Improving the quality of medical images

Wolfgang Stefan

Arizona State University

February 4, 2006
Physics of the Image Acquisition
  Positron Emission Tomography (PET)

Deblurring
  Forward Model
  Inverse Problem
  PET Examples
  Properties and Problems

Image Decomposition (u+v Decomposition)
  Introduction
  Brief History
  Application to Medical Images
Schema of a PET acquisition process

Annihilation

Coincidence Processing Unit

Sinogram/Listmode Data

Image Reconstruction
Example of a PET scan
Example of typical PET scan

Typical PET Images show
  ▶ High noise content
  ▶ High blurring
  ▶ Reconstruction artifacts
Signal degradation is modeled as a convolution

\[ g = f \ast h + n \]

- where \( g \) is the blurred signal
- \( f \) is the unknown signal
- \( h \) is the point spread function (PSF)
- \( n \) is noise
Forward Model Example

\[ g = f * h + n \]
Estimation of the Point Spread Function (PSF)

Estimations for the PSF come from:

- Phantom scans
Estimation of the Point Spread Function (PSF)

**Estimations** for the PSF come from:

- Phantom scans
- Rough estimation by a Gaussian
Estimation of the Point Spread Function (PSF)

**Estimations** for the PSF come from:

- Phantom scans
- Rough estimation by a Gaussian
- Blind Deconvolution
Inverse Problem

Find $f$ from $g = f * h + n$ given $g$ and $h$ with unknown $n$. 

Assuming normal distributed $n$ yields the estimator 

$$\hat{f} = \arg \min_f \{ \| g - f * h \|_2^2 \}$$

Reconstruction with $n$ normal distr. with $\sigma = 10^{-7}$
Inverse Problem

- Find $f$ from $g = f \ast h + n$ given $g$ and $h$ with unknown $n$.
- Assuming normal distributed $n$ yields the estimator

$$\hat{f} = \arg \min_f \{ \| g - f \ast h \|_2^2 \}$$
Inverse Problem

- Find $f$ from $g = f \ast h + n$ given $g$ and $h$ with unknown $n$.
- Assuming normal distributed $n$ yields the estimator

$$\hat{f} = \arg \min_f \{ \| g - f \ast h \|_2^2 \}$$

- Reconstruction with $n$ normal distr. with $\sigma = 10^{-7}$
Physiology of the Image Acquisition
Deblurring
Image Decomposition (u+v Decomposition)

Forward Model
Inverse Problem
PET Examples
Properties and Problems

(a)  

(b) 

(c) 

Wolfgang Stefan

Improving the quality of medical images
Regularization

- Add more information about the signal
Regularization

- Add more information about the signal
- e.g. statistical properties
Regularization

- Add more information about the signal
  - e.g. statistical properties
  - or information about the structure (e.g. sparse decon, or total variation decon)
Regularization

- Add more information about the signal
- e.g. statistical properties
- or information about the structure (e.g. sparse decon, or total variation decon)
- in latter case use a **penalty term**
Regularization

- Add more information about the signal
  - e.g. statistical properties
  - or information about the structure (e.g. sparse decon, or total variation decon)
- in latter case use a **penalty term**
- find
  \[
  \hat{f} = \arg\min_f \{ \|g - f \ast h\|_2^2 + \lambda R(f) \},
  \]
  
  where $R(f)$ is the penalty term and $\lambda$ is a penalty parameter.
Regularization Methods

- Common methods are Tikhonov (TK).

\[ R(f) = TK(f) = \int_{\Omega} |\nabla f(x)|^2 dx. \]
Regularization Methods

- Common methods are Tikhonov (TK).
  \[ R(f) = TK(f) = \int_{\Omega} |\nabla f(x)|^2 dx. \]

- Total Variation (TV)
  \[ R(f) = TV(f) = \int_{\Omega} |\nabla f(x)| dx. \]
Regularization Methods

- Common methods are Tikhonov (TK).
  \[ R(f) = TK(f) = \int_\Omega |\nabla f(x)|^2 dx. \]

- Total Variation (TV)
  \[ R(f) = TV(f) = \int_\Omega |\nabla f(x)| dx. \]

- Sparse deconvolution \((L^1)\)
  \[ R(f) = \|f\|_1 = \int_\Omega |f(x)| dx. \]
Improving the quality of medical images
Regularization Notes

\[ \hat{f} = \arg \min_f \{ \| g - f \ast h \|_2^2 + \lambda R(f) \} \]

- \( \lambda \) Governs the trade off between the fit to the data and the smoothness of the reconstruction and can be picked by the L-curve approach.
Regularization Notes

\[ \hat{f} = \arg \min_f \{ \| g - f \ast h \|_2^2 + \lambda R(f) \} \]

- \( \lambda \) governs the trade off between the fit to the data and the smoothness of the reconstruction and can be picked by the L-curve approach.
- TV yields a piece wise constant reconstruction and preserves the edges of the signal.
\[ \hat{f} = \arg \min_f \{ \| g - f * h \|_2^2 + \lambda R(f) \} \]

- \( \lambda \) governs the trade off between the fit to the data and the smoothness of the reconstruction and can be picked by the L-curve approach.
- TV yields a piecewise constant reconstruction and preserves the edges of the signal.
- TK yields a smooth reconstruction.
\[
\hat{f} = \arg\min_{f} \{\|g - f \ast h\|_2^2 + \lambda R(f)\}
\]

- \(\lambda\) Governs the trade off between the fit to the data and the smoothness of the reconstruction and can be picked by the L-curve approach.
- TV yields a piece wise constant reconstruction and preserves the edges of the signal.
- TK yields a smooth reconstruction.
\[ \hat{f} = \arg \min_{f} \{ \| g - f \ast h \|_2^2 + \lambda R(f) \} \]

- \( \lambda \) Governs the trade off between the fit to the data and the smoothness of the reconstruction and can be picked by the L-curve approach
- TV yields a piece wise constant reconstruction and preserves the edges of the signal
- TK yields a smooth reconstruction
- To find the minimum we use a limited memory BFGS method
Notes on the Optimization

- All the considered objective functions (OF) are convex
Notes on the Optimization

- All the considered objective functions (OF) are **convex**
- TK is a linear least squares (LS) problem

\[
\hat{f} = \arg \min_{f} \{ \| g - Hf \|^2_2 + \lambda \| Lf \|^2_2 \}
\]
Notes on the Optimization

- All the considered objective functions (OF) are convex
- TK is a linear least squares (LS) problem

\[
\hat{f} = \arg \min_f \{ \| g - Hf \|_2^2 + \lambda \| Lf \|_2^2 \} 
\]

- The TV objective function is non differentiable

\[
J(f) = \| g - Hf \|_2^2 + \lambda \| Lf \|_1 
\]
Notes on the Optimization

- All the considered objective functions (OF) are **convex**
- TK is a linear least squares (LS) problem
  \[ \hat{f} = \arg \min_f \{ \| g - Hf \|_2^2 + \lambda \| Lf \|_2^2 \} \]
- The TV objective function is **non differentiable**
  \[ J(f) = \| g - Hf \|_2^2 + \lambda \| Lf \|_1 \]
- The problems are **very large** (n order of 10000)
Notes on the Optimization

- All the considered objective functions (OF) are **convex**
- TK is a linear least squares (LS) problem
  \[
  \hat{f} = \arg \min_{f} \left\{ \| g - Hf \|_2^2 + \lambda \| Lf \|_2^2 \right\}
  \]
- The TV objective function is **non differentiable**
  \[
  J(f) = \| g - Hf \|_2^2 + \lambda \| Lf \|_1
  \]
- The problems are **very large** \((n\ \text{order of}\ 10000)\)
- **Evaluation** of the OF and its gradient is **cheap** \((\text{some FFTs and sparse matrix-vector multiplications})\)
Simulated PET

- Segmented data from an MRI scan is blurred using a Gaussian PSF.
Simulated PET

- Segmented data from an MRI scan is blurred using a **Gaussian PSF**.
- Simulated PET image also includes **Gauss distributed noise**.
Simulated PET

- Segmented data from an MRI scan is blurred using a Gaussian PSF
- Simulated PET image also includes Gauss distributed noise.

Note: The PSF is exactly known in this example, TV regularization
Real PET data

- Reconstruction done using Filtered Back Projection
- PSF estimated by a Gaussian
- TV regularization
Image improvement is possible even with a rough estimation of the PSF.
Image improvement is possible even with a rough estimation of the PSF.

Total Variation regularization (piecewise constant solution) is appropriate since the intensity levels depend on the tissue type.
Image improvement is possible even with a rough estimation of the PSF.

Total Variation regularization (piecewise constant solution) is appropriate since the intensity levels depend on the tissue type.

Improvement of these preliminary results when a better approximation of the PSF is available.
- Image improvement is possible even with a rough estimation of the PSF
- Total Variation regularization (piecewise constant solution) is appropriate since the intensity levels depend on the tissue type.
- Improvement of these preliminary results when a better approximation of the PSF is available
- Increased Artifacts and noise. (More post processing can improve this)
u+v decomposition

- decompose a signal

\[ f = u + v \]

such that \( u \) contains the wanted part i.e. the medical image and \( v \) the unwanted i.e. noise and artifacts.
u+v decomposition

- decompose a signal

\[ f = u + v \]

such that \( u \) contains the wanted part i.e. the medical image and \( v \) the unwanted i.e. noise and artifacts.

- The problem here is to identify \( u \) and \( v \) in appropriate Banach spaces.
u+v decomposition

- decompose a signal

\[ f = u + v \]

such that \( u \) contains the wanted part i.e. the medical image and \( v \) the unwanted i.e. noise and artifacts.

- The problem here is to identify \( u \) and \( v \) in appropriate Banach spaces.

- i.e. we seek Banach spaces that allow us to **measure** if a picture is the wanted picture or not.
Brief History

▶ Osher-Rudin Model 1992, 1994: Decompose image in a cartoon Part (piecewise constant) and a texture part.
Brief History

- Osher-Rudin Model 1992, 1994: Decompose image in a cartoon Part (piecewise constant) and a texture part.

- Osher and Rudin proposed $u \in BV$ i.e. the space of bounded Variation:

$$\|u\|_{BV} = \int_{\Omega} |\nabla u| \, dx < \infty,$$

where the derivative is to be understood in a distributional sense (weak).
Osher-Rudin Model 1992, 1994: Decompose image in a **cartoon** Part (piecewise constant) and a **texture** part.

Osher and Rudin proposed $u \in BV$ i.e. the space of bounded Variation:

$$\|u\|_{BV} = \int_{\Omega} |\nabla u| dx < \infty,$$

where the derivative is to be understood in a distributional sense (weak).

The residual $v = f - u$ was assumed to be in the Lebesgue space $L^2$. 

Wolfgang Stefan  
Improving the quality of medical images
Thus the solution to the decomposition problem is found by minimizing:

\[ E_{ROF}(u) = \| u \|_{BV} + \lambda \| f - u \|_{L^2}^2. \]

\( \lambda \) is a parameter to be chosen.


Leonid Rudin and Stanley Osher, Total variation based image restoration with free local constraints, Proceedings of the IEEE ICIP, Austin, USA 1 (1994), 31-35.
Thus the solution to the decomposition problem is found by minimizing:

$$E_{ROF}(u) = \|u\|_{BV} + \lambda \|f - u\|_{L^2}^2.$$ 

\(\lambda\) is a parameter to be chosen.
Thus the solution to the decomposition problem is found by minimizing:

\[ E_{ROF}(u) = \|u\|_{BV} + \lambda \|f - u\|_{L^2}^2. \]

- \( \lambda \) is a parameter to be chosen.


- Leonid Rudin and Stanley Osher, Total variation based image restoration with free local constraints, Proceedings of the IEEE ICIP, Austin, USA 1 (1994), 31-35.
Better Spaces for the Texture

- Different Spaces for the Texture have been proposed by Meyer
Different Spaces for the Texture have been proposed by Meyer.

And starting from there by Vese and Osher $v \in H^{-1}$,

resulting in the minimization of

$$E_{OSV}(u) = \|u\|_{BV} + \lambda \|f - u\|_{H^{-1}}^2$$
Better Spaces for the Texture

- Different Spaces for the Texture have been proposed by Meyer
- And starting from there by Vese and Osher \( v \in H^{-1} \),
- resulting in the minimization of

\[
E_{OSV}(u) = \|u\|_{BV} + \lambda \|f - u\|_{H^{-1}}^2
\]

- done usually by solving the PDE

\[
\dot{u} = -\nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) - 2\lambda \nabla^{-1}(f - u)
\]
Notes on the Vese and Osher Model

- Good separation of texture and Cartoon part
- $u$ contains the edges, though some edge information is lost to $v$ due to the loss of contrast.
- slow convergence i.e. very expensive
Barbara, typical $BV$, $H^{-1}$ Solution
Faster Algorithm using Wavelets

- Daubechies and Teschke:
- Replace $BV$ by the smaller Besov space $B_{1,1}^1$
Faster Algorithm using Wavelets

- Daubechies and Teschke:
- Replace \( BV \) by the smaller Besov space \( B_1^{1,1} \)
- Thus minimizing

\[
E(u, v) = \| f - (u + v) \|_{L^2} + 2\alpha \| u \|_{B_1^{1,1}} + \gamma \| v \|_{H^{-1}}
\]
Faster Algorithm using Wavelets

- Daubechies and Teschke:
- Replace $BV$ by the smaller Besov space $B_{1,1}^1$
- Thus minimizing

$$E(u, v) = \| f - (u + v) \|_{L^2} + 2\alpha \| u \|_{B_{1,1}^1} + \gamma \| v \|_{H^{-1}}$$

- $B_{1,1}^1$ and $H^{-1}$ have a Wavelet basis and an explicit solution exists (i.e. extremely fast in the order of a FFT)
Removal of Noise from Difference Images

**Application:** Two PET scans of the same patient at different times

**Question:** Are there any anatomical or functional changes?
Difference Image

- Scans from different days have to be aligned (Registration).
- The Registration is not perfect.
- Noise and artifacts change from scan to scan.
- Small changes are hard to locate in the difference image.
Difference Image

- Scans from different days have to be aligned (Registration).
- The Registration is not perfect.
- Noise and artifacts change from scan to scan.
- Small changes are hard to locate in the difference image.
  - try to enhance image by a $u + v$ decomposition
Decomposed Difference Image
Difference Image and u Part

Difference Image

u Part
u and v Part

\begin{align*}
\text{u Part} & \\
\text{v Part} & 
\end{align*}
Back to the deblurred PET scan
u+v decomposition of the deblurred PET scan

\[ u \in B^{1,1}_1 \text{ and } v \in H^{-1} \]
Changing the target space for $v$

- Recall: 2 Banach spaces: the wanted part $u \ (B_1^{1,1})$
- The noise and artifacts $v \ (H^{-1})$
Changing the target space for $v$

- Recall: 2 Banach spaces: the wanted part $u \ (B^{1,1}_1)$
- The noise and artifacts $v \ (H^{-1})$
- For the $v$ part: ”$v$ has to be in $L^2$ after taking one anti derivative”
- $v$ can be very ”wild” or ”irregular” and still have a small norm in $H^{-1}$
Changing the target space for $v$

- Recall: 2 Banach spaces: the wanted part $u (B_1^{1,1})$
- The noise and artifacts $v (H^{-1})$
- For the $v$ part: "$v$ has to be in $L^2$ after taking one anti derivative"
- $v$ can be very "wild" or "irregular" and still have a small norm in $H^{-1}$
- If we want more "regular" features to end up in $v$ we can change the target space for $v$
Changing the target space for $v$

- Recall: 2 Banach spaces: the wanted part $u$ ($B_1^{1,1}$)
- The noise and artifacts $v$ ($H^{-1}$)
- For the $v$ part: "$v$ has to be in $L^2$ after taking one anti derivative"
- $v$ can be very "wild" or "irregular" and still have a small norm in $H^{-1}$
- If we want more "regular" features to end up in $v$ we can change the target space for $v$
- e.g. from $H^{-1}$ to $H^{-0.01}$
$u+v$ decomposition of the deblurred PET scan

$u \in B_{1,1}^{1,1}$ and $v \in H^{-0.01}$
$u + v$ decomposition of the deblurred PET scan

$u \in B_1^{1,1}$ and $v \in H^{-1}$
Acknowledgment

- Thanks to my Advisor Rosie Renaut
- Haewon Nam and Kewe Chen for the data
- Svetlana Roudenko for discussion