6.801/6.866 Machine Vision

Syllabus

#	Date	Description	Readings	Assignments	Materials
11	10/10	Texture and Edges	Req: FP 9 Opt: FP 8	Exam #1 Assigned	Freeman Slides
	10/15	Columbus Day (NO LECTURE)			
12	10/17	Bayesian Analysis and Optic Flow	Req: H 12		
13	10/22	Direct SFM	Req: H 17	Exam #1 Due	
14	10/24	Affine Reconstruction	Req: FP 12	Pset #3 Assigned	

Today

- Edges
- Bayes
- Motion analysis

Mid-term exam

Problem set 3

- Open book, open web.
- Work by yourself. But you can ask us questions for clarification.
- Due Tuesday, Oct. 22 (in 5 days).

6.866 projects

- Proposals to us by Oct. 29 or earlier.
- We will ok them by Oct. 31
- 3 possible project types:
  - Original implementation of an existing algorithm
  - Rigorous evaluation of existing implementation.
  - Synthesis or comparison of several research papers.

6.866 projects, continued

- Some possible projects
  - Study conditions on shape and reflectance maps such that shape is interpretable from rendered image.
  - Pose and solve a problem: make an algorithm that detects broken glass, or that finds trash. Implement and evaluate it.
  - Evaluate accuracy of photometric stereo shape reconstructions.
  - Compare several motion estimation algorithms. Discuss how they're different, the benefits of each, etc. Put them in a common framework.

6.801/6.866 Machine Vision

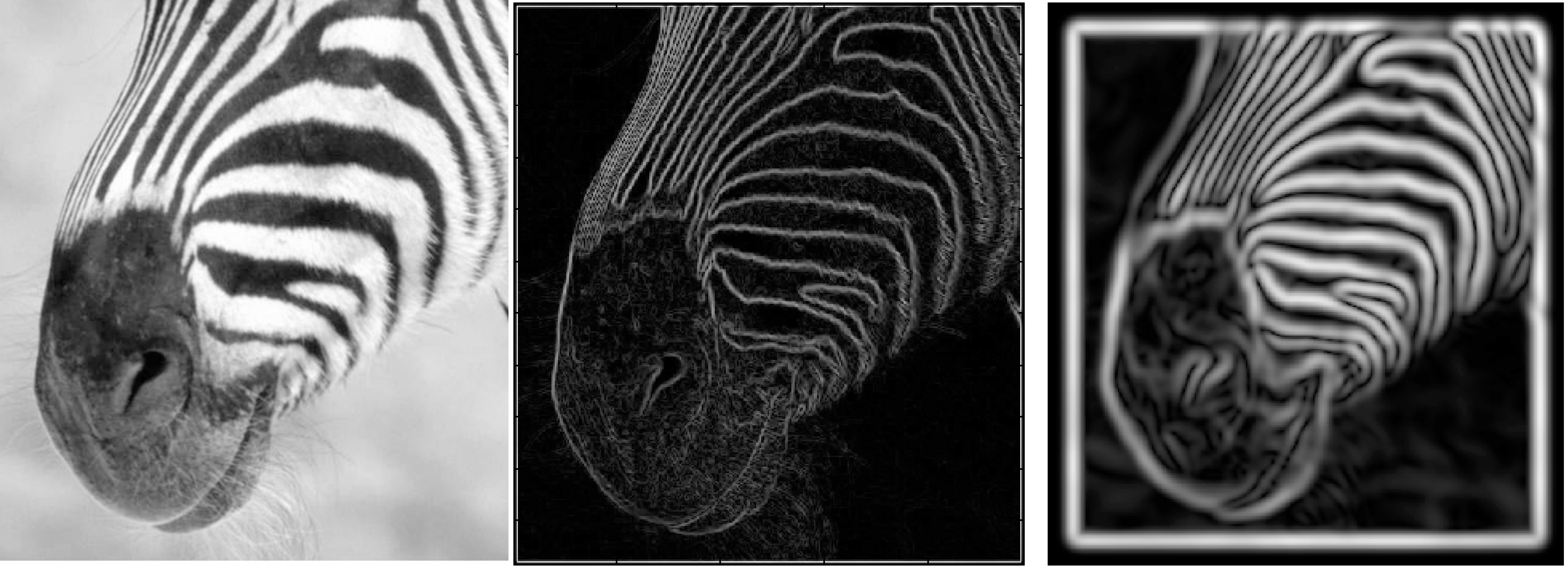
Syllabus

#	Date	Description
1	9/5	Course Introduction
2	9/10	Cameras, Lenses, and Sensors
3	9/12	Radiometry and Shading Models I
4	9/17	Radiometry and Shading Models II
5	9/19	Multiview Geometry
6	9/24	Stereo
7	9/26	Color
8	10/1	Shape from Shading
9	10/3	Image Filtering
10	10/8	Image Representations
11	10/10	Texture and Edges
	10/15	Columbus Day (NO LECTURE)
12	10/17	Bayesian Analysis and Optic Flow
13	10/22	Direct SFM
14	10/24	Affine Reconstruction
15	10/29	TBD
16	10/31	Statistical Classifiers I
17	11/5	Statistical Classifiers II
18	11/7	Projective Reconstruction
19	11/12	Clustering & Segmentation
20	11/14	EM
21	11/19	Hough Transforms, RANSAC, & Voting Methods
22	11/21	Tracking & Density Propagation
23	11/26	Model-Based Vision
	11/28	Thanksgiving (NO LECTURE)
24	12/3	Visual Surveillance / Activity Monitoring
25	12/5	Image Databases
26	12/10	Image-Based Rendering

Gradients and edges

- Points of sharp change in an image are interesting:
  - change in reflectance
  - change in object
  - change in illumination
  - noise
- Sometimes called **edge points**

- General strategy
  - determine image gradient
  - now mark points where gradient magnitude is particularly large wrt neighbours (ideally, curves of such points).



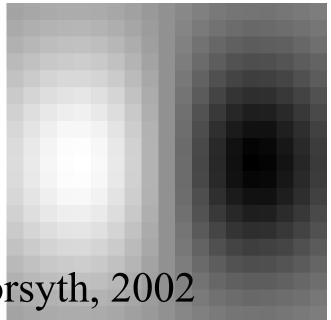
There are three major issues:

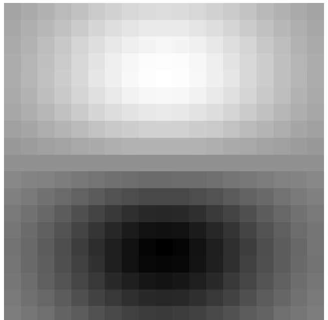
1) The gradient magnitude at different scales is different; which should we choose?
2) The gradient magnitude is large along thick trail; how do we identify the significant points?
3) How do we link the relevant points up into curves?

Forsyth, 2002

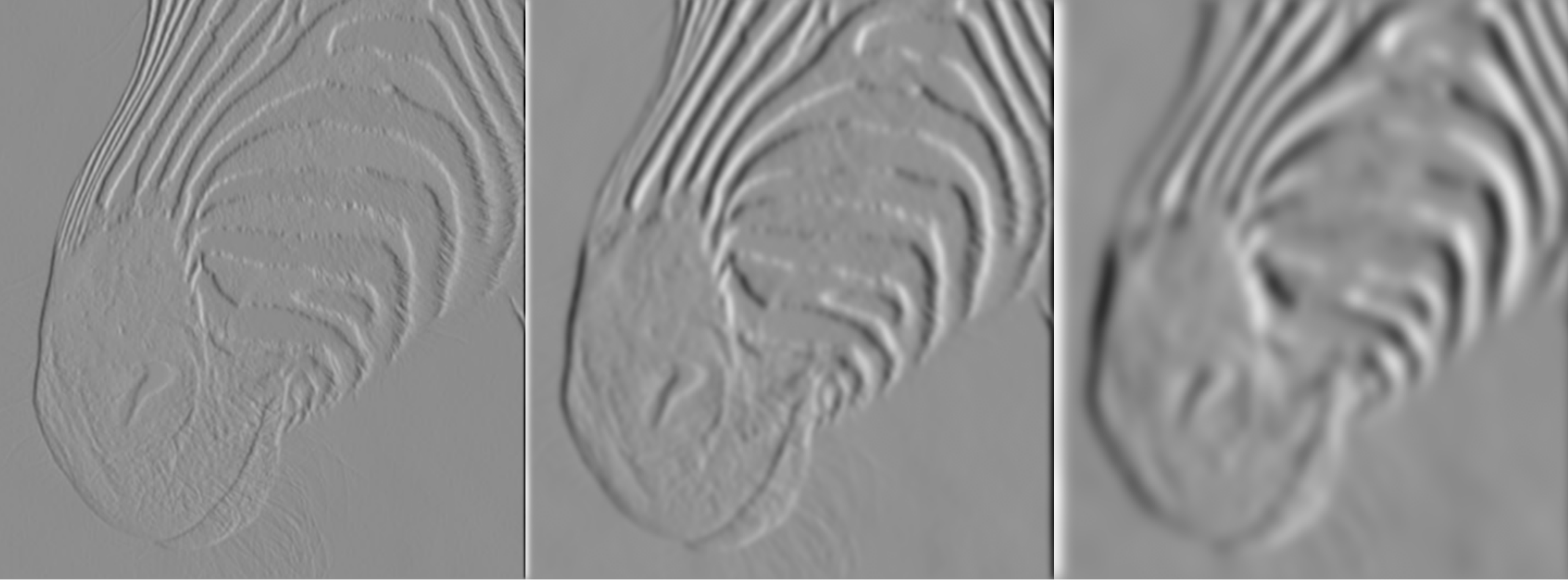
Smoothing and Differentiation

- Issue: noise
  - smooth before differentiation
  - two convolutions to smooth, then differentiate?
  - actually, no - we can use a derivative of Gaussian filter
    - because differentiation is convolution, and convolution is associative





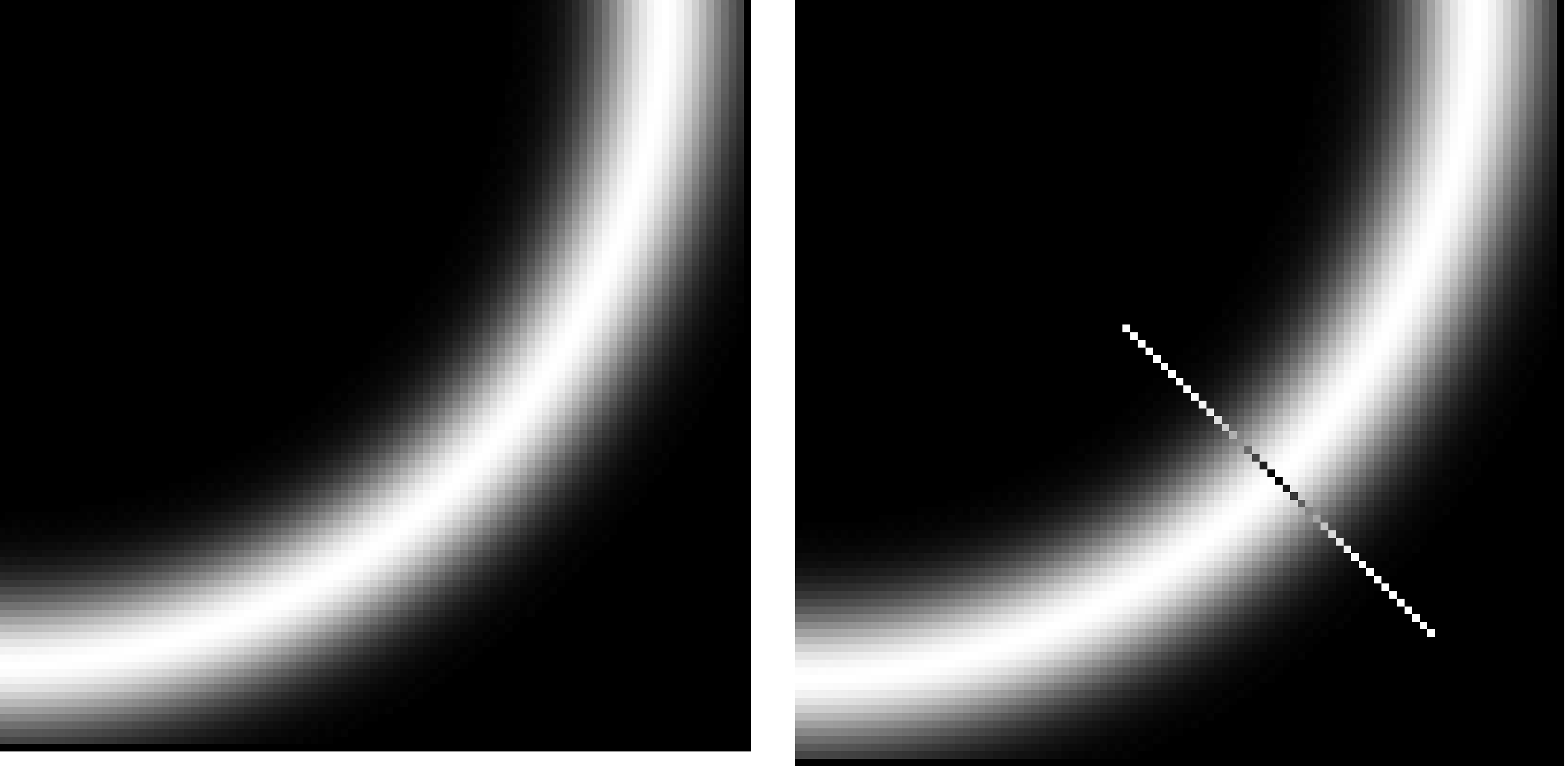
Forsyth, 2002



1 pixel　　3 pixels　　7 pixels

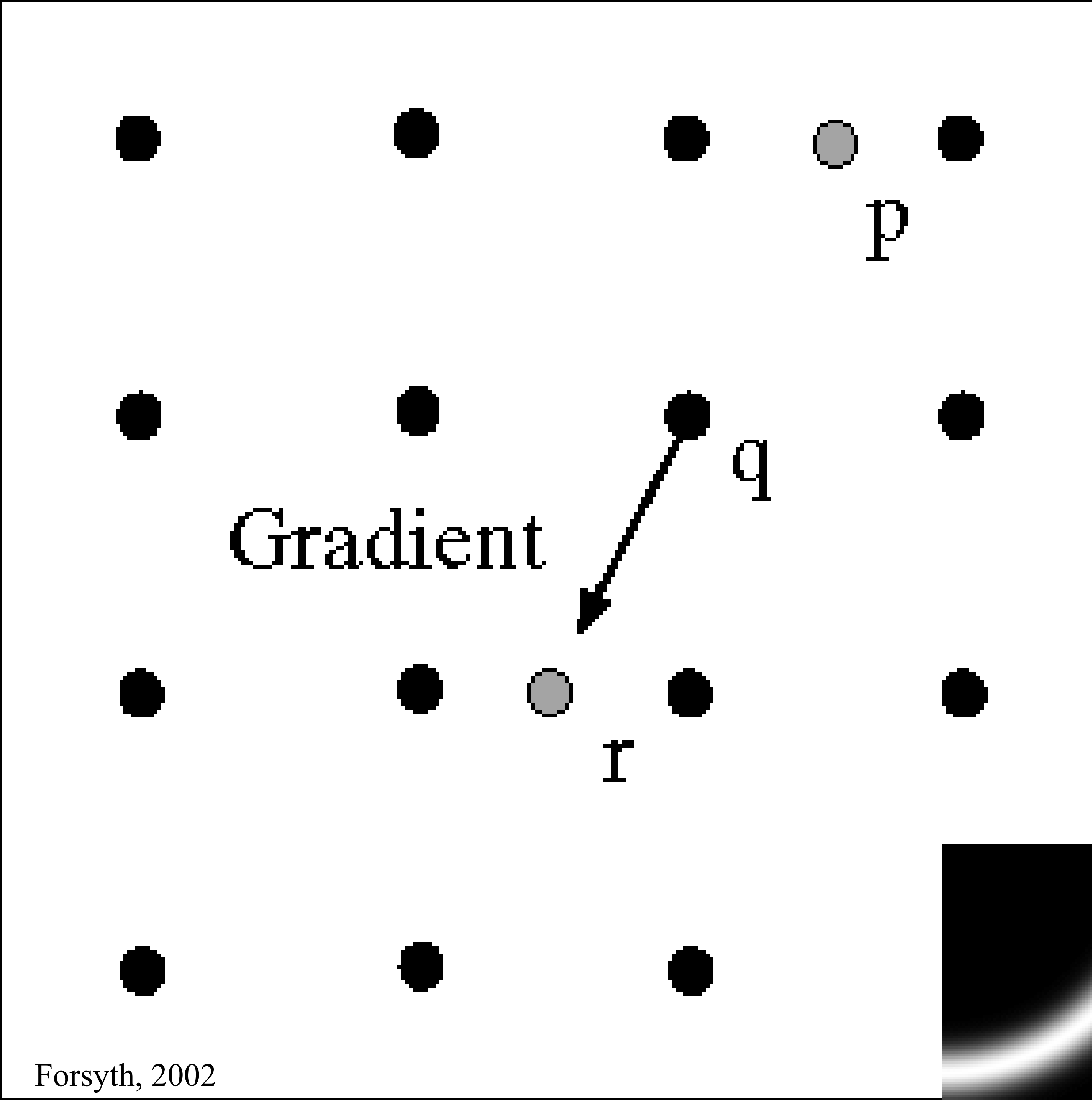
The scale of the smoothing filter affects derivative estimates, and also the semantics of the edges recovered.

Forsyth, 2002



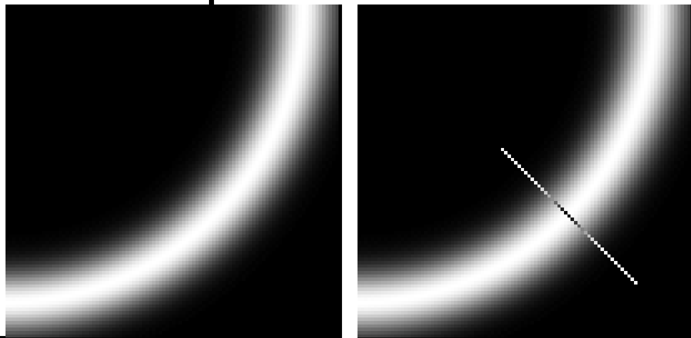
We wish to mark points along the curve where the magnitude is biggest. We can do this by looking for a maximum along a slice normal to the curve (non-maximum suppression). These points should form a curve. There are then two algorithmic issues: at which point is the maximum, and where is the next one?

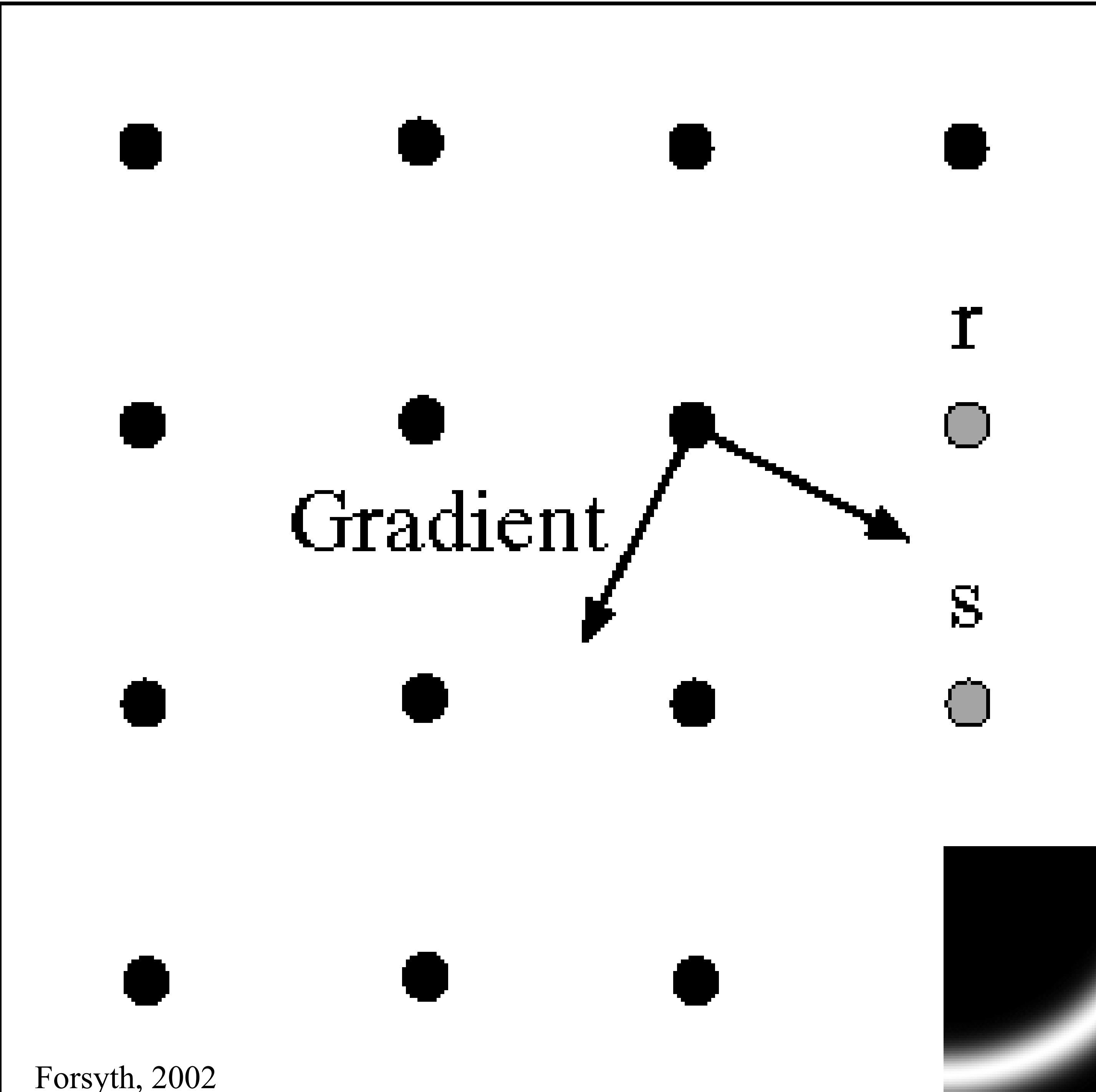
Forsyth, 2002



Non-maximum suppression

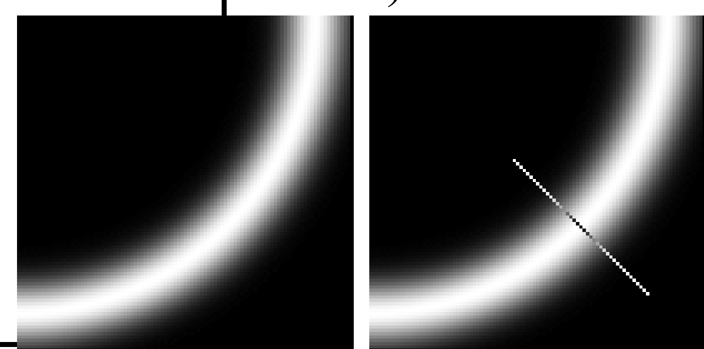
At q, we have a maximum if the value is larger than those at both p and at r. Interpolate to get these values.





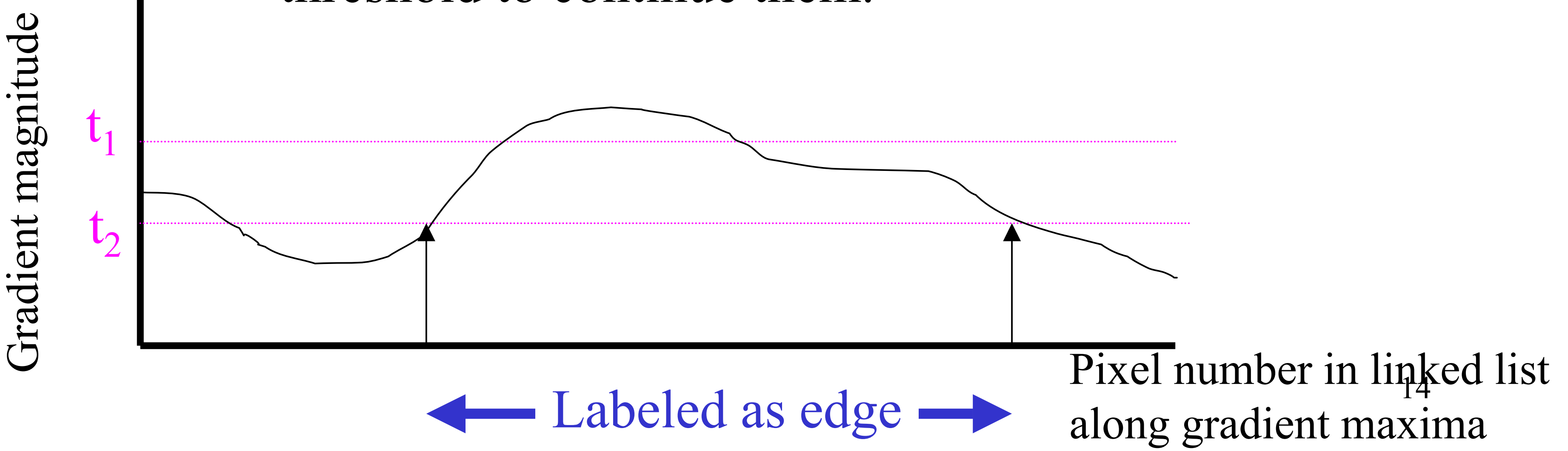
Predicting the next edge point

Assume the marked point is an edge point. Then we construct the tangent to the edge curve (which is normal to the gradient at that point) and use this to predict the next points (here either r or s).



Remaining issues

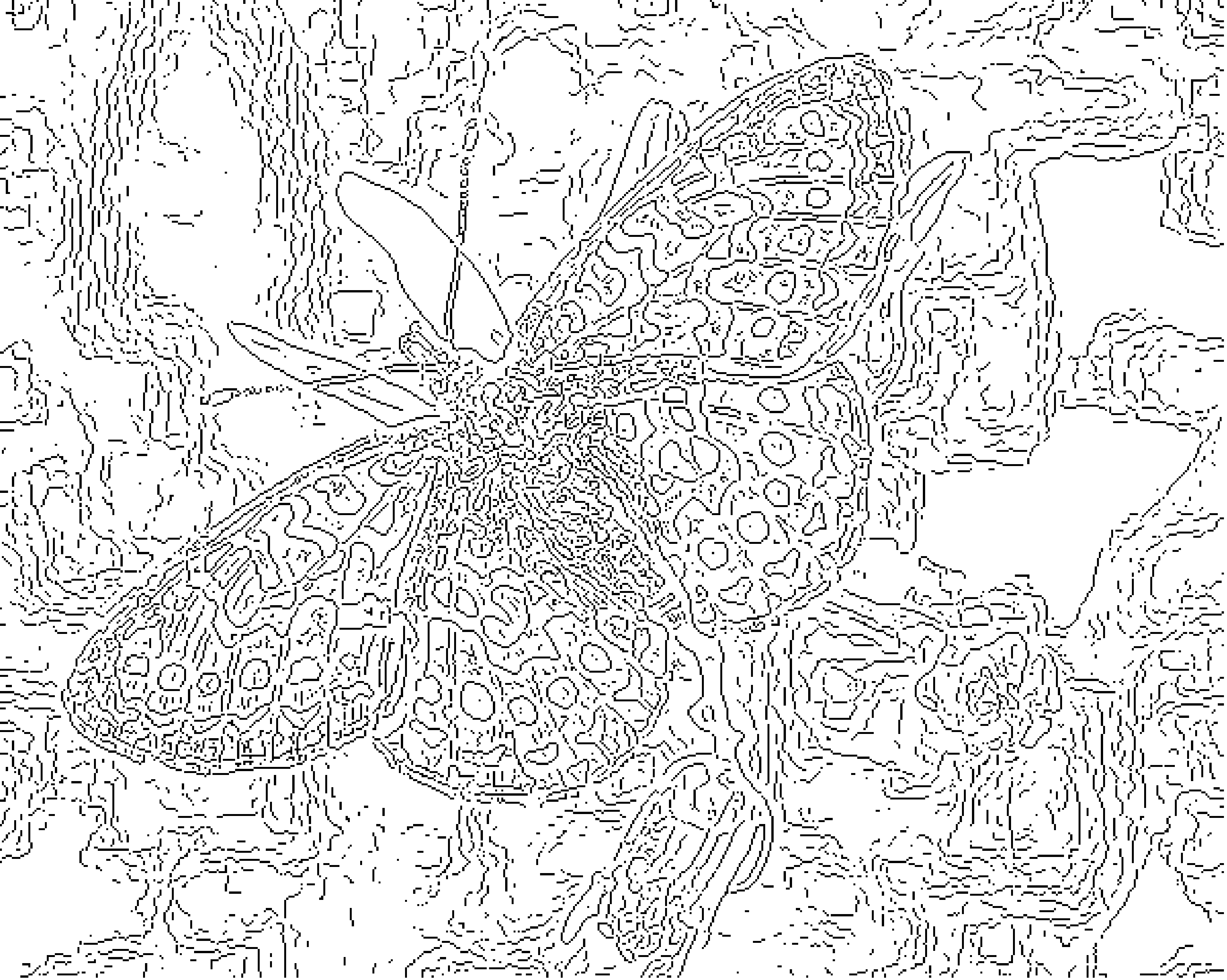
- Check that maximum value of gradient value is sufficiently large
  - drop-outs? use **hysteresis**
    - use a high threshold to start edge curves and a low threshold to continue them.



Notice

- Something nasty is happening at corners
- Scale affects contrast
- Edges aren’t bounding contours

Forsyth, 2002



fine scale
high
threshold

Forsyth, 2002

coarse
scale,
high
threshold

Forsyth, 2002

coarse
scale
low
threshold

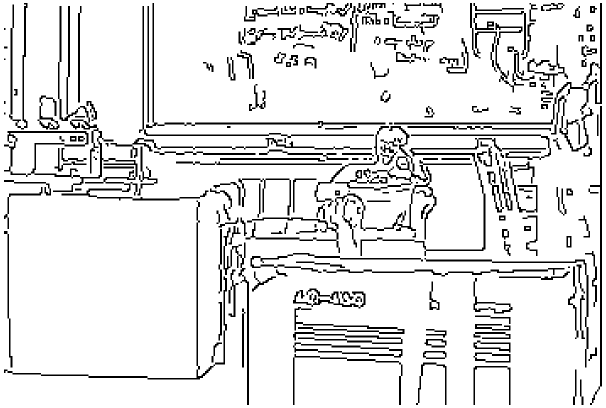
Forsyth, 2002

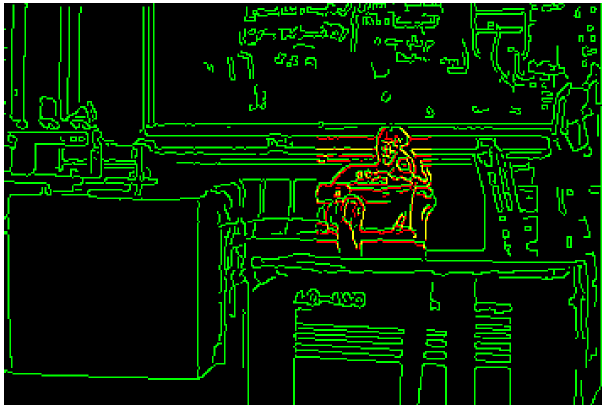
edges

- Issues:
  - On the one hand, what a useful thing: a marker for where something interesting is happening in the image.
  - On the other hand, isn’t it way to early to be thresholding, based on local, low-level pixel information alone?

Something useful with edges

 is a two-dimensional bitmap that serves as a model view of an object (Kevin sitting on a couch)

 is a two-dimensional bitmap (intensity edges) in which we want to locate this model, under the transformation of two-dimensional translation and scaling (that is the model is allowed to move in x and y, and also to scale separately in each of these dimensions, for a total of four transformation parameters).

 shows the best transformation (translation and scaling) of the model with respect to the image, in the sense that it maximizes the fraction of model edge points that lie near image edge points (within 1 pixel diagonally). The green points are image egdes, the red points are transformed model edges, and the yellow points are locations where both an image edge and a transformed model edge are coincident. Note that there are many red locations adjacent to green ones (which would not be detected by a method such as binary correlation).

Dan Huttenlocher

http://www.cs.cornell.edu/~dph/hausdorff/hausdorff1.html

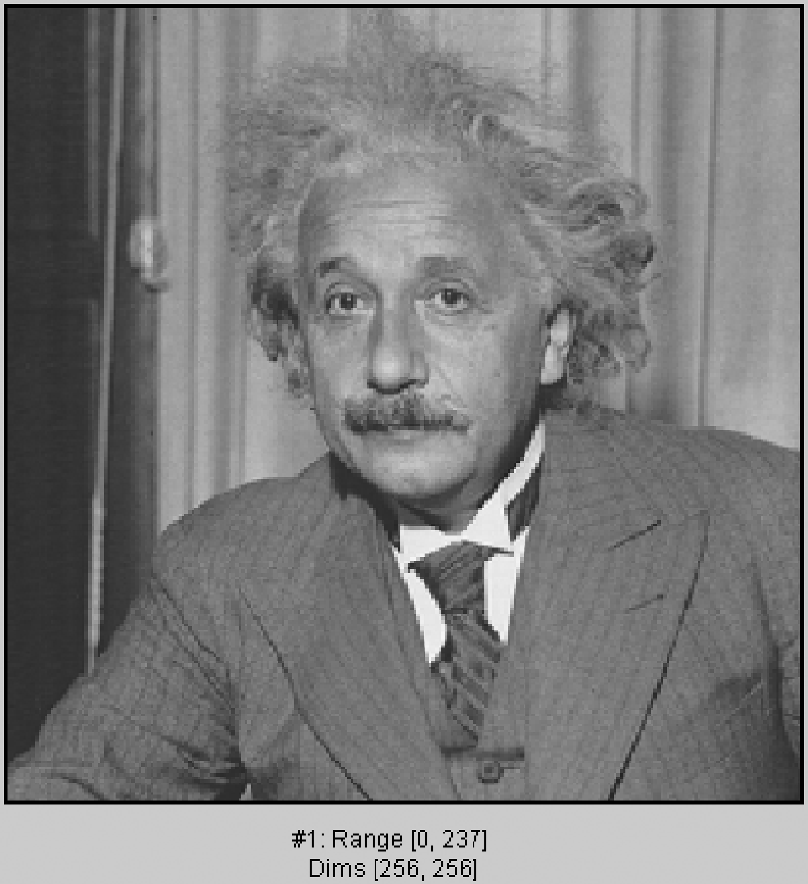
Another useful, bandpass-filter-based, non-linear operation: Contrast normalization

- Maintains more of the signal, but still does some gain control.
- Algorithm: bp = bandpassed image.

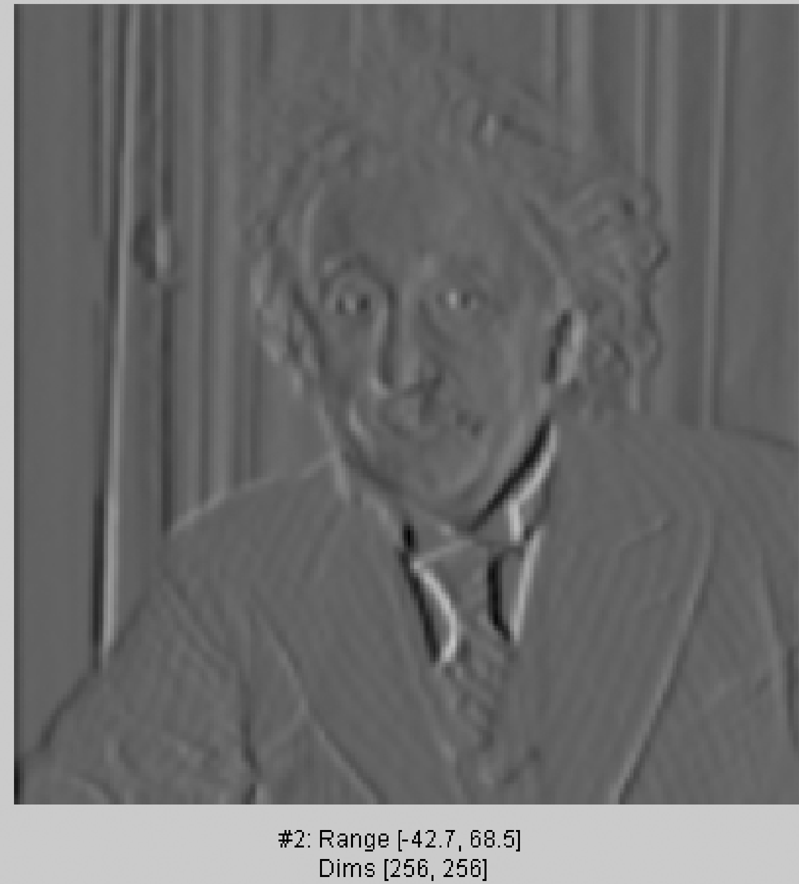
amplitude → `absval = abs(bp);`

local contrast → `avgAmplitude = upBlur(blurDn(absval, 2), 2);`

Contrast normalized output → `contrastNorm = bp ./ (avgAmplitude + const);`



Original image



Bandpass filtered
(deriv of gaussian)

Bandpass filtered

Absolute value

Blurred absolute value

#1: Range [-42.7, 68.5]
Dims [256, 256]

Bandpass filtered

#2: Range [-3.05, 3.42]
Dims [256, 256]

Bandpass filtered and contrast normalized

Bandpass filtered

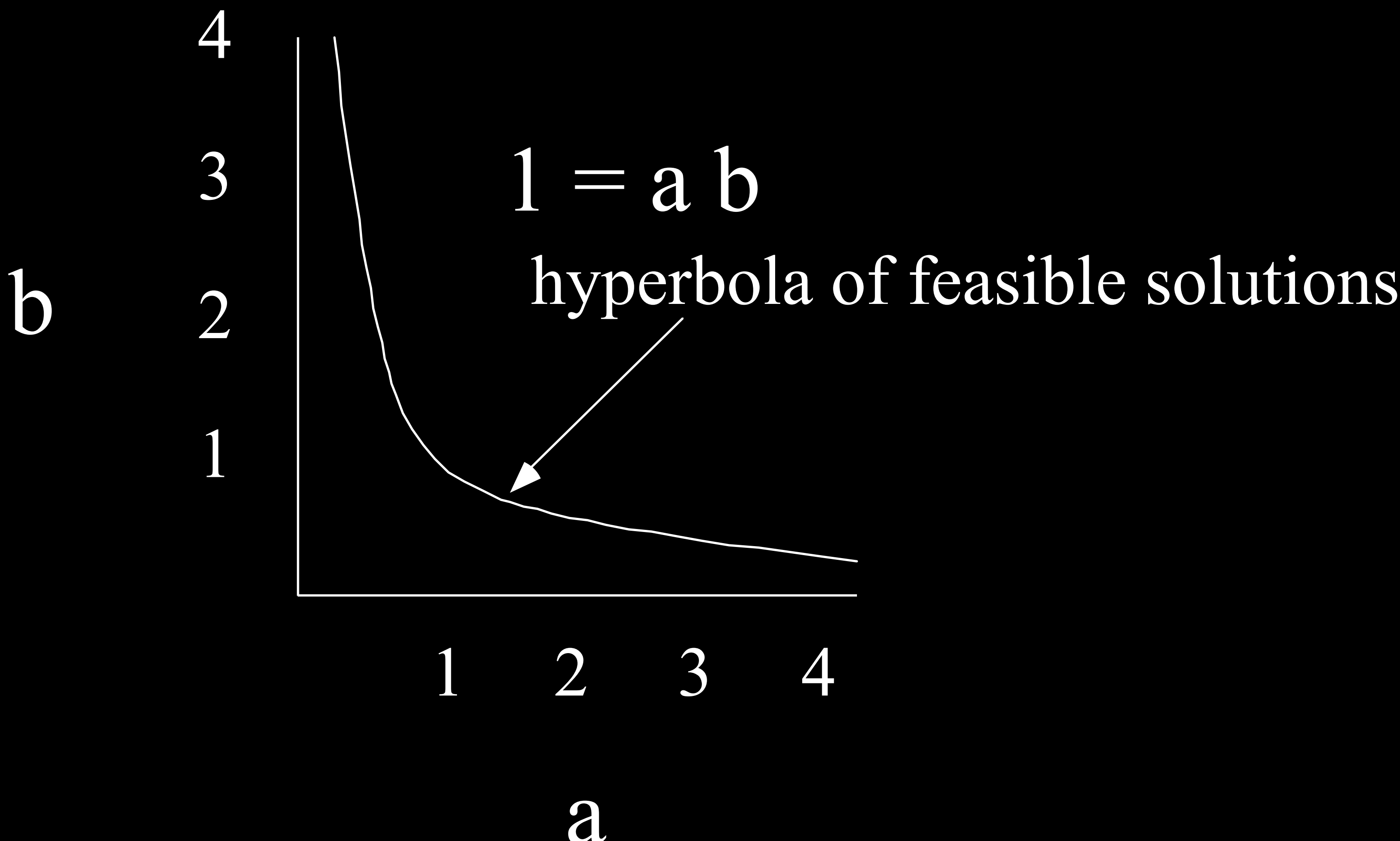
Bandpass filtered and contrast normalized

Bayesian methods

Simple, prototypical vision problem

- Observe some product of two numbers, say 1.0.
- What were those two numbers?
- Ie, $1 = ab$. Find a and b.

- Cf, simple prototypical graphics problem: here are two numbers; what's their product?



4
3
2
1
b
1 = a b
hyperbola of feasible solutions
1
2
3
4
a

Bayesian approach

- Want to calculate $P(a, b \mid y = 1)$.
- Use $P(a, b \mid y = 1) = k\, P(y=1 \mid a, b)\, P(a, b)$.

Likelihood function

Prior probability

Posterior probability

Likelihood function, P(obs|parms)

- The forward model, or rendering model, taking into account observation noise.
- Example: assume Gaussian observation noise. Then for this problem:

$$P(y=1 \mid a,b)=\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{(1-ab)^2}{2\sigma^2}}$$

Prior probability

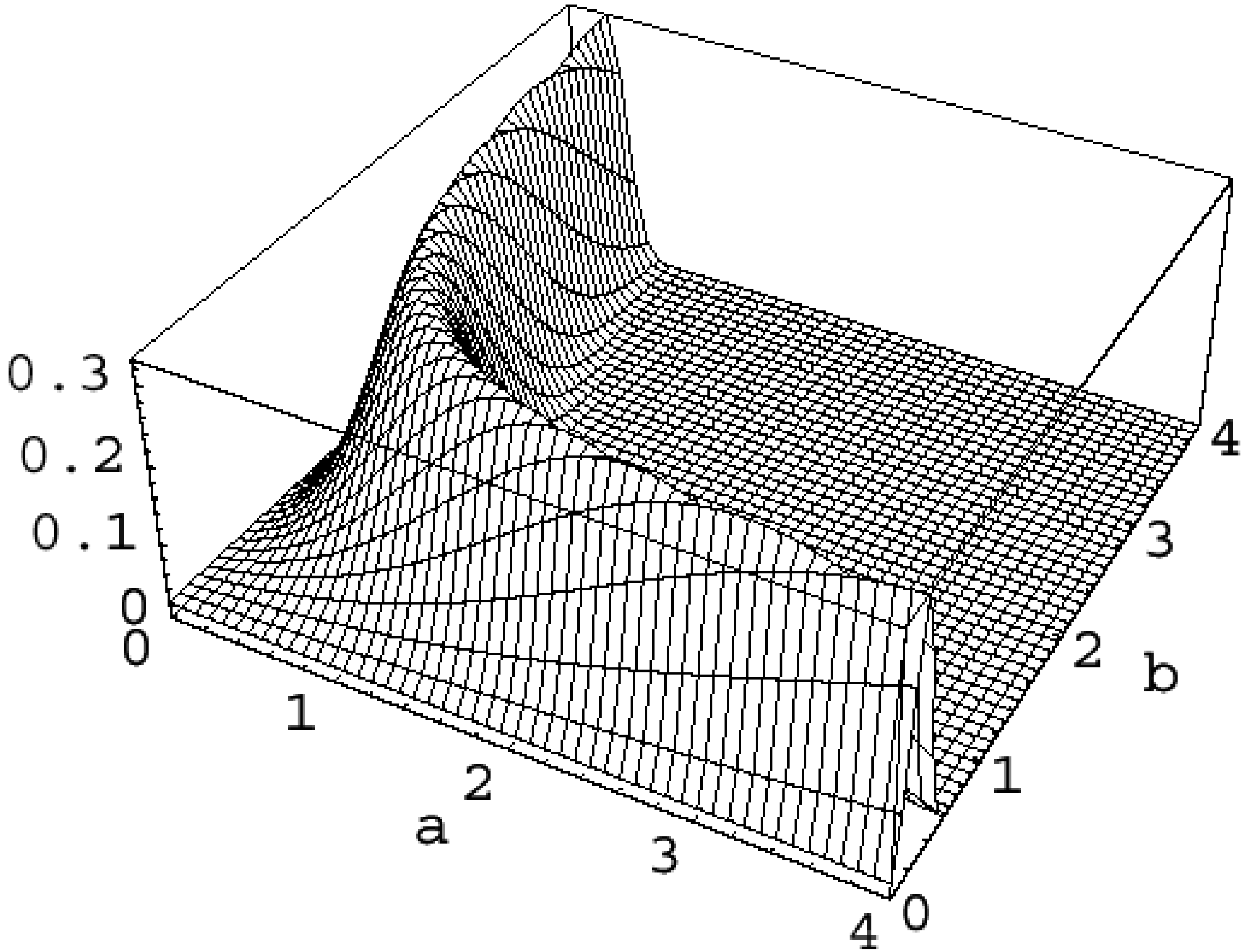
- In this case, we'll assume $P(a,b)=P(a)P(b)$, and $P(a) = P(b) = \text{const.}$, $0<a<4$.

Posterior probability

- Posterior = k likelihood prior

$$P(a,b \mid y=1) = ke^{-\frac{(1-ab)^2}{2\sigma^2}}$$

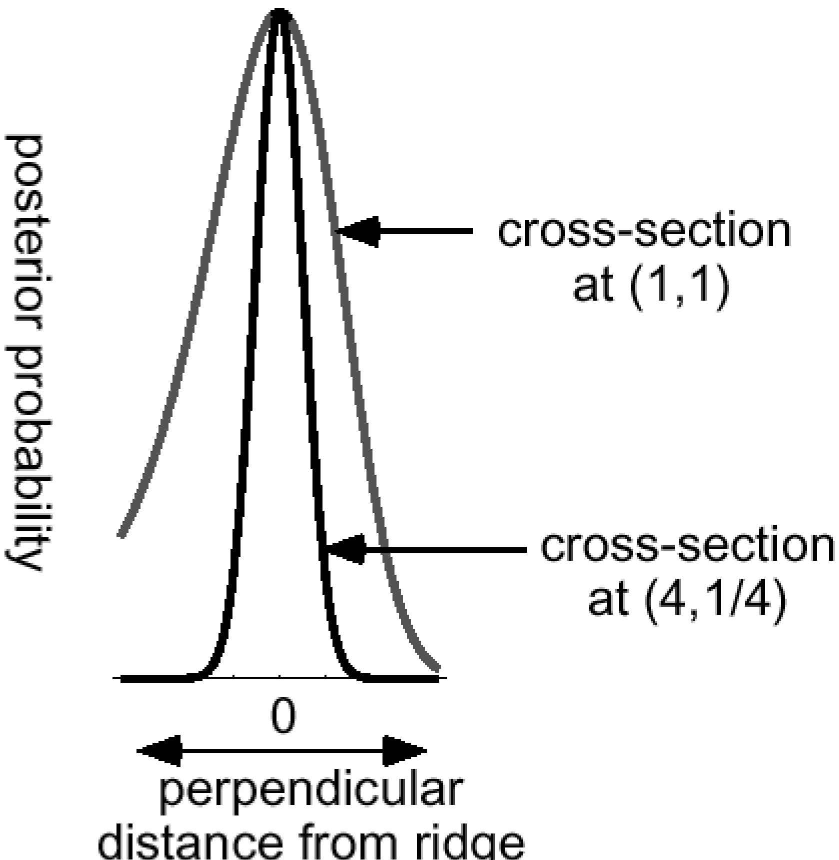
for 0 < a,b<4,
0 elsewhere



D. H. Brainard and W. T. Freeman, *Bayesian Color Constancy*, Journal of the Optical Society of America, A, 14(7), pp. 1393-1411, July, 1997

(a) Posterior Probability

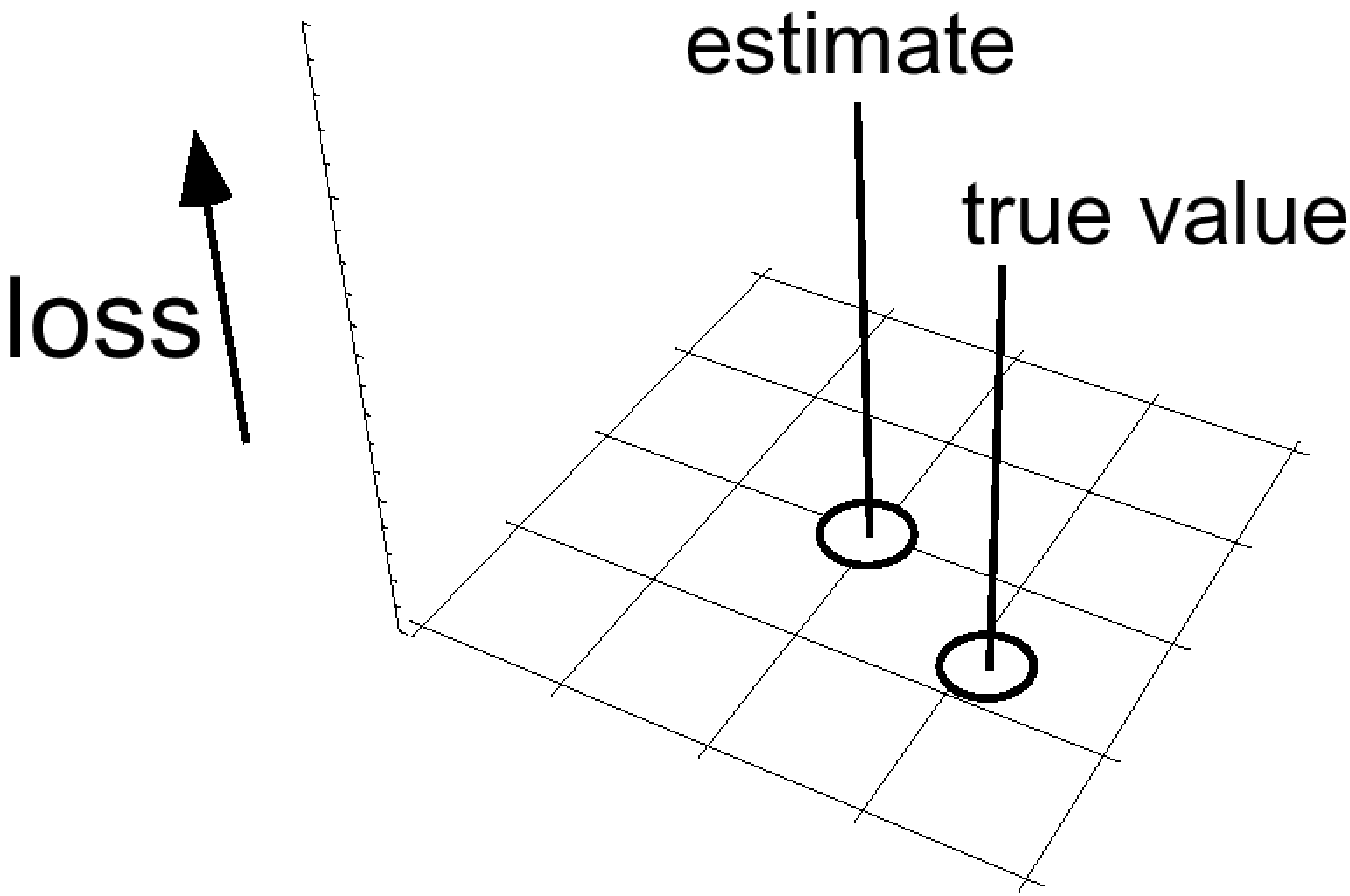
Loss functions



(b) Ridge Thickness Variations

Figure 1: Bayesian analysis of the problem $ab = 1$. Assuming uniform prior probabilities over the graphed region, (a) shows the posterior probability for gaussian observation noise of variance 0.18. The noise broadens the geometric solution into a hyperbola–shaped ridge of maximum probability. (b) Note the different thickness of the ridge; some parts have more local probability mass than others, even though the entire ridge has a constant maximum height.

D. H. Brainard and W. T. Freeman, *Bayesian Color Constancy*, Journal of the Optical Society of America, A, 14(7), pp. 1393-1411, July, 1997


estimate
true value
loss

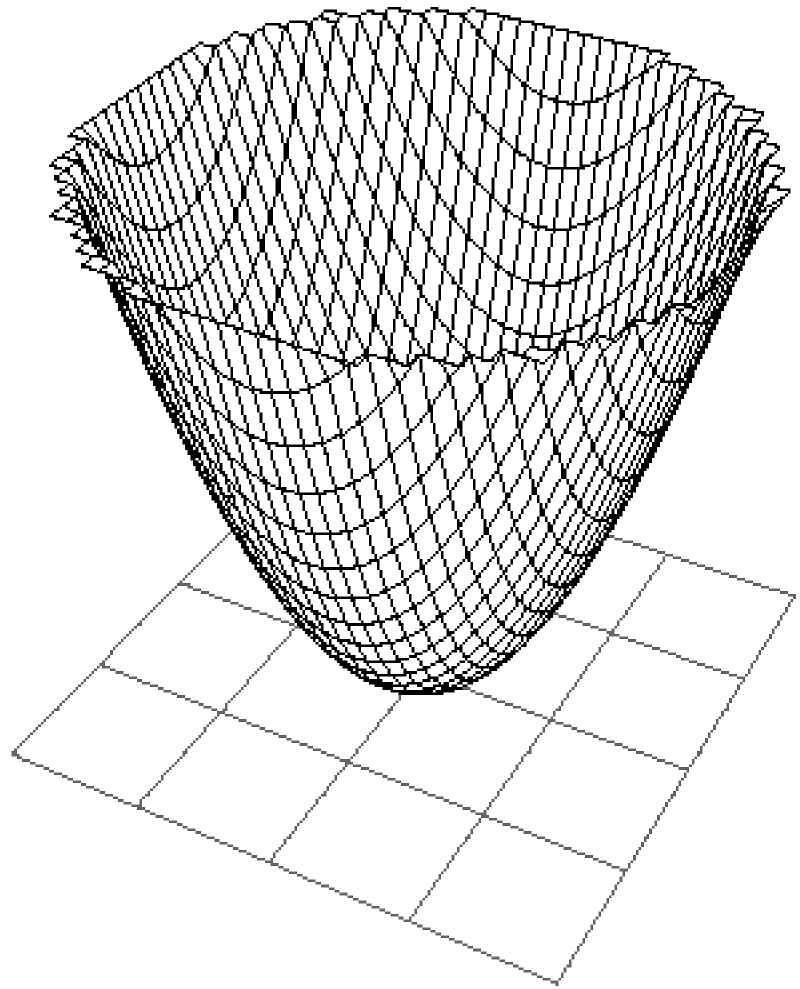
Bayesian decision theory

parameter variable, $\mathbf{z}$. A loss function $L(\mathbf{z}, \tilde{\mathbf{z}})$ specifies the penalty for estimating $\tilde{\mathbf{z}}$ when the true value is $\mathbf{z}$. Knowing the posterior probability, one can select the parameter values which minimize the expected loss for a particular loss function:

$$\begin{aligned} \text{[expected loss]} &= \int \text{[posterior] [loss function] d [parameters]} \\ R(\tilde{\mathbf{z}}|\mathbf{y}) &= -C \int [\exp[-\frac{\mathcal{T}}{2\sigma^2} \|\mathbf{y} - \mathbf{f}(\mathbf{z})\|^2] \, \mathbf{P}_{\mathbf{z}}(\mathbf{z})] \quad L(\mathbf{z}, \tilde{\mathbf{z}}) \quad d\mathbf{z}, \end{aligned} \tag{21}$$

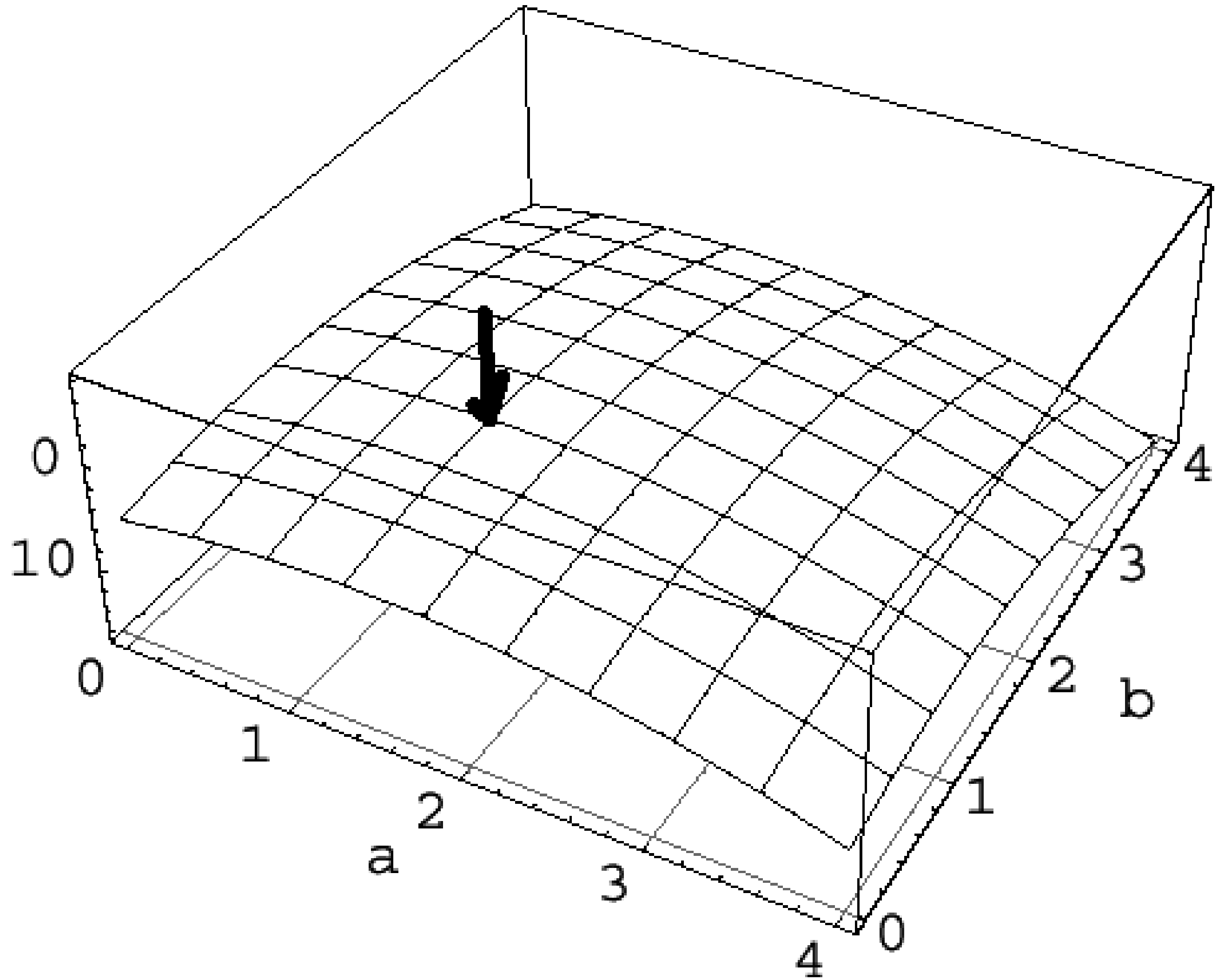
where we have substituted from Bayes' rule, Eq. (4), and the noise model, Eq. (3). The optimal estimate is the parameter $\tilde{\mathbf{z}}$ of minimum risk.

D. H. Brainard and W. T. Freeman, *Bayesian Color Constancy*, Journal of the Optical Society of America, A, 14(7), pp. 1393-1411, July, 1997



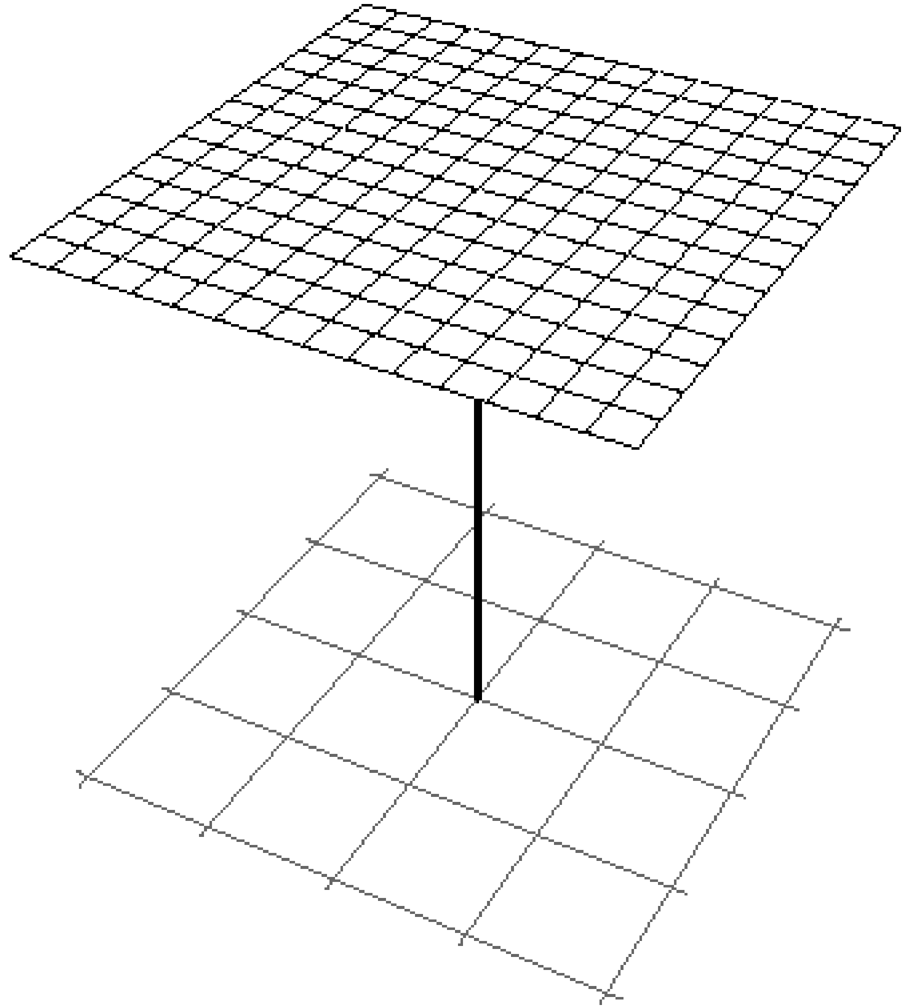
(b) MMSE loss fn.

D. H. Brainard and W. T. Freeman, *Bayesian Color Constancy*, Journal of the Optical Society of America, A, 14(7), pp. 1393-1411, July, 1997



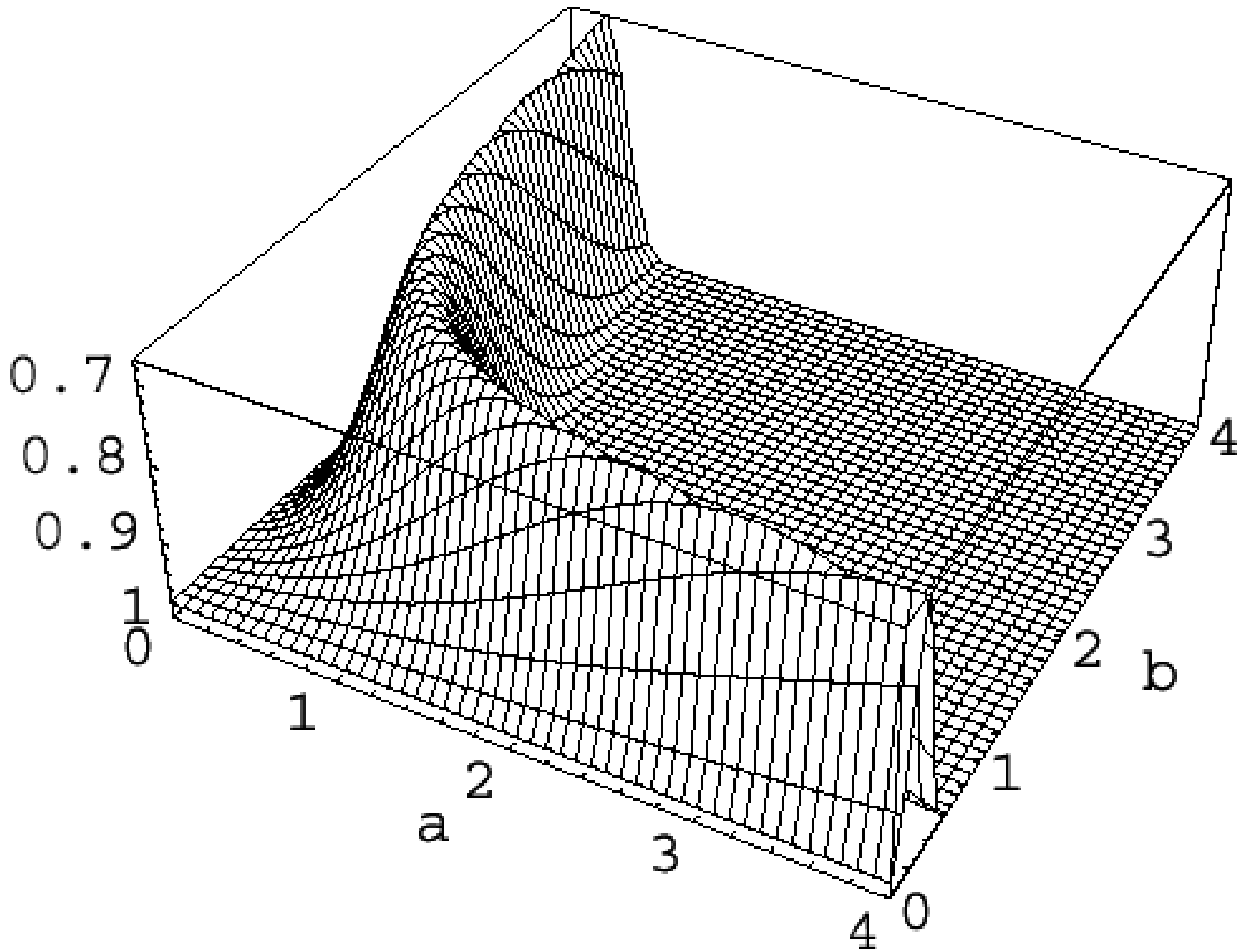
D. H. Brainard and W. T. Freeman, *Bayesian Color Constancy*, Journal of the Optical Society of America, A, 14(7), pp. 1393-1411, July, 1997

(e) (minus) MMSE risk



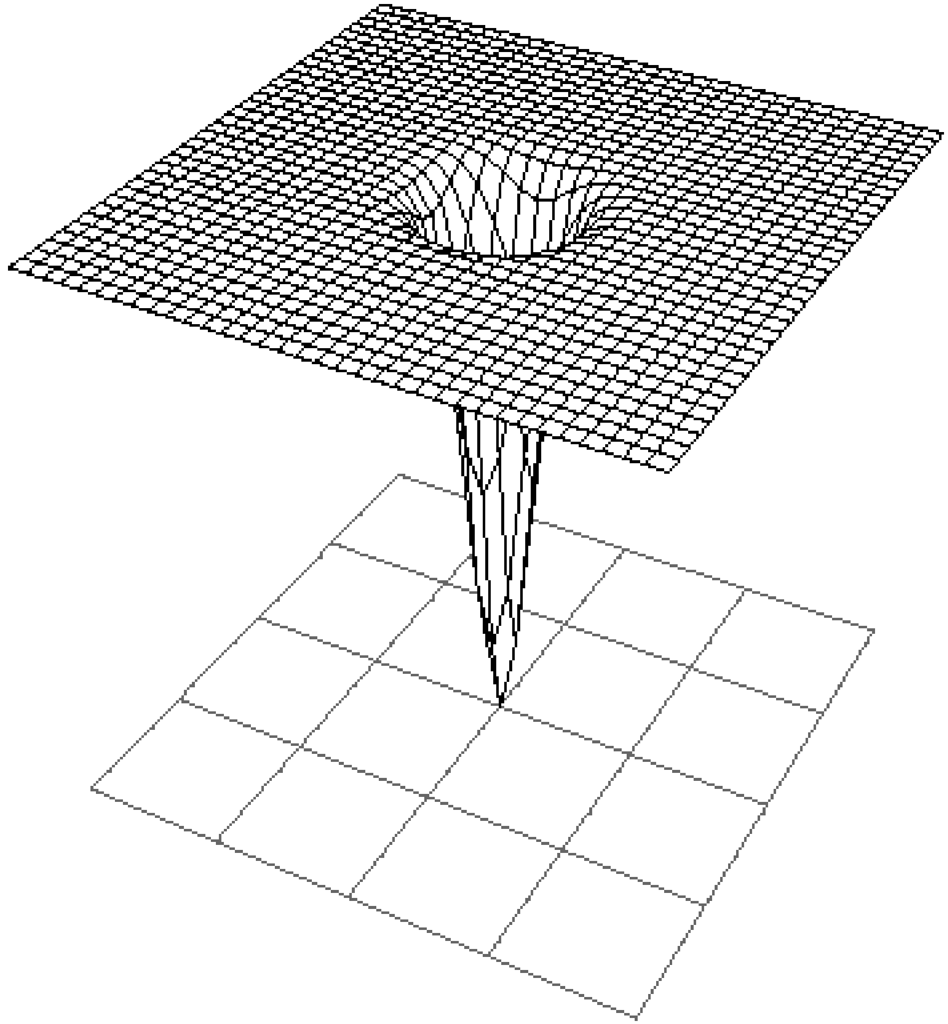
(a) MAP loss fn.

D. H. Brainard and W. T. Freeman, *Bayesian Color Constancy*, Journal of the Optical Society of America, A, 14(7), pp. 1393-1411, July, 1997



D. H. Brainard and W. T. Freeman, *Bayesian Color Constancy*, Journal of the Optical Society of America, A, 14(7), pp. 1393-1411, July, 1997

(d) (minus) MAP risk

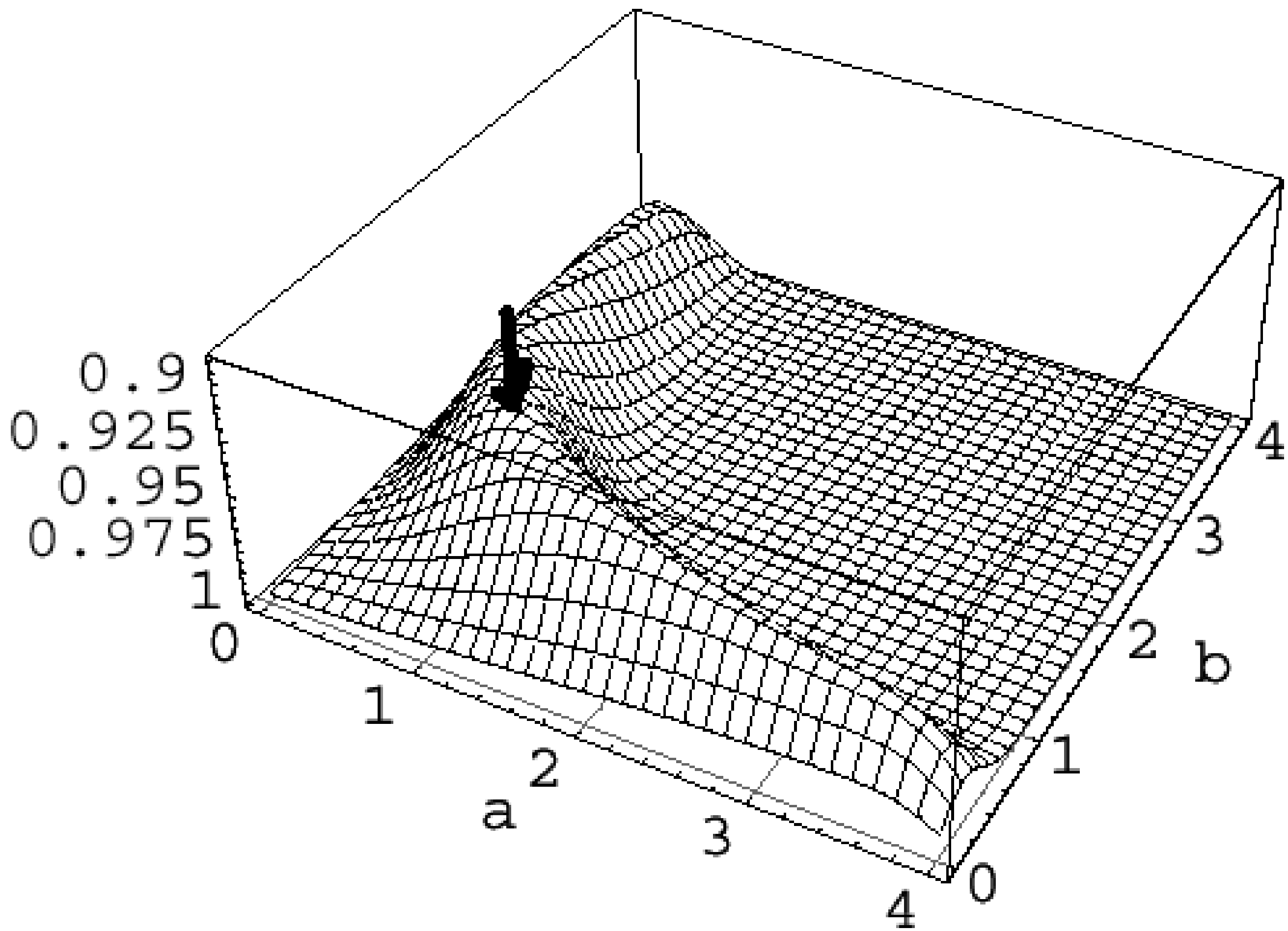


(c) MLM loss fn.

D. H. Brainard and W. T. Freeman, *Bayesian Color Constancy*, Journal of the Optical Society of America, A, 14(7), pp. 1393-1411, July, 1997

Local mass loss function may be useful model for perceptual tasks

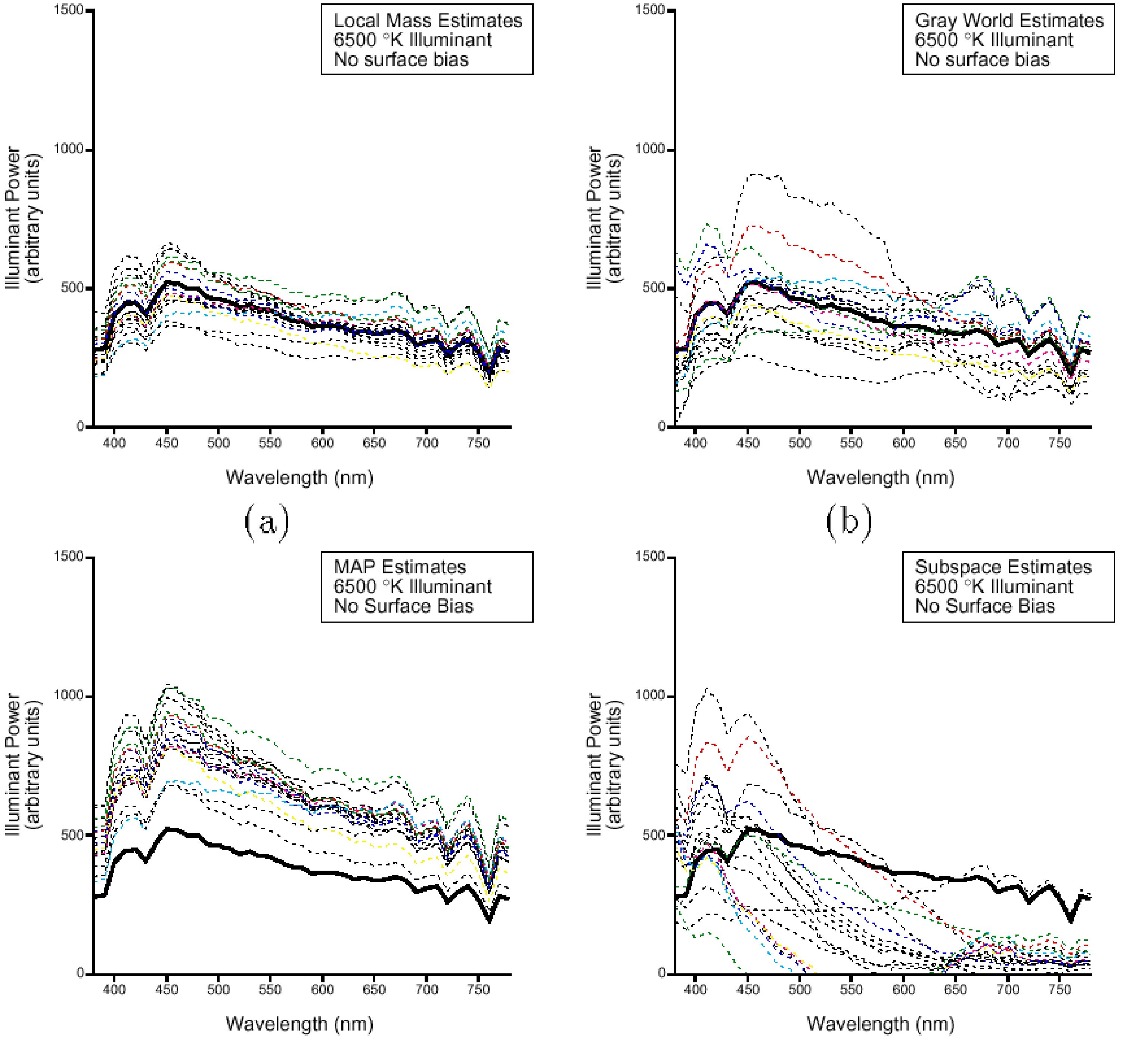
http://sportsillustrated.cnn.com/baseball/college/2000/college_world_series/news/2000/06/15/cws_notebook_ap/t1_borchard_ap_01.jpg



(f) (minus) MLM risk

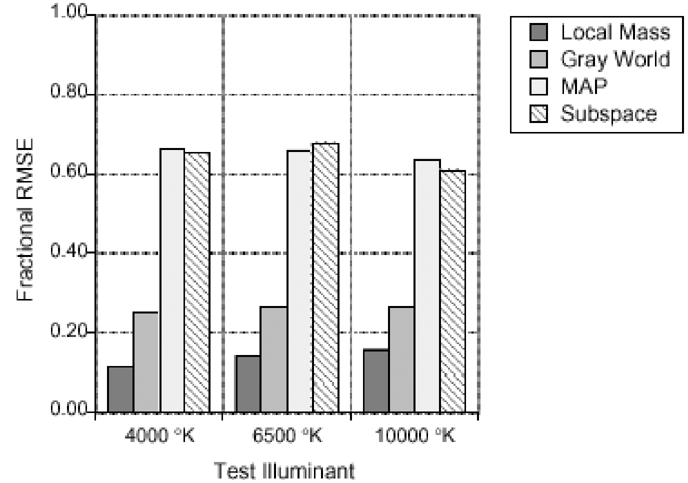
D. H. Brainard and W. T. Freeman, *Bayesian Color Constancy*, Journal of the Optical Society of America, A, 14(7), pp. 1393-1411, July, 1997

Figure 2: Left column: Three loss functions. Plots show penalty for guessing parameter values offset from the actual value, taken to be the plot center. (a) Minus delta function loss, assumed in MAP estimation. Only *precisely* the correct answer matters. (b) Squared error loss (a parabola), used in MMSE estimation. Very wrong guesses can carry inordinate influence. (c) Minus local mass loss function. Nearly correct answers are rewarded while all others carry nearly equal penalty. Right column: Corresponding expected loss, or Bayes risk, for the $y = ab$ problem. Note: loss *increases* vertically, to show extrema. (d) Expected loss for MAP estimator is minus the posterior probability. There is no unique point of minimum loss. (e) The minimum mean squared error estimate, (1.3, 1.3) (arrow) does not lie along the ridge of solutions to $ab = 1$. (f) The minus local mass loss favors the point (1.0, 1.0) (arrow), where the ridge of high probability is widest. There is the most probability mass in that local neighborhood.

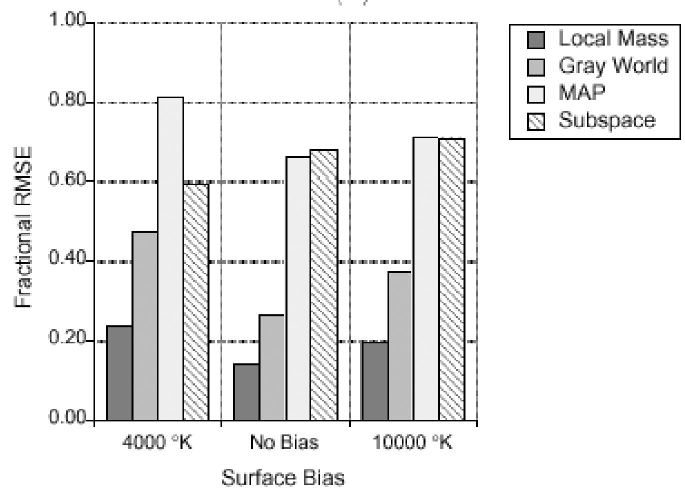


D. H. Brainard and W. T. Freeman, *Bayesian Color Constancy*, Journal of the Optical Society of America, A, 14(7), pp. 1393-1411, July, 1997

Figure 3: Visual comparison of illumination spectrum estimates for four color constancy algorithms: local mass, gray world, MAP and subspace. For a given illuminant, shown in dark line, a set of surfaces was drawn from the prior distribution 19 times. For each draw, each algorithm estimated the illuminant reflectance spectrum. The maximum local mass estimates, (a), are grouped closest to the actual illumination spectrum. The gray world algorithm estimates, (b), have wider variability. The MAP estimator, (c), ignores relevant information in the posterior distribution, which results in a systematic bias of its estimates. The subspace algorithm, (d), was not designed to work under the tested conditions, and performs poorly.



(a)



(b)

Figure 4: Summary results. (a) shows the performance of all four algorithms for three illuminants. (b) shows the performance of all four algorithms for three surface draw conditions. The performance measure is the average (over 19 individual runs) fractional root mean squared error (RMSE) between the estimate and true illuminant. For all conditions, the MLM estimate performs substantially better than the other algorithms. It is seen to be robust against these violations of its prior assumptions.

D. H. Brainard and W. T. Freeman, *Bayesian Color Constancy*, Journal of the Optical Society of America, A, 14(7), pp. 1393-1411, July, 1997

Regularization vs Bayesian interpretations

Regularization: minimize

$$(1-ab)^2+\lambda(a^2+b^2)$$

Bayes: maximize

$$e^{-\frac{(1-ab)^2}{2\sigma^2}}\, e^{-\lambda(a^2+b^2)}$$

likelihood

prior

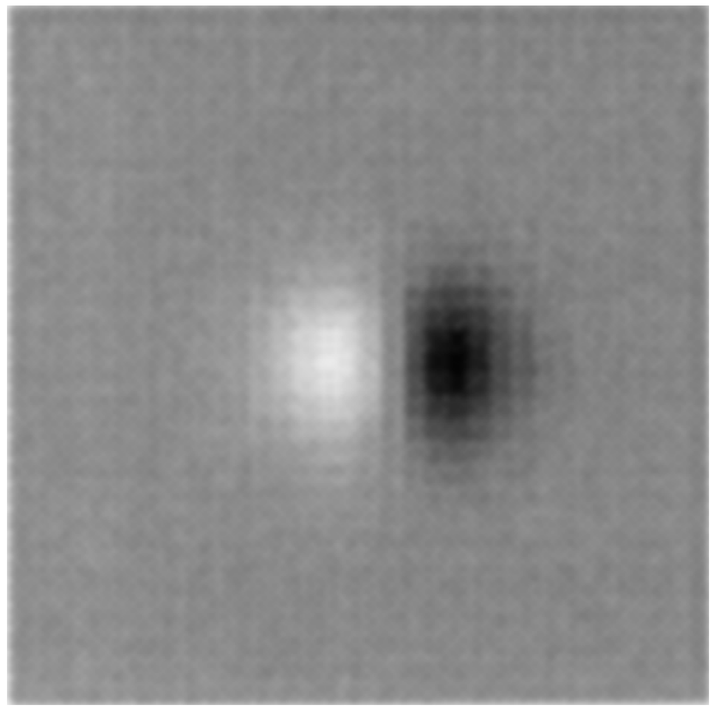
Bayesian interpretation of regularization approach

- For this example:
  - Assumes Gaussian random noise added before observation
  - Assumes a particular prior probability on a, b.
  - Uses MAP estimator (assumes delta fn loss).

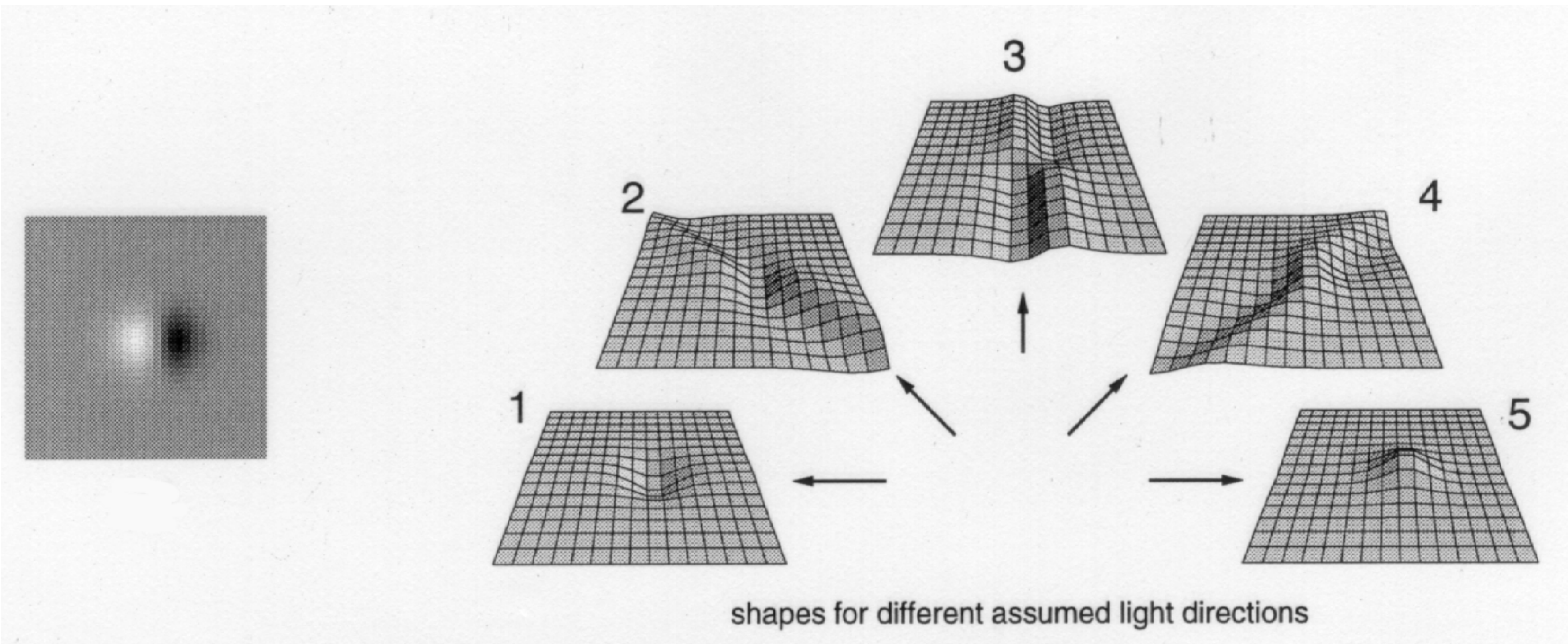
Why the difference matters

- Know what the things mean
- Speak with other modalities in language of probability
- Loss function
- Bayes also offers principled ways to choose between different models.

Example image

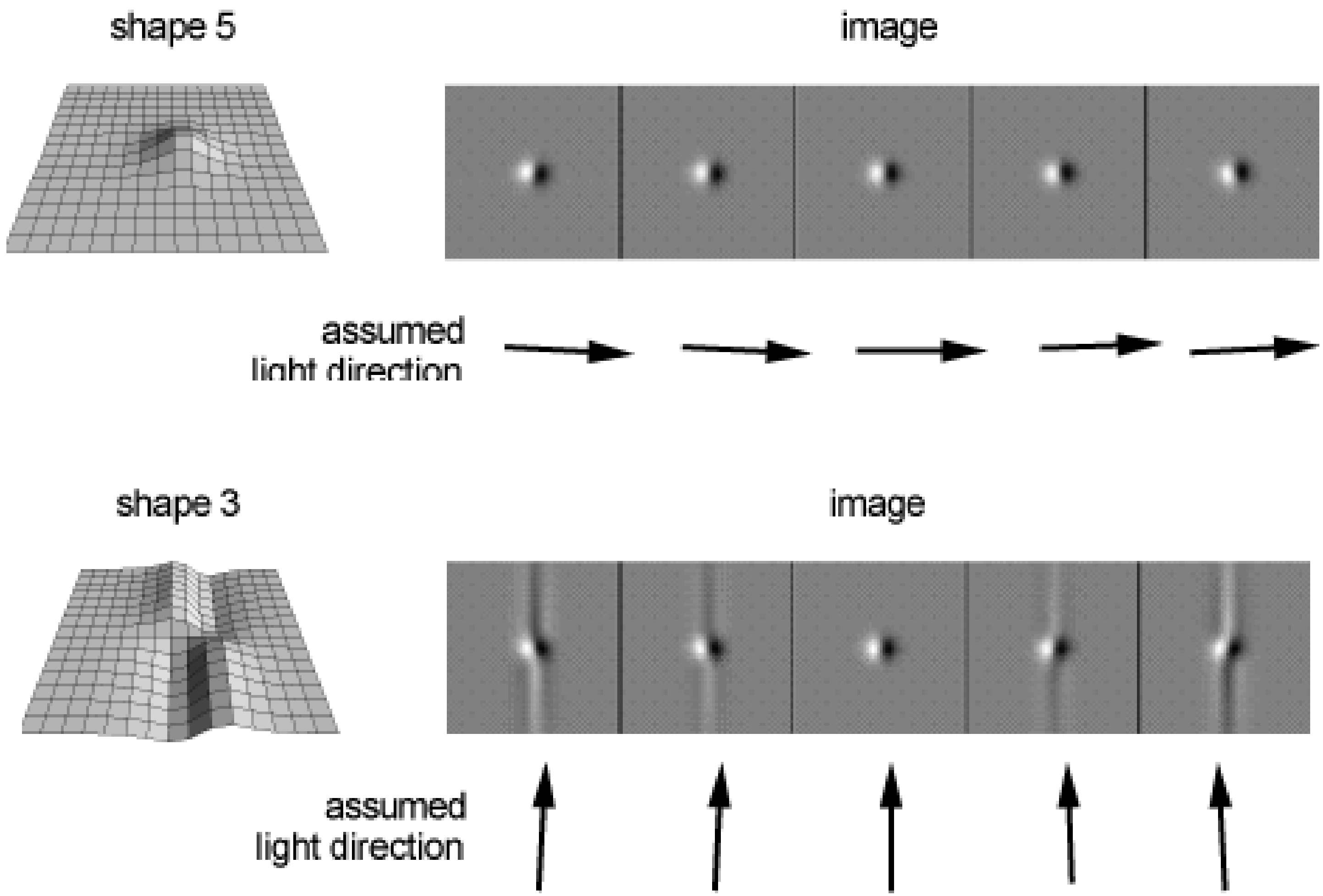


Multiple shape explanations



W. T. Freeman, *The generic viewpoint assumption in a framework for visual perception*, Nature, vol. 368, p. 542 - 545, April 7, 1994.

Generic shape interpretations render to the image over a range of light directions



W. T. Freeman, *The generic viewpoint assumption in a framework for visual perception*, Nature, vol. 368, p. 542 - 545, April 7, 1994.

Loss function

$$L(s,\theta \mid y) = \int P(s',\theta' \mid y) l(s,\theta,s',\theta') ds' s\theta'$$

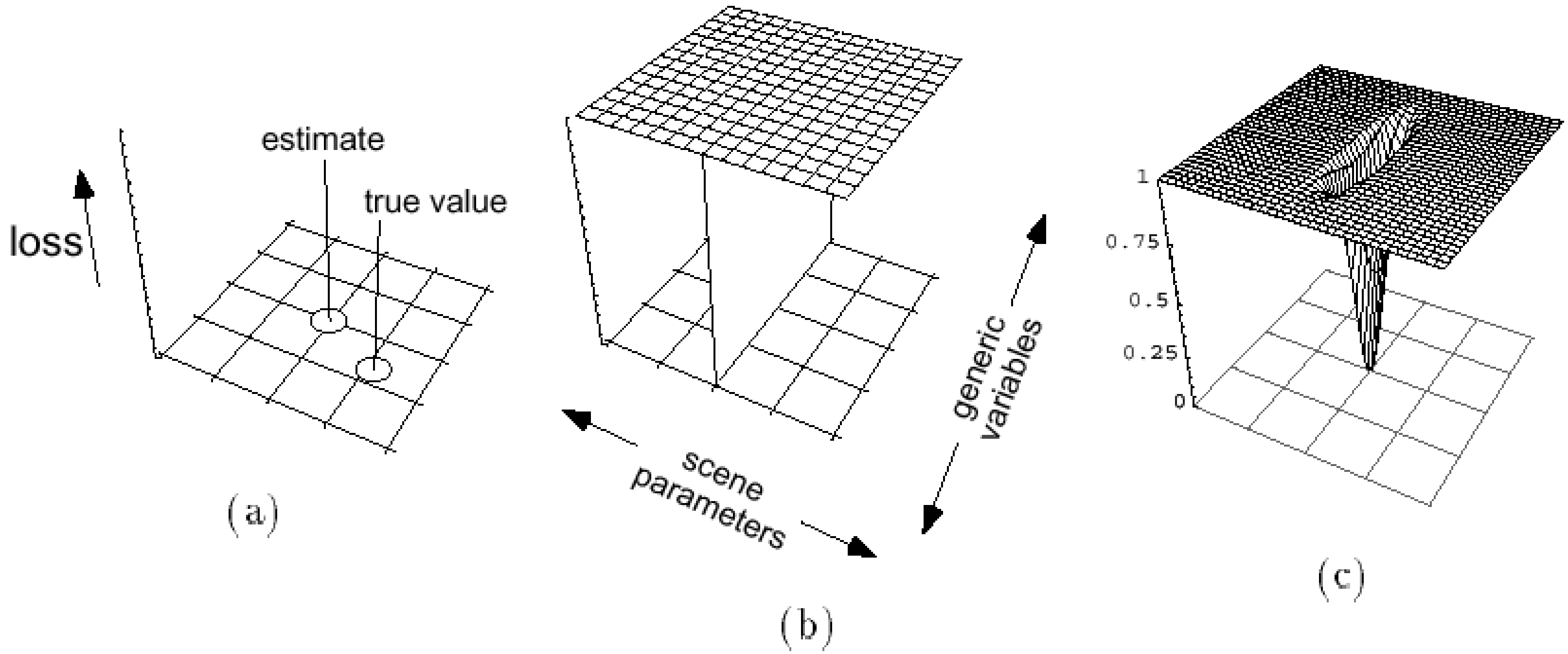
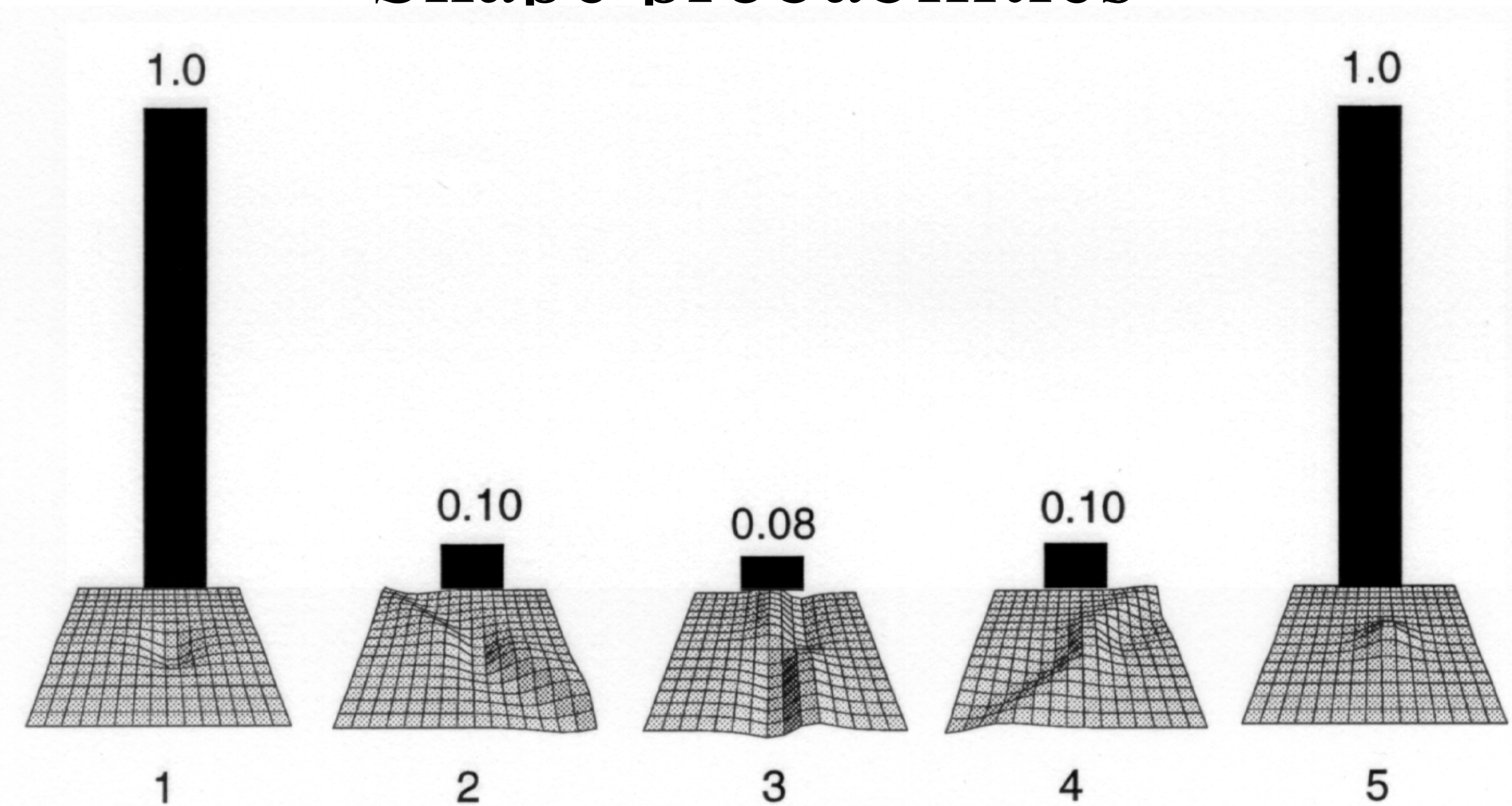


Figure 10: Loss function interpretation of generic viewpoint assumption. (a) shows the general form for a shift invariant loss function. The function $L(\mathbf{z}, \bar{\mathbf{z}})$ describes the penalty for guessing the parameter $\bar{\mathbf{z}}$ when the acutal value was $\mathbf{z}$. The marginalization over generic variables of Eq. (5) followed by MAP estimation is equivalent to using the loss function of (b). (c) Shows another possible form for the loss function, discussed in [11, 23, 24, 65].

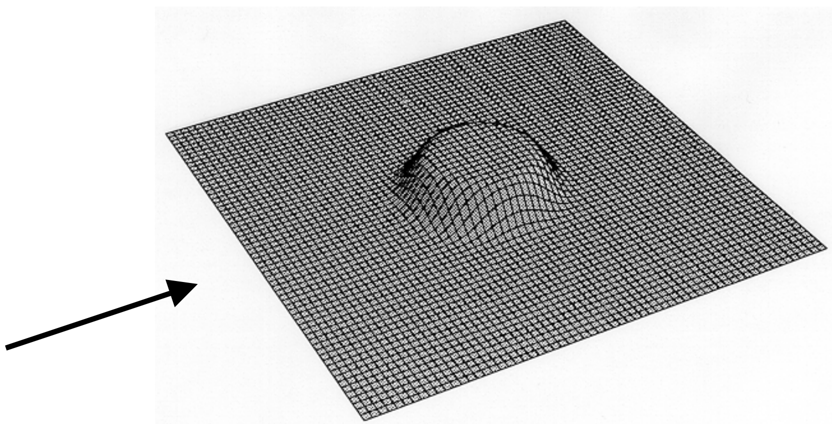
W. T. Freeman, *The generic viewpoint assumption in a framework for visual perception*, Nature, vol. 368, p. 542 - 545, April 7, 1994.

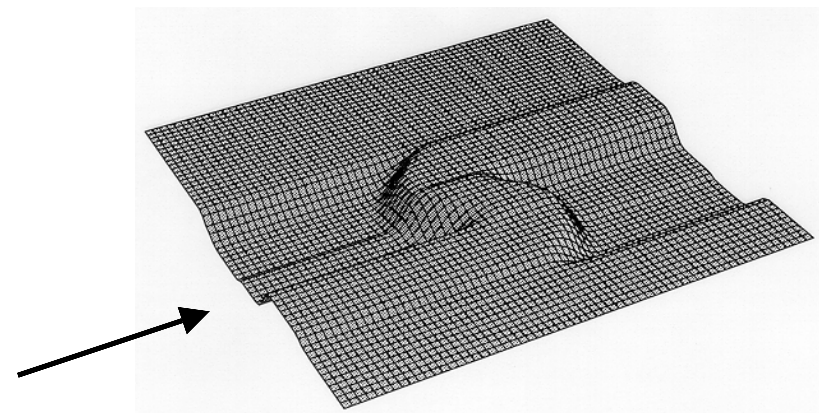
Shape probabilities



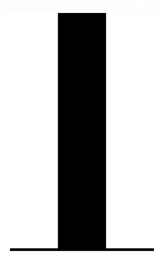
W. T. Freeman, *The generic viewpoint assumption in a framework for visual perception*, Nature, vol. 368, p. 542 - 545, April 7, 1994.

Comparison of shape explanations





- Lighting "genericity" of the shape explanation:



3.8

0.48