On Validation of Algorithms for Dynamic Medical Data Separation

Ondřej Tichý*

4th year of PGS, email: otichy@utia.cas.cz
Department of Mathematics
Faculty of Nuclear Sciences and Physical Engineering, CTU in Prague

advisor: Václav Šmídl, Institute of Information Theory and Automation, AS CR

Abstract. The problem of dynamic medical image sequence separation is studied. We introduced the state of the art algorithms for medical sequence decomposition together with those that are proposed by us. The validation and the comparison of the algorithms are nontrivial and challenging task. We propose to use a synthetic data where a ground truth is available so it is possible to compute a significant statistics for comparison reason. Moreover, we proposed a comparison on 99 real data from renal scintigraphy where relative renal functions are automatically computed and compared with those obtained by physician.

Keywords: blind source separation, deconvolution, scintigraphy, medical image sequence

Abstrakt. Tento příspěvek se zabývá zpracováním dynamických dat získaných metodou nukleární medicíny, scintigrafie. State of the art algoritmy společně s těmi, které předkládáme my, jsou představeny a diskutovány. Validace a srovnání těchto algoritmů je netriviální úloha. Nejprve navrhujeme srovnání pomocí generovaných dat, kde jsou k dispozici zdrojová data, díky kterým je možno napočítat základní statistické ukazatele výsledků. Předkládáme i srovnání algoritmů pomocí 99 reálných studií ze scintigrafie ledvin. Na těchto studiích automaticky počítáme relativní renální funkci, která může být srovnána s výsledky získanými zkušeným lékařem.

Klíčová slova: slepá separace, dekonvoluce, scintigrafie, obrazová sekvence

1 Introduction

Medical data postprocessing and analysis is important step in diagnostic medical examination. In many imaging modalities such as scintigraphy, the activity of tissues can be observed only via observing of the particles coming from radiopharmaceutical applied to the body. It can be seen the activity during the time in the respective tissues or part of the body using the method; however, several issues must be considered. Since the scintigraphical camera observed the body from one direction, the resulting image pixel is a sum of all underlying tissues. As a result, we observe a superposition of all tissues in respective region of interest (ROI). The task of medical image processing is to reconstruct the original sources of signal, i.e., tissues and their time-activity curves (TACs).

The problem is called blind source separation (BSS) and it is well described in a literature. The current methods used in practice is typically based on manual or semi-manual

*Institute of Information Theory and Automation, Department of Adaptive Systems, AS CR

selection of ROIs of the examined tissues and subtraction of the background activity [13]. More automated models can be based on model of a factor analysis (FA), [8, 7]; however, the solution of the FA is ambiguous and biological meaningfulness is not guaranteed. Other approach is based on modeling of fluid flow using compartment models such as in [5]; however, this could be too strict for biological processes and suffers from artifacts and computation tractability. In recent years, we proposed a number of probabilistic models based on FA model and solved using Variational Bayes (VB) method, [15]. The models are based on modeling both, images and TACs. We proposed (i) a modeling of TACs as results of convolutions of common input function and restricted convolution kernels, [10], (ii) modeling a probability mask on images reflecting that activity do not cover the whole image but only relatively small area [9], and (iii) model combining the advantages from both forcoming model and using the automatic relevance determination (ARD), [1], as a general principle, [11].

This paper summarize mentioned methods and focus on theirs validation and comparison methodology. The issue with validation of models is in no ground truth, no golden standard. Even physician have very different results in scintigraphy on each patient [3] or using different methodology [4]. The synthetic data can be used as an indicator of feasibility but it never reflects the nature. Comparison with physician results can be done but with consideration that manual results suffers from inaccuracy. We propose a comparison on a data from renal scintigraphy where relative renal function is automatically computed.

2 Mathematical Models

We summarize the used mathematical models in our study. All selected methods provides automatic results so they are comparable without biased interpretation.

The objective is to analyze a sequence of n images obtained at time $t = 1, \ldots, n$ and stored in vectors $\mathbf{d}_t$ with pixels stacked columnwise. The number of pixels in each image is p, thus $\mathbf{d}_t \in \mathbf{R}^p$. The important assumption is that every observed image is a linear combination of r factor images, stored in vectors $\mathbf{a}_k \in \mathbf{R}^p$, $k = 1, \ldots, r$, using the same order of pixels as in $\mathbf{d}_t$. The dimensions of the problem are typically ordered as $r < n \ll p$. Each source image has its respective time-activity curve stored in vector $\mathbf{x}_k \in \mathbf{R}^n$, $k = 1, \ldots, r$, $\mathbf{x}_k = [x_{1,k}, \ldots, x_{n,k}]'$, $\mathbf{x}'$ denotes transpose of vector $\mathbf{x}$.

We propose probabilistic formulations of this problem using several probabilistic models. The models are solved using Variational Bayes approximation [15]. The Bayes rule is given as

$$p\left(\theta|D\right) = \frac{p\left(\theta, D\right)}{p\left(D\right)} = \frac{p\left(D|\theta\right) p\left(\theta\right)}{\int p\left(D|\theta\right) p\left(\theta\right) d\theta}, \tag{1}$$

where D are observed data and θ are parameters of $p(D|\theta)$ with prior knowledge $p(\theta)$. Approximation of the Bayes rule via VB approximation can be reached as

$$p(\theta_i) \propto \exp\left(\mathsf{E}_{p(\theta_{/i})}\left(\ln\left(p\left(\theta, D\right)\right)\right)\right), i = 1, \ldots, n \tag{2}$$

Here, $\theta_{/i}$ denotes the complement of θ_i in θ and $\mathsf{E}_{p(\theta)}(g(\theta))$ denotes expected value of function $g(\theta)$ with respect to distribution $p(\theta)$. Equation (2) forms a set of implicit

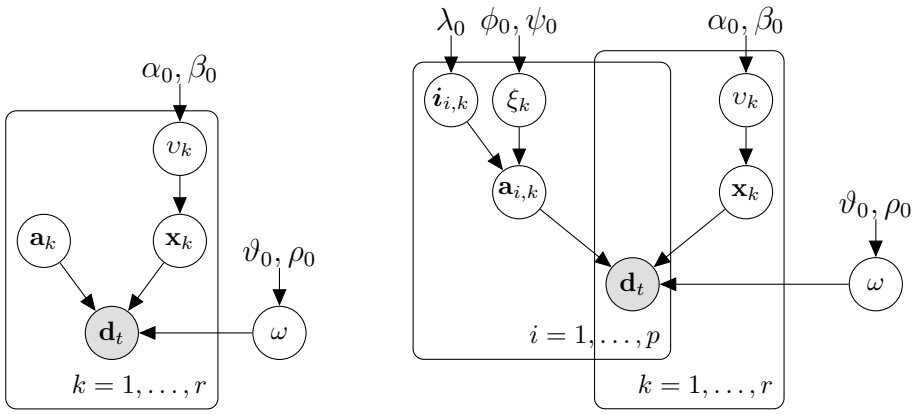

Figure 1: Hierarchical models of BSS+ (left) and FAROI (right).

equations which has to be solved iteratively.

2.1 Blind Source Separation Based on Factor Analysis

The described data sequence can be rewritten in terms of superposition, [7], as

$$\mathbf{d}_t = A\mathbf{x}_t, \tag{3}$$

where A is matrix of tissue images stored $\mathbf{a}_k$ as its columns. It is appropriate to set biologically motivated assumption such as (i) the observed data $\mathbf{d}_t$ are positive, (ii) the expected tissue images $\mathbf{a}_k$ and TACs $\mathbf{x}_k$ are also positive, (iii) the data $\mathbf{d}_t$ is strongly affected by a noise, and (iv) the number of relevant tissues, r, is unknown and should be estimated during the estimative procedure. These assumptions can be rewritten into the probabilistic model as:

$$f(\mathbf{d}_t|A, X, \omega) = \mathrm{tN}(A\mathbf{x}_t, \omega^{-1} I_p \otimes I_n), \tag{4}$$

$$f(\omega) = \mathrm{G}(\vartheta_0, \rho_0), \tag{5}$$

$$f(\mathbf{x}_k|\upsilon_k) = \mathrm{tN}(0_{n,1}, \upsilon_k^{-1} I_n), \tag{6}$$

$$f([\upsilon_1, \ldots, \upsilon_r]) = \prod_{k=1}^{r} \mathrm{G}(\alpha_{k,0}, \beta_{k,0}), \tag{7}$$

$$f(\mathbf{a}_k) = \mathrm{tN}(0_{p,1}, I_p), \tag{8}$$

where tN() denotes truncated normal distribution to positive values, G() denotes gamma distribution, I_p denotes identity matrix of the respective size, and symbol $\otimes$ denotes Kronecker product. The hierarchical model of this model is in Figure 1, left. The model will be denoted as the Blind Source Separation model with positivity constraints (BSS+).

2.2 Regions of Interest in Blind Source Separation

This model adopts the assumptions from section 2.1; however, it reflects the simple fact that tissues do not cover the whole scanned area but only a limited number of pixels.

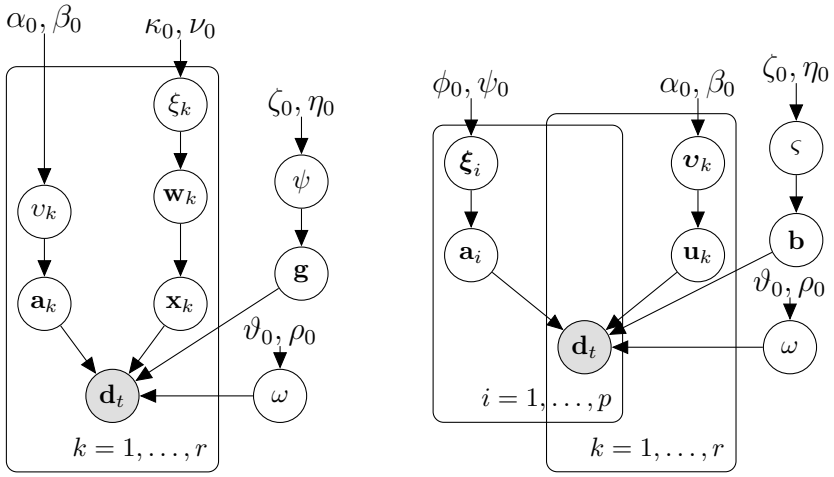

Figure 2: Hierarchical models of CFA (left) and S-BSS-DC (right).

Hence, we proposed a masking of each tissues image using indicator $\mathbf{i}$ of the same size as tissue image, [9]. This affects the model from section 2.1 as follows:

$$f(a_{i,k}|\mathbf{i}_{i,k}, \xi_k) = \mathrm{U}(0,1)^{\mathbf{i}_{i,k}} \mathrm{tN}(0, \xi_k^{-1})^{(1-\mathbf{i}_{i,k})}, \tag{9}$$

$$f(\xi_k) = \mathrm{G}(\phi_{k,0}, \psi_{k,0}), \tag{10}$$

$$f(\mathbf{i}_i, k) = \mathrm{tExp}(\lambda_{ik,0}, \langle 0, 1]), \tag{11}$$

where tExp() is truncated exponential distribution. The hierarchical model of this model is in Figure 1, right. The probabilistic masks $\mathbf{i}_k$ are estimated together with other parameters during the estimative procedure in VB method. This model will be denoted as the FAROI model (Factor Analysis with integrated ROI).

2.3 Convolution in Blind Source Separation

This model reflects the fact that each time-activity curve arise as a convolution of common input function and tissue-specific kernel, [6], such as

$$\mathbf{x}_k = \mathbf{b} * \mathbf{u}_k, \tag{12}$$

where $\mathbf{b} \in \mathbf{R}^{n \times 1}$ is input function, $\mathbf{u}_k \in \mathbf{R}^{n \times 1}$ is convolution kernel of the kth tissue, and $*$ denotes convolution. Both $\mathbf{b}$ and $\mathbf{u}_k$ are modeled as increases as vectors $\mathbf{g}$ and $\mathbf{w}_k$ respectively. This can be rewritten into the probabilistic model as [10, 12]:

$$f(\mathbf{w}_k|\xi_k) = \mathrm{tN}(M_{w_f}, \xi_k^{-1} I_n), \tag{13}$$

$$f(\xi_k) = \mathrm{G}(\kappa_{k,0}, \nu_{k,0}), \tag{14}$$

$$f(\mathbf{g}|\psi) = \mathrm{tN}(0_{n,1}, \psi^{-1} I_n), \tag{15}$$

$$f(\psi) = \mathrm{G}(\zeta_0, \eta_0), \tag{16}$$

where M_{w_f} is obtained in each iteration using clustering algorithm. The hierarchical model of this model is in Figure 2, left. This model will be denoted as the CFA model (Convolution with Factor Analysis).

2.4 Sparsity in Blind Source Separation and Deconvolution

Our latest model adopts ideas from the previous models from sections 2.1, 2.2, and 2.3. However, the assumptions of probabilistic masks, i.e. sparsity of tissue images, and of convolution are not so strict here. We use the Automatic Relevance Determination (ARD) principle, [1], to adopt the sparsity in both, tissue images and convolution kernels respectively. ARD principle is based on observation that variance of the redundant parameter tends to zero in VB solution.

The model can be written as [11]:

$$p(\mathbf{a}_i|\boldsymbol{\xi}_i) = \mathrm{tN}(\mathbf{0}_{1,r}, \mathrm{diag}(\boldsymbol{\xi}_i)^{-1}),\ i = 1, \ldots, p, \tag{17}$$

$$p(\boldsymbol{\xi}_i) = \prod_{k=1}^{r} \mathrm{G}(\phi_{ik,0}, \psi_{ik,0}), \tag{18}$$

$$p(\mathbf{b}|\varsigma) = \mathrm{tN}(0, \varsigma^{-1} I_n), \tag{19}$$

$$p(\varsigma) = \mathrm{G}(\zeta_0, \eta_0), \tag{20}$$

$$p(\mathbf{u}_k|\boldsymbol{\upsilon}_k) = \mathrm{tN}(0_{n,1}, \mathrm{diag}(\boldsymbol{\upsilon}_k)^{-1}), \tag{21}$$

$$p(\upsilon_{j,k}) = \mathrm{G}(\alpha_{jk,0}, \beta_{jk,0}),\ j = 1, \ldots, n, \tag{22}$$

where diag() denotes matrix with argument vector on its diagonal and zeros otherwise. The hierarchical model of this model is in Figure 2, right. This model will be denoted as the S-BSS-DC model (Sparsity in Blind Source Separation and Deconvolution).

2.5 CAM-CM algorithm

A complex compartment model for fMRI tumors imaging was described in [5] based on pharmacokinetic modeling using identifying representative pure pixels from each compartment in corners of cluster simplex. The algorithm is available online and is denoted as the CAM-CM algorithm.

3 Validation on Synthetic Data

Validation on synthetic data is widely used in cases when data with known ground truth are not available. This is the classical issue in the field of dynamic medical imaging including renal scintigraphy.

We propose synthetic data based on [5]. We adopt the image sources and generate our own TACs. It contains 3 image sources modeling the overlapping of all sources pairwise and shared overlap in the center. The size of images is 50×50 pixels; hence, $p = 2500$. The length of the generated sequence is 50 time steps; hence, $n = 50$. The image sources and theirs related TACs are in Figure 3, left.

We run each algorithm on this dataset. The number of expected tissues r is set to 3; hence, $r = 3$. The number of iteration is set to 100 which is reasonable for reach the convergence.

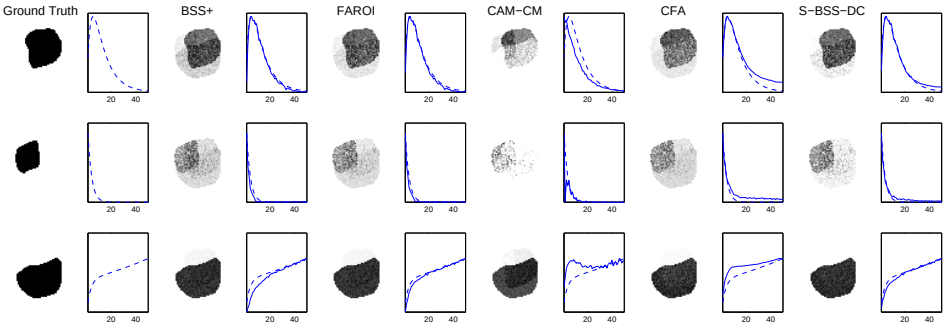

Figure 3: Results from the algorithms on synthetic dataset.

3.1 Results

The results from all algorithms are shown in Figure 3. The ground truth data are on the left and then results from algorithms (from left): BSS+, FAROI, CAM-CM, CFA, S-BSS-DC. The image sources are in the first column and the TACs are in the second. The dashed lines denotes ground truth and the full lines TACs estimated by algorithms. Note that the results are normed with respect to the activities of ground truth; hence, we study shapes, not amplitudes.

Since we have ground truth TACs, we can compute Mean square error (MSE), Mean absolute error (MAE), and Maximum error. The results is shown in Table 1. It can be seen that the computed statistics have significant explanatory value with S-BSS-DC algorithm being the best.

4 Validation on Real Data

Validation on real data is much more challenge then validation on synthetic data. Generally, we have no ground truth; hence, we can not compare results from algorithms with it. In renal scintigraphy, we have two main choices.

First, skilled operator can manually select regions contained each tissue and plot activities of the selected regions. Note that overlaps must be carefully considered. This task is extremely subjective and using of these types ground truths should be done with respect of this fact.

Second, diagnostic coefficients may be computed by a physician from the data. In renal scintigraphy, this task is very subjective too [3]. We are focused on computing of relative renal function (RRF) [2] which is a percentage of function of the left kidney and the right kidney. The RRF is estimated from the sum of activity in the left (L) and in the right (R) parenchyma during the uptake time. Then, $RRF_L = \frac{L}{L+R} \times 100\ \%$ and RRF_R can be computed analogically, both weighted by their time activity curves. Historically, the activity is taken only from the uptake time, the time when kidney accumulates activity only.

We propose a comparison on dataset [14] where RRF is computed by experienced physician. We select the sequences where both kidneys are present, i.e. 99 cases. The five

Mean Square Error					
	Algorithm				
Tissue no.	BSS+	FAROI	CAM-CM	CFA	S-BSS-DC
1	0.0061	0.0033	0.05	0.0135	0.0033
2	0.0047	0.0037	0.0205	0.0056	0.002
3	0.0455	0.0133	0.1420	0.0643	0.0095

Mean Absolute Error					
	Algorithm				
Tissue no.	BSS+	FAROI	CAM-CM	CFA	S-BSS-DC
1	0.0432	0.0416	0.1515	0.1017	0.0429
2	0.0321	0.0285	0.0363	0.0716	0.0374
3	0.1448	0.0737	0.2663	0.2208	0.0656

Maximum Error					
	Algorithm				
Tissue no.	BSS+	FAROI	CAM-CM	CFA	S-BSS-DC
1	0.4595	0.2827	0.7897	0.1684	0.2385
2	0.2651	0.2444	0.9516	0.1190	0.1589
3	0.5489	0.3569	0.8527	0.4362	0.2519

Table 1: Comparison of the algorithms on synthetic dataset is shown. The MSE, Mean error, and Maximum error are computed.

RRF estimation			
Algorithm	$<5\%$	$<10\%$	$\geqq 10\%$
BSS+	57.6%	78.8%	21.2%
FAROI	58.6%	83.8%	16.2%
CAM-CM	47.9%	63.8%	36.2%
CFA	59.6%	82.8%	17.2%
S-BSS-DC	68.7%	86.9%	13.1%

Table 2: Cumulative histogram of RRFs.

described algorithms will be compared via difference of their results of RRF computation from those provided by the experienced physician as a reference value. We will consider the automatic method that is closer to his results to be better [12].

4.1 Results

The results will be compared for BSS+, FAROI, CAM-CM, CFA, and S-BSS-DC algorithms. We use comparison over the cumulative histogram, see Table 2.

The results suggest the similar conclusion as results on synthetic data. The S-BSS-DC algorithm seems to outperform the other algorithms.

5 Conclusion

We study possibilities of comparison of algorithms for blind source separation of medical data sequence in this paper. We revise possible algorithms based on probabilistic modeling from base to more complex ones with additional assumptions. We discuss the way how to compare a performance of the algorithms. The synthetic data is proposed which provide a ground truth. It is possible to compute significant statistics using comparison of results with this ground truth. Comparison of the algorithms on real data from renal scintigraphy is more challenging task since no ground truth is available. We propose a comparison based on relative renal functions computation and comparison with those obtained from experienced physician.

We shown that the S-BSS-DC algorithm outperform other proposed algorithms in both synthetic and real data. In a future, we will prepare a comparison on directly manually selected tissue-images and related time-activity curves. It should prove the feasibility of algorithms in the best imaginable way.

References

[1] C. Bishop and M. Tipping. Variational relevance vector machines. In 'Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence', 46–53, (2000).

[2] M. Blaufox, M. Aurell, B. Bubeck, E. Fommei, A. Piepsz, C. Russell, A. Taylor, H. Thomsen, D. Volterrani, et al. *Report of the radionuclides in nephrourology com-*

mittee on renal clearance. Journal of nuclear medicine: official publication, Society of Nuclear Medicine **37** (1996), 1883.

[3] A. Brink, M. Šámal, and M. Mann. *The reproducibility of measurements of differential renal function in paediatric 99mtc-mag3 renography.* Nuclear medicine communications **33** (2012), 824–831.

[4] M. Caglar, G. Gedik, and E. Karabulut. *Differential renal function estimation by dynamic renal scintigraphy: influence of background definition and radiopharmaceutical.* Nuclear medicine communications **29** (2008), 1002.

[5] L. Chen, P. Choyke, T. Chan, C. Chi, G. Wang, and Y. Wang. *Tissue-specific compartmental analysis for dynamic contrast-enhanced mr imaging of complex tumors.* IEEE Transactions on Medical Imaging **30** (2011), 2044–2058.

[6] A. Kuruc, J. Caldicott, and S. Treves. *Improved Deconvolution Technique for the Calculation of Renal Retention Functions.* COMP. AND BIOMED. RES. **15** (1982), 46–56.

[7] J. Miskin. *Ensemble learning for independent component analysis.* PhD thesis, University of Cambridge, (2000).

[8] M. Šámal, C. Nimmon, K. Britton, and H. Bergmann. *Relative renal uptake and transit time measurements using functional factor images and fuzzy regions of interest.* European Journal of Nuclear Medicine and Molecular Imaging **25** (1997), 48–54.

[9] V. Šmídl and O. Tichý. Automatic Regions of Interest in Factor Analysis for Dynamic Medical Imaging. In '2012 IEEE International Symposium on Biomedical Imaging (ISBI)'. IEEE, (2012).

[10] V. Šmídl, O. Tichý, and M. Šámal. Factor Analysis of Scintigraphic Image Sequences with Integrated Convolution Model of Factor Curves. In 'Proceedings of the second international conference on Computational Bioscience'. IASTED, (2011).

[11] V. Šmídl and O. Tichý. *Sparsity in Bayesian Blind Source Separation and Deconvolution.* In 'Machine Learning and Knowledge Discovery in Databases', H. Blockeel et al., (ed.), volume 8189 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg (2013), 548–563.

[12] O. Tichý, V. Šmídl, and M. Šámal. Model-based Extraction of Input and Organ Functions in Dynamic Medical Imaging. In 'ECCOMAS Conference on Computational Vision and Medical Image Processing (VipImage 2013)'. Taylor and Francis, (2013). accepted.

[13] Y. Tomaru, T. Inoue, N. Oriuchi, K. Takahashi, and K. Endo. *Semi-automated renal region of interest selection method using the double-threshold technique: inter-operator variability in quantitating 99m tc-mag3 renal uptake.* European Journal of Nuclear Medicine and Molecular Imaging **25** (1997), 55–59.

[14] VFN Praha. Database of dynamic renal scintigraphy, (September 2013).

[15] V. Šmídl and A. Quinn. *The Variational Bayes Method in Signal Processing.* Springer, (2006).