Simultaneous estimation of attenuation and activity in positron tomography

Michel Defrise

Vrije Universiteit Brussel

Based on works with

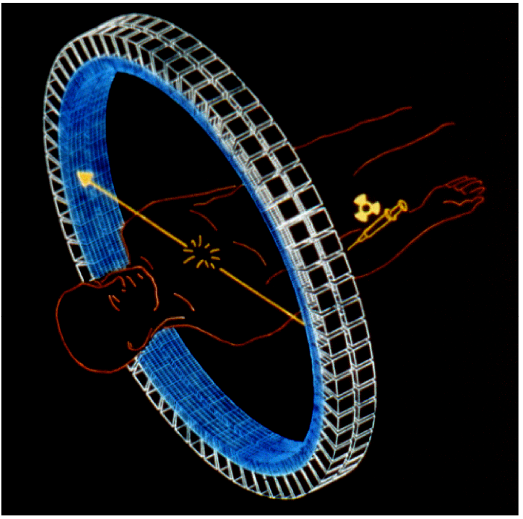
J. Nuyts, A. Rezaei (Katholieke Universiteit Leuven),

V. Panin, M. Casey, C. Michel, H. Bal, G. Bal, C. Watson, M. Conti (Siemens, Knoxville)

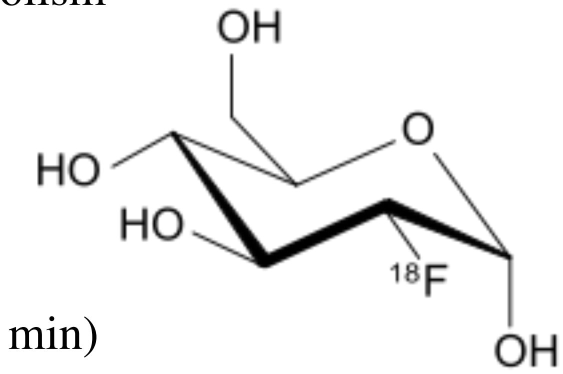
K. Salvo (Vrije Universiteit Brussel)

100 Years of the Radon Transform, RICAM, 2017

Positron tomography



In vivo imaging of glucose metabolism



^{18}F-déoxyglucose (FDG)

- Inject the tracer

- $^{18}F \rightarrow e^+ + {}^{18}O + \nu$ (T=120 min)

$e^+ + e^- \rightarrow \gamma + \gamma$ (2 photons emitted at 180 degree)

- Detect pairs of coincident γ

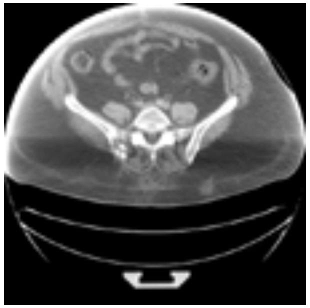
- Reconstruct the tracer spatial distribution $\lambda(x,y,z)$

Attenuation correction with hybrid PET-CT scanners

CT data: x-ray attenuation

Reconstruction of the organ density

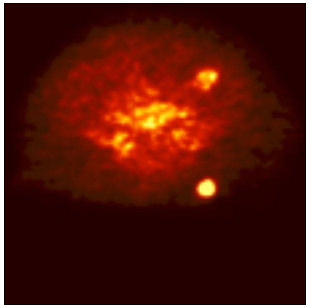
"μ map"



PET data : γγ coincidences

Reconstruction with correction for the γ attenuation

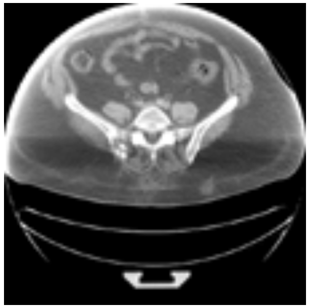
"Activity map λ": tracer's biodistribution



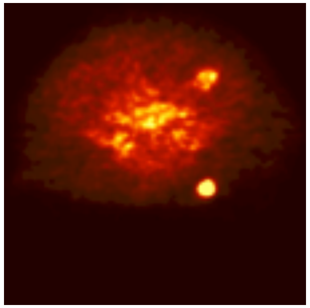
Simultaneous estimation of the activity and attenuation:

PET data y

“μ map"



"Activity λ map"



$$\text{Solve} \quad y = e^{-X\mu} \; X_{(TOF)}\lambda \quad \text{for } \mu, \lambda$$

Potential benefits:

- insensitive to motion between CT (MR) and PET acquisitions

- insensitive to CT or MR data truncation

Likely limitations:

- Estimated μ has no diagnostic value

- “Classical” non time-of-flight PET
- Time-of-flight PET:
    - Consistency conditions for the continuous model
    - Analytic solution for the continuous model
    - Maximum-likelihood estimation for a discrete model

Recent topical review: Y. Berker and Y. Li (Medical Physics 43, p. 807, 2016)

“Classical” non time-of-flight PET: 2D continuous model

$$\text{Solve} \quad y(\phi,s) = e^{-(X\mu)(\phi,s)}\ (X\lambda)(\phi,s) \quad |s| \le 1,\ \phi \in [0,\pi)$$

$$\text{for } \lambda(x), \mu(x), \quad x \in B_1 = \left\{ x \in R^2 \ \|x\| \le 1 \right\}.$$

- Attenuation is a nuisance parameter: we say that an object (λ,μ) can be identified if:

$$e^{-X\mu}\ X\lambda = e^{-X\mu'}\ X\lambda' \ \Rightarrow \lambda = \lambda'.$$

- A smooth radial object (λ,μ) is not identifiable because any smooth radial function satisfies the Helgason-Ludwig conditions hence

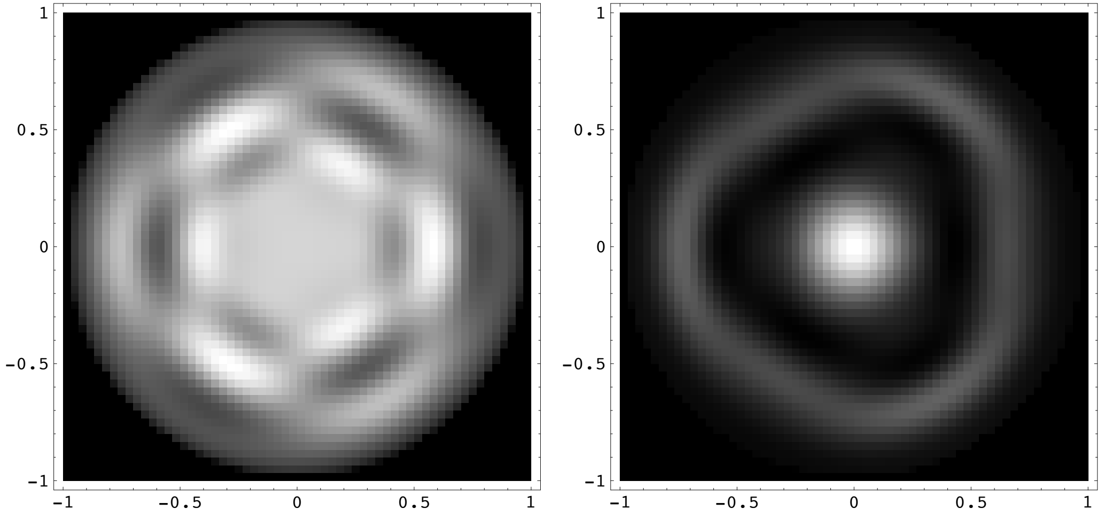
$$y = e^{-X\mu}\ X\lambda = X\lambda' \text{ with } \lambda'=\mathrm{X}^{-1}y$$

- For a non-linear problem one example of non-uniqueness does not imply that all objects are non-identifiable.

“Classical” non time-of-flight PET: 2D continuous model

- F. Natterer 1986, J. Boman 1990: the activity λ is determined by the attenuated PET data if λ consists of a finite number of point sources one of which is outside the support of a smooth attenuation μ.
- Example of two non-identifiable non-radial objects (built as a finite Fourier series constrained to satisfy HL):

The two activities λ (left) and λ’ (right) cannot be distinguished using PET data.
Both λ and λ’ are nonnegative.
Each image is scaled to its maximum.



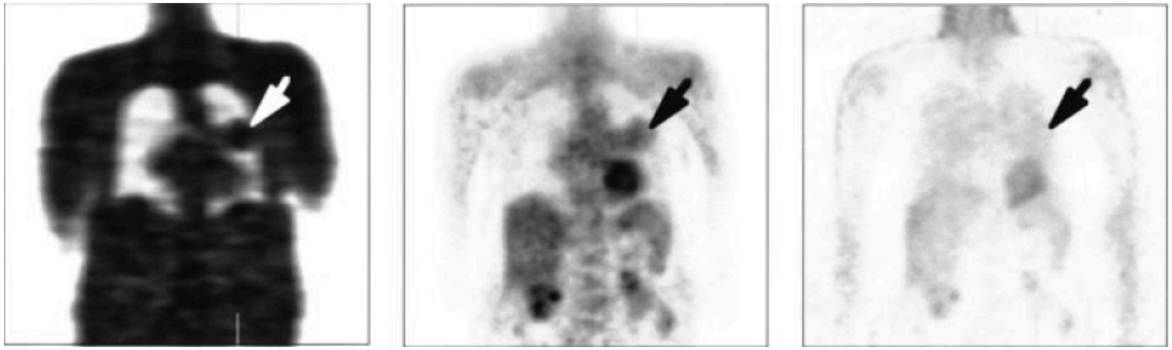
“Classical” non time-of-flight PET: 2D continuous model

- The following heuristic argument suggests that non-identifiability is the generic case for smooth functions:

Let λ and λ' be two smooth positive activity maps with support in the disk B_r, $r < 1$, and such that $(X\lambda)(\phi, s) = (X\lambda')(\phi, s)$ in a neighbourhood of $s = r$.

Define $h(\phi, s) = \log\left(\dfrac{(X\lambda)(\phi, s)}{(X\lambda')(\phi, s)}\right) \quad |s| \leq r.$

There is an attenuation map μ with support in $B_{1>r}$ such that $(X\mu)(\phi, s) = h(\phi, s)$ for $|s| \leq r$ because the range of the interior Radon transform contains all smooth functions (P. Maass, 1992) $\Rightarrow \quad e^{-X\mu}\, X\lambda = X\lambda'$



μ (transmission scan) λ, attenuation correction λ, no attenuation correction

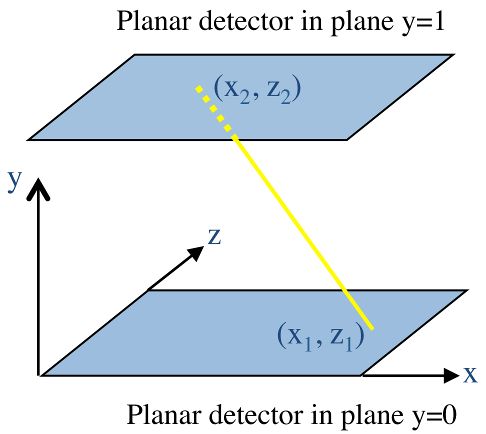
(C. Bai et al, J Nucl Med 2003)

“Classical” non time-of-flight PET: 3D continuous model

- Radial 3D objects cannot be identified using PET data.
- The range condition for the 3D x-ray transform (F. John, 1938) constrains the attenuated data much more than the HL conditions in 2D.

→ could the class of identifiable objects be larger in 3D ?

$$\frac{\partial^2 (X\lambda)(x_1,z_1,x_2,z_2)}{\partial x_1\,\partial z_2}=\frac{\partial^2 (X\lambda)(x_1,z_1,x_2,z_2)}{\partial z_1\,\partial x_2}$$



- Tests with data do not suggest a big improvement w.r.t. 2D (J. Nuyts)

“Classical” non time-of-flight PET: practical approaches

- Use the HL conditions with a sparse parameterization of μ:

- Welch A, Campbell C, Clackdoyle R, Natterer F, Hudson M, Bromiley A, Mikecz P, Chillcot F, Dodd M, Hopwood P, Craib S, Gullberg G T and Sharp P 1998 Attenuation correction in PET using consistency information, IEEE Trans. Nuclear Science

- Iterative methods for a discretized model:

- Censor Y, Gustafson D, Lent A, Tuy H, 1979, A new approach to ECT: simultaneous calculation of attenuation and activity coefficients, IEEE Trans. Nuclear Science.
- Nuyts J, Dupont P, Stroobants S, Benninck R, Mortelmans L, Suetens P 1999 Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms, IEEE Trans Med Imag (MLAA).
- F. Crepaldi, A. De Pierro, 2006, Activity and attenuation recovery from activity data only in emission computed tomography, Comput. Appl. Math.
- A. Bronnikov, 2000, Reconstruction of attenuation map using discrete consistency conditions, IEEE Trans Med Imag.
- V. Panin, G. Zeng, G. Gullberg., 2001, A method of attenuation map and emission activity reconstruction from emission data,

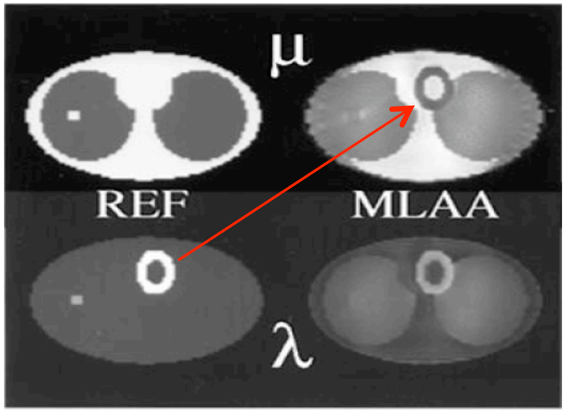


Image from Nuyts et al, IEEE Trans Med Imag 1999.

Note the "cross-talk" between activity and attenuation.

- “Classical” non time-of-flight PET

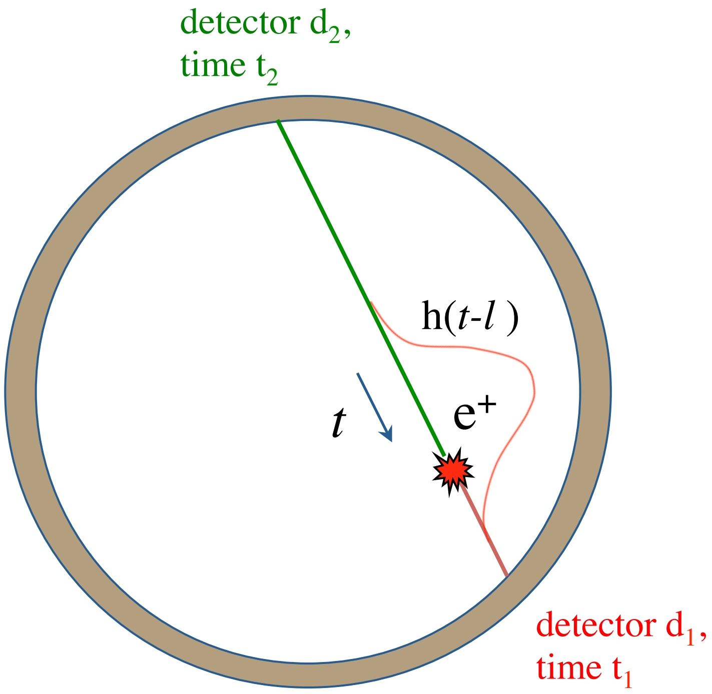
- Time-of-flight PET:

  Consistency conditions for the continuous model

  Analytic solution for the continuous model

  Maximum-likelihood estimation for a discrete model

Time-of-flight PET



- Measure the arrival time difference $t = c\ (t_2 - t_1\ )/2$

- Localize the positron decay along the line of response (LOR) with uncertainty profile h with width $\Delta l = c\ \Delta t\ /2$

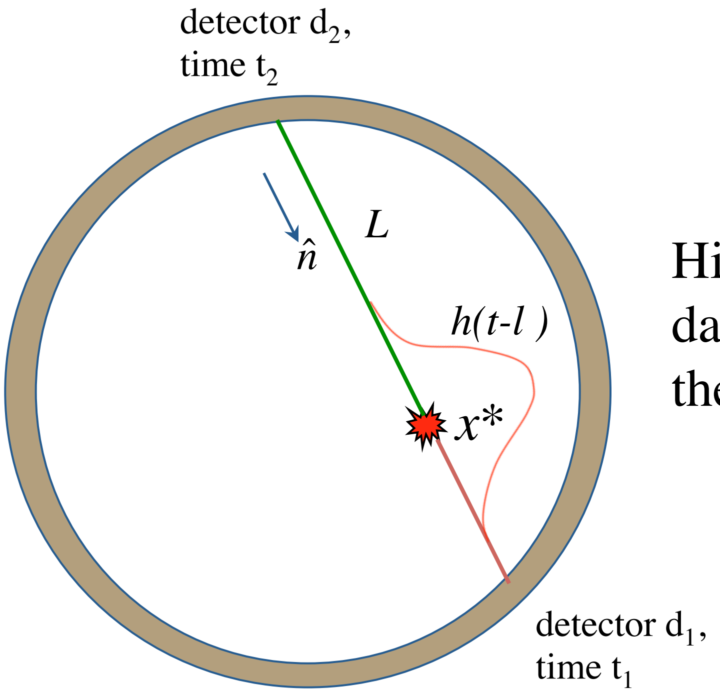
$\Delta t_{\text{FWHM}} \approx 300$ ps ⟹ $\Delta l_{\text{FWHM}} \approx 45$ mm

- TOF profile h is well approximated by a gaussian.

Each coincident event yields more information than without time-of-flight, by a factor ≈ patient diameter/Δl_{FWHM} ⟹ *improved SNR.*

Anger 1966; 1980's: Allemand et al, Ter Pogossian et al, Mullani et al, Lewellen et al
Snyder et al 1980's, Tomitani 1981, Watson 2007, Surti et al 2006/8, Popescu & Lewitt 2006, …

Continuous model



Histoprojections (S. Matej et al, 2009) parameterize data with the *most likely annihilation point x^** and the unit vector $\hat{n}$ along the line L.

attenuation image

TOF profile

activity image

$$y(L,t) = \underbrace{\exp(-\int_L dl'\,\mu(l'))}_{a(L)}\ \underbrace{\int_L dl\, h(t-l)\,\lambda(l)}_{p(L,t)}$$

measured data for all lines L and all t

attenuation factor

unattenuated data

3D data parameterization using *histoprojections*

Histoprojections are defined by

$$p(x^*,\hat{n}) = \int dl\ \ \lambda(x^* + l\ \hat{n})\, h(l) \qquad x^* \in R^3,\ \hat{n} \in S^2$$

- Consistency based on Fourier slice theorem:

$$\hat{p}(\omega,\hat{n})\ \hat{h}(-\omega\cdot\hat{n}') = \hat{p}(\omega,\hat{n}')\ \hat{h}(-\omega\cdot\hat{n}) \quad \omega \in R^3,\ \hat{n},\ \hat{n}' \in S^2$$

- Local consistency condition for a gaussian *h* with st. dev. σ :

$$(Dp)(x^*,\hat{n}) = \nabla_{\hat{n}} p(x^*,\hat{n}) - \sigma^2\ \nabla_{x^*}(\hat{n}\cdot\nabla_{x^*})\ p(x^*,\hat{n}) = 0$$

The characteristic curves are loci of constant *x** in the 5D data space.

Y. Li, M.D., S. Metzler, S. Matej, Phys Med Biol 2015, Inverse Problems 2016.

- “Classical” non time-of-flight PET

- Time-of-flight PET:

  Consistency conditions for the continuous model

  Analytic solution for the continuous model

  Maximum-likelihood estimation for a discrete model

Recall the measured attenuated data

$$y(x^*,\hat{n}) = a(x^*,\hat{n})\, p(x^*,\hat{n}) \qquad x^* \in R^3,\ \hat{n} \in S^2$$

with $a(x^*,\hat{n}) = \exp(-(X\mu)(x^*,\hat{n})) = \exp(-\int dl\ \mu(x^* + l\hat{n}))$

The corrected data are consistent ⟹ solve

$$\begin{cases} Dp = D\dfrac{y}{a} = \nabla_{\hat{n}} \dfrac{y}{a} - \sigma^2\ \nabla_{x^*}(\hat{n}\cdot\nabla_{x^*})\ \dfrac{y}{a} = 0 & \text{2 indept. eqns.} \\ \\ a(x^*,\hat{n}) = a(x^* + t\hat{n},\hat{n}) \quad t \in R & a \text{ is independent of } t \end{cases}$$

This yields for each LOR $(x_0,\hat{n})$ a separate set of linear equations:

$$y\ \nabla_{\hat{n}} \log a - \left(\sigma^2(\hat{n}\cdot\nabla_{x^*} y) + t\ \ y\right)\nabla_{x^*} \log a = Dy \qquad t\ \in R$$

with $\mathrm{y}(\mathrm{x}_0 + t\ \hat{n},\hat{n}),\ \ a(x_0,\hat{n})$

This yields for each LOR $(x_0,\hat{n})$ a separate set of linear equations:

$$y\ \nabla_{\hat{n}}\log a-\left(\sigma^2(\hat{n}\cdot\nabla_{x*}y)+t\ y\right)\nabla_{x*}\log a=Dy\quad t\in R$$

with $\mathrm{y}(\mathrm{x}_0+t\ \hat{n},\hat{n})$, $a(x_0,\hat{n})$

The gradient $\nabla_{\hat{n}}\log a,\nabla_{x*}\log a$ is determined by the data for all lines of response such that

- $y(x_0,\hat{n})>0$
- $\lambda(x)$ is not a single point source along the line $(x_0,\hat{n})$

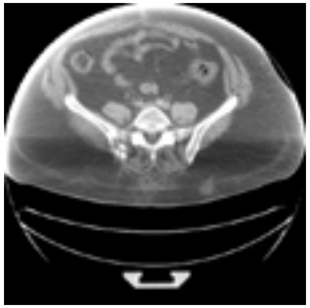
The attenuation factors is determined up to a global multiplicative constant for all LORs needed to reconstruct the activity.

Main result:

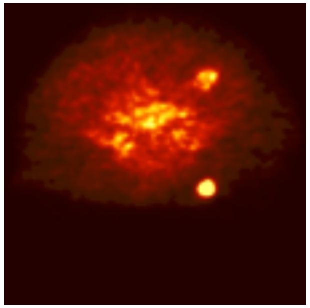
TOF-PET emission data determine the activity up to a constant.

TOF-PET data : $\gamma\gamma$ coincidences

"μ map"



”Activity λ map"



- “Classical” non time-of-flight PET
- Time-of-flight PET:
  - Consistency conditions for the continuous model
  - Analytic solution for the continuous model
  - Maximum-likelihood estimation for a discrete model
    - Estimation of λ and μ
    - Estimation of λ and a

Notations

- $*$ activity image $\lambda_j \quad j = 1,\ldots,M$
- $*$ attenuation image μ_j
- $*$ known background b_{it}
- $*$ data $y_{it} = \text{Poisson}\left(< y_{it} > = a_i \; p_{it} + b_{it} \right)$

  line index $i = 1,\ldots,N$ TOF index $t = 1,\ldots,T$

- $*$ expected non

attenuated data $p_{it} = \sum_j c_{ijt} \lambda_j \quad i = 1,\ldots,N$

- $*$ "system matrix" c_{ijt}
- $*$ attenuation factors $a_i = \exp(-\sum_j c_{ij} \mu_j)$

Typical dimensions: N = 400x168x621, T=13, M=400x400x109

Maximum likelihood estimation of activity and attenuation map

$$(\hat{\lambda},\hat{\mu}) = \underset{\lambda\geq 0,\ \mu\geq 0}{\operatorname{argmax}} L(y,\lambda,\mu)$$

$$L(y,\lambda,\mu) = \sum_{t=1}^{T} \sum_{i=1}^{N} \left\{ -\left\langle y_{it} \right\rangle + y_{it} \log\left\langle y_{it} \right\rangle \right\}$$

$$\left\langle y_{it} \right\rangle = e^{-\sum_{j'} c_{ij'}\mu_{j'}} \left(\sum_{j} c_{ijt}\lambda_j + b_{it} \right)$$

MLAA : alternate optimization w.r.t.

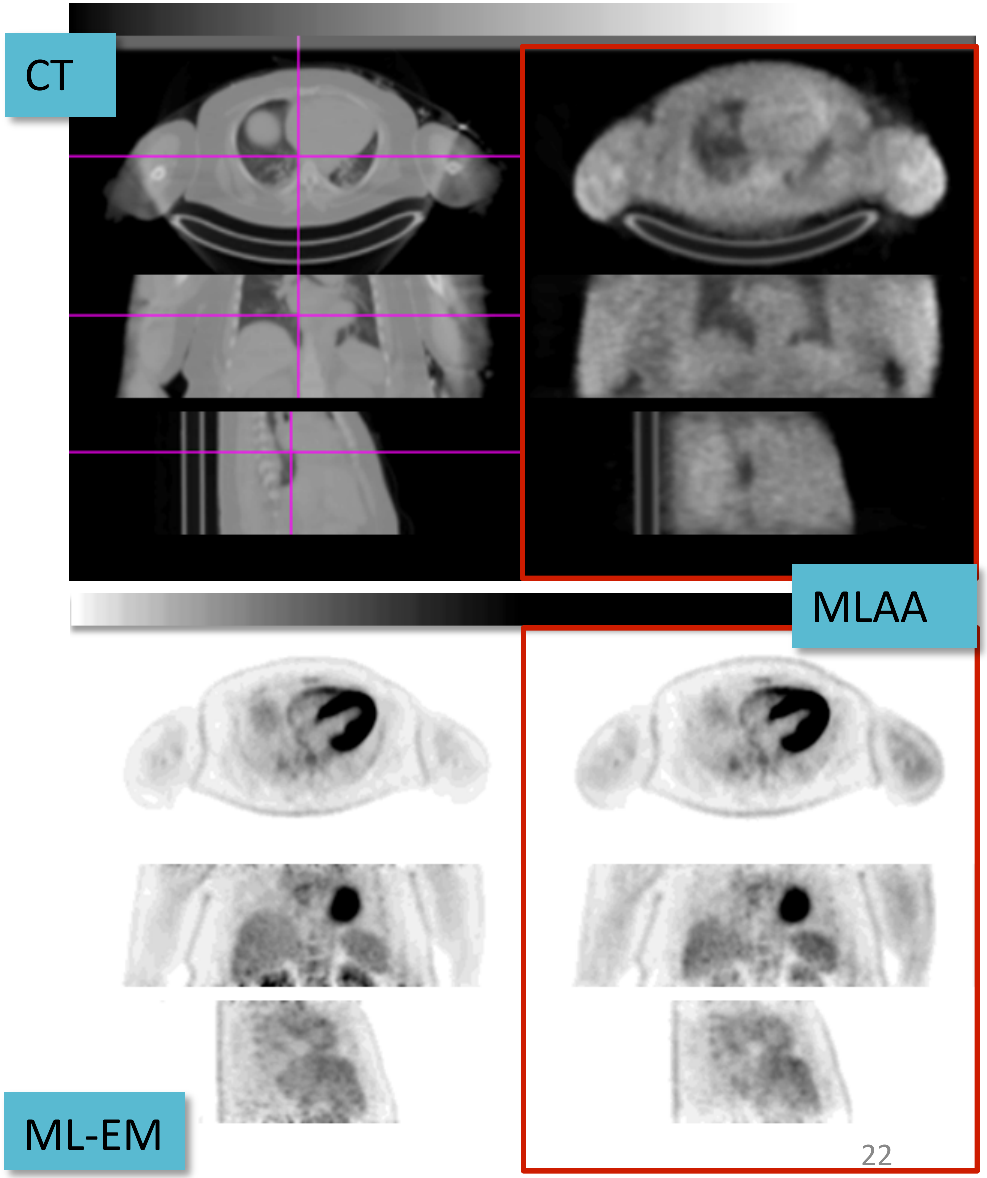
- λ : ML-EM
- μ : ML-TR (≈ linearized version of O’Sullivan & Benac’s algorithm)

(Nuyts et al 1999, Rezaei et al 2014, Boellaard et al 2014).

No result on the convergence.

CT-less TOF PET using MLAA

Rezaei et al Trans Med Imag 2012,
slide courtesy J. Nuyts



Thorax scan, 5 min, 570 MBq ^{18}F-FDG

Siemens Biograph mCT PET/CT

- TOF time resolution of 0.580ns
- 13 time bins of 0.312ns
- 168 projection angles over 180º
- 85 cm scanner diameter
- 70 cm FOV

- “Classical” non time-of-flight PET
- Time-of-flight PET:
  - Consistency conditions for the continuous model
  - Analytic solution for the continuous model
  - Maximum-likelihood estimation for a discrete model
    - Estimation of λ and μ
    - Estimation of λ and a

Maximum likelihood estimation of activity and attenuation factors

$$(\hat{\lambda},\hat{a}) = \underset{\lambda\geq 0,\ 1\geq a\geq 0}{\operatorname{argmax}} L(y,\lambda,a)$$

$$L(y,\lambda,a) = \sum_{t=1}^{T} \sum_{i=1}^{N} \left\{-\langle y_{it}\rangle + y_{it} \log\langle y_{it}\rangle\right\}$$

$$\langle y_{it}\rangle = a_i\left(\sum_{j} c_{ijt}\lambda_j + b_{it}\right)$$

- \+ : Simpler algorithms,
- \+ : Does not attempt to reconstruct μ outside the support of the activity,
- \- : Larger number of parameters in 3D PET (# LORs > # voxels),
- -: Impossible to use prior information (MR anatomy, or physical constraints on μ).

Maximum likelihood estimation of activity and attenuation factors

$$(\hat{\lambda}, \hat{a}) = \underset{\lambda \geq 0,\ 1 \geq a \geq 0}{\operatorname{argmax}} L(y, \lambda, a)$$

$$L(y, \lambda, a) = \sum_{t=1}^{T} \sum_{i=1}^{N} \left\{ -\left\langle y_{it} \right\rangle + y_{it} \log \left\langle y_{it} \right\rangle \right\}$$

$$\left\langle y_{it} \right\rangle = a_i \left(\sum_{j} c_{ijt} \lambda_j + b_{it} \right)$$

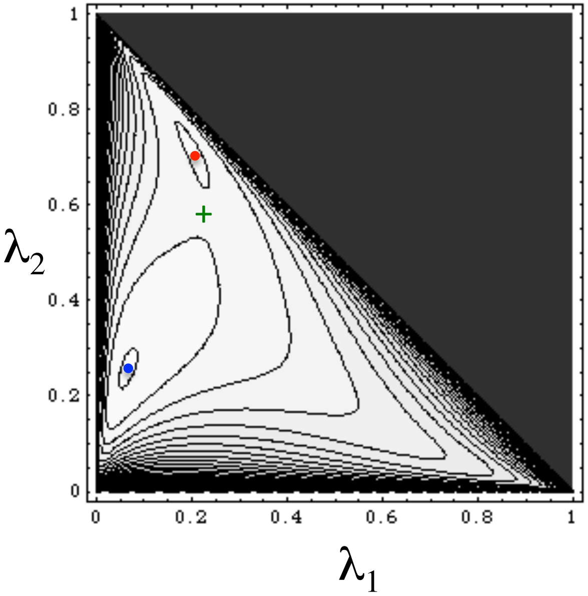
- L is bi-concave in a and λ,
- Unique maximum (up to scale factor) if $b=0$ and consistent data y,
- Scale invariance $L(y, \lambda, a) = L(y, \beta\lambda, a / \beta) \quad \beta > 0$.

Maximum likelihood estimation of activity and attenuation factors

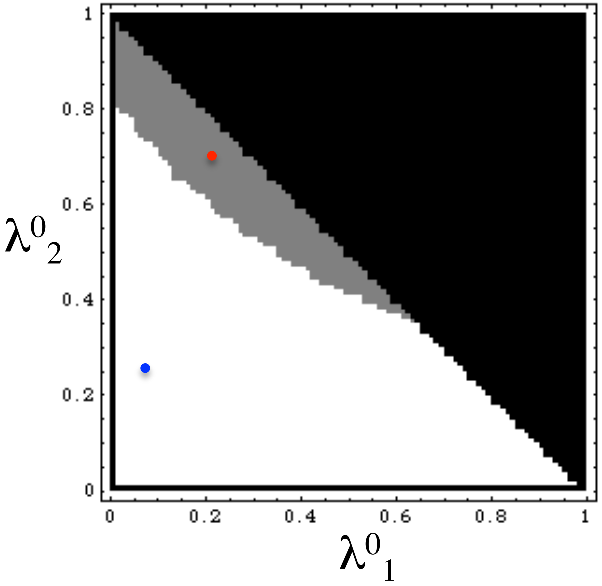
Checking the existence of local maxima of L(y, λ, a) on a toy problem, with non consistent data:

$$< \begin{Bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \end{Bmatrix} > = \begin{Bmatrix} a_1\lambda_1 \\ a_1(\lambda_1+\lambda_2) \\ a_2\lambda_2 \\ a_2(\lambda_2+\lambda_3) \\ a_3\lambda_3 \\ a_3(\lambda_1+\lambda_3) \end{Bmatrix}$$

$$M = 3, N = 3, T = 2$$



Level sets of the likelihood with two local maxima and a saddle point (+)



Convergence basins of MLACF (λ^0 initial iterate)

Scale set as $\lambda_1 + \lambda_2 + \lambda_3 = 1$.

Joint estimation with the MLACF algorithm

Alternate application of ML-EM for fixed attenuation factors and for fixed activity (A. Rezaei, M.D., J. Nuyts, 2014)

$$\lambda_j^{k+1} = \frac{\lambda_j^k}{\sum_i a_i^k c_{ij}} \sum_{i,t} a_i^k c_{ijt} \frac{y_{it}}{\left\langle y_{it} \middle| a^k, \lambda^k \right\rangle}$$

$$a_i^{k+1} = \frac{a_i^k}{\boxed{\sum_{i,t} p_{it}^{k+1}}} \sum_{i,t} p_{it}^{k+1} \frac{y_{it}}{\left\langle y_{it} \middle| a^k, \lambda^{k+1} \right\rangle}$$

with $p_{it}^k = \sum_j c_{ijt} \lambda_j^k, \; \left\langle y_{it} \middle| a^k, \lambda^k \right\rangle = a_i^k p_{it}^k + b_{it}$

+ special cases when
denominators
are close to 0 !

+ truncation 0≤a≤1

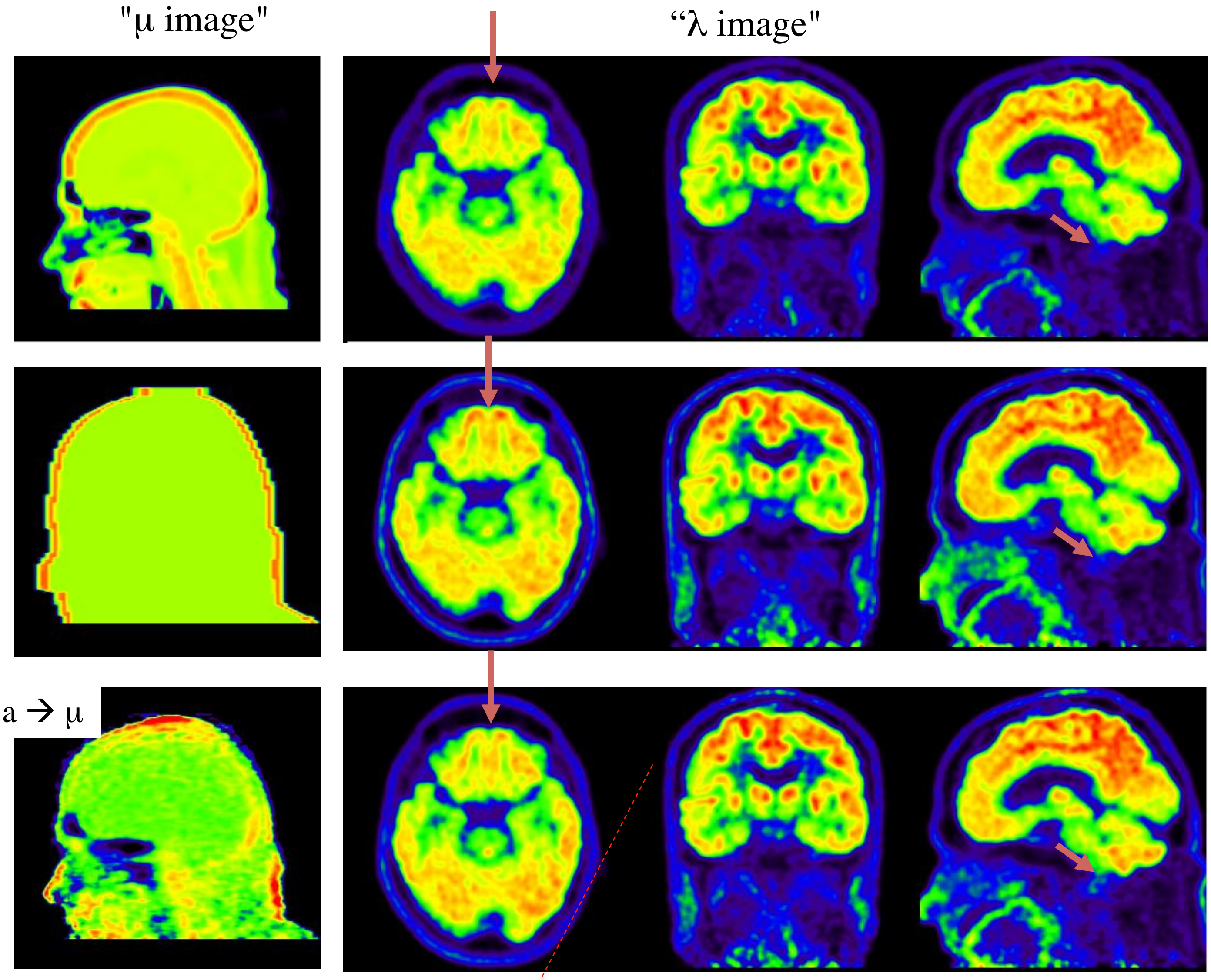
+ setting global scale

- Monotone: log-likelihood L is non-decreasing,
- The a update converges geometrically,
- Simplified algorithm if no background ($b=0$) or pre-corrected data (y-b).

(V. Panin et al, IEEE Med Conf 2012)

Example. High count FDG (196 10^6 trues)

Data and reconstruction: V. Panin, Siemens



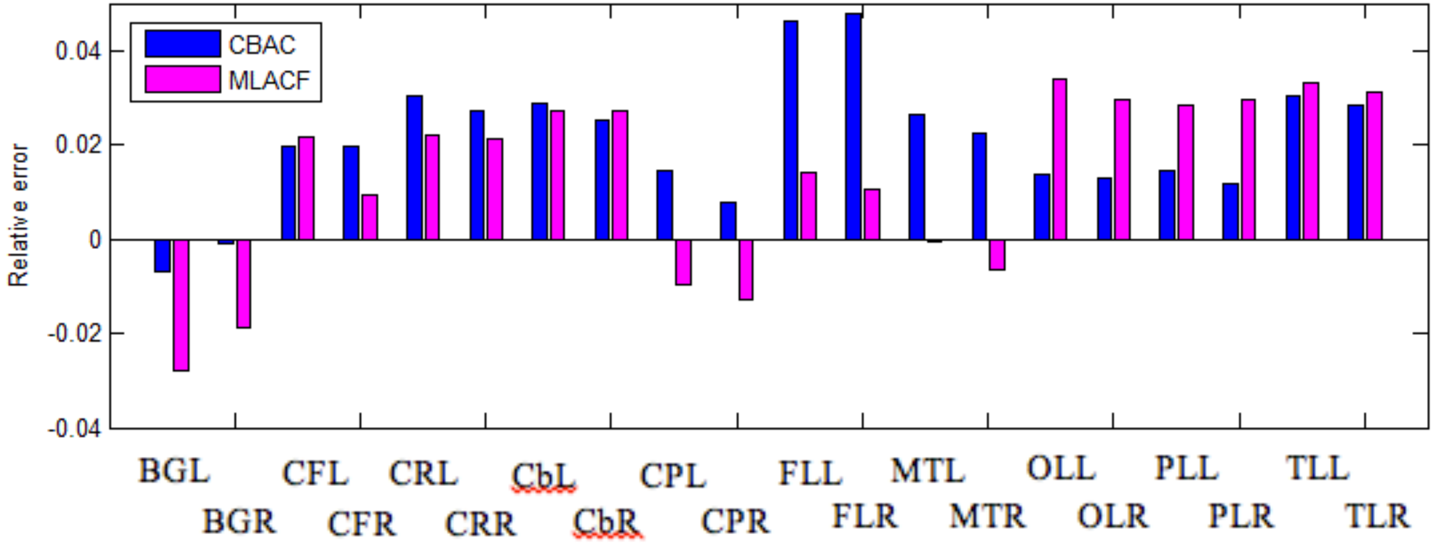
Using measured CT Density μ

Using a uniform model of the density μ + manual "skull"

Simultaneous MLACF estimation of a,λ

Evaluation of MLACF based calculated attenuation brain PET imaging for 57 FDG patient studies

H. Bal et al (Siemens Healthcare, VUB, Centre Hospitalier Princesse Grace Monaco) Phys Med Biol 2017.



Mean relative errors (a) and the standard deviation (b) in the standard ROIs of brain for CBAC and MLACF with respect to CT based reconstruction for pooled patient data.

BG – basal ganglia; CF – calcarine fissure and surrounding cortex; CR – central region; Cb – cerebellum; CP – cingulate and paracingulate gyri; FL – frontal lobe; MT – mesial temporal lobe; OL – occipital lobe; PL – parietal lobe; TL – temporal lobe. ”L” and “R” at the end of each ROI name indicates left and right respectively.

Joint estimation with a simultaneous update the expectation-maximization method

- Goal: could this lead to more results on convergence ?

- Based on the expectation-maximization method (Dempster, Laird, Rubin 1977) with appropriate complete (latent) variables (K. Salvo, 2016).

- Previous use of EM for joint estimation:
  - D. Politte and D. Snyder, 1991
  - J. Fessler, N. Clinthorne, W. Rogers, 1993

  - for CT: K. Lange and R. Carson, 1984
  - for joint estimation in PET: A. Mihlin and C. Levin, 2017

Joint estimation with a simultaneous update the expectation-maximization method

- Usual complete data for ML-EM (# events emitted in voxel *j* and detected in LOR *i* and TOF bin *t*).

$$x_{ijt} = Poisson(a_i c_{ijt} \lambda_j)$$

$$\Downarrow$$

(K. Salvo)

$$\begin{aligned} \{\ x_{ijt}^{(a)} &= Poisson((1-a_i)\ c_{ijt}\lambda_j), \\ x_{ijt}^{(b)} &= Poisson((a_i - a_{\min})\ c_{ijt}\lambda_j), \\ x_{ijt}^{(c)} &= Poisson(a_{\min} c_{ijt} \lambda_j)\ \} \end{aligned}$$

- Complete data for joint estimation.
- Automatically enforces bound 1 ≥ a ≥ $a_{\min}$.
- EM yields monotone algorithm with bounded estimates and asymptotic regularity.
- Slower convergence than alternate optimization.

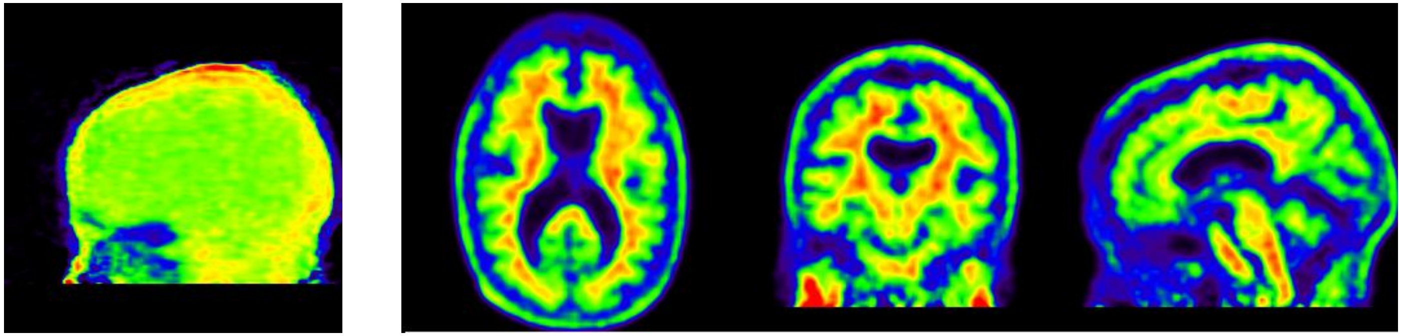
- “Classical” non time-of-flight PET

- Time-of-flight PET:
  - Consistency conditions for the continuous model
  - Analytic solution for the continuous model
  - Maximum-likelihood estimation for a discrete model
    - Estimation of λ and μ
    - Estimation of λ and a

- Conclusions

Conclusions

- Joint estimation in TOF PET has a unique solution for λ, up to a global multiplicative constant.
- Limited precision close to the edges of the patient.
- Few results on convergence.
- We only used the PET data but related works
    - Incorporate MR information (A. Salomon et al, IEEE Trans Med Imag 2011)
    - Jointly estimate λ and the deformation field of a CT scan (A. Rezaei et al, Phys Med Biol 2015; Bousse et al IEEE Trans Med Imag 2015)
    - Use partial CT data available in part of the field-of-view (V. Panin et al, IEEE Med Conf 2012, J. Nuyts et al, IEEE Trans Med Imag 2013)



Simultaneous
estimation of a, λ

Thank you !

Thanks to

J. Nuyts, A. Rezaei (Katholieke Universiteit Leuven),

V. Panin, M. Casey, C. Michel, G. Bal, C. Watson, M. Conti (Siemens Healthcare, Knoxville)

K. Salvo (Vrije Universiteit Brussel)