3d video

Past, Present, Future

Information Coding Group
Linköpings universitet

LiU expanding reality

Motivation

- James Camerons Avatar – first Hollywood fullength feature film to be filmed and produced completely in both 2D and 3D
  - But not first 3D movie in recent years!
  - Movies tailored to 3D (documentations e.g. Deep Sea, concerts e.g. U2 3D)
  - Feature films containing 3D scenes (e.g. Harry Potter and the Half-Blood Prince)
  - Animation movies (e.g. Pixar's Up); older 2D animation movies are planned to be rereleased in 3D

- 3D version of a movie generally more popular than its 2D counterpart

- Part of ongoing trend towards 3D

Why now?

- Price aspects
  - 3D systems became only recently cheap enough for broad usage
  - and common enough to be interesting for content producers
- Quality aspects

Bringing 3D home

- But still a lot to do to introduce 3D to private homes
- Cinema systems not applicable to living rooms
- Different usage scenario:
  - Living room is no cinema!
  - Event vs everyday life
  - Comfort much more important
  - Used as “background activity”

Outline

Quality and usage aspects

3D display system overview

3D video coding and rendering

Quality and usage aspects

- Motion parallax:
  - The change of the perspective in accordance to the occurring movement
  - Can be introduced e.g. by using eye-tracking
  - Most only consider horizontal motion parallax, what about vertical?

- Eye fatigue:
  - General fatigue, headache, pain in or around eyes, blurred or double-vision
  - Different causes:
    - Crosstalk
    - Flickering image
    - Cannot focus correctly

Quality and usage aspects

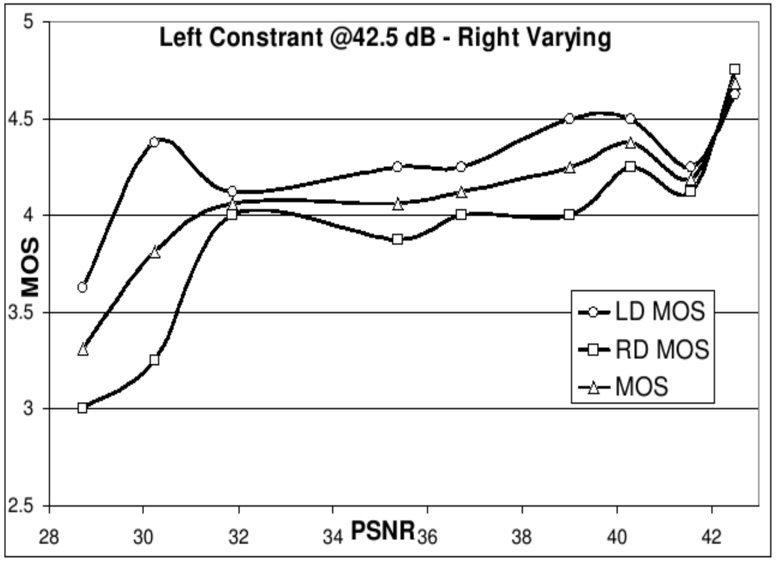
- Other effects:
  - Keystone Distortion
    - Wrong parallax due to incorrectly matched views
  - Puppet Theater Effect
    - Size and distance of an observed object don't seem to match
  - Cardboard Effect
    - Perceived depth to small for perceived size
  - Stereo-Inversion
    - Left eye receives right image and vice versa
  - Picket-fence Effect
    - Moiré effect caused by interference of the screen raster and the 3D filter

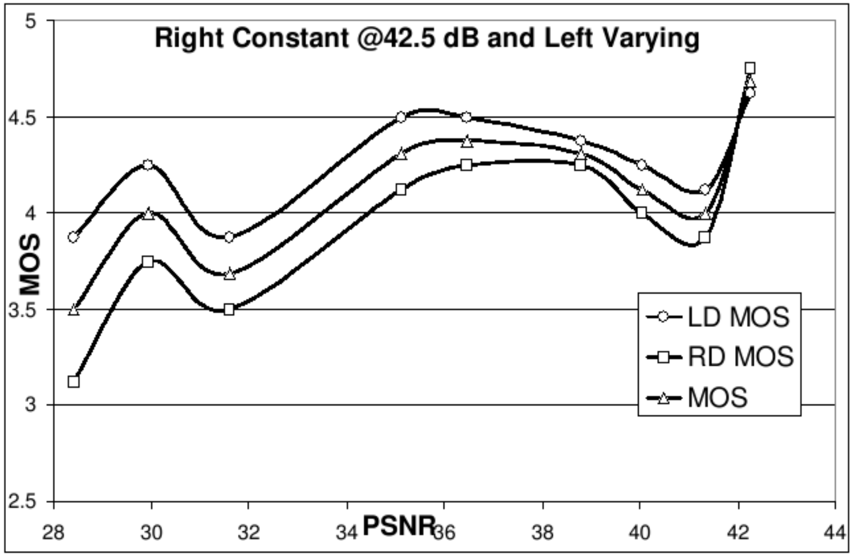
Quality and usage aspects

- Reducing of depth can reduce crosstalk and eye strain, but may not be visible
- Generally: artifacts and distortion less visible as in 2D, except for disparities and blockiness which have a greater impact on the subjective quality and can cause eye strain!

Quality and usage aspects

- Asymmetric coding: the higher quality signal blocks out artifacts and distortion the from the lower quality signal





Quality and usage aspects

- Usage aspects
  - Action movies, live events, sports & concerts
  - Depth impression and switch between 2D/3D key features
  - Main reason is to be entertained, experience of 3D itself less important
  - 3D leads to a higher immersion than 2D => less used as background activity?

Display techniques: Stereoscopic

- Presenting a different image to each eye
- Side-by-side method

Display techniques: Stereoscopic

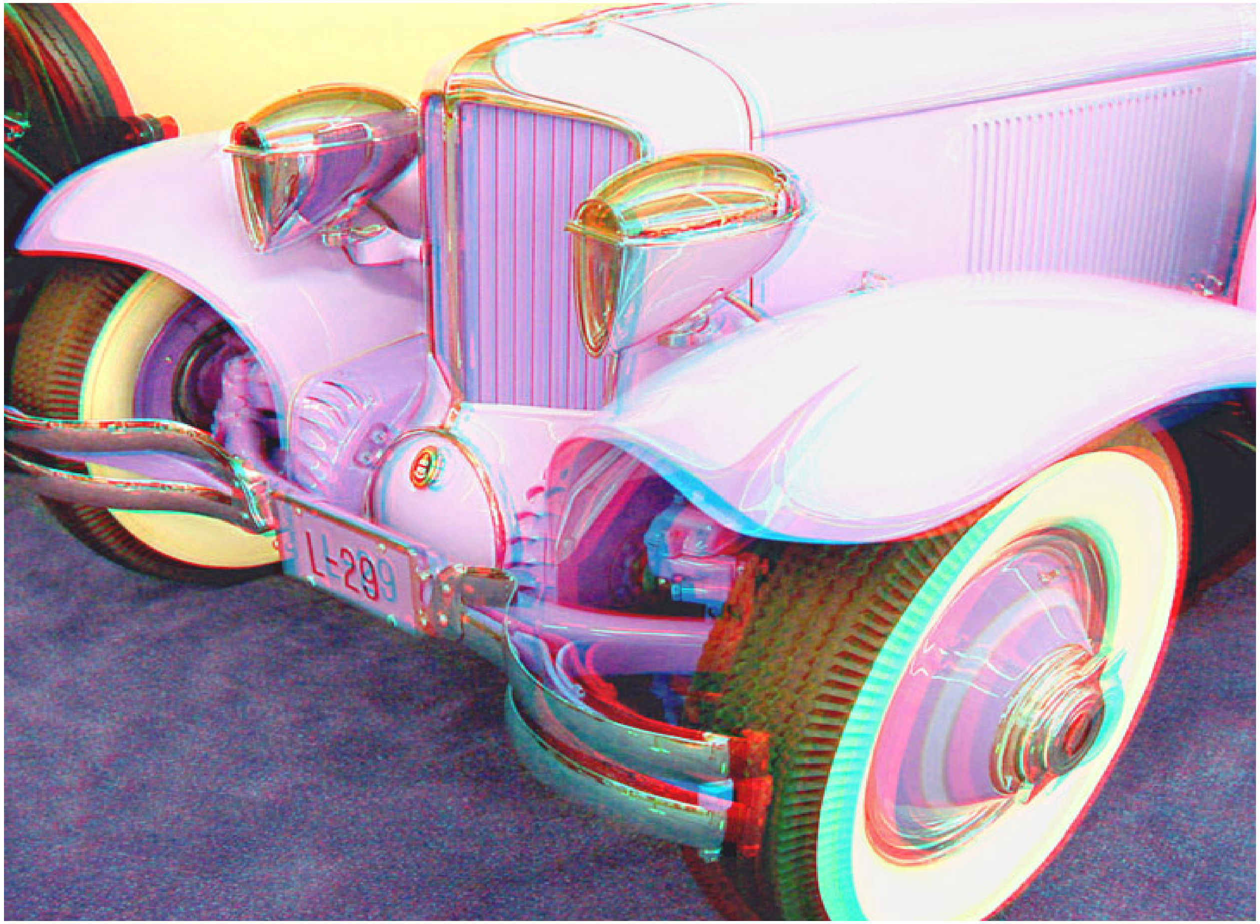
- Presenting a different image to each eye
- Side-by-side method (since 1840)

Display techniques: Stereoscopic

- Presenting a different image to each eye
- Side-by-side method (since 1840)
- Modern variant: VR systems

Display techniques: Stereoscopic (glasses)

- Anaglyph method



Display techniques: Stereoscopic (glasses)

- Anaglyph method
- “color multiplexing”

  - limited colorspace

  - highly susceptible to crosstalk

  + glasses very cheap

  + directly compatible to existing systems

Display techniques: Stereoscopic (glasses)

- Shutter glasses: time multiplex
- Needs synchronization with displaying system
- Need to introduce guardbands to avoid crosstalk
- Used in Nvidias 3D initiative

Display techniques: Stereoscopic (glasses)

- Shutter glasses: time multiplex
- Needs synchronization with displaying system
- Need to introduce guardbands to avoid crosstalk
- and in Segas Master System back in the 80s

Display techniques: Stereoscopic (glasses)

- Shutter glasses: time multiplex
- Needs synchronization with displaying system
- Need to introduce guardbands to avoid crosstalk

  - susceptible to crosstalk even with guardbands due to bad synchronization and after-images

  - expensive glasses due to synchronization logic, glasses battery powered

  - due to “black period” for each eye: flicker (if framerate is too low) and subjective brightness reduction possible

  + partly compatible to existing solutions

Display techniques: Stereoscopic (glasses)

- Polarization multiplex
- 1952: first 3D cinema movie in color

Display techniques: Stereoscopic (glasses)

- Polarization multiplex
- 1952: first 3D cinema movie in color
- Stilled used in IMAX 3D today, but:
- nowadays circular polarization used (to remove crosstalk)

  - reduces brightness by approx. half (depending on projector)

  - needs special, expensive screens (silver or aluminum alloy) (not suited for living rooms)

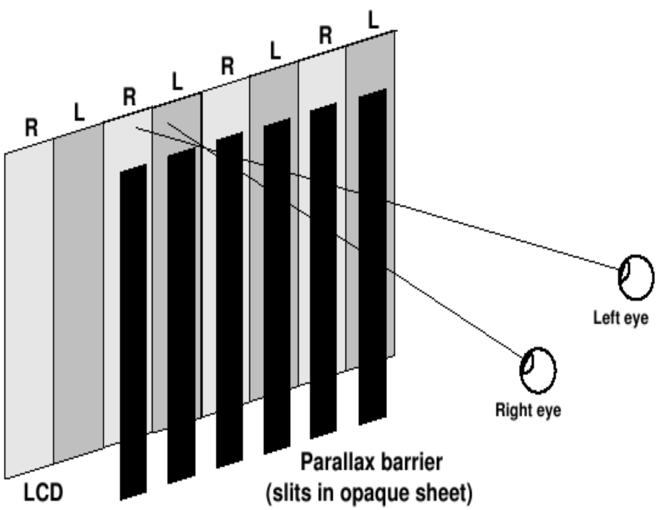
Display techniques: Stereoscopic (glasses)

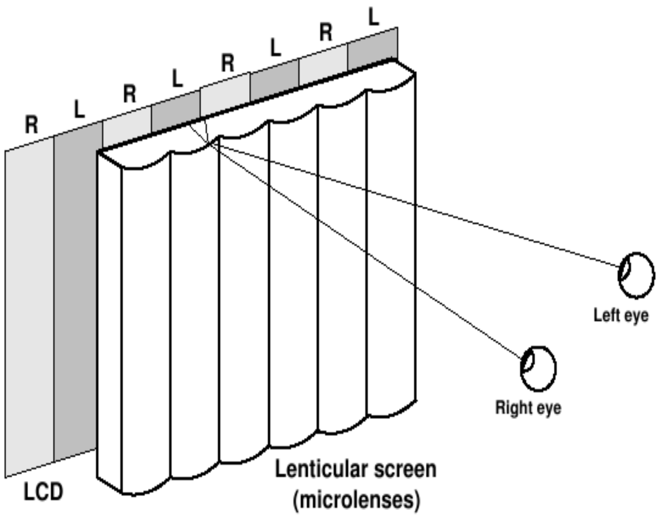
- For own experiments:

  Using cellophane to convert a liquid crystal display screen into a three dimensional display (3D laptop computer and 3D camera phone)

Display techniques: Stereoscopic (displays)

- Autostereoscopic displays
- Parallax barrier vs lenticular sheet





Display techniques: Stereoscopic (displays)

- Autostereoscopic displays
- Parallax barrier vs lenticular sheet
  - Parralax barrier simpler switch between 2D and 3D
  - Lenticular sheet less brightness loss
  - and bigger viewing window
- Brightness vs crosstalk

  - reduces resolution & brightness

  - susceptible to picket-fence effect

  - only one viewer possible, which needs to sit still in one position

Display techniques: Stereoscopic (displays)

- Autostereoscopic displays with eye-tracking
- Move parallax barrier / lenticular sheet or LEDs / projector according to head movement
- Could also be used to introduce motion parallax (in reality however seldom done)

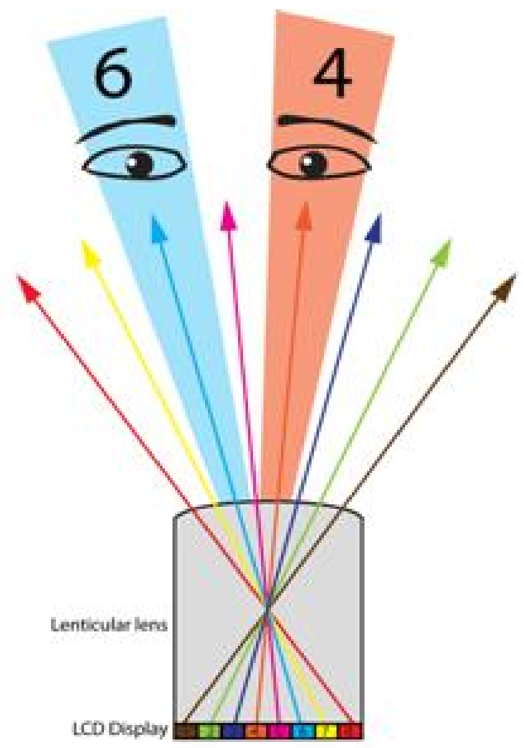
  - reduces resolution & brightness

  - still only one viewer

  + full motion parallax possible

Display techniques: Stereoscopic (displays)

- Autostereoscopic multiview displays
- Project several (e.g. 4, 5, 7 or 8) different views



Display techniques: Stereoscopic (displays)

- Autostereoscopic multiview displays
- Project several (e.g. 4, 5, 7 or 8) different views
- Easier to introduce using lenticular sheets
- Guard band needed to avoid stereo inversion
- Brightness vs crosstalk vs number of views
- Number of views vs resolution
- Already in usage for 3D-CAD, medical appliances but mainly advertisement

Display techniques: Stereoscopic (displays)

- Autostereoscopic multiview displays
- Project several (e.g. 4, 5, 7 or 8) different views

  - reduces significantly resolution and brightness as well

  - susceptible to picket-fence effect

  + partly motion parallax possible

Display techniques: Stereoscopic (displays)

- Autostereoscopic displays with super multiview
- Parallax interval needs to be less than diameter of the pupils of the viewer (ca. 3 to 8 mm)
- Volumetric display?

  - high number of views needed => needs very high DPI, reduces heavily resolution

  - artifacts?

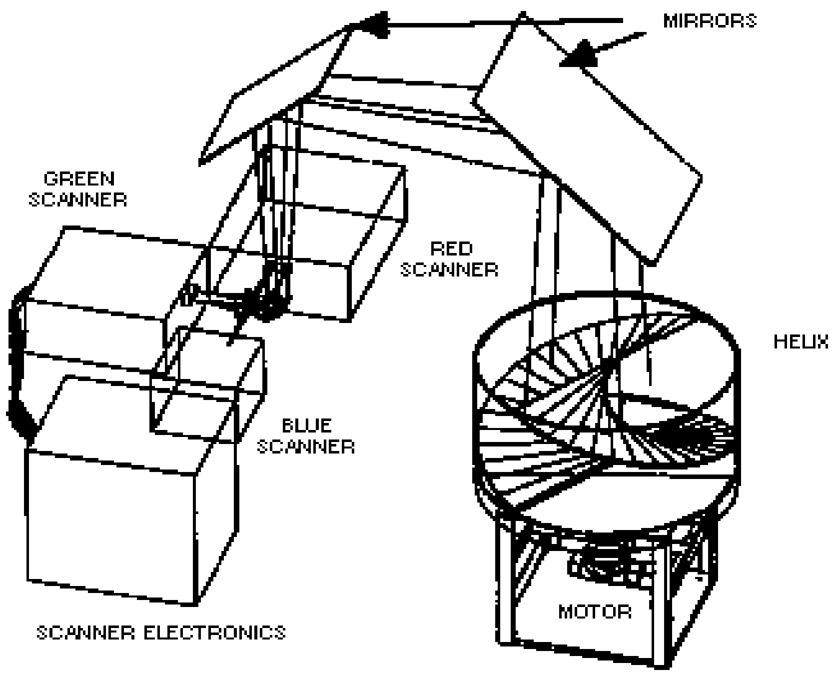
  + fully motion parallax

Display techniques: Volumetric displays

- Swept volume displays:
  - Voxel vs pixel
  - Susceptible to flicker
  - Occlusion hard, if not impossible to introduce

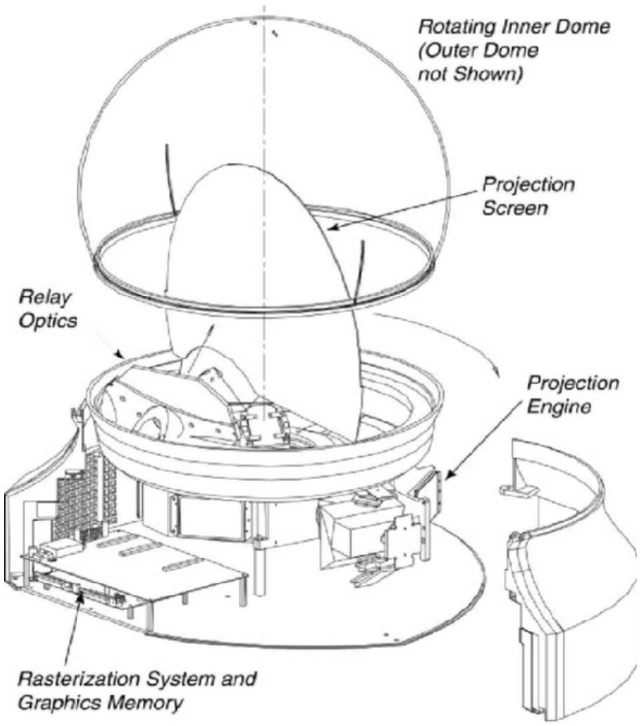
Display techniques: Volumetric displays

- Swept volume displays: projector + mirror



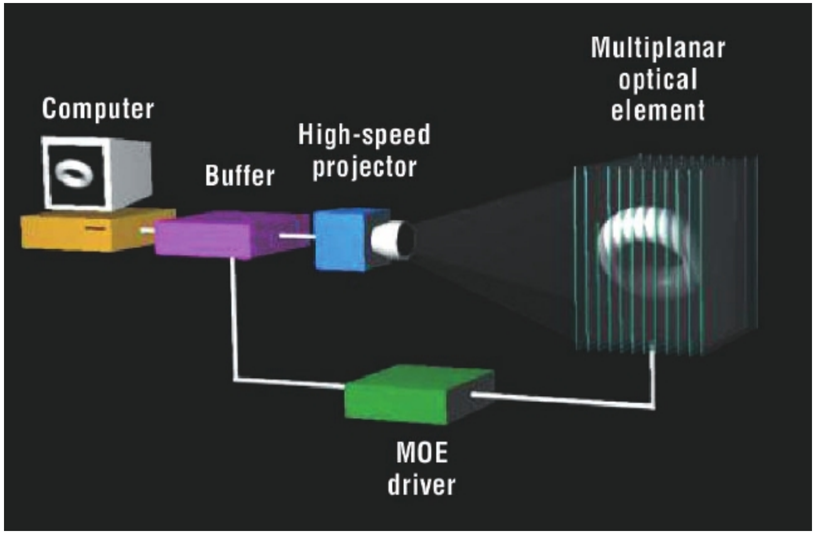
Display techniques: Volumetric displays

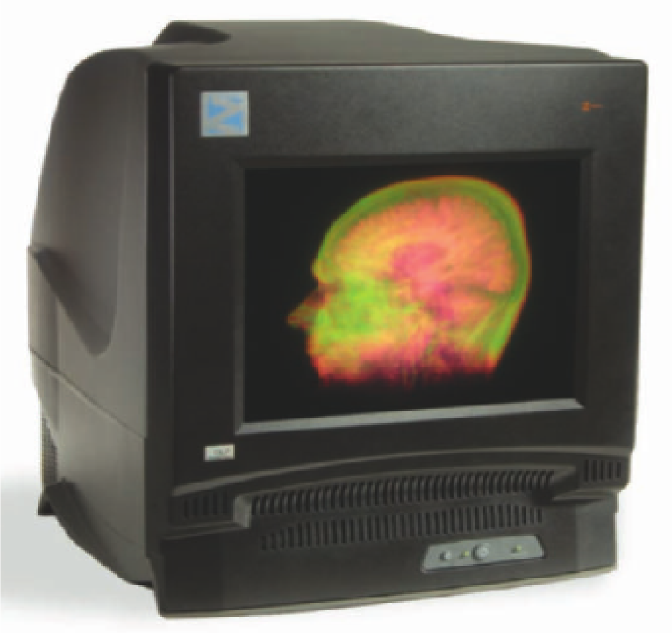
- Swept volume displays: moving display



Display techniques: Volumetric displays

- Swept volume displays: static display

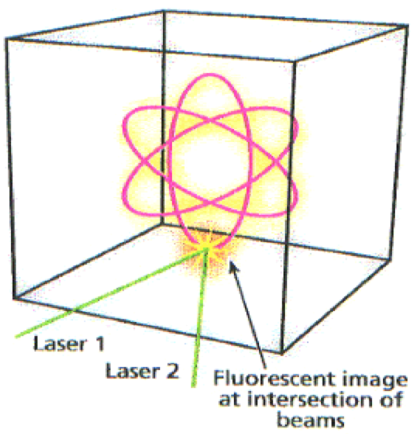


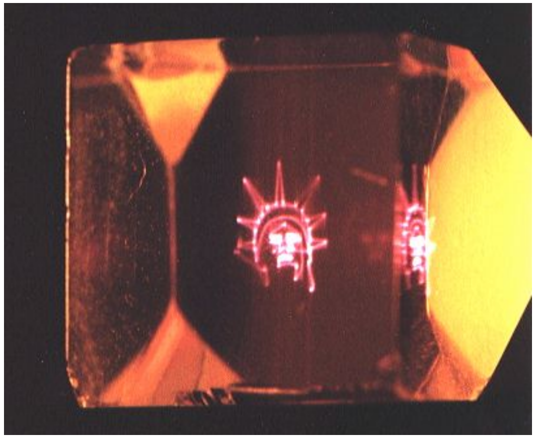


- cardboard effect vs depth resolution

Display techniques: Volumetric displays

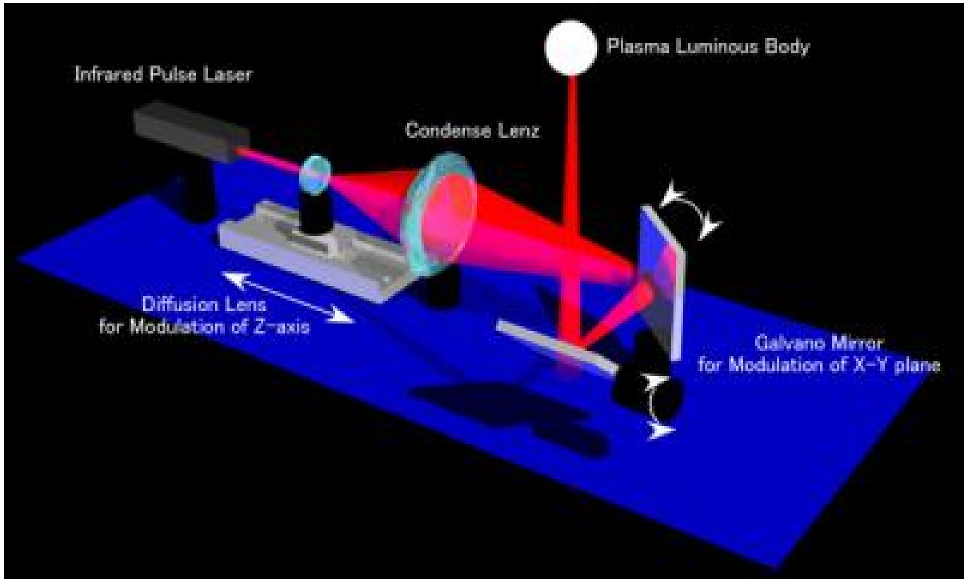
- Swept volume displays: static display
- Laser writing image
- Known since early 70s

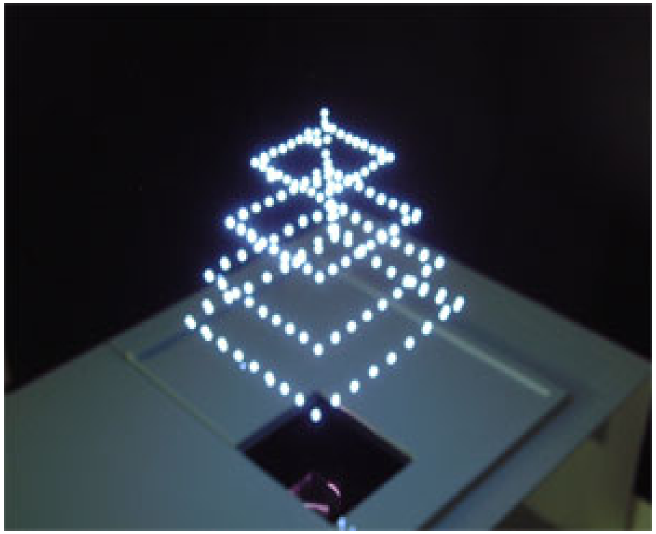




Display techniques: Volumetric displays

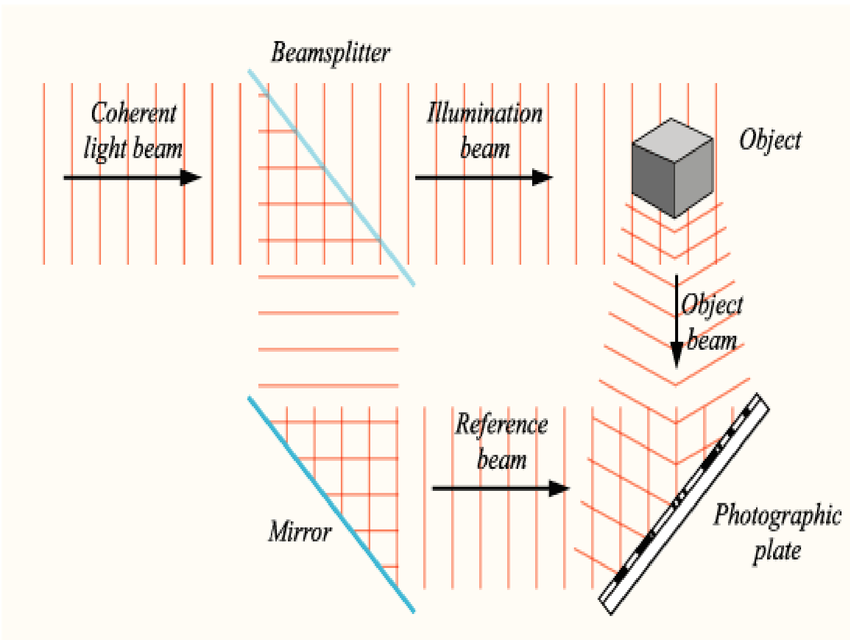
- Swept volume displays: static display
- Laser writing image
- Known since early 70s

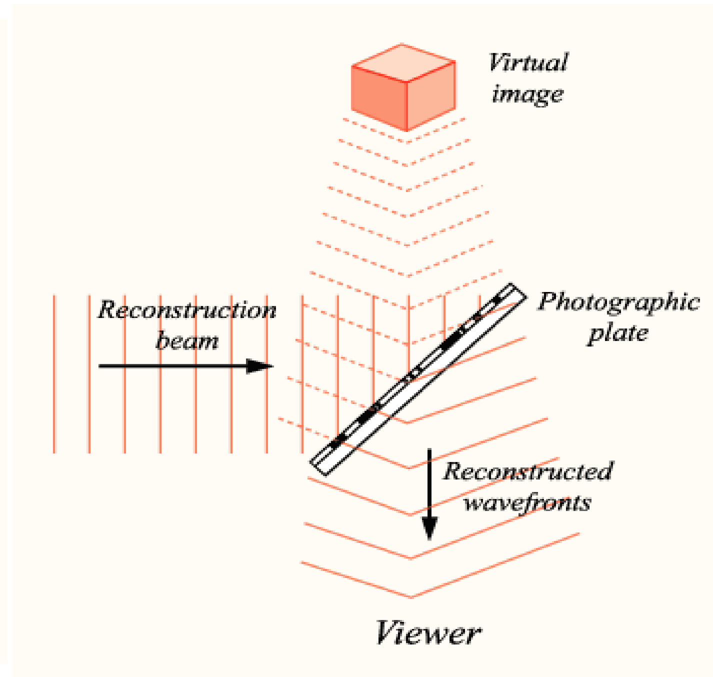




Display techniques: Volumetric displays

- Electroholographic displays

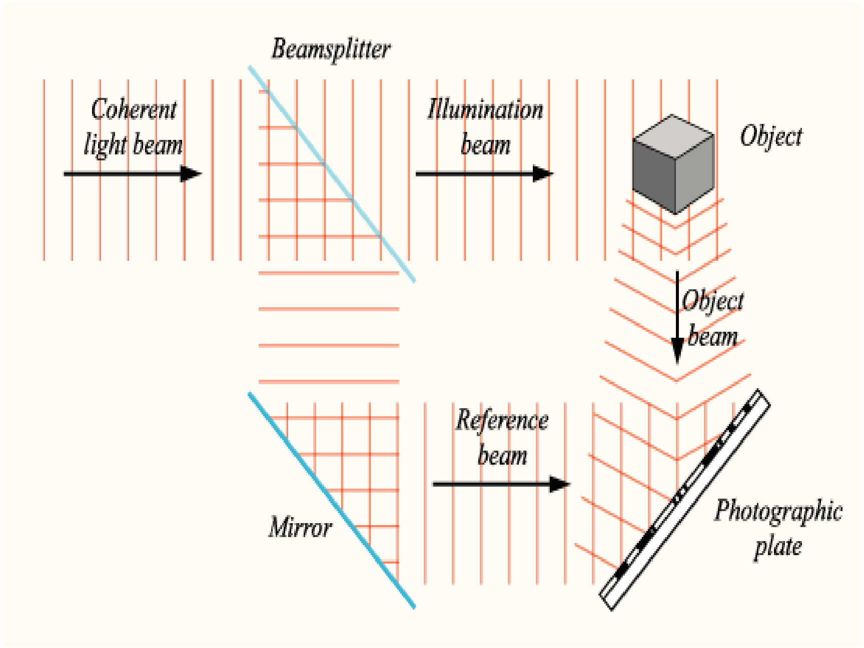


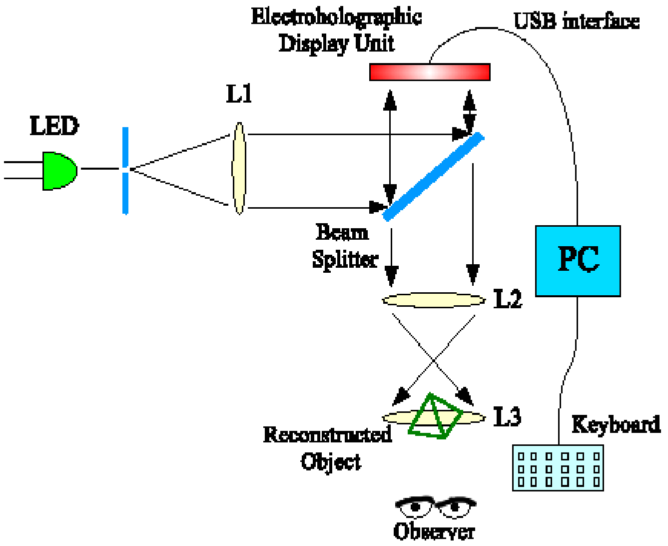


- Safes not only wavelength, but also angle and phase

Display techniques: Volumetric displays

- Electroholographic displays

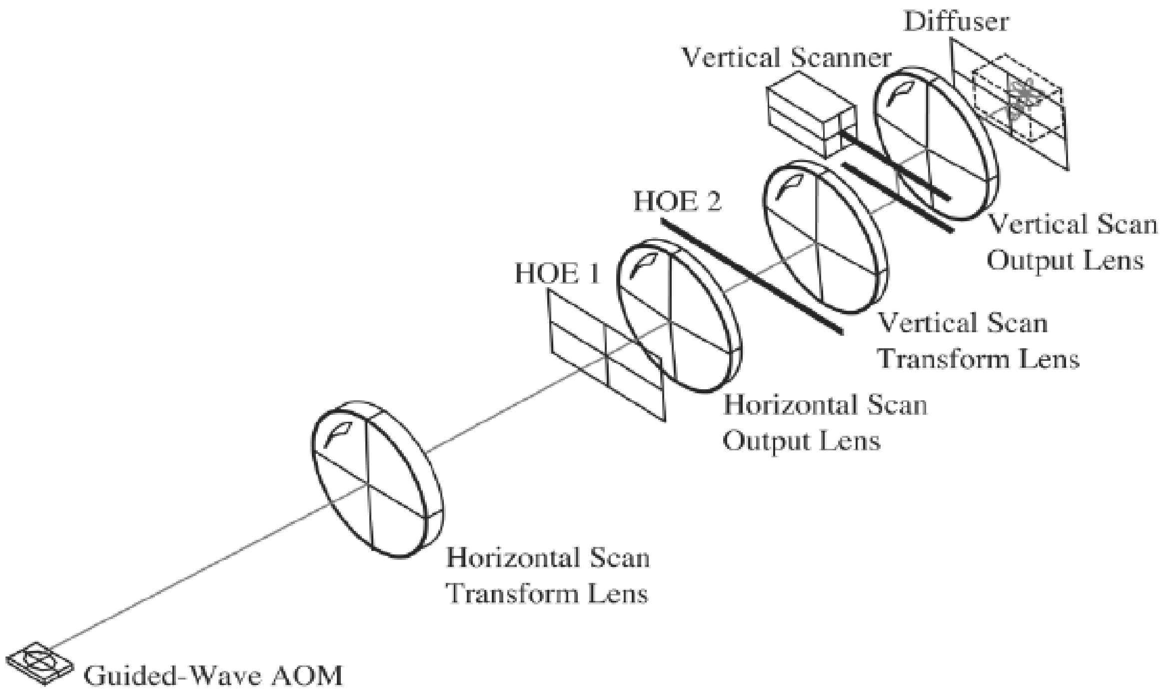




- Safes not only wavelength, but also angle and phase

Display techniques: Volumetric displays

- Holographic displays



- Use acousto-optic modulator: diffracts light, steerable via soundwaves (normal radio frequency are used)

Display techniques: Volumetric displays

- Products on the market, used in industry (3D cat) and for medical appliances

  - susceptible to puppet theater effect

  - many voxels => high demand on computation power and bandwidth

  + full motion parallax

The future?

- No mass acceptance of 3D home cinema in the next few years
- Autostereoscopic displays with supermultiview vs electroholographic displays
- Until then: increasing number of 3D systems among early adopters
- 3D content (movies) will be made available

3D vs multiview vs freeview

- Multiview: different videostreams of same motive but from different perspectives
- 3D: special case of multiview
- Freeview: viewpoint selectable

View stream vs video stream

- View stream: a stream containing all images associated with one particular view
- Video stream: contains one or more view streams

Independent compression of view-streams

- Each view stream is encapsulated in its own video stream

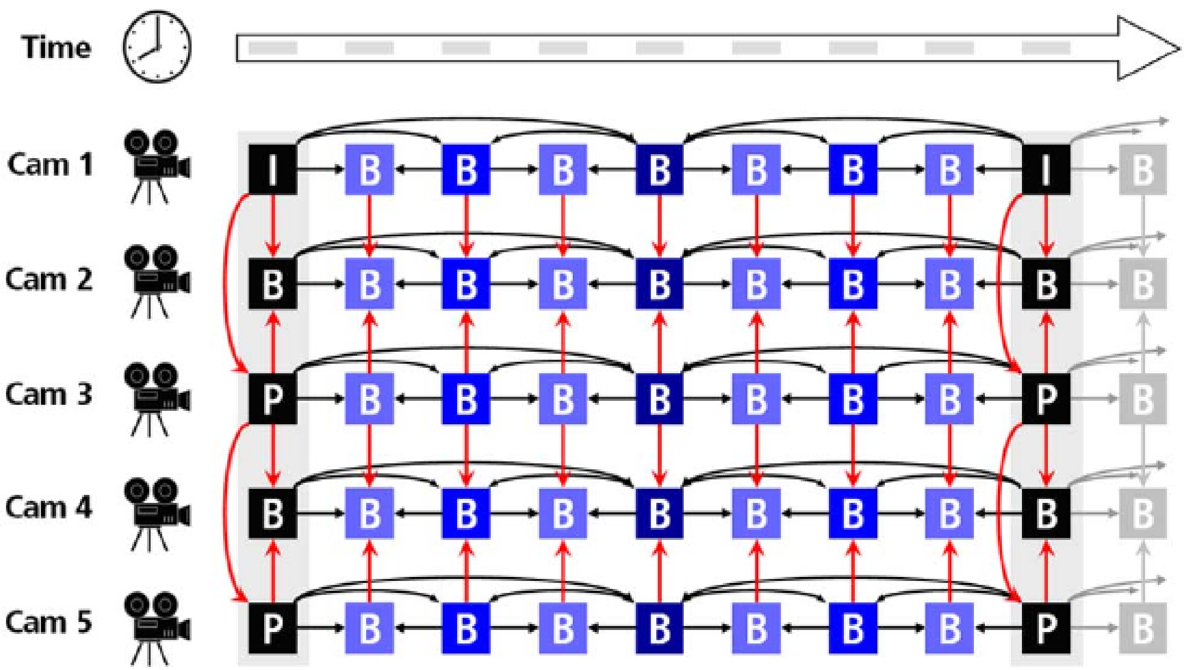
  - very inefficient compression

  + directly compatible to existing systems

  + errors from one view stream cannot spread to another

Combined video-stream

- Exploit spatial redundancies by predicting the images from one view stream by the images of another view stream



Combined video-stream

- Exploit spatial redundancies by predicting the images from one view stream by the images of another view stream
- Camera normalization might be added for further encoding improvement
- Best proposal (yet) for h.264/MVC

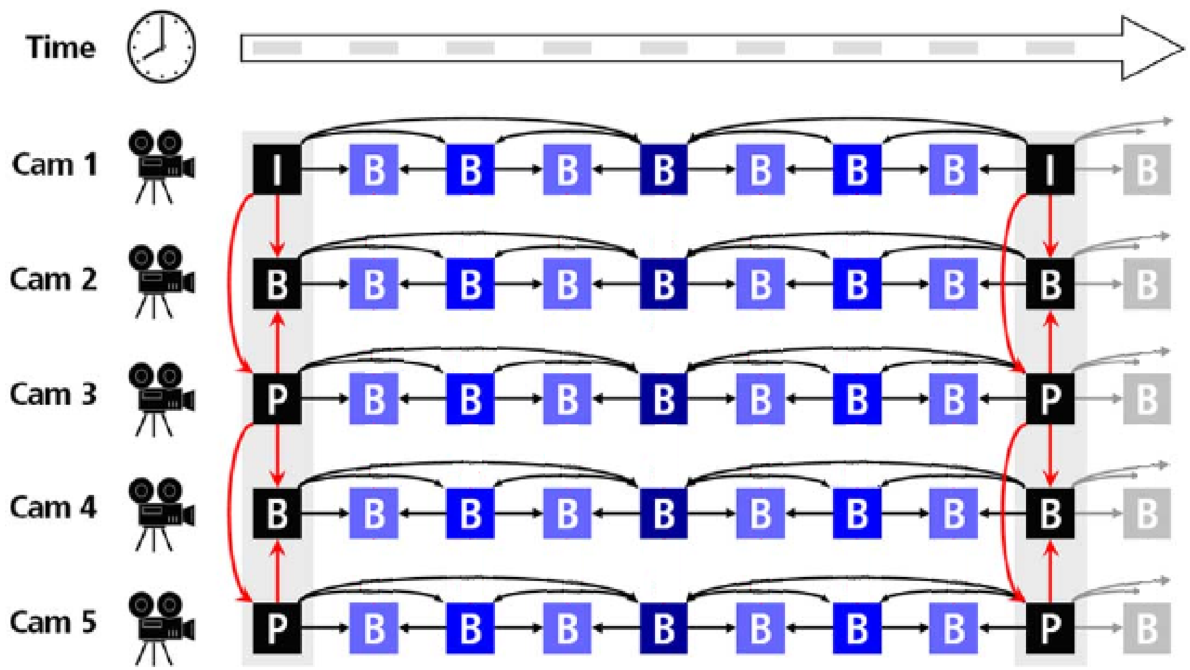
  - very difficult to decode

  + very good coding performance

  + compatible to existing solutions

Combined video-stream

- Most of coding gain from obsessive usage of b-frames and interspatial coding of keypictures
- Omit inter-view decoding for non-key pictures?



Combined video-stream

- Most of coding gain from obsessive usage of b-frames and interspatial coding of keypictures
- Omit inter-view decoding for non-key pictures?
- Much easier to decode
- Results?

Combined video-stream

- Several displays needs to calculate intermediate views (electroholographic systems, autostereoscopic displays with multiview)
- To be able to do that:
  - Camera parameters (3D-position, angle) have to be transmitted as well
- Other (computational) methods exists as well, but are imprecise and / or computational complex

View-stream & depth stream

- Use only one video stream, add a stream containing depth information

View-stream & depth stream

- Use only one video stream, add a stream containing depth information
- Typical depth resolution: 1 byte
- Chosen by European ATTEST project

  - problems with occlusion and reflections (though additional stream might be added carrying that information)

  + very efficient encoding

  + rendering of different view points easily possible

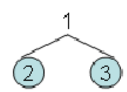
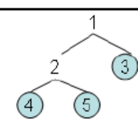
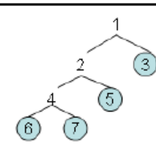
  + compatible to existing solutions

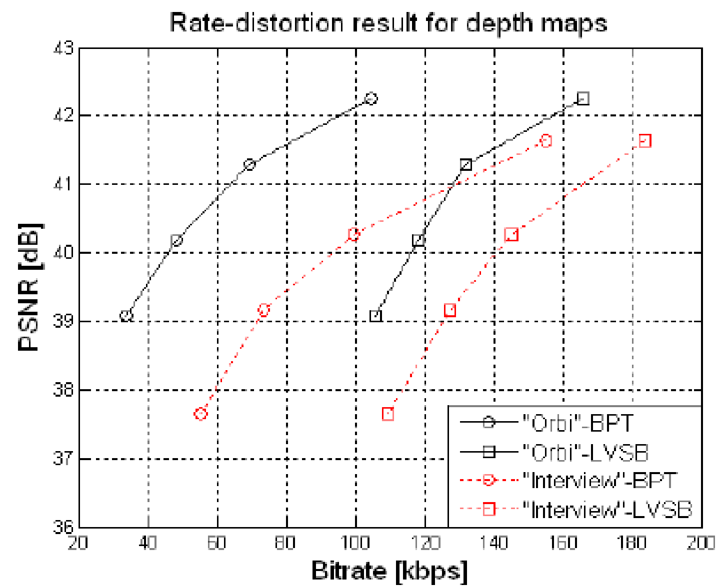
View-stream & depth stream

- How to encode depth stream?
- Possible to use a normal video codec (e.g. h.264)
- Prediction from depth image from the normal one (or vice versa)
- Sharing of motion vectors between view stream and depth image possible, too
- But: depth images other properties as normal ones
  - Consisting of large, smooth areas (less important)
  - and sharp edges (more important)
- Depth image coding profits from a variable blocksize

View-stream & depth stream

Flexible Motion Model with Variable Size Blocks for Depth Frames Coding in Colour-Depth Based 3D Video Coding

Tree-growing iterations	Original frame and its partitioning	BPT and the energy of motion compensated frame
0		 E = 53.39
1		 E = 51.77
2		 E = 42.37
3		 E = 38.02



Combined approaches

- Several view-streams and their accompanying depth streams
- Choose which views depending on occlusion in some views (ideally)
- Camera parameters might needed to be transmitted as well
- Results?

  - higher bitrate

  + suited to deal with different viewpoints and occlusion

Object based approaches

- Object based video coding has been a topic for quiet some time
  - Promises high data-rates
  - Proved to be problematic in practice

- Describe scene as background and several foreground objects

  + theoretically high data-rates

  + easy rendering of different viewpoints

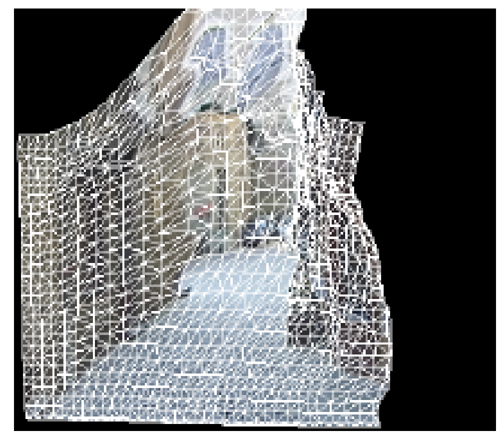
  - hard to do in practice

Object based approaches

- Scalable and Efficient Video Coding Using 3-D Modeling
  - Uses depth map to generate a 3D mesh, which is encoded using second-generation wavelets


a


b


c

Object based approaches

- Scalable and Efficient Video Coding Using 3-D Modeling
  - Uses depth map to generate a 3D mesh, which is encoded using second-generation wavelets
  - Mesh reused for several pictures, updated during transmission
  - Textures are encoded using EBCOT (blockbased coder using wavelets and arithmetic coding) and IPP scheme
  - Camera parameters are transmitted as well
  - Results?

4D Wavelet

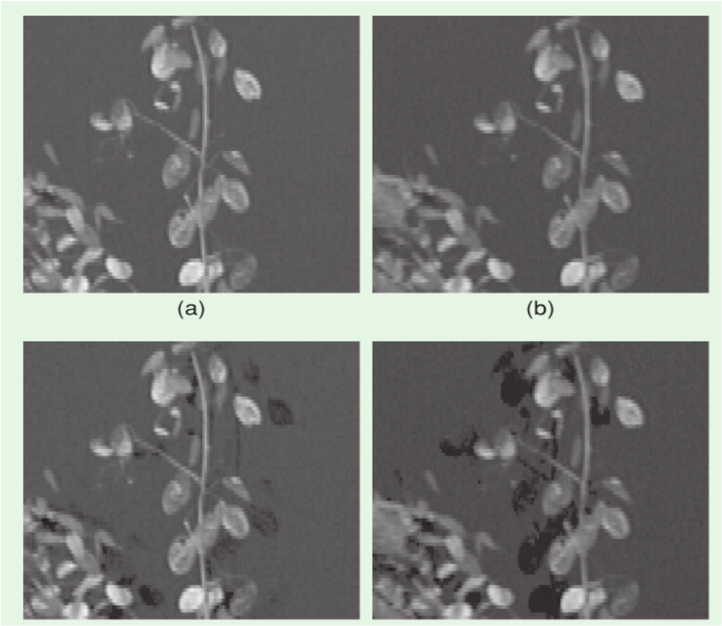
- Consider multiview video stream as 4D matrix of pixels
- Encode together using one single 4D wavelet
- But: spatial and temporal redundancy very different
- Practical approaches: do temporal and spatial prediction separately (using different techniques), a final wavelet on the data collected
- Coding performance of the approaches comparable, roughly the same as for the combined video approach
- Computational complexity?

Rendering Issues

- Hole filling
  - If occlusion occurs while rendering an intermediate view (most likely when using view-stream+depthmap)
  - Generally: use depth of the pixel around the hole which is most in the background
  - Use texture of the pixel whose depth was used: but only one color or whole patterns?
  - Possibility to use information of other images in the video stream?

Rendering Issues

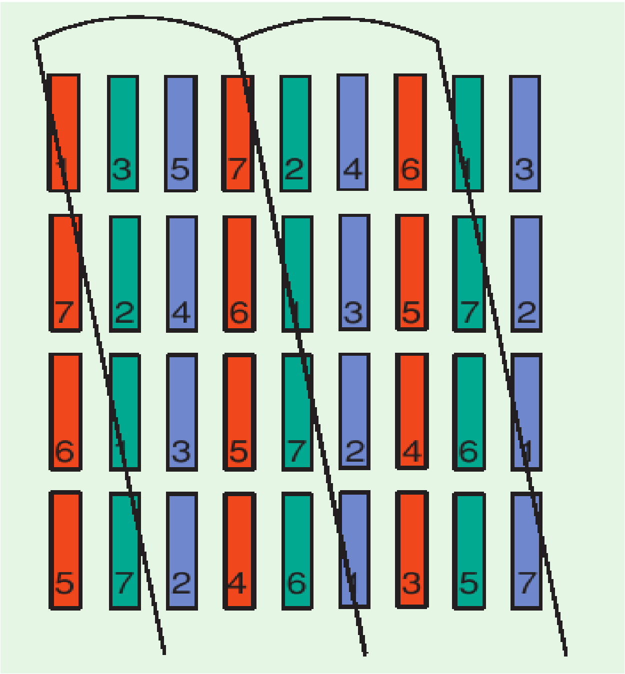
- Cross-talk compensation:
  - Add distortion supposed to cancel out “ghosts”



  - Need to know display parameters
  - Not possible for all ghosts

Rendering Issues

- Rendering for multiview autostereoscopic displays
  - Often: lenticules oriented at small angle



- Reduces picket-fence artifacts
- Balances resolution loss in both directions

Rendering Issues

- Rendering for multiview autostereoscopic displays
  - But how to map the view images to the screen?
  - Approximate by subsampling on a lattice or a union of lattices
  - Alt. approach: approximate in frequency domain
  - Some diodes may get no corresponding pixels
  - Anti-aliasing filter should be added

Rendering Issues

- Rendering for electroholographic displays
  - Need to calculate fringes which determines how to diffract the light
  - Direct computation of fringes too slow for realtime applications
  - Transform 3D scene to holographic plane
    - Generate hogels: small enough to appear to the views as a point, contains color and brightness information
    - Corresponding hogel-vector: contains all information necessary to generate the diffraction necessary for the hogel
  - Combine with precomputed basic fringes to create the physically usable fringes
  - Further compression possible

Rendering Issues

- Rendering for electroholographic displays
  - Example: RIP algorithm (for electroholographic displays)
    - Reconfigurable Image Projection
    - Projects one or more views through a holographic-reconstructed image plane

Thank you very much!

www.icg.isy.liu.se