Materials and weather

D.A. Forsyth, UIUC

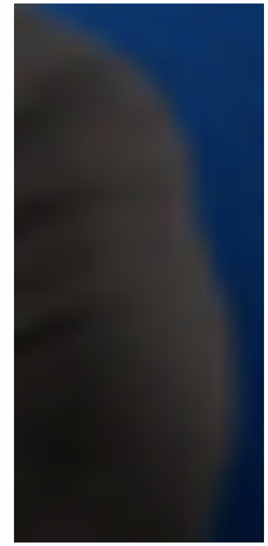
Key issues

- Physical
  - what makes a pixel take its brightness values?
  - Effects
    - at surfaces
    - in volume
  - Human: what can people do?
    - which suggests problems we might be able to solve
- Sensing
  - can we sense in ways that reduce significance of effects?
  - sensor fusion, etc.
- Inference
  - what can we recover from the world using sensed values?

Effects at surfaces

- We assume:
    - we see the world in a vacuum
        - or very clear air, no fog, nothing

nickwheeleroz

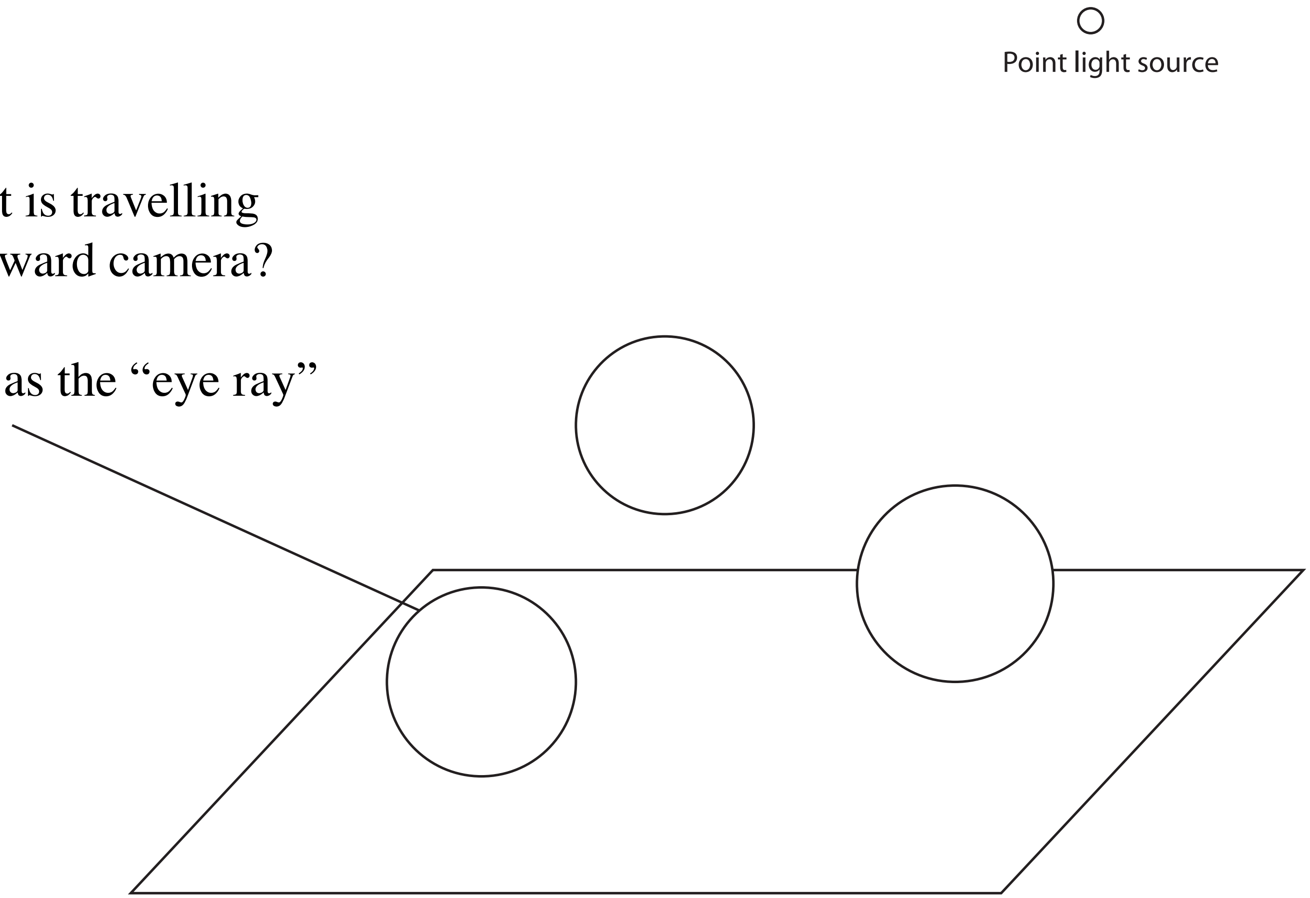


By nickwheeleroz, on Flickr

Very simple ray-tracing

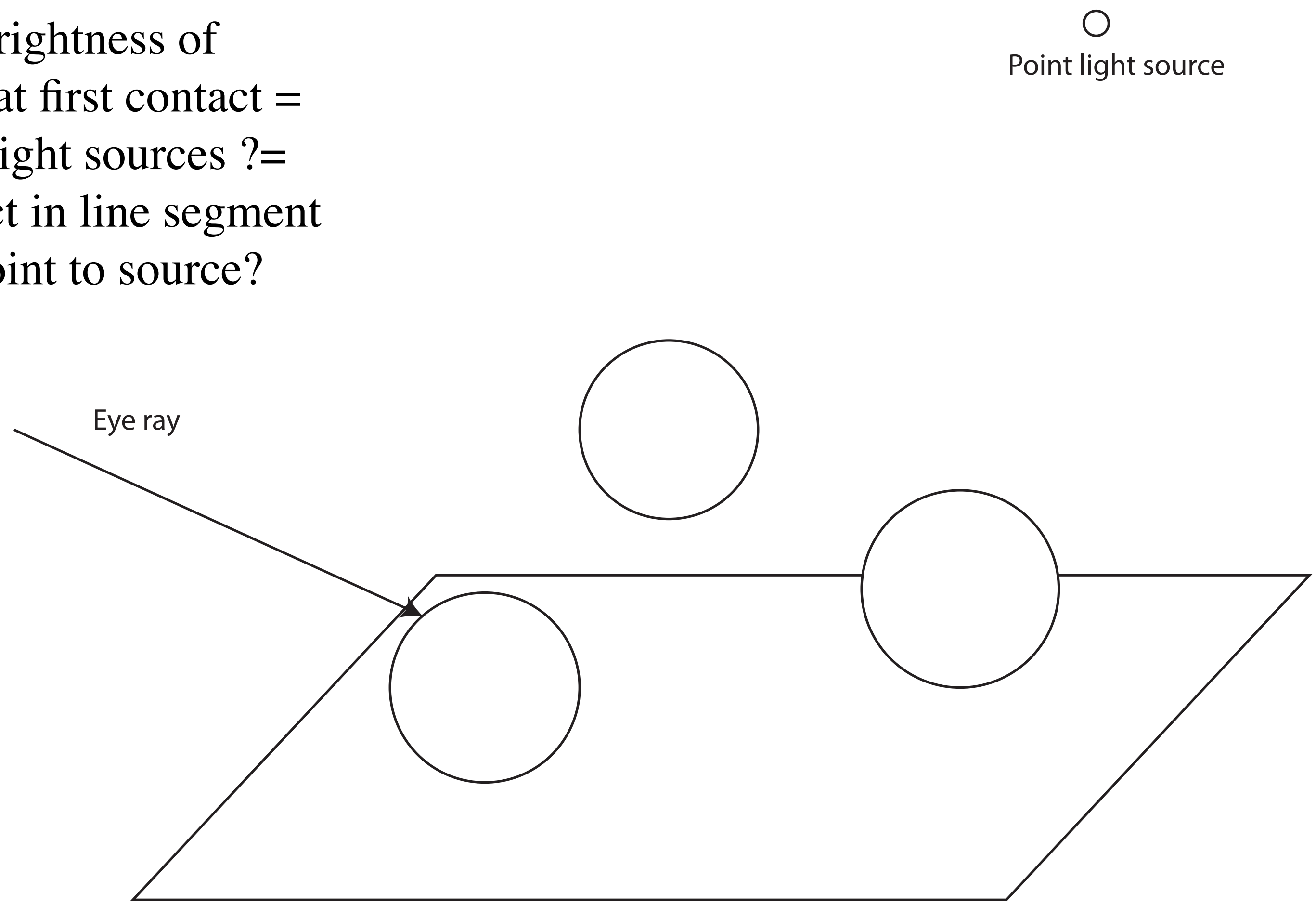
How much light is travelling
down this ray toward camera?

sometimes known as the “eye ray”



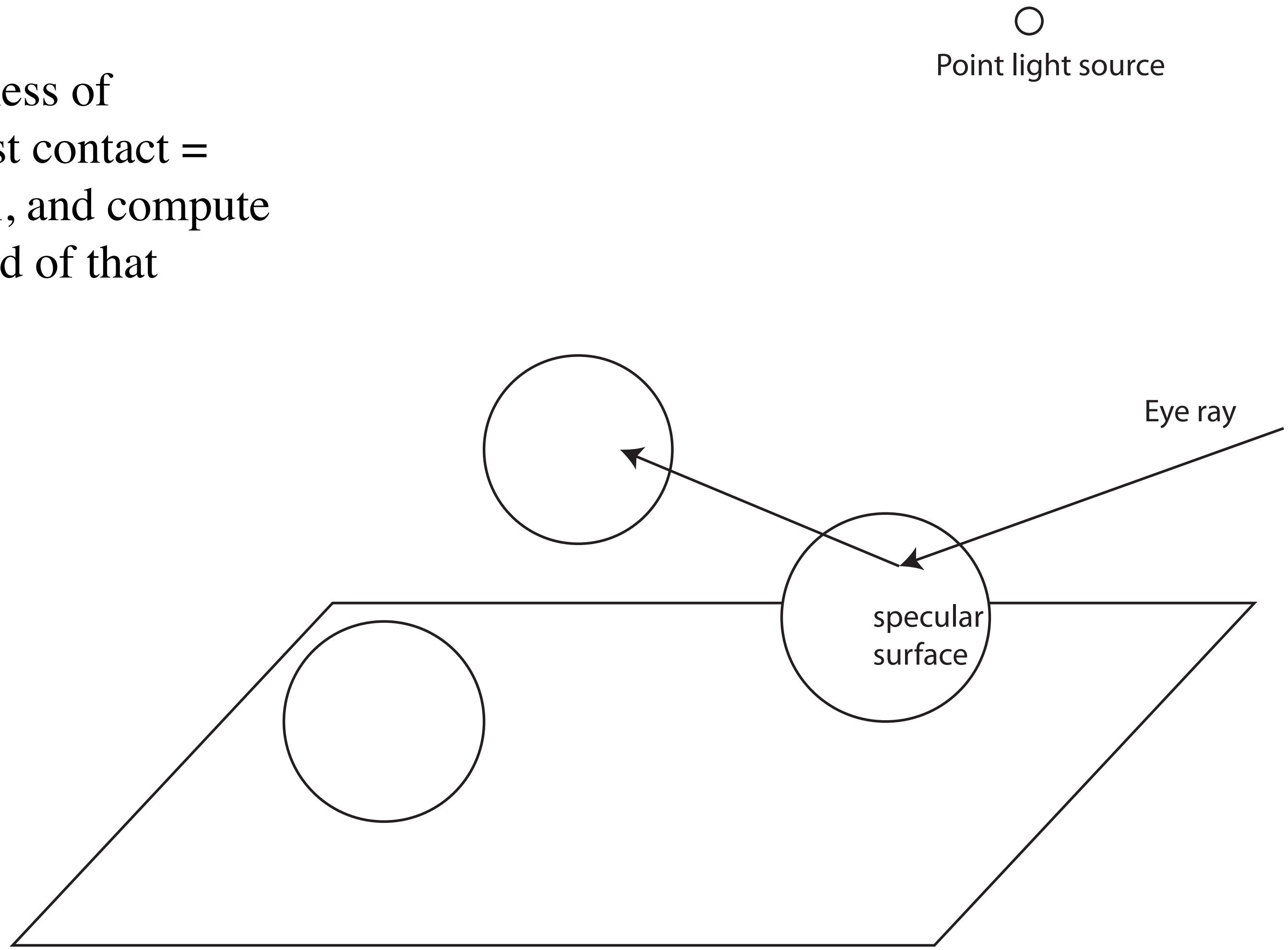
Eye ray strikes diffuse surface

Compute brightness of
diffuse surface at first contact =
Can it see the light sources ?=
Is there an object in line segment
connecting point to source?



Eye ray strikes specular surface

Compute brightness of
specular surface at first contact =
eye ray changes direction, and compute
brightness at the end of that



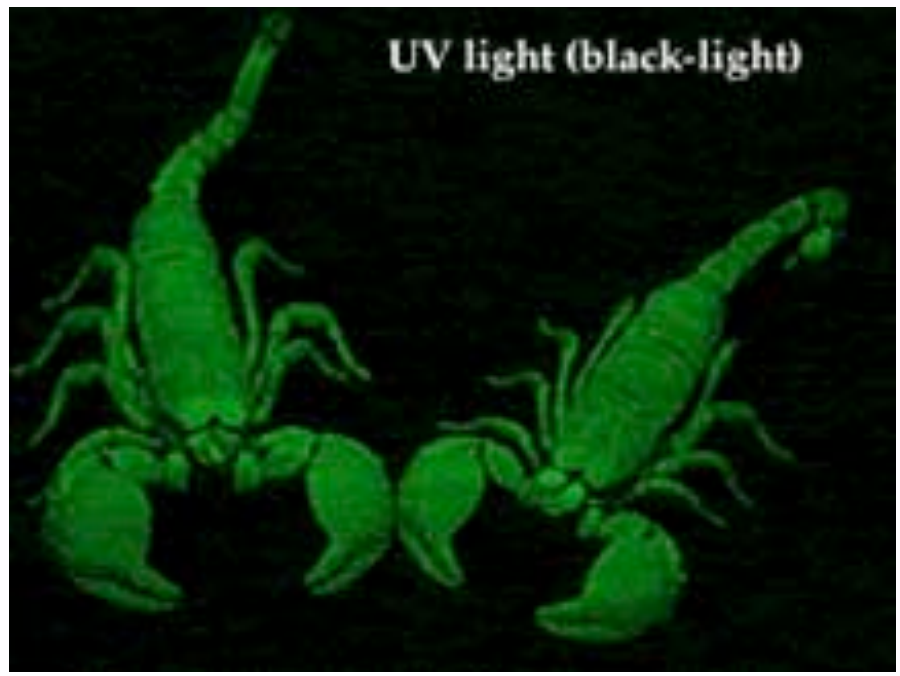
Lighting model

- Light arrives at a surface ONLY from a luminaire
  - this is an object that “makes light”
    - through chemical, mechanical, etc means
- Wild oversimplification, good for us right now
  - wait a few slides and it’ll get more complicated

Processes

- Cameras
  - film: non-linear
  - CCD: linear, with non-linearities made by electronics
- Light
  - is reflected from a surface
  - got there from a source
- Many effects when light strikes a surface -- could be:
  - absorbed; transmitted; reflected; scattered
  - Simplify
    - Assume that
      - surfaces don't fluoresce
      - surfaces don't emit light (i.e. are cool)
      - all the light leaving a point is due to that arriving at that point


White light (indoor)


UV light (black-light)

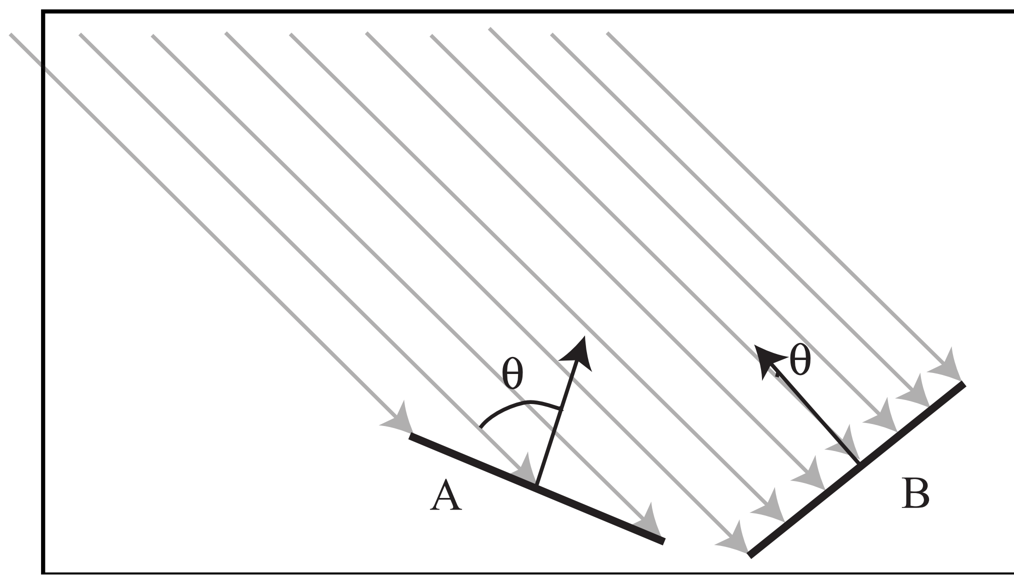
Diffuse reflection

- Light leaves the surface evenly in all directions
  - cotton cloth, carpets, matte paper, matte paints, etc.
  - most “rough” surfaces
  - Parameter: Albedo
    - percentage of light arriving that leaves
    - range 0-1
      - practical range is smaller
- Test:
  - surface has same apparent brightness when viewed from different dir’ns

Point source at infinity

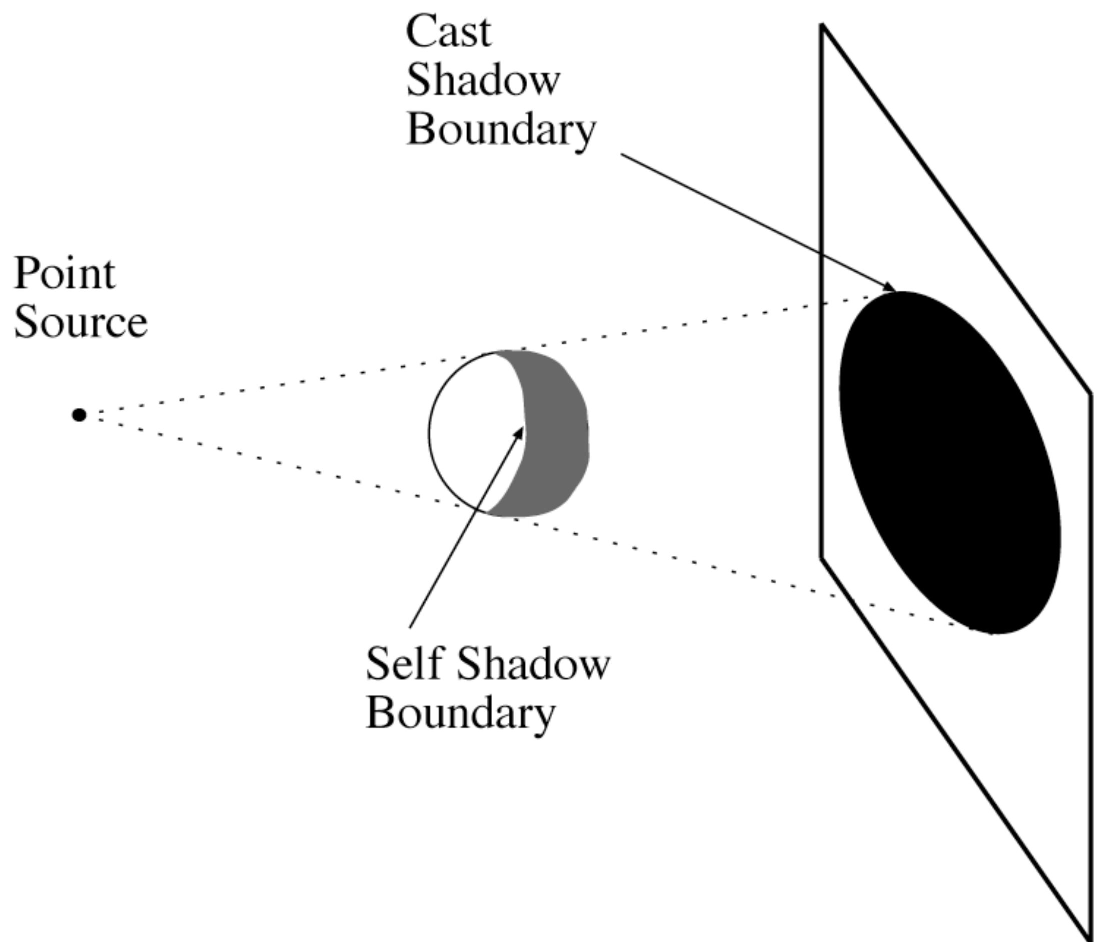
- E.g. the sun
  - energy travels in parallel rays
  - energy density received is proportional to cos theta
- Write:
  - p for albedo
  - S for source vector
    - from surface to source
    - length=intensity of source
  - N for normal
  - I for image intensity

$$I = \rho(N \cdot S)$$

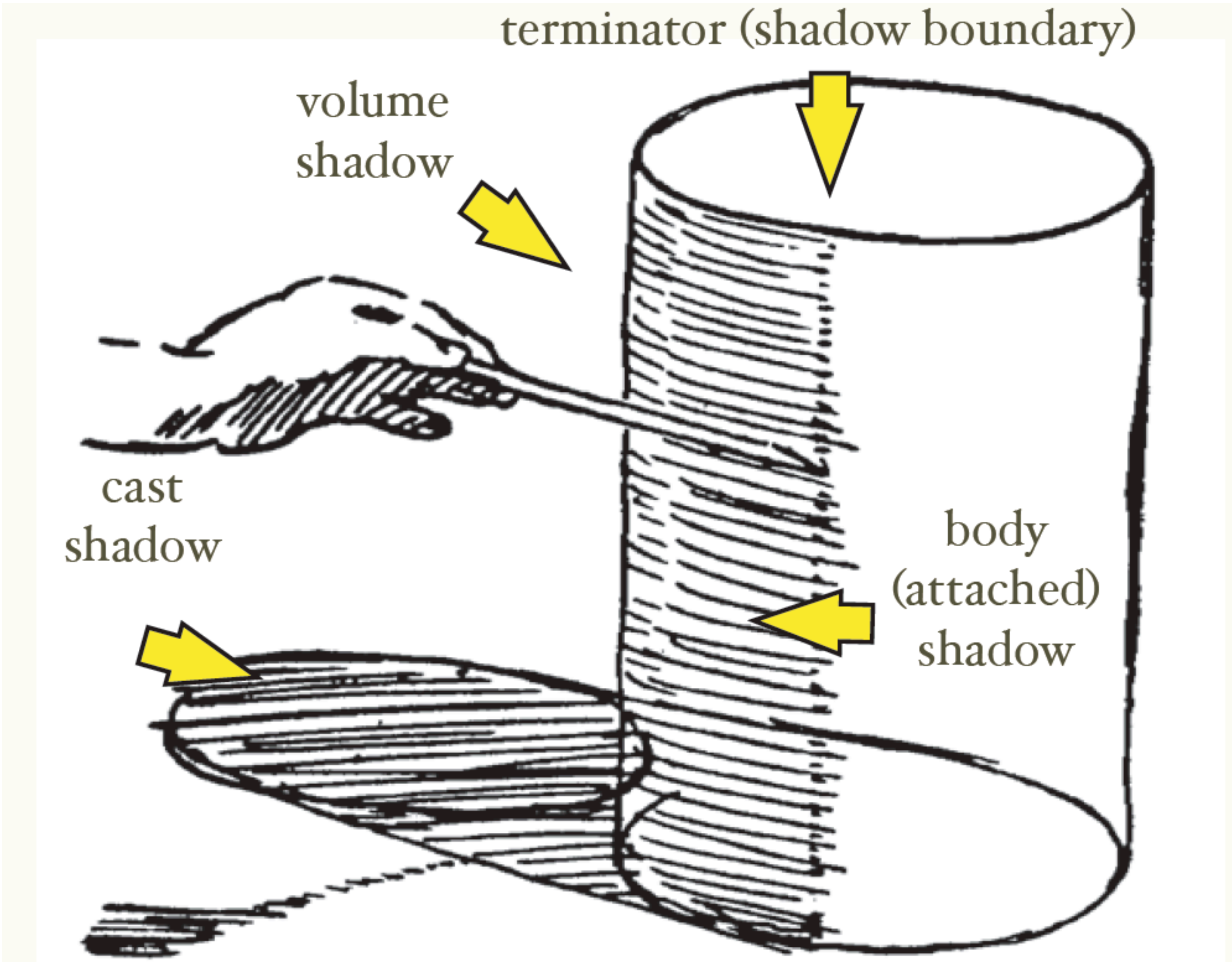


Shadows cast by a point source

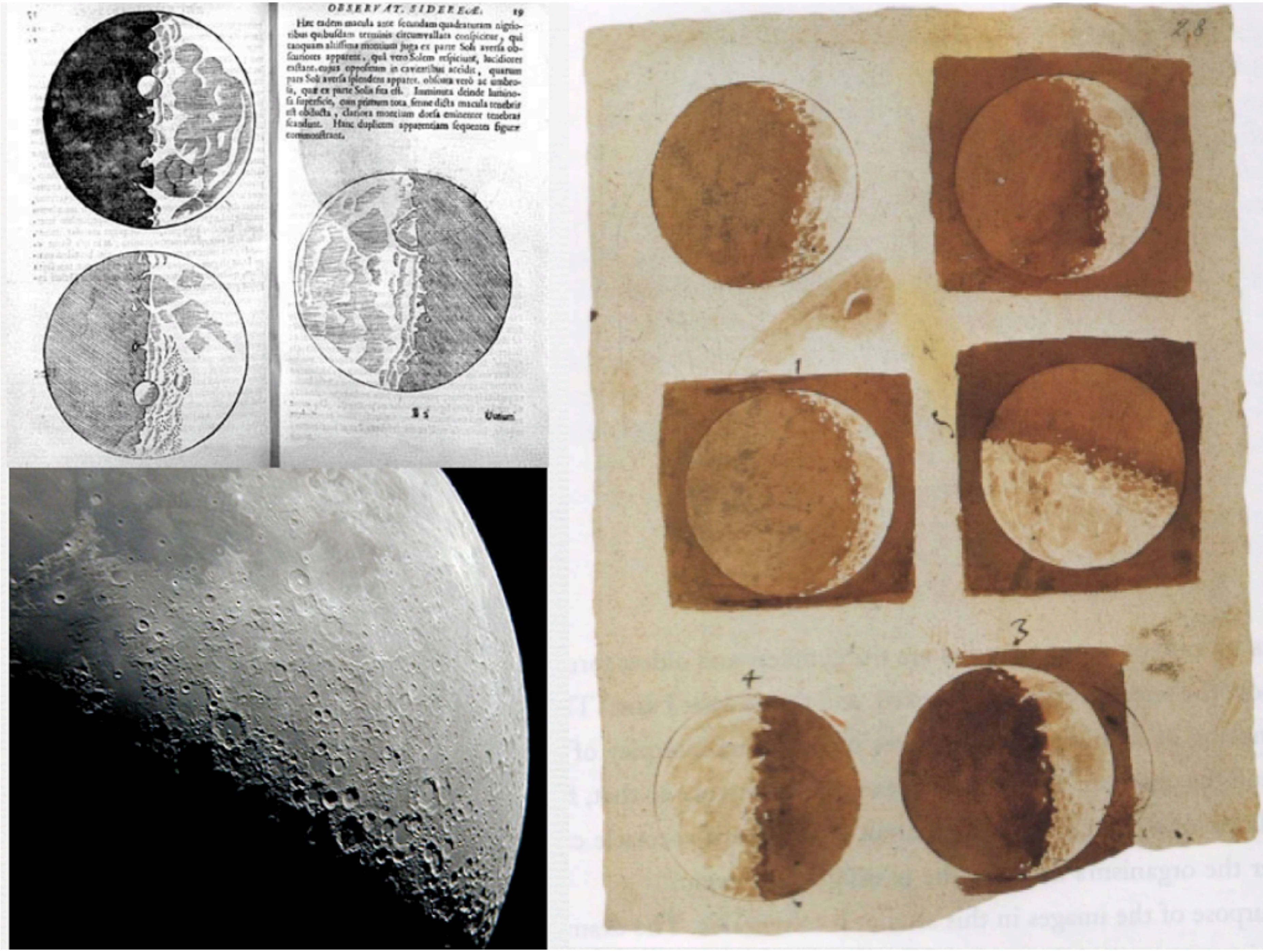
- A point that can’t see the source is in shadow
- For point sources, the geometry is simple



Cues to shape - shadows



From Koenderink slides on image texture and the flow of light



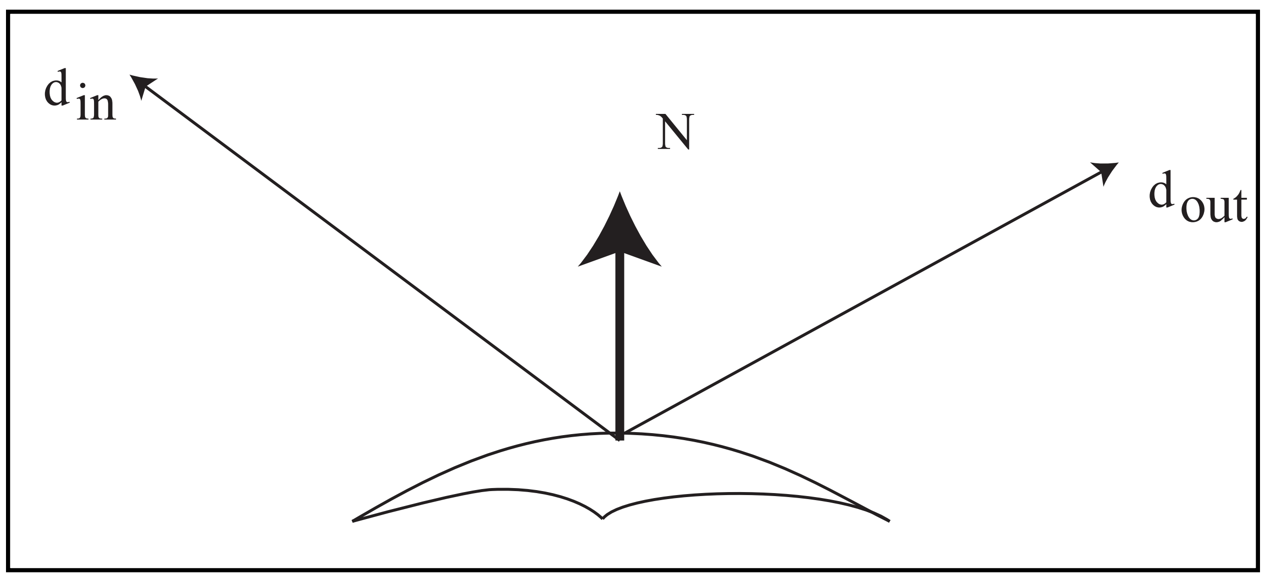
From Koenderink slides on image texture and the flow of light

Shadow geometry can be very nasty

From Hel Des, on Flickr

Specularities

- For some surfaces, reflection depends strongly on angle
  - mirrors (special case)
    - incoming direction, normal and outgoing direction are coplanar
    - angle din, normal and angle dout, normal are the same
  - specular surfaces
    - light reflected in a "lobe" of directions
    - eg slightly battered metal surface
    - can see light sources specularly reflected
      - specularities



Flickr, by suzysputnik

Flickr, by piratejohnny

- Specularities are relatively easy to detect
  - small and bright (usually)

Key idea - how bright is this point?

Radiometry

- Questions:
  - how “bright” will surfaces be?
  - what is “brightness”?
    - measuring light
    - interactions between light and surfaces
- Core idea - think about light arriving at a surface
  - around any point is a hemisphere of directions
  - what is important is what a source “looks like” to a receiver
    - receiver can’t know anything else about source

Radiance

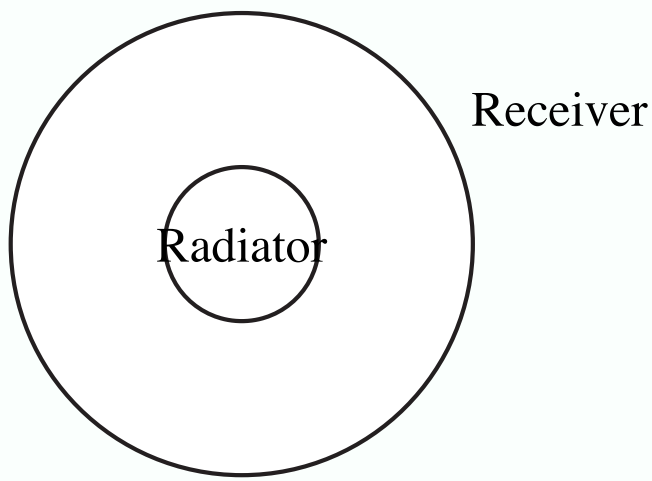
- Measure the “amount of light” at a point, in a direction

  the power (amount of energy per unit time) traveling at some point in a specified direction, per unit area *perpendicular to the direction of travel*, per unit solid angle.

- Units: watts per square meter per steradian (wm-2sr-1)
- Crucial property:
  - In a vacuum, radiance leaving p in the direction of q is the same as radiance arriving at q from p
  - hence the units

Why not watts/square meter?

- Consider sphere radiating 1 W into vacuum
  - Radius 1, center at origin
  - Vacuum neither creates nor consumes power
- There's another sphere around it
  - Radius R, center at origin
  - Area - 4 pi R^2
  - It can't collect more power than first sphere radiates so
    - watts/square meter must go down with distance….!!! (ew)



Surfaces and the BRDF

- Many effects when light strikes a surface -- could be:
  - absorbed; transmitted. reflected; scattered
- Assume that
  - surfaces don't fluoresce
  - surfaces don't emit light (i.e. are cool)
  - all the light leaving a point is due to that arriving at that point
- Can model this situation with the Bidirectional Reflectance Distribution Function (BRDF)
- the ratio of the radiance in the outgoing direction to the incident irradiance

$$\rho_{bd}(\underline{x},\vartheta_o,\varphi_o,\vartheta_i,\varphi_i,) = \frac{L_o(\underline{x},\vartheta_o,\varphi_o)}{L_i(\underline{x},\vartheta_i,\varphi_i)\cos\vartheta_i d\omega}$$

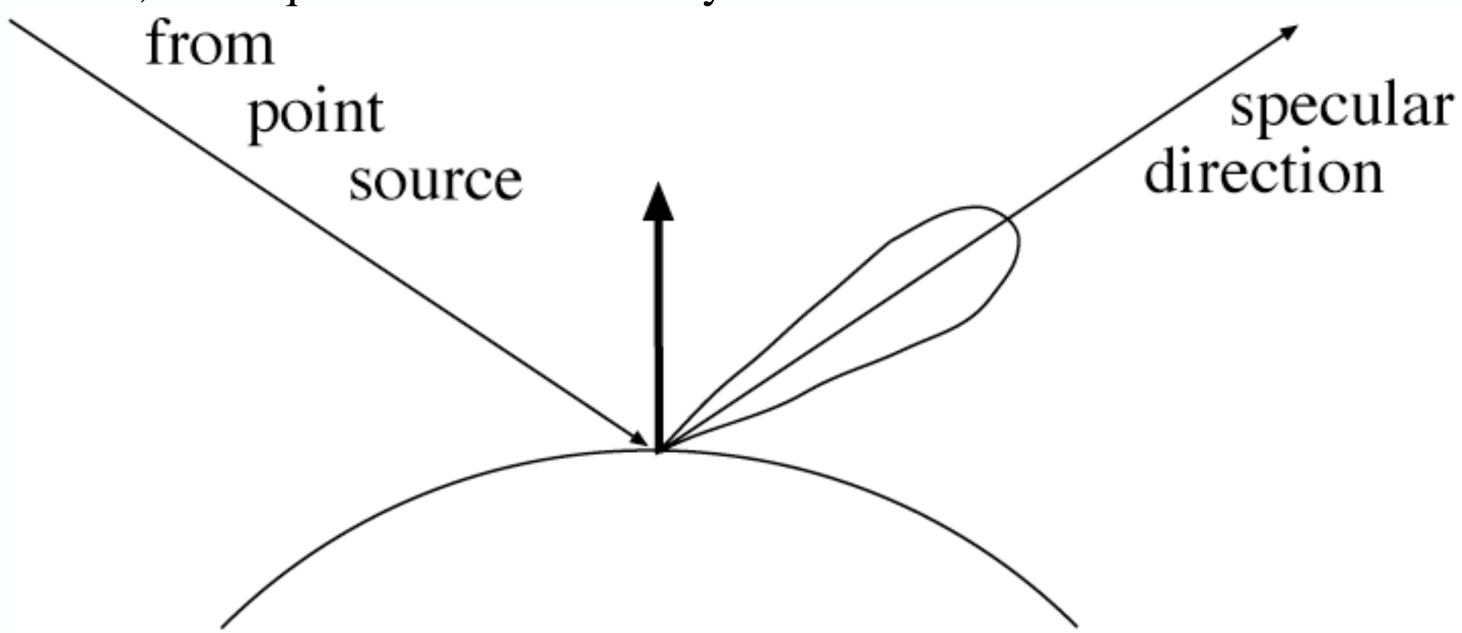
Lambertian surfaces and albedo

- For some surfaces, the BRDF is independent of direction
  - cotton cloth, carpets, matte paper, matte paints, etc.
  - radiance leaving the surface is independent of angle
  - Lambertian surfaces (same Lambert) or ideal diffuse surfaces
  - Use radiosity as a unit to describe light leaving the surface
  - percentage of incident light reflected is diffuse reflectance or albedo
- Useful fact:

$$\rho_{brdf} = \frac{\rho_d}{\pi}$$

Specular surfaces

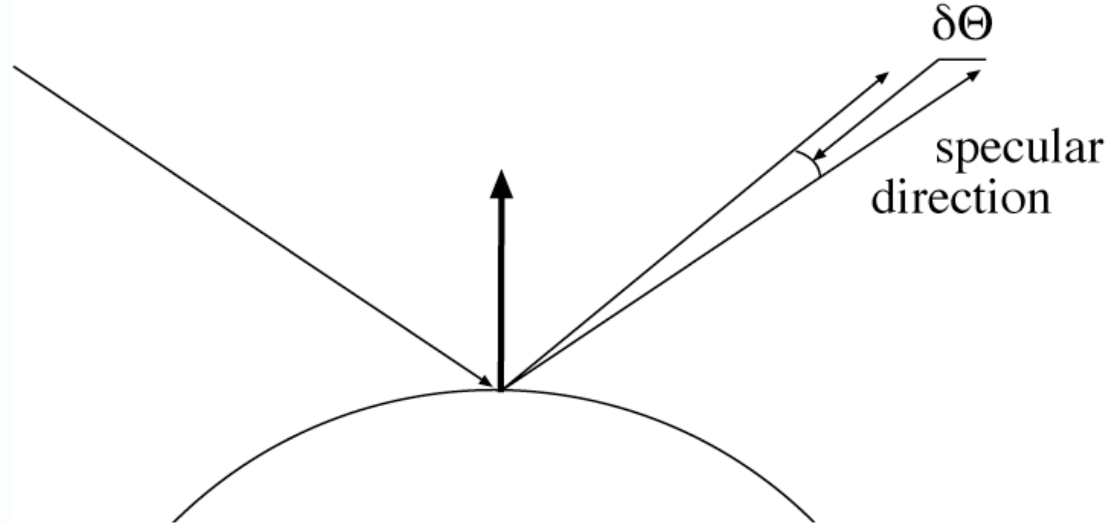
- Another important class of surfaces is specular, or mirror-like.
  - radiation arriving along a direction leaves along the specular direction
  - reflect about normal
  - some fraction is absorbed, some reflected
  - on real surfaces, energy usually goes into a lobe of directions
  - can write a BRDF, but requires the use of funny functions



Phong's model

- There are very few cases where the exact shape of the specular lobe matters.
- Typically:
  - very, very small --- mirror
  - small -- blurry mirror
  - bigger -- see only light sources as "specularities"
  - very big -- faint specularities
- Phong's model
  - reflected energy falls off with

$$\cos^n(\delta\vartheta)$$



Lambertian + specular

- Widespread model
  - all surfaces are Lambertian plus specular component
- Advantages
  - easy to manipulate
  - very often quite close true
- Disadvantages
  - some surfaces are not
    - e.g. underside of CD's, feathers of many birds, blue spots on many marine crustaceans and fish, most rough surfaces, oil films (skin!), wet surfaces
  - Generally, very little advantage in modelling behaviour of light at a surface in more detail -- it is quite difficult to understand behaviour of L+S surfaces

The Rendering Equation- 1

- We can now write

Angle between normal and incoming direction

$$L_o(\mathbf{x}, \omega_o) = L_e(\mathbf{x}, \omega_o) + \int_\Omega \rho_{bd}(\mathbf{x}, \omega_o, \omega_i) L_i(\mathbf{x}, \omega_i) \cos\theta_i d\omega_i$$

BRDF

Incoming radiance

Average over hemisphere

Radiance emitted from surface at that point in that direction

Radiance leaving a point in a direction

Radiance is constant along straight lines, so this is what we want to know

The Rendering Equation - II

- This balance works for
  - each wavelength,
  - at any time, so
- So

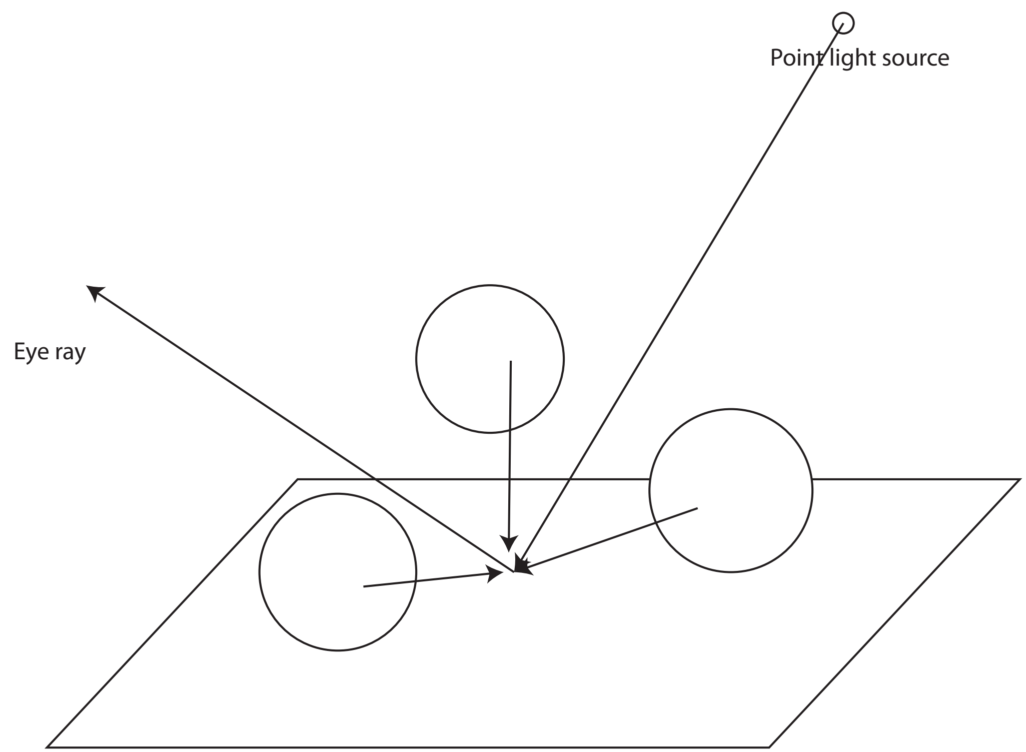
$$
\begin{aligned}
L_o(\mathbf{x}, \omega_o, \lambda, t) &= L_e(\mathbf{x}, \omega_o, \lambda, t) + \\
& \int_\Omega \rho_{bd}(\mathbf{x}, \omega_o, \omega_i, \lambda, t) L_i(\mathbf{x}, \omega_i, \lambda, t) \cos\theta_i d\omega_i
\end{aligned}
$$

Global illumination

$$L_o(\mathbf{x}, \omega_o) = L_e(\mathbf{x}, \omega_o) + \int_\Omega \rho_{bd}(\mathbf{x}, \omega_o, \omega_i) L_i(\mathbf{x}, \omega_i) \cos\theta_i d\omega_i$$

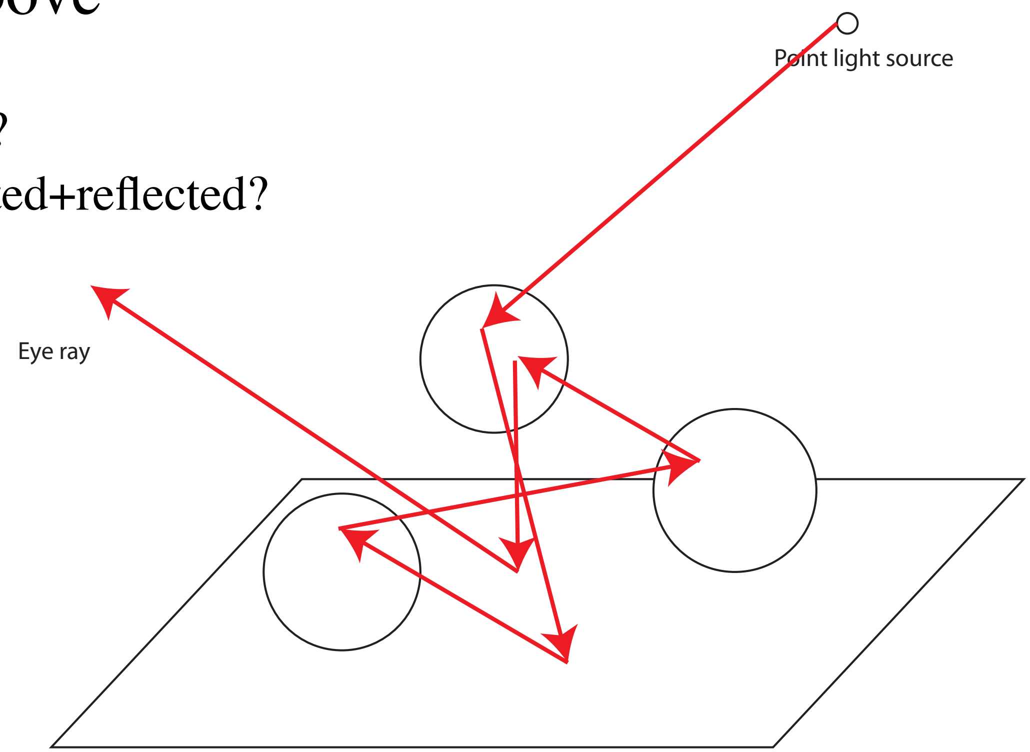
Incoming radiance

- Incoming radiance isn't just from luminaires
    - the reason you can see surfaces is they reflect light
    - other surfaces don't distinguish between reflected light and generated light



Light paths

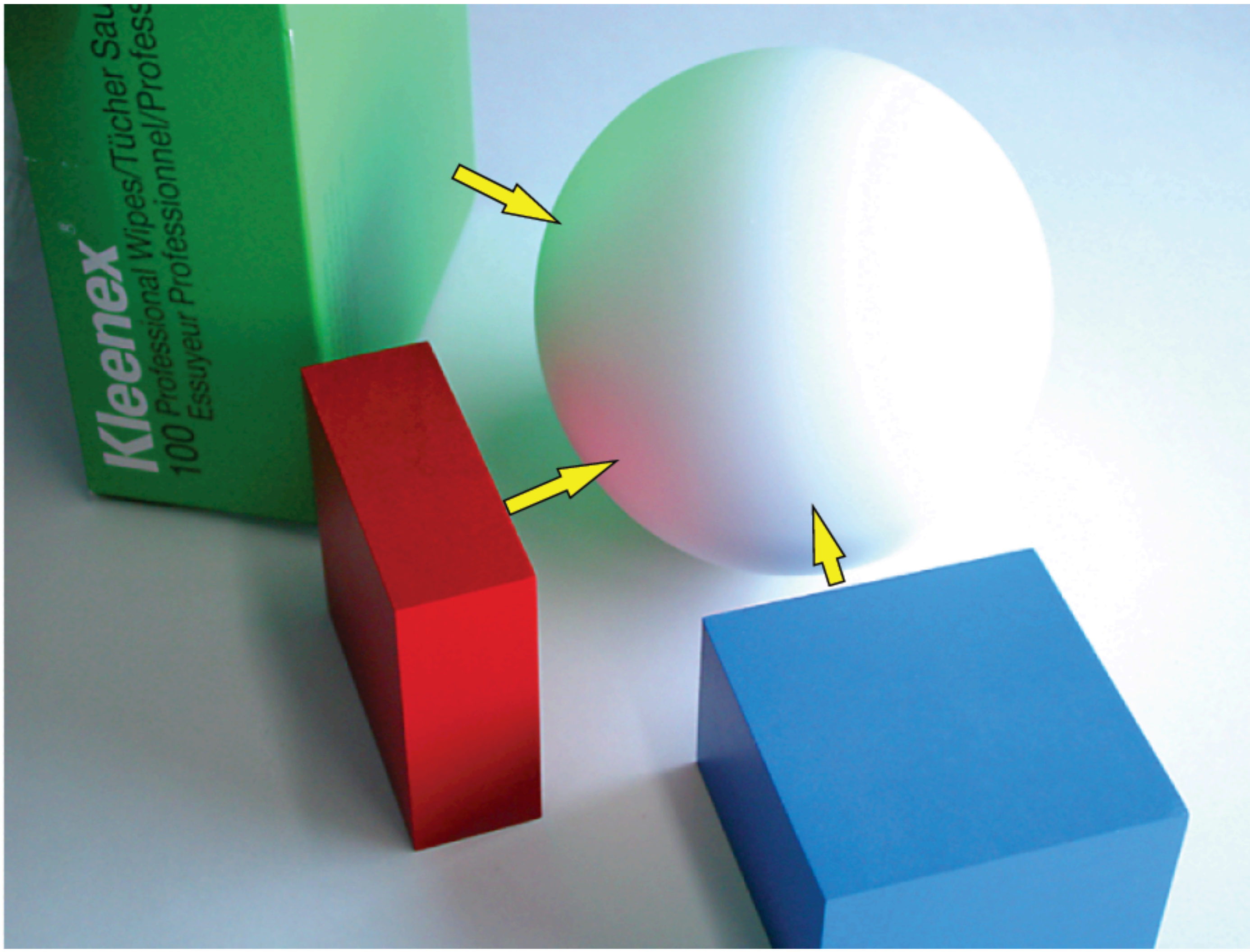
- Recursively expand, as above
  - sample the incoming directions
    - what radiance is coming in?
    - go to far end - what is emitted+reflected?
    - recur
-



Interreflections

- Issue:
  - local shading model is a poor description of physical processes that give rise to images
    - because surfaces reflect light onto one another
  - This is a major nuisance; the distribution of light (in principle) depends on the configuration of every radiator; big distant ones are as important as small nearby ones (solid angle)
  - The effects are easy to model
  - It appears to be hard to extract information from these models

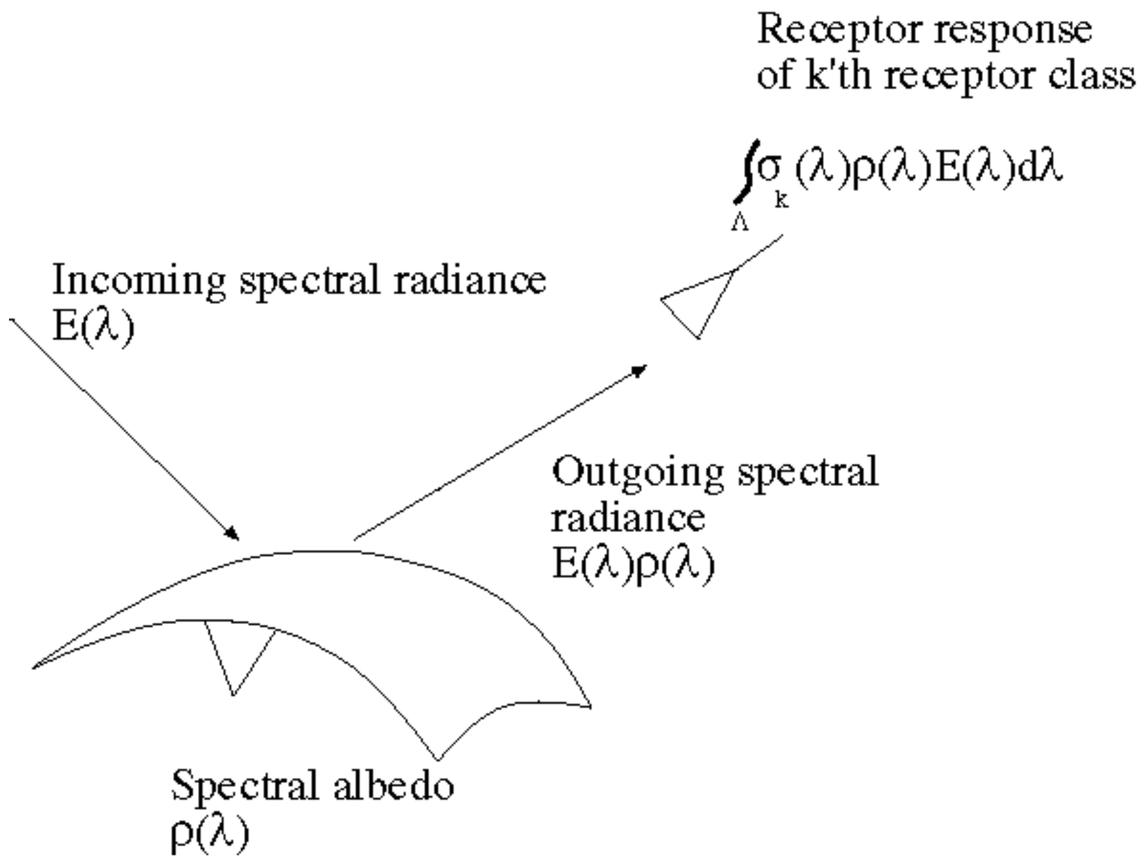
Interreflections



From Koenderink slides on image texture and the flow of light

The color of objects

- Colored light arriving at the camera involves two effects
  - The color of the light source
  - The color of the surface
  - Changes caused by different colored light sources can be large



Constancy

- You perceive objects in terms of their properties
  - rather than what they look like in an image
- Examples:
  - size constancy
    - distant objects are small in pictures, nearby objects bigger
      - but you don't think of them as changing size
  - lightness constancy
    - dark things in bright rooms can be brighter than light objects in dark rooms
      - but you perceive their lightness (=albedo)
  - color constancy
    - image color changes when lighting color changes
      - but you perceive the surface color
  - object constancy

Which fish is bigger?

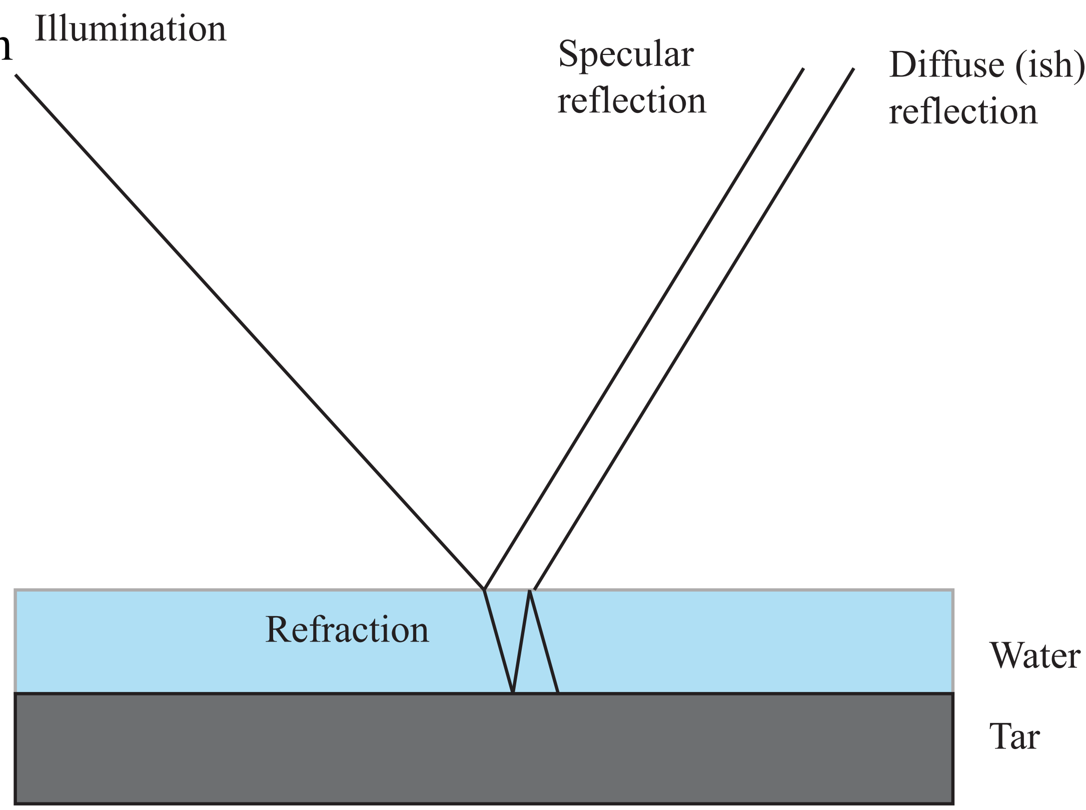
More complicated effects at surfaces

Refraction

- Light striking an interface changes direction
  - between translucent surfaces with different speed-of-light
  - (refraction)
- At critical angle, total internal reflection

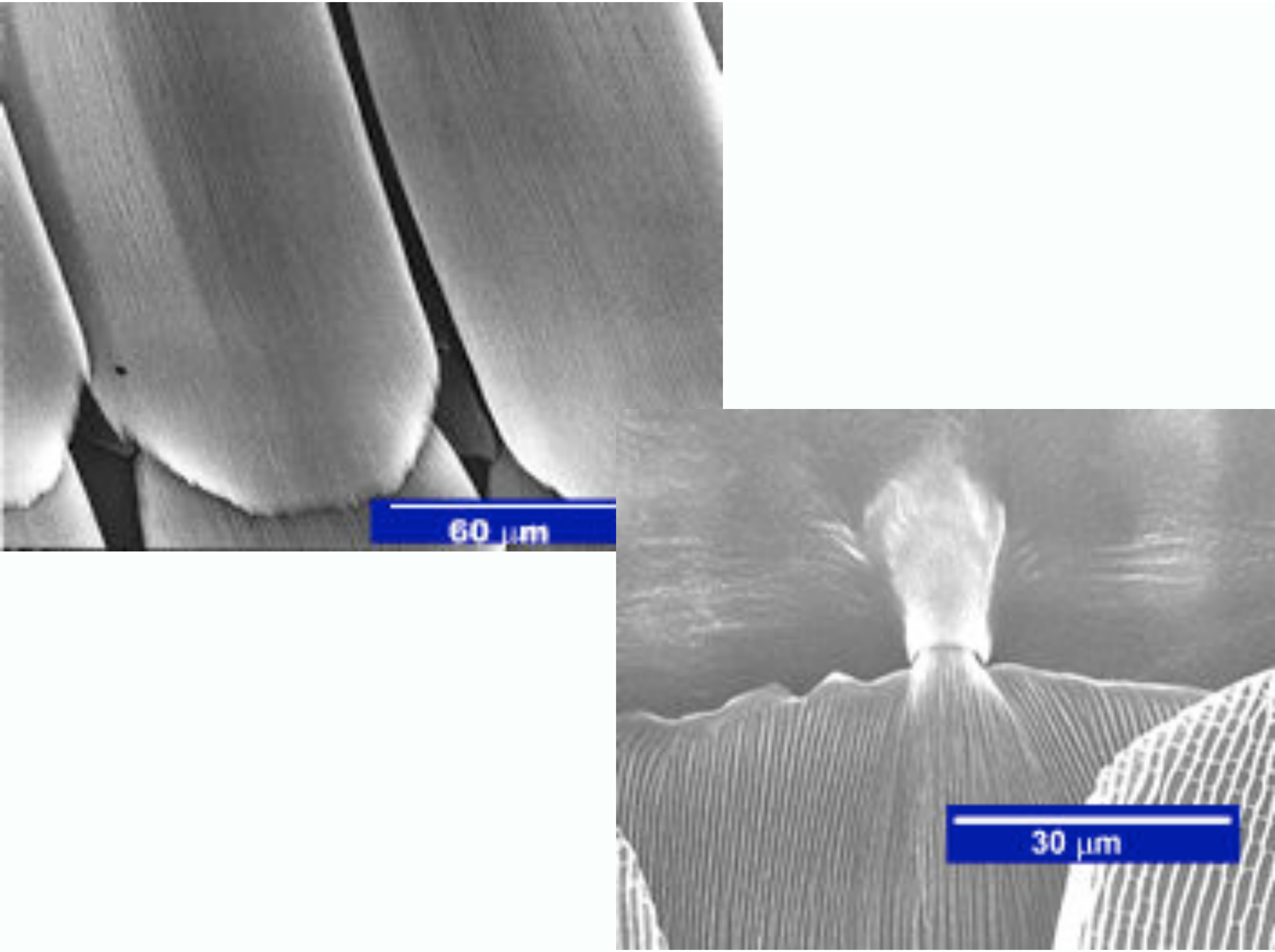
Films on surfaces

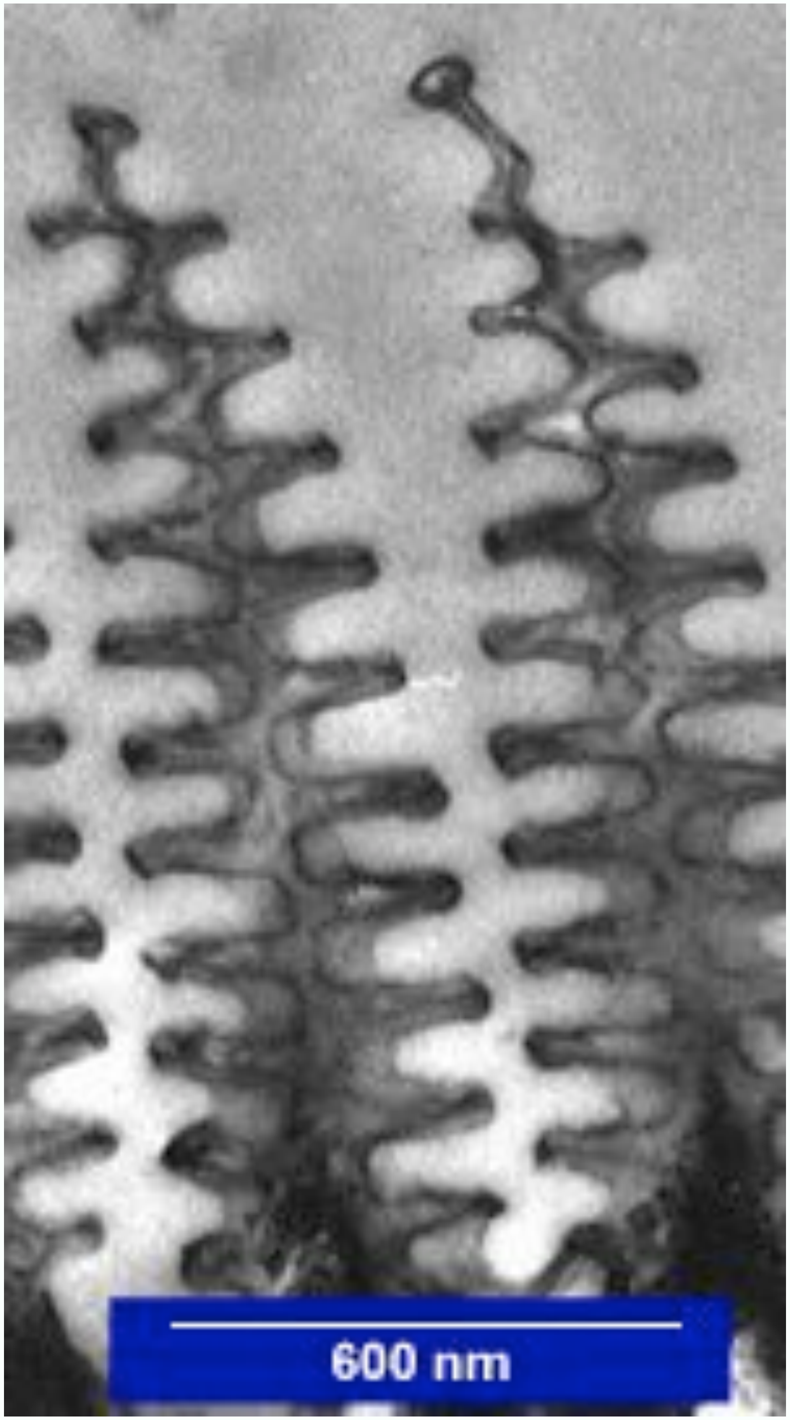
- eg water
- Assume:
  - film is thin
- You see:
  - specular reflection+diffuse term



Interference effects

- Sometimes seen on films
  - if the film is the right number of wavelengths thick
    - waves will interfere destructively (resp constructively)
    - can give rise to intense colors
      - oil films on water often do this


60 μm
30 μm

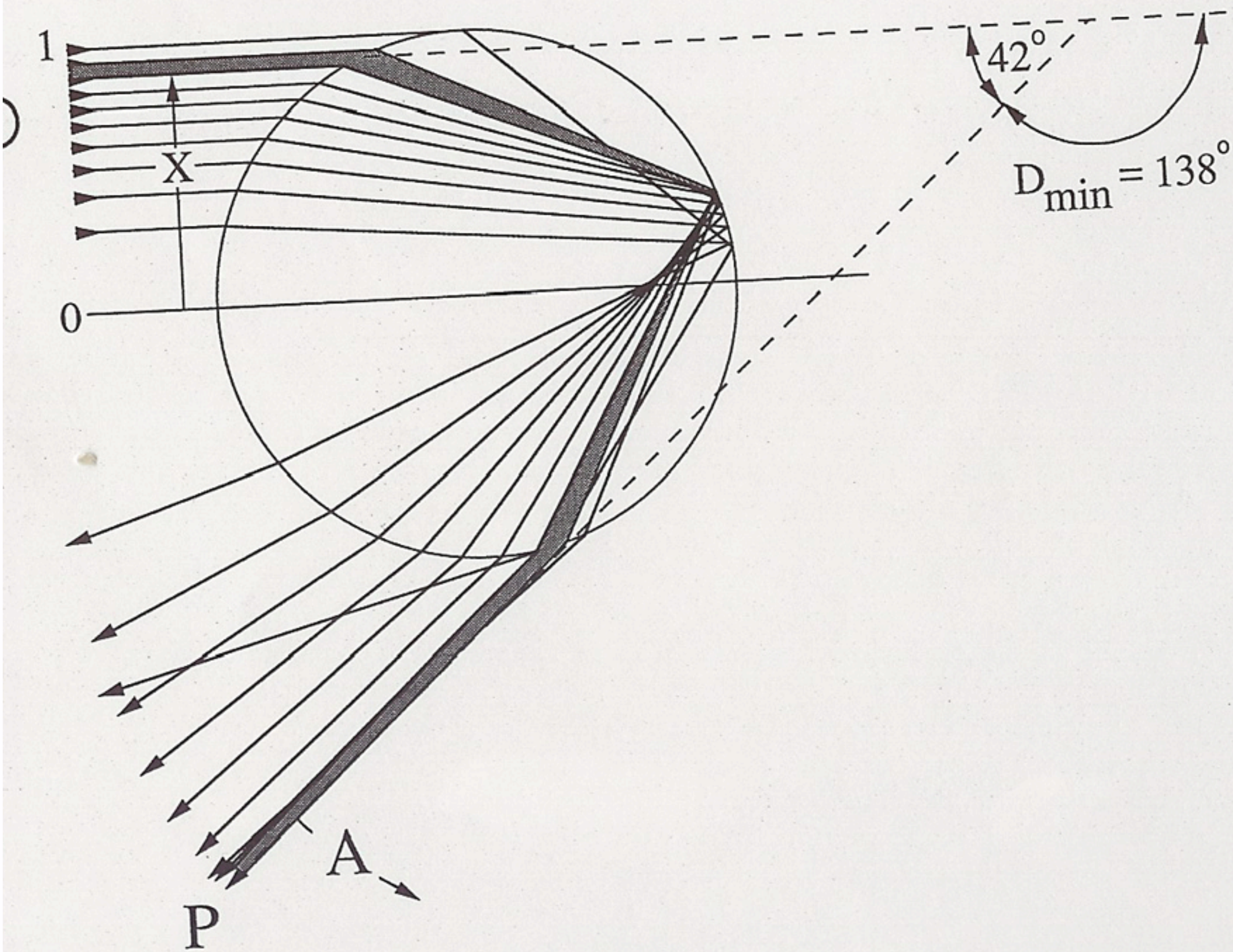

600 nm


Effects in air

Refraction

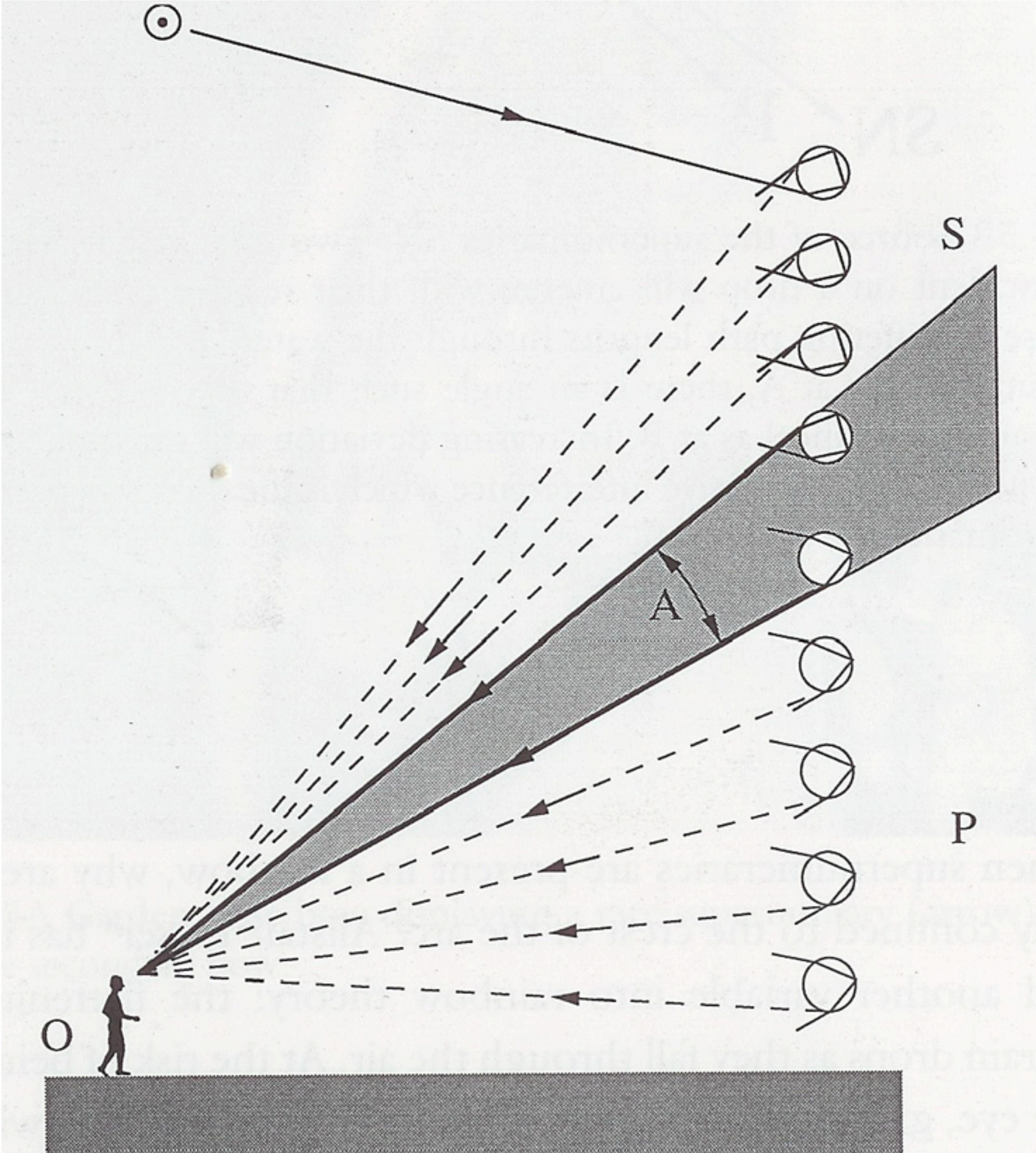
- Light striking an interface changes direction
  - between translucent surfaces with different speed-of-light
  - (refraction)
- At critical angle, total internal reflection

From Lynch and Livingstone, Color and Light in Nature



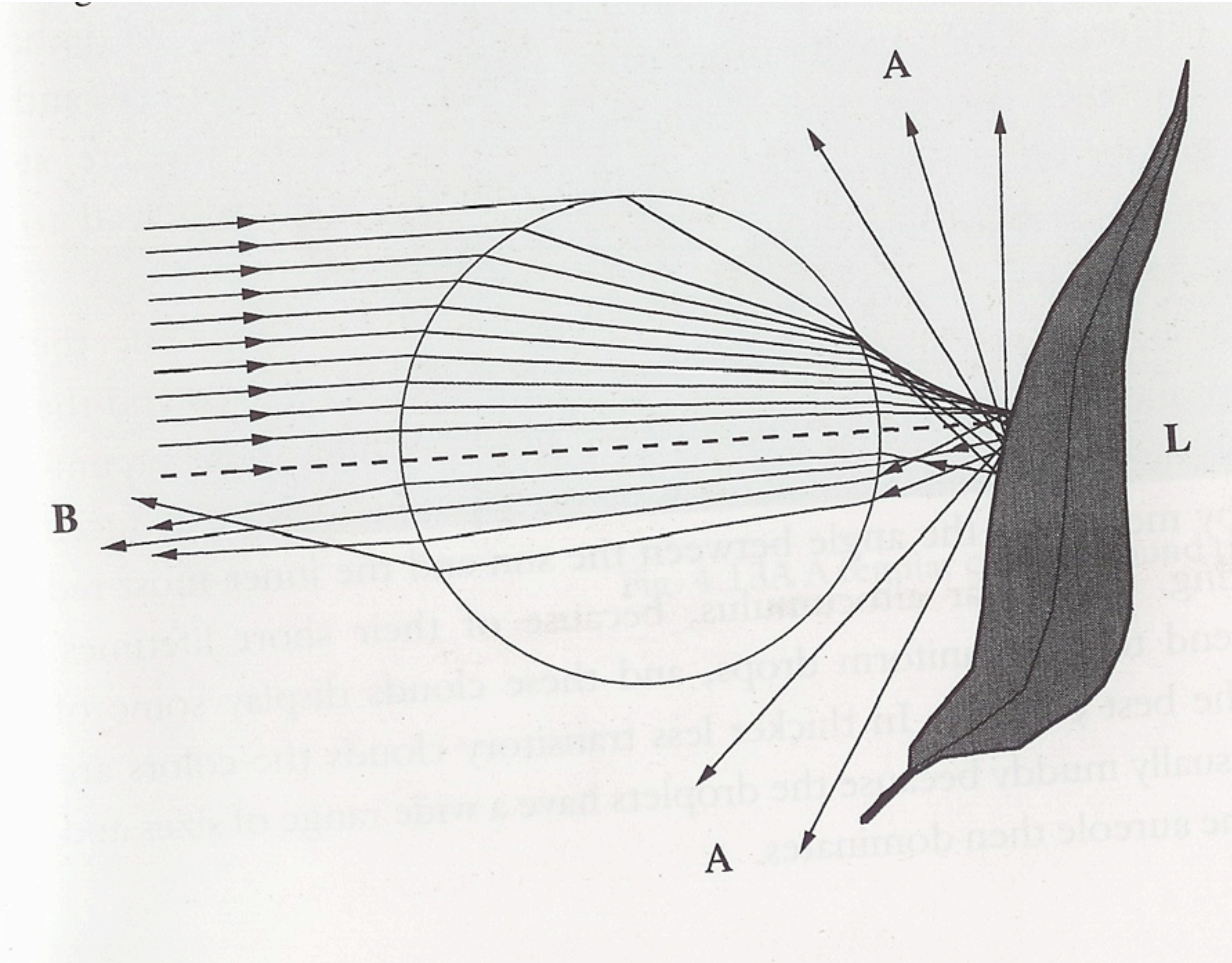
From Lynch and Livingstone, Color and Light in Nature

From Lynch and Livingstone, Color and Light in Nature

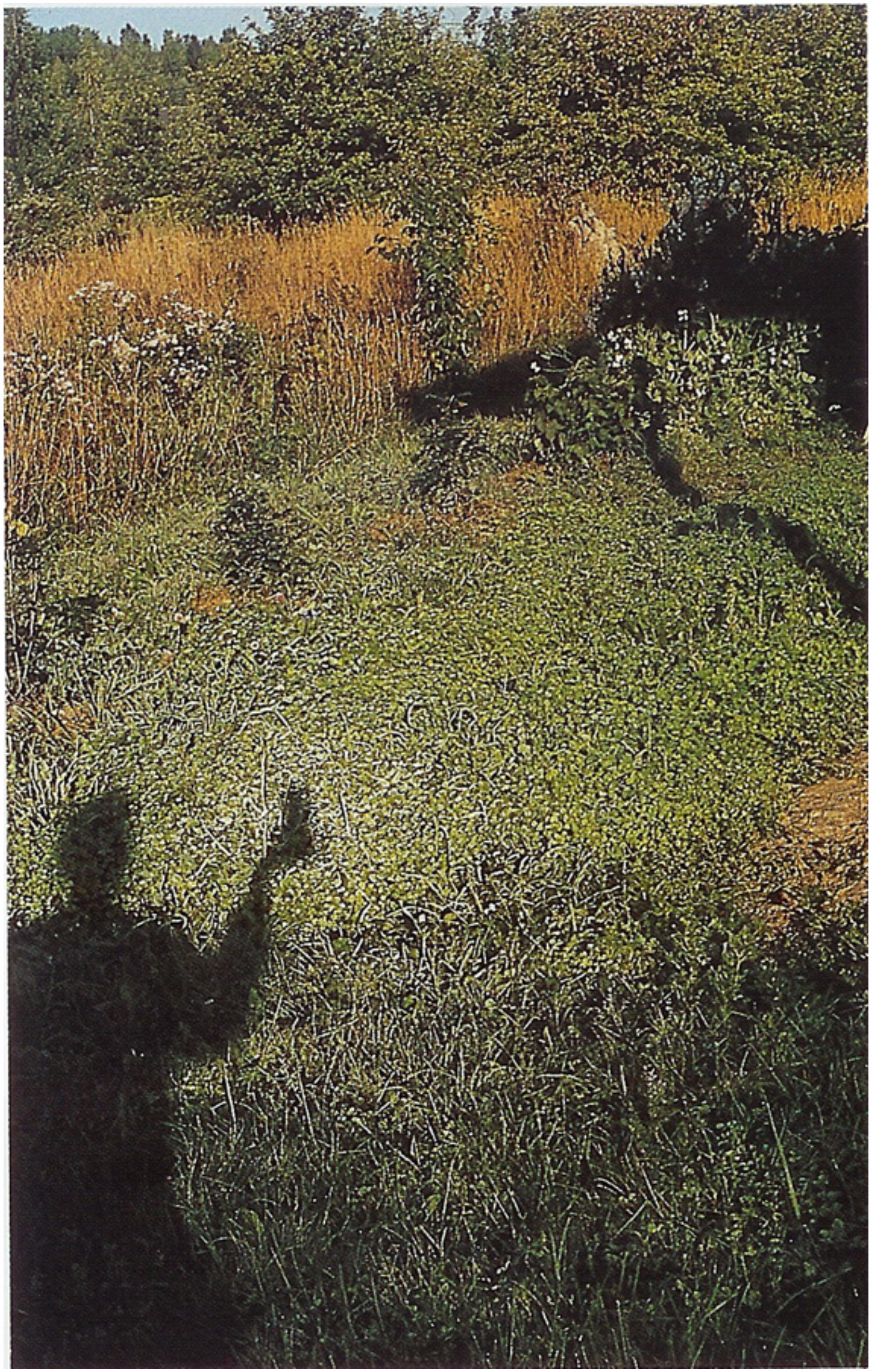


From Lynch and Livingstone, Color and Light in Nature

From Lynch and Livingstone, Color and Light in Nature



From Lynch and Livingstone, Color and Light in Nature



Minnaert, Light and Color in the outdoors
Heiligenschein

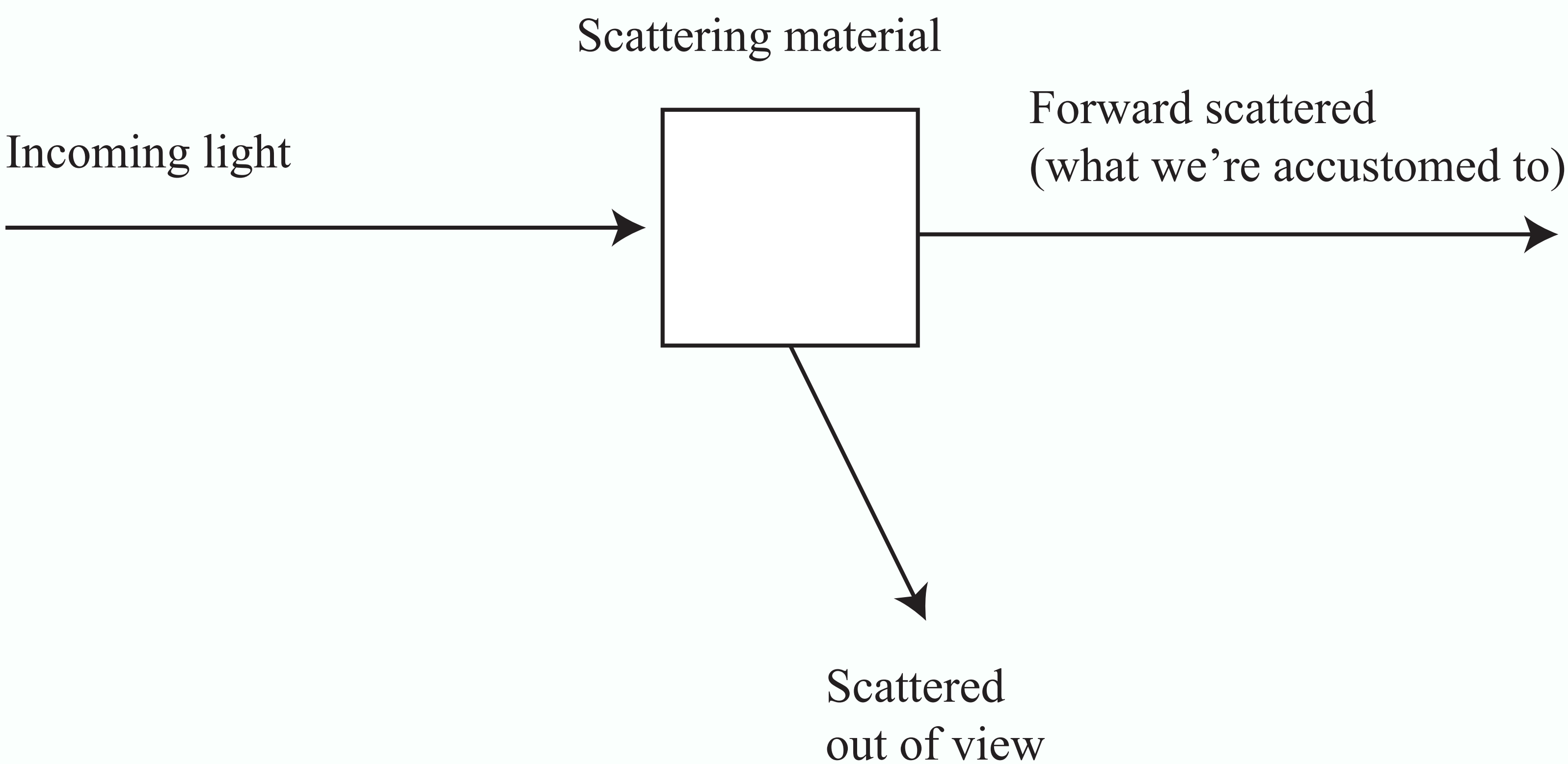
Scattering

- Fundamental mechanism of light/matter interactions
- Visually important for
  - slightly translucent materials (skin, milk, marble, etc.)
  - participating media

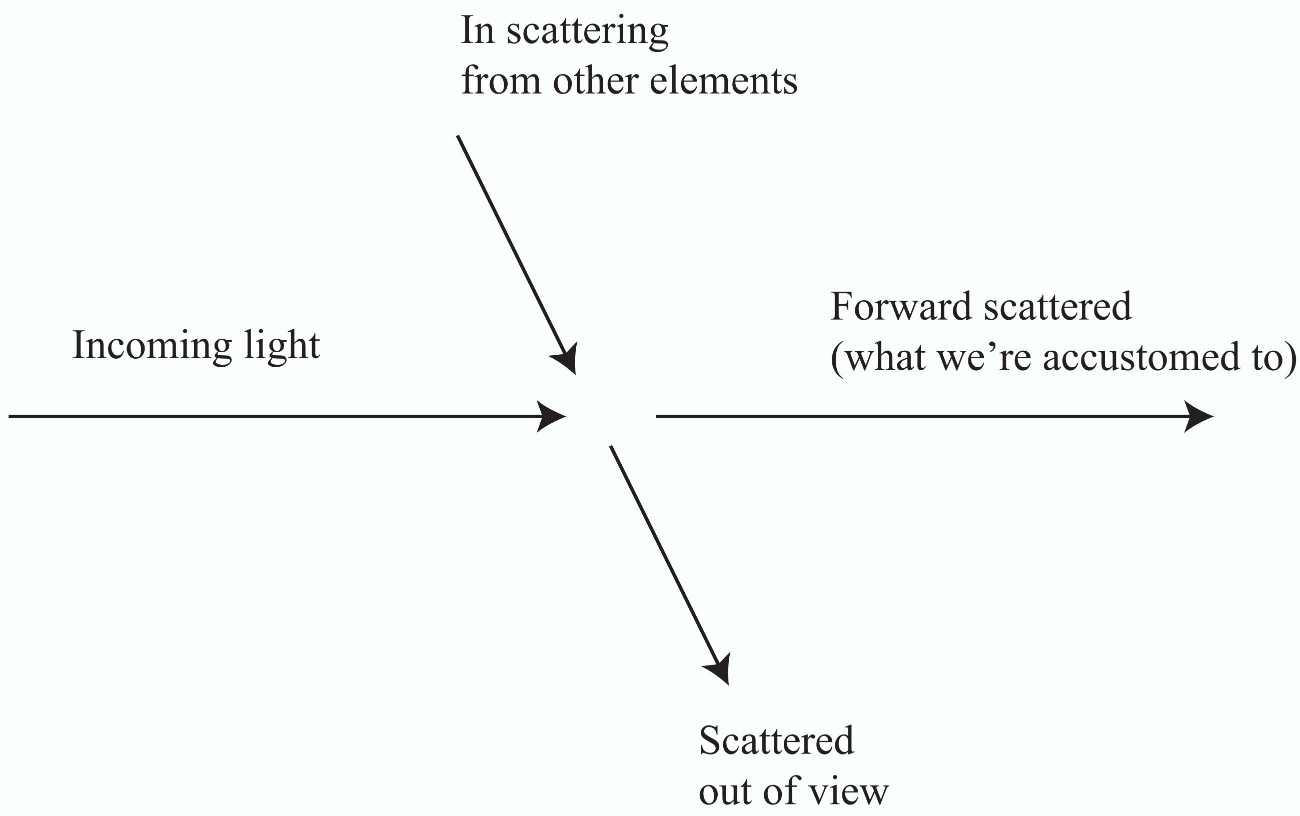
Participating media

- for example,
  - smoke,
  - wet air (mist, fog)
  - rain
  - dusty air
  - air at long scales
- Light leaves/enters a ray travelling through space
  - leaves because it is scattered out
  - enters because it is scattered in
- New visual effects

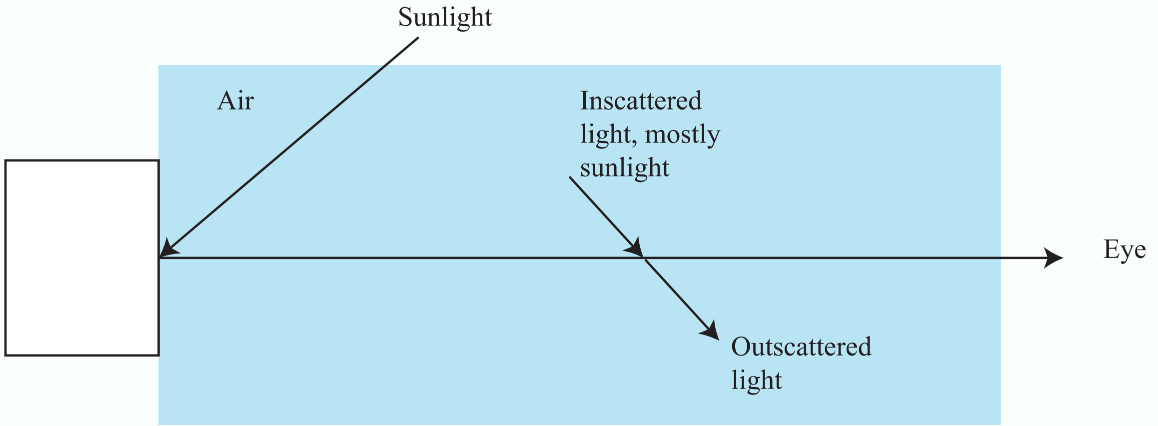
Light hits a small box of material



A ray passing through scattering material

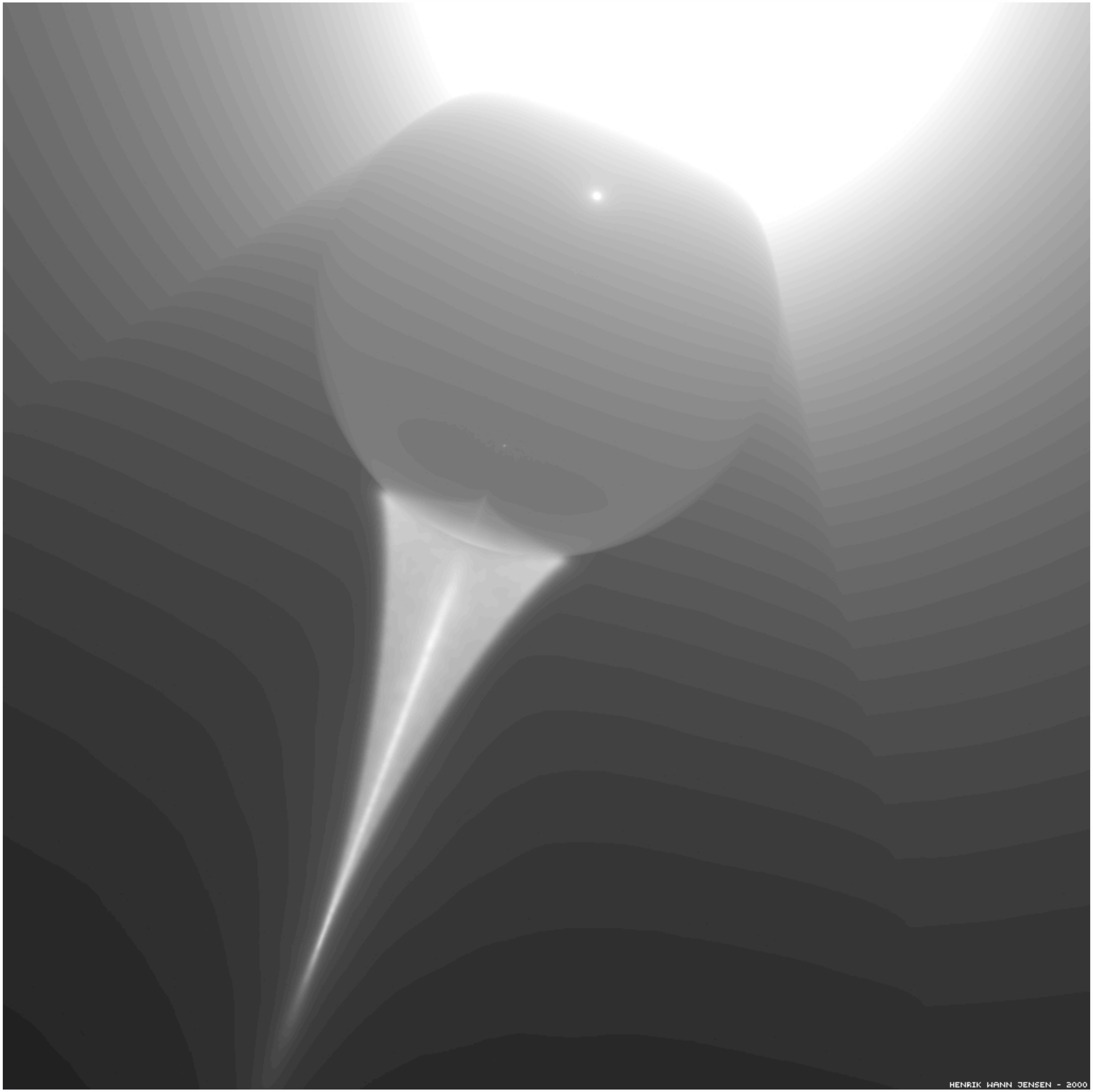


Airlight as a scattering effect



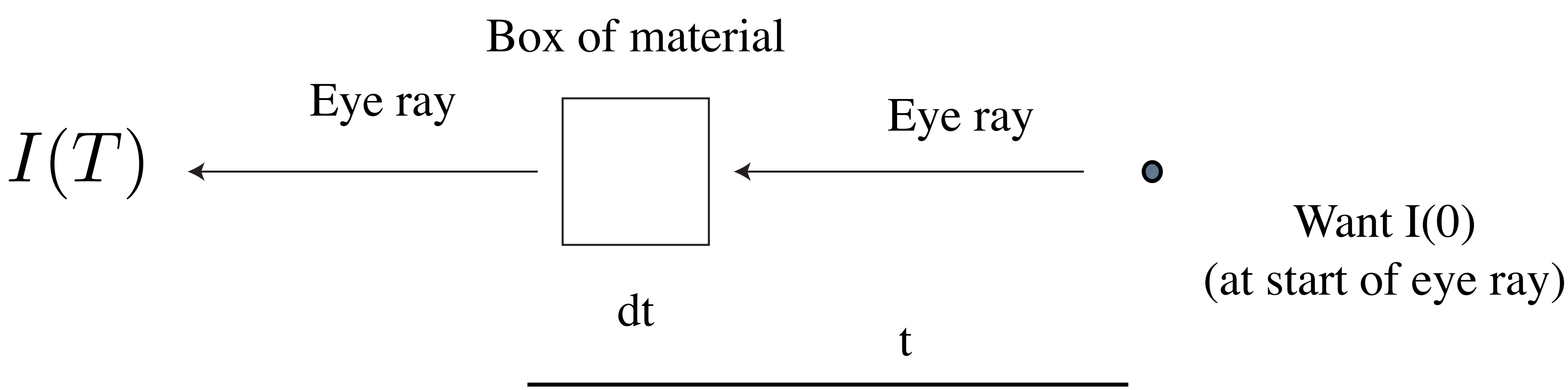

original unique filename: 20180329-141700_baie_des_fourmis.jpg
hosted by www.carto.net
photo ©André M. Winter

From Lynch and Livingstone, Color and Light in Nature

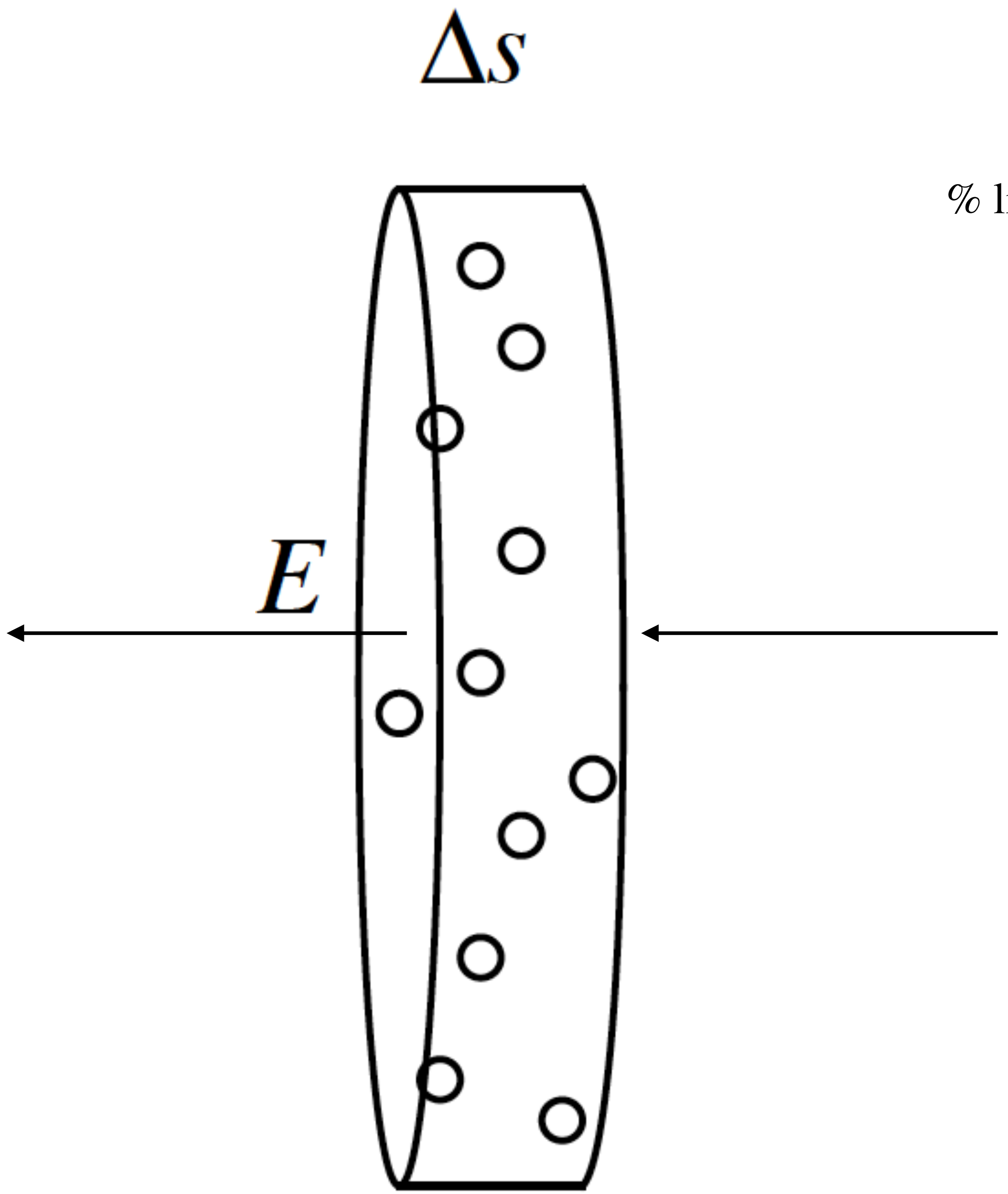

HENRIK WANN JENSEN - 2000

From Lynch and Livingstone, Color and Light in Nature

Absorption



- Ignore in-scattering
  - only account for forward scattering
- Assume there is a source at t=T
  - of intensity I(T)
  - what do we see at t=0?



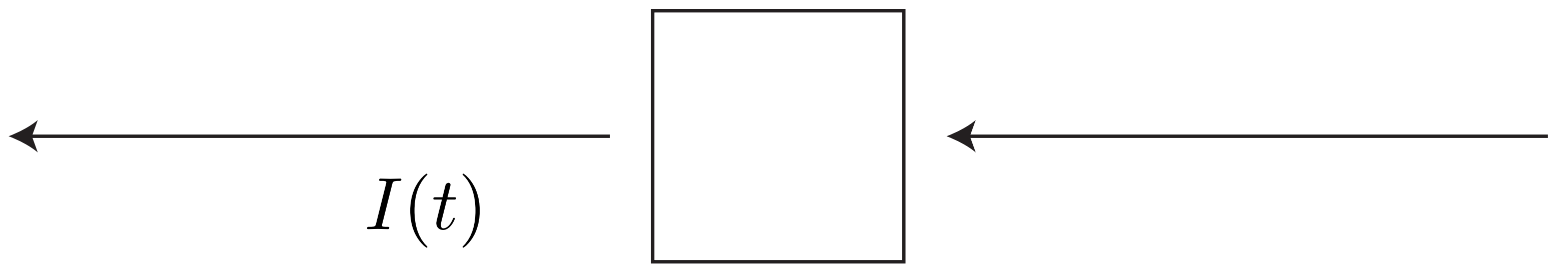
Cross sectional area of “slab” is E
Contains particles, radius r, density rho

Too few to overlap when projected

% light absorbed = (area of projected particles)/(area of slab)

This is:

$$\frac{(\rho E \Delta s)\pi r^2}{E} = \sigma(s)\Delta s$$



$$I(t - \delta t) = I(t) - \sigma(t)I(t)\delta(t)$$

Extinction
coefficient

$$\frac{dI}{dt} = \sigma(t)I(t)$$

$$\frac{d\log I}{dt} = \sigma(t)$$

$$I(T) = I(0)e^{\int_0^T \sigma(t)dt}$$

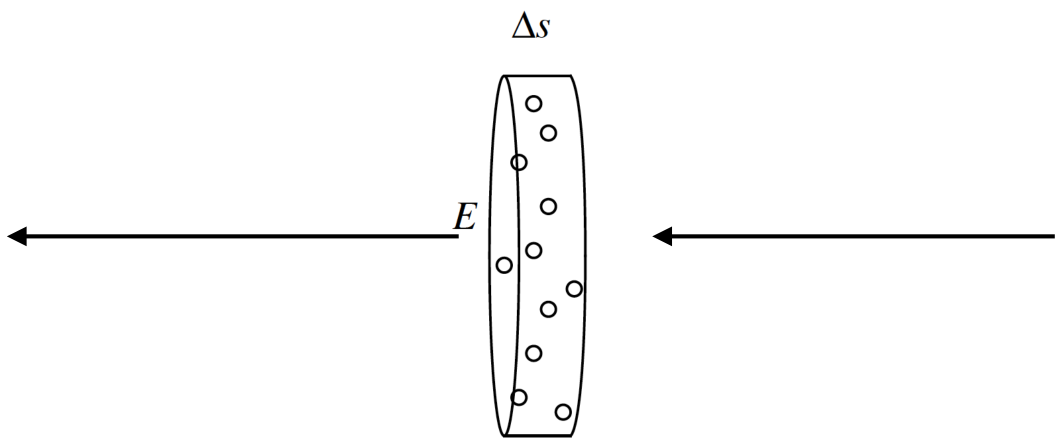
$$I(0) = I(T)e^{-\int_0^T \sigma(t)dt}$$

Eye is at 0

Intensity at T

More interesting…

- Intensity is "created along the ray"
  - by (say) airlight
  - Model - the particles glow with intensity C(x)



Cross sectional area of "slab" is E
Contains particles, radius r, density rho

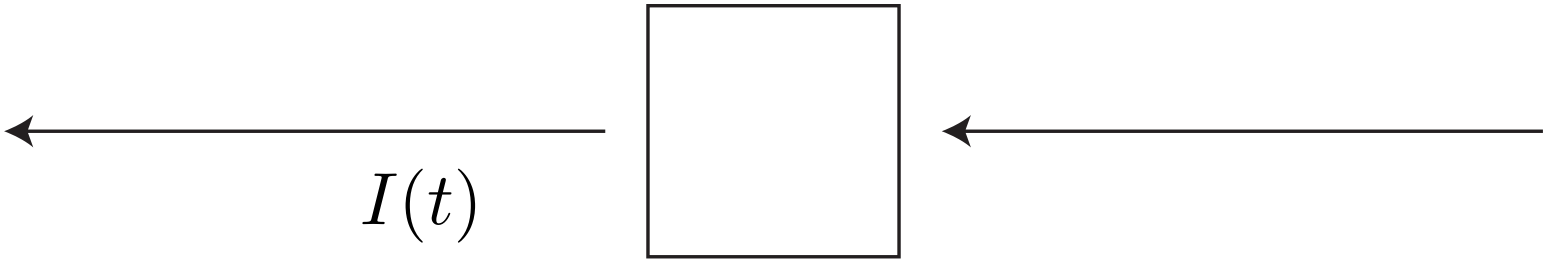
Too few to overlap when projected

Light out = Light in -
Light absorbed+
Light generated

Light generated: C x (area fraction
of proj. particles)

which is

$$C(\mathbf{x}(s))\frac{(\rho E \Delta s)\pi r^2}{E} = C(\mathbf{x}(s))\sigma(s)\Delta s$$



$$I(t - \delta t) = I(t) - \sigma(t)I(t)\delta t + \mathbf{c}(\mathbf{x}(t))\sigma(t)\delta t$$

Absorption

Generation

$$I(0) = \int_0^T \mathbf{c}(\mathbf{x}(s))\sigma(s)e^{-\int_0^s \sigma(u)du} ds$$

$$I(0) = \int_0^T \underbrace{\mathbf{c}(\mathbf{x}(s))\sigma(s)}_{\text{Made at s}} \underbrace{e^{-\int_0^s \sigma(u)du}}_{\text{Absorbed in transit from s to 0}} ds$$

Accumulate along ray

Dehazing and airlight

Airlight color at p

$$I(p) = J(p) \times T(p) + A(p) \times (1 - T(p))$$

Image color at p

Surface radiance color at p

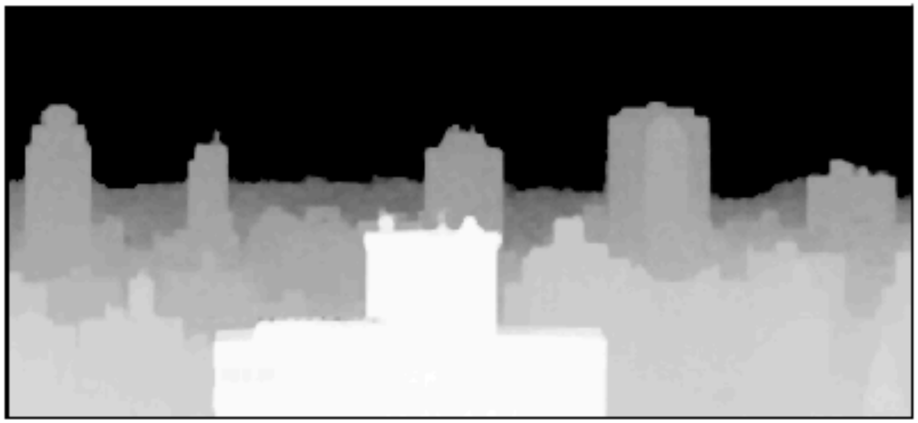
Absorption term, exponential in depth, at p

- Consequences
  - Brightness is a depth cue
  - Reasoning about airlight color yields dehazed image

Airlight yields a depth cue

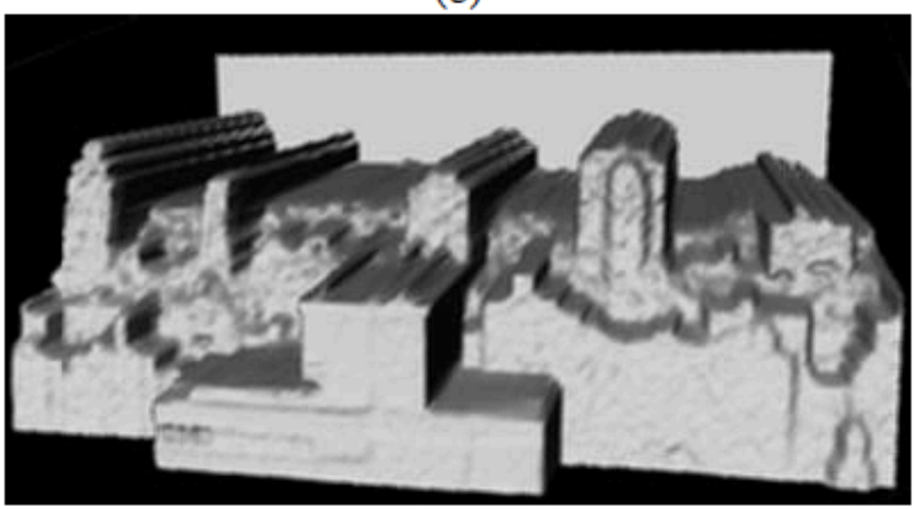
- Assume that airlight is dominant
  - (i.e. most of light arriving at camera is airlight)
  - then you can recover depth from a single image
- Disadvantages
  - requires significant fog (but not too much) or large scales

(a)



(b)

Nayar and Narasimhan, 1999



(c)

Model

Airlight color - same at all points

$$I(p) = J(p) \times T(p) + A(p) \times (1 - T(p))$$

Observed

Shading x albedo

Independent of shading

- With work, this yields
  - neighboring pixels with same albedo yield
    - constraints on shading and T
  - assume shading and T independent
    - estimate A to yield “most independent” shading and T
  - result: J(p)

Figure 1: Dehazing based on a single input image and the corresponding depth estimate.

Fattal, 08 - note depth map AND dehaze; note also slightly odd colors

Improved estimation by cleaner model

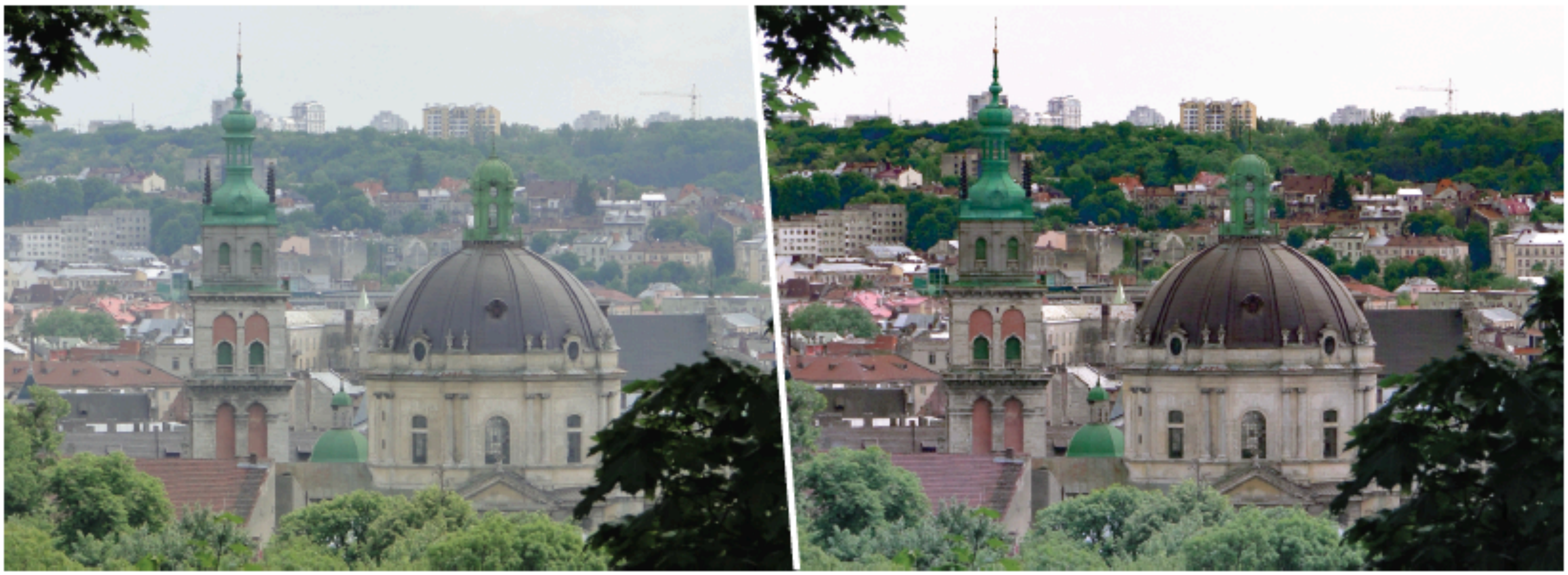


Fig. 1. Old Town of Lviv. Input image on the left, our result on the right.

Fattal, 08 - note depth map AND dehaze; note also slightly odd colors

Simple learning

- Idea:
  - recover transmission map from image
  - you could train by
    - get real images
    - make fake transmission maps, and combine
    - now have (hazy image, transmission map) pairs - train CNN

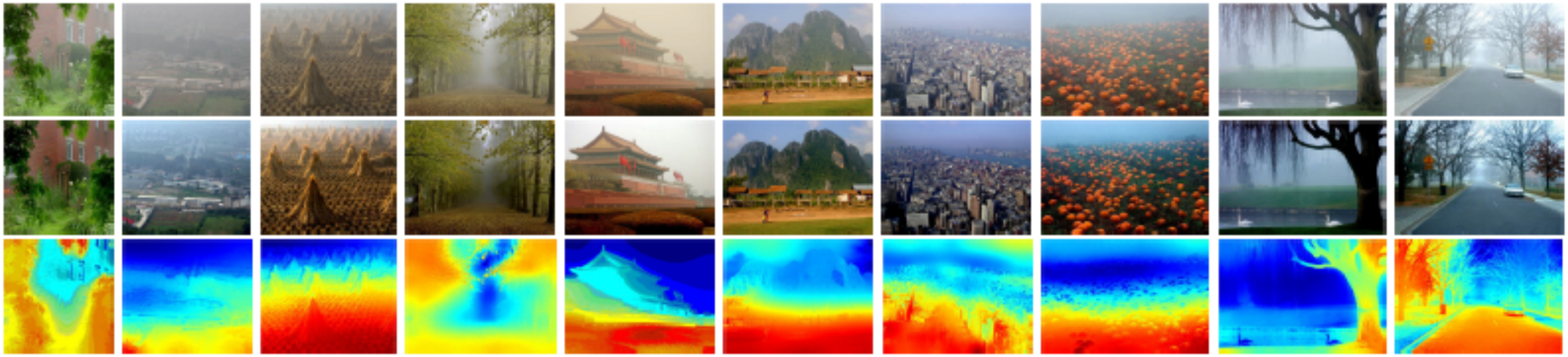


Fig. 11. The haze-free images and depth maps restored by DehazeNet

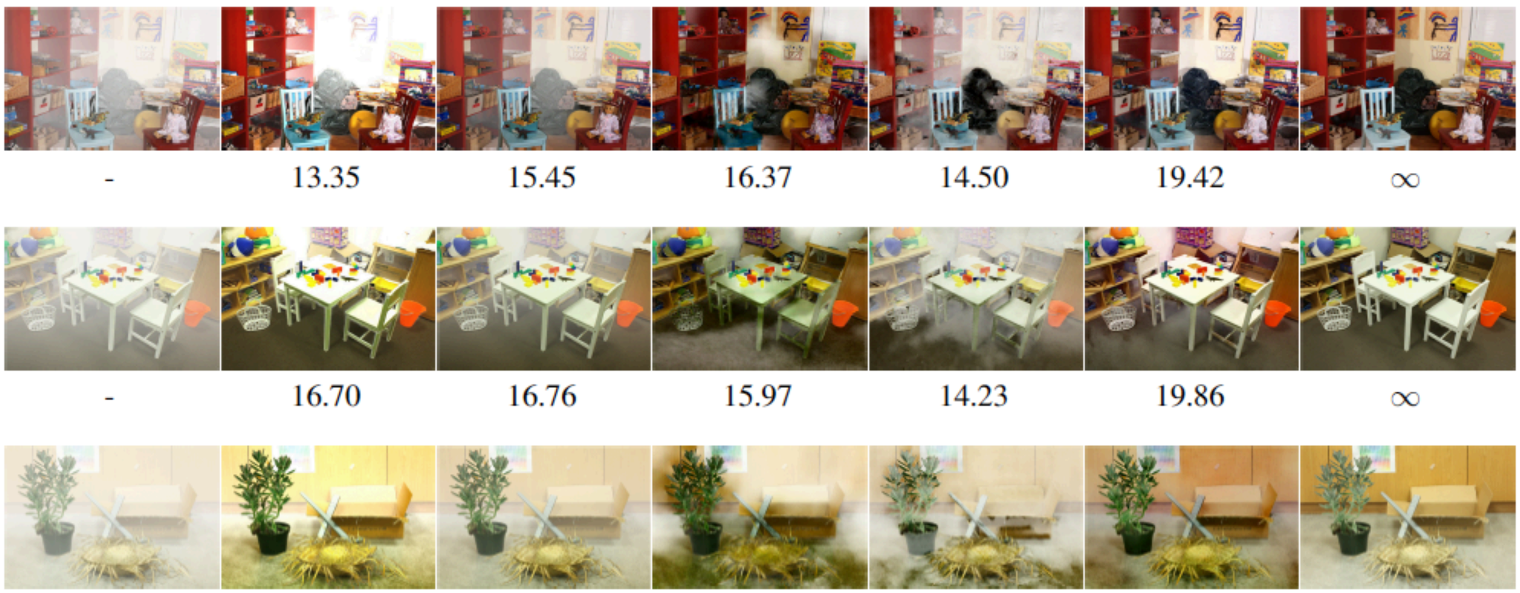
Cai et al 16 (DeHazeNet)

Paired datasets

- Idea:
  - obtain pairs (hazy image, clear image)
- Strategy:
  - Fake fog model on real image
    - Foggy cityscapes
      - https://people.ee.ethz.ch/~csakarid/SFSU_synthetic/
  - Render synthetic images fog/no-fog
    - RESIDE
      - https://arxiv.org/pdf/1712.04143.pdf
  - Take photos outdoors; introduce fog; repeat
    - NH-HAZE
      - https://data.vision.ee.ethz.ch/cvl/ntire20/nh-haze/

Single image dehazing

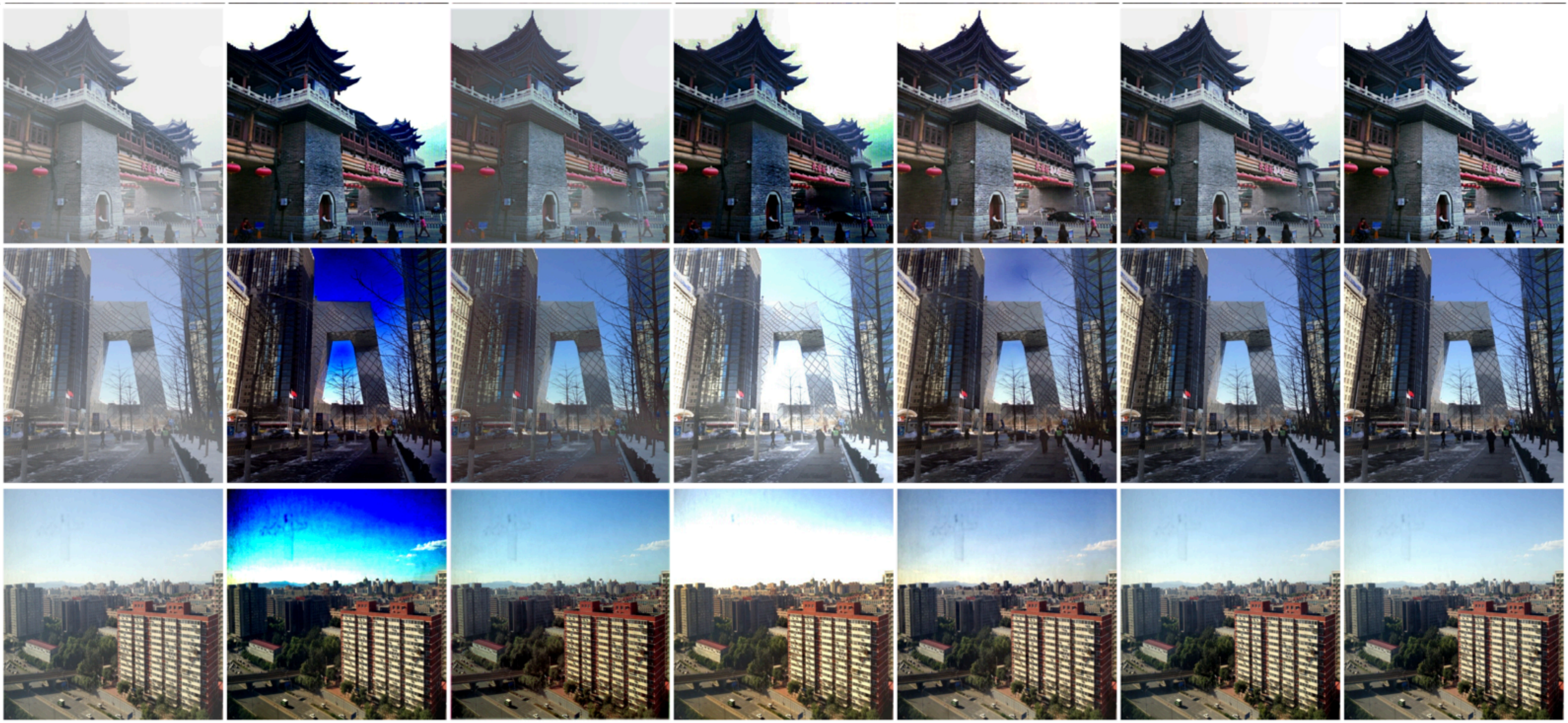
- Essentially
  - obtain images with/without haze (with haze by synthetic)
  - train network to reproduce without haze image from with haze



(a) HAZY	(b) DCP [15]	(c) AOD-Net [20]	(d) GRID-Net [24]	(e) FFA-Net [26]	(f) OURS	(g) GT
-	13.35	15.45	16.37	14.50	19.42	∞
-	16.70	16.76	15.97	14.23	19.86	∞
-	14.64	15.42	11.28	13.27	17.64	∞

Figure 6. Qualitative comparisons with different state-of-the-art dehazing methods for indoor synthesis hazy images. The top two rows are from SOTS, the third row is from TestA dataset and the bottom three rows are from MiddleBury dehazing dataset. The numbers below image are PSNR (dB) value of each image.

Shen et al 19

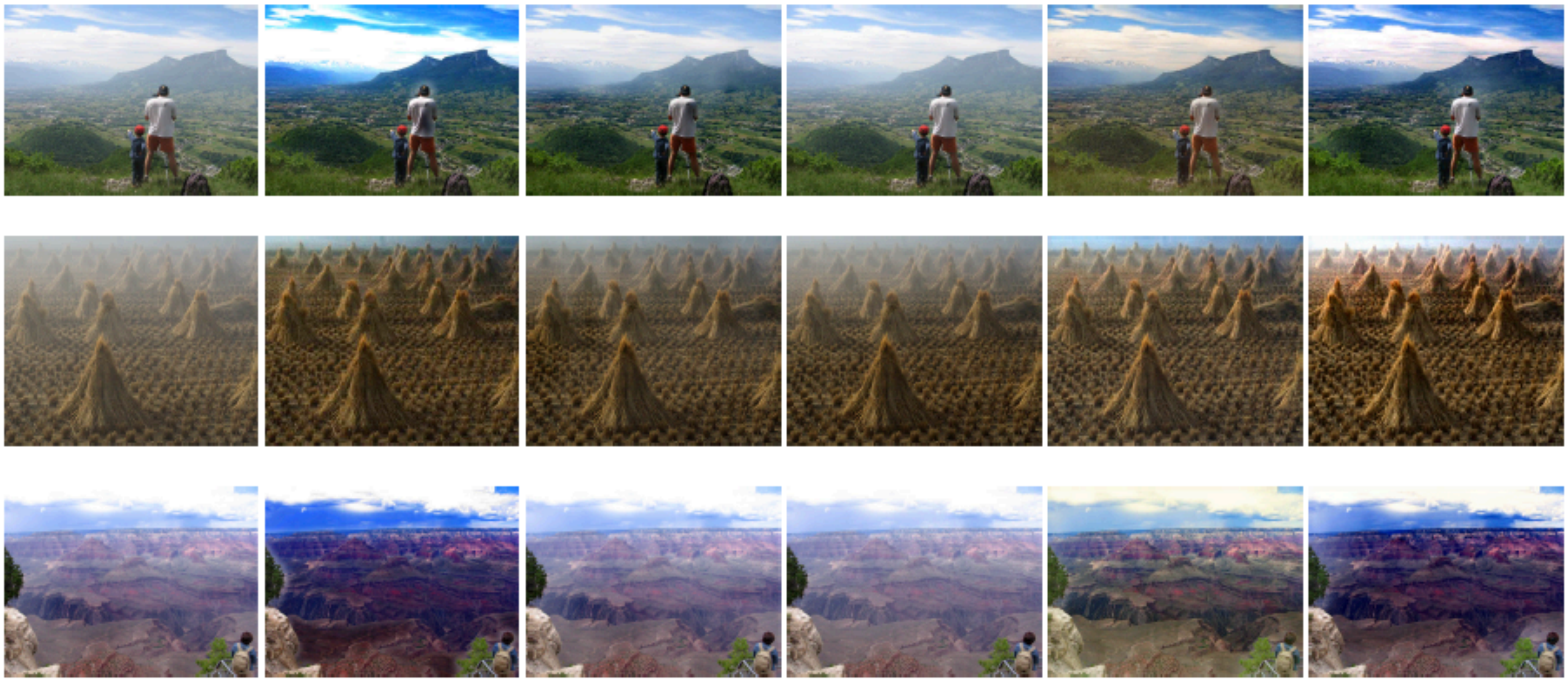


(a)Hazy inputs | (b)DCP | (c)AOD-Net | (d)DehazeNet | (e)GCANet | (f)Ours | (g) GT

Qin et al 19 - Use feature attention

Figure 5: Visual comparisons on real-world hazy images. Our model can generate more natural and visual pleasing dehazed results with less color distortion. Please see the details in red rectangles. Zoom in for best view.

Dong et al 21 - Use an adversarial loss



(a) INPUT (b) DCP [15] (c) GRID-Net [24] (d) FFA-Net [26] (e) FD-GAN [11] (f) OURS

Figure 7. Qualitative comparisons with different dehazing state-of-the-art methods for real hazy images.

Shen et al 20 - Use sequence model (resnet as implicit euler method)

Challenges

- NTIRE workshops and challenges
  - https://data.vision.ee.ethz.ch/cvl/ntire21/
  - https://data.vision.ee.ethz.ch/cvl/ntire20/

Haze and image interpretation

- Dehaze (derain; denighttime; etc) image, then apply
  - detector, segmenter, etc
  - issues:
    - dehazing (etc.) may create signal problems
- Simulate haze (rain; night; etc) existing labelled data, then train
  - detector, segmenter, etc.
  - issues:
    - simulator may not be accurate
- Collect paired data (good conditions/bad conditions)
- Multi-sensor fusion
  - different sensors are affected in different ways, so….

There's a review in Hnewa, 21

Paired data

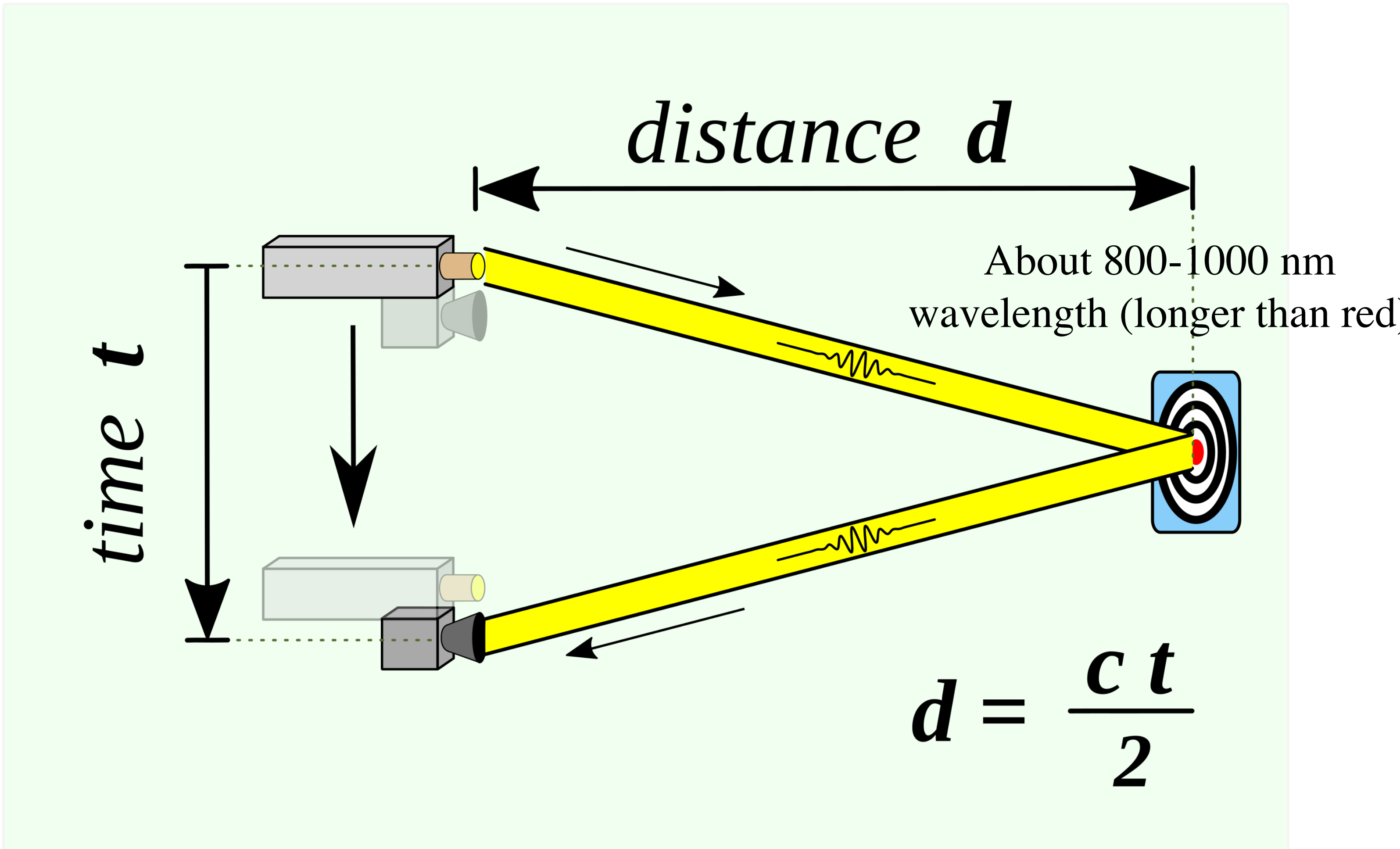
- Collect data on good days, bad days
  - along the same routes, w/ GPS
  - use dynamic programming, GPS to compute alignment at the image level
- Now label
  - annotator labels bad image round 1
  - compares to good image; fixes labelling round 2

(a) Input image I (b) Stage 1 annotation (draft) (c) Corresponding image I' (d) Stage 2 annotation (GT) (e) Invalid mask J

Figure 2. **Illustration of annotation protocol for ACDC.** The color coding of the semantic classes matches Fig. 1. All annotations in (b), (d) and (e) pertain to the input image I in (a). A white color in (b) and (d) denotes unlabeled pixels.

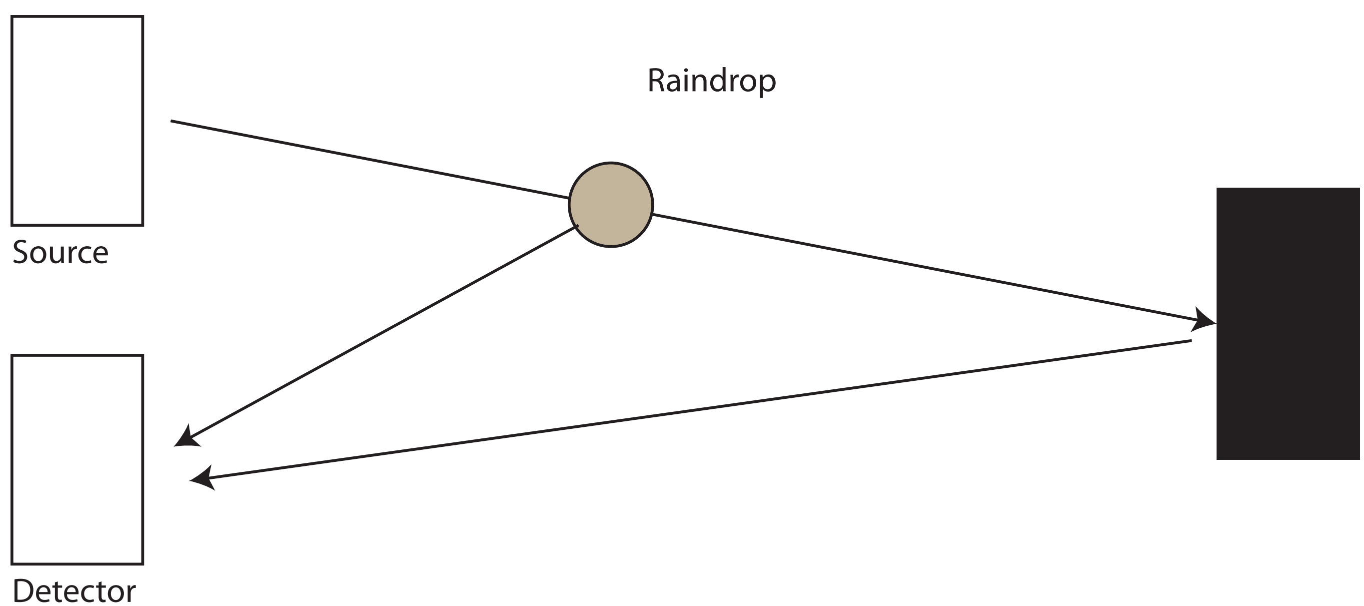
Sakaridis et al, 21

Fog and Lidar: Lidar

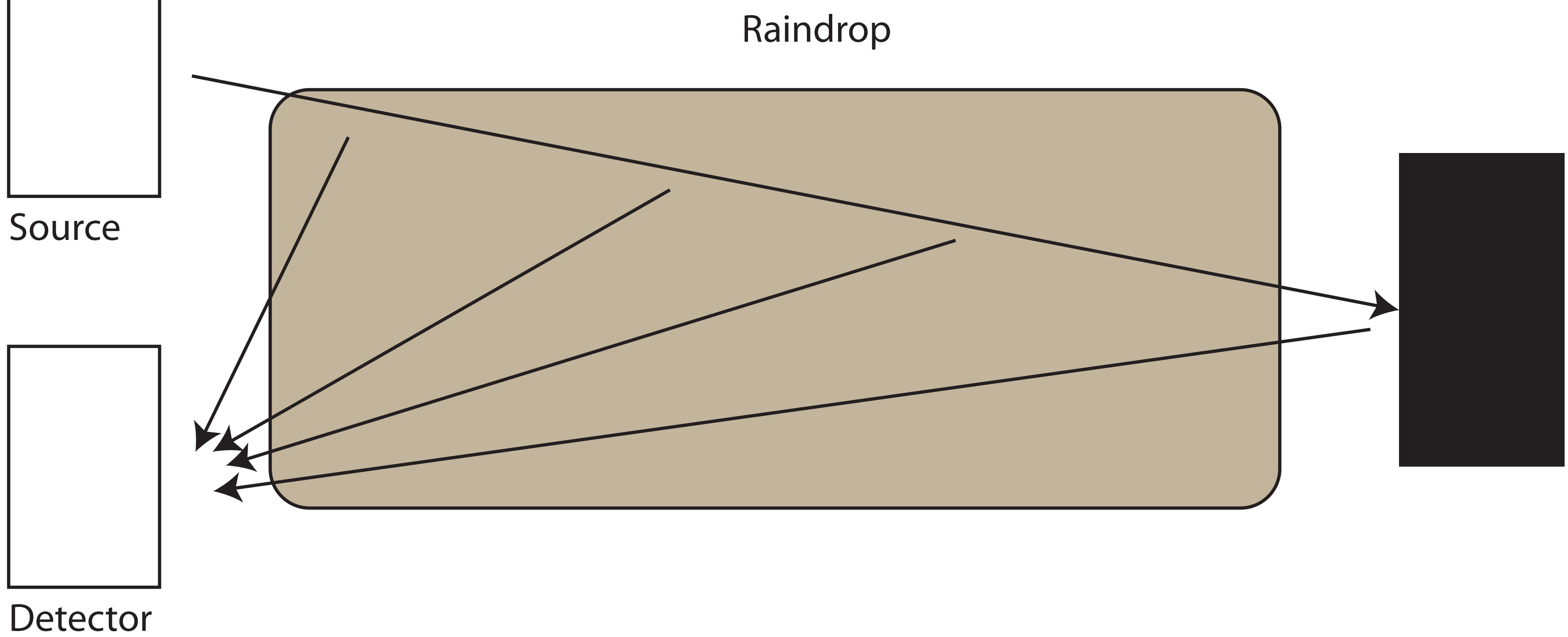


Wikipedia

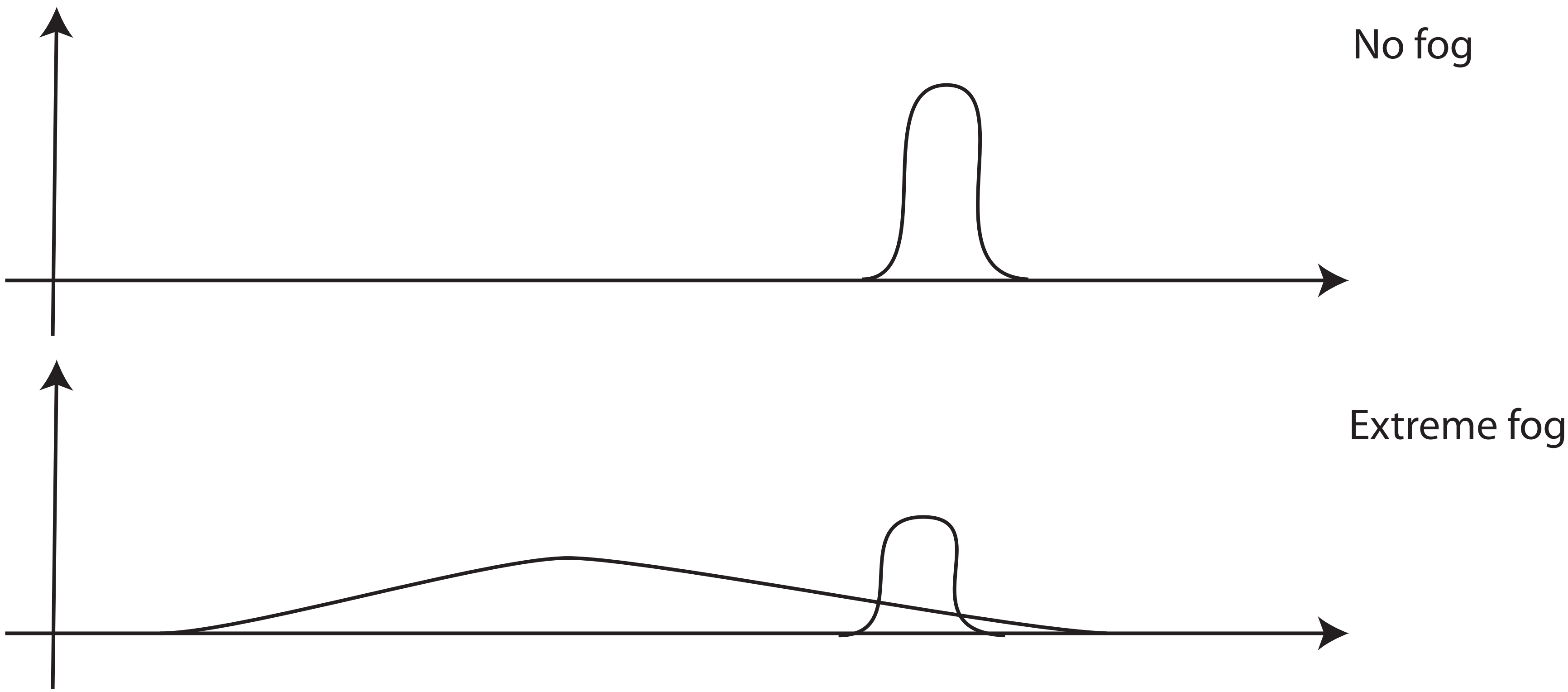
Raindrop backscatter



Fog scattering



What the sensor sees…



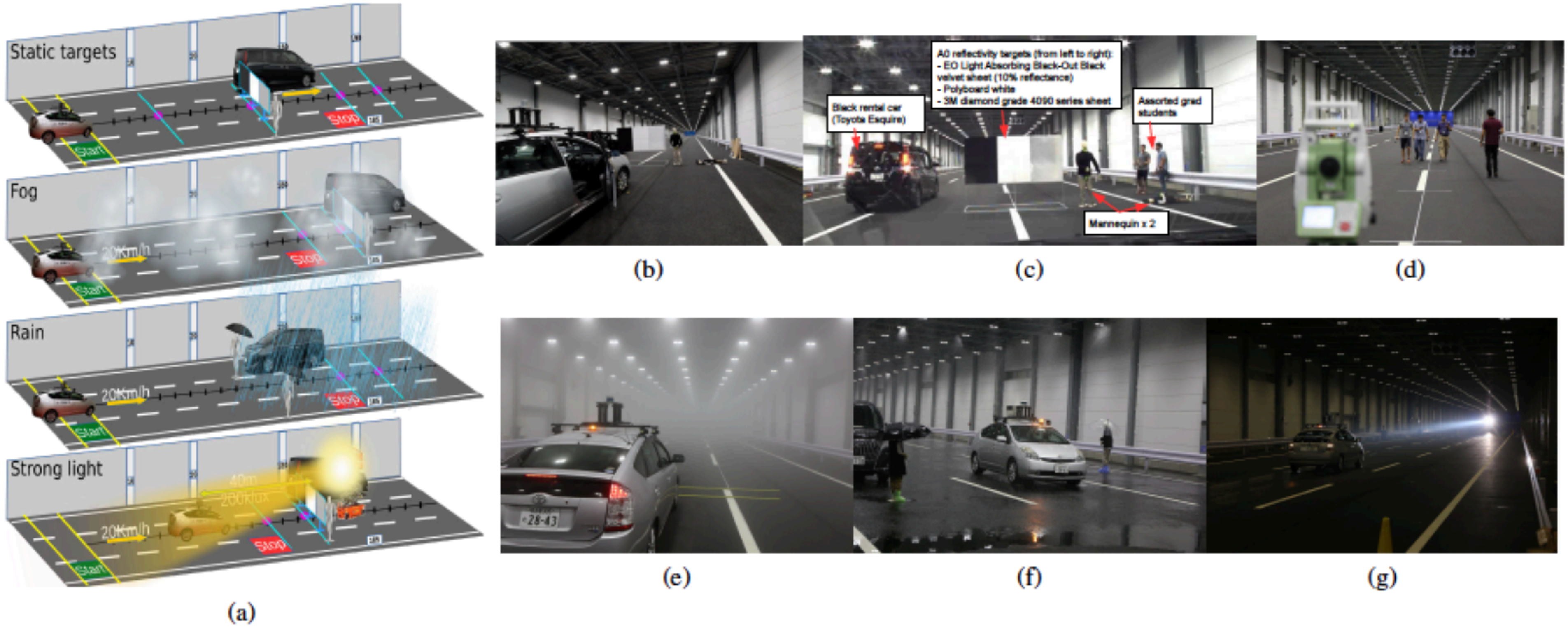


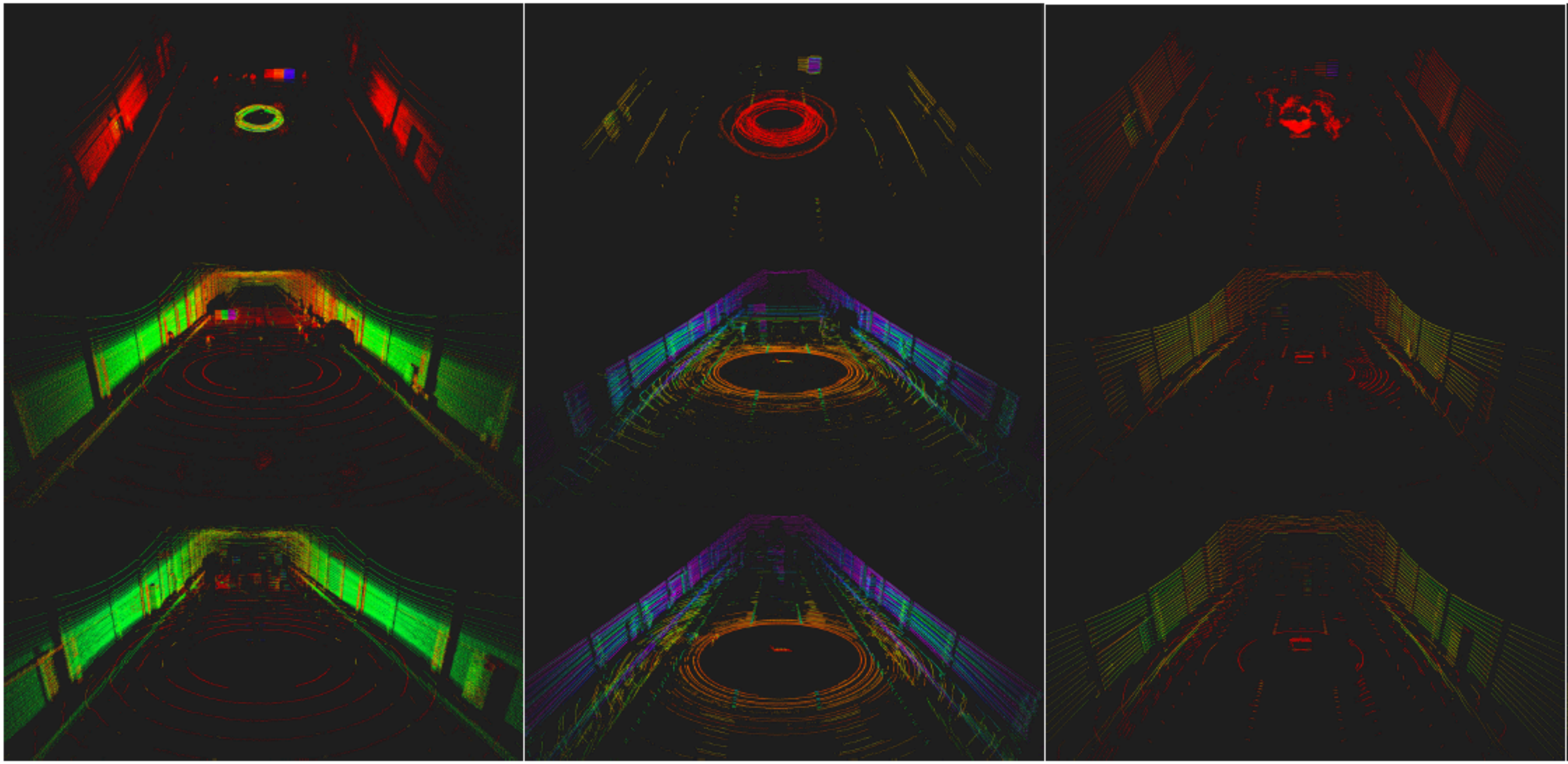
Fig. 5: Static targets and adverse weather experiments at JARI's weather chamber: (a) configuration of the different scenarios, (b) and (c) measurement, (e) to (g) sample adverse weather scenes, (d) setting up ground truth.

Carballo, 20

Fog

Rain

Very bright light



(a) VLS-128 (b) HDL-64S2 (c) HDL-32E

Carballo, 20

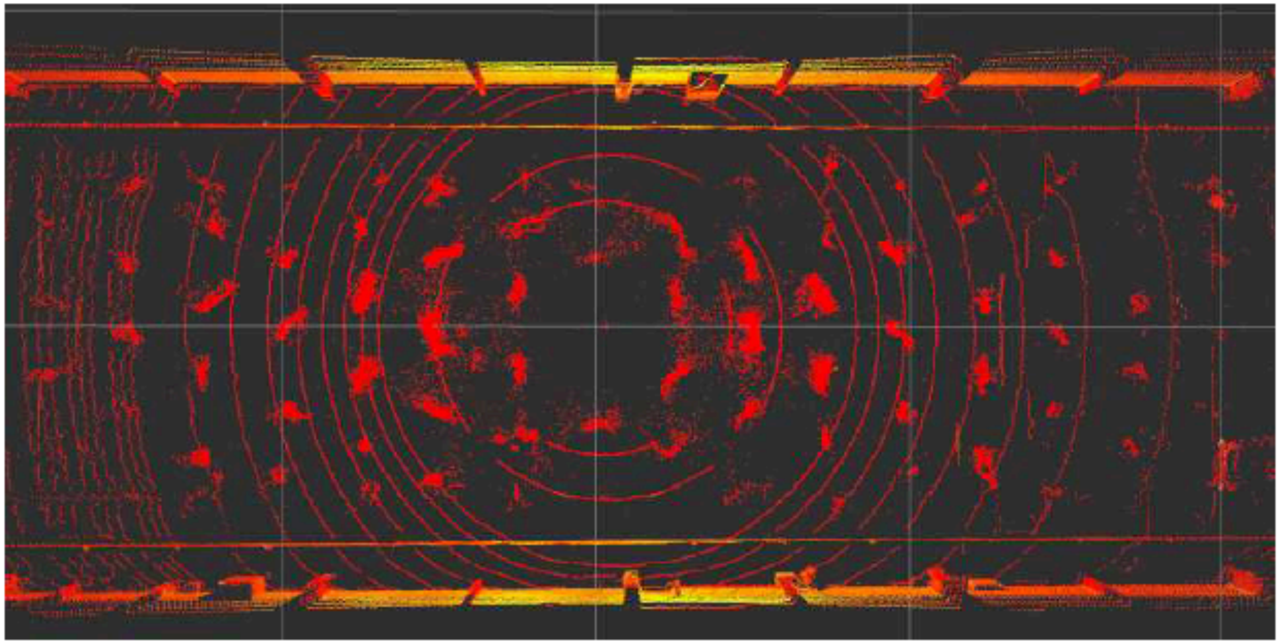
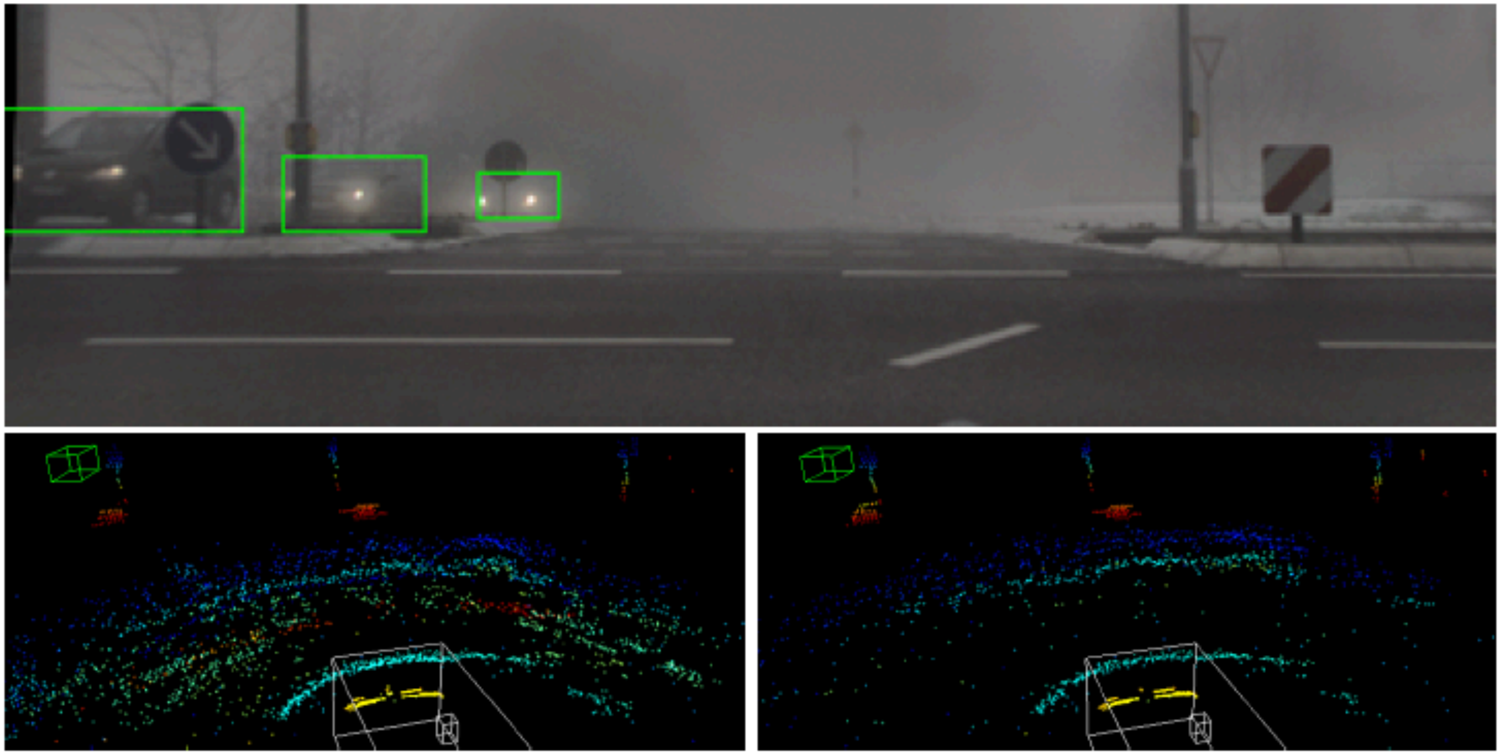


Fig. 9: “Rain pillars” as detected by a LiDAR.

- Qualitative effects
  - lost returns
  - fog torus
  - early returns
  - rain pillars
  - noise

Carballo, 20



(a) *strongest* returns (b) *last* returns

Figure 1: LiDAR returns caused by fog in the (top) scene. (a) shows the *strongest* returns and (b) the *last* returns, color coded by the LiDAR *channel*. The returns of the ground are removed for better visibility of the points introduced by fog. Best viewed in color (red $\widehat{=}$ low, cyan $\widehat{=}$ high, 3D bounding box annotation in green, ego vehicle dimensions in gray).

Hahner 21

Radar is unaffected

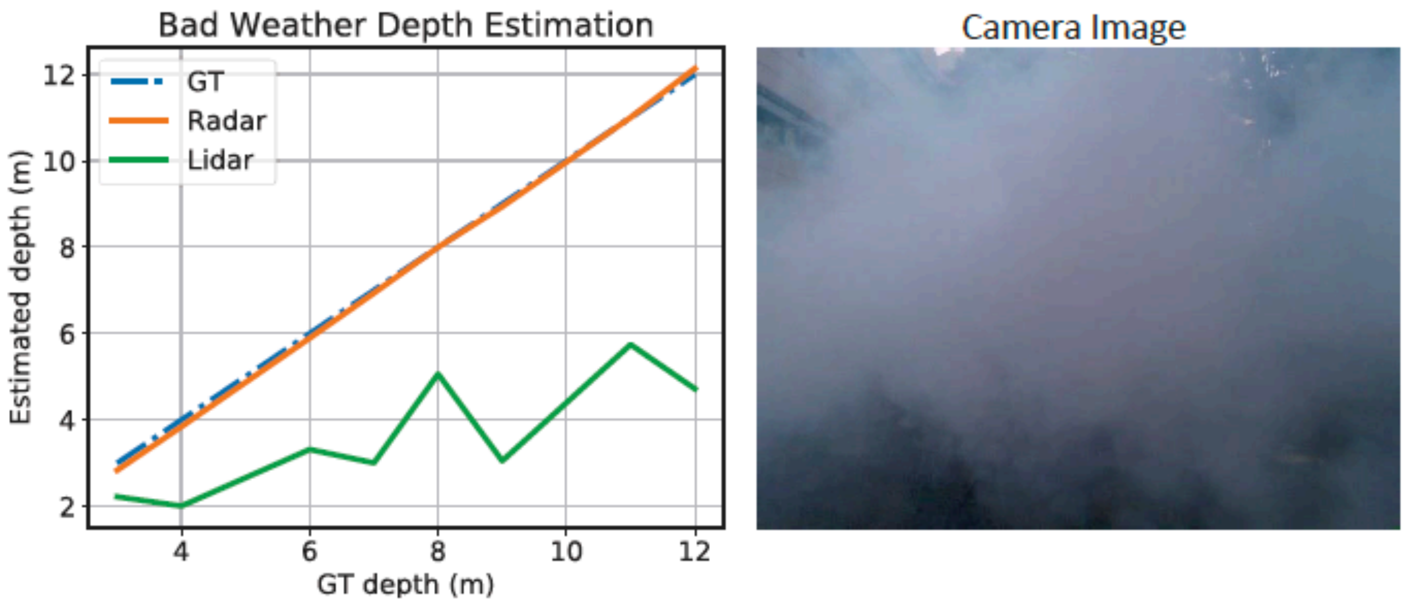
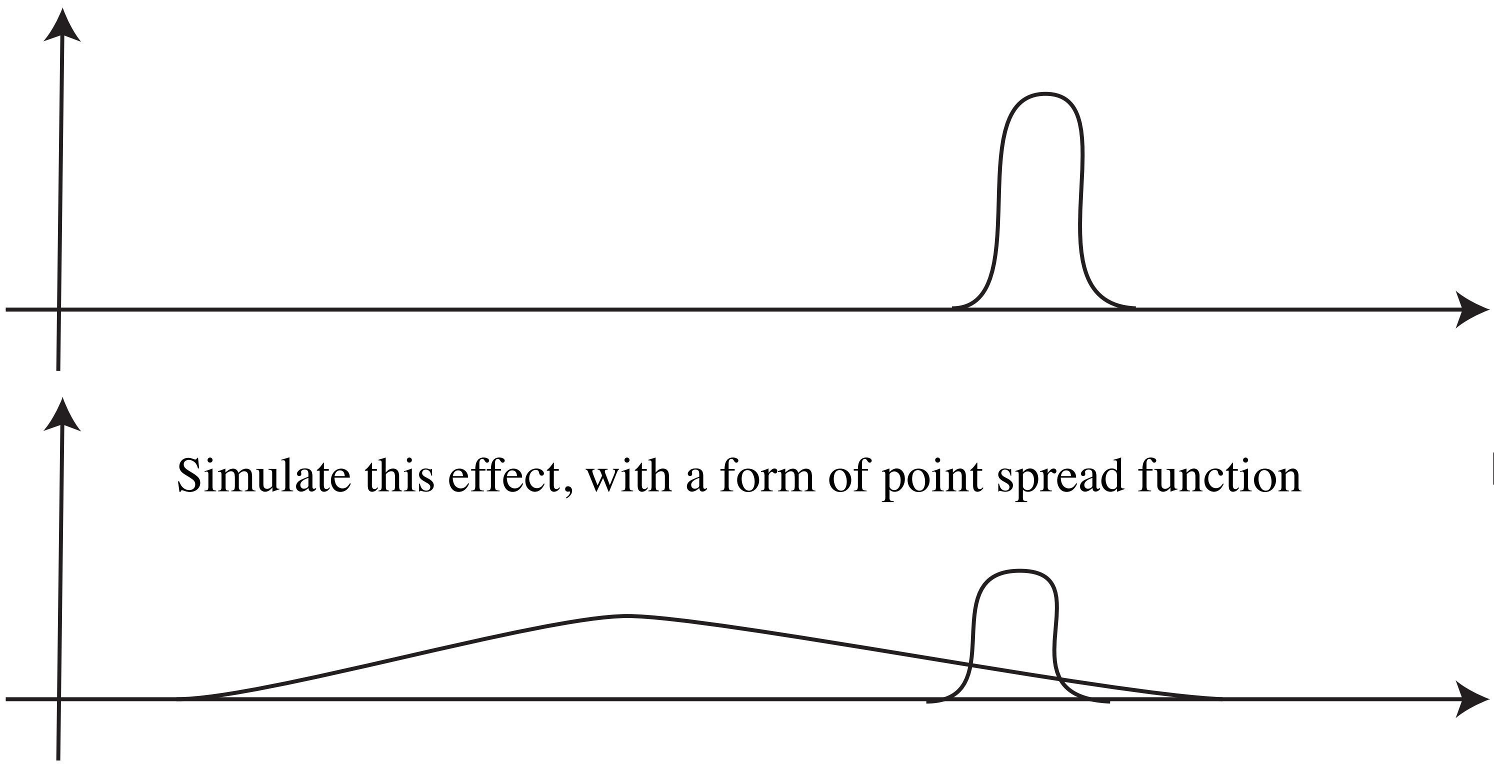


Figure 16: *Performance comparison of different sensors in the presence of adverse conditions. The left plot shows the depth estimation performance of Radar and LiDAR for an object directly in front of the sensor in the presence of fog. The right figure shows the camera image for the experiment.*

Bansal et al 20

What the sensor sees…

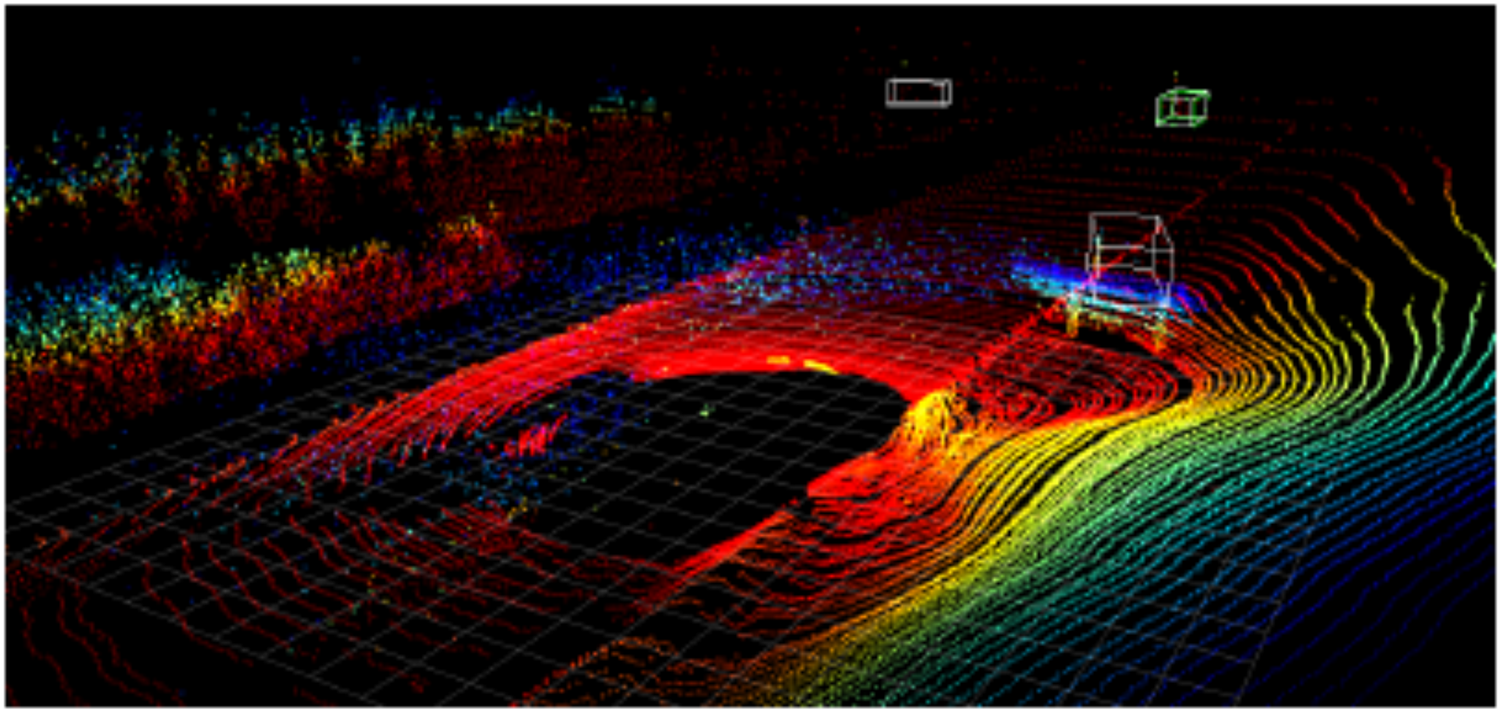
No fog



Extreme fog

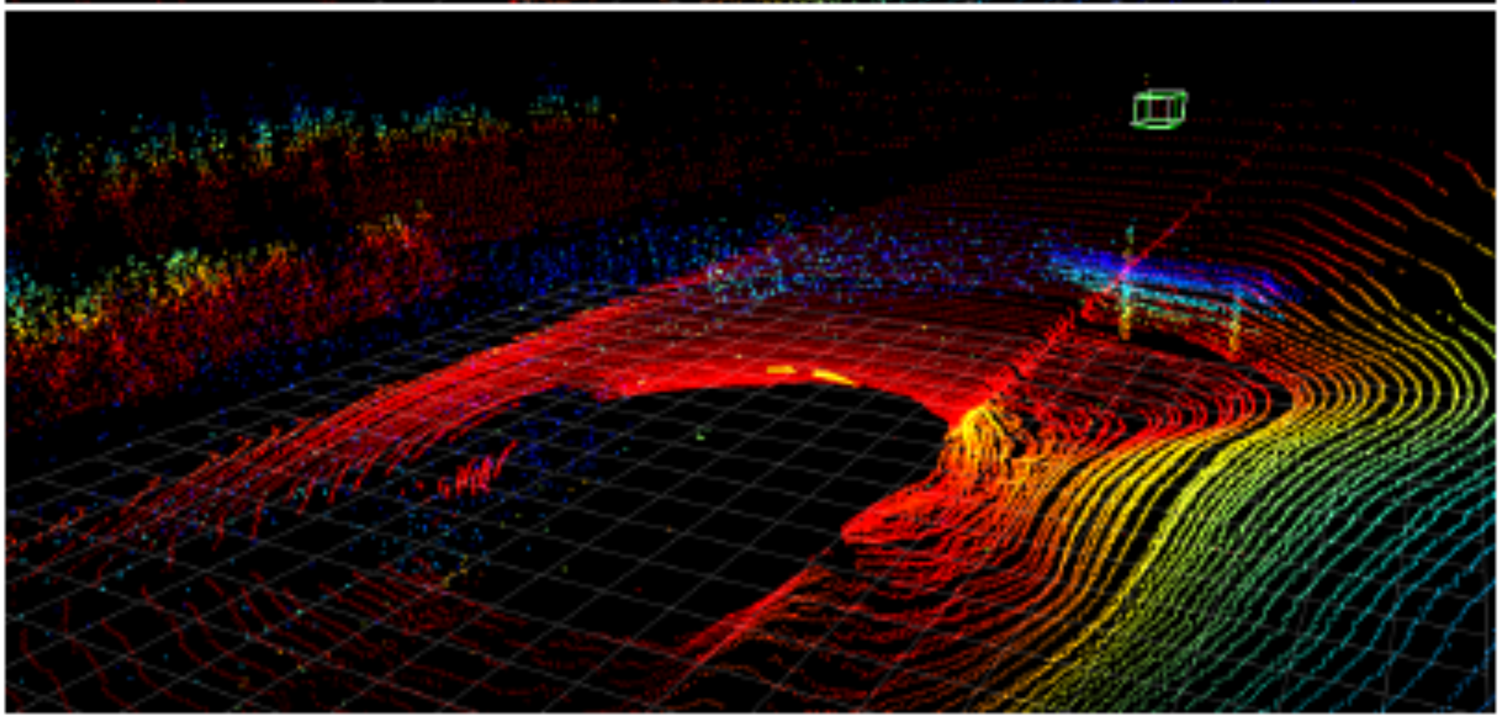
Simulate this effect, with a form of point spread function

Hahner 21



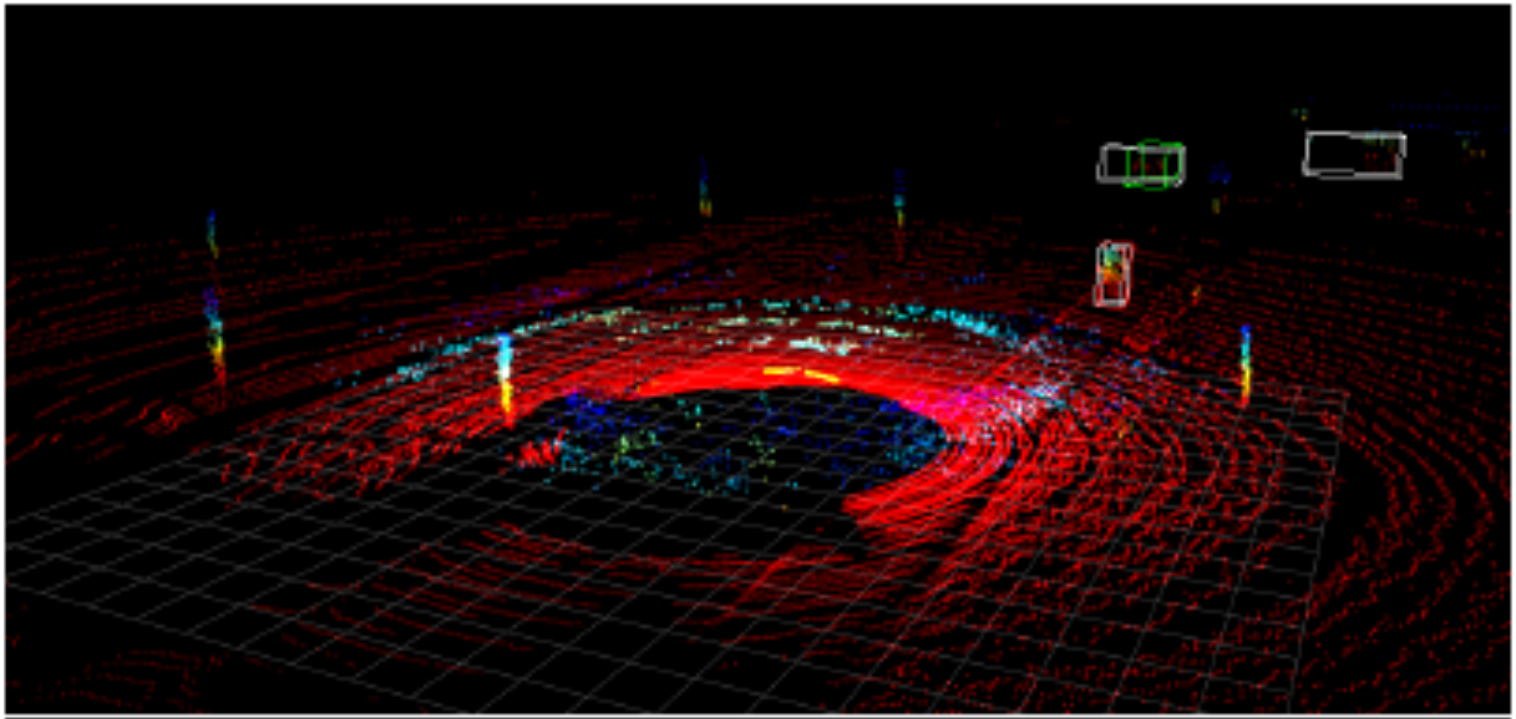
PV RCNN trained on good weather lidar returns only

Lidar captured in dense fog



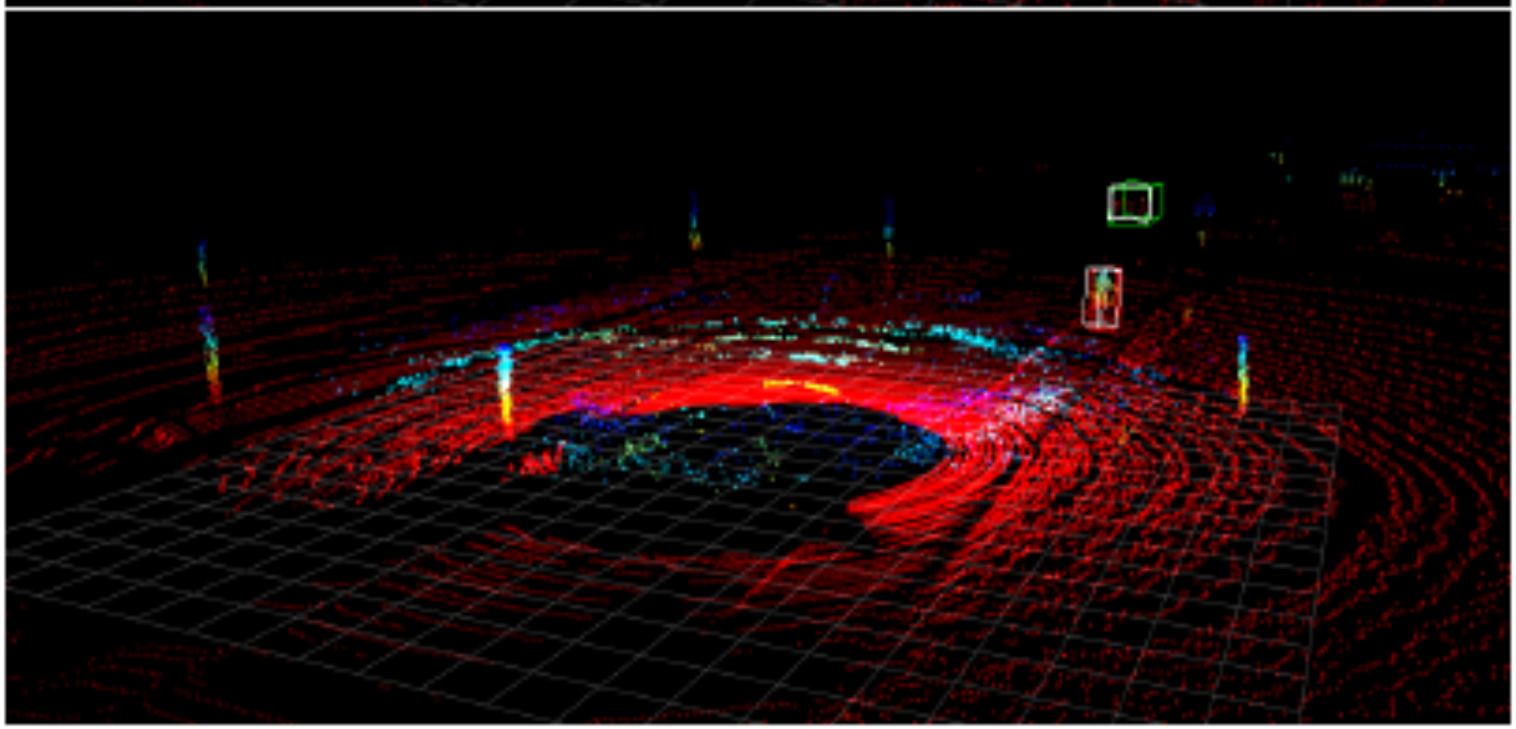
PV RCNN trained on good and simulated bad weather lidar returns only

Hahner 21



PV RCNN trained on good weather lidar returns only

Lidar captured in dense fog



PV RCNN trained on good and simulated bad weather lidar returns only

Hahner 21

Multi sensor methods

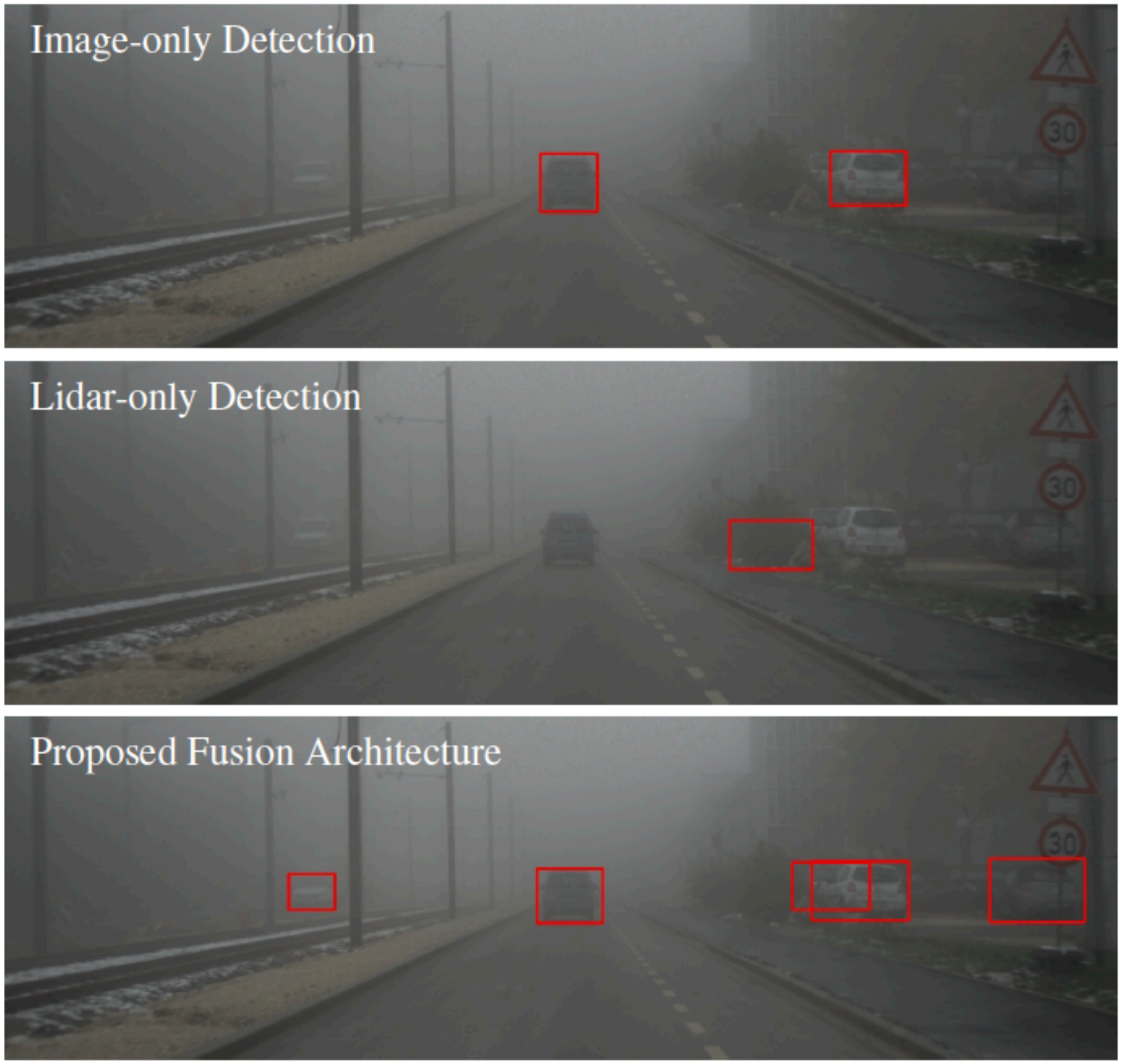
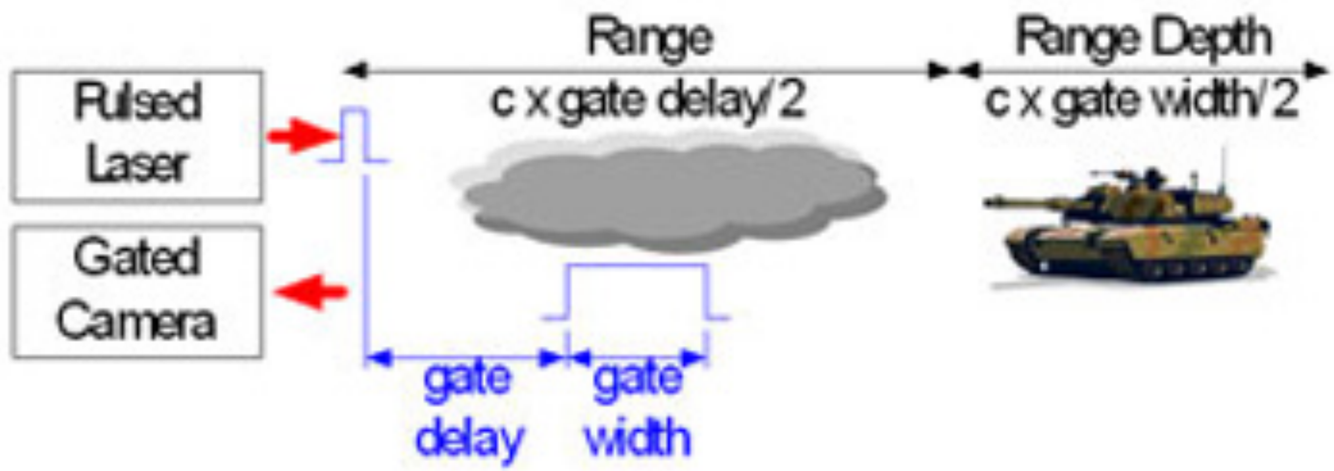


Figure 1: Existing object detection methods, including efficient Single-Shot detectors (SSD) [41], are trained on automotive datasets that are biased towards good weather conditions. While these methods work well in good conditions [19, 59], they fail in rare weather events (top). Lidar-only detectors, such as the same SSD model trained on projected lidar depth, might be distorted due to severe backscatter in fog or snow (center). These asymmetric distortions are a challenge for fusion methods, that rely on redundant information. The proposed method (bottom) learns to tackle unseen (potentially asymmetric) distortions in multimodal data without seeing training data of these rare scenarios.

Bijelic et al 20

Gated cameras



From sensors unlimited website

Multi sensor bad weather data

DATASET SENSOR SETUP	KITTI [19]	BDD [69]	Waymo [59]	NuScenes [6]	Ours
RGB CAMERAS	2	1	5	6	2
RGB RESOLUTION	1242×372	1280×720	1920×1080	1600x900	1920x1024
LIDAR SENSORS	1	✗	5	1	2
LIDAR RESOLUTION	64	0	64	32	64
RADAR SENSOR	✗	✗	✗	4	1
GATED CAMERA	✗	✗	✗	✗	1
FIR CAMERA	✗	✗	✗	✗	1
FRAME RATE	10 Hz	30 Hz	10 Hz	1 Hz/10 Hz	10 Hz
DATASET STATISTICS					
LABELED FRAMES	15K	100k	198k	40K	13.5K
LABELS	80k	1.47M	7.87M	1.4M	100K
SCENE TAGS	✗	✓	✗	✓	✓
NIGHT TIME	✗	✓	✓	✓	✓
LIGHT WEATHER	✗	✓	✗	✓	✓
HEAVY WEATHER	✗	✗	✗	✗	✓
FOG CHAMBER	✗	✗	✗	✗	✓

Table 1: Comparison of the proposed multimodal adverse weather dataset to existing automotive detection datasets.

Bijelic et al 20

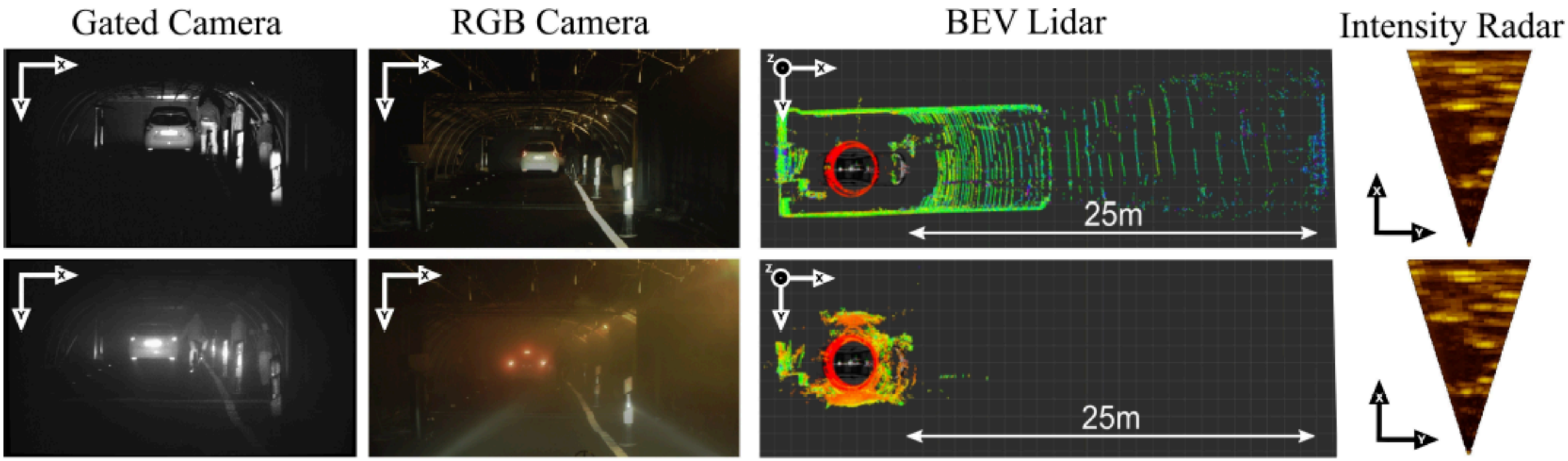


Figure 3: Multimodal sensor response of RGB camera, scanning lidar, gated camera, and radar in a fog chamber with dense fog. Reference recordings under clear conditions are shown in the first row, recordings in fog with visibility of 23 m are shown in the second row.

Bijelic et al 20

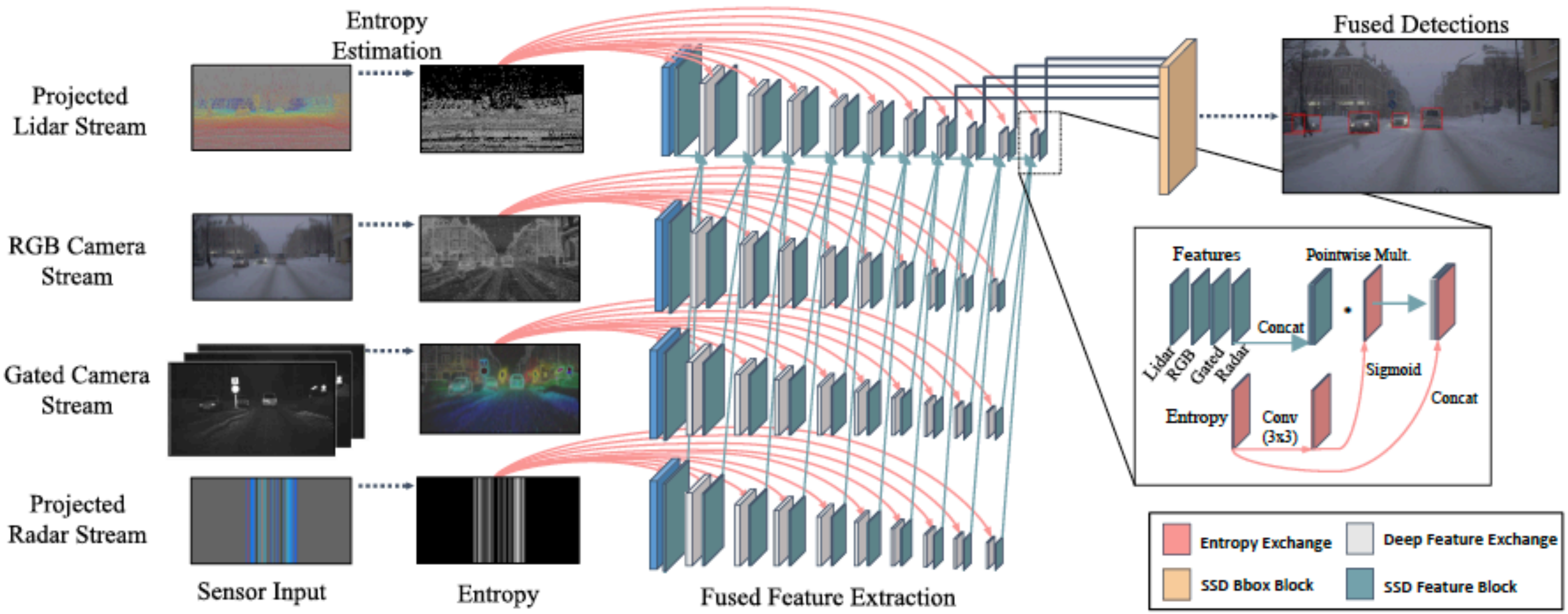


Figure 4: Overview of our architecture consisting of four single-shot detector branches with deep feature exchange and adaptive fusion of lidar, RGB camera, gated camera, and radar. All sensory data is projected into the camera coordinate system following Sec. 4.1. To steer fusion in-between sensors, the model relies on sensor entropy, which is provided to each feature exchange block *(red)*. The deep feature exchange blocks *(white)* interchange information *(blue)* with parallel feature extraction blocks. The fused feature maps are analyzed by SSD blocks *(orange)*.

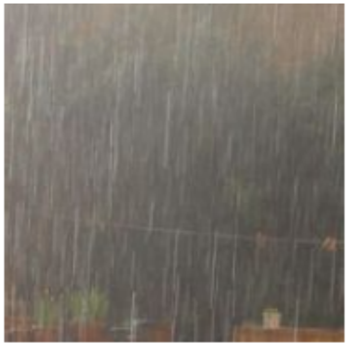
Bijelic et al 20

Rain has multiple interesting effects

Blur from wet air

Puddles



Color shifts

Streaks

These are often quite strongly coupled to scene geometry

Rain - multiple extrinsic phenomena, including smoothing, raindrops, loss of saturation, glossy/wet surfaces, etc. etc.

Rain - phenomena

Refraction causes each drop to contain a tiny image

(a) An image of a drop hanging from a pipette (b) Perspective views created from (a)

Figure 7. Looking at the world through a raindrop. (a) An image of a drop hanging from a pipette and a magnified version. (b) Near-perspective views computed using the geometric mapping due to refraction. Note that, in the perspective views, straight lines in the scene are mapped to straight lines in the image.

Garg and Nayar 07

Backscatter

- Refraction in drops causes backscatter of headlight light
  - makes driving in rain at night harder
- Neat trick
  - (Tamburo et al 14)
  - Do not illuminate raindrops by
    - having headlights that are highly steerable (multiple micro mirrors)
    - very fast exposure with usual illumination identifies raindrops
      - too fast for driver to resolve
    - now direct light between drops

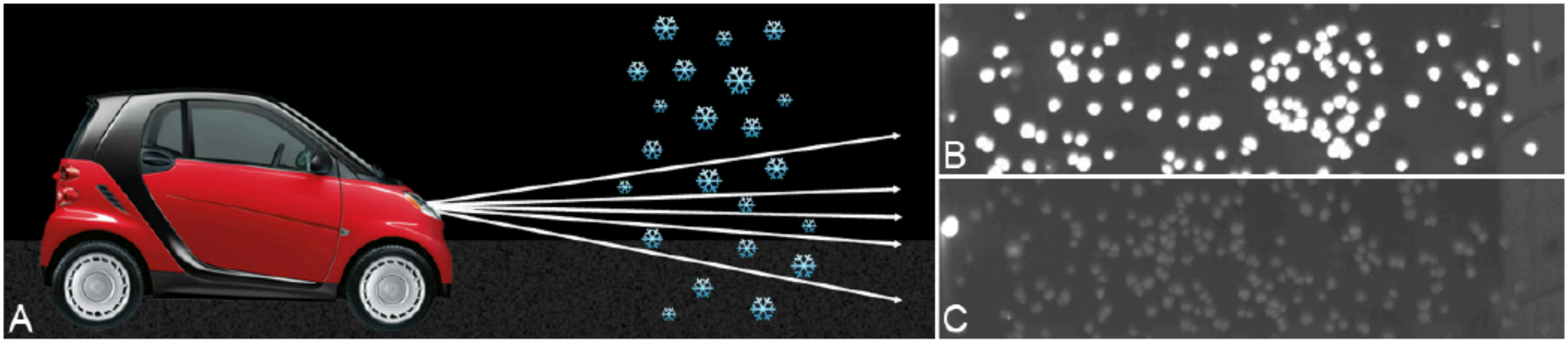
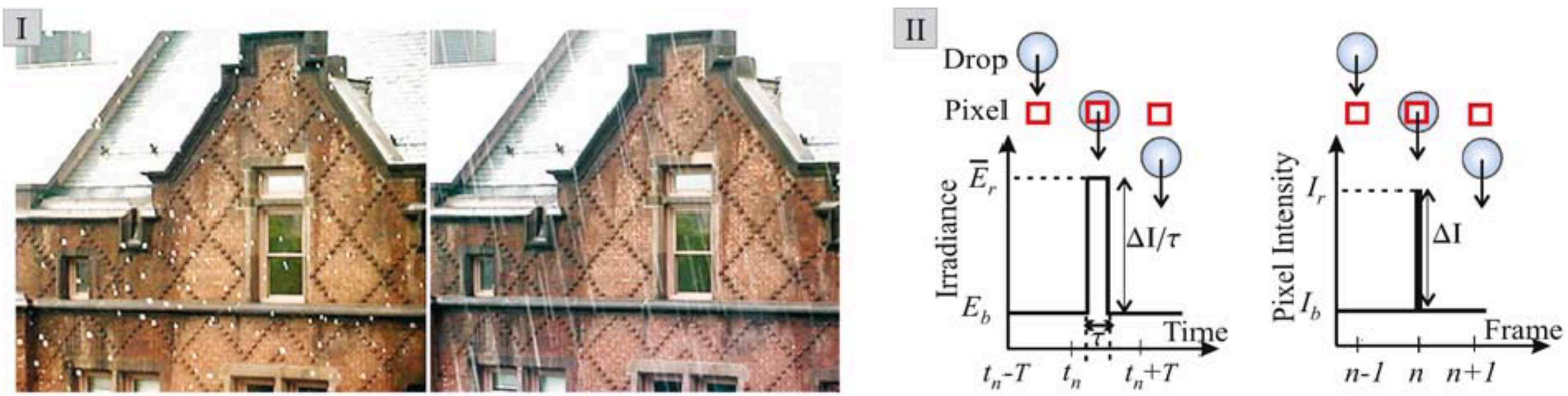


Fig. 7. A: Our headlight has unprecedented resolution over space and time so that beams of light may be sent in between the falling snow. Illustration adapted from [11]. B: Artificial snowflakes brightly illuminated by standard headlight. C: Our system avoids illuminating snowflakes making them much less visible.

Rain - phenomena

Drops move fast, and so create motion blur (streaks)



(a) Short exposure time (1 ms) (b) Normal exposure time (30 ms) (a) Average irradiance at a pixel (b) Intensity at a pixel

Figure 9. (I) Raindrops and motion-blur. An image of a scene taken in rain with (a) a short exposure time of 1 ms and (b) with typical exposure time of a camera (30 ms). (II) The intensities produced by motion-blurred raindrops. II (a) The average irradiance at the pixel due to the raindrop is $\bar{E}_r$ and that due to the background scene is E_b. Note that $\bar{E}_r > E_b$. The drop projects onto a pixel for time $\tau < 1.18$ ms, which is far less than the typical exposure time T of a camera. (b) Intensities of a pixel in three frames. A drop stays over the pixel in only a single frame and produces a positive intensity fluctuation of unit frame width.

Garg and Nayar 07

Rain - phenomena

Shallow free space - individual rain streaks
Deep free space - more bulk, fog-like effects

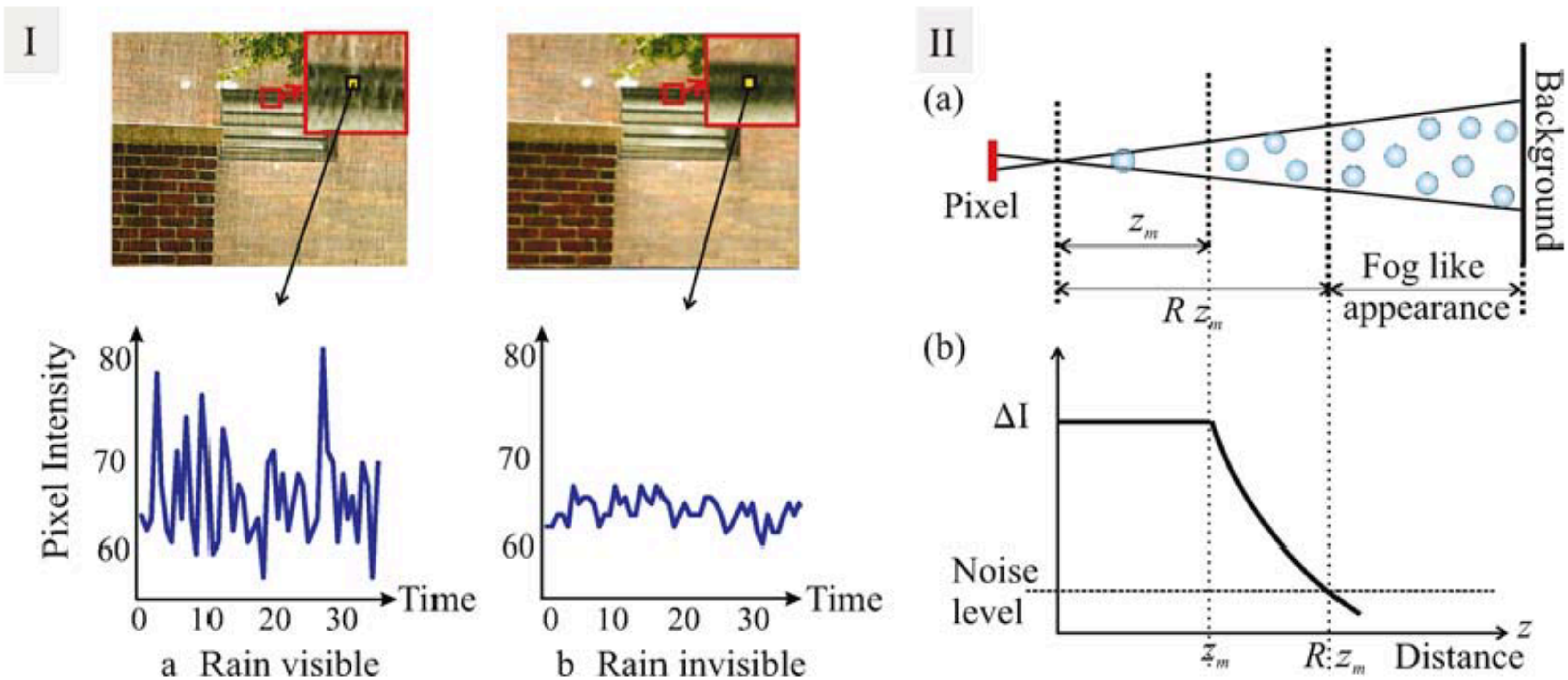
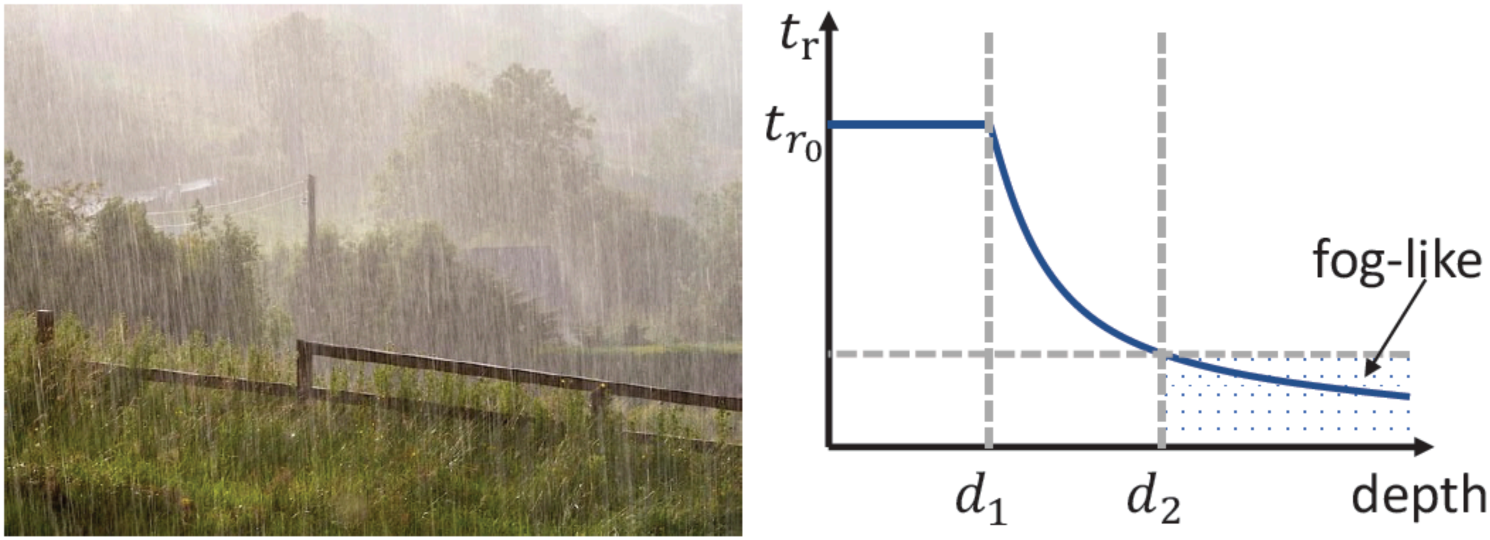


Figure 13. Dynamic weather and visibility: (I)(a) Frame from a video of a scene where rain is visible. The intensity variation due to rain is high. (b) Frame from a video of the same scene taken with camera parameters to reduce the visibility due to rain. The intensity at the same pixel shows low variance over time. (II) The change in intensity produced by a falling raindrop as a function of the drop's distance z from the camera. The change in intensity ΔI does not depend on z for drops that are close to the camera ($z < z_m$). While for raindrops far from the camera ($z > z_m$), ΔI decreases as $1/z$ and for distances greater than $R\,z_m$, ΔI is too small to be detected by the camera. Therefore, the visual effects of rain are only due to raindrops that lie close to the camera ($z < Rz_m$) which we refer to as the *rain visible region.*

Garg and Nayar 07

Rain - phenomena

Shallow free space - individual rain streaks
Deep free space - more bulk, fog-like effects

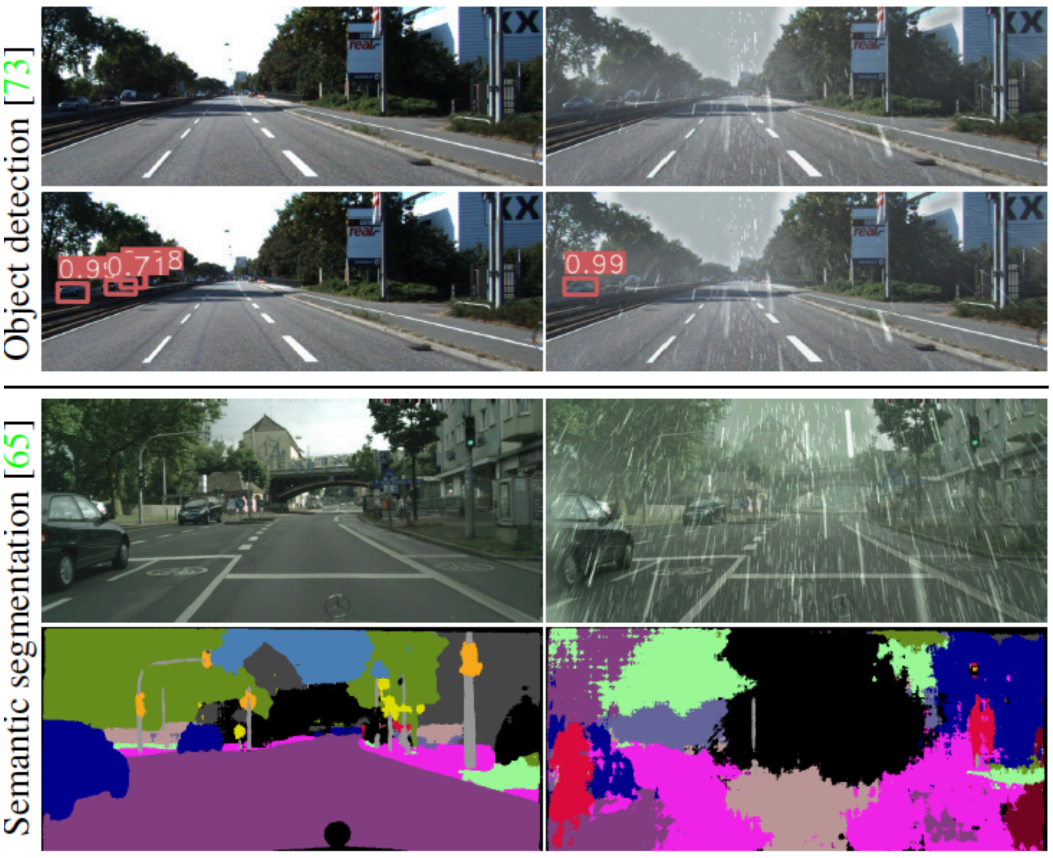


(a) input real photo (b) rain visibility & depth

Figure 1: (a) An example real photo that demonstrates the scene visibility variation with depth, and the presence of rain streaks and fog; and (b) a plot of rain streak intensity (t_r) against scene depth (d) based on the model in [13].

Hu et al 19

Simulating rain



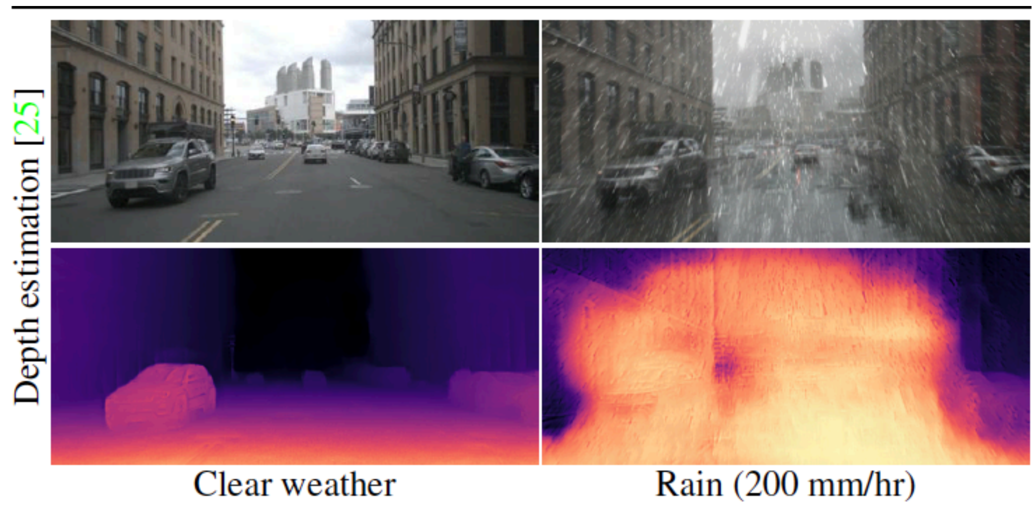


Fig. 1 Vision tasks in clear and rain-augmented images. Our synthetic rain rendering framework allows for the evaluation of computer vision algorithms in challenging bad weather scenarios. We render physically-based, realistic rain on images from the KITTI [23] (rows 1-2) and Cityscapes [13] (rows 3-4) datasets with object detection from mx-RCNN [73] (row 2), semantic segmentation from ESPNet [65] (row 4). We also present a combined data-driven and physic-based rain rendering approach which we apply to the nuScenes [9] (rows 5-6) dataset with depth estimation from Monodepth2 [25] (row 6). All algorithms are quite significantly affected by rainy conditions.

Tremblay et al 20

Simulating rain - issues

- Near field:
  - drops are bright, discrete, likely ballistic motion
    - how bright?
    - where?
    - how moving?
  - likely air is “wet”
    - so some fogging, depending on depth
- Far field:
  - fog like effects
- So we need to know
  - depth, environment map, falling drops, camera movement

Simulating rain

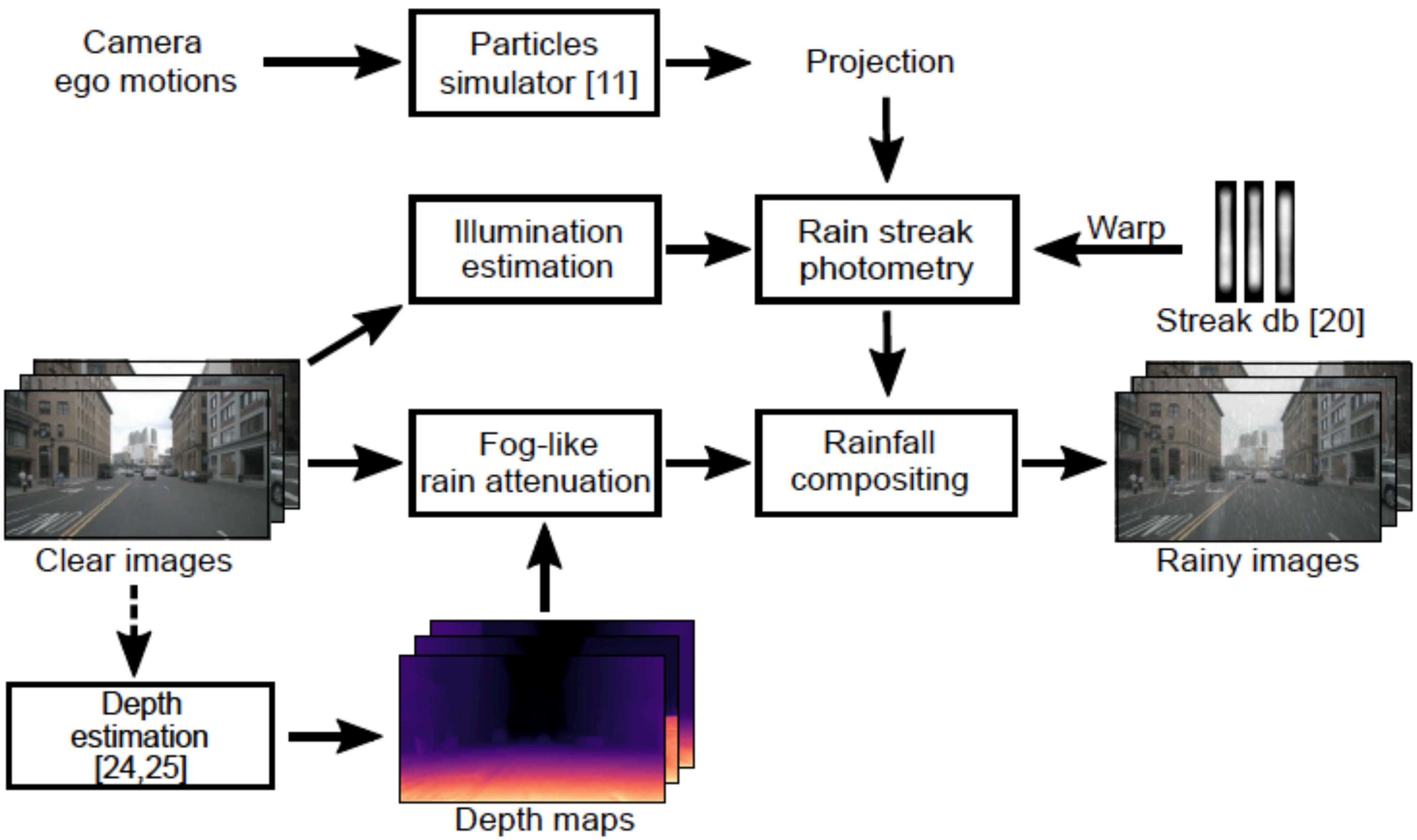
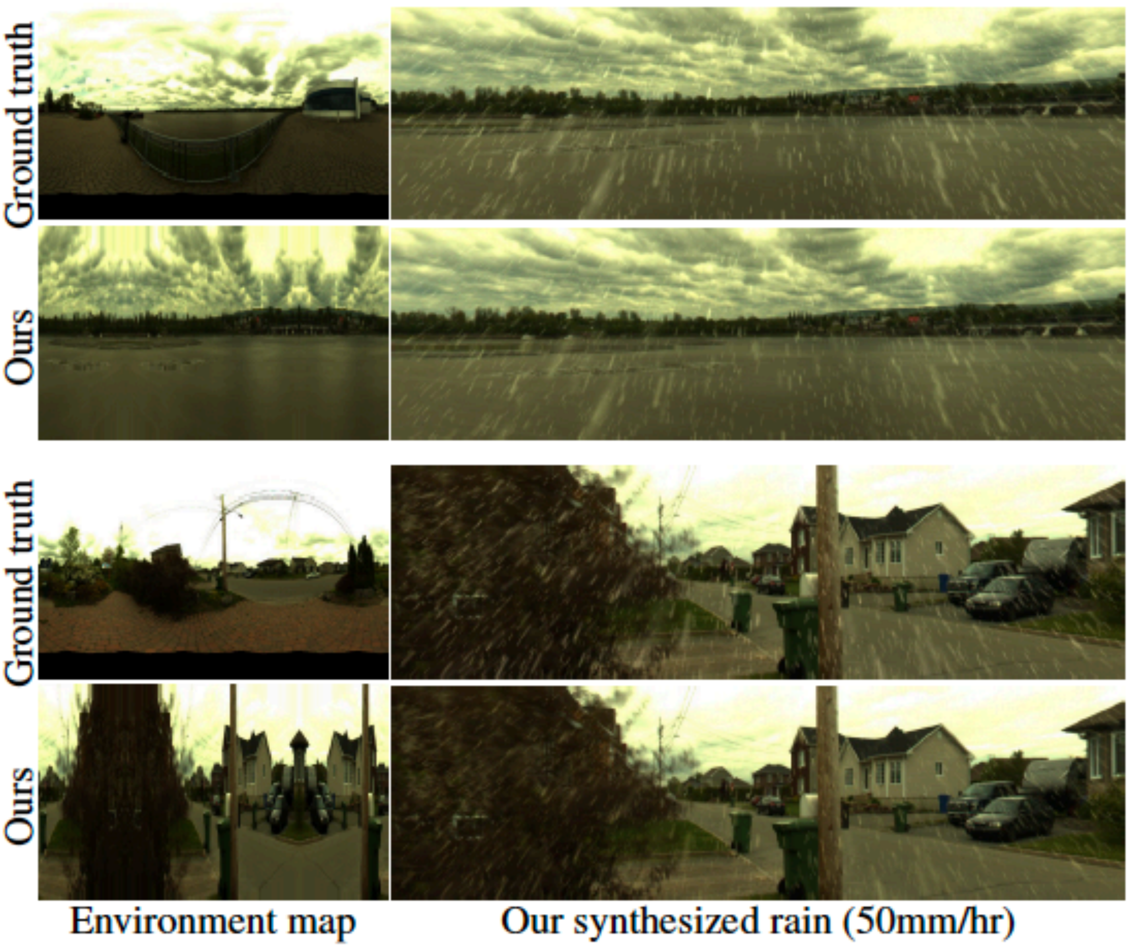


Fig. 2 Physics-Based Rendering for rain augmentation. We use particles simulation together with depth and illumination estimation to render arbitrarily controlled rainfall on clear images.

Tremblay et al 20

Simulating rain



Minor errors in environment map have no real effect on rain appearance

Tremblay et al 20

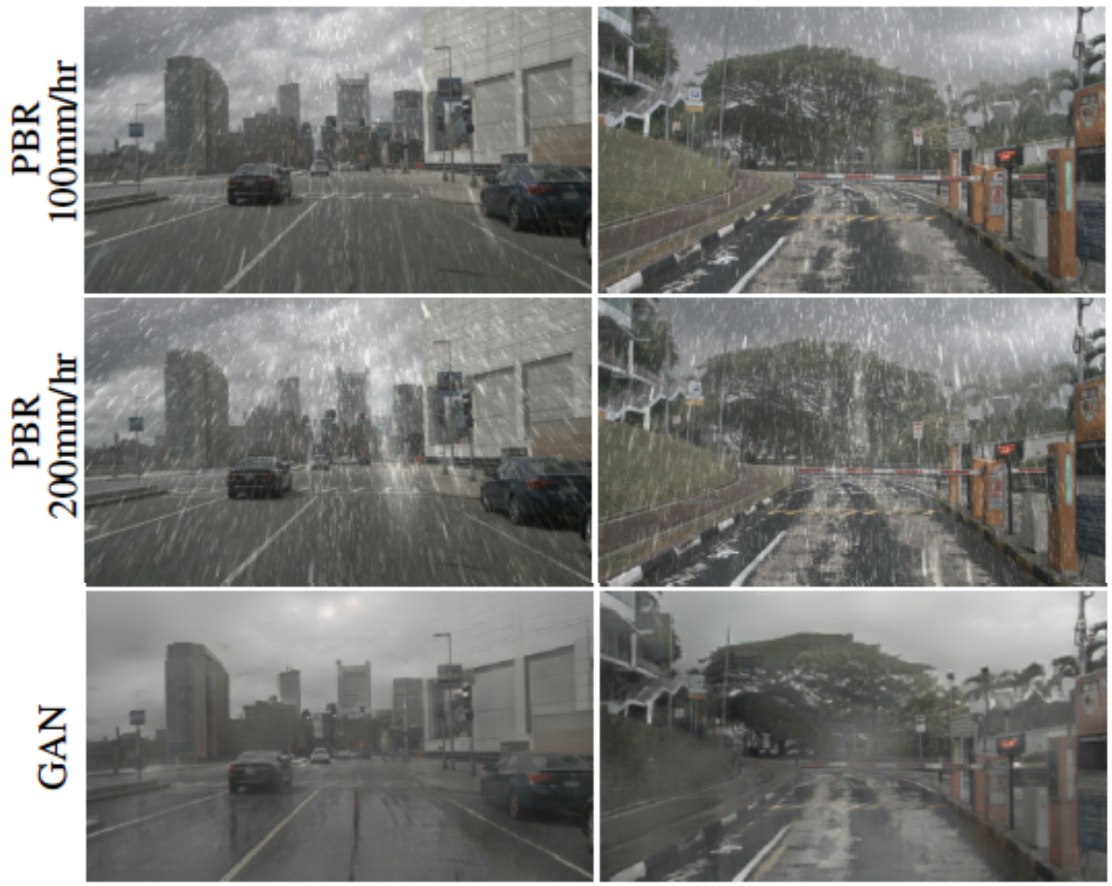
Simulating rain

- Trick:
  - rain causes color effects, specular effects etc.
    - CycleGAN is good at this, but bad at streaks
    - Physics based simulation is bad at this but good at streaks

Fig. 5 GAN+PBR rain-augmentation architecture. In this hybrid approach, clear images are first translated into rain with CycleGAN [83] and subsequently augmented with rain streaks with our PBR pipeline (see fig. 2).

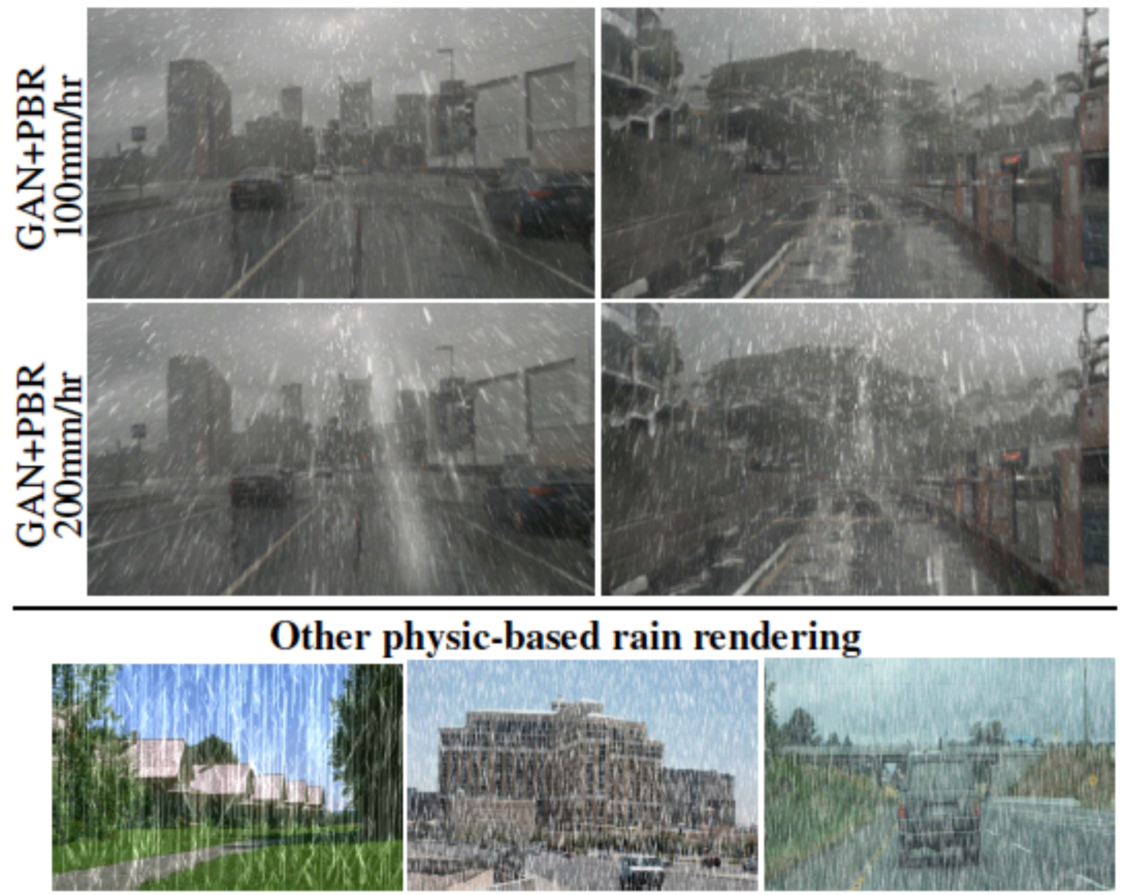
Tremblay et al 20

Rain photographs



Tremblay et al 20

Rain photographs



rain100H [74] rain800 [79] did-MDN [78]

Tremblay et al 20

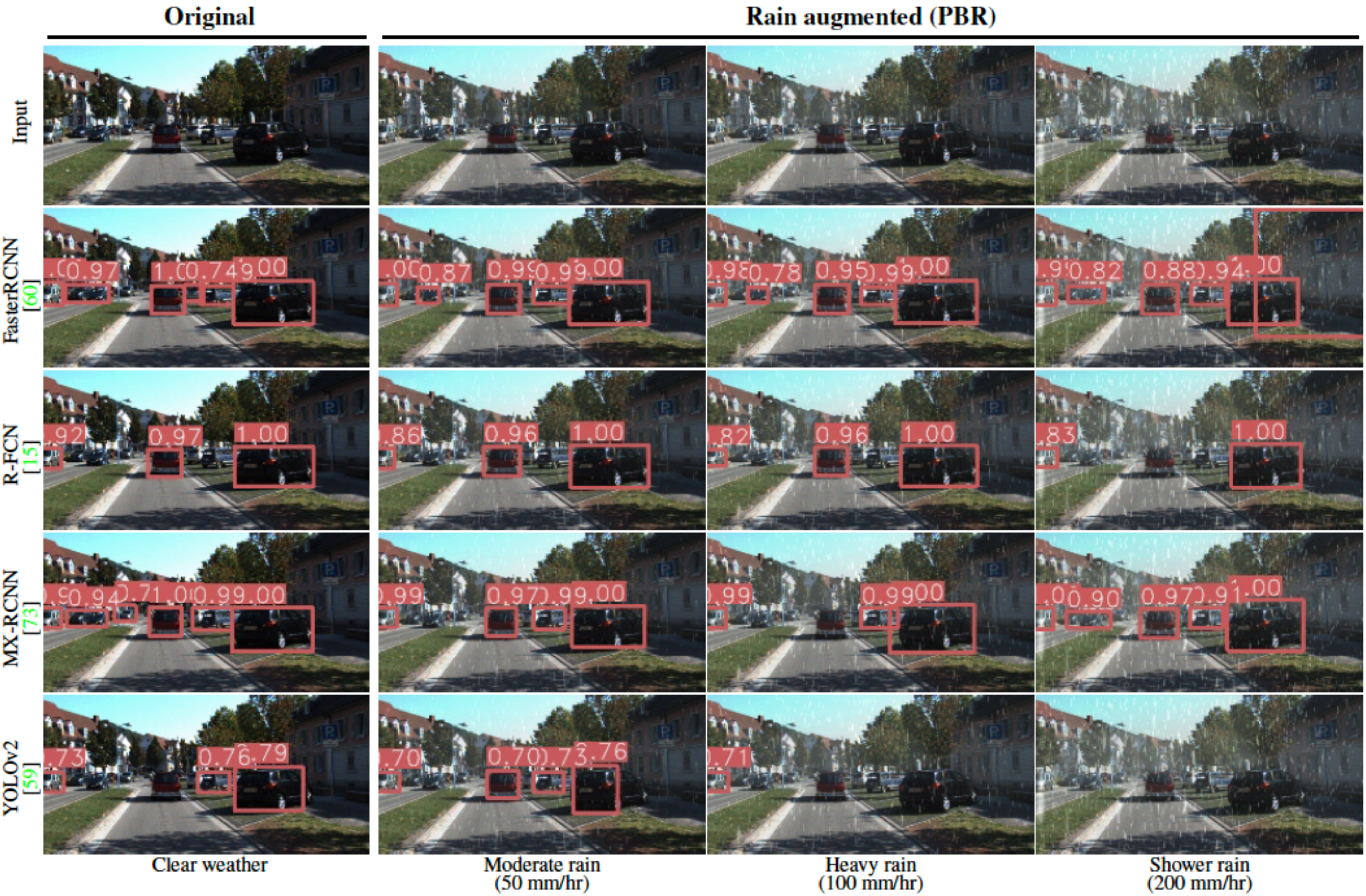


Fig. 11 Object detection on PBR rain augmentation of KITTI. From left to right, the original image (clear) and three PBR augmentations with varying rainfall rates. Images are cropped for visualization.

Tremblay et al 20

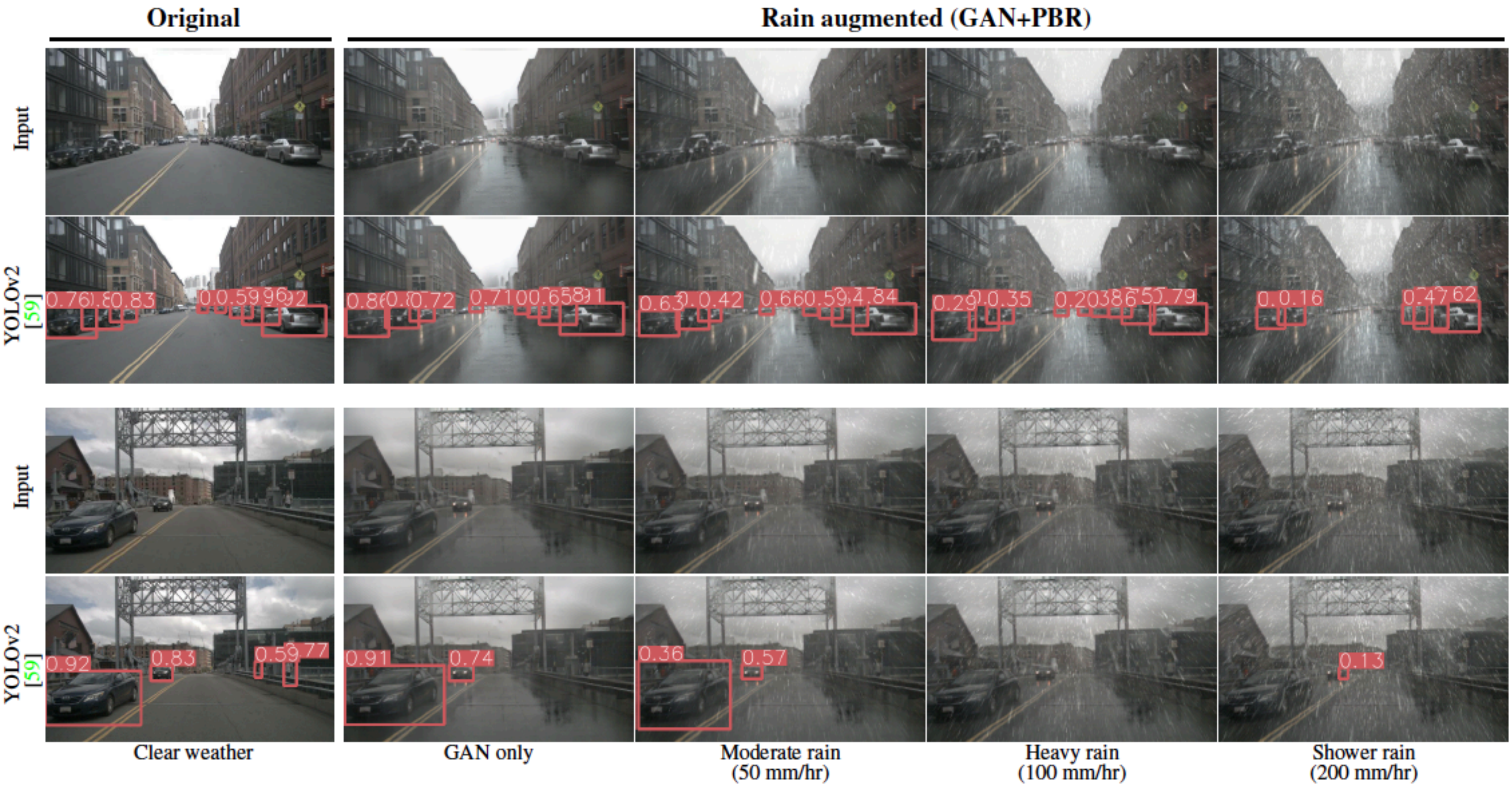


Fig. 15 Object detection on our GAN+PBR augmented nuScenes. From left to right, the original image (clear), the GAN augmented image and three GAN+PBR images.

Tremblay et al 20

Deraining - strategies

- Essentially
  - obtain images with/without rain (with rain by synthetic)
  - train network to reproduce without rain image from with rain
  - starts with Eigen et al 13

Figure 1. A photograph taken through a glass pane covered in rain, along with the output of our neural network model, trained to remove this type of corruption. The irregular size and appearance of the rain makes it difficult to remove with existing methods. This figure is best viewed in electronic form.

From Eigen et al. 13

Rainy windows

From Eigen et al. 13

Rain streaks

Figure 7: Visual comparison of different rain streak removal methods on real example images.

Li et al 16

Streaks

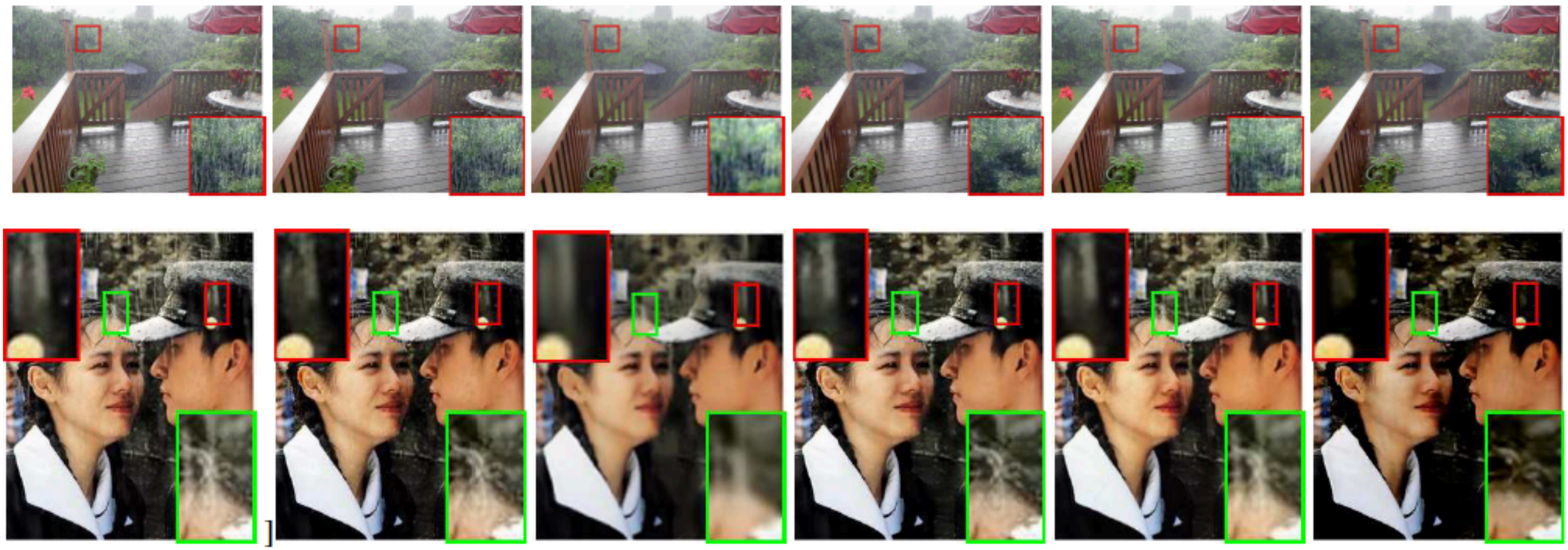


Figure 6: Real rain streaks removal experiments under different scenarios. From left to right are input image, results of DSC[26], LP [24], CNN [10], DID-MDN[31] and ours. Demarcated areas in each image are amplified at a 3 time larger scale.

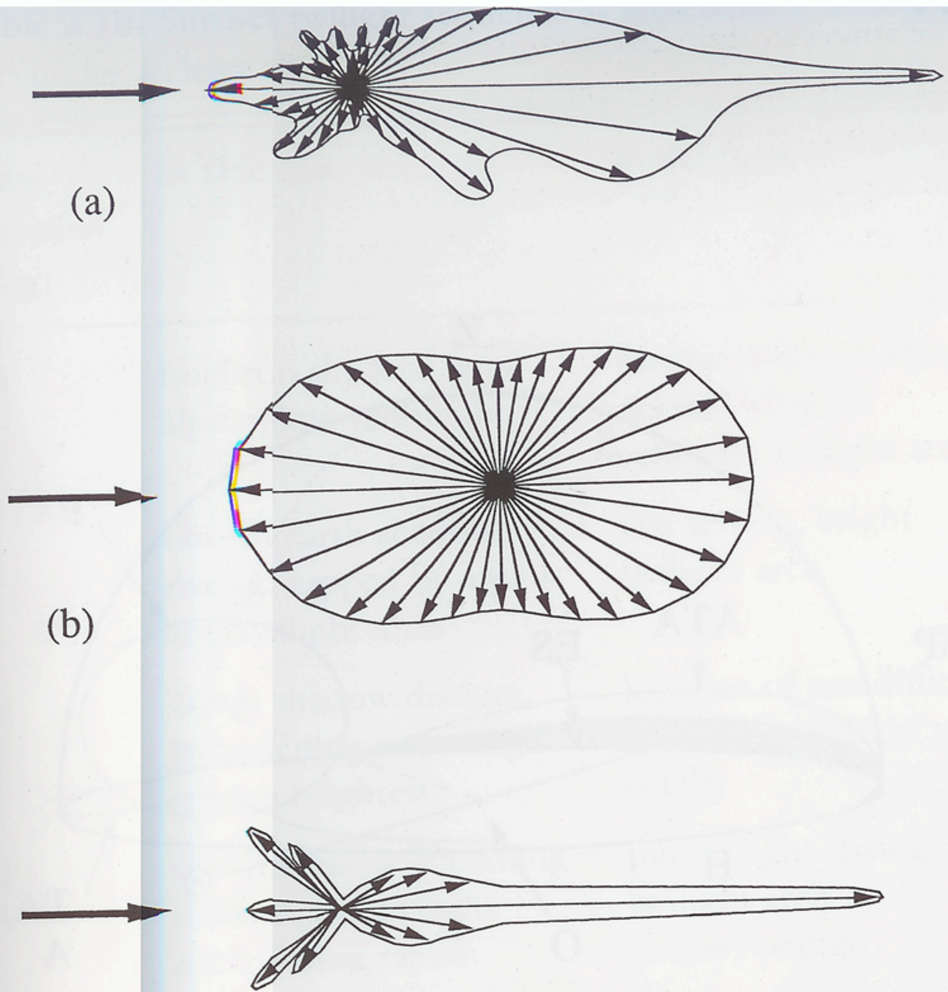
Wei et al 19

Both rain streaks and haze

Figure 7. Examples of JORDER-R-DEVEIL on heavy rain (left two images) and mist images (right two images).

Yang et al 17

Scattering profiles can be complicated



Big (dust, smoke)

Air molecules

Water drops

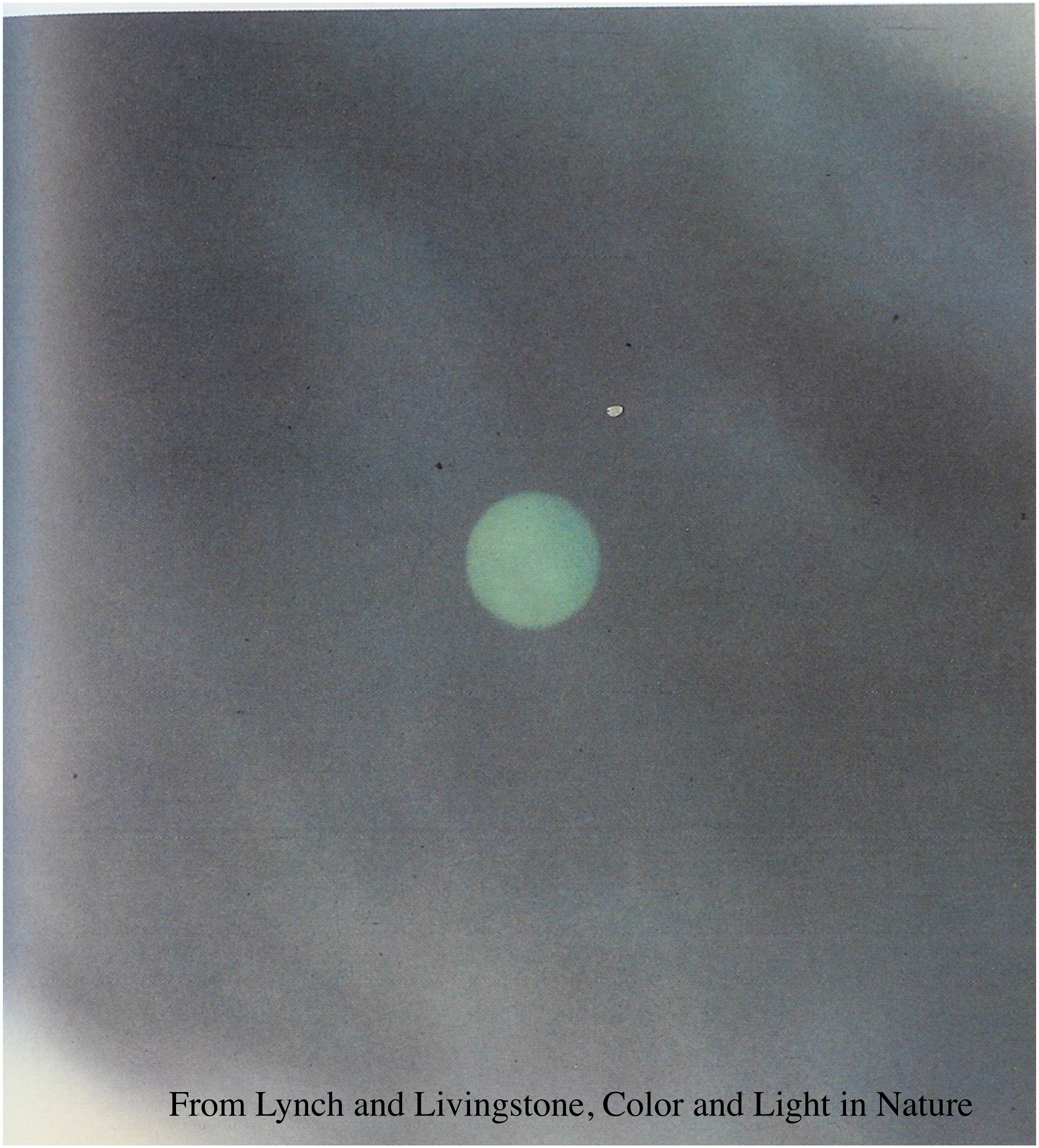
Fig. 2.7C Scattering patterns for different particles. (a) Large irregular particles, like those comprising dust and smoke, are irregular in the sense that they are not symmetric. They do, however, have a strong forward scattering peak and a smaller though still pronounced backscattering peak. (b) Air molecules have a scattering function that is symmetric fore and aft: they scatter the same amount of light in both the forward and backward directions but lack both the forward and backscattering peak. (c) Large water drops have a strong forward and backscattering peak and also show strong enhancements at the primary and secondary rainbow angles.

From Lynch and Livingstone, Color and Light in Nature

Fig. 2.7A (LEFT) Aureole around the sun. The sun is hidden by a street lamp. To the eye, the sky appeared clear.

Fig. 2.7B (RIGHT) The next day the sky was exceptionally clear and there was no aureole.

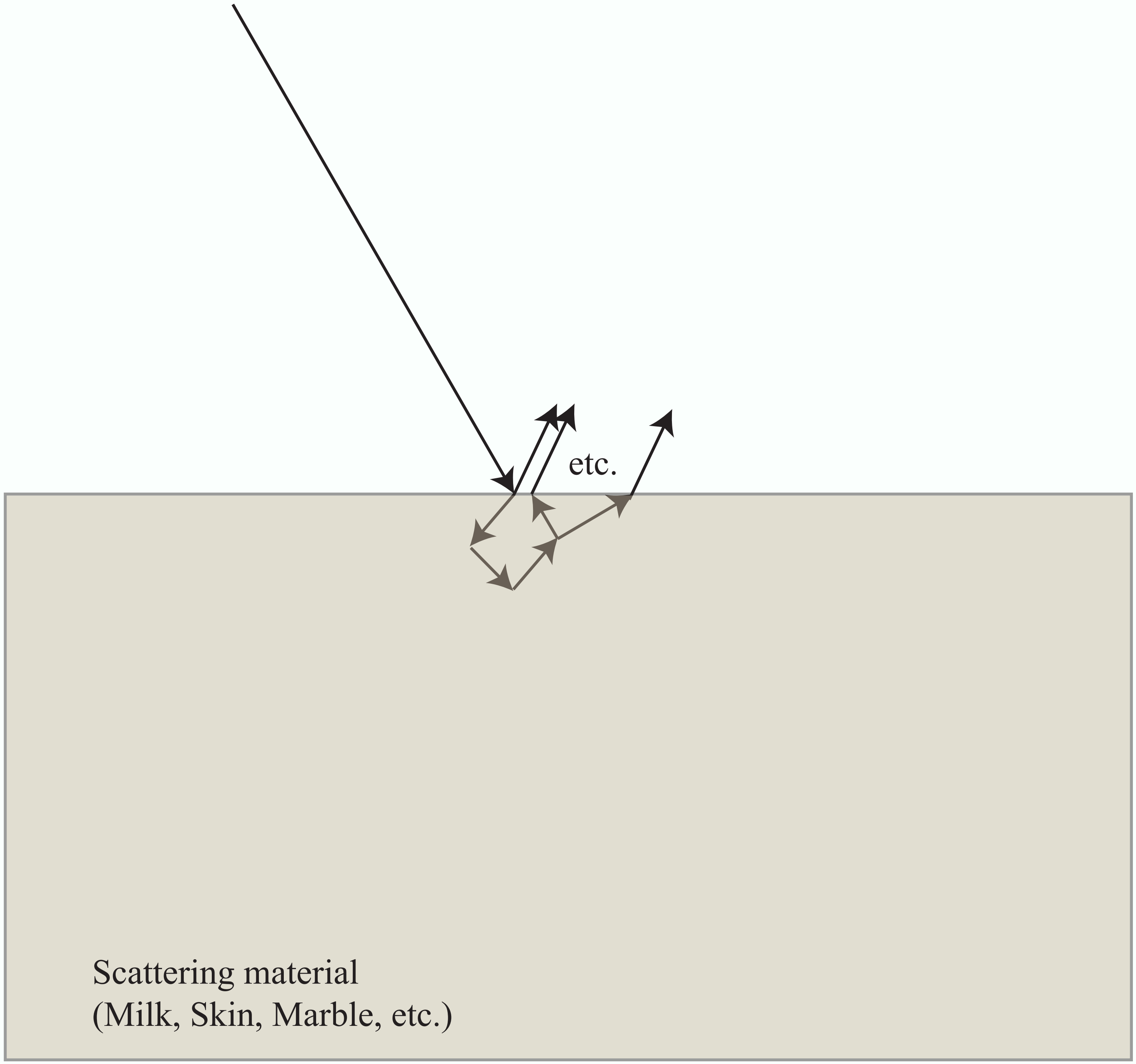
From Lynch and Livingstone, Color and Light in Nature



From Lynch and Livingstone, Color and Light in Nature

Minnaert, Light and Color in the outdoors

Notice flattened sun,
sparkles


etc.
Scattering material
(Milk, Skin, Marble, etc.)


RENDERED USING DALI - HENRIK WANN JENSEN 2001

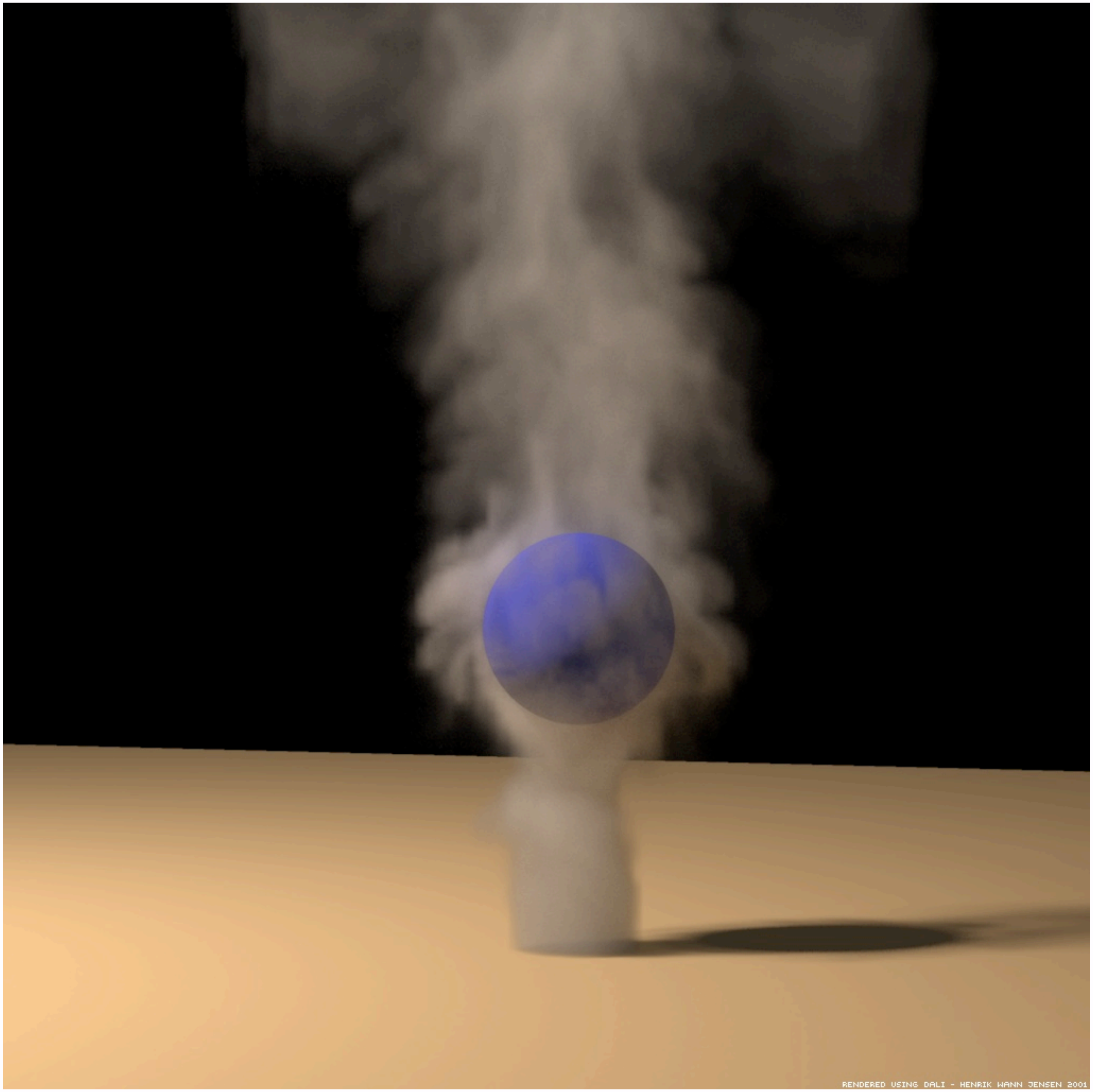

HENRIK WANN JENSEN 2000


HENRIK WANN JENSEN 2001

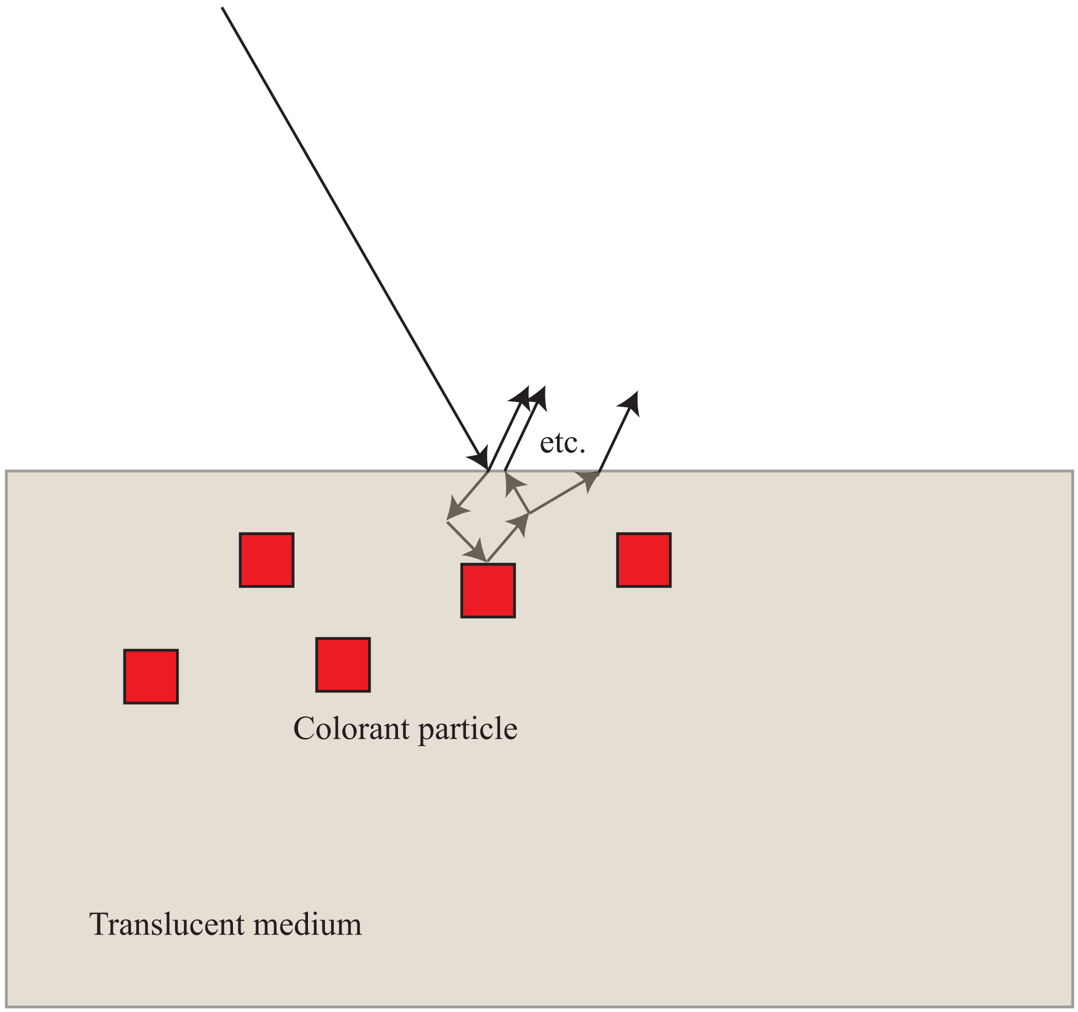

HENRIK WANN JENSEN 1999

subsurface scattering in skin (not rendered!)


RENDERED USING DALI - HENRIK WANN JENSEN 2001

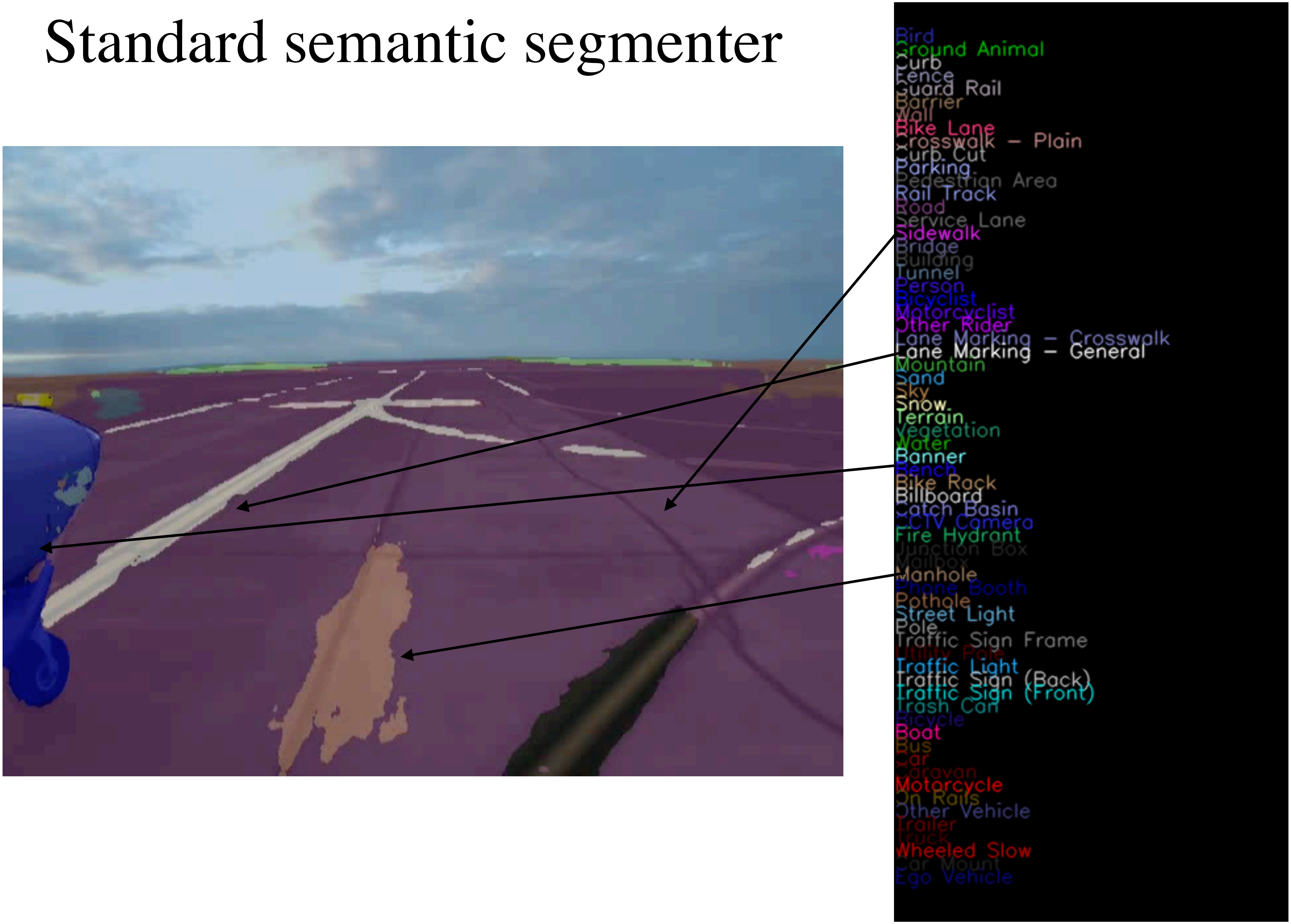
Paints are films with colored scatterers



Glowing paint from specular refl'ns

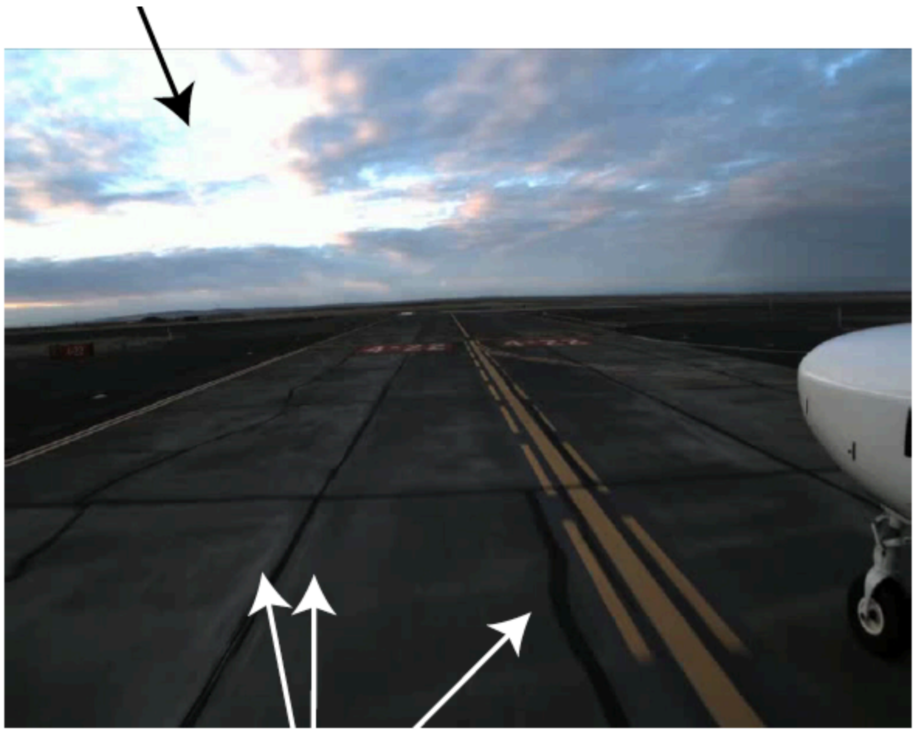

~~Boeing~~ Autonomy data

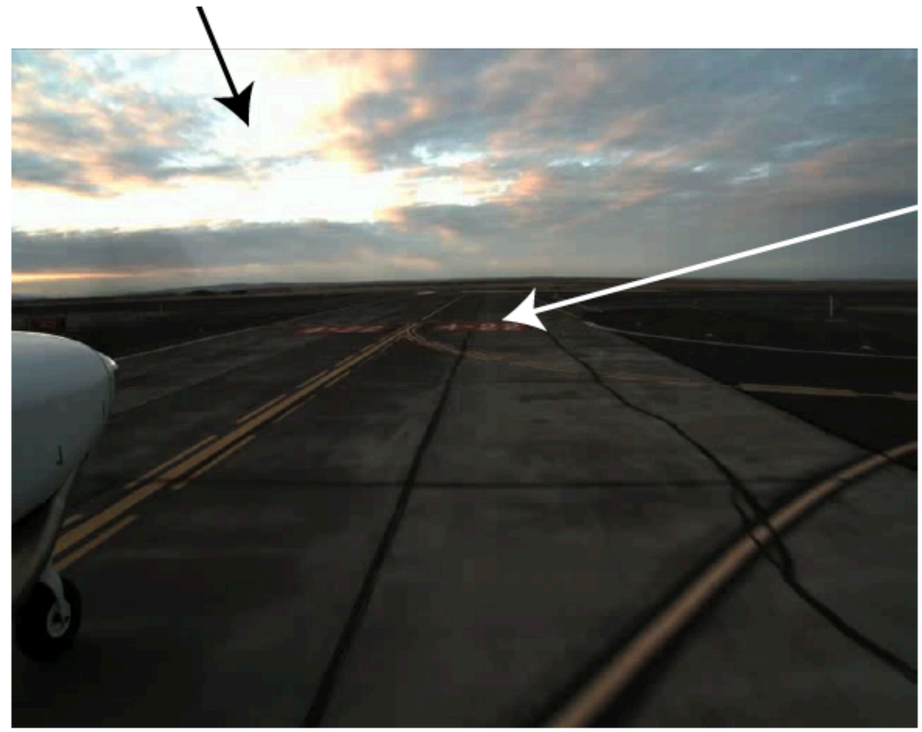
Standard semantic segmenter



Special features: rich appearance variation

Special features: rich appearance variation





Intrinsic images

- (Originally) Maps of an image that explain pixel values
  - Intrinsic properties:
    - independent of viewing; “object” or “world” properties
  - Extrinsic properties:
    - depend on viewing circumstances

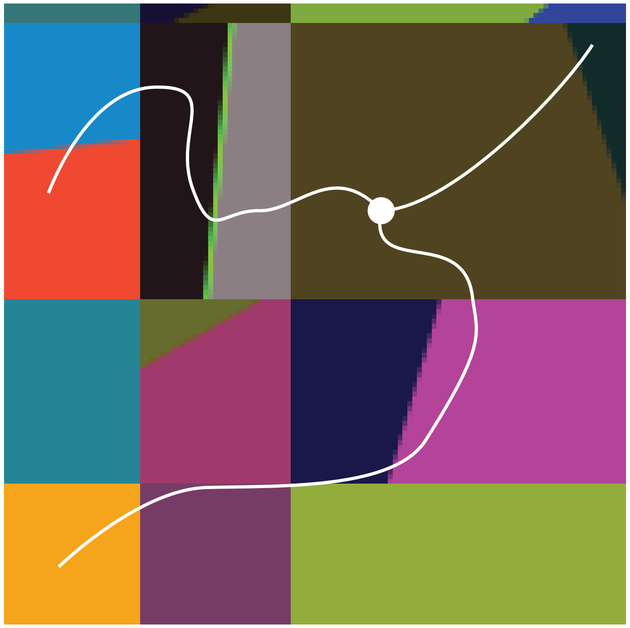
- (Later) Albedo/Shading maps
  - I=A x S
  - Albedo (A) is a natural intrinsic
  - Shading (S) is a natural extrinsic

No ground truth decompositions

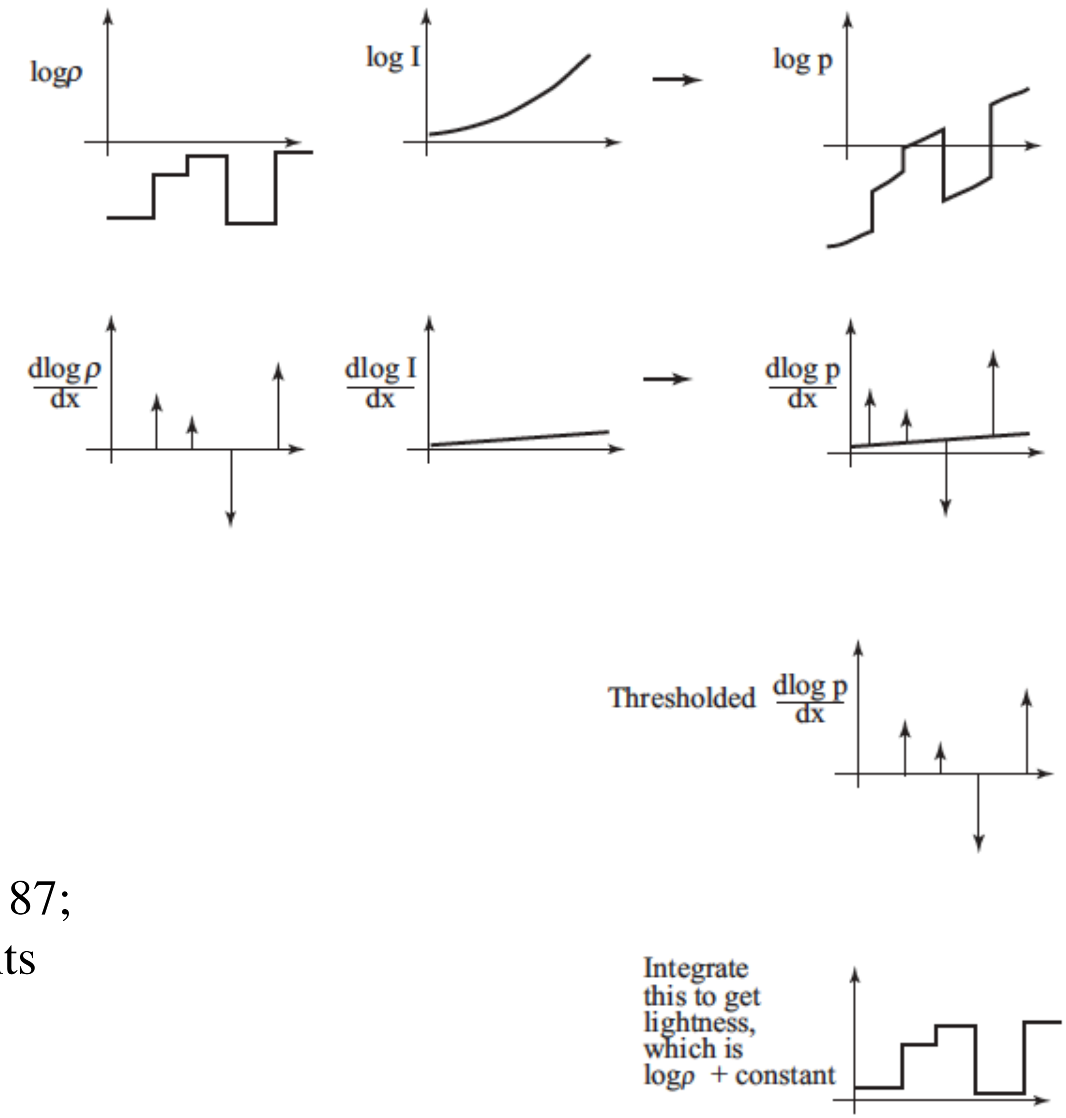
- And there never will be
  - rendering is do-able (but hard)
  - modelling is hopeless
- Q: how do you train an image decomposition method when you don’t know the right answer?
- Retinex provides clues - spatial statistics are the key

Albedo/shading and Retinex

- Spatial reasoning, Land (59, 59, 77); Land +McCann 71:
  - Surface color changes either quickly or not at all
  - Light color changes slowly
  - Retinex
    - large family of algorithms
    - quite hard to know what Retinex does (Brainard+Wandell, 86)



Computer vision versions of Retinex

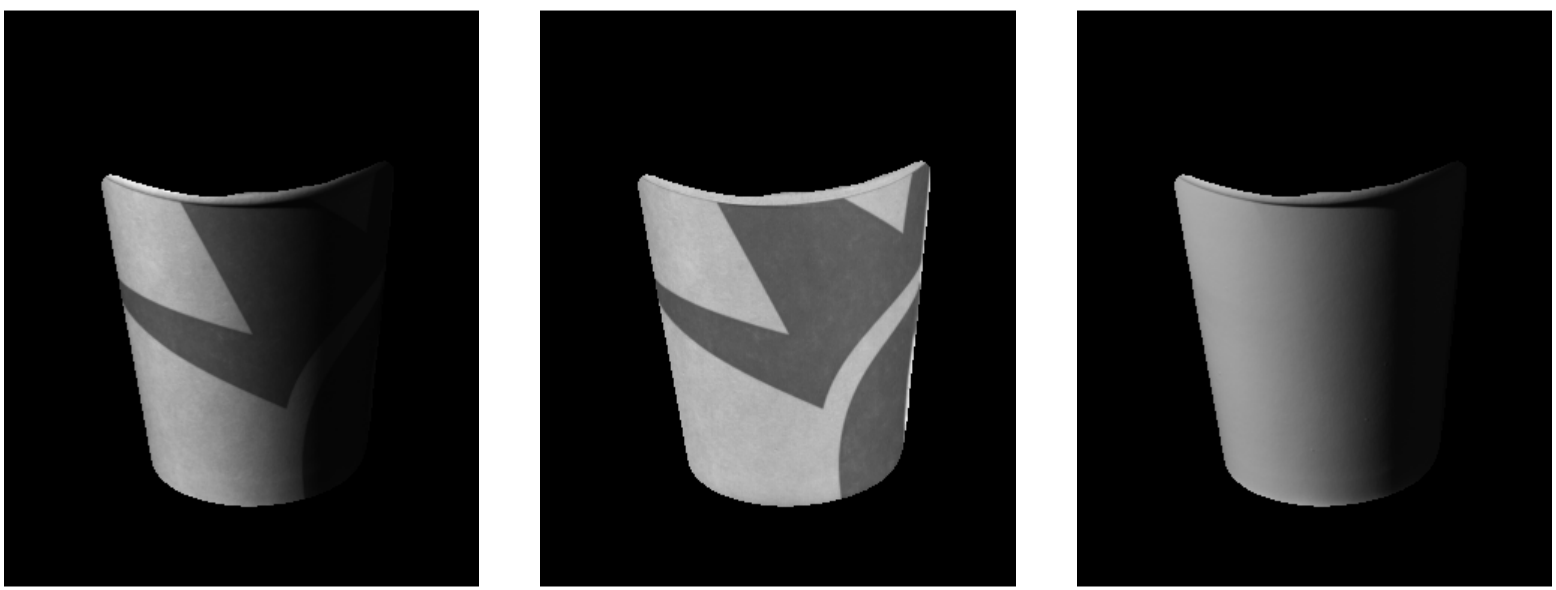


Horn, 73; 74
Brelstaff+Blake, 87;
multiple variants

Real data is hard to collect

- spraypaint, multiple images, etc…

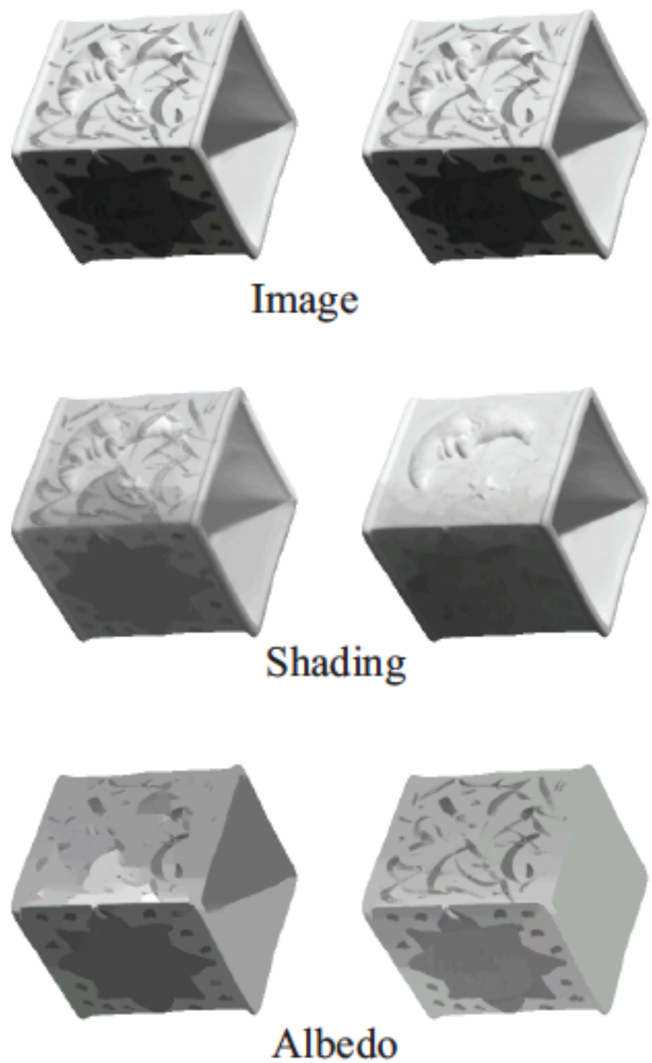
Images from dataset of Gosse et al. 09

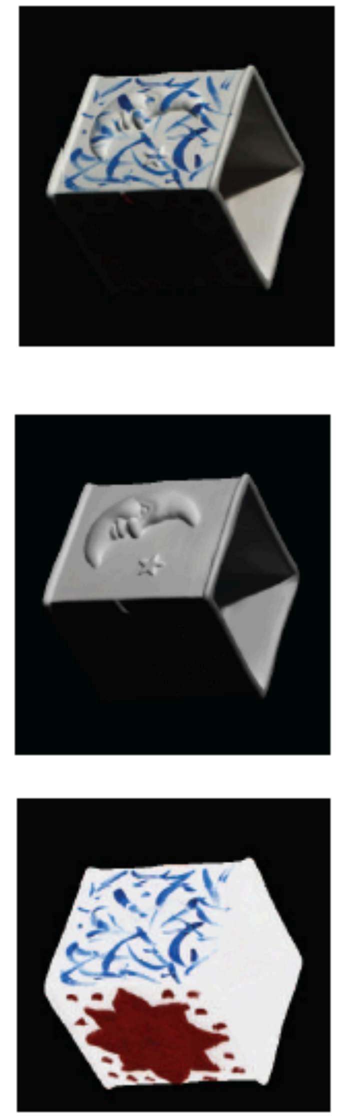


Retinex is really quite good

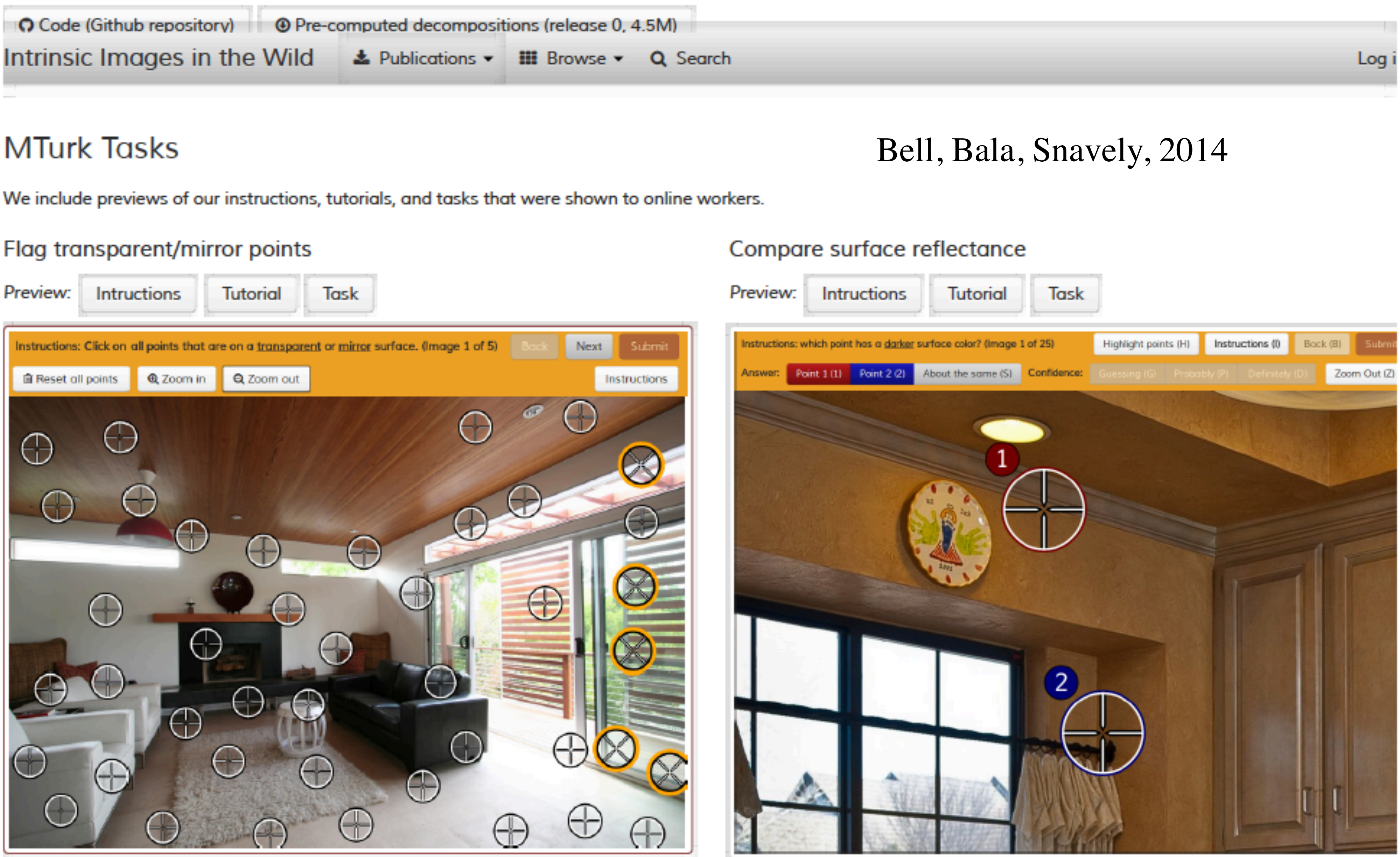
Implementation of Retinex
due to Kevin Karsch

Ground truth
images from dataset of Gosse et al. 09





Human judgements are easier



Bell, Bala, Snavely, 2014

This gives an evaluation task

- WHDR=Weighted Human Disagreement Ratio
  - compute lightness from intrinsic image representation at points
  - predict
    - A lighter than B
    - B lighter than A
    - Lightness match
  - compute weighted estimate of accuracy
    - weights low where human judgements are uncertain, high otherwise

- There are issues, but allows evaluation
  - and competition

Modern strategies - Optimization

- Apply the priors that
  - albedo is piecewise constant
  - there are “few” albedo values
  - albedo and shading explain image
- Solve
  - eg Bell 14, Nestmeyer 17, Bi 15

Modern strategies - Regression

- Regression of ground truth against image
  - use training set from WHDR data (Narihira et al 2015)
    - and perhaps rendered data
  - surprisingly, rendered data is very helpful
    - Li et al 18; Bi et al 18; Fan et al 18; etc
- Surprising because
  - Albedo in renderings isn't like albedo in the world
  - Illumination in renderings *really* isn't like illumination in the world

Recent history

Method	Source	Training uses IIW labels	Training uses CG	Flattening	Test WHDR
Shi *et al.* '17 [26]	[27]	N	Y	N	54.44
Zhou *et al* '15 [28]	[27]	Y	N	Y	19.95
Narihira *et al* [29]	ibid	N	N	N	18.1
Bi *et al* '18 [27]	ibid	N	Y	Y	17.18
Zhou *et al* '15 [30]	ibid	Y	N	Y	15.7
Li and Snavely '18 [31]	ibid	Y	Y	Y	14.8
Fan *et al* '18 [32]	ibid	Y	N	Y	14.45
*Zhao *et al.* '12 [14]	[29]	N	N	N	26.4
Shen and Yeo '11 [23]	[29]	N	N	N	26.1
Yu and Smith '19 [33]	ibid	N	N	N	21.4 (a)
Retinex (rescaled; color/gray)	[29]	N	N	N	19.5*/18.69*
Bell *et al* '14 [34]	[29]	N	N	Y	18.6
Liu *et al* '20 [35]	ibid	N	Y+	N	18.69
Bi *et al* '15 [36]	ibid	N	N	Y	18.1
Bi *et al* '15 [36]	[27]	N	N	Y	17.69
[redacted]					

TABLE 1

*Summary comparison to recent high performing supervised (above) and unsupervised (below) methods, all evaluated on the standard IIW test set; sources indicated. We distinguish between training with IIW and threshold selection using IIW. WHDR values computed for Retinex use the most favorable scaling, using the rescaling experiments of [29]. For our method, we report the held-out threshold value of WHDR. We report two figures for [36], because we found two distinct figures in the literature. Key: * - method uses IIW training data to set scale or threshold ONLY. + - [35] build models of albedo and shading from CGI, but does not use them for direct supervision. a - [33] use patches of registered images from MegaDepth.*

WHDR is tricky - I

Narihira et al 15

From Fan 18

Methods	WHDR (mean)
Baseline (const shading)	51.37
Baseline (const reflectance)	36.54
Shen *et al.* 2011 [17]	36.90
Retinex (color) [11]	26.89
Retinex (gray) [11]	26.84
Garces *et al.* 2012 [9]	25.46
Zhao *et al.* 2012 [20]	23.20
L_1 flattening [3]	20.94
Bell *et al.* 2014 [2]	20.64
Zhou *et al.* 2015 [21]	19.95
Nestmeyer *et al.* 2017 (CNN) [16]	19.49
Zoran *et al.* 2015* [22]	17.85
Nestmeyer *et al.* 2017 [16]	17.69
Bi *et al.* 2015 [3]	17.67
Ours w/o D-Filter	15.40
Ours w/o joint training	14.52
Ours	**14.45**

Table 1. Quantitative results on the IIW benchmark. All the results are evaluated on the test split of [15], except for the one marked with * which is evaluated on their own test split and is not directly comparable with other methods.

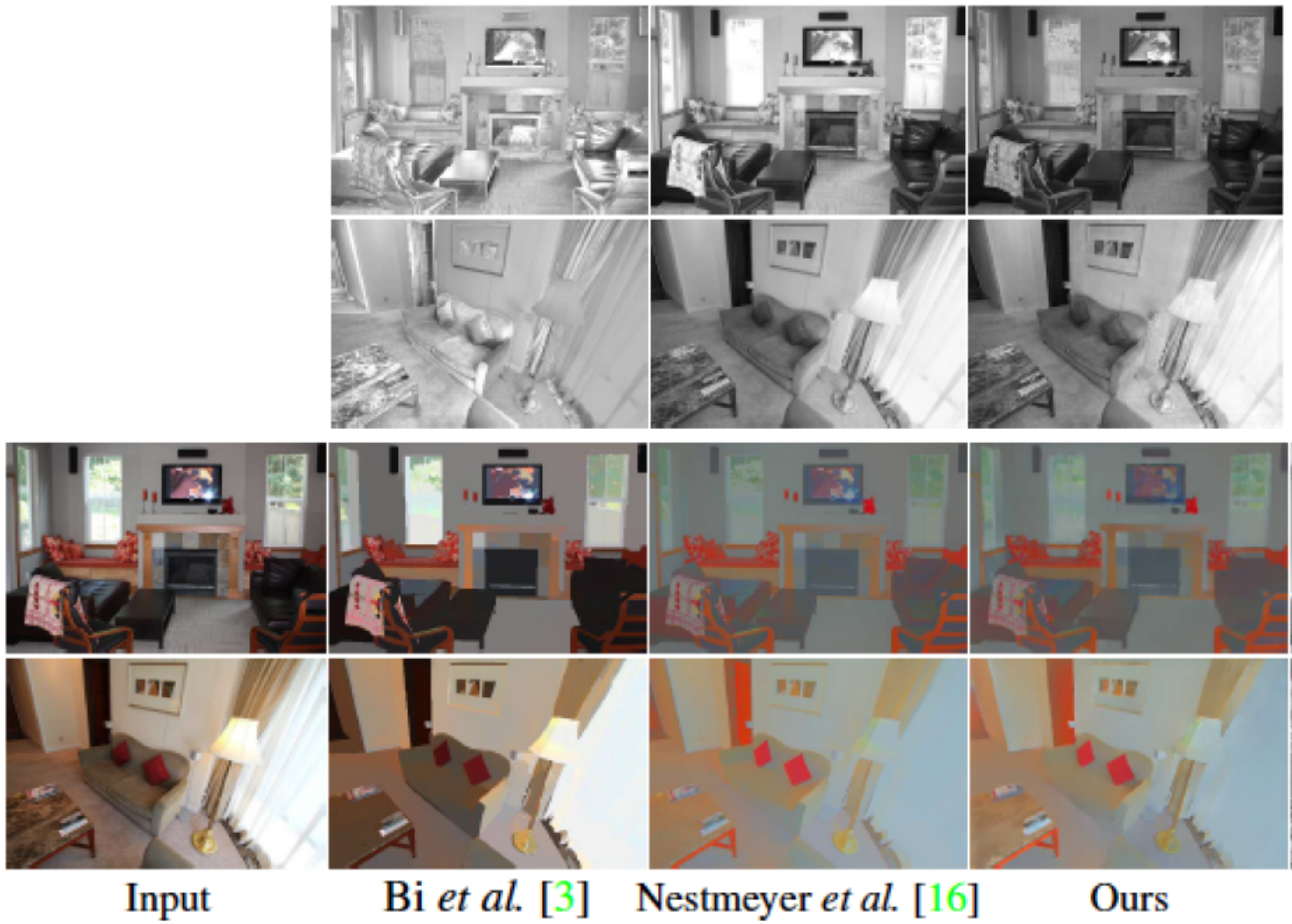
	WHDR (%)	Error Rate (%)
Ours (HSC)	20.9	24.5
Ours (CNN)	18.3	22.3
Ours (CNN-ImageNet)	**18.1**	**22.0**
CRF [4] (rescaled)	18.6	22.3
Retinex-Color [10] (rescaled)	19.5	23.3
Retinex-Gray [10] (rescaled)	19.8	23.8
Shen and Yeo [22] (rescaled)	23.2	26.1
Zhao *et al.* [26] (rescaled)	22.8	26.4
CRF [4]	20.6	25.6
Retinex-Color [10]	26.9	32.4
Retinex-Gray [10]	26.8	32.3
Shen and Yeo [22]	32.5	35.1
Zhao *et al.* [26]	23.8	28.2

Table 1. **Intrinsic Images in the Wild benchmark results.** For each algorithm, we display the weighted human disagreement rate (WHDR, lower is better), as well as the error rate on classifying the sign of lightness change between pairs of points labeled in the ground-truth. We include our own re-evaluation of competing methods, which closely matches the performance reported in [4]. In addition, we report performance of a rescaled version of competing methods, which specifically optimizes their output for the pairwise classification task. Our algorithm is on par with the CRF approach developed by [4] for state-of-the-art performance. We refer the reader to [4] for comparison to an expanded set of prior work.

WHDR is tricky - II

- Predict by
  - $f(m1, m2) > t$ -> 1 is lighter
  - $-t < f(m1, m2) < t$ -> same
  - $f(m1, m2) < -t$ -> 2 is lighter
- Issues:
  - choice of f
    - m1 - m2
    - log(m1/m2)-1
  - choice of m
    - lightness potential
    - predicted albedo
  - choice of threshold
    - interacts with scale

WHDR is tricky - III



Fan 18 - current SOTA WHDR of 14.45%

WHDR is tricky - IV

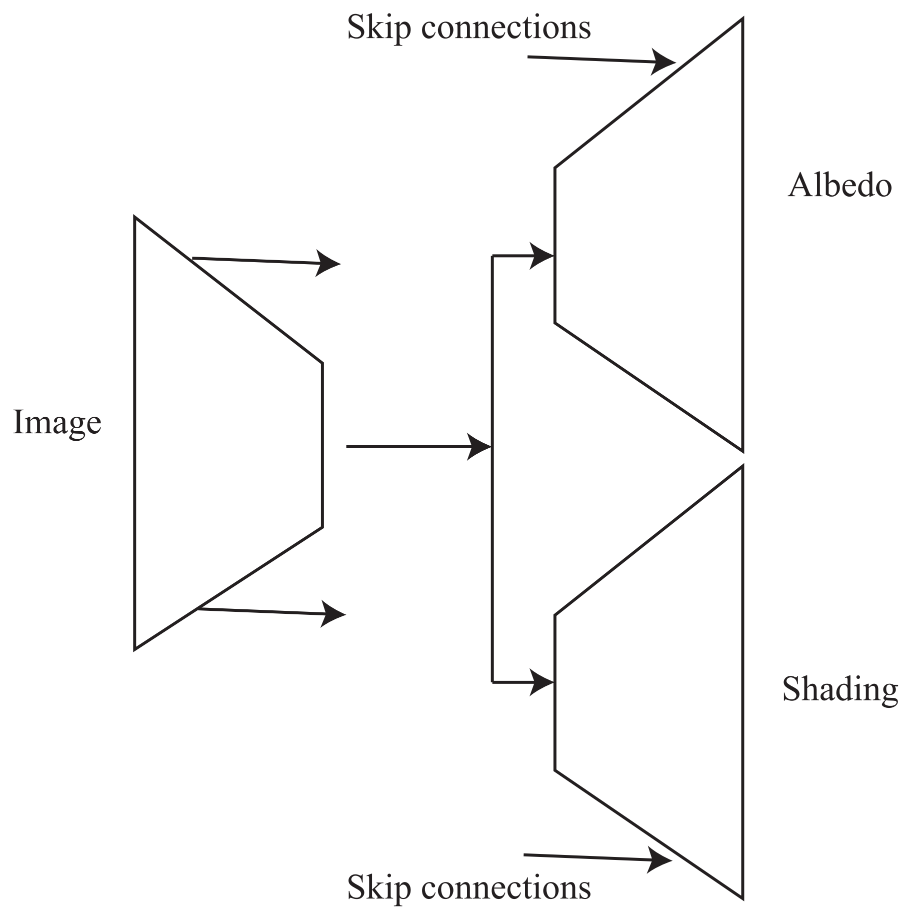
Bi et al, 2018 - this image WHDR 6.61%

WHDR: 75.70% Shi et al. [2017] | WHDR: 36.03% Narihira et al. [2015] | WHDR: 11.48% Zhou et al. [2015]

WHDR: 7.35% Nestmeyer et al. [2017] | WHDR: 6.61% Bi et al 2018

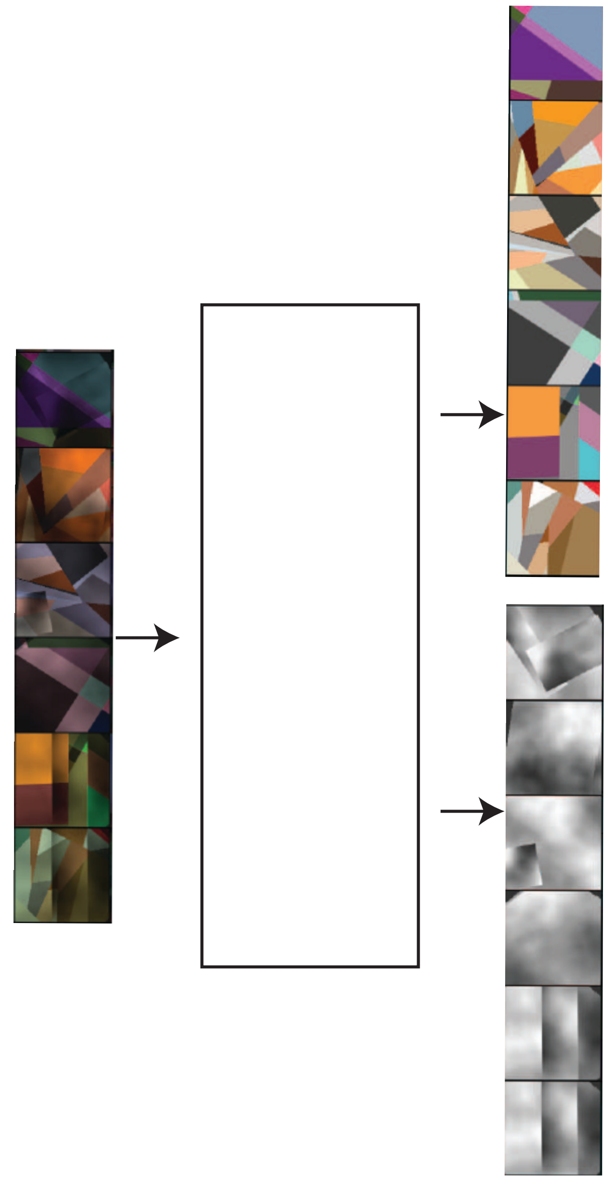
- Note:
  - odd colors
  - "colored paper" effect
  - "indecision"

One approach (local!)

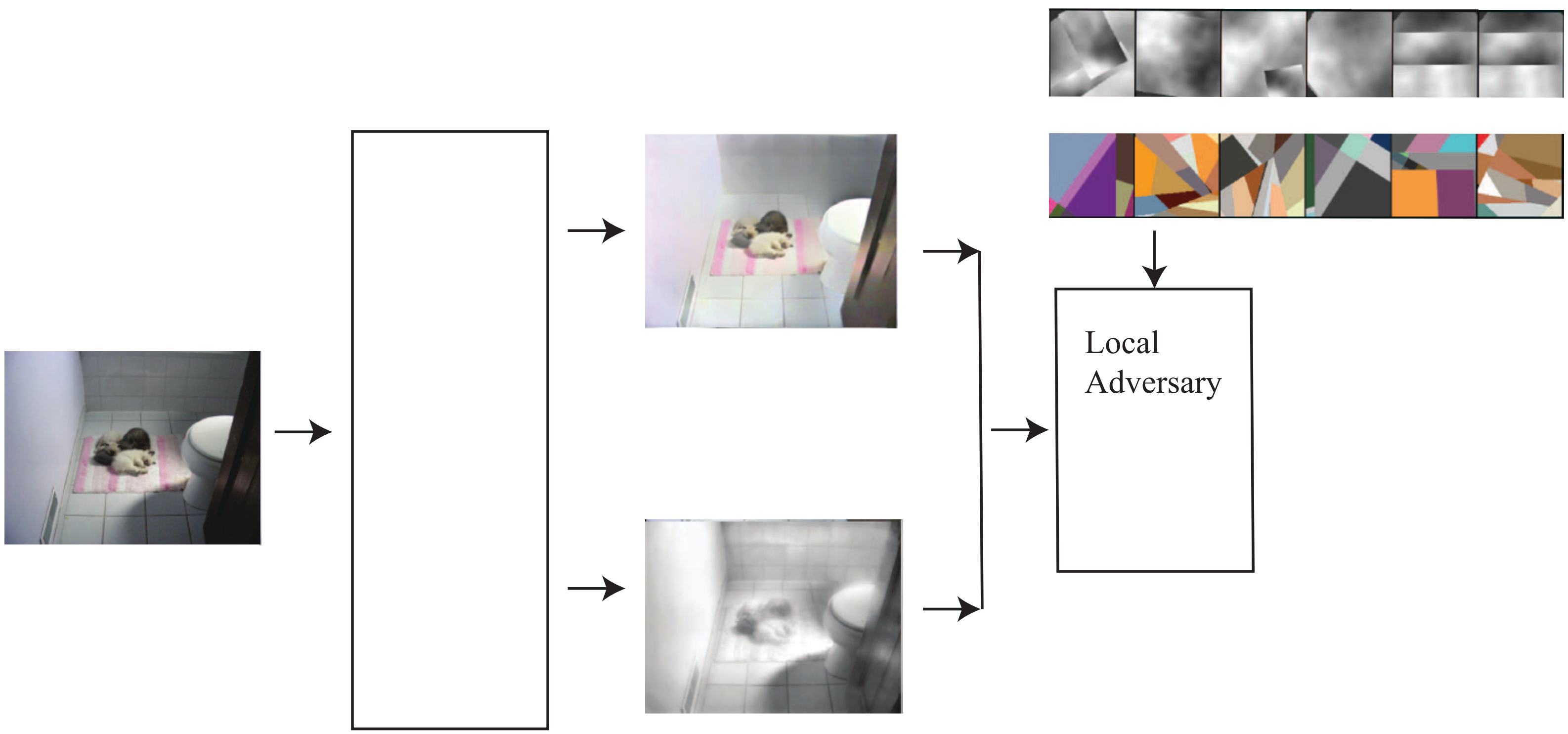


Training - I

Our albedo paradigm uses a surface color model and a spatial model. The qualitative properties it is intended to capture are: albedoes are piecewise constant; the color distribution should reflect likely surface colors; there should be a profusion of edges with no strong orientation bias; there should be at least some vertices with degree greater than three. Surface color is modelled by drawing color samples uniformly and at random from the IIW training set. These must be adjusted for presumed illumination. We do so by assuming the range of illumination intensity is approximately the same as the range of lightnesses, and so dividing by the square root of intensity.



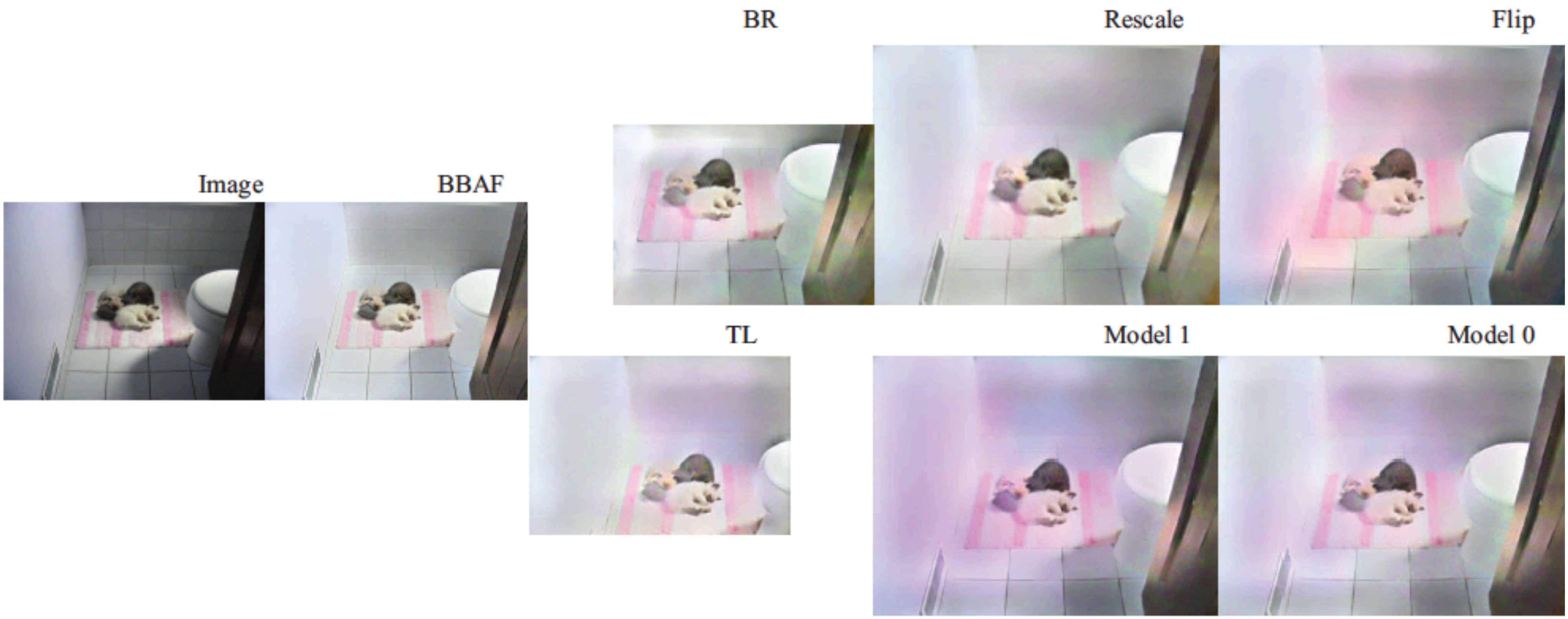
Training - II



Inference

- Network is trained on 128 x 128 tiles of image
- We want equivariance properties from albedo, shading
  - eg translate, rotate, scale image
    - albedo for translated (etc) image should be translated albedo
    - shading for translated (etc) image should be translated shading
- This doesn't come naturally

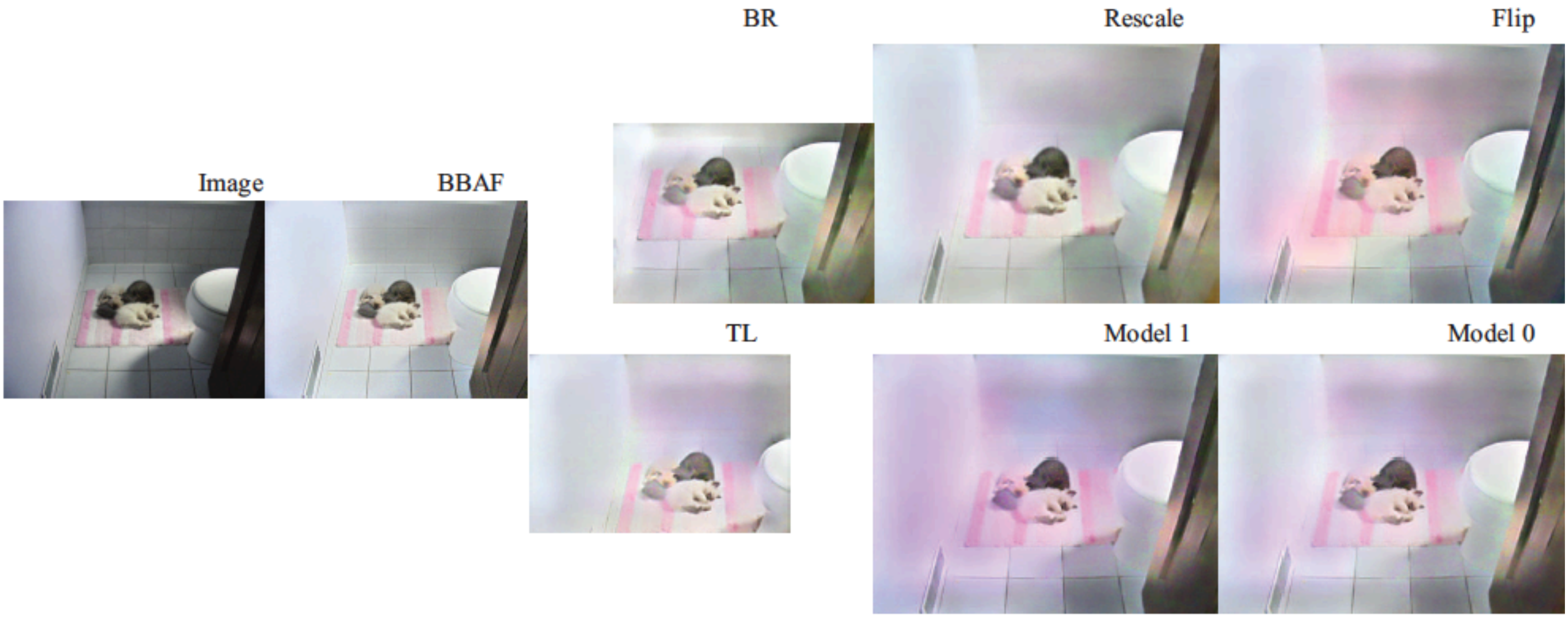
Equivariance must be imposed



Imposing equivariance

- Translation:
  - cover image with many, shifted, overlapping tiles
  - for each, recover albedo, shading
    - albedo at pixel is weighted average of all overlapping tiles
- Scale:
  - rescale image up, down
    - for each, recover albedo/shading using translation averaging
    - then rescale back
  - average results
- Rotation
  - average estimates from above over 8 flips

Averaging very strongly suppresses error



Results

Method	Source	Training uses IIW labels	Training uses CG	Flattening	Test WHDR
Shi *et al.* '17 [26]	[27]	N	Y	N	54.44
Zhou *et al* '15 [28]	[27]	Y	N	Y	19.95
Narihira *et al* [29]	ibid	N	N	N	18.1
Bi *et al* '18 [27]	ibid	N	Y	Y	17.18
Zhou *et al* '15 [30]	ibid	Y	N	Y	15.7
Li and Snavely '18 [31]	ibid	Y	Y	Y	14.8
Fan *et al* '18 [32]	ibid	Y	N	Y	14.45
*Zhao *et al.* '12 [14]	[29]	N	N	N	26.4
Shen and Yeo '11 [23]	[29]	N	N	N	26.1
Yu and Smith '19 [33]	ibid	N	N	N	21.4 (a)
Retinex (rescaled; color/gray)	[29]	N	N	N	19.5*/18.69*
Bell *et al* '14 [34]	[29]	N	N	Y	18.6
Liu *et al* '20 [35]	ibid	N	Y+	N	18.69
Bi *et al* '15 [36]	ibid	N	N	Y	18.1
Bi *et al* '15 [36]	[27]	N	N	Y	17.69
Our **BBA**		N	N	N	17.04*
Our **BBAF**		N	N	N	17.11*

TABLE 1

*Summary comparison to recent high performing supervised (above) and unsupervised (below) methods, all evaluated on the standard IIW test set; sources indicated. We distinguish between training with IIW and threshold selection using IIW. WHDR values computed for Retinex use the most favorable scaling, using the rescaling experiments of [29]. For our method, we report the held-out threshold value of WHDR. We report two figures for [36], because we found two distinct figures in the literature. Key: * - method uses IIW training data to set scale or threshold ONLY. + - [35] build models of albedo and shading from CGI, but does not use them for direct supervision. a - [33] use patches of registered images from MegaDepth.*

Fig. 2. *Qualitative examples, from our best model (*BBAF*), showing (L to R): suppression of indoor shadows; suppression of backscatter from shiny bathroom fittings; suppression of fast shading effects from clothing folds; correctly handled dark shadow (couch back).*

Fig. 6. *Qualitative comparison to [27], [26], [48], [45] and [62], using parts of Figure 1 of [27]. As [27] remark, the methods of [26] and [48] are trained on rendered data alone, and face difficulties due to the difference between rendered data and real images. As [27] remark, the methods of [48] and [45] face difficulties due to the deep shadows in the scene. The albedo produced by our method does not show the "colored paper" effect seen in other methods and does not produce odd colors; this is an advantage (text). Our method reports albedo and shading up to image boundaries, that of [27] appears not to (the crop of the figures is as in the original paper; for our method, we show the whole image).*

Smoothing is important

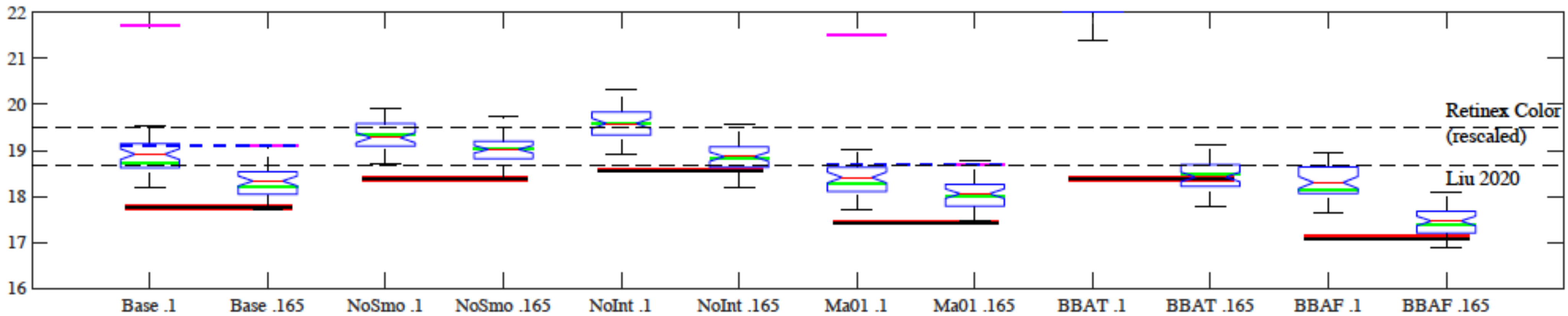


Fig. 8. *Smoothing, averaging and postprocessing are important. Without adversarial smoothing (***NoSmo***), performance is comparable to Retinex. Adversarial smoothing* alone *(***NoInt***) is surprisingly well behaved. Averaging makes a very significant difference (compare blue/black bars and purple/green bars) and averaging over a larger number of tiles is better (cf.* **BBA** *and* **Base***). Discrete image averaging results in improvements (cf.* **BBA** *and* **BBAF***), and is clearly better than discrete tile averaging (cf.* **BBAF** *and* **BBAT***).* **Key:** Fixed thresholds: *shown in boxplots of WHDR values for 50 simulated test sets for the two fixed thresholds, and green bars are the value for the standard test set.* Oracle thresholds: *heavy black bar.* Held out threshold: *heavy red bar.* Oracle threshold without smoothing: *heavy blue dashed bar.* Fixed threshold without smoothing: *heavy purple bar.* Boxplots: *horizontal bar = median; notch = fraction of interquartile range outside which a difference in medians is significant; bottom and top of the box = 25 and 75 percentiles resp.; whiskers extend to the most extreme data points that are not outliers; outliers – greater than 1.5 times the interquartile range outside top and bottom – are '+'.* Best viewed in color.

Paradigms beat graphics

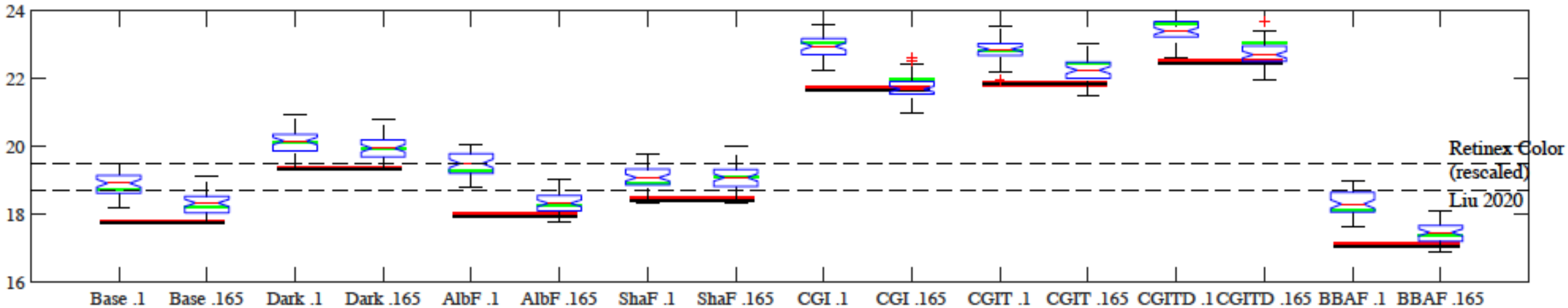


Fig. 9. *Varying the details of the paradigm has some effect; a* Dark *shading paradigm creates notable difficulties, but varying the size of shading (*ShaF*) and albedo (*ShaF*) fragments seems to have only minor effects. Using tiles excerpted from CGIntrinsics [47] leads to significant fall off in performance (*CGI *– tiles extracted from CGIntrinsics at original scale;* CGIT *– extracted from images shrunk so that tiles contain more detaile;* CGITD *– dependency between shading and albedo preserved). Graphical conventions as in Figure 5.* Best viewed in color.

Scale matters

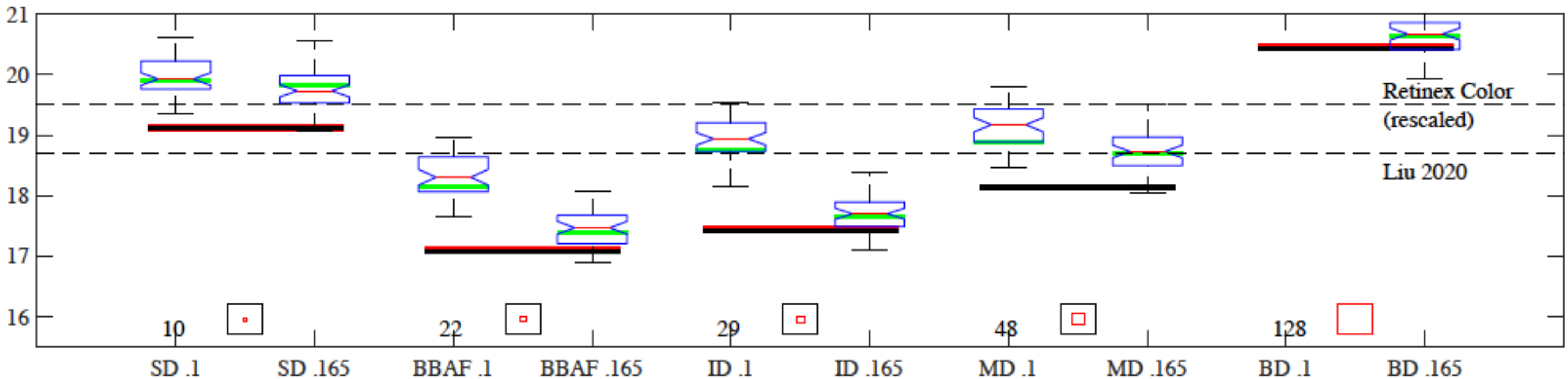


Fig. 10. *Varying the scale of the discriminator has an important effect on performance.* **SD** *the discriminator sees* 10×10 *patches;* **BBAF** *as in other figures our best model,* 22×22*;* **ID** 29×29*;* **MD** 48×48*; and* **BD** 128×128*. The scale of* **ID** *was chosen by interpolating oracle WHDR for the others, then choosing the scale that produced the best predicted WHDR. The red boxes show the scale of the discriminator patches with respect to the tile (black boxes) for each model. Graphical conventions as in Figure 5.* Best viewed in color.

Indecisiveness remains (aargh!)

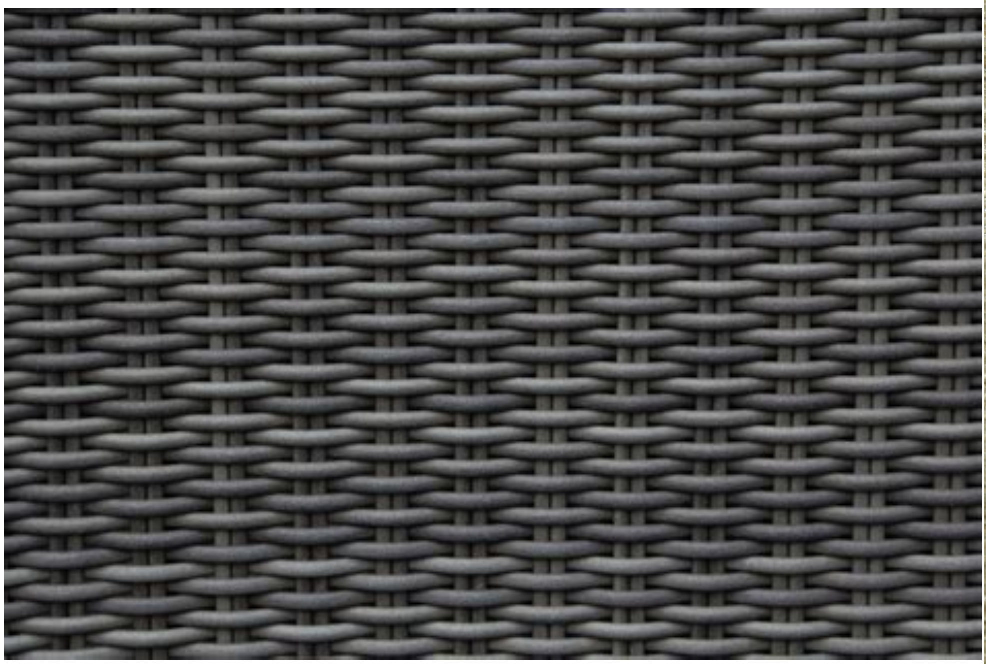
Fig. 13. *Our method suffers indecisiveness, as do others; this is a persistent problem in intrinsic image methods. Figures show a decomposition of an outdoor image, using our method. Note the pronounced shadow leaves effects in both albedo and shading fields; versions of this effect for other methods can be seen in Figure 6.* Best viewed in color.

Other Possible Intrinsics

- Surface relief and material properties
  - and perhaps many of them
- Surface mechanical properties
- Surface glossiness
- Texture flow

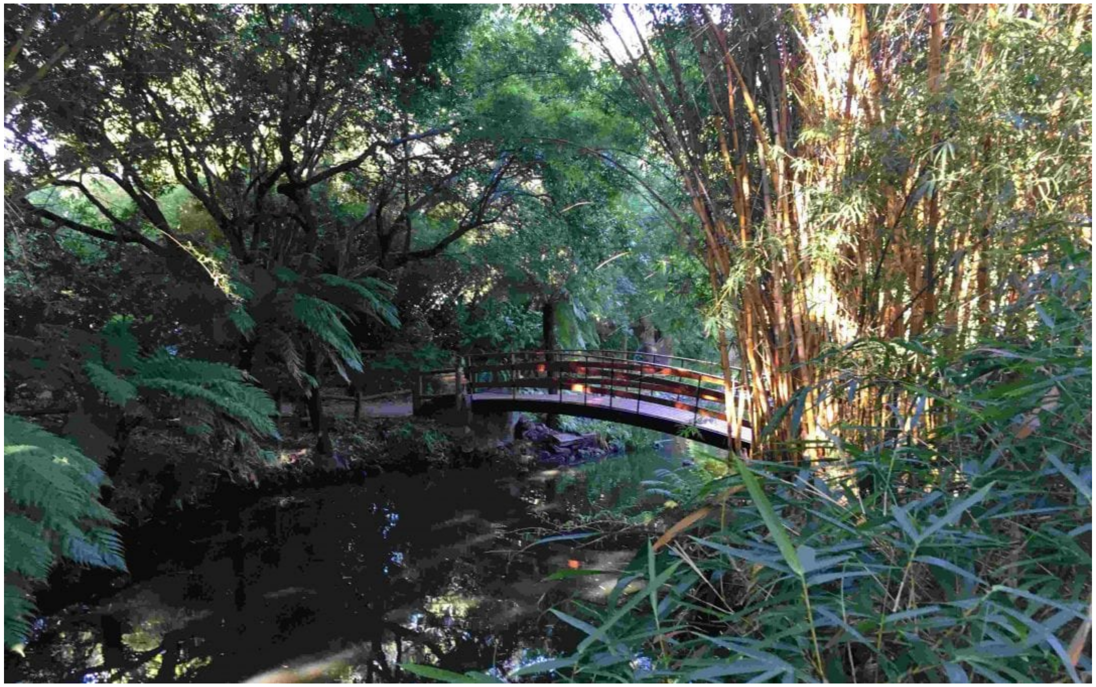
Relief - intrinsic, because
small local shadows do not
move with illumination
(at least Koenderink+Van Doorn, 77)

Relief - intrinsic, because
small local shadows do not
move with illumination
(at least Koenderink+Van Doorn, 77)



Fur - intrinsic, because
small local shadows do not
move with illumination
(at least Koenderink+Van Doorn, 77)

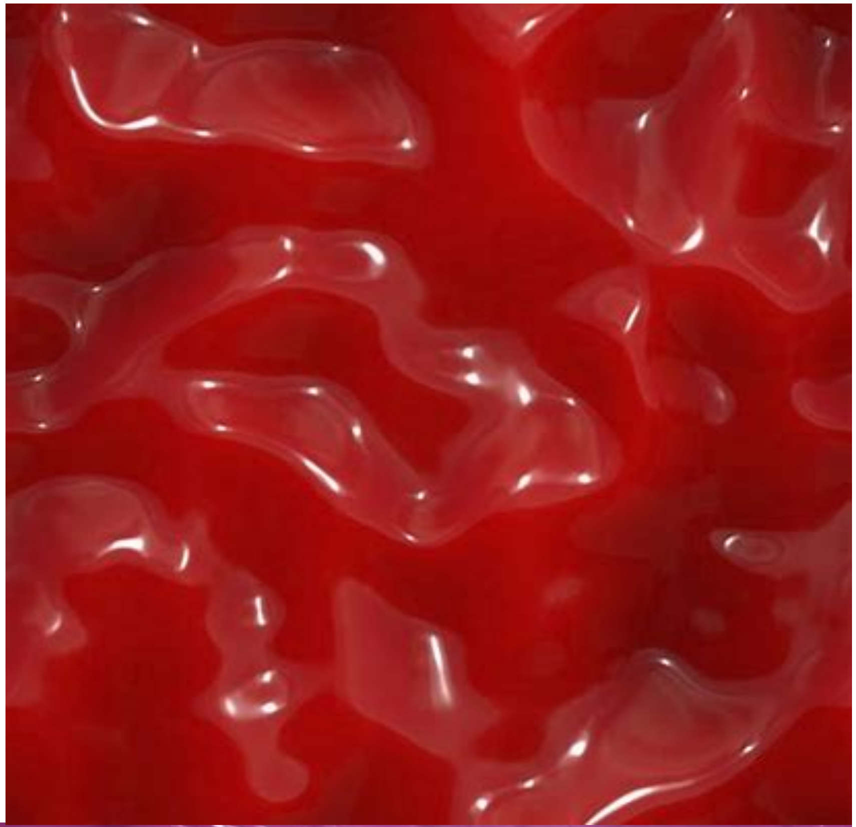
Relief - intrinsic (at least at this scale), because small local shadows do not move with illumination (at least Koenderink+Van Doorn, 77)

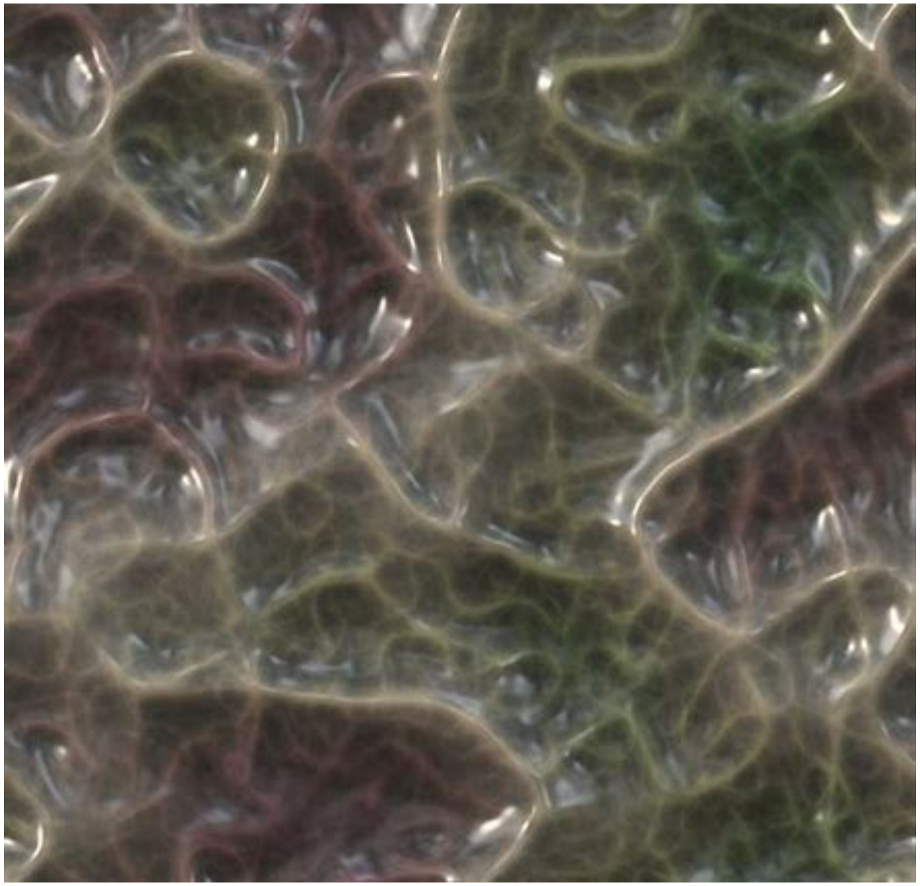


??? - intrinsic, because
mostly not a property of viewing
circumstances (?)

Iridescence
creating intrinsic gloss effects
intrinsic because the color effects will be
there for almost all illumination

??? - intrinsic, the specularities move but are always there





??? - intrinsic, the specularities move but are always there

Other Possible Extrinsics

- Glossy reflected component
- Luminaires
- Lens flare
- Rain effects
- etc.

Gloss/specular - clearly extrinsic, when the light moves, this moves

Lens flares - clearly intrinsic, product of viewing circumstances

Luminaires -
extrinsic or intrinsic?
worth knowing about, anyhow

Rain - multiple extrinsic phenomena,
including smoothing, raindrops, loss of saturation,
glossy/wet surfaces, etc. etc.

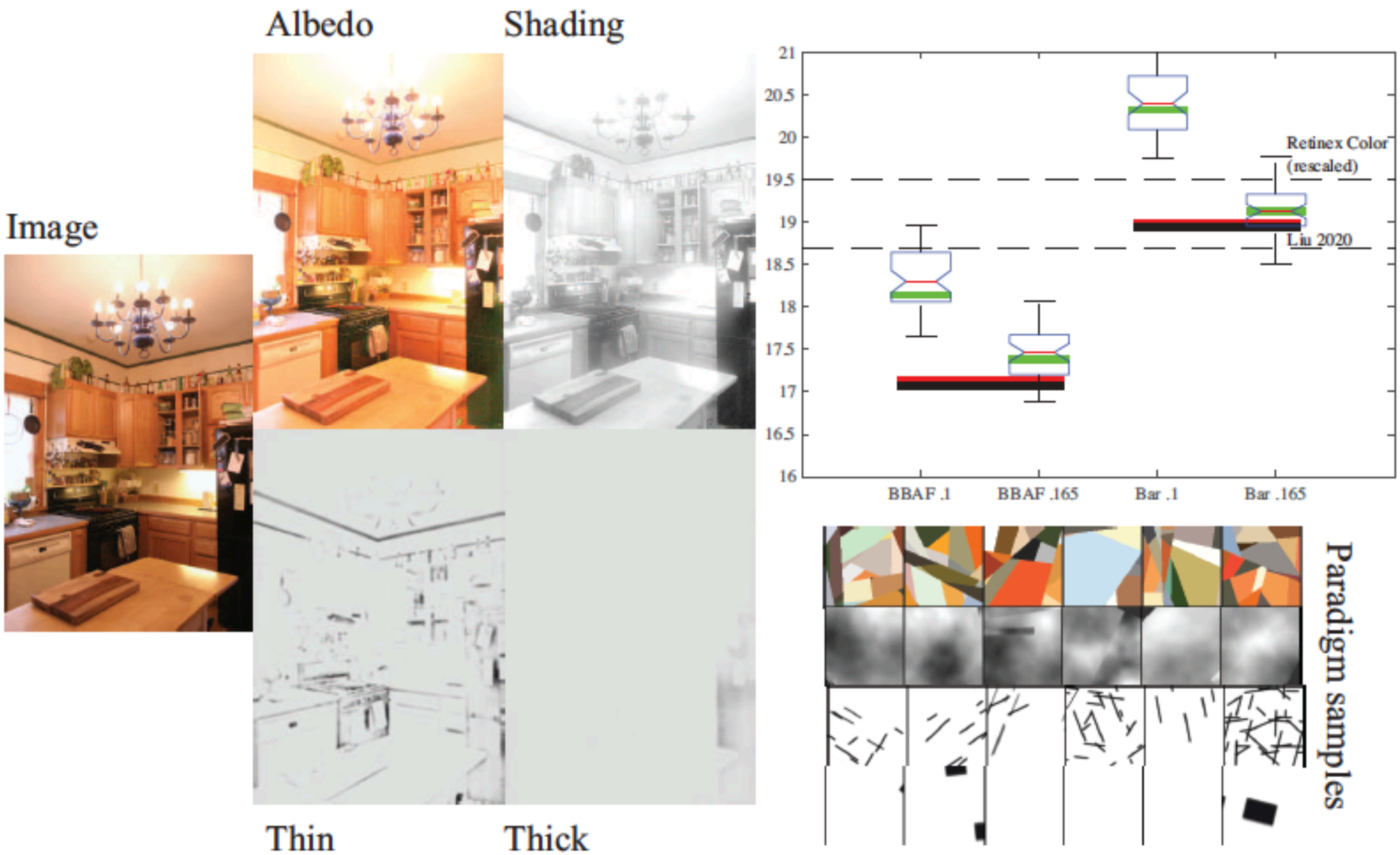


Fig. 12. *The method can be extended to capture thin and thick bars of darkness by extending the decomposer to have four heads (albedo, shading, thin bars, thick bars), and extending the paradigms (bottom left shows examples). The advantage of doing so is that a decomposition will then capture the thin bars of darkness associated with grooves separately from albedo (example decomposition shown here). Qualitatively, these thin bars do appear to be associated with grooves (but note the thin dark paint bars on the ceiling, which also appear in this map). The cost in WHDR (top right compares to BBAF) is noticeable, but may be tolerable in some applications.* Best viewed in color.

No ground truth decompositions

- And there never will be
  - rendering is do-able (but hard)
  - modelling is hopeless
- Q: how do you train an image decomposition method when you don’t know the right answer?
- Retinex provides clues - spatial statistics are the key