Optimal control problems with sparsity constraints in image processing

Kristian Bredies, Dirk Lorenz, Peter Maass

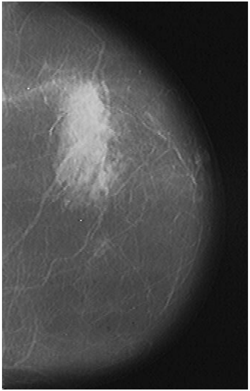
RPWAFT 2006

January 25, 2006

Outline

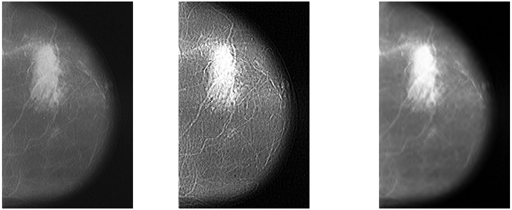
Outline

Mammography screening



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

The examination: a sweep across scales



y_0 original y_f fine scales y_c coarse scale

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

$$y(0) = y_0, \quad y(.5) = y_f, \quad y(1) = y_c$$

Easier: Just two images y_0 and y^* with

$$y(0) = y_0, \quad y(1) = y^*$$

Outline

Use an optimal controlled PDE for interpolation

An initial value problem

$$y_t = D(y)$$
$$y(0) = y_0$$

But we desire: $y(1) \stackrel{!}{=} y^*$

An optimal controlled initial value problem

$$\text{Minimize} \int |y(1) - y^*|^2 dx + \alpha \|u\|^2$$

subject to $y_t = D(u, y), \ y(0) = y_0$

What PDE to use for modelling the sweep?

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

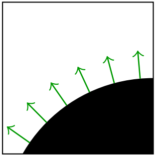
remove textures

One particular model: Source term and diffusion tensor

$$\min_{u,P,y} \int |y(1) - y^*|^2 dx + \alpha_u \|u\|^2 + \alpha_p \Phi_p(p)$$

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

$$P = (I - \sigma(|p|) \tfrac{p}{|p|} \otimes \tfrac{p}{|p|})$$

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

Outline

1 Introduction: Optimize presentation of mammography images

2 Modelling: Optimal control of PDEs

3 Algorithms: Generalized conditional gradient methods and the method of surrogate functionals

4 Application

From optimal control to inverse problems

The control-to-state operator

$$A : u \overset{\text{Solve PDE}}{\longmapsto} y \overset{\text{restict to endpoint}}{\longmapsto} y(1)$$

The reduced cost functional aka Tikhonov regularization

With the help of the control to state mapping

$$\int |y(1) - y^*|^2 dx + \alpha \|u\|^2$$

becomes

$$\|Au - y^*\|^2 + \alpha \|u\|^2.$$

Constraint or regularization?

$$\|Au - y^*\|^2 + \alpha\|u\|^2$$

Constraint or regularization?

$$\|Au - y^*\|^2 + \alpha\Phi(u)$$

Constraint or regularization?

$$\|Au - y^*\|^2 + \alpha\Phi(u)$$

u smooth: $\Phi(u) = \|\nabla u\|^2$

Constraint or regularization?

$$\|Au - y^*\|^2 + \alpha\Phi(u)$$

u smooth: $\Phi(u) = \|\nabla u\|^2$

u sparse: $\Phi(u) = \sum |\langle \psi_k | u\rangle|^p,\ p < 2$

$\Phi(u) = \#\{|\langle \psi_k | u\rangle| \neq 0\}$

Constraint or regularization?

$$\|Au - y^*\|^2 + \alpha\Phi(u)$$

u smooth: $\Phi(u) = \|\nabla u\|^2$

u sparse: $\Phi(u) = \sum |\langle\psi_k|u\rangle|^p,\ p < 2$

$\Phi(u) = \#\{|\langle\psi_k|u\rangle| \neq 0\}$

Regularization:

Make the problem $Au = y^*$ well-posed.

Constraint or regularization?

$$\|Au - y^*\|^2 + \alpha\Phi(u)$$

u smooth: $\Phi(u) = \|\nabla u\|^2$

u sparse: $\Phi(u) = \sum |\langle\psi_k|u\rangle|^p,\ p < 2$

$\Phi(u) = \#\{|\langle\psi_k|u\rangle| \neq 0\}$

Regularization:

Make the problem $Au = y^*$ well-posed.

Constraint:

Solve $Au = y^*$ while u has further properties.

Tool from constraint optimization: the conditional gradient method

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

Tool from constraint optimization: the conditional gradient method

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

by

1. directional derivative v_n by

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

2. step size s_n by

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

3. update

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

Dunn, 1980

Tool from constraint optimization: the conditional gradient method

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

1. directional derivative v_n by
$$\min \langle F'(u_n) | v \rangle + \Phi(u)$$
2. step size s_n by
$$\min_{s \in [0,1]} (F + \Phi)(u_n + s(v_n - u_n))$$
3. update
$$u_{n+1} = u_n + s_n(v_n - u_n)$$

K. Bredies

The generalized conditional gradient method converges

Theorem

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

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

Then: convergence to a stationary point of $F + \Phi$.

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

Remark: F need not to be convex.

The generalized conditional gradient method is iterated shrinkage

Proposition

The minimization of

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

with the generalized conditional gradient method for

$$F(u) = \|Au - y^*\|^2 - \|u\|^2, \quad \Phi(u) = \|u\|^2 + \alpha \sum w_j |\langle u|\psi_j\rangle|^p$$

gives the iteration

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

where **S** is an elementwise shrinkage operation.

Remarks on iterated shrinkage

Remarks

- The iterated shrinkage is also motivated by the methods of surrogate functionals (Daubechies, Defries, DeMol, 2004).
- The above proposition also applies to nonlinear operators A:

$$\min \ \|A(u) - y^*\|^2 + \alpha \sum w_j |\langle u|\psi_j\rangle|^p$$

gives

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

Brief sketch of the surrogate method

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

Brief sketch of the surrogate method

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

Decouple A and $|\cdot|^p$

Brief sketch of the surrogate method

$$\|Au - y^*\|^2 + \alpha \sum w_j |\langle u|\psi_j\rangle|^p + \left(\|u - a\|^2 - \|Au - Aa\|^2\right) = J(u,a)$$

Decouple A and $|\cdot|^p$

Brief sketch of the surrogate method

$$\|Au - y^*\|^2 + \alpha \sum w_j |\langle u|\psi_j\rangle|^p + \left(\|u - a\|^2 - \|Au - Aa\|^2\right) = J(u, a)$$

Decouple A and $|\cdot|^p$
alternate minimization for u, a

Brief sketch of the surrogate method

$$\|Au - y^*\|^2 + \alpha \sum w_j |\langle u|\psi_j\rangle|^p + \left(\|u - a\|^2 - \|Au - Aa\|^2\right)$$
$$= J(u, a)$$

Decouple A and $|\cdot|^p$
alternate minimization for u, a

Theorem

$$u_{n+1} = \operatorname{argmin}_u J(u, u_n) = \mathbf{S}_{\alpha w_j, p}(u_n - A^*(Au_n - y^*))$$

converges strongly to a minimizer of $\|Au - y^*\|^2 + \alpha \sum w_j |\langle u|\psi_j\rangle|^p$.

Daubechies, Defries, De Mol, 2004.

Trying to get "extremely sparse"

$$\Phi(u) = \#\{\langle u|\psi_j\rangle \neq 0\}$$

Trying to get "extremely sparse"

$$\Phi(u) = \sum w_j |\langle u|\psi_j\rangle|^p$$

Trying to get "extremely sparse"

$$\Phi(u) = \sum w_j |\langle u|\psi_j\rangle|^p$$

For $p < 1$ the problem is non-convex.

Trying to get "extremely sparse"

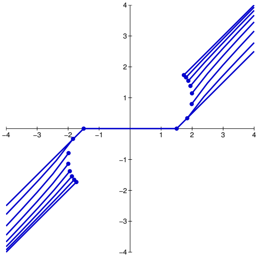
$$\Phi(u) = \sum w_j |\langle u|\psi_j\rangle|^p$$

For $p < 1$ the problem is non-convex.
Special case $A = I$:

$$\|u - y^*\|^2 + \alpha \sum w_j |\langle u|\psi_j\rangle|^p$$

has a "unique" minimizer.

$$\bar{u} = \mathbf{S}_{\alpha,w_j,p}(y^*)$$
$$= \sum S_{\alpha,w_j,p}(\langle y^*|\psi_j\rangle)\psi_j$$



Non-convex penalties in the surrogate method again gives shrinkage

Proposition ($p = 0$)

- The iteration
$$u_{n+1} = \operatorname{argmin}_u \|Au - y^*\|^2 + \alpha\#\{\langle u|\psi_j\rangle \neq 0\} + \big(\|u - u_n\|^2 - \|Au - Au_n\|^2\big)$$
gives the iterated hard shrinkage
$$u_{n+1} = \mathbf{S}_{\sqrt{2\alpha}}(u_n - A^*(Au_n - y^*)).$$
- If the sequence u_n converges, it converges to a minimizer of
$$\|Au - y^*\|^2 + \alpha\#\{\langle u|\psi_j\rangle \neq 0\}.$$

Outline

Sparse reconstruction under the presence of noise

(ψ_j): Two ONBs (Haar wavelet packet basis + spike basis)

$\bar{y}$: finite linear combination of ψ_j's

y^*: $\bar{y}$ + noise

$Hu = \sum u_j \psi_j$: Reconstruction operator.

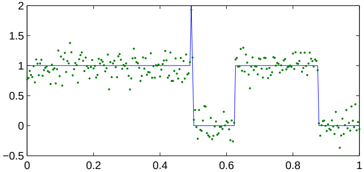
Try to recover the sparse representation of $\bar{y}$ from y^*.
minimize $\|Hu - y^*\|^2 + \alpha \sum |u_j|^p$

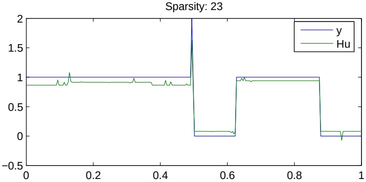
Sparse reconstruction under the presence of noise

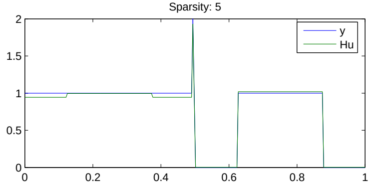
Try to recover the sparse representation of $\bar{y}$ from y^*.
minimize $\|Hu - y^*\|^2 + \alpha \sum |u_j|^p$

Sparse reconstruction under the presence of noise

Try to recover the sparse representation of $\bar{y}$ from y^*.
minimize $\|Hu - y^*\|^2 + \alpha \sum |u_j|^p$







Summary

- Optimization of the presentation of mammography images by optimal control of PDEs
- The generalized conditional gradient method converges even for non-convex functionals
- The method of surrogate functionals is a generalized conditional gradient descent
- Non-convex sparsity constraints seem to be handable by the method of surrogate functionals