HHS Public Access

Author manuscript

*Med Image Anal*. Author manuscript; available in PMC 2016 December 01.

Published in final edited form as:

*Med Image Anal*. 2015 December ; 26(1): 1–18. doi:10.1016/j.media.2015.06.009.

# Abdominal multi-organ segmentation from CT images using conditional shape–location and unsupervised intensity priors

**Toshiyuki Okada**[a], **Marius George Linguraru**[b,c], **Masatoshi Hori**[d], **Ronald M Summers**[e], **Noriyuki Tomiyama**[d], and **Yoshinobu Sato**[f]

Toshiyuki Okada: okada.t@md.tsukuba.ac.jp; Yoshinobu Sato: yoshi@is.naist.jp

[a]Department of Surgery, Faculty of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8575, Japan

[b]Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Medical Center, Washington DC 20010, USA

[c]Departments of Radiology and Pediatrics, School of Medicine and Health Sciences, George Washington University, Washington, DC 20037, USA

[d]Department of Radiology, Graduate School of Medicine Osaka University, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan

[e]National Institutes of Health, Clinical Center, Radiology and Imaging Sciences, 10 Center Drive Bethesda, MD 20892, USA

[f]Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma, Nara, 630-0192, Japan

## Abstract

This paper addresses the automated segmentation of multiple organs in upper abdominal computed tomography (CT) data. The aim of our study is to develop methods to effectively construct the conditional priors and use their prediction power for more accurate segmentation as well as easy adaptation to various imaging conditions in CT images, as observed in clinical practice. We propose a general framework of multi-organ segmentation which effectively incorporates interrelations among multiple organs and easily adapts to various imaging conditions without the need for supervised intensity information. The features of the framework are as follows: (1) A method for modeling conditional shape and location (shape–location) priors, which we call *prediction-based* priors, is developed to derive accurate priors specific to each subject, which enables the estimation of intensity priors without the need for supervised intensity information. (2) Organ correlation graph is introduced, which defines how the conditional priors are constructed and segmentation processes of multiple organs are executed. In our framework, predictor organs, whose segmentation is sufficiently accurate by using conventional single-organ segmentation methods, are pre-segmented, and the remaining organs are hierarchically segmented

Correspondence to: Toshiyuki Okada, okada.t@md.tsukuba.ac.jp; Yoshinobu Sato, yoshi@is.naist.jp.

using conditional shape–location priors. The proposed framework was evaluated through the segmentation of eight abdominal organs (liver, spleen, left and right kidneys, pancreas, gallbladder, aorta, and inferior vena cava) from 134 CT data from 86 patients obtained under six imaging conditions at two hospitals. The experimental results show the effectiveness of the proposed prediction-based priors and the applicability to various imaging conditions without the need for supervised intensity information. Average Dice coefficients for the liver, spleen, and kidneys were more than 92%, and were around 73% and 67% for the pancreas and gallbladder, respectively.

**Keywords**

subject-specific priors; shape prediction; partial least squares regression; computational anatomy; statistical shape model; probabilistic atlas

## 1. Introduction

Organ segmentation from medical images is an important preprocess for computer-aided diagnosis and therapy. In recent literature, several approaches to multiple organ segmentation from 3D medical images have been proposed. These approaches commonly utilize a number of radiological images with manual tracing of organs, called atlases, as training data, and can be classified as multi-atlas label fusion, machine learning, and statistical atlas approaches. In this paper, we address segmentation of multiple organs in upper-abdominal CT images. The upper abdominal organs are not only spatially but also functionally and physically interrelated, and thus their patient-specific shape anatomy has potentially wide applications in diagnostic and therapeutic assistance, including one-stop shop diagnosis, radiotherapy planning, patient-specific surgical simulation, and so on.

Multi-atlas label fusion is generally applicable to the segmentation of various organs through the preparation of a sufficient number of atlases (Aljabar et al., 2009; Isgum et al., 2009). It consists of intensity-based nonrigid registration between a target image and the original image in each of the atlases, and subsequent label fusion. Its applications to abdominal multi-organ segmentation have been addressed by establishing hierarchies that accommodate large inter-subject variability (Wolz et al., 2013; Chu et al., 2013). However, these works used the sum of squared differences (SSD) in intensity-based nonrigid registration, which means that the target images will be required to have similar contrast patterns to original computed tomography (CT) images in the atlases. Therefore, their performance for CT data acquired by different imaging conditions (ICs) may be limited. While multi-atlas label fusion can deal with contrast variations due to different ICs by using, for example, mutual information as a similarity measure instead of SSD, it may cause instability in nonrigid registration. Although a recent work addressed cross-modality multi-atlas segmentation (Iglesias et al., 2013), its usefulness was shown only to brain magnetic resonance (MR) images and it was not applied to abdominal organs.

Machine learning approaches, such as decision forests, also provide general segmentation frameworks (Seifert et al., 2009; Montillo et al., 2011; Glocker et al., 2012; Criminisi et al., 2013). While these methods were validated for various abdominal organs, such as the liver,

spleen, kidneys, and aorta, they were not evaluated for organs with large inter-subject variations in shape and location, such as the pancreas and gallbladder (GB). In addition, the dependence of the performance to ICs is not clear.

Statistical atlas approaches have been most commonly applied to abdominal organ segmentation. Explicit prior models constructed from atlases, such as the probabilistic atlas (PA) (Park et al., 2003; Zhou et al., 2006) and statistical shape models (SSMs) (Lamecker et al., 2004; Heimann and Meinzer, 2009), are used in these approaches. In an early study, Park et al. (2003) used PA for abdominal organ segmentation. Their original formulation assumed unknown ICs, and intensity priors were estimated from the target data to be segmented. However, manual specification of control points was required. In addition, organs with large inter-subject variations were not evaluated. Regarding single-organ segmentation, recent works on statistical atlas have proved that segmentation of the liver is stable and accurate even for unknown ICs, where intensity priors are constructed from target data only without the need for intensity information obtained from manually traced training data, which we call *supervised intensity information* (Zhou et al., 2006; Okada et al., 2008; Linguraru et al., 2012b,a). Freiman et al. (2010) addressed kidney segmentation, for which the estimation of both shape and intensity priors was performed by EM algorithm. In initialization for the estimation of the intensity prior, intensity-based nonrigid registration between target and CT data in atlases was required (although only target data was used without atlases after the initialization). That means that it may partly depend on supervised intensity information and cause a limitation in its generalization. A framework for adapting to contrast-enhancements was also proposed by Linguraru et al. (2010), but only for arterial and portal venous protocols in the liver and spleen.

There have been two main statistical atlas approaches to multi-organ segmentation. One approach is *joint modeling* of multiple structures. Linguraru et al. (2012b) addressed joint modeling of the liver, spleen, and kidneys, which are relatively stable in shape and location. Shimizu et al. (2007) studied 12 organs in the abdomen, but the segmentation accuracy required further improvement. Other works on joint modeling addressed multi-organ interrelations and multi-resolution hierarchy. Uzunbaş et al. (2010) and Yang et al. (2004) improved segmentation accuracy by modeling not only the shape of each structure but also the interrelation between structures. Cerrolaza et al. (2012, 2013) and Bagci et al. (2012) proposed multi-resolution modeling approaches, in which segmentation stability was improved by combining multiple structures into one structure at low resolution levels. In contrast, the other approach is *conditional modeling* of multiple structures. Conditional modeling of the target structure, given pre-segmented other structures, will be also regarded as one method for embedding multi-organ relations. This approach was shown to be effective for pancreas segmentation from CT images (Shimizu et al., 2010; Hammon et al., 2013). In this approach, the constraints provided by pre-segmented structures were incorporated in shape and location priors (hereafter, we call it *shape– location priors*). However, these existing methods were designed specific to the pancreas imaged with particular ICs.

Overall, general frameworks of multi-organ segmentation suitable for abdominal CT data with various ICs, as seen in routine clinical practice, have not been established so far.

Therefore, we propose a framework of conditional multi-organ segmentation which is adaptable to the large inter-subject variability and various ICs without supervised intensity information. The novel features of the framework are as follows:

1. A method for modeling conditional shape–location priors, which we call *prediction-based* priors, is developed. We utilize partial least squares (PLS) prediction of target organs from pre-segmented predictor organs to derive more accurate priors specific to each subject. This approach enables estimation of intensity priors from only target data and optionally a number of untraced CT data of the same IC as the target data, that is, from unsupervised intensity information.
2. An organ correlation graph (OCG) is introduced, which embeds the spatial correlations among organs inherent in human anatomy. The graph defines how the predictor organs in the prediction-based priors are related to the target organ, and how the conditional segmentation processes of multiple organs are executed. We describe a method for systematic construction and utilization of OCG.

Although our preliminary reports have been presented (Okada et al., 2012b,a, 2013), this paper extends them with respect to the following points: (1) Detailed formulations of single-organ segmentation, prediction-based prior construction, unsupervised intensity model estimation, and OCG construction are provided. (2) Descriptions of parameter optimizations involved in the methods are provided. (3) More detailed experimental results are shown including the use of additional CT data with another IC.

The organization of the paper is as follows. In Section 2, a method for single-organ segmentation as a combination of conventional methods is firstly described, which is extended to accommodate multiple organs by replacing shape–location priors. Then, our proposed methods are described, including prediction-based shape–location priors, unsupervised intensity priors, and multi-organ segmentation formulations using OCG. In Section 3, we evaluated the proposed methods through the segmentation of eight abdominal organs (liver, spleen, left kidney (L-kidney), right kidney (R-kidney), pancreas, GB, aorta, and inferior vena cava (IVC)) using 134 abdominal CT datasets obtained from 86 patients using six ICs at two hospitals. We discuss the work in Section 4 and conclude it in Section 5.

## 2. Methods

The fundamental idea is to incorporate inter-organ spatial correlations to attain stable and accurate segmentation of target organs. We first perform the segmentation of relatively stable organs in their position and shape (that is, predictor organs), and then hierarchically segment other less stable organs whose positions and shapes are expected to be well-predicted by pre-segmented predictor organs. In order to realize this concept, the liver is regarded as an anchor organ and assumed to be initially segmented. Given the segmented liver region, the prediction-based priors of the remaining organs are obtained by incorporating the constraints on shape and location provided by partial least squares (PLS) prediction from the liver region. The prediction-based prior of each organ in the reference space is constructed from training dataset of the predictor and target organs in the training

phase, and mapped to the patient space during segmentation processes using the segmented regions of the predictor organs. At the first stage of multi-organ segmentation, each of the remaining organs is segmented using the prediction-based shape–location prior. Once the segmented regions of these organs are available, stronger constraints are provided from the multiple organs in addition to the liver. The multi-organ interrelations optimizing the PLS prediction are embedded in the organ correlation graph (OCG), which is done in the training phase. At later stages, segmentation of each organ is performed using the improved prediction-based priors based on the multi-organ interrelations. As a result, we can expect improvement of segmentation accuracy as the stage is progressing. In the following subsections, details of method formulations are described. Commonly used acronyms and notations in the formulations are given in Tables 1 and 2, respectively.

### 2.1. Single-organ segmentation method

We utilize SSM and PA for shape–location priors (Okada et al., 2008), and an intensity model (IM) represented as a histogram of the intensity distribution for intensity prior. We assume that our target CT data cover upper abdominal organs. Figure 1(a) shows an example of a coronal view of our target CT data. First of all, CT data is spatially transformed to the normalized space defined by the liver dome top and the circumscribing planes of bone tissue regions (Okada et al, 2008). Figure 1(b) shows our definition of the abdominal normalized space. Given an abdominal CT data, the abdominal normalized space is determined using the axial reference plane tangential to the top of the liver dome and the bounding coronal and sagittal reference planes of the bone tissue regions. We adjust the xyz-translation and xy-scaling so that these reference planes are aligned among all the CT data. See Appendix A for an automated method for the above-described spatial normalization of upper abdominal CT data.

Let $I(\mathbf{x})$ be the target CT image data after spatial normalization. That is, $I(\mathbf{x})$ denotes the CT value (intensity) at 3D coordinates $\mathbf{x}$. Let $k$ be organ index, and $L_k(\mathbf{x})$ be the segmented label image of organ $k$, where $L_k(\mathbf{x}) = l \in \{0, 1\}$, in which $l = 1$ denotes object, $l = 0$ background. A set of shape–location and intensity priors for organ $k$ are given by

$$A_k = (P_k(\mathbf{x}), \mathbf{s}_k(\mathbf{b}), H_k(I)), \quad (1)$$

where $P_k(\mathbf{x})$ denotes a PA, $\mathbf{s}_k(\mathbf{b}) = \bar{\mathbf{s}}_k + \Phi_k\mathbf{b}$ a SSM, and $H_k(I) = \{H_{Ok}(I), H_{Bk}(I)\}$ an IM for organ $k$, in which $\mathbf{b} = (b_1\ b_2\ \cdots\ b_m)^t$ is the shape parameter vector, $\bar{\mathbf{s}}_k$ and $\Phi_k$ the average shape and the $m$ eigenvectors of the SSM of organ $k$, respectively, and $H_{Ok}(I)$ and $H_{Bk}(I)$ the normalized histograms of organ $k$ and its background regions, respectively. $H_{Ok}(I)$ and $H_{Bk}(I)$ are normalized so that their integral is 1. Given $A_k$, segmentation of organ $k$ from $I(\mathbf{x})$ is denoted by

$$L_k(\mathbf{x}) = S\ eg(I(\mathbf{x}); A_k), \quad (2)$$

where $S\ eg(\cdot)$ denotes the segmentation operation.

The single-organ segmentation method consists of three sub-modules as follows:

1. PA-based maximum a posteriori (MAP) segmentation
2. SSM-based refinement
3. Graph-cut refinement

Figure 2 shows the block diagram of the single-organ segmentation method. Details of the above three submodules of single organ segmentation are described in Appendix B.

## 2.2. Prediction-based shape–location priors

**2.2.1. Statistical shape prediction using PLS regression (PLSR)—**Let $\mathbf{s}_0$ and $\mathbf{s}_k$ ($k$ = 1,…, $n$) be the surface shape vectors of the target organ and its predictor organs, respectively, which are represented by the concatenation of the 3D coordinates on the surface of the organ, where $n$ is the number of organs used to predict a target (For example, the target organ is the pancreas while the predictor organs are the liver and spleen as explained in Figs. 3 and 4.). Given the shape vectors of $n$ + 1 organs of $N$ cases (subjects) $\{\mathbf{s}_{ki}\}$ ($k$ = 0,…, $n$, $i$ = 1,…, $N$) as training datasets, the regression model predicting the shape vector of the target organ from $n$ predictor organs is calculated using PLSR (Geladi and Kowalski, 1986). Unlike other standard least squares methods, PLSR works even when the predictor dataset is highly collinear. Yang et al. (2008) successfully applied PLSR to shape prediction of anatomical structures. Let $\tilde{\mathbf{s}}_i = (\mathbf{s}_{1i}^t \mathbf{s}_{2i}^t \dots \mathbf{s}_{ni}^t)^t$ be the predictor vector of case $i$. Given training datasets $T_\mathrm{p} = (\tilde{\mathbf{s}_1}\ \tilde{\mathbf{s}_2} \cdots \tilde{\mathbf{s}_N})^t$ and $T_\mathrm{t} = (\mathbf{s}_{01}\ \mathbf{s}_{02} \cdots \mathbf{s}_{0N})^t$, the PLSR model to predict target vector $\hat{\mathbf{s}}_0$ from predictor vector $\tilde{\mathbf{s}}$ is given by

$$\hat{\mathbf{s}}_0 = PLS\ R(\tilde{\mathbf{s}}; T_\mathrm{p}, T_\mathrm{t}), \quad (3)$$

where $\tilde{\mathbf{s}}$ is the predictor vector for a data not included in the training dataset. The details of the training of the prediction model $PLS\ R(\tilde{\mathbf{s}}; T_\mathrm{p}, T_t)$ are described in Appendix C.

**2.2.2. Construction of prediction-based shape–location priors—**Let $\{\mathbf{s}_k^*\}$($k$ = 1, …,$n$) be the segmented shape vectors of $n$ predictor organs and $\tilde{\mathbf{s}}^* = (\mathbf{s}_1^{*t} \mathbf{s}_2^{*t} \dots \mathbf{s}_n^{*t})^t$ be the predictor vector, where the shape vector with superscript *, e.g. $\mathbf{s}_k^*$, denotes the automatically segmented shape vector for target data, while the shape vector without superscript *, e.g. $\mathbf{s}_k$, denotes the shape vector for manual tracing of the training dataset. $\mathbf{s}_k^*$ is obtained by fitting the SSM of organ $k$ to the segmentation result $L_k(\mathbf{x})$. The prediction function is defined as

$$\mathbf{s}_0 = \hat{\mathbf{s}}_0^* + \mathbf{r}, \quad (4)$$

where

$$\hat{\mathbf{s}}_0^* = PLS\ R(\mathbf{s}^*; T_\mathrm{p}, T_\mathrm{t}) \quad (5)$$

and $\mathbf{r}$ denotes the residual after the prediction which is the difference between predicted and true shapes. In this work, the residual $\mathbf{r}$ is represented using PA and SSM as prediction-based shape–location priors.

Figure 3 shows the construction process for the prediction-based shape–location priors. To obtain PA and SSM of **r**, we perform the shape prediction of the target organ within the training dataset in a leave-one-out manner, and use the average shape of $T_t$ as the reference shape $\mathbf{s}_0'$. Let $\mathbf{r}_i = \mathbf{s}_{0i} - \hat{\mathbf{s}}_{0_i}$ be the residual of case *i* in the prediction whose model is trained without data of case *i*, where $\mathbf{s}_{0i}$ and $\hat{\mathbf{s}}_{0_i}$ are the true and predicted target organs of case *i*, respectively. The dense 3D deformation field from the original space to the reference space is determined by thin-plate spline interpolation using correspondences from $\hat{\mathbf{s}}_{0_i}$ to $\mathbf{s}_0'$. These correspondences are known because they are represented by the same SSM. The true shape $\mathbf{s}_{0i}$ in the original space is mapped to the reference space using the determined deformation field, and the mapped version $\mathbf{s}_{0i}'$ of the true shape $\mathbf{s}_{0i}$ is given by $\mathbf{s}_{0i}' = F(\mathbf{s}_{0i}; \hat{\mathbf{s}}_{0_i}, \mathbf{s}_0')$, where $F(\cdot\,; \mathbf{a}, \mathbf{b})$ denotes the mapping function defined by the deformation field determined by thin-plate spline interpolation from shape vectors **a** to **b**. The difference between predicted and true shapes in the original space can be represented in the reference space by this mapping.

Now, the residuals $\mathbf{r}_i'$ in the reference space are given by $\mathbf{r}_i' = \mathbf{s}_{0i}' - \mathbf{s}_0' (i = 1, \ldots, N)$. SSM and PA of $\{\mathbf{r}_i'\}$ are constructed in the reference space. Let $L_i'(\mathbf{x})$ be a label image which is voxelized from the surface of shape vector $\mathbf{s}_{0i}'$. PA $P_{\text{res}}'(\mathbf{x})$ and SSM $\mathbf{s}_{\text{res}}'(\mathbf{b}_{\text{res}})$ in the reference space are given by

$$P_{\text{res}}'(\mathbf{x}) = \frac{1}{N}\sum_{i=1}^{N} L_i'(\mathbf{x}) \quad (6)$$

$$\mathbf{s}_{\text{res}}'(\mathbf{b}_{\text{res}}) = \mathbf{s}_0' + \Phi_{\text{res}}\mathbf{b}_{\text{res}}, \quad (7)$$

where $\Phi_{\text{res}}$ denotes the principal components of the residual and $\mathbf{b}_{\text{res}}$ their coefficients. Because $L_i'(\mathbf{x})$ is the label image of the predicted organ shape of each case and regarded as the composition of reference shape $\mathbf{s}_0'$ and residual $\mathbf{r}_i'$, smaller residuals result in less ambiguous PA.

Figure 4 shows the generation process of subject-specific shape–location priors. When prediction-based priors are utilized for segmentation of a target data, their subject-specific versions are obtained as follows. By using the mapping through the deformation field from the reference space to the target data space, $F(\cdot; \mathbf{s}_0', \hat{\mathbf{s}}_0^*)$, the prediction-based priors in the target data space are given by

$$P_{\text{res}}(\mathbf{x}) = P_{\text{res}}'\left(F^{-1}(\mathbf{x}; \mathbf{s}_0', \hat{\mathbf{s}}_0^*)\right) \quad (8)$$

$$\mathbf{s}_{\text{res}}(\mathbf{b}_{\text{res}}) = F\left(\mathbf{s}_{\text{res}}'(\mathbf{b}_{\text{res}}); \mathbf{s}_0', \hat{\mathbf{s}}_0^*\right), \quad (9)$$

where $F^{-1}(\cdot)$ indicates the inverse mapping of $F(\cdot)$. In Eq. (8), it is necessary to use the inverse mapping $F^{-1}(\cdot)$ in order to obtain the coordinates in the reference space corresponding to **x** since **x** is the coordinates in the target data space.

Given the set of segmented shape vectors of the predictor organs, $S^*=\{\mathbf{s}_k^*\}$, the set of shape–location and intensity priors is given by

$$A(S^*)=(P_{\mathrm{res}}(\mathbf{x}),\mathbf{s}_{\mathrm{res}}(\mathbf{b}_{\mathrm{res}}),H(I)). \quad (10)$$

Segmentation is performed using the method described in Section 2.1 by replacing $A_k$ with $A_k(S^*)$ in Eq. (2). That is, the generated subject-specific versions of the prediction-based PA and SSM are utilized for segmentation by replacing the conventional PA and SSM in Fig. 2 with them.

Figure 5 shows prediction-based PAs for the R-kidney, GB, and pancreas (in comparison with conventional ones), which demonstrates that the prediction-based PA varies depending on the shapes of predictor organs. In Fig. 5, the predictor organs of the R-kidney, the GB, and the pancreas are the liver, the liver, and the liver and spleen, respectively. These relationships can be represented using directed graphs. The organs are represented as nodes of a graph, and predictor-target relations as directed edges. Using this graph representation, Figs. 5(a), (b), and (c) are represented as Figs. 5(d), (e), and (f), respectively. In Section 2.4, we formulate this graph representation as OCG.

### 2.3. Estimating intensity model (IM) from untraced CT data

Segmentation is performed using the methods described in Section 2.1 by replacing $H_{Ok}(I)$ and $H_{Bk}(I)$ with the IMs described below. Three types of IMs are constructed as follows:

1. Target-data specific IM (TD-IM): Only one CT data, which is the segmentation target, is used. Preparation of any training data is unnecessary for intensity modeling. Therefore, this model is immediately adaptable to any unknown imaging condition (IC) as far as shape–location priors are available.
2. Imaging-condition specific IM from *untraced* CT data (unsupervised IC-IM): In addition to the target CT data to be segmented, a sufficient number of CT data from different patients acquired by the same IC (protocol) as the target data are used. Laborious manual tracings are not required for intensity modeling of different ICs, and just original CT data are required for training. These training data are easy to collect from clinical image repositories.
3. Imaging-condition specific IM from *traced* CT data (supervised IC-IM): A sufficient number of CT data, which were acquired using the same IC as the target data and associated with manual tracings of the target organ, are used (Linguraru et al., 2012b; Okada et al., 2012a). This model requires a sufficient number of manual tracings for each different IC.

**2.3.1. TD-IM—**TD-IM is constructed as follows. Let $I_0(\mathbf{x})$ be the target CT data and $P(\mathbf{x})$ the PA. Figure 6 shows the construction processes for TD-IM. First, $P(\mathbf{x})$ is binarized using

the threshold value determined by the 1st percentile in the probability distribution of $P(\mathbf{x})$. Let the binarized region be $R_{PA}(\mathbf{x})$, which is used as the initial region for intensity modeling. Because the initial region $R_{PA}(\mathbf{x})$ should be highly likely to belong to the target organ region, only the 1st percentile is used. The initial histogram $H_{O0}$ is obtained from the intensity distribution of $I_0(\mathbf{x})$ within $R_{PA}(\mathbf{x})$, and $H_{O0}$ is normalized so that its integral is 1. The refined region $R(\mathbf{x})$ is then obtained by applying the MAP segmentation (described in Appendix B.1) to $I_0(\mathbf{x})$ using $P(\mathbf{x})$, $H_{O0}$, and $H_{B0}$, where $H_{B0}$ is the histogram for background, which is obtained from $I_0(\mathbf{x})$ within the region where the probability value of $P(\mathbf{x})$ is less than 5%. The processes obtaining $H_{O0}$ with $R(\mathbf{x})$ and updating $R(\mathbf{x})$ by MAP segmentation are iterated three times. Finally, TD-IM $H_O$ and $H_B$ are obtained as $H_{O0}$ and $H_{B0}$, respectively, after the iterations. The prediction-based PA can be used as $P(\mathbf{x})$ if the pre-segmented organ regions are available.

**2.3.2. IC specific IM from untraced CT data (unsupervised IC-IM)—**The unsupervised IC-IM is constructed as follows. Let $I = \{I_i\}(i = 1,\ldots, N)$ be $N$ additional CT data from different patients acquired using the same IC as $I_0$ with the assumption that manual tracings are unavailable for these CT data. TD-IM is obtained for each $I_i(i = 0,1,\ldots, N)$. Note that segmentation target data $I_0$ is also involved in intensity modeling. Let $H_O = \{H_{Oi}\}$ and $H_B = \{H_{Bi}\}$ $(i = 0,1, \ldots, N)$ be the histograms for object and background, respectively. Unsupervised IC-IM $\bar{H} = \{\bar{H}_O, \bar{H}_B\}$ are calculated as

$$\bar{H}_O = \frac{1}{Z_O}\sum_{i=0}^{N} H_{Oi} \quad (11)$$

$$\bar{H}_B = \frac{1}{Z_B}\sum_{i=0}^{N} H_{Bi}, \quad (12)$$

where $Z_O$ and $Z_B$ are normalization factors so that integral is 1.

**2.3.3. IC specific IM from traced CT data (supervised IC-IM)—**The supervised IC-IM is a conventional IM based on manual tracings in the training data, and is constructed as follows. Given $I = \{I_i\}$, which are acquired using the same IC as $I_0$, with their corresponding manual tracings $L = \{L_i\}$ $(i = 1,\ldots, N)$, the supervised IC-IM $\bar{H}_L = \{\bar{H}_{O,L}, \bar{H}_{B,L}\}$ are given by

$$\bar{H}_{O,L} = \frac{1}{Z_O}\sum_{i=1}^{N} H(I; I_i, L_i) \quad (13)$$

$$\bar{H}_{B,L} = \frac{1}{Z_B}\sum_{i=1}^{N} H(I; I_i, L_i^c), \quad (14)$$

where $H(I; I_i, L_i)$ is the histograms of intensity distributions of $I_i$ within $L_i$ and $L_i^c$ the complement of $L_i$.

### 2.4. Multi-organ segmentation based on organ correlation graph (OCG)

In this section, we formulate the predictor-target relations among the abdominal organs described in Section 2.2, by which the prediction-based priors are constructed, as the OCG. The OCG is defined as a set of nodes and directed edges. Each node corresponds to an organ and each directed edge denotes the correlation between one organ and another. Examples of simple OCGs are shown in Figs. 5(d), (e), and (f). The basic assumption is that automated segmentation of organs with out-edges in the OCG is sufficiently stable for use of these organs as predictors, while the organs with in-edges are effectively predicted by organs with out-edges. In the following, we describe how the predictor-target relations among the eight organs which we consider, that is, the liver, spleen, L-kidney, R-kidney, GB, aorta, IVC, and pancreas, are embedded into one OCG and how the conditional segmentation procedures are derived from the OCG.

Among the eight organs, segmentation of the liver has been studied intensively, and is now sufficiently accurate and stable by itself (Bagci et al., 2012; Linguraru et al., 2012b; Wolz et al., 2013; Montillo et al., 2011; Suzuki et al., 2012a; Seifert et al., 2009; Ling et al, 2008; Kohlberger et al., 2009; Okada et al., 2008; Zhang et al, 2010; Kainmüller et al., 2007; Massoptier and Casciaro, 2007; Shimizu et al, 2007; Lamecker et al., 2004; Park et al, 2003; Camara et al., 2004; Glocker et al., 2012). Segmentation of the spleen and kidneys has also been studied by a sufficient number of works, which showed sufficient accuracy (Bagci et al., 2012; Linguraru et al., 2012b; Wolz et al., 2013; Suzuki et al., 2012a; Montillo et al, 2011; Seifert et al., 2009; Shimizu et al, 2007; Glocker et al, 2012; Cuingnet et al., 2012; Freiman et al, 2010; Khalifa et al., 2011; Chen et al., 2012; Park et al, 2003; Camara et al, 2004). Finally, segmentation of other organs has not been well-established or shown to be accurate enough (Hammon et al., 2013; Wolz et al, 2013; Suzuki et al., 2012a; Shimizu et al., 2010, 2007; Zhou et al., 2010; Yoshida et al., 2011; Kurkure et al., 2008; Duquette et al., 2012; Montillo et al., 2011; Sanchez-Castro et al, 2010). Based on the above observations, we classify the eight organs or nodes into the following three types:

Type 1: Only out-edges are defined. Because no in-edges are defined, the segmentation of this type of organs is performed unconditionally. The liver is assigned to this type. The liver is particularly large compared with other abdominal organs. Especially in the prediction of small organs, only the local shape of the liver may be closely related to them. Thus, for effective prediction, we divide the liver surface into four sub-shapes. Figure 7 shows the four sub-shapes of the average liver, and three different patient liver shapes. The division method based on canonical correlation analysis developed by Yokota et al. (2013) was used for dividing the liver surface. The division method is detailed in Appendix D. The set of (sub-)organs of Type 1 is denoted as $V_1 = \{liver_1, liver_2, liver_3, liver_4\}$.

Type 2: Both in- and out-edges are defined. Analysis of these organs is performed under the condition that the segmentation is completed for at least one node connected by an in-edge. In addition, the organ can be a predictor for nodes connected by out-edges. The set of organs of Type 2 is denoted as $V_2 = \{spleen, R\text{-}kidney, L\text{-}kidney\}$.

Type 3: Only in-edges are defined. Segmentation of these organs is performed conditionally like for $V_2$. This type cannot be a predictor. The set of organs of Type 3 is denoted as $V_3$ = {*pancreas*, *GB*, *aorta*, *IVC*}.

Table 3 summarizes the organ classification described above.

Given the constraints on the above three types of nodes, edge connections representing organ correlations are automatically defined based on shape predictability by PLSR described in Section 2.2.1. Let $V_p = V_1 \cup V_2$ be a set of nodes for predictors, $V_t = V_2 \cup V_3$ be a set of nodes for target organs, and $V_p(v_t)$ be the set of predictors of target organ $v_t \in V_t$. $V_p(v_t)$ is determined by selecting the set of predictors which minimize the prediction error for $v_t$ among all possible combinations of nodes in $V_p$. Let $V_{\text{all}} = \cup_{k \in \{1,\ldots,|V_p|\}} \begin{pmatrix} V_p \\ k \end{pmatrix}$ be all possible combinations of nodes in $V_p$, where $\begin{pmatrix} V_p \\ k \end{pmatrix}$ denotes the set of all $k$-combinations in $V_p$. Let $T_V$ and $T_{v_t}$ be the training data of organ shapes of $V \in V_{\text{all}}$ and $v_t \in V_t$, respectively, and $\varepsilon(V; v_t)$ be the prediction error of $v_t$ using the prediction model *PLS R*($\mathbf{s}; \tilde{T}_V, T_{v_t}$) (see Appendix C). Then the set of predictor nodes $V_p(v_t)$ for $v_t$ is defined as

$$V_p(v_t) = \arg\min_{V \in V_{\text{all}}} \varepsilon(V; v_t). \quad (15)$$

The OCG is defined as *OCG* =< *V*, *E* > with

$$V = V_1 \cup V_2 \cup V_3 \quad (16)$$

$$E = \bigcup_{v_t \in V_t} E(v_t), \quad (17)$$

where

$$E(v_t) = \{(v, v_t) | v \in V_p(v_t)\} \quad (18)$$

denotes a set of directed edges toward $v_t$. The relations between each target and its predictors in PLSR that minimizes the prediction error are associated with the edges of OCG. Figure 8(a) shows the OCG of the eight organs constructed from our training data.

The synchronized processes of multi-stage updating in multi-organ segmentation based on OCG are formulated in Algorithm 1. Let $v$ be a node (that is, an organ) to be segmented, $V_{\text{out}}(v)$ be a set of nodes with the edge from $v$, and $V_{\text{in}}(v)$ be a set of nodes with the edge toward $v$. Figure 8(b) shows examples of $V_{\text{out}}(v)$ and $V_{\text{in}}(v)$. Here, we assume that one or more nodes in $V_{\text{in}}(v)$ have already been segmented at the current stage. Segmentation of node $v$ is performed using prediction-based shape–location priors constructed from segmented nodes in $V_{\text{in}}(v)$. Note that the segmentation result of a target organ is not directly used as the initial region in its segmentation processes at the next stage. It is used only for

updating prediction-based priors at the next stage. After segmentation of $v$, all nodes in $V_{\text{out}}(v)$ become ready for segmentation at the next stage. Let $j$ be the index of the segmentation stage. Node $v$ is ready for segmentation at stage $j$ after at least one node in $V_{\text{in}}(v)$ has been segmented during stage $j-1$. If $V_{\text{in}}(v)$ is empty, e.g. $V_1$ (the liver in this study) does not have in-edges, segmentation of $v$ is completed using conventional shape–location priors without predictor organs. That is, segmentation of organ $v$ in $V_1$ can be performed at the initial stage (Stage 0) of the segmentation processes. Figure 9 shows the states of OCG at different stages in multi-organ segmentation processes. Let $j_{\max}$ be the maximum number of stages of the segmentation processes, $V^j_{\text{target}}$ be a set of nodes ready for segmentation at Stage $j$, and $V_{\text{segmented}}$ be a set of segmented nodes. Multi-stage updating of the segmentation processes is repeated at least until $V^j_{\text{target}}$ is unchanged. Figure 10 shows examples of prediction-based PAs at different stages in multi-organ segmentation processes for three cases.

**Algorithm 1** Multi-organ segmentation based on OCG

**Input:** Target image $I(\mathbf{x})$, $OCG =< V, E >$, number of stages $j_{\max}$

**Output:** Label images of each organ $\{L_v(\mathbf{x}) \mid v \in V\}$

1: $V_{\text{segmented}} \leftarrow \emptyset$

2: $V^0_{\text{target}} \leftarrow V_1$

3: **for** $j = 0$ **to** $j_{\max}$ **do**

4: **for all** $v \in V^j_{\text{target}}$ **do**

5: $S^* \leftarrow \{\mathbf{s}^*_k | k \in V_{\text{in}}(v) \cap V_{\text{segmented}}\}$

6: $L_v(\mathbf{x}) = S\,eg(I(\mathbf{x}); A(S^*))$

7: $\mathbf{s}^*_v \leftarrow$ shape vector calculated by fitting SSM to $L_v(\mathbf{x})$

8: **end for**

9: $V^{j+1}_{\text{target}} \leftarrow \cup_{v \in V^j_{\text{target}}} V_{\text{out}}(v)$

10: $V_{\text{segmented}} \leftarrow V_{\text{segmented}} \cup V^j_{\text{target}}$

11: **end for**

## 3. Experimental Results

### 3.1. Dataset

We tested the proposed methods using 134 abdominal CT data from 86 patients (cases) obtained under six different ICs at two hospitals: Osaka University Hospital, Japan (OUH) and National Institutes of Health, USA (NIH). Table 4 shows the details of six ICs. Figure 11 shows examples of CT data of six ICs. CT data obtained at OUH were two-phase CT data of 49 cases. The contrast phases used in image acquisition were early and late arterial phases (which were Datasets A and B, respectively). Data for 10 cases (Datasets A-1 and

B-1) were acquired using a 4-slice Light Speed QX/i (GE Healthcare) scanner while data for the other 39 cases (Datasets A-2 and B-2) were acquired using a 64-slice Light Speed VCT (GE Healthcare) scanner, among which only the late arterial phase was available for one case. CT data obtained at NIH included 25 cases from the portal venous phase (Dataset C) and 12 cases from the non-contrast CT data (Dataset D). These data were acquired using a 4-slice HighSpeed QX/I (GE Healthcare), an 8-slice Light-Speed Ultra (GE Healthcare), a 64-slice Brilliance 64 (Philips Healthcare), a 16-slice Mx8000 IDT 16 (Philips Healthcare), and a dual source 256-slice Definition Flash (Siemens Healthcare). Eight organs (the liver, spleen, L-kidney, R-kidney, GB, aorta, IVC, and pancreas) were manually segmented from all data under the supervision of experienced radiologists.

Because the liver was segmented using only the conventional single-organ segmentation method, the descriptions of results focuses on the other seven organs.

### 3.2. Experimental method

Two-fold cross validation was performed. The datasets from each IC was randomly separated into two datasets, that is, training and test datasets. Shape–location priors were constructed using the training dataset from all ICs. Note the unsupervised IC-IM was estimated using the test dataset from each IC, and thus was not based on any intensity information combined with manual tracings.

Segmentation using conventional shape–location priors and/or supervised IC-IM were also performed for comparison purpose. We did not use the training dataset but used the test dataset for constructing the supervised IC-IM, which was also used for the unsupervised IC-IM, because the difference in performance due to the method of IM construction will not be properly evaluated if the different datasets were used for constructing IM. Optimization of segmentation parameters was performed using only the training dataset. Note that the test dataset was totally separated from parameter optimization. The details of the parameter optimization are described in the next sub-section. Initially, segmentation of the liver was performed using conventional shape–location priors and TD-IM (Okada et al., 2008). Subsequent segmentations of other organs were performed using three IMs (TD-IM, unsupervised IC-IM, and supervised IC-IM). The maximum number of stages of multi-organ segmentation was $j_{max} = 4$. Two-fold cross validation was performed for five different patterns of two-fold separations of the dataset, which were randomly generated, and the results were averaged. Dice Coefficient (DC) (Dice, 1945), Jaccard Index (JI) (Jaccard, 1901), and Average Symmetric Surface Distance (ASD) (Heimann et al., 2009) were used for quantitative evaluation.

### 3.3. Parameter optimization

Parameter optimization was performed using leave-one-out cross validation only within a subset of the training dataset for each organ. In order to avoid over-tuning in the parameter optimization, we selected the training data so as to exclude CT data of the same IC as the test data. Because CT data acquired from OUH did not include the same ICs as those from NIH, and vice versa, parameter optimization was performed for one institution using training dataset from the other institution. Segmentation using conventional shape–location

priors (described in Section 2.1) was performed, and the set of parameter values with which the best segmentation accuracy achieved were selected as follows (Details of the algorithms including these parameters are described in Appendix B.):

- Radius $\rho$ of the structuring element in opening and closing at the final step of PA-based segmentation: The optimal value of $\rho$ was selected so as to maximize the PA-based segmentation accuracy.
- Weight parameter $\lambda$ in SSM-based refinement: The optimal value $\lambda$ was selected so as to maximize the SSM-based refinement accuracy, given PA-based segmentation results using the optimal $\rho$.
- Threshold values $p_O$ and $p_B$ to determine regions of hard constraint in graph-cut refinement: Optimal values of $p_O$ and $p_B$ were selected so as to maximize the graph-cut refinement accuracy, given the segmentation results obtained using PA-based segmentation with the optimal $\rho$ followed by SSM-based refinement with the optimal $\lambda$.

We used Jaccard index (JI) as the segmentation accuracy measure in the above procedures.

During cross-validations for parameter optimization, the most frequently selected values for radius $\rho$ in PA-based segmentation, weight parameter $\lambda$ in SSM-based refinement, values $p_O$ and $p_B$ in graph-cut refinement were $\rho = 1$, $\lambda = 0.5$, $p_O = 0.1$, and $p_B = 2$, respectively.

### 3.4. Results

Figure 12 summarizes the results of comparisons of segmentation accuracy between methods using the proposed prediction-based and conventional shape–location priors. In these comparisons, the unsupervised IC-IM was used for intensity modeling. As shown in Fig. 12, significant accuracy improvement by using the proposed priors was observed for all seven organs in comparison with the conventional priors. Significance was determined using paired T-test with significance levels of 0.01 and 0.05. For most organs and evaluation measures, significant improvements in accuracy with a significant level of 0.01 were observed, excepting that the R-kidney in DC and the spleen in ASD showed a significance level of 0.05.

Figure 13 shows the accuracy improvement of each segmentation submodule, which corresponds to each block of Fig. 2. The shape-location and intensity priors were the same as used in the proposed method in Fig. 12. The average shape of the prediction-based SSM (conventional SSM only for the liver) was used to evaluate the accuracy of the spatial normalization. The segmentation accuracy was monotonically increased as subsequent submodules were combined. Significant accuracy improvements were observed among all submodules for all organs ($p < 0.01$).

Figure 14 shows the effects of the multi-stage updating in the proposed multi-organ segmentation based on OCG, where average DC at each stage was plotted for each organ. Our multistage updating method does not guarantee the convergence. To confirm the behavior of multi-stage updating, we performed experiments up to Stage 8. Because similar results were obtained in Stage 4 or later, the results up to Stage 4 are shown in Fig. 14. The

accuracies at Stage 3 or later (using all available predictor organs) were significantly improved in comparison with Stage 1 (using only the liver as the predictor organ) for the L-kidney ($p < 0.05$). The accuracies at Stage 2 or later were significantly improved in comparison with Stage 1 for the pancreas ($p < 0.05$) and aorta ($p < 0.01$). For the IVC, significant improvement was observed between Stage 1 and Stage 2 ($p < 0.05$). For the spleen, however, no significant improvement in accuracy was observed (the range of $p$-value was from 0.084 to 0.337 among Stage 1 and later stages). For the R-kidney and GB, segmentation accuracy was not changed through Stages 1 to 8 because only the liver was used as the predictor organ. The accuracy variations from Stages 3 to 8 were calculated as the difference of the maximum and minimum DCs. The maximum variation was 0.32%, which was observed in L-kidney.

The proposed method assumes that the liver is initially segmented. Hence, failure of liver segmentation may affect subsequent predictions and segmentation of other organs. In order to examine the effects of segmentation accuracy of the liver on other organs, the datasets used in the experiments were divided into two groups based on the accuracy of liver segmentation, that is, those below and above the lower quartile of DC of liver segmentation. Table 5 shows average DCs of the seven organs for the two groups. Although significant difference was not observed in all seven organs between the two groups, the average DCs of R-kidney, GB, and IVC, which are adjacent to the liver, decreased by 11.0, 19.3, and 9.7 percentage points, respectively, in the group below the lower quartile. For other organs, the accuracy decrease was 0 to 5 percentage points.

In order to evaluate the accuracy of the unsupervised IC-IM estimation rather than final segmentation accuracy, similarities between unsupervised and supervised IC-IMs were calculated using normalized cross correlation (NCC). Figure 15 shows average NCCs of the six ICs for each organ. The average NCCs were improved and greater than 0.9 for all organs except for the GB by using the prediction-based PA, while the improvement was prominent especially in the GB.

Figure 16 summarizes segmentation accuracies for the three IMs using the proposed prediction-based shape–location priors. We first focus on the results in DC in this paragraph. The average DC of the liver, in which only TD-IM was used, was 94.1%. Average DCs of the spleen, R-kidney, and L-kidney were more than 90% for all three IMs. For the spleen and R-kidney, statistical significance was not observed among three IMs, although the standard deviation (SD) of TD-IM in the R-kidney was larger than the other two IMs. For the L-kidney, statistical significance was observed between TD-IM and supervised IC-IM ($p < 0.05$). Nevertheless, the decrease in accuracy was not obvious. The average DC of the pancreas was around 73% for all three IMs and statistical significance was not observed among them. However, the SD of TD-IM was slightly larger than other IMs. For the GB, the unsupervised IC-IM was significantly more accurate than TD-IM ($p < 0.01$). Furthermore, the supervised IC-IM was significantly more accurate than unsupervised IC-IM ($p < 0.01$). As shown in Fig. 15, the similarity of the estimated IMs between the supervised and unsupervised IC-IMs was lower in the GB than in other organs. Note that the GB is small and has large inter-subject location variability, which makes it difficult to ensure a sufficient amount of high probability area even in the prediction-based PA. For the aorta, TD-IM was

significantly less accurate than unsupervised and supervised IC-IMs ($p < 0.05$ and $p < 0.01$, respectively), and the SD of TD-IM was larger than other IMs. However, statistical significance was not observed between unsupervised and supervised IC-IMs. For the IVC, statistical significance was not observed among three IMs.

Similar results were obtained for evaluations by JI and ASD in Fig. 16. For the kidneys, accuracy in the TD-IM was significantly decreased in comparison with unsupervised and supervised IC-IMs in ASD ($p < 0.05$). However, statistical significance was not observed between unsupervised and supervised IC-IMs in ASD. For the aorta, supervised IC-IM was significantly more accurate than TD-IM and unsupervised IC-IM in ASD ($p < 0.05$). The results in Fig. 16 are summarized as follows. Unsupervised IC-IM was regarded as mostly comparable to supervised IC-IM excepting the GB. TD-IM was regarded as less accurate in the kidneys, aorta, and GB, but the accuracy decrease in the kidneys and aorta was not large.

Figure 17 shows illustrative three examples of segmentation results, which effectively demonstrate usefulness of the proposed methods. DC of each organ in these three cases is plotted in Fig. 12. In Fig. 17, the failed regions by the conventional prior are indicated by white arrows. In Fig. 17(a), the pancreas and IVC were better segmented by using the proposed prediction-based prior with different IMs. In the pancreas, the head of the pancreas was better segmented by using TD-IM than the other IMs (the improved part in TD-IM is shown by a black arrow). In Fig. 17(b), the pancreas, GB, aorta, and IVC were missed by the conventional prior while the prediction-based priors were effective. The segmentation result of the R-kidney was leaked into the liver by the conventional prior. In this case, the segmented liver region was leaked into the R-kidney, GB, and IVC. However, these organs were successfully segmented by using the prediction-based priors even when the leaked liver was used for the predictor organ. The regions of the liver leakage are indicated by white arrow heads in Fig. 17(b). In Fig. 17(c), the spleen, L-kidney, pancreas, GB, aorta, and IVC were better segmented by using the proposed prediction-based priors. For the GB, unsupervised IC-IM was worse than TD-IM and supervised IC-IM by the prediction-based prior (the failed part in unsupervised IC-IM is shown by a black arrow). In TD-IM, the pancreas was largely missed by the prediction-based prior while the other IMs were effective. As shown in Fig. 16, the SD of TD-IM for the pancreas was larger than other IMs. Figs. 17(a) and (c) indicate typical results which caused higher SD in TD-IM. In the illustrated three cases, the unsupervised IM combined with the prediction-based priors was superior or equivalent to the conventional priors for the segmentation of all seven organs.

## 4. Discussion

### 4.1. Summary of the proposed method

In this paper, we have described a general framework of multi-organ segmentation which effectively incorporates interrelations among multiple organs. The main contribution of this paper is that a method for modeling prediction-based conditional shape–location priors based on OCG is developed. In our framework, predictor organs, such as the liver, whose segmentation is sufficiently accurate by using conventional single-organ segmentation methods, are pre-segmented, and the remaining organs are hierarchically segmented using conditional shape–location priors. The prediction-based priors enabled sufficiently accurate

estimation of intensity priors without using supervised intensity information. As shown in Fig. 12, the prediction-based priors were a significantly more effective way to estimate intensity priors and use them for segmentation without supervised intensity information in comparison to conventional priors.

As shown in Fig. 16, segmentation accuracy using TD-IM was slightly but significantly less accurate than supervised IC-IM in DC for the L-kidney and aorta. However, it became comparable by using unsupervised IC-IM for these organs. For the GB, the proposed IMs (TD-IM and unsupervised IC-IM) were significantly less accurate than supervised IC-IM, but unsupervised IC-IM was significantly more accurate than TD-IM. From these results that significant improvements were observed in GB (in DC, JI, and ASD), kidneys (in ASD), and aorta (in DC and JI) for unsupervised IC-IM compared to TD-IM, unsupervised IC-IM is considered to be a reasonable way of intensity modeling from a number of untraced CT data. In contrast, significant difference was not observed in the spleen, pancreas, and IVC among the three intensity models (in all of DC, JI, and ASD). Intensity variations are considered to be fairly large in these organs even within the same IC. Therefore, unsupervised and supervised IC-IMs were not so effective in these organs and TD-IM was comparable to them.

### 4.2. Effectiveness of the proposed method

As shown in Fig. 12, the largest improvement by using the proposed prediction-based priors was observed in the GB. Although the GB is a small organ and its inter-subject variation in location is relatively large, it is in contact with the liver. Therefore, the effect of the prediction of the GB from the segmented liver region was substantial. The improvement of unsupervised IC-IM estimation (Fig. 15) by the proposed prediction-based priors was particularly prominent for the GB.

As shown in Fig. 13, significant accuracy improvements were observed among all submodules for all organs. Among the submodules, improvements due to SSM-based refinement look relatively small (although statistical significance was observed). Therefore, additional experiments were conducted, in which graph-cut refinement was performed directly after PA-based MAP segmentation (without SSM-based refinement). The experimental results showed that the segmentation accuracy without SSM-based refinement was significantly decreased in comparison with that with SSM-based refinement ($p < 0.01$ for all organs). Therefore, all segmentation submodules were effective.

As shown in Table 5, the segmentation accuracy of the liver affected the accuracy of organs adjacent to the liver. In the experiments, the segmented liver region was often leaked into the GB and IVC. However, segmentation accuracy using prediction-based shape–location priors was improved in comparison with conventional ones even when this leakage occurred. In Fig. 17(b), the R-kidney, GB, and IVC were successfully segmented by using the proposed prediction-based priors although the segmented liver was leaked into them. Even when SSM fitting was performed to mis-segmented liver shape due to leakage in order to generate subject-specific versions of the prediction-based shape–location priors, sufficiently well-approximated liver shape for the predictor organ was recovered since SSM of the liver has specificity for the liver shape. Thus, appropriate predicted shapes of target

organs and their prediction-based PAs were obtained (GB and IVC in Fig. 17(b)). Subsequent SSM-based and graph-cut refinements also contributed to the recovery from the failure in PA-based MAP segmentation (R-kidney in Fig. 17(b)). The behaviors of the proposed method as described above were observed in a few cases. The liver segmentation method used in this work was based on our previous work (Okada et al., 2008). It can be replaced by recently developed improved methods (Tomoshige et al., 2014; Umetsu et al., 2014), which showed good performance even for pathological and largely deformed livers. Therefore, improvement of liver segmentation accuracy would be expected.

Multi-stage updating in the proposed method was demonstrated to be useful (Fig. 14). Significant accuracy improvement in most of organs interrelated with multiple predictor organs was attained at Stage 4 (using multiple organs as the predictors) in comparison with Stage 1 (using only the liver), as described in the second paragraph of Section 3.4. Only in the spleen, significant improvement was not observed (although average accuracy increased). In addition, the accuracy decreased from Stage 3 to Stage 4 in the spleen. Our multi-stage updating method does not guarantee the convergence. Therefore, instability may be caused during the iteration. However, accuracy variations were within 0.32% in DC after Stage 3 and did not affect the results of significance test. Regarding the number of iterations, we consider that additional one or two iterations will be sufficient after the OCG reaches a steady state shown in Fig. 9(d). Accuracy variations during the iteration were mainly observed in the spleen and left kidney. Its potential cause is that these organs are the predictor and target organs each other (as shown in Fig. 9) and oscillation phenomena occurred in a few cases. When the OCG is constructed, the constraint to avoid such relations may be effective.

Among the eight organs, the GB was the organ for which supervised intensity information was significantly useful. However, if CT cholangiography data are given, in which the fluid in the GB is contrast-enhanced, our method may be easily adaptable while the conventional method requires a number of manual tracings to obtain supervised intensity information. With respect to the GB, its supervised IC-IM in non-contrast CT data will be generally applicable to most contrast-enhancement protocols and phases excepting CT cholangiography because the GB is usually unaffected by contrast media via blood vessels. Therefore, its supervised IC-IM will be worth being constructed through time-consuming efforts.

The experimental results show the effectiveness of the proposed unsupervised IC-IM although the number of input CT datasets for constructing IMs were about 20 at most. It is considered that collecting 20 or more CT datasets without manual tracing is easy in clinical practice. In the construction of the unsupervised IM, threshold value determined by the $1^{st}$ percentile in the probability distribution of PA were used for binarization of PA. The number of voxels of the region determined by the $1^{st}$ percentile were more than several thousand voxels by using prediction-based PA. It is considered that it was sufficient for the estimation of the initial histogram to be used for further MAP segmentation. Once core datasets with manual tracing are prepared for constructing shape–location priors, the proposed method can be applied to CT data of various ICs by collecting untraced CT datasets for the required IC. This feature is important for use in routine clinical practice. The

number of training data for shape–location priors was 43 in our two-fold cross validation. It is expected that the segmentation accuracy will be improved by increasing the number of training data. Investigation of the effects of the number of training data on segmentation accuracy will be a future work.

### 4.3. Comparison with existing methods

We compare the segmentation accuracy of the proposed method with those reported in the previous literatures (Wolz et al., 2013; Linguraru et al., 2012b; Shimizu et al., 2010; Hammon et al., 2013) although the datasets used for evaluation were different. These previous works used supervised intensity information and were tested for one contrast-enhancement pattern of CT data. In the spleen and kidneys, the segmentation accuracy of the proposed method (92.7 to 93.8% in DC) was comparable to or slightly better than state-of-the-art methods using one phase of the contrast-enhanced CT data with supervised intensity information (92.0 to 92.5%, 150 cases (Wolz et al., 2013) and 90.9 to 92.9%, 28 cases (Linguraru et al., 2012b) in DC). In the pancreas, the proposed method (60.0% in JI) was also comparable to or slightly better than the methods using supervised intensity information (55.5%, 150 cases (Wolz et al., 2013), 57.9%, 20 cases (Shimizu et al., 2010), 61.2%, 40 cases (Hammon et al., 2013) in JI from contrast enhanced CT data) in spite that supervised intensity information was not used in the proposed methods. In the GB, the segmentation accuracy of the supervised IC-IM (69.4% in JI) was slightly better and that of unsupervised IC-IM (57.9% in JI) was slightly worse than the multi-organ segmentation method developed by Shimizu et al. (2007) (around 62% in JI), which utilized supervised intensity information inherent in non-contrast CT data and was tested only for 10 cases. Average JIs of the aorta (80.4%) and IVC (50.1%) in the proposed method with unsupervised IC-IM were comparable to or slightly better than the Shimizu's method (around 77% in aorta and 48% in IVC) (Shimizu et al., 2007). From the literatures, segmentation accuracy of the proposed method was summarized as mostly comparable to or slightly better than the state-of-the-art methods in spite that supervised intensity information was not used in it as well as mixture sets of contrast-enhanced and non-contrast CT data were tested.

From the methodological viewpoints, a particular feature which differentiates the proposed method from existing methods is explicit modeling and utilization of spatial interrelations among multiple organ shapes for prediction of the target organ regions. That is, the targets are predicted from neighboring organ regions. In contrast, the machine-learning approaches have typically used low-level image features such as anatomical landmarks for the prediction (Seifert et al., 2009; Montillo et al., 2011; Glocker et al., 2012; Criminisi et al., 2013). Its advantages would be that the salient features are automatically determined during learning processes. That is, the targets are predicted from low-level features. In the statistical and multi-atlas approaches, multi-resolution modeling of multiple organs has been mainly used for the prediction in order to gradually increase the prediction accuracy in a robust manner (Wolz et al., 2013; Iglesias et al., 2013; Cerrolaza et al., 2012, 2013; Bagci et al., 2012). That is, the targets are predicted from lower-resolution organ regions. Overall, each approach uses different context information for the prediction, which is suitable for its methodological characteristics. The proposed method provides a different source of the

context information. Approaches to integrate different context information sources will be an important future work, one example of which is described next in Section 4.4.

### 4.4. Limitations and future work

In this method, spatial normalization is performed based on bone detection by simple thresholding and liver top detection by rough segmentation of the lung. Therefore, our current method for spatial normalization requires that these features and regions are imaged in the field of view (FOV) of input CT data, that is, the whole liver is included in the FOV. One possible method to overcome this limitation would be to use anatomical landmarks (Seifert et al., 2009; Hanaoka et al., 2011). We are currently considering a machine learning approach to perform spatial normalization based on detected anatomical landmarks.

As future work, it is desirable that the method is applicable to CT data with pathologies or missing organs. We are now extending the proposed method in the following points; (1) adding other abdominal organs, such as the gastrointestinal tract, to OCG (Hirayama et al., 2013), and (2) application to organ segmentation under pathological conditions (Fukuda et al., 2013). In addition, the proposed method obviously faces a problem when an organ is missing. A method for detecting missing organs, as proposed in Suzuki et al. (2012a), needs to be incorporated. Because specific patterns of organ shifts are statistically modeled after surgical removal of organs, as shown in Suzuki et al. (2012b), prediction-based shape–location priors would be constructed for each pattern of missing organs in order to reduce instability. With respect to intensity priors, our intensity modeling is currently based on intensity histograms of the organ and background regions. Incorporation of texture information inside the organ regions will improve the segmentation accuracy when neighboring organs have similar intensity histograms but different textures.

## 5. Conclusion

We have presented a general framework for multi-organ segmentation which is adaptable to various ICs without the need for supervised intensity information. In this paper, an OCG, which encodes the spatial correlations among organs, was introduced. By constructing and utilizing subject-specific shape– location priors based on shape prediction from pre-segmented organs, the segmentation accuracy of abdominal organs was largely improved. The initialization of organ locations and shapes becomes accurate enough to estimate IM in an unsupervised manner by using prediction-based shape–location priors. Average DCs of the liver, spleen, and kidneys were more than 92%. Average DCs of the pancreas, GB, and IVC were around 73%, 67%, and 65%, respectively. These results were comparable to state-of-the-art segmentation methods using supervised intensity information.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was partly supported by MEXT/JSPS Grand-in-Aid for Scientific Research on Innovative Areas No. 21103003 and No. 26108004, and the Intramural Research Program of the National Institutes of Health, Clinical Center, and a philanthropic gift from the Government of Abu Dhabi to Children's National Medical Center.

## Appendix A. Automated transformation of upper abdominal CT data to the abdominal normalized space

Figure A.1 shows the processes of automated abdominal spatial normalization. First, bone tissue regions and the liver dome (right lobe) are automatically extracted from the target CT data, and then spatial normalization is performed based on them. The procedure is as follows. Step 1: Bone tissue regions are extracted by thresholding the CT data within a fixed range along the caudal direction from the top axial plane of the 3D field of view of the CT data. The fixed range was the average z-length of the bounding box of the liver in the training data. The bone tissue regions determine their circumscribed two coronal and two sagittal planes of the abdominal space. Step 2: The top of the liver dome is determined by fitting SSM of the lung bases, which is spatially normalized by the liver top axial plane as well as the circumscribed coronal and sagittal planes, to the lung tissue regions extracted by thresholding. In this SSM fitting, the z-coordinate is also variable (in addition to the shape parameter vector of the SSM) and its estimated value is used to determine the liver top axial plane. Step 3: The determined axial, two sagittal, and two coronal reference planes are used to align the target CT data to the normalized space, in which the PAs and SSMs of all the eight organs are defined.

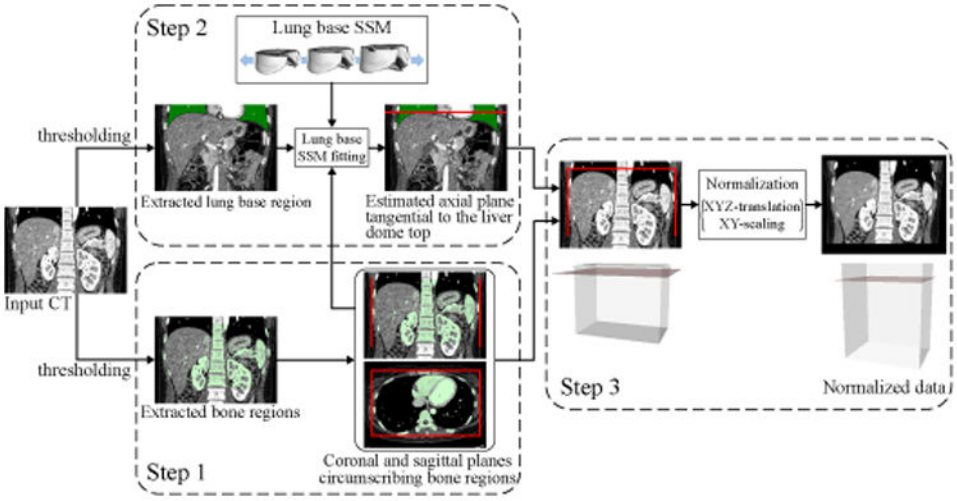



**Figure A.1.**
Processes of automated abdominal spatial normalization.

## Appendix B. Details of three submodules in single organ segmentation

For simplicity, the organ index *k* is omitted in the following descriptions.

## 1. PA-based MAP segmentation

Given target image data $I(\mathbf{x})$, $\hat{L}(\mathbf{x})$ maximizing the posterior probability Prob($L \mid I$) is calculated using PA $P(\mathbf{x})$ and IM $H(I) = \{H_O(I), H_B(I)\}$. The problem is formulated as binary MAP estimation for the target organ and background regions. Using Bayes' theorem, the posterior probability Prob($L \mid I$) is proportional to the product of likelihood Prob($I \mid L$) and the prior probability Prob($L$). Using $P(\mathbf{x})$, Prob($L$) is defined as

$$\mathrm{Prob}(L(\mathbf{x})=l)=\begin{cases} P(\mathbf{x}) & l=1 \\ 1-P(\mathbf{x}) & l=0 \end{cases}. \quad \text{(B.1)}$$

We assume that each $I(\mathbf{x})$ is conditionally independent. The likelihood Prob($I(\mathbf{x}) \mid L(\mathbf{x})$) is defined as

$$\mathrm{Prop}(I(\mathbf{x})|L(\mathbf{x})=l)=\begin{cases} H_{\mathrm{O}}(I(\mathbf{x})) & l=1 \\ H_{\mathrm{B}}(I(\mathbf{x})) & l=0 \end{cases}. \quad \text{(B.2)}$$

$\hat{L}(\mathbf{x})$is calculated by

$$\hat{L}(\mathbf{x})=\arg\min_{l}\mathrm{Prob}(I(\mathbf{x})|L(\mathbf{x})=l)\cdot\mathrm{Prob}(L(\mathbf{x})=l)=\begin{cases} 1 & H_{\mathrm{O}}(I(\mathbf{x}))\cdot P(\mathbf{x}) \geq H_{\mathrm{B}}(I(\mathbf{x}))\cdot(1-P(\mathbf{x})) \\ 0 & otherwise \end{cases} \quad \text{(B.3)}$$

After opening and closing using spherical structuring element of radius $\rho$ are applied to $\hat{L}(\mathbf{x})$, we obtain the PA-based segmentation result, which is denoted as $L_{\mathrm{PA}}(\mathbf{x})$. Optimization of radius $\rho$ is described in Section 3.3.

## 2. SSM-based refinement

SSM-based refinement of $L_{\mathrm{PA}}(\mathbf{x})$ is performed by repeating edge point detection on the binarized likelihood image and fitting SSM to the detected edge points (Okada et al., 2008). As a threshold value of the likelihood image, the value was calculated to satisfy the following: the intensity distribution within the region of binarized likelihood image include the intensity distribution of target organ and not include that of the background region as possible. To do so, the threshold value of likelihood image is determined by following way. Figure B.2 shows the method for determining the threshold value for the likelihood image. Let $p$ be the probability value of $H_O(I)$. The true positive fraction $TPF(p)$ and false positive fraction $FPF(p)$ of $H_O(I)$ at the probability $p$ are defined as $TPF(p) = \Sigma_i H_O(I = i)$ $[H_O(I = i) \geq p]$ and $FPF(p) = \Sigma_i H_B(I = i)[H_O(I = i) \geq p]$, respectively, where $[\cdot]$ denotes an Iverson bracket, which is equal to 1 when the logical condition enclosed is true and 0 otherwise (Fig. B.2). The threshold value for the likelihood image $p$ îs defined as

$$\hat{p}=\arg\min_{p}(TPF(p)-FPF(p)), \quad \text{(B.4)}$$

where Youden's index ($TPF(p) - FPF(p)$) (Youden, 1950) is maximized.

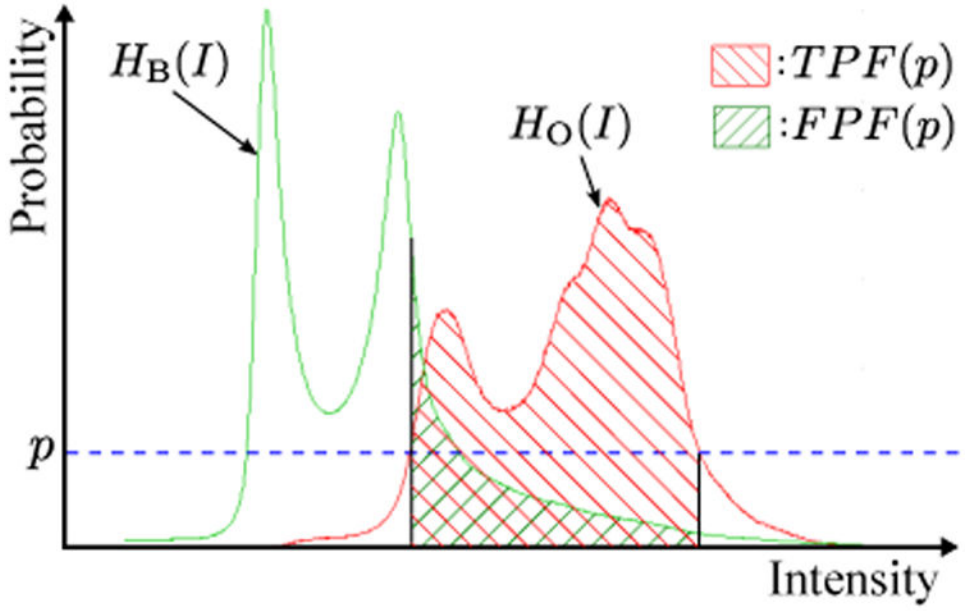



**Figure B.2.**
Determination of threshold value $p$ for likelihood image. Red and green plots are the histograms of object $H_O(I)$ and background $H_B(I)$, respectively. Given a probability value $p$, $TPF(p)$, $FPF(p)$, and Youden's index can be calculated. The threshold value for likelihood image $p$ is determined so that Youden's index is maximized.

Let $M(\mathbf{b})$ be the surface model generated by SSM $\mathbf{s}(\mathbf{b})$ and $X_{\text{likelihood}}$ the contour points of the binarized likelihood image by the threshold value $\hat{p}$. We define edge points $X^j$ detected at the $j$-th iteration as

$$X^j = \{\mathbf{x} | \mathbf{x} \in X_{\text{likehood}}, D(\mathbf{x}; M(\mathbf{b}^{j-1})) \leq d_{\max}\}, \quad \text{(B.5)}$$

where $D(\mathbf{x}; M)$ is the Euclidean distance between point $\mathbf{x}$ and the nearest point of $\mathbf{x}$ on the surface $M$, $\mathbf{b}^{j-1}$ is the shape parameter vector of SSM at the $(j-1)$-th iteration, and $d_{\max}$ is the maximum distance to limit the search area for edges. In this paper, we used $d_{\max} = 20$ mm. The maximum distance to limit the search area was also used by Lamecker et al. (2004) to reduce the influence of outliers. The contour points of $L_{\text{PA}}(\mathbf{x})$ are used for $X^1$.

Given edge points $X^j$, the shape parameter vector $\mathbf{b}^j$ of SSM is defined as

$$\mathbf{b}^j = \arg\min_{\mathbf{b}} \left( C_D(\mathbf{b}; M(\mathbf{b}), X^j) + \lambda C_M(\mathbf{b}) \right), \quad \text{(B.6)}$$

where $\lambda$ is a weight parameter, and $C_D(\mathbf{b}; M(\mathbf{b}), X^j)$ and $C_M(\mathbf{b})$ are defined as

$$C_{\mathrm{D}}(\mathbf{b};M(\mathbf{b}),X^j)=\frac{1}{|M(\mathbf{b})|}\sum_{\mathrm{y}\in M(\mathbf{b})}\min_{\mathbf{x}\in X^j}\|\mathbf{x}-\mathbf{y}\|^2+\frac{1}{|X^j|}\sum_{\mathrm{x}\in X^j}D^2(\mathbf{x};M(\mathbf{b})) \quad \text{(B.7)}$$

and

$$C_{\mathrm{M}}(\mathbf{b})=\sum_{i=1}^{m}b_i^2, \quad \text{(B.8)}$$

respectively, in which $|M|$ denotes the number of vertices of surface $M$, $\|\mathbf{v}\|$ the Euclidean norm of vector $\mathbf{v}$, and $m$ the number of principal modes of SSM. The Levenberg-Marquardt algorithm is used for minimization in Eq. (B.6). Optimization of weight parameter $\lambda$ is described in Section 3.3. The segmentation result of SSM refinement $L_{\mathrm{SSM}}(\mathbf{x})$ is obtained by voxelization of the shape model, which is obtained by a fixed number of iterations of edge point detection and fitting to the detected points.

## 3. Graph-cut refinement

Graph-cut refinement (Boykov and Jolly, 2001) is performed for $L_{\mathrm{SSM}}$. The boundary term and hard constraint in the original formulation (Boykov and Jolly, 2001) are used while the regional term is not used. Let $R = \{\mathbf{x} \mid L_{\mathrm{SSM}}(\mathbf{x}) = 1\}$ be the extracted region in $L_{\mathrm{SSM}}(\mathbf{x})$. The boundary term $B_{\{\mathbf{x},\mathbf{y}\}}$ is defined as

$$B_{\{\mathbf{x},\mathbf{y}\}}=\exp\left(-\frac{(I(\mathbf{x})-I(\mathbf{y}))^2}{2\sigma^2}\right)\cdot\frac{1}{\|\mathbf{x}-\mathbf{y}\|}, \quad \text{(B.9)}$$

where $\sigma^2=\frac{1}{n_p-1}\sum_{\mathbf{x}\in R}\sum_{\mathbf{y}\in N(\mathbf{x})\cap R}(I(\mathbf{x})-I(\mathbf{y}))^2$ is the variance of the intensity value between adjacent voxels, $n_p$ the total number of pairs of adjacent voxels within $R$, and $N(\mathbf{x})$ a neighborhood of $\mathbf{x}$. Hard constraint regions for object and background are automatically determined based on segmented region $R$ (while they are manually specified by a user in the original method developed by Boykov and Jolly (2001)). Let $H(I; R)$ be the probability distribution of the intensity histogram within $R$. The hard constraint regions for object and background $\mathcal{O}$ and $\mathcal{B}$ are respectively defined as

$$\mathcal{O}=\{\mathbf{x}|\mathbf{x}\in R, H(I(\mathbf{x});R)\geq t(p_{\mathrm{O}};H(I;R))\} \quad \text{(B.10)}$$

$$\mathcal{B}=\{\mathbf{x}|\mathbf{x}\in R^c, H(I(\mathbf{x});R)\leq t(p_{\mathrm{B}};H(I;R))\}, \quad \text{(B.11)}$$

where $R^c$ is the complement of $R$, $p_{\mathrm{O}}$ and $p_{\mathrm{B}}$ the values to determine the hard constraint regions for object and background, respectively, and $t(p; H(I; R)) = \max_t\{t \mid \Sigma_i H(I = i; R) [H(I = i; R) \geq t] \geq Z(p)\}$, in which $Z(p)$ is the probability of a random variable with standard normal distribution $\mathcal{N}(0, 1)$ falling in the range $[-p, p]$. Optimization of $p_{\mathrm{O}}$ and $p_{\mathrm{B}}$ is described in Section 3.3.

## Appendix C. Training of prediction model using PLSR

Let $X_c$ and $Y_c$ be the mean centered matrices of $X$ and $Y$, respectively. Using PLSR, $Y_c$ can be written as

$$Y_c = X_c B_{pls}, \quad \text{(C.1)}$$

where $B_{pls}$ is the regression coefficient matrix. This matrix can be calculated using the NIPALS algorithm (Yang et al., 2008; Geladi and Kowalski, 1986). The target vector $\hat{\mathbf{y}}$ predicted from the predictor vector $\mathbf{x}$ is defined as

$$\hat{\mathbf{y}} = B_{pls}^t(\mathbf{x} - \mu_X) + \mu_Y, \quad \text{(C.2)}$$

where $\mu_X$ and $\mu_Y$ are the mean vectors of $X$ and $Y$, respectively. Let $\tilde{X_i}$ and $\tilde{Y_i}$ be the datasets except for case $i$, $\tilde{B}_{pls}^{i}$ be the regression coefficient matrix constructed from $\tilde{X_i}$ and $\tilde{Y_i}$, and $B_{pls,m}$ be the matrix using the first $m$ columns of $B_{pls}$. The prediction error $\varepsilon^m$ is defined as

$$\varepsilon^m = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{y}_i - \hat{\mathbf{y}}_i^m\|^2, \quad \text{(C.3)}$$

where

$$\hat{\mathbf{y}}_i^m = \tilde{\boldsymbol{B}}_{pls,m}^{it}(\mathbf{x}_i - \mu_{\tilde{X}_i}) + \mu_{\tilde{Y}_i}. \quad \text{(C.4)}$$

The optimal number of factors of $B_{pls}$ is defined as

$$\hat{m} = \arg\min_m \varepsilon^m. \quad \text{(C.5)}$$

Given the predictor vector $\mathbf{x}$, the predicted target vector $\hat{\mathbf{y}}$ is defined as

$$\hat{\mathbf{y}} = PLS\ R(\mathbf{x}; X, Y) = B_{pls,\hat{m}}^t(\mathbf{x} - \mu_X) + \mu_Y. \quad \text{(C.6)}$$

## Appendix D. Automated surface subdivision based on canonical correlation analysis

Organ surfaces are recursively subdivided based on a significance level of point-pair correlation within the surface. We use canonical correlation analysis (CCA) to calculate the significance of correlations between two points (Fillard et al, 2007).

Let $V_0 = \{\mathbf{v}_i\}$ be the vertices of the surface $S$ to be divided and $p(\mathbf{u}, \mathbf{v})$ be the significance of correlations between the two points $\mathbf{u}$ and $\mathbf{v}$ based on CCA. Given the maximum level of division $k_{\max}$, the processes of the automated surface subdivision are as follows:

1. $p(\mathbf{v}_i, \mathbf{v}_j)$ is calculated among all possible pairs of points $(\mathbf{v}_i, \mathbf{v}_j)(i \neq j)$in $V_0$.
2. The level of division $k \leftarrow 0$
3. The most uncorrelated point-pair $(\mathbf{v}_{V_1}, \mathbf{v}_{V_2})$ is selected as

$$(\mathbf{v}_{V_1}, \mathbf{v}_{V_2}) = \underset{\mathbf{v}_i, \mathbf{v}_j \in V_0}{\arg\min} p(\mathbf{v}_i, \mathbf{v}_j). \quad \text{(D.1)}$$

4. All vertices in $V_0$ are separated into the two sets of vertices $V_1$ and $V_2$ based on the significance of the correlations with $\mathbf{v}_{V_1}$ and $\mathbf{v}_{V_2}$ as follows:

$$V_1 = \{\mathbf{v}_{V_1}\} \cup \{\mathbf{v}_i | \mathbf{v}_i \in V_0, p(\mathbf{v}_i, \mathbf{v}_{V_1}) < p(\mathbf{v}_i, \mathbf{v}_{V_2})\} \quad \text{(D.2)}$$

$$V_2 = \{\mathbf{v}_{V_2}\} \cup \{\mathbf{v}_i | \mathbf{v}_i \in V_0, p(\mathbf{v}_i, \mathbf{v}_{V_1}) > p(\mathbf{v}_i, \mathbf{v}_{V_2})\}. \quad \text{(D.3)}$$

5. The sub-shapes $S_1$ and $S_2$ containing only the vertices in$V_1$ and $V_2$ are extracted from $S$, respectively.
6. $k \leftarrow k + 1$
7. The processes 3. to 6. are recursively applied to $S_i(i = 1, 2)$ until the level of division reaches $k_{\max}$.

Once the surface subdivision is performed for the average shape, the surface model represented by SSM can be automatically subdivided using known correspondences.

## References


Aljabar P, Heckemann RA, Hammers A, Hajnal JV, Rueckert D. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. NeuroImage. 2009; 46:726–38. [PubMed: 19245840]

Bagci U, Chen X, Udupa JK. Hierarchical Scale-Based Multiobject Recognition of 3-D Anatomical Structures. IEEE Trans Med Imaging. 2012; 31:777–89. [PubMed: 22203704]

Boykov Y, Jolly MP. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. Proc Eighth IEEE Int Conf Comput Vis. 2001:105–12.

Camara O, Colliot O, Bloch I. Computational modeling of thoracic and abdominal anatomy using spatial relationships for image segmentation. Real-Time Imaging. 2004; 10:263–73.

Cerrolaza, JJ.; Herrezuelo, NC.; Villanueva, A.; Cabeza, R.; González Ballester, MA.; Linguraru, MG. Multiresolution Hierarchical Shape Models in 3D Subcortical Brain Structures. In: Mori, K.; Sakuma, I.; Sato, Y.; Barillot, C.; Navab, N., editors. Med Image Comput Comput Assist Interv. Vol. 8150. Springer; Berlin Heidelberg: 2013. p. 641-8.

Cerrolaza JJ, Villanueva A, Cabeza R. Hierarchical statistical shape models of multiobject anatomical structures: application to brain MRI. IEEE Trans Med Imaging. 2012; 31:713–24. [PubMed: 22194238]

Chen X, Summers RM, Cho M, Bagci U, Yao J. An Automatic Method for Renal Cortex Segmentation on CT Images. Acad Radiol. 2012; 19:562–70. [PubMed: 22341876]

Chu, C.; Oda, M.; Kitasaka, T.; Misawa, K.; Fujiwara, M.; Hayashi, Y.; Nimura, Y.; Rueckert, D.; Mori, K. Multi-organ Segmentation Based on Spatially-Divided Probabilistic Atlas from 3D Abdominal CT Images. In: Mori, K.; Sakuma, I.; Sato, Y.; Barillot, C.; Navab, N., editors. Med Image Comput Comput Assist Interv. Vol. 8150. Springer; Berlin Heidelberg: 2013. p. 165-72.

Criminisi A, Robertson D, Konukoglu E, Shotton J, Pathak S, White S, Siddiqui K. Regression forests for efficient anatomy detection and localization in computed tomography scans. Med Image Anal. 2013; 17:1293–1303. [PubMed: 23410511]

Cuingnet, R.; Prevost, R.; Lesage, D.; Cohen, LD.; Mory, B.; Ardon, R. Automatic Detection and Segmentation of Kidneys in 3D CT Images Using Random Forests. In: Ayache, N.; Delingette, H.; Golland, P.; Mori, K., editors. Med Image Comput Comput Assist Interv. Vol. 7512. Springer; Berlin Heidelberg: 2012. p. 66-74.

Dice LR. Measures of the Amount of Ecologic Association Between Species. Ecology. 1945; 26:297–302.

Duquette AA, Jodoin PM, Bouchot O, Lalande A. 3D segmentation of abdominal aorta from CT-scan and MR images. Comput Med Imaging Graph. 2012; 36:294–303. [PubMed: 22257909]

Fillard P, Pennec X, Thompson P, Ayache N. Evaluating Brain Anatomical Correlations via Canonical Correlation Analysis of Sulcal Lines. Proceedings of MICCAI 2007 workshop on statistical registration: Pair-wise and Group-wise Alignment and Atlas Formation. 2007

Freiman, M.; Kronman, A.; Esses, S.; Joskowicz, L.; Sosna, J. Non-parametric Iterative Model Constraint Graph min-cut for Automatic Kidney Segmentation. In: Jiang, T.; Navab, N.; Pluim, J.; Viergever, M., editors. Med Image Comput Comput Assist Interv. Vol. 6363. Springer; Berlin Heidelberg: 2010. p. 73-80.

Fukuda N, Nakamoto M, Okada T, Shoji S, Ukimura O, Gill IS, Hori M, Sato Y. Automated CT segmentation of cancerous kidneys by combining tumor detectors. JAMIT Annual Meeting 2013. 2013 in Japanese.

Geladi P, Kowalski BR. Partial least-squares regression: a tutorial. Anal Chim Acta. 1986; 185:1–17.

Glocker, B.; Pauly, O.; Konukoglu, E.; Criminisi, A. Joint Classification-Regression Forests for Spatially Structured Multi-object Segmentation. In: Fitzgibbon, A.; Lazebnik, S.; Perona, P.; Sato, Y.; Schmid, C., editors. Comput Vis - ECCV 2012. Vol. 7575. Springer; Berlin Heidelberg: 2012. p. 870-81.

Hammon M, Cavallaro A, Erdt M, Dankerl P, Kirschner M, Drechsler K, Wesarg S, Uder M, Janka R. Model-Based Pancreas Segmentation in Portal Venous Phase Contrast-Enhanced CT Images. J Digit Imaging. 2013; 26:1082–90. [PubMed: 23471751]

Hanaoka, S.; Masutani, Y.; Nemoto, M.; Nomura, Y.; Yoshikawa, T.; Hayashi, N.; Yoshioka, N.; Ohtomo, K. Probabilistic Modeling of Landmark Distances and Structure for Anomaly-proof Landmark Detection. In: Pennec, X.; Joshi, S.; Nielsen, M., editors. Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy - Geometrical and Statistical Methods for Modelling Biological Shape Variability. 2011. p. 159-69.

Heimann T, van Ginneken B, Styner MA, Arzhaeva Y, Aurich V, Bauer C, Beck A, Becker C, Beichel R, Bekes G, Bello F, Binnig G, Bischof H, Bornik A, Cashman PMM, Chi Y, Cordova A, Dawant BM, Fidrich M, Furst JD, Furukawa D, Grenacher L, Hornegger J, Kainmüller D, Kitney RI, Kobatake H, Lamecker H, Lange T, Lee J, Lennon B, Li R, Li S, Meinzer HP, Nemeth G, Raicu DS, Rau AM, van Rikxoort EM, Rousson M, Rusko L, Saddi KA, Schmidt G, Seghers D, Shimizu A, Slagmolen P, Sorantin E, Soza G, Susomboon R, Waite JM, Wimmer A, Wolf I. Comparison and evaluation of methods for liver segmentation from CT datasets. IEEE Trans Med Imaging. 2009; 28:1251–65. [PubMed: 19211338]

Heimann T, Meinzer HP. Statistical shape models for 3D medical image segmentation: a review. Med Image Anal. 2009; 13:543–63. [PubMed: 19525140]

Hirayama S, Okada T, Hori M, Tomiyama N, Sato Y. Automated segmentation of the upper digestive tract from abdominal contrast-enhanced CT data using hierarchical statistical modeling of organ inter-relations. Inst Electron Inf Commun Eng, Tech Rep. 2013; 112:149–154. in Japanese.

Iglesias JE, Sabuncu MR, Leemput KV. A unified framework for cross-modality multi-atlas segmentation of brain MRI. Med Image Anal. 2013; 17:1181–91. [PubMed: 24001931]

Isgum I, Staring M, Rutten A, Prokop M, Viergever MA, van Ginneken B. Multi-atlas-based segmentation with local decision fusion– application to cardiac and aortic segmentation in CT scans. IEEE Trans Med Imaging. 2009; 28:1000–10. [PubMed: 19131298]

Jaccard P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bull Soc Vaud Sci Nat. 1901; 37:547–79.

Kainmüller D, Lange T, Lamecker H. Shape constrained automatic segmentation of the liver based on a heuristic intensity model. Proc MICCAI Workshop on 3-D Segmentation in the Clinic: A Grand Challenge. 2007:109–16.

Khalifa F, Gimel'farb G, El-Ghar M, Sokhadze G, Manning S, McClure P, Ouseph R, El-Baz A. A new deformable model-based segmentation approach for accurate extraction of the kidney from abdominal CT images. 18th IEEE Int Conf Image Process (ICIP). 2011:3393–6.

Kohlberger, T.; Uzunbaş, MG.; Alvino, C.; Kadir, T.; Slosman, DO.; Funka-Lea, G. Organ Segmentation with Level Sets Using Local Shape and Appearance Priors. In: Yang, GZ.; Hawkes, D.; Rueckert, D.; Noble, A.; Taylor, C., editors. Med Image Comput Comput Assist Interv. Vol. 5762. Springer; Berlin Heidelberg: 2009. p. 34-42.

Kurkure U, Avila-Montes O, Kakadiaris I. Automated segmentation of thoracic aorta in non-contrast CT images. 5th IEEE Int Symp Biomed Imaging: From Nano to Macro. 2008:29–32.

Lamecker, H.; Lange, T.; Seebaß, M. Technical report. Zuse Institue; Berlin: 2004. Segmentation of the liver using a 3D statistical shape model.

Ling H, Zhou S, Zheng Y, Georgescu B, Suehling M, Comaniciu D. Hierarchical, learning-based automatic liver segmentation. IEEE Conf Comput Vis Pattern Recognit. 2008:1–8.

Linguraru M, Richbourg W, Liu J, Watt J, Pamulapati V, Wang S, Summers RM. Tumor Burden Analysis from CT Data of Diseased Patients via Automated Liver and Tumor Segmentation. IEEE Trans Med Imaging. 2012a; 31:1965–76. [PubMed: 22893379]

Linguraru M, Sandberg J, Li Z, Shah F, Summers R. Automated Segmentation and Quanitification of Liver and Spleen from CT Images using Normalized Probabilistic Atlases and Enhancement Estimation. Med Phys. 2010; 37:771–83. [PubMed: 20229887]

Linguraru MG, Pura Ja, Pamulapati V, Summers RM. Statistical 4D graphs for multi-organ abdominal segmentation from multiphase CT. Med Image Anal. 2012b; 16:904–14. [PubMed: 22377657]

Massoptier L, Casciaro S. Fully Automatic Liver Segmentation through Graph-Cut Technique. 29th Annu Int Conf IEEE Eng Med Biol Soc. 2007:5243–6.

Montillo, A.; Shotton, J.; Winn, J.; Iglesias, JE.; Metaxas, D.; Criminisi, A. Entangled Decision Forests and Their Application for Semantic Segmentation of CT Images. In: Székely, G.; Hahn, H., editors. Inf Process Med Imaging. Vol. 6801. Springer; Berlin Heidelberg: 2011. p. 184-96.

Okada, T.; Linguraru, MG.; Hori, M.; Summers, RM.; Tomiyama, N.; Sato, Y. Abdominal Multi-organ CT Segmentation Using Organ Correlation Graph and Prediction-Based Shape and Location Priors. In: Mori, K.; Sakuma, I.; Sato, Y.; Barillot, C.; Navab, N., editors. Med Image Comput Comput Assist Interv. Vol. 8151. Springer; Berlin Heidelberg: 2013. p. 275-82.

Okada T, Linguraru MG, Hori M, Suzuki Y, Summers RM, Tomiyama N, Sato Y. Multi-organ segmentation in abdominal CT images. IEEE Annu Int Conf Eng Med Biol Soc, IEEE. 2012a: 3986–9.

Okada, T.; Linguraru, MG.; Yoshida, Y.; Hori, M.; Summers, R.; Chen, YW.; Tomiyama, N.; Sato, Y. Abdominal Multi-Organ Segmentation of CT Images Based on Hierarchical Spatial Modeling of Organ Interrelations. In: Yoshida, H.; Sakas, G.; Linguraru, MG., editors. Abdom Imaging Comput Clin Appl. Vol. 7029. Springer; Berlin Heidelberg: 2012b. p. 173-80.

Okada T, Shimada R, Hori M, Nakamoto M, Chen YW, Nakamura H, Sato Y. Automated segmentation of the liver from 3D CT images using probabilistic atlas and multilevel statistical shape model. Acad Radiol. 2008; 15:1390–403. [PubMed: 18995190]

Park H, Bland PH, Meyer CR. Construction of an abdominal probabilistic atlas and its application in segmentation. IEEE Trans Med Imaging. 2003; 22:483–92. [PubMed: 12774894]

Sanchez-Castro F, Thierry L, Mory B. Automatic Inferior Vena Cava Segmentation for Hepatic Surgery Planning in Contrast-Enhanced CT Images. Int J Comput Assist Radiol Surg. 2010:S118–S119.

Seifert S, Barbu A, Zhou SK, Liu D, Feulner J, Huber M, Suehling M, Cavallaro A, Comaniciu D. Hierarchical parsing and semantic navigation of full body CT data. Proceedings of SPIE. 2009:725902–725902–8.

Shimizu A, Kimoto T, Kobatake H, Nawano S, Shinozaki K. Automated pancreas segmentation from three-dimensional contrast-enhanced computed tomography. Int J Comput Assist Radiol Surg. 2010; 5:85–98. [PubMed: 20033509]

Shimizu A, Ohno R, Ikegami T, Kobatake H, Nawano S, Smutek D. Segmentation of multiple organs in non-contrast 3D abdominal CT images. Int J Comput Assist Radiol Surg. 2007; 2:135–42.

Suzuki, M.; Linguraru, M.; Okada, K. Multi-Organ Segmentation with Missing Organs in Abdominal CT Images. In: Ayache, N.; Delingette, H.; Golland, P.; Mori, K., editors. Med Image Comput Comput Assist Interv. Vol. 7512. Springer; Berlin Heidelberg: 2012a. p. 418-25.

Suzuki, M.; Linguraru, M.; Summers, R.; Okada, K. Analyses of missing organs in abdominal multi-organ segmentation. In: Yoshida, H.; Sakas, G.; Linguraru, M., editors. Abdom Imaging Comput Clin Appl. Vol. 7029. Springer; Berlin Heidelberg: 2012b. p. 256-263.

Tomoshige S, Oost E, Shimizu A, Watanabe H, Nawano S. A conditional statistical shape model with integrated error estimation of the conditions; Application to liver segmentation in non-contrast CT images. Med Image Anal. 2014; 18:130–143. [PubMed: 24184436]

Umetsu S, Shimizu A, Watanabe H, Kobatake H, Nawano S. An automated segmentation algorithm for ct volumes of livers with atypical shapes and large pathological lesions. IEICE Trans Inf Syst. 2014; 97:951–963.

Uzunbaş MG, Soldea O, Unay D, Cetin M, Unal G, Erçil A, Ekin A. Coupled nonparametric shape and moment-based intershape pose priors for multiple basal ganglia structure segmentation. IEEE Trans Med Imaging. 2010; 29:1959–78. [PubMed: 21118755]

Wolz R, Chu C, Misawa K, Fujiwara M, Mori K, Rueckert D. Automated abdominal multi-organ segmentation with subject-specific atlas generation. IEEE Trans Med Imaging. 2013; 32:1723–30. [PubMed: 23744670]

Yang J, Staib L, Duncan J. Neighbor-constrained segmentation with level set based 3-D deformable models. IEEE Trans Med Imaging. 2004; 23:940–8. [PubMed: 15338728]

Yang YM, Rueckert D, Bull AMJ. Predicting the shapes of bones at a joint: application to the shoulder. Comput Methods Biomech Biomed Eng. 2008; 11:19–30.

Yokota, F.; Okada, T.; Takao, M.; Sugano, N.; Tada, Y.; Tomiyama, N.; Sato, Y. Automated CT segmentation of diseased hip using hierarchical and conditional statistical shape models. In: Mori, K.; Sakuma, I.; Sato, Y.; Barillot, C.; Navab, N., editors. Med Image Comput Comput Assist Interv. Vol. 8151. Springer; Berlin Heidelberg: 2013. p. 190-7.

Yoshida Y, Okada T, Hori M, Chen YW, Tomiyama N, Sato Y. Hierarchical spatial normalization based on multi-organ interrelation for atlas-based segmentation: application to gallbladder segmentation from abdominal CT data. Int J Comput Assist Radiol Surg. 2011:S326–S327.

Youden WJ. Index for rating diagnostic tests. Cancer. 1950; 3:32–35. [PubMed: 15405679]

Zhang X, Tian J, Deng K, Wu Y, Li X. Automatic Liver Segmentation Using a Statistical Shape Model With Optimal Surface Detection. IEEE Trans Biomed Eng. 2010; 57:2622–6. [PubMed: 20615804]

Zhou J, Huang W, Zhang J, Yang T, Liu J, Chui CK, Chang S. Segmentation of gallbladder from CT images for a surgical training system. 3rd Int Conf Biomed Eng Inform (BMEI). 2010:536–40.

Zhou, X.; Kitagawa, T.; Hara, T.; Fujita, H.; Zhang, X.; Yokoyama, R.; Kondo, H.; Kanematsu, M.; Hoshi, H. Constructing a Probabilistic Model for Automated Liver Region Segmentation Using Non-contrast X-Ray Torso CT images. In: Larsen, R.; Nielsen, M.; Sporring, J., editors. Med Image Comput Comput Assist Interv. Vol. 4191. Springer; Berlin Heidelberg: 2006. p. 856-63.

## Highlights

- Multiple abdominal organs were automatically segmented from CT images.
- A method for modeling prediction-based conditional shape-location priors is developed.
- Organ correlation graph is introduced, which embeds the spatial correlations among organs.
- Proposed priors enables the estimation of intensity priors from unsupervised intensity information.
- The results were comparable to state-of-the-art methods using supervised intensity information

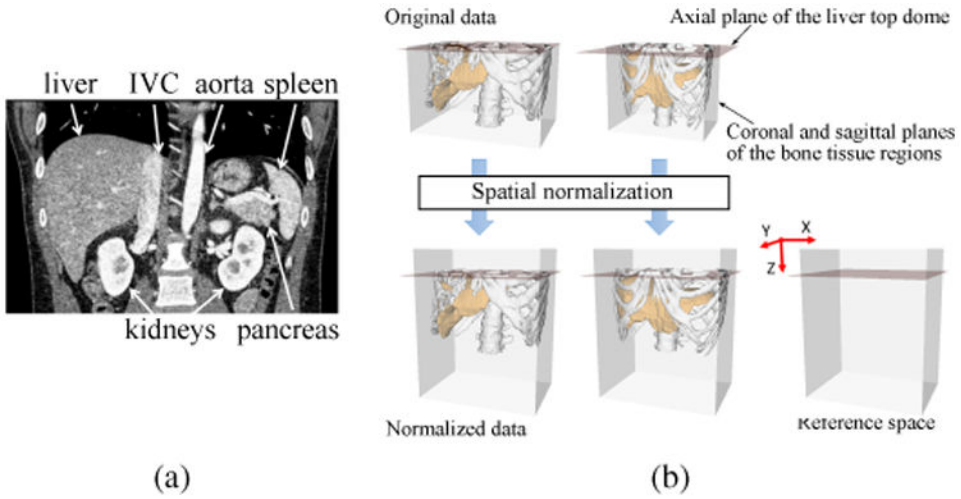



**Figure 1.**

Spatial normalization. (a) Coronal view of typical our CT data. (b) Definition of the abdominal normalized space. In spatial normalization, the xyz-translation and xy-scaling are adjusted so that the reference planes are aligned among all the CT data. The z-scaling is not performed because stable image features were not available for the z-scale determination. In addition, strong correlation was not found between the xy- and z-scaling at least for the liver shape. Although only sagittal planes are shown in the figures, coronal planes are similarly defined. See Appendix A for an automated method for the spatial normalization.

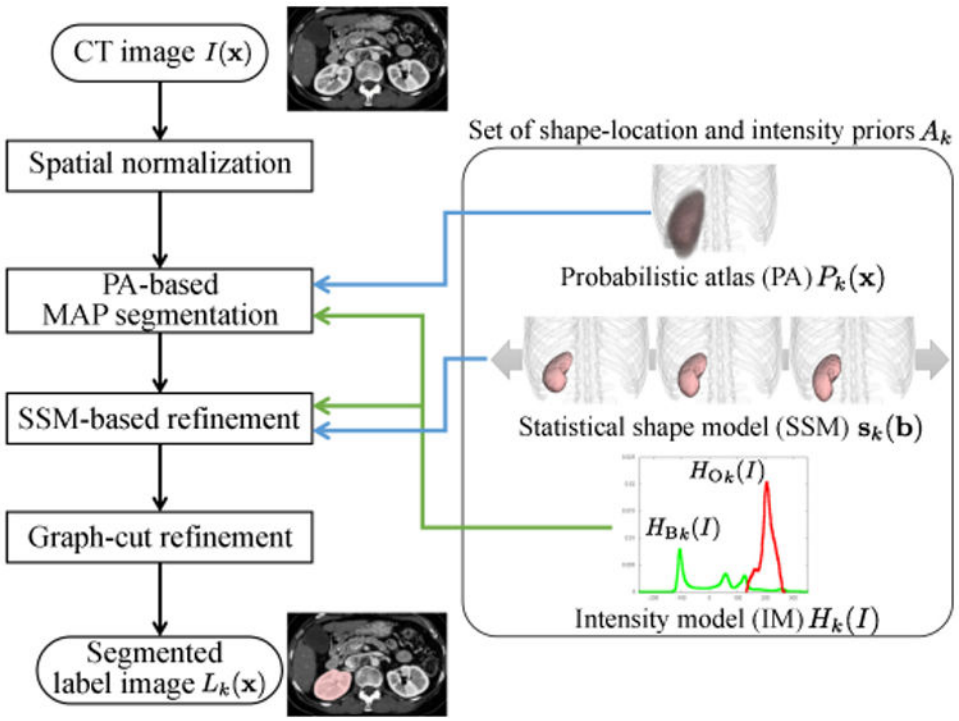



**Figure 2.**

Block diagram of single-organ segmentation method. As an example, the segmentation target is R-kidney here. This is utilized as a basic module for multi-organ segmentation, where (conventional) PA and SSM will be replaced by prediction-based conditional ones described in Section 2.2. Three methods for constructing intensity models will be considered in Section 2.3.

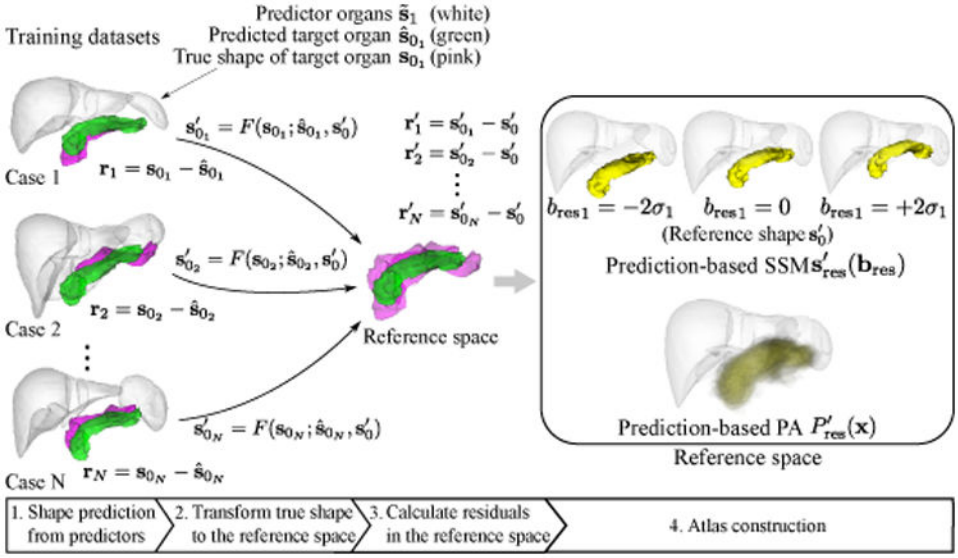



**Figure 3.**

Construction of prediction-based shape–location priors in reference space. Green shapes in training dataset and reference space indicate the predicted and reference shapes of the target organ (which is the pancreas in this figure), respectively. Pink shapes in training dataset and reference space indicate the true shapes and transformed true shapes of the target organ, respectively. White shapes indicate the predictor organs, which are the liver and spleen in this case. After predicting the target organ shape from predictor organs, true shapes are transformed to reference space. Then SSM and PA of the residuals are constructed in the reference space.

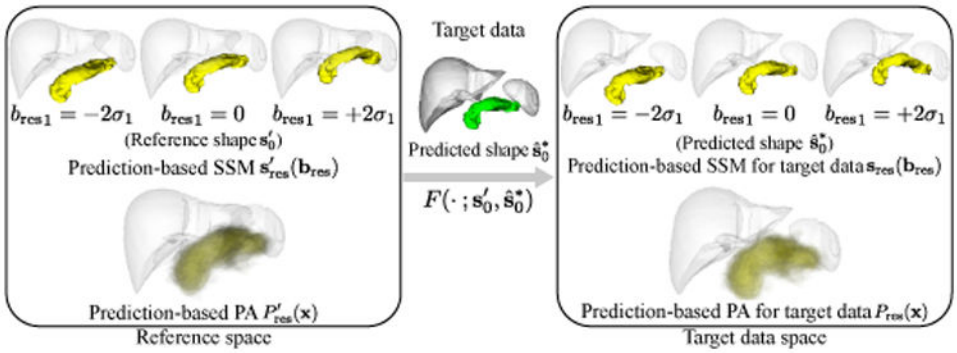



**Figure 4.**

Generating prediction-based shape–location priors for segmentation of target data. Given segmented predictor organs of the target data, the target organ shape $\hat{\mathbf{s}}_{0t}$ is predicted. The mapping defined by the deformation field from reference shape $\mathbf{s}'_0$ to predicted shape, $\hat{\mathbf{s}}^*_0$, $F(\cdot;\mathbf{s}'_0, \hat{\mathbf{s}}^*_0)$, is applied to the prediction-based SSM and PA in the reference space to obtain those in the target data space.

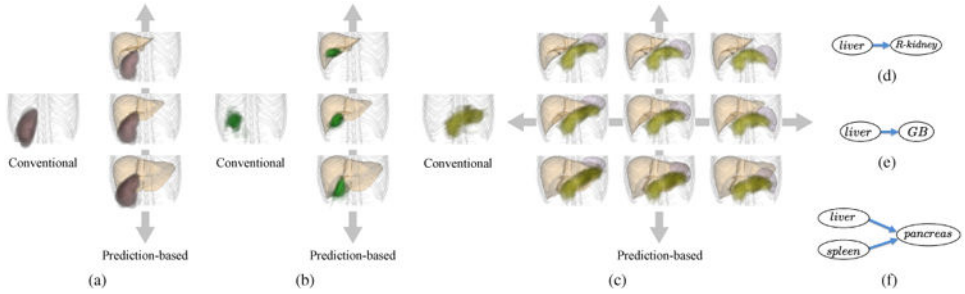



**Figure 5.**

Conventional and proposed prediction-based PAs. (a) R-kidney. (b) GB. (c) Pancreas. Left: Conventional PA. Right: Proposed prediction-based PA using the predictor organs (the liver for R-kidney and GB, and the liver & spleen for the pancreas). The variations of the prediction-based PA due to the deformations of the predictor organs are shown. (d), (e), and (f) Graph representations of (a), (b), and (c), respectively.

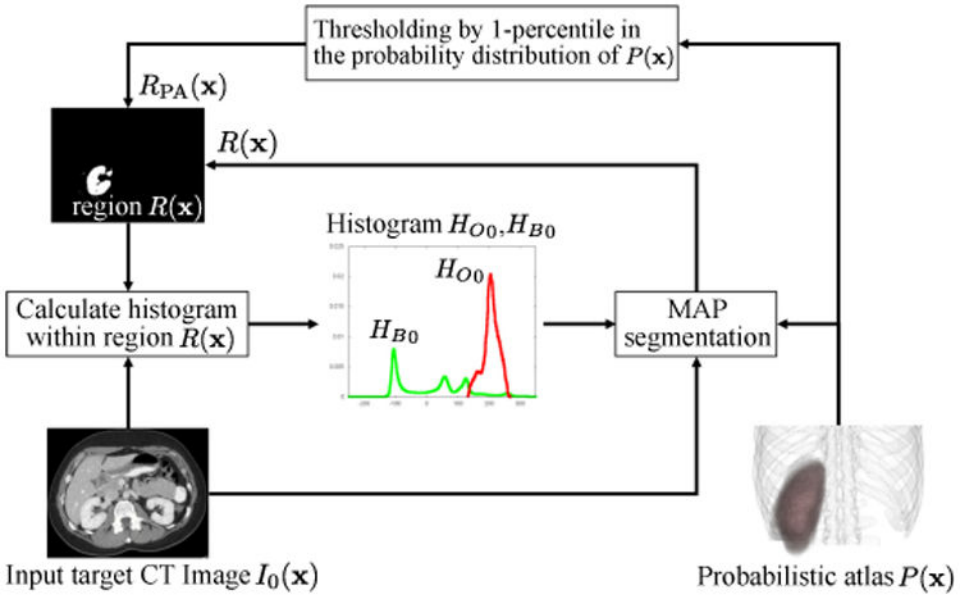



**Figure 6.**

Process of construction of target-data specific intensity model (TD-IM). TD-IM $H_{\mathrm{O}}$ and $H_{\mathrm{B}}$ are obtained by iterating the processes obtaining $H_{\mathrm{O0}}$ with $R(\mathbf{x})$ and updating $R(\mathbf{x})$ by MAP segmentation.

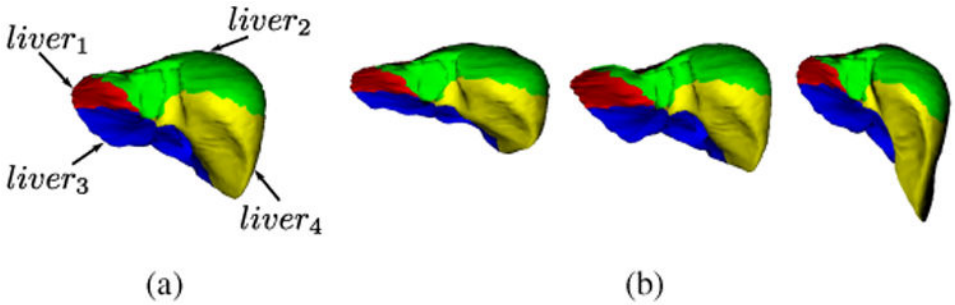



**Figure 7.**

Four sub-shapes of the liver shape. (a) Average shape. (b) Three different patient liver shapes. Red, green, blue, and yellow sub-shapes are $liver_1$, $liver_2$, $liver_3$, and $liver_4$, respectively.

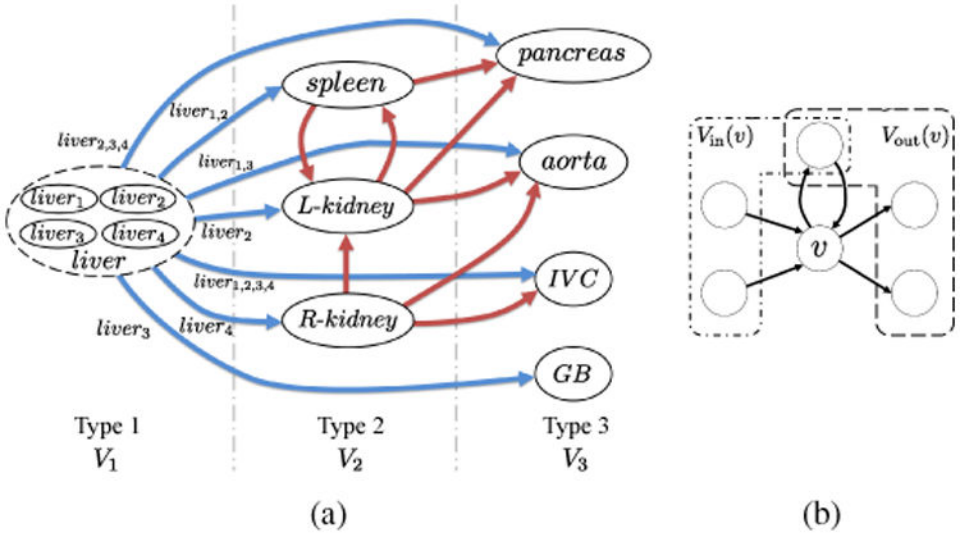



**Figure 8.**

Organ correlation graph (OCG). (a) OCG constructed from eight abdominal organs. Blue and red edges indicate the edges from nodes of Types 1 and 2, respectively. For simplicity, the multiple edges from the sub-shapes of the liver to a target organ are indicated by one edge from the liver to it and the identifiers of sub-shapes are indicated by the labels on the edges. In the graph, a node having in-edges is predicted from nodes having out-edges. (b) Examples of $V_{\text{out}}(v)$ and $V_{\text{in}}(v)$. As shown in this figure, some nodes can be included in both $V_{\text{out}}(v)$ and $V_{\text{in}}(v)$.

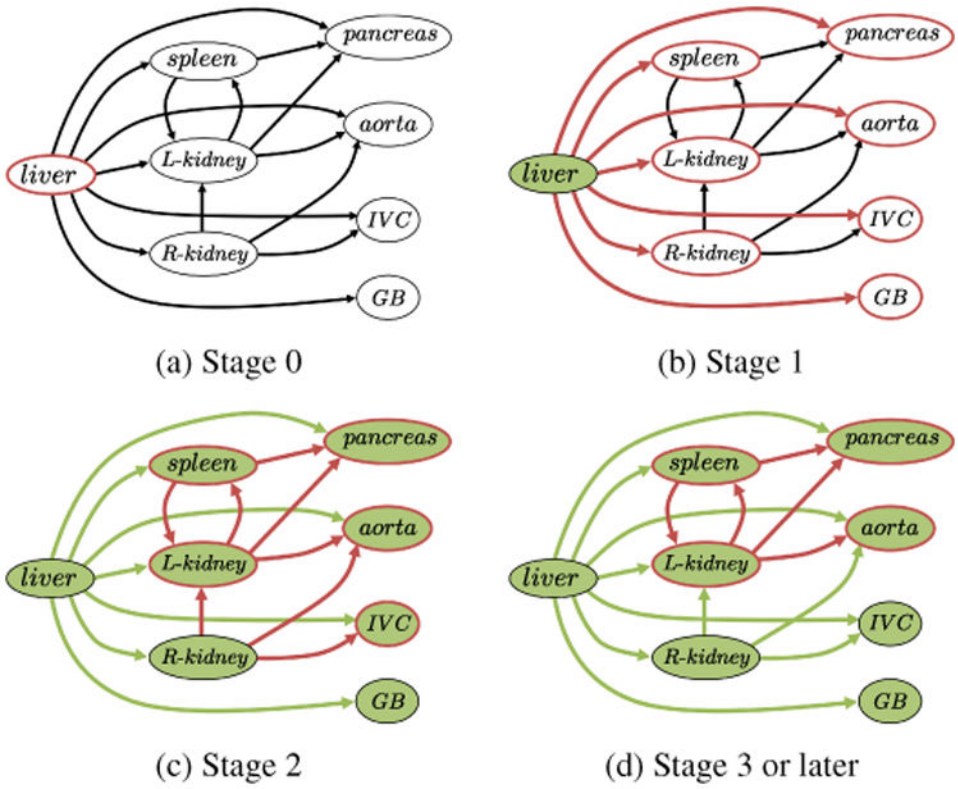



**Figure 9.**

Processes of multi-stage updating in multi-organ segmentation based on OCG. Red contoured nodes indicate those to be segmented. Green-colored nodes indicate those that have already been segmented. (a) Stage 0. The liver (only one organ in $V_1$) is firstly segmented. (b) Stage 1. The organs connected from the sub-shapes of the liver are segmented. (c) Stage 2. The organs connected from those segmented at the previous stage are segmented. Note that the R-kidney and GB are not segmented at this stage because the segmentation results of predictors (only the liver in this case) do not change. (d) Stage 3 or later. Further segmentations are performed using updated results of the predictors.

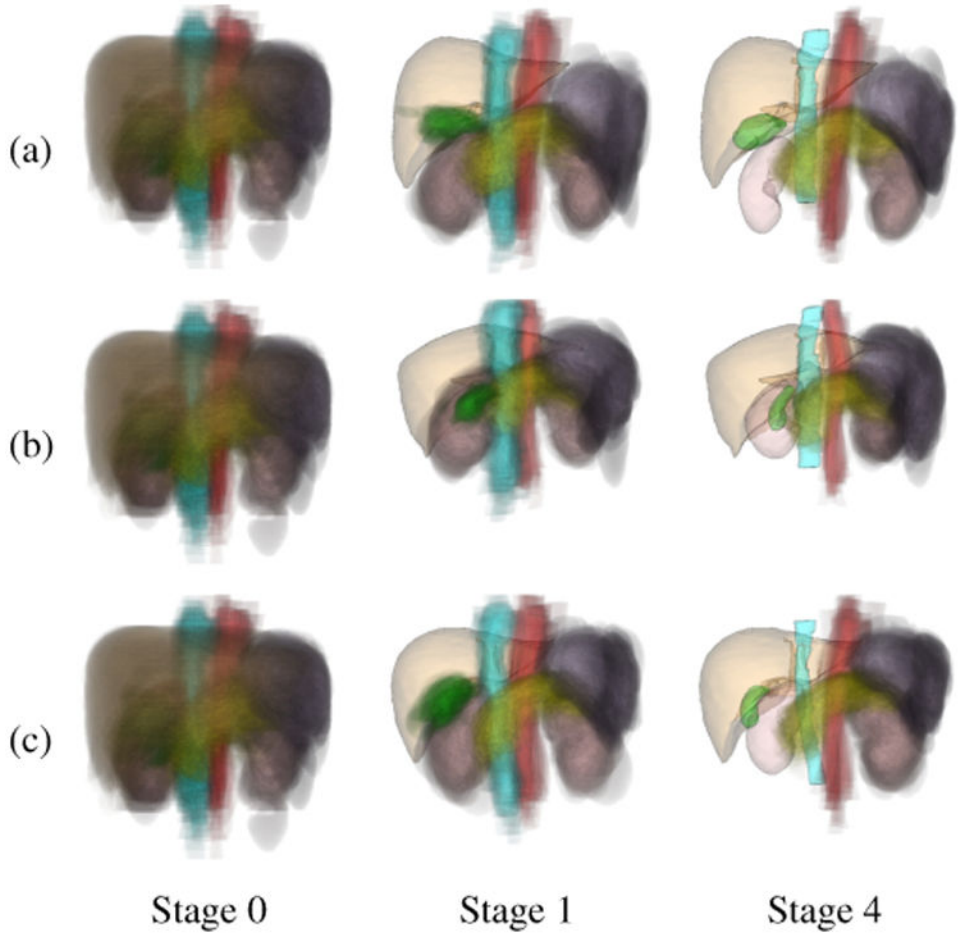



**Figure 10.**

Examples of prediction-based PAs at different stages of multi-organ segmentation. (a) Case 1. (b) Case 2. (c) Case 3. Left: Conventional PA of eight organs (Stage 0). Middle: Prediction-based PAs using the segmented liver (Stage 1). Right: Prediction-based PAs using the segmented liver, spleen, and kidneys (Stage 4). Orange indicates the liver, purple the spleen, pink the kidneys, yellow the pancreas, green the GB, red the aorta, and cyan the IVC. PAs are illustrated by cloud-like semi-transparent volume rendering. The segmented organ shapes are represented by semi-transparent surface rendering. At Stage 4, segmentation of R-kidney, gallbladder, and IVC (in addition to the liver) stops updating, and thus their segmented surfaces are shown instead of PAs.

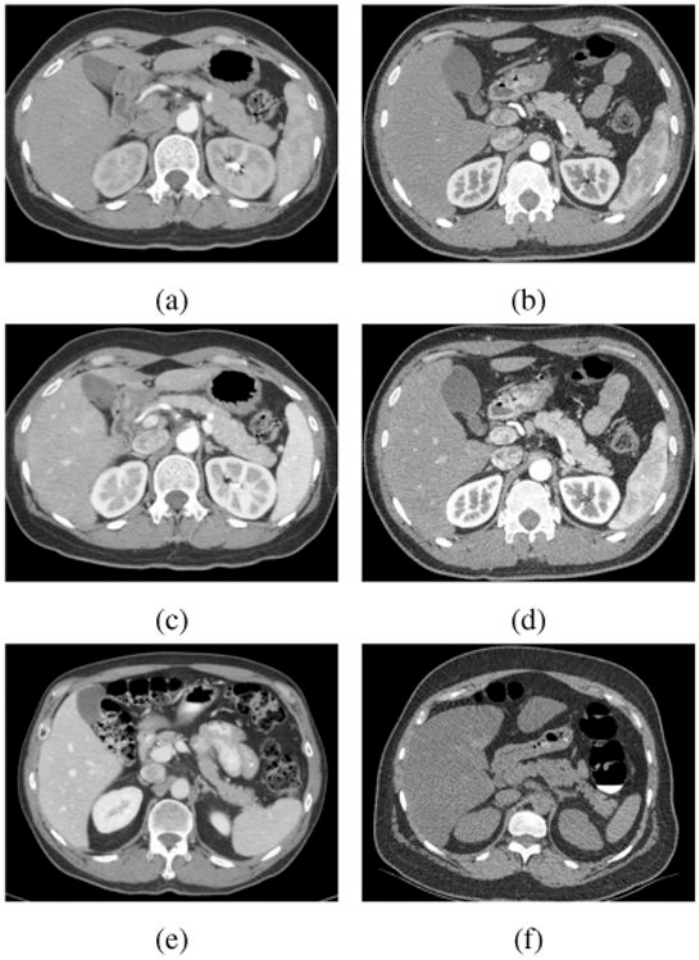
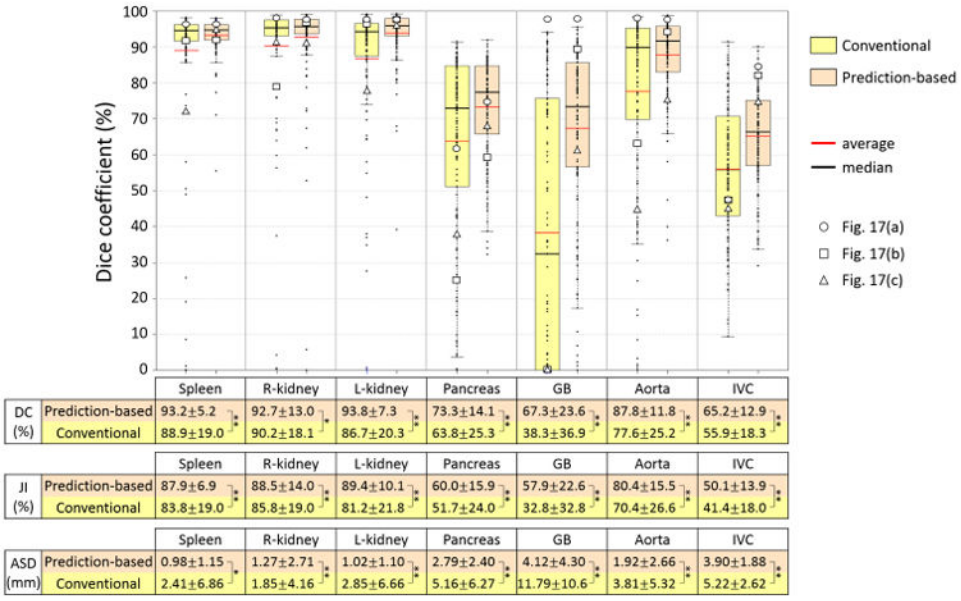



**Figure 11.**
Examples of cross-sectional CT images of our data. See Table 4 for the descriptions of each IC. (a) A-1, (b) A-2, (c) B-1 (d) B-2, (e) C, and (f) D.




| | | Spleen | R-kidney | L-kidney | Pancreas | GB | Aorta | IVC |
|---|---|---|---|---|---|---|---|---|
| DC (%) | Prediction-based | 93.2±5.2 | 92.7±13.0 | 93.8±7.3 | 73.3±14.1 | 67.3±23.6 | 87.8±11.8 | 65.2±12.9 |
| | Conventional | 88.9±19.0 | 90.2±18.1 | 86.7±20.3 | 63.8±25.3 | 38.3±36.9 | 77.6±25.2 | 55.9±18.3 |

| | | Spleen | R-kidney | L-kidney | Pancreas | GB | Aorta | IVC |
|---|---|---|---|---|---|---|---|---|
| JI (%) | Prediction-based | 87.9±6.9 | 88.5±14.0 | 89.4±10.1 | 60.0±15.9 | 57.9±22.6 | 80.4±15.5 | 50.1±13.9 |
| | Conventional | 83.8±19.0 | 85.8±19.0 | 81.2±21.8 | 51.7±24.0 | 32.8±32.8 | 70.4±26.6 | 41.4±18.0 |

| | | Spleen | R-kidney | L-kidney | Pancreas | GB | Aorta | IVC |
|---|---|---|---|---|---|---|---|---|
| ASD (mm) | Prediction-based | 0.98±1.15 | 1.27±2.71 | 1.02±1.10 | 2.79±2.40 | 4.12±4.30 | 1.92±2.66 | 3.90±1.88 |
| | Conventional | 2.41±6.86 | 1.85±4.16 | 2.85±6.66 | 5.16±6.27 | 11.79±10.6 | 3.81±5.32 | 5.22±2.62 |

**Figure 12.**

Summary of accuracy evaluation for proposed prediction-based and conventional shape–location priors when unsupervised IC-IM was used for intensity model. Dice coefficients (DCs) of 134 CT data using prediction-based and conventional shape–location priors for each of seven organs were plotted. The averages of DC, Jaccard index (JI), and average symmetric surface distance (ASD) as well as statistical significance are also shown in a table format. In the table, * and ** indicate that significant accuracy improvement with significance levels of 0.05 and 0.01 was observed, respectively. DCs of three cases indicated in Figs. 17(a), (b), and (c) are plotted by ○,□, and Δ, respectively.

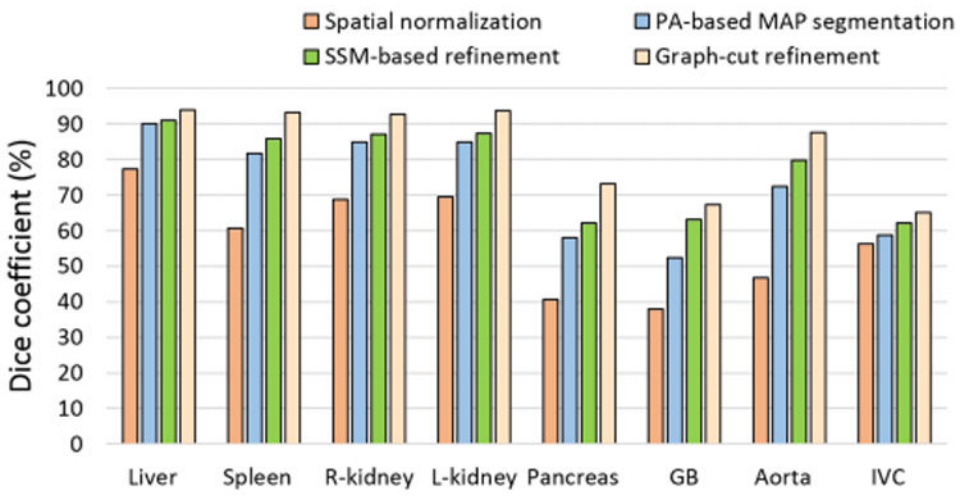



**Figure 13.**

Accuracy improvement of each segmentation submodule corresponding to each block in Fig. 2. The same shape-location and intensity priors were used as the proposed method in Fig. 12. Statistical significance was observed among all submodules for all organs ($p < 0.01$).

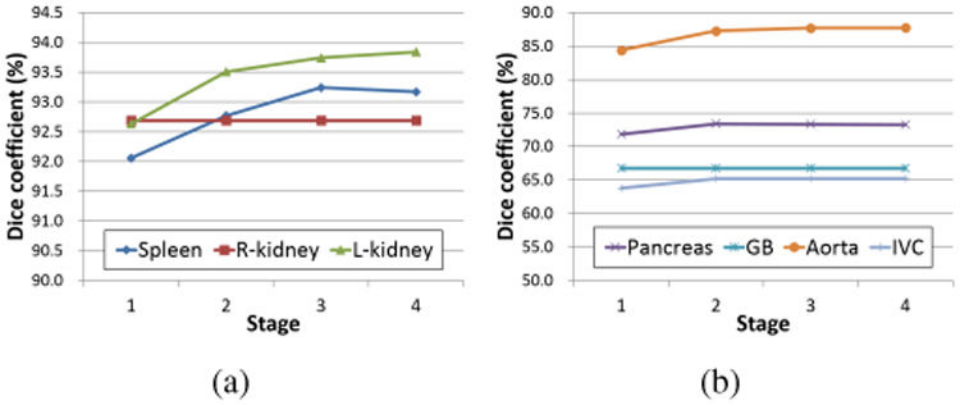



**Figure 14.**

Effects of multi-stage updating in proposed multi-organ segmentation based on OCG. Average DC of 134 CT data at each stage is plotted for each organ. (a) Plots for the spleen, R-kidney, and L-kidney. (b) Plots for the pancreas, GB, aorta, and IVC. Only the liver was used as the predictor at Stage 1. The spleen, R-kidney, and L-kidney were used as the predictors in addition to the liver at Stages 2 to 4.

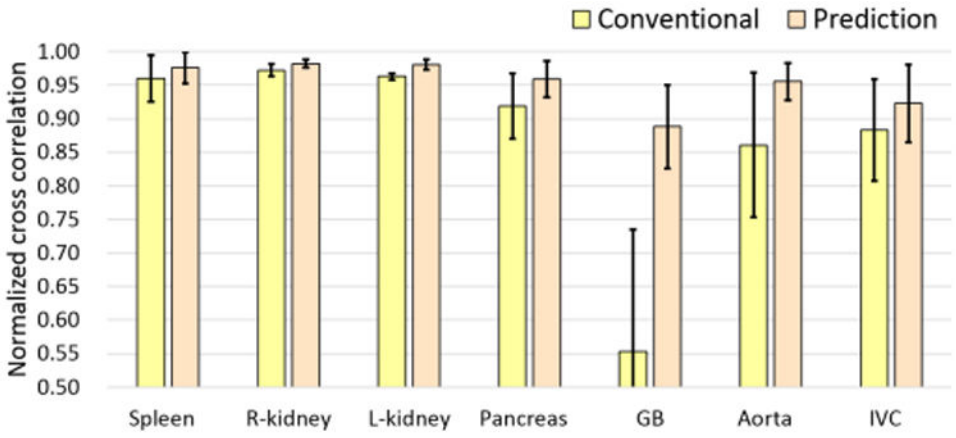
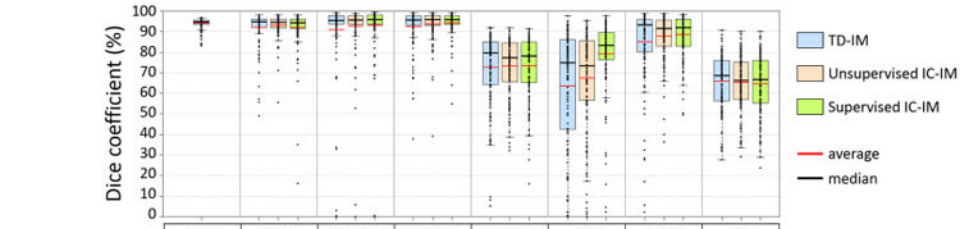



**Figure 15.**

Similarity between supervised and unsupervised IC-IMs. Normalized cross correlations (NCCs) between supervised IC-IM and unsupervised IC-IM using conventional (yellow bar) and prediction-based PAs (orange bar) were shown. NCCs of the six imaging conditions were averaged.



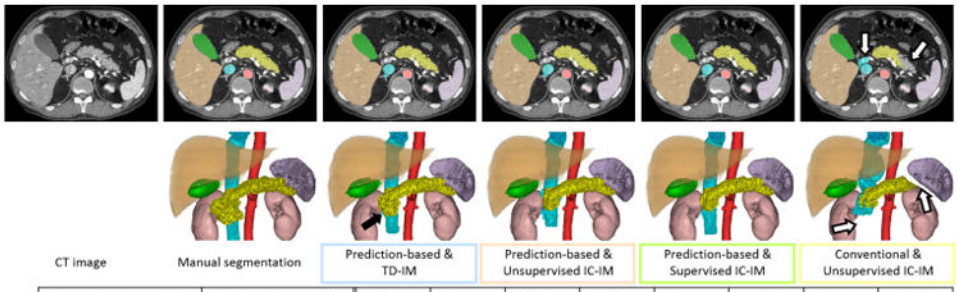

| | | Liver | Spleen | R-Kidney | L-Kidney | Pancreas | GB | Aorta | IVC |
|---|---|---|---|---|---|---|---|---|---|
| DC (%) | TD-IM | 94.1±2.4 | 92.1±8.1 | 90.9±18.4 | 92.3±10.1 | 72.5±17.6 | 63.5±28.5 | 85.0±18.6 | 65.8±13.3 |
| | Unsupervised IC-IM | - | 93.2±5.2 | 92.7±13.0 | 93.8±7.3 | 73.3±14.1 | 67.3±23.6 | 87.8±11.8 | 65.2±12.9 |
| | Supervised IC-IM | - | 92.0±9.7 | 93.1±13.0 | 94.4±6.1 | 73.4±15.1 | 79.2±17.5 | 88.5±10.3 | 64.5±14.4 |

| | | Liver | Spleen | R-Kidney | L-Kidney | Pancreas | GB | Aorta | IVC |
|---|---|---|---|---|---|---|---|---|---|
| JI (%) | TD-IM | 88.9±4.1 | 87.0±9.6 | 87.1±18.8 | 88.2±11.7 | 60.1±18.2 | 54.9±26.9 | 78.1±20.6 | 51.0±14.3 |
| | Unsupervised IC-IM | - | 87.9±6.9 | 88.5±14.0 | 89.4±10.1 | 60.0±15.9 | 57.9±22.6 | 80.4±15.5 | 50.1±13.9 |
| | Supervised IC-IM | - | 86.4±11.5 | 89.1±14.1 | 90.1±8.7 | 60.4±16.7 | 69.4±18.5 | 81.2±14.1 | 49.6±15.2 |

| | | Liver | Spleen | R-Kidney | L-Kidney | Pancreas | GB | Aorta | IVC |
|---|---|---|---|---|---|---|---|---|---|
| ASD (mm) | TD-IM | 1.73±0.92 | 1.20±1.61 | 1.66±3.91 | 1.33±1.85 | 3.03±3.13 | 5.15±5.49 | 2.11±3.27 | 4.00±1.82 |
| | Unsupervised IC-IM | - | 0.98±1.15 | 1.27±2.71 | 1.02±1.10 | 2.79±2.40 | 4.12±4.30 | 1.92±2.66 | 3.90±1.88 |
| | Supervised IC-IM | - | 1.26±2.43 | 1.12±2.41 | 0.90±0.90 | 2.77±2.40 | 2.30±3.15 | 1.60±2.08 | 3.84±1.88 |

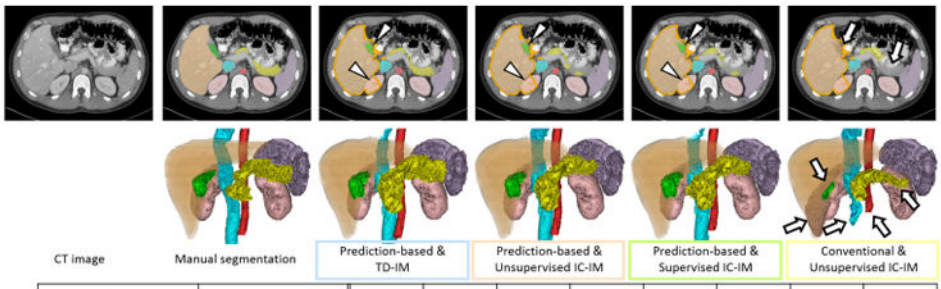
**Figure 16.**

Summary of accuracy evaluation for three intensity models when the proposed prediction-based shape–location priors were utilized. DCs of 134 CT data using TD-IM, unsupervised IC-IM, and supervised IC-IM for each of eight organs were plotted. With respect to the liver, only the result using TD-IM is shown. Average DC, JI, and ASD values and statistical significance are also shown below the plots in a table format. In the table, * and ** indicate that significant accuracy improvement with significance levels of 0.05 and 0.01 was observed, respectively.



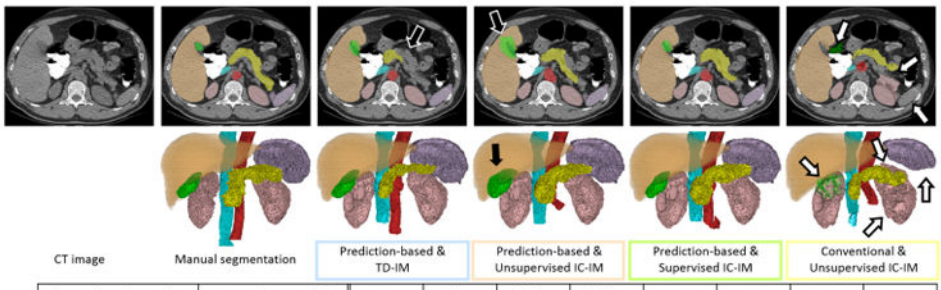

| | Shape-location priors | Intensity model | Liver | Spleen | R-kidney | L-kidney | Pancreas | GB | Aorta | IVC |
|---|---|---|---|---|---|---|---|---|---|---|
| DC (%) | Prediction-based | TD-IM | 96.1 | 96.1 | 98.0 | 97.4 | 83.7 | 97.5 | 96.5 | 85.1 |
| | | Unsupervised IC-IM | - | 96.1 | 97.6 | 97.1 | 74.4 | 97.7 | 97.8 | 84.4 |
| | | Supervised IC-IM | - | 96.1 | 98.0 | 96.8 | 74.0 | 97.4 | 98.0 | 81.7 |
| | Conventional | Unsupervised IC-IM | - | 96.2 | 98.3 | 97.0 | 61.3 | 97.6 | 98.1 | 47.7 |

(a) Case in Dataset B-2




| | Shape-location priors | Intensity model | Liver | Spleen | R-kidney | L-kidney | Pancreas | GB | Aorta | IVC |
|---|---|---|---|---|---|---|---|---|---|---|
| DC (%) | Prediction-based | TD-IM | 89.7 | 93.1 | 96.5 | 97.0 | 61.6 | 89.3 | 93.7 | 76.2 |
| | | Unsupervised IC-IM | - | 91.3 | 97.0 | 97.1 | 59.8 | 89.6 | 93.6 | 82.0 |
| | | Supervised IC-IM | - | 91.3 | 96.9 | 97.0 | 59.4 | 87.9 | 91.9 | 76.8 |
| | Conventional | Unsupervised IC-IM | - | 89.6 | 78.1 | 96.5 | 25.3 | 0.0 | 63.1 | 47.5 |

(b) Case in Dataset C




| | Shape-location priors | Intensity model | Liver | Spleen | R-kidney | L-kidney | Pancreas | GB | Aorta | IVC |
|---|---|---|---|---|---|---|---|---|---|---|
| DC (%) | Prediction-based | TD-IM | 95.8 | 93.7 | 93.0 | 95.5 | 32.5 | 82.5 | 88.6 | 72.5 |
| | | Unsupervised IC-IM | - | 94.9 | 91.0 | 95.6 | 68.2 | 61.8 | 75.9 | 74.2 |
| | | Supervised IC-IM | - | 94.1 | 89.4 | 95.6 | 79.5 | 89.5 | 83.9 | 72.7 |
| | Conventional | Unsupervised IC-IM | - | 73.2 | 91.6 | 77.6 | 38.0 | 0.0 | 45.4 | 45.7 |

(c) Case in Dataset D

**Figure 17.**

Illustrative segmentation results (a) Case in Dataset B-2 (arterial phase enhancement). (b) Case in Dataset C (portal venous phase enhancement). (c) Case in Dataset D (non-contrast). From the left to the right, input CT image, manual segmentation, and results using the proposed prediction-based shape–location priors & TD-IM (blue), the proposed prediction-

based priors & unsupervised IC-IM (orange), the proposed priors & supervised IC-IM (green), and the conventional priors & unsupervised IC-IM (yellow). Upper Axial cross-section. Lower 3D rendering. Orange, purple, pink, yellow, green, red, and cyan volumes and regions indicate the liver, spleen, kidneys, pancreas, GB, aorta, and IVC, respectively. Dice coefficient (DC) of each segmented organ is also shown in a table format. In (b), orange contour indicates the contour of the liver region to show liver leakage. White arrows show the failed region by the conventional priors. Black arrows show the improved or failed regions by the prediction-based priors. Arrow heads show the regions of liver leakage.

**Table 1**

Commonly used acronyms.

| | Acronym | Description |
|---|---|---|
| Shape–location prior | OCG | Organ Correlation Graph |
| | PA | Probabilistic Atlas |
| | SSM | Statistical Shape Model |
| Intensity prior | IC | Imaging Condition |
| | IM | Intensity Model |
| | TD-IM | Target-data specific IM |
| | IC-IM | Imaging-condition specific IM |
| Evaluation measure of segmentation results | DC | Dice Coefficient |
| | JI | Jaccard Index |
| | ASD | Average Symmetric Surface Distance |
| Organ | R-kidney | Right kidney |
| | L-kidney | Left kidney |
| | GB | Gallbladder |
| | IVC | Inferior Vena Cava |

**Table 2**

Commonly used notations.

| Symbols | Description |
|---|---|
| $I(\mathbf{x})$ | CT image |
| $\mathbf{x}$ | 3D coordinates |
| $k$ | Organ index |
| $A_k$ | Set of shape–location and intensity priors of organ $k$ |
| $P_k(\mathbf{x})$ | Probabilistic atlas (PA) of organ $k$ |
| $\mathbf{s}_k(\mathbf{b})$ | Statistical shape model (SSM) of organ $k$ |
| $H_k(I)$ | Intensity model (IM) of organ $k$ |
| $H_{Ok}(I)$ | Normalized histogram of organ $k$ |
| $H_{Bk}(I)$ | Normalized histogram of background of organ $k$ |
| $\mathbf{b}$ | Shape parameter vector |
| $Seg(\cdot)$ | Segmentation operation |
| $\mathbf{s}_0$ | Surface shape vector (concatenation of 3D coordinates) of target organ |
| $\mathbf{s}_k$ | Surface shape vector of predictor organ $k$ |
| $\tilde{\mathbf{s}_i}$ | Predictor vectors (set of surface shape vectors of predictor organs) of case $i$ |
| $T_p$ | Training datasets of predictor organs for partial least squares regression (PLSR) |
| $T_t$ | Training datasets of target organ for PLSR |
| $PLSR(\tilde{\mathbf{s}}; T_p, T_t)$ | PLSR model |
| $\mathbf{r}_i$ | Residual of case $i$ after PLSR prediction, which is the difference between predicted and true shapes |

**Table 3**

Organ classification. Type 1 organs are segmented unconditionally and are always predictor organs. Type 2 organs are segmented conditionally and can be predictor organs. Type 3 organs are segmented conditionally and cannot be predictor organs.

| Type | Set of organs | Role |
|---|---|---|
| Type 1 | $V_1 = \{liver_1, liver_2, liver_3, liver_4\}$ | predictor |
| Type 2 | $V_2 = \{spleen, R\text{-}kidney, L\text{-}kidney\}$ | predictor and target |
| Type 3 | $V_3 = \{pancreas, GB, aorta, IVC\}$ | target |

**Table 4**

Details of CT scans used in experiments.

| Dataset | | Institution | Phase | # of cases | Slice thickness [mm] |
|---|---|---|---|---|---|
| A | A-1 | Osaka Univ. Hospital (OUH) | Early arterial phase | 10 | 2.5 |
| | A-2 | | | 38* | 0.625 |
| B | B-1 | | Late arterial phase | 10 | 2.5 |
| | B-2 | | | 39* | 0.625 |
| C | | NIH | Portal venous phase | 25 | 0.68 – 1.25 |
| D | | | Non-contrast | 12 | 1.0 |

*: Only late arterial phase was available in one case.

**Table 5**

Average DCs of two groups of datasets whose accuracies of liver segmentation were below and above the lower quartile (LQ).

| | Liver segmentation accuracy | Spleen | R-kidney | L-kidney | Pancreas | GB | Aorta | IVC |
|---|---|---|---|---|---|---|---|---|
| DC [%] | Below the LQ (Low accuracy) | 91.8 | 84.5 | 91.9 | 69.2 | 52.9 | 87.7 | 58.0 |
| | Above the LQ (Medium and high accuracy) | 93.7 | 95.5 | 94.5 | 74.7 | 72.2 | 87.8 | 67.7 |