Imperial College London

An Introduction to Image Based Rendering and Plenoptic Layers

By Jesse Berent

January 25, 2008

Imperial College London

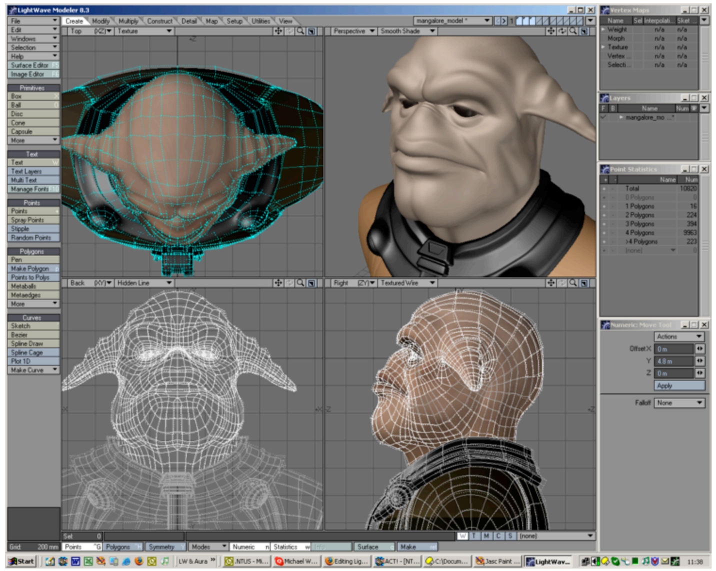
Talk Outline

1. Image based rendering and traditional graphics
2. The plenoptic function: sampling and interpolation
3. Layers and plenoptic layers
4. Experimental results
5. Conclusion and future insights

Traditional graphics rendering

- Detailed geometry and texture/reflectance maps are available or estimated using stereo modeling methods
- New views are synthesized by projecting the objects onto the virtual camera planes
- Source description

[Obtained from middlebury stereo vision]



[obtained from wikipedia]

Exact geometry is difficult to obtain in complex scenes !

Imperial College London

Image Based Rendering

- Vision systems detect electro-magnetic radiation
- Image based rendering uses many images of the scene (100-1000 !!!)
- New views are obtained by interpolating intensities from nearby images - Don’t need to model anything
- Enables photorealistic rendering of complicated environments (we all remember the Matrix scene…)
- appearance description

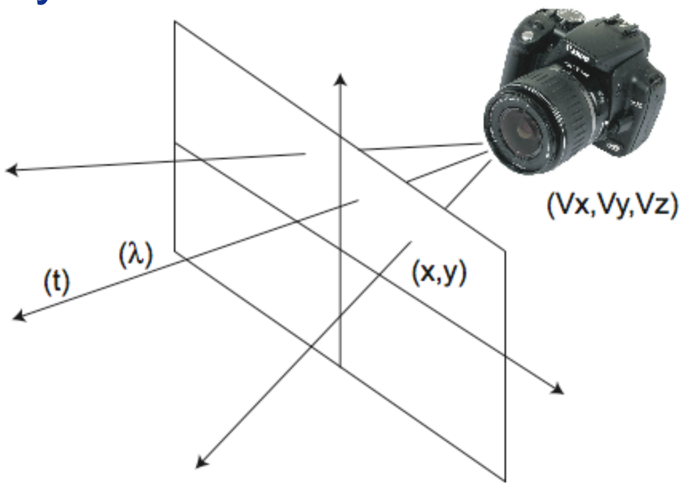
The Plenoptic Function

- 7D function that describes the intensity of each light ray that reaches a point in space [AdelsonB:91]

$$I = I_7(x, y, \lambda, t, v_x, v_y, v_z)$$

- Assumptions can be made to reduce the high number of dimensions
  - Wavelength remains constant (unless ray is occluded)
  - 3 channels for RGB or 1 channel for grayscale
  - Static scenes
  - Viewing position constraints

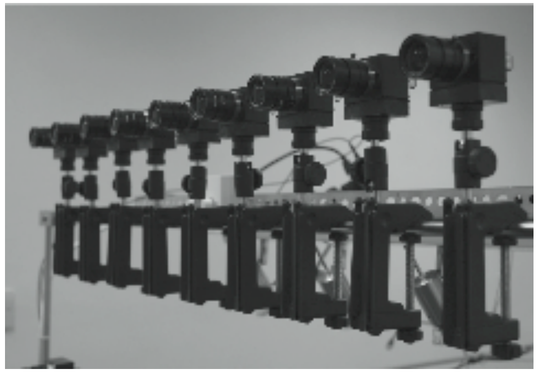
“the sole communication link between physical objects and their
corresponding retinal images” - Adelson/Bergen



Different camera setups

3D (x,y,t)

3D (x,y,vx)

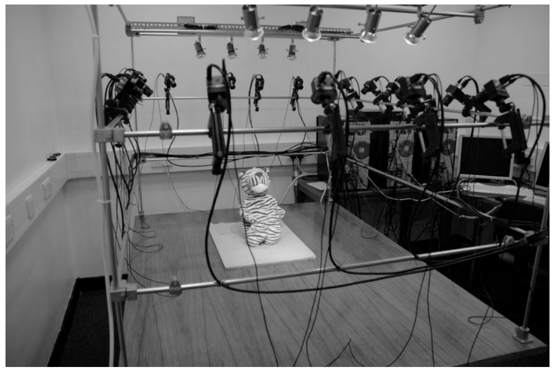


2D (x,y)

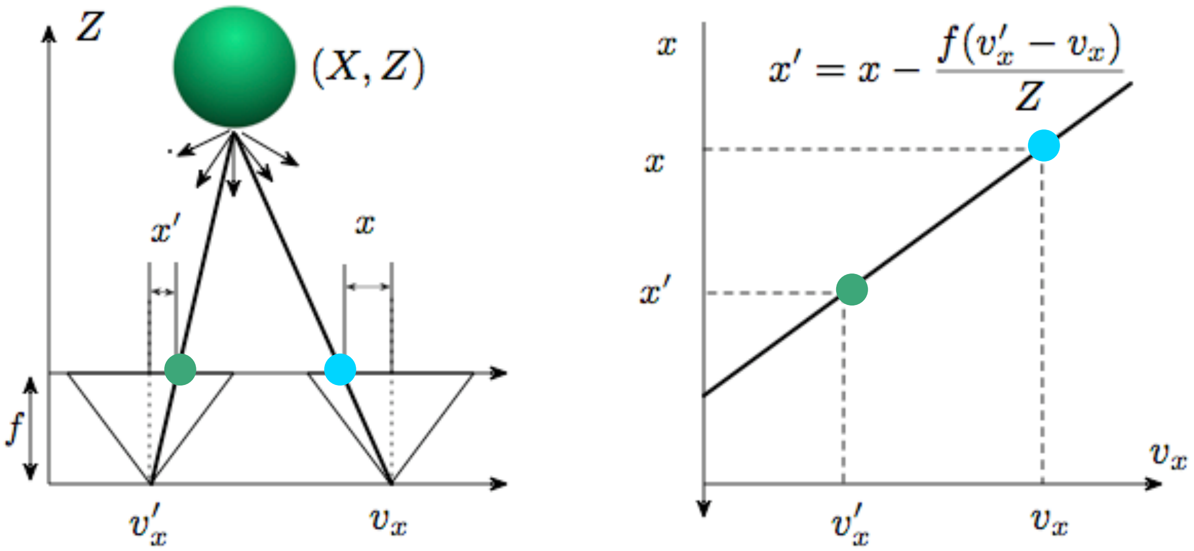
4D (x,y,vx,vy)

4D (x,y,vx,t)

5D (x,y,vx,vy,vz)



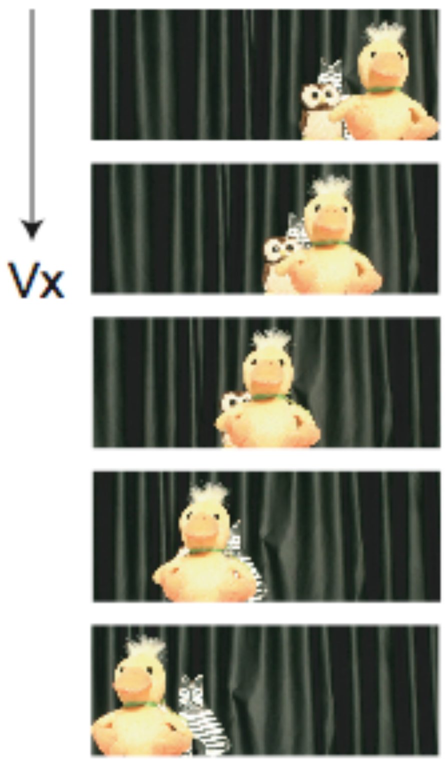
The Epipolar Plane Image (EPI)

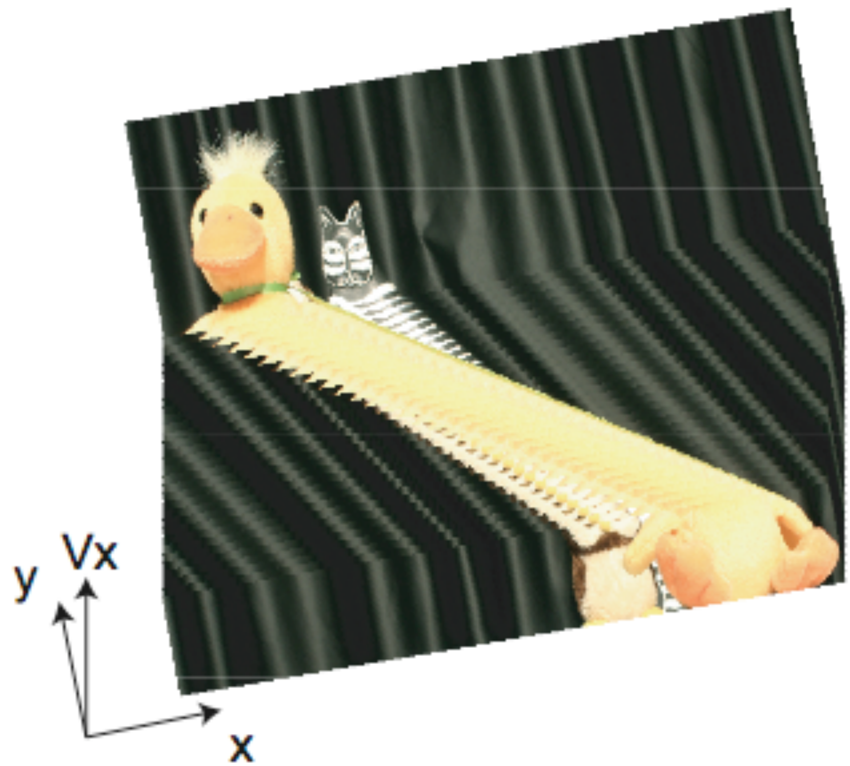


(a) (b)

- First introduced in [BollesBM:87]
- 3D parameterization of the plenoptic function
- Particular structure: points are mapped onto lines in the EPI
- Intensity along the line is approximately constant
- The slope of the line ∝ 1/depth

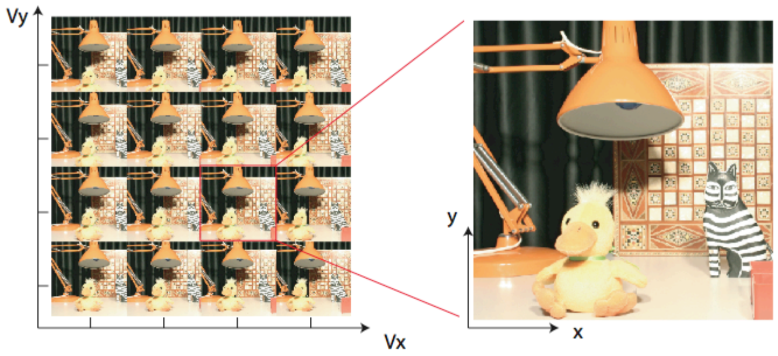
The Epipolar Plane Image (EPI)



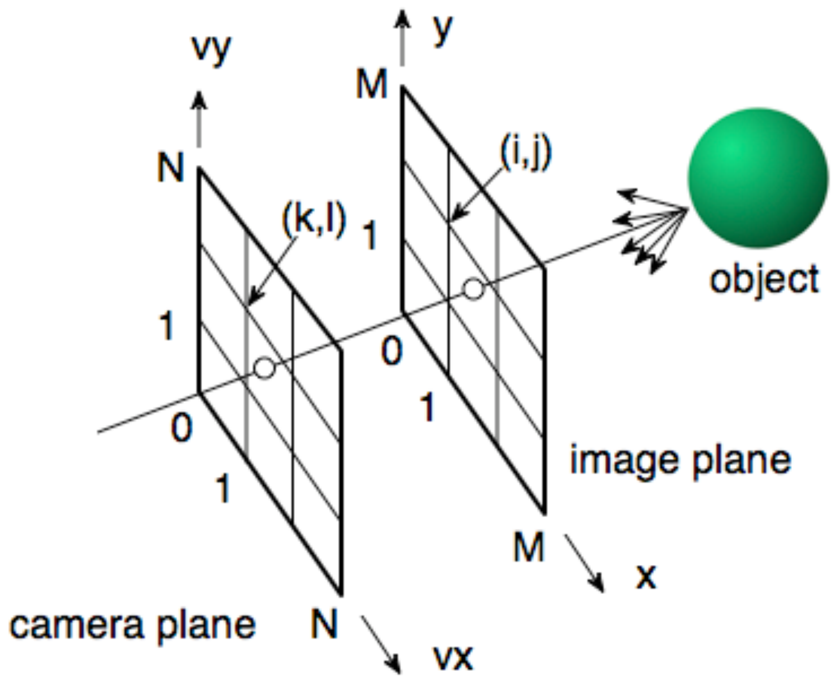


The Light Field

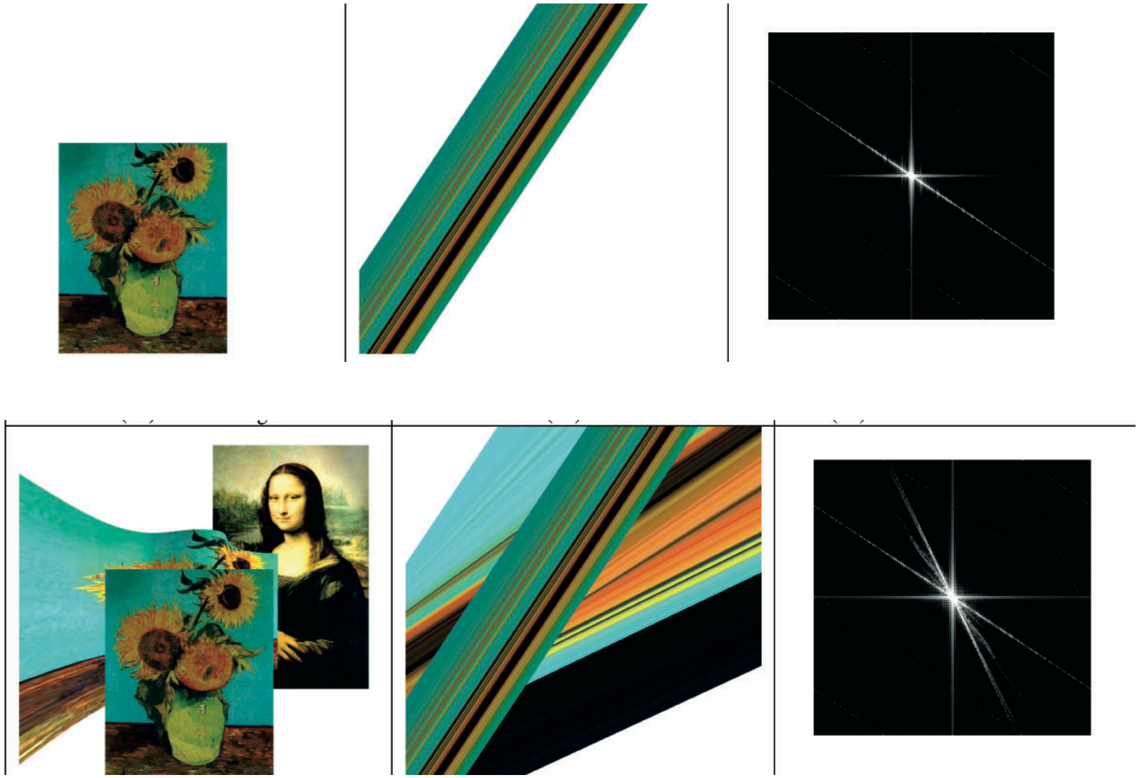
- First introduced in [LevoyH96]
- Light rays are parameterized by their intersection with the camera plane (vx,vy) and the image plane (x,y)
- 4D parameterization of the plenoptic function



a.k.a ray space



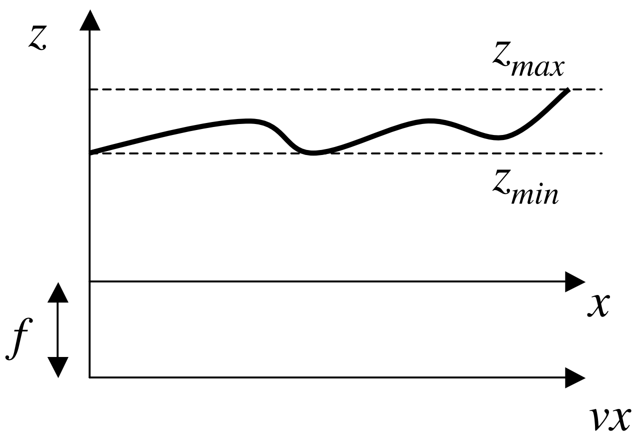
Spectral Analysis

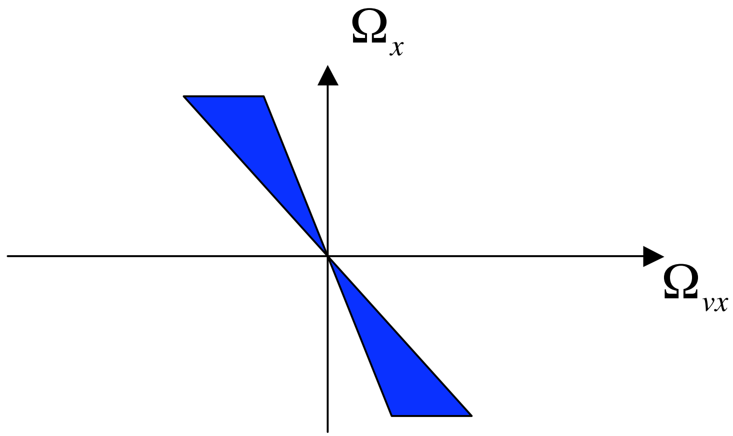


Images from: Plenoptic Sampling, Chai et al., International Conference on Computer Graphics and Interactive Techniques, 2000

Spectral Analysis of Light Fields

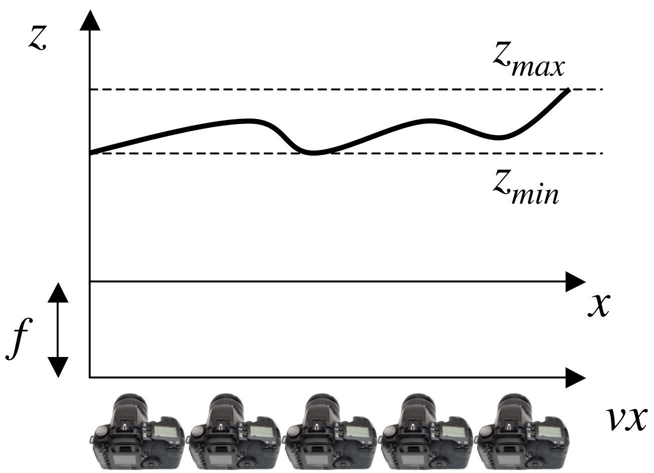
- Sampling and interpolation in a traditional sense: Look at the spectrum
- The light field is approximately bandlimited [ChaiCST:00, ZhangC:03]
- Bound by minimum and maximum depth in the scene
- Conclusion: Sampling and interpolation is possible

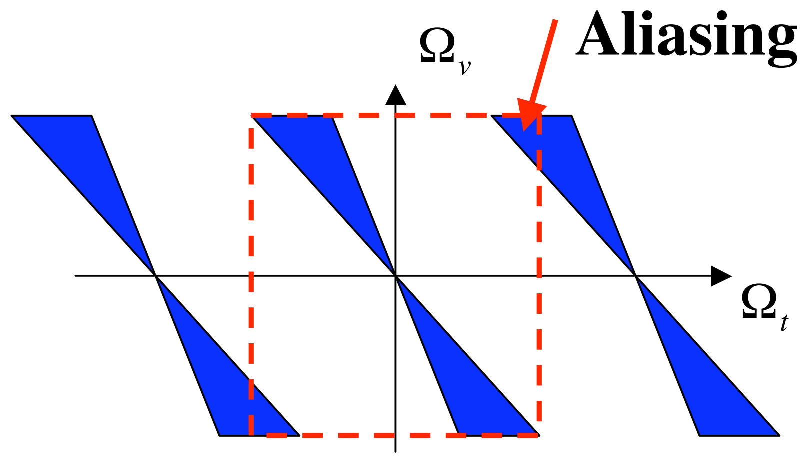




Spectral Analysis of Light Fields

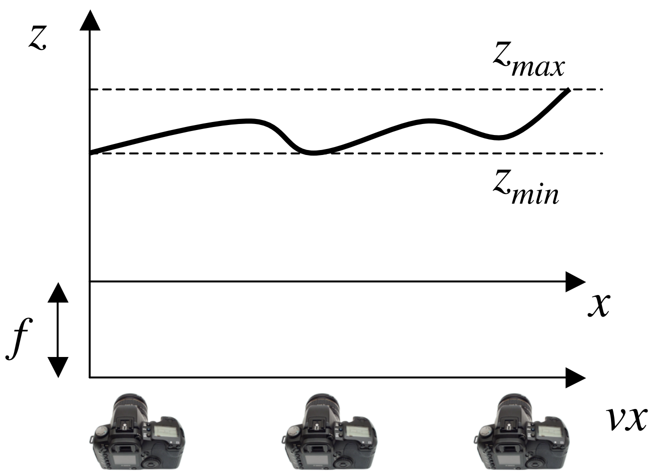
- Sampling and interpolation in a traditional sense: Look at the spectrum
- The light field is approximately bandlimited [ChaiCST:00, ZhangC:03]
- Bound by minimum and maximum depth in the scene
- Conclusion: Sampling and interpolation is possible

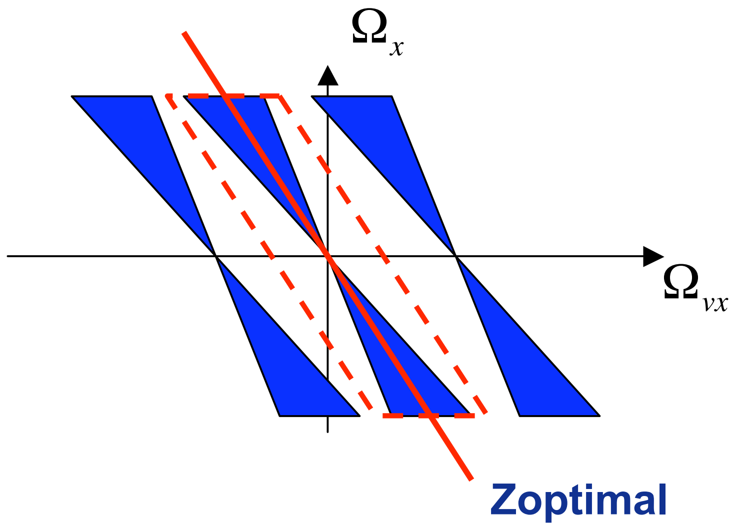




Spectral Analysis of Light Fields

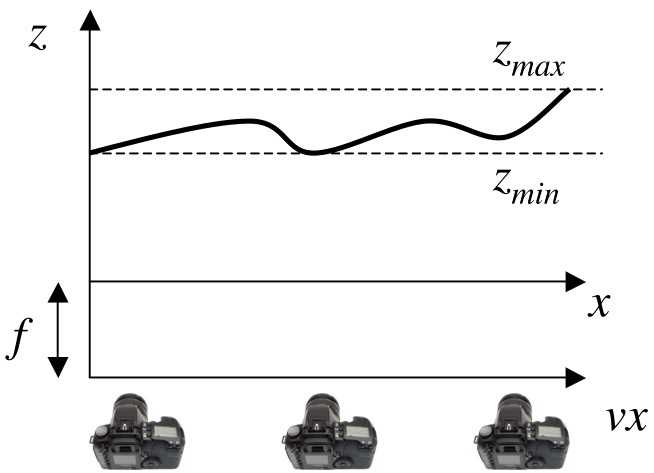
- Sampling and interpolation in a traditional sense: Look at the spectrum
- The light field is approximately bandlimited [ChaiCST:00, ZhangC:03]
- Bound by minimum and maximum depth in the scene
- Conclusion: Sampling and interpolation is possible

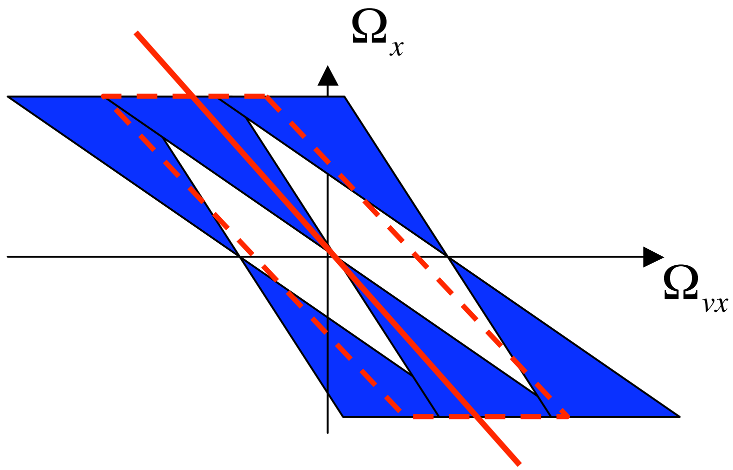




Spectral Analysis of Light Fields

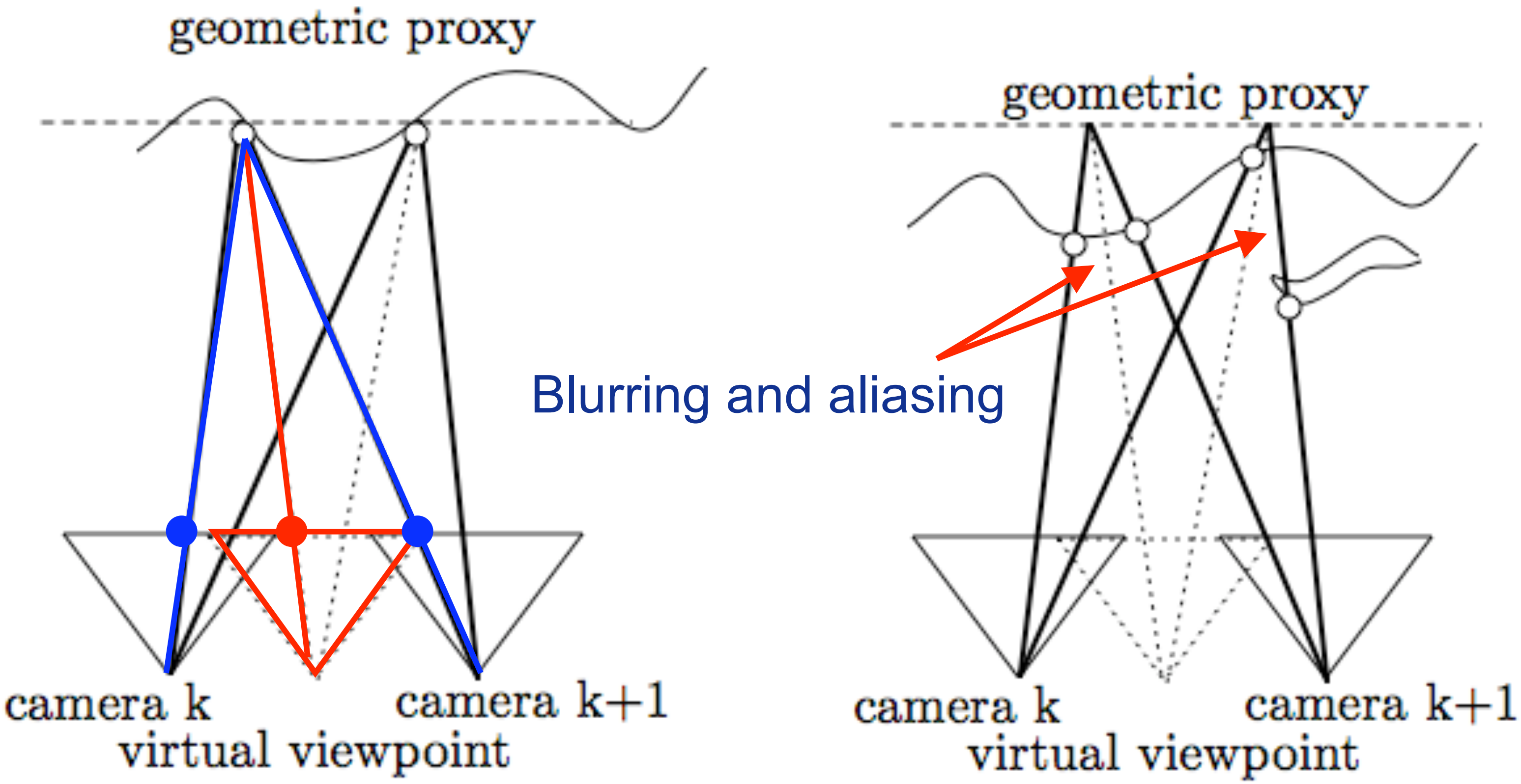
- Sampling and interpolation in a traditional sense: Look at the spectrum
- The light field is approximately bandlimited [ChaiCST:00, ZhangC:03]
- Bound by minimum and maximum depth in the scene
- Conclusion: Sampling and interpolation is possible





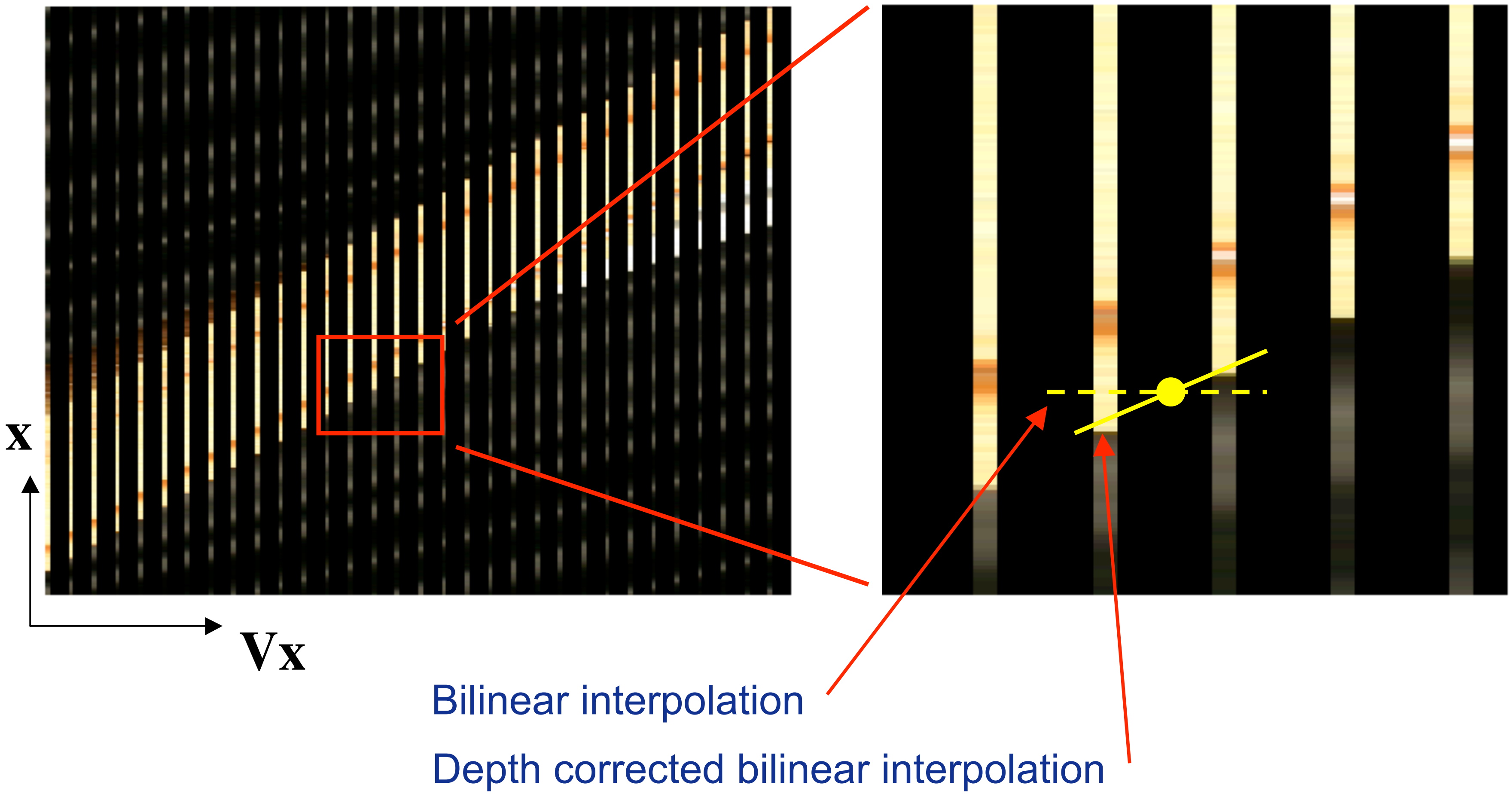
Larger depth variations

Spatial analysis



- Representing the scene's geometry with a single depth plane

Depth corrected interpolation



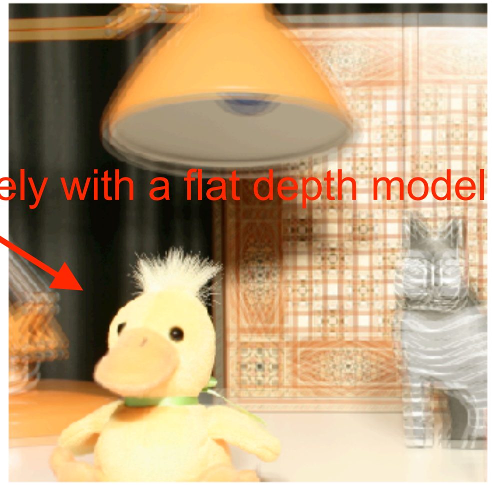
Linear interpolation with a sparse LF

D = infinity

D = back

D = dopt

Duck is not flat
but rendered nicely with a flat depth model



D = front

Preliminary conclusion

- Image based rendering allows photorealistic rendering of complicated scenes
- Little or no scene geometry is used
- Light Fields are approximately bandlimited - Sampling and interpolation in a traditional sense is possible
- However scenes with large depth variations and occlusions have a wider spectrum and result in interpolated viewpoints that are aliased
  - Solution 1: Add more sample images…
  - Solution 2: decompose the scene into regions with small depth variation and find occlusion boundaries

Imperial College London

Image based rendering vs traditional graphics

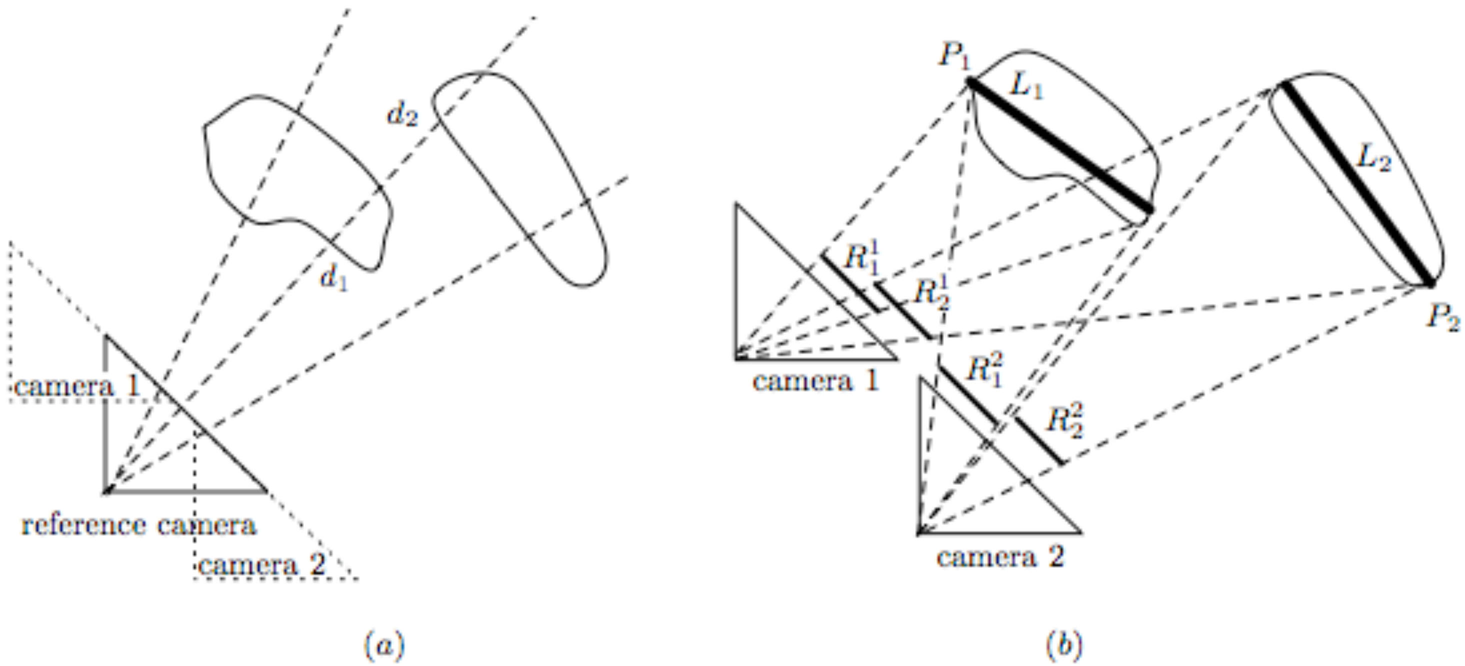
- At one end: Source description
- Small amount of images
- very accurate geometry of the scene (can be difficult to obtain)

- Plenoptic layers: Somewhere in the middle

- At the other end: Appearance description
- no geometry
- Huge amount of images (can be impractical to capture, store, …)

Imperial College London

Layered representations

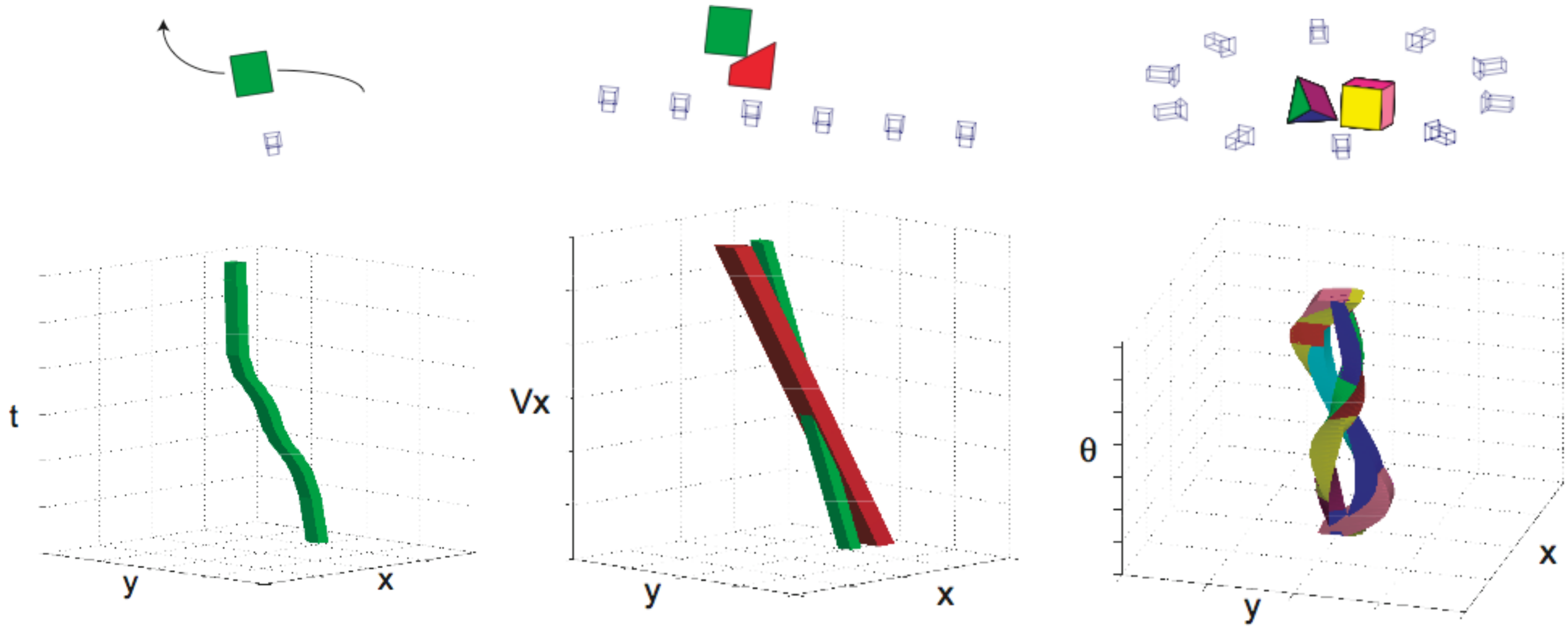


Layered depth images [ShadeGHS,98]:
- Accurate geometry
-One reference view
-Warping is used to interpolate

Coherent layers - Plenoptic layers:
- Approximate geometry
- No reference view
-Nearby images are used to interpolate

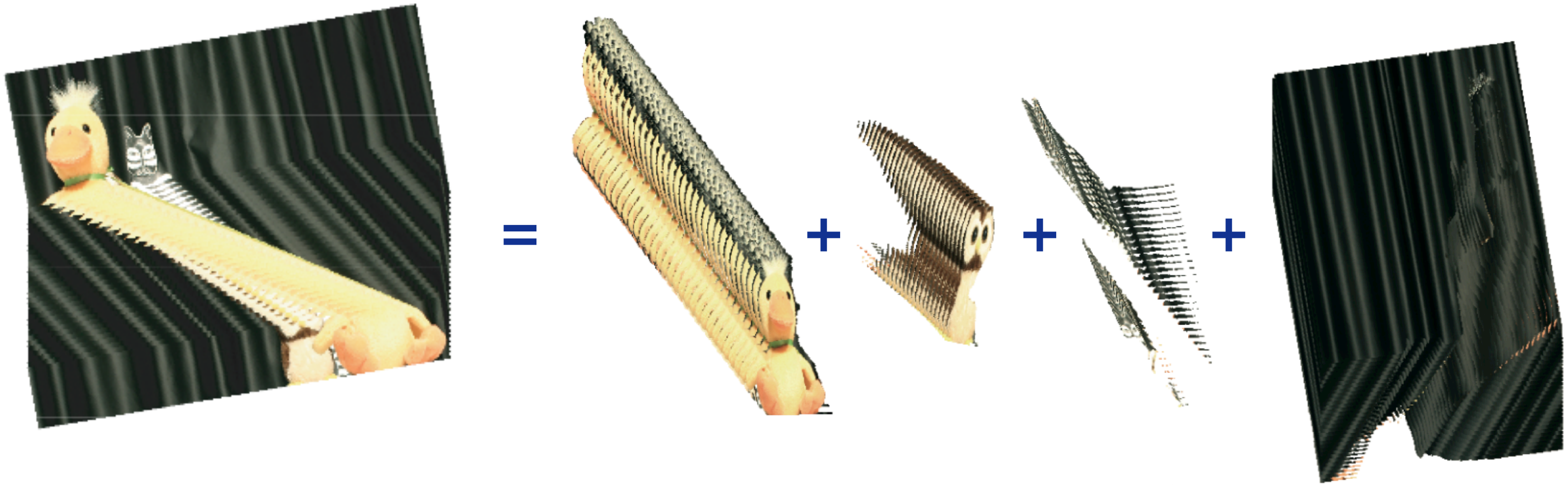
Plenoptic layers

- Plenoptic function is made of a collection of plenoptic layers [Coherent layers: ShumSYLY04, BerentD07]
- Shape and occlusions are constrained by the camera setup



Imperial College London

Plenoptic layers



- Recomposing plenoptic layers reconstructs perfectly the data
- Each layer is modeled with a simple geometry (i.e. a plane)

Imperial College London

A semi-parametric approach to segment light field data

- Segment the light field into regions that can be rendered free of aliasing:
  - Global for added robustness and handling of occlusions
  - Take into account the particular structure of the data
  - Modular in terms of the number of plenoptic layers used to represent the light field:
    - Few layers: less computation, less rendering quality
    - Many layers: more computation, better rendering quality

Segmentation of light fields

- Plenoptic layer carves out a 4D hypervolume in the light field
- Use a 4D active contour method to minimize

$$E_{tot}(\tau) = \sum_{n=1}^{N} \int_{\mathcal{H}_n^{\perp}(\tau)} d_n(\vec{x}) d\vec{x}$$

$d_n(\vec{x})$ is a measure of consistency with hypervolume n

- Minimize the function subject to plenoptic constraints

Imperial College London

Active Contours

- Consider a cost function of the type:

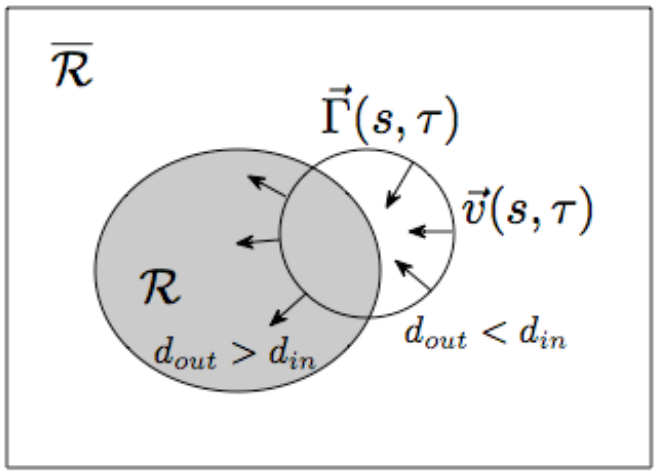
$$E(\Gamma) = \int_{\mathcal{R}} d_{in}(\vec{x}) d\vec{x} + \int_{\overline{\mathcal{R}}} d_{out}(\vec{x}) d\vec{x} + \int_{\Gamma} \lambda d\vec{s}$$

- Gradient [KassWT:88, CasellesKS:97, Jehan-BessonBA:01]

$$dE(\Gamma, \vec{v}) = \int_{\Gamma} [d_{in}(\vec{x}) - d_{out}(\vec{x}) - \lambda\kappa](\vec{v} \cdot \vec{N}) d\vec{s}$$

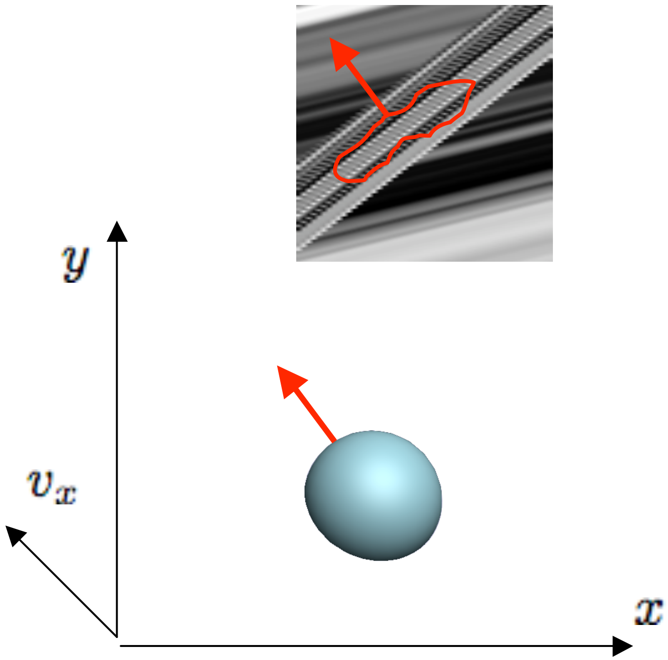
- Steepest descent

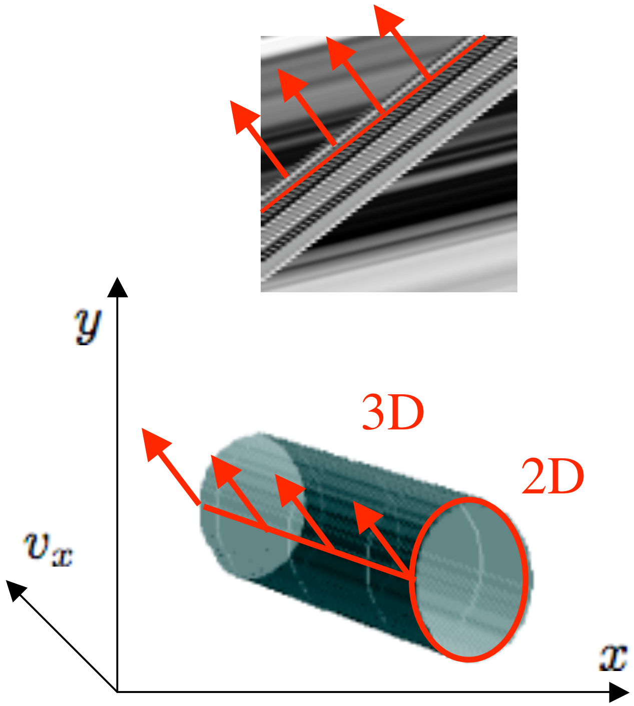
$$\vec{v} = [d_{out}(\vec{x}) - d_{in}(\vec{x}) + \lambda\kappa]\vec{N}$$



Shape constraints

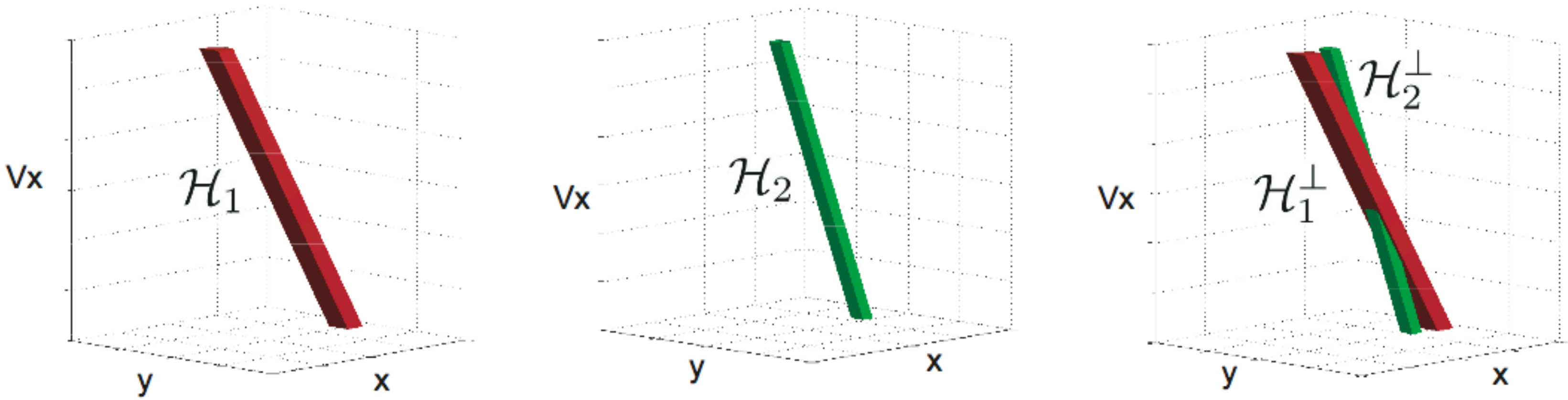
- The structure of the camera array is assumed to be known
- The shape of the plenoptic layer is constrained
- Leads to constrained surface evolution that can be implemented in a 2D subspace - reduces computational complexity





Occlusions

- A line with a larger slope will always occlude a line with a smaller one
- Occlusions occur at line intersections
- Occlusions are explicit



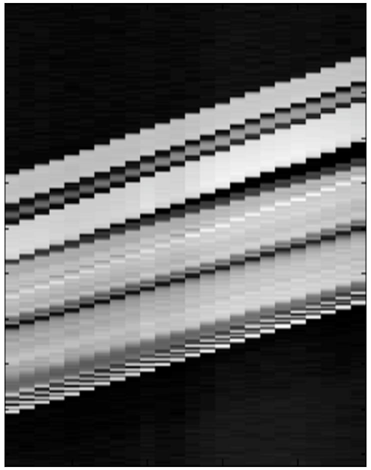
$$\mathcal{H}_1^{\perp} = \mathcal{H}_1 \qquad \mathcal{H}_2^{\perp} = \mathcal{H}_2 \cap \overline{\mathcal{H}_1^{\perp}}$$

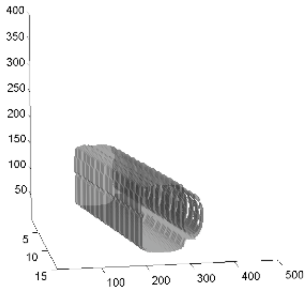
Overall optimization

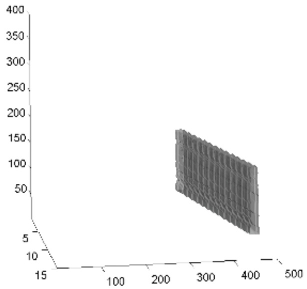
- Initialize
  - Chose number of layers
  - Use available depth or stereo method to initialize layers
- Iteratively alternate
  - Segmentation given layer depth maps
    - Evolve each contour iteratively with the level set method
  - Estimation of depth maps given segmentation and update occlusion ordering
- End when there is no significant decrease in energy or after a predetermined number of iterations

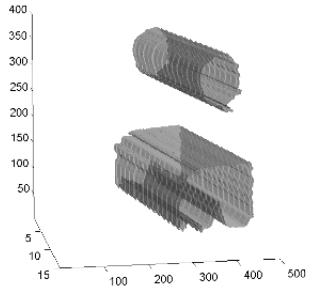
Experimental Results

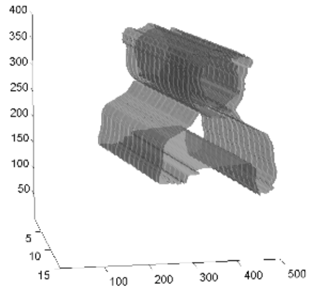
- Tiger image sequence (15 images)

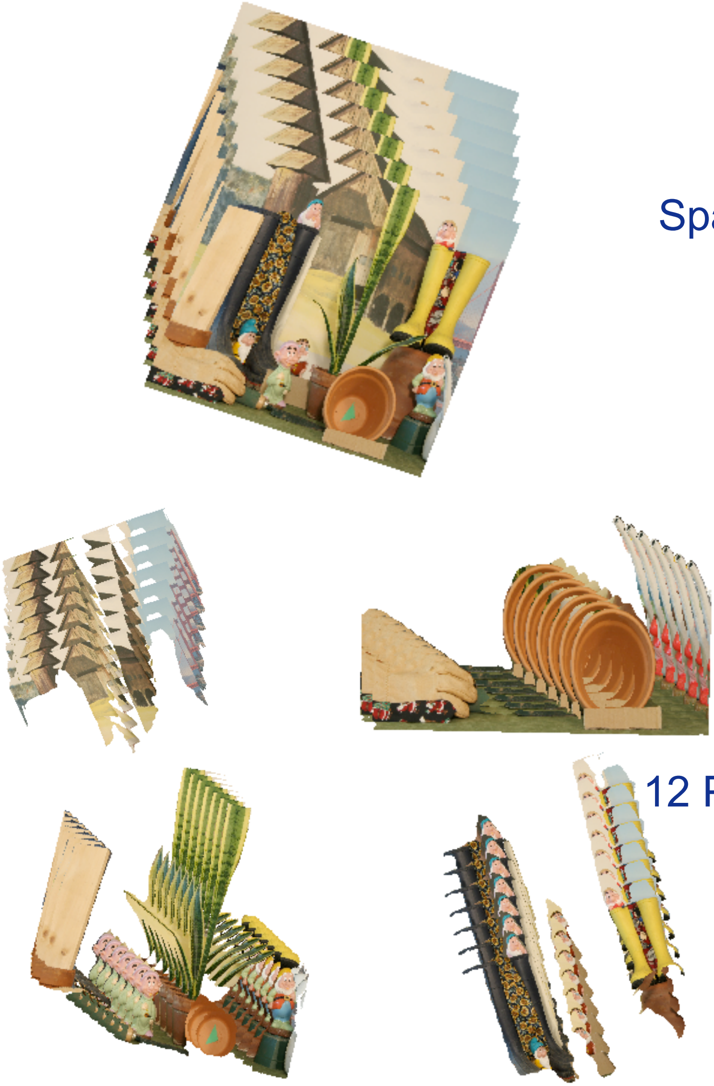









Dwarves dataset



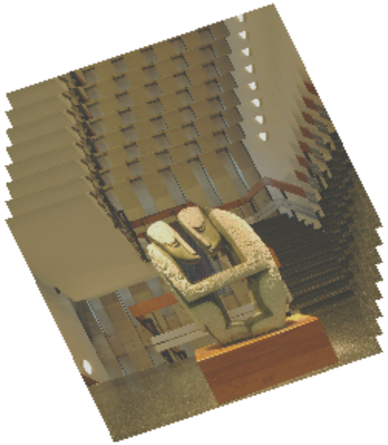
Sparse LF (6 x 1)

12 Plenoptic layers

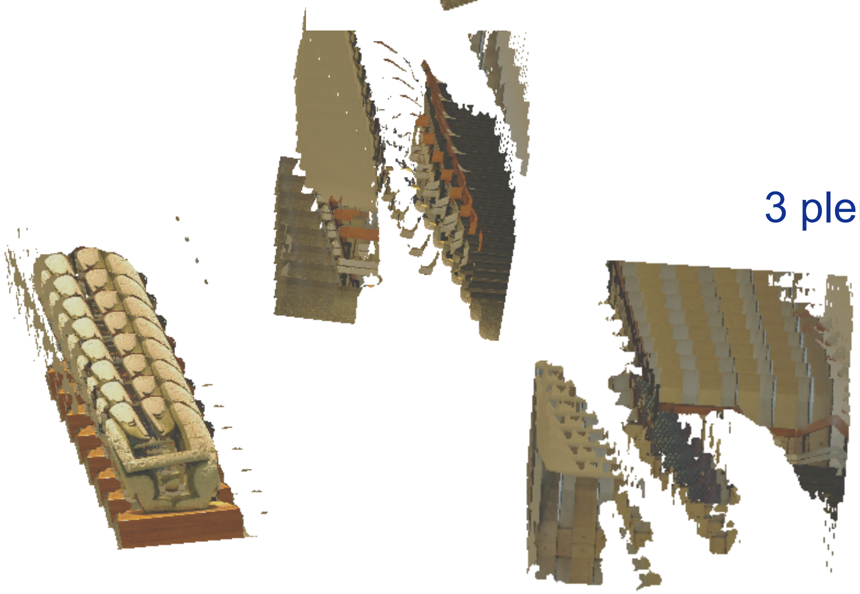
[Dataset courtesy of middlebury stereo vision]

Imperial College London

EE department lobby

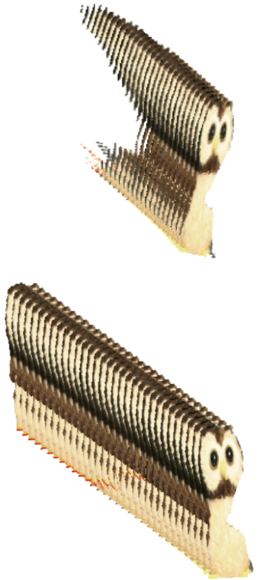


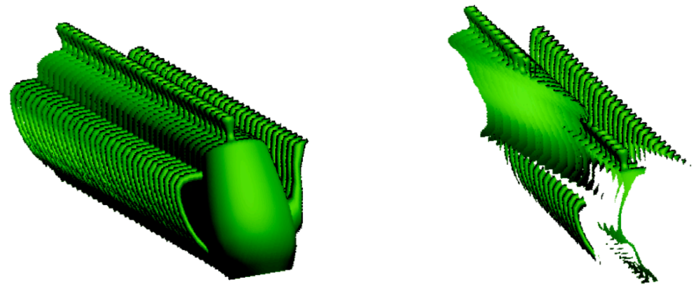
Sparse LF (5 by 1)

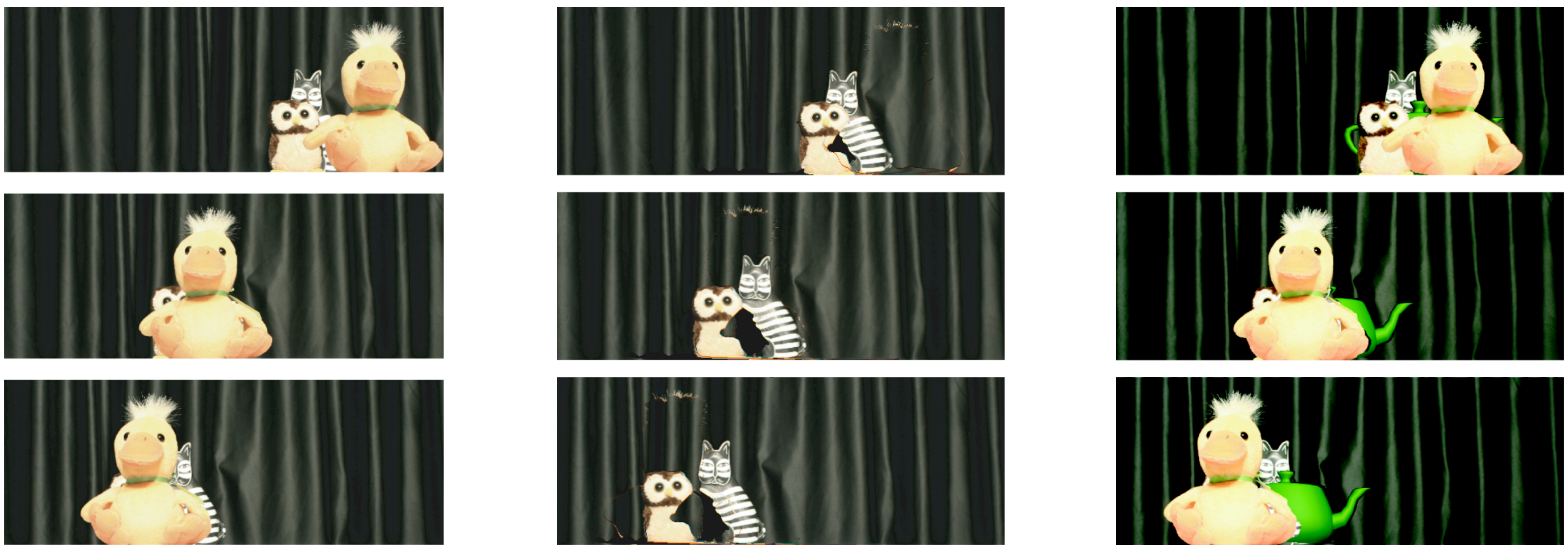


3 plenoptic layers

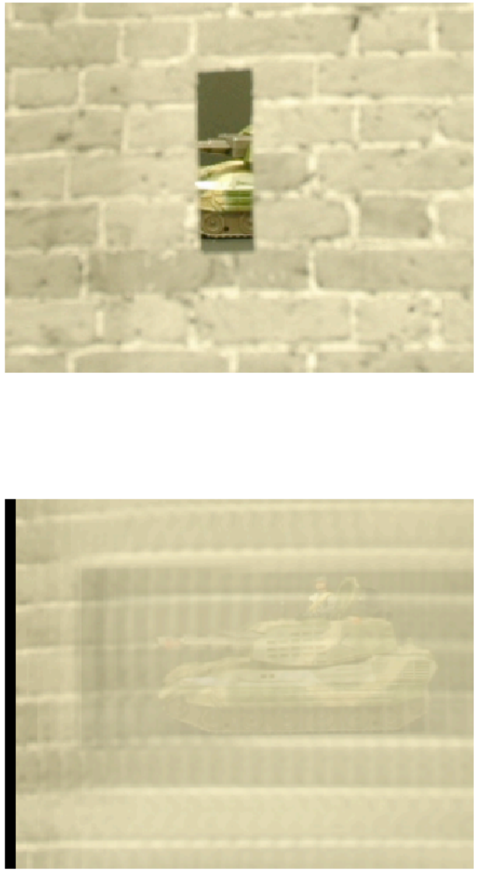
Scene manipulation



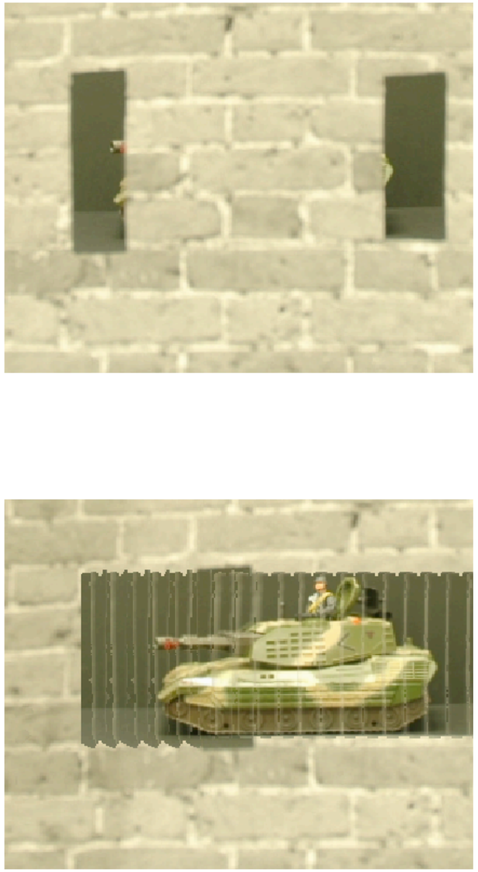




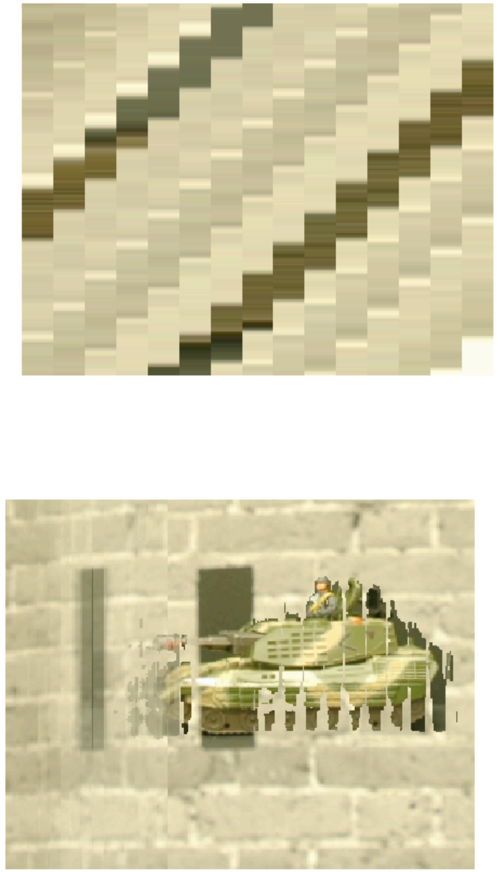
Occlusion removal



Synthetic aperture



Plenoptic layers



State-of-the-art stereo

Conclusions

- The plenoptic function provides a nice framework for multiview image analysis and IBR
- Plenoptic layers capture the coherence of the plenoptic function and enable good rendering quality with only a few layers
- Segmentation scheme:
  - Constrained surface evolution (uses knowledge of camera setup for added robustness)
  - Takes into account all the images simultaneously
  - Handles occlusions
  - Is scalable to higher dimensions (i.e. more general plenoptic functions such as dynamic light fields)
  - However, requires a structured and constrained camera setup

What's in store for the future

- Freeviewpoint imaging and 3DTV is a hot topic both in academia and industry
- The problem is still not completely solved…
- Real-time freeviewpoint video requires a lot of processing power and memory!

- For more information: There was an IEEE Signal Processing Magazine special issue on Multiview Imaging and 3DTV in November 2007

Imperial College London

Questions?