An Automated Segmentation Algorithm for CT Volumes of Livers with Atypical Shapes and Large Pathological Lesions

Shun UMETSU††, Nonmember, Akinobu SHIMIZU†††, Member, Hidefumi WATANABE††, Nonmember, Hidefumi KOBATAKE†††, Fellow, and Shigeru NAWANO††††, Nonmember

SUMMARY This paper presents a novel liver segmentation algorithm that achieves higher performance than conventional algorithms in the segmentation of cases with unusual liver shapes and/or large liver lesions. An L1 norm was introduced to the mean squared difference to find the most relevant cases with an input case from a training dataset. A patient-specific probabilistic atlas was generated from the retrieved cases to compensate for livers with unusual shapes, which accounts for liver shape more specifically than a conventional probabilistic atlas that is averaged over a number of training cases. To make the above process robust against large pathological lesions, we incorporated a novel term based on a set of “lesion bases” proposed in this study that account for the differences from normal liver parenchyma. Subsequently, the patient-specific probabilistic atlas was forwarded to a graph-cuts-based fine segmentation step, in which a penalty function was computed from the probabilistic atlas. A leave-one-out test using clinical abdominal CT volumes was conducted to validate the performance, and proved that the proposed segmentation algorithm with the proposed patient-specific atlas reinforced by the lesion bases outperformed the conventional algorithm with a statistically significant difference.

key words: CT volume, liver segmentation, lasso, lesion, probabilistic atlas

1. Introduction

Liver segmentation from a CT volume is one of the most popular segmentation problems in the field of medical image analysis, and it is useful for computer-aided diagnosis and computer-aided surgery. There are a number of papers on liver segmentation [1]–[5]. The most sophisticated algorithm [2] was presented by a winner of the first MICCAI Grand Challenge in 2007 [4]. This algorithm is a fully automated algorithm and proved the usefulness of statistical shape information for the liver [6]. Multiphase contrast-enhanced CT volumes would add useful information to enhance liver segmentation performance, particularly in cases with lesions [5]. However, only single-phase contrast-enhanced imaging is available in many clinical situations because of the dose problem. Liver segmentation from a single-phase contrast-enhanced CT volume is still a challenging problem, mainly because of livers with large pathological lesions and livers with atypical shapes that lie in peripheral areas of the liver shape distribution, which are difficult to represent using a statistical shape model.

A recent topic of importance is multiple abdominal organ segmentation [7]–[11]. A well-known advantage of multiple organ segmentation is the ability to use the relationships between neighboring organs, which can solve inconsistencies between neighboring organs, leading to higher segmentation performance. A combination of statistical shape information and multiple organ segmentation is a promising method to boost segmentation performance even further [9], [10]. Current disadvantages of the above methods are high computational costs, and organs outside the imaging area. For example, a multiple level set-based segmentation algorithm [8] used a cluster computer where a CPU was assigned to each organ, but the computational cost remained high owing to the cost incurred when exchanging shape and location information between CPUs of neighboring organs. In addition, some of the target organs might not be scanned in a CT volume either in whole or in part, which violates an implicit or explicit assumption employed by most algorithms that whole target organs are recorded in the input CT volume.

This study focuses on liver segmentation from a single-phase contrast-enhanced CT volume, or a portal venous phase CT volume, which is popular in the diagnosis of metastasis. We propose a novel method in this paper that is robust to large pathological lesions in liver and/or livers with an unusual shape. Pathological lesions, such as metastatic tumors, change CT values, as shown in Fig. 1. They make the liver segmentation problem difficult because of the large differences in CT values from that of normal liver parenchyma, resulting in a low likelihood of liver, as shown in Fig. 1 (c). In addition, the shape of this liver is atypical, and might not be precisely accounted for by a statistical shape model and a probabilistic atlas, which means that both a statistical model and an atlas-based method might not work in such a case. Figure 2 (a) is a 3D rendering of the liver from Fig. 1. Figures 2 (b) and 2 (c) show the average shape of the level set distribution model [12], or signed distance-based statistical shape model, and a probabilistic atlas with the true boundary marked on an axial slice, respectively. Figure 3 is a scatter plot of the training shapes, and the shape in Fig. 2 (a) in the eigenshape space of the level set distribution model. The plot of this case lies in a peripheral region of the shape distribution and the Mahalanobis distance of principal component scores from
Fig. 1 A liver with large lesions and its likelihood of liver: Panels (a) and (b) depict an axial slice showing the true boundary (red) and a coronal slice from a liver CT volume, respectively. (c) The likelihood of liver. Since the CT values of lesions are different from those of normal parenchyma, the likelihood decreases in the liver lesions.

Fig. 2 The liver shape from Fig. 1, an average shape, a conventional probabilistic atlas, and the segmentation result: (a) A 3D rendering of the liver shape from Fig. 1. (b) An average shape from the level set distribution model. (c) An axial slice from a conventional probabilistic atlas (with the true boundary marked in red) of (a). (d) A maximum a posteriori-based segmentation result using the probabilistic atlas from (c) and the likelihood from Fig. 1 (c). The difference in shape from the average shape causes inappropriate prior probability for conventional probabilistic atlases, resulting in segmentation failures. In addition, a part of the lesion was identified as other tissue because of the difference in CT values from normal liver parenchyma.

Fig. 3 A scatter plot of the liver shape from Fig. 2 (a) and training shapes (black dots) in an eigenshape space, constructed using 113 training liver shapes. The liver shape from Fig. 2 (a) is located in a peripheral region of the shape distribution as indicated by the red box and cross, which means that the shape is different from those of the training shapes, or an atypical shape.

This paper presents a novel liver segmentation algorithm that achieves higher performance in the segmentation of cases with large lesions and/or atypical shapes. The three main contributions of this paper are as follows.

First, a lasso (least absolute shrinkage and selection operator) [13] based patient-specific probabilistic atlas is proposed. An L1 norm is added to the mean squared difference so as to select the most relevant case as an input case from a training dataset, called a dictionary in this paper. The lasso algorithm minimizes the total cost function and a patient-specific probabilistic atlas is constructed from the selected cases.

Second, this paper presents a method that makes the lasso algorithm robust against pathological lesions. A set of “lesion bases” that was inspired by error bases [14] is incorporated in the minimization process of the lasso, leading to a segmentation result robust against lesions. The patient-specific probabilistic atlas reinforced by the lesion bases was incorporated in the graph-cuts-based fine segmentation.

Third, the performance of the proposed algorithm was validated in a leave-one-out fashion using clinical abdominal CT volumes of a portal venous phase taken in daily clinical routine. All volumes contain large pathological lesions of liver, and some in which the liver has an atypical shape.

In the remainder of this paper, we explain the details of our proposed method and demonstrate its efficacy using abdominal CT volumes. The quantitative evaluation was performed using the sum of the accuracy in the liver segmentation and sensitivity to a lesion. This is followed by a discussion of the efficacy of our proposed method.
2. Proposed Liver Segmentation Algorithm

2.1 Overview

The proposed algorithm consists of a probabilistic atlas-based rough segmentation and a graph-cuts-based fine segmentation. The rough segmentation process computes posterior probabilities of liver and other tissues based on the proposed patient-specific probabilistic atlas constructed using selected label volumes from a training label dataset, called a dictionary throughout this paper. The graph-cuts-based fine segmentation process finds the optimal solution of a submodular energy function with a shape prior term estimated from the rough segmentation result [15]. This section will explain the details of each process.

2.2 Patient-Specific Probabilistic Atlas via a Sparse Representation

A patient-specific probabilistic atlas is constructed from label volumes in a dictionary. To make the probabilistic atlas adaptive to not only an average shape but also to an atypical shape, a few liver label volumes are selected instead of all labels. The selection process is conducted using a pair of an input CT volume and a liver likelihood volume that has values close to one inside the liver and zero outside the liver. A CT volume is used to make the selection result robust against the difference between the likelihood and liver label. Intuitively speaking, we select pairs of CT volumes and liver label volumes from a dictionary whose linear combination approximates the input pair. The L1 penalty term of the coefficients for the linear combination is introduced to focus on the relevant cases in terms of CT values and labels. Consequently, we formulate the selection process as the following minimization.

\[
\min_{\beta, \beta_0} \left\{ \frac{1}{2N} \| y - X\beta - \beta_0 \|_2^2 + \lambda \| \beta \|_1 \right\}
\]  

(1)

where \( y \) and \( \beta \) are vectors whose sizes are \( 2N \times 1 \) and \( p \times 1 \), and whose elements are \( y_i \) and \( \beta_j \), respectively. \( X \) is a matrix of size \( 2N \times p \) and an element is \( x_{ij} \). Symbol \( y_i \) denotes the \( i \)th element, or \( i \)th voxel, of a vector that concatenates a vectorized CT volume and a vectorized liver likelihood volume of a given test case, and symbol \( x_{ij} \) indicates the \( j \)th element of a vector of the \( j \)th vector in a dictionary that concatenates a vectorized CT volume and a vectorized label volume. Constant \( N \) is equal to the number of voxels in a slab, as explained in the next paragraph, and constant \( p \) is the number of training vectors, or the total number of slabs from all training cases (\( K \) cases in total), respectively. Coefficients \( \beta_j \) and \( \lambda \) correspond to the \( j \)th vector in a dictionary and a weight for the L1 penalty term. Setting a large \( \lambda \) causes shrinkage, resulting in most of the coefficients equaling zero. Otherwise, it solves a conventional least squares problem, leading to many non-zero coefficients. The reason for using the L1 penalty term is to make the probabilistic atlas more specific to the patient, leading to higher performance in segmentation. \( \beta_0 \) shows an intercept of the approximation, and \( I \) is a vector of size \( 2N \times 1 \), in which all elements are 1.

In the remainder of Sect. 2.2, we describe a procedure to generate vectors in a dictionary as well as an input vector, followed by a procedure to construct a patient-specific probabilistic atlas.

To generate vectors in our dictionary, we downsampled an original CT volume as well as the corresponding label volume, and divided them into several slabs, each of which consists of several slices. This was for the purpose of reducing the memory required for the minimization process. Note that the slabs in the dictionary overlap each other, resulting in a sufficient number of label volumes to boost the ability of representation via lasso [14]. The detailed process is as follows (see also Fig. 4).

i. Spatial standardization of CT volumes and liver label volumes to suppress the variation in shape and location of livers [8].

ii. Downsampling the volumes from \( 512 \times 512 \times (\text{original z-size}) \) to \( 32 \times 32 \times (\text{original z-size})/16 \).

iii. Division of the volumes into several slabs, each of which consists of \( n \) axial slices. Note that the overlap between neighboring slabs is \( n - 1 \) axial slices.

iv. Normalization of CT values so that the variance and mean in a slab are equal to one and zero, respectively. This process is carried out after limiting the range of CT values from \(-20\) to \(200\) HU, so as to focus on the liver.

v. Vectorization of the slab of normalized CT values and that of the corresponding labels followed by concatenation of the two vectors. Note that \( N = 32 \times 32 \times n \) and \( p = \sum_{k=1}^{K} m_k \) where \( m_k \) is the number of slabs from case \( k \).

The procedure for generating a vector for an input test case is the same as that for vectors in a dictionary, except that an input test CT volume is divided into several slabs without overlap (see Fig. 5). Because a set of non-overlapping slabs from an input test volume are sufficient for segmentation. In addition, a likelihood volume is used instead of a label volume, because a test case lacks a label volume. The likelihood of each voxel is calculated under the assumption that the distribution of CT values in a CT volume can be described by a mixture of five Gaussians; two Gaussians for the liver and three Gaussians for surrounding tissues, or heart, right kidney and others. All statistical parameters for the mixture of Gaussians are estimated from an atlas-guided EM algorithm [8], and the likelihood of liver is defined as a mixture of two Gaussians of liver. In addition, we performed a normalization of the likelihood such that the sum of likelihoods of different tissues is equal to one at each voxel. Therefore, the likelihood will be similar to the label volume when no pathological lesion is present in the liver. Finally, a slab and a likelihood volume are transformed into vectors and concatenated into a vector, which is approximated by vectors in the dictionary via the minimization of
Fig. 4  A procedure to generate vectors in a dictionary from a CT volume and the corresponding liver label volume. In this study, the procedure is applied to all training data, or pairs of CT volumes and the corresponding liver label volumes.

Fig. 5  A procedure to generate a vector from an input test case consisting of pairs of CT volumes and liver likelihood volumes.
(1) using the lasso algorithm. Note that the lasso algorithm is applied to each input vector independently, and it outputs a set of $\beta j$ corresponding to the $j^{th}$ vector in the dictionary.

A patient-specific probabilistic atlas is constructed from the selected label slabs, or label slabs with non-zero $\beta j$ values. We combined pairs of $\beta j$ and the $j^{th}$ label ($j = 1, \ldots, p$) in the dictionary to estimate the probability of the existence of liver $\hat{p}_{\text{liver}}(i)$ using the following equation.

$$
\hat{p}_{\text{liver}}(i) = \begin{cases} 
0 & \text{if } \hat{y}^*(i) < 0 \\
\hat{y}^*(i) & \text{if } 0 \leq \hat{y}^*(i) \leq 1 \\
1 & \text{if } \hat{y}^*(i) > 1 
\end{cases}
$$

(2)

where $\hat{y}^*(i) = \beta_0 + \sum_{j=1}^{p} \beta_i x_{ij}$

where symbol $x_{ij}$ indicates of the $i^{th}$ element, or $i^{th}$ voxel, of the $j^{th}$ vectorized label volume in a dictionary. The computation of the $\beta j$ coefficients and construction of a patient-specific probabilistic atlas are performed for each slab independently, and the patient-specific probabilistic atlases of all slabs are concatenated to construct a 3D probabilistic atlas whose size is equal to that of an input CT volume (Fig. 5).

2.3 Simultaneous Estimation of Probabilities of Normal Liver Parenchyma and Pathological Lesions

Pathological lesions in a liver change the likelihood of liver in an input CT volume. The likelihood in a lesion becomes lower than that of normal liver parenchyma, as shown in Fig. 1 (c), leading to failure in selecting label slabs from the dictionary via the lasso algorithm. Inappropriate label sets result in incorrect prior probabilities for a patient-specific probabilistic atlas, which can be confirmed in Fig. 6.

To compensate for this problem, we introduced a set of "lesion bases" that were inspired by error bases [14], which can depict lesions in the minimization process using the lasso algorithm. We assume a vector $y$ from an input case with pathological lesions can be better approximated by the right hand side of the following equation.

$$
y \approx (X \ L) \left( \begin{array}{c} \beta \\ \beta_e \end{array} \right) + \beta_0 1
$$

(3)

where the symbols $y$, $X$, $\beta$, $\beta_0$, and $1$ are the same in (1) and $L$ is the proposed lesion basis matrix of size $2N \times N$ that concatenates an $N \times N$ zero matrix and an $N \times N$ diagonal matrix where an element $l_{ii}$ ($i = 1, \ldots, N$) is one when it meets the condition of lesion, and zero otherwise. Here, the lesion condition is defined by CT values from the portal venous phase. As we focus on metastatic liver tumors, a voxel with CT values larger than or equal to 20 HU, and less than or equal to the average CT value of liver minus one standard deviation (estimated in Sect. 2.2) is considered a voxel in a lesion. In this paper, column vectors in the bottom half of the lesion basis matrix are named lesion bases. Coefficient vector $\beta_e$ is a vector of size $2N \times 1$, each element of which represents a weight for a lesion basis. In accordance with the assumptions of (3), we minimize the following equation instead of (1).

$$
\min_{\beta, \beta_e, \beta_0} \left\{ \frac{1}{2N} \left\| y - (X \ L) \left( \begin{array}{c} \beta \\ \beta_e \end{array} \right) - \beta_0 1 \right\|^2 + \lambda (\|\beta\|_1 + \|\beta_e\|_1) \right\}
$$

(4)

Note that (4) includes not only $\beta$ but also $\beta_e$, which correspond to lesion bases, and which allows simultaneous estimation of normal liver regions by $\beta$ and pathological lesions by $\beta_e$ as follows. The probability of existence of normal liver parenchyma can be estimated by (2), and that of pathological lesions can be estimated by (5) (see Fig. 7).
\[ \hat{p}_{\text{lesion}}(i) = \begin{cases} 0 & \text{if } \hat{y}_{\text{lesion}}(i) < 0 \\ \hat{y}_{\text{lesion}}(i) & \text{if } 1 \geq \hat{y}_{\text{lesion}}(i) \geq 0 \\ 1 & \text{if } \hat{y}_{\text{lesion}}(i) > 1 \end{cases} \]

where \(\hat{y}_{\text{lesion}}(i) = -\beta_{v,i}\)

where \(\beta_{v,i}\) is the coefficient of the \(i\)th element. This equation was derived based on the fact that lesion bases were used to reduce the error in a lesion caused by an undesirable decrease in the likelihood; hence, \(\beta_{v,i}\) has negative values in the lesion. As a result, \(\hat{y}_{\text{lesion}}(i)\) was estimated by flipping the sign of \(\beta_{v,i}\). Eventually, \(\hat{p}_{\text{liver}}(i)\) and \(\hat{p}_{\text{lesion}}(i)\) are used to compute posterior probabilities of normal liver parenchyma as well as that of lesions, and will be forwarded to the next graph-cuts-based fine segmentation. Note that prior probability is normalized so that the sum of prior probabilities of all tissues is equal to one at each voxel.

Figure 8 (a) presents patient-specific probabilistic atlases of both liver and pathological lesions. Note that the areas with high prior probabilities in the lesion probabilistic atlas correspond to lesions. Figure 8 (b) shows the likelihoods of normal liver parenchyma and lesion computed using the parameters estimated by the EM algorithm (see Sect. 2.2). The images in Fig. 8 (c) correspond to the posterior probability \(P(l|I)\) of normal liver parenchyma and that of lesion, respectively, both of which are computed from Figs. 8 (a) and 8 (b) using (6).

\[ P(l|I) = P(I|l)P(l)/Z \]

where \(P(I|l)\) represents the likelihood of a CT value \(I\), \(P(l)\) is the prior probability of organ \(l\) given by the proposed probabilistic atlas, and the denominator \(Z\) is a normalization term. After combining both probabilities using a maximization operation at each voxel, we have a maximum a posteriori volume of the whole liver, as shown in Fig. 8 (d), which compensates for the decrease in posterior probability caused by lesions in the left image of Fig. 8 (c).

### 2.4 Graph-Cuts-Based Liver Segmentation

We performed a graph-cuts-based fine segmentation using the proposed patient-specific probabilistic atlas [15]–[17]. The energy function employed in this study is given below.

\[ E(A) = \lambda_g \sum_{v \in V} R_v(A_v) + \sum_{\{v,w\} \in \eta} (B_{v,w} + S_{v,w})\delta_{A_v \neq A_w} \]

\[ R_v(A_v) = \max_{l \in \{\text{liver}, \text{lesion}\}} P(I_v|l)P(l)/Z \]

\[ = \begin{cases} - & \max_{l \in \{\text{liver}\}} P(I_v|l)P(l)/Z \\ \cdots (A_v = \text{liver}) \\ \max_{l \in \{\text{heart}, \text{right kidney}, \text{others}\}} P(I_v|l)P(l)/Z \\ \cdots (A_v = \text{background}) \end{cases} \]

where the first term \(R_v(A_v)\), called the regional term, expresses a penalty for assigning label \(A_v\) to voxel \(v\). The symbol \(V\) represents the set of all voxels in a CT volume. This study employs the maximum a posteriori probability multiplied by \((-1)\) as the regional term (see (8)). The denominator \(Z\) is a normalization term. The maximum a posteriori probability of liver and that of the background, or heart, right kidney, and others are shown in Figs. 9 (a) and (b). Higher probability is represented by brighter colors. Here, the maximum a posteriori probability of background is calculated based on the likelihood derived from the mixture of Gaussians estimated in Sect. 2.2. and conventional probabilistic atlases. Figure 9 (c) shows the difference between the two maximum a posteriori images. It should be noted that the graph-cuts-based segmentation identifies a voxel with a brighter color in Fig. 9 (c) as liver. It was found that the regional term has high performance in discrimination of the whole liver, including lesions from other surrounding tissues.

The boundary term \(B_{v,w}\) and the shape term \(S_{v,w}\) are given below:
two probabilities, in which brighter voxels will be assigned to a liver label and tissues, heart, right kidney, or others, and (c) the difference between them with the true boundary marked in red: (a) The maximum a posteriori probability of liver, where voxels with brighter colors have a higher probability of representing whole liver, (b) the maximum a posteriori probability of surrounding tissues, heart, right kidney, or others, and (c) the difference between the two probabilities, in which brighter voxels will be assigned to a liver label in the graph-cuts-based segmentation.

\[ B_{v,w} = \exp\left(\frac{(I_v - I_w)^2}{2\sigma^2}\right) \cdot \frac{1}{\|v - w\|} \]  
\[ S_{v,w} = \sqrt{1 - \cos(\theta)} \]  

where \( B_{v,w} \) expresses a penalty for assigning labels \( A_v \) and \( A_w \) to the two neighboring voxels \( v \) and \( w \). The symbol \( \theta \) represents the angle between a vector connecting voxels \( v \) and \( w \), and a gradient vector of the signed distance \( \phi(v) \) from a boundary estimated by combining the statistical shape model with the maximum a posteriori-based rough segmentation of the liver [15]. This term encourages the energy to be low when the direction of a vector connecting \( v \) and \( w \) is similar to that of the gradient vector of \( \phi(v) \). As a result, the surface normals of an extracted region tend to be parallel to those of the estimated shape, and the extracted region is similar to the shape.

Symbol \( \eta \) in (7) is a set of neighboring voxel pairs, and the function \( \delta \) is 1 if \( A_v \neq A_w \), and 0 otherwise. Coefficient \( \lambda_s \) is a constant value in the equation balancing the terms. A detailed explanation of the boundary term \( B_{v,w} \) and shape term \( S_{v,w} \) can be found in [15].

3. Experiments

3.1 Materials and Performance Indices

Sixteen portal venous phase CT volumes with large lesions whose sphere-equivalent diameter was larger than 30 mm (43.5 ± 14.7 mm, max: 84.2 mm), as well as 16 liver label volumes manually delineated by experts were used for the experiments. All volumes were anonymized and approved by the IRB of the International University of Health and Welfare. The specifications of the original CT volumes are listed below.

- Size of volume [voxel]: 512 × 512 × (191–701)
- Voxel size [mm]: (0.625–0.735) × (0.625–0.735) × (0.5–1.0)

Isotropic transformation of the voxel size by linear interpolation was carried out before all processing. In addition, we performed spatial standardization and downsampling from 512×512×(original z-size) to 32×32×(original z-size)/16, as explained in Sect. 2.2. The lasso algorithm-based patient-specific probabilistic atlas construction was performed on the downsampled volume, and computation of the posterior probability as well as the graph-cuts-based fine segmentation were carried out in the 256 × 256 × (original z-size)/2 volume.

The validation study was conducted in a leave-one-out manner, in which 15 pairs of CT volumes and liver labels were used for a dictionary and one CT volume served as testing data. This was repeated such that all CT volumes were used for testing once. To reinforce the dictionary, we added an additional 27 pairs of portal venous phase CT volumes with small lesions and liver labels to the dictionary. The initial values of parameters for the CT value distribution and atlas-guided EM algorithm (see Sect. 2.2) were derived from the dictionary. Note that the training data in the dictionary and the test case were completely separated from each other in each round of leave-one-out testing.

The number of axial slices \( n \) in a slab in the dictionary and the number of overlapping slices between neighboring slabs were set throughout the experiment to three and two, respectively. The statistical shape model for computing the shape term of the graph-cuts-based segmentation was constructed using another 113 liver labels that were different from above 43 (= 16 + 27) cases. The number of axes and \( \lambda_s \) from (7) were set throughout the experiment to five and one, respectively [15]. The parameter \( \lambda \) from (1) and (4) was optimized experimentally between 0.001 and 0.1 using the training data from the leave-one-out test. The performance of the liver segmentation process was evaluated by computing the Jaccard indices, as well as the sensitivity to pathological lesions.

\[ JI = \frac{\#(A \cap B)}{\#(A \cup B)} \]  

where \( A \) and \( B \) denote a true region and an extracted region, respectively. \( \#(A) \) means the volume of \( A \). The sensitivity is defined as the ratio of the extracted region to the lesions [17]. A higher index indicates better performance.

3.2 Results

Figure 10 presents results based on a conventional probabilistic atlas, the patient-specific probabilistic atlas constructed using the lasso algorithm, and the patient-specific probabilistic atlas with lesion probabilistic atlas constructed...
by the lasso and lesion bases. The top row shows the probabilistic atlases and the second row consists of the likelihoods of liver and lesion. The third row presents a posteriori probabilities of whole liver by combining the probabilistic atlases and the likelihoods. The fourth row corresponds to differences in the regional terms of the liver and that of the background in the graph cuts, in which brighter represents a lower penalty for assigning a liver label. The fifth row shows the segmentation results of the graph cuts. The numerals in the bottom row indicate the JI and sensitivity to lesions, and the farthest right figure in the bottom row is the original CT volume.

Another case that is not atypical in shape, but contains a large lesion is shown in Fig. 11. Although the conventional probabilistic atlas describes the liver shape well, the decreased likelihood cannot be recovered by probabilistic atlases, resulting in poor segmentation performance. In contrast, the proposed method succeeded in generating effective regional terms for the graph–cuts-based segmentation where the lesion probabilistic atlas played an important role in generating the high posterior probability in the lesion.

These figures demonstrated that the proposed patient-specific probabilistic atlases outperformed a conventional probabilistic atlas, and the lasso algorithm with lesion bases achieved the best segmentation performance for a liver with atypical shape and large lesions.

Figure 12 illustrates boxplots of JIs, sensitivity, and the average of the two indices for all segmentation results, with the statistical test (Wilcoxon signed-rank test) results of the leave-one-out test. These figures suggest that segmentation performance with a conventional probabilistic atlas was improved by the patient-specific probabilistic atlas with a statistically significant difference. The lasso with lesion bases was significantly superior to that without lesion bases in terms of being particularly sensitive to lesions, resulting in high JIs.

4. Discussion

Segmentation of livers with atypical shapes and large lesions is a challenging problem. This paper proposed two methods, a lasso-based probabilistic construction and lasso with lesion bases to compensate for the deterioration of the posterior probability of liver due to pathological lesions.

The shape of the liver in Fig. 10 has an atypical shape, as indicated in Figs. 2 and 3, where the ventral part of left lobe is larger than average. The patient-specific atlas in the first row and second column of Fig. 10 proved that the lasso could construct a patient-specific probabilistic atlas that fitted the atypical shape. The result in the fifth row, however, indicated that the patient-specific atlas was not enough to compensate for the decrease in posterior probability of liver.
caused by the generally low CT values of lesions. In contrast, the lasso with lesion bases could generate the lesion probabilistic atlas and output high a posteriori probability in the lesion, resulting in high JIs for the segmentation (see the panel in the third column of the fifth row). The advantage of the lasso with lesion bases method was also proved by the segmentation results of the liver in Fig. 11, whose shape is close to an average shape, but includes a large lesion. This figure also indicated that the lasso with the lesion bases outperformed other methods. The superiority of the proposed method can be explained by the lesion probabilistic atlas. Although the likelihood of lesions enhanced not only lesions, but also surrounding tissues, the multiplication of the lesion probabilistic atlas and likelihood yielded an a posteriori probability that was high in pathological lesions and low in surrounding tissues.

Figure 12 summarizes the performance of all test cases evaluated in a leave-one-out fashion. The performance difference between a conventional probabilistic atlas and a patient-specific probabilistic atlas created using lasso seems to be small. However, the statistical difference was confirmed with \( p < 0.05 \) or \( p < 0.01 \). In contrast, the performance of the patient-specific atlas with the lesion probabilistic atlas was enormously superior to the other results. The difference from other results was statistically significant with very small p-values (\( p < 0.01 \)). The major part of the difference resulted from the difference in sensitivity to lesions. Successful segmentation of large lesions based on the lasso with lesion bases resulted in the success of the whole liver segmentation. In other words, high sensitivities to lesions led to high JIs of the whole liver. Figure 13 (a) shows the relationship between sensitivity and lesion diameter. All lesions were successfully detected by the proposed method (sensitivity > 0.9), except for the lesion indicated by a red arrow. Figure 13 (b) and (c) is an original CT image of the failure case and its segmentation result (yellow) with the true boundary marked in red, in which a part of the lesion close to boundary of the liver was not extracted because of the high irregularity of CT values in the lesion.

We introduced the concept of atypical-ness, which is the Mahalanobis distance from the mean shape in the eigen-shape space with five eigen axes. The maximum atypical-ness in our study was 4.19 and its JI was 0.889. Figure 14 presents the relationship between JI and atypical-ness. There is no significant correlation (cross correlation coefficient = −0.386 (\( p > 0.05 \)), which means that our lasso-based segmentation algorithm is robust against liver atypical-ness. Note that the case indicated by a red arrow is a failure case in which a part of the heart was overextracted resulting from a failure in the spatial standardization in Sect. 2.2 that made the location of the heart much lower than average.

Another concern of readers might be effectiveness of the additional 27 CT volumes with labels that were added to
Fig. 12  JIs (a), sensitivity (b) and the average of the two indices (c) evaluated using a leave-one-out test. "*" means a statistical difference with a p-value less than 0.05 and "**" with less than 0.01 in a Wilcoxon test, in which null hypothesis is "there is no statistical difference between the two methods".

Fig. 13  Relationship between sensitivity and lesion diameter (a) the original CT image (b) and its segmentation result (yellow) with the true boundary marked in red (c).

Fig. 14  Relationship between JI and Mahalanobis distance from an average shape (atypical-ness).

reinforce the dictionary, as explained in Sect. 3.1. We carried out experiments with and without an additional 27 pairs of portal venous phase CT volumes with small lesions and liver labels that were different from the test CT volumes. The JI improved from 0.899 to 0.906, sensitivity improved from 0.967 to 0.979, and the average of both indices improved from 0.933 to 0.942, on average. Although there was no statistical difference in the sensitivities, the JI and average of both indices were boosted significantly (p < 0.05). Thus, it can be concluded that additional CT volumes with labels were beneficial in enhancing performance.

The limitations of the proposed methods warrant attention. Figure 15 is a case with an atypical shape where the ventral part of liver is very small compared to the average shape, and the segmentation results are summarized in Fig. 16. As in the leftmost panel in the top row, the average probabilistic atlas protrudes outside the liver, as shown by the red line, while patient-specific atlases using the proposed methods in the second and third columns fit the liver regions. Although the combination of a probabilistic atlas and the likelihood of liver results in a high posterior probability in normal liver parenchyma, a decrease in the posterior probability in the lesions was observed in the panels in the leftmost and second columns in the third row. The proposed lasso with lesion bases compensated for the decrease, and
output high probabilities for the lesions, while it wrongly enhanced the gallbladder attached to the liver. The failure was caused by the high similarity in CT values between the gallbladder and the lesions. The high posterior probability in the gallbladder ended up labeled as liver, or as a false positive, in the graph-cuts-based segmentation, as denoted by a red arrow. Improvement of lesion bases using higher-order texture features or multi-organ segmentation that extracts both liver and gallbladder simultaneously might be a possible solution, and important future work.

5. Conclusion

This paper presented a liver segmentation algorithm to compensate for segmentation problems encountered in livers with atypical shapes and large lesions that have not been fully solved by the conventional methods. Although liver segmentation from a portal venous CT volume is one of the most popular topics in medical image analysis [4], there is no paper that focuses on livers with atypical shapes and/or large lesions in a single-phase contrast-enhanced CT volume, as well as the sensitivity to lesions in the validation process, which was evaluated in our study. Recently, an important paper was presented [18], in which a non-rigid ICP
algorithm was combined with a statistical shape model for the purpose of liver segmentation with pathological lesions attached to the liver surface. An L1 term was incorporated to detect surface outliers caused by pathological lesions. The results showed effectiveness and robustness against the pathological lesions. There are, however, a number of differences from their paper in terms of the statistical shape model employed, output of the objective function, and its purpose for use. Their statistical shape model was constructed based on statistics of the transformation matrices of the surface vertices, and directly fit the vertices to the liver surface by minimizing the objective function, including the proposed L1 norm. In contrast, our model was a level set distribution model, and the minimization of our objective function yielded a patient-specific probabilistic atlas and a lesion probabilistic atlas that were used in the graph-cuts-based fine segmentation. In addition, they did not mention applications to a liver with an atypical shape in the results of their algorithm, which is our concern in this study.

The experimental results in this paper showed that the proposed patient-specific probabilistic atlas statistically outperformed a conventional probabilistic atlas, and the probabilistic atlas constructed by the lasso with the lesion bases boosted the performance significantly. Important contributions of this paper are summarized as follows:

1) The first contribution is the proposal of the lasso-based liver segmentation algorithm that constructs a patient-specific probabilistic atlas to compensate for a liver with an atypical shape.

2) The patient-specific probabilistic atlas was reinforced by a lesion probabilistic atlas that was constructed by the lesion bases. The combination of a patient-specific atlas and a lesion probabilistic atlas yielded better regional terms for the graph-cuts-based fine segmentation.

3) We validated the proposed algorithms using anonymized abdominal CT volumes of the portal venous phase, in which large pathological lesions (diameters greater than 30 mm) were contained in all volumes, and some volumes included livers with atypical shapes. The leave-one-out test was employed to validate the proposed methods. Statistical tests among a conventional probabilistic atlas, the patient-specific probabilistic atlas, and the patient-specific probabilistic atlas with lesion probabilistic atlas were conducted. All test results supported the conclusion that the proposed probabilistic atlas with lesion probabilistic atlas statistically outperformed other atlases.

Our future work is detailed below:

1) Further improvement of the lesion bases method is scheduled to reduce false positives from surrounding tissues. This paper mainly focuses on CT values of a metastatic tumor in a portal venous phase volume. Higher-order texture information inside and neighboring a lesion might improve the performance. In addition, extensions to other lesions in other contrast phase volumes are also planned, such as a hepatocellular carcinoma in an early phase volume.

2) A validation study with a severely deformed liver is planned. Shapes of postoperative or postmortem livers [19] are essentially different from preoperative living livers. Segmentation of postoperative or postmortem livers is a challenging topic in this field.

3) Applications to other organs should be an interesting topic for future studies. There are several organs, such as the kidney and pancreas, whose shape could be strongly deformed by pathology, resulting in an outlier in a shape distribution.

4) Extensions to multi-organ segmentation might boost the segmentation performance. As in the discussion section, simultaneous segmentation of the liver and gallbladder might solve the false positive problem encountered in our study.

5) Validation of the proposed algorithm using CT volumes with normal livers and livers in which a portion lies outside the volume are interesting future research directions.

Acknowledgments

This study was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

References

[1] L. Soler, H. Delingette, G. Malandain, J. Montagnat, N. Ayache, C. Koehl, O. Dourthe, B. Malassagne, M. Smith, D. Mutter, and J. Marescaux, “Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery,” Comput. Aided Surg., vol.6, no.3, pp.131–142, May 2001.

[2] D. Kainmüller, T. Lange, and H. Lamecker, “Shape constrained automatic segmentation of the liver based on a heuristic intensity model,” Proc. MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge, 2007, pp.109–116, 2007.

[3] T. Okada, R. Shimada, Y. Sato, M. Hori, K. Yokota, M. Nakamoto, Y. Chen, and H. Nakamura, “Automated segmentation of the liver from 3D CT images using probabilistic atlas and multi-level statistical shape model,” Proc. MICCAI, 2007, pp.86–93, 2007.

[4] T. Heimann, B. Ginneken, M.A. Styner, Y. Arzhaeva, V. Aurich, C. Bauer, A. Beck, C. Becker, R. Beichel, G. Bekes, F. Bello, G. Binnig, H. Bischof, A. Bornik, P.M. Cashman, Y. Chi, A. Cordova, B.M. Dawant, M. Fidrich, J.D. Furst, D. Furukawa, L. Grenacher, J. Hornegger, D. Kainmüller, R.I. Kitney, H. Kobatake, H. Lamecker, T. Lange, J. Lee, B. Lennon, R. Li, S. Li, H.P. Meinzer, G. Nemeth, D.S. Raicu, A.M. Rau, E.M. Rikxoort, M. Rousson, L. Rusko, K.A. Saddi, G. Schmidt, D. Seghers, A. Shimizu, P. Slagmolen, E. Sorantin, G. Soza, R. Susonboon, J.M. Waite, A. Wimmer, and I. Wolf, “Comparison and evaluation of methods for liver segmentation from CT datasets,” IEEE Trans. Med. Imaging, vol.28, no.8, pp.1251–1265, May 2009.

[5] L. Ruskó, G. Bekes, and M. Fidrich, Automatic segmentation of the liver from multi- and single-phase contrast-enhanced CT images/Medical Image Analysis, vol.13, pp.871–882, July 2009.

[6] T. Heimann and H. Meinzer, “Statistical shape models for 3D medical image segmentation: A review,” Medical Image Analysis, vol.13, no.4, pp.543–563, May 2009.
[7] H. Park, P. Bland, and C. Meyer, “Construction of an abdominal probabilistic atlas and its application in segmentation,” IEEE Trans. Med. Imaging, vol.22, no.4, pp.483–492, April 2003.

[8] A. Shimizu, R. Ohno, T. Ikegami, H. Kobatake, S. Nawano, and D. Smutek, “Segmentation of multiple organs in non-contrast 3D abdominal CT images,” International Journal of Computer Assisted Radiology and Surgery, 2, pp.135–142, Nov. 2007.

[9] T. Okada, M.G. Linguraru, and Y. Sato, “Abdominal multi-organ segmentation of CT images based on hierarchical spatial modeling of organ interrelations,” Proc. Third International Workshop of Abdominal Imaging Computational and Clinical Applications, pp.173–180, 2011.

[10] M.G. Linguraru, J.A. Pura, V. Pamulapati, and R.M. Summers, “Statistical 4D graphs for multi-organ abdominal segmentation from multiphase CT,” Medical Image Analysis, vol.16, no.4, pp.904–914, May 2012.

[11] M. Suzuki, M.G. Linguraru, and K. Okada, “Multi-organ segmentation with missing organs in abdominal CT images,” Proc. MICCAI, 2012, pp.418–425, 2012.

[12] M. Leventon, O. Faugeras, W.E.L. Grimson, and W.M. Wells III, “Level set based segmentation with intensity and curvature priors,” Proc. IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, pp.4–11, 2000.

[13] R. Tibshirani, “Regression shrinkage and selection via the lasso,” J. the Royal Statistical Society, Series B, 58, pp.267–288, Jan. 1996.

[14] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma, “Robust face recognition via sparse representation,” IEEE Trans. Pattern Anal. Mach. Intell., vol.31, no.2, pp.210–227, Feb. 2009.

[15] A. Shimizu, K. Nakagomi, T. Narihira, H. Kobatake, S. Nawano, K. Shinozaki, K. Ishizu, and K. Togashi, “Automated segmentation of 3D CT images based on statistical atlas and graph cuts,” Proc. Workshop on Medical Computer Vision, 2010, pp.127–138, 2010.

[16] Y. Boykov and G. Funka-Lea, “Segmentation of dynamic N-D data sets via graph cuts using Markov model,” Proc. ICCV, 2001, I, pp.105–112, 2001.

[17] K. Nakagomi, A. Shimizu, H. Kobatake, M. Yakami, K. Fujimoto, and K. Togashi, “Multi-shape graph cuts with neighbor prior constraints and its application to lung segmentation from a chest CT volume,” Medical Image Analysis, vol.17, pp.62–77, Jan. 2013.

[18] H. Hontani, T. Matsuno, and Y. Sawada, “Robust nonrigid ICP using outlier-sparsity regularization,” Proc. CVPR, 2012, pp.1740–1747, 2012.

[19] A. Saito, A. Shimizu, H. Watanabe, S. Yamamoto, S. Nawano, and H. Kobatake, “Automated liver segmentation from a CT volume of a cadaver using a statistical shape model,” Proc. International Journal of Computer Assisted Radiology and Surgery, vol.8, supplement1, 2013, pp.S48–S49, 2013.

**Shun Umetsu** received the B.E. degrees from Tokyo University of Agriculture and Technology in 2012. He is now with Graduate School of Engineering of Tokyo University of Agriculture and Technology.

**Akinobu Shimizu** received the B.E., M.E. and Ph.D. degrees from Graduate School of Engineering, Nagoya University in 1989, 1991 and 1995, respectively. From 1994 to 1998, he was a research associate at Nagoya University. Since 1998, he has been an associate professor at Graduate School of Engineering of Tokyo University of Agriculture and Technology. His research activities are in the areas of image processing, pattern recognition and its application to medical image analysis. He is the member of the IEEE, the Information Processing Society in Japan, the Japanese Society for Medical and Biological Engineering and the Japanese Society of Medical Imaging Technology.

**Shigeru Nawano** graduated from the School of Medicine, Chiba University, in 1981. He became a research associate at Chiba University Hospital in 1982. He moved to the Diagnostic Radiology Division, National Cancer Center Hospital, in 1986, where he was chief of the Radiology Division, National Cancer Center Hospital East. Since 2007, he joined International University of Health and Welfare, where he is a professor. His research activities include computed radiography, mammography, MRI, CT and computer-aided diagnosis. He is the member of Japan Radiological Society, the Japanese Society of Nuclear Medicine, Radiological Society of North America and the Japanese Society of Medical Imaging Technology. He holds M.D./Ph.D. degrees.