Abstract—In this paper, we propose a novel method for single-channel blind separation of non-overlapped sources and, to the best of our knowledge, apply it for the first time to automatic segmentation of lung tumors in Positron Emission Tomography (PET) images. Our approach first converts 3D PET image into a pseudo multichannel image. Afterwards, regularization free sparseness constrained nonnegative matrix factorization is used to separate tumor from other tissues. By using complexity based criterion, we select tumor component as the one with minimal complexity. We have compared the proposed method with threshold on 40% and 50% maximum standardized uptake value (SUV), graph cuts (GC), random walks (RW) and affinity propagation (AP) algorithms on 18 non-small cell lung cancer datasets with respect to ground truth provided by two radiologists. Dice similarity coefficient averaged with respect to two ground truths is: 0.78±0.12 by the proposed algorithm, 0.78±0.01 by GC, 0.77±0.13 by AP, 0.77±0.07 by RW, and 0.75±0.13 by 50% maximum SUV threshold. Since the proposed method achieved performance comparable with interactive methods, considering the unique challenges of lung tumor segmentation from PET images, our findings support possibility of using our fully automated method in routine clinics. The source codes will be available at www.mipav.net/English/research/research.html.

Index Terms—Single-channel blind source separation, nonnegative matrix factorization, sparseness, lung tumor delineation, positron emission tomography (PET).

I. INTRODUCTION

LUng tumor is the main cause of cancer death in men and the second main cause of cancer death in women [1]. The overall five-year survival rate of lung cancer for all stages is only approximately 15%. However, if the primary tumor is resectable, survival rate is improved [2], [3]. Accurate staging is critical to characterize tumors such as understanding the stage of the tumor or its suitability for resection. To this end, positron emission tomography (PET)-computed tomography (CT) are widely used [4]. CT provides anatomical information but it lacks metabolic information [5]. PET yields increased uptake of fluorodeoxyglucose (FDG) in tumor tissues relative to surrounding normal tissues. This is typical for non-small cell lung cancer (NSCLC) [3], [6]. On the other hand, NSCLC tumors show high regulation of glucose metabolism such that labeling glucose with FDG enables detectability of NSCLC [7]. Increased FDG uptake is often referred to as a "hot spot". Due to limited spatial resolution of PET images, the boundary of a tumor lesion usually appears fuzzy and indistinct. That is why tumor delineation using sole PET images is still considered suboptimal [8], [9], [10]. In this regard, Figure 1 illustrates difficulties associated with segmentation of lung tumor with complex shape and fuzzy boundary.

Fig. 1. Comparative results for a slice of PET study. Top row from left to right: input PET image, ground truth 1, ground truth 2, segmentation result obtained by random walks, affinity propagation, graph cuts and thresholding based on 50% of maximum tumor FDG uptake.

Nevertheless, PET can detect sites of disease in non-enlarged lymph nodes within the thorax and in structures outside the thorax, which may not appear abnormal in CT images [3]. There has been increased research undertaken to improve accuracy of lung tumor delineation using PET [9], [11], [12]. One way to improve boundary definition is to combine PET and CT information and jointly segment boundaries from CT and PET [3], [8], [9], [10], [13]. Compared with segmentation results obtained by random walks, affinity propagation, graph cuts (GC) [15], affinity propagation (AP) [16], [17] and thresholding based on 50% of the maximum tumor FDG uptake (TH50), the
method proposed herein achieves the same performance as GC and better performance than RW which are interactive methods. It is also better than AP method that requires postprocessing. Proposed method yields comparable results as RW and GC co-segmentation method in [13] that uses both PET and CT modalities. However, while method in [13] is interactive, proposed method is fully automatic.

There is no PET image segmentation method that is optimal for all applications [9]. Methods that segment PET image only are divided in the following groups: manual segmentation, thresholding-based, region-based, stochastic and learning-based [9]. Manual segmentation is inefficient and highly subjective with substantial intra- and inter-operator disagreement rates [18], [19], [9]. Thresholding is widely used in PET images due to its simplicity and good contrast of the PET images. Because of the large variability of pathologies, low resolution and uncertainty in fuzzy object boundaries, there is no consensus on the selection of a thresholding level [9]. The range of suggested thresholds is commonly between 40%-50% of the maximum tumor FDG uptake [19]-[22]. However, not always primary NSCLCs display highest FDG uptake and benign conditions such as pneumonia can also have high FDG uptake [6]. That is why a single threshold method is not considered satisfactory [23], [24], [9]. Many adaptive thresholding methods [25], [26], [9] and iterative thresholding strategies [27], [9] were developed to address the challenges of single thresholding methods. Although improvements were obtained, none of these strategies were generalized due to the need for developing more advanced and robust techniques.

Stochastic methods, as an alternative and more powerful methods, exploit differences between uptake regions and surrounding tissues statistically. In this category, three groups of methods are distinguished: mixture models based methods [28], [29], fuzzy locally adaptive Bayesian method [29], and clustering/classification of PET image intensities methods [16], [17], [30], [31]. Since clustering methods are grouping voxels according to chosen similarity metric, they are geometry independent and are, therefore, particularly useful when the shapes of the uptakes regions are non-convex with a heterogeneous background. Non-convex regions are quite common and that is cause of growing interest in the use of clustering methods to segment complex shapes uptake regions [9]. In particular, recent AP method was demonstrated to address difficult problem related to segmentation of the multi-focal uptake pattern in the PET image [16], [17]. The method proposed herein belongs to the class of clustering methods.

The region-based segmentation methods are using homogeneity of the intensities in the image for determining object boundaries [9]. In this group, graph-based methods were shown particularly more suitable for segmentation of PET images than other methods due to incorporation of a priori knowledge into segmentation process. That is accomplished by using foreground and background seeds specified by the user, or automatically, to locate objects in the image [9]. The two most prominent graph-based methods used for PET image segmentation are GC [15] and RW [16]. The advantage of RW over GC is to be more robust against noise and weak boundaries. Despite automated seeding process [10], it is believed that human incorporation may still be necessary for some extreme cases [9]. A comprehensive review for PET segmentation methods can be found in [9].

Herein, we propose a method for automatic delineation of lung tumor from PET images. That is motivated with efforts to develop automatic segmentation methods as a substitute for time-consuming manual delineation performed by radiologists [32]. It is also motivated by the clinical cases which may not appear abnormal in CT images and only PET can detect such sites [3]. Furthermore, it is motivated by the non-convex shape of NSCLC what is hard to segment. As opposed to multivariate data analysis methods such as principal component analysis (PCA) [33] and independent component analysis (ICA) [34] which yield separated components with physically not interpretable mixed signs, the proposed method yields physically meaningful nonnegative separated components. Compared with sparseness regularized nonnegative matrix factorization (NMF) such as [35], the proposed method is free of regularization constant and that is important for automatic segmentation. The same comment applies to subspace clustering algorithms such as normalized cut (Ncut) [36], since Ncut can be implemented through orthogonality constrained NMF [37]. The automatic segmentation method proposed by us is based on blind source separation (BSS) approach [38]. It performs decomposition of 3D PET images into tumor and non-tumor parts. Since BSS is multivariate data analysis methodology, it assumes availability of a multichannel image. Because PET is a single-channel image, we first derive nonlinear transform that maps a single-channel PET image into pseudo multichannel. Afterwards, sparseness constrained nonnegative factorization of the pseudo multichannel image is performed to separate lung tumor from non-tumor tissues. In this regard, the proposed method belongs to class of clustering methods, where PET image is partitioned into sets of pixels (segments) corresponding to distinct objects (organs, tissues, etc.) based on their uptake level. Objects with distinct uptakes are represented by distinct segments. Lung tumor-related segment is identified automatically by means of criterion that combines measures of smoothness and entropy. Segmented image of the lung tumor is obtained through thresholding/binarization of the separated tumor component. Since lung tumor is separated from non-tumor tissues, a binarization threshold can be set to a low value to remove remnant parts of the non-tumor tissues, which are still present in separated tumor component. In particular, we demonstrate through leave-one-out cross validation on 18 NSCLCs patient datasets with manual delineation by two radiologists, that optimal values exist for both order of the nonlinear mapping transform and binarization threshold. Learned parameters can be set to predefined values, which makes the proposed 3D lung tumor delineation method automatic.

The remainder of the paper is organized as follows. Section II presents an overview of BSS approaches to multichannel image decomposition as well as overview of single-channel BSS algorithms. A novel single-channel BSS algorithm for
automatic 3D PET image decomposition is presented in Section III. Experimental results, in particular parameterization and comparative performance analysis of the proposed algorithm, are presented in Section IV. Discussion is presented in Section V, while conclusions are given in Section VI.

II. BACKGROUND AND RELATED WORK

A. BSS Approach to Multichannel Image Decomposition

The BSS approach to automatic image decomposition is based on multivariate data analysis [38]. We implicitly assume a multichannel image: \( X \in R^{N \times T} \) comprised of \( N \) vectorized intensity images, also known as (a.k.a.) mixtures in the BSS vocabulary, and \( T \) pixels (voxels), a.k.a. samples in the BSS vocabulary. BSS decomposition is based on multiplicative linear mixture model (LMM):

\[
X = AS,
\]

(1)

where \( A \in R^{M \times N} \) stands for a basis (or mixing) matrix and \( S \in R^{N \times T} \) stands for a matrix of encoding coefficients. \( M \) rows of \( S \) matrix, a.k.a. sources in the BSS vocabulary, are interpreted as partitions of the image occupied by particular tissue/organ. Thereby, presence of source \( m \) at pixel \( t \) is regulated by the encoding coefficients \( \{ s_{mt} \in R_{N \times T} \} \). BSS decomposition effectively partitions an image into its constituent components. The term blind implies that decomposition of a multichannel image \( X \) into the basis matrix \( A \) and the source matrix \( S \) is based on the multichannel image \( X \) only. Linear representation (1) is customarily used to model a multichannel image in multispectral/hyperspectral imaging [39], [40], [41], [42], magnetic resonance imaging [43], and multi-phase CT imaging [44], [45]. Given a linear representation (1), and depending on constraints imposed on \( A \) and \( S \), various methods are used to perform BSS-based image decomposition. PCA [33] constraints sources to be uncorrelated and Gaussian. ICA constraints sources to be statistically independent and non-Gaussian (at most one source is allowed to be Gaussian) [34], [46], [43], [44], [41]. PCA and ICA address over- and determined BSS problems in which number of sources cannot exceed number of mixtures available. Sparse component analysis (SCA) [47], [42], [48] constraints sources to be sparse. Thus, underdetermined BSS problems, characterized with more sources than mixtures, can be solved by SCA. Nonnegative matrix factorization (NMF) [49], [50] constraints sources to be nonnegative. As shown in [51], non-negativity is a natural constraint when applying factorization techniques to image related problems. NMF can also be combined with sparseness constraint [52], [35], [53] which enables solution of nonnegative underdetermined BSS problems.

B. Single-channel BSS

LMM-based representation of a single-channel image on the pixel-level becomes:

\[
x_t = \sum_{m=1}^{M} a_{mt} s_{mt}, \quad \forall t = 1,...,T
\]

(2)

where \( \{a_{mt}\} \) represent the mixing coefficients and \( \{s_{mt}\} \) represent the encoding coefficients. Due to requirements of multiple mixtures, BSS methods are not applicable to single-channel. Single-channel BSS can be seen as limited (degenerative) case of underdetermined BSS where only one mixture is available [54]. In this regard, single-channel BSS is a highly ill-posed problem and hard constraints have to be imposed on source signals to enable their separation from a single-channel mixture. A single-channel mixture has to be transformed into pseudo multichannel mixture before some existing BSS algorithm such as ICA or NMF can be used. Several techniques were developed to convert single-channel mixture to pseudo multichannel [54]-[64] and, afterwards, to use some existing multivariate data analysis algorithms to perform BSS. For instance, Davies and James [54] assumed that source signals have disjoint spectral support and their method partitions single-channel time series to yield pseudo multichannel mixture, where an ICA algorithm was then applied to extract sources. Mijović et. al. [56] used empirical mode decomposition [65] to decompose single-channel mixture into intrinsic mode functions (IMFs) that stand for pseudo multichannel mixture. In order to be separated by ICA algorithms, sources of interest are required to be IMFs. In [57], wavelet transform is used to generate pseudo multichannel mixture from single-channel one. Thereby, mother wavelets have to be non-orthogonal and have to correspond to shapes of the sources we are interested in. Thus, this wavelet-ICA method is applicable to the separation of the specific source signals such as vibration signals [57], [58]. Many of single-channel BSS algorithms are derived for separation of acoustic signals by factorizing nonnegative spectrogram (magnitude of the short time Fourier transform) [59]-[63]. Ma et. al. [64] applied singular spectrum analysis (SSA) technique [66] to convert single-channel to pseudo multichannel mixture assuming that scalar time series is stationary. The BSS algorithm based on minimization of mutual information is applied to pseudo multi-channel mixture to separate source signals. The critical step in the SSA algorithm is determination of the length of the window in partitioning of scalar time series.

III. THE THRESHOLDED NONLINEAR NONNEGATIVE MATRIX UNDERAPPROXIMATION ALGORITHM

In this paper we propose a single-channel BSS algorithm for decomposition of vectorized single-channel intensity image into its constituent components. The flow chart diagram of proposed thresholded nonlinear nonnegative matrix underapproximation (TNNMU) algorithm is shown in Figure 2.
It is assumed by the TNNMU algorithm that components have dominantly disjoint spatial support. Within the context of the present paper, we are interested in decomposition of a 3D PET image tensor $X \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ comprised of $I_3$ slices of the size $I_1 \times I_2$ pixels. Thus, a single-channel PET image is obtained by unfolding 3D tensor $X$ into row vector $x \in \mathbb{R}^{I_1 I_2 I_3}$ comprised of $T=I_1 I_2 I_3$ pixels. Thereby, $x_t$ denotes intensity value of the pixel $t=1,...,T$. Within the context of the present paper we assume:

A1) $s_m s_n = \delta(n-m) \quad \forall m,n=1,...,M \quad \forall t=1,...,T$

A1) implies that each pixel (voxel) is dominantly occupied by one organ/tissue as well as that the encoding coefficients have binary $\{0, 1\}$ values. A1) is justified because we are interested in solving image segmentation problem. In this case, $s_m=0$ implies that source $m$ is not present at pixel $t$ while $s_m=1$ implies that source $m$ is present at pixel $t$. In case of PET image, coefficients $\{a_m\}_{m=1}^M$ correspond to uptakes of organs/tissues present in the PET image. Thus, it is justified to assume that $\{a_m\}_{m=1}^M$ are distinct:

A2) $a_m \neq a_n \quad \forall m,n=1,...,M \quad \text{and} \quad m \neq n$.

A. Nonlinear mapping to pseudo-multichannel PET image

We propose a pixel-wise nonlinear mapping of order $D$:

$\{x_t \mapsto \phi_d(x_t) \in \mathbb{R}^{(D+1)I_3}\}_{t=1}^T$

such that:

$\phi_d(x_t) = \begin{bmatrix} 1 & x_t & x_t^2 & \cdots & x_t^D \end{bmatrix}^T$

By using (2) and using multinomial theorem, we obtain:

$x'_t = \left(\sum_{a=0}^M a_m x_t^a\right)^D$

$= \sum_{d=0}^D \frac{d!}{d_1! \cdots d_M!} \prod_{i=1}^M (a_m x_t^a)^{d_i} \quad d_0,...,D$

(4)

Now by using A1), equation (4) is simplified as:

$x'_d = \left(\sum_{a=0}^M a_m s_m x_t^a\right)^d = \sum_{a=0}^M a'_m s_m x_t^a \quad d = 0,...,D$

(5)

Thus, using equation (3) and equation (5) we obtain $\forall t=1,...,T$:

$\phi_d(x_t) \approx \begin{bmatrix} 1 & x_t & x_t^2 & \cdots & x_t^D \end{bmatrix}^T$

$= \begin{bmatrix} 1 & 1 & a_1 & a_2 & \cdots & a_M \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_M \end{bmatrix}$

(6)

or on the matrix level:

$\phi_d(x) \approx VS$

(7)

where $V$ stands for mixing or basis matrix in mapped space and $S$ stands for matrix of encoding coefficients in mapped space and it is, due to A1), equivalent to the encoding coefficients in (2). As an example, we show in Figure 3 pseudo multi-channel image in log10 scale generated by (3) for $D=7$.

B. Nonnegative matrix underapproximation based decomposition of pseudo-multichannel PET image

The basis matrix $V$ is a special matrix known as Vandermonde matrix [67]. The important property of the Vandermonde matrix $V$ is that under assumption A2) and for $D=M-1$ it is nonsingular, while for $D>M-1$ the matrix $V$ is a full column rank [68]. Thus, the LMM (7) is identifiable. Regarding the order of the mapping $D$, it does not have to be very large since in case of PET imaging, the number of organs/tissues $M$
that are present in the image is small. Then, sparseness and nonnegativity constrained BSS algorithm, such as NMF algorithm [51], can be used to separate $S$ from (7) even if the LMM is possibly underdetermined.

Nonnegativity and sparseness constrained factorization is applied to solve BSS problem implied by equation (7). The nonnegative matrix underapproximation (NNU) algorithm [52], with MATLAB code available at [69], has been used for nonnegative matrix underapproximation (NMU) algorithm applied to solve BSS problem implied by equation (7). The Nonnegativity and sparseness constrained factorization is performed on the residue matrix $\phi_0(x) \leftarrow \phi_0(x) - v_i s_i$. Herein, $v_i$ denotes column vector and $s_i$ denotes row vector. To preserve non-negativity of $\phi_0(x)$, an underapproximation constraint is imposed on $V$ and $S$: $VS \leq \phi_0(x)$. This constraint yields localized parts-based decomposition, where different basis elements $\{v_n\}_{n=1}^{M}$ describe disjoint parts of the input data $\phi_0(x)$. It has been proven in theorem 1 in [52] that, due to underapproximation constraint, number of non-zero entries of $V$ and $S$ is less than number of non-zero entries of $\phi_0(x)$. That, as opposed to other sparseness constrained NMF algorithms, enables sparseness constrained factorization without regularization constant and that would require a tuning procedure. When performing NMU-based factorization of matrix $\phi_0(x)$ in (7), the unknown number of sources $M$ needs to be set to $M_{NMU}$ for the NMU algorithm as an input. Our strategy is to set $M=M+1$ if, possibly, source related to lung tumor component is not extracted, $M$ can be increased. However, our experiments executed on all clinical data show that $D=8$ and $M_{NMU}+1$ is sufficient. Thus, estimate of the sources $S$ is obtained as:

$$S = NMU (\phi_0(x), M+1)$$

(8)

where, $M+1$ denotes number of sources (tissue components) to be extracted.

C. Selection of tumor component

After $M$ source components are estimated by (8), it is necessary to automatically select the lung tumor component. For this reason, we formulate criterion that combines measures of smoothness and entropy. Smoothness measure reflects the fact that pixels belonging to the lung tumor component are locally more correlated than it is the case with other components. Thus, vectorized lung tumor component will be more predictable than vectorized non-tumor components. For the predictability (smoothness) measure we use [70], [71]:

$$F(\{s_\alpha(t_k)\}) = \log \frac{W_\alpha}{U_\alpha} = \log \frac{\sum_{k=1}^{k_{max}} (\sum_{l=1}^{L} \frac{(s_\alpha(t_k) - \hat{s}_\alpha(t_k))}{L})^2}{\sum_{k=1}^{k_{max}} (s_\alpha(t_k) - \hat{s}_\alpha(t_k))^2}$$

(9)

where $\hat{s}_\alpha(t_k)$ is the value of separated source component $m$ at pixel $t_k$, $W_\alpha$ reflects the extent to which $\hat{s}_\alpha(t_k)$ is predicted by a long term moving average $\hat{s}_\alpha(t_k)$ of values in $\hat{s}_\alpha(t_k)$, and $U_\alpha$ reflects the extent to which $\hat{s}_\alpha(t_k)$ is predicted by a short term moving average $\hat{s}_\alpha(t_k)$ of values in $\hat{s}_\alpha(t_k)$. The predicted values $\hat{s}_\alpha(t_k)$ and $\hat{s}_\alpha(t_k)$ of $\hat{s}_\alpha(t_k)$ are exponentially weighted sums of signal values measured up to shift $(k_{max}-1)$, where $k_{max}$ stands for maximal shift, see also eq. (1.1) and (1.2) in [70]. The value of $k_{max}$ is not critical. That is, the lung tumor component is identified for $k_{max} \in \{1, 2, 3, 4, 5\}$. In reported experimental results we have set $k_{max}=5$ pixels. The predictability measures $\{F(\hat{s}_\alpha(t_k))\}_{s_{\alpha=1}}^{M}$ estimated by (9) are scaled to interval $[0, 1]$ through division of estimated values by the maximal one. It is expected that the tumor related component $\hat{s}_\alpha$ has the highest value of predictability. The estimated amounts of entropy, $H(\hat{s}_\alpha)$, contained in each separated component $\hat{s}_\alpha$ are also scaled to interval $[0, 1]$ in a way analogous to predictability measure. Entropy is computed with the MATLAB command entropy. It is defined as

$$H(\hat{s}_\alpha) = -\sum_{n=1}^{M_{max}} p_n(\hat{s}_\alpha) \log_2(p_n(\hat{s}_\alpha))$$

where, $\{p_n(\hat{s}_\alpha)\}_{s_{\alpha=1}}^{M}$ are probabilities associated with the bins of the histogram of $\hat{s}_\alpha$. Thus, the tumor related component $\hat{s}_\alpha$ with "almost" constant values is expected to have the lowest value of entropy. Thus, we formulate the lung tumor component selection criterion as complexity measure:

$$m^* = \arg \min_{m} \left[ \alpha \left(1-F(\hat{s}_\alpha(t_k))\right) + (1-\alpha)H(\hat{s}_\alpha) \right]$$

(10)

such that $\alpha \in [0.4, 0.5]$. The selection process is illustrated in Figure 4, whereat the minimum of criterion (10) corresponds with the component 2 that also has the largest Dice coefficient. As can be seen either predictability or entropy alone are not enough to select the tumor component.

![Fig. 4. The lung tumor component selection process. Component with the lowest complexity, eq. (10), has largest Dice coefficient.](image-url)
D. Thresholding of the tumor component

In order to segment the lung tumor, corresponding source component extracted by the NMU algorithm has to be binarized. We obtain segmented lung tumor by thresholding each entry of \( \hat{s}_m \) according to:

\[
\tilde{s}_{m,t} = \begin{cases} 
1 & \text{if } \tilde{s}_m \geq \tau (s_{m}^{\text{max}} - s_{m}^{\text{min}}) \\
0 & \text{otherwise}
\end{cases} \quad \forall t = 1, \ldots, T 
\tag{11}
\]

where \( \tau \in [0, 1] \) stands for threshold, \( s_{m}^{\text{max}} \) and \( s_{m}^{\text{min}} \) stand for maximal value and minimal value of \( \tilde{s}_m \), respectively. Thresholding (11) reflects the fact that the NMU-based decomposition (8) does not yield perfect separation of components \( \{ \tilde{s}_m \}_{m=1}^M \). Thus, remnant parts of non-tumor tissues will, up to some extent, still be present in \( \tilde{s}_m \). However, due to NMU-based separation, their contribution is expected to be low. Consequently, the threshold value \( \tau \) in (11) ought to be low and that should enable delineation of lung tumor in the region of low uptake value. Indeed, cross-validation experiments executed on clinical data in Section IV yield \( \tau = 0.15 \) when \( D = 8 \). The proposed TNNMU algorithm is summarized in the Algorithm 1 as below.

**Algorithm 1.** The TNNMU algorithm based single-channel PET image decomposition. 

**Inputs:** \( X \in R_{I_1 \times I_2 \times I_3}^{I_1 \times I_2 \times I_3} \) 3D PET image comprised of \( I_1 \) slices of the size \( I_1 \times I_2 \) pixels (voxels). Order of the mapping in (3): \( D = 8 \). Segmentation threshold in (11): \( \tau = 0.15 \).

**Step 1.** Unfold 3D PET image \( X \in R_{I_1 \times I_2 \times I_3}^{I_1 \times I_2 \times I_3} \) into vector \( x \in R_{3D}^{I_1 \times I_2 \times I_3} \), where \( T = I_1 I_2 I_3 \).

**Step 2.** Perform nonlinear mapping of single-channel PET image according to (3)/(6).

**Step 3.** Perform the NMU decomposition according to (8).

**Step 4.** Select the lung tumor related source component according to (10).

**Step 5.** Segment selected lung tumor component \( \tilde{s}_{m,t} \) according to (11).

**Step 6.** Fold \( \tilde{s}_{m,t} \) into \( \hat{s}_{m,t} \in R_{I_1 \times I_2 \times I_3}^{I_1 \times I_2 \times I_3} \).

**Output:** \( \hat{s}_{m,t} \) segmented 3D lung tumor component.

IV. EXPERIMENTAL RESULTS

A. Data

For validation we used 18 PET-CT studies from patients with NSCLC. Patients were scanned on GE Discovery STE 16 PET-CT scanner. PET images were reconstructed into 120 \( \times \) 120 matrices with the voxel size of 5.47 \( \times \) 5.47 \( \times \) 3.27 mm\(^3\). We evaluated the performance of the TNNMU algorithm in comparison with manual delineation by two radiologists, having more than ten years of experience, abbreviated as BZ and XW. The manual delineations were used as "ground truth" (GT) and referred to as GT-1 and GT-2 from now on. The size of the tumor varied between 8.05 cm\(^3\) (82.28 voxels) to 506.52 cm\(^3\) (5177 voxels) according to GT-1, and between 6.75 cm\(^3\) (68.99 voxels) to 448.11 cm\(^3\) (4580 voxels) according to GT-2. Figure 5 shows estimated tumor size per patient according to two ground truths. Inter-observer variability calculated as a relative difference between aggregate volume of 18 NSCLCs according to GT-1 and GT-2 is 1.33%. Figure 6a shows result of linear regression analysis for tumor size in cm\(^3\) estimated by GT-1 and GT-2: GT-2=0.9618\times GT-1+4.624. Pearson correlation coefficient between the two estimates is 0.9969. Figure 6b shows Bland-Altman plot for GT-1 vs. GT-2, where difference in tumor size between GT-1 and GT-2 is normalized with respect to the average: (GT-1+GT-2)/2.

**B. Other Delineation Approaches for Comparisons**

We compared the TNNMU algorithm with five delineation methods which were: 1) a threshold at 40% of maximum SUV, 2) a threshold at 50% of maximum SUV, 3) AP [16], [17], 4) GC [15], and 5) RW [14]. The thresholding methods are referred as TH40 and TH50 while the proposed method is referred as TNNMU. The maximum SUV was estimated within the entire volume. The TH40, TH50, AP and TNNMU algorithms were implemented in MATLAB\(^{\circledR}\) 2011b (The MathWorks, Inc., Natick, MA) script language. The MATLAB code for AP method is available for free download at [72].

**C. Performance Measure**

We have used Dice similarity coefficient (DSC) to measure the degree of overlap between each of the five delineation results and ground truths. DSC is calculated as:

\[
\text{DSC}(V_1, V_2) = \frac{2|V_1 \cap V_2|}{|V_1| + |V_2|}
\]

where \( V_1 \) stands for ground truth volume, and \( V_2 \) stands for the volume obtained from each delineation approach. Thereby, \( \text{DSC}(V_1, V_2) \in [0, 1] \). A DSC value of 0 indicates no overlap and a value of 1 indicates perfect overlap.
D. Parameters of the TNNMU Algorithm

The TNNMU algorithm includes two main parameters: order of nonlinear mapping $D$ in (3) and threshold $\tau$ used in (11) to segment the lung tumor component separated by the NMU algorithm in (8). Thus, we have calculated DSC between ground truths and delineation results obtained by the TNNMU($D, \tau$) from 18 NSCLCs PET studies, whereas $D \in \{3, ..., 9\}$ and $\tau \in \{0.1, 0.15, 0.2, 0.25, 0.3\}$. Thereby, we have used leave-one-out cross validation to estimate optimal values for $D$ and $\tau$, as well as to estimate performance of the TNNMU algorithm. That is, one image was left as test image and optimal values for $D$ and $\tau$ were selected based on the rest and applied to segmentation of the test image. The values of $D$ and $\tau$ that correspond to the highest value of Dice coefficient are shown in Figure 7. The mean value of $D$ is 8.11 for GT-1 and 8.06 for GT-2. The mean value of $\tau$ is 0.144 for GT-1 and 0.147 for GT-2. That justifies selection of $D=8$ and $\tau=0.15$ as approximately optimal. Figure 8 shows segmentation results for some characteristic combinations of $D$ and $\tau$. Visually, it is evident that proposed combination $D=8$ and $\tau=0.15$ corresponds best with the GT-1.
E. Results of Comparative Performance Analysis

Delineation results obtained by the TNNMU, TH50, GC, RW and AP algorithms are quantified in terms of DSC calculated for 18 PET image studies with respect to GT-1 and GT-2, and shown respectively in Fig. 9a) and 9b). Thereby, results for the TNNMU algorithm were obtained by leave-one-out cross validation. Aggregate results for all 18 PET image studies in term of mean value of DSC (± standard deviation) are shown in Table I. Table II shows p values estimated pairwise by repeated measures ANOVA tests on DSC obtained by the TNNMU method and TH50, TH40, AP, GC and RW methods. We have used anova_rm function available at MATLAB File Exchange Central, [78], to run the test. The automatic TNNMU method achieves the same performance as GC. It is better than the rest of methods, although not statistically significantly. However, the TNNMU method is fully automatic, while GC and RW are interactive. In regard to the AP, that is supposed to be automatic, combination of several segmented layers has to be found to achieve best result. That requires postprocessing. In case of results presented herein, 3 layers had to be combined to achieve reported results. Layer selection is analogous to the component selection performed by the TNNMU method. However, that is executed automatically by using criterion (10). We also want to emphasize that RW-GC integrated co-segmentation of PET-CT image method in [13] yielded result in term of DSC value equal to 0.84±0.06, but with the help of CT information. Thus, given the fact that TNNMU method is fully automatic and is using information from the PET image only, result obtained by the TNNMU can be considered good. Hence, results presented in Tables I and II suggest that the proposed TNNMU algorithm can be method of choice for lung tumor delineation from sole PET image. We showed in Figure 10 segmentation results for one PET study obtained by TH50, GC, RW, AP and TNNMU algorithms. Ground truths 1 and 2 as well as input PET images are also shown. Obtained mean value of DSC (± one standard deviation) for GT-1 for TH50, RW, AP, GC and TNNMU methods is given as: 0.71±0.08, 0.79±0.11, 0.84±0.08, 0.80±0.12 and 0.90±0.05. For GT-2, mean DSC values (± one standard deviation) are given in the same order as: 0.71±0.08, 0.79±0.09, 0.85±0.09, 0.80±0.10, 0.93±0.08. To validate precision (reproducibility) of the TNNMU algorithm, it has been applied repeatedly 100 times on the 3D PET images of 18 patients. The aggregate result, in terms of Dice coefficient, with respect to GT-1 for all 18 patients and all 100 segmentations is 0.78±0.12 which coincides with the value presented in Table I. Furthermore, Figure 11 shows segmentation results for the PET study, whereas demarcation lines, corresponding with GC, TNNMU and GT-1, are superimposed on input PET images. Reported segmentation experiments were executed in MATLAB script language running under 64-bit Windows operating system on desktop computer with clock speed of 2.4GHz and 24GB of RAM. It takes 63 s for the TNNMU algorithm to segment PET image comprised of 50 slices. Direct comparison with GC and RW is not possible due to their interactive nature. AP algorithm is very fast and takes 0.68 s to segment the same PET image once proper combination of segmented layers is found. However, it takes multiple runs to find such a combination. This also raises question of overfitting.

Fig. 9. Mean value (± one standard deviation) of DSC of delineation results obtained by the TNNMU (squares), TH50 (stars), AP (circles), GC (pentagram) and RW (triangle down) algorithms on 3D segmentation of 18 PET studies of patients diagnosed with NSCLC. DSC calculation with respect to: (a) GT_1, and (b) GT_2.
TABLE I
MEAN VALUES OF DSC (± ONE STANDARD DEVIATION) OBTAINED BY THE TNNMU, TH50, TH40, AP, GC AND RW ALGORITHMS OVER 18 PET IMAGE STUDIES AND AVERAGED OVER TWO GROUND TRUTHS. MAPPING ORDER D=8. SEGMENTATION THRESHOLD \( \tau = 0.15 \). THE BEST RESULT IS IN BOLD FONT.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Averaged GT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNNMU</td>
<td>0.78±0.12</td>
</tr>
<tr>
<td>TH50</td>
<td>0.75±0.13</td>
</tr>
<tr>
<td>TH40</td>
<td>0.60±0.17</td>
</tr>
<tr>
<td>AP</td>
<td>0.77±0.13</td>
</tr>
<tr>
<td>GC</td>
<td>0.78±0.1</td>
</tr>
<tr>
<td>RW</td>
<td>0.77±0.07</td>
</tr>
</tbody>
</table>

TABLE II
P VALUES ESTIMATED BY REPEATED MEASURES ANOVA TEST BETWEEN THE TNNMU AND TH50, TH40, AP, GC AND RW ALGORITHMS OVER 18 PET IMAGE STUDIES AND TWO GROUND TRUTHS. MAPPING ORDER D=8. SEGMENTATION THRESHOLD \( \tau = 0.15 \).

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>GT-1</th>
<th>GT-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TH50</td>
<td>0.1509</td>
<td>0.1860</td>
</tr>
<tr>
<td>TH40</td>
<td>1.5±10^-6</td>
<td>0</td>
</tr>
<tr>
<td>AP</td>
<td>0.4015</td>
<td>0.3921</td>
</tr>
<tr>
<td>GC</td>
<td>0.6062</td>
<td>0.6685</td>
</tr>
<tr>
<td>RW</td>
<td>0.8229</td>
<td>0.3519</td>
</tr>
</tbody>
</table>

Fig. 10. Comparative results for one PET study. Shown area between pixels 53:73 and 42:64. From left to right: input PET image, segmentation results obtained by TH50, RW, AP, GC and TNNMU methods, ground truths 1 and 2. From top to bottom: slices 1 to 7. Mean value of DSC (± one standard deviation) for GT_1: TH50: 0.71 ±0.08, RW: 0.79±0.11, AP: 0.84±0.08, GC: 0.80±0.12, TNNMU: 0.90±0.05 and for GT_2: TH50: 0.71±0.08, RW: 0.79±0.09, AP: 0.85±0.09, GC: 0.80±0.10, TNNMU: 0.93±0.08.

V. DISCUSSION
The proposed single-channel BSS method for segmentation of the lung tumor component from PET image combines nonlinear mapping of PET image to pseudo-multichannel image with NMU-based factorization of mapped image. Underapproximation constraint of the NMU algorithm incorporates implicitly sparseness constraints on the encoding coefficients matrix \( S \) in (7). That is achieved without sparseness regularization tradeoff parameter that would required tuning.

That, in combination with predefined values of mapping order and binarization threshold, makes proposed method fully automatic. That is a big advantage of the proposed method with respect to interactive methods GC and RW as well as with respect to the subspace clustering methods such as Ncut [36], that is equivalent to orthogonality regularized NMF and therefore requires tuning of the regularization constant. However, implicit underapproximation based sparseness constraint has a disadvantage of not yielding the sparsest decomposition possible. Having no direct control over degree of sparseness of \( S \) can decrease quality of delineation when two components have very similar uptake values. That happens if the lung tumor is not homogenous but contains also a necrotic part. Quality of delineation in such a case may deteriorate since uniqueness of the solution of (7) depends on the amount of sparseness of \( \{ s_i \in \mathbb{R}^{M_i \times 1} \}_{i=1}^{N} \) and mutual coherence, \( \mu(V) \), of \( V \) in (6)/(7) through [73]:

\[
\| V \|_F \leq \sqrt{\frac{1}{2}(1+\frac{1}{\mu(V)})}
\]

where \( \| \cdot \|_F \) stands for the \( \ell_2 \)-quasi-norm that counts number of non-zero coefficients in \( s \). Thus, it is beneficial to have more direct control over the degree of sparseness of \( s \) . When two components have very similar uptake values the corresponding basis vectors in the LMM (6) will, due to A2), be highly correlated. Thus, \( \mu(V) \) will be close to 1. Hence, the required amount of sparseness imposed on \( s \) by (12) is \( \| s \|_F \leq 1 \).

Thanks to A1) that condition is fulfilled. It is envisioned that using the NMF algorithm with sparseness constraint based on the \( \ell_1 \), [74], [75] or the \( \ell_0 \)-norm, [76], of \( S \) to solve factorization problem (7) would provide more accurate demarcation of the necrotic part from the rest of the tumor than using the NMU algorithm. In particular, for the \( \ell_0 \)-norm constrained NMF algorithm [76] explicit sparseness constraint can be imposed on \( \| s \|_F \) to be 1. As demonstrated in [77], that enabled separation of tissue components present in the image of unstained histopathological specimen.

VI. CONCLUSIONS
In this paper, we have proposed a novel method for single-channel nonnegative sparse BSS as well as its application to automatic 3D delineation of the lung tumor from PET image. Proposed method was compared with thresholds of 40% and 50% of maximum SUV, AP, GC and RW on 18 NSCLC datasets with respect to manual delineation performed by two clinical experts. Since the proposed method achieved performance comparable with interactive methods, considering the unique challenges of lung tumor segmentation from PET images, our findings support possibility of using our fully automated method in routine clinicals. The source codes will be available at www.mipav.net/English/research/research.html .
ACKNOWLEDGMENTS

We want to thank Professors Milan Sonka for proofreading the paper and giving us useful suggestions.

REFERENCES


