Optimal Control of PDEs Applied in Image Processing

Analysis of a Generalized Conditional Gradient Method

Dirk Lorenz, Peter Maass

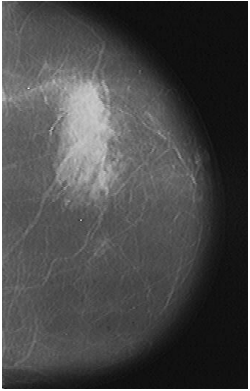
Applied Math Seminar, Tel Aviv University

November 15, 2005

Overview

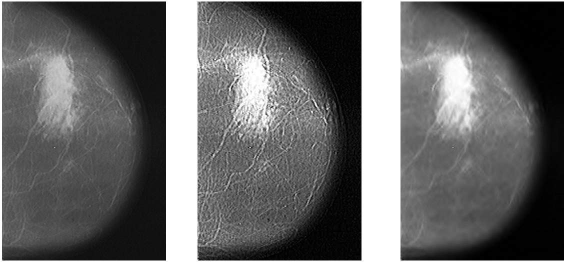
Outline

Mammography screening



- Examination of X-ray scans
- Several thousand scans per year
- Mammographie Projekt Bremen

The examination: a sweep across scales

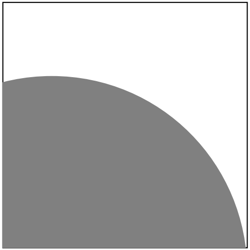


y_0 original $\quad$ y_f fine scales $\quad$ y_c coarse scale

Display: $y(t)$, $t \in [0,1]$, $y(0) = y_0$

$$y^*(t) = \begin{cases} y_f, & t = .3 \\ y_c, & t = .6 \end{cases}$$

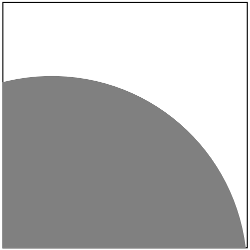
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a ‘natural’ way.

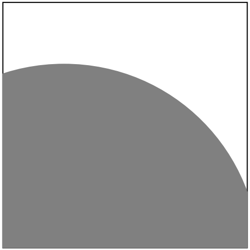
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a 'natural' way.

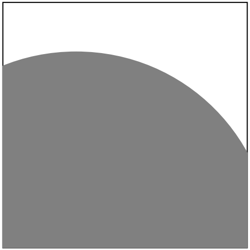
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a 'natural' way.

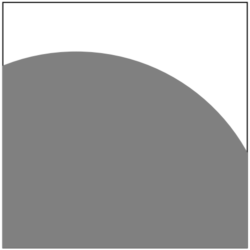
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a 'natural' way.

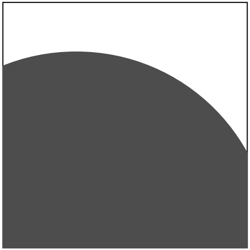
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a 'natural' way.

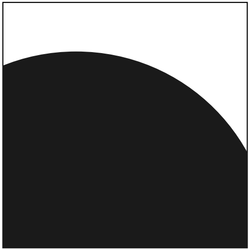
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a 'natural' way.

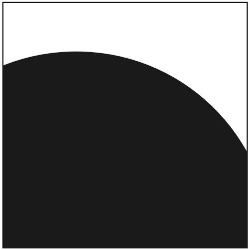
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a ‘natural’ way.

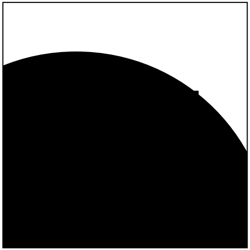
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a 'natural' way.

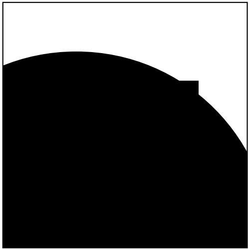
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a ‘natural’ way.

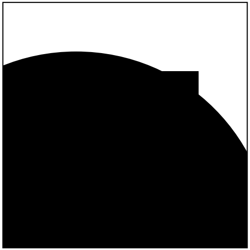
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a 'natural' way.

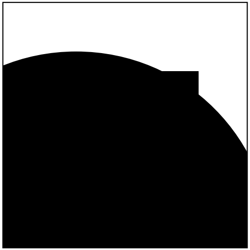
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a ‘natural’ way.

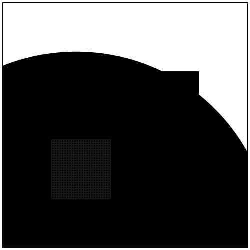
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a 'natural' way.

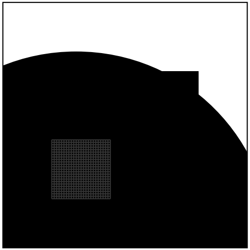
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a 'natural' way.

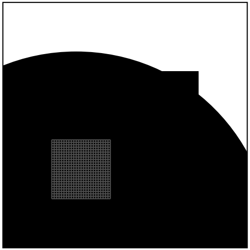
Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a 'natural' way.

Modeling differences between images



- Edges move
- Contrast changes
- Details created
- Texture changes

Given two images:
Interpolate in a ‘natural’ way.

Outline

Examples for PDEs in images processing

Remove structure The heat equation

$$y_t = \Delta y$$

Examples for PDEs in images processing

Remove structure The heat equation

$$y_t = \Delta y$$

Keep and enhance edges Anisotropic diffusion

$$y_t = \operatorname{div}(A(\nabla y)\nabla y)$$

Perona-Malik, Weickert

Examples for PDEs in images processing

Remove structure The heat equation

$$y_t = \Delta y$$

Keep and enhance edges Anisotropic diffusion

$$y_t = \operatorname{div}(A(\nabla y)\nabla y)$$

Perona-Malik, Weickert

Move edges Mean curvature motion, erosion, dilation

$$y_t = |\nabla y| \operatorname{div}\left(\frac{\nabla y}{|\nabla y|}\right), \; y_t = \pm|\nabla y|$$

Preusser-Rumpf, Alvarez-Guichard-Morel

A PDE for our purpose

Source term
allow contrast changes
allow formation of details

$$y_t - \operatorname{div}(P\nabla y) = u$$

Diffusion tensor
keep edges sharp,
allow movement of edges
remove textures

Outline

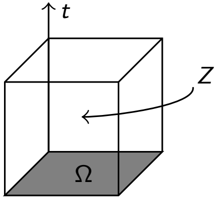
First model: Control via source terms

From one image y_0 to another y_*

Minimize the cost functional

$$\int |y(1) - y^*|^2 dx + \alpha \|u\|^2_{L^2(Z)}$$

where y solves the PDE



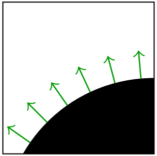
$$\begin{aligned} y_t - \Delta y &= u && \text{in } Z = \Omega \times [0,1] \\ \partial_\nu y &= 0 && \text{on } \partial\Omega \\ y(0,x) &= y_0(x). \end{aligned}$$

Second model: Source term and diffusion tensor

$$\min_u \int |y(1) - y^*|^2 dx + \alpha_u \|u\|_{L^2(Z)}^2 + \alpha_p \Phi_p(p)$$

where y solves

$$\begin{aligned} y_t - \operatorname{div}(P\nabla y) &= u \\ \partial_{\nu_p} y &= 0 \\ y(0,x) &= y_0(x). \end{aligned}$$



with diffusion tensor

$$P = (I - p \otimes p)$$

($|p| = 1$: projection on $p^\perp$)

Outline

Mapping control onto state

Solving the PDE

$$y_t + \Delta y = u$$

Solution operator

$$\begin{array}{rccc} & u & \mapsto & y \\ S: & L^2(Z) & \to & C([0,1], L^2(\Omega)) \end{array}$$

Mapping control onto state

Solving the PDE

$$y_t + \Delta y = u$$

Solution operator

$$\begin{array}{cccc} & u & \mapsto & y \\ S: & L^2(Z) & \to & C([0,1], L^2(\Omega)) \end{array}$$

The control-to-state operator

$$\begin{array}{cccc} & u & \overset{S}{\longmapsto} y \overset{R}{\longmapsto} & y(1) \\ A: & L^2(Z) & \longrightarrow & L^2(\Omega) \end{array}$$

A is compact.

The reduced cost functional aka Tikhonov regularization

The reduced cost functional

With the help of the control to state mapping

$$\int |y(1) - y^*|^2 dx + \alpha \|u\|^2_{L^2(Z)}$$

becomes

$$\|Au - y^*\|^2_{L^2(\Omega)} + \alpha \|u\|^2_{L^2(Z)}.$$

Different tasks in different areas

$$\|Au - y^*\|_{L^2(\Omega)}^2 + \alpha\|u\|_{L^2(Z)}^2$$

Control problems:

- A (semi-)linear PDE with source terms
- y^* desired state
- usually $u \in C$

Inverse problems:

- A any compact operator
- y^* noisy measurement
- regularity of u

Closer to image processing

$$\|Au - y^*\|^2 + \alpha |u|^p_{B^s_{p,p}}$$

$p \geq 1$, $s \geq 0$
Daubechies, Defries, De Mol, 2004

$|\cdot|^p$ for $p = 1$ non differentiable

Classical cases:

1. $p = 2$, A compact (regularization theory)
2. $p \neq 2$, $A = I$ (denoising in image processing)

Outline

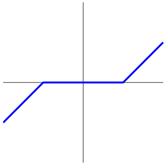
Tool for minimization: shrinkage

$\mathbf{S}_\tau$: shrinkage operation on a basis expansion

$$u = \sum_j \langle u|\psi_j\rangle \psi_j$$

$$\mathbf{S}_\tau(u) = \sum_j S_\tau(\langle u|\psi_j\rangle)\psi_j$$

$$S_\tau(z) = \operatorname{sign}(z)[|z| - \tau]_+$$



soft shrinkage

Explicit minimization for $A = I$

ψ_j wavelet base:

$$|u|^p_{B^s_{p,p}} \asymp \sum_j w_{p,j} |\langle u|\psi_j\rangle|^p$$

$$\|u - y^*\|^2 + \alpha |u|^p_{B^s_{p,p}} = \sum_j (\langle u|\psi_j\rangle - \langle y^*|\psi_j\rangle)^2 + \alpha w_{p,j} |\langle u|\psi_j\rangle|^p$$

minimization per coefficient, 1D problem

$$u = \mathbf{S}_{\alpha w_{p,j},p}(y^*), \quad \langle u|\psi_j\rangle = S_{\alpha w_{p,j},p}(\langle y^*|\psi_j\rangle)$$

$p = 1$: soft shrinkage, $p > 1$: generalized shrinkage function (Chambolle, Lucier 1998, L. 2003)

Iterative minimization for $A \neq I$

$$\|Au - y^*\|^2 + \alpha |u|^p_{B^s_{p,p}} = \sum_j (\langle Au|\psi_j\rangle - \langle y^*|\psi_j\rangle)^2 + \alpha w_{p,j} |\langle u|\psi_j\rangle|^p$$

$\rightsquigarrow$ coupled, nonlinear equations
Decouple A and $|\cdot|^p$, surrogate functional:

$$\|Au - y^*\|^2 + \alpha |u|^p_{B^s_{p,p}} + \left(\|u - a\|^2 - \|Au - Aa\|^2\right)$$

alternate minimization for u, a

Theorem

$$u_{n+1} = \mathbf{S}_{\alpha w_{p,j}, p}(u_n - A^*(Au_n - y^*))$$

converges strongly to a minimizer of $\|Au - y^*\|^2 + \alpha |u|^p_{B^s_{p,p}}$.

Daubechies, Defries, De Mol, 2004.

Outline

Classical: The conditional gradient method

$\min_{u \in C} F(u)$ by

Classical: The conditional gradient method

$\min_{u\in C} F(u)$ by

1. directional derivative

$$\min_{v\in C}\langle F'(u_n)|v\rangle$$

2. line search

$$\min_{s\in[0,1]} F(u_n + s(v_n - u_n))$$

3. update

$$u_{n+1} = u_n + s(v_n - u_n)$$

Dunn (1980), F convex, differentiable, e. g. $F(u) = \|Au - y^*\|^2$.

Preparing for generalization

$$I_C(u) = \begin{cases} 0 & u \in C \\ \infty & u \notin C \end{cases} \quad \rightarrow \quad \begin{array}{c} \min F(u) + \Phi(u) \\ \text{with } \Phi(u) = I_C(u) \end{array}$$

F: smooth, minimization hard
main ingredient

Φ: not differentiable, minimization easy,
influence rather small

$$\min_{v \in C} \langle F'(u_n) | v \rangle \quad \leftrightarrow \quad \min_{v} \langle F'(u_n) | v \rangle + \Phi(v)$$

Outline

Now: Generalized Conditional gradient method

$\min F(u) + \Phi(u)$ by

1. directional derivative

$$\min_v \langle F'(u_n) | v \rangle + \Phi(v)$$

2. line search

$$\min_{s \in [0,1]} (F + \Phi)(u_n + s(v_n - u_n))$$

3. update

$$u_{n+1} = u_n + s(v_n - u_n)$$

Convergence for the generalized conditional gradient method

Theorem

Φ proper, convex, lsc.
F continuously differentiable, $F + \Phi \geq 0$ and

$$E_t = \{\Phi(u) \leq t\} \text{ compact for every } t.$$

Then: convergence to a stationary point.

K. Bredies, D. L., P. Maass, 2005

Remark: F need not to be convex. Sketch of proof

Equivalence with the method of surrogate functionals

Corollary

The minimization of

$$\|Au - y^*\|^2 + \alpha \sum w_p, j |\langle u|\psi_j\rangle|^p$$

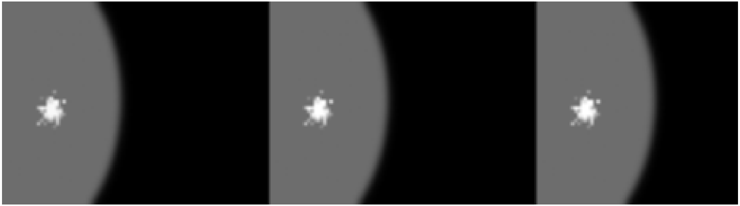
by means of the surrogate method is the same as the application of the generalized conditional gradient method for

$$F(u) = \|Au - y^*\|^2 - \lambda\|u\|^2, \quad \Phi(u) = \lambda\|u\|^2 + \sum w_{p,j}|\langle u|\psi_j\rangle|^p$$

for $\lambda = 1$ *without line search.*

→ The method of surrogate functionals is a gradient descent.

Illustration



Summary

1. Presentation of mammography images can be improved by optimal control of PDEs.
2. Non standard penalty terms arise from image processing.
3. The generalized conditional gradient method converges even for non-convex functionals.
4. The surrogate method from inverse problems is equivalent to a generalized conditional gradient method.

Convergence of the generalized conditional gradient method

Sketch of proof.

1. Necessary first order condition:

$$\forall v : \langle F'(u)|u\rangle + \Phi(u) = \min_v \langle F'(u)|v\rangle + \Phi(v)$$

2. Condition not fulfilled: $(F+\Phi)(u^{n+1}) < (F+\Phi)(u^n)$
3. $\Psi(u) := \langle F'(u)|u\rangle + \Phi(u) - \big(\min_v \langle F'(u)|v\rangle + \Phi(v)\big)$ is lsc
4. $\Psi(u^n) \to 0$

□

Back to theorem