Methods for joint image reconstruction from multiple modalities

Simon R. Arridge[1]

[1]Department of Computer Science, University College London, UK

Synergistic Reconstruction Symposium,
Chester, 3rd-6th November2019

Outline

Outline

Introduction

Methods for joint image reconstruction from multiple modalities

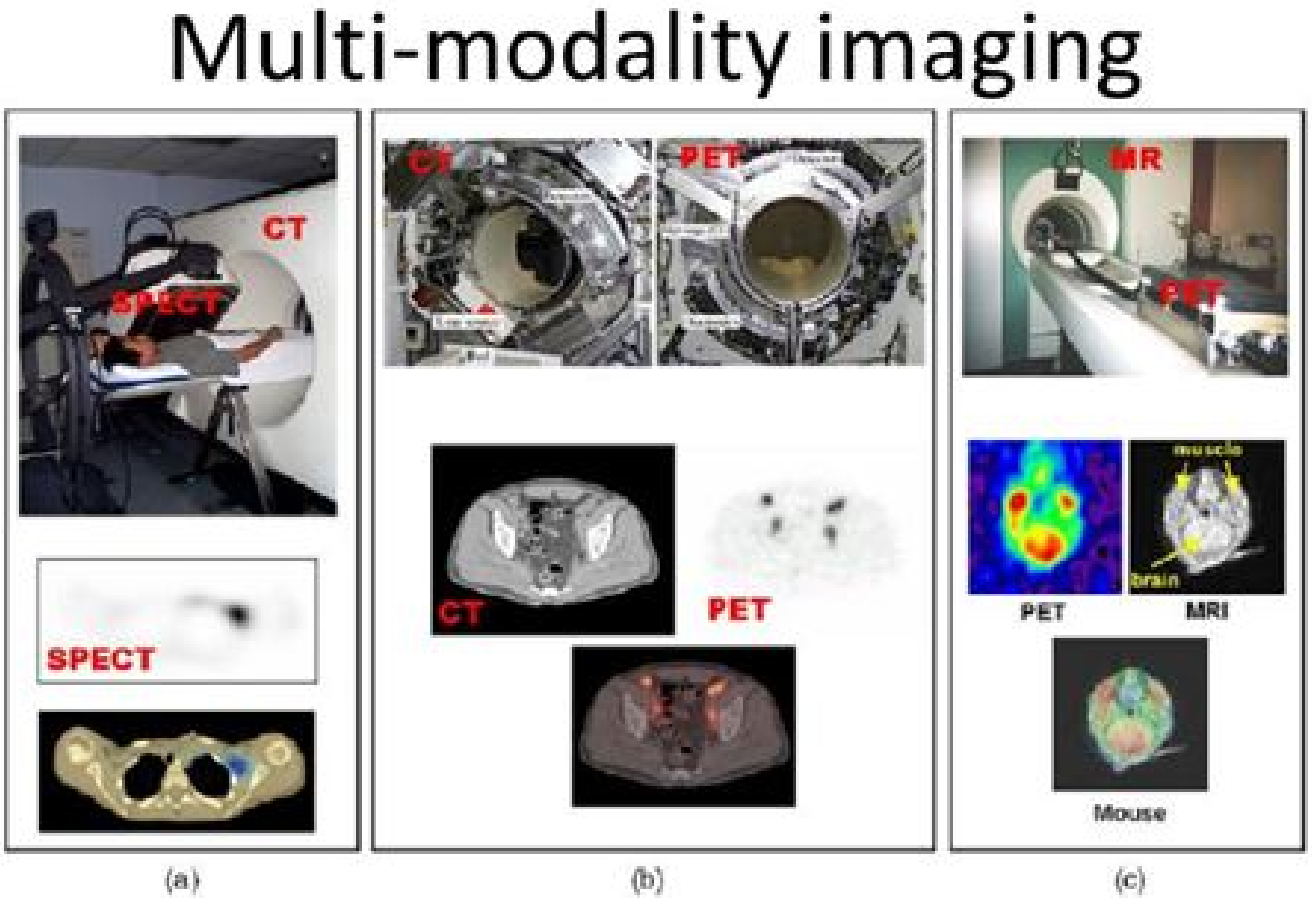


Figure 1. The first multimodality prototypes developed in the mid 1990s: (a) a SPECT/CT scanner (photos courtesy of Bruce Hasegawa, PhD, UCSF), (b) a PET/CT scanner evaluated clinically at the University of Pittsburgh and (c) an MR/PET scanner for small animals (photos courtesy of Simon Cherry, PhD, UC Davis).

Townsend DW, *Phys Med Biol* **53**:R1-R39, 2008

Introduction

examples

- PET-MR (Two modalities, linear)
  e.g. [M. Ehrhardt et al 2015], [J. Rasch, E. M. Brinkmann, and M. Burger (2018), Diedda et al 2019]
- US and PAT (Two modalities, nonlinear)
  e.g. [T. P. Matthews and M. A. Anastasio, 2017]
- multispectral CT (Multiple channels, linear/nonlinear)
  e.g. [S. Niu et al. 2018], [D. Kazantsev et al. 2018], [J. F. P. J. Abascal, N. Ducros, and F. Peyrin 2018], [Y . Hu et al. 2019],
- Multiparametric MR, MR fingerprinting (multiple channels, non-linear)
  e.g. [S.Tang et al 2018], [R. Botnar 2019]
- dynamic fMRI with structural MRI (two modalities, multiple channels, linear)
  e.g. [J. Rasch et al. , 2018]
- GeoPhysics (multiple modalities, linear/nonlinear)
  e.g. [Haber and Holtzman-Gazit, 2013], [D. Rovetta and D. Colombo, 2018]

Introduction

Challenges

- multiple likelihoods - how to scale them ?
- multiple priors - how to scale them ?
- one known image as prior for others ?
- alternating/cycling through priors ?
- joint all at once ?

Modality	Forward Mapping	Data fit (negative log likelihood)
1	$A^{(u)} : u \mapsto g_u$	$\underbrace{\ell(g_u, A^{(u)}(u))}_{-\log P(g_u \vert u)}$
2	$A^{(v)} : v \mapsto g_v$	$\underbrace{\ell(g_v, A^{(v)}(v))}_{-\log P(g_v \vert v)}$
combined	$\mathbf{A} : \mathbf{f} \mapsto \mathbf{g}$	$\underbrace{\ell(\mathbf{g}, \mathbf{A}(\mathbf{f}))}_{-\log P(g_u, g_v \vert u, v)}$

where

$$\mathbf{A} \equiv \begin{pmatrix} A^{(u)} \\ A^{(v)} \end{pmatrix} \quad \mathbf{f} \equiv \begin{pmatrix} u \\ v \end{pmatrix} \quad \mathbf{g} \equiv \begin{pmatrix} g_u \\ g_v \end{pmatrix}$$

Each modalitity maybe independently : linear/non-linear, well-posed/weakly illposed/strongly illposed

Joint Maximum *a posteriori* (JMAP) estimate

$$\min_{u,v} \left[-\log P(u, v|g_u, g_v) := \ell(g_u, A^{(u)}(u)) + \ell(g_v, A^{(v)}(v)) + \Psi(u, v) \right]$$

Compare to "one-sided" reconstruction with fixed auxiliary image

$$\min_{u} \left[-\log P(u|g_u, v) := \ell(g_u, A^{(u)}(u)) + \Psi(u|v) \right]$$

Descent methods

$$\mathbf{f}^{(n+1)} = \mathbf{f}^{(n)} + \alpha \mathbf{h}^{(n)}$$

where $\mathbf{h}^{(n)}$ is a descent direction, and α is a step length.
E.g steepest descent

$$\mathbf{h}^{(n)} = -\left(\frac{\partial \ell}{\partial \mathbf{f}} + \frac{\partial \Psi}{\partial \mathbf{f}}\right)$$

or Quasi-Newton methods

$$\mathbf{h}^{(n)} \rightarrow \tilde{\mathbf{h}}^{(n)} = Q\mathbf{h}^{(n)}$$

with Q an approximation to the Hessian of the negative log posterior.

Rather than classical optimisation methods, it is possible to split into two steps:

Rather than classical optimisation methods, it is possible to split into two steps:

$$\text{likelihood update step} \quad \mathbf{f}^{(n+1/2)} = \mathbf{f}^{(n)} - \alpha Q \frac{\partial \ell}{\partial \mathbf{f}}$$

Rather than classical optimisation methods, it is possible to split into two steps:

$$\text{likelihood update step} \quad \mathbf{f}^{(n+1/2)} = \mathbf{f}^{(n)} - \alpha Q \frac{\partial \ell}{\partial \mathbf{f}}$$

$$\text{“Image Processing” step} \quad \mathbf{f}^{(n+1)} = \min_{\mathbf{f}} \Phi(\mathbf{f}^{(n)}, \mathbf{f}^{(n+1/2)}, \mathbf{g})$$

Rather than classical optimisation methods, it is possible to split into two steps:

$$\begin{aligned} \text{likelihood update step} \quad \mathbf{f}^{(n+1/2)} &= \mathbf{f}^{(n)} - \alpha Q \frac{\partial \ell}{\partial \mathbf{f}} \\ \text{“Image Processing” step} \quad \mathbf{f}^{(n+1)} &= \min_{\mathbf{f}} \Phi(\mathbf{f}^{(n)}, \mathbf{f}^{(n+1/2)}, \mathbf{g}) \end{aligned}$$

The second step may be evolution of a PDE induced by the form of the regularisation functional, [Scherzer et.al. “Variational Methods in Imaging” 2008].

$$\frac{\partial \mathbf{f}}{\partial t} = \mathcal{L}\mathbf{f}$$

but, crucially, *may be an image enhancement step not explicitly derived as the variation of a function.*

Rather than classical optimisation methods, it is possible to split into two steps:

$$\begin{aligned} \text{likelihood update step} \quad \mathbf{f}^{(n+1/2)} &= \mathbf{f}^{(n)} - \alpha Q \frac{\partial \ell}{\partial \mathbf{f}} \\ \text{“Image Processing” step} \quad \mathbf{f}^{(n+1)} &= \min_{\mathbf{f}} \Phi(\mathbf{f}^{(n)}, \mathbf{f}^{(n+1/2)}, \mathbf{g}) \end{aligned}$$

The second step may be evolution of a PDE induced by the form of the regularisation functional, [Scherzer et.al. “Variational Methods in Imaging” 2008].

$$\frac{\partial \mathbf{f}}{\partial t} = \mathcal{L}\mathbf{f}$$

but, crucially, *may be an image enhancement step not explicitly derived as the variation of a function.*

Remark : the possible disadvantage of using image flows that are not the variation of a functional is that they lose the Bayesian interpetation of the maximisation of a posterior.

- The Markov Random Field idea :

$$\Psi(u) = \frac{1}{p}\sum_i \sum_{j\in\mathcal{N}(i)} w_{ij}|u_i - u_j|^p$$

$$\frac{\partial \Psi}{\partial u_i} = \sum_{j\in\mathcal{N}(i)} w_{ij}|u_i - u_j|^{p-1}$$

Note : $p = 2 \rightarrow \Psi(u) = \langle u, Lu\rangle$ with $L_{ij} = \begin{cases} w_{ij} & j \neq i \\ -\sum_{j\neq i} w_{ij} & i = j \end{cases}$
E.g. 4-connected neighbourhood, $w_{ij} = 1 \Rightarrow L \simeq -\nabla^2$
Extend to global image : L becomes *Graph Laplacian*
In between : local (weighted) neighbourhood/patch-based approach.

- Basic idea (one-sided prior) : design local weightings w_{ij} to reflect expected influence of v on u

- Sparsity idea : Under some transform

$$\xi = \mathbb{E}f \quad \leftrightarrow \quad f = \mathbb{D}\xi$$

  the transformed parameters have many/mostly zero or close to zero values.
- The sparsity can be described by $|\xi|_0$ which simply counts the non-zero components
- since optimisation of L_0 norm leads to non-convex problem, it is replaced by "convex relaxation" L_1-norm.
- Transform possibilities include finite differences (gradient regularisation), wavelets, generalised kernels, neural nets (encoder/decoder).
- Typical inverse problem synthesis approach) :

$$\mathcal{D}(g, A\mathbb{D}(\xi)) + \alpha|\xi|_1$$

- Joint recon extension : *joint sparsity*.
- Alternatively : *Low-Rank plus sparse*.

Outline

Some Variational Regulariation Methods

Scalar Valued Case

Regularisation functionals have first variation (Euler-Lagrange equation) that defines a PDE for image flow

$$\Psi(u) := \int_\Omega \psi(|\nabla u|)\mathrm{d}x \quad \rightarrow \quad \Psi'(u) = \nabla \cdot \underbrace{\left(\frac{\psi'(|\nabla u|)}{|\nabla u|}\right)}_{\kappa} \nabla u$$

	$\psi(\lvert\nabla u\rvert)$	κ
TV	$\lvert\nabla u\rvert$	$\frac{1}{\lvert\nabla u\rvert}$
Smoothed TV	$T\sqrt{\lvert\nabla u\rvert^2 + T^2} - T$	$\frac{T}{\sqrt{\lvert\nabla u\rvert^2+T^2}}$
Perona-Malik (1)	$\frac{T^2}{2}\log\left[1 + \frac{\lvert\nabla u\rvert^2}{T^2}\right]$	$\frac{T^2}{T^2+\lvert\nabla u\rvert^2}$
Perona-Malik (2)	$\frac{T^2}{2}\left[1 - \exp\left(-\frac{\lvert\nabla u\rvert^2}{T^2}\right)\right]$	$\exp\left(-\frac{\lvert\nabla u\rvert^2}{T^2}\right)$
Huber	$\begin{cases} T\lvert\nabla u\rvert - \frac{T^2}{2} & \lvert\nabla u\rvert > T \\ \frac{\lvert\nabla u\rvert^2}{2} & \lvert\nabla u\rvert \le T \end{cases}$	$\begin{cases} \frac{T}{\sqrt{\lvert\nabla u\rvert^2+T^2}} & \lvert\nabla u\rvert > T \\ 1 & \lvert\nabla u\rvert \le T \end{cases}$
Tukey	$\begin{cases} \frac{T^2}{6}\left[1 - \left(1 - \frac{\lvert\nabla u\rvert^2}{T}\right)^3\right] & \lvert\nabla u\rvert < T \\ \frac{T^2}{6} & \lvert\nabla u\rvert \ge T \end{cases}$	$\begin{cases} \left[1 - \frac{\lvert\nabla u\rvert^2}{T}\right]^2 & \lvert\nabla u\rvert < T \\ 0 & \lvert\nabla u\rvert \ge T \end{cases}$

We can define a flow without it being the variation of a functional form Weikert proposed using a tensor

$$\frac{\partial u}{\partial t} = \nabla \cdot (\mathrm{D}(\mathrm{J}_\rho(\nabla u_\sigma))\nabla u$$

where the (symmetric, positive semi-definite) structure tensor is constructed as

$$\mathrm{J}_\rho(\nabla u_\sigma)(\mathrm{x}) = G_\rho * (\nabla u_\sigma \nabla u_\sigma^T) = \begin{bmatrix} G_\rho * u_x^2 & G_\rho * u_y u_x \\ G_\rho * u_x u_y & G_\rho * u_y^2 \end{bmatrix},$$

with eigensystem $\{\eta_k, \mathbf{v}_k\}$ and

$$\mathrm{D} = [\mathbf{v}_1\ \mathbf{v}_2] \begin{bmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^\mathrm{T} \\ \mathbf{v}_2^\mathrm{T} \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{bmatrix}.$$

Edge Enhancing Diffusion (EED) approach, reduces smoothing in the normal direction by the (Perona-Malik) diffusivity function:

$$\gamma_1 = \kappa_\epsilon(|\nabla u|) = \exp\left(-|\nabla u|^2/\epsilon^2\right); \gamma_2 = 1$$

Tensor based anisotropic diffusion filtering

Scalar Valued Case

The diffusion tensor $\mathbb{D}$ rotates and scales the flux in order to adapt for the underlying geometrical configurations of the image

- G_σ, G_ρ are Gaussian smoothing kernels and $u_\sigma = G_\sigma * u$
- σ is a *differentiation scale* introduced to avoid false edge detections due to noise, and introduce a non-local generalization in the estimation of the dominant directions of ∇u
- ρ is an *integration* scale which controls the size of the neighbourhood $O(\rho)$.
- The eigenvalues η_k convey the level of contrast along the given eigenvectors $\mathbf{v}_k$ within the neighbourhood of size $O(\rho)$.
- Note that:

$$\lim_{\rho\to 0} \mathbf{v}_1 = \hat{\nu} = \frac{\nabla u_\sigma}{|\nabla u_\sigma|} \quad ; \quad \lim_{\rho\to 0} \mathbf{v}_2 = \hat{\tau} = \frac{\nabla u_\sigma^\perp}{|\nabla u_\sigma|}$$

- The threshold parameter ϵ controls the strength of diffusion.

Outline

Methods Based on Joint Probability

Information Theoretic Regularisation

Joint entropy $H(u, v)$ is a measure of randomness characterizing the *joint system* of two random variables u and v. It is given by:

$$H(u, v) = -\int_{\Omega(u)} \int_{\Omega(v)} P(u, v) \log(P(u, v) \mathrm{d}x_u \, \mathrm{d}x_v$$

where $P(u, v)$ is the joint probability density function of the joint system defined by r.v. u and v, denoting the probability of a joint event

Methods Based on Joint Probability

Information Theoretic Regularisation

Joint entropy $H(u, v)$ is a measure of randomness characterizing the *joint system* of two random variables u and v. It is given by:

$$H(u, v) = -\int_{\Omega(u)} \int_{\Omega(v)} P(u, v) \log(P(u, v) \mathrm{d}x_u \, \mathrm{d}x_v$$

where $P(u, v)$ is the joint probability density function of the joint system defined by r.v. u and v, denoting the probability of a joint event
Joint Entropy and Mutual Information are routinely used in image registration.

Methods Based on Joint Probability

Information Theoretic Regularisation

Joint entropy $H(u, v)$ is a measure of randomness characterizing the *joint system* of two random variables u and v. It is given by:

$$H(u, v) = -\int_{\Omega(u)} \int_{\Omega(v)} P(u, v) \log(P(u, v) \mathrm{d}x_u \, \mathrm{d}x_v$$

where $P(u, v)$ is the joint probability density function of the joint system defined by r.v. u and v, denoting the probability of a joint event
Joint Entropy and Mutual Information are routinely used in image registration.
Putting

$$\Psi(u, v) = H(u, v)$$

imposes a regularisation scheme that minimises the joint entropy between u and v.

Methods Based on Joint Probability

Information Theoretic Regularisation

Joint entropy $H(u, v)$ is a measure of randomness characterizing the *joint system* of two random variables u and v. It is given by:

$$H(u, v) = -\int_{\Omega(u)} \int_{\Omega(v)} P(u, v) \log(P(u, v) \mathrm{d}x_u \, \mathrm{d}x_v$$

where $P(u, v)$ is the joint probability density function of the joint system defined by r.v. u and v, denoting the probability of a joint event
Joint Entropy and Mutual Information are routinely used in image registration.
Putting

$$\Psi(u, v) = H(u, v)$$

imposes a regularisation scheme that minimises the joint entropy between u and v.
Differentiation of $H(u, v)$ can be made computationally, efficient using Parzen kernel density estimators to develop a continuous function based on the sample pixels in u, v [Panagiotou, ..., A., 2009]

Methods Based on Joint Probability

Joint Entropy Regularisation

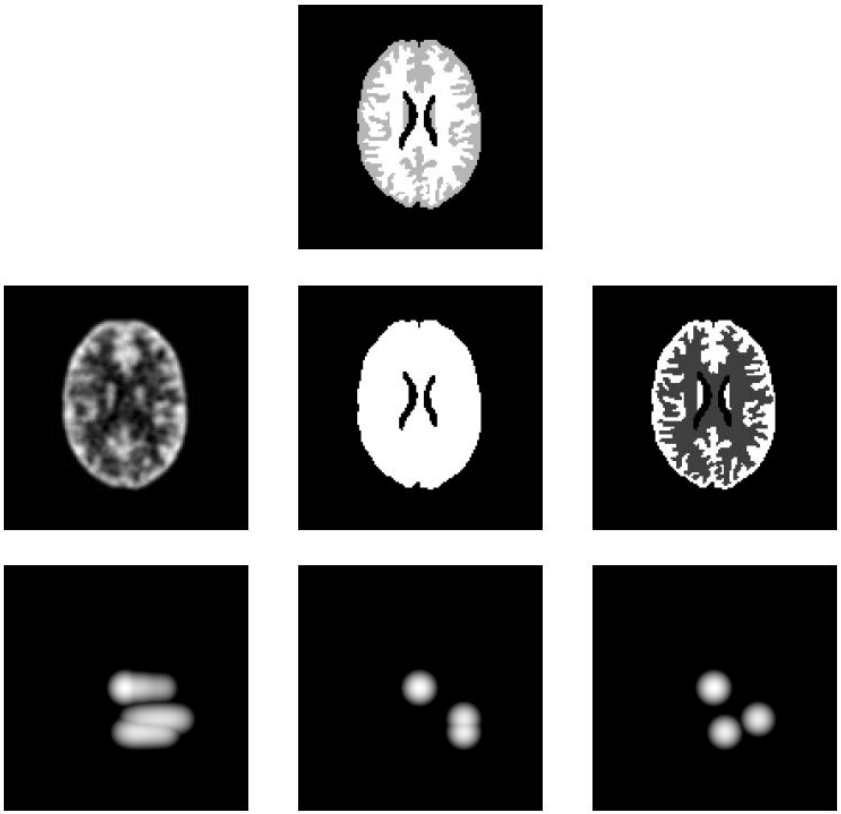


Fig. 1. Log of estimated joint probability distribution functions: Top: Reference image, Middle: Example recomstructed images (L to R): Blurred image, image with same activity in two distinct anatomical regions, and image with identical structure as reference image. Bottom: Estimated joint pdf of the images in the middle row and anatomical image.

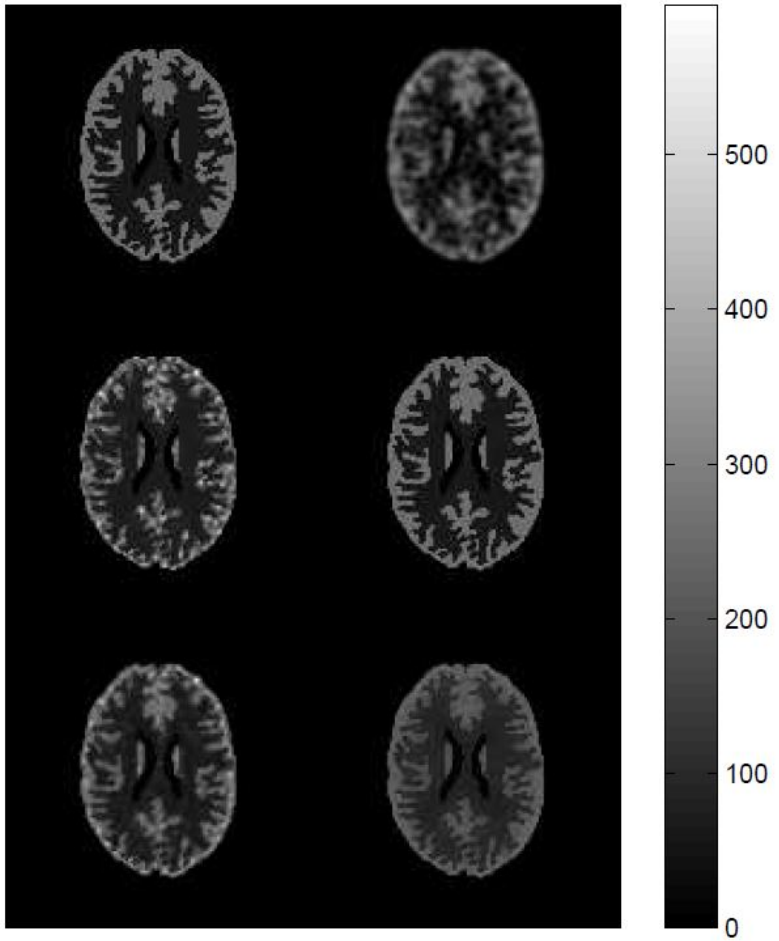
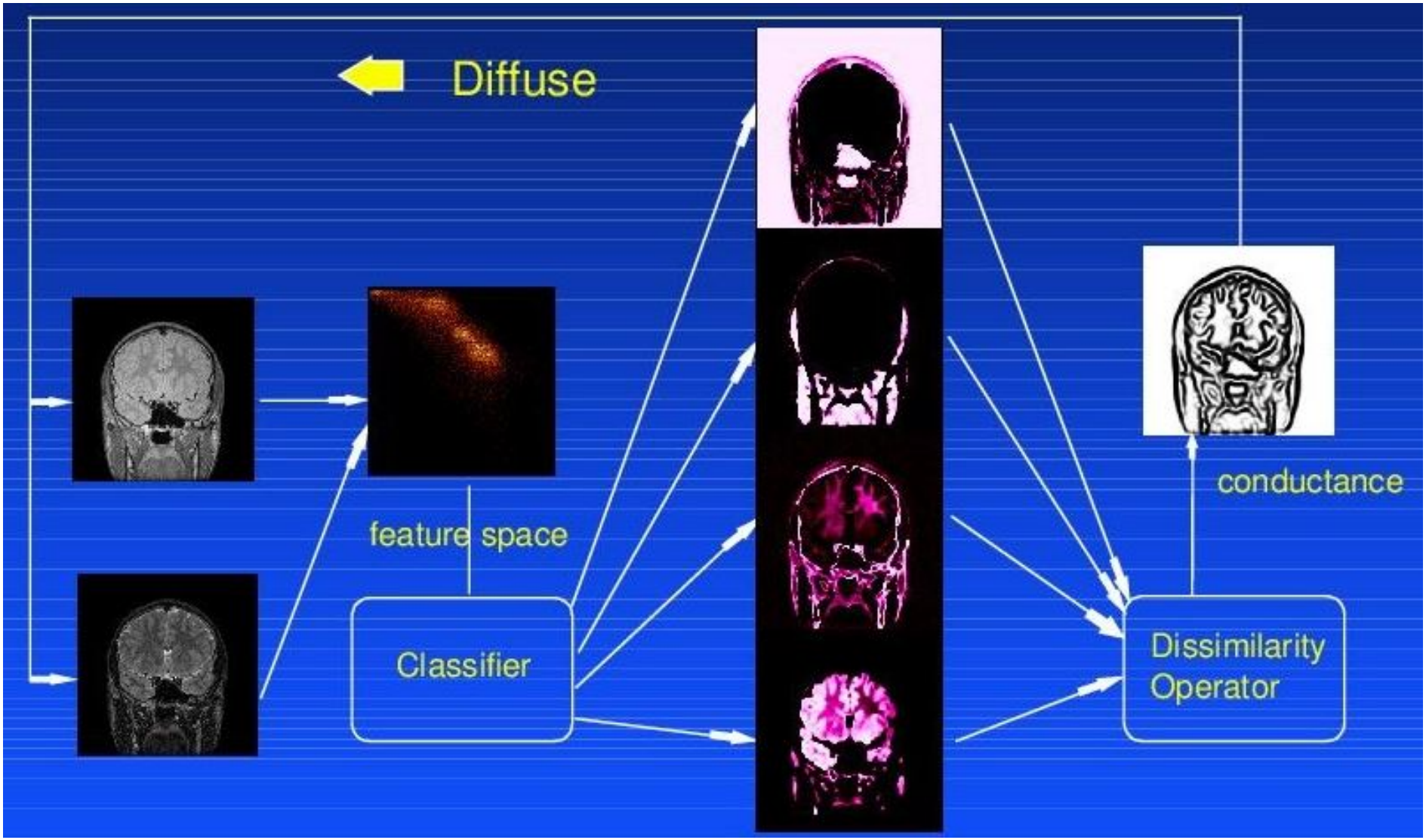


Fig. 4. True and reconstructed images: Top: True PET image (left), QP reconstruction with $\beta = 0.5$ (right). Middle: MI reconstruction for $\mu = 10000$ (left), JE reconstruction for $\mu = 5000$ (right). Bottom: MI-scale reconstruction for $\mu = 2500$ (left), JE-scale reconstruction for $\mu = 2500$.

Methods Based on Joint Probability

Multispectral Probabilistic Diffusion



Outline

Incorporating structural Information

Local Orientation Tensor Field

[Kaipio, Kolehmainen, Vauhkonen, Somersalo, IP, 1999] :

$$\Psi(u) = \int_{\Omega} |\nabla u|^2_{B(v)} \mathrm{d}x = \int_{\Omega} |L^{\mathrm{T}} \nabla u|^2 \mathrm{d}x = \int_{\Omega} \tilde{\nabla} \cdot \Lambda \tilde{\nabla} u \mathrm{d}x$$

$B(v)$ is a fixed tensor field, incorporating structure from another image v.

Incorporating structural Information

Local Orientation Tensor Field

[Kaipio, Kolehmainen, Vauhkonen, Somersalo, IP, 1999] :

$$\Psi(u) = \int_{\Omega} |\nabla u|^2_{B(v)} \mathrm{d}x = \int_{\Omega} |L^{\mathrm{T}} \nabla u|^2 \mathrm{d}x = \int_{\Omega} \tilde{\nabla} \cdot \Lambda \tilde{\nabla} u \mathrm{d}x$$

$B(v)$ is a fixed tensor field, incorporating structure from another image v. E.g.

$$B = I - (1-\gamma)\hat{\nu}\hat{\nu}^{\mathrm{T}} = \hat{\tau}\hat{\tau}^{\mathrm{T}} + \gamma\hat{\nu}\hat{\nu}^{\mathrm{T}} = R\Lambda R^{\mathrm{T}} = LL^{\mathrm{T}}$$

Incorporating structural Information

Local Orientation Tensor Field

[Kaipio, Kolehmainen, Vauhkonen, Somersalo, IP, 1999] :

$$\Psi(u) = \int_\Omega |\nabla u|^2_{B(v)} \mathrm{d}x = \int_\Omega |L^{\mathrm{T}} \nabla u|^2 \mathrm{d}x = \int_\Omega \tilde{\nabla} \cdot \Lambda \tilde{\nabla} u \mathrm{d}x$$

$B(v)$ is a fixed tensor field, incorporating structure from another image v. E.g.

$$B = I - (1-\gamma)\hat{\nu}\hat{\nu}^{\mathrm{T}} = \hat{\tau}\hat{\tau}^{\mathrm{T}} + \gamma\hat{\nu}\hat{\nu}^{\mathrm{T}} = R\Lambda R^{\mathrm{T}} = LL^{\mathrm{T}}$$

- $R = [\hat{\nu}\,\hat{\tau}] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is a rotation matrix,

Incorporating structural Information

Local Orientation Tensor Field

[Kaipio, Kolehmainen, Vauhkonen, Somersalo, IP, 1999] :

$$\Psi(u) = \int_\Omega |\nabla u|^2_{B(v)} \mathrm{d}x = \int_\Omega |L^{\mathrm{T}} \nabla u|^2 \mathrm{d}x = \int_\Omega \tilde{\nabla} \cdot \Lambda \tilde{\nabla} u \mathrm{d}x$$

$B(v)$ is a fixed tensor field, incorporating structure from another image v. E.g.

$$B = I - (1-\gamma)\hat{\nu}\hat{\nu}^{\mathrm{T}} = \hat{\tau}\hat{\tau}^{\mathrm{T}} + \gamma\hat{\nu}\hat{\nu}^{\mathrm{T}} = R\Lambda R^{\mathrm{T}} = LL^{\mathrm{T}}$$

- $R = [\hat{\nu}\,\hat{\tau}] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is a rotation matrix,
- $\Lambda = \begin{pmatrix} \gamma & 0 \\ 0 & 1 \end{pmatrix}$ is an anisotropy matrix, $L = R\Lambda^{1/2}$

Incorporating structural Information

Local Orientation Tensor Field

[Kaipio, Kolehmainen, Vauhkonen, Somersalo, IP, 1999] :

$$\Psi(u) = \int_\Omega |\nabla u|^2_{B(v)} \mathrm{d}x = \int_\Omega |L^{\mathrm{T}} \nabla u|^2 \mathrm{d}x = \int_\Omega \tilde{\nabla} \cdot \Lambda \tilde{\nabla} u \mathrm{d}x$$

$B(v)$ is a fixed tensor field, incorporating structure from another image v. E.g.

$$B = I - (1-\gamma)\hat{\nu}\hat{\nu}^{\mathrm{T}} = \hat{\tau}\hat{\tau}^{\mathrm{T}} + \gamma\hat{\nu}\hat{\nu}^{\mathrm{T}} = R\Lambda R^{\mathrm{T}} = LL^{\mathrm{T}}$$

- $R = [\hat{\nu}\,\hat{\tau}] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is a rotation matrix,
- $\Lambda = \begin{pmatrix} \gamma & 0 \\ 0 & 1 \end{pmatrix}$ is an anisotropy matrix, $L = R\Lambda^{1/2}$
- $0 \leq \gamma \leq 1$ is an edge indicator. E.g. $\gamma = \exp\left(-\frac{|\nabla v|^2}{T^2}\right)$

Incorporating structural Information

Local Orientation Tensor Field

[Kaipio, Kolehmainen, Vauhkonen, Somersalo, IP, 1999] :

$$\Psi(u) = \int_\Omega |\nabla u|^2_{B(v)} \mathrm{d}x = \int_\Omega |L^{\mathrm{T}} \nabla u|^2 \mathrm{d}x = \int_\Omega \tilde{\nabla} \cdot \Lambda \tilde{\nabla} u \mathrm{d}x$$

$B(v)$ is a fixed tensor field, incorporating structure from another image v. E.g. $B = I - (1-\gamma)\hat{\nu}\hat{\nu}^{\mathrm{T}} = \hat{\tau}\hat{\tau}^{\mathrm{T}} + \gamma\hat{\nu}\hat{\nu}^{\mathrm{T}} = R\Lambda R^{\mathrm{T}} = LL^{\mathrm{T}}$

- $R = [\hat{\nu}\,\hat{\tau}] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is a rotation matrix,
- $\Lambda = \begin{pmatrix} \gamma & 0 \\ 0 & 1 \end{pmatrix}$ is an anisotropy matrix, $L = R\Lambda^{1/2}$
- $0 \leq \gamma \leq 1$ is an edge indicator. E.g. $\gamma = \exp\left(-\frac{|\nabla v|^2}{T^2}\right)$
- $\tilde{\nabla} = R^{\mathrm{T}}\nabla = \begin{pmatrix} \frac{\partial}{\partial\hat{\nu}} \\ \frac{\partial}{\partial\hat{\tau}} \end{pmatrix}$ is gradient in local "gauge" coordinates

Incorporating structural Information

Local Orientation Tensor Field

[Kaipio, Kolehmainen, Vauhkonen, Somersalo, IP, 1999] :

$$\Psi(u) = \int_\Omega |\nabla u|^2_{B(v)} \mathrm{d}x = \int_\Omega |L^{\mathrm{T}} \nabla u|^2 \mathrm{d}x = \int_\Omega \tilde{\nabla} \cdot \Lambda \tilde{\nabla} u \mathrm{d}x$$

$B(v)$ is a fixed tensor field, incorporating structure from another image v. E.g.

$$B = I - (1-\gamma)\hat{\nu}\hat{\nu}^{\mathrm{T}} = \hat{\tau}\hat{\tau}^{\mathrm{T}} + \gamma\hat{\nu}\hat{\nu}^{\mathrm{T}} = R\Lambda R^{\mathrm{T}} = LL^{\mathrm{T}}$$

- $R = [\hat{\nu}\,\hat{\tau}] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is a rotation matrix,
- $\Lambda = \begin{pmatrix} \gamma & 0 \\ 0 & 1 \end{pmatrix}$ is an anisotropy matrix, $L = R\Lambda^{1/2}$
- $0 \leq \gamma \leq 1$ is an edge indicator. E.g. $\gamma = \exp\left(-\frac{|\nabla v|^2}{T^2}\right)$
- $\tilde{\nabla} = R^{\mathrm{T}}\nabla = \begin{pmatrix} \frac{\partial}{\partial\hat{\nu}} \\ \frac{\partial}{\partial\hat{\tau}} \end{pmatrix}$ is gradient in local “gauge” coordinates

Diffusion flow is

$$\frac{\partial u}{\partial t} = \nabla \cdot R\Lambda R^{\mathrm{T}} \nabla u$$

Weighted norm

$$\Psi(u) := \int_\Omega w(v)\psi(|\nabla u|)\mathrm{d}x\,,$$

$$\rightarrow \quad \Psi'(u) = \nabla\cdot\left(w(v)\frac{\psi'(|\nabla u|)}{|\nabla u|}\right)\nabla u = \nabla\cdot w(v)\kappa(u)\nabla u$$

Incorporating structural Information

Locally weighted norm

Weighted norm

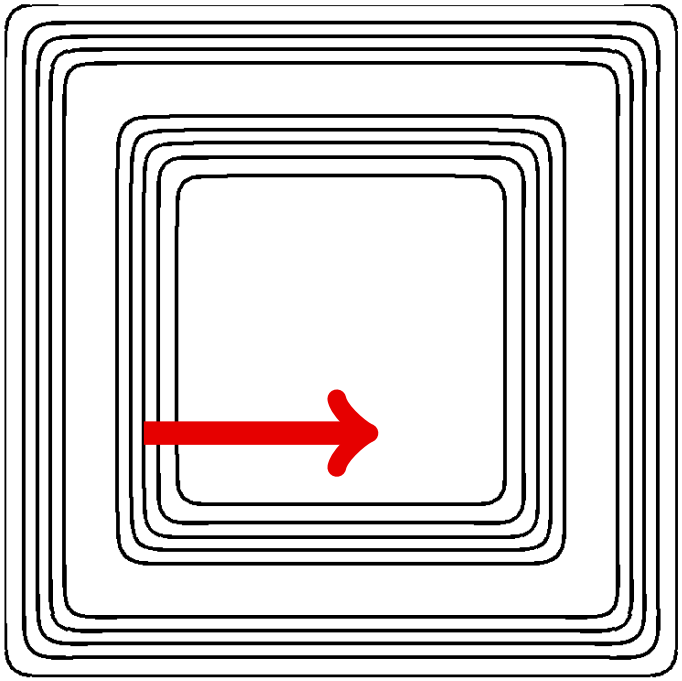
$$\begin{aligned} \Psi(u) &:= \int_\Omega w(v)\psi(|\nabla u|)\mathrm{d}x\,, \\ \rightarrow \quad \Psi'(u) &= \nabla\cdot\left(w(v)\frac{\psi'(|\nabla u|)}{|\nabla u|}\right)\nabla u = \nabla\cdot w(v)\kappa(u)\nabla u \end{aligned}$$

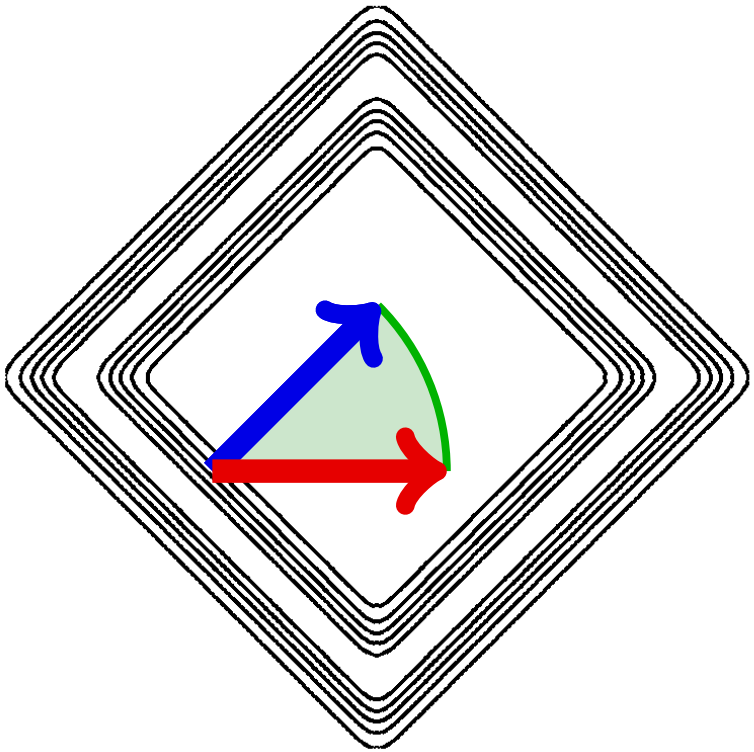
Generalising

$$\begin{aligned} \Psi(u) &:= \int_\Omega \psi(|L^{\mathrm{T}}\nabla u|)\mathrm{d}x\,, \\ \rightarrow \quad \Psi'(u) &= \nabla\cdot L\left(\frac{\psi'(|L^{\mathrm{T}}\nabla u|)}{|L^{\mathrm{T}}\nabla u|}\right)L^{\mathrm{T}}\nabla u \\ &= \tilde{\nabla}\cdot\underbrace{\Lambda^{1/2}\left(\frac{\psi'(|\nabla u|_B)}{|\nabla u|_B}\right)\Lambda^{1/2}}_{K}\tilde{\nabla}u \end{aligned}$$

Outline

Similar Structures by Level Sets





$$\langle \nabla u, \nabla v \rangle = \cos(\theta) \, \|\nabla u\| \, \|\nabla v\|$$

Measure of Parallelism II

Recall $$d(\nabla u, \nabla v) = \varphi\Big(\psi\left(\|\nabla u\|\,\|\nabla v\|\right) - \psi\left(|\langle \nabla u, \nabla v\rangle|\right)\Big).$$

Parallel Level Set Prior [2]

$$\mathcal{S}(u, v) = \int_\Omega \varphi\Big(\psi\left(\|\nabla u\|\,\|\nabla v\|\right) - \psi\left(|\langle \nabla u, \nabla v\rangle|\right)\Big)$$

Special case [3,4]: $\varphi(x) = x$ and $\psi(x) = x^2$

[2] Ehrhardt and A., IEEE Trans Im Proc, 2014
[3] Gallardo and Meju, Geophysical Research Letters, 2003
[4] Haber and Holtzman-Gazit, Surveys in Geophysics, 2013

Generated Diffusion

Diffusivity of Parallel Level Sets [1]

Let $R[u]$ and $R[v]$ be local coordinates for u and v. Then the derivative of $\mathcal{S}_\beta$ can be written as

$$D\mathcal{S}_\beta[u] = -\operatorname{div}\left(R[v]K(u,v)\Lambda[u,v]R[v]^T\nabla u\right)$$

$$D\mathcal{S}_\beta[v] = -\operatorname{div}\left(R[u]K(v,u)\Lambda[v,u]R[u]^T\nabla v\right)$$

with $\Lambda[u,v] = \operatorname{diag}(\gamma^{\perp}(u,v),\gamma^{\parallel}(u,v),\ldots,\gamma^{\parallel}(u,v))$.

[1] Ehrhardt ...A., Inverse Problems, 2015.

Generated Diffusion

Diffusivity of Parallel Level Sets [1]

Let $R[u]$ and $R[v]$ be local coordinates for u and v. Then the derivative of $\mathcal{S}_\beta$ can be written as

$$D\mathcal{S}_\beta[u] = -\operatorname{div}\left(R[v]K(u,v)\Lambda[u,v]R[v]^T\nabla u\right)$$

$$D\mathcal{S}_\beta[v] = -\operatorname{div}\left(R[u]K(v,u)\Lambda[v,u]R[u]^T\nabla v\right)$$

with $\Lambda[u,v] = \operatorname{diag}(\gamma^{\perp}(u,v),\gamma^{\parallel}(u,v),\ldots,\gamma^{\parallel}(u,v))$.

[1] Ehrhardt ...A., Inverse Problems, 2015.

Generated Diffusion

Diffusivity of Parallel Level Sets [1]

Let $R[u]$ and $R[v]$ be local coordinates for u and v. Then the derivative of $\mathcal{S}_\beta$ can be written as

$$D\mathcal{S}_\beta[u] = -\operatorname{div}\left(R[v]K(u,v)\Lambda[u,v]R[v]^T\nabla u\right)$$

$$D\mathcal{S}_\beta[v] = -\operatorname{div}\left(R[u]K(v,u)\Lambda[v,u]R[u]^T\nabla v\right)$$

with $\Lambda[u,v] = \operatorname{diag}(\gamma^\perp(u,v),\gamma^\parallel(u,v),\ldots,\gamma^\parallel(u,v))$.

- form of the derivative independent of φ and ψ

[1] Ehrhardt ...A., Inverse Problems, 2015.

Generated Diffusion

Diffusivity of Parallel Level Sets [1]

Let $R[u]$ and $R[v]$ be local coordinates for u and v. Then the derivative of $\mathcal{S}_\beta$ can be written as

$$D\mathcal{S}_\beta[u] = -\operatorname{div}\left(R[v]K(u,v)\Lambda[u,v]R[v]^T\nabla u\right)$$

$$D\mathcal{S}_\beta[v] = -\operatorname{div}\left(R[u]K(v,u)\Lambda[v,u]R[u]^T\nabla v\right)$$

with $\Lambda[u,v] = \operatorname{diag}(\gamma^\perp(u,v),\gamma^\parallel(u,v),\ldots,\gamma^\parallel(u,v))$.

- form of the derivative independent of φ and ψ
- $\gamma^\parallel > 0$, Linear PLS: $\gamma^\perp \in \mathbb{R}$, Quadratic PLS: $\gamma^\perp > 0$

[1] Ehrhardt ...A., Inverse Problems, 2015.

Parallel Level Sets

Results

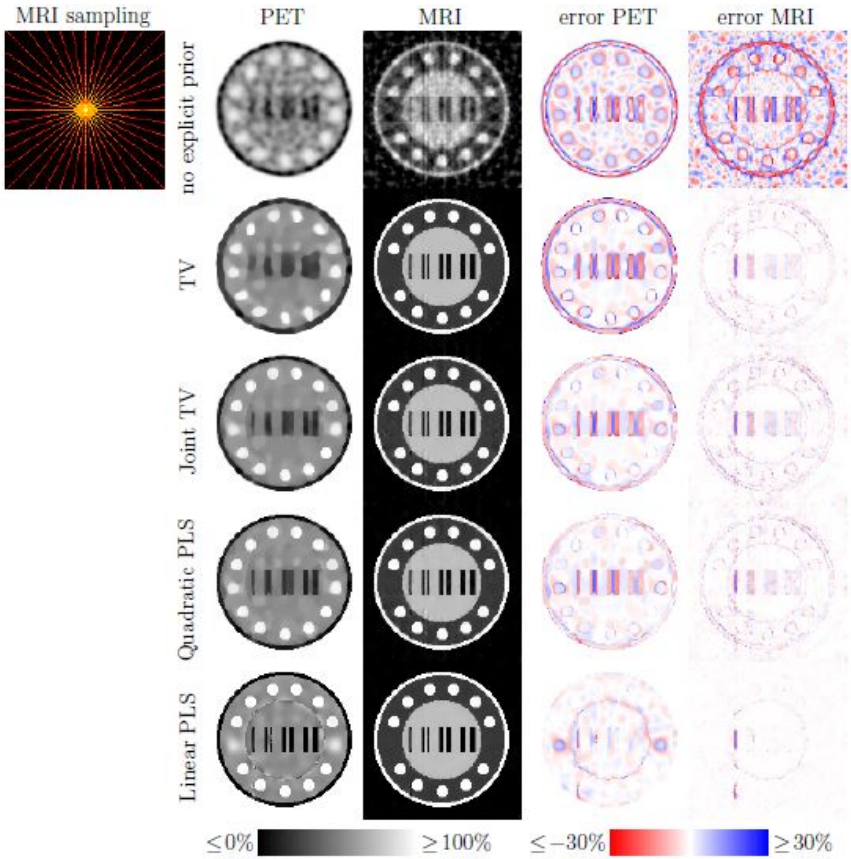


Figure 5: Reconstruction results for the case `radial25`, i.e. MRI is sampled on 25 radial lines. The reconstructed images are shown in gray on the left and the error images in red and blue on the right where blue indicates over and red under estimation. The methods' abbreviations are TV = Total Variation, PLS = Parallel Level Sets.

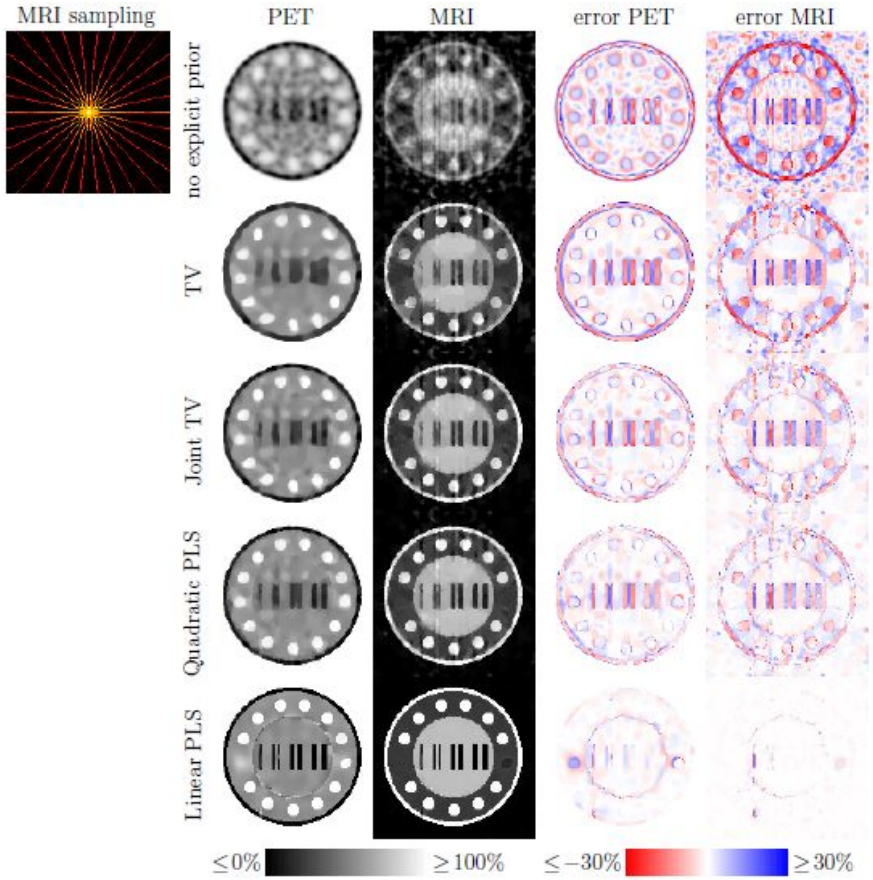


Figure 6: Reconstructions for `radial15` in gray with the error images shown in blue and red on the right. The methods' abbreviations are TV = Total Variation, PLS = Parallel Level Sets.

Summary and Conclusions

- Priors in joint image reconstruction can exploit both statistical and geometric cross-information
- Regularisation based on variation of a functional, or on an explicit PDE
- How best to combine statistics and geometry ?
- Role of machine learning

References

[1] Ehrhardt, Thielemans, Pizarro, Atkinson, Ourselin, Hutton and A. Joint Reconstruction of PET-MRI by Exploiting Structural Similarity. Inverse Problems, 31(1), 015001, 2015.

[2] Ehrhardt and A, Vector-Valued Image Processing by Parallel Level Sets. IEEE Transactions on Image Processing, 23(1), 9-18, 2014.

[3] Gallardo and Meju, Characterization of Heterogeneous Near-Surface Materials by Joint 2D Inversion of DC Resistivity and Seismic Data. Geophysical Research Letters, 30(13), 2003.

[4] Haber and Holtzman-Gazit, Model Fusion and Joint Inversion. Surveys in Geophysics, (34), 675-695, 2013.

[5] Sochen, Kimmel and Malladi. A General Framework for Low Level Vision. IEEE Transactions on Image Processing, 7(3), 310-318, 1998.

[6] Kaipio, Kolehmainen, Vauhkonen and Somersalo. Inverse Problems with Structural Prior Information. Inverse Problems, 15, 713-729, 1999.

[7] J. Rasch et al. Dynamic MRI reconstruction from undersampled data with an anatomical prescan". Inverse Problems 34.7 (2018).

[8] D. Deidda et al. Hybrid PET-MR list-mode kernelized expectation maximization reconstruction". Inverse Problems 35.4 (2019).

References

[9] S. Niu et al. Nonlocal low-rank and sparse matrix decomposition for spectral CT reconstruction". Inverse Problems 34.2 (2018).

[10] D. Kazantsev et al. Joint image reconstruction method with correlative multi-channel prior for x-ray spectral computed tomography". Inverse Problems 34.6 (2018).

[11] Y. Hu et al. Nonlinear optimization for mixed attenuation polyenergetic image reconstruction". Inverse Problems 35.6 (2019).

[12] J. F. P. J. Abascal, N. Ducros, and F. Peyrin. Nonlinear material decomposition using a regularized iterative scheme based on the Bregman distance". Inverse Problems 34.12 (2018).

[13] L. Bungert et al. Blind image fusion for hyperspectral imaging with the directional total variation". Inverse Problems 34.4 (2018).

[14] D. Rovetta and D. Colombo. Analysis of inter-domain coupling constraints for multi-physics joint inversion". Inverse Problems 34.12 (2018).

[15] B. Crestel, G. Stadler, and O. Ghattas. A comparative study of structural similarity and regularization for joint inverse problems governed by PDEs". Inverse Problems 35.2 (2019).

References

[16] T. P. Matthews and M. A. Anastasio. Joint reconstruction of the initial pressure and speed of sound distributions from combined photoacoustic and ultrasound tomography measurements". Inverse Problems 33.12 (2017).

[17] J. Rasch, E. M. Brinkmann, and M. Burger. Joint reconstruction via coupled Bregman iterations with applications to PET-MR imaging". Inverse Problems 34.1 (2018).

[18] S. Tang et al. Multicompartment magnetic resonance fingerprinting". Inverse Problems 34.9 (2018).

[19] M. Hintermüller, M. Holler, and K. Papafitsoros. A function space framework for structural total variation regularization with applications in inverse problems". Inverse Problems 34.6 (2018).

[20] M. Holler, R. Huber, and F. Knoll. Coupled regularization with multiple data discrepancies". Inverse Problems 34.8 (2018).

[21] R. Tovey et al. Directional sinogram inpainting for limited angle tomography". Inverse Problems 35.2 (2019).

References

[22] Abolfazl Mehranian and Habib Zaidi. Joint estimation of activity and attenuation in whole-body TOF PET/MRI using constrained Gaussian mixture models. IEEE transactions on medical imaging 34(9):1808–1821 (2015).

[23] Abolfazl Mehranian, Habib Zaidi, and Andrew J. Reader. MR-guided joint reconstruction of activity and attenuation in brain PET-MR. NeuroImage, 162:276–288 (2017).

[24] Abolfazl Mehranian, Martin A Belzunce, Claudia Prieto, Alexander Hammers, and Andrew J Reader. Synergistic PET and SENSE MR image reconstruction using joint sparsity regularization. IEEE transactions on medical imaging, 37(1):20–34 (2017).

[24] Abolfazl Mehranian, Martin A Belzunce, Flavia Niccolini, Marios Politis, Claudia Prieto, Federico Turkheimer, Alexander Hammers, and Andrew J Reader. PET image reconstruction using multi-parametric anato-functional priors. Physics in Medicine & Biology, 62(15):5975 (2017).

Outline

Acknowledgements

- **Collaborators :**
  - UCL : Matthias Ehrhardt, Brian Hutton, Daniil Kazantsev, Christos Panagiotou, Kris Thielmens,
  - USC : Richard Leahy
- **Funding**
  - This work was supported by the EPSRC (EP/K005278/1) and supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.