Virtual Viewpoint Reality
NTT: Visit
1/7/99

Overview of VVR Meeting

- Motivation from MIT …
- Discuss current and related work
  - Video Activity Monitoring and Recognition
  - 3D Modeling
  - Demonstrations
- Related NTT Efforts
- Discussion of collaboration
- Future work
- Lunch

Motivating Scenario

- Construct a system that will allow a user to observe any viewpoint of a sporting event.
  - From behind the goal
  - Along the path of the ball
  - As a participating player
- Provide high level commentary/statistics
  - Analyze plays
  - Flag goals/fouls/offsides/strikes

Given a number of fixed cameras... Can we simulate any other?

A Virtual Reality Spectator Environment

- Build an exciting, fun, high-profile system
  - Sports: Soccer, Hockey, Tennis, Basketball
  - Drama, Dance, Ballet
- Leverage MIT technology in:
  - Vision/Video Analysis
    - Tracking, Calibration, Action Recognition
    - Image/Video Databases
  - Graphics
- Build a system that provides data available nowhere else...
  - Record/Study Human movements and actions
  - Motion Capture / Motion Generation

Factor 1: Window of Opportunity

- 20-50 cameras in a stadium
  - Soon there will be many more
- HDTV is digital
  - Flexible, very high bandwidth transmissions
- Future Televisions will be Computers
  - Plenty of extra computation available
  - 3D Graphics hardware will be integrated
- Economics of sports
  - Dollar investments by broadcasters is huge (Billions)
- Computation is getting cheaper

Factor 2: Research

- Calibration
  - How to automatically calibrate 100 moving cameras?
- Tracking
  - How to detect and represent 30 moving entities?
- Resolution
  - Assuming moveable/zoomable cameras: How to direct cameras towards the important events?
- Action Understanding
  - Can we automatically detect significant events - fouls, goals, defensive/offensive plays?
  - Can we direct the user towards points of interest?
  - Can we learn from user feedback?

Factor 3: Research

- Learning / Statistics
  - Estimating the shape of complex objects like human beings is hard. How can we effectively use prior models?
  - Can we develop statistical models for human motions?
    - For the actions of an entire team?
- Graphics
  - What are the most efficient/effective representations for the immersive video stream?
  - What is the best scheme for rendering it?
  - How to combine conflicting information into a single graphical image?

Factor 4: Enabling Other Applications

- *Cyberware Room*
  - A room that records the shape of everything in it.
  - Every action and motion.
- Provide Unprecedented Information
  - Study human motion
    - Build a model to synthesize motions (Movies)
  - Study sports activities
    - Provide constructive feedback
  - Study ballet and dance
    - Critique?
  - Study drama and acting

Factor 5: NTT Interest and Involvement

- NTT has expertise:
  - Networking and information transmission
  - Computer Vision
  - Human Interfaces
- We would like your feedback here!

Overview of VVR Meeting

- Motivation from MIT …
- Discuss current and related work (MIT)
  - Video Activity Monitoring and Recognition
  - 3D Modeling
  - Demonstrations
- Related NTT Efforts
- Discussion of collaboration
- Future work
- Lunch

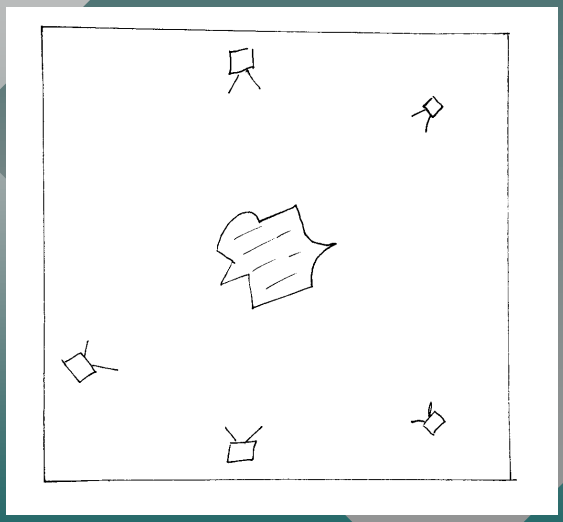
Progress on 3D Reconstruction

- Simple intersection of silhouettes
  - Efficient but limited.
- Tomographic reconstruction
  - Based on medical reconstructions.
- Probabilistic Voxel Analysis (Poxels)
  - Handles transparency.

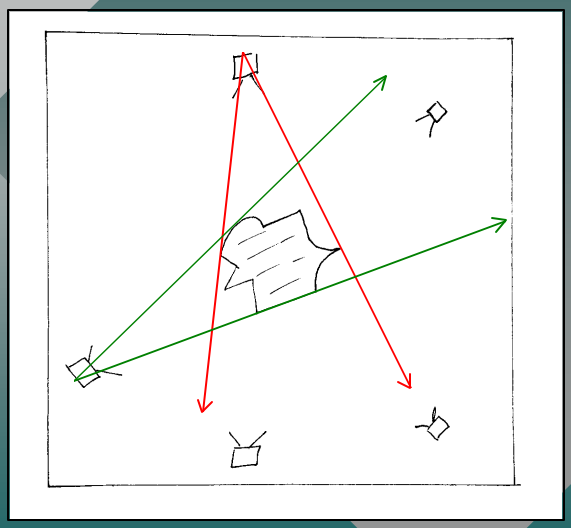
Simple Technical Approach

- 1 Integration/Calibration of Multiple Cameras
- 2: Segmentation of Actors from Field
  - Yields silhouettes -> FRUSTA
- 3: Build Coarse 3D Models
  - Intersection of FRUSTA
- 4: Refine Coarse 3D Models
  - Wide baseline stereo

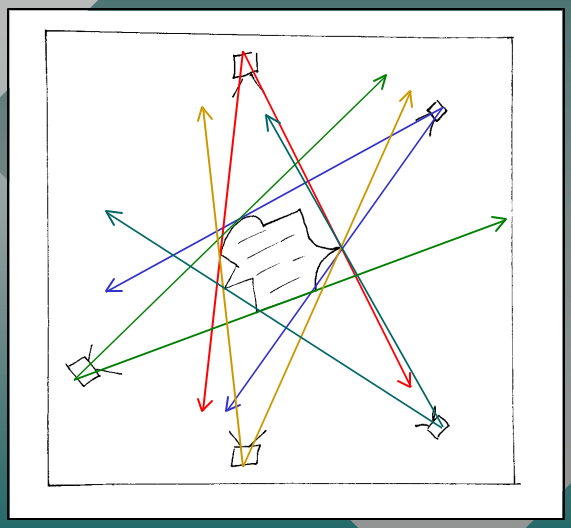
Idea in 2D



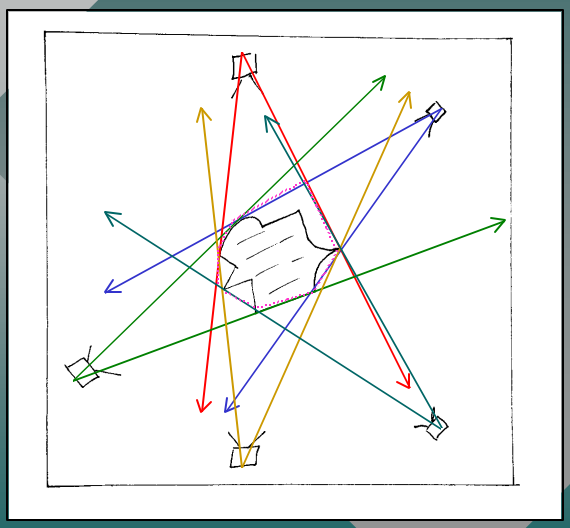
Idea in 2D: Segment



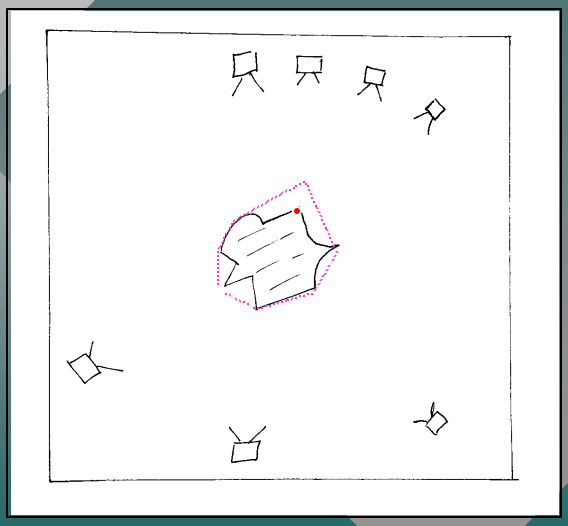
Idea in 2D: Segment



Idea in 2D: Intersection



Coarse Shape

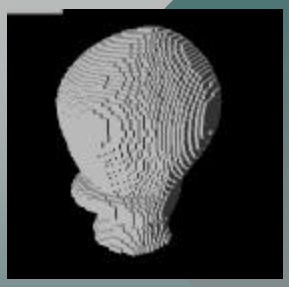


Real Data: Tweety



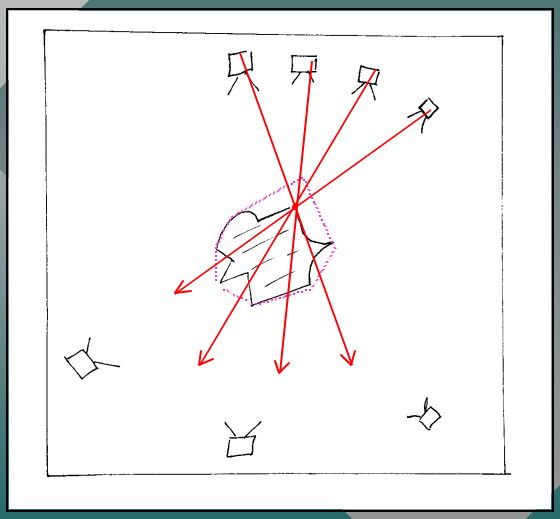
- Data acquired on a turntable
  - 180 views are available... not all are used.

Intersection of Frusta



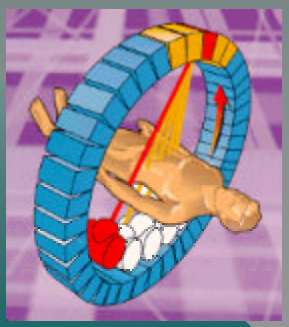

- Intersection of 18 frusta
  - Computations are very fast
    - perhaps real-time

Agreement provides additional information

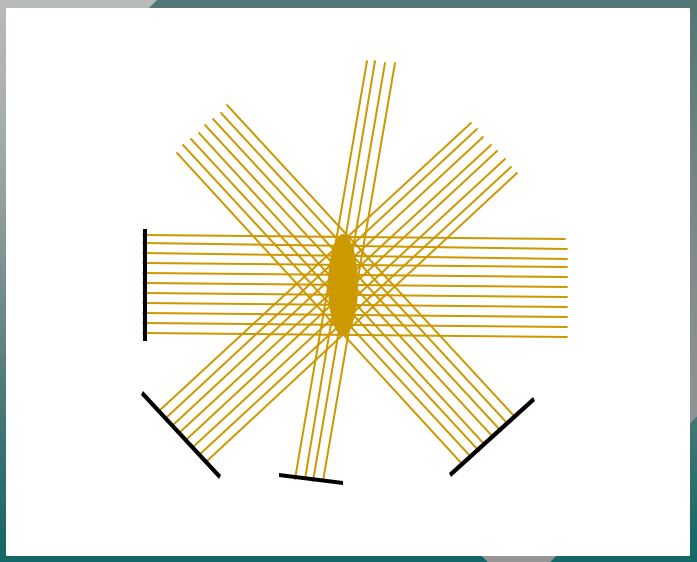


Tomographic Reconstruction

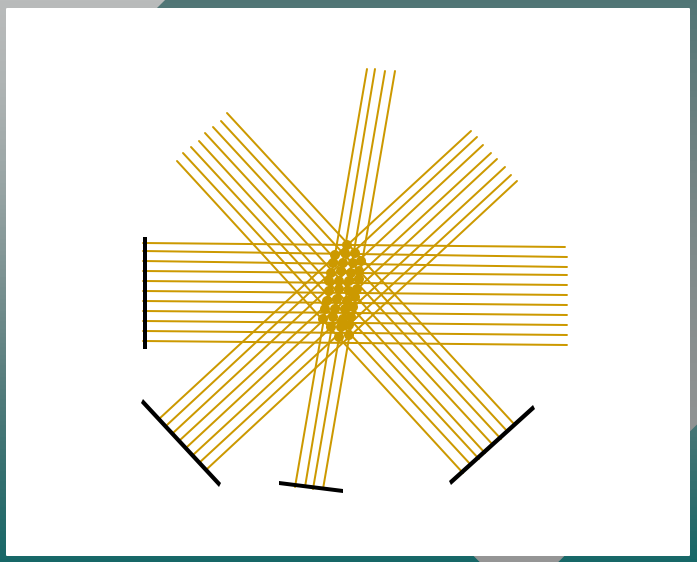
- Motivated by medical imaging
  - CT - Computed Tomography
  - Measurements are line integrals in a volume
  - Reconstruction is by back-projection & deconvolution



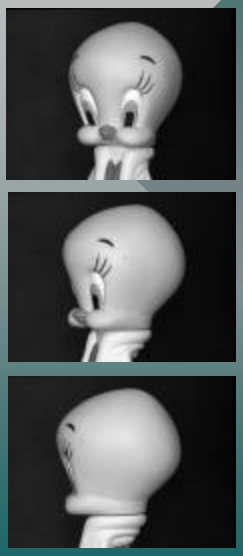
Acquiring Multiple Images (2D)

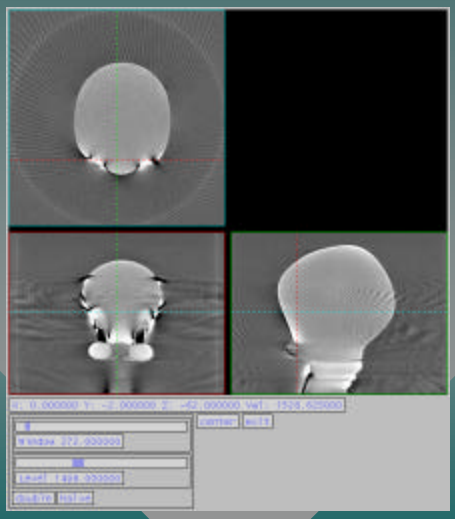


Backprojecting Rays

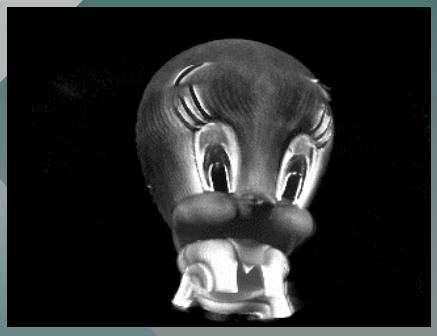


Back-projection of image intensities

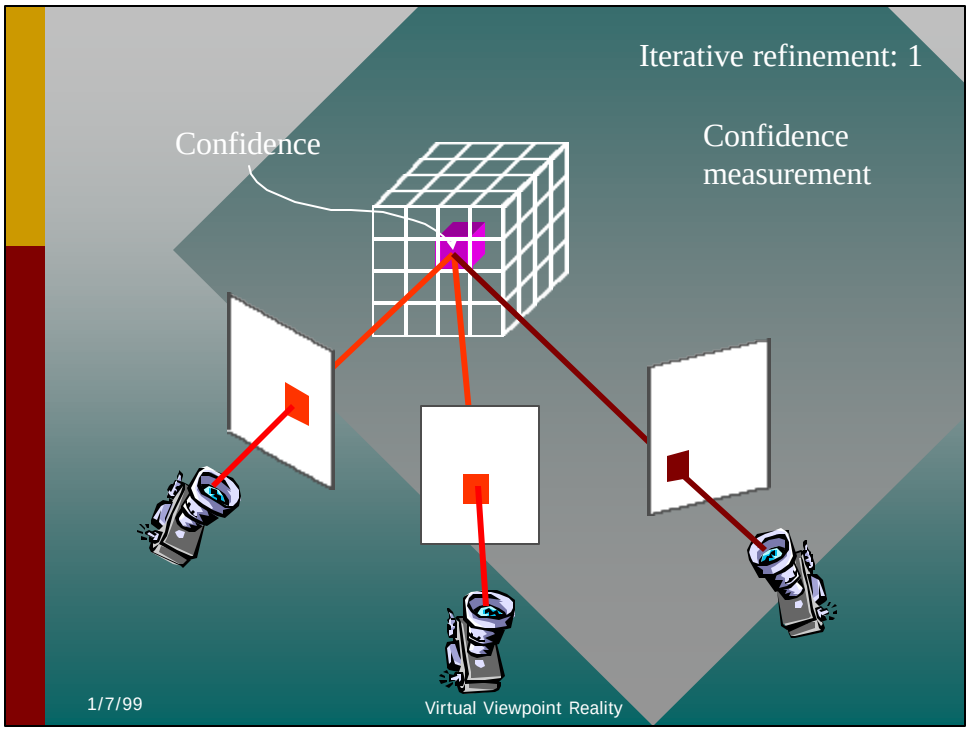


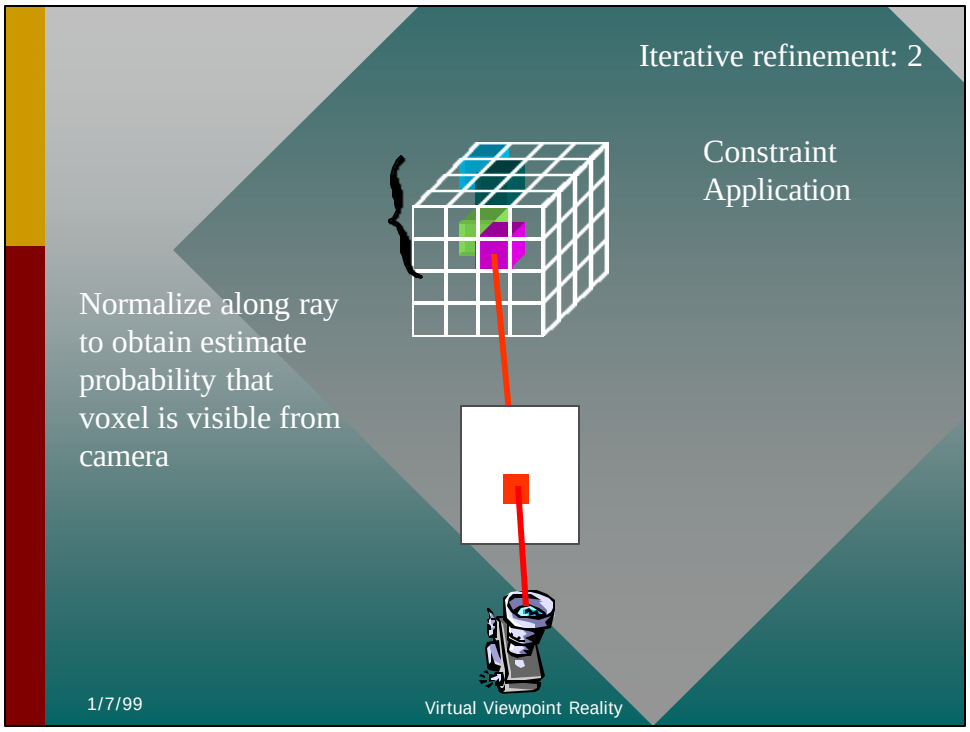


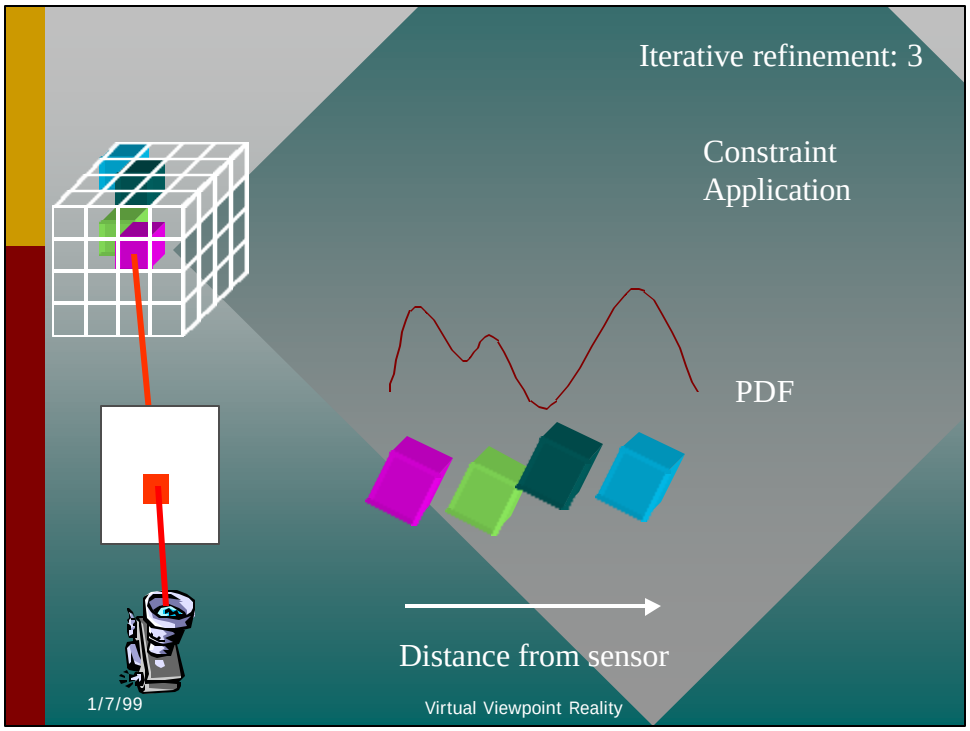
Volume Render…

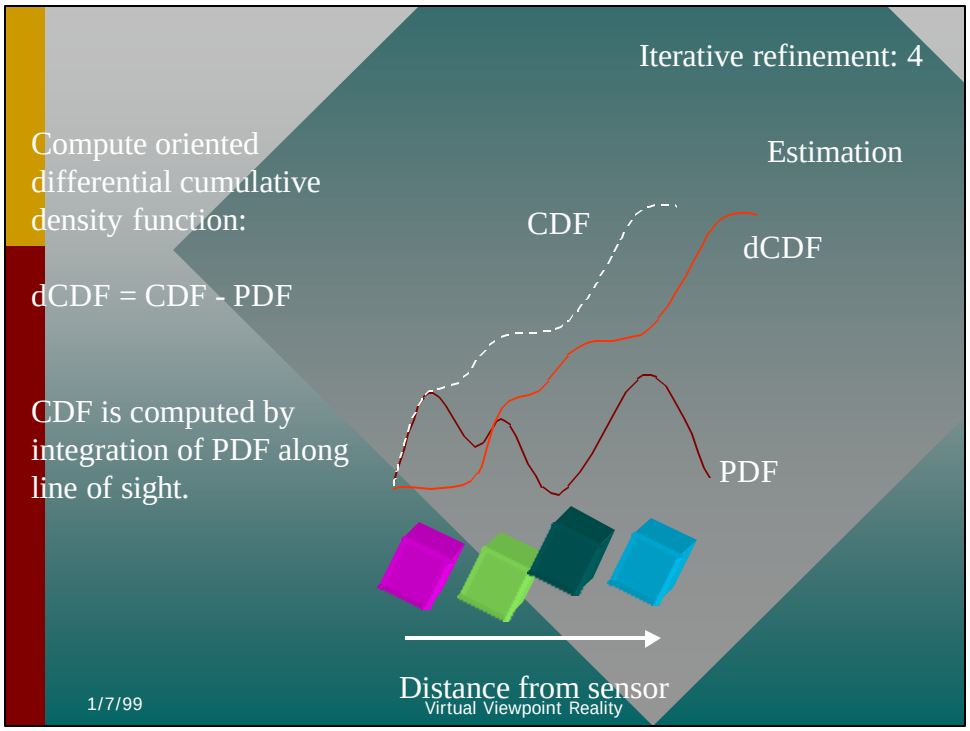


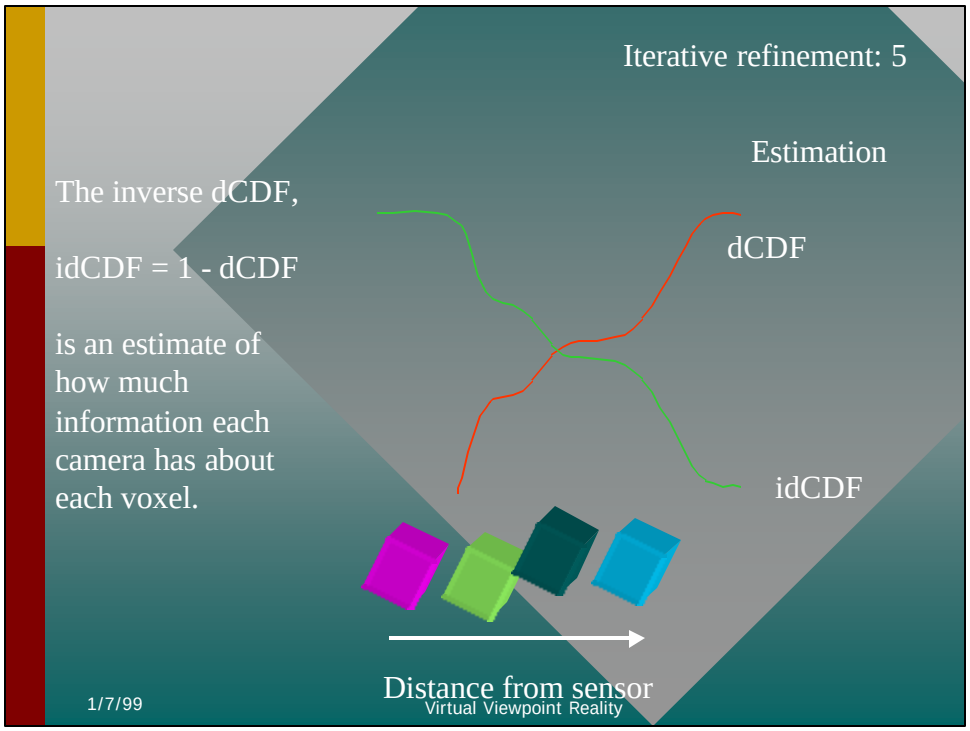
- Captures shape very well
- Intensities are not perfect

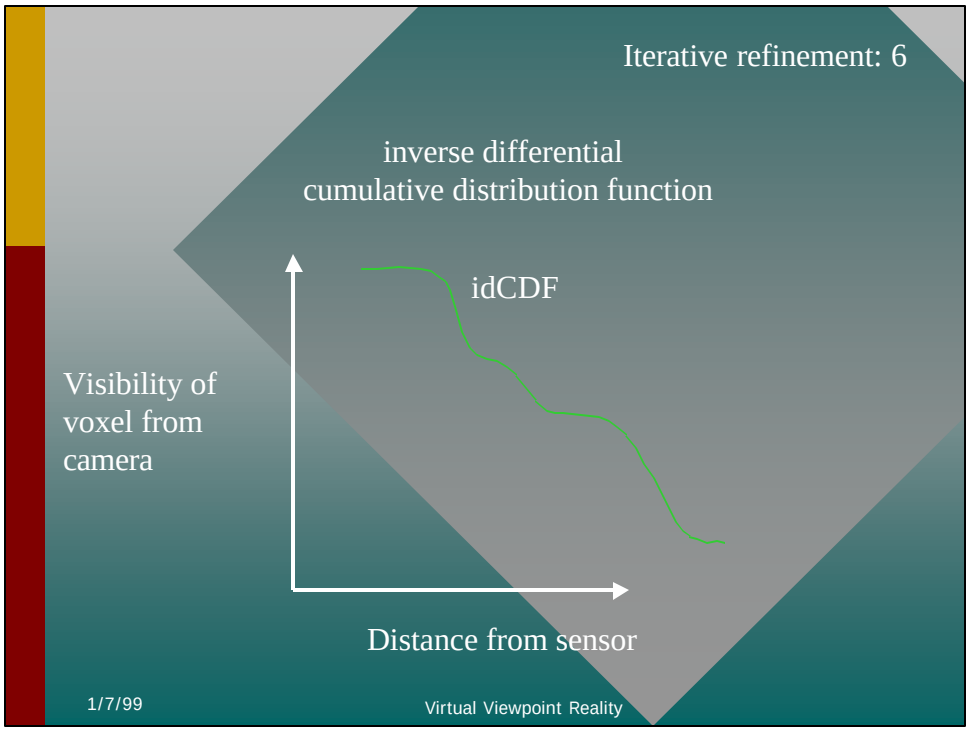

Iterative refinement: 1
Confidence
Confidence measurement
1/7/99
Virtual Viewpoint Reality

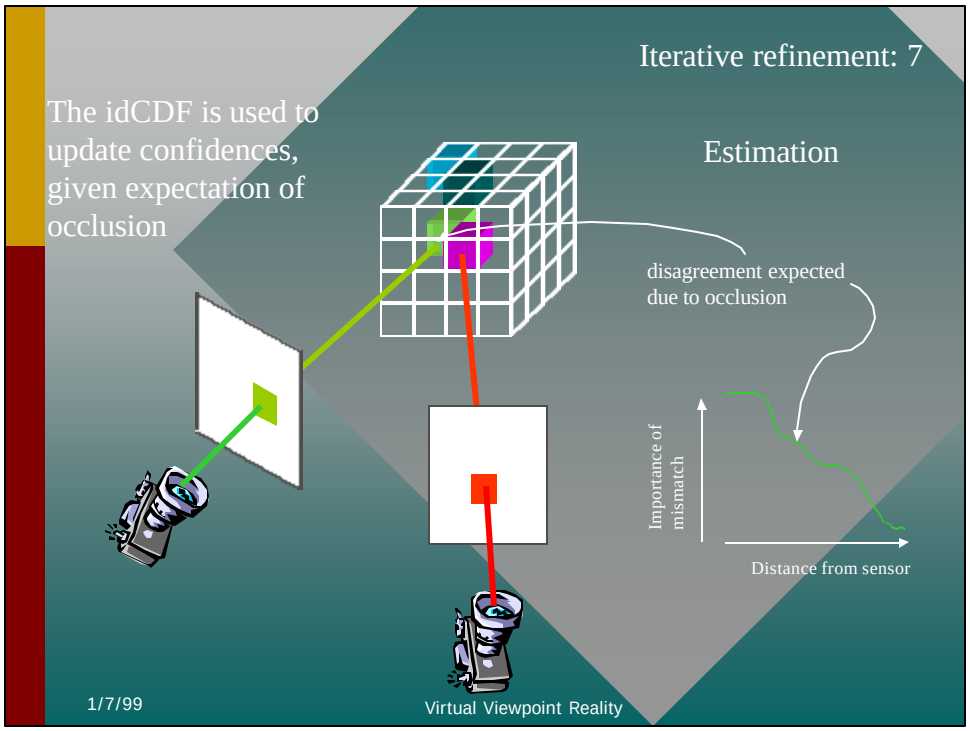

Iterative refinement: 2
Constraint Application
Normalize along ray to obtain estimate probability that voxel is visible from camera
1/7/99
Virtual Viewpoint Reality

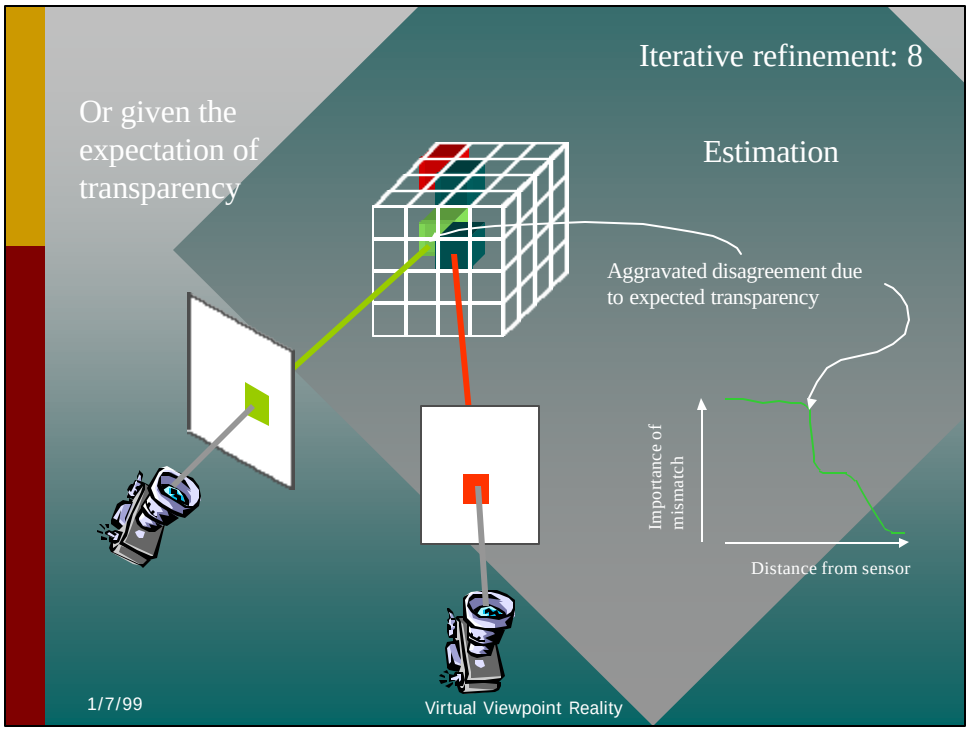

Iterative refinement: 3
Constraint
Application
PDF
Distance from sensor
1/7/99
Virtual Viewpoint Reality

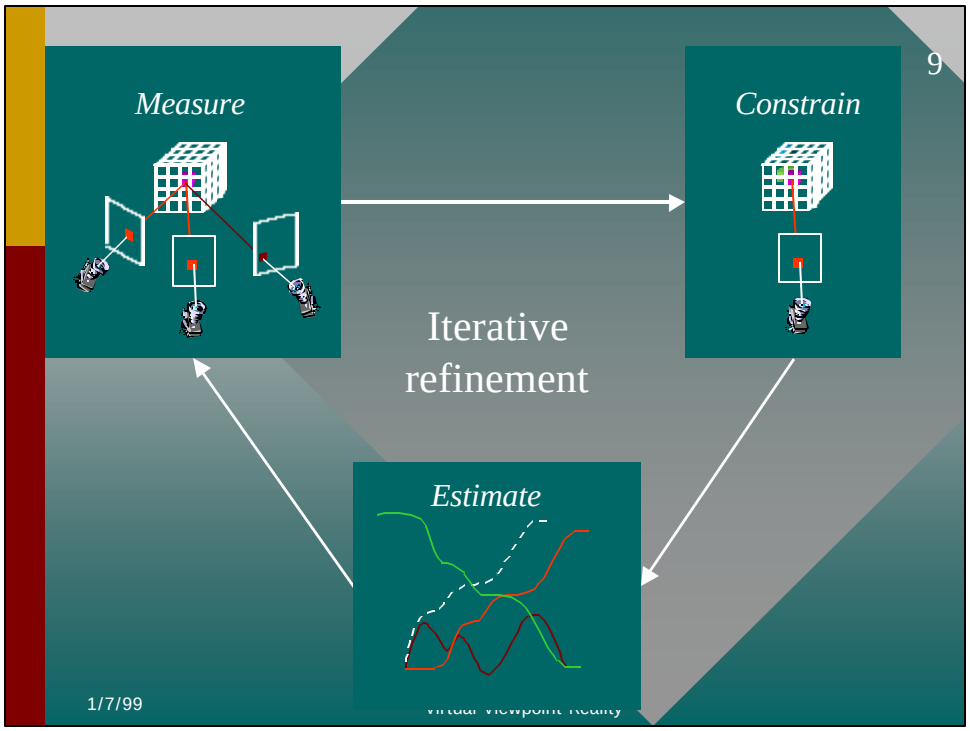

Iterative refinement: 4
Compute oriented differential cumulative density function:
dCDF = CDF - PDF
CDF is computed by integration of PDF along line of sight.
Estimation
CDF
dCDF
PDF
Distance from sensor
1/7/99
Virtual Viewpoint Reality

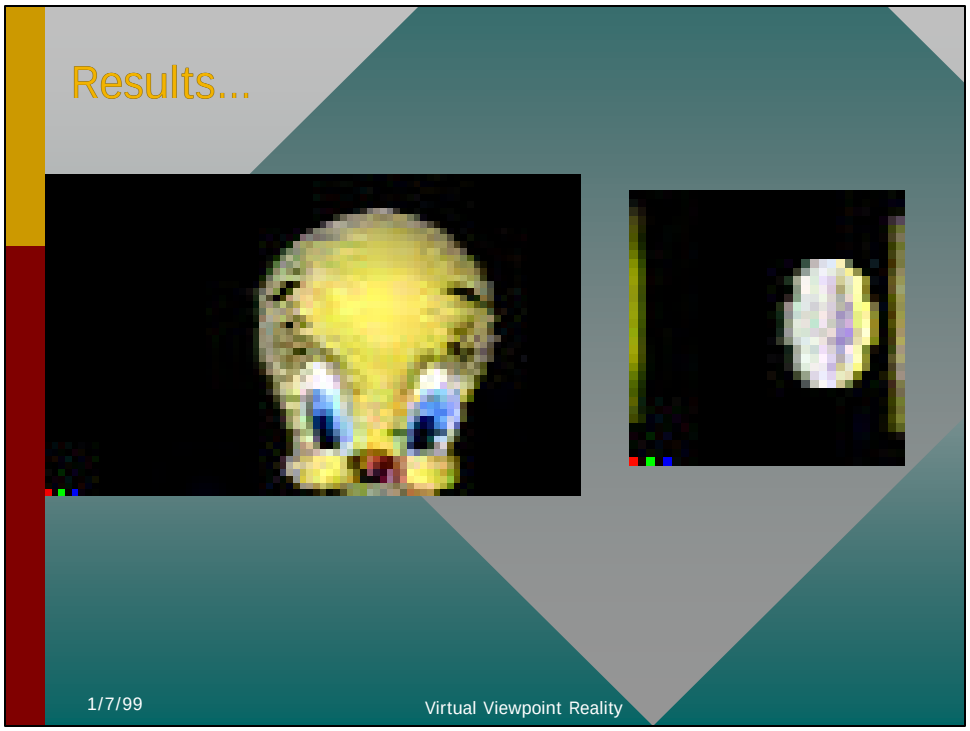

Iterative refinement: 5
Estimation
The inverse dCDF,
idCDF = 1 - dCDF
is an estimate of how much information each camera has about each voxel.
dCDF
idCDF
Distance from sensor
1/7/99
Virtual Viewpoint Reality


Iterative refinement: 6
inverse differential cumulative distribution function
idCDF
Visibility of voxel from camera
Distance from sensor
1/7/99
Virtual Viewpoint Reality


Iterative refinement: 7
The idCDF is used to update confidences, given expectation of occlusion
Estimation
disagreement expected due to occlusion
Importance of mismatch
Distance from sensor
1/7/99
Virtual Viewpoint Reality


Iterative refinement: 8
Or given the expectation of transparency
Estimation
Aggravated disagreement due to expected transparency
Importance of mismatch
Distance from sensor
1/7/99
Virtual Viewpoint Reality


9
Measure
Constrain
Iterative
refinement
Estimate
1/7/99


Results...
1/7/99
Virtual Viewpoint Reality

Overview of VVR Meeting

- Motivation from MIT …
- Discuss current and related work (MIT)
  - Video Activity Monitoring and Recognition
  - 3D Modeling
  - Demonstrations
- Related NTT Efforts
- Discussion of collaboration
- Future work
- Lunch