Cat herding for fun and profit

D.A. Forsyth, UIUC

What every vision person should know

Vision is useful, hard, and poorly understood

- Some math
  - linear algebra, calculus, geometry, probability
  - how to learn new stuff
- Some skepticism
  - revolutionary new solutions to vision come every five years or so
- Much opportunism
  - vision is viciously difficult, because
    - you need to know so much
    - deep knowledge of one thing doesn't seem to help much
- General principles of vision, as can be deduced from
  - evolutionary examples
  - what has been successful in computer vision

Three problems of shading

- Shading reveals how much light a surface gets, so its shape
  - (due to Horn, 1970; traditional view is it got solved and its too hard)

traditionally distinct from

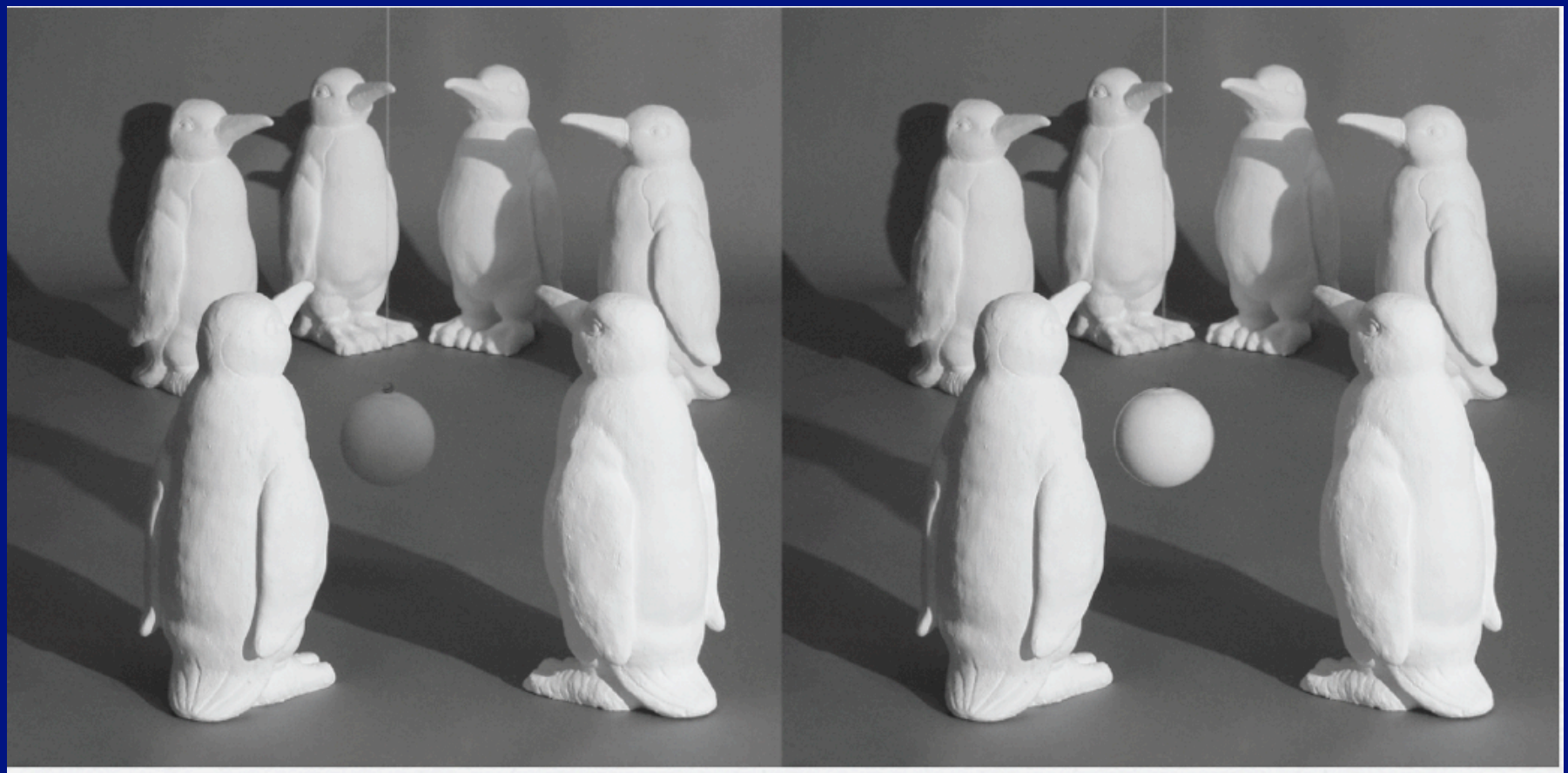
- Shading reveals how light a surface is
  - (Land, 1971; traditional view it got solved and anyhow too hard to get better)

- Shading reveals how much light there is in space

traditional view: problem doesn't exist

Stage Lighting

- Could tell us material properties
  - roughness
  - shininess
  - lightness
  - surface color
- Shape (!?!)



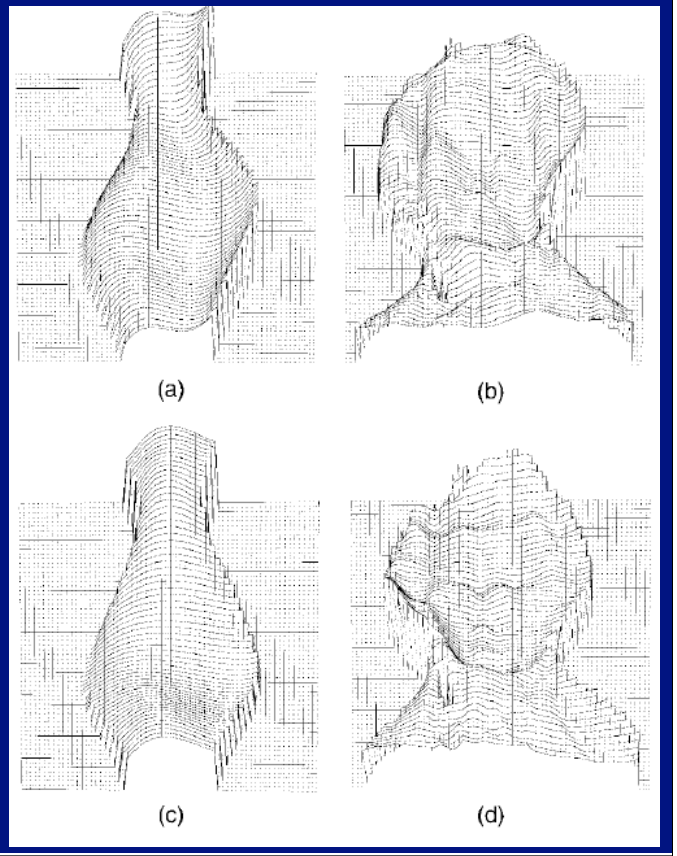
Human observers turn out to be remarkably sensitive to the light field, both to direction and diffuseness.

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

Current state of shading analysis 1, 2

- Lightness is hard
  - not much has changed in performance since Retinex (eg Grosse 09)
- Shape from shading doesn't work
  - ample evidence
    - Zhang 99 - doesn't work
    - Daniel Durou 08 - better than 99, but

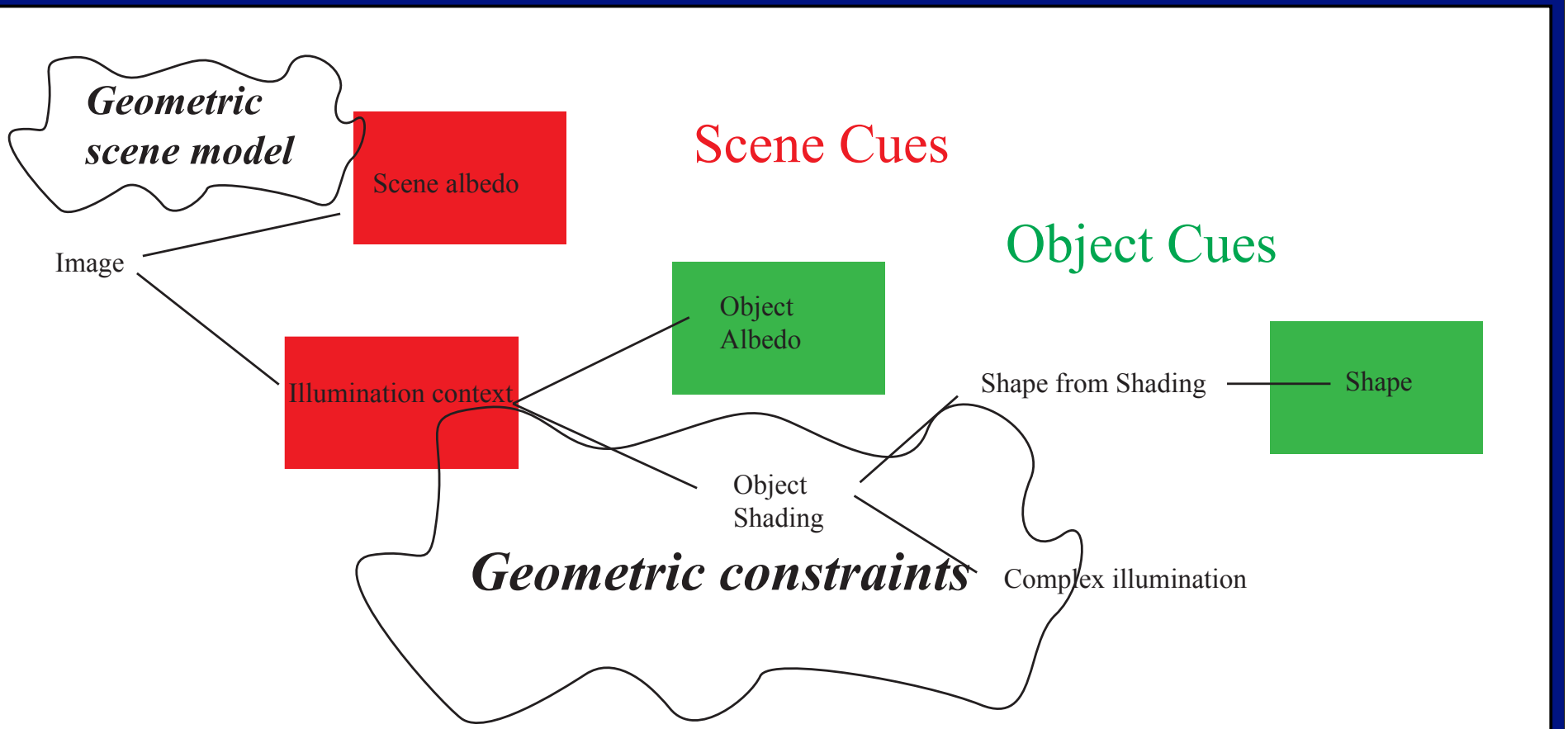


From Zhang ea, 99

Making the pieces play nice together

- Knowing a little about one thing helps another
  - eg scene geometry
    - how much?
      - plane (Hoiem ea, 06); box (Hedau ea, 09); etc.
    - for what?
      - pedestrian detection (Hoiem ea, 06); human affordances (Gupta 11); furniture detection (Hedau ea, 10)
  - context/object
- Absent from list
  - geometry/shading
  - familiar/unfamiliar
  - segmentation/shape
  - shape/object identity
  - segmentation/detection

Better picture of where shading sits



Reconstruction from shading

Usually difficult/impossible to guarantee existence

$$R(p, q; \mathbf{S}) = I(x, y)$$

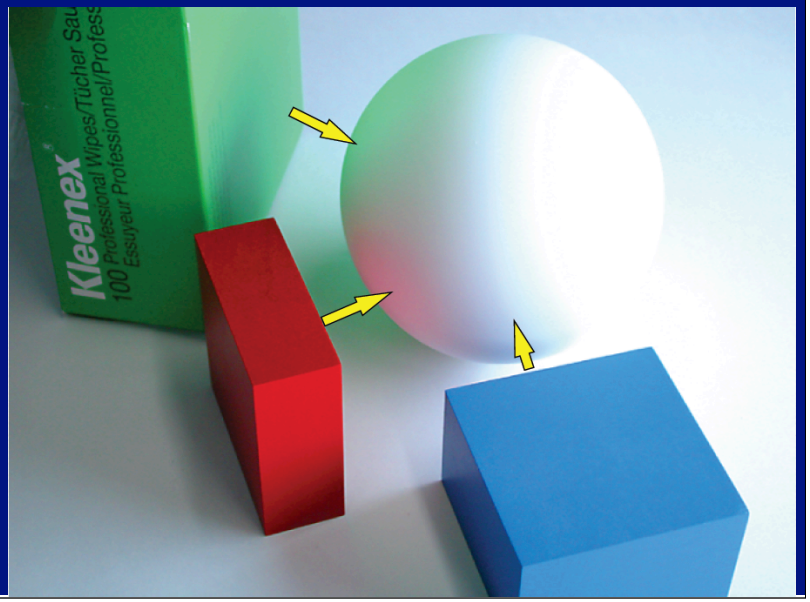
Reflectance Map Image intensity

- Local model
  - Points with the same normal get the same shading value
- The Image Irradiance Equation (IIE)
  - Horn, 1970 and lots of later papers by lots of authors
- This is a PDE
  - First order, non-linear, Eikonal equation in special case,
    - general is Hamilton Jacobi
    - inspired a huge literature

Physical Critiques

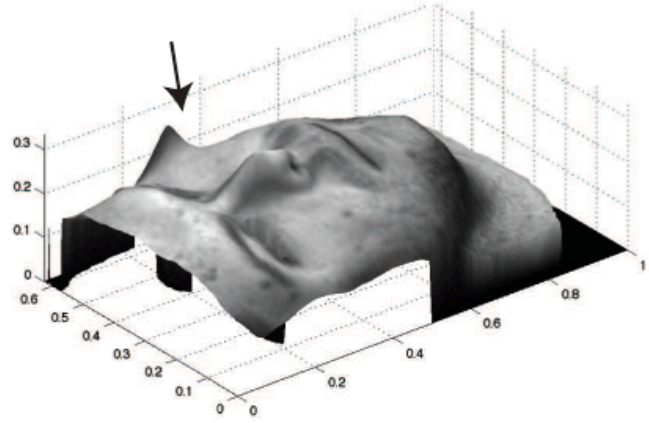
- Real shading is not local
  - interreflections
    - points with the same normal get different shading values
- Devastating
  - because a physically exact formulation is unmanageable
    - (it has been tried, Nayar et al 91)
    - cannot account for distant radiators we can't see

From Koenderink slides on image texture and the flow of light

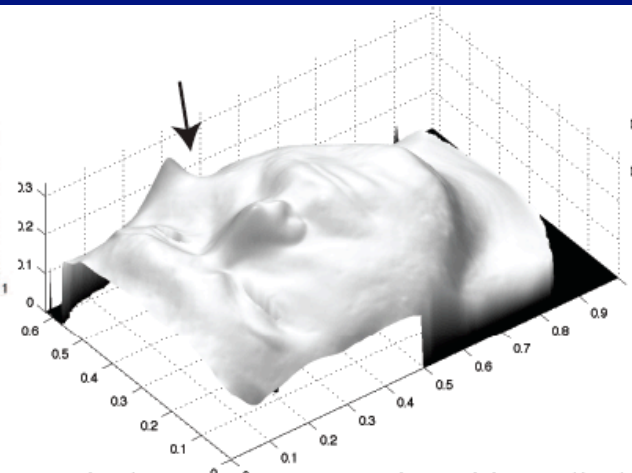


Hope for SFS - 1

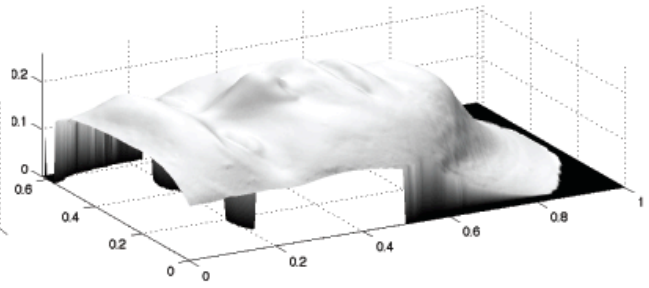
- Reconstructions of moderate accuracy
  - with existence, clean mathematics
  - From more complex illumination model+richer geometric constraints



Single source reconstruction with albedo



Single source reconstruction without albedo

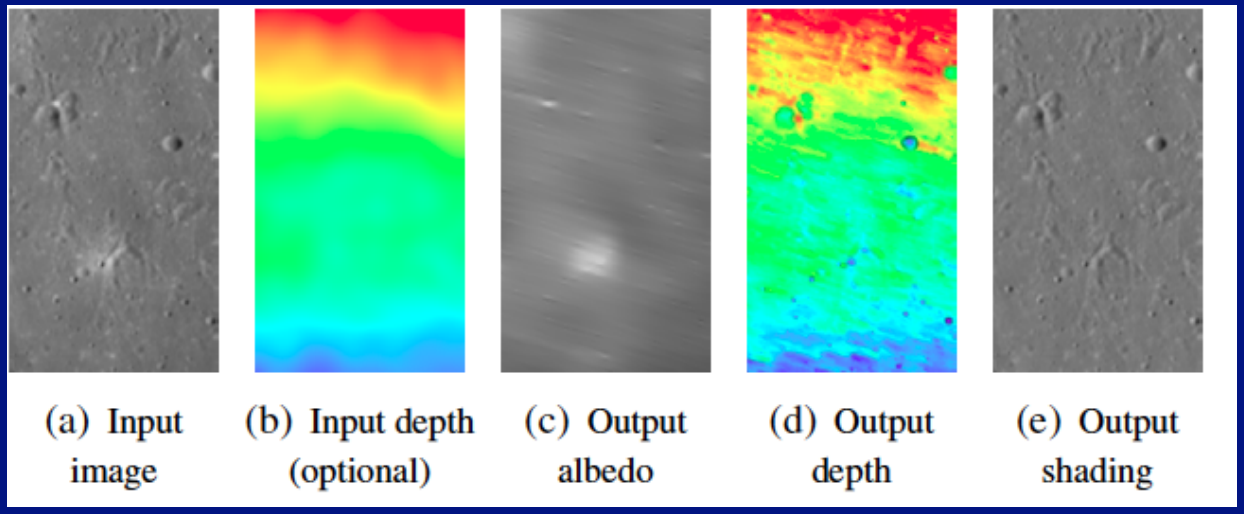


Reference photometric stereo reconstruction

Forsyth, 2011

Hope for SFS - 2

- Reconstruct detail + albedo with tolerable accuracy
  - from initial low frequency shape estimate+model



Barron Malik 2011

Rooms as boxes

V. Hedau et al ’09, many other papers

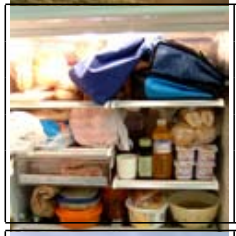
Karsch ea 11

Karsch ea 11

What we need

- Selection
  - (the cart and people are worth talking about; the buildings are not)
- Attributes
  - try to describe unfamiliar things in familiar terms
- Geometric representations that generalize
  - eg carts
- Situating objects in space with respect to one another
  - contact; potential; etc
- Predicting who can do what
  - so we notice when they don't
- Some form of narrative structure
  - in terms of goals, intentions, etc.

Examples

	(pet, sleep, ground) (dog, sleep, ground) (animal, sleep, ground) (animal, stand, ground) (goat, stand, ground)	see something unexpected. Cow in the grassfield. Beautiful scenery surrounds a fluffly sheep. Dog hearding sheep in open terrain. Cattle feeding at a trough.
	(furniture, place, furniture) (furniture, place, room) (furniture, place, home) (bottle, place, table) (display, place, table)	Refrigerator almost empty. Foods and utensils. Eatables in the refrigerator. The inside of a refrigerator apples, cottage cheese, tupperwares and lunch bags. Squash apenny white store with a hand statue, picnic tables in front of the building.
	(transportation, move, track) (bike, ride, track) (transportation, move, road) (pet, sleep, ground) (bike, ride, road)	A man stands next to a train on a cloudy day A backpacker stands beside a green train This is a picture of a man standing next to a green train There are two men standing on a rocky beach, smiling at the camera. This is a person laying down in the grass next to their bike in front of a strange white building.
	(display, place, table) (furniture, place, furniture) (furniture, place, furniture) (bottle, place, table) (furniture, place, home)	This is a lot of technology. Somebody's screensaver of a pumpkin A black laptop is connected to a black Dell monitor This is a dual monitor setup Old school Computer monitor with way to many stickers on it

Farhadi ea 10

Attribute predictions for unknown objects

Farhadi et al 09; cf Lampert et al 09

What we need

- Selection
  - (the cart and people are worth talking about; the buildings are not)
- Attributes
  - try to describe unfamiliar things in familiar terms

G
- Geometric representations that generalize
  - eg carts

G
- Situating objects in space with respect to one another
  - contact; potential; etc

S
- Predicting whøwhat can do what
  - so we notice when they don't

S
- Some form of narrative structure
  - in terms of goals, intentions, etc.