Image Composition and Relighting

D.A. Forsyth, UIUC

Inserting things into pictures

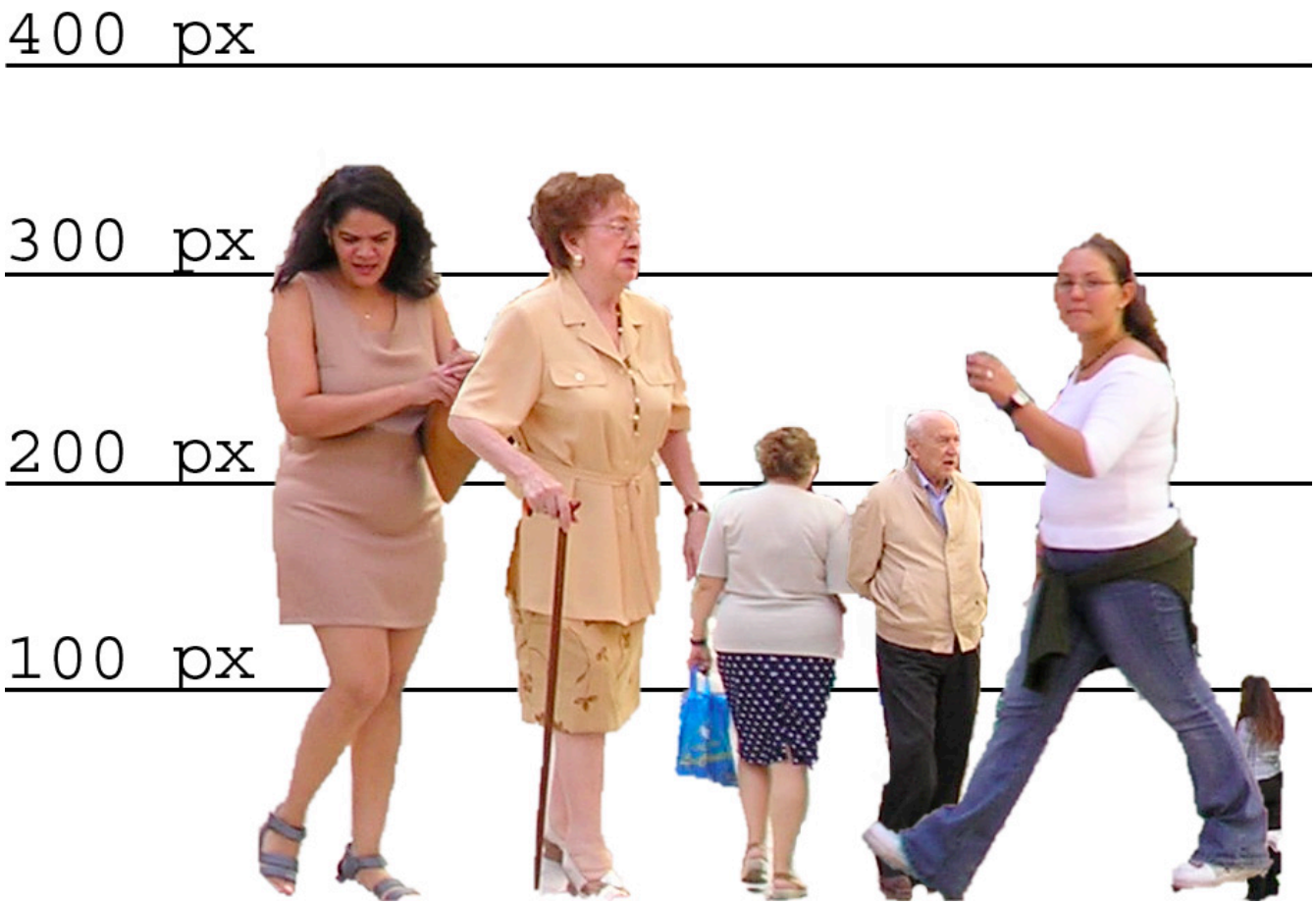
- Insert geometric models or animations
  - Applications:
    - sales, marketing
    - motion pictures
    - commercial art
- Insert image fragments
  - Applications:
    - commercial art

Inserting fragments

- Algorithm
  - build a dictionary of image fragments, ideally tagged
    - for these fragments, estimate height using ground plane
  - artist chooses image
    - system estimates horizon, ground plane
    - this gives foreshortening
  - artist searches with tag, chooses fragment
  - places on image
- Q: what if the light is wrong?
  - A: don’t use that fragment

Height issues

Height issues



Height issues

Inserting fragments

Inserting fragments

Lalonde+Efros, 07

Inserting fragments

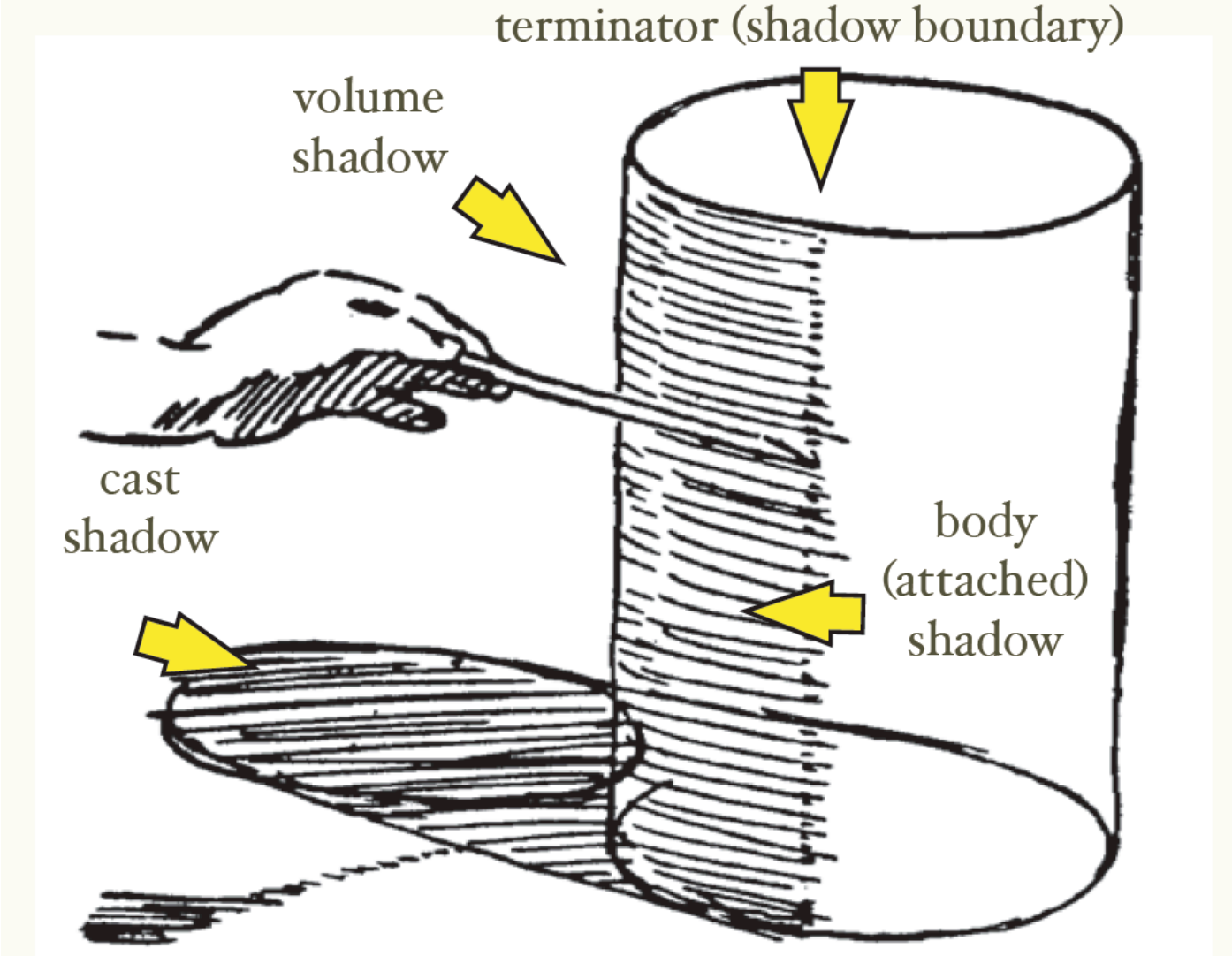
Inserting fragments

Illumination issues

Illumination issues: good match

Illumination issues: bad match

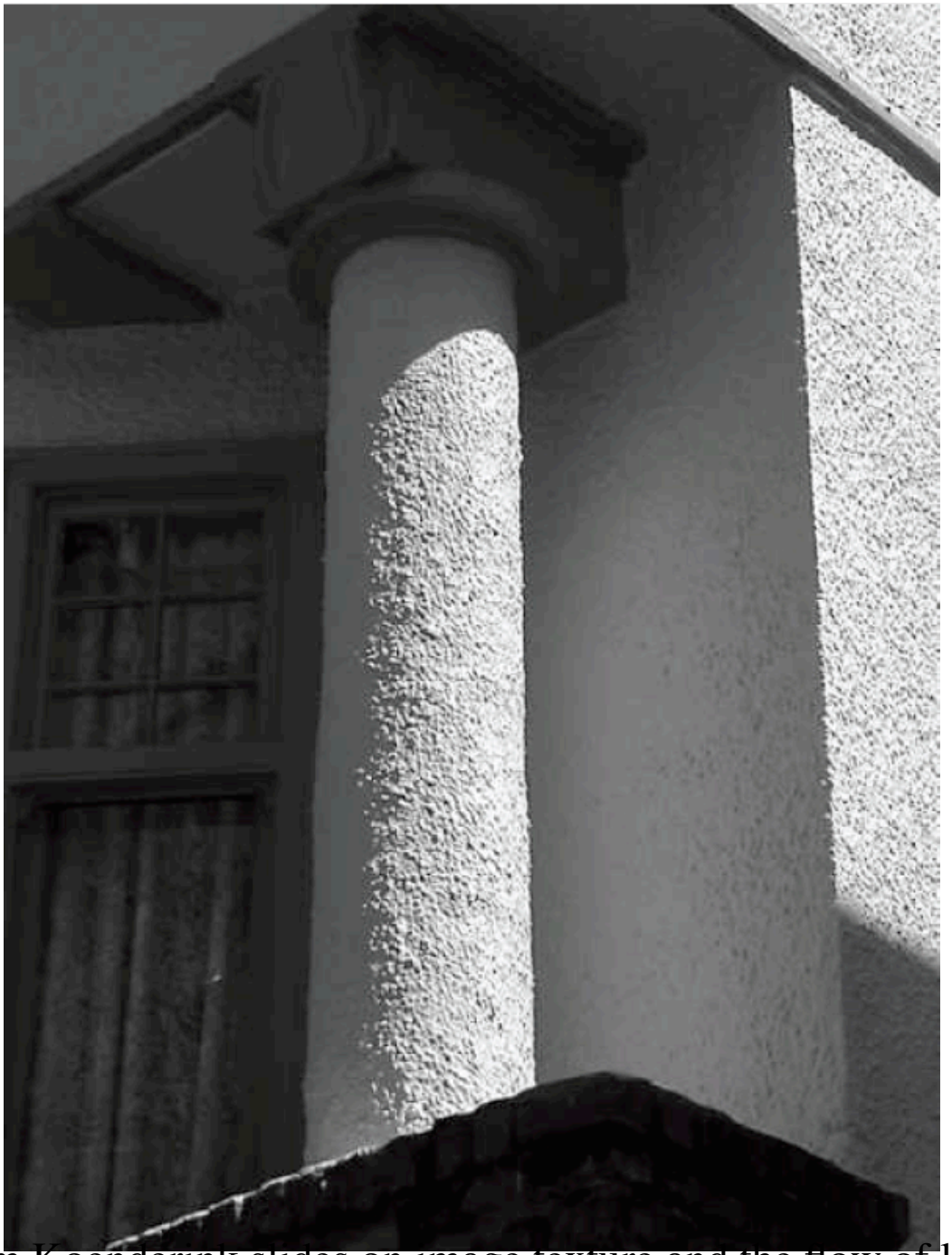
Cues to shape - shadows



From Koenderink slides on image texture and the flow of light

Shadow
Bright

From Koenderink slides on image texture and the flow of light



From Koenderink slides on image texture and the flow of light

Shadow geometry can be very nasty

From Hel Des, on Flickr

From Koenderink slides on image texture and the flow of light

Shape from shading

- Recover a shape representation from the shading field
  - people seem to be able to do it
  - Qn's:
    - what shape representation?
    - how?
  - there is a story in computer vision, but we know it's wrong


Original with Tracking
Retextured Video
Computer Vision
Research Task Force

By Technicolour Yawp, on Flickr

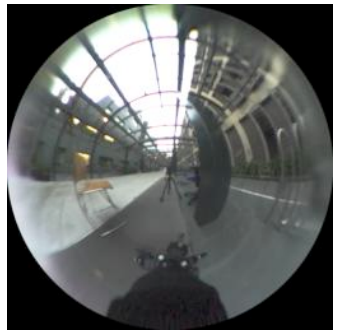
From Koenderink slides on image texture and the flow of light

Open issues

- Which I haven't discussed (yet)...
  - how one extracts fragments really accurately (matteing)
  - how one mattes fragments into the image to get good blends
    - eg hair

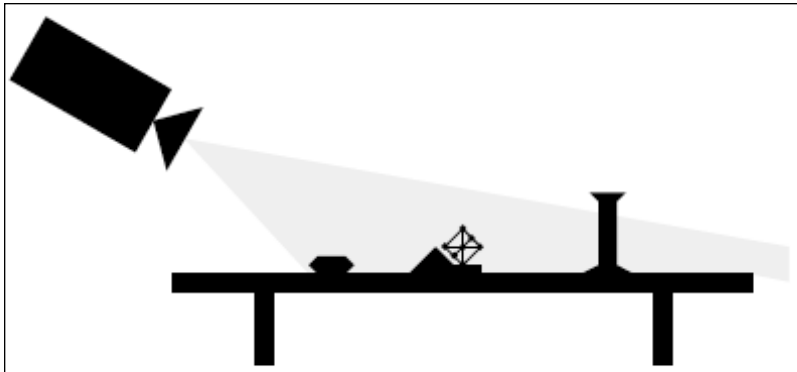
(a) Input image | (b) GrabCut [Rother et al '04] | (c) Poisson Editing [Perez et al '03] | (d) Blending Mask - Our Method | (e) Our Method | (f) Digital Photomontage [Agarwala et al '04] | (g) Drag-and-Drop Pasting [Jia et al '06]

If you have access to the scene...

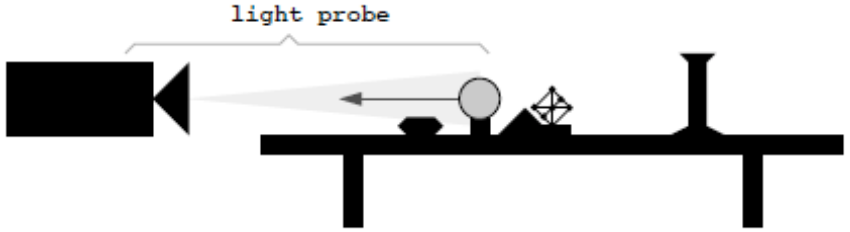


- Use a light probe
  - shiny metal ball - yields measurements of radiance in directions
- Procedure:
  - take a photograph of scene from direction you will use
  - put in light probe, calibration object, take a high dynamic range picture
  - build a local geometric model for scene objects near to new objects
  - recover material properties for those objects using light probe
  - light that model with light probe, compute global illumination
  - composite the lit model with the image

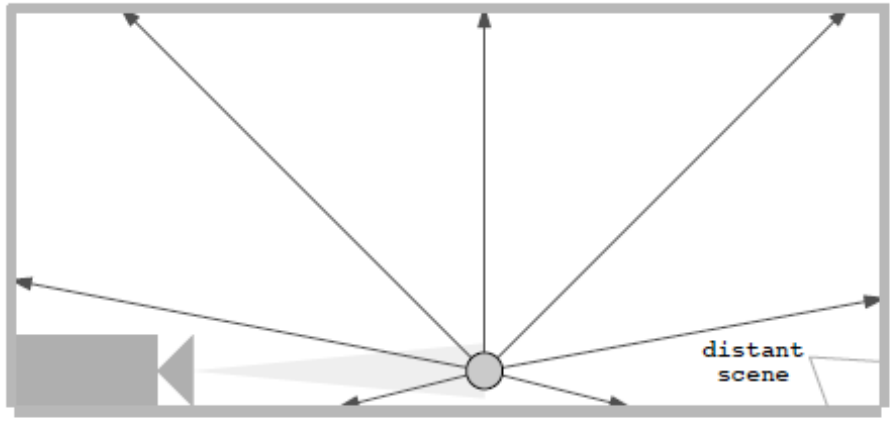
Debevec, 1998



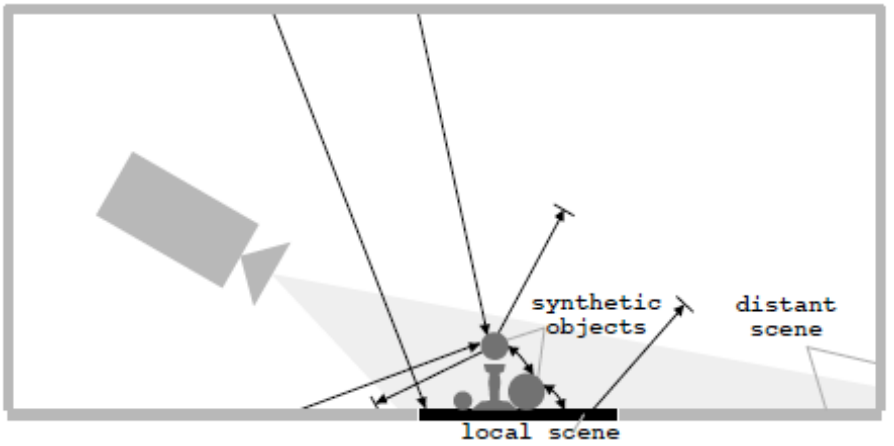
(a) Acquiring the background photograph



(b) Using the light probe



(c) Constructing the light-based model



(d) Computing the global illumination solution

Debevec, 1998

Light probe steps

(a) Background photograph

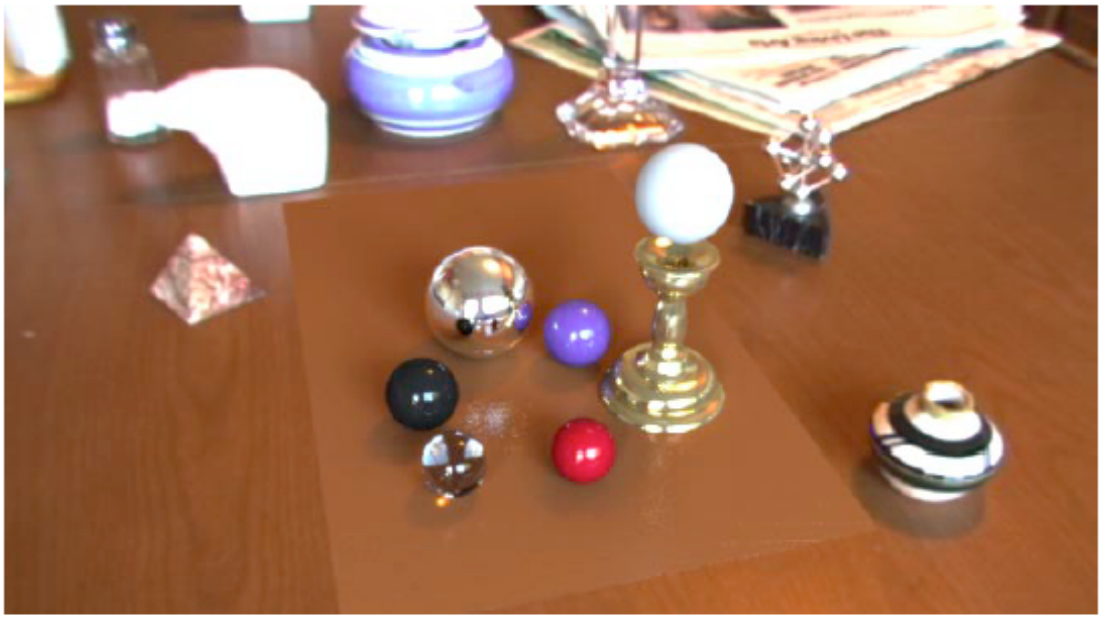
Debevec, 1998

Light probe steps

(b) Camera calibration grid and light probe

Debevec, 1998

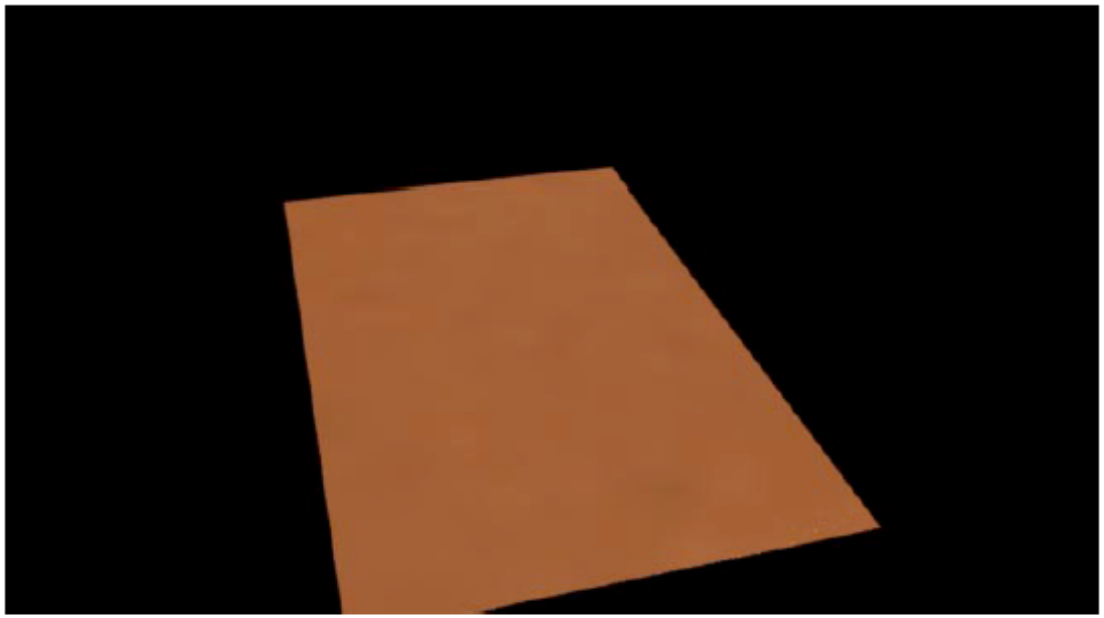
Light probe steps



(**c**) Objects and local scene matched to background

Debevec, 1998

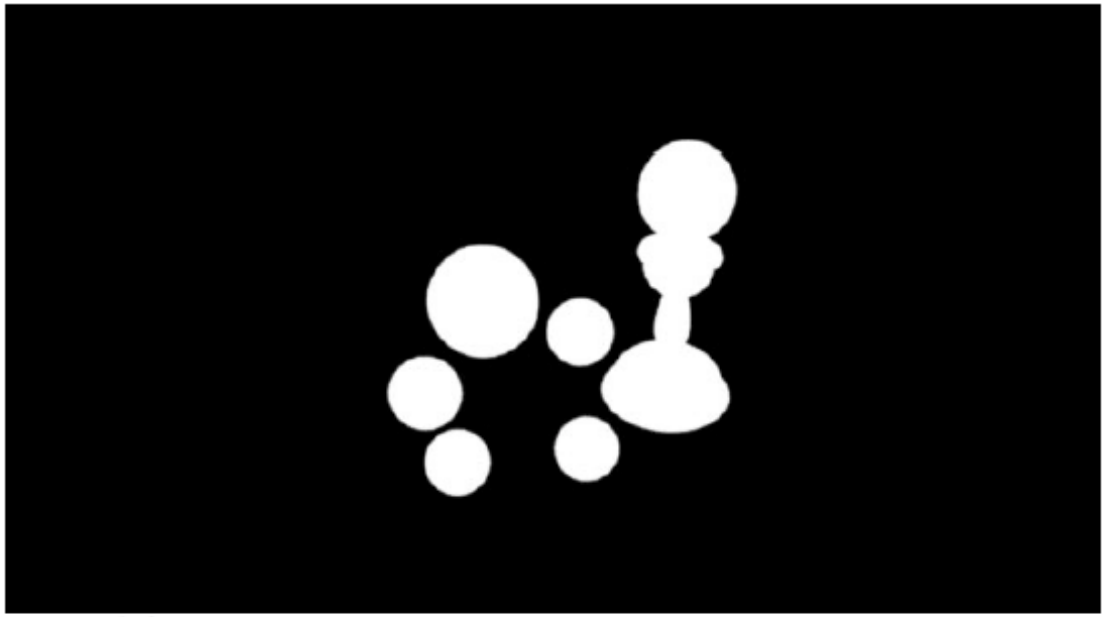
Light probe steps



(d) Local scene, without objects, lit by the model

Debevec, 1998

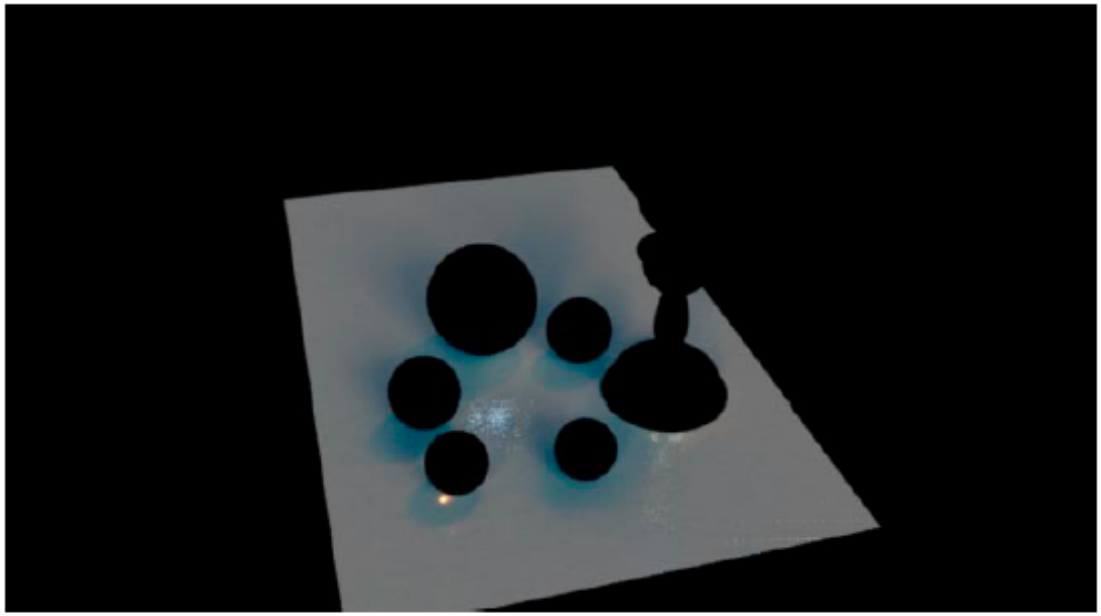
Light probe steps



(e) Object matte

Debevec, 1998

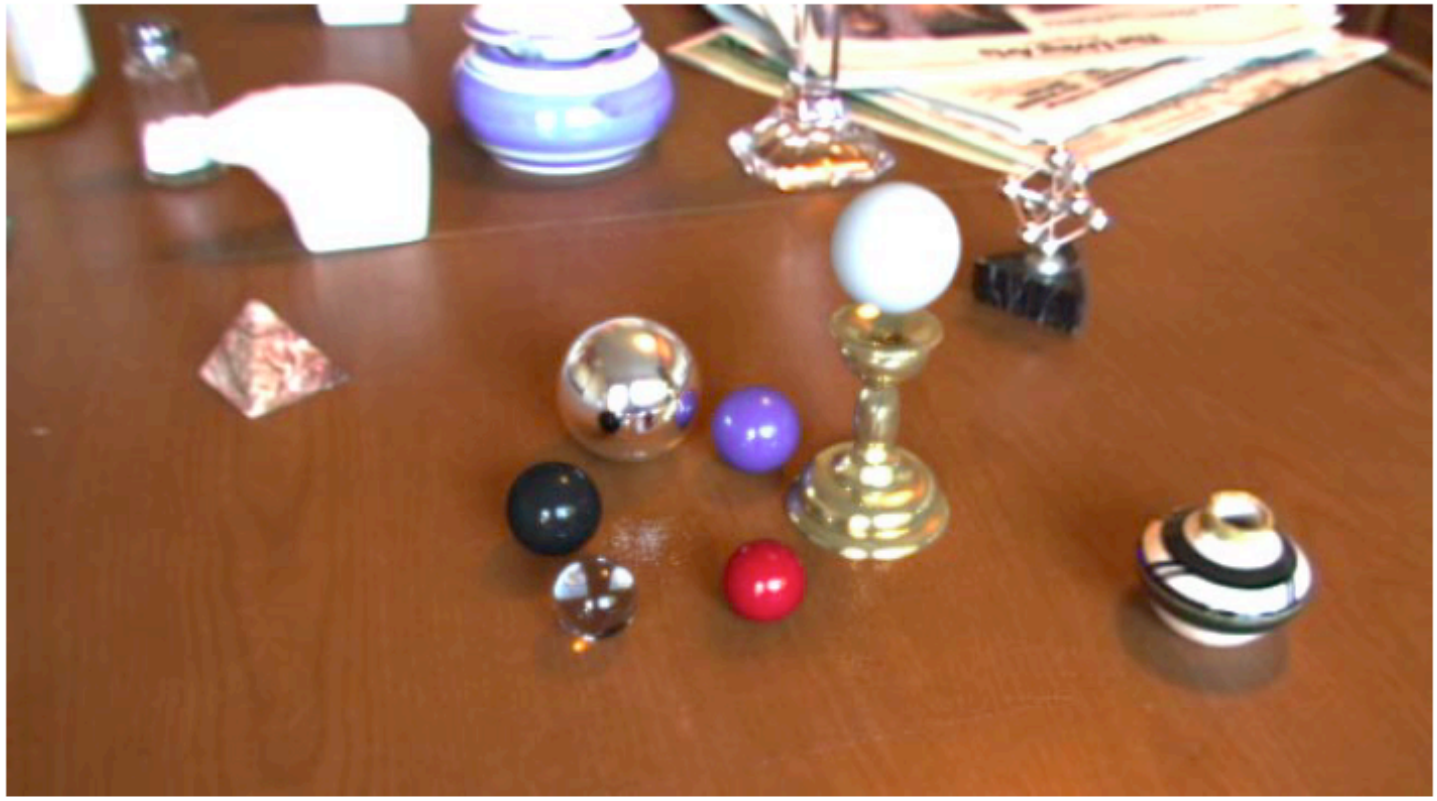
Light probe steps



(f) Difference in local scene between **c** and **d**

Debevec, 1998

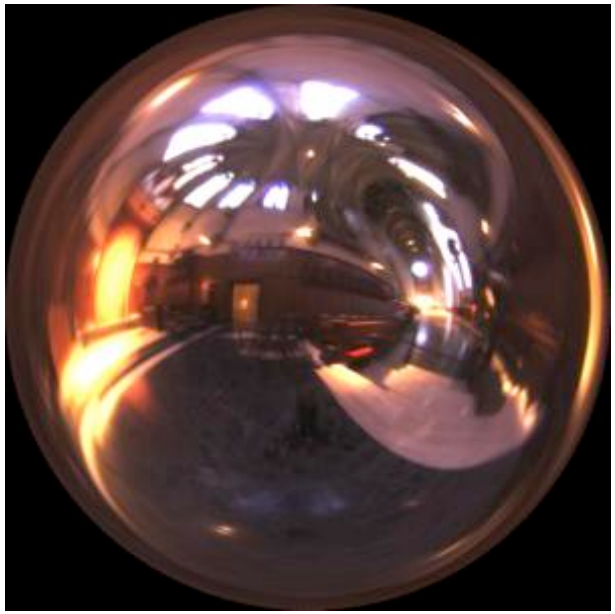
Light probe steps



(g) Final result with differential rendering

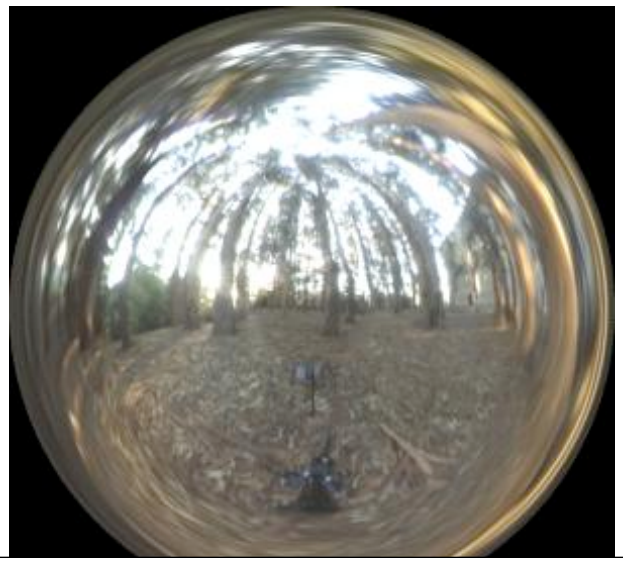
Debevec, 1998

Example light probe images



Grace cathedral

Images from
http://www.pauldebevec.com/Probes/



Eucalyptus grove, UC Berkeley

Example light probe images

Galileo's tomb

Images from
http://www.pauldebevec.com/Probes/

Skull on tomb, rendered with tomb light probe

Light probe

- This is the standard method in motion pictures
  - lots of experience using it
  - works really well
- Issues
  - It's a performance - getting good results takes a lot of work
  - You must have access, to stick in the probe
    - probe should be photographed from multiple directions
  - You might need quite a lot of probes, particularly for moving objects
  - Probe image must be high dynamic range

If you don’t have access....

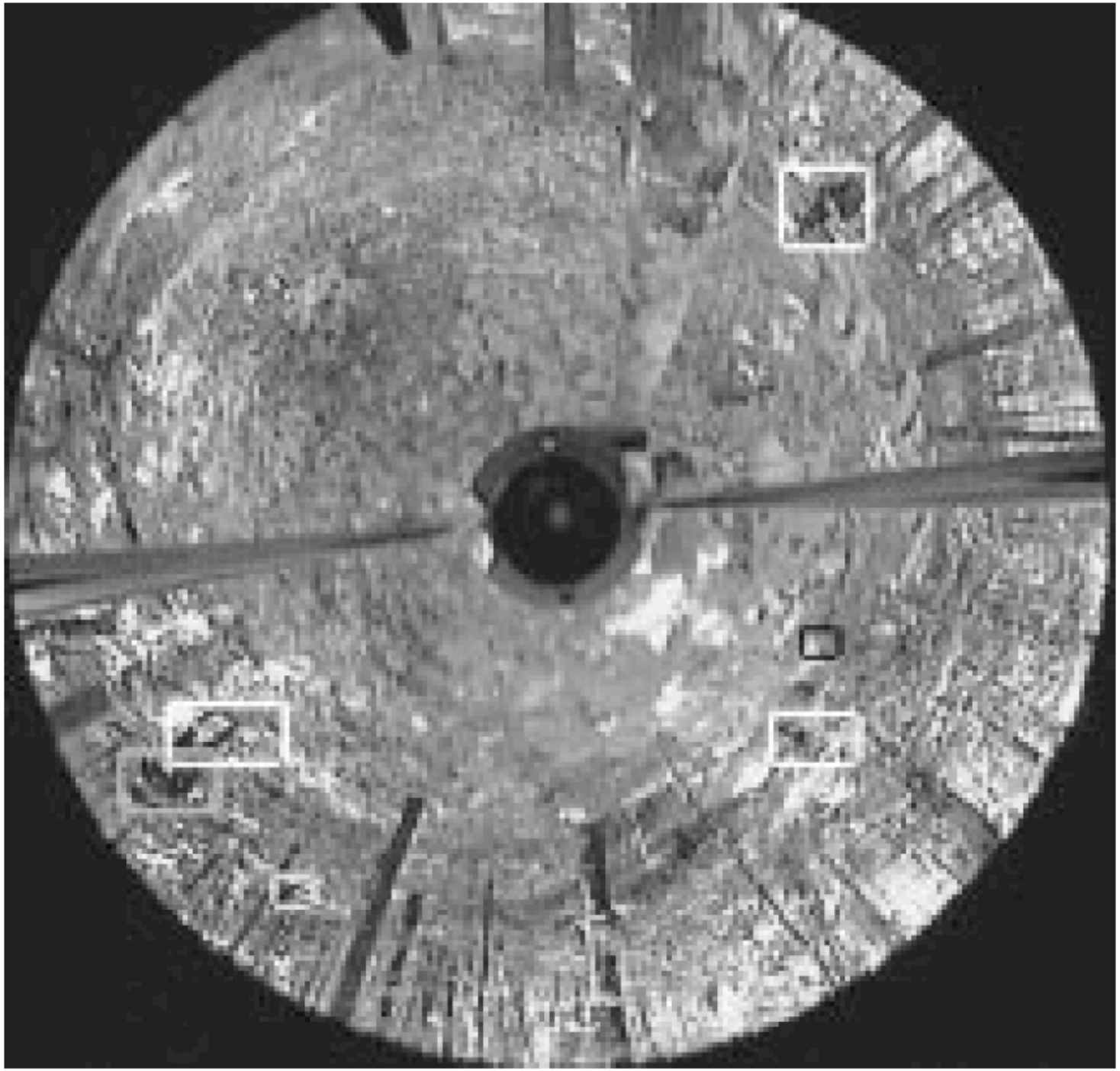
- Strategy
  - build a model of scene
    - geometry
    - materials
    - light
  - render using that model
  - composite

Conclusions

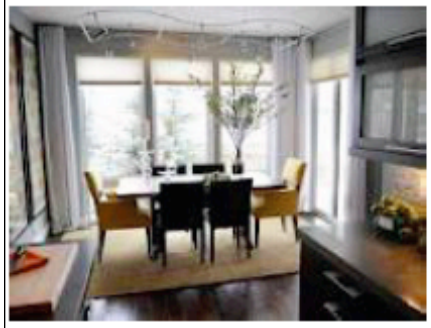
- Rooms are important
- Rough 3D representations help recognition
- Free space has motion "potential"
  - it tells you where you can move
  - and we can recover descriptions of free space that are good for motion
- Free space has light "potential"
  - light lives there
  - and we can recover descriptions of free space that are good for rendering
- Little 3D scene information -> very neat applications

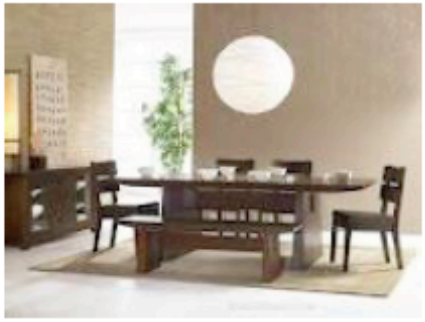
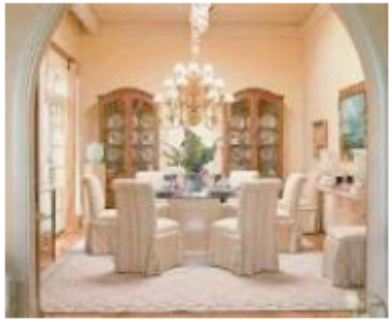
Rooms are important

- Because people live in them
  - and different behaviors occur in different rooms
- Because they admit useful geometric approximations
  - boxes; or others
- Because other spaces are often like rooms
  - examples later



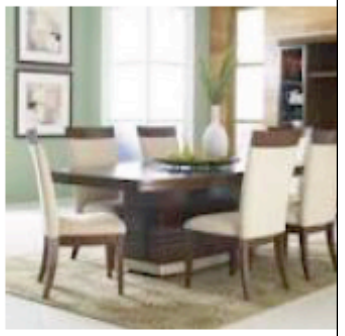
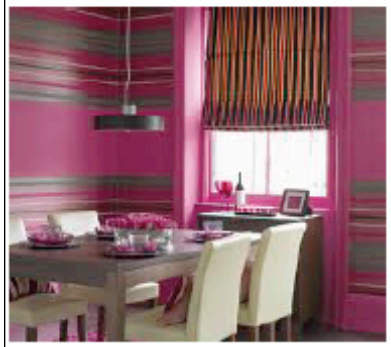
Where you are can suggest
you are doing something
you shouldn't be
Boult 2001





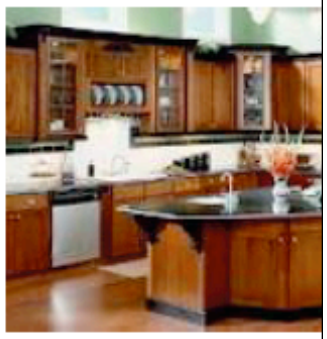


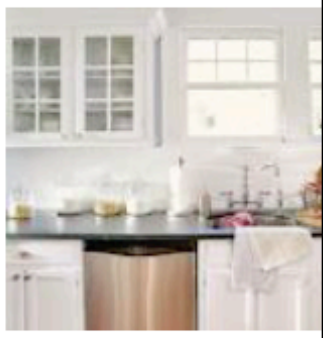


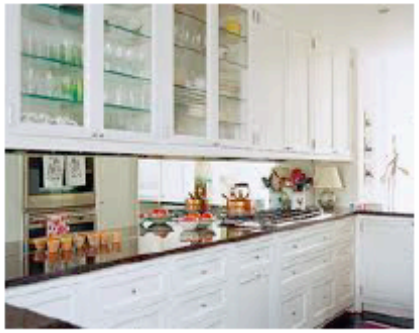
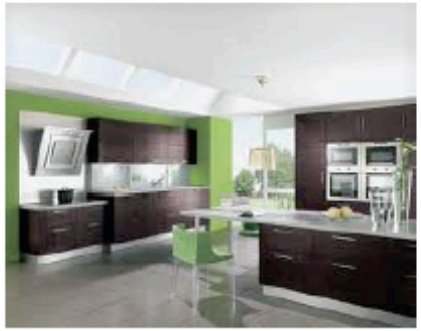
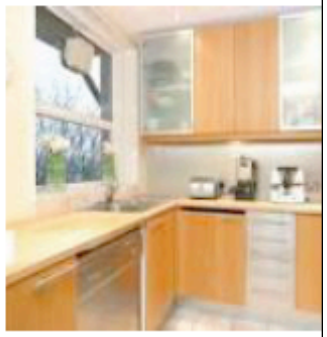




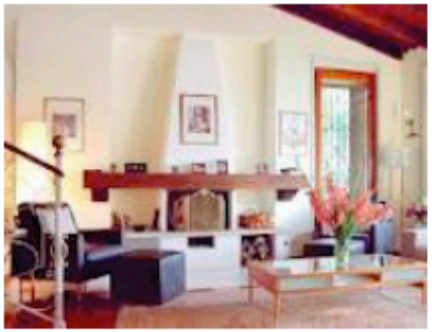
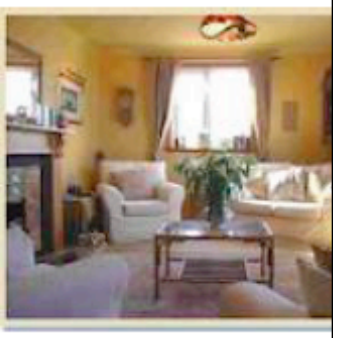


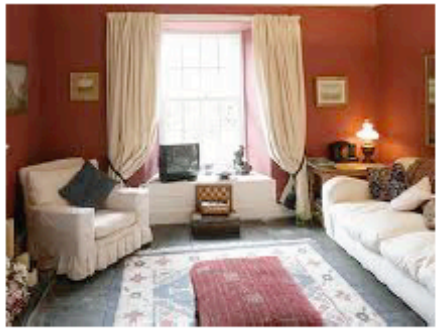
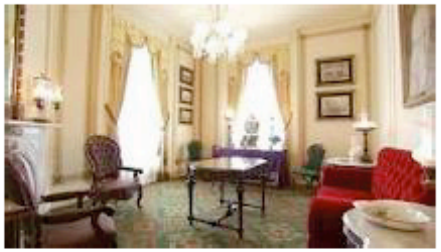
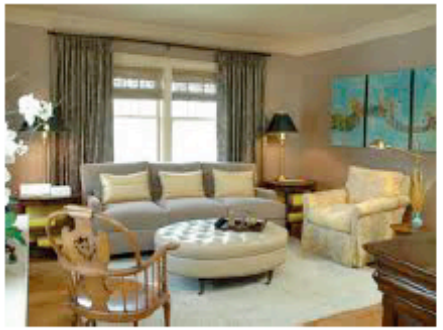
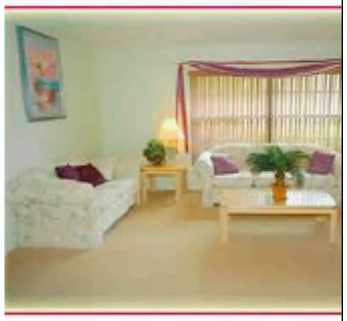
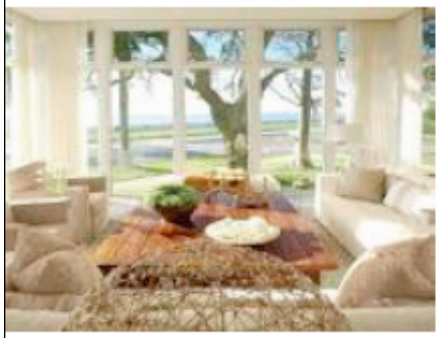


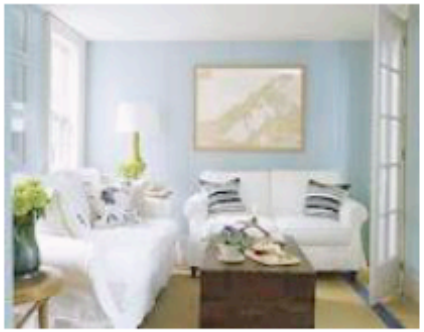



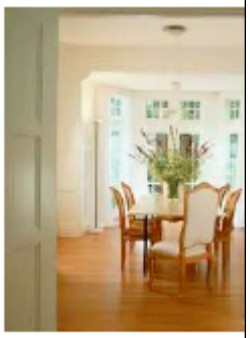







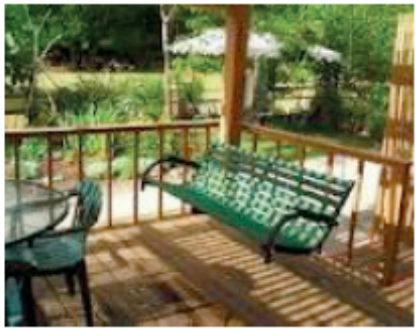








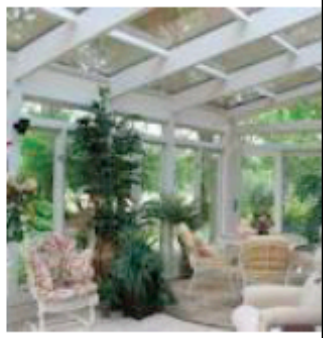












Conclusions

- Rooms are important
- Rough 3D representations help recognition
- Free space has motion "potential"
  - it tells you where you can move
  - and we can recover descriptions of free space that are good for motion
- Free space has light "potential"
  - light lives there
  - and we can recover descriptions of free space that are good for rendering
- Little 3D scene information -> very neat applications

Rooms as boxes

- Assume: rooms are boxes
  - camera is orthographic, known aspect ratio
  - camera is inside the box
  - aspect ratio of box is unknown
- Estimate
  - rotation of box --- by vanishing points
  - location of corners (= translation, aspect ratio) --- search
- Reestimate
  - remove clutter --- doesn't look like faces
  - now estimate again

Approximate estimation of 3D repns, Hedau et al, '09; see also: Barinova et al, '08; Delage et al '06; Hoiem et al '06; Lee et al, '09, '10; Nedovic et al '07; Saxena et al '06;

Vanishing points

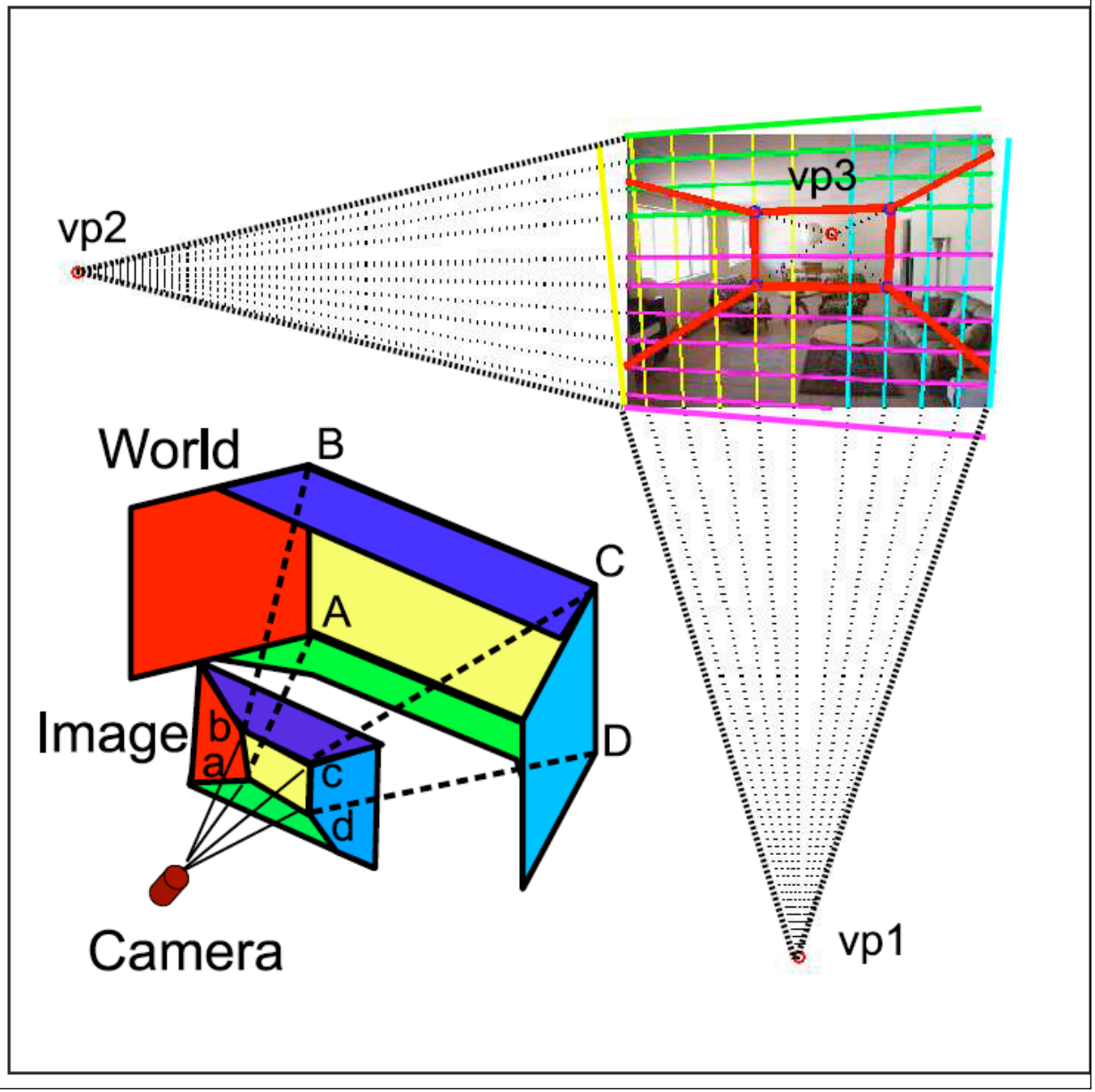
- Parallel lines in the world intersect in a vanishing point
  - so if we have a box, there are three important vanishing points
    - could get automatically, or ask user to mark
- Once we know these, we need to know 4 more DOF
  - user could mark these too, or there are automatic methods

Vanishing points

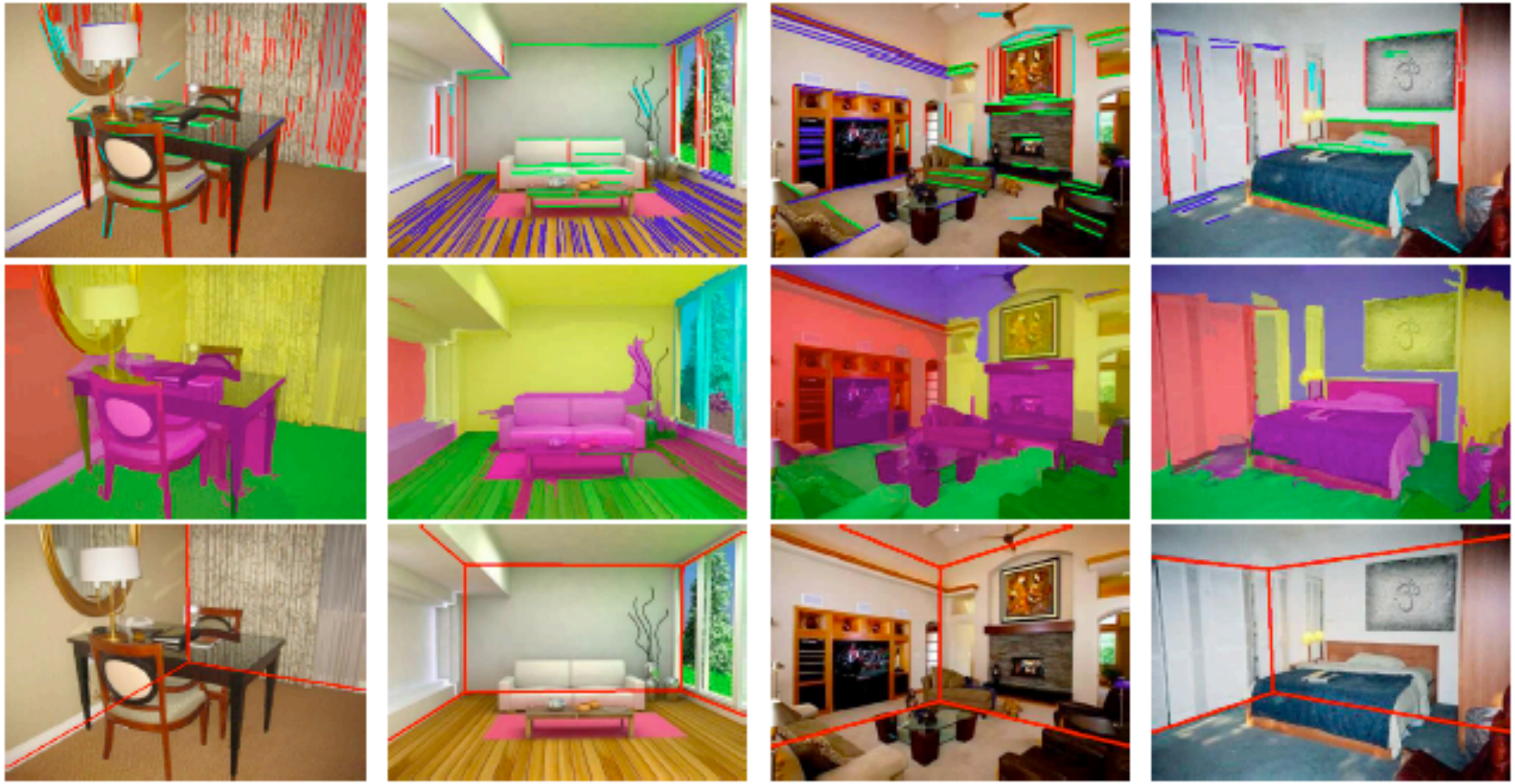
- Cluster long straight edges into three clusters (after Rother, 02)

Estimating layout

- Choice of layout= 4DOF in image
- Search cost function
  - learned from examples



Clutter maps



Conclusions

- Rooms are important
- Rough 3D representations help recognition
- Free space has motion "potential"
  - it tells you where you can move
  - and we can recover descriptions of free space that are good for motion
- Free space has light "potential"
  - light lives there
  - and we can recover descriptions of free space that are good for rendering
- Little 3D scene information -> very neat applications

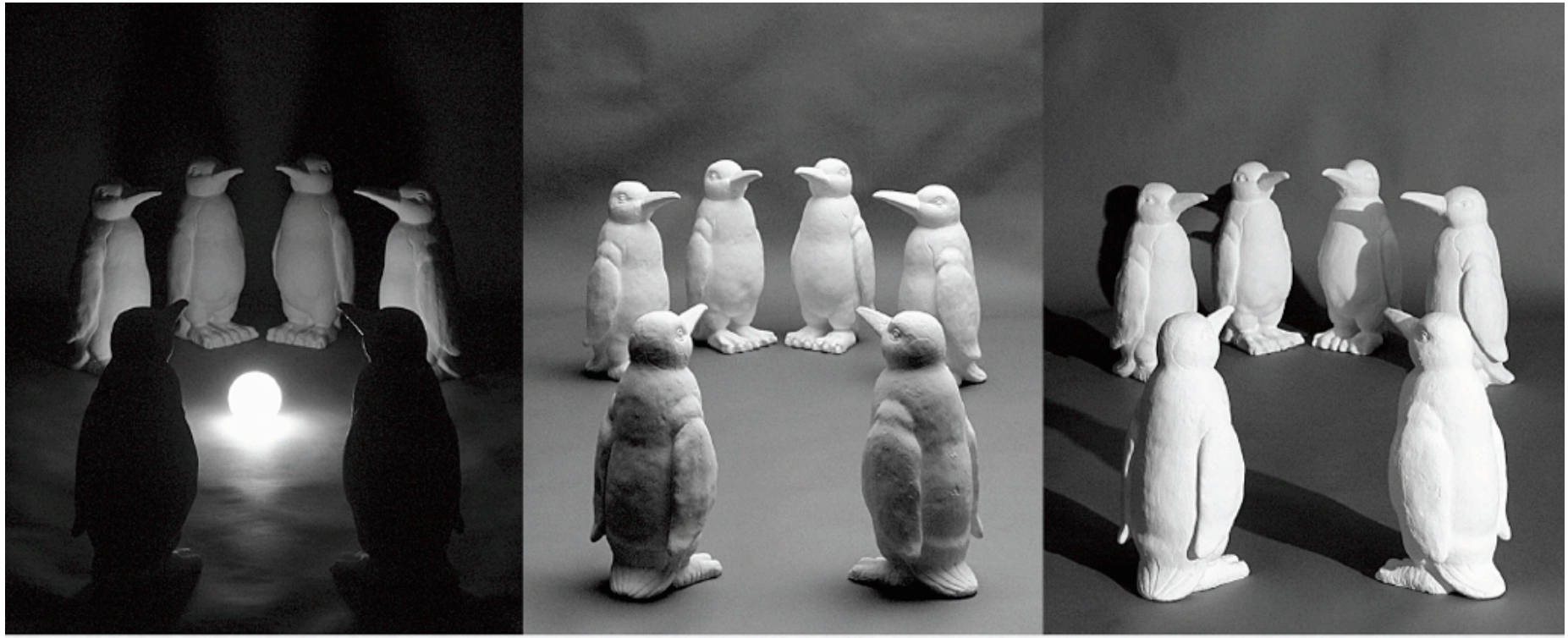
Stage lighting

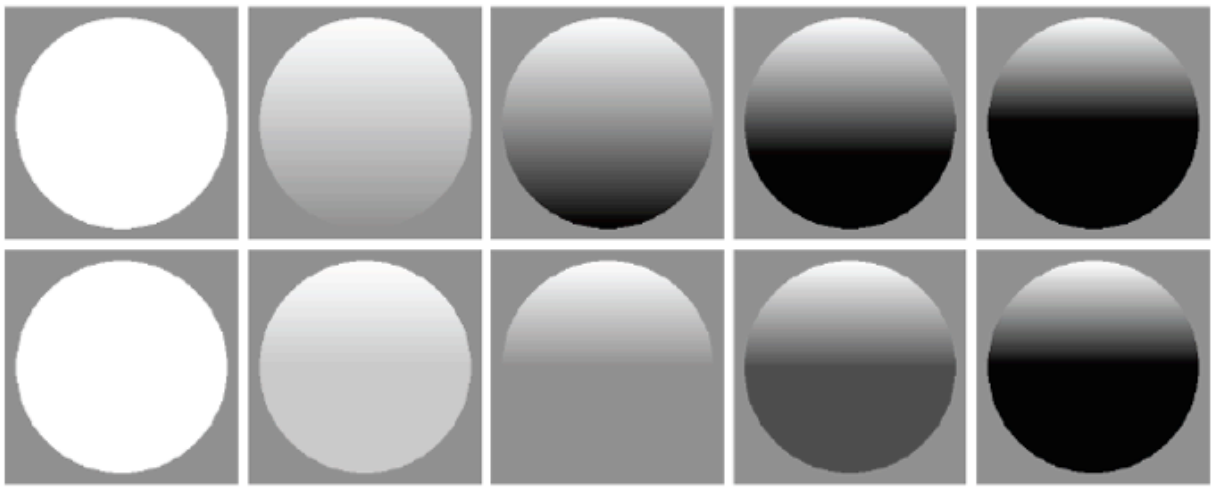
From Koenderink slides on image texture and the flow of light

Stage lighting is hard

Adolph von Menzel:
Das Balkonzimmer, 1845

From Koenderink slides on image texture and the flow of light





scenes ↑

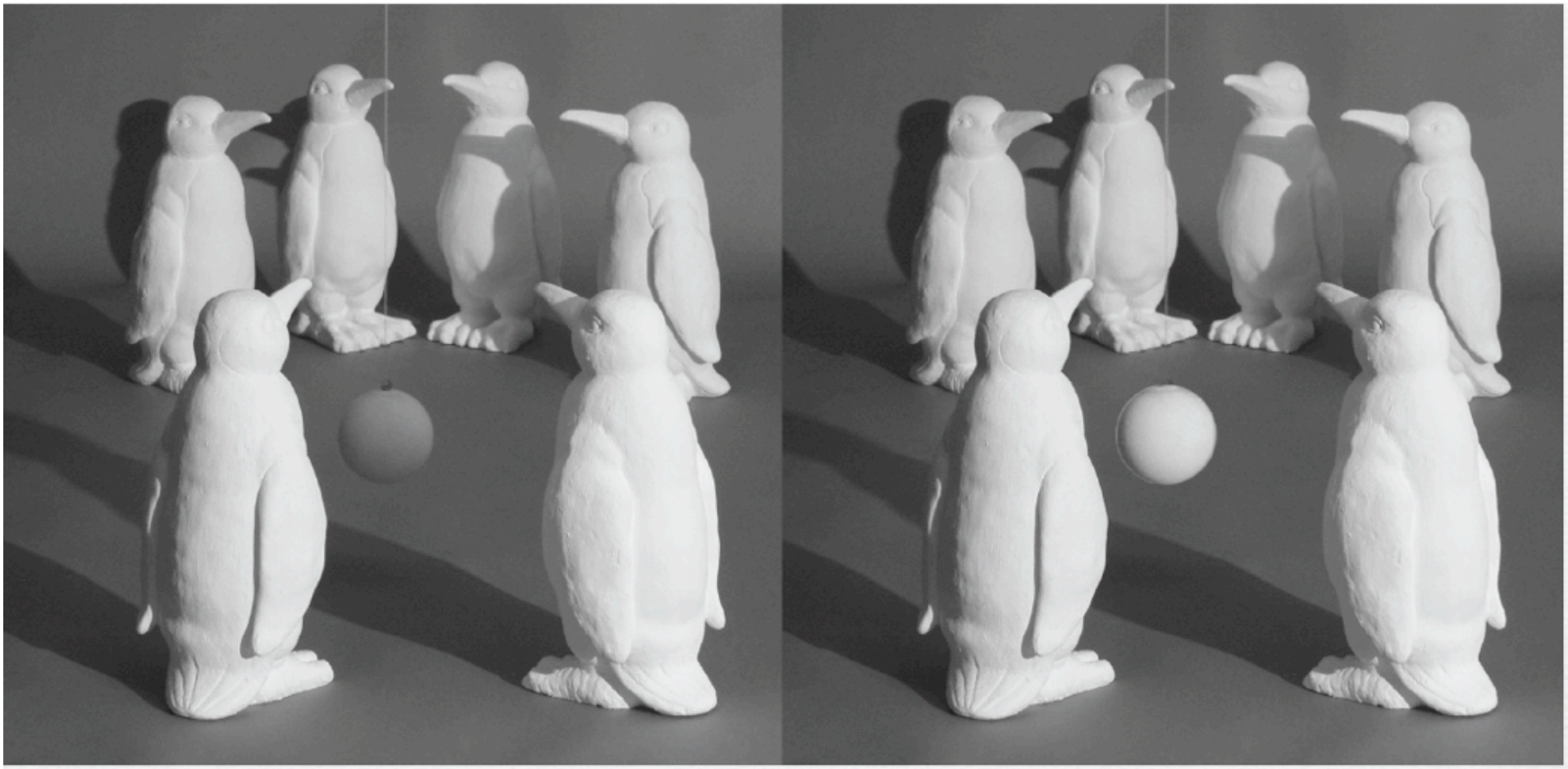
Psychophysics of "light fields".

← *probes*

22

From Koenderink slides on image texture and the flow of light

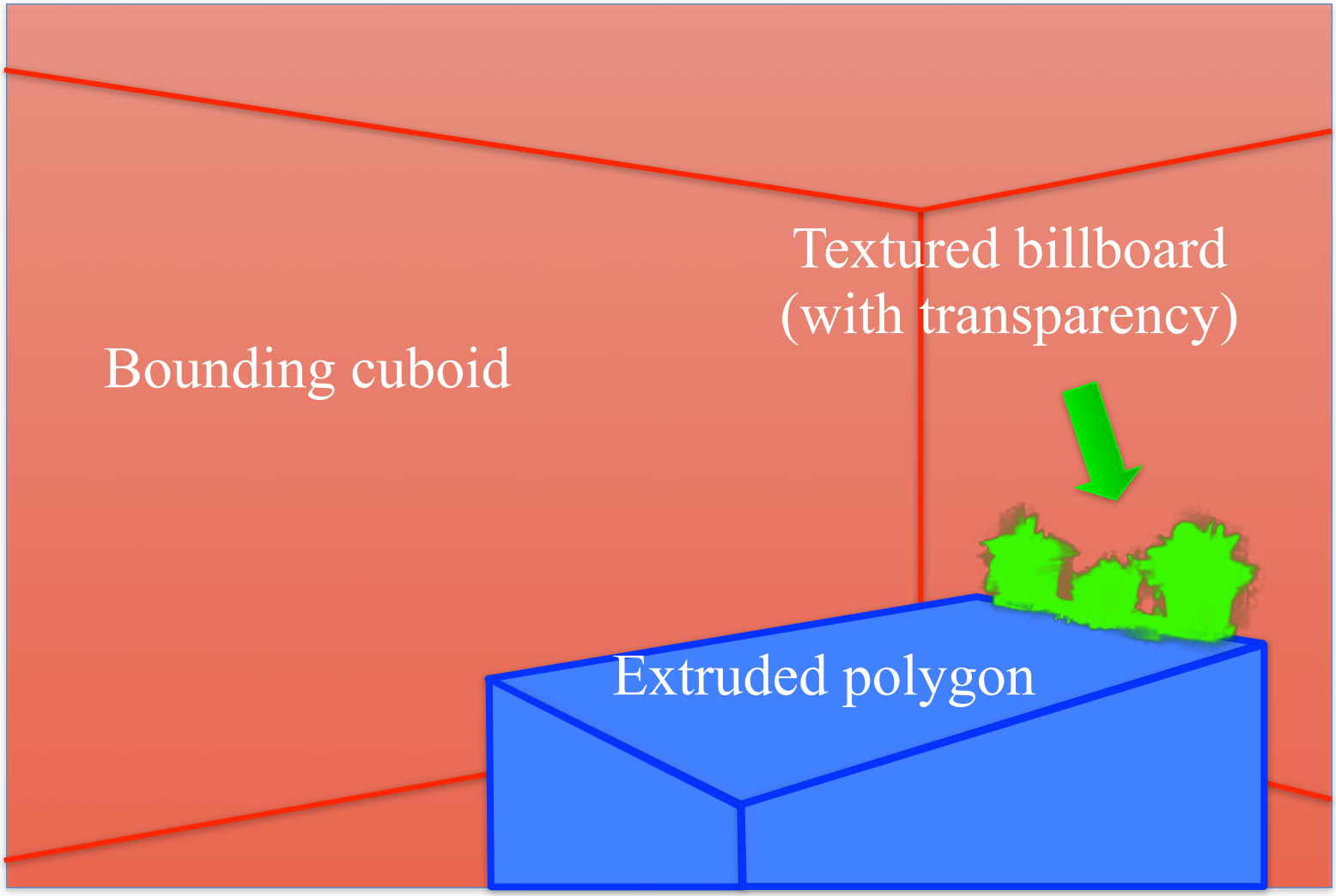
Koenderink et al 07

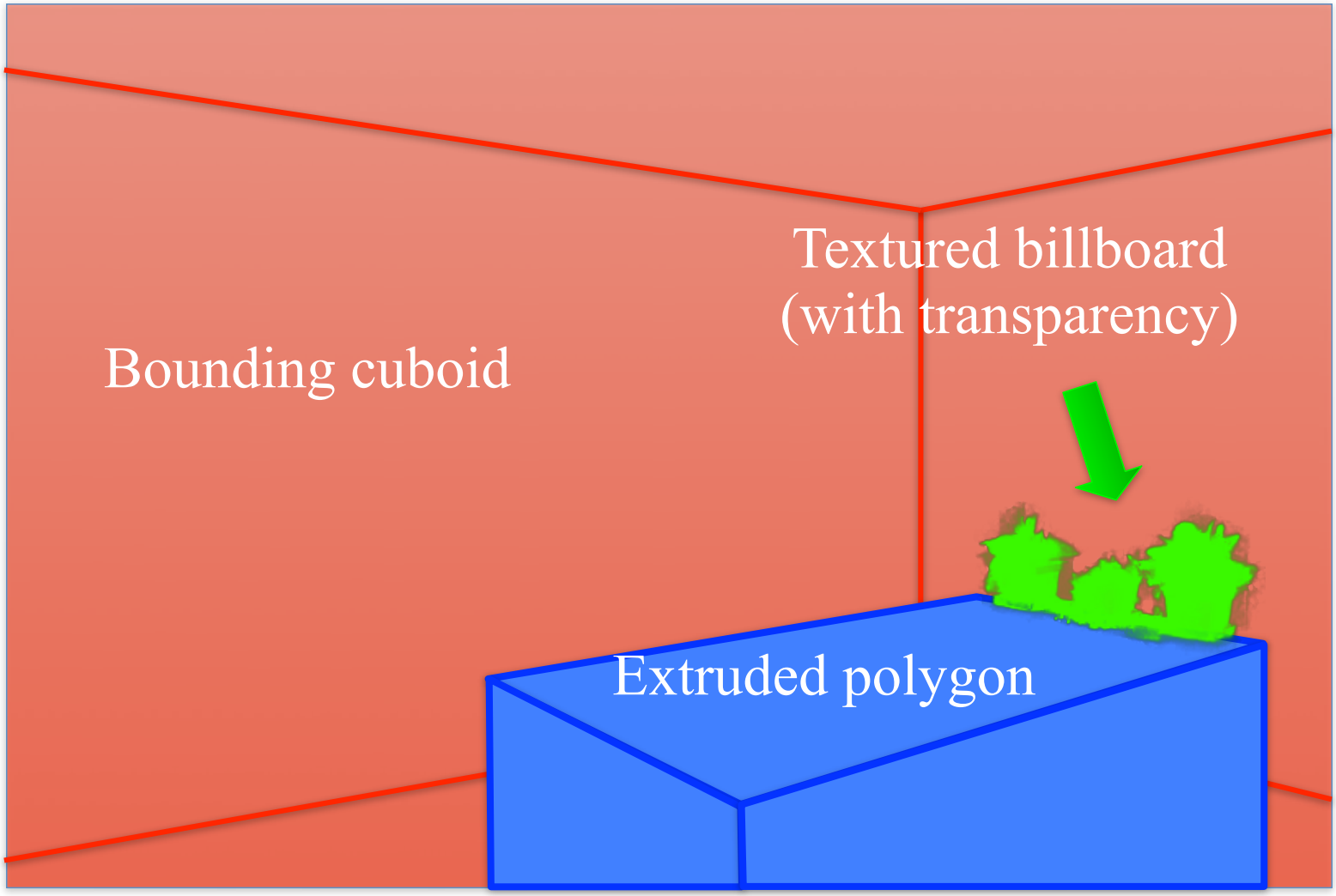


Human observers turn out to be remarkably sensitive to the light field, both to direction and diffuseness.

One exception: all observers "missed" the effect of volume shadow (ground truth – *left*) and produced a non-physical setting – *right*. Cast shadow volumes are ignored.

From Koenderink slides on image texture and the flow of light Koenderink et al 07

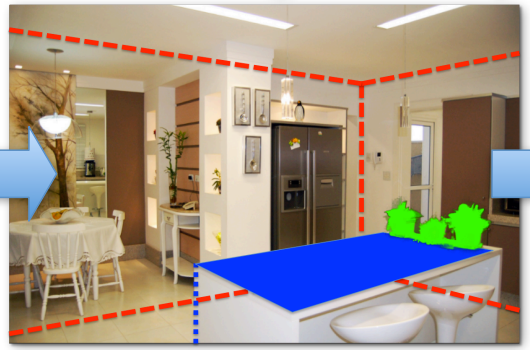

Textured billboard
(with transparency)
Bounding cuboid
Extruded polygon


Textured billboard
(with transparency)
Bounding cuboid
Extruded polygon

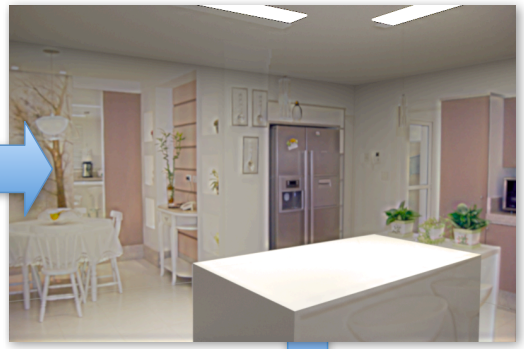
System Overview

Input image

Estimate geometry



Estimate materials



Estimate lighting

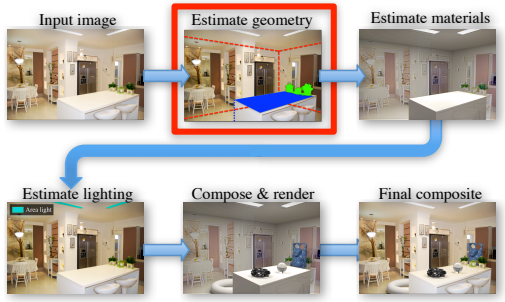
Compose & render

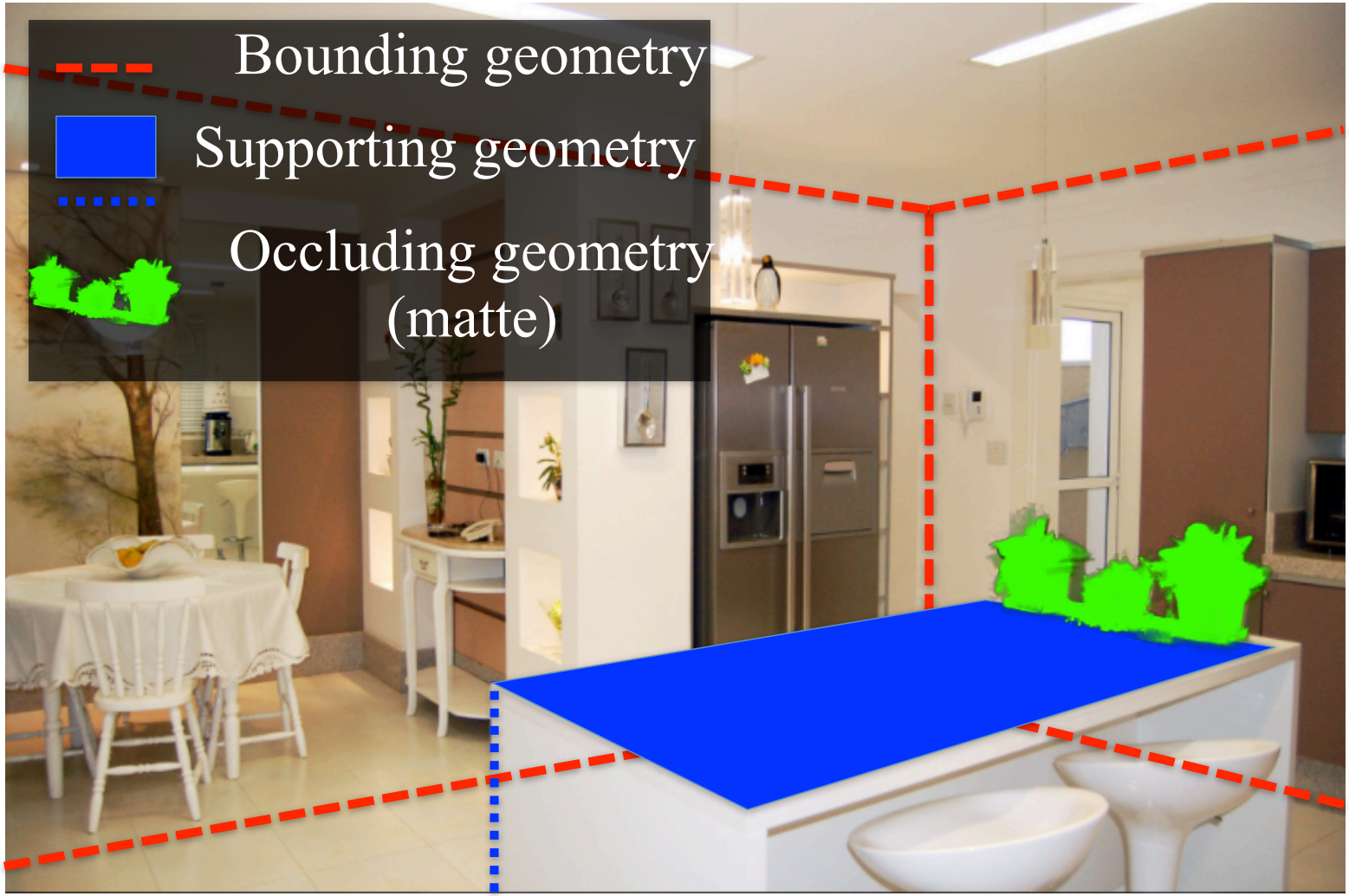
Final composite

Secret sauce: Consistency

Secret sauce: Physical renderer

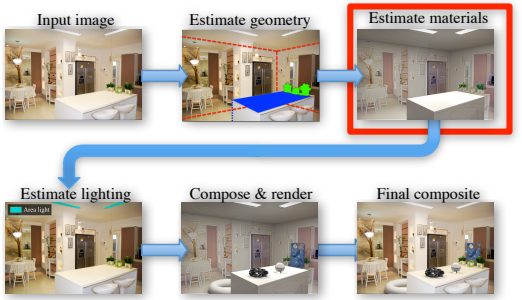
Geometry estimation

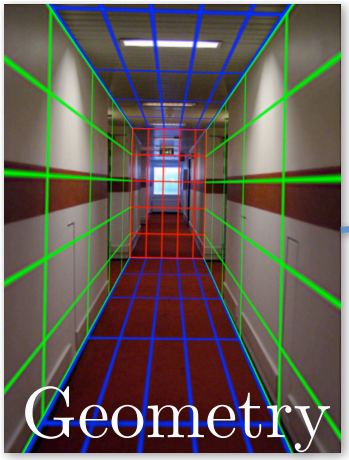




Some automatic, mostly just markup

Material estimation





Retinex-like decomposition

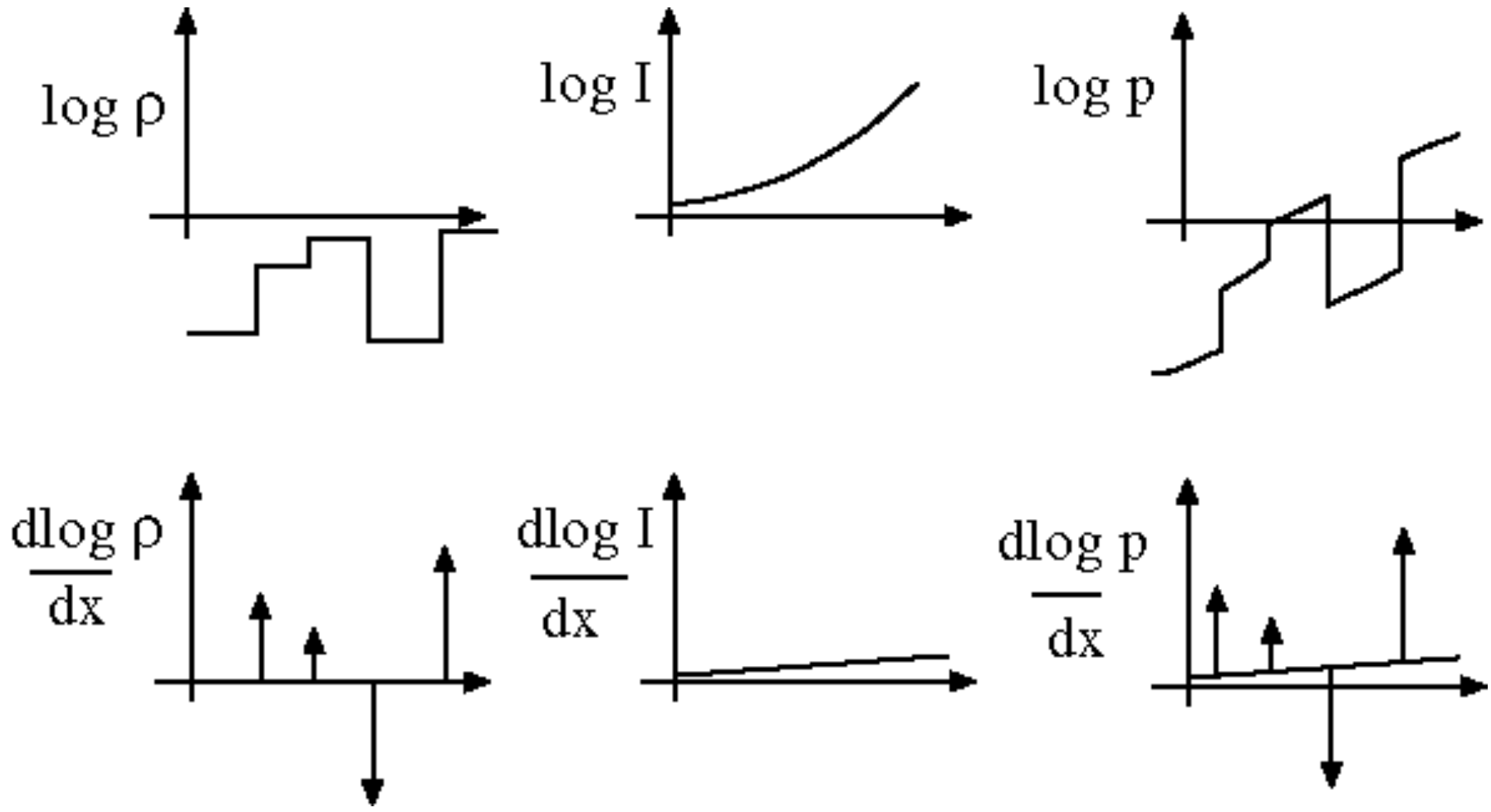
$$\underset{D,\rho}{\operatorname{argmin}}\, E_{\text{smooth}}(D) + E_{\text{piecewise}}(\rho)$$

Smooth shading

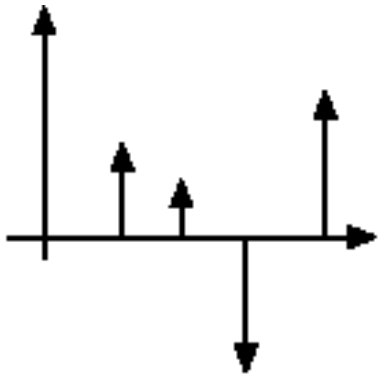
Piecewise constant reflectance

Lightness Constancy

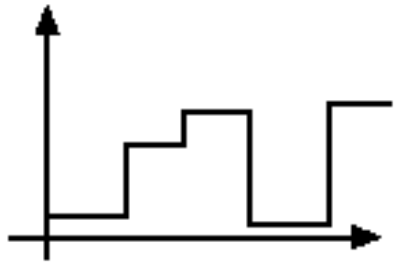
- Lightness constancy
  - how light is the surface, independent of the brightness of the illuminant
  - issues
    - spatial variation in illumination
    - absolute standard
  - Human lightness constancy is very good
- Assume
  - frontal 1D “Surface”
  - slowly varying illumination
  - quickly varying surface reflectance

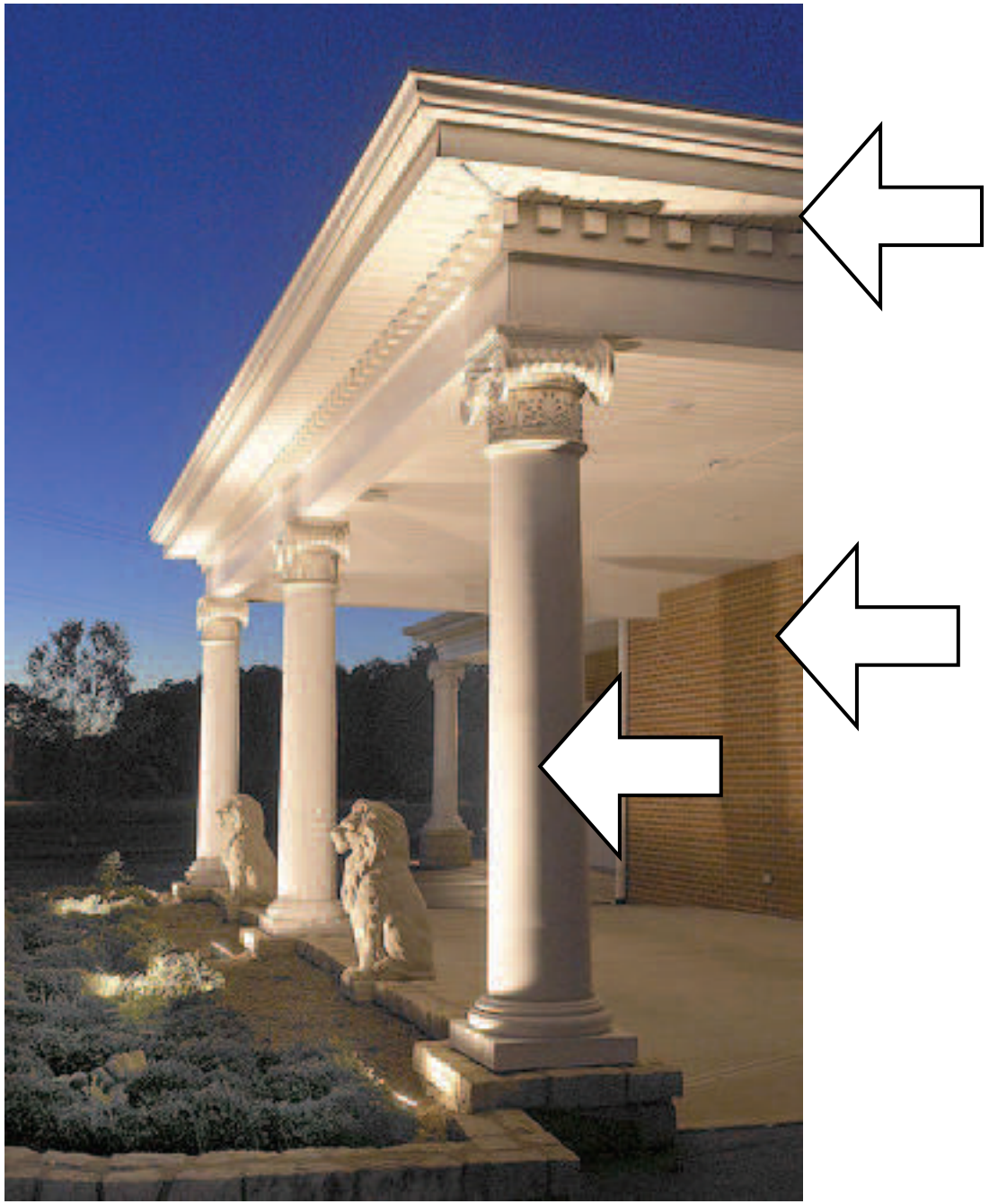

log ρ
log I
log p
dlog ρ / dx
dlog I / dx
dlog p / dx

Thresholded $\frac{d \log p}{dx}$

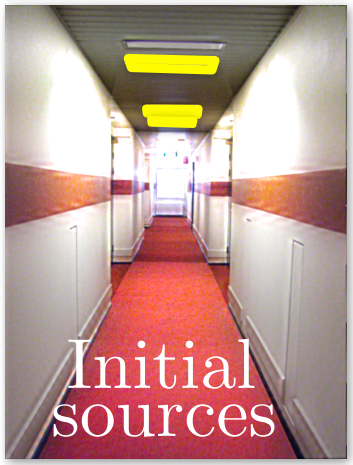


Integrate
This to get





Light Markup



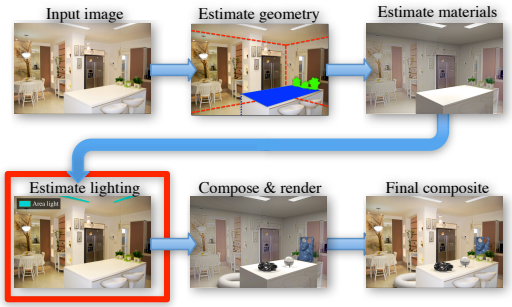
Notice long tradition of light source estimation, under various conditions, typically single/ multiple point sources; Wang+Samaras, 03; other refs in Panagopoulos' thesis

We now have ...

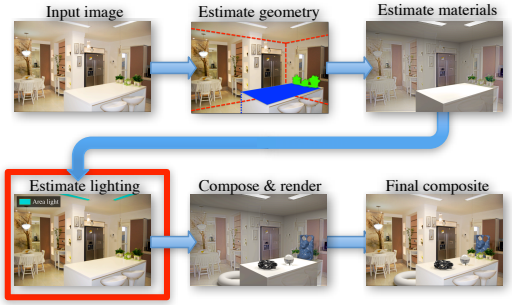
- A geometric estimate that is slightly wrong
  - inevitably
- An albedo estimate that is slightly wrong
  - inevitably
- A light markup that is slightly wrong
  - inevitably

- Adjust lights
  - so that geometry+albedo rendered with lights looks like original image

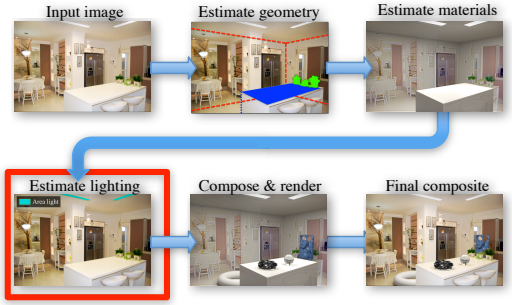
Initial light parameters

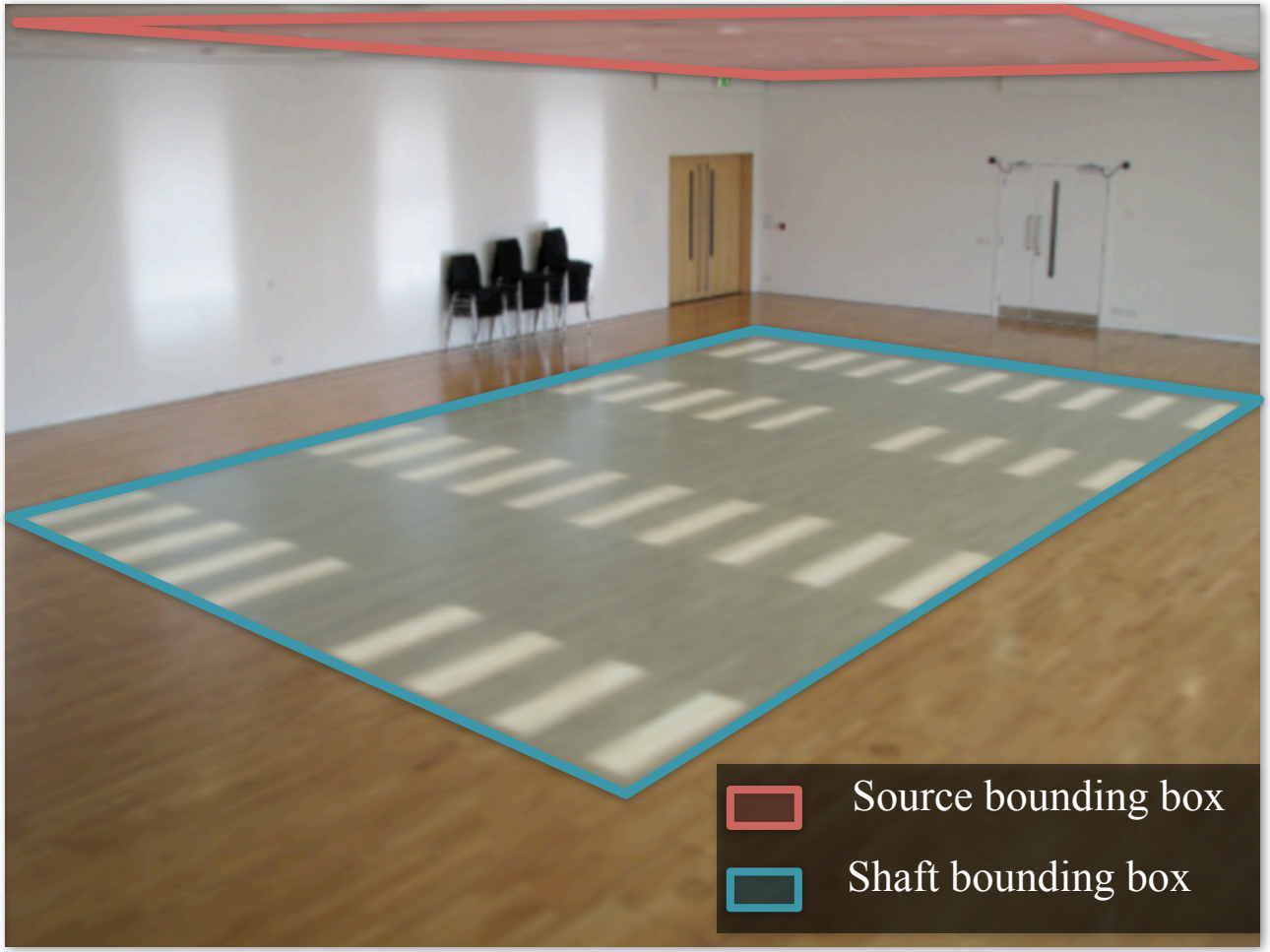


Refined light parameters



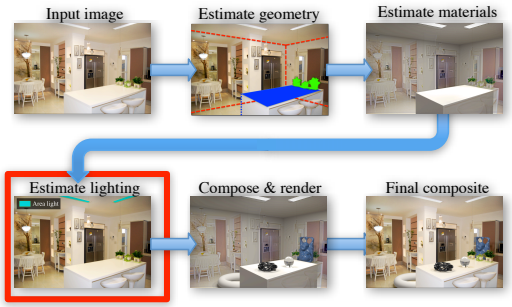
External light shafts

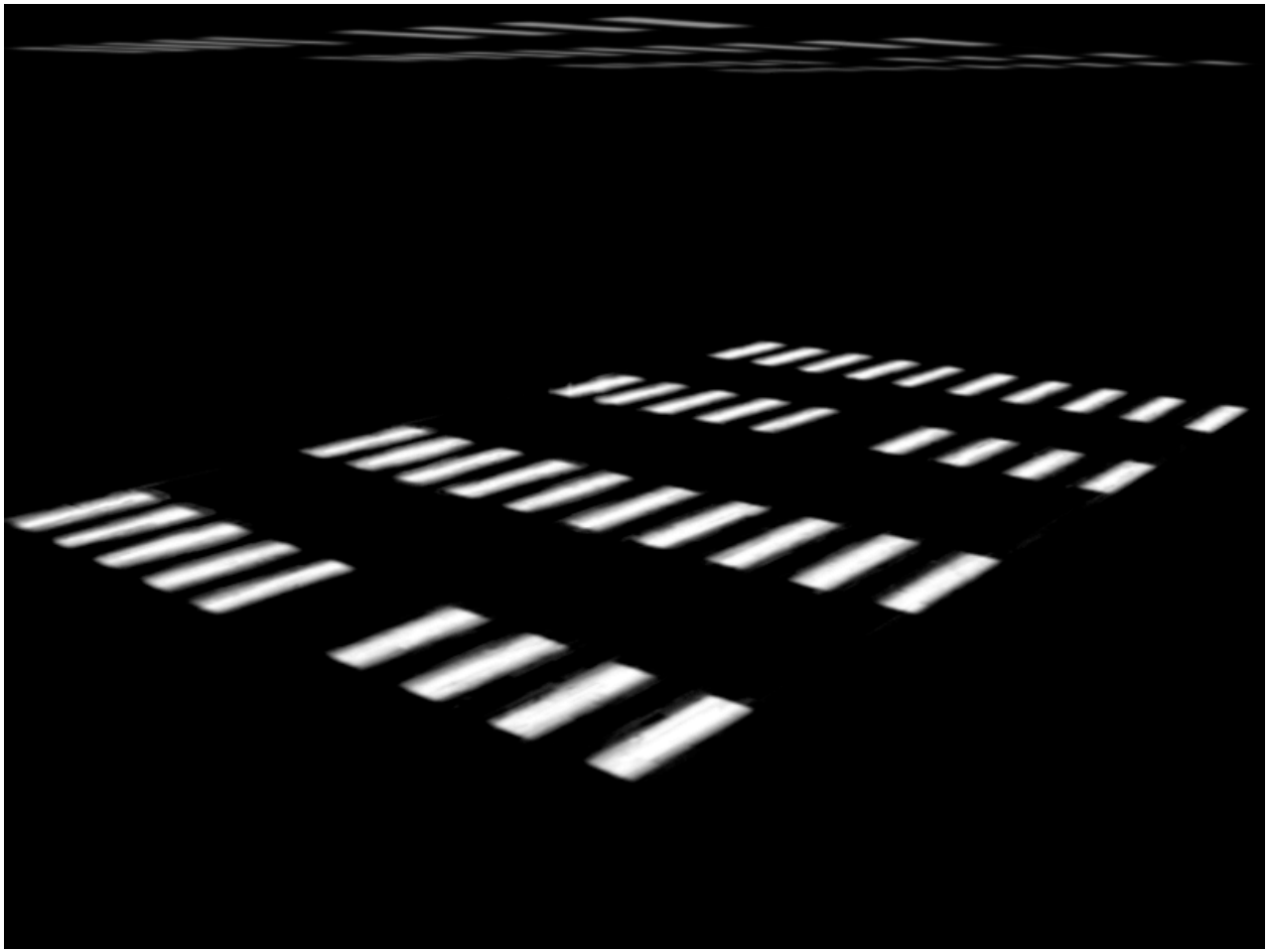




Shafts are “inverse shadows”

External light shafts

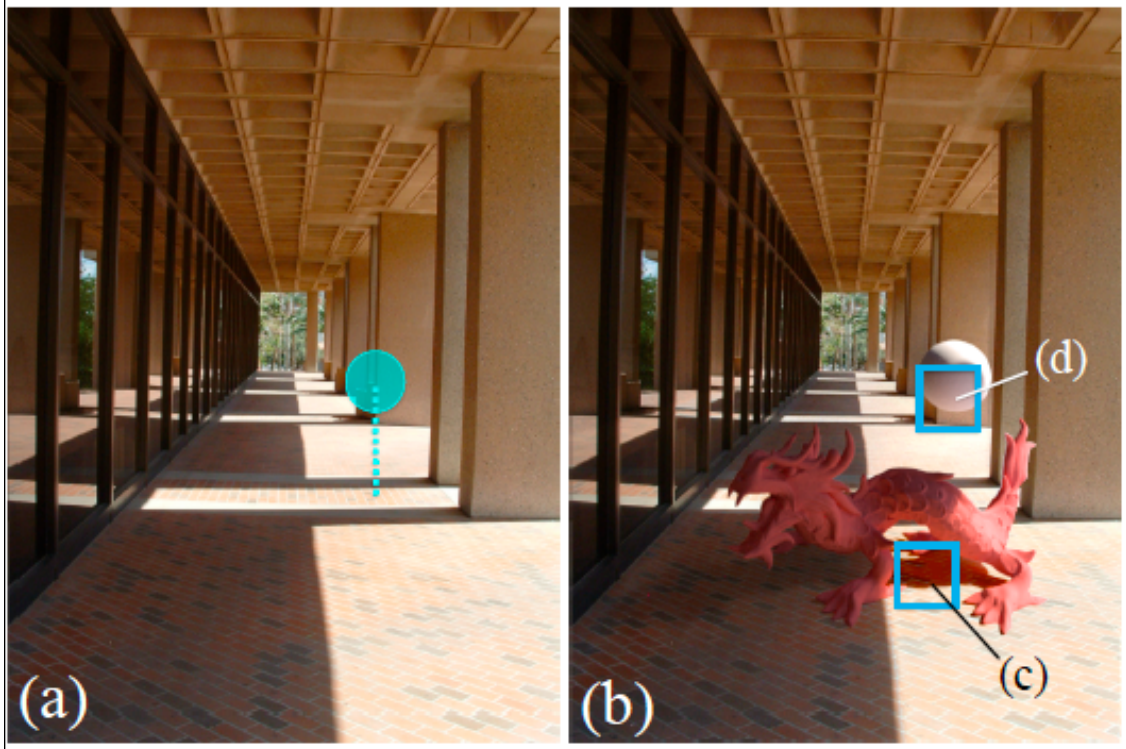




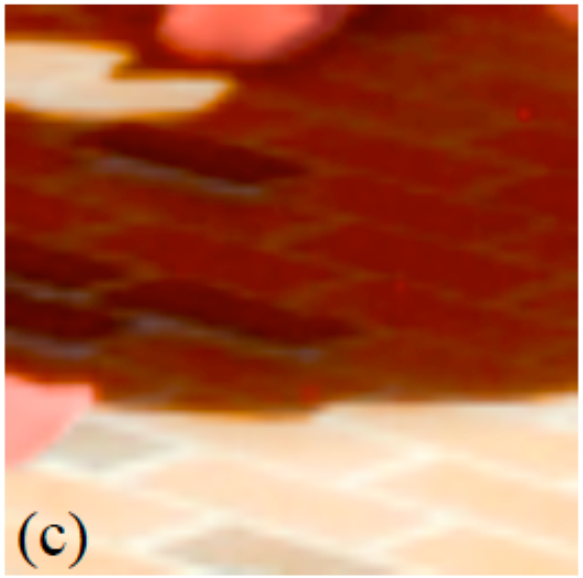
Shadow matting via Guo et al. [2011]
Many other shadow detectors, eg Panagoploulos ea 09, 10

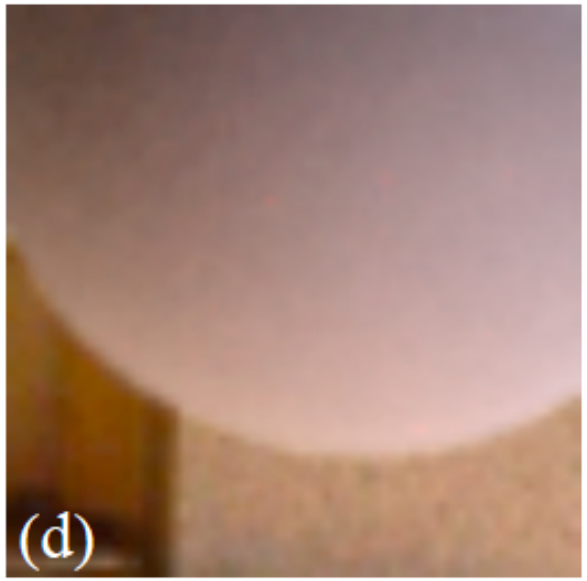
Render and composite

- Rendering
  - drop all into existing physically based renderer (LuxRender)
    - room model with materials
    - object model
    - illumination model (= glowing polygons)
    - shaft model (= distant sources)
- Compositing
  - standard method, due to Debevec, 98



Physical renderers
are wonderful!





Results

Results

Results

Results

User study

- Methods compared:
  - Our technique
  - Light probe variant [Debevec ‘98]
  - Our technique with naïve lighting (baseline)
- Illumination is dependent variable
- Geometry and materials constant
  - Estimated using our technique
- Several variants tested in real vs synthetic comparison

User study


let's do
we can have
and send them
we can put all
and they will be
Donating not


let's do
we can have
and send them
we can put all
and they will be
Donating not

User study

User study

- All three methods are highly realistic

Percentage of times users chose synthetic over real

$N = 30$	ours	light probe	baseline	total
none	20	30	13.3	21.1
monochrome	36.7	23.3	16.7	26.6
clutter	30	36.7	16.7	27.8
cropped	43.3	23.3	20	28.9
spotlight	40	23.3	N/A	31.7
total	34	27.3	16.7	26.7

- Our method preferred over light probe method 67% of image pairs
  - Possibly skewed due to light probe implementation

Results

Results

Results

Results

Results

Results

Conclusions

- Rooms are important
- Rough 3D representations help recognition
- Free space has motion "potential"
  - it tells you where you can move
  - and we can recover descriptions of free space that are good for motion
- Free space has light "potential"
  - light lives there
  - and we can recover descriptions of free space that are good for rendering
- Little 3D scene information -> very neat applications

Material inference - what is it made of?

- Think of a room as a weird goniometer
  - we know the illumination field falling on the object
  - for some objects, we know approximate shapes (eg, boxy objects; spheres)
  - we see reflected light
- Infer
  - BRDF - by comparing reflected light to illumination field
    - apply priors
    - currently in progress
  - coefficient of refraction
    - by search, matching

Coefficient of refraction

1.2

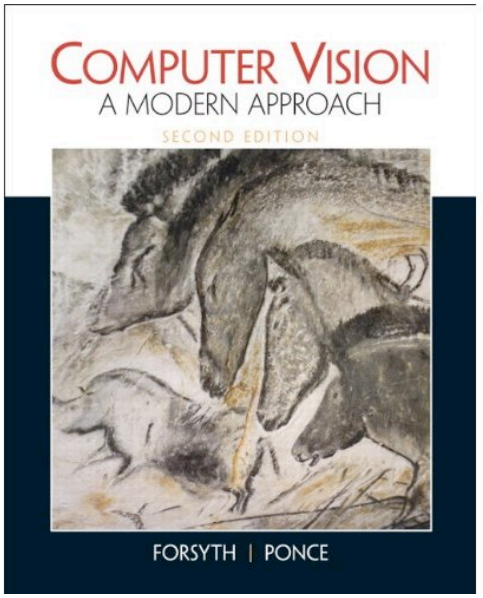
1.7

2.4

Results

Results

Conclusions



- Rooms are important
- Rough 3D representations help recognition
- Free space has motion "potential"
  - it tells you where you can move
  - and we can recover descriptions of free space that are good for motion
- Free space has light "potential"
  - light lives there
  - and we can recover descriptions of free space that are good for rendering
- Little 3D scene information -> very neat applications