Carnegie Mellon

Re-live the Movie "Matrix":
From Harry Nyquist to Image-Based Rendering

Tsuhan Chen
tsuhan@cmu.edu
Carnegie Mellon University
Pittsburgh, USA

Advanced Multimedia Processing Lab

Carnegie Mellon

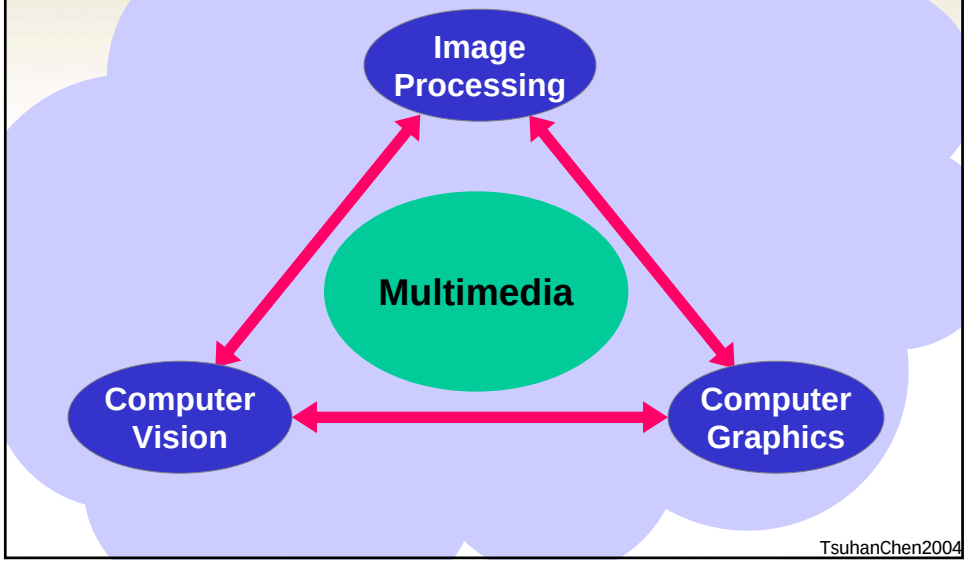
Some History...

- IEEE Multimedia Signal Processing (MMSP) Technical Committee, 1996~
- IEEE MMSP Workshops
  - Princeton 1997, Los Angeles 1998, Copenhagen 1999, Cannes 2001, St. Thomas 2002, Siena 2004,...
- IEEE International Conf. on Multimedia and Expo. (ICME)
  - New York 2000, Tokyo 2001, Lausanne 2002, Baltimore 2003, Taipei 2004,...
- *Proceedings of IEEE*, Special Issue on MMSP, 1998
- *IEEE Transactions on Multimedia*, March 1999~
  - Special issues: networked multimedia 2001, multimedia database 2002, multimodal interface 2003, streaming media 2004,...

Advanced Multimedia Processing Lab

Carnegie Mellon

Convergence...



TsuhanChen2004

Carnegie Mellon

Convergence of Image, Vision, and Graphics

- Yah, right... The truth is...
- Between vision and graphics communities
  - "Vision is more solid research than graphics"
  - "But graphics people make more money!"
- Graphics and vision communities say:
  - "Image processing is low-level processing"
- Image processing community says:
  - "Graphics is only some fancy toys..."
  - "Vision is things that don't work!"

AMP Advanced Multimedia Processing Lab

TsuhanChen2004

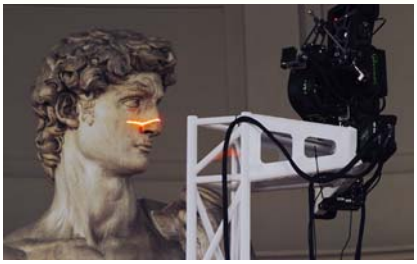
Some Attempts

- MPEG-4 promised the convergence
  - Started out as model-based coding
    - Encoding: analysis using vision
    - Decoding: synthesis using graphics
  - Well, almost...
    - Settled with 2D shape-based coding
    - Model-based for limited content, e.g., faces
    - AFX (Animated Framework Extension)
- Next wave is image-based rendering...

Fundamentals of Image-Based Rendering

Carnegie Mellon

Model-Based Rendering vs. Image-Based Rendering

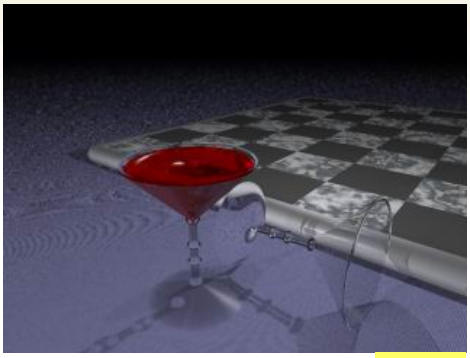


[Digital Michelangelo Project, Stanford]

Light field of Michelangelo's statue of Night

Carnegie Mellon

Example



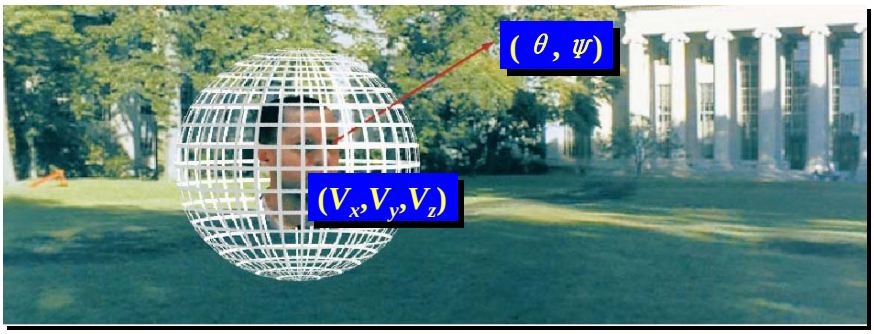
Demo

Carnegie Mellon

7D Plenoptic Function

$$f(V_x, V_y, V_z, \theta, \psi, \lambda, t)$$

[Adelson'91]



Carnegie Mellon

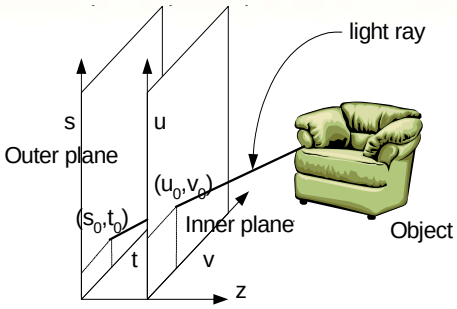
Representations

- 7D: $f(V_x, V_y, V_z, \theta, \psi, \lambda, t)$ [Adelson'91]
- 5D: Stationary and monochrome [McMillan'95]
- 4D: Scene inside a bounded region
  - Lumigraph [Gortler'96]
  - Lightfield [Levoy'96]
  - EyeVision [Kanade'01]
- 3D: Viewpoint on a plane
  - Concentric Mosaics [Shum'99]
  - BulletTime ["Matrix"]
- 2D: Viewpoint at a single point
  - Panorama [Chen'95], QuickTime VR

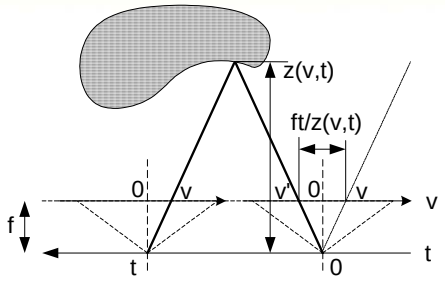
Carnegie Mellon

Lumigraph/Lightfield

[Gortler'96] [Levoy'96]



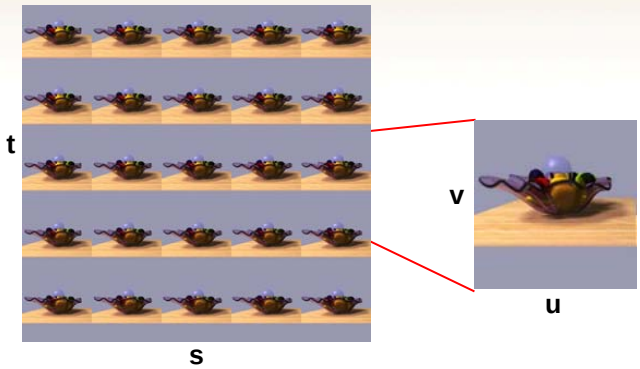
→ 4D



2D Simplification

Carnegie Mellon

Captured Images



4D Image Array

Carnegie Mellon

Example: EyeVision

→4D (incl. time)

[Kanade'01]

Before Correction

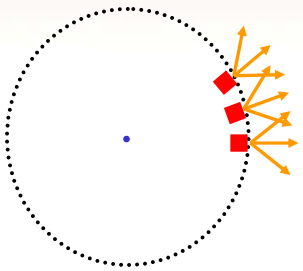
After Correction

SuperBowl

Carnegie Mellon

Concentric Mosaics

[Shum'99]



→ 3D

Carnegie Mellon

3D Examples

Lobby (1350x320x240)

Kids (1462x352x288)

Demo

Carnegie Mellon

IBR: A Sampling Problem

- "Given a set of discrete samples (complete or incomplete) from the plenoptic function, the goal of image based rendering is to generate a continuous representation of that function" [McMillan'95]
- Q: This is a sampling problems
  - How many samples and where?
- A: Need "Nyquist Sampling Theorem" for IBR

Sampling Theorem

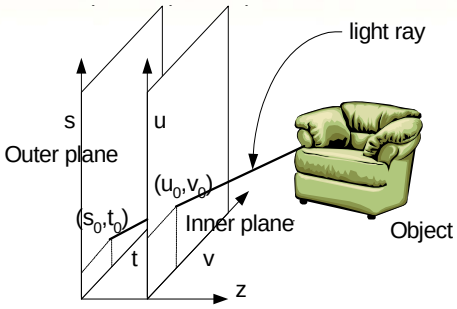
$$\omega_S > 2\omega_M$$ ← Nyquist Rate

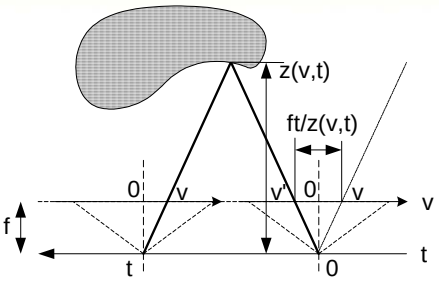
Shannon, 1949 (in communication theory)
Whittaker, 1964 (in math)

2WT numbers (in Fourier series) to represent a function of duration *T* and highest frequency *W*

Nyquist, 1928 ←
Gabor, 1946

Recall u,v-s,t Parameterization...

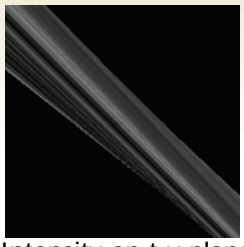




→ Need multidimensional spectral analysis

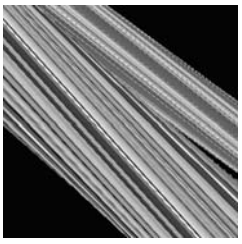
Carnegie Mellon

Spectral Analysis on t-v Plane



Intensity on t-v plane

Spectrum

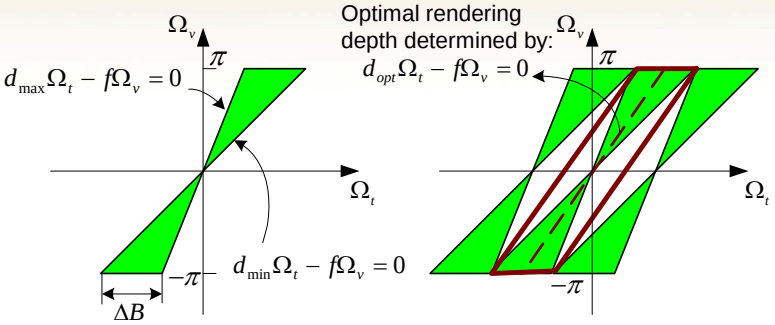


Intensity on t-v plane

Spectrum

Carnegie Mellon

Sampling for IBR



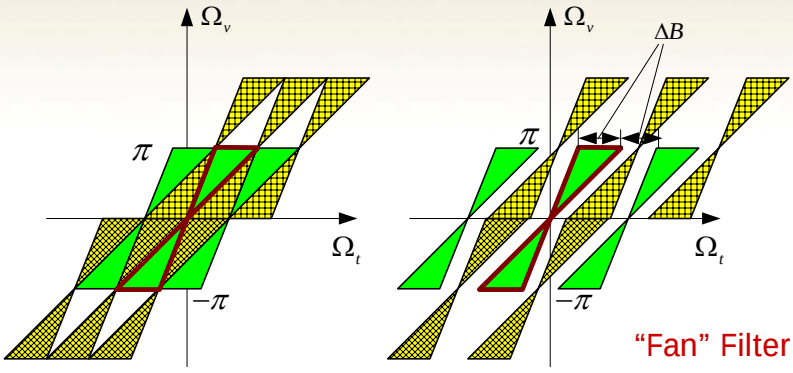
Optimal rendering depth determined by: $d_{opt}\Omega_t - f\Omega_v = 0$

- Lambertian surface
- No occlusion
- "Truncating window" analysis [Chai et. al, 2000]

Lowpass Filter

Carnegie Mellon

Optimal Sampling [Zhang and Chen, 2003]



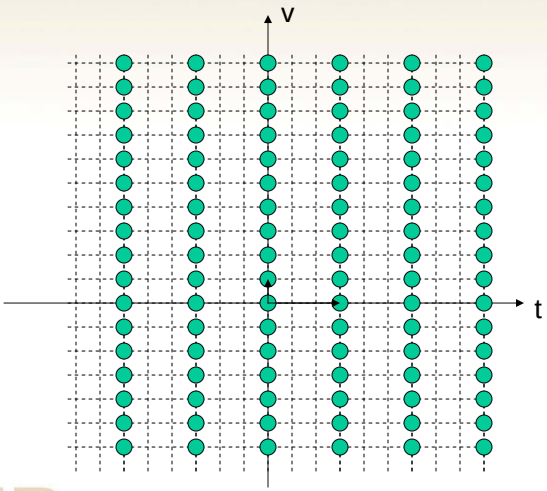
"Fan" Filter

2x more compact
50% fewer samples

Same rate as rectangular sampling
Easier to design the filter

Carnegie Mellon

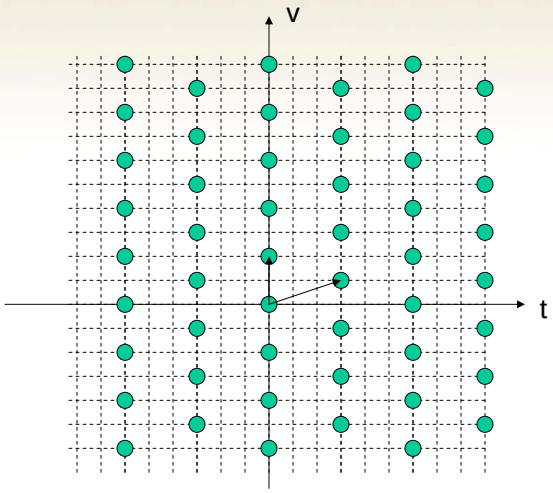
Rectangular Sampling



$$V = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

Carnegie Mellon

Optimal Sampling



$$V = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}$$

Hexagonal/Quincunx Sampling!!!

Carnegie Mellon

Beating Nyquist...

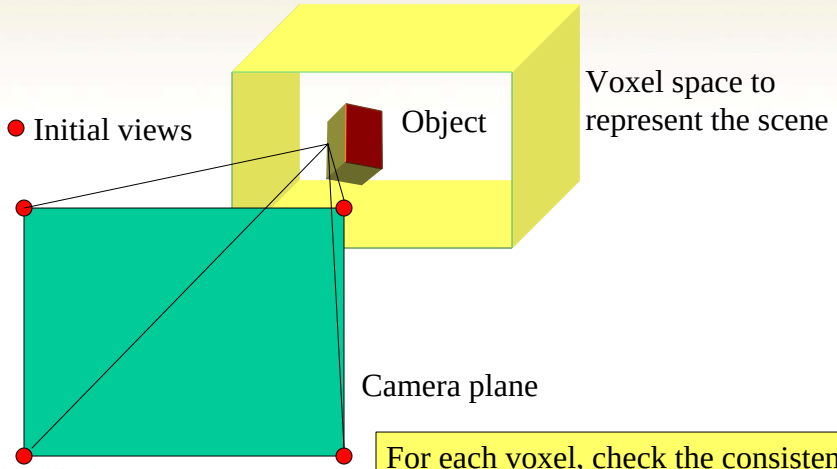
Carnegie Mellon

Beating Nyquist...

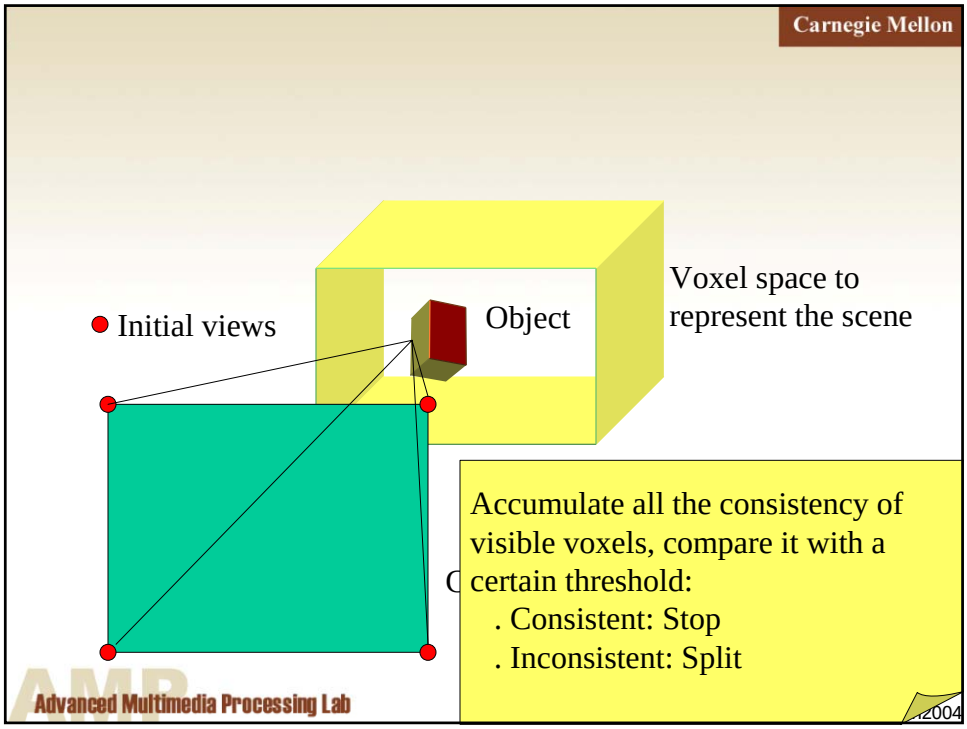
- Plenoptic functions are non-stationary
  - Non-Lambertian surfaces
  - Occlusion
- Non-uniform sampling is preferred
- Active IBR
  - Determine where to capture the images
  - Resulting a non-uniform sampling scheme

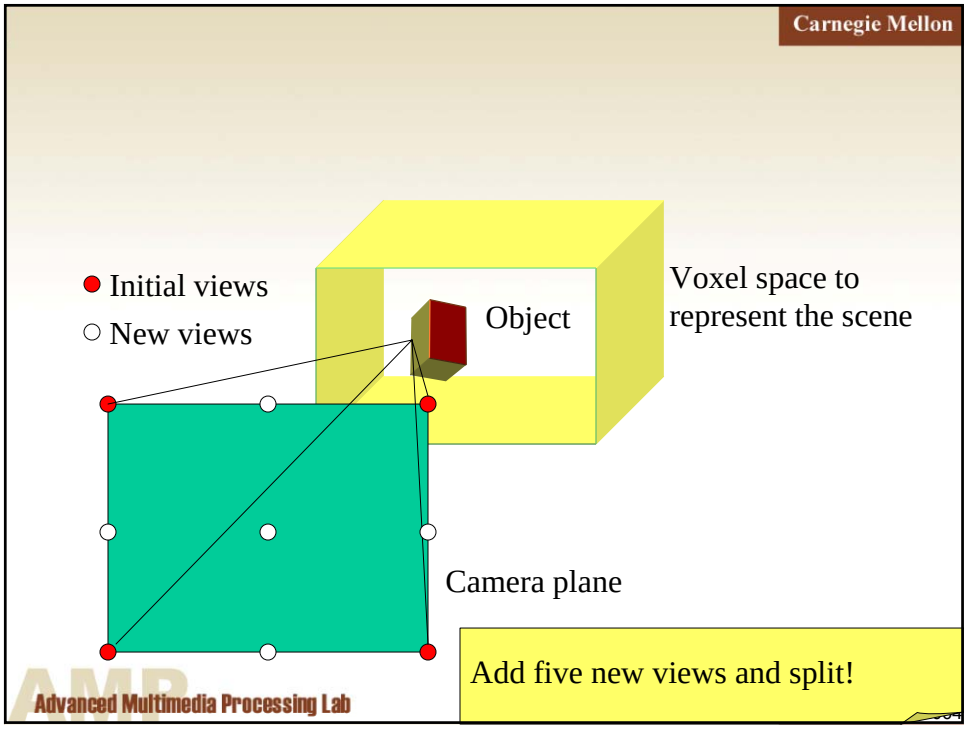
Carnegie Mellon

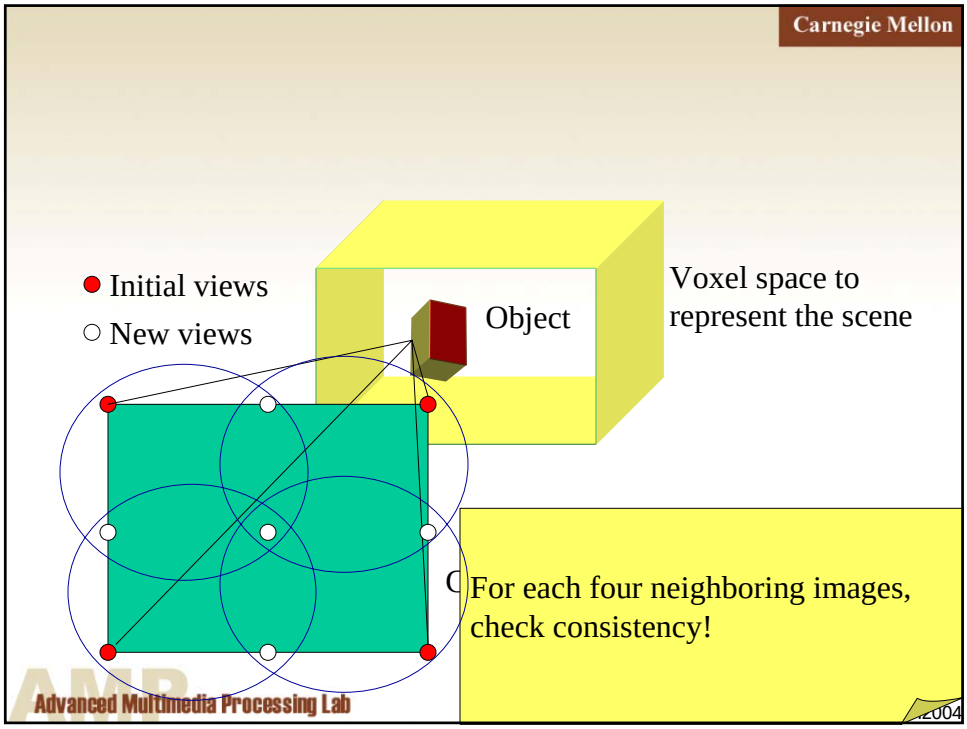
A Naive Approach

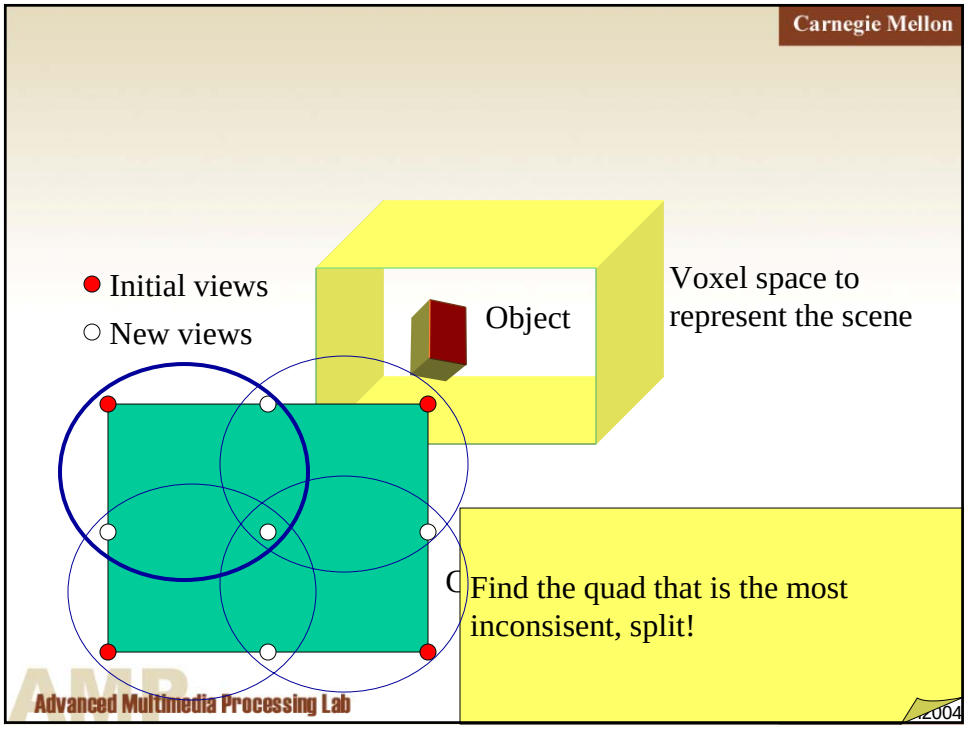


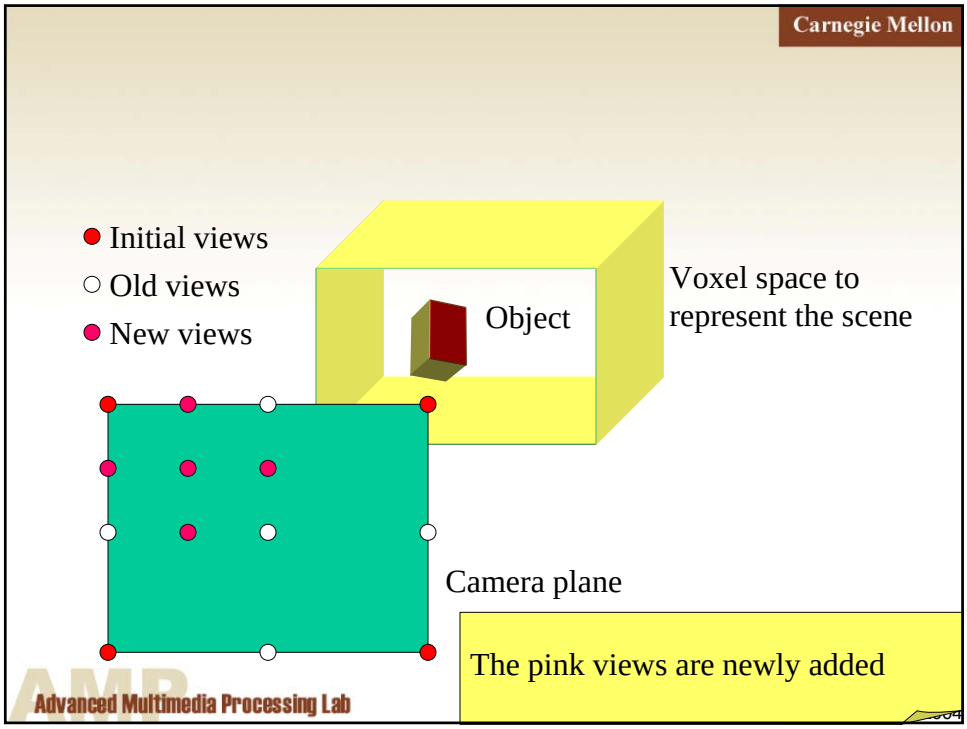
For each voxel, check the consistency to the initial views

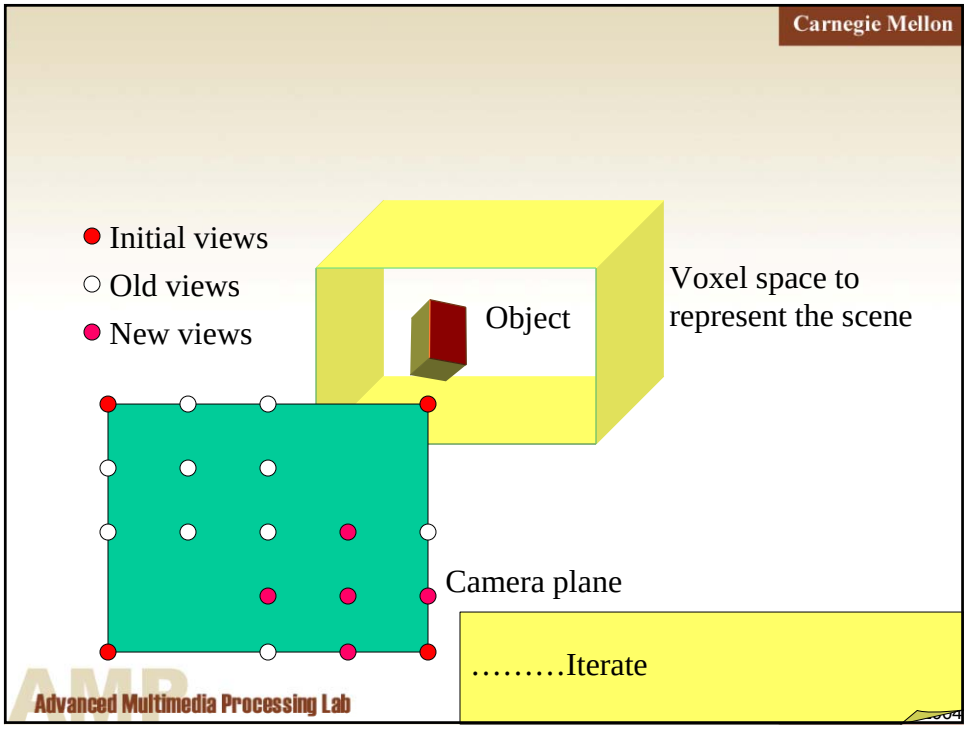

Carnegie Mellon
Initial views
Object
Voxel space to represent the scene
Accumulate all the consistency of visible voxels, compare it with a certain threshold:
. Consistent: Stop
. Inconsistent: Split
Advanced Multimedia Processing Lab

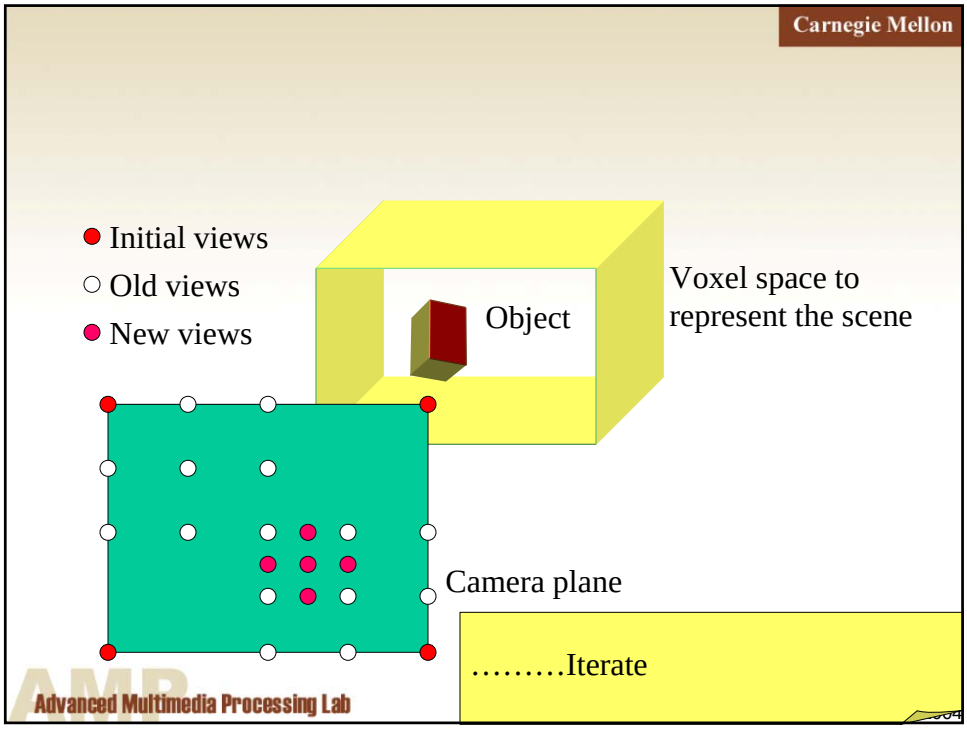

Carnegie Mellon
Initial views
New views
Object
Voxel space to represent the scene
Camera plane
Add five new views and split!
Advanced Multimedia Processing Lab

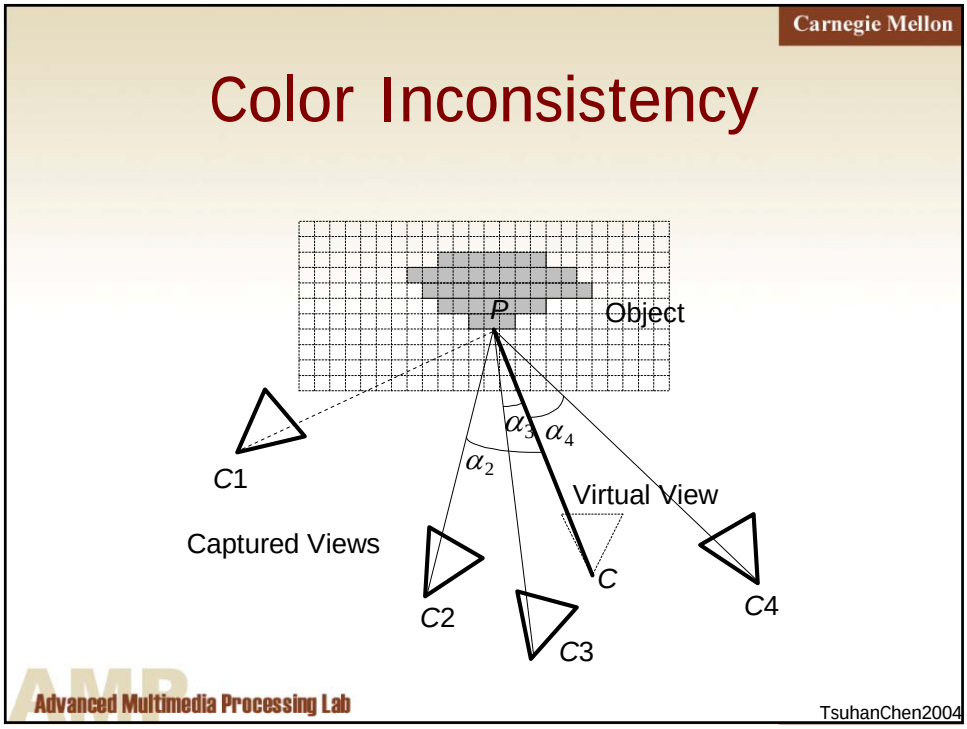

Carnegie Mellon
Initial views
New views
Object
Voxel space to represent the scene
For each four neighboring images, check consistency!
Advanced Multimedia Processing Lab


Carnegie Mellon
Initial views
New views
Object
Voxel space to represent the scene
Find the quad that is the most inconsisent, split!
Advanced Multimedia Processing Lab


Carnegie Mellon
Initial views
Old views
New views
Object
Voxel space to represent the scene
Camera plane
The pink views are newly added
AMP
Advanced Multimedia Processing Lab

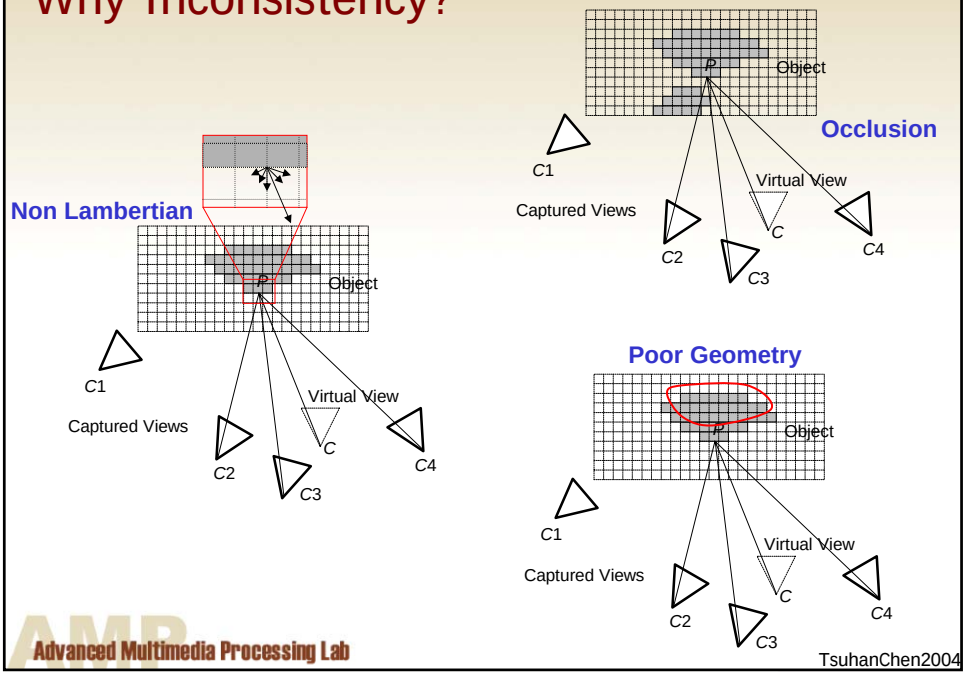

Carnegie Mellon
Initial views
Old views
New views
Object
Voxel space to represent the scene
Camera plane
……Iterate
AMP
Advanced Multimedia Processing Lab


Carnegie Mellon
Initial views
Old views
New views
Object
Voxel space to represent the scene
Camera plane
……Iterate
AMP
Advanced Multimedia Processing Lab


Carnegie Mellon
Color Inconsistency
P
Object
C1
Captured Views
C2
C3
C4
C
Virtual View
AMP
Advanced Multimedia Processing Lab
TsuhanChen2004

Why Inconsistency?

Carnegie Mellon

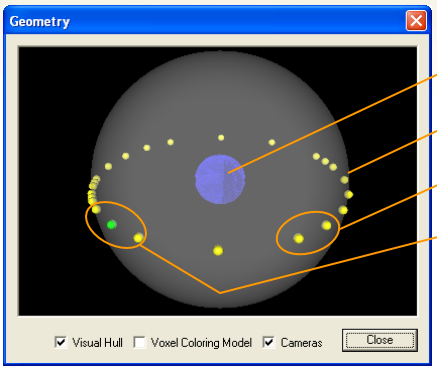


Advanced Multimedia Processing Lab

TsuhanChen2004

Progressive Capturing (PCAP)

Carnegie Mellon



Object

Camera positions

Image pair

A split of the image pair

Q: Where to split, i.e., add one more image?

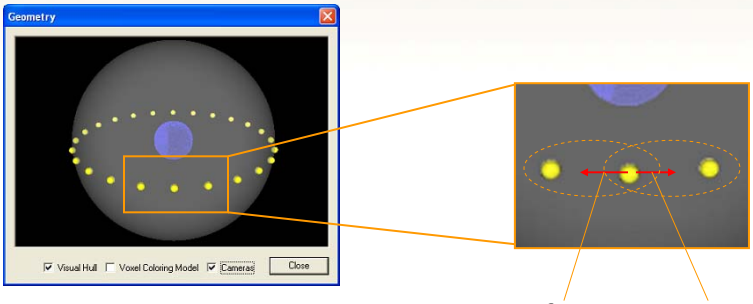
A: The image pair with highest inconsistency

Advanced Multimedia Processing Lab

TsuhanChen2004

Carnegie Mellon

Rearranged Capturing (RCAP)



Force from the left

Force from the right

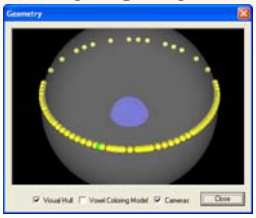
Force proportional to inconsistency

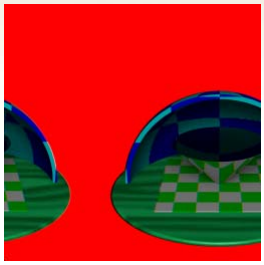
TsuhanChen2004

Carnegie Mellon

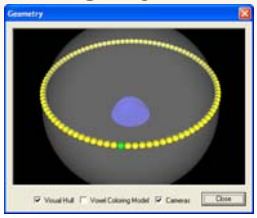
Test Scene I: Capturing

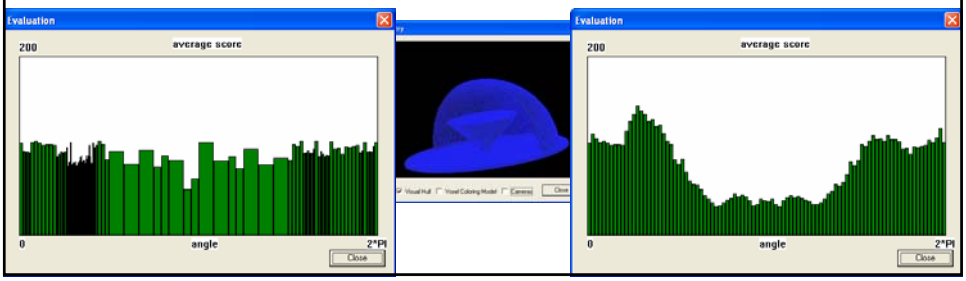
Non-Uniform





Uniform

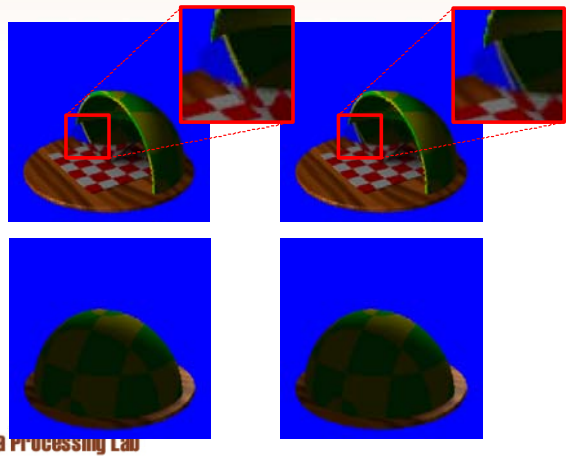




Carnegie Mellon

Test Scene I: Rendering Results (I)

Non-Uniform Uniform



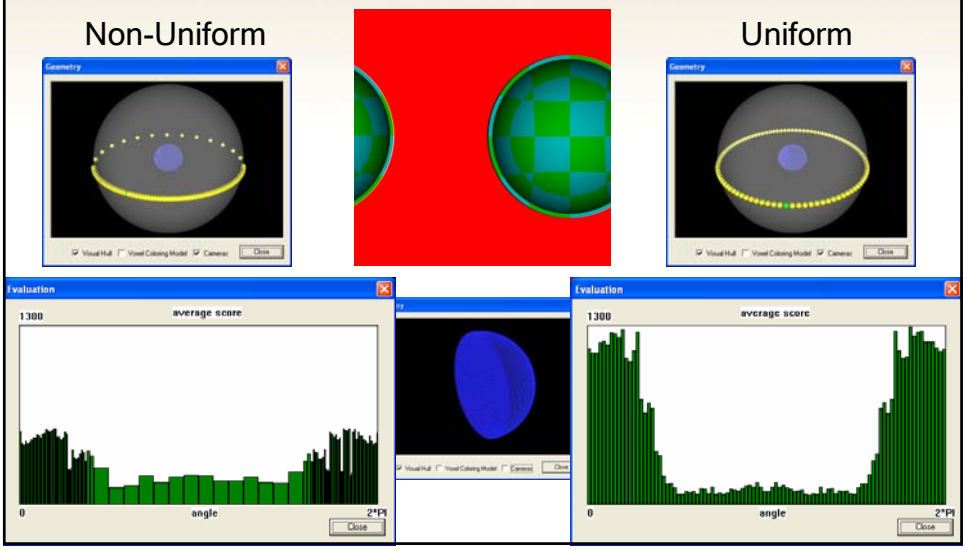
Advanced Multimedia Processing Lab

TsuhanChen2004

Carnegie Mellon

Test Scene II: Capturing

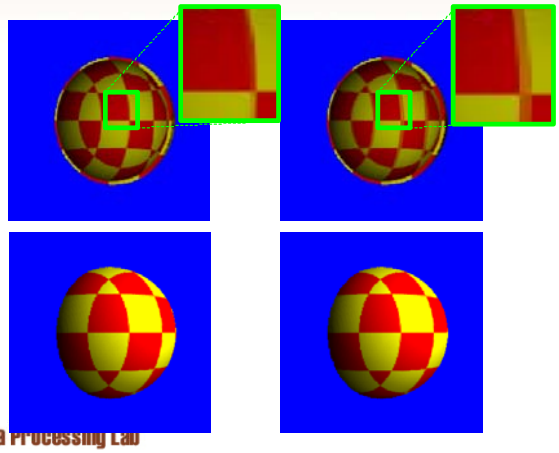
Non-Uniform Uniform



Carnegie Mellon

Test Scene II: Rendering Results (I)

Non-Uniform Uniform



Advanced Multimedia Processing Lab

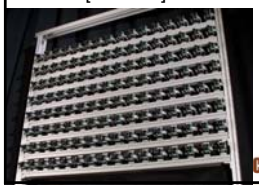
TsuhanChen2004

Carnegie Mellon

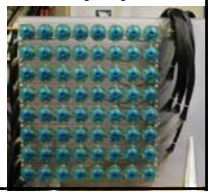
Self-Reconfigurable Camera Array

[Zhang and Chen, CMU]

[Stanford]



[MIT]



cessing Lab

Carnegie Mellon

Details

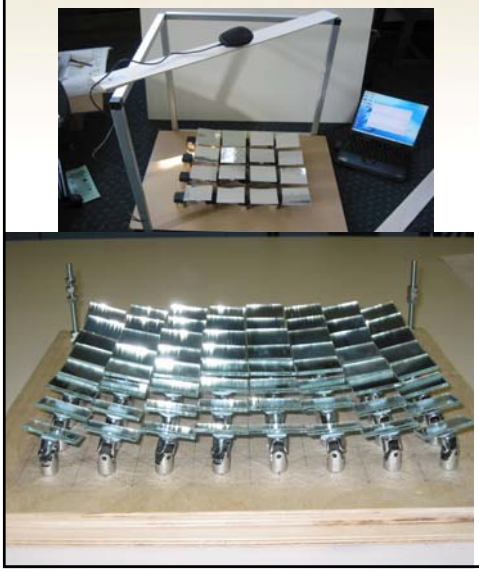
- 48 webcams (embedded processors)
- 2 step-motors each for translation and pan
- **Real-time** capturing/calibration/rendering
  - New architecture may be needed

AMP Advanced Multimedia Processing Lab

TsuhanChen2004

Carnegie Mellon

Future: Mirror/Lens Array

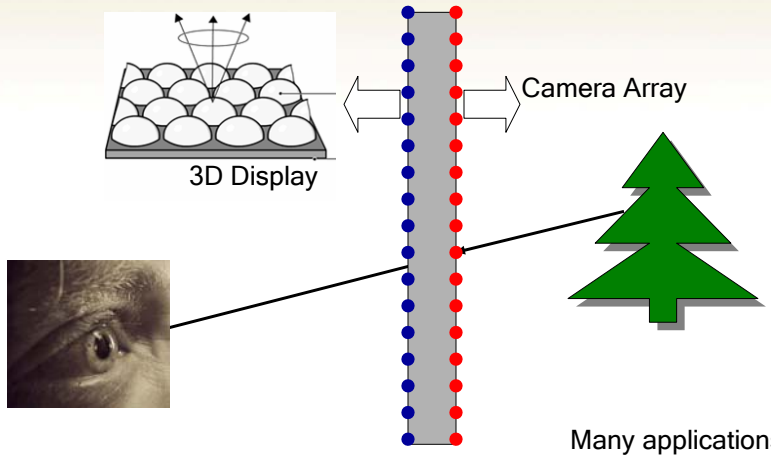


Many applications…

TsuhanChen2004

Carnegie Mellon

Future: “Transparent Material”



Many applications…

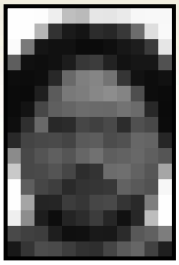
Carnegie Mellon

Future Research Directions

- IBR composition
  - Merging objects/IBR, alignment, relighting
- IBR compression
  - Sampling/compressing IBR
- IBR communication
  - Streaming IBR, error resilience, etc.

Carnegie Mellon

We can beat Nyquist if we can...



Reconstruct

One Single Image

[Baker and Kanade, "Hallucinating Faces"]

Number of all possible 16×12 images $= 2^{16 \times 12 \times 8}$

$>> 30 \times 60 \times 60 \times 24 \times 365 \times$ human history $\times$ world population

>> number of all possible face images

→ We can beat Nyquist with **prior**

Carnegie Mellon

What does this say?

[http://www.palmyra.demon.co.uk]

Human is the best sampler!!!

Carnegie Mellon

Advanced Multimedia Processing Lab

Please visit us at:

http://amp.ece.cmu.edu

AMP Advanced Multimedia Processing Lab

TsuhanChen2004