BUILDING TELEPRESENCE SYSTEMS:
Translating Science Fiction Ideas into Reality

Henry Fuchs

University of North Carolina at Chapel Hill (USA)

and

NSF Science and Technology Center for

Computer Graphics and Scientific Visualization

Major support is gratefully acknowledged

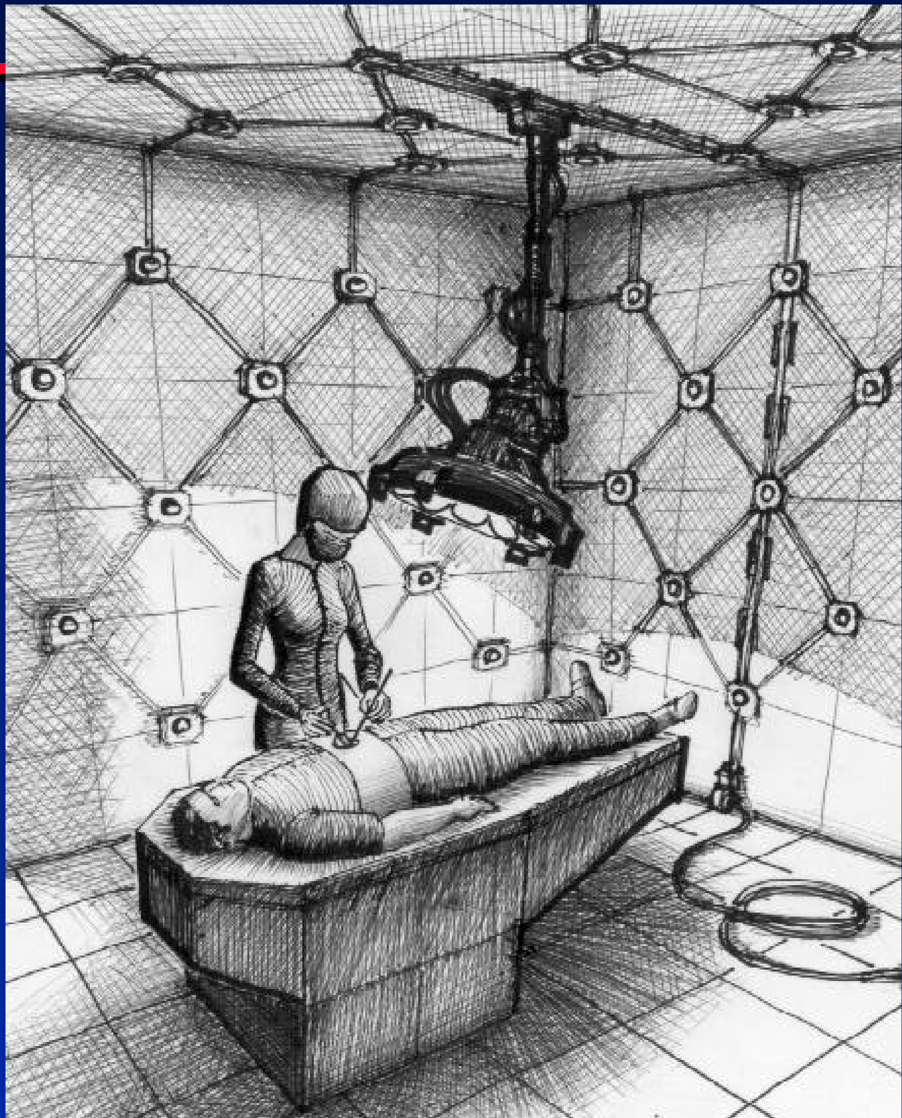
from the U.S. National Science Foundation and

from the U.S. Defense Advanced Research Projects Agency

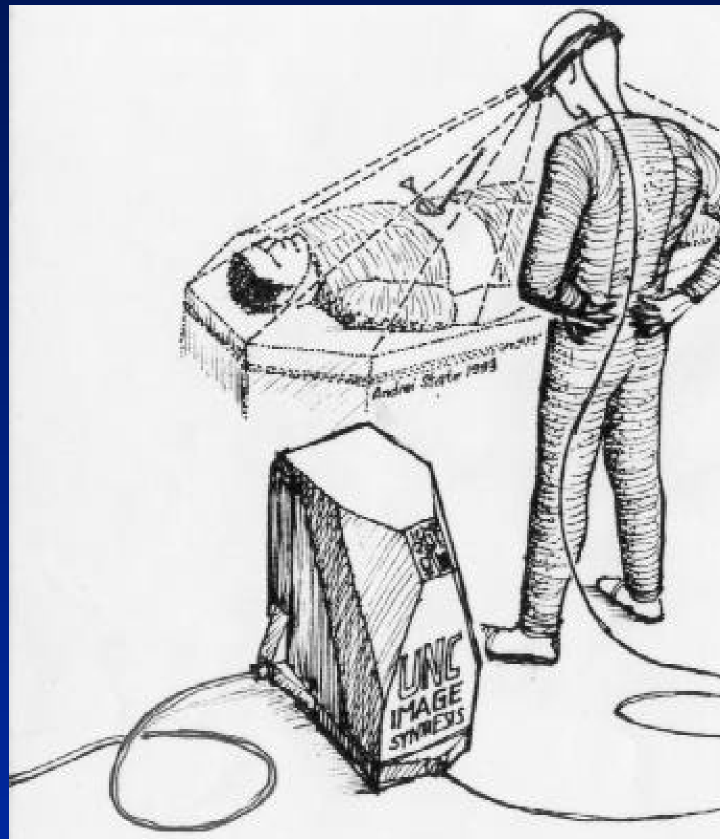
Introduction

- The dominant grand challenge of graphics in the past 30 years has been realism, esp. photorealism
- Briefly, in the past decade, being in a virtual world, captured the public imagination
- Next, to be immersed (at least partly), in a far-away place/ with far-away people
- Driving examples: telemedicine, telecollaboration (MCAD) and laparoscopic surgery

Initial Concepts:Visual Telepresence (1993)

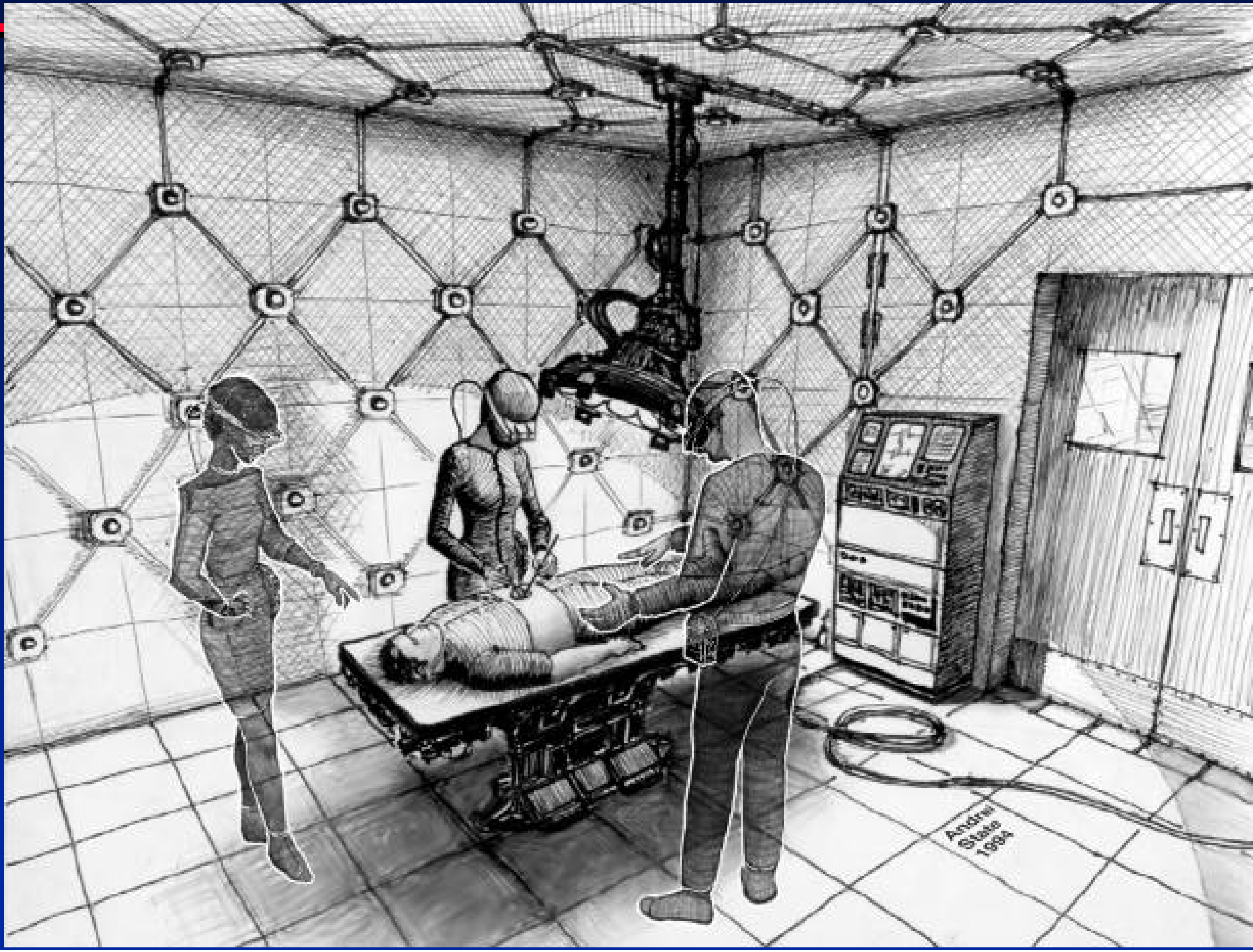


Local MD & Patient



Remote Consultant

Medical Consultants "Together" with Local MD & Patient



Problem: Too Difficult (for now)

- Real-time 3D scene capture at each of the sites
- Presentation of local plus remote scenes on head-mounted displays to each of the participants
- Other tasks: image generation, head and hand tracking, etc.   are easy by comparison

Solution: Work on an easier problem first

- Eliminate need for head-mounted display
- Reduce need of scene capture
  - to smaller regions of the rooms
  - to reconstruction (and viewing) from fewer places
- New, easier problem:

  Advancing teleconference-based TELECOLLABORATIONS toward TELEPRESENCE

Our Vision of Telecollaboration: A Normal Office

Our Vision of Telecollaboration: Overlapped Projected Displays

Our Vision of Telecollaboration: Seeing and Manipulating Objects

Stereo via shutter / prescription glasses

Displays can light up and cover the entire room

Our Vision of Telecollaboration: "Being There" Together

What Will It Take: major areas

- Displays: fixed, not head-mounted
- 3D scene capture
- Image generation system
- Tracking system

Displays: New opportunities with Micromirror-based displays

- Physical micro-mirrors on custom IC
  - 800 x 600 pixel resolution typical
  - Commercial product from Texas Instruments
- One bit of memory behind each mirror
- Consider it as part of the memory-space of the graphics system, not as a separate projector
- Use for both display and for (lighting to aid) scene capture

Displays: Fixed, large visual area

- Adapt to user's own environment
- Large area
- High resolution
- Bright / High contrast
- (Increased demand on image generation)
  - Lot more pixels
  - Adapt to custom screen geometries

Video of “3D Talking Heads”

- Rapid depth extraction via ‘structured light’
  - reduces problem of finding corresponding points in multiple camera images
- Light patterns made nearly imperceptible by projecting complementary patterns very rapidly

From Images to Geometry: A New Paradigm for 3D Computer Graphics

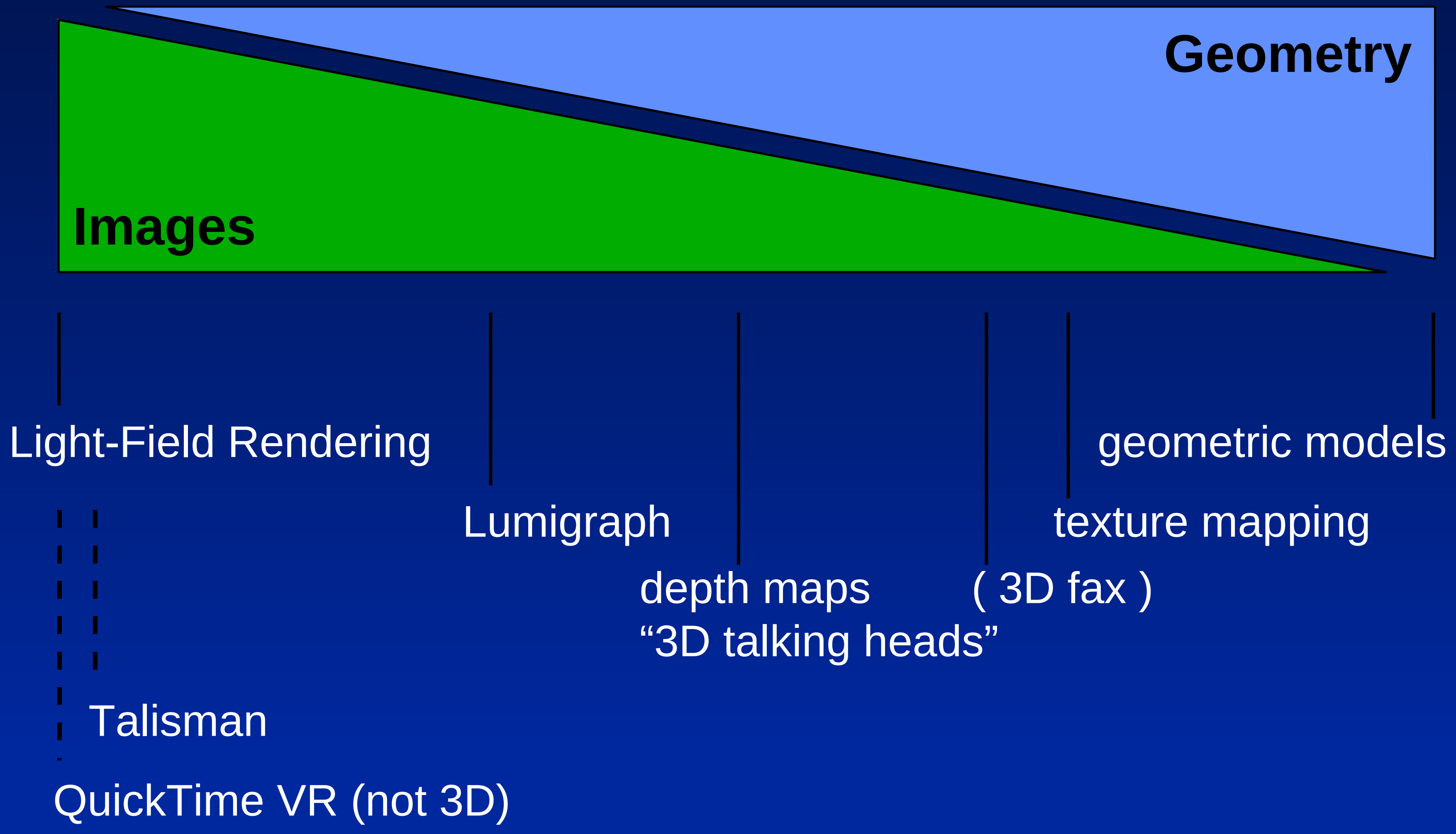
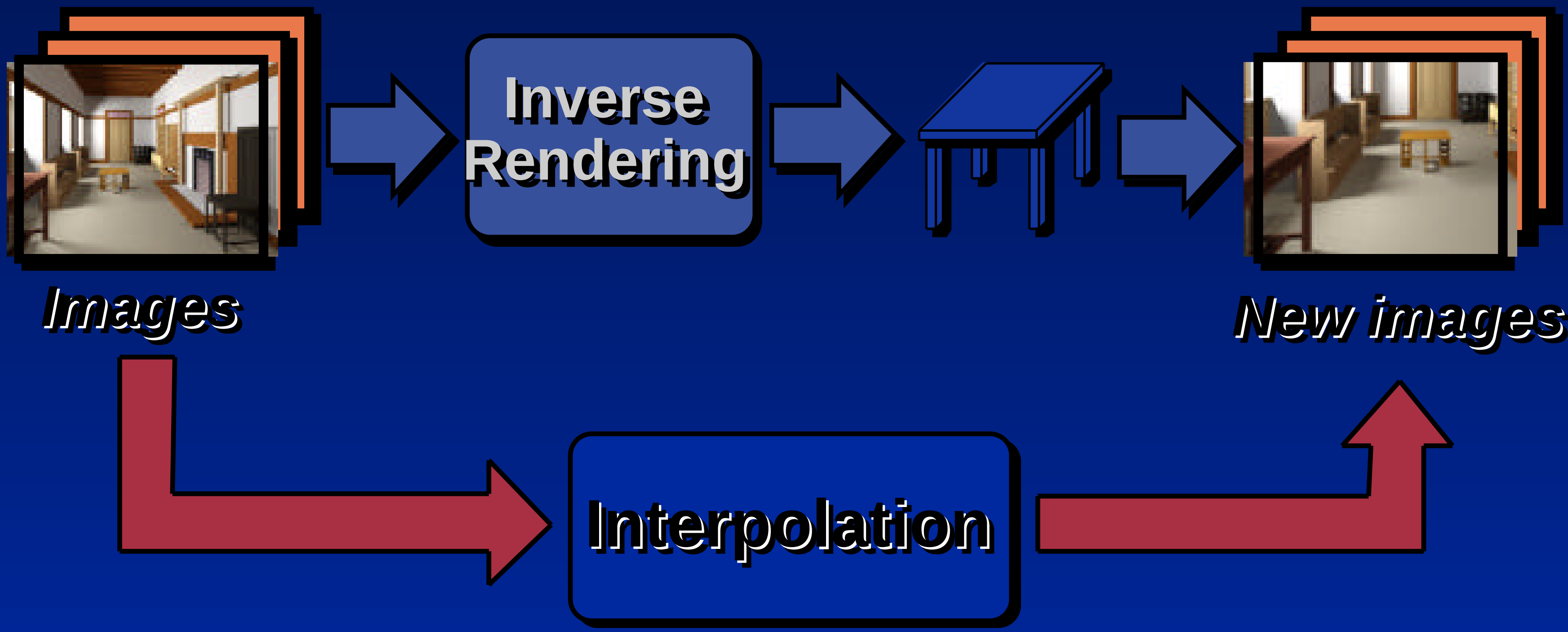
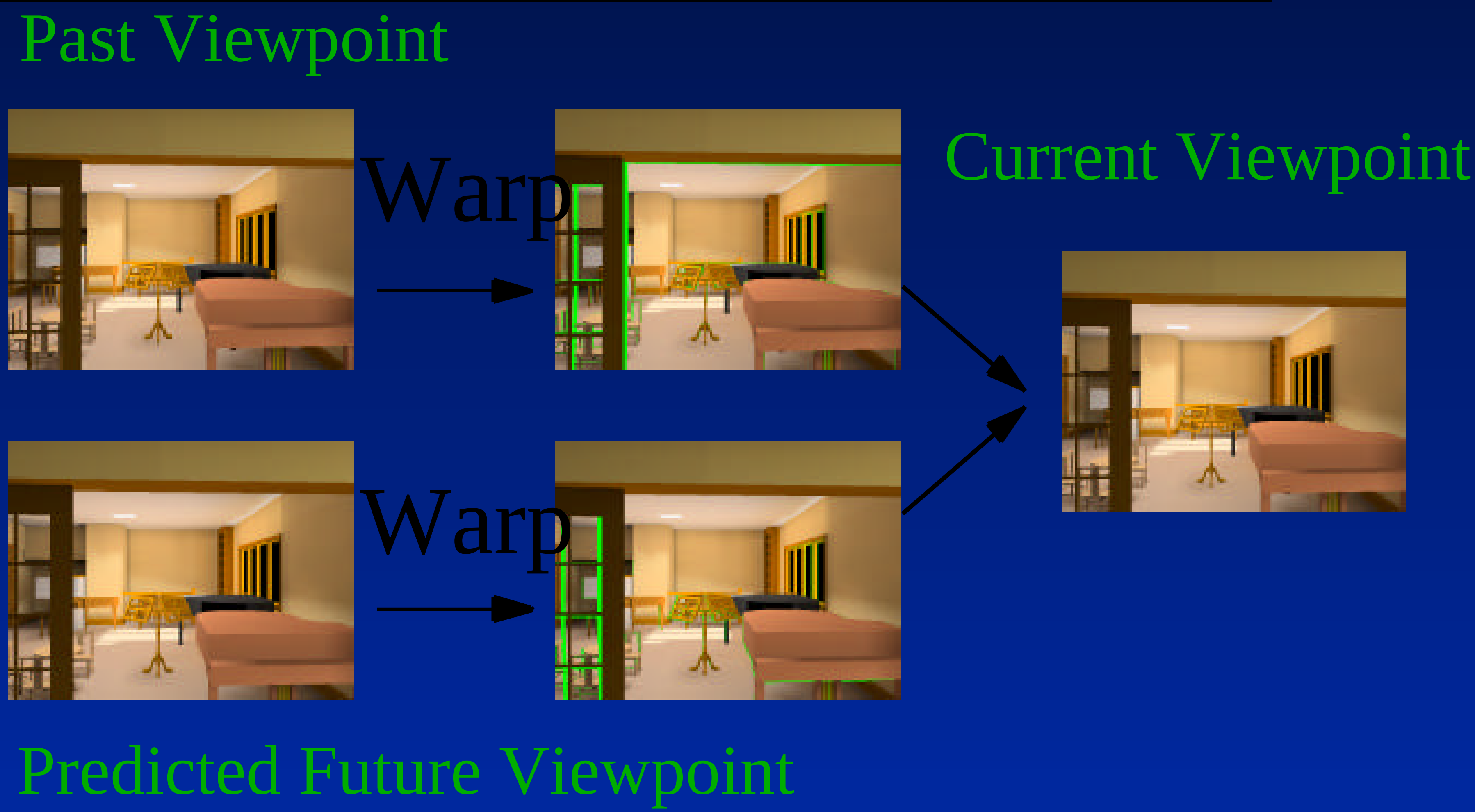


Image-Based Rendering, Inverse Rendering (from J.Arvo)



New images can be obtained by interpolating

Post-Rendering Warp



Video of Post Rendering Warp

•

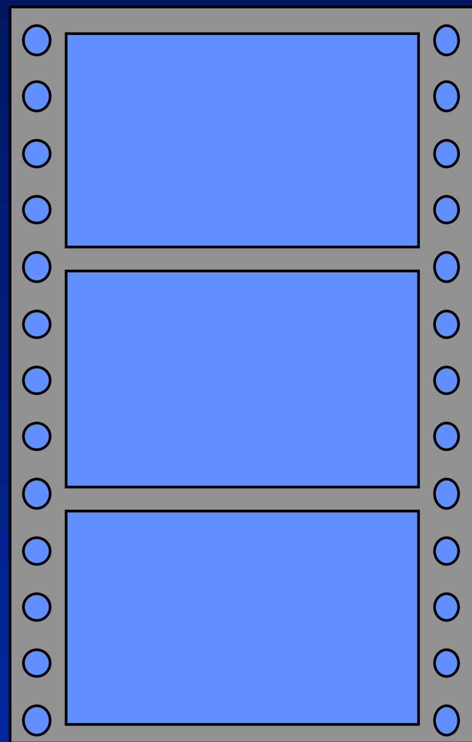
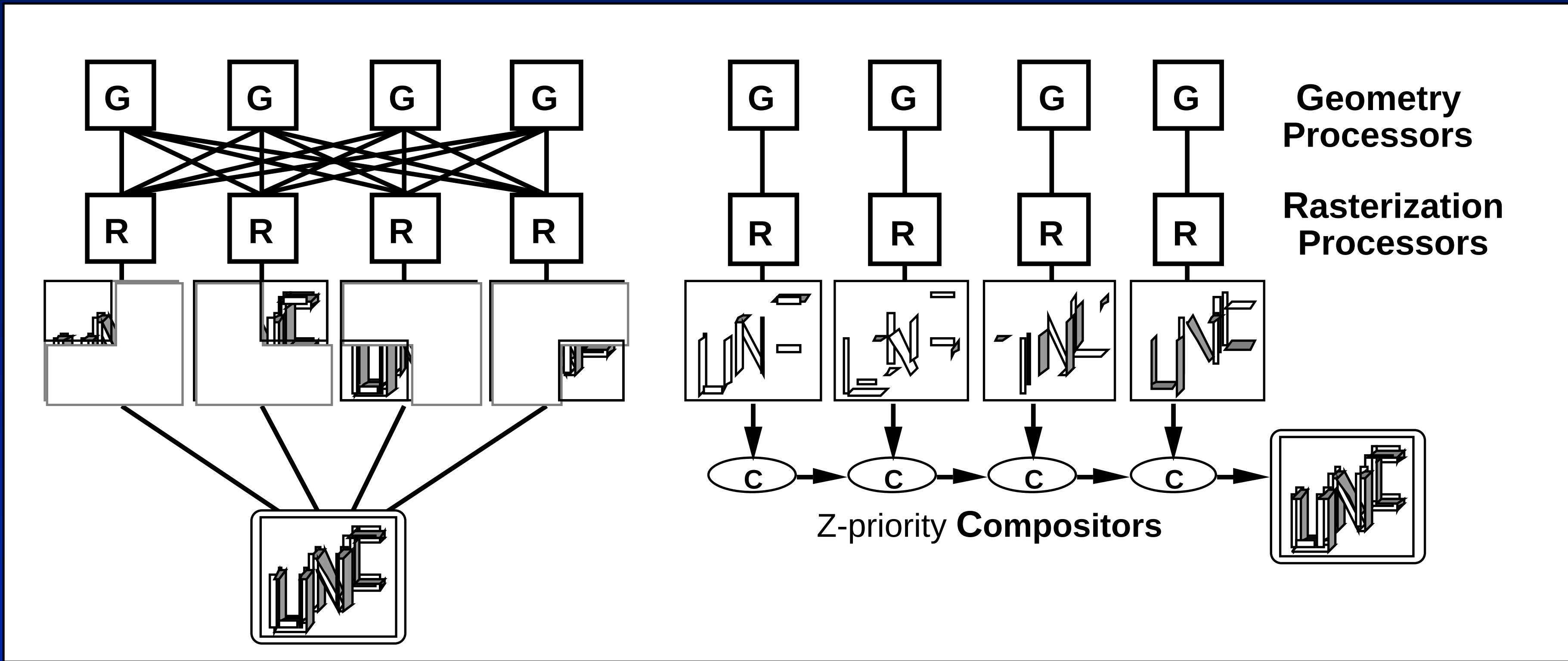


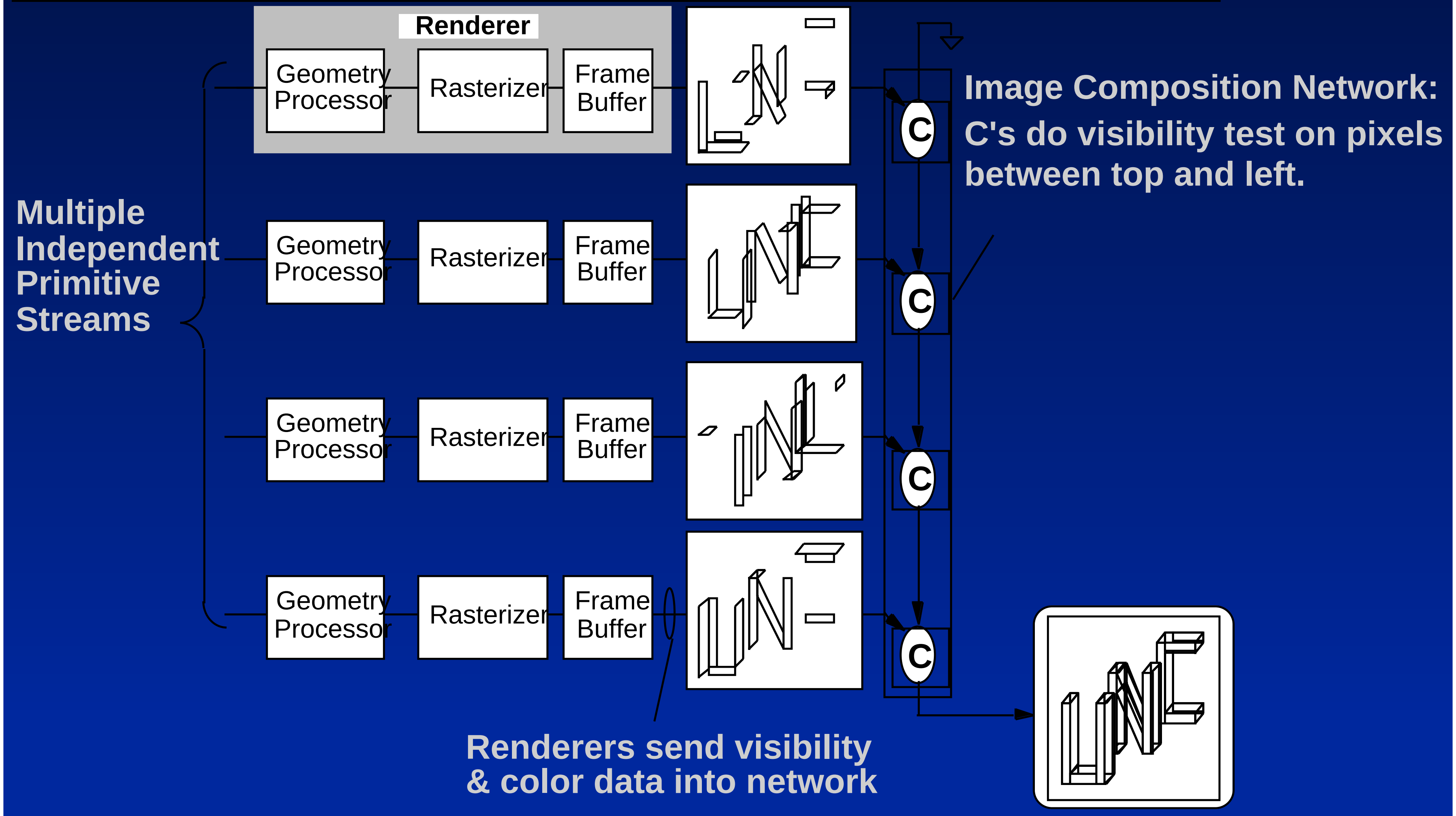
Image Generation: High-performance Graphics Computers

Now: screen subdivision (~SGI RE, UNC Pixel-Planes 5)

Next: z-compositing final images (UNC PixelFlow)



Object Parallel by Image Composition



ImageFlow

- Departure from polygon-based rendering
- History of rendering: lines, polygons, texture, depth (our belief)
- Warp images based on depth value at each image sample
- Input from cameras
- Preliminary design begun

Tracking user's head

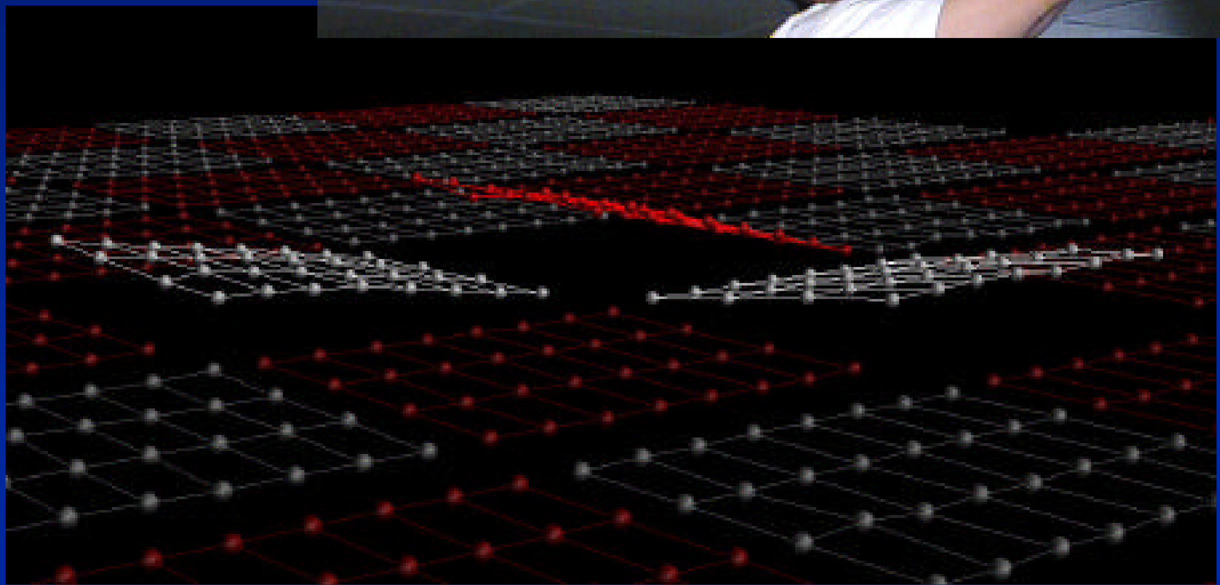
- Difficulty with commercial trackers
- Image-based tracking / hybrid tracking
  - new difficulty: keeping tracking targets in view
- Predictive tracking (Ron Azuma: possible for 50-60ms)

The HiBall Tracker

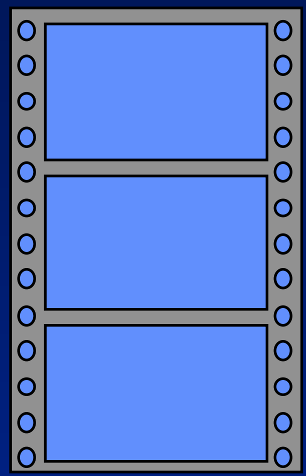
- Second generation ceiling tracker
- Ceiling tiles completed and installed
  - simple drop-in “acoustical” tiles
  - enabled by Brown, Caltech, UNC collaboration
- Design and fabrication at UNC and Utah
- System functioning
  - 2KHz estimates (0.5 ms latency)
  - 0.1 mm RMS position noise
  - 0.02 degree RMS orientation noise

HiBall: photo

SCAAT Autocalibration: ceiling photo & LED calibrations



Video of UNC HiBall Tracker



“Being There”: in 5-10 years

- Key: Acquire and Display **<u>EVERY mm EVERY sec</u>**

(end)

Video of “Walking around Leonard’s Yard”

- Illustrated photographic “feel” of rendering from image input
- For each pixel in 360-degree panorama gets a depth/disparity value

New Custom Head-mounted Display (D Colucci, K Keller, R Fish@U.Utah)

- Video see-through to correctly merge real & synthetic parts of the scene (esp. occlusion)
- Video cameras optically at user’s eye positions
- Unobstructed view except for display
- Flip-up / flip-down

Displays (2 of 2): Head-mounted

- Comfortable
- See-through (usually)
  - Optical: cheap, easy; can’t combine real and virtual
  - Video: bulky, esp. for wide field of view
- Field of view
- Brightness / Resolution

Laparoscopic Visualization

(with Anthony Meyer, MD, PhD)

- Goal: view from surgeon's normal point of view -- as with open surgery
- Key challenge: extract 3D range image from laparoscopic camera
- Initial experiment
  - Pre-experiment: mechanically scanned 3D surface of medical model
  - During experiment: mapped live "laparoscopic" camera video onto the 3D surface

Our Vision of Telecollaboration: "Being There" Together

CAD & Visualization

Interaction

odeling

Rendering

ge Generation

Scene Acquisition

isplays

Networking

Tracking

Needing to Know Depth within Video Images

- Needed to extract 3D remote environment
- Postrendering warp: widely applicable to speed-up image generation frame rate