Introduction to Variational Methods

for

Computer Vision and Image Processing

Dr. Juan Fco. Garamendi

Universitat Pompeu Fabra

jf.garamendi@upf.edu

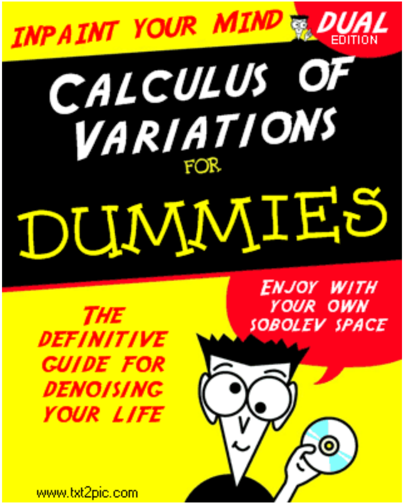
Optimization and Inference Techniques for Computer Vision

Master in Computer Vision Barcelona

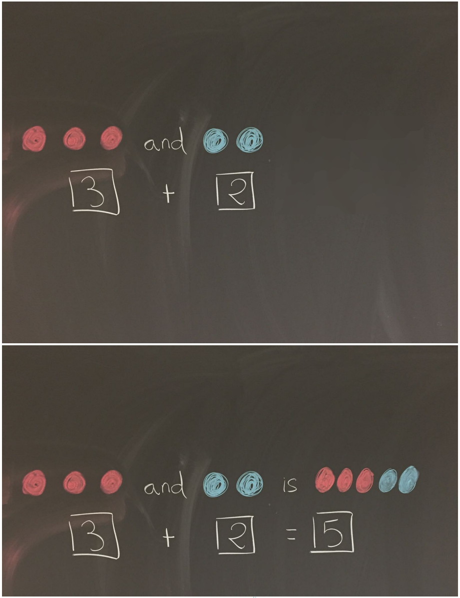
October 4. 2021

This is a conceptual (and not formal) introduction to understand

- Relationship between digital image processing, mathematical analysis and linear algebra.
- Problem modelling as energy functional minimization.
- Partial differential equations (PDEs) and why they appear.
- Algebraic system of equations and why they appear.
- Notation.



Mathematical Modelling



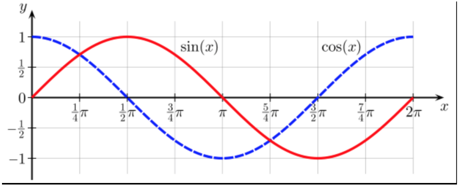
- A mathematical model is a description of a system using mathematical concepts and language.
- A Model may help to explain a system and to study the effects of different components, and to make

A (mathematical) concept we will need

Function

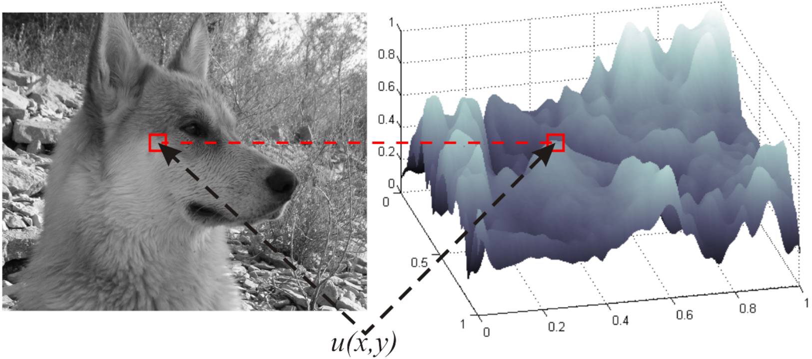
- Relation between a set of inputs and a set of permissible outputs.
- Each input is related to exactly one output.
- There are many ways to describe or represent a function
  - By a formula. $f(x) = x^2$
  - By an algorithm that tells how to compute the output.
  - **By a picture, called the graph of a function.**
  - By a table,
  - Implicitly as a solution of a (partial) differential equation.

(Wikipedia)



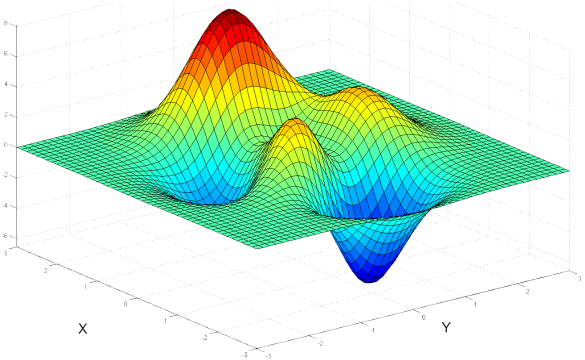
Modelling Images as Functions

For each position (x, y) there is a value $u(x, y)$.



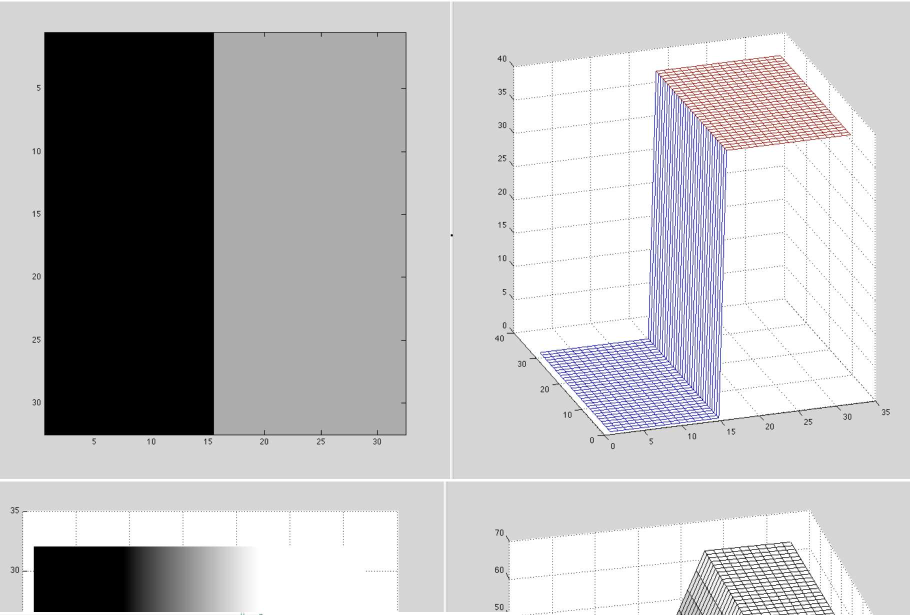
- Image u(x,y) with $u : \Omega \rightarrow \mathbb{R}$
- $\Omega \subset \mathbb{R}^2$, $\Omega = [0, 1]x[0, 1]$
- $(x, y) \in \Omega$ denotes pixel location

Modelling Images as Functions



Now forget where you come from and apply everything you know about functions.... If you know something

Modelling Images as Functions

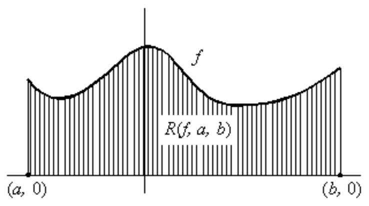


What can we do with functions?

What can we do with functions?

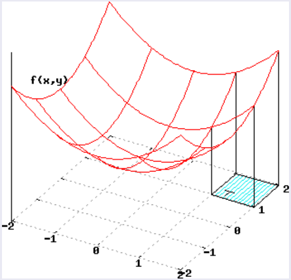
Integral

f(x)



$$\int_a^b f(x)dx$$

f(x,y)

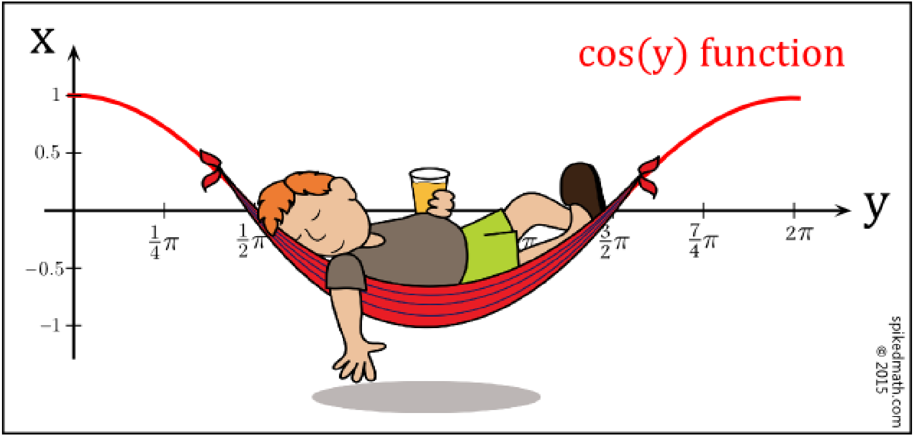


$$\int_a^b \int_c^d f(x,y)dxdy$$

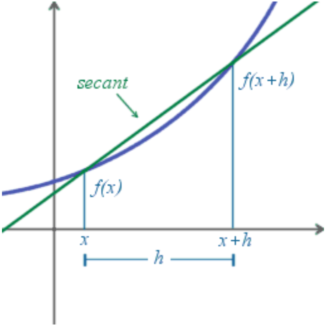
$$\int_\Omega f(x,y)dxdy$$

n Integral: The area under the

What can we do with functions?



$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$



The derivative is a measure of how a function (locally) **changes** as its input changes by a small amount h. Given a function, the result is another function. It is a LOCAL measure.

What can we do with functions?

The scalar case: Derivative

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Let's be h very, very small, very close to 0, then

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

$$hf'(x) \approx f(x+h) - f(x)$$

$$f(x+h) \approx f(x) + hf'(x)$$

Notice the last expression is the first order Taylor's approximation of $f(x+h)$ at point x. Keep this in mind, you will need for the **Optical Flow** problem.

What can we do with functions?

The Vectorial case: Directional Derivative

Suppose that $f : \mathbb{R}^2 \to \mathbb{R}$. Now we have infinite directions over we can measure the rate of change, just choose one given by the unitary vector $\vec{h}$, then

$$D_{\vec{h}} f(\vec{X}) = \lim_{\epsilon \to 0} \frac{f(\vec{X} + \epsilon \vec{h}) - f(\vec{X})}{\epsilon}$$

where $\vec{X} = (x_1, x_2)$
Special case: $\vec{h}$ is a vector of the basis, then we call it **Partial Derivative**

Partial derivatives

For $\vec{h} = (1, 0)$

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = \lim_{\epsilon \to 0} \frac{f(x_1 + \epsilon, x_2) - f(x_1, x_2)}{\epsilon}$$

For $\vec{h} = (0, 1)$

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = \lim_{\epsilon \to 0} \frac{f(x_1, x_2 + \epsilon) - f(x_1, x_2)}{\epsilon}$$

Gradient

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right)$$

Relationship between Gradient and Directional derivative

$$D_{\vec{h}} f(\vec{X}) = < \nabla f, \vec{h} >$$

What can we do with functions?

Some (boring but useful) definitions

Gradient Module

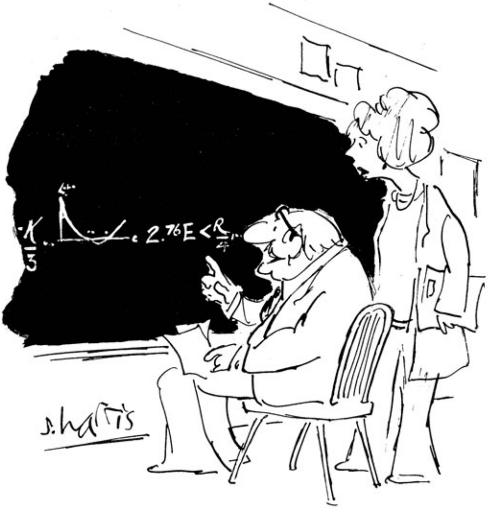
$$|\nabla f| = \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2}$$

Laplacian

$$\Delta f = (\nabla \cdot \nabla) f = \nabla^2 f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2}$$

For more (basic) properties.....Wikipedia is your friend

On Images



"THE BEAUTY OF THIS IS THAT IT IS ONLY OF THEORETICAL IMPORTANCE, AND THERE IS NO WAY IT CAN BE OF ANY PRACTICAL USE WHATSOEVER."

What can we do with functions?

Example of Integral

$$\int_{\Omega} |u(x,y) - f(x,y)|^2 dxdy$$

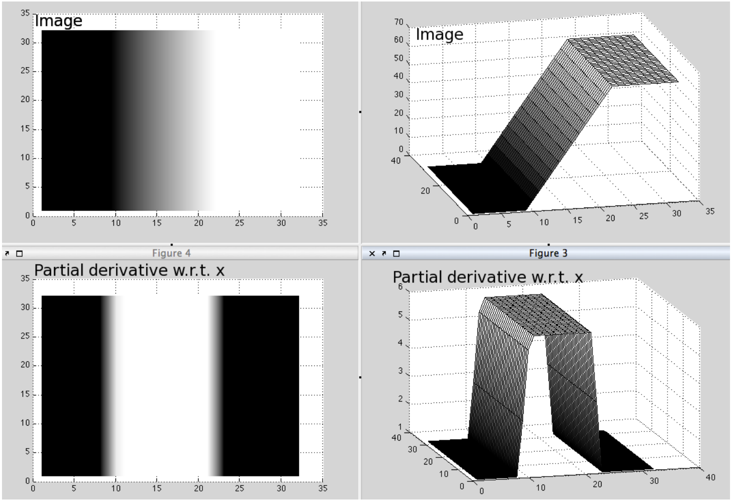
$$\int_{\Omega} |u - f|^2 dxdy$$

- $f(x,y) : \Omega \to \mathbb{R}$, Given image
- $u(x,y) : \Omega \to \mathbb{R}$, Given Image

What is it measured by this functional? It gives a number indicating a kind of difference between two images. We can know how much they are different but we do not know where.

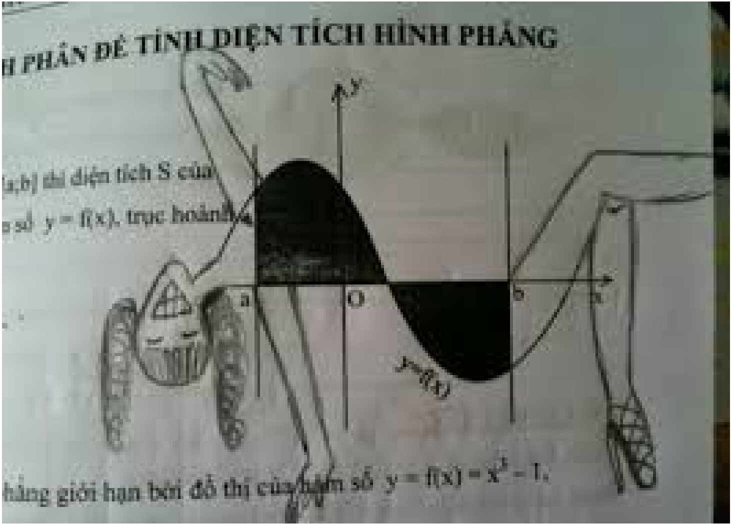
What can we do with functions?

Example of Derivative



Could someone tell me how the Partial Derivative w.r.t. y is?

What can we do with functions?



What can we do with functions?

Functional

- We can build a function of functions
- By a functional, we mean a correspondence which assigns a finite (real) number to each **function** belonging to some class.
- Thus, one might say that a functional is a kind of function, where the independent variable is itself a function.

Example 1

$$J(u) = \int_{\Omega} u(x, y) dxdy$$

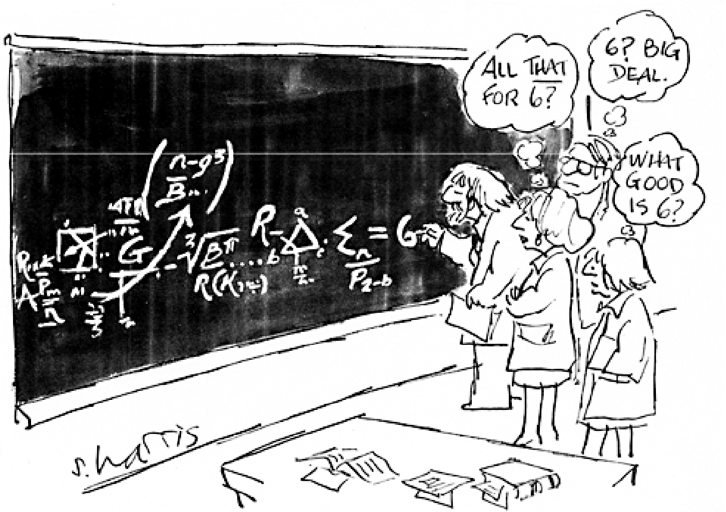
Example 2

$$J(u, f) = \int_{\Omega} |u(x, y) - f(x, y)|^2 dxdy$$

Example 3

$$\int \quad y)| dxdy$$

Example 4: Putting it all together

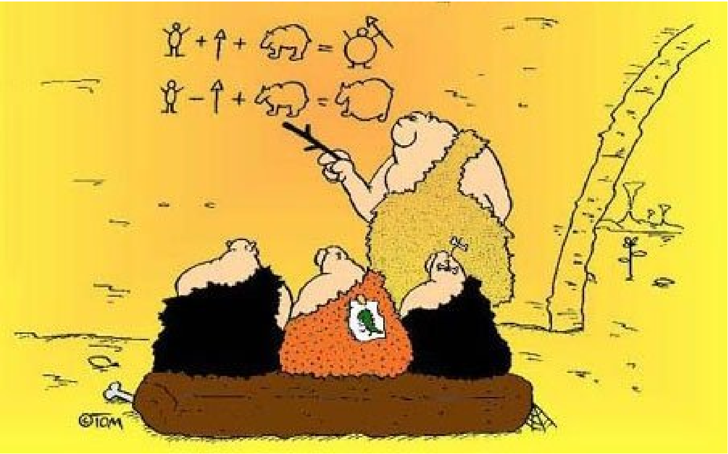


Example

$$\int_{\Omega} |\nabla u| dx + \int_{\Omega} |u - f|^2 dx$$

This energy functional measures similarity between *u* and *f* plus the overall change of u, i.e. smoothness.

Modelling Real World Problems as Minimization Problems



Modelling: Image restoration

Image Restoration: Image Restoration is the operation of taking a corrupt/**noisy** image and estimating the clean, original image.

Modelling: Image restoration

Given f, the (damaged) image to restore, we describe the characteristics of the (unknown) restored image u, the solution.

- Similar to the original image f: Overall difference between u and f is small.
- Without noise: The noise changes very quickly from one pixel to the next one. So the solution should be "smooth".

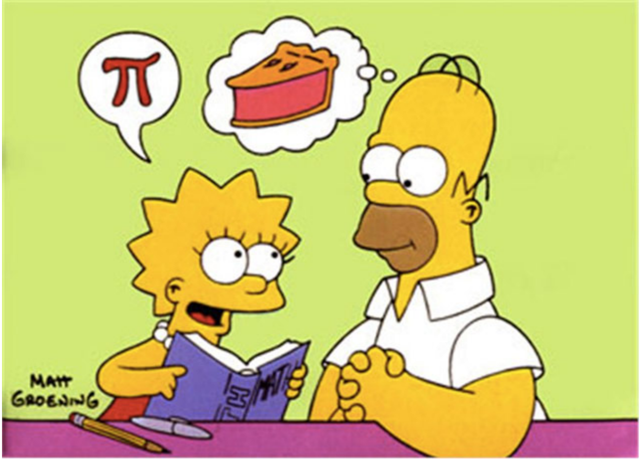
Functional

$$J(u) = \int_\Omega |\nabla u|^2 dx + \lambda \int_\Omega |u - f|^2 dx$$

$$J(u) = \underbrace{\int_\Omega |\nabla u|^2 dx}_{\text{Regularity of solution}} + \underbrace{\lambda \int_\Omega |u - f|^2 dxdy}_{\text{Fidelity to data}}$$

- $f(x) : \Omega \to \mathbb{R}$, **KNOWN**. The **given** noisy image
- $u(x) : \Omega \to \mathbb{R}$, **UNKNOWN**. The restored image
- ge that minimizes $J(u)$

Formally



Let

$$\begin{aligned} J: \quad V &\to \mathbb{R}, \\ u &\mapsto J(u) = \int_\Omega \mathcal{F}(x, u(x), \nabla u(x))\, dx \end{aligned}$$

be a functional over functions u, where

- V is a suitable space of functions
- $\Omega \in \mathbb{R}^D$ is a bounded open domain on the D dimensional space $\mathbb{R}^D$.
- $u \in V : \Omega \to \mathbb{R}^N$ is a function defined on a Ω.
- ∇ is the gradient operator and $x \in \Omega$ is the spatial variable.

The fundamental problem of the calculus of variations is to find the extremum (maximum or minimum) of the functional $J(u)$.

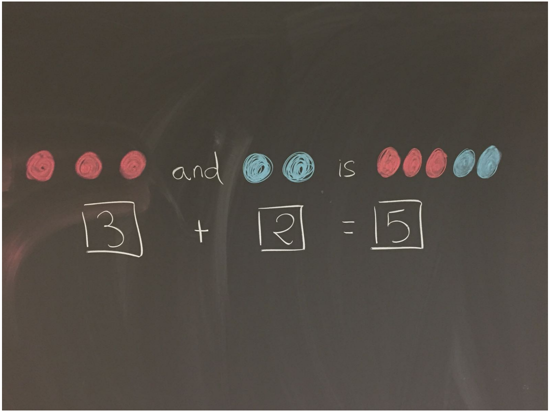
$$\arg\min_{u \in V} J(u)$$

Advantages

Modelling as Minization

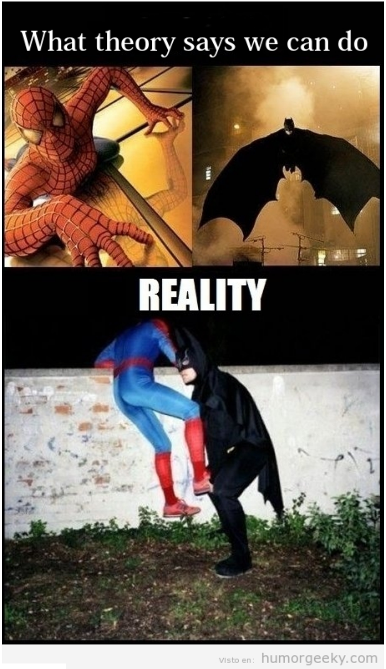
Advantage

Once the problem in the real world (denoising, restoration, segmentation, optical flow.....) is modelled as a minimization problem, we can apply all the well established mathematical theory and tools to solve the problem.



Examples

(wroten in a little bit formal way)



Given the image $f \in L^{\infty}(\Omega)$ solve

$$u = \underset{u \in W^{1,2}(\Omega)}{\arg\min} \left\{ \int_{\Omega} |\nabla u|^2 dxdy + \frac{1}{2\lambda} \int_{\Omega} |u - f|^2 dxdy \right\}$$

Where

- $W^{1,2}(\Omega) = \{u \in L^2(\Omega);\ \nabla u \in L^2(\Omega)^2\}$ is the space for u is well-defined.
  - $u \in L^2(\Omega)$ means $\int_{\Omega} u^2 dx < \infty$.
  - $\nabla u \in L^2(\Omega)^2$ means all derivatives up to order 1 belong to $L^2(\Omega)$
- λ Given parameter that controls the trade-off between data fidelity term and smoothness term.

Image Restoration

Given noisy image f

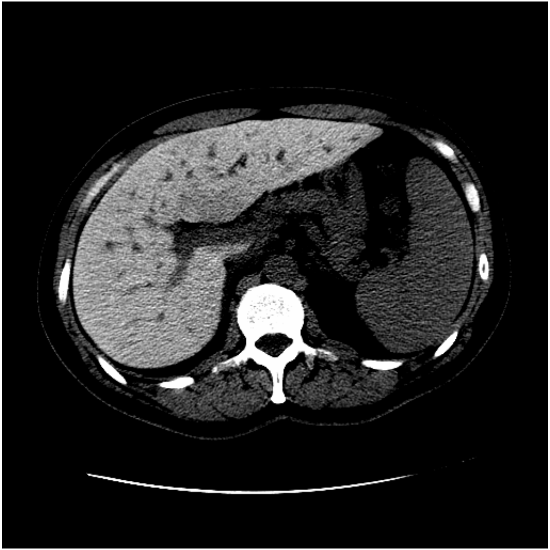


Image Restoration

Restored image u (minimum of $J(u)$)

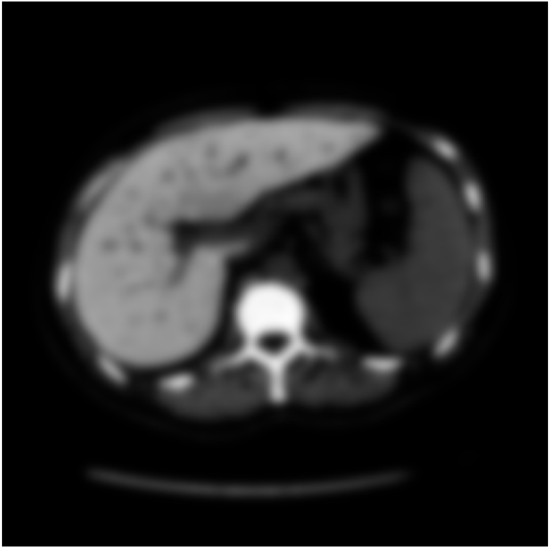


Image Restoration

Is u the solution to our mathematical problem (finding the minimum of the proposed energy functional)?

YES!!

but Is the solution u a good candidate for our restoration problem (the real world problem)?

NO!!

The boundaries of the objects have been lost.

Question

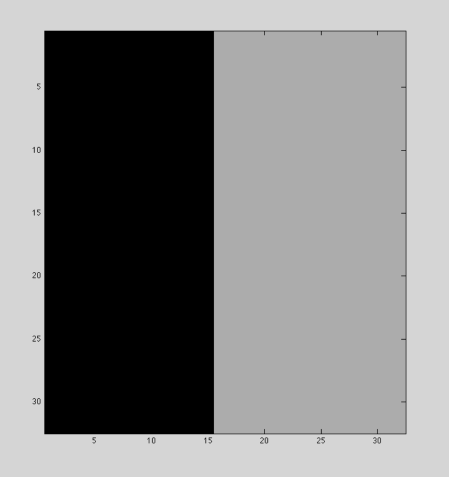
Which characteristic about regularity has the $W^{1,2}(\Omega)$ functional space?

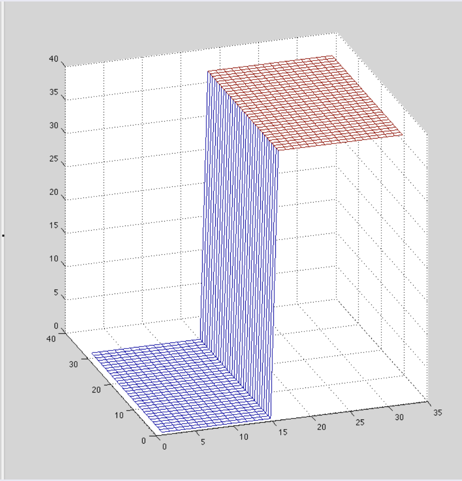
Answer

Functions with discontinuities does not belong to $W^{1,2}(\Omega)$ functional space, so it is too regular to model natural images.

Image Restoration

Could be this image a solution?





Another (and more intuitive) explanation: The L^p norm $|\cdot|^p$, with $p = 2$ of the gradient, $\int_\Omega |\nabla u|^2 dx$, penalizes too much the gradients corresponding to edges.
One should then decrease p in order to preserve the edges as much as possible.

Image Restoration

Total Variation, $p = 1$

$$\int_\Omega |\nabla u|^p dxdy$$

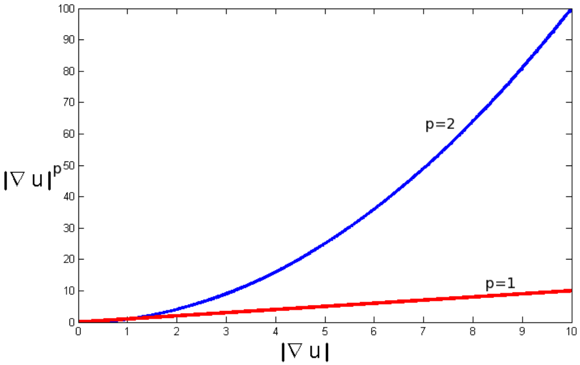
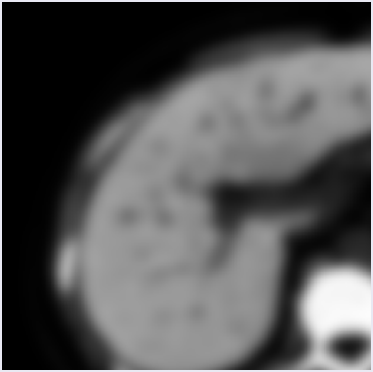
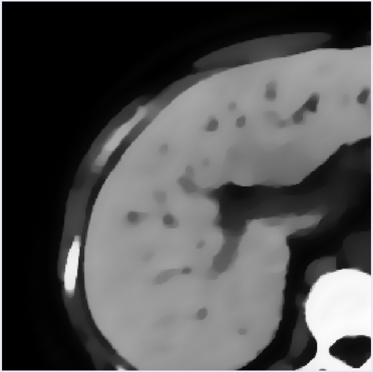


Image Restoration

p=2



p=1



Take home message

- A problem can be modelled in many different ways.
- The solution of the model may be is not the solution to your problem.

What is Inpainting?

Detail of "Cornelia, Mother of the Gracchi" by J. Suvee (Louvre). Courtesy of Emile-Male "The Restorer's Handbook of easel painting".

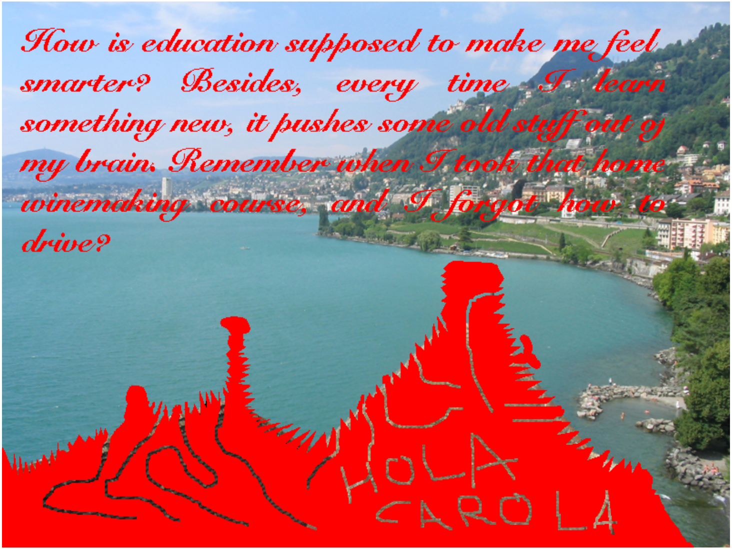
What is Inpainting?

What is Inpainting?

What is Inpainting?

What is not Inpainting?

Your Project

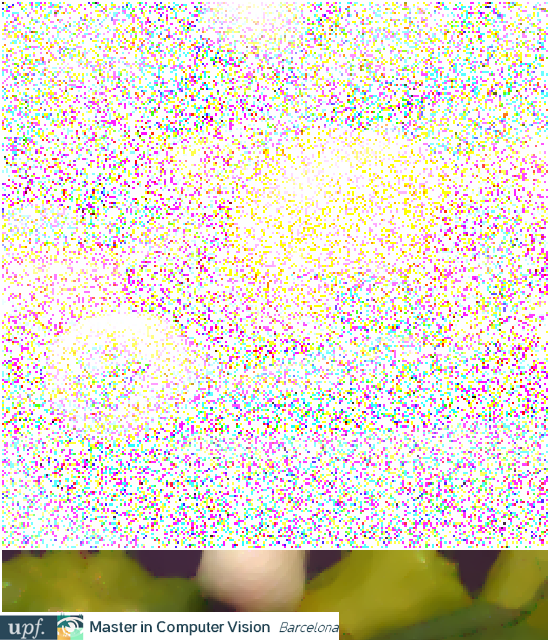


Inpainting using Laplace's Equation

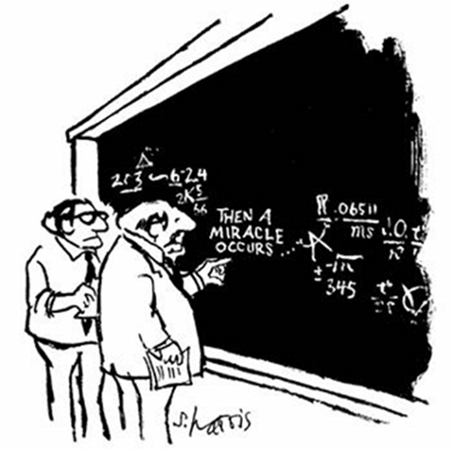
Let $f : \Omega \to \mathbb{R}$ be a given grayscale image and let $D \subset \Omega$ be an open set representing the region to be inpainted.
It is supposed that f is known in $\Omega \backslash D := \{x \in \Omega : x \notin D\}$
The inpainting solution u can be found as the minimum of

$$\begin{cases} \displaystyle\arg\min_{u \in W^{1,2}(\Omega)} \int_D |\nabla u(x)|^2 dx, \\ u|_{\partial D} = f \end{cases}$$

Inpainting



Understanding the Miracle: From the minimization problem to image solution



"I think you should be more explicit here in step two."

The Mathematical Problem to Solve

For a given energy functional

$$J(u) = \int_{\Omega} \mathcal{F}\left(x, u(x), \nabla u(x)\right) dx$$

The fundamental problem of the calculus of variations is to find the extremum (maximum or minimum) of the functional $J(u)$.

Solving the Mathematical Problem

Among others, there are two main strategies

- First strategy: Based on Gradient Descent:
  - Start from some initial solution u_0.

$$u^{[0]} = u_0$$

  - Iteratively go to another solution $u^{[k+1]}$, following a direction *dir* that is far from the previous solution $u^{[k]}$ by a distance τ and has lower energy.

$$u^{[k+1]} = u^{[k]} + \tau dir$$

  - Remember, gradient points to the direction of maximum change. Do previous step following the (some kind of generalization) gradient descent direction of your energy functional.

$$\begin{cases} u^{[0]} & = & u_0 \\ u^{[k+1]} & = & u^{[k]} - \tau \frac{d\,J(u)}{d\,u} \end{cases}$$

Solving the Mathematical Problem

- Second strategy: Finding the necessary condition for the extremum
  - The change of the functional with respect to the unkown/s has/have to be zero. i.e, it's derivative w.r.t. u has to be zero.
  - So, under some assumptions, if u is a extremum, u has to satisfy the equation.

$$\frac{dJ(u)}{du} = 0$$

  - More difficult to implement than Gradient Descent, but usually faster.

On both strategies

We have to measure the rate of change (i.e. compute the derivative) of the energy functional $J(u)$ with respect to the unknown u.
Have you noticed that u is a function?

It’s a complex business to derive w.r.t. functions!!

Remember Definition of Derivative

1D Derivative

$$f'(x) = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon}$$

Multidimensional Derivative

$$D_{\vec{h}} f(\vec{X}) = \lim_{\epsilon \to 0} \frac{f(\vec{X} + \epsilon \vec{h}) - f(\vec{X})}{\epsilon}$$

For functionals

We do something similar: The Gâteaux Derivative

$$\left.\frac{dJ}{du}\right|_h = \lim_{\alpha \to 0} \frac{J(u + \alpha h) - J(u)}{\alpha}$$

The Gâteaux Derivative

The Gâteaux derivative extends the concept of directional derivative (in differential calculus) to infinite-dimensional spaces.
The derivative w.r.t. u of the functional J(u) in the direction $h(x)$ is defined as

$$\left.\frac{dJ}{du}\right|_h = \lim_{\alpha \to 0} \frac{J(u + \alpha h) - J(u)}{\alpha}$$

IMPORTANT: Now, the direction is a FUNCTION!!!!!!

The Gâteaux Derivative

Let $u \in \Omega \to \mathbb{R}$ with $\Omega = [a, b] \subset \mathbb{R}$ being an open bounded subset of $\mathbb{R}$ (i.e. u is 1D). The canonical form of a functional is

$$J(u) = \int_\Omega \mathcal{F}(u, u')dx$$

and the Gâteaux Derivative is given by

$$\left.\frac{dJ}{du}\right|_h = \lim_{\alpha \to 0} \frac{1}{\alpha}\left(J(u + \alpha h) - J(u)\right)$$

$$= \lim_{\alpha \to 0} \frac{1}{\alpha}\left(\int_\Omega \mathcal{F}(u + \alpha h, (u + \alpha h)')dx - \int_\Omega \mathcal{F}(u, u')dx\right)$$

$$= \lim_{\alpha \to 0} \frac{1}{\alpha}\int_\Omega \mathcal{F}(u + \alpha h, (u + \alpha h)') - \mathcal{F}(u, u')dx$$

$$= \lim_{\alpha \to 0} \frac{1}{\alpha}\int_\Omega \mathcal{F}(u + \alpha h, u' + \alpha h') - \mathcal{F}(u, u')dx$$

Because of $\alpha \to 0$, we use taylor expansion of $\mathcal{F}(u + \alpha h, u' + \alpha h')$ in (u, u')

The Gâteaux Derivative

$$= \lim_{\alpha \to 0} \frac{1}{\alpha} \int_\Omega \mathcal{F}(u + \alpha h, u' + \alpha h') - \mathcal{F}(u, u') dx$$

Because of $\alpha \to 0$, we use taylor expansion of $\mathcal{F}(u + \alpha h, u' + \alpha h')$ in (u, u')

$$= \lim_{\alpha \to 0} \frac{1}{\alpha} \int_\Omega \mathcal{F}(u, u') + \frac{\partial \mathcal{F}}{\partial u} \alpha h + \frac{\partial \mathcal{F}}{\partial u'} \alpha h' + O(\alpha^2) - \mathcal{F}(u, u') dx$$

$$= \lim_{\alpha \to 0} \frac{1}{\alpha} \int_\Omega \cancel{\mathcal{F}(u, u')} + \frac{\partial \mathcal{F}}{\partial u} \alpha h + \frac{\partial \mathcal{F}}{\partial u'} \alpha h' + O(\alpha^2) - \cancel{\mathcal{F}(u, u')} dx$$

$$= \lim_{\alpha \to 0} \frac{1}{\cancel{\alpha}} \int_\Omega \cancel{\mathcal{F}(u, u')} + \frac{\partial \mathcal{F}}{\partial u} \cancel{\alpha} h + \frac{\partial \mathcal{F}}{\partial u'} \cancel{\alpha} h' + O(\alpha^{\cancel{2}}) - \cancel{\mathcal{F}(u, u')} dx$$

$$= \lim_{\alpha \to 0} \frac{1}{\cancel{\alpha}} \int_\Omega \cancel{\mathcal{F}(u, u')} + \frac{\partial \mathcal{F}}{\partial u} \cancel{\alpha} h + \frac{\partial \mathcal{F}}{\partial u'} \cancel{\alpha} h' + \cancel{O(\alpha^{\cancel{2}})} \to 0 - \cancel{\mathcal{F}(u, u')} dx$$

$$= \int_\Omega \frac{\partial \mathcal{F}}{\partial u} h + \frac{\partial \mathcal{F}}{\partial u'} h' dx$$

Integrition by parts $\int_a^b u dv = \underbrace{uv|_a^b}_{u(b)v(b) - u(a)v(a)} - \int_a^b v du$ over the second term

The Gâteaux Derivative

$$\left.\frac{dJ}{du}\right|_h = ... = \int_\Omega \frac{\partial \mathcal{F}}{\partial u} h + \frac{\partial \mathcal{F}}{\partial u'} h' dx$$

Integration by parts $\int_a^b udv = \underbrace{uv|_a^b}_{u(b)v(b)-u(a)v(a)} - \int_a^b vdu$ over the second term, choosing

$$u = \frac{\partial \mathcal{F}}{\partial u'} \qquad dv = h' dx$$

we have

$$du = \frac{\partial}{\partial x}\frac{\partial \mathcal{F}}{\partial u'} dx \qquad v = h$$

$$\left.\frac{dJ}{du}\right|_h = ... = \int_\Omega \left(\frac{\partial \mathcal{F}}{\partial u} h - \frac{\partial}{\partial x}\frac{\partial \mathcal{F}}{\partial u'} h\right) dx + \left.\frac{\partial \mathcal{F}}{\partial u'} h\right|_a^b$$

$$= \int_\Omega \left(\frac{\partial \mathcal{F}}{\partial u} - \frac{\partial}{\partial x}\frac{\partial \mathcal{F}}{\partial u'}\right) hdx + \left.\frac{\partial \mathcal{F}}{\partial u'} h\right|_a^b = 0$$

The Gâteaux Derivative

$$\left.\frac{dJ}{du}\right|_h = ... = \int_\Omega \left(\frac{\partial \mathcal{F}}{\partial u} - \frac{\partial}{\partial x}\frac{\partial \mathcal{F}}{\partial u'} \right) h dx + \left.\frac{\partial \mathcal{F}}{\partial u'} h \right|_a^b = 0$$

We can impose (at the same time) two conditions

$$\begin{cases} 1) \quad \dfrac{\partial \mathcal{F}}{\partial u} - \dfrac{\partial}{\partial x}\dfrac{\partial \mathcal{F}}{\partial u'} & = 0 \\[2ex] 2) \quad \left.\dfrac{\partial \mathcal{F}}{\partial u'} h \right|_a^b & = 0 \end{cases}$$

Let us to study the second condition

The Gâteaux Derivative

Options on the boundary (boundary conditions):

1. $h(x)|_{\partial\Omega} = 0$, means that $u(x)|_{\partial\Omega}$ is fixed to some value. These are the **Dirichlet boundary conditions.** When $u(x)|_{\partial\Omega} = 0$, they are called homogeneous Dirichlet boundary conditions.
2. $h(x)|_{\partial\Omega} \neq 0$, means that one allows variations of $u(x)$ on the boundary, but then, following condition has to be fullfilled.

$$\left.\frac{\partial \mathcal{F}}{\partial u'}\right|_{\partial\Omega} = 0$$

   These are called **Von Neumann boundary conditions**

(note: $\partial\Omega$ means boundary of Ω)

The Gâteaux Derivative

Let me remind you something

$$\left.\frac{d\,J(u)}{d\,u}\right|_h = \left\langle \frac{d\,J(u)}{d\,u}, h \right\rangle = \int \frac{d\,J(u)}{d\,u} h dx$$

And we aldeady prooved that

$$\left.\frac{d\,J(u)}{d\,u}\right|_h = \int_\Omega \left(\frac{\partial \mathcal{F}}{\partial u} - \frac{\partial}{\partial x}\frac{\partial \mathcal{F}}{\partial u'} \right) h dx$$

So

$$\frac{d\,J(u)}{d\,u} = \left(\frac{\partial \mathcal{F}}{\partial u} - \frac{\partial}{\partial x}\frac{\partial \mathcal{F}}{\partial u'} \right)$$

Euler-Lagrange Equations

For $\Omega \subset \mathbb{R}$
Dirichlet Boundary Conditions

$$\frac{\partial \mathcal{F}}{\partial u} - \frac{\partial}{\partial x}\frac{\partial \mathcal{F}}{\partial u'} = 0 \qquad u(x)|_{\partial\Omega} = u_0(x)$$

Von Neumann Boundary Conditions

$$\frac{\partial \mathcal{F}}{\partial u} - \frac{\partial}{\partial x}\frac{\partial \mathcal{F}}{\partial u'} = 0 \qquad \left.\frac{\partial \mathcal{F}}{\partial u'}\right|_{\partial\Omega} = 0$$

For $\Omega \subset \mathbb{R}^D$

$$\boxed{\underbrace{\frac{\partial \mathcal{F}}{\partial u} - \sum_{i=1}^{D}\frac{\partial}{\partial x_i}\frac{\partial \mathcal{F}}{\partial u_{x_i}}}_{\frac{d\,J(u)}{d\,u}} = 0}$$

where $\nabla u = (u_{x_1}, .., u_{x_D})$. The equation is completed with the corresponding boundary conditions

The Fundamental Problem of Variational Calculus

For a given energy functional

$$J(u) = \int_\Omega \mathcal{F}\left(x, u(x), \nabla u(x)\right) dx$$

We saw that the fundamental problem of the calculus of variations is to find the extremum (maximum or minimum) of the functional $J(u)$.

Necessary condition for the extremum

Under some assumptions over $\mathcal{F}$ and for the derivatives, if function u is a extremum of $J(u)$, then u is solution of the Euler-Lagrange (partial) differential equation

$$\frac{d\,J}{d\,u} = -\sum_{i=1}^{D} \frac{\partial}{\partial x_i} \frac{\partial \mathcal{F}}{\partial u_{x_i}} + \frac{\partial \mathcal{F}}{\partial u} = 0$$

where $\nabla u(x) = (u_{x_1}, ..u_{x_D})$ and D is the dimension .

Example:

Denoising with the L^2 norm regularity term

Example

Denoising with the L^2 norm regularity term

The problem of denoising a given image is modelled as the minimization of

$$\frac{1}{2}\int_{\Omega} |\nabla u|^2 dxdy + \frac{1}{2\lambda}\int_{\Omega} |u - f|^2 dxdy$$

We can find the solution using gradient descent.

$$\begin{cases} u^{[0]} & = & u_0 \\ u^{[k+1]} & = & u^{[k]} - \tau \frac{d\,J(u)}{d\,u} \end{cases}$$

(stop when $(u^{[k+1]} - u^{[k]}) < \epsilon$)

Example

Denoising with the L^2 norm regularity term

Let's compute $\frac{d\,J(u)}{d\,u}$

$$\mathcal{F}\left(x, u(x), \nabla u(x)\right) = \frac{1}{2}|\nabla u|^2 + \frac{1}{2\lambda}|u - f|^2$$

Notice that $\nabla u = (u_x, u_y)$, so $\mathcal{F}$ can be written as

$$\mathcal{F}\left(x, u, \nabla u\right) = \frac{1}{2}\underbrace{\sqrt{u_x^2 + u_y^2}^{\,2}}_{u_x^2+u_y^2} + \frac{1}{2\lambda}|u - f|^2$$

Compute the needed elements

$$\frac{\partial \mathcal{F}}{\partial u} = \frac{1}{\cancel{2}\lambda}\cancel{2}(u - f)$$

$$\frac{\partial \mathcal{F}}{\partial u_x} = \frac{1}{\cancel{2}}\cancel{2}(u_x) \qquad \frac{\partial \mathcal{F}}{\partial u_y} = \frac{1}{\cancel{2}}\cancel{2}(u_y)$$

Example

Denoising with the L^2 norm regularity term

So

$$\frac{dJ(u)}{du} = \frac{1}{\lambda}(u - f) - \left(\frac{\partial}{\partial x}u_x + \frac{\partial}{\partial y}u_y\right)$$

$$\frac{dJ(u)}{du} = \frac{1}{\lambda}(u - f) - \underbrace{(u_{xx} + u_{yy})}_{\Delta u}$$

$$\frac{dJ(u)}{du} = \frac{1}{\lambda}(u - f) - \Delta u$$

And the gradient descent scheme is

$$\begin{cases} u^{[0]} & = & u_0 \\ u^{[k+1]} & = & u^{[k]} - \tau\left(\frac{1}{\lambda}(u - f) - \Delta u\right) \end{cases}$$

Inconvenient: There are u without superscript.... Easy: Choose the superscript you want!!!!

Choosing superscripts

Choose superscript for u (advice: $[k+1]$)

$$u^{[k+1]} = u^{[k]} - \tau \left(\frac{1}{\lambda} (u^{[k+1]} - f) - \Delta u \right)$$

$$u^{[k+1]} + \frac{\tau}{\lambda} u^{[k+1]} = u^{[k]} + \frac{\tau}{\lambda} f + \tau \Delta u$$

$$u^{[k+1]} (1 + \frac{\tau}{\lambda}) = u^{[k]} + \frac{\tau}{\lambda} f + \tau \Delta u$$

$$u^{[k+1]} = \frac{1}{(1 + \frac{\tau}{\lambda})} \left(u^{[k]} + \frac{\tau}{\lambda} f + \tau \Delta u \right)$$

Choosing superscripts

Choose superscript for Δu. Two possible choices (easy and hard)

- Easy: $[k]$. **Explicit methods** (everything on the right side is known)

$$u^{[k+1]} = \frac{1}{(1 + \frac{\tau}{\lambda})} \left(u^{[k]} + \frac{\tau}{\lambda} f + \tau \Delta u^{[k]} \right)$$

- Hard: $[k+1]$. **Implicit methods.** You have to compute derivatives of some function you do not know!!

$$u^{[k+1]} = \frac{1}{(1 + \frac{\tau}{\lambda})} \left(u^{[k]} + \frac{\tau}{\lambda} f + \tau \Delta u^{[k+1]} \right)$$

STOP!!!! Let's Recap

1. Model images as functions.
2. Model your image processing problem as a minimization of an energy functional.
3. Compute the derivative of the energy functional w.r.t the unknown.
4. Choose a (explicit or implicit) gradient descent method or solve the Euler-Lagrange equation.

Solving the Partial Differential Equation

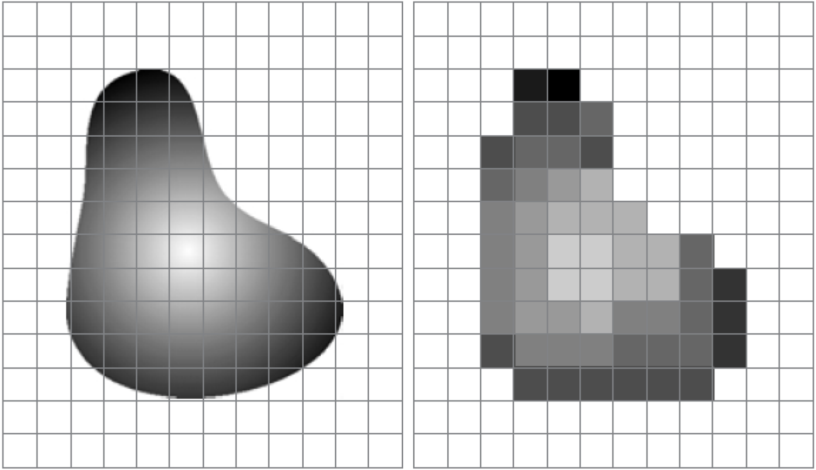
- We are not yet in the pixel level.
- How can we computationally solve the equation? Numerical Analysis

Numerical analysis is the study of algorithms that use numerical ***approximation*** (as opposed to general symbolic manipulations) for the problems of mathematical analysis (as distinguished from discrete mathematics).

wikipedia

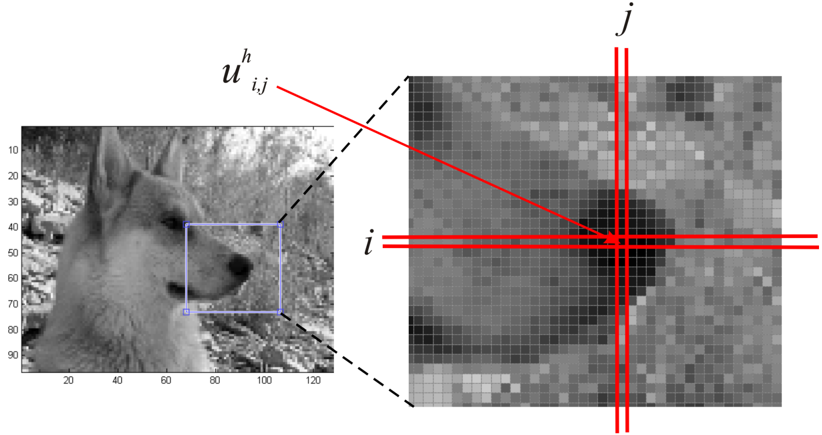
Discretization

Discretization concerns the process of transferring continuous functions, models, and equations into discrete counterparts



Discretization

Discretization concerns the process of transferring continuous functions, models, and equations into discrete counterparts



$$u(x, y) \approx u_{ij}$$

h: Distance between pixels.

Finite differences

We can approximate the derivative using Taylor as (there are many other possibilities)

$$\frac{\partial u(x,y)}{\partial x} \approx \frac{u(x+h,y)-u(x,y)}{h}$$

$$\frac{\partial u(x,y)}{\partial x} \approx \frac{u_{i+1,j}-u_{i,j}}{h}$$

$$\frac{\partial u(x,y)}{\partial y} \approx \frac{u_{i,j+1}-u_{i,j}}{h}$$

$$\frac{\partial^2 u(x,y)}{\partial x^2} \approx \frac{u_{i+1,j}-2u_{i,j}+u_{i-1,j}}{h^2}$$

$$\frac{\partial^2 u(x,y)}{\partial y^2} \approx \frac{u_{ij+1}-2u_{ij}+u_{ij-1}}{h^2}$$

Finite differences

take h=1

Gradient Descent Example

Energy Functional

$$J(u) = \int_{\Omega} \frac{1}{2}|\nabla u|^2 + \frac{1}{2\lambda}|u - f|^2 dx$$

Rate of change w.r.t u

$$\frac{dJ}{du} = \frac{1}{\lambda}(u - f) - \Delta u$$

Gradient descent scheme

$$u^{[k+1]} = u^{[k]} - \tau \left(\frac{1}{\lambda}(u - f) - \Delta u \right)$$

$$u^{[k+1]} = \frac{1}{(1 + \frac{\tau}{\lambda})} \left(u^{[k]} + \frac{\tau}{\lambda} f + \tau \Delta u \right)$$

Explicit gradient descent scheme:

$$u_{i,j}^{[k+1]} = \frac{1}{(1 + \frac{\tau}{\lambda})} \left[u_{i,j}^{[k]} + \frac{\tau}{\lambda} f_{i,j} + \tau \left(u_{i+1,j}^{[k]} - 2u_{i,j}^{[k]} + u_{i-1,j}^{[k]} + u_{ij+1}^{[k]} - 2u_{ij}^{[k]} + u_{ij-1}^{[k]}) \right) \right]$$

Implicit gradient descent scheme :

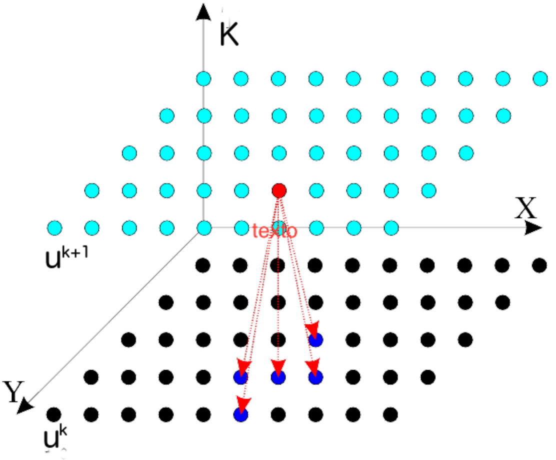
$$\left(4\tau + 1 + \frac{\tau}{\lambda}\right) u_{i,j}^{[k+1]} - \tau \left(u_{i+1,j}^{[k+1]} + u_{i-1,j}^{[k+1]} + u_{ij+1}^{[k+1]} + u_{ij-1}^{[k+1]} \right) = u_{i,j}^{[k]} + \frac{\tau}{\lambda} f_{ij}$$

Numerical Aspects

Gradient Descent

Explicit schemes

$$u_{i,j}^{[k+1]} = \frac{1}{(1+\frac{\tau}{\lambda})}\left[u_{i,j}^{[k]} + \frac{\tau}{\lambda} f_{i,j} + \tau\left(u_{i+1,j}^{[k]} - 2u_{i,j}^{[k]} + u_{i-1,j}^{[k]} + u_{ij+1}^{[k]} - 2u_{ij}^{[k]} + u_{ij-1}^{[k]})\right)\right]$$



Implicit schemes

With implicit methods, an algebraic system of equations $A\vec{x} = \vec{b}$ has to be solved where (usually A is sparse)

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \qquad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

or leads to an algebraic system of non-linear equations $A(x) = b$

Example:

$$\left(4\tau + 1 + \frac{\tau}{\lambda}\right) u_{i,j}^{[k+1]} - \tau\left(u_{i+1,j}^{[k+1]} + u_{i-1,j}^{[k+1]} + u_{ij+1}^{[k+1]} + u_{ij-1}^{[k+1]}\right) = u_{i,j}^{[k]} + \frac{\tau}{\lambda} f_{ij}$$

Ordering the pixels lexycographically by columns:

$$\begin{aligned} a_{ii} &= \left(4\tau + 1 + \tfrac{\tau}{\lambda}\right) \\ a_{i,i-1} &= -\tau \\ a_{i,i+11} &= -\tau \\ a_{i,i+n_x} &= -\tau \\ a_{i,i-n_x} &= -\tau \end{aligned} \qquad \vec{x} = \begin{bmatrix} u_{11}^{[k+1]} \\ u_{21}^{[k+1]} \\ \vdots \\ u_{12}^{[k+1]} \\ u_{22}^{[k+1]} \\ \vdots \\ u_{n_x,n_y}^{[k+1]} \end{bmatrix}, \qquad b = \begin{bmatrix} u_{11}^{[k]} + \frac{\tau}{\lambda} f_{11} \\ u_{21}^{[k]} + \frac{\tau}{\lambda} f_{21} \\ \vdots \\ u_{12}^{[k]} + \frac{\tau}{\lambda} f_{12} \\ u_{22}^{[k]} + \frac{\tau}{\lambda} f_{22} \\ \vdots \\ u^{[k]} + \frac{\tau}{\lambda} f_{n_x,n_y} \end{bmatrix}$$

Being n_x and n_y the number of pixels on x and y directions.

- Explicit gradient descent is very easy to implement.
- Explicit gradient descent is naturally parallel.
- Explicit gradient descent is conditionally stable, so τ has to be small $\rightarrow$ large number of iterations to converge.
- Implicit gradient descent use to be unconditionally stable so τ can be bigger $\rightarrow$ less number of iterations to converge.
- Implicit Gradient descent needs less iterations but an algebraic system of equations has to be solved.

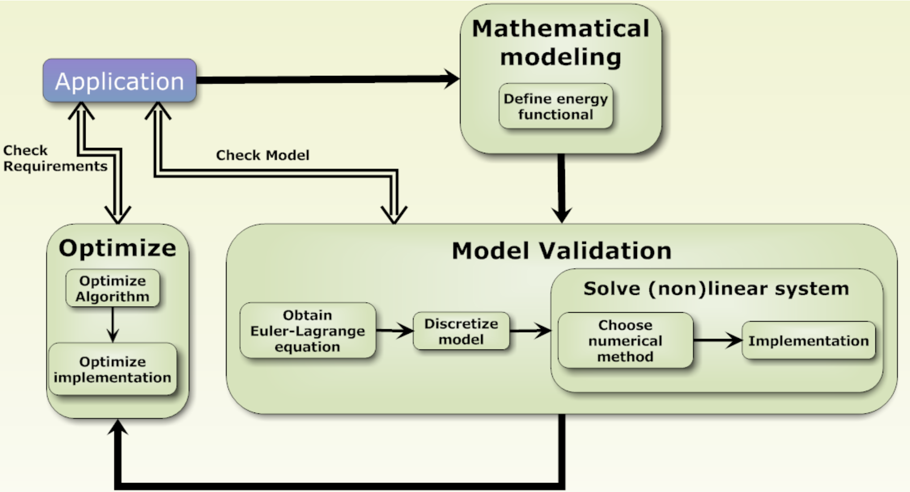
Boundary conditions

What happens on the boundary of the images? We do not have $u_{i-1,j}$, $u_{i+1,j}$, $u_{i,j-1}$, $u_{i,j+1}$ pixels!!!!!!!
We impose ANOTHER completely different condition, the boundary conditions:

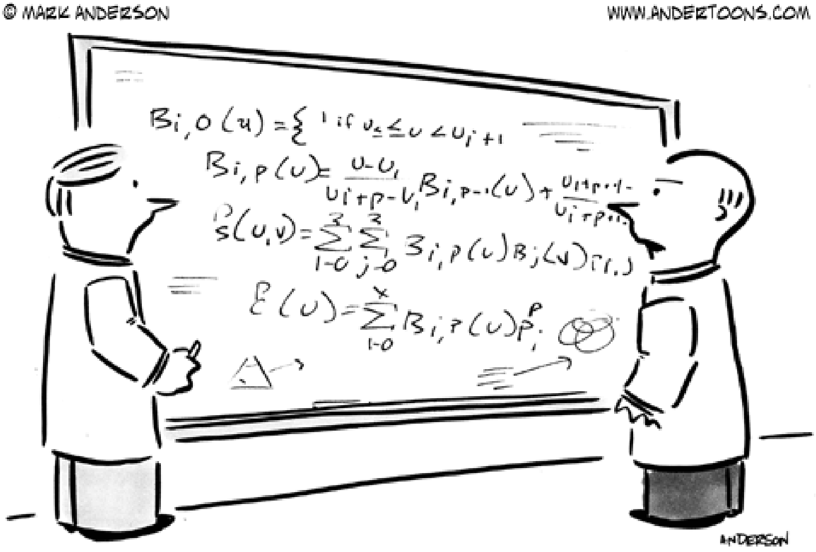
- Dirichlet: We know a priory the value of the **solution** u at the boundary, so we know the values for $u_{i-1,j}$, $u_{i+1,j}$, $u_{i,j-1}$, $u_{i,j+1}$
- Neumann: We know a priory the **rate of change** u at the boundary, so we know that

$$\left.\frac{\partial u}{\partial \bar{n}}\right|_{\partial\Omega} = \textit{some value}$$

Briefly


Mathematical modeling
Define energy functional
Application
Check Requirements
Check Model
Optimize
Optimize Algorithm
Optimize implementation
Model Validation
Solve (non)linear system
Obtain Euler-Lagrange equation
Discretize model
Choose numerical method
Implementation

Questions



"What the hell is *that* supposed to mean?!"