Mumford-Shah Energy Functional

Zoltan Kato

http://www.cab.u-szeged.hu/~kato/variational/

Introduction

- Proposed in their influential paper by
  - David Mumford
    - http://www.dam.brown.edu/people/mumford/
  - Jayant Shah
    - http://www.math.neu.edu/~shah/

Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems. *Communications on Pure and Applied Mathematics, Vol. XLII, pp 577-685, 1989*

Images as functions

- A gray-level image represents the light intensity recorded in a plan domain ***R***
  - We may introduce coordinates ***x***, ***y***
  - Let ***g(x,y)*** denote the intensity recorded at the point ***(x,y)*** of ***R***
  - The function ***g(x,y)*** defined on the domain ***R*** is called an image.

What kind of function is g?

- The light reflected by the surfaces S_i of various objects O_i will reach the domain R in various open subsets R_i
- When O_1 appears as the background to the sides of O_2 then the open sets R_1 and R_2 will have a common boundary (**edge**)
- One usually expects $g(x,y)$ to be discontinuous along this boundary

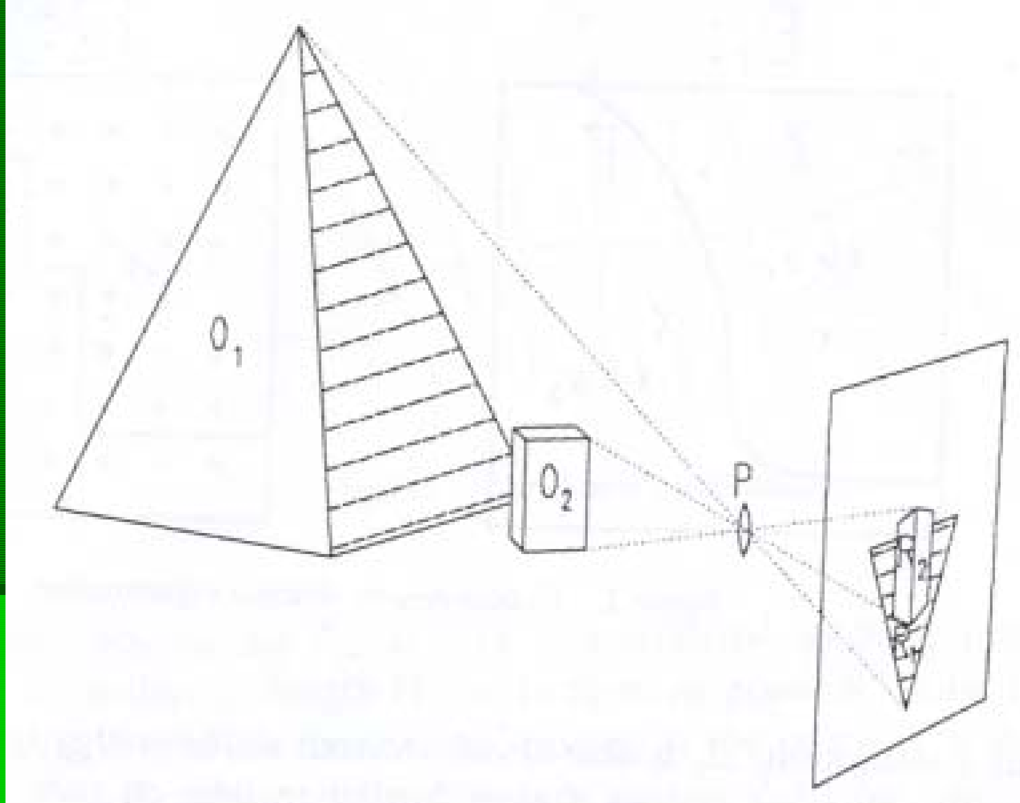
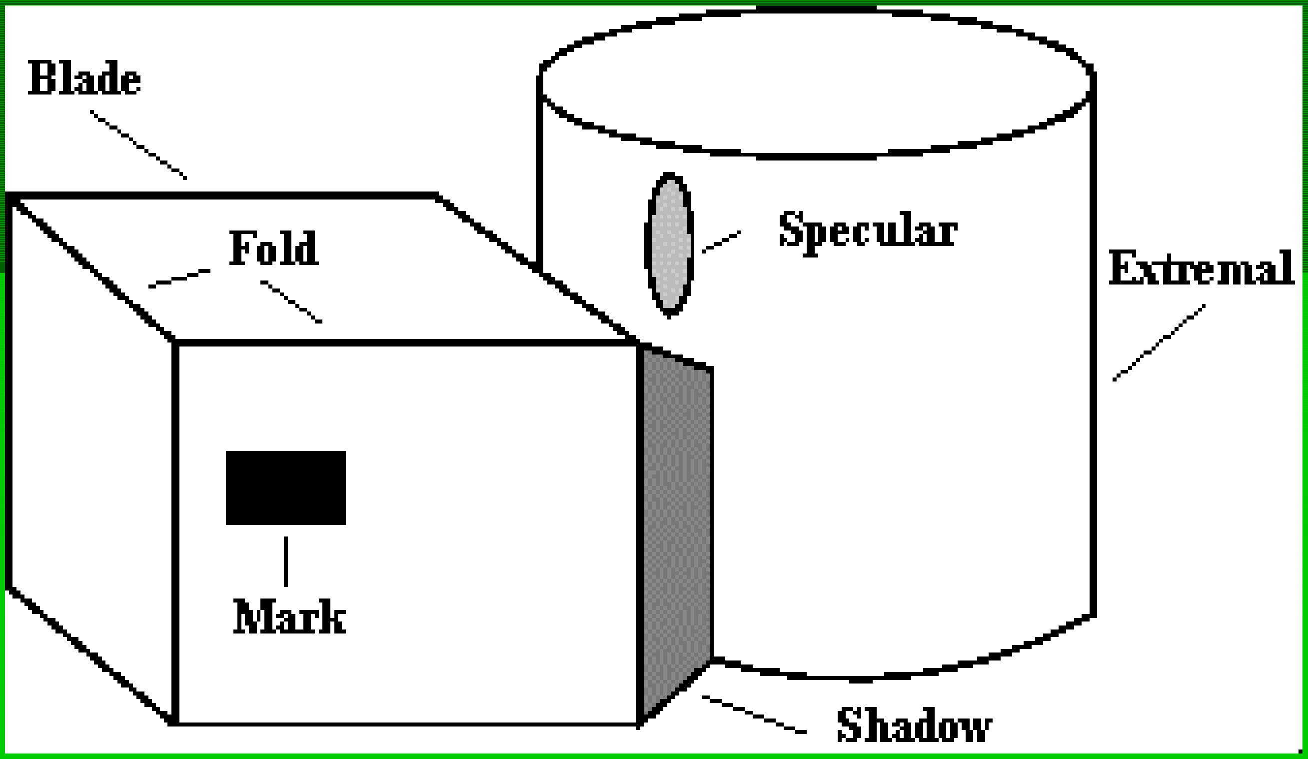


Figure 1. An *image* of a 3D scene.

Figure from D. Mumford & J. Shah: **Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems.** *Communications on Pure and Applied Mathematics, Vol. XLII, pp 577-685, 1989*

Other discontinuities

- Surface orientation of visible objects (cube)
- Surface markings
- Illumination (shadows, uneven light)



Piece-wise smooth g

- In all cases, we expect ***g(x,y)*** to be piece-wise smooth to the first approximation.
- It is well modelled by a set of smooth functions f_i defined on a set of disjoint regions R_i covering R.
- Problems:
  - Textured objects (regions perceived homogeneous but lots of discontinuities in intensity)
  - Sahdows are not true discontinuities
  - Partially transparent objects
  - Noise
- Still widely and succesfully applied model!

Segmentation problem

- Consists in computing a decomposition of the domain of the image ***g(x,y)***

$$R = \bigcup_{i=1}^{n} R_i$$

1. ***g*** varies smootly and/or slowly within R_i
2. ***g*** varies discontinuously and/or rapidly across most of the boundary Γ between regions R_i

Optimal approximation

- Segmentation problem may be restated as
  - finding optimal approximations of a general function ***g***
  - by piece-wise smooth functions ***f***, whose restrictions f_i to the regions R_i are **differentiable**
- **Many other applications:**
  - **Speech recognition**
  - **Sonar, radar or laser range data**
  - **CAT scans**
  - **etc…**

Optimal segmentation

- Mumford and Shah studied 3 functionals which measure the degree of match between an image ***g(x,y)*** and a segmentation.
- First, they defined a general functional ***E*** (the famous Mumford-Shah functional):
  - R_i will be disjoint connected open subsets of the planar domain ***R***, each one with a piece-wise smooth boundary
  - Γ will be the union of the boundaries.

$$R = \coprod_{i=1}^{n} R_i \coprod \Gamma$$

Mumford-Shah functional

- Let ***f*** differentiable on $\cup R_i$ and allowed to be discontinuous across Γ.

$$E(f,\Gamma) = \mu^2 \iint_R (f-g)^2 dxdy + \iint_{R-\Gamma} \|\nabla f\|^2 dxdy + \nu|\Gamma|$$

- The smaller ***E***, the better ***(f, Γ)*** segments ***g***

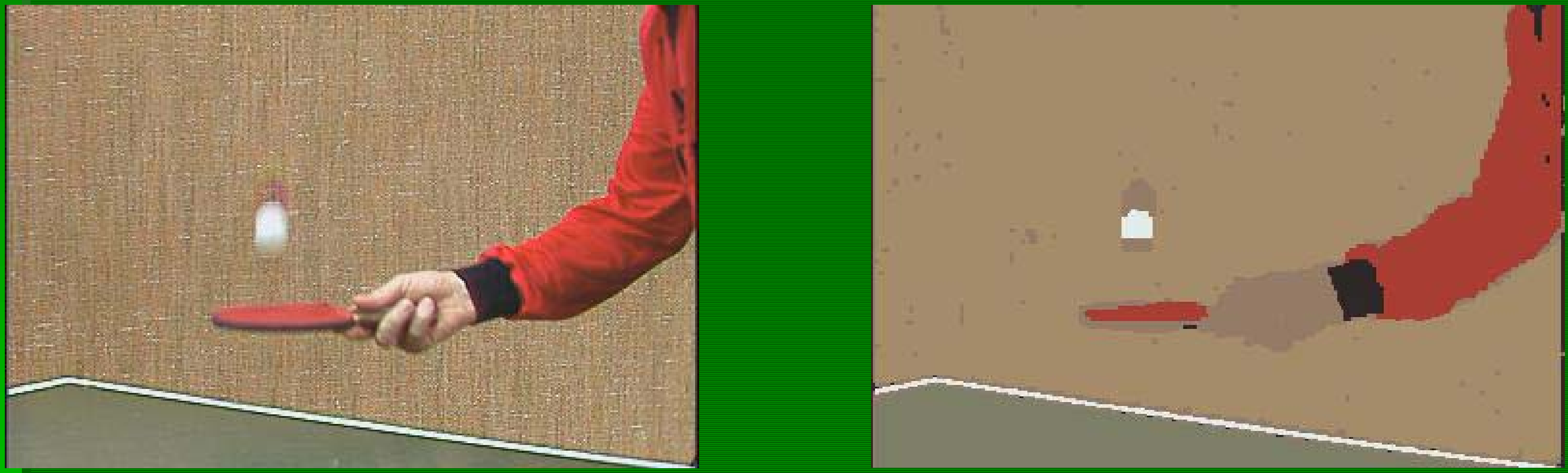
1. ***f*** approximates ***g***
2. ***f*** (hence ***g***) does not vary much on R_is
3. The boundary Γ be as short as possible.

- Dropping any term would cause ***inf E=0***.

Cartoon image

- ***(f, Γ)*** is simply a cartoon of the original image ***g***.
  - Basically ***f*** is a new image with edges drawn sharply.
  - The objects are drawn smootly without texture
  - ***(f, Γ)*** is essentially an idealization of ***g*** by the sort of image created by an artist.
  - Such cartoons are perceived correctly as representing the same scane as ***g*** → ***f*** is a simplification of the scene containing most of its essential features.

Cartoon image example



Related problems

- **D. Geman & S. Geman**: Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans. on PAMI 6, pp 721-741, 1984.*
  - MRF model
- **A. Blake & A. Zisserman**: Visual Reconstruction. *MIT Press, 1987*
  - Weak membrane model
- **M. Kass, A. Witkin & D. Terzopoulos**: Snakes: Active contour Models. *International Journal of Computer Vision, vol. 1, pp 321-332, 1988*.
  - Active contour model

Piecewise constant approximation

- A special case of ***E*** where ***f=a$_i$*** is constant on each open set ***R$_i$***.

$$\mu^{-2}E(f,\Gamma)=\sum_i \iint_{R_i}(g-a_i)^2\,dxdy+\frac{\nu}{\mu^2}|\Gamma|$$

- Obviously, it is minimized in ***a$_i$*** by setting ***a$_i$*** to the mean of ***g*** in ***R$_i$***:

$$a_i = mean_{R_i}(g) = \iint_{R_i} g\,dxdy \Big/ area(R_i)$$

Piecewise constant approximation

$$E_0(\Gamma) = \sum_i \iint_{R_i} (g - mean_{R_i}(g))^2 \, dxdy + \frac{\nu}{\mu^2} |\Gamma|$$

- It can be proven that minimizing ***E_0*** is well posed:
  - For any continuous ***g***, there exists a Γ made up of finit number of singular points joined by a finit number of arcs on which ***E_0*** atteins a minimum.
- It can also be shown that ***E_0*** is the natural limit functional of ***E*** as ***$\mu \to 0$***

Relation to the Ising model

- If we further restrict ***f*** to
  - take only values of +/-1,
  - assume that ***g*** and ***f*** are defined on a lattice
  - then E_0 becomes the energy of the **Ising model**.

Relation to the Ising model

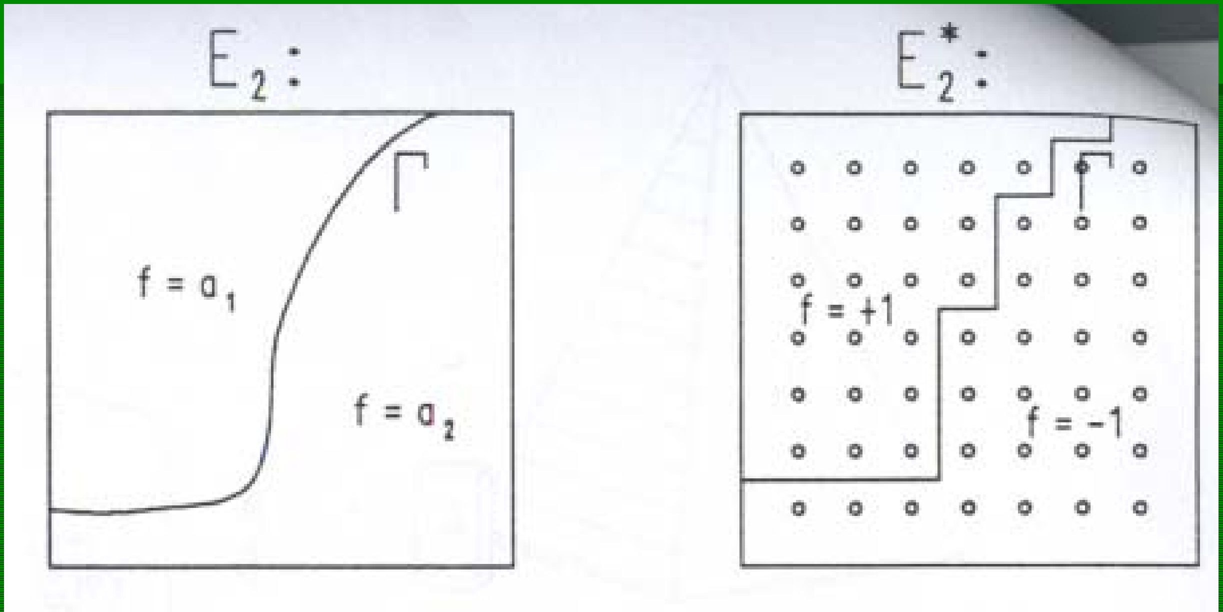


Figure 2. Continuous vs. discrete segmentation.

Figure from D. Mumford & J. Shah: Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems. *Communications on Pure and Applied Mathematics, Vol. XLII, pp 577-685, 1989*

- Γ is the path between all pairs of lattice points on which ***f*** changes sign:

$$E_0^*(f) = \sum_{i,j} (f(i,j) - g(i,j))^2 + \frac{\nu}{\mu^2} \sum_{(i,j),(k,l)} (f(i,j) - f(k,l))^2$$

Weak string

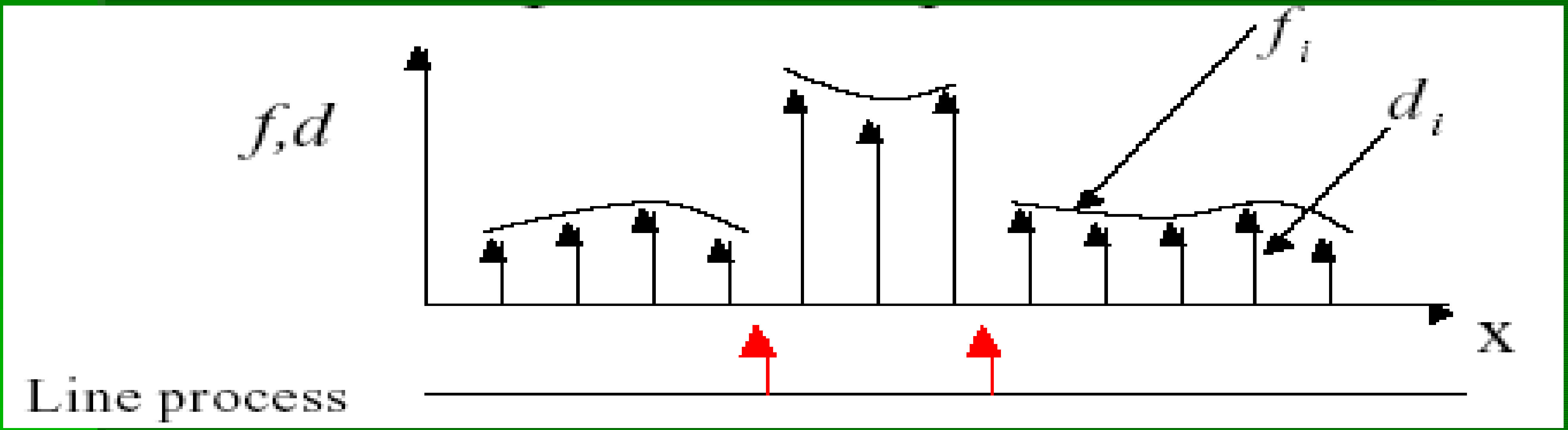


Image from *CMU 15-385 Computer Vision course, Spring 2002 by Tai Sing Lee*

- Fitting an elastic spline with possible breaks (line process or local edges)
  - Remove noise
  - Approximate with smooth curves
  - Breaks where smoothness is not satisfied

Energy of a weak string

$$E(f) = \sum_i (f_i - d_i)^2 + \lambda \sum_i (f_{i+1} - f_i)^2 (1 - l_i) + \alpha \sum_i l_i$$

- ***α*** is the cost of inserting a break (local edge element) ***l_i***
- ***l_i*** may take binary values [0,1]
- ***l_i*** is turned on when ***$(f_{i+1}-f_i)^2>\alpha/\lambda$***

Weak membrane model

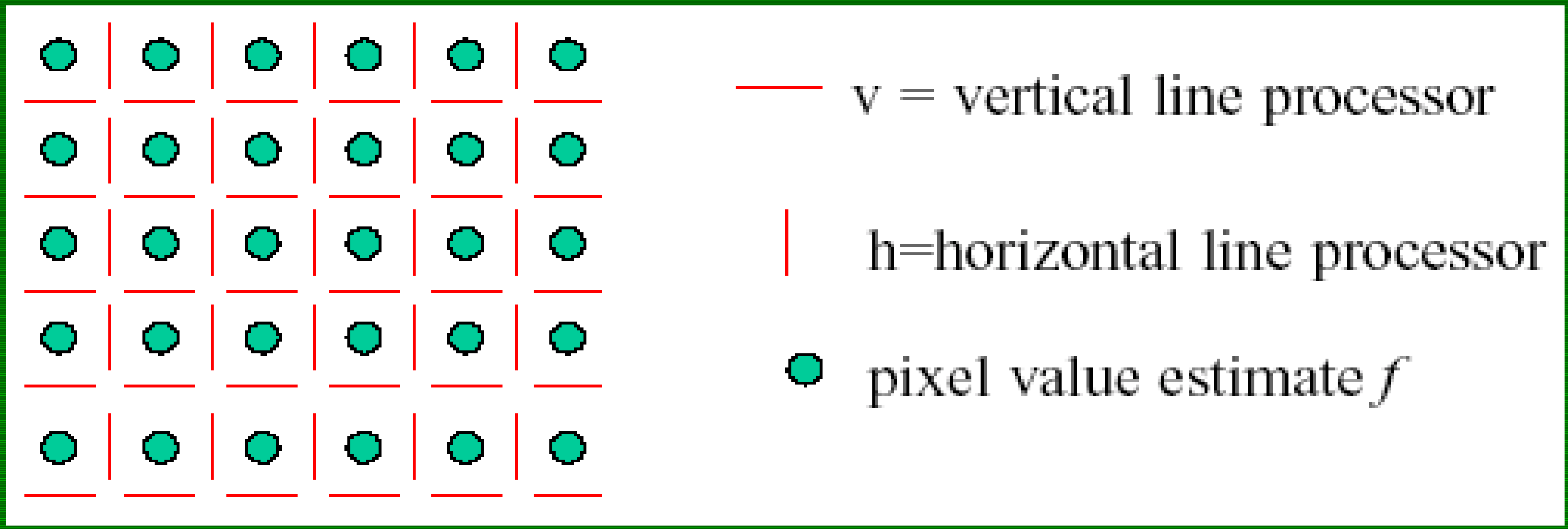


Image from CMU 15-385 Computer Vision course, Spring 2002 by Tai Sing Lee

$$E(f_{i,j},h_{i,j},v_{i,j}) = \sum_{i,j}(f_{i,j}-d_{i,j})^2 + \sum_{i,j}(f_{i+1,j}-f_{i,j})^2(1-h_{i,j}) + \lambda\sum_{i,j}(f_{i,j+1}-f_{i,j})^2(1-v_{i,j}) + \alpha\sum_{i,j}(v_{i,j}+h_{i,j}) + \kappa\sum_{i,j}V_c(i,j)$$

Contour continuity constraint

- ***$V_c(i,j)$*** energy term:
  - Low for 0, 2 lines
  - Medium for 3 lines
  - High for 1, 4 lines

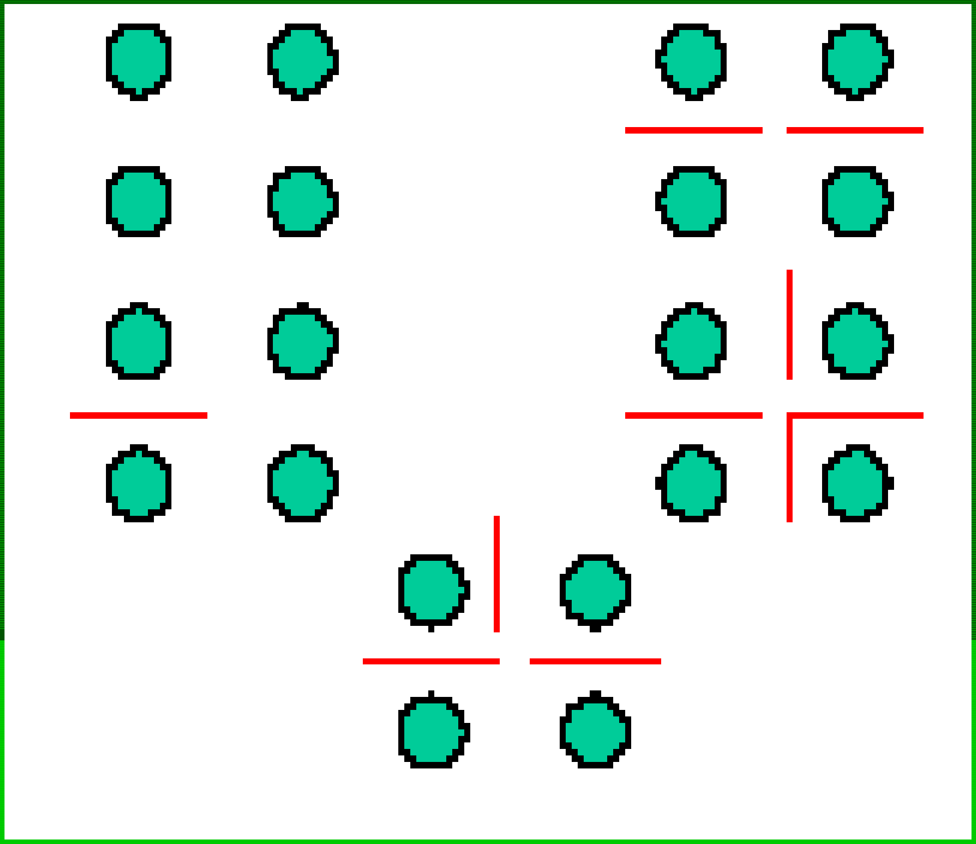


Image from *CMU 15-385 Computer Vision course, Spring 2002 by Tai Sing Lee*

First variation and the Euler equation

- The extrema of a function ***f(x)*** are attained where ***f' = 0***
- Similarly, the extrema of the functional ***E(u)*** are obtained where ***E' = 0***.
  - ***E = (∂E / ∂u)*** is the ***first variation***.
  - Assuming a common formulation where ***u(x):[0,1]→R***, ***u(0)=a*** and ***u(1)=b***, the basic problem is to minimize:

$$E(u) = \int_0^1 F(u, u')\,dx$$

  - The necessary condition for ***u*** to be an extremum of ***E(u)*** is the ***Euler equation*** of a one dimensional problem.

$$\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right) = 0$$

Energy minimization: Gradient descent

- The Euler equation can be solved by numerically solving (1).
- Can be formulated as an evolution equation (***t*** – time):
  - For example, updating ***f*** while ***l*** is fixed (string):

$$\frac{df_i}{dt} = -\frac{\partial E}{\partial f_i} \qquad (1)$$

$$\frac{\partial E}{\partial f_i} = 2(f_i - d_i) - 2\lambda(f_{i+1} - f_i)(1 - l_i) \qquad (2)$$

$$\Delta f_i = f_i^{t+1} - f_i^t = -\mu \frac{\partial E}{\partial f_i} \qquad (3)$$

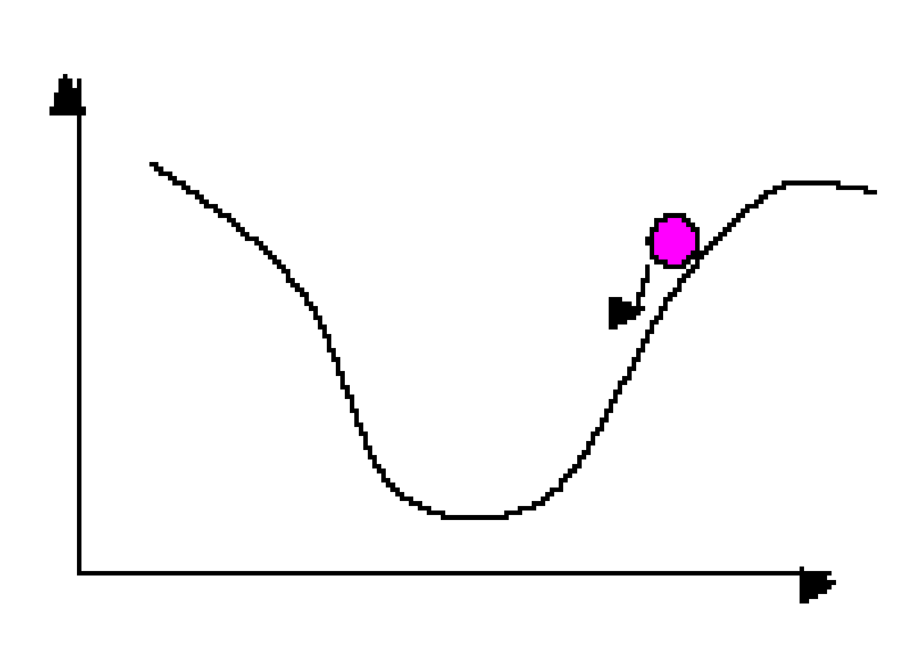


Image from *CMU 15-385 Computer Vision course, Spring 2002 by Tai Sing Lee*

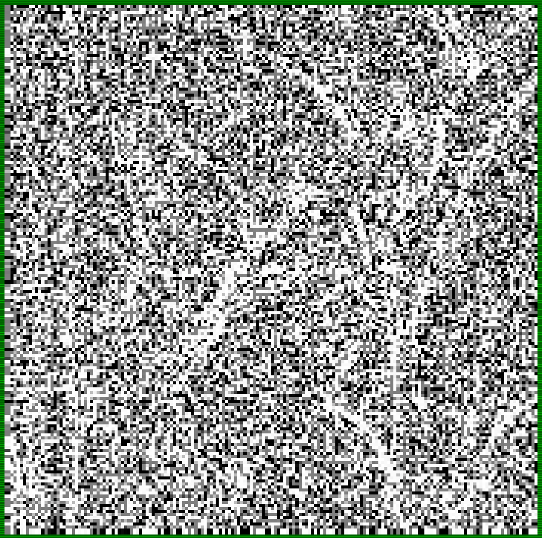
Works only for convex functions!

Energy minimization: Simulated Annealing example

- Works for non-convex energy functions
  - ***λ=6***, ***α=0.04***

Images from *CMU 15-385 Computer Vision course, Spring 2002 by Tai Sing Lee*

Line process (***l***)



Surface signal (***f***)

Energy minimization: Graduated non-convexity

- Proposed by Blake & Zisserman for the weak membrane's energy
- Basic idea:

1. Approximate the original energy functional by a convex one
2. Do a gradient descent on the approximation
3. Gradually morph back the approximation into the original energy while repeating step 2.

- In case of a weak membrane energy, the morphing can be parametrized!

Graduated non-convexity

- GNC runs downhill on each of a sequence of functions
- It reaches a global optimum assuming a seuence of approximating and locally convex funtions exist.

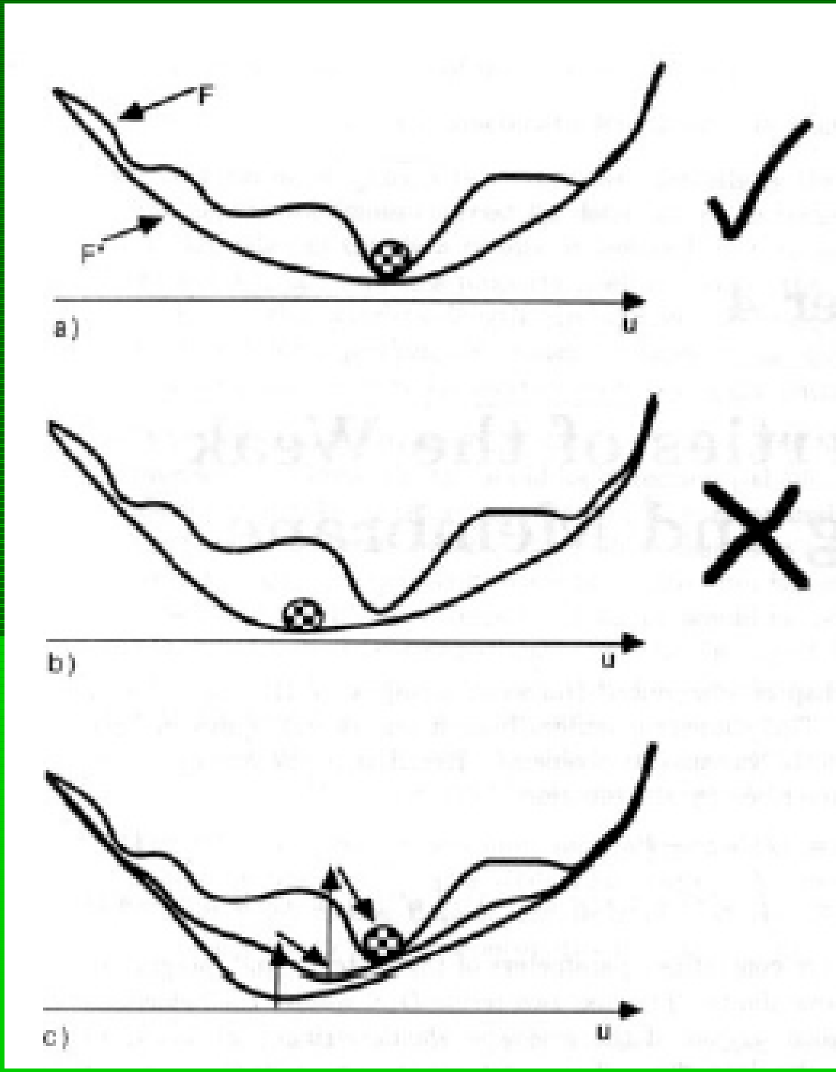


Image from *CMU 15-385 Computer Vision course, Spring 2002 by Tai Sing Lee*

Convex approximation of the weak membrane energy

$$F^{(p)} = \sum_i (f_i - d_i)^2 + \sum_i g^{(p)}(f_i - f_{i-1})$$

- ***F^(0)=E*** the original functional,
- ***F^(1)=F**** the convex approximation

$$c = c^*/p, \qquad c^* = \begin{cases} 1/2 & string \\ 1/4 & membrane \end{cases}$$

$$r^2 = \left(\frac{2}{c} + \frac{1}{\lambda}\right) \qquad q = \frac{\alpha}{\lambda^2 r}$$

Contours are defined as the set of ***i*** for which ***$|f_i - f_{i-1}| > 0$***

$$g^{(p)}(t) = \begin{cases} \lambda^2 t^2 & if \quad |t| < q \\ \alpha - c(|t| - r^2)^2/2 & if \quad q \le |t| < r \\ \alpha & if \quad |t| \ge r \end{cases}$$

Convex approximation of the weak membrane energy

- It is shown that $F^{(p)}$ is convex for $p \geq 1$
  - $F^{(1)}$ can be minimized using gradient descent
- As $p \to 0$
  - Increased localization of boundaries (l)
  - Gradual anisotropic smoothing of surface (f)
- Parameters used for the test: $\lambda=6$, $\alpha=0.03$

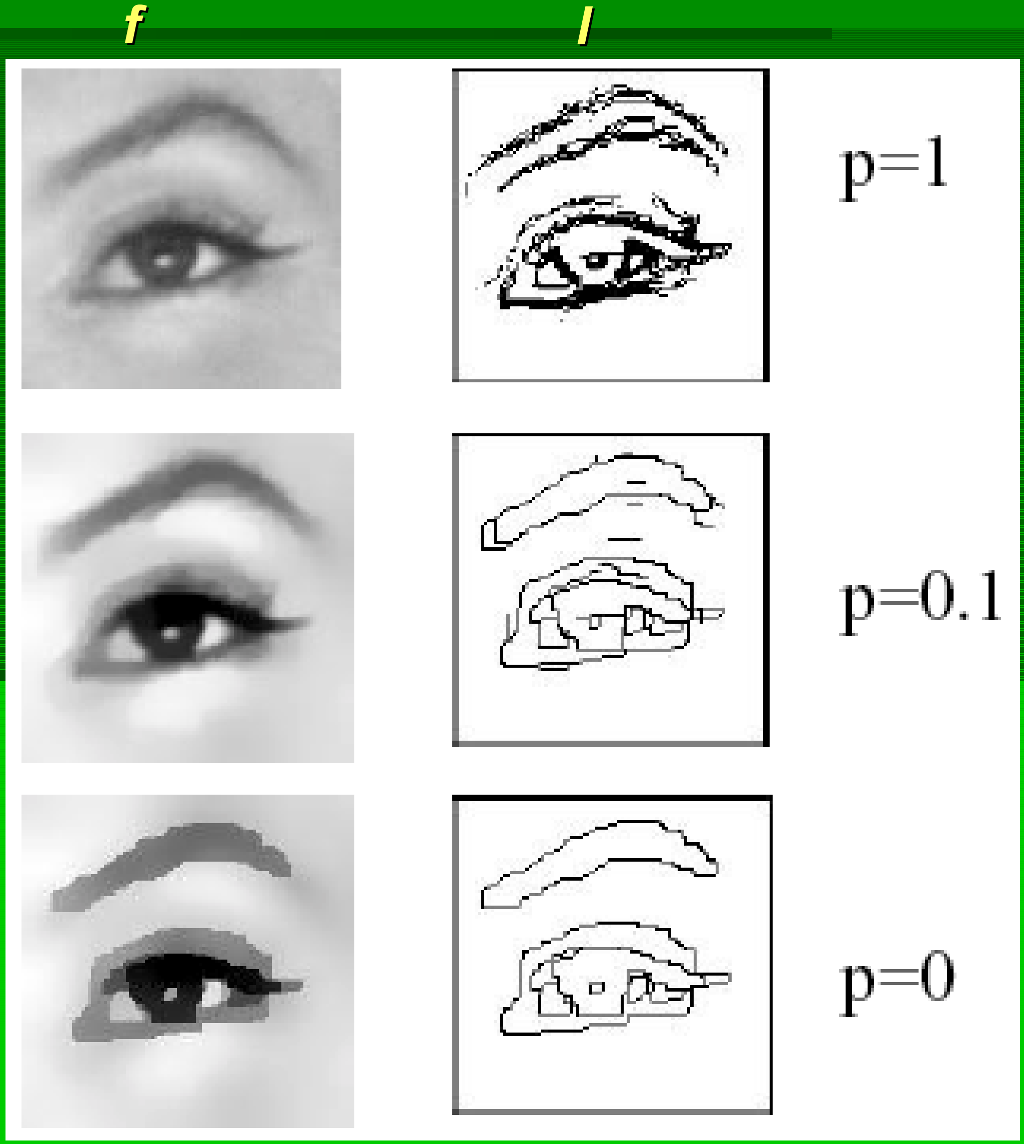


Image from *CMU 15-385 Computer Vision course, Spring 2002 by Tai Sing Lee*

Energy functional – MRF equivalence

- Formal equivalence between the two approaches. For example:

$$E(f) = \sum_i (f_i - d_i)^2 + \lambda \sum_i (f_{i+1} - f_i)^2 (1 - l_i) + \alpha \sum_i l_i$$

- Taking exponential (~ Hammersley-Clifford)
  - ***T*** ~ uncertainty („*temperature*”)

$$e^{-E(f)/T} = \underbrace{\prod_i e^{-(f_i - d_i)^2/T}}_{\text{data term (Gaussian)}} \underbrace{\prod_i e^{-(\lambda (f_{i+1} - f_i)^2 (1-l_i) + \alpha l_i)/T}}_{\text{smoothness prior (MRF)}}$$

Energy functional – MRF equivalence

- ***Size of the neighborhood*** in the MRF (or Gibbs field) corresponds to the ***degree of derivatives*** in the energy functional
  - Membrane: $(f_{i+1}-f_i)^2$
  - Thin plate: $(f_{i+1}-2f_i+f_{i-1})^2$