IEEE VR 2003 tutorial 1

Recent Methods for Image-based Modeling and Rendering

Darius Burschka
Johns Hopkins University

Greg Hager
Johns Hopkins University

Dana Cobzas
University of Alberta

Martin Jagersand
University of Alberta

Zach Dodds
Harvey Mudd College

Keith Yerex
Virtual Universe Corporation

Image-based Modeling Rendering

IBR/IBM: Label on a wide range of techniques

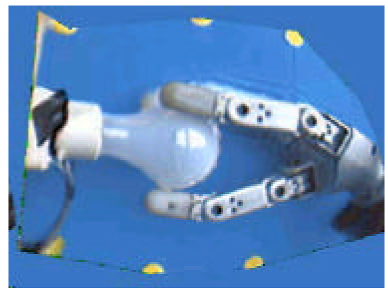
Promising for various reasons, e.g.:

1. Cameras are cheap/common while 3D laser range sensors are expensive and manual modeling time consuming.
2. Achieving photo-realism is easier if we start with real photos.
3. Speed up graphics rendering by warping and blending whole images instead of building them from components in each frame.

- Common trait: Images serve important role. Partially or wholly replaces geometry and modeling.

Image-based Models from consumer cameras

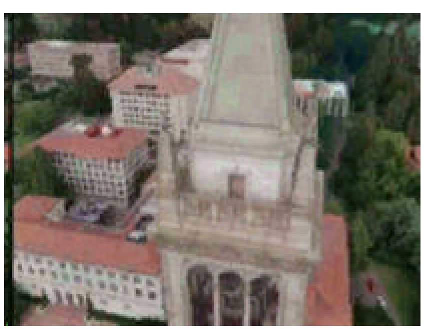
- Rendering of models obtained using a $100 web cam and a home PC (Cobzas, Yerex, Jagersand 2002)



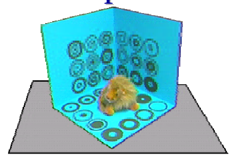
We'll learn how to do this in the lab this afternoon…

Photo-Realism from images

1. Geometry+images
   (Debevec – Camillo Façade)

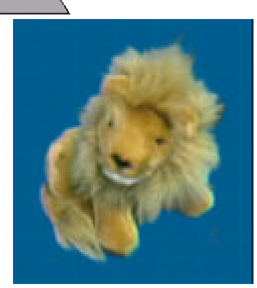


2. Set of all light rays – Plenoptic function



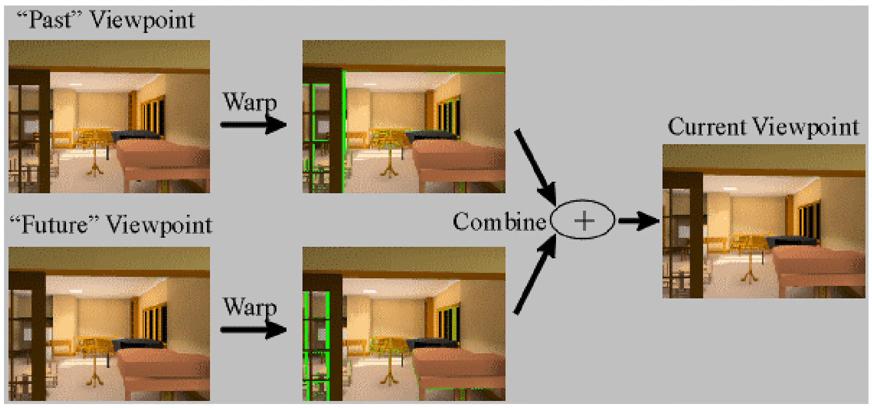
Capture

Render new views



Rendering speed-up

Post-warping images (Mark and Bishop 1998)



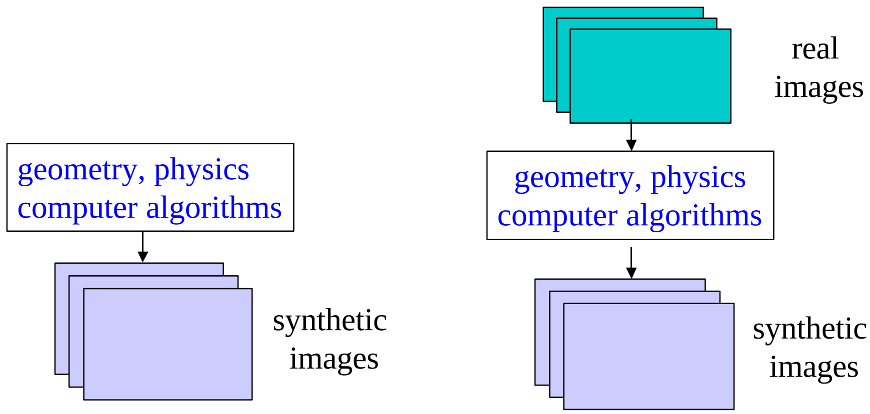
Rendering speed-up

Blending a light basis (Yerex, Jagersand)

Modeling: Two Complementary Approaches

- Conventional graphics

Image-based modeling and rendering



Confluence of Computer Graphics and Vision

- Traditional computer graphics (image synthesis, forward modeling)
  - Creating artificial images and videos from scratch
- Computer vision & image processing (image analysis & transformation, inverse modeling)
  - Analyzing photographs & videos of the real world

- Both fields rely on the same physical & mathematical principles and a common set of representations
- They mainly differ on how these representations are built

Object & Environment Modeling

- Basic techniques from the conventional (hand) modeling perspective:
  - Declarative: write it down (e.g. typical graphics course)
  - Interactive: sculpt it (Maya, Blender …)
  - Programmatic: let it grow (L-systems for plants, Fish motion control)
- Basic techniques from the image-based perspective:
  - Collect many pictures of a real object/environment; rely on image analysis to unfold the picture formation process (principled)
  - Collect one or more pictures of a real object/environment; manipulate them to achieve the desired effect (heuristic)

Rendering

- Traditional rendering
  1. Input: 3D description of 3D scene & camera
  2. Solve light transport through environment
  3. Project to camera’s viewpoint
  4. Perform ray-tracing

- Image-based rendering
  1. Collect one or more images of a real scene
  2. Warp, morph, or interpolate between these images to obtain new views

Important Issues in Image-Based Modeling and Rendering

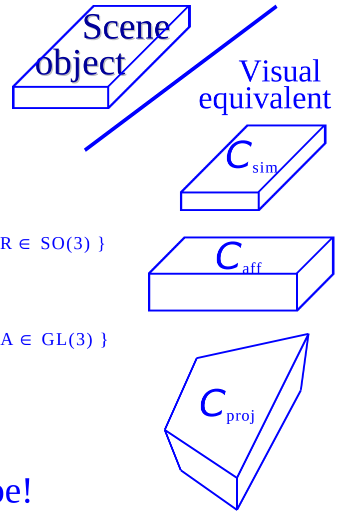
- What are theoretical limits on the information obtained from one or multiple images? (Geometry)
- How to stably and reliably compute properties of the real word from image data? (Comp Vision)
- How to efficiently represent image-based objects and merge multiple objects into new scenes? (CG)
- How to efficiently render new views and animate motion in scenes? (IBR)

Information obtained from images

- Viewing geometry describes global properties of the scene structure and camera motion
  - Traditional Euclidean geometry
  - Past decade surge in applying non-Euclidean (projective, affine) geometry to describe camera imaging
- Differential properties in the intensity image gives clues to local shape and motion.
  - Shape from shading, texture, small motion

Viewing Geometry and Camera Models

(Zach Dodds PhD thesis 2000)



Viewing Geometry

Shape invariant transform

- Euclidean
  - Calibrated camera

  $\{ g \mid g \in GL(4),\ g = \begin{pmatrix} \lambda R & t \\ 0 & 1 \end{pmatrix},\ R \in SO(3) \}$

- Affine
  - “Infinite” camera

  $\{ g \mid g \in GL(4),\ g = \begin{pmatrix} A & t \\ 0 & 1 \end{pmatrix},\ A \in GL(3) \}$

- Projective
  - Uncalibrated cam

  $\{ g \mid g \in GL(4) \}$

Possibly ambigous shape!

Intensity-based Information

- We get information only when there is intesity difference (Baker et.al. 2003)
- Hence there are often local ambiguities

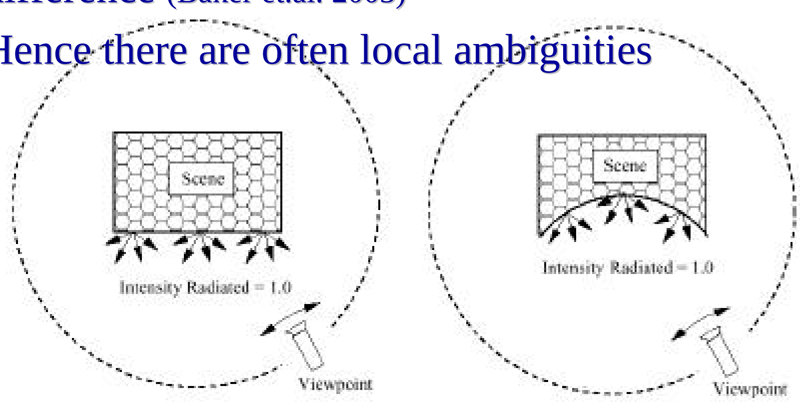
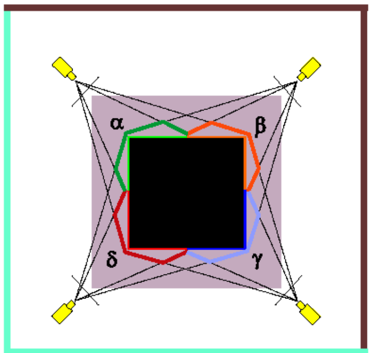
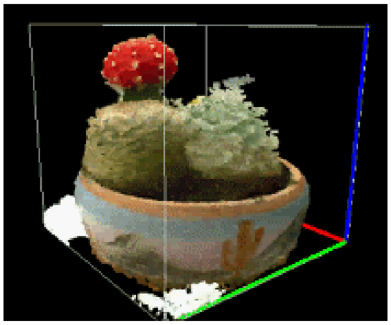


Photo-Consistent Hull

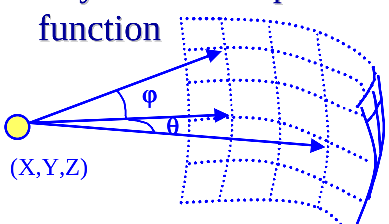
- In cases of structural ambiguity it is possible to define a photo-consistent shape – "visual hull"
  (Kutulakos and Seitz 2001)





Two main representations in Image-Based Modeling

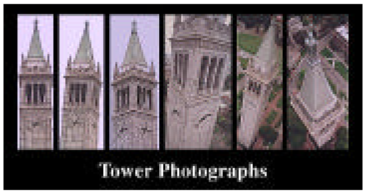
- Ray set = Plenoptic function



Represents

the intensity of light rays passing through the camera center at every location, at every possible viewing angle (5D)"

- Geometry and texture



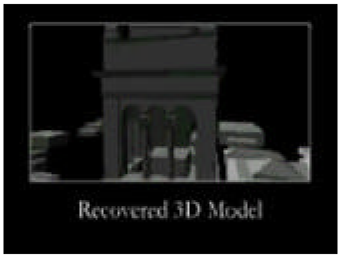
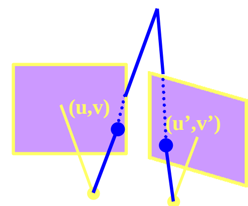


Image Mosaics

- When images sample a planar surface or are taken from the same point of view, they are related by a linear projective transformation (homography).

$$\mathbf{m}=[u,v]^T \qquad \mathbf{m'}=[u',v']^T$$

$$s\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$



- So … images can be mosaicked into a larger image
- 3D plenoptic function.

Cylindrical Panorama Mosaics

- Quicktime VR: Warps from cylindrical panorama to create new planar view (from same viewpoint)

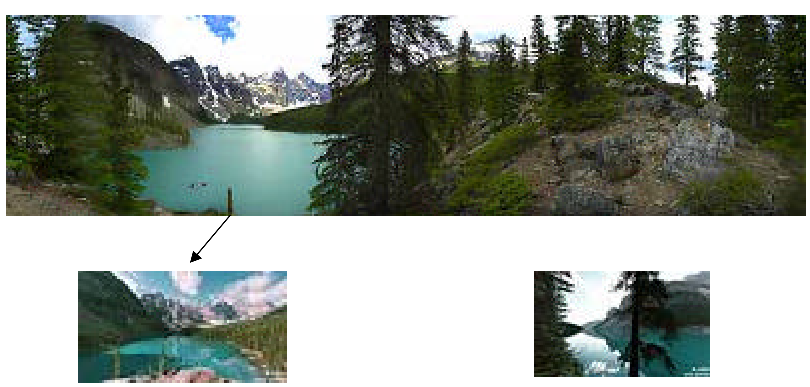
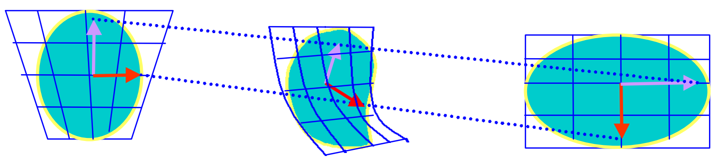


Image and View Morphig

Generate intermediate views by image/ view/ flow-field interpolation.



- Can produce geometrically incorrect images

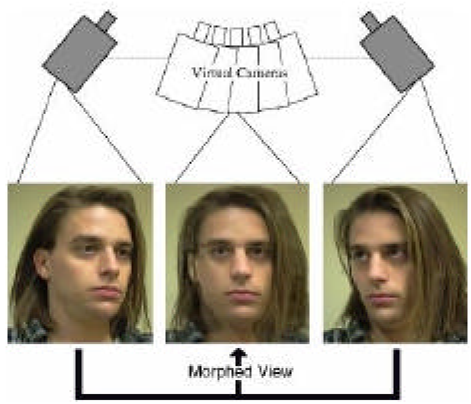
Image and View Morphing - *Examples*

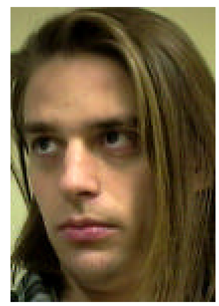
Beier &Neely – *"Feature-Based Image Metamorphosis"*

- Image processing technique used as an animation tool for metamorphosis from one image to another.
- Specify correspondence between source and destination using a set of line segments pairs

View Morphing along a line

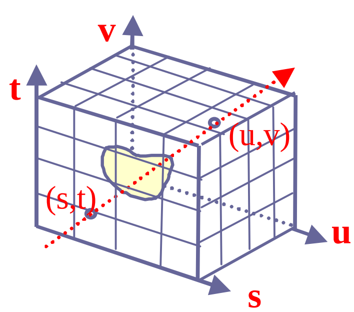
- Generate new views that represent a physically-correct transition between two reference images. (Seitz & Dyer)



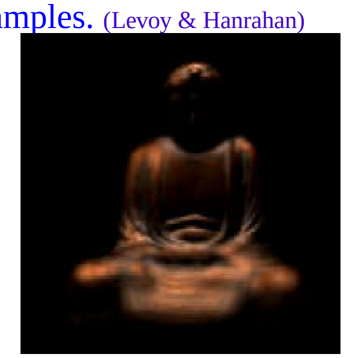


Light Field Rendering

Sample a 4D plenoptic function if the scene can be constrained to a bounding box



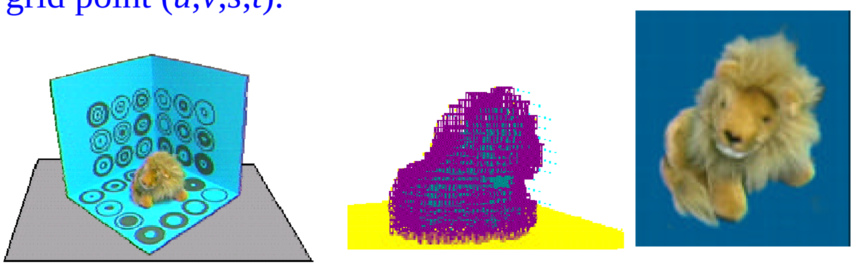
Approximate the resampling process by interpolating the 4D function from nearest samples. (Levoy & Hanrahan)



The Lumigraph

Gortler and *al*.; Microsoft

Lumigraph is reconstructed by a linear sum of the product between a basis function and the value at each grid point (*u*,*v*,*s*,*t*).

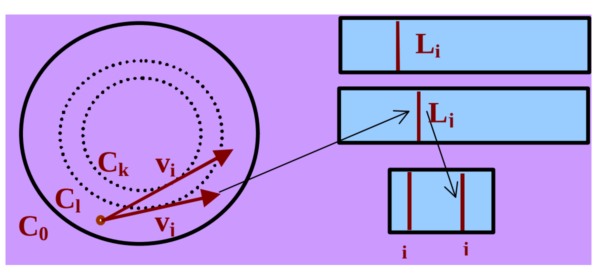


acquisition stage | volumetric model | novel view

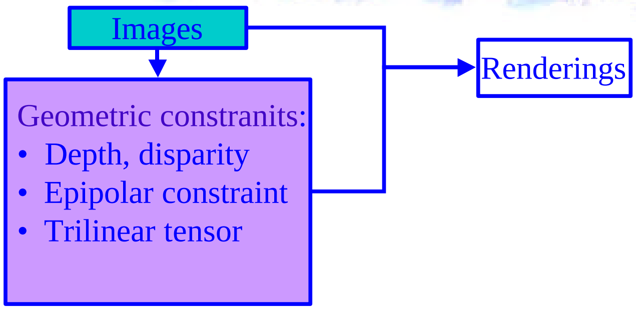
Concentric Mosaics

H-Y Shum, L-W He; Microsoft

Sample a 3D plenoptic function when camera motion is restricted to planar concentric circles.



Pixel Reprojection Using Scene Geometry



Laveau and Faugeras:

Use a collection of images (reference views) and the disparities between images to compute a novel view using a raytracing process.

Plenoptic Modeling

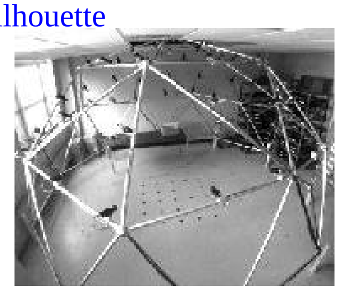
McMillan and Bishop:

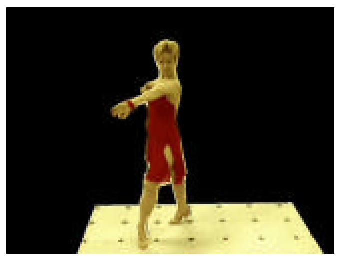
Plenoptic modeling (5D plenoptic function): compute new views from cylindrical panoramic images.

Virtualized Reality

T. Kanade -CMU

- 49 cameras for images and six uniformly spaced microphones for sound
- 3D reconstruction: volumetric method called Shape from Silhouette

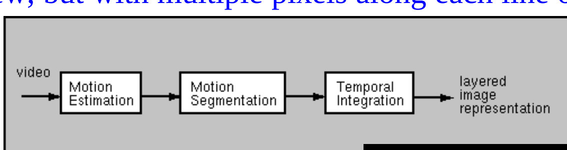




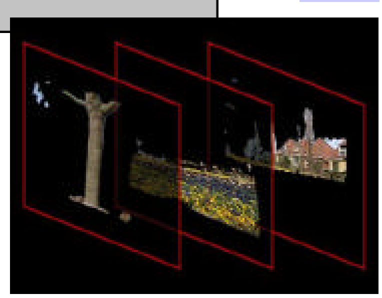
Layer Depth Images

Shade et. al.

LDI is a view of the scene from a single input camera view, but with multiple pixels along each line of sight.

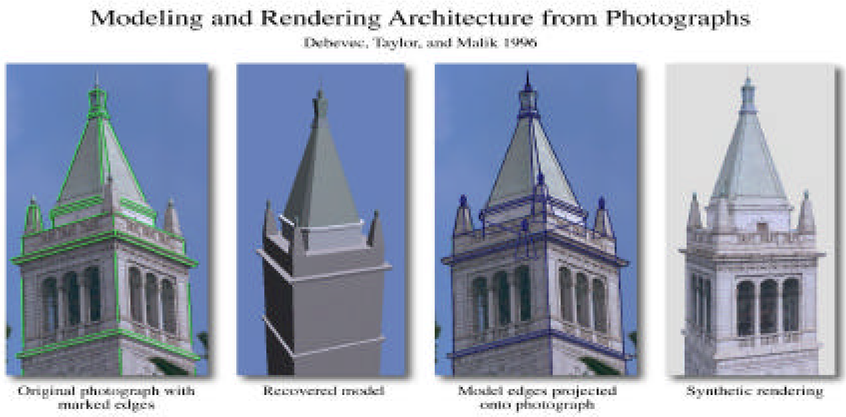


movie



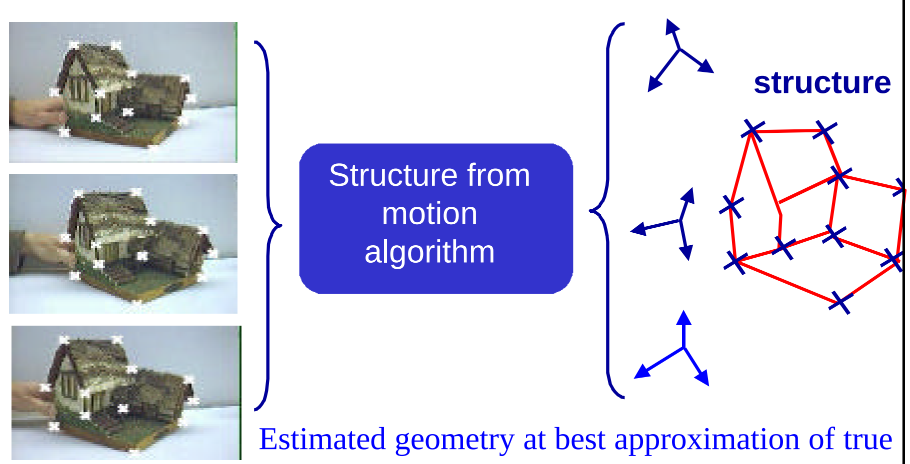
Rendering Architecture from Photographs

Combine both image-based and geometry based techniques. “Façade” (Debevec et. al.)



Structure from motion

Tracked features **poses**



Estimated geometry at best approximation of true

Geometric re-projection errors

Texturing:

static

dynamic

(Cobzas, Jagersand ECCV 2002)

Spatial Basis Intro

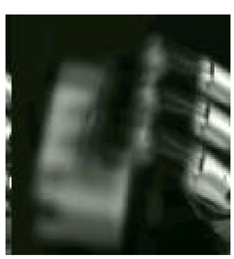
1. Moving sine wave can be modeled:

$$I = \sin(u + at) = \sin(u)\cos(at) + \cos(u)\sin(at) = \sin(u)y_1 + cos(u)y_2$$

2. Small image motion

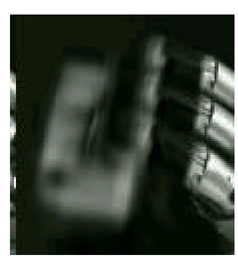
$$I = I_0 + \frac{\partial I}{\partial u}\Delta u + \frac{\partial I}{\partial v}\Delta v$$

Spatially fixed basis



2 basis vectors

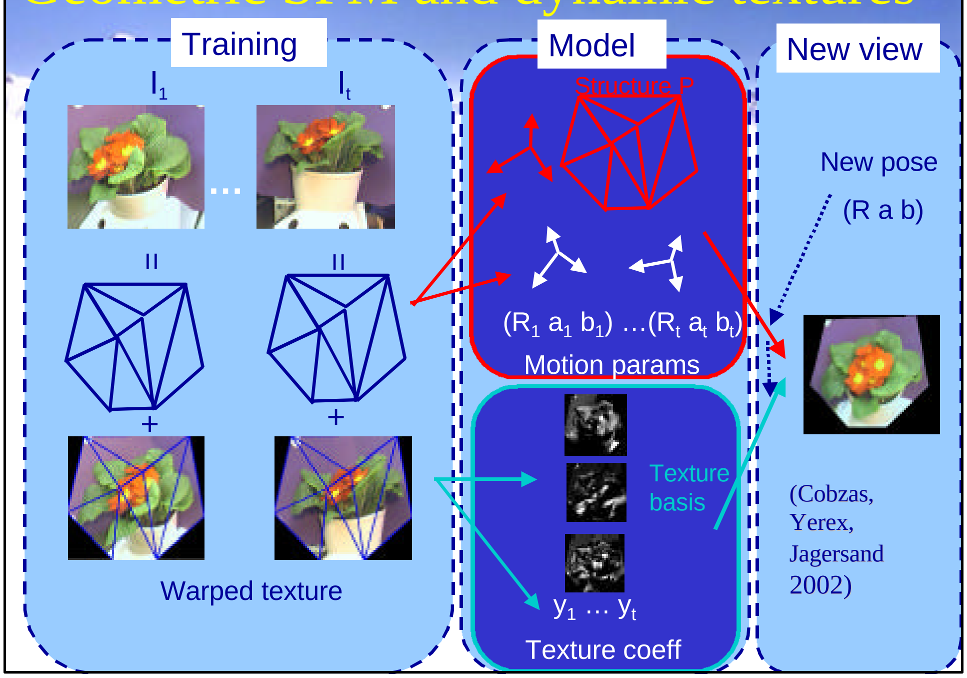
(Jagersand 1997)



6 basis vectors

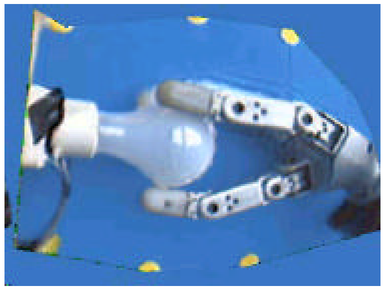
Example: Spatial basis for Light variation

Geometric SFM and dynamic textures



Geometric SFM and dynamic textures Example Renderings

- Rendering of models obtained using a $100 web cam and a home PC (Cobzas, Yerex, Jagersand 2002)



We'll learn how to do this in the lab this afternoon…

Summary - IBMR

	Technique	Input data	Rendered images	+/-
Image and view morphing	Interpolation	2 images	Interpolate the reference images	+ easy to generate images - nonrealistic
Interpolation from dense samples	4D plenoptic function of a constrained scene	Samples of the plenoptic function	Interpolate the 4D function	+ easy to generate renderings - Need exact cam. Cal. - mostly synthetic scenes - large amount of data
Geometrically valid pixel reprojection	Use geometric constraints	2,3, more images taken from the same scene	Pixel reprojection	+ low amount o data + geometrically correct renderings - requires depth/ disparity
Geometric SFM + Dynamic texture	Obtain coarse geometry from images	Many (100+) images from the same scene	Geometric projection and texture mapping	+ geometrically correct renderings + integrates with standard computer graphics scenes -large amount of data.

IEEE Virtual Reality 2003
Next Lectures

1. Single view geometry and camera calibration
2. Plenoptic function and light field rendering
3. Multiple view projective, affine and Eucl. geometry
4. Scene and object modeling from images
5. Real-time visual tracking and video processing
6. Differential image variability and dynamic textures
7. Hard-ware accelerated image-based rendering
8. Software system and hands-on lab