

# 3-D Lung Segmentation by Incremental Constrained Nonnegative Matrix Factorization

Ehsan Hosseini-Asl, *Member, IEEE*, Jacek M. Zurada\*, *Life Fellow, IEEE*, Georgy Gimel’farb, and Ayman El-Baz, *Senior Member, IEEE*

**Abstract**—Accurate lung segmentation from large-size 3-D chest-computed tomography images is crucial for computer-assisted cancer diagnostics. To efficiently segment a 3-D lung, we extract voxel-wise features of spatial image contexts by unsupervised learning with a proposed incremental constrained nonnegative matrix factorization (ICNMF). The method applies smoothness constraints to learn the features, which are more robust to lung tissue inhomogeneities, and thus, help to better segment internal lung pathologies than the known state-of-the-art techniques. Compared to the latter, the ICNMF depends less on the domain expert knowledge and is more easily tuned due to only a few control parameters. Also, the proposed slice-wise incremental learning with due regard for interslice signal dependencies decreases the computational complexity of the NMF-based segmentation and is scalable to very large 3-D lung images. The method is quantitatively validated on simulated realistic lung phantoms that mimic different lung pathologies (seven datasets), *in vivo* datasets for 17 subjects, and 55 datasets from the Lobe and Lung Analysis 2011 (LOLA11) study. For the *in vivo* data, the accuracy of our segmentation w.r.t. the ground truth is 0.96 by the Dice similarity coefficient, 9.0 mm by the modified Hausdorff distance, and 0.87% by the absolute lung volume difference, which is significantly better than for the NMF-based segmentation. In spite of not being designed for lungs with severe pathologies and of no agreement between radiologists on the ground truth in such cases, the ICNMF with its total accuracy of 0.965 was ranked fifth among all others in the LOLA11. After excluding the nine too pathological cases from the LOLA11 dataset, the ICNMF accuracy increased to 0.986.

**Index Terms**—Constrained nonnegative matrix factorization, incremental learning, lung segmentation.

## I. INTRODUCTION

ACCURATE automated segmentation of lung tissues from computed tomographic (CT) images is of profound importance for developing noninvasive computer-assisted system for early diagnosis of lung cancer and other pulmonary

Manuscript received April 10, 2015; revised May 26, 2015 and August 27, 2015; accepted September 18, 2015. Date of publication September 25, 2015; date of current version May 19, 2016. This work was supported in part by Research Scholar Grant 120556-RSG-11-266-01-CCE from the American Cancer Society. Asterisk indicates corresponding author.

E. Hosseini-Asl is with the Department of Electrical and Computer Engineering, University of Louisville.

\*J. M. Zurada is with the Department of Electrical and Computer Engineering, University of Louisville Louisville, KY 40292 USA and also with the Information Technology Institute, University of Social Science (e-mail: jmzura02@louisville.edu).

G. Gimel’farb is with the Department of Computer Science, University of Auckland.

A. El-Baz is with the Bioengineering Department, University of Louisville. Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBME.2015.2482387

diseases [1]–[5]. This problem is challenging because discriminatory pixel/voxel-wise signals for lung tissues are often inhomogeneous and lack strong differences from tissues surrounding the lungs and pulmonary structures, such as arteries, veins, bronchi, bronchioles, etc.

The success of a particular segmentation technique is measured in terms of its accuracy, processing time, and automation level. Most of the known lung segmentation techniques employ and combine in various ways multiple image processing and analysis tools, such as, e.g., signal thresholding, region growing, deformable parametric or geometric (level-set-based) boundaries that evolve toward a goal region border under external and internal forces, depending at each step on image signals and a current boundary shape, prior models of shapes of goal regions, and/or image signals classifiers and region maps refiners employing Markov–Gibbs random field (MGRF) models of images and maps. Adding the learned shape prior model to image signals improves the segmentation accuracy, but requires to coaligned an input CT image to the prior before starting the segmentation. In particular, Itai *et al.* [6] extracted the lungs with a 2-D parametric deformable boundary guided by an external force depending on deviations from anticipated lung borders on the image. Sluimer *et al.* [7] segmented a pathological lung using the shape prior for the normal lungs. A two-stage MGRF-based segmentation by El-Baz *et al.* [8] separates the lungs from a low-dose CT image using first a voxel-wise Bayesian classifier. A linear combination of discrete Gaussians that closely approximates the empirical marginal probability distribution of image signals is separated into the lung and background models. Then regions after the classification are refined by searching for the closest local minimum of energy for a joint MGRF combining the first-order conditional independent random field of images, given a region map, and the unconditional MGRF of region maps with second-order spatial voxel interactions. Many today’s lung segmentation techniques are surveyed in [9].

The recent “LObe and Lung Analysis 2011 challenge” (LOLA11) yielded a number of state-of-the-art techniques for segmenting the chest CT scans. Lassen *et al.* [10] initialize a lung region using a fixed signal threshold, grow the region to coarsely extract pulmonary airspaces; approximate a lung bounding box using a low-resolution lung mask, and then select an appropriate lung segmentation threshold by analyzing the empirical marginal signal distribution in the box. A tracheobronchial tree is segmented in [11] growing the left and right lung regions with two guiding signal thresholds; filling holes in region maps for axial 2-D CT slices, and using the morphological closing to include sub-pleural abnormalities and concavities to the lungs.

Mansoor *et al.* [12] propose a fuzzy connectedness principle to extract initial lung parenchyma and estimate the lung volume directly from the ribcage information without explicitly delineating the lungs. Lo *et al.* [13] segment the lungs by the 3-D region growing and morphological operations. The region growing is also employed in [14] to extract large airways, e.g., trachea and main bronchi; segment the lungs using an optimal upper threshold [15], and improve each lung separately by filling 3-D holes and using the morphological closing with a spherical structuring element to include blood vessels and smooth the borders. Sun *et al.* [16] detect a ribcage to initialize and roughly outline the lungs with a deformable shape model and adapt the outlined surface to the image.

However, a vast majority of these techniques do not account explicitly for characteristic spatial dependencies between image signals in the lungs and their background, and thus, have limited abilities in adapting to and retaining acceptable accuracies in the presence of significant natural lung inhomogeneities. The conventional first-order signal models have gained wide-spread acceptance only because accurate modeling of higher-order spatial dependencies and computationally feasible embedding of such models into segmentation process are generally challenging problems. Nonetheless, recent applications of the nonnegative matrix factorization (NMF) to extract voxel-wise vectorial features quantifying high-order spatial signal dependencies for image segmentation [17]–[20] hold much promise for better object discrimination.

#### A. Nonnegative Matrix Factorization

Let  $\mathbb{R}_+$  denote the set of nonnegative real numbers and  $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_n] \in \mathbb{R}_+^{m \times n}$  be a nonnegative  $m \times n$  data matrix, with columns,  $\mathbf{a}_j = [a_{1,j}, \dots, a_{m,j}]^\top; j = 1, \dots, n$ , representing  $m$ -dimensional data vectors.

The NMF [21] approximates the data matrix with a product  $\mathbf{A} \approx \mathbf{WH}$  of two lower-rank nonnegative matrices: the  $m \times r$  basis matrix  $\mathbf{W} \in \mathbb{R}_+^{m \times r}$  and the  $r \times n$  projection matrix  $\mathbf{H} \in \mathbb{R}_+^{r \times n}$ . The columns of  $\mathbf{W}$ , i.e., the  $m$ -dimensional vectors  $\mathbf{w}_q; q = 1, \dots, r$ , form the basis of a reduced  $r$ -dimensional feature space;  $r < m$ :  $\mathbf{w}_q = [w_{1,q}, \dots, w_{m,q}]^\top$  where  $\top$  denotes the vector or matrix transposition. The columns of  $\mathbf{H}$  are the  $r$ -dimensional vectors  $\mathbf{h}_j = [h_{1,j}, \dots, h_{r,j}]^\top; j = 1, \dots, n$ , representing projections of the initial data vectors  $\mathbf{a}_j$  to the  $r$ -dimensional feature space, i.e.,  $\mathbf{a}_j \approx \sum_{q=1}^r h_{q,j} \mathbf{w}_q; j = 1, \dots, n$ . In other words, the NMF reduces the initial  $m$ -dimensional data vectors  $\mathbf{a}_j$  to the  $r$ -dimensional ones  $\mathbf{h}_j$  in  $\mathbf{H}; j = 1, \dots, n$ .

The factors  $\mathbf{W}$  and  $\mathbf{H}$  are estimated by minimizing the reconstruction error [22], measured often with the squared Frobenius norm  $\|\mathbf{D}\|^2$  of the difference  $\mathbf{D} = \mathbf{A} - \mathbf{WH}$ :

$$F_{\text{re}} = \|\mathbf{A} - \mathbf{WH}\|^2 \equiv \sum_{i=1}^m \sum_{j=1}^n (a_{ij} - (\mathbf{WH})_{ij})^2. \quad (1)$$

Here and below, for brevity,  $(\mathbf{WH})_{ij}$  denotes the  $(i, j)$ -element of the product matrix  $\mathbf{WH}$ :

$$(\mathbf{WH})_{ij} = \sum_{k=1}^r w_{ik} h_{kj}; \quad i = 1, \dots, m; \quad j = 1, \dots, n.$$

In some cases, the NMF is able to produce factors  $\mathbf{W}$  and/or  $\mathbf{H}$ , which can be considered basic data constituents due to only a small number of nonzero elements. Sparsity and smoothness constraints on the factors  $\mathbf{W}$  and/or  $\mathbf{H}$  [23]–[27], in particular, on their Frobenius norms [27], [28]:

$$F = F_{\text{re}} + \lambda_{\mathbf{W}} \|\mathbf{W}\|^2 + \lambda_{\mathbf{H}} \|\mathbf{H}\|^2 \quad (2)$$

amplify this ability, making the NMF a major focus of attention. The empirically selected weights  $\lambda_{\mathbf{W}}$  and  $\lambda_{\mathbf{H}}$  control the smoothness of  $\mathbf{W}$  and the sparseness of  $\mathbf{H}$ , respectively.

To perform the NMF, a local minimum of the unconstrained [(1)] or constrained reconstruction error, such as, e.g., in (2), is searched for with an optimization algorithm such as, e.g., the alternating least square (ALS), active set-based ALS (ALS-AS), or multiplicative update rules (MUR) [22], [29]–[31]. The equivalence between the squared Frobenius norm  $\|\mathbf{D}\|^2$  of an arbitrary matrix  $\mathbf{D}$  and the trace  $\text{Tr}(\mathbf{DD}^\top)$  of the square matrix  $\mathbf{DD}^\top$ :

$$\|\mathbf{D}\|^2 \equiv \text{Tr}(\mathbf{DD}^\top)$$

simplifies obtaining matrix derivatives [32] for minimizing the errors of (1) and (2) by expanding (1):

$$\begin{aligned} F_{\text{re}} &= \text{Tr}((\mathbf{A} - \mathbf{WH})(\mathbf{A} - \mathbf{WH})^\top) \\ &= \text{Tr}(\mathbf{AA}^\top) - 2\text{Tr}(\mathbf{AH}^\top \mathbf{W}) + \text{Tr}(\mathbf{WH} \mathbf{H}^\top \mathbf{W}^\top) \end{aligned} \quad (3)$$

and taking its derivatives:

$$\begin{aligned} \frac{\partial F_{\text{re}}}{\partial \mathbf{W}} &= -2(\mathbf{A} - \mathbf{WH})\mathbf{H}^\top; \\ \frac{\partial F_{\text{re}}}{\partial \mathbf{H}} &= -2\mathbf{W}^\top(\mathbf{A} - \mathbf{WH}); \quad \frac{\partial \|\mathbf{V}\|^2}{\partial \mathbf{V}} = 2\mathbf{V} \end{aligned} \quad (4)$$

where  $\mathbf{V}$  stands for either  $\mathbf{W}$  or  $\mathbf{H}$ . However, if the matrix  $\mathbf{A}$  is very large, the straightforward error minimization becomes computationally too expensive.

#### B. NMF for Image Segmentation

The NMF is often used as an efficient dimensionality reduction tool in machine learning and image analysis by feature extraction, clustering, and classification. However, due to relative novelty, it is rarely used for segmentation, especially, in application to medical images, in spite of its advantages, such as dealing with piecewise-homogeneous vectorial properties of pixels or voxels, possible unsupervised learning of basis features, and using only a few control parameters. After extracting one or more characteristic basis features for each region of interest, the initial pixel/voxel-wise vectors are projected onto the feature space and every projection is associated with the most relevant feature(s). Xie *et al.* [17] extracted features of spinal cord, corpus callosum, and hippocampus from diffusion tensor Images of rat brains with the NMF and stratified all the the pixels by the  $K$ -means clustering [33] of their projections in the matrix  $\mathbf{H}$ . For segmenting a multispectral barley grain cross section, Lazar *et al.* [18] decorrelated an image dataset with the principal component analysis and applied the NMF to the decorrelated data. The number  $r$  of the basis features in

$\mathbf{W}$  was equal to the number of the most significant principal components, responsible for the main part of the data variance. Then the data samples were classified by their closeness to the features found, which was evaluated by the maximum component of the corresponding projection vector in  $\mathbf{H}$ . Sandler and Lindenbaum [19] divided a texture mosaic into nonoverlapping rectangular blocks and described each block with a vector of outputs of a bank of linear Gabor, or wavelet filters. The characteristic basis features were found with the NMF employing the Earth mover's distance, and the Bayesian maximum a posteriori decision rule was applied to classify each pixel.

Because 3-D lung CT images are extremely large and make the ALS computationally too expensive, the ALS-AS-based NMF was applied in our previous study [20] separately to each axial 2-D CT slice converted into a classical context image [34]. Each voxel of the latter is represented by a context vector, containing the original voxel-wise signal and signals from its selected neighbors, e.g., the 27-vector for the nearest  $3 \times 3 \times 3$  volume centered on the voxel. The obtained projections  $\mathbf{h}_j$  are used as characteristic voxel-wise descriptors for segmentation and associated with the darker lung and brighter chest tissues by their total intensity: the lower the Frobenius norm  $\|\mathbf{h}_j\|$  (the Cartesian distance from  $\mathbf{h}_j$  to the origin 0), the darker the tissue. However, that the separate slice-wise segmentation accounts for the interslice signal dependencies only implicitly, via the context, decreases separability and smoothness of the chest and lung manifolds in the space of h-descriptors, and hence, decreases the segmentation accuracy. This drawback is overcome by using the introduced below incremental constrained NMF (ICNMF), combining basic ideas of the earlier incremental NMF (INMF) [35] and constrained NMF (CNMF) [27], [28]. Voxel-by-voxel decomposition of the entire 3-D context image with the INMF improves capturing the interslice spatial dependencies, while the lung and chest manifolds in the h-space learned by the CNMF are smoother and more discriminable than in [20]. To reduce the computational complexity, the large data matrix  $\mathbf{A}$  for a 3-D CT image is decomposed with the ICNMF slice-after-slice, so that factorizing each next axial CT slice is initialized with the basis and projection matrices, having been already obtained from all the preceding slices.

The proposed ICNMF and its application to the 3-D CT lung image segmentation are detailed in Section II. Section III presents and discusses experimental results, and conclusions are given in Section IV.

## II. METHODS

Sections II-A and II-B below describe embedding constraints into the INMF in order to build the ICNMF and the ICNMF-based lung segmentation algorithm, respectively. To minimize the reconstruction error of (1) for a large matrix  $\mathbf{A}$  in a computationally feasible way, the INMF builds  $\mathbf{W}$  and  $\mathbf{H}$  incrementally, using at each step the iterative MUR converging to the closest local minimum [35]. After adding each next data vector to  $\mathbf{A}$ , the already computed  $\mathbf{W}$  and  $\mathbf{H}$  initialize the next step, thus reducing the overall computational complexity. However, the INMF by itself does not guarantee a smooth and sparse data

representation, such that the CNMF obtains due to constraining the matrix factors. The ICNMF combines the main ideas of the INMF and CNMF.

### A. Incremental Constrained NMF

Let  $\mathbf{A}_k = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_k]$ ;  $\mathbf{W}_k; \mathbf{H}_k$ ;  $F_{\text{re}:k}$ , and  $F_k$ , denote the data matrix with the initial  $k$  data samples, the corresponding basis and projection matrices, the reconstruction error, and the constrained reconstruction error, respectively:

$$\begin{aligned} F_{\text{re}:k} &= \|\mathbf{A}_k - \mathbf{W}_k \mathbf{H}_k\|^2; \\ F_k &= F_{\text{re}:k} + \lambda_{\mathbf{W}} \|\mathbf{W}_k\|^2 + \lambda_{\mathbf{H}} \|\mathbf{H}_k\|^2. \end{aligned} \quad (5)$$

The INMF [35] assumes that every new sample,  $\mathbf{a}_{k+1}$ , does not significantly affect the current basis  $\mathbf{W}_k$ , optimized for the previous  $k$  samples, so that their projection vectors  $\mathbf{H}_k$  need not be changed. Then the first  $k$  columns of  $\mathbf{H}_{k+1}$  remain equal to  $\mathbf{H}_k$ , i.e.,  $\mathbf{H}_{k+1} = [\mathbf{H}_k \ \mathbf{h}_{k+1}]$ , and only the basis,  $\mathbf{W}_{k+1}$  and the last projection vector-column  $\mathbf{h}_{k+1}$  have to be updated [35]. To reach a local minimum of the constrained reconstruction error of (5) after adding the new sample  $\mathbf{a}_{k+1}$  the updates:

$$\begin{aligned} F_{\text{re}:k+1} &= \|\mathbf{A}_{k+1} - \mathbf{W}_{k+1} \mathbf{H}_{k+1}\|^2; \\ F_{k+1} &= F_{\text{re}:k+1} + \lambda_{\mathbf{W}} \|\mathbf{W}_{k+1}\|^2 + \lambda_{\mathbf{H}} \|\mathbf{H}_{k+1}\|^2 \end{aligned} \quad (6)$$

are converted into an incremental form separating the previous samples from the new one:

$$\begin{aligned} F_{k+1} &\cong \widehat{F}_{\text{re}:k} + \|\mathbf{a}_{k+1} - \mathbf{W}_{k+1} \mathbf{h}_{k+1}\|^2 \\ &\quad + \lambda_{\mathbf{W}} \|\mathbf{W}_{k+1}\|^2 + \lambda_{\mathbf{H}} \|\mathbf{H}_{k+1}\|^2. \end{aligned} \quad (7)$$

Here,  $\widehat{F}_{\text{re}:k}$  is the approximate reconstruction error for the first  $k$  samples:

$$\begin{aligned} \widehat{F}_{\text{re}:k} &= \sum_{i=1}^m \sum_{j=1}^k (a_{ij} - (\mathbf{W}_{k+1} \mathbf{H}_{k+1})_{ij})^2 \\ &\cong \sum_{i=1}^m \sum_{j=1}^k (a_{ij} - (\mathbf{W}_{k+1} \mathbf{H}_k)_{ij})^2. \end{aligned} \quad (8)$$

Therefore, after adding the next sample  $\mathbf{a}_{k+1}$ , the conditional minimization of (7) is rewritten to

$$\begin{aligned} F_{k+1} &= \|\mathbf{A}_k - \mathbf{W}_{k+1} \mathbf{H}_k\|^2 \\ &\quad + \|\mathbf{a}_{k+1} - \mathbf{W}_{k+1} \mathbf{h}_{k+1}\|^2 \\ &\quad + \lambda_{\mathbf{W}} \|\mathbf{W}_{k+1}\|^2 \\ &\quad + \lambda_{\mathbf{H}} (\|\mathbf{H}_k\|^2 \|\mathbf{h}_{k+1}\|^2) \end{aligned} \quad (9)$$

which optimizes the basis  $\mathbf{W}_{k+1}$  and adds the new column  $\mathbf{h}_{k+1}$  to the projection matrix,  $\mathbf{H}_{k+1} = [\mathbf{H}_k \ \mathbf{h}_{k+1}]$ . A conditional local minimum of the error in (9) is found (see Appendix B) by a gradient-like iterative search [36]:

$$\begin{aligned} (\mathbf{h}_{k+1})_q &\leftarrow (\mathbf{h}_{k+1})_q - \alpha_q \frac{\partial F_{k+1}}{\partial (\mathbf{h}_{k+1})_q}; \\ (\mathbf{W}_{k+1})_{iq} &\leftarrow (\mathbf{W}_{k+1})_{iq} - \beta_{iq} \frac{\partial F_{k+1}}{\partial (\mathbf{W}_{k+1})_{iq}}; \\ q &= 1, \dots, r; \quad i = 1, \dots, m \end{aligned} \quad (10)$$

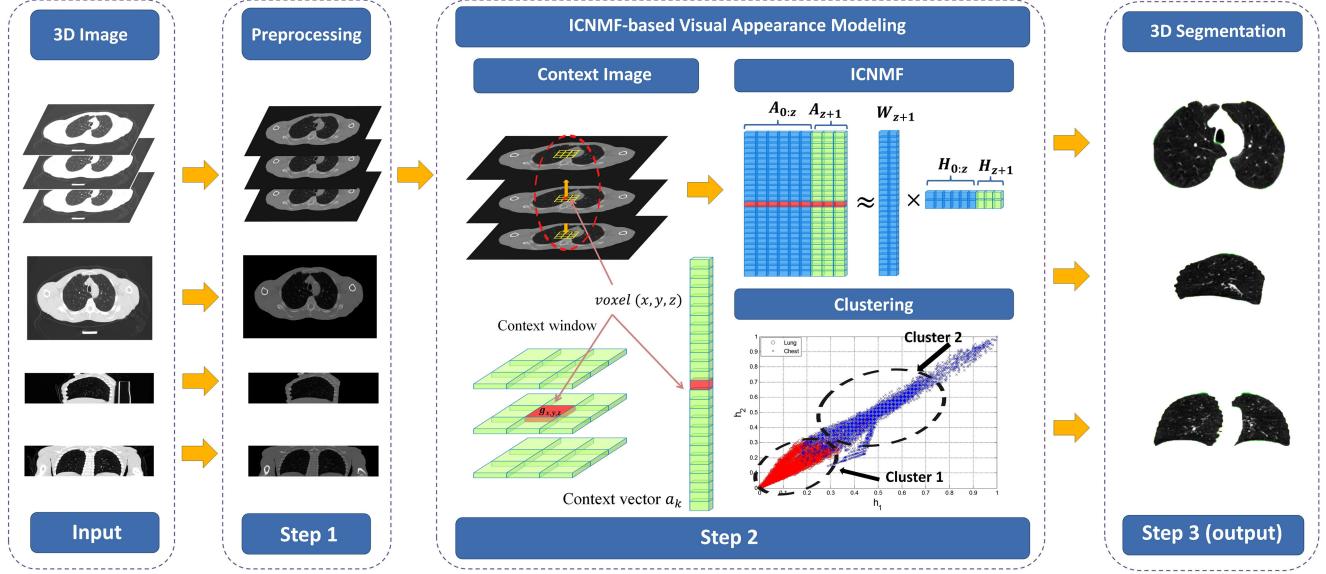


Fig. 1. Proposed 3-D lung segmentation: preprocessing removes the background by 3-D region growing from a seed at the 3-D image corner.

---

**Algorithm 1** 3-D Lung Segmentation by ICNMF-based Visual Appearance Modeling.

---

- 1: *Preprocessing*: Remove background of an input 3D CT image  $\mathbf{g}$  by conventional 3-D region growing.
  - 2: Apply Algorithm 2 to describe visual appearance of the remaining 3-D context image with  $r$ -dimensional voxel-wise projection vectors in the matrix  $\mathbf{H}$  (the value  $r$  is determined empirically).
  - 3: Assign the voxel-wise descriptors to a prescribed number,  $K$ , of objects by the  $K$ -means clustering [33].
  - 4: Discriminate between the lung and chest clusters by characterizing their relative brightness with the Frobenius norms of their centroids in the  $\mathbf{h}$ -space.
  - 5: Refine the segmented 3-D lung regions by analyzing 3-D connected components.
- 

where  $\alpha_q$  and  $\beta_{iq}$  are specific steps for updating the elements  $(\mathbf{h}_{k+1})_q$  and  $(\mathbf{W}_{k+1})_{iq}$ , respectively, and the partial derivatives follow from (4):

$$\begin{aligned} \frac{\partial F_{k+1}}{\partial \mathbf{h}_{k+1}} &= -2\mathbf{W}_{k+1}^T (\mathbf{a}_{k+1} - \mathbf{W}_{k+1}\mathbf{h}_{k+1}) \\ &\quad + 2\lambda_{\mathbf{H}}\mathbf{h}_{k+1}; \\ \frac{\partial F_{k+1}}{\partial \mathbf{W}_{k+1}} &= -2(\mathbf{A}_k - \mathbf{W}_{k+1}\mathbf{H}_k)\mathbf{H}_k^T \\ &\quad - 2(\mathbf{a}_{k+1} - \mathbf{W}_{k+1}\mathbf{h}_{k+1})\mathbf{h}_{k+1}^T \\ &\quad + 2\lambda_{\mathbf{W}}\mathbf{W}_{k+1}. \end{aligned} \quad (11)$$

As shown in [36], the required adaptive steps result in the multiplicative updates, ensuring that factors  $\mathbf{W}_k$  and  $\mathbf{H}_k$ , which initially (for  $k = 1$ ) were nonnegative, remain nonnegative at every updating iteration, and hence, at every step,  $k = 2, \dots, n$ :

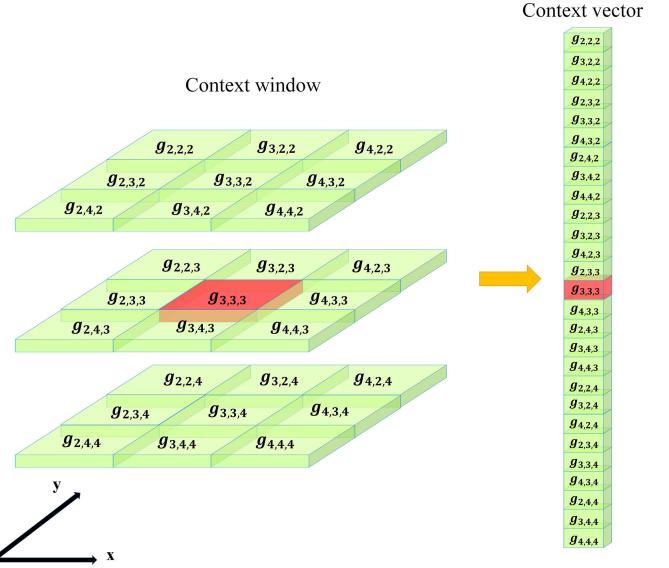


Fig. 2. Context vector for the voxel (3, 3, 3) and its nearest 26-neighborhood.

for  $q = 1, \dots, r$  and  $i = 1, \dots, m$

$$\begin{aligned} \alpha_q &= \frac{(\mathbf{h}_{k+1})_q}{(\mathbf{W}_{k+1}^T \mathbf{W}_{k+1} \mathbf{h}_{k+1})_q}; \\ (\mathbf{h}_{k+1})_q &\leftarrow (\mathbf{h}_{k+1})_q \frac{(\mathbf{W}_{k+1}^T \mathbf{a}_{k+1})_q}{(\mathbf{W}_{k+1}^T \mathbf{W}_{k+1} \mathbf{h}_{k+1} + \lambda_{\mathbf{H}} \mathbf{h}_{k+1})_q}; \\ \beta_{iq} &= \frac{(\mathbf{W}_{k+1})_{iq}}{S_{iq}}; \\ (\mathbf{W}_{k+1})_{iq} &\leftarrow (\mathbf{W}_{k+1})_{iq} \frac{(\mathbf{A}_k \mathbf{H}_k^T + \mathbf{a}_{k+1} \mathbf{h}_{k+1}^T)_{iq}}{S_{iq}} \end{aligned} \quad (12)$$

where  $S_{iq} = (\mathbf{W}_{k+1} \mathbf{H}_k \mathbf{H}_k^T + \mathbf{W}_{k+1} \mathbf{h}_{k+1} \mathbf{h}_{k+1}^T + \lambda_{\mathbf{W}} \mathbf{W}_{k+1})_{iq}$  and  $\mathbf{W}_{k+1}$  is initialized with  $\mathbf{W}_k$ , when the new sample  $\mathbf{a}_{k+1}$

**Algorithm 2** Describing Visual Appearance with the ICNMF.1: **Initialization**

- (i) Given a 3D image  $\mathbf{g}$ , form the data matrix  $\mathbf{A}_0$  from the context vectors of the first slice  $z = 0$ .
- (ii) Set the number  $r$  of voxel-wise features characterizing visual appearance of the object-of-interest.
- (iii) Initialize randomly the factors  $\mathbf{W}_0$  and  $\mathbf{H}_0$ .
- (iv) Update  $\mathbf{H}_0$  and  $\mathbf{W}_0$  iteratively until convergence to a local minimum of the reconstruction error:

$$(\mathbf{H}_0)_{qj} \leftarrow (\mathbf{H}_0)_{qj} \frac{(\mathbf{W}_0^\top \mathbf{A}_0)_{qj}}{(\mathbf{W}_0^\top \mathbf{W}_0 \mathbf{H}_0 + \lambda_{\mathbf{H}} \mathbf{H}_0)_{qj}}; \quad (\mathbf{W}_0)_{iq} \leftarrow (\mathbf{W}_0)_{iq} \frac{(\mathbf{A}_0 \mathbf{H}_0^\top)_{iq}}{(\mathbf{W}_0 (\mathbf{H}_0 \mathbf{H}_0^\top) + \lambda_{\mathbf{W}} \mathbf{W}_0)_{iq}}. \quad (13)$$

(v) Set the initial joint data matrix  $\mathbf{A}_{0:0} = \mathbf{A}_0$  and projection matrix  $\mathbf{H}_{0:0} = \mathbf{H}_0$ .

2: **Slice-wise increments:** For each slice  $z = [1, \dots, Z]$ ,

- (i) Form the matrix  $\mathbf{A}_z$  from the context vectors of the slice  $z$  and extend the joint data matrix  $\mathbf{A}_{0:z} = [\mathbf{A}_{0:z-1} \ \mathbf{A}_z]$ .
- (ii) Initialize the slice-wise projection matrix  $\mathbf{H}_z$  randomly and extend the joint projection matrix  $\mathbf{H}_{0:z} = [\mathbf{H}_{0:z-1} \ \mathbf{H}_z]$ .
- (iii) Initialize the basis matrix  $\mathbf{W}_z \leftarrow \mathbf{W}_{z-1}$  and update  $\mathbf{H}_z$  and  $\mathbf{W}_z$  iteratively until convergence to a local minimum of the reconstruction error:

$$\begin{aligned} (\mathbf{H}_z)_{qj} &\leftarrow (\mathbf{H}_z)_{qj} \frac{(\mathbf{W}_z^\top \mathbf{A}_z)_{qj}}{(\mathbf{W}_z^\top \mathbf{W}_z \mathbf{H}_z + \lambda_{\mathbf{H}} \mathbf{H}_z)_{qj}}; \\ (\mathbf{W}_z)_{iq} &\leftarrow (\mathbf{W}_z)_{iq} \frac{(\mathbf{A}_{0:z-1} \mathbf{H}_{0:z-1}^\top + \mathbf{A}_z \mathbf{H}_z^\top)_{iq}}{(\mathbf{W}_z (\mathbf{H}_{0:z-1} \mathbf{H}_{0:z-1}^\top + \mathbf{H}_z \mathbf{H}_z^\top) + \lambda_{\mathbf{W}} \mathbf{W}_z)_{iq}}. \end{aligned} \quad (14)$$

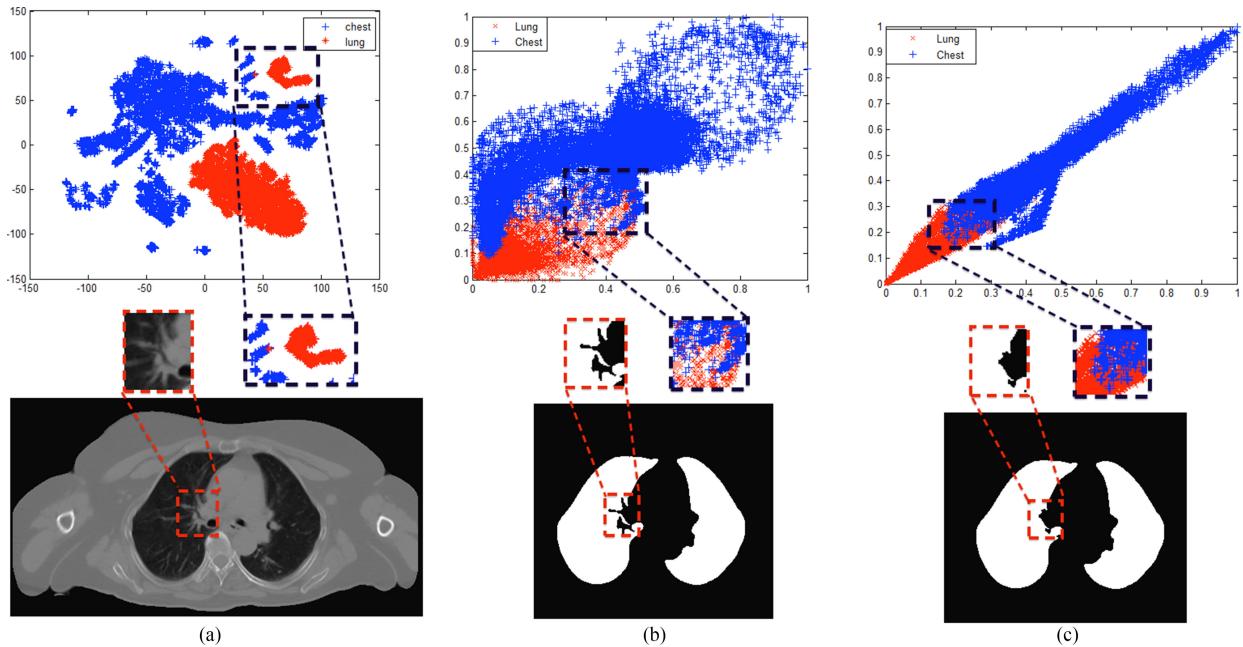
3: **Output:** The joint projection matrix  $\mathbf{H}_{0:Z}$  describing visual appearance of the image  $\mathbf{g}$  by approximation of the voxel neighborhoods with linear combinations of the  $r$  features specified by the basis matrix  $\mathbf{W}_Z$ .

Fig. 3. Signal distributions for segmented voxels in the original 27-dimensional space (a) and the  $r$ -dimensional spaces reduced with the NMF (b) and ICNMF (c) visualized using the t-SNE projection [37]. The better ICNMF performance is exemplified by signal distributions and segmentation results for pathologies on the lung-chest border.

is added. The updates of (12) provably guarantee that iterations at every step  $k$  converge to a local minimum of the reconstruction error [22], [36]. For completeness, Appendix B shows how the multiplicative updates are derived using the conditional Lagrange optimization.

Computational complexity of the ICNMF (like the INMF) is  $O(mr^2)$  per iteration, comparing with  $O(nmr)$  for the NMF. Because the running time depends linearly on the number  $n$  of

samples for the NMF, but is independent of this number for the ICNMF and INMF, the latter are more suitable for learning the factors  $\mathbf{W}$  and  $\mathbf{H}$  to represent a large collection of data samples. Moreover, as follows from (12), the learning computations can be simplified because both the matrices  $\mathbf{A}_k$  and  $\mathbf{H}_k$  do not change after adding every new data sample  $\mathbf{a}_{k+1}$ . Thus, instead of keeping separately the growing matrices  $\mathbf{A}_k$  and  $\mathbf{H}_k$ , only their fixed-size products  $\mathbf{A}_k \mathbf{H}_k^\top$  and  $\mathbf{H}_k \mathbf{H}_k^\top$  have to be stored.

TABLE I  
SEGMENTATION ACCURACY FOR DIFFERENT WEIGHTS IN (2): THE BEST COLUMN/ROW SCORES ARE BOLD-FACED/ITALICIZED

$\lambda_W$	$\lambda_H$	0	0.1	0.3	1	3	10	30	100	300	1000
0		0.821	0.930	0.892	0.898	<b>0.961</b>	<b>0.955</b>	0.947	0.941	0.955	0.955
0.1		0.850	0.942	0.940	0.935	0.953	0.936	0.928	0.953	0.955	0.955
0.3		0.839	<b>0.953</b>	0.952	0.940	0.953	0.953	0.953	0.953	0.959	0.954
1		0.847	<b>0.953</b>	<b>0.953</b>	<b>0.953</b>	0.952	0.947	0.953	<b>0.958</b>	<b>0.962</b>	0.959
3		<b>0.851</b>	<b>0.953</b>	0.952	0.951	0.953	0.953	0.952	0.953	0.957	0.954
10		0.848	<b>0.953</b>	0.951	<b>0.953</b>	0.952	0.952	0.953	0.953	0.955	0.954
30		0.848	0.952	0.952	<b>0.953</b>	0.953	0.953	<b>0.954</b>	0.951	0.955	0.955
100		0.849	0.949	<b>0.953</b>	0.952	0.953	0.952	0.953	0.952	0.957	0.955

The above updating process holds (with mostly notational changes), if the matrix  $\mathbf{A}_k$  is appended at each step  $k+1$  with not a single vector  $\mathbf{a}_{k+1}$  but a small-size data matrix, e.g., a context image for the next 2-D slice in the lung segmentation.

### B. ICNMF-Based Lung Segmentation

Fig. 1 and Algorithm 1 outline the proposed segmentation in a 3-D CT chest image by removing image background; describing visual appearance of the remaining chest-lung image with the ICNMF, and extracting 3-D lung voxels by data clustering and cleaning the region map. The last two stages are detailed below.

Let  $\mathbb{R}$  and  $\mathbb{Q} = \{0, \dots, Q\}$  denote a finite arithmetic lattice supporting 3-D digital images and their region maps:

$$\mathbb{R} = \{(x, y, z) : x = 0, \dots, X; y = 0, \dots, Y; z = 0, \dots, Z\}$$

and a finite set of integer voxel-wise intensities, or gray values, respectively. To describe visual appearance of objects-of-interest with the ICNMF, the context image [34] is built from an original 3-D image  $\mathbf{g} = \{g_{x,y,z} : (x, y, z) \in \mathbb{R}; g_{x,y,z} \in \mathbb{Q}\}$ . Each voxel  $(x, y, z)$  of the context image supports the context vector  $\mathbf{a}_{x,y,z}$ , containing the intensities for this voxel and its nearest 3-D neighbors in the original image  $\mathbf{g}$ , e.g., the 27 intensities in total for the nearest  $3 \times 3 \times 3$  neighborhood. Each context vector, exemplified in Fig. 2, forms a column of the data matrix  $\mathbf{A}$ .

To minimize the constrained reconstruction error of (7) after adding each next context vector  $\mathbf{a}_{k+1}$  (see Section II-A), the ICNMF uses the iterative multiplicative algorithm of (12) producing the goal descriptor  $\mathbf{h}_{k+1}$  of this voxel and updating the basis matrix  $\mathbf{W}$ . However, repeating such computations for all the voxels of a large 3-D CT image is too expensive. To reduce the complexity, the CT slices are processed sequentially, but all the voxels of every next slice are added to the matrix  $\mathbf{A}$  simultaneously. The above ICNMF algorithm remains almost the same, apart of considering the appended data sample  $\mathbf{a}_{k+1}$  and its projection  $\mathbf{h}_{k+1}$  matrices, rather than vectors. In this case, the already computed optimal basis  $\mathbf{W}_z$  of the previous  $z$  slices initializes updating the basis  $\mathbf{W}_{z+1}$  for the  $z+1$  slices. Algorithm 2 outlines the proposed version of the ICNMF.

After projecting the context vectors to the feature space with the ICNMF, the  $K$ -means clustering is applied to the voxel-wise vectorial projections in the joint matrix  $\mathbf{H}_{0:Z}$  in order to form a prescribed number  $K$  of data clusters ( $K=2$  for separating the

lung and chest tissues). The Frobenius norm of centroid of each cluster helps to identify the lungs by their relative brightness in the image: the darker the voxel and its neighborhood, i.e., the smaller their context vector norm, the closer its  $\mathbf{h}$ -descriptor to the origin. Then, the segmentation is refined by analyzing 3-D connected components to keep connectivity inside the lung region. In so doing, an initial seed voxel is chosen outside the segmented lung.

To highlight capabilities of the ICNMF in revealing characteristic inter-voxel dependencies, Fig. 3 compares distributions of the original voxel-wise context vectors in the 27-dimensional signal space for the nearest  $3 \times 3 \times 3$  neighborhoods and their projections to the reduced  $r$ -dimensional  $\mathbf{h}$ -spaces formed by the NMF and ICNMF. Due to much better separation of the lung and chest voxels than in the original space, both the NMF- and ICNMF-based  $\mathbf{h}$ -spaces allow for more accurate data clustering and segmentation. However, the ICNMF constrains the signal distributions more effectively than the NMF, and thus, forms smoother and tighter lung and chest manifolds, which are easier for separation. As shown in Fig. 3, selecting the neighborhood size in accord with characteristic spatial dependencies between the lung voxels allows the ICNMF to accurately capture inhomogeneities and pathologies at the lung-chest border, where the conventional 3-D connected component analysis often fails.

### III. EXPERIMENTS WITH SYNTHETIC AND REAL DATA

The proposed ICNMF-based segmentation was evaluated on both real (*in-vivo*) and synthetic data using three common performance metrics detailed, for completeness, in Appendix A: the Dice similarity coefficient (DSC) [38], the modified 95-percentile Hausdorff distance (MHD) [39], and the absolute lung volume difference (ALVD).

Synthetic 3-D phantoms [40] simplify initial performance testing because accurate lung borders on real CT images are very difficult to obtain manually due to the observers' variability. The phantom images mimic visual appearance of the real 3-D CT data by Gibbs sampling of a learned generalized 3-D Gauss-Markov random field model [41].

The *in-vivo* 3-D CT image datasets for 17 patients (each set of the size of  $512 \times 512 \times 390$  voxels) have been acquired with a multidetector GE Light Speed Plus scanner (General Electric, Milwaukee, WI, USA) using the following scanning protocol: the 2.5-mm-thick slices reconstructed every 5 mm; the scanning pitch of 1.5; 140 KV; 100 MA; and F.O.V of 36 cm. The segmentation separated two objects: the darker lung tissues and the brighter chest tissues.

Table I shows impacts of the sparseness and smoothness constraints of (2) on the segmentation accuracy in terms of the DSC for a real dataset. The INMF-based segmentation with zero constraining weights,  $\lambda_W = \lambda_H = 0$ , in (5) is distinctly less accurate than the proposed ICNMF-based one. The sparseness of  $\mathbf{H}$  affects the accuracy more, than the smoothness of  $\mathbf{W}$ , but the accuracy varies insignificantly and nonmonotonously for a variety of the weight combinations, so that selecting the best pair calls for a too long experimentation. However, the DSCs in Table I indicate that increasing the sparseness weight  $\lambda_H$

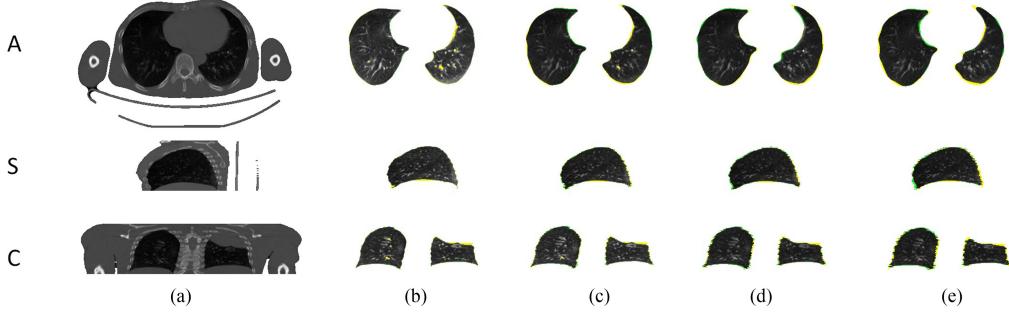


Fig. 4. Original 2-D CT slices (a) and segmented lungs for  $3 \times 3 \times 3$  (b),  $7 \times 7 \times 3$  (c),  $11 \times 11 \times 3$  (d), and  $15 \times 15 \times 3$  (e) neighborhoods in the axial (A), sagittal (S), and coronal (C) planes. False positive (FP) and negative (FN) voxels are highlighted in green and yellow, respectively.

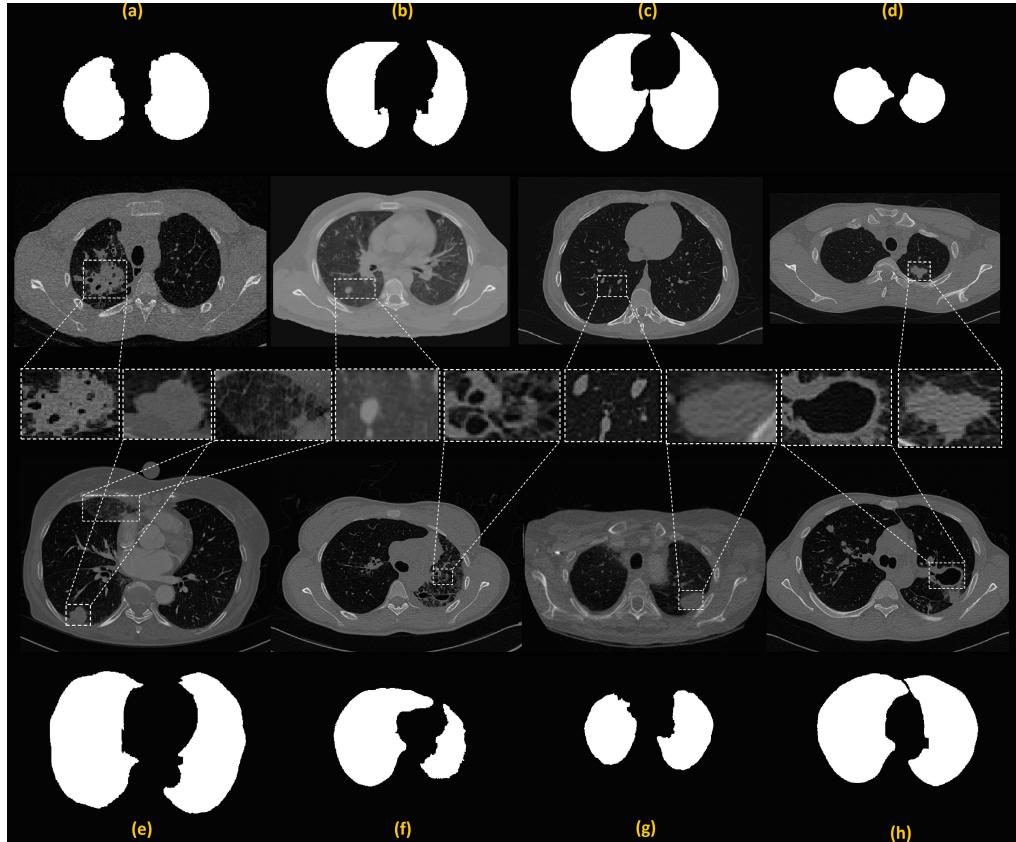


Fig. 5. (a) Lungs segmented with Algorithm 1 using the  $3 \times 3 \times 3$  neighborhood on pulmonary CT in the cases of airspace consolidation, (b) tree-in-bud and micro-nodules, (c) usual nodules, (d) cancer, (e) ground-glass opacity and juxtapleural nodules, (f) honeycomb, (g) diffuse consolidation, and (h) cavity.

for both zero and nonzero smoothness weight  $\lambda_W$  improves the segmentation accuracy, although the improvement becomes more stable when both the smoothness and sparseness are constrained. Constraining the encoding vectors in the  $h$ -space has to be limited because the accuracy decreases for over-smoothed and/or too sparse voxel manifolds. According to the experiments in Table I, the pair  $\lambda_W = 1$  and  $\lambda_H = 300$  yields the most stable top accuracy of 0.962. For the nonzero smoothness weights  $\lambda_W > 0$  in these experiments, the lower sparseness weight ( $\lambda_H < 300$ ) gradually decreases the DSC, whereas the larger one ( $\lambda_H > 300$ ) either does not affect, or also decreases the accuracy.

The segmentation accuracy was also tested for different numbers  $r \in \{2, 4, 6, 8\}$  of the basis vectors, the best result

TABLE II  
COMPARATIVE ACCURACY (DSC) OF OUR ALGORITHM 1 WITH THE  $3 \times 3 \times 3$  NEIGHBORHOODS ON THE *in vivo* DATASETS

Algorithm	Mean $\pm$ std	<i>p</i> -value
Algorithm 1	$0.969 \pm 0.010$	
MRS [42]	$0.783 \pm 0.078$	0.0001
IT [15]	$0.816 \pm 0.091$	0.0001
GVF [43]	$0.848 \pm 0.087$	0.0003

having been achieved for  $r = 4$ . Obviously, the accuracy depends also on the neighborhood size and shape. Comparative experiments with the  $(3 \times 3 \times 3)$ ,  $(7 \times 7 \times 3)$ ,  $(11 \times 11 \times 3)$ , and  $(15 \times 15 \times 3)$  voxel neighborhoods in Fig. 4 have shown

TABLE III  
ACCURACY (mean $\pm$ std [p-VALUE]) OF ALGORITHM 1 WITH THE  $(3 \times 3 \times 3)$  NEIGHBORHOODS W.R.T. THE NMF-BASED SEGMENTATION [20] ON SYNTHETIC AND REAL (*in vivo*) DATASETS

	DSC		ALVD		MHD		Time(sec)	
	Real data	Synthetic data	Real data	Synthetic data	Real data	Synthetic data	Real data	Synthetic data
Algorithm 1	<b>0.96<math>\pm</math>0.01</b>	<b>0.97<math>\pm</math>0.01</b>	<b>0.87<math>\pm</math>0.62</b>	<b>0.51<math>\pm</math>0.07</b>	<b>9.0<math>\pm</math>0.001</b>	<b>4.8<math>\pm</math>0.006</b>	<b>180</b>	<b>140</b>
NMF [20]	0.95 $\pm$ 0.02 [0.027]	0.96 $\pm$ 0.01 [0.50]	2.4 $\pm$ 1.1 [ $< 0.0001$ ]	2.4 $\pm$ 1.1 [0.004]	9.7 $\pm$ 0.010 [ $< 0.0001$ ]	5.9 $\pm$ 0.005 [ $< 0.0001$ ]	365	310

that the more expanded the neighborhood, the lesser the segmentation accuracy. The increasing segmentation errors highlighted in green and yellow can be explained in part by higher similarity between the larger neighborhoods for the adjacent voxels along the lung-chest boundary. Also, a number of the CT scans of patients with different lung diseases [12] were segmented in order to evaluate the performance of Algorithm 1 in the case of severe lung pathologies. The eight CT scans selected in Fig. 5 demonstrate diverse pulmonary patterns, such as, e.g., caused by airspace or diffuse consolidation, cancer, different nodules, including juxtapleural ones, etc. Algorithm 1 adapts successfully to such pathologies.

Table II shows the higher accuracy of Algorithm 1 w.r.t. three other segmentation algorithms, by comparing their mean values and standard deviations of the DSCs using the statistical paired *t*-test. The algorithms for comparisons include the multiresolution segmentation [42]; segmentation with iterative thresholding (IT) followed by morphological operations [15], and segmentation with a deformable boundary guided by the gradient vector flow (GVF) [43].

Table III compares Algorithm 1 with our earlier NMF-based segmentation [20] on the 17 real and 7 synthetic datasets, using all three performance metrics. By the DSC, the algorithms differ insignificantly for the synthetic data, whereas Algorithm 1 demonstrates a small, but statistically significant improvement for the real data (the mean and standard deviation  $0.96\pm0.01$  versus  $0.95\pm0.02$ ). But by the ALVD and MHD, the ICNMF-based Algorithm 1 outperforms the NMF-based one on both the real and synthetic datasets.

The sensitivity of Algorithm 1 against selecting the constraining weights, number of the basis vectors, and neighborhood size was evaluated using statistics of the receiver operating characteristic curves, in particular, the area under the curve (AUC). The AUC for  $\lambda_H$ ,  $\lambda_W$ ,  $r$  and the neighborhood size ( $N_{x,y,z}$ ) was, respectively, 0.98, 0.98, 0.97, and 0.99, demonstrating both good performance and low sensitivity of the algorithm. Also, as was shown experimentally, different arrangements of the 3-D context voxels in the data vector  $a$ , do not affect the segmentation results.

To demonstrate its applicability to the data collected by various scanning protocols, our Algorithm 1 was also tested on 55 real chest 3-D CT scans provided by the LOLA11 challenge ([www.lola11.com](http://www.lola11.com)) and acquired at different places with several scanners, scanning protocols, and reconstruction parameters. To evaluate our results using the LOLA11 dataset, we removed the trachea and main bronchi, and if needed, separated the lung by finding a maximum cost path in connected

TABLE IV  
OVERLAPS WITH THE TRUE LEFT (LL) AND RIGHT LUNGS (RL) FOR ALGORITHM 1 (A1) W.R.T. A HUMAN EXPERT AND THE MOST (RANK 1), MEDIAN (RANK 7), AND LEAST ACCURATE (RANK 13) LOLA11 SEGMENTATION OF THE 55 DATASETS : THE MEAN, STANDARD DEVIATION (std), MINIMUM (min), 25%-QUARTILE (Q1), MEDIAN (med), 75%-QUARTILE (Q3), AND MAXIMUM (max) OVERLAPS

Algorithm	Object	mean	std	min	Q1	med	Q3	max
A1 [rank 5]	Average score: <b>0.965</b>							
	LL	0.965	0.108	0.205	0.981	0.988	0.992	0.998
	RL	0.964	0.133	0.010	0.982	0.988	0.991	0.997
Human	Average score: <b>0.984</b>							
	LL	0.984	0.031	0.782	0.987	0.992	0.996	0.998
	RL	0.984	0.047	0.662	0.988	0.995	0.997	0.999
Rank 1 [10]	Average score: <b>0.973</b>							
	LL	0.974	0.097	0.277	0.987	0.992	0.995	0.999
	RL	0.972	0.135	0.000	0.991	0.994	0.996	0.999
Rank 7 [12]	Average score: <b>0.955</b>							
	LL	0.957	0.137	0.034	0.979	0.987	0.995	0.999
	RL	0.952	0.151	0.000	0.984	0.990	0.997	0.999
Rank 13 [45]	Average score: <b>0.939</b>							
	LL	0.929	0.154	0.083	0.945	0.974	0.983	0.995
	RL	0.950	0.121	0.150	0.960	0.978	0.988	0.994

axial slices as in [44]. Table IV presents the results in comparison with the best, median, and worst results for the 13 lung segmentation algorithms participated in the LOLA11 challenge for 2011–2014. Selected examples of the lung region maps obtained by Algorithm 1 are shown in Fig. 6. To further demonstrate the algorithm’s performance, Fig. 7 visualizes the segmented 3-D lungs for selected subjects.

It should be noted that several pathologies in the LOLA11 dataset are far outside capabilities of the proposed simple visual appearance descriptors accounting for only the nearest-neighbor relations of the voxels. Accurate segmentation of such pathological lungs requires much more profound lung and chest models. Moreover, there is no consensus of the medical imaging community on whether the pleural fluid should be considered as a part of the lung field as it is done in the LOLA11 ground truth [12]. Because our ICNMF-based Algorithm 1 does not include the pleural fluid to the lungs, by the overall accuracy of 0.965 (the relative overlap with the ground truth) it has the fifth rank among all the LOLA11 contestants. However, without the relevant nine pathological subjects from the LOLA11 dataset it achieves the top-rank accuracy of 0.986 for the remaining 46 subjects.

In terms of the algorithm complexity, the proposed ICNMF-based segmentation extracts voxel-wise features in a completely unsupervised mode, using only a few parameters, such as, e.g.,  $\lambda_W$ ,  $\lambda_H$ ,  $r$ , and  $N_{x,y,z}$ , whereas the conventional top-ranked techniques [10]–[12] comprise specific feature

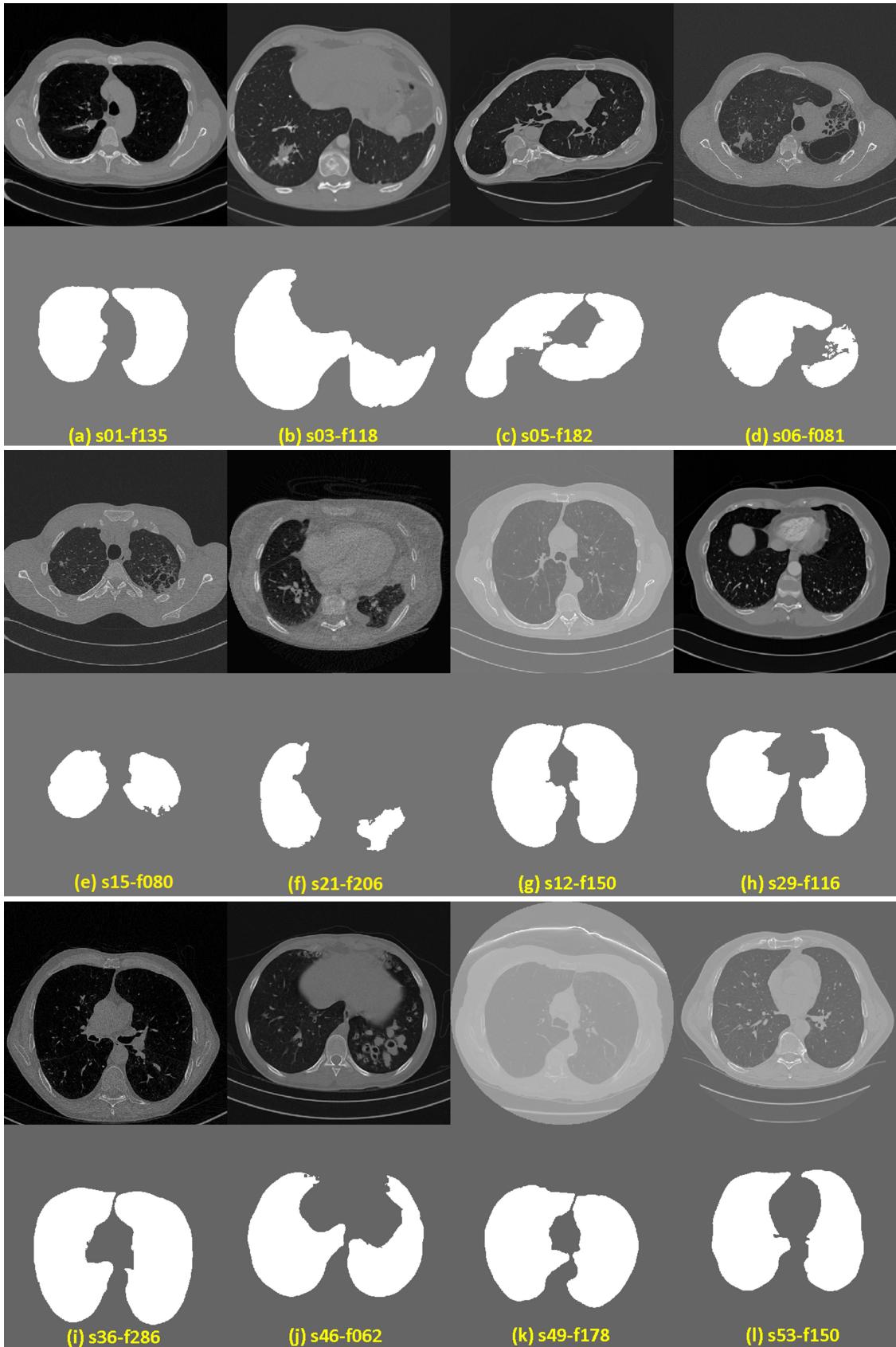


Fig. 6. 3-D lungs segmented with Algorithm 1 using the  $3 \times 3 \times 3$  neighborhoods: the CT slice (“s” and “f” indicate the scan and slice numbers, respectively) versus the corresponding lung region map (white) in the LOLA11 dataset.

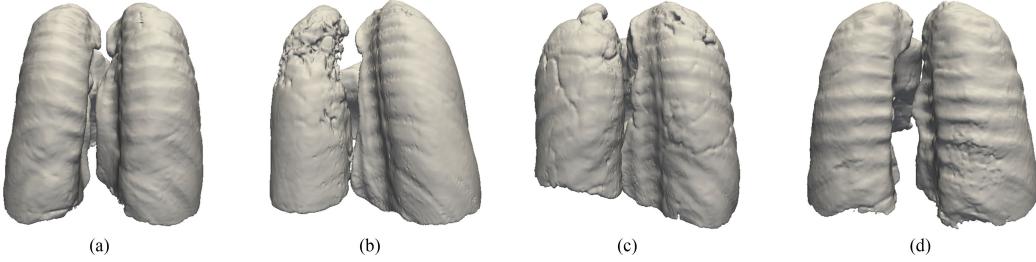


Fig. 7. 3-D visualization of segmented lungs from selected LOLA11 subjects (“s” indicates the scan) (a) s1 (b) s4 (c) s6 (d) s11.

engineering steps, which require proper initialization and parameter tuning.

#### IV. CONCLUSION

Our experiments with both *in-vivo* and synthetic images confirmed that the proposed ICNMF-based 3-D lung segmentation outperforms the existing segmentation approaches, including the NMF-based one, by the DSC, ALVD, and MHD measures of accuracy. Its better performance stems from the more accurate compressed description of characteristic spatial signal dependencies in every input image. The ICNMF reveals the robust features (columns of the basis nonnegative matrix  $\mathbf{W}$ ) for encoding the voxel neighborhoods (contexts) with smooth descriptors—columns of the nonnegative projection matrix  $\mathbf{H}$ . To make the description process computationally feasible for a typically very large 3-D CT image, the ICNMF combines the conventional INMF and CNMF. The optimal basis and projection matrices are estimated incrementally, while constraining their Frobenius norms enforce their smoothness and sparseness, respectively. Testing on the 3-D chest CT images provided by the LOLA11 challenge, collected with different scanners, scanning protocols, and reconstruction parameters, scores our algorithm sufficiently high among the 13 known state-of-the-art methods.

#### APPENDIX A SEGMENTATION PERFORMANCE METRICS

The DSC [38] measures an overlap between the segmented object and its ground truth:

$$\text{DSC} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (15)$$

where TP, FP, and FN denote the true positive, FP, and FN correspondences, respectively. The higher the DSC, the better the segmentation, i.e., the closer the match to the ground truth: DSC = 0 indicates no overlap, and DSC = 1 characterizes the ideal agreement.

The ALVD is the percentage volume difference between the segmented region and its ground truth:

$$\text{ALVD} = \frac{\text{FN} + \text{FP}}{\text{FN} + \text{TP}} \times 100\%.$$

The MHD characterizes linear distances (in voxel units) between the segmented object and ground truth. The HD [39] from a point set  $A_1$  to a point set  $A_2$  is the maximum distance from the points of the set  $A_1$  to their nearest points in the set  $A_2$ :

$$H(A_1, A_2) = \max_{p \in A_1} \left\{ \min_{r \in A_2} \{D(p, r)\} \right\}$$

where  $D(p, r)$  is the Cartesian distance between the points  $p$  and  $r$ . The bidirectional HD is defined as  $H_{\text{bi}}(A_1, A_2) = \max \{H(A_1, A_2), H(A_2, A_1)\}$ . The MHD, used in this paper to measure the segmentation accuracy, is defined as the 95th-percentile bidirectional HD between the segmented region and its ground truth, such that selecting the maximum distance in the HD is replaced by selecting the 95th-percentile one.

#### APPENDIX B UPDATE RULE DERIVATION USING LAGRANGE MULTIPLIERS

To minimize the constrained reconstruction error  $F_{k+1}$ , (7) can be rewritten as

$$\begin{aligned} F_{k+1} = & \text{Tr}((\mathbf{A}_k - \mathbf{W}_{k+1}\mathbf{H}_k)(\mathbf{A}_k - \mathbf{W}_{k+1}\mathbf{H}_k)^T) \\ & + \text{Tr}((\mathbf{a}_{k+1} - \mathbf{W}_{k+1}\mathbf{h}_{k+1})(\mathbf{a}_{k+1} - \mathbf{W}_{k+1}\mathbf{h}_{k+1})^T) \\ & + \lambda_{\mathbf{W}} \text{Tr}(\mathbf{W}_{k+1}\mathbf{W}_{k+1}^T) + \lambda_{\mathbf{H}} \text{Tr}(\mathbf{H}_k\mathbf{H}_k^T) \\ & + \lambda_{\mathbf{H}} \text{Tr}(\mathbf{h}_{k+1}\mathbf{h}_{k+1}^T) \end{aligned} \quad (16)$$

and simplified further by the trace properties  $\text{Tr}(\mathbf{YZ}) = \text{Tr}(\mathbf{ZY})$  and  $\text{Tr}(\mathbf{Z}) = \text{Tr}(\mathbf{Z}^T)$ :

$$\begin{aligned} F_{k+1} = & \text{Tr}(\mathbf{A}_k\mathbf{A}_k^T) - 2\text{Tr}(\mathbf{W}_{k+1}\mathbf{H}_k\mathbf{A}_k^T) \\ & - \text{Tr}(\mathbf{W}_{k+1}\mathbf{H}_k\mathbf{H}_k^T\mathbf{W}_{k+1}^T) + \text{Tr}(\mathbf{a}_{k+1}\mathbf{a}_{k+1}^T) \\ & - 2\text{Tr}(\mathbf{a}_{k+1}\mathbf{h}_{k+1}^T\mathbf{W}_{k+1}^T) \\ & - \text{Tr}(\mathbf{W}_{k+1}\mathbf{h}_{k+1}\mathbf{h}_{k+1}^T\mathbf{W}_{k+1}^T) \\ & + \lambda_{\mathbf{W}} \text{Tr}(\mathbf{W}_{k+1}\mathbf{W}_{k+1}^T) + \lambda_{\mathbf{H}} \text{Tr}(\mathbf{H}_k\mathbf{H}_k^T) \\ & + \lambda_{\mathbf{H}} \text{Tr}(\mathbf{h}_{k+1}\mathbf{h}_{k+1}^T). \end{aligned} \quad (17)$$

Let  $\Psi = [\psi_{ik}]$  and  $\Phi = [\phi_{jk}]$  be matrices of the Lagrange multipliers  $\psi_{ik}$  and  $\phi_{jk}$  for the constraints  $w_{ik} \geq 0$  and  $h_{jk} \geq 0$ , respectively. Then, the matrix properties  $\text{Tr}(\mathbf{AB}) = \text{Tr}(\mathbf{BA})$  and  $\text{Tr}(\mathbf{A}) = \text{Tr}(\mathbf{A}^T)$  lead to the following Lagrangian  $\mathcal{L}$  for (17):

$$\begin{aligned} \mathcal{L} = & \text{Tr}(\mathbf{A}_k\mathbf{A}_k^T) - 2\text{Tr}(\mathbf{W}_{k+1}\mathbf{H}_k\mathbf{A}_k^T) \\ & - \text{Tr}(\mathbf{W}_{k+1}\mathbf{H}_k\mathbf{H}_k^T\mathbf{W}_{k+1}^T) \\ & + \text{Tr}(\mathbf{a}_{k+1}\mathbf{a}_{k+1}^T) - 2\text{Tr}(\mathbf{a}_{k+1}\mathbf{h}_{k+1}^T\mathbf{W}_{k+1}^T) \\ & - \text{Tr}(\mathbf{W}_{k+1}\mathbf{h}_{k+1}\mathbf{h}_{k+1}^T\mathbf{W}_{k+1}^T) \\ & + \lambda_{\mathbf{W}} \text{Tr}(\mathbf{W}_{k+1}\mathbf{W}_{k+1}^T) \\ & + \lambda_{\mathbf{H}} \text{Tr}(\mathbf{H}_k\mathbf{H}_k^T) + \lambda_{\mathbf{H}} \text{Tr}(\mathbf{h}_{k+1}\mathbf{h}_{k+1}^T) \\ & + \text{Tr}(\Phi\mathbf{H}_{k+1}^T) + \text{Tr}(\Psi\mathbf{W}_{k+1}^T). \end{aligned} \quad (18)$$

The partial derivatives of  $\mathcal{L}$  with respect to  $\mathbf{W}$  and  $\mathbf{H}$  are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \mathbf{W}_{k+1}} = & -2\mathbf{A}_k \mathbf{H}_k^\top + 2\mathbf{W}_{k+1} \mathbf{H}_k \mathbf{H}_k^\top \\ & -2\mathbf{a}_{k+1} \mathbf{h}_{k+1}^\top + 2\mathbf{W}_{k+1} \mathbf{h}_{k+1} \mathbf{h}_{k+1}^\top \\ & + 2\lambda_{\mathbf{W}} \mathbf{W}_{k+1} + \Psi \end{aligned} \quad (19)$$

and

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \mathbf{h}_{k+1}} = & -2\mathbf{W}_{k+1}^\top \mathbf{a}_{k+1} + 2\mathbf{W}_{k+1}^\top \mathbf{W}_{k+1} \mathbf{h}_{k+1} \\ & + 2\lambda_{\mathbf{H}} \mathbf{h}_{k+1} + \phi_{k+1} \end{aligned} \quad (20)$$

where  $\phi_{k+1}$  denotes the Lagrange multipliers for  $\mathbf{h}_{k+1}$ . Applying the classical Karush–Kuhn–Tucker conditions  $\phi_{iq} w_{iq} = 0$  and  $\psi_{qj} h_{qj} = 0$  yields the following equations for  $w_{iq}$  and  $h_{qj}$ :

$$\begin{aligned} & -(\mathbf{A}_k \mathbf{H}_k^\top + \mathbf{a}_{k+1} \mathbf{h}_{k+1}^\top)_{iq} w_{iq} \\ & + (\mathbf{W}_{k+1} \mathbf{H}_k \mathbf{H}_k^\top + \mathbf{W}_{k+1} \mathbf{h}_{k+1} \mathbf{h}_{k+1}^\top)_{iq} w_{iq} \\ & + \lambda_{\mathbf{W}} \mathbf{W}_{k+1})_{iq} w_{iq} = 0 \end{aligned} \quad (21)$$

and

$$\begin{aligned} & -(\mathbf{W}_{k+1}^\top \mathbf{a}_{k+1} + \lambda_{\mathbf{H}} \mathbf{h}_{k+1})_q h_q \\ & + (\mathbf{W}_{k+1}^\top \mathbf{W}_{k+1} \mathbf{h}_{k+1})_q h_q = 0. \end{aligned} \quad (22)$$

These equations lead to the following updating rules:

$$(\mathbf{W}_{k+1})_{iq} \leftarrow (\mathbf{W}_{k+1})_{iq} \frac{(\mathbf{A}_k \mathbf{H}_k^\top + \mathbf{a}_{k+1} \mathbf{h}_{k+1}^\top)_{iq}}{S_{iq}} \quad (23)$$

and

$$(\mathbf{h}_{k+1})_q \leftarrow (\mathbf{h}_{k+1})_q \frac{(\mathbf{W}_{k+1}^\top \mathbf{a}_{k+1})_q}{(\mathbf{W}_{k+1}^\top \mathbf{W}_{k+1} \mathbf{h}_{k+1})_q + \lambda_{\mathbf{H}} \mathbf{h}_{k+1})_q} \quad (24)$$

where  $S_{iq} = (\mathbf{W}_{k+1} (\mathbf{H}_k \mathbf{H}_k^\top + \mathbf{h}_{k+1} \mathbf{h}_{k+1}^\top) + \lambda_{\mathbf{W}} \mathbf{W}_{k+1})_{iq}$ . (23) and (24) prove that the gradient descent updates in Section II-A are correct.

## REFERENCES

- [1] A. El-Baz and J. S. Suri, *Lung Imaging and Computer Aided Diagnosis*, Boca Raton, FL, USA: CRC Press, 2011.
- [2] A. El-Baz *et al.*, “Automatic analysis of 3-D low dose CT images for early diagnosis of lung cancer,” *Pattern Recognit.*, vol. 42, no. 6, pp. 1041–1051, 2009.
- [3] Q. Wei and Y. Hu, “A hybrid approach to segmentation of diseased lung lobes,” *IEEE J. Biomed. Health Informat.*, vol. 18, no. 5, pp. 1696–1706, Sep. 2014.
- [4] P. Korfiatis *et al.*, “Texture-based identification and characterization of interstitial pneumonia patterns in lung multidetector CT,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 675–680, May 2010.
- [5] Q. Wei *et al.*, “Segmentation of lung lobes in high-resolution isotropic ct images,” *IEEE Trans. Biomed. Eng.*, vol. 56, no. 5, pp. 1383–1393, May 2009.
- [6] Y. Itai *et al.*, “A segmentation method of lung areas by using snakes and automatic detection of abnormal shadow on the areas,” *Int. J. Innovative Comput. Inf. Control*, vol. 3, no. 2, pp. 277–284, 2007.
- [7] I. Sluimer *et al.*, “Toward automated segmentation of the pathological lung in CT,” *IEEE Trans. Med. Imag.*, vol. 24, no. 8, pp. 1025–1038, Aug. 2005.
- [8] A. El-Baz *et al.*, “A new stochastic framework for accurate lung segmentation,” in *Proc. Med. Image Comput. Comput.-Assisted Intervention*, 2008, vol. 5241, pp. 322–330.
- [9] A. El-Baz *et al.*, “Computer-aided diagnosis systems for lung cancer: Challenges and methodologies,” *Int. J. Biomed. Imag.*, vol. 2013, art. no. 942353 (46 pages), 2013.
- [10] B. Lassen *et al.*, “Lung and lung lobe segmentation methods at Fraunhofer MEVIS,” in *Proc. 4th Int. MICCAI Workshop Pulmonary Image Anal.*, Toronto, ON, Canada, 2011, pp. 185–199.
- [11] O. Weinheimer *et al.*, “Automatic lung segmentation in MDCT images,” in *Proc. 4th Int. MICCAI Workshop Pulmonary Image Anal.*, Toronto, ON, Canada, 2011, pp. 241–255.
- [12] A. Mansoor *et al.*, “A generic approach to pathological lung segmentation,” *IEEE Trans. Med. Imag.*, vol. 33, no. 12, pp. 2293–2310, Dec. 2014.
- [13] P. Lo *et al.*, “Historic automated lung segmentation method: Performance on LOLA11 data set,” in *Proc. 4th Int. MICCAI Workshop Pulmonary Image Anal.*, Toronto, ON, Canada, 2011, pp. 257–260.
- [14] E. van Rikxoort and B. van Ginneken, “Automatic segmentation of the lungs and lobes from thoracic CT scans,” in *Proc. 4th Int. MICCAI Workshop Pulmonary Image Anal.*, Toronto, ON, Canada, 2011, pp. 261–268.
- [15] S. Hu *et al.*, “Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images,” *IEEE Trans. Med. Imag.*, vol. 20, no. 6, pp. 490–498, Jun. 2001.
- [16] S. Sun *et al.*, “Robust active shape model based lung segmentation in CT scans,” in *Proc. 4th Int. MICCAI Workshop Pulmonary Image Anal.*, Toronto, ON, Canada, 2011, pp. 213–224.
- [17] Y. Xie *et al.*, “Nonnegative factorization of diffusion tensor images and its applications,” in *Proc. Inf. Process. Med. Imag.*, 2011, vol. 6801, pp. 550–561.
- [18] C. Lazar *et al.*, “Non negative matrix factorisation clustering capabilities; application on multivariate image segmentation,” *Int. J. Bus. Intell. Data Mining*, vol. 5, no. 3, pp. 285–296, Jun. 2010.
- [19] R. Sandler and M. Lindenbaum, “Nonnegative matrix factorization with Earth mover’s distance metric for image analysis,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1590–1602, Aug. 2011.
- [20] E. Hosseini-Asl *et al.*, “Lung segmentation based on nonnegative matrix factorization,” in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2014, pp. 877–881.
- [21] D. D. Lee and H. S. Seung, “Learning the parts of objects by non-negative matrix factorization,” *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [22] D. D. Lee and H. S. Seung, “Algorithms for non-negative matrix factorization,” in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, 2000, pp. 556–562.
- [23] P. O. Hoyer, “Non-negative sparse coding,” in *Proc. 12th IEEE Workshop Neural Netw. Signal Process.*, 2002, pp. 557–565.
- [24] P. O. Hoyer, “Non-negative matrix factorization with sparseness constraints,” *J. Mach. Learning Res.*, vol. 5, pp. 1457–1469, 2004.
- [25] J. Eggert and E. Korner, “Sparse coding and NMF,” in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jul. 2004, vol. 4, pp. 2529–2533.
- [26] A. Cichocki *et al.*, “Flexible component analysis for sparse, smooth, nonnegative coding or representation,” in *Proc. 14th Int. Conf. Neural Inf. Process.*, 2008, vol. 4984, pp. 811–820.
- [27] V. P. Paucu *et al.*, “Nonnegative matrix factorization for spectral data analysis,” *Linear Algebra Its Appl.*, vol. 416, no. 1, pp. 29–47, 2006.
- [28] H. Kim and H. Park, “Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis,” *Bioinformatics*, vol. 23, no. 12, pp. 1495–1502, 2007.
- [29] M. W. Berry *et al.*, “Algorithms and applications for approximate non-negative matrix factorization,” *Comput. Statist. Data Anal.*, vol. 52, no. 1, pp. 155–173, 2007.
- [30] H. Kim and H. Park, “Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method,” *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 2, pp. 713–730, 2008.
- [31] E. Hosseini-Asl and J. M. Zurada, “Nonnegative matrix factorization for document clustering: A survey,” in *Proc. 13th Int. Conf. Artif. Intell. Soft Comput.*, 2014, vol. 8468, pp. 726–737.
- [32] K. B. Petersen and M. S. Pedersen. (2012). *The Matrix Cookbook*. [Online]. Available: <http://matrixcookbook.com>
- [33] T. Pang-Ning *et al.*, *Introduction to Data Mining*, 1st ed. Reading, MA, USA: Addison-Wesley, May 2005.
- [34] N. Ahuja *et al.*, “Neighbor gray levels as features in pixel classification,” *Pattern Recognit.*, vol. 12, no. 4, pp. 251–260, 1980.
- [35] S. S. Bucak and B. Gunsel, “Incremental subspace learning via non-negative matrix factorization,” *Pattern Recognit.*, vol. 42, no. 5, pp. 788–797, 2009.
- [36] D. Cai *et al.*, “Graph regularized nonnegative matrix factorization for data representation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1548–1560, Aug. 2011.

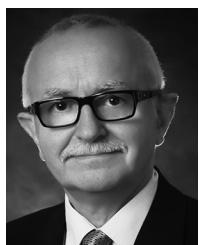
- [37] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learning Res.*, vol. 9, no. 11, pp. 2579–2605, 2008.
- [38] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [39] G. Gerig *et al.*, "Valmet: A new validation tool for assessing and improving 3D object segmentation," in *Proc. Med. Image Comput. Comput.-Assisted Intervention*, 2001, vol. 2208, pp. 516–523.
- [40] A. Soliman *et al.*, "Segmentation of lung region based on using parallel implementation of joint MGRF: Validation on 3-D realistic lung phantoms," in *Proc. 10th Int. Symp. Biomed. Imag.*, Apr. 2013, pp. 864–867.
- [41] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Process.*, vol. 2, no. 3, pp. 296–310, Jul. 1993.
- [42] B. Abdollahi *et al.*, "A novel Gaussian scale space-based joint MGRF framework for precise lung segmentation," in *Proc. 19th IEEE Intern. Conf. Image Process.*, Sep. 2012, pp. 2029–2032.
- [43] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 359–369, Mar. 1998.
- [44] M. S. Brown *et al.*, "Method for segmenting chest CT image data using an anatomical model: preliminary results," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 828–839, Dec. 1997.
- [45] J. Pu *et al.*, "Adaptive border marching algorithm: Automatic lung segmentation on chest CT images," *Comput. Med. Imag. Graph.*, vol. 32, no. 6, pp. 452–462, 2008.



**Ehsan Hosseini-Asl** (M'12) received the M.Sc. degree in electrical engineering (automation and instrumentation systems) from the Petroleum University of Technology, Tehran, Iran. Since 2014, he has been working toward the Ph.D. degree at the University of Louisville, Louisville, KY, USA.

He coauthored six conference papers published in the proceedings, two journal papers, and several papers in submission. His current research interests include unsupervised feature learning and deep learning techniques and their applications representation learning of 2-D and 3-D images, biomedical image segmentation, tissue classification, and disease detection.

Mr. Hosseini-Asl received the University Scholarship from the University of Louisville, in 2012, where he is currently a Research and Teaching Assistant. He also received the Graduate Student Research Grant from IEEE Computational Intelligence Society for his deep learning research in 2015.



**Jacek M. Zurada** (M'82–SM'83–F'96–LF'14) received the Ph.D. degree from the Gdansk Institute of Technology, Gdansk, Poland.

He is currently a Professor of electrical and computer engineering at the University of Louisville, Louisville, KY, USA. He authored or coauthored several books and more than 380 papers in computational intelligence, neural networks, machine learning, logic rule extraction, and bioinformatics, and delivered more than 100 presentations throughout the world. His work has been cited more than 8500 times.

Dr. Zurada served as an IEEE V-President, Technical Activities (TAB Chair) in 2014. In 2015, he chaired the IEEE TAB Strategic Planning Committee. He chaired the IEEE TAB Periodicals Committee, and TAB Periodicals Review and Advisory Committee, respectively, and was the Editor-in-Chief of the IEEE TRANSACTIONS ON NEURAL NETWORKS (1997–2003), an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, *Neural Networks* and of THE PROCEEDINGS OF THE IEEE. In 2004–2005, he was the President of the IEEE Computational Intelligence Society. He is an Associate Editor of *Neurocomputing* and several other journals. He has received numerous distinctions, including the 2013 Joe Desch Innovation Award and 2015 Distinguished Service Award, and four honorary professorships. He is a Member of the Polish Academy of Sciences and has been a Board Member of the IEEE, IEEE CIS, and IJCNN.



**Georgy Gimel'farb** received the Candidate of Engineering Sciences (equivalent to the Ph.D.) degree from the Academy of Sciences of the Ukrainian SSR, Kiev, Ukraine, in 1969, and the advanced Doctor of Engineering Sciences degree from the Higher Certifying Commission of the USSR, Moscow, USSR, in 1991.

After the long-term work at the V. M. Glushkov's Institute of Cybernetics of the Ukrainian Academy of Sciences, he joined The University of Auckland, Auckland, New Zealand, in 1997, where he is currently a Professor in computer science and a codirector of the Intelligent Vision System Laboratory. His research interests include probabilistic image modeling, statistical pattern recognition, texture analysis in medical and remotely sensed images, and computer stereo vision. He authored or coauthored more than 400 publications, including six books and 66 journal articles.

Dr. Gimel'farb received the prestigious State Premium of Ukraine in Science and Technology for 1997 for the fundamental and applied research results in image and signal recognition and development of intelligent information technologies and systems.



**Ayman El-Baz** (SM'15) received the Bachelor's and Master's degrees in electrical engineering in 1997 and 2001, and Ph.D. degree in electrical engineering all from the University of Louisville, Louisville, KY, USA, in 2006.

He is an Associate Professor and a University Scholar in the Department of Bioengineering, University of Louisville.

He has 13 years of hands-on experience in the fields of bioimaging modeling and noninvasive computer-assisted diagnosis systems. He has been the PI of several research projects funded by the Coulter Foundation and the American Cancer Society, and coPI on two R01 projects received by the NIH. Based on these awards, he developed a new CAD system for early diagnosis of lung cancer. The system has been licensed by PulmoCADx, Inc. (St. Louis, MO, USA). He has authored or coauthored more than 300 publications including 72 journal articles, 8 books, 30 book chapters, 135 refereed-conference papers, and 12 US patents.

Dr. El-Baz was named a Coulter Fellow for his contribution in the biomedical translational research in 2009. His work related to novel image analysis techniques for autism, dyslexia, and lung cancer has earned multiple awards, including the Wallace H. Coulter Foundation Early Career Translational Research Award in Biomedical Engineering Phase I & Phase II, a Research Scholar Grant from the American Cancer Society, and first place at the annual Research Louisville 2002, 2005, 2006, 2008, 2009, 2010, 2011, and 2012 meetings.