

# A novel marker-less lung tumor localization strategy on low-rank fluoroscopic images with similarity learning

Wei Huang · Jing Li · Peng Zhang · Min Wan ·  
Can Fang · Minmin Shen

Received: 26 January 2014 / Revised: 25 June 2014 / Accepted: 9 July 2014 /  
Published online: 24 July 2014  
© Springer Science+Business Media New York 2014

**Abstract** Fluoroscopic images depicting the movement of lung tumor lesions along with patients' respirations are essential in contemporary image-guided lung cancer radiotherapy, as the accurate delivery of radiation dose on lung tumor lesions can be facilitated with the help of fluoroscopic images. However, the quality of fluoroscopic images is often not high, and several factors including image noise, artifact, ribs occlusion often prevent the tumor lesion from being accurately localized. In this study, a novel marker-less lung tumor localization strategy is proposed. Unlike conventional lung tumor localization strategies, it doesn't require placing external surrogates on patients or implanting internal fiducial markers in patients. Thus ambiguous movement correlations between moving tumor lesions and surrogates as well as the risk of patients pneumothorax can be totally avoided. In this new strategy, fluoroscopic images are first decomposed into low-rank and sparse components

---

W. Huang (✉) · J. Li · M. Wan  
School of Information Engineering, Nanchang University, Nanchang, China  
e-mail: huangwei@ncu.edu.cn

J. Li  
e-mail: jing.li.2003@gmail.com

M. Wan  
e-mail: wanmin1983@gmail.com

P. Zhang (✉)  
School of Computer Science, Northwestern Polytechnical University, Xi'an, China  
e-mail: zh0036ng@nwpu.edu.cn

C. Fang  
School of Computer and Information Science, Southwestern University, Chongqing, China  
e-mail: canfang@swu.edu.cn

M. Shen  
School of Software Engineering, South China University of Technology, Guangzhou, China  
e-mail: minmin.shen@uni-konstanz.de

M. Shen  
INCIDE Center, University of Konstanz, Konstanz, Germany

via the split Bregman method, and then spectral clustering techniques are incorporated for similarity learning to realize the tumor localization task. Clinical data obtained from 60 patients with lung tumor lesions is utilized for experimental evaluation, and promising results obtained by the new strategy are demonstrated from the statistical point of view.

**Keywords** Tumor localization · Low-rank and sparse decomposition · Similarity learning · Spectral clustering

## 1 Introduction

Lung cancer, the most common cause of cancer-related death, is responsible for over 1.38 million deaths annually in worldwide population [8]. It is also widely acknowledged that, accurate diagnosis of lung cancer at its early stage and timely treatments are essential to enhance the survival time of lung cancer patients, or even cure the disease [21]. There are diverse effective treatment manners for lung cancer to date, including surgery, radiotherapy, chemotherapy, palliative care, etc [9, 21]. Among them, radiotherapy, whose main purpose is to eliminate malignant cells via ionizing radiation, is often indispensable in contemporary lung cancer treatments [21].

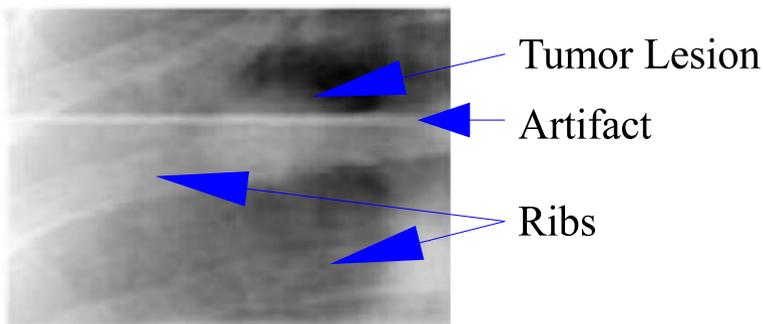
For lung cancer radiotherapy, accurate predictions on positions of tumor lesions along with patients' respiratory cycles (i.e., the procedure also known as lung tumor localization or lung tumor tracking) is highly demanded, as high-dose-rate radiation beam needs to be concentrated on moving tumor lesions, and meanwhile radiation exposure towards normal tissues surrounding tumor lesions should be kept as low as possible [21]. In order to realize the above task, conventional lung tumor localization strategies in radiotherapy often rely on markers, which include either external surrogates placed on the abdomen of patients [13] or internal fiducial markers implanted within patients via surgeries [27]. However, their disadvantages also obvious: for external surrogates, it is often ambiguous to correlate the complicate movement between external surrogates and moving tumor lesions inside patients, making the tumor position derivation lack of accuracy; for internal surrogates, patients with percutaneous marker implantations are likely to suffer from the risk of pneumothorax [18]. Therefore, marker-less localization strategies, which are conducted without the aid of external or internal markers, become more favored in lung tumor radiotherapy nowadays.

There are already several studies proposed to realize marker-less tumor localization in lung tumor radiotherapy in recent years [14, 16, 17, 19, 23, 26, 29, 30]. For instance, potential regions in lung image sequences containing discriminative tumor feature are shortlisted via a principal component analysis (PCA) model in [16]. In [19], nonlinear manifold learning methods, including locally linear embedding [25], local tangent space alignment [32], and Laplacian eigenmap [2], are incorporated similarly to [16] but replacing the role of PCA for tumor position derivation. In [17, 23, 26, 29, 30], techniques of artificial neural network (ANN), support vector machine (SVM), and linear/non-linear regression are incorporated to realize the lung tumor localization task. It can be concluded from existing studies that, popular pattern recognition tools are widely utilized in contemporary marker-less lung tumor localization studies.

For most marker-less lung tumor localization studies, medical images are often adopted to reveal movements of lung tumor lesions for radiation beams delivery [14, 16, 17, 19, 23, 26, 29, 30]. Among a number of medical imaging facilities, fluoroscopy, which is an effective and affordable medical imaging manner to obtain fluoroscopic image sequences

inspecting internal structures of patients, is widely incorporated [20]. However, it is worth to note that the image quality of contemporary fluoroscopy is not very high. For one thing, fluoroscopy is often equipped with x-ray image intensifiers, which make fluoroscopy suffers from several spatial blurring factors (e.g., Lubberts effect, K-fluorescence reabsorption, electron range, etc), which nearly all x-ray imaging devices have. For another, fluoroscopic systems also experience temporal blurring due to system lag, which may create motion blurring for moving objects. Thus, processing fluoroscopic images is more challenging than processing images obtained by other contemporary imaging facilitates of higher spatial resolution and better imaging quality. A case in point is illustrated in Fig. 1. It can be observed that, the spatial resolution of the example fluoroscopic image is not high, and the tumor lesion is also intersected with unfavorable artifact as well as obscured by ribs. Meanwhile, non-tumor tissues surrounding tumor lesions often move together with tumor lesions in patients respiratory cycles, making it more challenging to differentiate the target tumor lesion and localize its position precisely. The reason to widely adopt fluoroscopic systems in ordinary clinical lung tumor diagnosis nowadays is because, the x-ray exposure needed to be put on patients when producing fluoroscopic images is low, compared with other contemporary radiography facilitates such as CT. Therefore, it complies well with patients' safety issues.

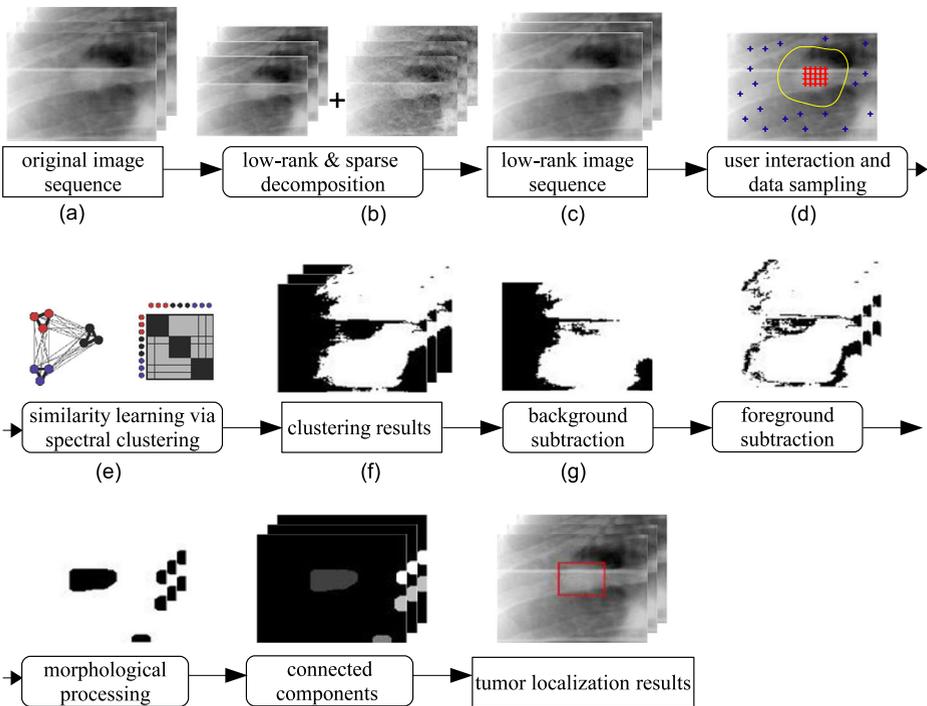
For most contemporary marker-less lung tumor localization studies based on fluoroscopic images, they often apply diverse image processing or pattern recognition methods directly on original fluoroscopic images to realize the tumor localization task [16, 17, 19, 23, 26, 29, 30]. However, the poor quality of original fluoroscopic images often prevent those methods from achieving satisfactory localization performance. Therefore, it inspires us to find more “clean” fluoroscopic images, in which the bad influence brought by image noise and other degraded factors could be reduced as low as possible. Another problem in contemporary marker-less tumor localization studies is that, in order to discern the target tumor lesion from other non-tumor tissues, image pixels are either classified into different tissue categories or clustered into different groups, according to certain criteria in their utilized classification or clustering methods [16, 17, 19, 23, 26, 29, 30]. Finding a proper criterion measuring the similarity between pixels to differentiate the target tumor lesion from other non-tumor tissues is often of great importance in those methods. However, for many of these methods, such an important criterion is often determined empirically. Since medical imaging data often have large varying statistical properties across different patients, either making an assumption about the adopted similarity beforehand, or learning it based on training images obtained from some patients and adopting learned results on images of



**Fig. 1** Example of a fluoroscopic image

other patients afterwards (as the conventional pattern recognition way) may not fit the nature of the tumor localization problem on diverse patients well. Therefore, a method which can learn unique similarity for each particular patient will fit the problem better. It is necessary to conduct similarity learning case by case in this marker-less tumor localization study, and such a problem has not been taken into consideration in previous related researches [14, 16, 17, 19, 23, 26, 29, 30].

Therefore, with the above requirements of both finding more “clean” fluoroscopic images and conducting case-by-case similarity learning in lung tumor localization into consideration, a novel marker-less tumor localization strategy on fluoroscopic images for image-guided lung cancer radiotherapy is presented in this paper. The flowchart of the introduced localization strategy is illustrated in Fig. 2. It can be observed that, a low-rank & sparse decomposition step differentiating an original fluoroscopic image into a sparse fluoroscopic image with major information of non-tumor tissues and a low-rank fluoroscopic image with major information of the target tumor tissue is conducted firstly in the newly introduced localization strategy. Low-rank images obtained through such a decomposition are more “clean” compared with original fluoroscopic images, thus they are more suitable for tumor localization. Afterwards, a spectral clustering method is utilized to fulfill pixel-wise similarity learning in low-rank fluoroscopic images. A series of pattern recognition and image processing techniques, including K-means clustering, morphological processing, connected component analysis, are incorporated to accomplish the tumor localization task therein. It is necessary to point out that, the newly introduced strategy is not fully automatic



**Fig. 2** Flowchart of the introduced marker-less lung tumor localization strategy with low-rank & sparse decomposition and similarity learning via spectral clustering

since a pre-requisite user interaction step is incorporated into the essential similarity learning step via spectral clustering.

The organization of the paper is as follows. In Section 2, a low-rank & sparse decomposition method based on robust-PCA is introduced to decompose original fluoroscopic images into low-rank and sparse components via the split Bregman method. Section 3 elaborates steps to conduct similarity learning via spectral clustering based on decomposed low-rank fluoroscopic images. Section 4 presents a series of pattern recognition and image processing techniques, which are incorporated to accomplish the lung tumor localization task based on previous decomposition (Section 2) and learning (Sections 3) results. In Section 5, clinical data obtained from 60 patients with lung tumor lesions are utilized to evaluate the performance of the newly introduced marker-less lung tumor localization strategy. The superiority of adopting low-rank & sparse decomposition as well as case-by-case similarity learning via spectral clustering in the newly introduced lung tumor localization strategy are demonstrated through various experiments and comprehensive statistical analysis, compared with other conventional tumor localization strategies. In Section 6, the conclusion of this study is drawn.

## 2 Low-rank & sparse decomposition on original fluoroscopic images

As introduced in Section 1, it is challenging to localize tumor lesion precisely on original fluoroscopic images, as their image quality is often poor and non-tumor tissues surrounding the tumor lesion also move simultaneously, making the tumor and non-tumor differentiation even harder. A general intuition to tackle the problem is that, the tumor localization task should become more convenient to handle, if an original fluoroscopic image can be decomposed into an image component with major tumor movement over an ideal stationary background, as well as another image component with major non-tumor tissues movement. Conducting tumor localization on the first component of images should be easier than doing so directly on original images.

The above intuition is similar towards matrix decomposition in mathematics, in which a large data matrix  $M$  can be decomposed as  $M = L + S$ , where the matrix  $L$  has low rank and the matrix  $S$  is sparse. In order to solve  $L$  and  $S$ , the classical principal component analysis (PCA) can be adopted, provided the assumption that the matrix  $S$  is small and independent and identically distributed Gaussian. However, in many real-life applications, entries in  $S$  can have arbitrarily large magnitude, and the general assumption in classical PCA may not be always true.

Inspired by the recent prominent progress of robust-PCA [7], whose main idea is to recover a low-rank measurement and a sparse measurement from a corrupted original measurement provided arbitrary errors in magnitude are supported in the accompanying sparse measurement, such a low-rank & sparse decomposition on original fluoroscopic images can be explicitly formulated via the following optimization:

$$\arg \min_{(\chi_1, \chi_2)} \|(\chi_1 + \chi_2) - \chi\|_2^2 + \lambda_* \|\chi_1\|_* + \lambda \|\chi_2\|_1 \quad (1)$$

where,  $\chi$ ,  $\chi_1$  and  $\chi_2$  depict an original fluoroscopic image, its decomposed low-rank component, and its decomposed sparse component, respectively;  $\|\cdot\|_1$ ,  $\|\cdot\|_2$  and  $\|\cdot\|_*$  denote the  $\ell_1$ -,  $\ell_2$ -, and nuclear norm, respectively;  $\|\chi_1\|_*$  penalizes the rank of  $\chi_1$  defined as the sum of its singular values with regularizing coefficient  $\lambda_*$ ;  $\|\chi_2\|_1$  is for promoting the sparsity of  $\chi_2$  with regularizing coefficient  $\lambda$ . In this way, major information of the moving tumor

lesion is decomposed and contained in the low-rank component  $\chi_1$ , while other information including noise, outliers, as well as movements of non-tumor tissues are mainly sparse and they are included in the sparse component  $\chi_2$ . It is worth to notice that, such a decomposition via robust-PCA begins to attract increasing attentions in general tracking studies in computer vision in recent years [15, 34, 35], but its utilization in medical applications is still rare.

In this study, a split Bregman method [11] is incorporated to solve  $\chi_1$  and  $\chi_2$  in the above optimization problem. The split Bregman method is often implemented via iterations. To be specific, (1) can be splitted into the following three sub-problems at iteration time  $k$ :

$$\text{Sub-problem 1: } \chi_1^k = \min_{\chi_1} \|(\chi_1 + \chi_2^{k-1}) - \chi^{k-1}\|_2^2 + \lambda_* \|\chi_1\|_* \quad (2)$$

$$\text{Sub-problem 2: } \chi_2^k = \min_{\chi_2} \|(\chi_1^k + \chi_2) - \chi^{k-1}\|_2^2 + \lambda \|\chi_2\|_1 \quad (3)$$

$$\text{Sub-problem 3: } \chi^k = \chi^{k-1} - (\chi_1^k + \chi_2^k) \quad (4)$$

For initialization ( $k=0$ ),  $\chi^0$  is equivalent to the original fluoroscopic image  $\chi$ . At each iteration time  $k$ , sub-problem 1 can be solved via a singular value thresholding (SVT) algorithm at a low computational cost [6]. According to [11], the optimal solution of sub-problem 2 can be rapidly obtained using a shrinkage operator  $\chi_2^k = \text{shrink}(\chi^{k-1} - \chi_1^k, \frac{\lambda}{2})$ . After solving sub-problems 1 and 2,  $\chi$  is updated in sub-problem 3 for the next iteration time ( $k + 1$ ). The whole iteration terminates when the following criterion meets:  $\max(\|\chi_1^{k+1} - \chi_1^k\|_\infty, \|\chi_2^{k+1} - \chi_2^k\|_\infty) \leq \text{tol}$ , in which  $\text{tol}$  denotes an enough small change between two decomposed images within two consecutive iteration times and  $\|\cdot\|_\infty$  represents the spectral norm.

After all these subproblems are solved, a low-rank fluoroscopic image  $\chi_1$  mainly containing information of the moving tumor lesion is produced, and it will be utilized in the following similarity learning step to fulfill the tumor localization task, instead of the original fluoroscopic image  $\chi$ . It is worth to notice that, the low-rank and sparse decomposition through robust-PCA has already received much popularity in general object tracking and motion segmentation studies in computer vision in recent years [15, 33, 35]. However, it has not been incorporated in medical imaging applications before. Thus, it is novel to do the low-rank and sparse decomposition on fluoroscopic images for tumor localization in this study.

### 3 Similarity learning via spectral clustering on low-rank fluoroscopic images

After obtaining low-rank fluoroscopic images, it is necessary to differentiate tumor pixels from non-tumor pixels on low-rank fluoroscopic images for tumor localization. Generally speaking, in conventional studies where classification or clustering models are incorporated [16, 17, 19, 23, 26, 29, 30], it is often highly demanded to determine a “good” similarity measure, which is able to assign high similarity to pixels from the same group (i.e., either tumor or non-tumor group) and low similarity to pixels from different groups. Since medical imaging data are often with largely varying statistical properties across different patients, either making an assumption about the adopted similarity beforehand, or incorporating some patient data for training and then applying learned results on other patients for testing (i.e., the conventional pattern recognition way) may not fit the nature of the problem well. Therefore, it is essential to conduct similarity learning case by case in tumor localization.

In this section, pixel-wise similarity learning via spectral clustering is introduced. It is also the first attempt to adopt spectral clustering techniques on fluoroscopic images for marker-less tumor localization.

### 3.1 Data sampling for similarity learning

In order to obtain training data for similarity learning, a pre-requisite user interaction step needs to be incorporated. In this study, clinicians are allowed to draw their own region-of-interest (ROI), in which they assume that the tumor lesion is to be enclosed, on the first low-rank fluoroscopic image of the obtained low-rank fluoroscopic image sequence of every patient. Such a ROI can be in any arbitrary shape, and the enclosure does not have to be very close to the tumor lesion's boundary on the first image. The ROI drawing is conducted only once for each patient (illustrated as yellow in Fig. 2a). It is necessary to point out that, the concept of user interaction here is different from the previously mentioned “marker-less”, which is described in the medical domain.

Using such a ROI, training data from one patient for her/his own pixel-wise similarity learning can be extracted conveniently: points inside the enclosed ROI can be sampled as positive training samples (i.e., from the tumor tissue, illustrated as red in Fig. 2a), while points outside the enclosed ROI are to be sampled as negative training samples (i.e., from non-tumor tissues, illustrated as blue in Fig. 2a). For implementation, positive samples often sharing similar visual attributes on low-rank fluoroscopic images are chosen from the center area of the ROI.

For negative samples, they are usually highly variable in visual attributes, and are often more difficult to be sampled. The conventional random sampling strategy is likely to under-represent negative sub-populations (a.k.a. stratum [24], which is the collection of points sharing similar visual characteristics), making chosen negative samples less representative. Also, less representative data will tend to bias learning outcomes, deteriorating the whole performance therein in conventional pattern recognition studies. Therefore, in order to make negative samples more representative, a stratified random sampling strategy [24] is incorporated to sample negative points in this study. The optimal stratum allocation scheme used here is the Neyman allocation [24], in which the number of sampled negative points  $n_l$  from a stratum  $l$  is determined by:

$$n_l = n \cdot \frac{R_l \cdot \sigma_l}{\sum_{k=1}^L R_k \cdot \sigma_k} \quad R_l = \frac{N_l}{N_1 + \dots + N_L} \quad (5)$$

where,  $n$  is the total number of negative points to be sampled;  $\sigma_l$  is the standard deviation of negative points located at stratum  $l$ ;  $R_l$  denotes the fraction provided by the number of negative points in the stratum  $l$  ( $N_l$ ) over the number of all negative points from all  $L$  strata. It can be observed from (5) that, Neyman allocation intends to allow more negative samples to be taken from one stratum, if it contains a larger fraction of negative points with largely varying visual characteristics. In this way, more representative negative training data can be sampled for the next similarity learning.

### 3.2 Supervised similarity learning via spectral clustering

After obtaining samples for similarity learning, an explicit form of the pixel-wise similarity utilized in this study is to be defined. A spatially weighted metric-based similarity  $d(x_i, x_j)$

reflecting the similarity between pixels  $x_i$  and  $x_j$  in a low-rank fluoroscopic image is as follows:

$$d(x_i, x_j) = P(x_i, x_j) \cdot Q(x_i, x_j) = \exp\left(-\frac{\|p_i - p_j\|^2}{\sigma_p^2}\right) \cdot \exp\left(-(s_i - s_j)^T A (s_i - s_j)\right) \quad (6)$$

where,  $\sigma_p$  is a scalar and  $A$  is a full matrix;  $p_i$  ( $p_j$ ) is the normalized spatial coordinates of pixel  $x_i$  ( $x_j$ ), and  $s_i$  ( $s_j$ ) is the extracted low-level visual feature of pixel  $x_i$  ( $x_j$ ). Obviously,  $P(x_i, x_j)$  reflects the spatial similarity between  $x_i$  and  $x_j$ , which is constructed as an isotropically-scaled Gaussian.  $Q(x_i, x_j)$  reveals the similarity of low-level visual features between  $x_i$  and  $x_j$  using a Mahalanobis metric. The similarity measure in (6) emphasizes on spatial localization, by which the similarity between pairwise pixels  $x_i$  and  $x_j$  decreases with the increase of their in-between distance. Therefore, two spatially nearby pixels will have more dominant influence on the measured similarity than two pixels that are far apart. It is worth to notice that, the spatial information can be constructed with visual features and only one term in (6) (e.g., term  $Q$ ) is necessary for similarity calculation. However, since the unknown  $A$  to be learned is a full matrix, more elements need to be determined (as the size of the matrix  $A$  increases) and the efficiency of the whole learning will be influenced. In our utilized similarity (6), the unknown associated with spatial information  $\sigma$  is a scalar and it is more efficient to be determined within the proposed tumor localization strategy.

The newly adopted similarity in (6) is also different from those utilized in other related studies [22, 28]. In the Normalized Cut algorithm [28], Shi et al. did not include an explicit form of the term  $P$  of (6) in their similarity and assumed the matrix  $A$  in the term  $Q$  to be a diagonal matrix, which ignores the correlation among different dimensions of extracted features. In the Ng-Jordan-Weiss algorithm [22], only term  $Q$  of (6) is utilized in their similarity, and  $s_i$  in their work only represents extracted low-level feature of pixel  $x_i$ , which ignores spatial information. In order to ensure distances used in (6) as metrics, they should meet four axioms of a metric (i.e., non-negativity, identity of indiscernibles, symmetry and triangle inequality), and the matrix  $A$  in term  $Q$  should be at least positive semi-definite (i.e.,  $A \geq 0$ ). The purpose of learning a similarity is to find proper parameters of it, so that data from the same or different categories can be well grouped and differentiated therein. In this study, the purpose of learning such a similarity involves finding proper values for parameters  $A$  and  $\sigma_p^2$ . Hence, unknown parameters of (6) are determined algorithmically, not empirically.

In order to fulfill the learning task, spectral clustering is incorporated. Spectral clustering is a special clustering technique, which can be considered as an approximate solution to the classic graph-cut problem achieving partitioning of an image by cutting weak links between graph nodes to separate an image into various portions [5]. The pixel-wise similarity learning task here is realized via a spectral clustering algorithm, whose main steps based on [1] are elaborated in Table 1. In [1], similarity learning via spectral clustering aims to learn the Gaussian width  $\sigma^2$  in its simple Gaussian RBF (a.k.a. radial basis function) similarity measure. Since the adopted similarity in (6) is more sophisticated, simply applying the existing learning algorithm directly on (6) is not reasonable. Therefore, an optimization function with the Frobenius norm in (7) of Table 1 is utilized for learning unknowns in (6). Through a gradient descent method with constraints in (7), unknown parameters in (6) can be solved. Hence, a learned spatially weighted metric-based similarity can be determined via spectral clustering techniques.

### 3.3 Unsupervised out-of-sample extension

The next step is to perform tumor and non-tumor differentiation on pixels other than those sampled for similarity learning (in Section 3.1) on the low-rank fluoroscopic images by spectral clustering. It is also acknowledged that, the main computational burden of spectral clustering resides in the eigen-decomposition step of the graph Laplacian matrix  $L$  of size  $n \times n$ , where  $n$  is the total number of pixels. In the previous learning step, this is not a problem since the number of sampled points to form the training set  $S$  in Table 1 is still small. If spectral clustering is used to group all other points in the ROI, the size of matrix  $L$  will become extremely large and causes memory problems (e.g., for a  $200 \times 200$  pixel-wise ROI, the constructed graph Laplacian  $L$  will be of the size  $40,000 \times 40,000$ , which is often difficult to load and handle for ordinary computers) as well as computational burden in eigen-decomposition (i.e., its computational cost is around  $O(n^3)$ ). Therefore, an unsupervised out-of-sample extension method [3] is incorporated. The purpose of applying out-of-sample extension here is to map points into the spectral domain directly using a mapping function without performing the ordinary eigen-decomposition step. Therefore, the main computational burden of spectral clustering can be avoided.

For the rest pixels in ROI, the similarity is calculated with respect to each sample in the training set  $S$  using (6) with the learned similarity function. In this study, since the prior knowledge about the location of tumor lesion on the first fluoroscopic image is incorporated by drawing a ROI, points outside the ROI should not be considered as candidate points for tumor lesion. Thus, a weighted adjacency matrix  $W'$  of the size: (number of remaining

**Table 1** An algorithm of similarity learning via spectral clustering

<b>Input:</b>	a set of $n$ pixels $S = \{s_1, s_2, \dots, s_n\}$ in $\mathbf{R}^d$
<b>Algorithm:</b>	<ol style="list-style-type: none"> <li>1. Initialize unknown parameters <math>A</math> and <math>\sigma_p^2</math> of the adopted similarity in (6).</li> <li>2. Calculate the pixel-wise similarity <math>d(x_i, x_j)</math> in (6) for each pixel pair. Use those computed <math>d(x_i, x_j)</math> as elements to construct an affinity matrix <math>\mathbf{D} \in \mathbf{R}^{n \times n}</math> for all <math>n</math> pixels in <math>S</math>.</li> <li>3. Form a new diagonal matrix <math>\mathbf{C}</math>, whose <math>(i, i)</math> element is the sum of the <math>i</math>-th row of <math>\mathbf{D}</math>.</li> <li>4. Construct a graph Laplacian matrix <math>\mathbf{L} = \mathbf{C}^{-1/2} \cdot \mathbf{D} \cdot \mathbf{C}^{-1/2}</math>, and find the <math>k</math> largest eigenvalues and their corresponding eigenvectors.</li> <li>5. Form a new matrix <math>\mathbf{X} \in \mathbf{R}^{n \times k}</math> by stacking the <math>k</math> extracted eigenvectors in columns.</li> <li>6. Solving <math>A</math> and <math>\sigma_p^2</math>: form an optimization problem in the Frobenius norm: <div style="text-align: right; margin-top: 10px;"> <math display="block">\min J(\mathbf{A}, \sigma_p^2) = \frac{1}{2} \left\  \mathbf{X}\mathbf{X}^T - \mathbf{X}_{part}\mathbf{X}_{part}^T \right\ _F^2 \text{ s.t. } \mathbf{A} \succeq 0 \tag{7}</math> </div> </li> </ol> <p style="margin-left: 20px;">where, <math>\mathbf{X}_{part} = \mathbf{C}^{-1/2}\mathbf{E}(\mathbf{E}^T\mathbf{C}\mathbf{E})^{-1/2}\mathbf{B}</math>; <math>\mathbf{E}</math> is an indicator matrix of data set partitions and <math>\mathbf{B}</math> is an arbitrary orthogonal matrix. Applying <math>\mathbf{A} \succeq 0</math> as constraints to the optimization problem and find optimal solutions of <math>\mathbf{A}</math> and <math>\sigma_p^2</math> through gradient descent.</p>
<b>Output:</b>	Learned parameters $\mathbf{A}$ and $\sigma_p^2$ in (6).

samples in ROI)  $\times$  (number of samples in  $S$ ) can be obtained. After normalization, the affinity matrix  $W'_{norm}$  is given by

$$W'_{norm(i,j)} = \frac{W'_{ij}}{\sqrt{\sum_{x \in S} W'_{ix} \sum_{x' \in S} W'_{jx'}}} \quad (8)$$

The spectral embedding of pixels in ROI can be computed using the following mapping function directly:

$$X_{mapped} = W'_{norm} \cdot X \cdot V^{-1} \quad (9)$$

where,  $X$  and  $V$  are matrices of eigenvectors and eigenvalues obtained from the training data set  $S$ , respectively. After obtaining  $X_{mapped}$ , conventional clustering algorithms can be utilized to differentiate the tumor lesion unsupervisedly. In this work, the classic K-means algorithm [4] is incorporated in this step, and the pre-defined number of groups equals to 2 ( $K = 2$ ) for tumor and non-tumor groups.

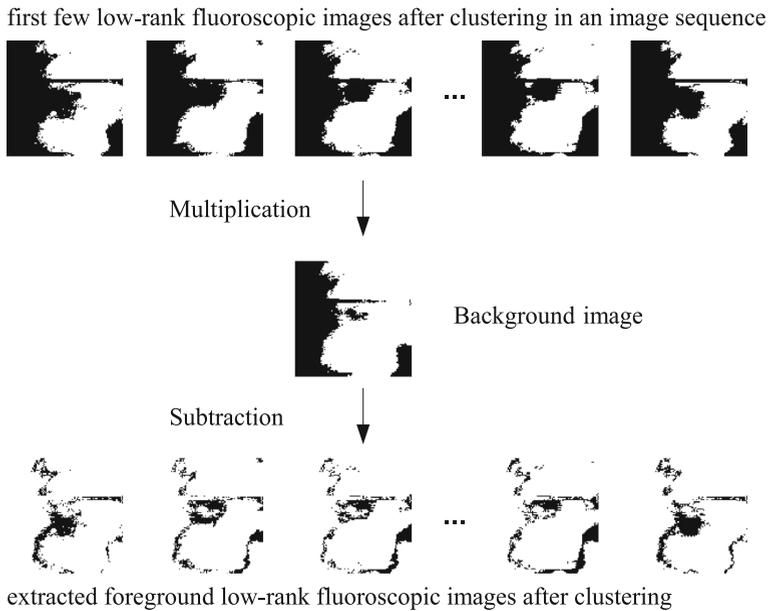
For other low-rank fluoroscopic images besides the first one in an image sequence of one patient, the computation is similar. For each of them, the affinity matrix in (8) is calculated with respect to all pixels on that image using the learned similarity measure in (6). After that, the spectral embedding of pixels is realized using (9), and K-means is also implemented to differentiate tumor and non-tumor on that low-rank fluoroscopic image. Examples of those clustering results are illustrated in Fig. 2b, in which a binary image obtained via K-means clustering on a low-rank fluoroscopic image is demonstrated. Pixels with value of 1 (i.e., black) represent the tumor group, while pixels with value of 0 (i.e., white) belong to the non-tumor group. In this way, the tumor localization result on each fluoroscopic image can be roughly obtained, and the next step is to refine rough localization results.

#### 4 Tumor localization on low-rank fluoroscopic images

Although major moving non-tumor tissues surrounding the tumor lesion can be removed from low-rank fluoroscopic image  $\chi_1$  via the low-rank & sparse decomposition step in (1), some non-tumor tissues still exist after the above clustering step (i.e., for instance, black areas except for the target tumor lesion in Fig. 2b), as their movement is not much and they are generally regarded as stationary background in the previous decomposition step. Therefore, background subtraction and foreground extraction steps are necessary to refine tumor localization results. In this study, since the movement of non-tumor tissues in low-rank fluoroscopic images after clustering is small, a background image is obtained via the multiplication of clustering results of the first few low-rank fluoroscopic images of each patient. The foreground image excluding non-tumor tissues with less movement is extracted via the subtraction between obtained clustering results of low-rank fluoroscopic images and their background image. An illustration of the above process is in Fig. 3.

For extracted foreground images, they are likely to be deteriorated by thin line structures and holes on tumor lesions (as shown in Fig. 3), which are caused by image artifact and ribs occlusion already existing in original fluoroscopic images. Hence, a morphological processing step [12] is applied on extracted foreground images to remove unnecessary fractions afterwards. Morphological processing steps in this study are composed of an opening operation and a closing operation:

$$\text{opening: } f \circ s = (f \ominus s) \oplus s \quad \text{closing: } f \bullet s = (f \oplus s) \ominus s \quad (10)$$



**Fig. 3** An illustration of background subtraction and foreground extraction

where,  $f$  indicates extracted foreground images;  $s$  is a disk-shaped structuring element of radius 5 in this study;  $\oplus$  and  $\ominus$  represent dilation and erosion operations, respectively [12]. Generally speaking, too big structuring element will blur the target tumor lesion, and too small structuring element will not help to remove unnecessary isolated points, holes, and thin line structures in tumor localization results. Thus, the shape and size of the structuring element utilized in this study are determined based on trial-and-errors. An example after applying the morphological processing step is illustrated in Fig. 2e.

After morphological processing, several tumor lesion candidate regions are available (as shown in Fig. 2e). An automatic shortlisting strategy for the target tumor lesion is incorporated here via connected component analysis (CCA) [12], which is capable to find uniquely labeled connected components on binary images. As illustrated in Fig. 2f, different tumor candidate regions are labeled differently on the example image after morphological processing. The target tumor lesion in the current frame is selected as the one with the highest spatial correlation towards the determined tumor lesion in the previous frame, given the fact that lung tumor does not move rapidly frame by frame in the whole fluoroscopic image sequence along with patients' respirations. After all above steps are performed, the tumor position in one fluoroscopic image can be localized using a minimum rectangle enclosing the detected target tumor lesion as illustrated in Fig. 2g.

## 5 Experiments and discussion

### 5.1 Data description and methods implementation

The performance of the newly introduced marker-less tumor localization strategy has been evaluated by fluoroscopic images obtained from 60 real patients with lung tumor cancer. All

images were obtained in the affiliated hospital of Nanchang University, and informed consent to access those data for academic purpose was obtained from all patients. The average duration of each image sequence of each patient is around 2 mins, in which 24 to 40 respiratory cycles exist, and the spatial resolution of an original fluoroscopic image is  $1024 \times 768$ . For extracted feature from obtained low-rank fluoroscopic images, normalized horizontal and vertical coordinates of each pixel  $i$  are incorporated as  $p_i$  in (6). Normalized intensity, normalized entropy value, normalized local range and normalized local standard deviation in a neighborhood of size  $3 \times 3$  of each pixel  $i$  are adopted to construct  $s_i$  in (6). For unknown parameters  $A$  and  $\sigma_p$  in (6), they are set as an identity matrix and 0.1, respectively as initializations in Table 1. For sampled training data, totally 50 pixels per patient are sampled to construct her/his training data set (i.e., input  $S$  in Table 1), in consideration of both the efficiency of implementation in clinical practice as well as the effectiveness of the introduced tumor localization strategy. Among 50 sampled pixels, 25 pixels are positive training samples taken within ROI and the rest 25 pixels are negative training samples taken outside of ROI. For negative training samples, 5 stratum of candidates are utilized in the applied stratified random sampling strategy via Neyman allocation (i.e.,  $L = 5$  in (5)).

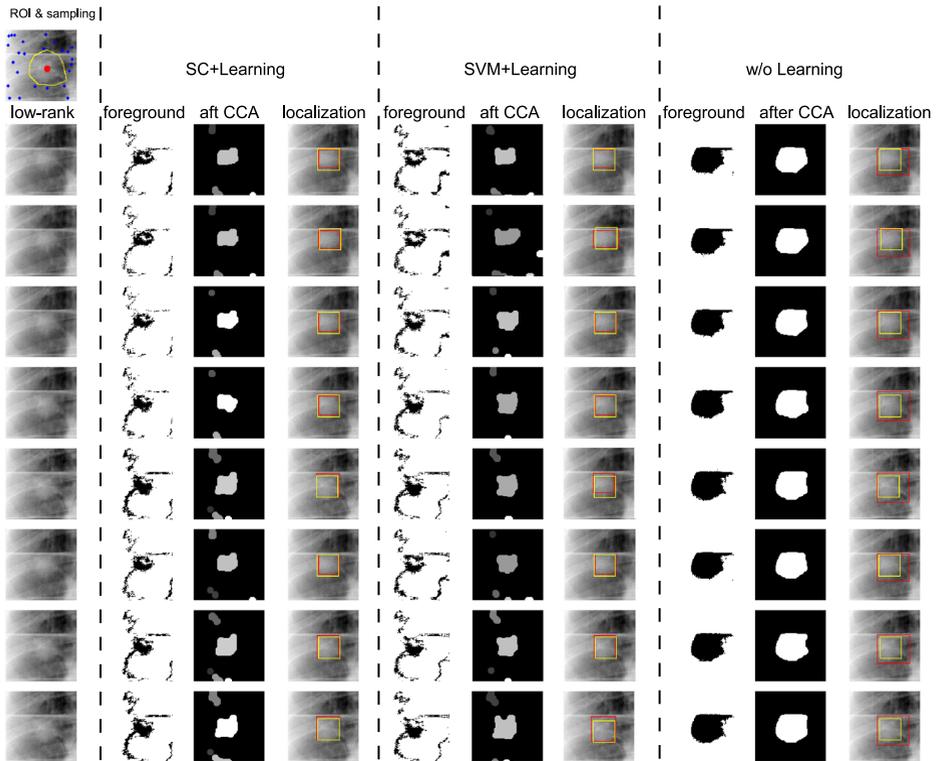
In order to demonstrate merits of incorporating low-rank & sparse decomposition as well as similarity learning via spectral clustering in the introduced lung tumor localization strategy, experiments are categorized into two folds: one is to demonstrate the advantage of similarity learning (Section 5.2) and the other is for the merit of low-rank & sparse decomposition (Section 5.3). In each fold, dozens of comparison experiments as well as comprehensive statistical analysis are employed. An additional discussion is also conducted in Section 5.4.

## 5.2 Experiments and analysis on similarity learning

Besides the introduced strategy (denoted as “SC+Learning”), two other tumor localization strategies are implemented for comparison. One is the same as the introduced strategy but to replace spectral clustering with support vector machine for similarity learning (denoted as “SVM+Learning”), the other is to apply all other steps as the introduced strategy but without the similarity learning step (denoted as “w/o Learning”). The purpose here is to demonstrate the superiority of incorporating similarity learning via spectral clustering in “SC+Learning”, compared with the classic learning paradigm (“SVM+Learning”) and non-learning (“w/o Learning”).

For the above three methods, all patient data are utilized and their tumor localization results are evaluated. For SVM+Learning, a Gaussian radius basis function (Gaussian-RBF) is incorporated, and the Gaussian width in such a Gaussian-RBF is the parameter to learn. A radius-margin bound method [31] is implemented for learning the Gaussian width. In radius-margin bound method, an upper bound specifying the number of classification errors in a leave-one-out procedure is defined using the margin of the SVM classifier as well as the radius of a sphere, which includes all transformed feature vectors in the high-dimensional feature space (a.k.a “kernel trick”) [31]. Hence, parameters in the compared “SVM+Learning” strategy are also determined algorithmically. In this study, pre-defined parameters of “SVM+Learning” are determined via trial-and-errors for optimal tumor localization performance as suggested, including the trade-off between training error and margin set as 0.01, the initial Gaussian width set as 0.1, and the cost-factor by which training errors on positive samples out-weight errors on negative examples set as 1.

Figure 4 demonstrates tumor localization results from an obtained low-rank fluoroscopic image sequence of one patient by incorporating our introduced strategy as well as the two



**Fig. 4** An illustration of tumor localization results provided by three compared strategies on example low-rank fluoroscopic images

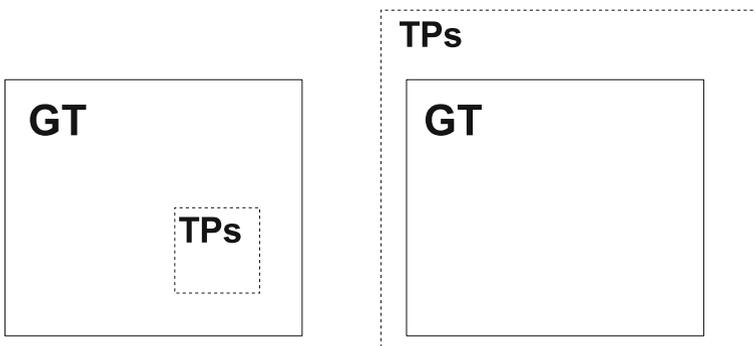
compared strategies. ROI delineation (in yellow) and sampling data (in red and blue) on the first low-rank fluoroscopic image of that sequence are depicted on the top left corner of Fig. 4. The same ROI and sampling data are utilized in all three strategies. For the rest 8 images below the ROI image at the same column, they depict different low-rank fluoroscopic images from that image sequence. In each row of the 8 images, the foreground, the intermediate tumor localization result after CCA, and the final tumor localization result highlighted by bounding boxes are demonstrated for “SC+Learning”, “SVM+Learning”, and “w/o Learning”, respectively.

In each final tumor localization result, a red rectangle representing the smallest rectangle containing the detected tumor region after CCA is compared with a yellow rectangle, which was determined as the tumor ground truth by our senior clinicians by consensus. Localization results with the highest matching towards the ground truth will be considered to be the best, as the highest concentration of radiation beam to the tumor as well as the lowest radiation exposure towards surrounding non-tumor tissues could be obtained for it. It can be noticed in Fig. 4 that, the introduced “SC+Learning” strategy has the most similar localization results to the ground truth among all three strategies. For “w/o Learning”, similarity learning is not incorporated and foreground extraction results are often not precise. It will badly influence the following CCA step, resulting in degraded tumor localization results. For “SVM+Learning”, although similarity learning is adopted, its localization results are

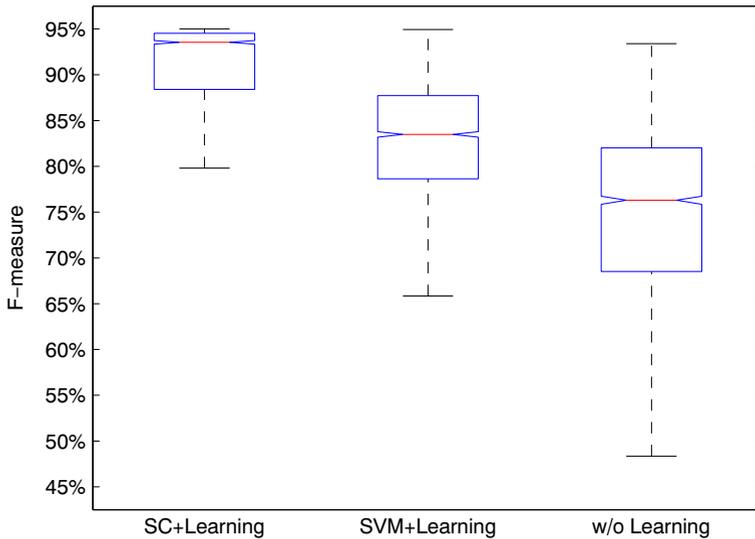
not as precise as those of the introduced “SC+Learning” strategy (e.g., matchings are not as good as ours in the 2nd, 5nd, 8nd images).

A quantitative analysis is conducted based on all obtained tumor localization results from all compared strategies with statistical perspectives. F-measure, the harmonic mean of precision and recall measures [10], is adopted here to evaluate the tumor localization performance quantitatively. The reason of adopting F-measure in this study is explained as follows. Provided the following annotations: TPs - true positives (i.e., tumor correctly identified); FPs - false positives (i.e., non-tumor wrongly identified as tumor); GT - ground truth. Definitions of precision and recall are as follows:  $precision = TPs / (TPs + FPs)$ ,  $recall = TPs / GT$ . Generally speaking, precision can be biased by the situation of under-segmentation, in which the segmentation result is only a tiny portion of the whole GT (i.e., in this case, FPs=0 and the precision value equals to 1); while recall can also be biased by the situation of over-segmentation, in which the segmentation result is a large overlapping on GT (i.e., in this case, TPs=GT and the recall value equals to 1). An illustration of the two above situations are displayed in Fig. 5. Hence, individual usage of precision or recall in evaluating the tumor localization performance in this study is not reasonable. Thus, F-measure, which combines precision and recall as their harmonic mean, is adopted in this study for quantitative evaluation of tumor localization performance. The definition of F-measure is:  $2 \times (precision \times recall) / (precision + recall)$ . Therefore, the two above mentioned biased conditions can be well tackled by F-measure, and F-measure can evaluate the tumor localization performance objectively.

Figure 6 illustrates a box-and-whisker plot, in which statistical boxes of F-measure results calculated based on tumor localization results and corresponding ground truth from all patients data are presented for all compared localization strategies. In each box, a horizontal red line is drawn across the box at the median; the upper and lower quartiles are blue lines above and below it. A vertical dashed line is drawn up from the upper quartile and down from the lower quartile to their most extreme data points, which are within a distance of 1.5 inter-quartile range [24]. It can be noticed that, the box of “SC+Learning” is obviously higher than those of other strategies, suggesting that the tumor localization performance of “SC+Learning” is superior to others. For “w/o Learning”, the range between upper and lower quartiles of its box is larger than others, indicating that pre-defined similarity (without learning) may work well for some cases, but not for all. Also, the very range of “SC+Learning” is the smallest among all compared strategies, suggesting that the



**Fig. 5** An illustration of the under-segmentation situation (left) and the over-segmentation situation (right)



**Fig. 6** Box-and-whisker plot of F-measure among all three compared strategies based on tumor localization results from all patients data

introduced similarity learning method via spectral clustering may provide the most stable tumor localization performance among all compared strategies.

A detailed statistical test made up of one-way analysis of variance (ANOVA) followed by post-hoc multiple comparison tests [24] is utilized for further statistical evaluation. In one-way ANOVA, F-measure results from all strategies are compared to test a hypothesis ( $H_0$ ) that, F-measure means of various strategies are equivalent, against the general alternative that these means cannot be all the same. P-value is used here as an indicator to reveal whether  $H_0$  holds or not. In this study, P-value calculated from F-measure results from all strategies are nearly 0, which is a strong indication that all these strategies cannot share the same F-measure mean. Therefore, the next step is to conduct more detailed paired comparison. The reason to do so is because the alternative against  $H_0$  is too general. Information about which method is superior cannot be perceived by one-way ANOVA alone. Therefore, tests that can provide such information are needed and they are multiple comparison tests.

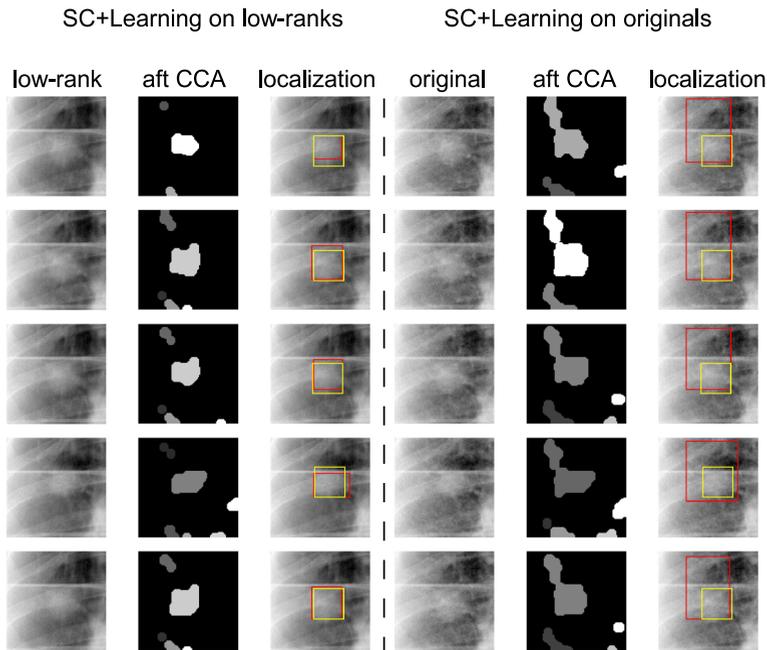
There are two kinds of evaluation results after applying multiple comparison tests on F-measure outcomes obtained by all three compared strategies. One is the estimated mean difference, which is a single-value estimator of F-measure mean difference between any two compared strategies. Another is a 95 % confidence interval (CI), which is a special form of interval estimator for a parameter (e.g. F-measure mean difference in this study). Generally speaking, instead of estimating the parameter by a single value, CI provides an estimated range, which is likely to include the estimated parameter from statistical perspectives. To be specific, for the comparison between “SC+Learning” and “SVM+Learning”, the F-measure of “SC+Learning” is 0.1532 higher than that of “SVM+Learning”, and the F-measure mean difference (using “SC+Learning” minus “SVM+Learning”) is likely to fall within a 95 % CI between 0.1423 and 0.1641. Since the upper and lower bounds of the CI are both positive, it gives a strong indication (> 95 %) that “SC+Learning” is superior to

“SVM+Learning” in terms of F-measure from the statistical point of view. For the comparison between “SC+Learning” and “w/o Learning”, the analysis is similar. “SC+Learning” is 0.3269 higher than “w/o Learning”. The F-measure mean difference (using “SC+Learning” minus “w/o Learning”) is likely to fall within a 95 % CI between 0.3160 and 0.3378. Since the upper and lower bounds of the CI are both positive as well, it gives a strong indication (> 95 %) that “SC+Learning” is superior to “w/o Learning” in terms of F-measure from the statistical point of view.

To sum up, based on the above quantitative F-measure results and their corresponding comprehensive statistical analysis, it can be concluded that “SC+Learning” outperforms compared strategies based on tumor localization results from all patients data in terms of F-measure from the statistical point of view. The above quantitative statistical outcomes also substantiate the qualitative observation of boxes from Fig. 6 regarding the three compared strategies.

### 5.3 Experiments and analysis on low-rank & sparse decomposition

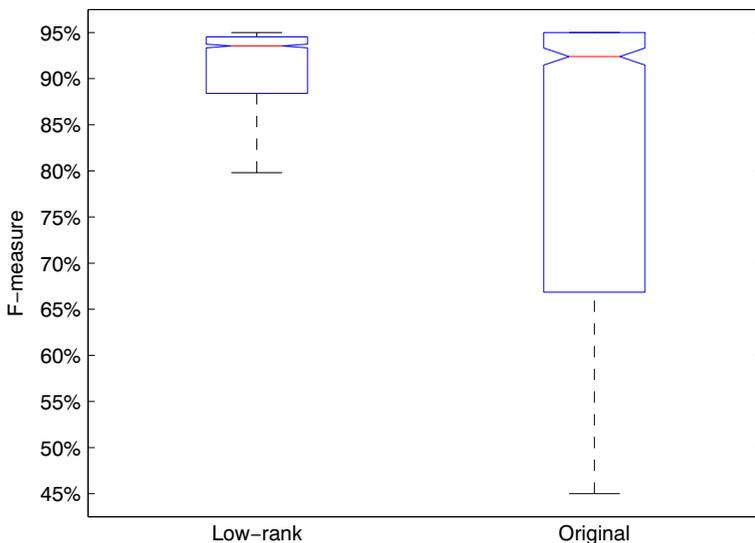
In order to demonstrate the superiority of adopted low-rank fluoroscopic images in the marker-less tumor localization study, the same steps described in Sections 3 and 4 are implemented on original fluoroscopic images for comparison (i.e., without the low-rank & sparse decomposition introduced in Section 2). Figure 7 depicts tumor localization results obtained by incorporating the introduced “SC+Learning” strategy on several example original fluoroscopic images (rows of Fig. 7 on the right hand side of the dashed line) and their corresponding decomposed low-rank fluoroscopic images (rows of Fig. 7 on the left



**Fig. 7** An illustration of tumor localization results provided by the introduced strategy on low-rank and original fluoroscopic images for comparison

hand side of the dashed line), respectively. It can be clearly observed that, although original fluoroscopic images (i.e., 4th column of Fig. 7) do not demonstrate much visual difference from their corresponding low-rank images (i.e., 1st column of Fig. 7), intermediate tumor localization results after CCA vary so much (i.e., 2nd and 5th column of Fig. 7, the same row for comparison). The reason is because that, low-rank fluoroscopic images can be less influenced by non-tumor tissue, which do not move in a high amplitude and is likely to be wrongly regarded as tumor lesion (i.e., revealed via bigger detected tumor areas on 5th column of Fig. 7). Hence, accompanying tumor localization results vary so much accordingly (i.e., 3rd and 6th column of Fig. 7). It can be concluded from examples in Fig. 7 that, applying “SC+Learning” on low-rank fluoroscopic images can provide better matchings between localized tumor (in red) and the ground truth (in yellow), compared with applying “SC+Learning” directly on original fluoroscopic images.

Similar statistical analysis has been conducted based on calculated F-measure results obtained by incorporating “SC+Learning” on low-rank and original fluoroscopic images using all patients data. A box-and-whisker plot is shown in Fig. 8. It can be noticed that, applying the introduced tumor localization strategy on low-rank fluoroscopic images outperforms applying the same strategy directly on original fluoroscopic images in terms of F-measure. Also, “SC+Learning” on low-rank images is 0.1987 higher than “SC+Learning” on original images. The F-measure mean difference (using “low-rank” minus “original”) is likely to fall within a 95 % CI between 0.1836 and 0.2139. Since the upper and lower bounds of the CI are both positive, it gives a strong indication (> 95 %) that, “SC+Learning” on low-rank images is superior to “SC+Learning” on original images in terms of F-measure from the statistical point of view. Thus, the superiority of adopting low-rank & sparse decomposition in this tumor localization study is also revealed.

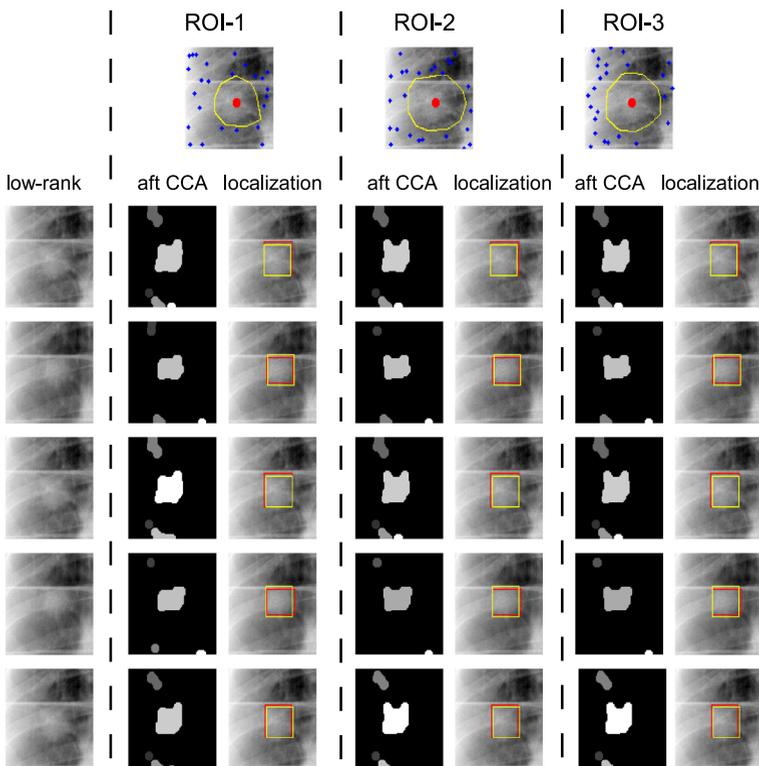


**Fig. 8** Box-and-whisker plot of F-measure outcomes of tumor localization results obtained by incorporating “SC+Learning” on low-rank and original fluoroscopic images of all patients

## 5.4 Discussion

In Section 3.1, a pre-requisite user interaction step is adopted to incorporate clinicians' prior knowledge about the location of the tumor lesion on the 1st low-rank fluoroscopic image of each image sequence of one patient. Its main purpose is to sample pixels for similarity learning. In this section, the influence of different ROI drawings on tumor localization results in this study is discussed.

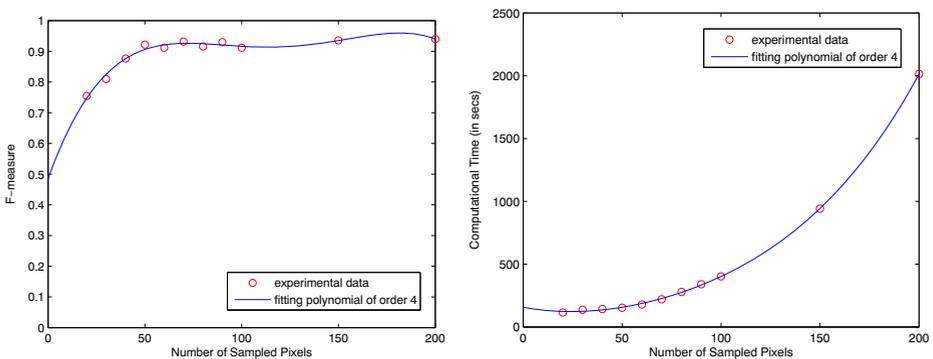
Figure 9 illustrates tumor localization results obtained by applying the introduced strategy on low-rank fluoroscopic images with 3 different delineated ROIs for instance. It can be noticed that, initial drawings vary much from ROI-1 to ROI-3, and distributions of sampled training data are quite diverse brought by different ROIs therein. Based on their tumor localization results, it is interesting to see that although ROIs and data samplings are different, detected tumor areas after CCA does not vary much on each row corresponding to the same low-rank fluoroscopic image. Therefore, tumor localization results are similar as well. To be specific, the average F-measure of the whole image sequence of the patient is 0.9398, 0.9402, 0.9387 for ROI-1, ROI-2, and ROI-3, respectively. It can also be observed that, tumor localization results (in red) and their ground-truth (in yellow) are well matched. Another quantitative test of adopting 3 different ROIs on each of all patients within the



**Fig. 9** An illustration of tumor localization results provided by the introduced tumor localization strategy on example low-rank fluoroscopic images with different delineated ROIs

introduced tumor localization strategy also suggests that, the mean F-measure difference percentage is 7.5464 %. It is worth to note that, the main function of ROI drawings is to provide positive and negative samples for similarity learning in tumor localization. It is observed that detected tumor lesions do not vary much based on different delineated ROIs for each patient. The reason is because that, representative samples can be picked following the adopted sampling strategy even if different ROIs are delineated. Learning results will not be deteriorated when representative samples are chosen following the incorporated sampling strategy. Therefore, ROIs are not required to be drawn to closely match the tumor lesion. Also, the stability of the introduced strategy will not be influenced much by the pre-requisite user interaction step.

Another interesting thing to note is that, totally 50 sampled pixels (equally divided into positive and negative samples) are picked for similarity learning. The number of sampled pixels is determined with consideration of both effectiveness and efficiency of the tumor localization strategy. Generally speaking, more sampled pixels would bring more information that could help to improve learning results. However, computational time spent during the learning process (i.e., especially cost in the eigen-decomposition of the utilized graph Laplacian in Step 4 of Table 1) would increase largely. The increasing sampled pixels would somehow become “redundant”, if less significant improvement on the tumor localization performance can be obtained. In our experimental design, several numbers of samples are tested using all patients data, and their obtained F-measure changes as well as computational time cost for the tumor localization task are evaluated. The left sub-figure of Fig. 10 is related to F-measure and number of sampled points. Red hollow points represent sampling numbers that are tested (including 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200. Positive and negative samples are equal) and their obtained F-measure results based on all patients data. A blue polynomial of order 4 is fit towards those isolated hollow points, to reveal the trend. It can be noticed that, the tumor localization performance increases dramatically from 20 to 50 (i.e., F-measure increases from 0.7545 to 0.9217), and is stable when more samples are added in thereafter (i.e., F-measure from 50 to 200 resides in the range [0.9115, 0.9398]). Thus, more samples would not help much after 50. The right sub-figure of Fig. 10 is related to the computational time of tumor localization and number



**Fig. 10** (Left) F-measure obtained by different numbers of samples on all patients data. (Right) Computational time (in secs) spent using different numbers of samples on all patients data

of sampled points. Red hollow points represent sampling numbers that are tested (including 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200). Positive and negative samples are equal and their computational time spent. A blue polynomial of order 4 is also fit towards those isolated hollow points, to reveal the trend. It can be noticed that, the computational time doesn't vary much when the number of samples is no more than 100, but will increase dramatically after that. Therefore, the number of 50 incorporated data sampling of the similarity learning in this study is determined with consideration of both effectiveness and efficiency.

## 6 Conclusion

In this study, a novel marker-less tumor localization strategy on low-rank fluoroscopic images is introduced. Compared with other general data-driven methods, the proposed strategy has several merits. First, fluoroscopic images are often of poor quality and robust-PCA is to be incorporated into this medical tracking application for “cleaner” images, inspired by other recent tracking studies in computer vision. Second, the conventional pattern recognition way to conduct similarity learning cannot tackle the problem in real-life patients well since imaging data of patients vary so much. Thus, similarity learning based on unique data from each particular patient is realized in this proposed strategy via spectral clustering. The superiority of low-rank & sparse decomposition as well as similarity learning via spectral clustering utilized in this strategy is verified by dozens of experiments together with comprehensive statistical analysis. Promising results are demonstrated on real patient data. Future studies will be continued following the introduced tumor localization framework, and more sophisticated tumor localization techniques will be investigated. The localization strategy will also be studied for other types of tumor lesions.

**Acknowledgments** This work is supported by 61363046, 61301194, and 61302121 approved by National Natural Science Foundation China, 20142BBE50023, 20142BAB217033 and 20142BAB217030 approved by Jiangxi Provincial Department of Science and Technology, as well as NWPU grant 3102014JSJ0014.

## References

1. Bach F, Jordan M (2006) Learning spectral clustering, with application to speech separation. *J Mach Learn Res* 7:1963–2001
2. Belkin M, Niyogi P (2003) Laplacian Eigenmaps for dimensionality reduction and data representation. *Neural Comput* 15(6):1373–1396
3. Bengio Y, Paëment J, Vincent P, Delalleau O, Roux N, Ouïmet M (2003) Out-of-sample extension for LLE, Isomap, MDS, Eigenmaps, and spectral clustering. *Adv Neural Inf Process Syst* 857–863
4. Bishop C (2007) *Pattern recognition and machine learning*, 1st edn. Springer
5. Boykov Y, Veksler O, Zabih R (2001) Fast approximate energy minimization via graph cuts. *IEEE Trans Pattern Anal Mach Intell* 23:1222–1239
6. Cai J, Candes E, Shen Z (2010) A singular value thresholding algorithm for matrix completion. *SIAM J Optim* 20(4):1956–1982
7. Candes E, Li X, Ma Y, Wright J (2009) Robust principal component analysis. *Cornell University Library (arXiv.org)* 0912:3599:1–39
8. Ferlay J, Shin H, Bray F (2010) Estimates of worldwide burden of cancer in 2008: Globocan 2008. *Int J Cancer* 127(12):2893–2917

9. Ferrell B, Koczywas M, Grannis F, Harrington A (2011) Palliative care in lung cancer. *Surg Clin N Am* 91(2):403–417
10. Fowlkes C (2010) Segmentation consistency measures. [http://www.cs.berkeley.edu/fowlkes/tutorial/metrics\\_thesis\\_03.pdf](http://www.cs.berkeley.edu/fowlkes/tutorial/metrics_thesis_03.pdf)
11. Goldstein T, Osher S (2009) The split Bregman method for L1-regularized problems. *SIAM J Imaging Sci* 2(2):323–343
12. Gonzalez R, Wintz P, Woods R (2002) *Digital image processing*, 2nd edn. Prentice Hall Press
13. Hoisak J, Sixel K, Tirona R, Cheung P, Pignol J (2006) Correlation of lung tumor motion with external surrogate indicators of respiration. *Int J Radiat Oncol Biol Phys* 60(4):1298–1306
14. Huang W, Li J, Zhang P, Wan M (2013) A novel marker-less tumor tracking strategy on low-rank fluoroscopic images for image-guided lung cancer radiotherapy. *Int Conf Image Process*:1399–1403
15. Hsu D, Kakade S, Zhang T (2011) Robust matrix decomposition with sparse corruptions. *IEEE Trans Inf Theory* 57(11):7221–7234
16. Ionascu D, Park S, Killoran J, Allen A, Berbeco R (2008) Application of principal component analysis for marker-less lung tumor tracking with beam's-eye-view epid images. *Med Phys* 35(6):2893. (1 page)
17. Isaksson M, Jalden J, Murphy M (2005) On using an adaptive neural network to predict lung tumor motion during respiration for radiotherapy applications. *Med Phys* 32(12):3801–3809
18. Kothary N, Dieterich S, Louie J, Chang D, Hofmann L, Sze D (2009) Percutaneous implantation of fiducial markers for imaging-guided radiation therapy. *Am J Roentgenol* 192(4):1090–1096
19. Li X, Xu H, Mukhopadhyay S, Balakrishnan N, Sawant A, Iyengar P (2012) Toward more precise radiotherapy treatment of lung tumors. *IEEE Comput* 45(1):59–65
20. McNair H, Kavanagh A, Powell C, Symonds-Taylor J, Brada M, Evans P (2012) Fluoroscopy as a Surrogate for lung tumour motion. *Br J Radiol* 85:168–175
21. Mountain C (1997) Revisions in the international system for staging lung cancer. *Chest* 111(6):1710–1717
22. Ng A, Jordan M, Weiss Y (2002) On spectral clustering: analysis and an algorithm. *Adv Neural Inf Process Syst* 64–72
23. Riaz N, Shanker P, Wiersma R, Gudmundsson O, Mao W, Widrow B, Xing L (2009) Predicting respiratory tumor motion with multi-dimensional adaptive filters and support vector regression. *Phys Med Biol* 54(19):5735–5748
24. Rice J (2006) *Mathematical statistics and data analysis*, cengage learning, 3rd edn
25. Roweis S, Saul L (2000) Nonlinear dimensionality reduction by locally linear embedding. *Science* 290(5500):2323–2326
26. Ruan D, Fessler J, Balter J (2007) Real-time prediction of respiratory motion based on local regression methods. *Phys Med Biol* 52(23):7137–7152
27. Seppenwoolde Y, Shirato H, Kitamura K, Shimizu S, van Herk M, Lebesque J, Miyasaka K (2002) Precise and real-time measurement of 3d tumor motion in lung due to breathing and heartbeat measured during radiotherapy. *Int J Radiat Oncol Biol Phys* 53(4):882–834
28. Shi J, Malik J (2000) Normalized cuts and image segmentation. *IEEE Trans Pattern Anal Mach Intell* 22:888–905
29. Shimizu S, Shirato H, Ogura S, Dosaka-Akita H, Kitamura K, Nishioka T, Kagei K, Nishimura M, Miyasaka K (2001) Detection of lung tumor movement in real-time tumor-tracking radiotherapy. *Int J Radiat Oncol Biol Phys* 51(2):304–310
30. Shirato H, Shimizu S, Kitamura K, Nishioka T, Kagei K, Hashimoto S, Aoyama H, Kunieda T, Shinohara N, Dosaka-Akita H, Miyasaka K (2000) Four-dimensional treatment planning and fluoroscopic real-time tumor tracking radiotherapy for moving tumor. *Int J Radiat Oncol Biol Phys* 48(2):435–442
31. Vapnik V (1998) *Statistical learning theory*. Wiley
32. Zhang Z, Zha H (2004) Principal manifolds and nonlinear dimension reduction via local tangent space alignment. *SIAM J Sci Comput* 26(1):313–338
33. Zhang T, Ghanem B, Liu S, Ahuja N (2012) Low-rank sparse learning for robust visual tracking. *Lect Notes Comput Sci* 7577:470–484
34. Zhou Z, Li X, Wright J, Candes E, Ma Y (2010) *Stable Principal Component Pursuit*, Cornell University Library (arXiv.org) 1001:2363:1–5
35. Zhou T, Tao D (2013) Shifted subspaces tracking on sparse outlier for motion segmentation. *Proc Int Jt Conf Artif Intell* 1946–1952



**Wei Huang** obtained his B.Eng. and M.Eng. degrees from the Harbin Institute of Technology, China in 2004 and 2006, respectively. He obtained his Ph.D. degree from the Nanyang Technological University, Singapore in 2011. Before joining the Nanchang University as an Associate Professor, he worked in the University of California San Diego, USA, as well as the Agency for Science Technology and Research, Singapore as Research Staffs. Dr. Huang's recent research interests mainly include but not limited to medical image processing, pattern recognition, and computer vision.



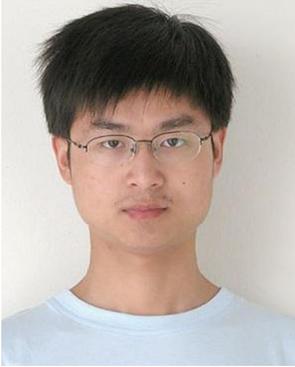
**Jing Li** received the Ph.D. degree from the Department of Electronic and Electrical Engineering at the University of Sheffield, U.K., in 2012. She is currently an Associate Professor with the School of Information Engineering at Nanchang University, China. Her research interests include content-based image retrieval, pattern recognition, and video analysis. She has authored or coauthored in various peer-reviewed journals, such as IEEE Transactions on Industrial Informatics, Information Sciences (Elsevier), etc.



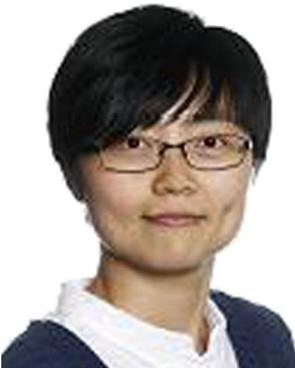
**Peng Zhang** received the B.E. degree from the Xi'an Jiaotong University, China in 2001. He received his Ph.D. from Nanyang Technological University, Singapore in 2011. He is now an Associate Professor in School of Computer Science, Northwestern Polytechnical University, China. His current research interests include image processing, signal processing, and computer vision. He is a member of IEEE. He is the TPC member and reviewer of several high ranked international conferences and journals.



**Min Wan** received B.Eng. and M.Eng. degrees from Beijing University of Post and Telecom as well as China Academy of Telecom Technology, Beijing, China in 2004 and 2008, respectively, and was awarded the Ph.D degree by Nanyang Technological University in 2012. He worked in National Heart Centre Singapore as Research Fellow from 2012 to 2013, and then a scientist in Institute of High Performance Computing, Agency for Science, Technology and Research (A\*STAR) from 2013 to 2014. He is now an Associate Professor in Nanchang University. His research interests include medical imaging, computational geometry, computer graphics, and mesh generation.



**Can Fang** received his B. Eng. degree from Xi'an Jiaotong University, China in 2001, and Ph.D. from Nanyang Technological University, Singapore in 2009. He is currently the faculty of Southwestern University, China. Dr. Fang has strong industrial and academic experience in computer science, optics and electronics. His research interests include image processing, sensor networks, wireless and mobile communications, optics and electronic devices. Dr. Fang is holding several patents in U.S and China; he has served as reviewer for several journals and conferences.



**Minmin Shen** received the B.Sci. degree in information engineering from the Shanghai Jiao Tong University, Shanghai, China, in 2006, and the Ph.D. degree from the Nanyang Technological University, Singapore, in 2011. She was a Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University from 2010 to 2011. She is currently a Postdoctoral Fellow at the Department of Computer and Information Science, University of Konstanz, Germany. She is also with School of Software Engineering, South China University of Technology. Her current research interests include biological image/video processing, computer vision, super-resolution, video coding, and video/image enhancement.