# Theory and Practice for HDTV-quality Stereo Animations in the Sciences

***Joel R. Stiles & Stuart M. Pomerantz***

***Arthur Wetzel, Greg Hood, David Deerfield***

PITTSBURGH SUPERCOMPUTING CENTER

*Biomedical Supercomputing Initiative*

## Stereo Imaging - The Goal

***Viewers perceive a real scene with spatial depth***

Corollaries:

- Arbitrarily complex scenes (scientific models)
- High resolution (HDTV)
- Still images or full speed animations (30 fps)
- Full color (24-bit)
- Comfortable, safe viewing
- Seamless integration with non-stereo content
- Easy and inexpensive for large audiences

***To perceive stereo depth, different images must be viewed simultaneously by the left and right eyes.***

## Stereo Imaging - The Problem

"Whether it's been attempted in film, television, or PC games, stereoscopic viewing has been relegated to novelty status. The technology has elicited both curiosity and nausea, but it has never taken hold among consumers as a useful innovation."

PC Magazine. After Hours. Dec 3, 2002, p. 190

***The same statements hold equally well for the scientific, business and high performance computing communities.***

## Stereo Imaging - The Problem (cont.)

***Why has it not taken hold?***

1. Display technology limitations
2. High quality stereo content is difficult to create (perceptual and rendering issues)
3. *#2 can't be solved until #1 is solved*

## Movie Demonstrations

- Aldehyde Dehydrogenase
- OpenDX Tutorial Dataset
- MCell Tutorial (Diffusion)
- Visible Human
- Synaptic Pathophysiology
- Large-scale Synaptic Reconstruction

## Overview of the Presentation

- Display Technology
- Stereo Content Creation
- Encoding HDTV Stereo Animations
- PSC Playback & Display System
- Movie Demonstrations

# Display Technologies

**(1) *Alternating display* of left and right images**

- Requires synchronized shutters in front of the eyes (so-called "active stereo")
- Stresses display device speed/resolution
- Shutter glasses are complex, heavy, expensive
- Flickering images are not good for the visual system

# Display Technologies (cont.)

**(2) *Continuous display* of left and right images**

- Side-by-Side
  - Mechanical or Optical views (*the viewmaster*)
  - Free Fusion
    - Right/Left – cross-eyed viewing
    - Left/Right – wall-eyed viewing
- Superimposed
  - Anaglyph – left and right images in contrasting colors, viewed through two correspondingly colored filters (*simple but limited quality*)
  - Polarized – left and right images projected using polarized light, viewed through two correspondingly polarized filters (so-called "passive stereo", *simple, safe, underutilized*)

# Display Technologies – Summation

***Historical Recap – Stereo visualization has been driven primarily by the gaming and supercomputing communities.***

- Gaming – focus is on single user, interactive applications ➔ active stereo, single display ($)
- Supercomputing – focus is on immense data sets, interactive rendering, ultrahigh resolution (64 Mpixel) ➔ parallel computing, active or passive stereo, tiled displays (alignment issues), considerable manpower ($$$$$)

***Neither scenario satisfies the practical goals of generalized stereo visualization, for which the best choice is a single passive display at HDTV resolution (1-3 Mpixels; $$$ and decreasing).***

# Stereo Content Creation (1)

***Depth Perception – how do we see in stereo?***

*Parallax* is "the apparent change in the direction of an object, caused by a change in observational position that provides a new line of sight."

The American Heritage Dictionary of the English Language, 4th Ed. (2000)

***Computer generated (stereo) images must duplicate the parallax of real-world binocular vision.***

# Stereo Content Creation (2)

***Depth Perception – how do we see in stereo?***

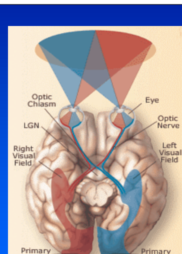
Binocular vision – overlapping views with horizontal parallax

Convergence of gaze

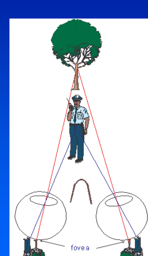
Accommodation (focusing through depth)

Depth of field

Fusion of left & right images in visual cortex



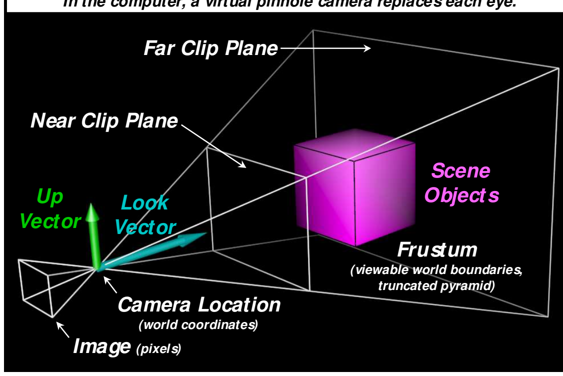

http://www.hhmi.org/senses/popups/images/b410_1p.gif

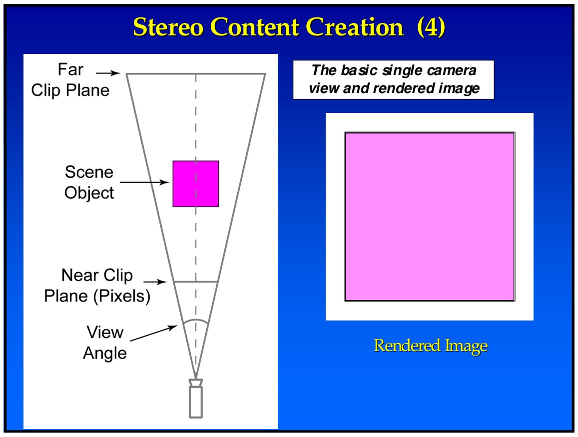


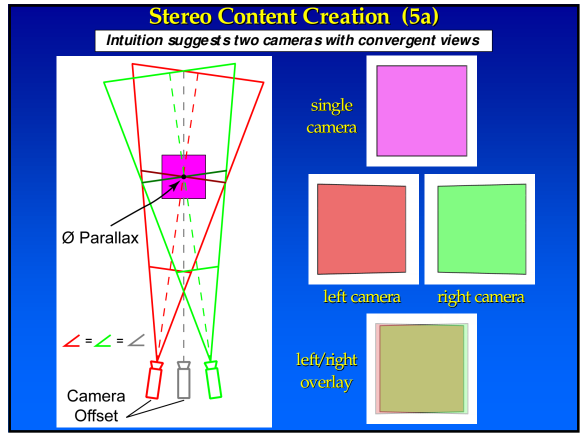

http://admin.vmi.edu/ir/Dictionary/B/b.htm5.gif

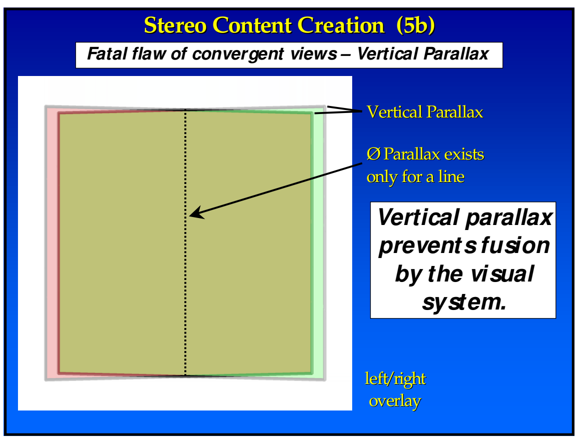
# Stereo Content Creation (3)

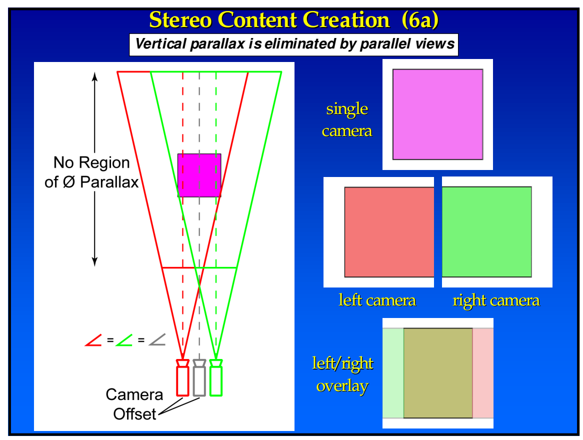
***In the computer, a virtual pinhole camera replaces each eye.***

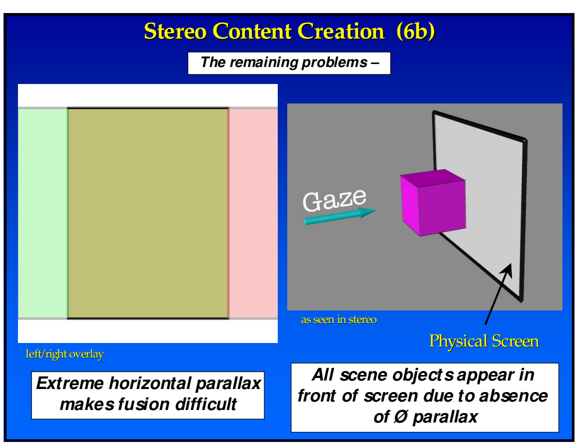


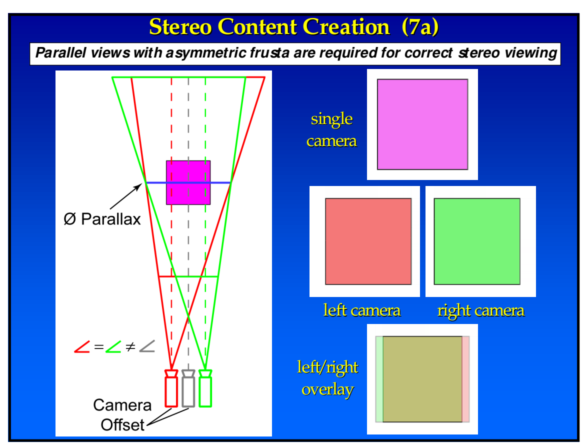

Stereo Content Creation (4)
Far Clip Plane
Scene Object
Near Clip Plane (Pixels)
View Angle
The basic single camera view and rendered image
Rendered Image

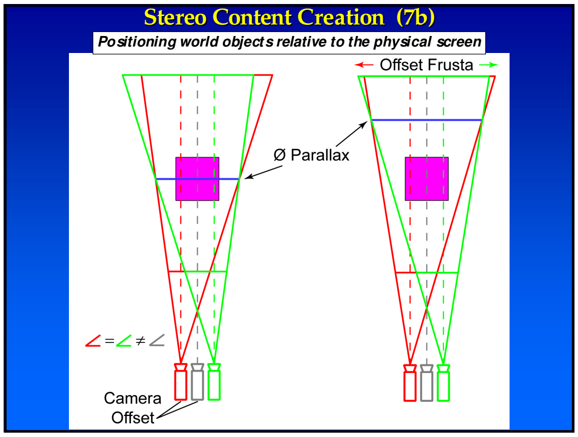

Stereo Content Creation (5a)
Intuition suggests two cameras with convergent views
Ø Parallax
Camera Offset
single camera
left camera
right camera
left/right overlay

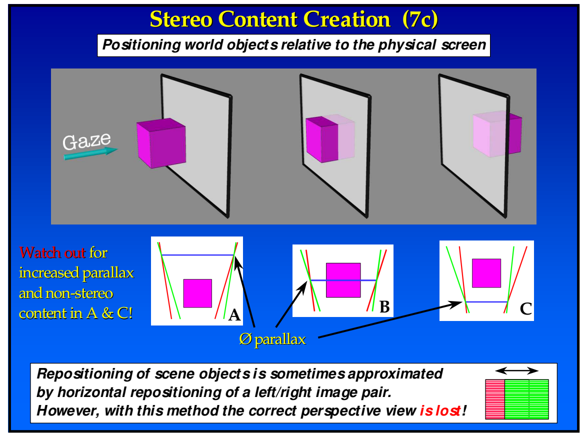

Stereo Content Creation (5b)
Fatal flaw of convergent views – Vertical Parallax
Vertical Parallax
Ø Parallax exists only for a line
Vertical parallax prevents fusion by the visual system.
left/right overlay

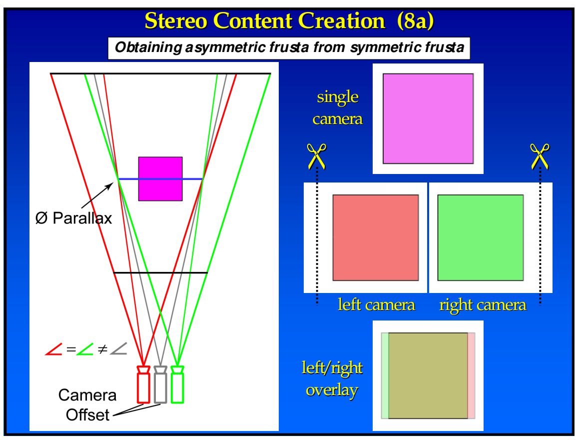

Stereo Content Creation (6a)
Vertical parallax is eliminated by parallel views
No Region of Ø Parallax
Camera Offset
single camera
left camera
right camera
left/right overlay

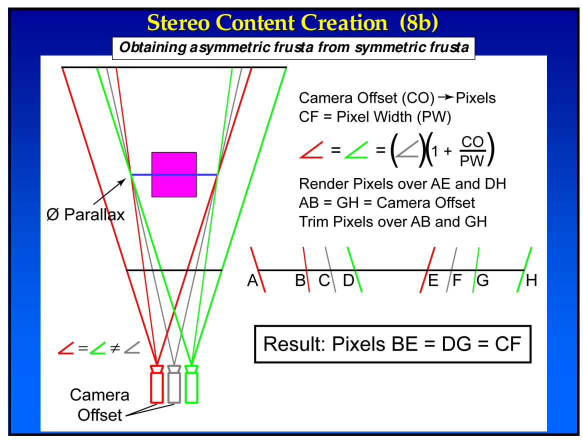

Stereo Content Creation (6b)
The remaining problems –
Gaze
as seen in stereo
Physical Screen
left/right overlay
Extreme horizontal parallax makes fusion difficult
All scene objects appear in front of screen due to absence of Ø parallax

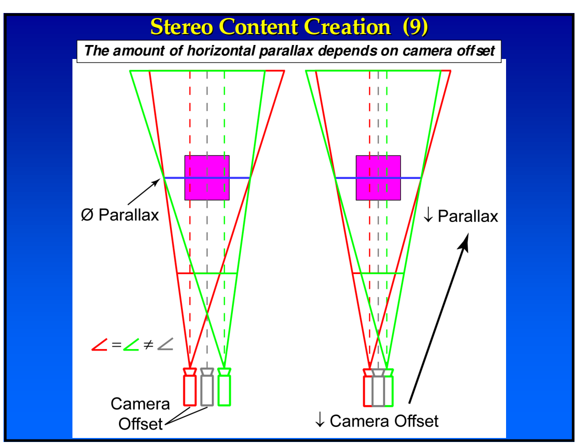

Stereo Content Creation (7a)
Parallel views with asymmetric frusta are required for correct stereo viewing
Ø Parallax
Camera Offset
single camera
left camera
right camera
left/right overlay

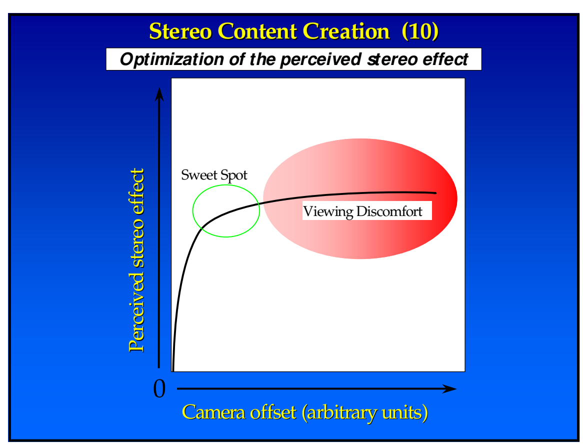

Stereo Content Creation (7b)
Positioning world objects relative to the physical screen
Offset Frusta
Ø Parallax
Camera
Offset


Stereo Content Creation (7c)
Positioning world objects relative to the physical screen
Gaze
Watch out for increased parallax and non-stereo content in A & C!
A
B
C
Ø parallax
Repositioning of scene objects is sometimes approximated by horizontal repositioning of a left/right image pair.
However, with this method the correct perspective view is lost!


Stereo Content Creation (8a)
Obtaining asymmetric frusta from symmetric frusta
Ø Parallax
single camera
left camera
right camera
left/right overlay
Camera
Offset


Stereo Content Creation (8b)
Obtaining asymmetric frusta from symmetric frusta
Camera Offset (CO) → Pixels
CF = Pixel Width (PW)
(1 + CO/PW)
Render Pixels over AE and DH
AB = GH = Camera Offset
Trim Pixels over AB and GH
A B C D E F G H
Result: Pixels BE = DG = CF
Ø Parallax
Camera
Offset


Stereo Content Creation (9)
The amount of horizontal parallax depends on camera offset
Ø Parallax
↓ Parallax
Camera
Offset
↓ Camera Offset


Stereo Content Creation (10)
Optimization of the perceived stereo effect
Perceived stereo effect
Sweet Spot
Viewing Discomfort
0
Camera offset (arbitrary units)

## Stereo Content Creation – Summary

- Convergent view method is wrong (vertical parallax)
- Horizontal repositioning of images is wrong (incorrect perspective)
- Parallel views and asymmetric frusta are required
- Frusta can be offset to reposition world objects with respect to the physical screen (*Caveats*: may be difficult to fuse, increases non-stereo image content)
- Camera offset controls the absolute amount of horizontal parallax
- Use camera offset to find the sweet spot for perceived stereo effect

## Encoding HDTV Stereo Animations (1)

**Example data rate for uncompressed movie:**

**1024 x 1024 pixels = 1 Mpixel per frame**
**1 Mpixel x 3 bytes = 3 MBytes per frame**
**3 MBytes x 30 fps = 90 MBytes per sec**
**90 MBytes x 2 channels = 180 MBytes per sec**

- ***Need to use compression (codec) to reduce data rate***
- ***Codec must preserve quality (e.g., color gradients)***
- ***Decoding must be full speed (30 fps) for two channels***

## Encoding HDTV Stereo Animations (2)

**The MPEG (Moving Pictures Experts Group) Codecs:**

MPEG encoding is *lossy*. Picture quality *decreases* as compression *increases*.

MPEG-1 and -2 use interframe compression, i.e., some frames are encoded based on changes from previously encoded frames(s).

MPEG-1 (ISO/IEC 11172) - 1.5 Mbits/sec; suitable for CD-ROMs and VideoCD applications (typically up to 352 x 240 pixels for NTSC)

MPEG-2 (ISO/IEC 13818) - 2 to 10 Mbits/sec (and up); DVD, computer video, digital satellite, HDTV systems. For example, DVD supports up to 9.8 Mbits/sec, but generally uses a variable rate ~4.7 Mbits/sec (720 x 480 pixels for NTSC)

## Encoding HDTV Stereo Animations (3)

**MPEG-2 Profiles and Levels**

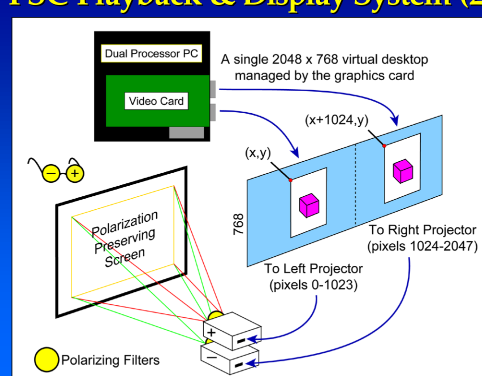
| | Simple | Main | SNR | Spatial | High |
|---|---|---|---|---|---|
| **Low** | | 352 x 288<br>4 Mbps | 352 x 288<br>4 Mbps | | |
| **Main** | 720 x 576<br>15 Mbps | 720 x 576<br>15 Mbps | 720 x 576<br>15 Mbps | | 720 x 576<br>20 Mbps |
| **High 1440** | | 1440 x 1152<br>60 Mbps | | 1440 x 1152<br>60 Mbps | 1440 x 1152<br>80 Mbps |
| **High** | | 1920 x 1152<br>80 Mbps | *How to decode ??* | | 1920 x 1152<br>100 Mbps |

## PSC Playback & Display System (1)

**Goal: Read *two* MPEG-2 streams that may be larger than available RAM, decode, and display the synchronized images at 30 fps.**

1. Read Data
   a) Large files (up to 16 Terabytes per eye) require secondary buffering
   b) Networked video also requires secondary buffering
   c) Asynchronous
2. Decoding Software
   a) Bypasses video hardware – uses dual CPUs
   b) Thread safe
   c) Asynchronous
   d) Interleaved 2D and stereo content in both channels
   e) Supports concatenated files
   f) Optional video profiling

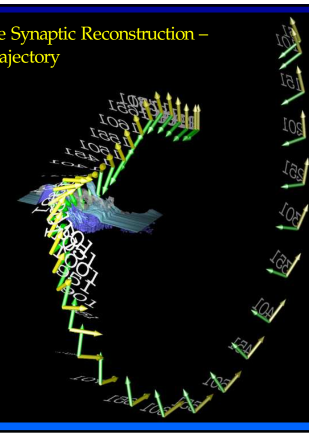
## PSC Playback & Display System (2)



## Movie Demonstrations

- Aldehyde Dehydrogenase
- OpenDX Tutorial Dataset
- MCell Tutorial (Diffusion)
- Visible Human
- Synaptic Pathophysiology
- Large-scale Synaptic Reconstruction

Large-scale Synaptic Reconstruction –
Camera Trajectory



Large-scale Synaptic Reconstruction –
Camera Offset