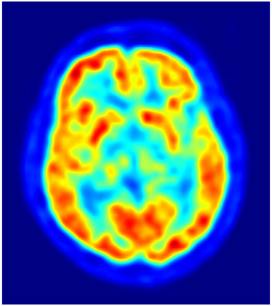
PET provides useful metabolic information



Widely used in

- diagnosis
- monitoring
- follow-up of the patient

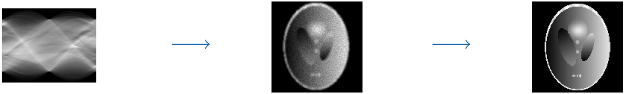
Accurate delineation of tumours is a challenging task

PET images suffer from two main limitations:

- low spatial resolution
- high level of multiplicative noise

We can act on two levels:
During the reconstruction phase:

During a post-processing phase:



Contribution of the work

Choice: inverse problem approach in a post-processing phase

Assumption: the noise after reconstruction is still multiplicative (Poisson in first approximation)

The method will

- be specific to the restoration of PET images;
- introduce the total generalized variation (TGV) as a regularization term;
- take into account the Poisson statistics of the noise.

Post-Reconstruction Deconvolution of PET Images by Total Generalized Variation Regularization

Stéphanie Guérit

10th March 2015

Outline

Observations are a corrupted version of the reality

Forward model

Let

- $\boldsymbol{u}_0 \in \mathbb{R}^N$ be the original PET image;
- $\boldsymbol{K}$ be a blur operator;
- $\mathcal{N}_\nu$ be a noise corruption of parameter ν.

Observations $\boldsymbol{z}$ are associated with $\boldsymbol{u}_0$ through

$$\boldsymbol{z} = \mathcal{N}_\nu(\boldsymbol{K}(\boldsymbol{u}_0)).$$

Observations are a corrupted version of the reality

Forward model

Let

- $\boldsymbol{u}_0 \in \mathbb{R}^N$ be the original PET image;
- $\boldsymbol{K}$ be a blur operator;
- $\mathcal{N}_\nu$ be a noise corruption of parameter ν.

Observations $\boldsymbol{z}$ are associated with $\boldsymbol{u}_0$ through

$$\boldsymbol{z} = \mathcal{N}_\nu(\boldsymbol{K}(\boldsymbol{u}_0)).$$

Assumptions:

- PSF near the center of the field of view is uniform, Gaussian and isotropic;
- Noise follows a Poissonian distribution.

Observations are a corrupted version of the reality

Forward model (with assumptions)

Let

- $\boldsymbol{u}_0 \in \mathbb{R}^N$ be the original PET image;
- $\boldsymbol{K}$ be a *linear* and *bounded* blur operator $\in \mathbb{R}^{N \times N}$;
- $\mathcal{P}$ be a *Poisson noise* with mean $(\boldsymbol{K}\boldsymbol{u}_0)$.

Observations $\boldsymbol{z}$ are associated with $\boldsymbol{u}_0$ through

$$\boldsymbol{z} = \mathcal{P}(\boldsymbol{K}\boldsymbol{u}_0),$$

Observations are a corrupted version of the reality

Forward model (with assumptions)

Let

- $\boldsymbol{u}_0 \in \mathbb{R}^N$ be the original PET image;
- $\boldsymbol{K}$ be a *linear* and *bounded* blur operator $\in \mathbb{R}^{N \times N}$;
- $\mathcal{P}$ be a *Poisson noise* with mean $(\boldsymbol{K}\boldsymbol{u}_0)$.

Observations $\boldsymbol{z}$ are associated with $\boldsymbol{u}_0$ through

$$\boldsymbol{z} = \mathcal{P}(\boldsymbol{K}\boldsymbol{u}_0),$$

Two additional constraints specific to PET imaging:

Observations are a corrupted version of the reality

Forward model (with assumptions)

Let

- $\boldsymbol{u}_0 \in \mathbb{R}^N$ be the original PET image;
- $\boldsymbol{K}$ be a *linear* and *bounded* blur operator $\in \mathbb{R}^{N \times N}$;
- $\mathcal{P}$ be a *Poisson noise* with mean $(\boldsymbol{K}\boldsymbol{u}_0)$.

Observations $\boldsymbol{z}$ are associated with $\boldsymbol{u}_0$ through

$$\boldsymbol{z} = \mathcal{P}(\boldsymbol{K}\boldsymbol{u}_0),$$

Two additional constraints specific to PET imaging:

- Positivity: $\boldsymbol{u}_0 \succeq \mathbf{0}$

Observations are a corrupted version of the reality

Forward model (with assumptions)

Let

- $\boldsymbol{u}_0 \in \mathbb{R}^N$ be the original PET image;
- $\boldsymbol{K}$ be a *linear* and *bounded* blur operator $\in \mathbb{R}^{N \times N}$;
- $\mathcal{P}$ be a *Poisson noise* with mean $(\boldsymbol{K}\boldsymbol{u}_0)$.

Observations $\boldsymbol{z}$ are associated with $\boldsymbol{u}_0$ through

$$\boldsymbol{z} = \mathcal{P}(\boldsymbol{K}\boldsymbol{u}_0),$$

Two additional constraints specific to PET imaging:

- Positivity: $\boldsymbol{u}_0 \succeq \mathbf{0}$
- Photometry invariance: $\sum_{i=1}^{N} (\boldsymbol{u}_0)_i \approx \sum_{i=1}^{N} z_i$

Outline

The problem of finding an estimator of $\boldsymbol{u}_0$ is ill-posed

Our goal is to find an estimator $\hat{\boldsymbol{u}}_0$ of the original image $\boldsymbol{u}_0$ from observations $\boldsymbol{z}$.

This problem is ill-posed due to

- the noise;
- the low-pass filter effect of the PSF;
- etc.

$\Rightarrow$ need for a regularization

The assumption of piecewise smoothness by TGV

Let $\boldsymbol{x} \in \mathbb{R}^N$.
Total variation used to promote piecewise constant images:

$$\mathsf{TV}(\boldsymbol{x}) = \|\nabla \boldsymbol{x}\|_{2,1}$$

Total generalized variation of order 2 (Bredies 2010):

$$\mathsf{TGV}_\alpha^2(\boldsymbol{x}) = \min_{\boldsymbol{w} \in \mathbb{R}^{N \times 2}} \|\nabla \boldsymbol{x} - \boldsymbol{w}\|_{2,1} + \alpha \|\varepsilon(\boldsymbol{w})\|_{2,1},$$

with $\alpha \in \mathbb{R}$.

The assumption of piecewise smoothness by TGV

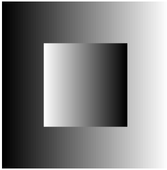
Let $\boldsymbol{x} \in \mathbb{R}^N$.
Total variation used to promote piecewise constant images:

$$\mathrm{TV}(\boldsymbol{x}) = \|\nabla \boldsymbol{x}\|_{2,1}$$

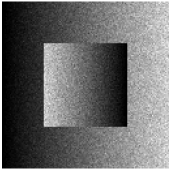
Total generalized variation of order 2 (Bredies 2010):

$$\mathrm{TGV}^2_\alpha(\boldsymbol{x}) = \min_{\boldsymbol{w} \in \mathbb{R}^{\boldsymbol{N} \times \mathbf{2}}} \|\nabla \boldsymbol{x} - \boldsymbol{w}\|_{2,1} + \alpha \|\varepsilon(\boldsymbol{w})\|_{2,1},$$

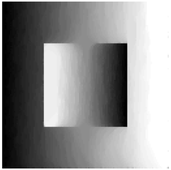
with $\alpha \in \mathbb{R}$.



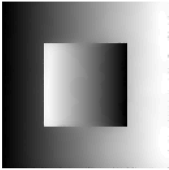
Original



Noisy



TV



TGV

The maximum *a posteriori* estimator is the best one

Bayesian approach

$$
\begin{aligned}
\hat{\boldsymbol{u}}_0 &= \arg\max_{\boldsymbol{u}\in\mathbb{R}^{\boldsymbol{N}}} p(\boldsymbol{u}|\boldsymbol{z}) \\
&= \arg\max_{\boldsymbol{u}\in\mathbb{R}^{\boldsymbol{N}}} p(\boldsymbol{z}|\boldsymbol{u})p(\boldsymbol{u}) \quad \text{(Bayes)} \\
&= \arg\min_{\boldsymbol{u}\in\mathbb{R}^{\boldsymbol{N}}} \; -\log p(\boldsymbol{z}|\boldsymbol{u}) \; -\log p(\boldsymbol{u})
\end{aligned}
$$

- Fidelity term → $-\log p(\boldsymbol{z}|\boldsymbol{u})$
- Regularization term → $-\log p(\boldsymbol{u})$

Data fidelity term for Poisson noise (Anthoine 2012):

$$-\log p(\boldsymbol{z}|\boldsymbol{u}) = \sum_{i=1}^{N} (\boldsymbol{K}\boldsymbol{u} - \boldsymbol{z}\cdot f(\boldsymbol{K}\boldsymbol{u}))_i + r(\boldsymbol{z}),$$

Constraints and prior are integrated to the MAP estimator

From

$$\hat{\boldsymbol{u}}_0 = \underset{\boldsymbol{u}\in\mathbb{R}^{\boldsymbol{N}}}{\arg\min} \; -\log p(\boldsymbol{z}|\boldsymbol{u}) - \log p(\boldsymbol{u}),$$

we get the inverse problem formulation.

Inverse problem

$$\hat{\boldsymbol{u}}_\lambda = \underset{\substack{\boldsymbol{u}\in\mathbb{R}^{\boldsymbol{N}},\\ \boldsymbol{w}\in\mathbb{R}^{\boldsymbol{N}\times 2}}}{\arg\min} \; \lambda \sum_{i=1}^{N} (\boldsymbol{K}\boldsymbol{u} - \boldsymbol{z}\cdot f(\boldsymbol{K}\boldsymbol{u}))_i + \|\nabla \boldsymbol{u} - \boldsymbol{w}\|_{2,1} + \alpha\|\varepsilon(\boldsymbol{w})\|_{2,1} + \imath_{\mathbb{R}_+^{\boldsymbol{N}}}(\boldsymbol{u}) + \imath_{\mathcal{C}}(\boldsymbol{u}),$$

with $f(t) = \log t$ if $t > 0$ and 0 otherwise and $\lambda > 0$ is the regularization parameter. Indicator functions are defined as

$$\imath_{\mathcal{D}}(\boldsymbol{x}) = \begin{cases} 0 & \text{if } \boldsymbol{x}\in\mathcal{D} \\ +\infty & \text{otherwise} \end{cases}$$

Outline

The Chambolle-Pock primal-dual algorithm

Chambolle-Pock (CP) primal-dual algorithm (Chambolle 2010)

Let $\boldsymbol{L} : \mathcal{X} \to \mathcal{Y}$ be a linear continuous operator with a norm defined by $\|\boldsymbol{L}\|_2 = \max\{\|\boldsymbol{L}\boldsymbol{x}\|_2 \mid \boldsymbol{x} \in \mathcal{X} \text{ with } \|\boldsymbol{x}\|_2 \leq 1\}$ and $G : \mathcal{X} \to [0, +\infty[$ and $F^\star : \mathcal{Y} \to [0, +\infty[$ be proper, convex, and lower-semicontinuous functions. The CP primal-dual algorithm is designed to solve the following saddle-point problem

$$\min_{\boldsymbol{x} \in \mathcal{X}} \max_{\boldsymbol{y} \in \mathcal{Y}} \quad \langle \boldsymbol{L}\boldsymbol{x}, \boldsymbol{y} \rangle - F^\star(\boldsymbol{y}) + G(\boldsymbol{x}),$$

which is a primal-dual formulation of the primal problem $\min_{\boldsymbol{x} \in \mathcal{X}} \; F(\boldsymbol{L}\boldsymbol{x}) + G(\boldsymbol{x})$.

The choice of CP algorithm is adapted to the inverse formulation

Inverse problem

$$\hat{\boldsymbol{u}}_\lambda = \underset{\substack{\boldsymbol{u}\in\mathbb{R}^{\boldsymbol{N}},\\ \boldsymbol{w}\in\mathbb{R}^{\boldsymbol{N}\times\boldsymbol{2}}}}{\arg\min}\ \underbrace{\lambda \sum_{i=1}^{N} (\boldsymbol{K}\boldsymbol{u} - \boldsymbol{z}\cdot f(\boldsymbol{K}\boldsymbol{u}))_i}_{F_1(\boldsymbol{K}\boldsymbol{u})} + \underbrace{\|\nabla \boldsymbol{u} - \boldsymbol{w}\|_{2,1}}_{F_2(\nabla\boldsymbol{u}-\boldsymbol{w})}$$

$$+ \underbrace{\alpha\|\varepsilon(\boldsymbol{w})\|_{2,1}}_{F_3(\varepsilon(\boldsymbol{w}))} + \underbrace{\imath_{\mathbb{R}_+^{\boldsymbol{N}}}(\boldsymbol{u}) + \imath_{\mathcal{C}}(\boldsymbol{u})}_{G(\boldsymbol{u})}$$

CP algorithm is a proximal algorithm

Inverse problem

$$\hat{\boldsymbol{u}}_\lambda = \underset{\substack{\boldsymbol{u}\in\mathbb{R}^{\boldsymbol{N}},\\ \boldsymbol{w}\in\mathbb{R}^{\boldsymbol{N}\times 2}}}{\arg\min}\ \underbrace{\lambda \sum_{i=1}^{N} (\boldsymbol{K}\boldsymbol{u} - \boldsymbol{z}\cdot f(\boldsymbol{K}\boldsymbol{u}))_i}_{F_1(\boldsymbol{K}\boldsymbol{u})} + \underbrace{\|\nabla\boldsymbol{u} - \boldsymbol{w}\|_{2,1}}_{F_2(\nabla\boldsymbol{u}-\boldsymbol{w})}$$

$$+ \underbrace{\alpha\|\varepsilon(\boldsymbol{w})\|_{2,1}}_{F_3(\varepsilon(\boldsymbol{w}))} + \underbrace{\imath_{\mathbb{R}_+^{\boldsymbol{N}}}(\boldsymbol{u}) + \imath_{\mathcal{C}}(\boldsymbol{u})}_{G(\boldsymbol{u})}$$

The proximal operator of $\varphi : \mathcal{X} \to \mathcal{X}$ evaluated in $\tilde{\boldsymbol{x}} \in \mathcal{X}$ is unique and defined as (Parikh 2013)

$$\mathrm{prox}_\varphi(\tilde{\boldsymbol{x}}) := \underset{\boldsymbol{x}\in\mathcal{X}}{\arg\min}\ \tfrac{1}{2}\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_2^2 + \varphi(\boldsymbol{x}).$$

In our case, proximal operators of $F_1, F_1^\star, F_2, F_2^\star, F_3, F_3^\star$ and G are easy to compute (closed form).

CP algorithm applied to PET images deconvolution

Algorithm 1 for TGV denoising and deblurring of PET images.

initialization: $n = 0$; $\boldsymbol{u}^{(0)} = \bar{\boldsymbol{u}}^{(0)} = \boldsymbol{y} \in \mathbb{R}^N$; $\boldsymbol{w}^{(0)} = \bar{\boldsymbol{w}}^{(0)} = 0 \in \mathbb{R}^{N\times 2}$; $\boldsymbol{p}^{(0)} = 0 \in \mathbb{R}^N$; $\boldsymbol{q}^{(0)} = 0 \in \mathbb{R}^{N\times 2}$; $\boldsymbol{r}^{(0)} = 0 \in \mathbb{R}^{N\times 4}$; choose $\tau^{(0)} = \sigma^{(0)} = 0.9/\|\boldsymbol{L}\|_2$.

repeat

$\quad \boldsymbol{p}^{(n+1)} = \mathrm{prox}_{\sigma^{(n)} F_1^\star}\left(\boldsymbol{p}^{(n)} + \sigma^{(n)} \boldsymbol{K} \bar{\boldsymbol{u}}^{(n)}\right)$

$\quad \boldsymbol{q}^{(n+1)} = \mathrm{prox}_{\sigma^{(n)} F_2^\star}\left(\boldsymbol{q}^{(n)} + \sigma^{(n)} (\nabla \bar{\boldsymbol{u}}^{(n)} - \bar{\boldsymbol{w}}^{(n)})\right)$

$\quad \boldsymbol{r}^{(n+1)} = \mathrm{prox}_{\sigma^{(n)} F_3^\star}\left(\boldsymbol{r}^{(n)} + \sigma^{(n)} \varepsilon(\bar{\boldsymbol{w}}^{(n)})\right)$

$\quad \boldsymbol{u}^{(n+1)} = \mathrm{prox}_{\tau^{(n)} G}\left(\boldsymbol{u}^{(n)} + \tau^{(n)} (\mathrm{div}[\boldsymbol{q}^{(n)}] - \boldsymbol{K}^* \boldsymbol{p}^{(n)})\right)$

$\quad \boldsymbol{w}^{(n+1)} = \boldsymbol{w}^{(n)} - \tau^{(n)} (-\boldsymbol{q}^{(n)} - \mathrm{div}[\boldsymbol{r}^{(n)}])$

$\quad \bar{\boldsymbol{u}}^{(n+1)} = 2\boldsymbol{u}^{(n+1)} - \boldsymbol{u}^{(n)}$

$\quad \bar{\boldsymbol{w}}^{(n+1)} = 2\boldsymbol{w}^{(n+1)} - \boldsymbol{w}^{(n)}$

until convergence of $\boldsymbol{u}$

Outline

The discrepancy Principle is one strategy to select λ

The Morozov's discrepancy principle for Poisson noise (Bertero 2009):

$$\mathrm{KL}(\boldsymbol{z}, \boldsymbol{K}\hat{\boldsymbol{u}}_\lambda) \approx \frac{N}{2},$$

where

$$\mathrm{KL}(\boldsymbol{z}, \boldsymbol{K}\hat{\boldsymbol{u}}_\lambda) = \sum_{i=1}^{N} (\boldsymbol{K}\hat{\boldsymbol{u}}_\lambda - \boldsymbol{z} + \boldsymbol{z} \cdot f(\boldsymbol{z}) - \boldsymbol{z} \cdot f(\boldsymbol{K}\hat{\boldsymbol{u}}_\lambda))_i$$

is the Kullback-Leibler difference.

This principle can be adapted to images with nul background

The Morozov's discrepancy principle for Poisson noise in images with low intensity and no background (Carlavan 2011):

$$\mathrm{KL}(\boldsymbol{z}, \boldsymbol{K}\hat{\boldsymbol{u}}_\lambda) \approx \frac{M}{2},$$

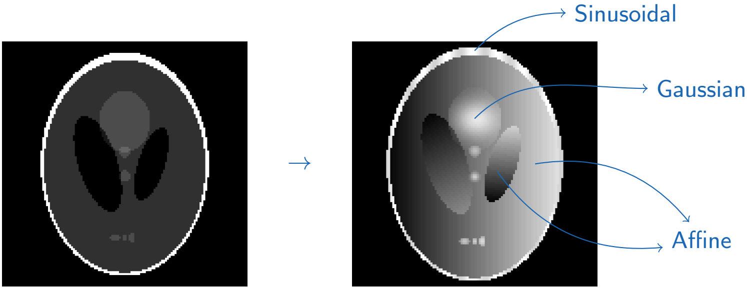
where $M \leq N$ is the number of non zero pixels.

Updating rule for tuning parameter λ at iteration l

$$\lambda^{(l+1)} = \lambda^{(l)} \frac{\mathrm{KL}(\boldsymbol{z}, \boldsymbol{K}\hat{\boldsymbol{u}}_{\lambda^{(l)}})}{M/2}$$

Outline

A modified Shepp-Logan phantom without constant areas



Original phantom → Modified phantom

Image properties:

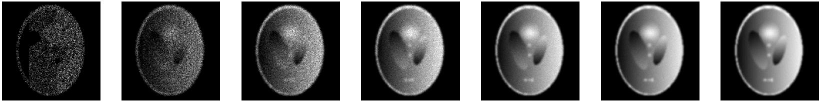
- Size: 128×128 pixels;
- Pixel values: $[0, 255]$;
- Null background;

Experiments realized for 13 different levels of noise

Synthetic data model

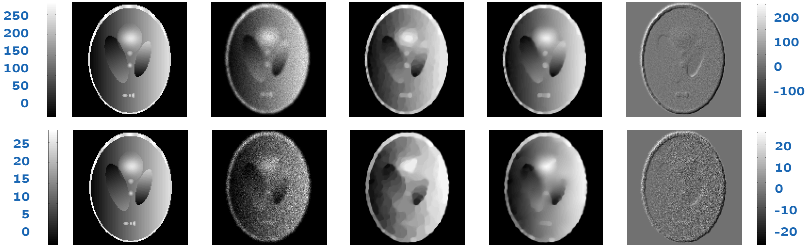
$$\boldsymbol{z}_\beta = \mathcal{P}(\ \beta\ \ \boldsymbol{K}\ \boldsymbol{u}_0)$$

- Influence of acquisition time
- Gaussian kernel with $\sigma = 1.17$ pixel



$\beta = 0.01$ $\beta = 0.05$ $\beta = 0.2$ $\beta = 1$ $\beta = 5$ $\beta = 20$ $\beta = 100$

TGV restoration is visually good, without staircasing artifact



Original

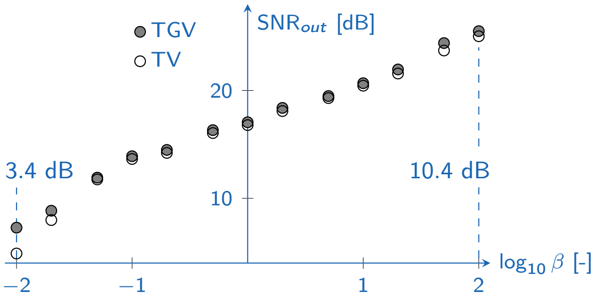
Corrupted ($\beta_{\mathbf{top}} = 1$, $\beta_{\mathbf{bot}} = 0.1$)

TV restoration

TGV restoration

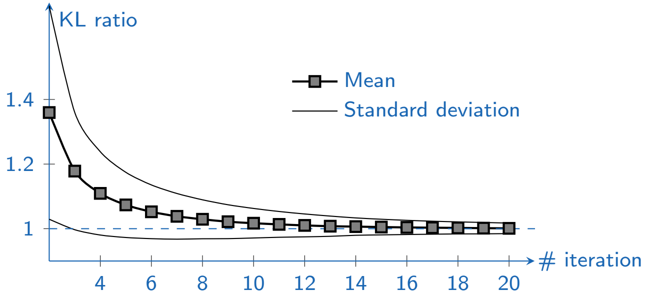
TGV residual

In average, SNR_{out} of TGV is 0.5 dB above TV's



$$\text{SNR}_{out}(\hat{\boldsymbol{u}}_\beta, \boldsymbol{u}_0) = 20\log_{10}\left(\frac{\|\hat{\boldsymbol{u}}_\beta\|_2}{\|\hat{\boldsymbol{u}}_\beta - \boldsymbol{u}_0\|_2}\right)$$

The criterion for automatic selection of λ_{opt} converges to 1



$$\lambda^{(l+1)} = \lambda^{(l)} \underbrace{\frac{\mathrm{KL}(\boldsymbol{z}, \boldsymbol{K}\hat{\boldsymbol{u}}_\lambda)}{M/2}}_{\text{KL ratio}}, \quad l \in \mathbb{N}$$

PET images of patients with tumour in the head-neck region

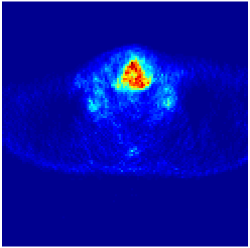
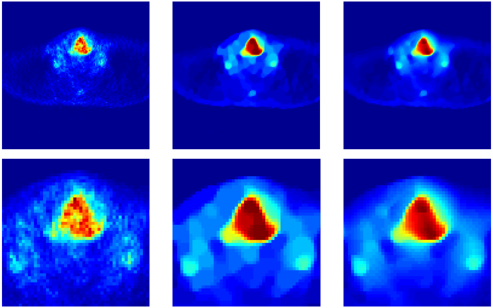


Image properties:

- Size: 128×128 pixels with a resolution of 2.2×2.2 mm^2;
- PSF measured experimentally: Gaussian and isotropic with 6 mm of FWHM, in first approximation.

The updating rule for the choice of λ_{opt} does not work...

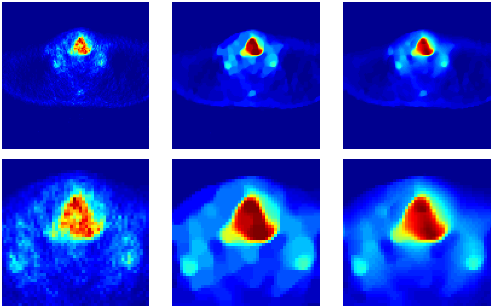


$\lambda = 5$

The assumption of pure Poisson noise is not verified:

$$\boldsymbol{z} = \gamma\, \mathcal{P}(\beta \boldsymbol{K} \boldsymbol{u}_0), \quad \gamma \in \mathbb{R}_0^+$$

The updating rule for the choice of λ_{opt} does not work...



$\lambda = 5$

The assumption of pure Poisson noise is not verified:

$$\boldsymbol{z} = \gamma\, \mathcal{P}(\beta \boldsymbol{K} \boldsymbol{u}_0), \quad \gamma \in \mathbb{R}_0^+$$

Is it possible to find an optimal value for λ? Is it useful?

Outline

What is next?

Some good things...

- TGV is more adapted to real medical data;
- the updating rule of λ is validated for synthetic data.

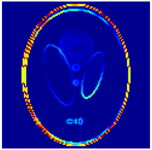
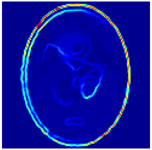
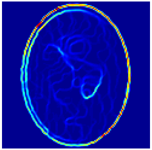
What is next?

Some good things...

- TGV is more adapted to real medical data;
- the updating rule of λ is validated for synthetic data.

... but challenges remain!

- a deeper understanding of the nature of the noise;
- an updating rule more flexible to noises of different natures;
- a connection between restoration and segmentation.

Post-Reconstruction Deconvolution of PET Images by Total Generalized Variation Regularization

Stéphanie Guérit

10th March 2015