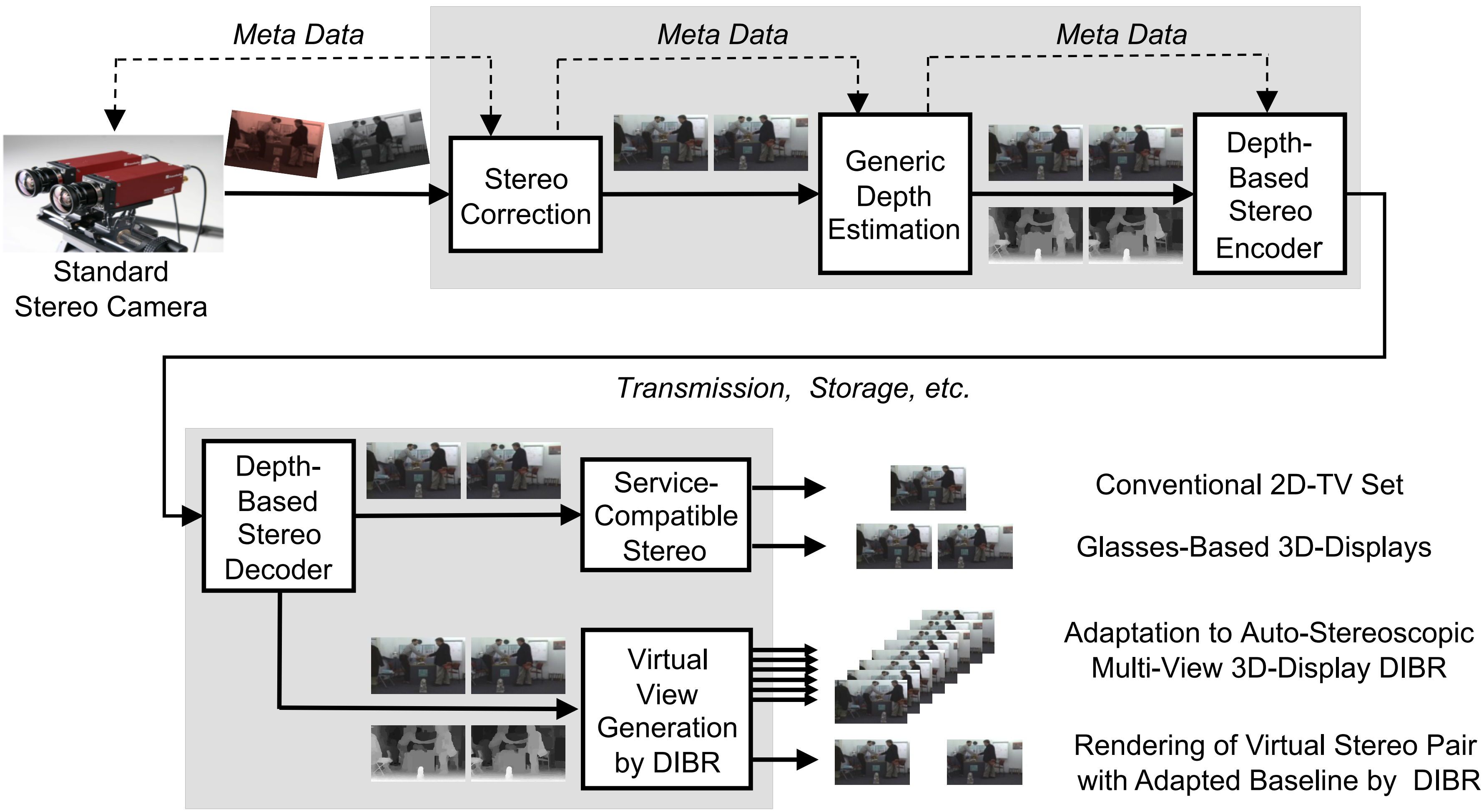


Acquisition: Plenoptic Sampling

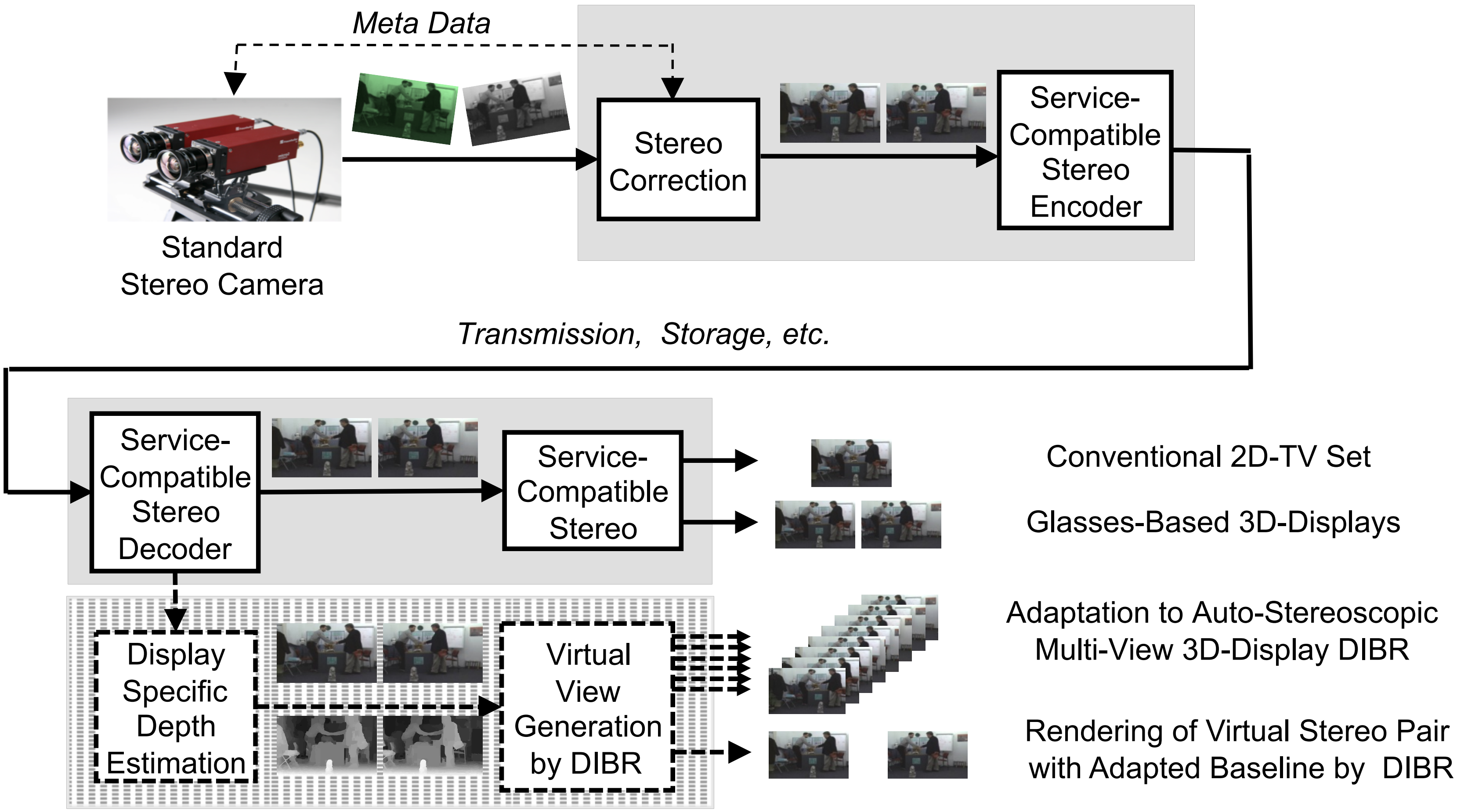


Depth Maps

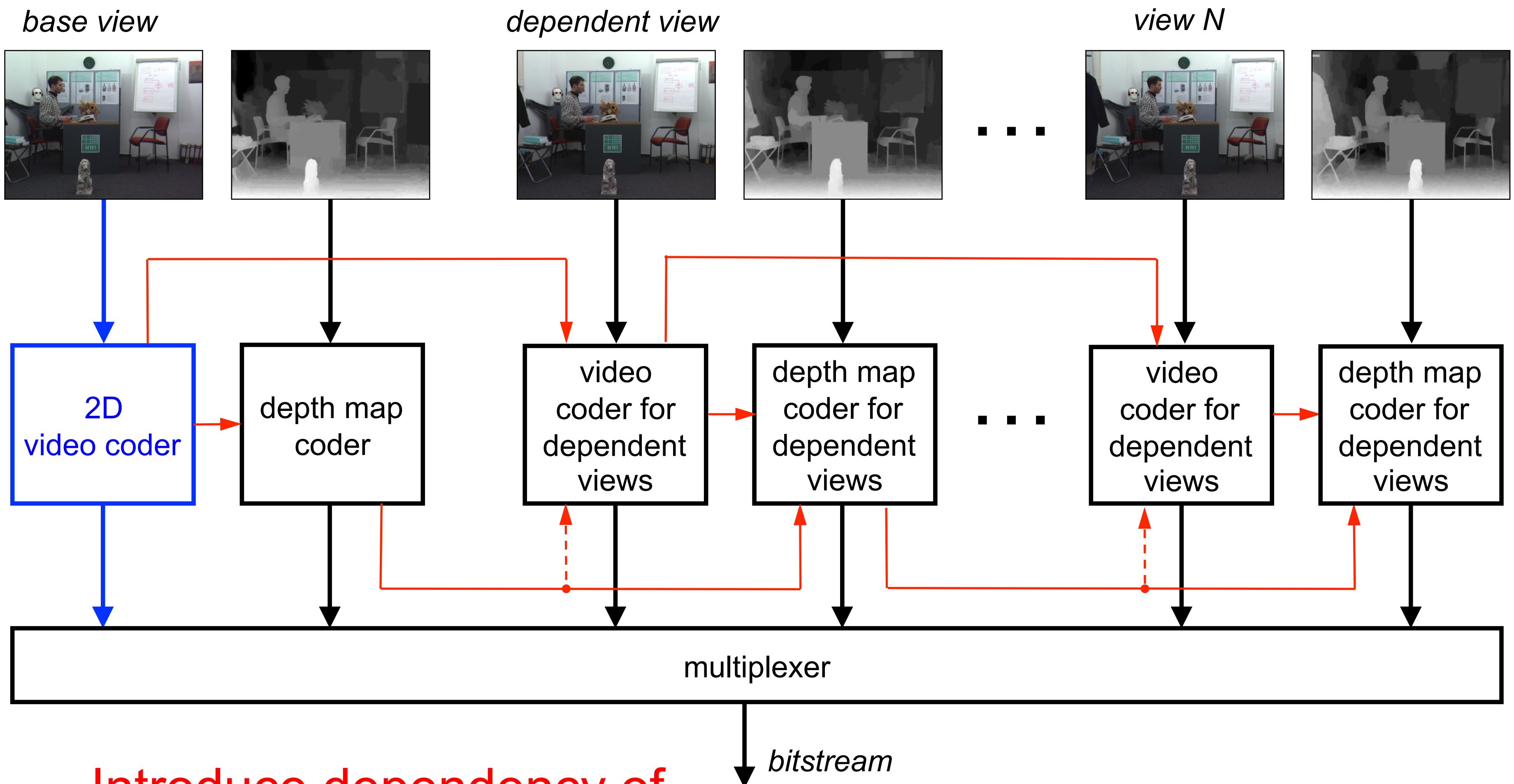
3DV System A: Creating Depth Maps at the Sender



3DV System B: Creating Depth Maps at the Receiver

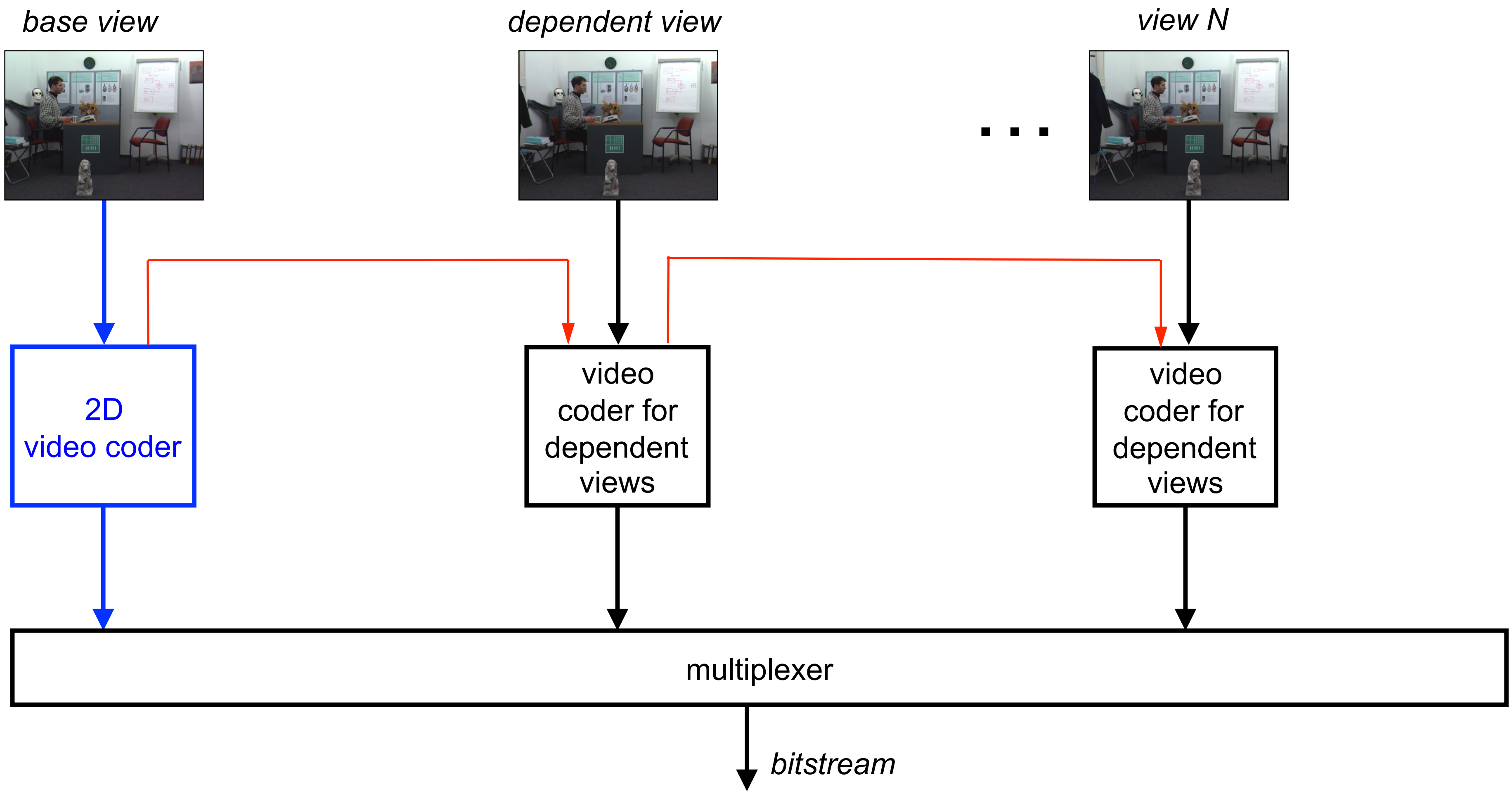


3D Video Encoder



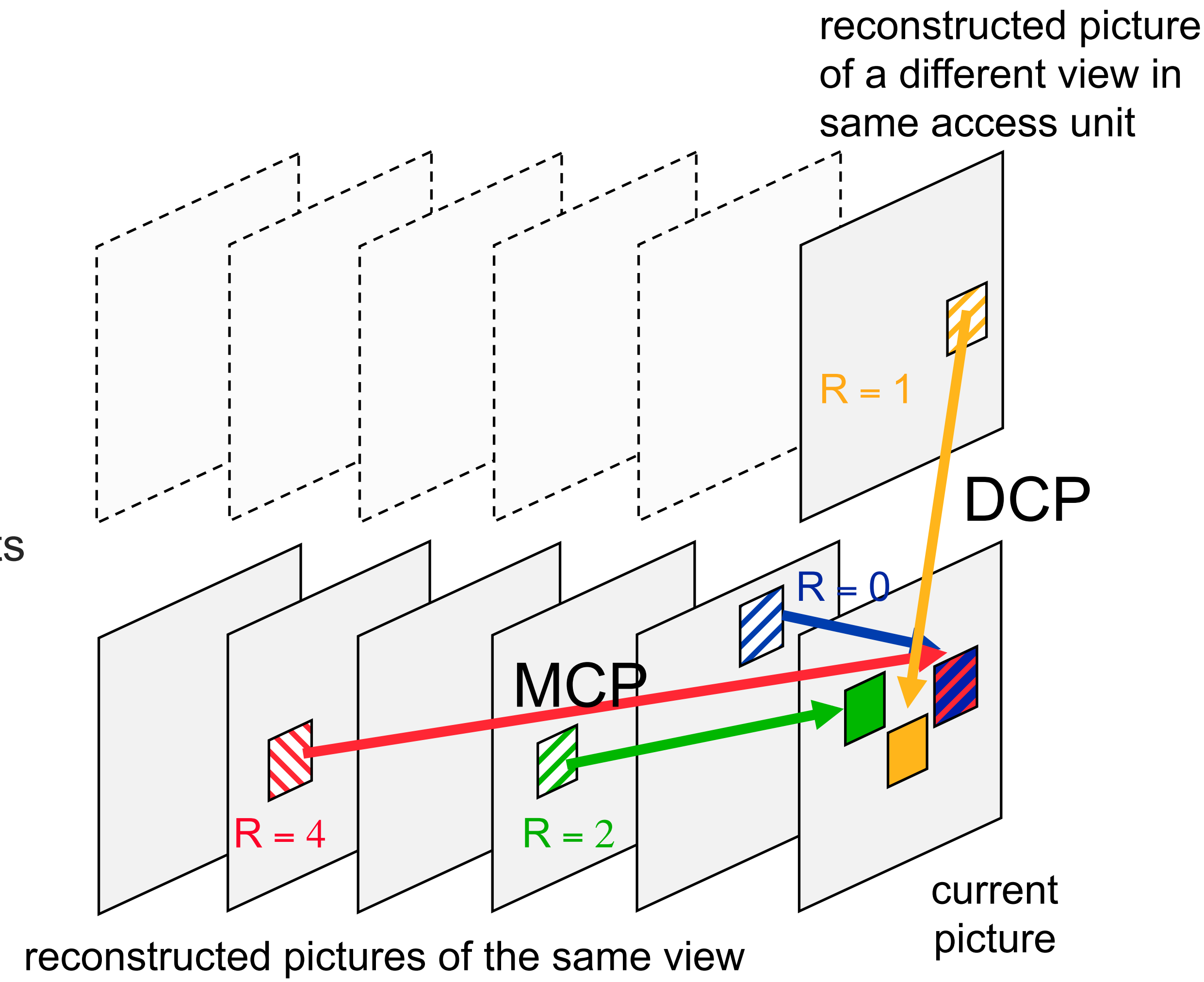
Introduce dependency of video signal from depth signal?

Coding of Dependent Views



Disparity-Compensated Prediction

- Use previously decoded pictures as additional reference pictures
- Only construction of reference picture lists is modified



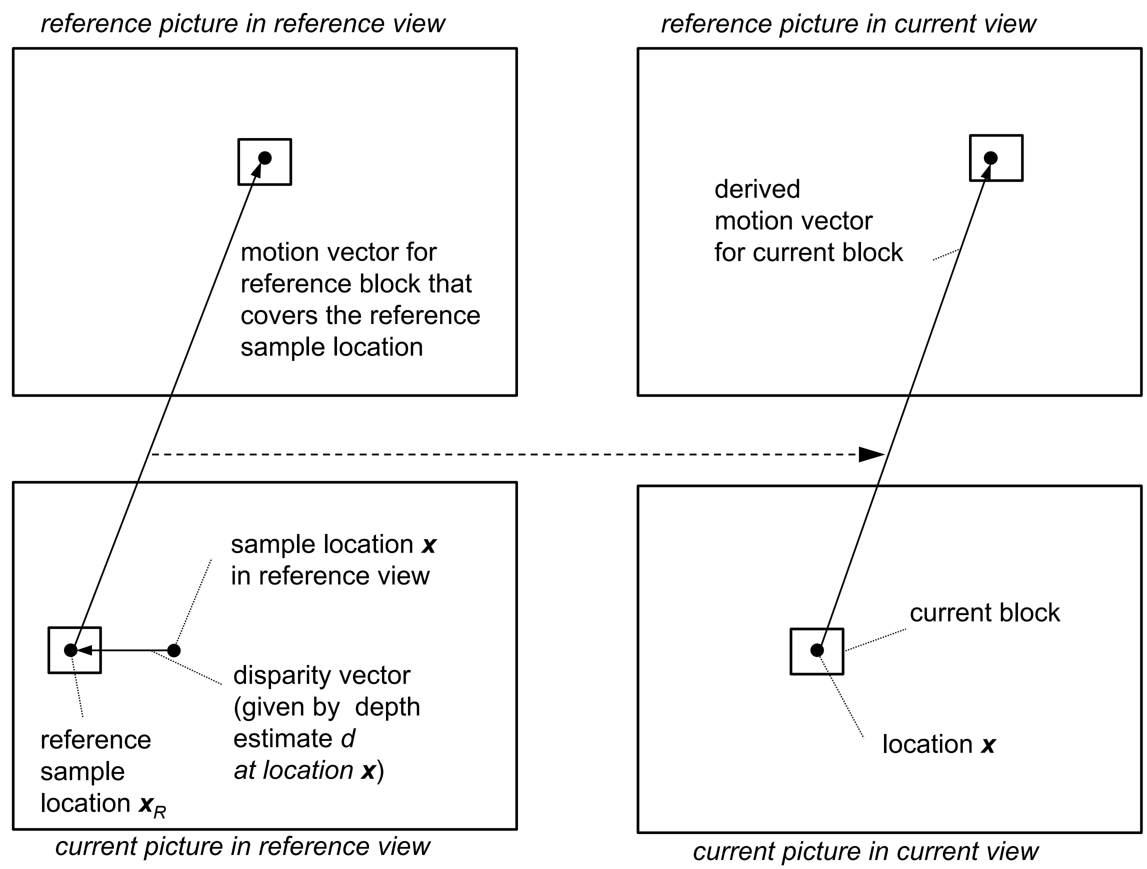
Inter-View Motion Parameter Prediction

Basic idea:

- Motion is similar in different views
- Create an efficient motion vector predictor

Uses an estimate of the current depth map:

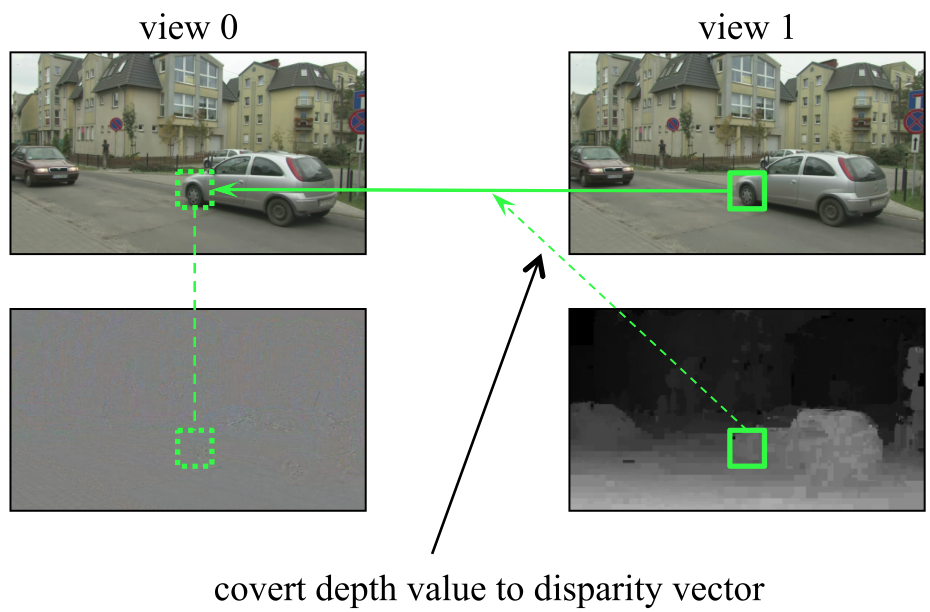
- Not yet coded
- How to derive it?



Inter-View Residual Prediction

Residual prediction:

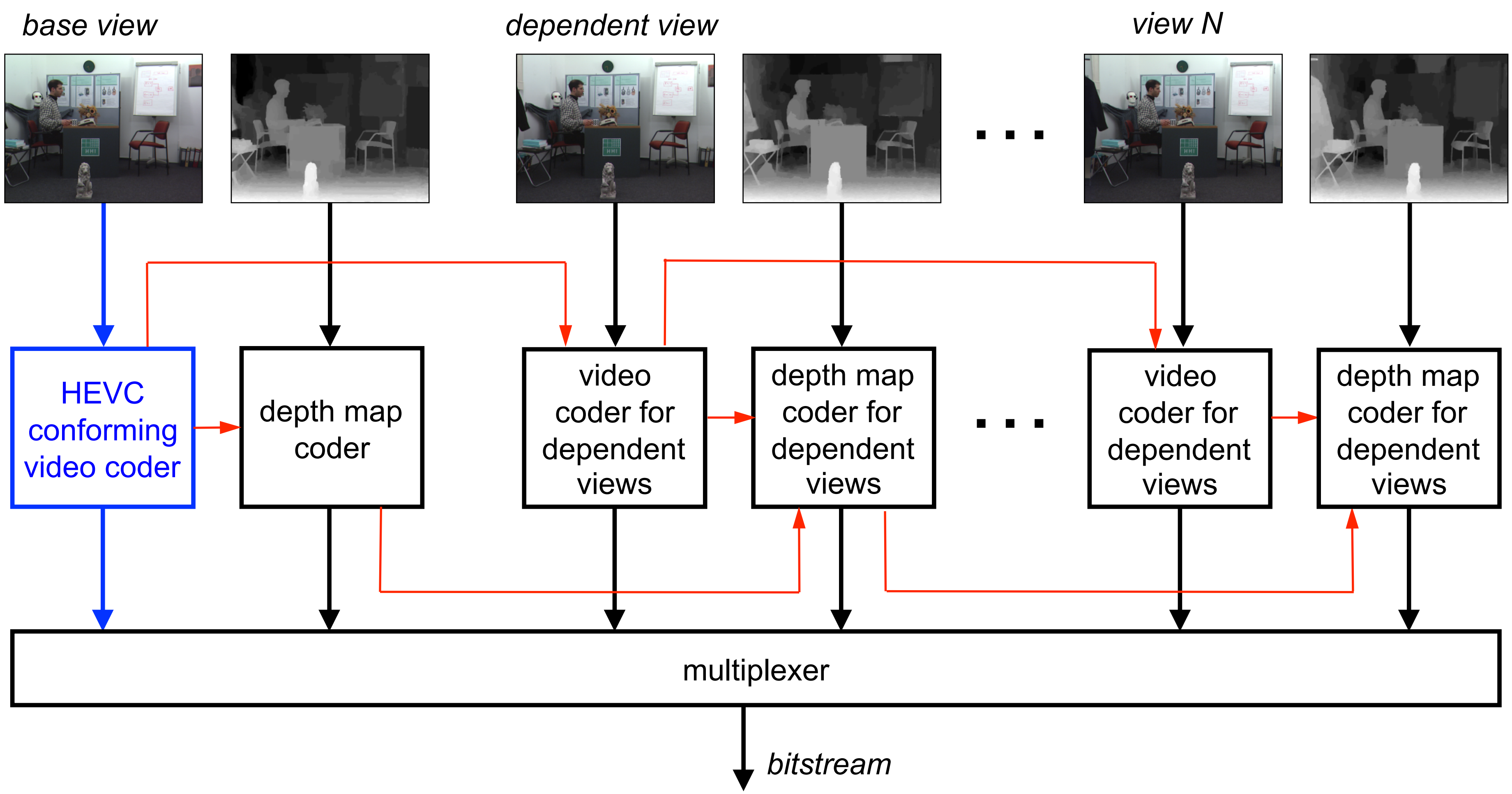
- Predict residual of a current block using a coded residual block in a reference view
- Determined disparity using the estimated depth maps (same depth as for inter-view motion parameter prediction)
- Disparity-compensated prediction of the current residual
- Bilinear filtering for sub-sample interpolation



Signalling:

- Transmit residual prediction flag for motion-compensated blocks

Coding of Depth Maps



Depth Map Coding: Simplified Video Codec

- Integer-pel motion vector accuracy
  - Motion compensation with full-sample accuracy
  - No interpolation filter
  - Coding of motion vector differences with full-sample accuracy
- Disabling of in-loop filtering
- Disparity-compensated prediction
  - For dependent views
  - Same concept as for video pictures

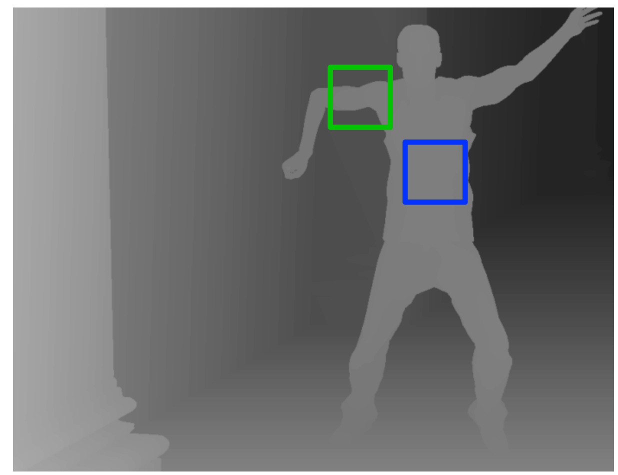
Depth Modelling Modes

Depth map properties:

- Sharp edges representing object borders
- Large areas of slowly varying values representing object areas
- Edges in depth maps are correlated with edges in video pictures

Coding ideas specific to depth maps:

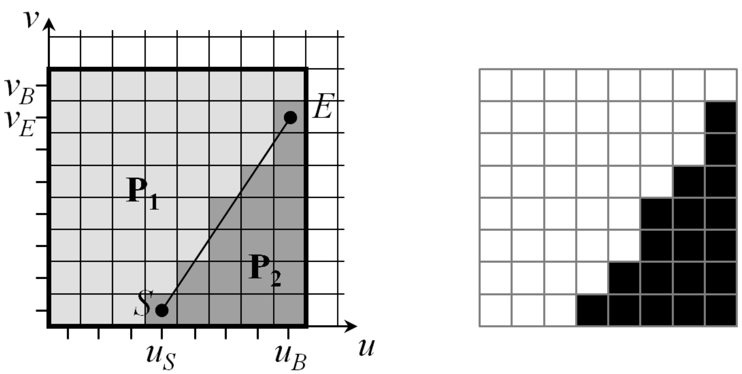
- Representation of depth edges
- Partition block into two regions with constant sample values
- Prediction based on co-located texture block
- Optional transform coding of residual

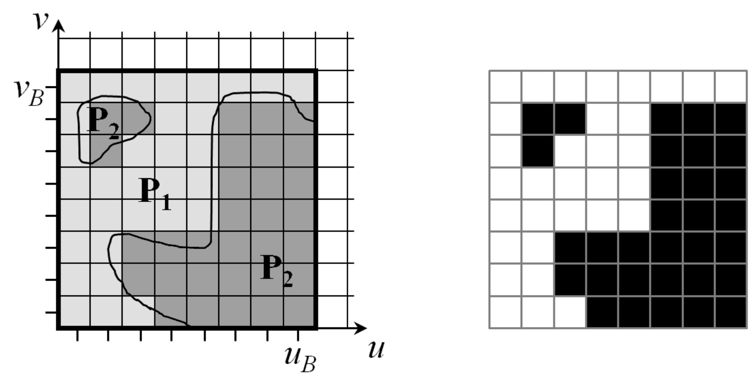


Intra Partitioning Modes

Non-rectangular block partitions

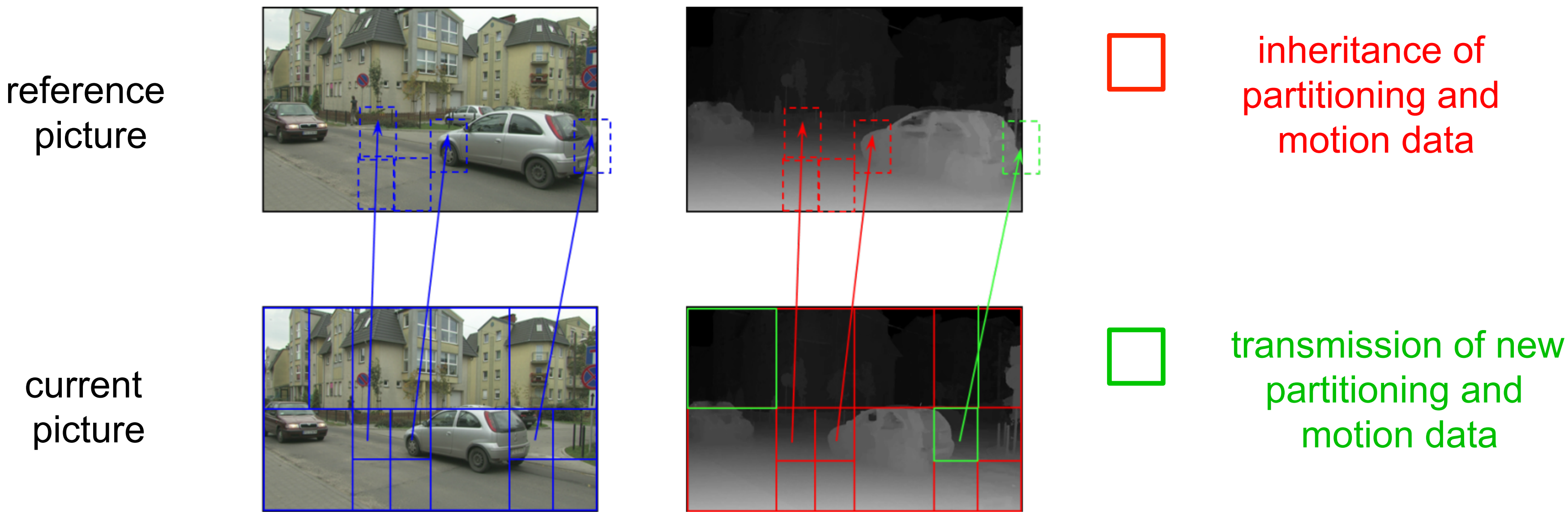
- Approximating the signal of a depth block by a model that partitions the area into two non-rectangular regions
- Each region is represented by a constant value



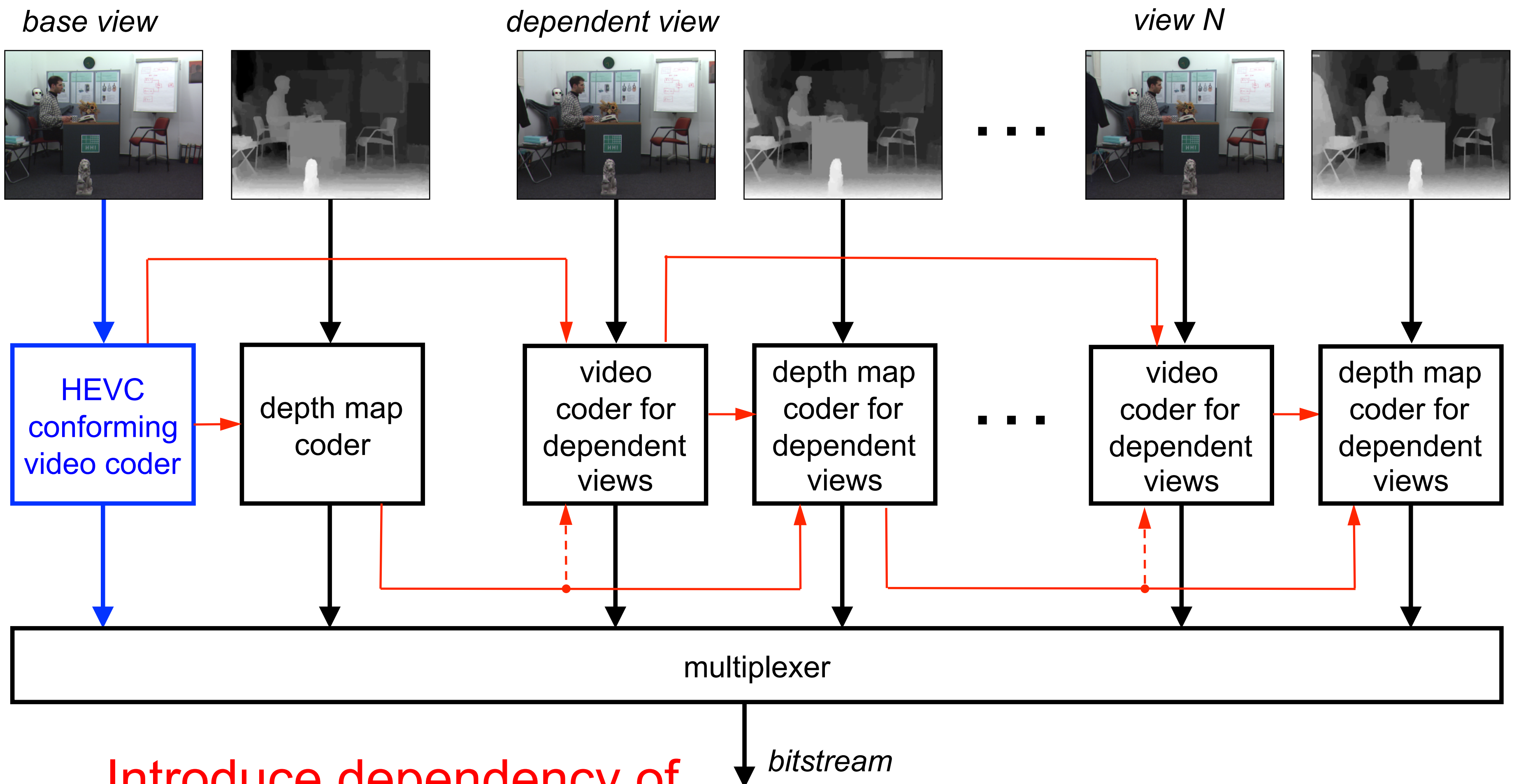


(even several parts)

Motion Parameter Inheritance



How to Control All of This?



Introduce dependency of video signal from depth signal?

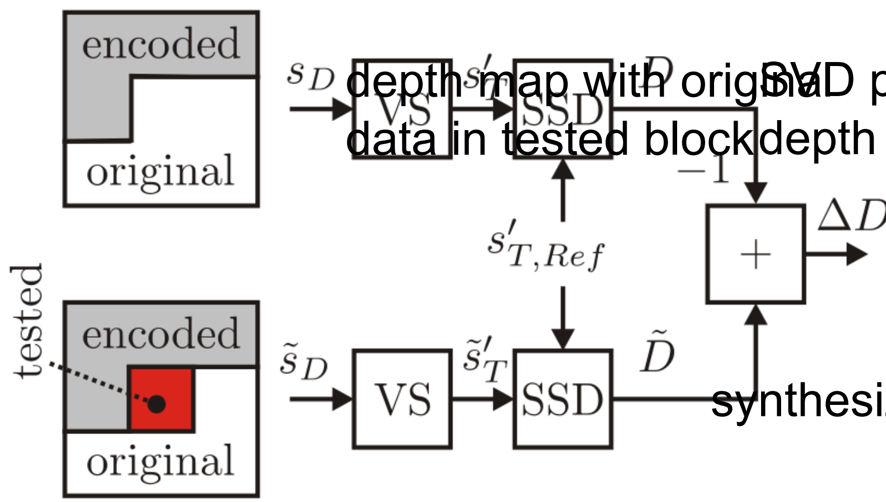
General Encoder Control

- Video picture encoding
  - No changes to HM encoder control
  - Decisions based on Lagrangian costs $D + \lambda \cdot R$
    - Motion estimation and mode decision
    - Quantization
  - Fast motion search and fast mode decision of HM
  - Optionally: Depth-aware coder control for dependent views
- Depth Maps
  - Basically same encoder control as for video pictures
  - For mode decision: **View synthesis optimization**
    - Measure distortion in synthesized views (instead of in depth domain)
    - Includes rendering capabilities in encoder

View Synthesis Optimization

Distortion measure: Synthesized view distortion change (SVDC)

- Difference in distortion due to depth coding with a particular mode
- Additive distortion measure
- Reference: Original view or synthesized view obtained with original data



depth map with original data in tested block

SVD prior to depth change

SVDC:
used as distortion measure in RDO

synthesized view distortion (SVD)

Depth map:
- Coded data in encoded blocks
- Original data in not yet coded blocks
- Coded data in current block

view synthesis with reconstructed depth map

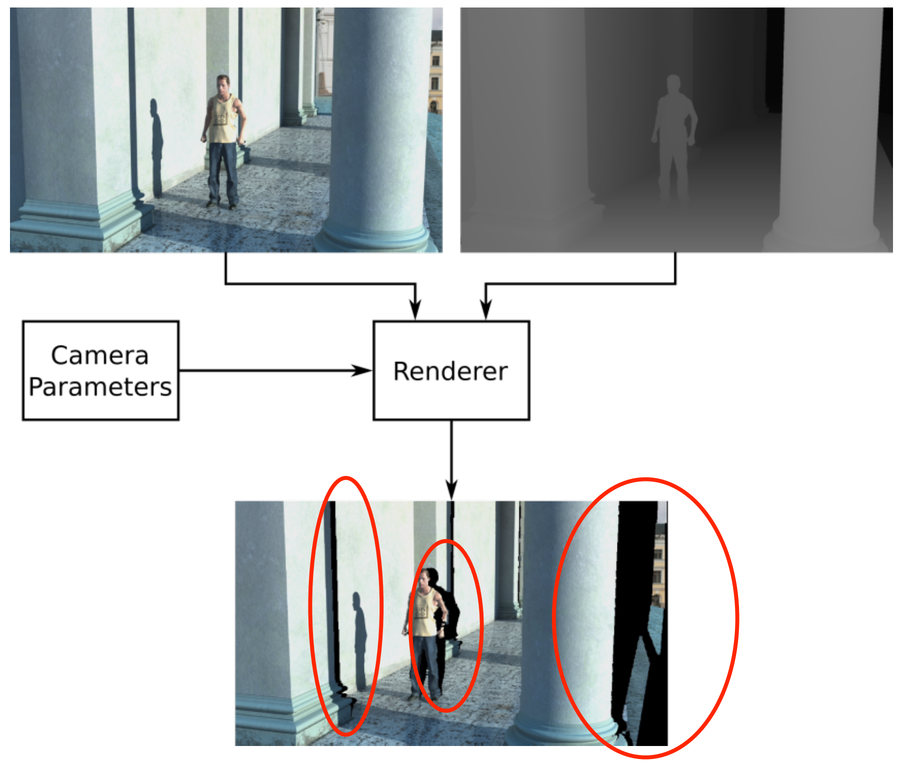
distortion computation comparing to reference rendered from original data

Depth-Aware Encoder Control

- Regions in one view can be rendered given the data for other views
- Encode such regions only once with high quality (in base view)
- Render these regions after decoding

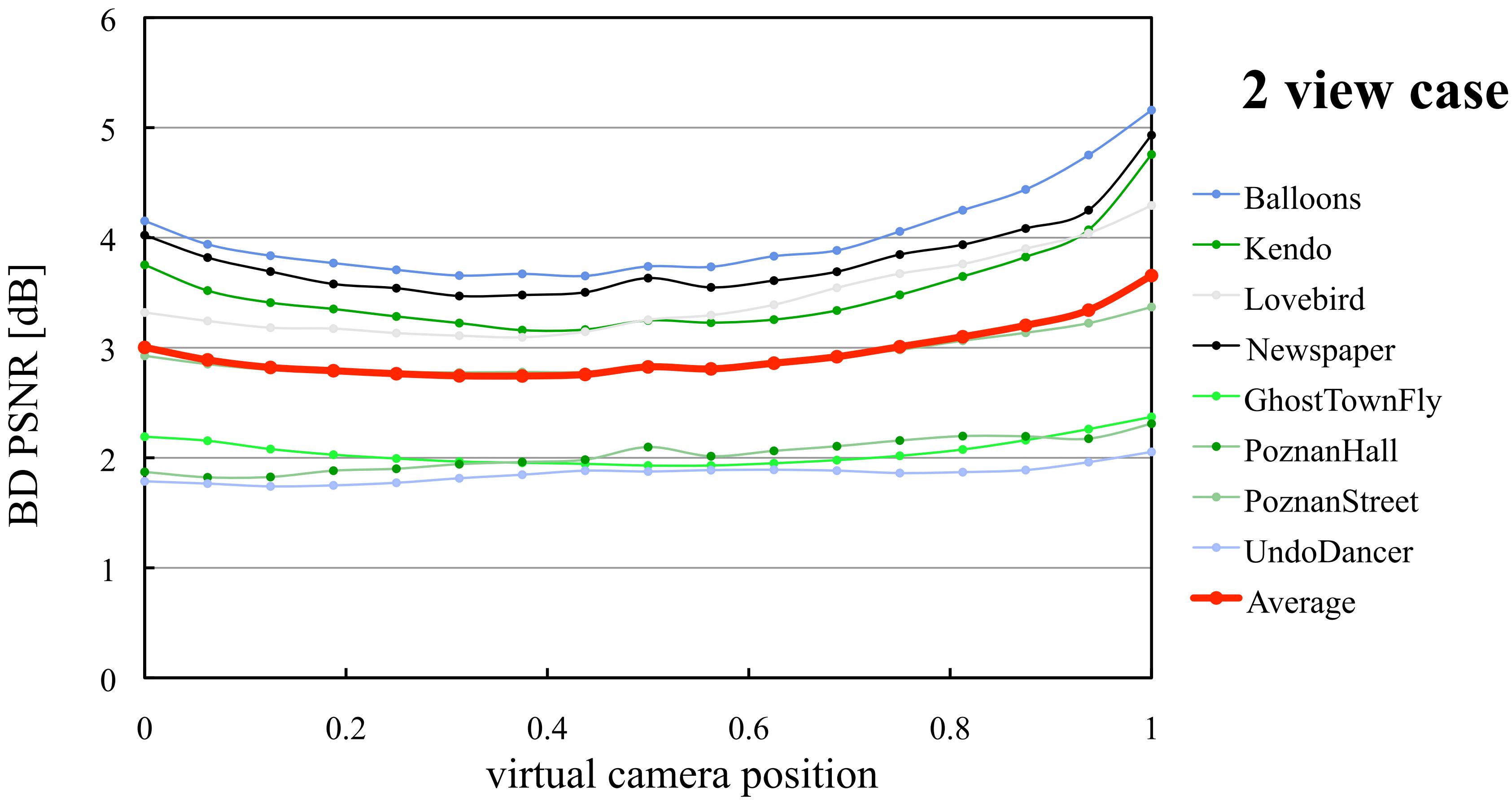
Only encode occluded parts of the second view

Encoder only technique



Coding Efficiency - Stereo

- PSNR in synthesized views against synthesized pictures with original data
- Calculate Bjøntegaard Delta (BD) PSNR relative to simulcast



Summary

- 3D is perceived by multiple clues
- Glasses provide control over what each eye sees: stereo video
- Efficient stereo codec: disparity compensation, motion and residual prediction
- Efficient multiview plus depth codec: depth prediction using video contours, inheritance of motion vectors
- Need for transmission of depth data not clear