Přednáška 8: Bilineární model a bayesovský přístup k maticové dekompozici, odhad počtu komponent, ukázka aplikací

Ondřej Tichý

Bayesovské metody v machine learningu (BML)

April 13, 2018

Obsah přednášky

- opakování
- bayesovská formulace maticové dekompozice
  - odhad počtu komponent
- další předpoklady a přístupy
- praktické aplikace
  - vyhodnocení dat ze scintigrafie ledvin
  - pozitronová emisní tomografie (PET)
  - hyperspektrální zobrazování

Formulace modelu

Maticová formulace

- každý obrázek (sloupec matic D) $\mathbf{d}_t$ je lineární kombinace jednotlivých zdrojových obrázků:

$$\mathbf{d}_t = \mathbf{a}_1 x_{1,t} + \mathbf{a}_2 x_{2,t} + \cdots + \mathbf{a}_r x_{r,t} \qquad (1)$$

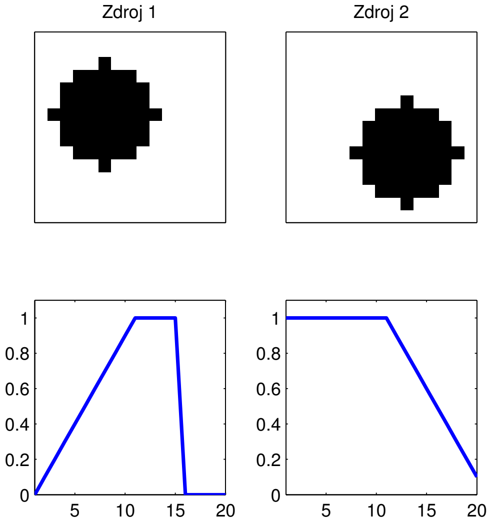
  - měřená data: $D = [\mathbf{d}_1, \ldots, \mathbf{d}_t, \ldots, \mathbf{d}_n]$
  - zdrojové obrázky: $A = [\mathbf{a}_1, \ldots, \mathbf{a}_r]$
  - váhy zdrojových obrázků: $X = \left[\underline{\mathbf{x}}_1^T, \ldots, \underline{\mathbf{x}}_t^T, \ldots, \underline{\mathbf{x}}_n^T\right]^T$
- maticový zápis:

$$D = AX^T. \qquad (2)$$

pozn.: $D_{i,j} = \sum_{k=1}^{r} A_{i,k} X_{k,j}$

Cvičná data

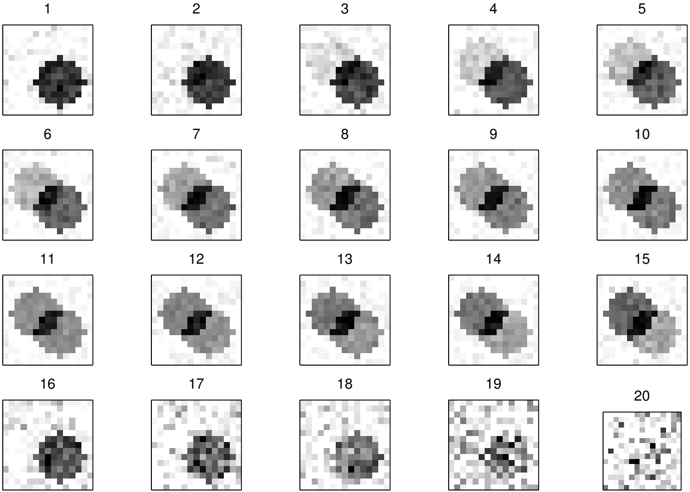
Simulované hodnoty



Cvičná data

Data

- vynásobíme AX^T a přidáme šum:



Opakování

- PCA a SVD
- NMF
- Variační PCA
- Variační PCA s pozitivitou

- Variační PCA

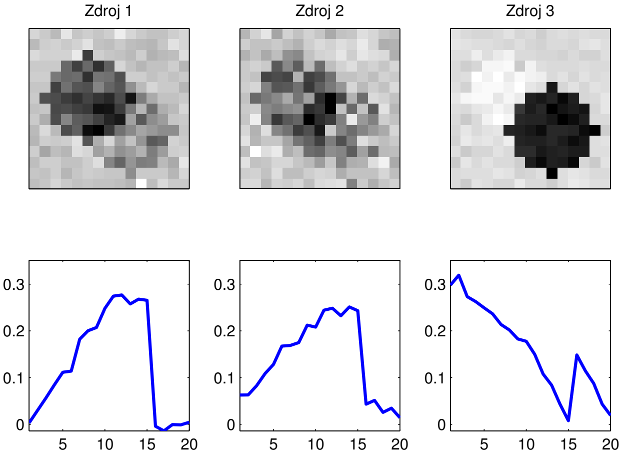
$$
\begin{aligned}
f\left(D|A,X,\omega\right) =& \mathcal{N}\left(AX^T, \omega^{-1}I_p \otimes I_n\right), && (3)\\
f(\omega) =& \mathcal{G}\left(\vartheta_0, \rho_0\right), && (4)\\
f\left(A\right) =& \mathcal{N}\left(\mathbf{0}, I_p \otimes I_r\right), && (5)\\
f(X) =& \mathcal{N}\left(\mathbf{0}, I_n \otimes I_r\right). && (6)
\end{aligned}
$$

- odhad matice A: $\tilde{f}(A|D,r) = \mathcal{N}_A(\mu_A, I_p \otimes \Sigma_A)$, kde

$$
\Sigma_A = \left(\widehat{\omega}\widehat{X^T X} + I_r\right)^{-1}, \; \mu_A = \widehat{\omega} D \widehat{X} \Sigma_A. \tag{7}
$$

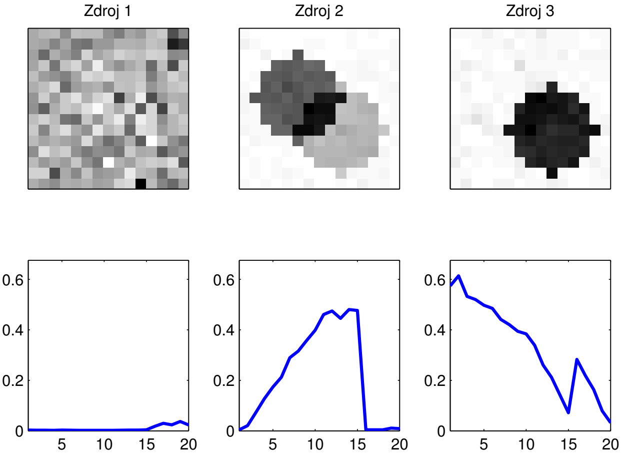
Opakování

Variační PCA



Opakování

Variační PCA s pozitivitou



Variační PCA s odhadem počtu komponent

Formulace

- model pozorování
$$D = AX^T + E, \tag{8}$$
kde $E_{i,j} = \mathcal{N}(0, \omega^{-1}), \ \forall i, \forall j.$
- model dat
$$f\left(D|A, X, \omega\right) = \mathcal{N}\left(AX^T, \omega^{-1} I_p \otimes I_n\right), \tag{9}$$
kde ω je (neznámá) variance šumu s apriorním modelem
$$f(\omega) = \mathcal{G}\left(\vartheta_0, \rho_0\right). \tag{10}$$
- pozn.:
$$\mathcal{N}_A(\mu_A, \Sigma_p \otimes \Phi_r) = \frac{1}{(2\pi)^{\frac{pr}{2}} |\Sigma_p|^{\frac{r}{2}} |\Phi_r|^{\frac{p}{2}}} \times \\ \times \exp\left(-\frac{1}{2}\mathrm{tr}\left[\Sigma_p^{-1}(A-\mu_A)(\Phi_r^{-1})^T(A-\mu_A)^T\right]\right), \tag{11}$$

Variační PCA s odhadem počtu komponent

Formulace

- předpokládejme známé (zvolené, nadsazené) r
- apriorní modely pro A jsou voleny jako

$$f(A) = \mathcal{N}\left(\mathbf{0}, I_p \otimes V^{-1}\right), \;\; V = \mathrm{diag}(\mathbf{v}), \;\; v = [v_1, \ldots, v_r], \tag{12}$$

  a apriorní model pro prvky vektoru **v** volíme jako

$$f(v_k) = \mathcal{G}(\alpha_0, \beta_0), \;\; \forall k \tag{13}$$

- model pro X necháme

$$f(X) = \mathcal{N}\left(\mathbf{0}, I_n \otimes I_r\right). \tag{14}$$

- máme model dat $f(\mathbf{y}|\theta)$,
- máme q podmíněně nezávislých proměnných, tzn. pro jejich arpiorna platí

$$f(\theta_1, \theta_2, \ldots, \theta_q) = \prod_{i=1}^{q} f(\theta_i), \tag{15}$$

- pak aposteriorní rozdělení $\tilde{f}(\theta_i|\mathbf{y})$ pro $i$tý parametr získáme VB aproximací jako

$$\tilde{f}(\theta_i|\mathbf{y}) \propto \exp\left[\mathrm{E}_{\tilde{f}(\theta_{\setminus i}|\mathbf{y})}\left(\ln f(\theta_1, \theta_2, \ldots, \theta_q, \mathbf{y})\right)\right], \ \forall i. \tag{16}$$

Variační PCA s odhadem počtu komponent

Řešení

- odhad variance šumu: $\tilde{f}(\omega|D,r) = \mathcal{G}_\omega(\vartheta, \rho)$, kde

$$\vartheta = \vartheta_0 + \frac{np}{2}, \tag{17}$$

$$\rho = \rho_0 + \frac{1}{2}\mathrm{tr}\left(DD^T - \widehat{A}\widehat{X}^T D^T - D\widehat{X}\widehat{A}^T\right) + \frac{1}{2}\mathrm{tr}\left(\widehat{A^T A}\widehat{X^T X}\right) \tag{18}$$

s momentem $\widehat{\omega} = \frac{\vartheta}{\rho}$.

- ohad matice X: $\tilde{f}(X|D,r) = \mathcal{N}_X(\mu_X, I_n \otimes \Sigma_X)$, kde

$$\Sigma_X = \left(\widehat{\omega}\widehat{A^T A} + I_r\right)^{-1}, \ \mu_X = \widehat{\omega} D^T \widehat{A} \Sigma_X. \tag{19}$$

Variační PCA s odhadem počtu komponent

Řešení

- ohad matice A: $\tilde{f}(A|D,r) = \mathcal{N}_A(\mu_A, I_p \otimes \Sigma_A)$, kde

$$\Sigma_A = \left(\widehat{\omega}\widehat{X^T X} + \widehat{V}\right)^{-1}, \; \mu_A = \widehat{\omega} D \widehat{X} \Sigma_A. \tag{20}$$

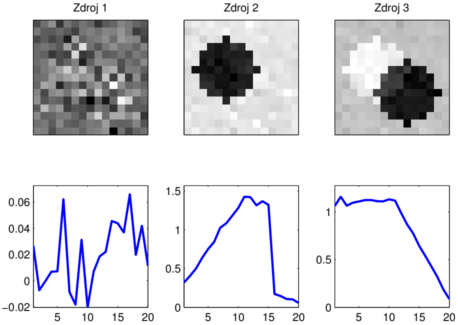
- odhad vektoru **v**: $\tilde{f}(v_k|D,r) = \mathcal{G}_{v_k}(\alpha_k, \beta_k)$,kde

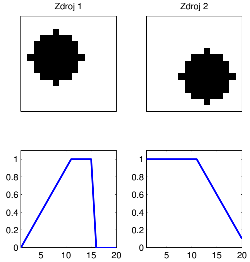
$$\alpha_k = \alpha_0 + \frac{p}{2}, \; \beta_k = \beta_0 + \frac{1}{2}\left(\widehat{X^T X}\right)_{k,k}, \tag{21}$$

Variační PCA s odhadem počtu komponent

Cvičný příklad

- ukážeme si případ pro $r = 3$ a $r = 6$:

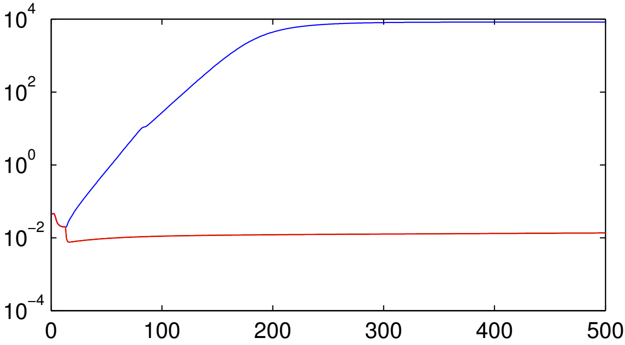




Variační PCA s odhadem počtu komponent

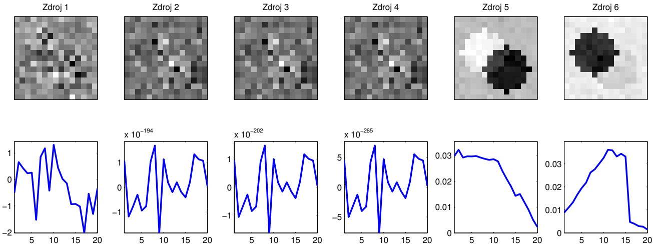
Cvičný příklad

- vektor **v** v průběhu iterací



Variační PCA s odhadem počtu komponent

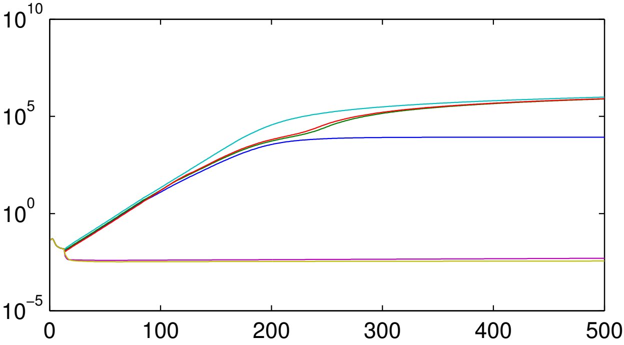
Cvičný příklad



Variační PCA s odhadem počtu komponent

Cvičný příklad

- vektor **v** v průběhu iterací



Další možnosti

Jen zmínka...

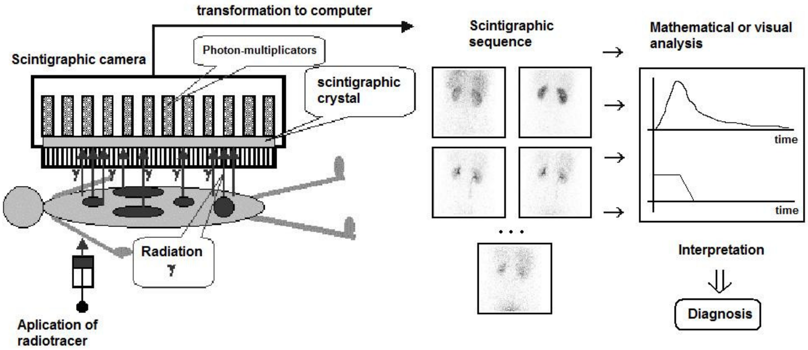
- řídkost na A
- řídkost na X
- modely šumu
- fyzikální modely křivek

Praktické ukázky

- vyhodnocení dat ze scintigrafie ledvin
- pozitronová emisní tomografie (PET)
- hyperspektrální obrázky

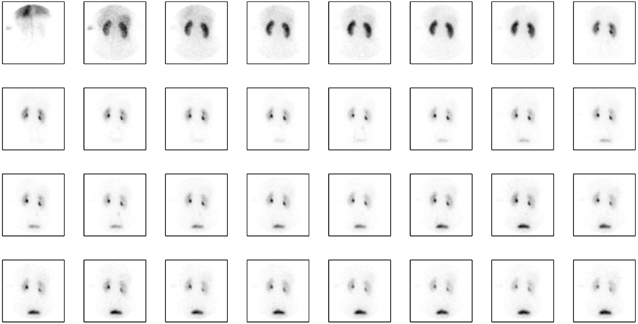
Scintigrafie ledvin

Vznik obrazu



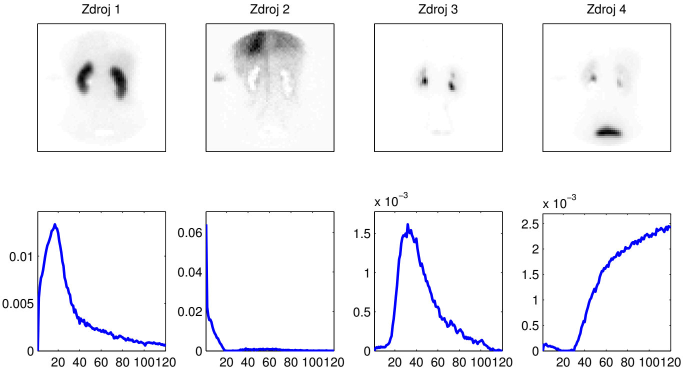
Scintigrafie ledvin

Obrazová sekvence



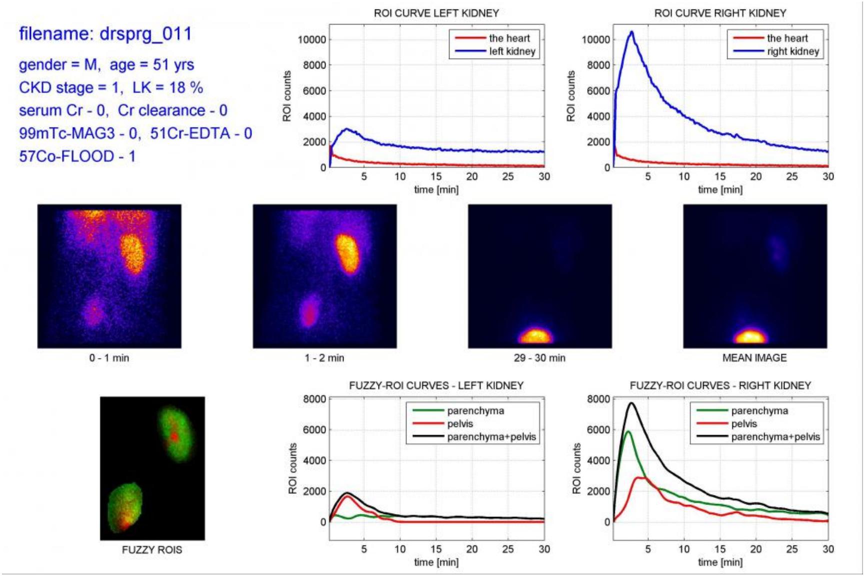
Scintigrafie ledvin

Separace



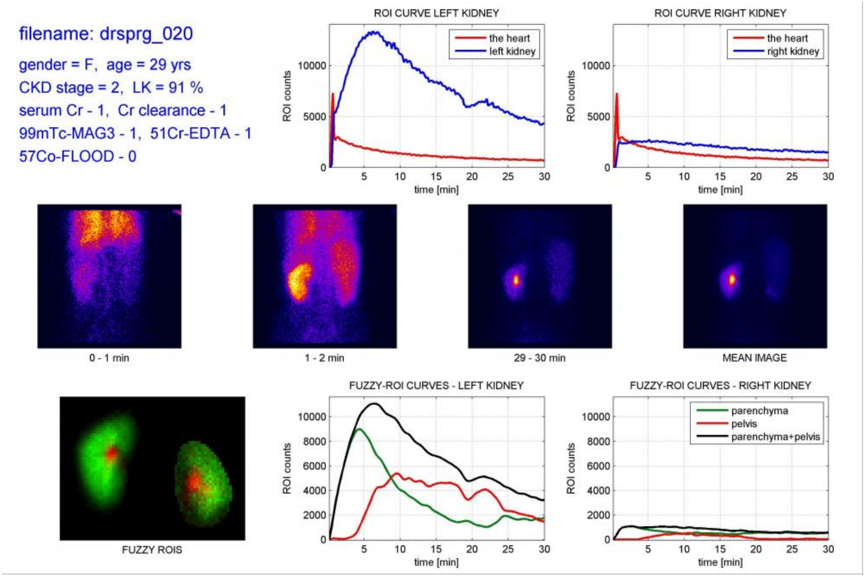
Scintigrafie ledvin

Případy



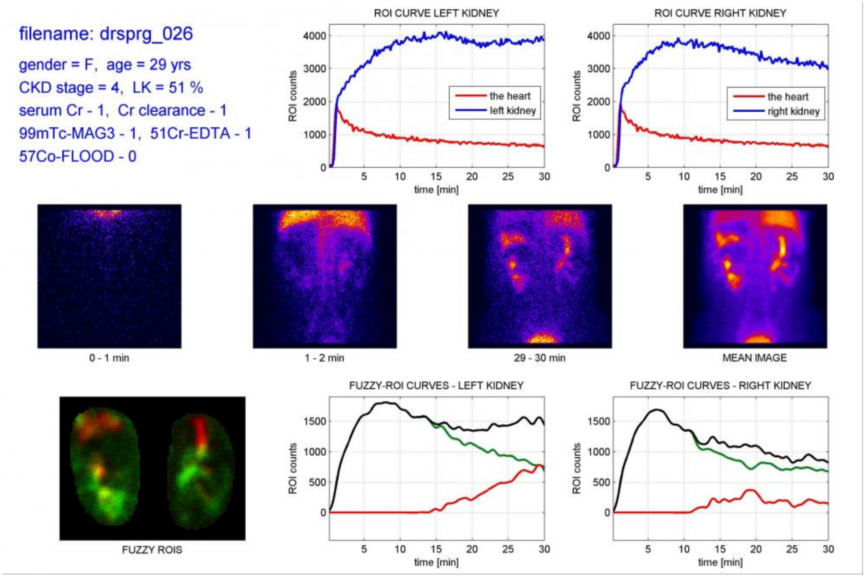
Scintigrafie ledvin

Případy



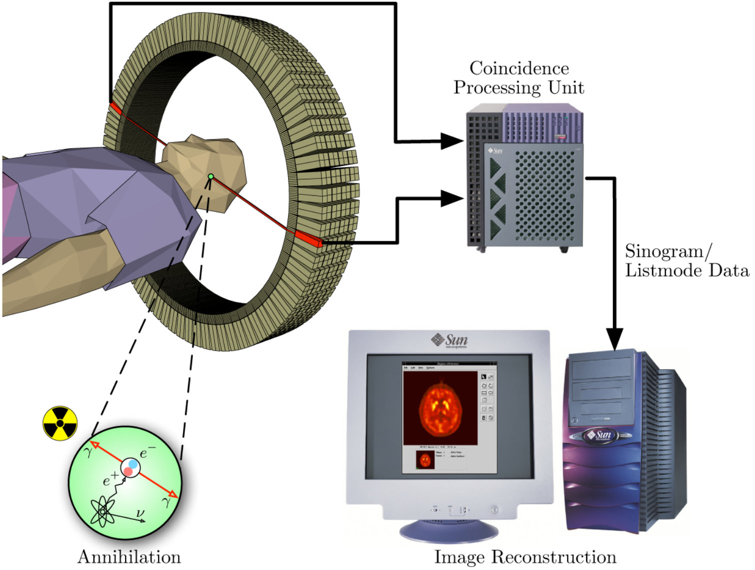
Scintigrafie ledvin

Případy



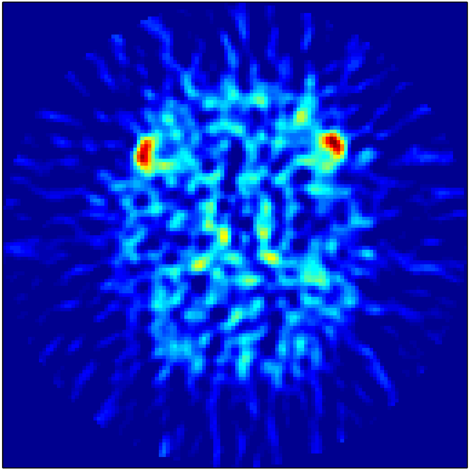
Pozitronová emisní tomografie

Vznik obrazu



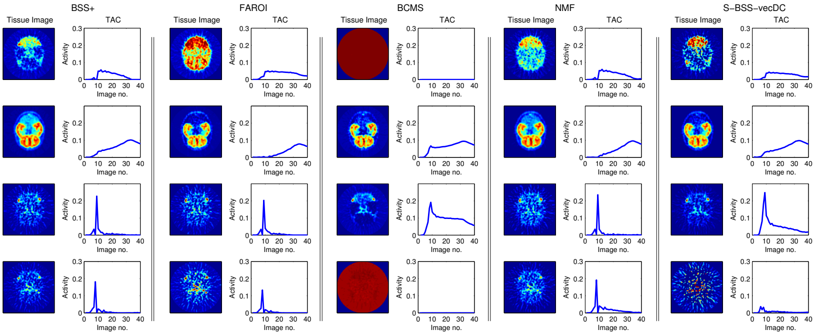
Pozitronová emisní tomografie

1 řez



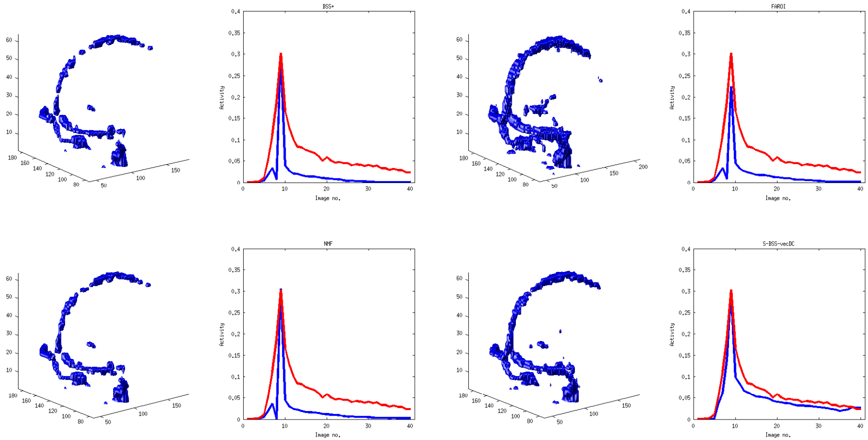
Pozitronová emisní tomografie

Separace 1 řezu



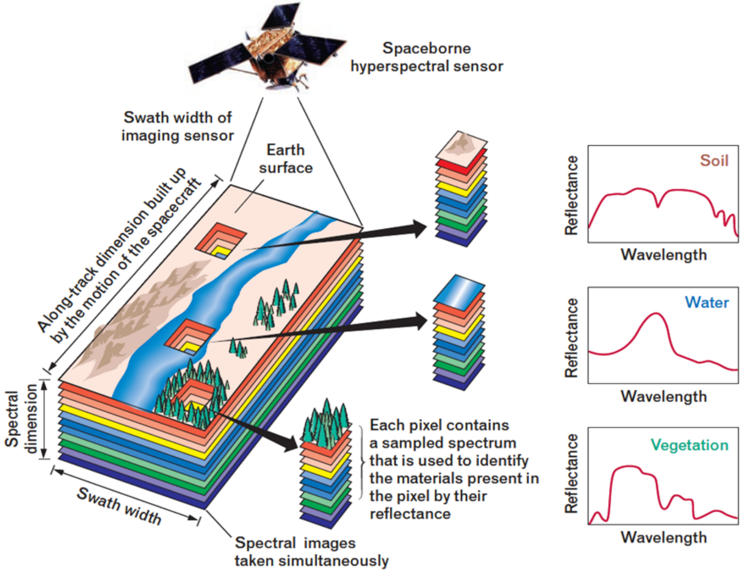
Pozitronová emisní tomografie

Separace objemu



Hyper-spektrální snímkování

Princip



Hyper-spektrální snímkování

Separace