Paediatric Liver Segmentation for Low-Contrast CT Images

Mariusz Bajger[1,2], Gobert Lee[1,2(✉)], and Martin Caon[1,3]

[1] Medical Device Research Institute, Flinders University, Adelaide, Australia
{Mariusz.Bajger,Gobert.Lee,Martin.Caon}@flinders.edu.au
[2] College of Science and Engineering, Flinders University, Adelaide, Australia
[3] College of Nursing and Health Sciences, Flinders University, Adelaide, Australia

Abstract. CT images from combined PET-CT scanners are of low contrast. Automatic organ segmentation on these images are challenging. This paper proposed an adaptive kernel-based Statistical Region Merging (SRM) algorithm for paediatric liver segmentation in low contrast PET-CT images. The results are compared to that from the original SRM. The average dice index is 0.79 for SRM and 0.85 for the adaptive kernel-based SRM. In addition, the proposed method was successful in segmenting all 37 CT images while SRM failed in 5 images.

Keywords: Low contrast CT · PET-CT · Adaptive-kernel

1 Introduction

Children are sensitive to radiation dose. The use of ionizing radiation such as x-ray on children needs extra care. In some settings, segmenting of the liver from x-ray images may be required. Conventional CT images from dedicated CT scanners are typically involved due to high image quality but radiation dose is a concern. Combined PET-CT scanners generate CT images with lower radiation dose but the image has lower contrast. Automatic segmentation of organs from these low contrast CT images is challenging. A review of liver segmentation techniques can be found in [1,2].

A probabilistic atlas has been widely used for liver segmentation [3–6] and has produced good segmentation outcomes. In [3] Linguraru et al. have investigated the use of a probabilistic atlas for liver segmentation in low-contrast CT images. Based on a 20 patients' datasets (10 for training and 10 for testing) a Dice index of 88.2 ± 3.7 was achieved. The use of probabilistic atlas information can be augmented with other information. Li et al. [6] supplements the probabilistic atlas with primary liver shape and localization obtained from the PET scans. The probabilistic atlas was built using 60 CT studies from dedicated CT scanners. The PET-guided probabilistic atlas approach was applied on 35 PET-CT studies with a volume overlap percentage (VOP) of $92.9\% \pm 2.1$. In this approach, larger number of training data lead to improved segmentation results [3,6,7].

© Springer Nature Switzerland AG 2018
A. Melbourne and R. Licandro et al. (Eds.): DATRA/PIPPI 2018, LNCS 11076, pp. 169–178, 2018.
https://doi.org/10.1007/978-3-030-00807-9_17

In more recent studies, statistical shape model (SSM) has attracted a lot of interest [8–13]. A review of the statistical shape models for 3D medical image segmentation can be found in [14]. The results are promising but they also require large datasets for the SSM. For instance, in [13], 120 cases were used to develop the SSM. Annotation of the liver in the large training set is laborious.

The Statistical Region Merging (SRM) technique [15] is founded on Probability and Statistical theory and has been proposed for natural scene image segmentaion. The technique merges pixels into statistically homogenous regions (superpixels) to be regrouped into target objects/organs. Lee et al. [16] employed the SRM method [15] for multi-organ segmentation on non-contrast CT images. The technique has also been extended to 3D-SRM [17] for the spatial connectivity of volume CT data. Medical image segmentation based on the SRM method does not require large dataset for developing probabilistic atlas or statistical shape model. It does employ a prior knowledge of shape and location but primary segmentation of the liver on PET scans such as in [10] or the use of a simple model [18] suffices. In this paper, an adaptive kernel-based SRM (*kernel*-SRM) is proposed for segmentation of low contrast CT images. The method uses a kernel regressor and employs regional statistics. Results are compared with that of the SRM method.

2 Proposed Method

2.1 Adaptive *Kernel*-SRM

Consider a gray level intensity image of size $M \times N$

$$I : \{1, 2, \ldots, M\} \times \{1, 2, \ldots N\} \rightarrow [0, 255)$$

where $I(m, n) = f(m, n) + \epsilon$, with f being the true intensity value and ϵ the noise. The task is to estimate the unknown function f. In 1964 Nadaraya [19] (and also Watson [20]) proposed the following non-parametric estimator of the regression function

$$f(x, y) = E(I(X, Y)|(X, Y) = (x, y)), \tag{1}$$

where E denotes (conditional) expectation, and $((X, Y), I(X, Y))$ is the observed couple of random variables.

In order to estimate the regression function in Eq. (1), a non-parametric kernel-based estimator of Nadaraya-Watson type, which combines estimation and smoothing of the regression function, is commonly used. In this study, we consider a local version of the estimator defined for a given region in the image. The Nadaraya-Watson *local estimator* of the regression function $f(m, n)$, for a given region $\mathcal{R}$, can be defined as

$$\widehat{f}(m, n) = \sum_{(m_i, n_i) \in \mathcal{R}} w_i I(m_i, n_i), \tag{2}$$

$$w_i = \frac{K\left(I_i^{(m,n)}\right)}{\sum_{(m_i,n_i)\in\mathcal{R}} K\left(I_i^{(m,n)}\right)}, \tag{3}$$

$$I_i^{(m,n)} = \frac{I(m,n) - I(m_i,n_i)}{h_I}, \tag{4}$$

where $(m_i, n_i) \in M \times N$, K is a kernel function and h_I is the smoothing parameter for a given image I. Observe that (3) gives a weighted contribution from $I(m_i, n_i)$ to the estimated (true) value at (m_i, n_i).

As an example, consider the normal distribution $N_{\mu,\sigma}(m,n)$ with mean μ and standard deviation σ

$$N_{\mu,\sigma}(m,n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{I(m,n)-\mu}{\sigma}\right)^2\right). \tag{5}$$

The kernel is then defined as

$$K(I_i^{(m,n)}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{|I_i^{(m,n)}|-\mu}{\sigma}\right)^2\right) \tag{6}$$

where $|.|$ denotes absolute value.

From the definition of weights (3), one can observe that for a given image pixel the pixel's estimated (new) value will be most influenced by those local pixels whose intensities differ from the given one by the expected value of the intensity across the region. In other words, pixels having intensities very different from the given pixel intensity (with difference significantly bigger or smaller than the average intensity of the region) - noise pixels - will have little or no impact on the estimated intensity value (value of the regression function). Hence, assuming that the noise pixels have the distribution following the kernel function distribution the formula (2) can be effectively used to reduce noise in the image.

Using the notation in [15], the Statistical Region Merging (SRM) method allows merging of two regions R, R' if

$$|\bar{R} - \bar{R}'| \leq \sqrt{b^2(R) + b^2(R')} \tag{7}$$

where

$$b(R) = g\sqrt{\frac{1}{2Q|R|}\ln\frac{2}{\delta}}, \tag{8}$$

$|.|$ denotes cardinality of a set, $\bar{R}$ denotes the average intensity across the region R, Q is a parameter which controls coarseness of the segmentation, $\delta = \frac{1}{6\,|I|^2}$ and g is the number of image intensity levels.

By incorporating an appropriate kernel function into the regional expectation $\bar{R}$ one can alleviate the effect of noise. Each time two pixels are considered for merging Eq. (2) is used to modify intensities in spatial neighbourhoods of these pixels. The radius for this neighbourhood was fixed to 2 pixels. This proposed method shall be called Adaptive Kernel-Based Statistical Region Merging method, and abbreviated as "*kernel*-SRM".

2.2 Determination of the Kernel Function

As described in the previous section, in order to successfully alleviate noise the kernel function corresponding to the noise distribution has to be determined. This is achieved by defining a structure-free region outside the human body on the CT image for image noise estimation. The primary variation of the image intensity in this region is due to noise. The histograms are built using the long-standing Sturges' rule [21] to estimate the number of bins $k = 1 + \log_2(n)$, where n is the number of data points.

The best fit probability density function was selected from the following range of distributions: Rayleigh, normal, Poisson, gamma and the generalized extreme value (GEV) distributions. In determining the most appropriate distribution, the Mean Square Error (MSE) was calculated for each fit and the distribution with the smallest MSE value was selected as the best fit.

3 Experiments and Results

3.1 Data

Thirty-seven paediatric liver CT images acquired from combined PET-CT scanner were included in this retrospective study. The images were de-identified and were obtained from a hospital in Sydney, Australia with ethics approval. The Siemens Emotion Duo scanner was used in acquiring the images with pixel size 0.98×0.98 mm and slice separation 0.34 mm. Ground truths of the liver regions were delineated by an expert in human anatomy and physiology (MC). CT images acquired from combined PET-CT scanners are of low image contrast and high noise level when compare with CT images acquired from dedicated CT scanners.

3.2 *Kernel*-SRM Segmentation

Image Pre-processing. Adaptive kernel-based segmentation is a computationally demanding process. To decrease the processing time, each CT image was automatically cropped to the area of the patient body. The CT images were then subsampled by 2 using the nearest-neighbour method to further reduce computational time.

Kernel Function/Image Noise Distribution Estimation. The kernel function (Sect. 2.2) of each CT image was estimated by automatically analysing the noise distribution in that image. In each image, the region comprising the top 120 rows of pixels across the full-width of the image is designated for noise distribution estimation. This region is outside the human body and is structure-free (anatomy-free). The histogram of the pixel intensities in this region was analysed (Sect. 2.2). Table 1 shows that in all but one case, the noise distribution was best estimated using normal distribution. For the remaining single case, the image

Table 1. Determination of the kernel function/image noise distribution estimation. The distributions Rayleigh, normal, Poisson, gamma and GEV were considered for the CT image noise estimation. The best fit with the smallest Mean Square Error (MSE) for individual CT are shown in below.

Index	Type	Param 1	Param 2	Param 3	Index	Type	Param 1	Param 2	Param 3
1	normal	25.01	9.60	–	20	normal	25.16	8.94	–
2	normal	24.47	9.96	–	21	normal	24.61	8.58	–
3	normal	25.10	9.72	–	22	normal	24.18	8.48	–
4	gev	−0.19	8.80	19.92	23	normal	24.58	8.98	–
5	normal	25.14	9.89	–	24	normal	24.73	8.57	–
6	normal	24.09	9.70	–	25	normal	24.66	8.93	–
7	normal	24.38	9.87	–	26	normal	24.09	8.50	–
8	normal	25.11	9.22	–	27	normal	24.96	8.04	–
9	normal	24.13	9.33	–	28	normal	24.38	8.35	–
10	normal	24.67	9.66	–	29	normal	23.94	7.82	–
11	normal	25.06	9.26	–	30	normal	25.06	8.11	–
12	normal	24.36	9.24	–	31	normal	24.52	8.89	–
13	normal	24.71	9.17	–	32	normal	24.10	8.53	–
14	normal	23.00	8.63	–	33	normal	24.99	8.45	–
15	normal	24.75	8.98	–	34	normal	24.58	8.36	–
16	normal	24.29	8.49	–	35	normal	24.23	8.37	–
17	normal	24.87	8.51	–	36	normal	24.63	8.87	–
18	normal	24.99	9.23	–	37	normal	24.90	8.39	–
19	normal	24.23	8.74	–					

noise has a generalized extreme value (GEV) distribution. In the proposed kernel-SRM method, the kernel function for each CT image was determined based on the estimated noise distribution of that image. Further analysis of these noise distributions shows that, for the normal distributions, the average of the mean parameter is 24.6 (std 0.45; range 23–25.2) and the average of the standard deviation parameter is 8.9 (std 0.6; range 7.8–9.9). One image has the generalized extreme value distribution with parameter $(\mu, \sigma, \xi) = (-0.19, 8.80, 19.92)$. Guided by these results, the normal (Gaussian) distribution with $\mu = 24$ and $\sigma = 9$ was selected as the kernel function.

Kernel Bandwidth Optimizing. The smoothing parameter h_I (Eq. 4), also known as the kernel bandwidth, was determined experimentally by searching over a wide range $[1, 30]$. Table 2 shows that the best segmentation result is achieved for $h_I = 3$. In addition, for $h_I \in \{2, 3, 4, 5\}$, the segmentation outcomes are similarly good with the average Dice index and average Hausdorff distance (in pixels) over all CT images being (0.84, 0.85, 0.84, 0.84) and (10.04, 8.77, 9.98, 10.48), respectively. This suggests that the performance of the proposed kernel-SRM is robust to small changes in the parameter h_I. The last column

Table 2. Optimization of the kernel bandwidth h_I. The parameter h_I was searched over the range [0.6, 30]. For each value of h_I, the average Dice index and the average Hausdorf value over all 37 kernel-SRM segmented livers are shown. The last column shows the number of CT images in which the kernel-SRM failed to segment the liver. The value of h_I producing the best results is boxed.

h_I value	Ave Dice	Ave Hausdorff	Failures
0.6	0.79	16.20	2
1	0.83	12.49	–
2	0.84	10.04	–
$\boxed{3}$	0.85	8.77	–
4	0.84	9.98	–
5	0.84	10.48	–
6	0.83	11.57	–
7	0.84	11.93	1
8	0.84	11.85	1
9	0.84	11.81	1
10	0.84	12.15	1
15	0.82	12.53	1
20	0.82	12.75	1
30	0.81	14.86	2

in Table 2 reports the number of CT images in which the proposed kernel-SRM method failed to generate a segmentation of the liver (Sect. 3.3).

3.3 Segmentation Representation and Evaluation Measures

As in the Statistical Region Merging (SRM) method in [15], through the employment of an appropriate value of the parameter Q (Eq. 8), the *kernel*-SRM was set to over-segmented the CT images, thereby, partitioned the images into statistically homogeneous regions (superpixels). These superpixels are non-overlapping. The union of these superpixels gives the exact image.

Over-Segmentation and Eligible Superpixels. In a perfect segmentation, the segmented liver would be ideally represented by a single superpixel. However, anatomical structures on CT images are not homogeneous. As such, the representation of the liver (target organ/tissue) is relaxed such that the segmented liver is represented by the union of one or more statistically homogeneous regions (superpixels). For a superpixel to be included in the segmented liver, over 50% of the superpixel must overlap with the ground truth which is unknown but can be estimated using different approaches such as the model-based approach [18]. In order to evaluate the performance of the proposed *kernel*-SRM against that

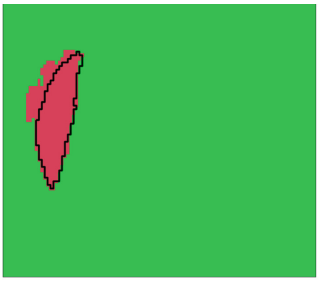
of the SRM method without the interaction with ground truth estimated, the ground truth is used in this experiment. A superpixel satisfying this condition of >50% overlap is called an 'eligible' superpixel. Thus, the *kernel*-SRM/SRM segmented liver is the union of these 'eligible' superpixels. Though the number of the eligible superpixels in the segmented liver does not directly associate with the accuracy of the segmentation, when two segmentation outcomes are of similar accuracies, the one with a smaller number of eligible superpixels is of lower complexity and is a preferred solution. Figures 1a and b show two examples of the *kernel*-SRM segmentation outcomes. The ground truth was outlined in black and the eligible superpixel(s) that contributed to the *kernel*-SRM segmentation results are shown in patch(es) of color (false color for visualization). In the first example (Fig. 1a), only one eligible superpixel with a major (>50%) of the superpixel in the ground truth was found. In the second example (Fig. 1b), three eligible superpixels were found. Though two of them are small, over 50% of each the superpixel is in the ground truth. The *kernel*-SRM segmented liver is the union of these superpixels.

Failure - Resulting in Empty Set. If the liver is presented on a CT image but no eligible superpixels was found, this means that the liver segmentation on that image failed. The generated segmentation of liver is an empty set and no segmentation result was presented. Figures 1e and f show the SRM segmentation outcomes in two examples. The SRM statistically homogeneous regions (superpixels) were presented in false color for visualization. In Fig. 1e, less than 50% of the green superpixel overlap with the ground truth whereas in Fig. 1f, less than 50% of the purple superpixel overlap with the ground truth. Thus, the segmentation outputs are empty sets in both examples.

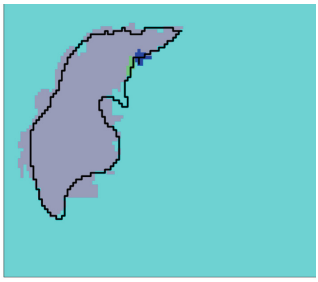
Segmentation Accuracy - Dice Index and Hausdorff Distance. For segmentation accuracy, Dice index and Hausdorff distance are measured on the segmentation outcome (union of the eligible superpixels), if eligible superpixel(s) is/are found. Dice index measures the agreement between the machine segmentation and the ground truth whereas Hausdorff distance measures the largest deviation between the two. If a segmentation outcome has no eligible superpixel, i.e. the segmentation failed, it follows that Dice index = 0 and the Hausdorff distance cannot be calculated.

3.4 Results and Discussions

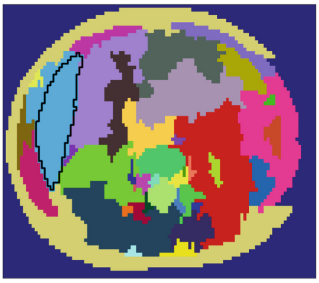
Segmentation results of the proposed *kernel*-SRM and that of the original SRM are compared in this section. The results are generated with the kernel function a normal distribution with mean $\mu = 24$ and standard deviation $\sigma = 9$ (Sect. 3.2), an optimal bandwidth $h_I = 3$ (Sect. 3.2), and a g value of $g = 256$ for 256 grayscale level images and a Q value (Eq. 8) of $Q = 256$ determined empirically. Table 3 shows the average Dice index, average Hausdorff distance and the number of failures (Sect. 3.3). *Kernel*-SRM performs better than the SRM. The average



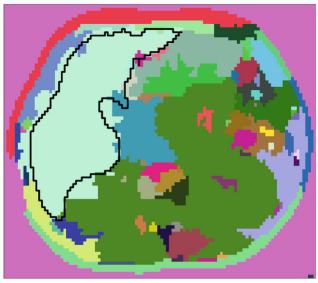
(a) Example 1: 1 eligible superpixel found



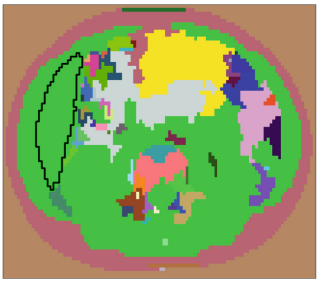
(b) Example 2: 3 eligible superpixels found



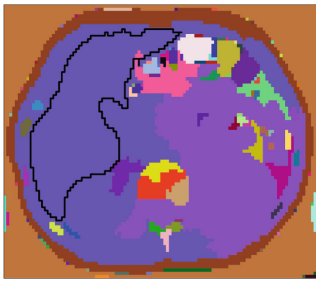
(c) Example 1:*Kernel*-SRM (proposed)



(d) Example 2:*Kernel*-SRM (proposed)



(e) Example 1:SRM



(f) Example 2:SRM

Fig. 1. Two Examples. The ground truth is delineation in black in all panels. Example 1 is presented in (a, c, e) and example 2 is presented in (b, d, f). Using the *kernel*-SRM, (a) 1 eligible superpixel (red) overlap with the liver was found in Example 1 and (b) 3 eligible superpixels (grey, green and dark blue) were found. The green and dark blue superpixels are small but over 50% is inside the ground truth, making them eligible. (c, d) *kernel*-SRM segmentation. The Dice index for (c) Example 1 is 0.88 and that for (d) Example 2 is 0.83. (e, f) SRM segmentation failed to produce any eligible superpixels in both examples as less than 50% of the green superpixel in example 1 and the purple superpixel in example 2 are inside the ground truth. Dice = 0 for both examples. (False color for visualization) (Color figure online)

Table 3. Comparison of SRM and *kernel*-SRM segmentation results.

	Ave Dice	Ave Hausdorff (in pixel)	Failure
SRM [15]	0.79	19.06[a]	14%(5/37)
Kernel-SRM (proposed)	0.85	8.77	0%(0/37)

[a]Excluding the 5 failure cases that Hausdorff distance cannot be calculated

Dice index and average Hausdorff distance over all 37 CTs were 0.79 and 19.06 for SRM and 0.85 and 8.77 for *kernel*-SRM, respectively. Moreover, SRM failed to segment (produced empty sets) in 5 cases while *kernel*-SRM was successful (no empty sets) in segmenting all images. Figures 2 and 3 show the detail of SRM and *kernel*-SRM comparisons in Dice index and Hausdorff distance, respectively. The five SRM failed examples are identified with Dice index = 0 in Fig. 2. Hausdorff distance cannot be calculated for the 5 failures and are shown with infinite lines in Fig. 3. Both Figs. 2 and 3 show that the *kernel*-SRM performs better in almost all cases.

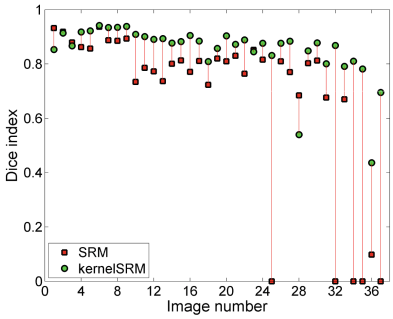


Fig. 2. Dice index - SRM vs. *kernel*-SRM liver segmentation results.

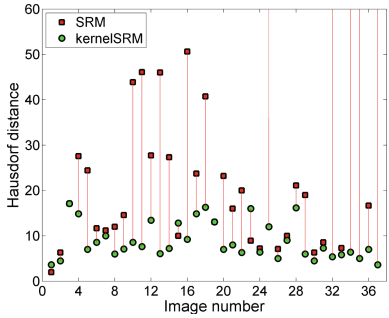


Fig. 3. Hausdorff distance (in pixel)- SRM vs. *kernel*-SRM results.

4 Conclusion

Segmentation of abdominal organs in low image contrast CT images generated from combined PET-CT scanners is challenging. This paper extended the well founded statistical region merging (SRM) method with a built-in kernel that handles the high level of image noise adaptively for every pair of regions to be considered for merging. Results showed that the proposed adaptive kernel-based statistical region merging (*kernel*-SRM) performs significantly better when compared with the original SRM method. The results, however, were found using a small dataset. Future work in validating the results with a larger dataset is required.

References

1. Gotra, A., Sivakumaran, L., et al.: Liver segmentation: indications, techniques and future directions. Insights Imag. **8**, 377–392 (2017)
2. Moghbel, M., Mashohor, S., et al.: Review of liver segmentation and computer assisted detection/diagnosis methods in computed tomography. Artif. Intell. Rev. (2017). https://doi.org/10.1007/s10462-017-9550-x
3. Linguraru, M.G., Li, Z., et al.: Automated liver segmentation using a normalized probabilistic atlas. In: Proceedings of SPIE, Medical Imaging, vol. 7262 (2009)
4. Linguraru, M.G., Sandberg, J.A., et al.: Atlas-based automated segmentation of spleen and liver using adaptive enhancement estimation. Med. Phys. **37**(2), 771–783 (2010)
5. Linguraru, M.G., Pura, J.A., et al.: Statistical 4D graphs for multi-organ abdominal segmentation from multiphase CT. Med. Image Anal. **16**, 904–914 (2012)
6. Li, C., Wang, X., et al.: Automated PET-guided liver segmentation from low-contrast CT volumes using probabilistic atlas. CMPB **107**, 164–174 (2012)
7. Zhou, X., et al.: Constructing a probabilistic model for automated liver region segmentation using non-contrast X-ray torso CT images. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 856–863. Springer, Heidelberg (2006). https://doi.org/10.1007/11866763_105
8. Okada, T., Linguraru, M.G., et al.: Abdominal multi-organ segmentation from CT images using conditional shape-location and unsupervised intensity priors. Med. Image Anal. **26**, 1–18 (2015)
9. Saito, A., Nawano, S., Shimizu, A.: Joint optimization of segmentation and shape prior from level-set-based statistical shape model, and its application to the automated segmentation of abdominal organs. Med. Image Anal. **28**, 46–65 (2016)
10. Li, G., Chen, X., et al.: Automatic liver segmentation based on shape constraints and deformable graph cut in CT images. IEEE TIP **14**(12), 5315–5329 (2015)
11. Shi, C., Cheng, Y., et al.: A hierarchical local region-based sparse shape composition for liver segmentation in CT scans. Pattern Recognit. **50**, 88–106 (2016)
12. Wang, X., Zheng, Y., et al.: Liver segmentation from CT images using a sparse priori statistical shape model (SP-SSM). PLoS ONE **12**(10), e0185249 (2017)
13. Tomoshige, S., Oost, E., et al.: A conditional statistical shape model with integrated error estimation of the conditions: application to live segmentation in non-contrast CT images. Med. Image Anal. **18**, 130–142 (2014)
14. Heimann, T., Meinzer, H.-P.: Statistical shape models for 3D medical image segmentation: a review. Med. Image Anal. **13**, 543–563 (2009)
15. Nock, R., Nielsen, F.: Statistical region merging. IEEE Trans. Pattern Anal. Mach. Intell. **26**(11), 1452–1458 (2004)
16. Lee, G., Bajger, M., Caon, M.: Multi-organ segmentation of CT images using statistical region merging. In: Proceedings of the 9th IASTED International Conference on Biomedical Engineering, pp. 199–206 (2012)
17. Bajger, M., Lee, G., Caon, M.: 3D segmentation for multi-organs in CT images. Electron. Lett. Comput. Vis. Image Anal. **12**(2), 13–27 (2013)
18. Sedlar, J., Bajger, M., et al.: Model-guided segmentation of liver in CT and PET-CT images of child patients based on statistical region merging. In: Proceedings of the DICTA, pp. 156–163 (2016)
19. Nadaraya, E.A.: On estimating regression. Theory Probab. Appl. **10**, 186–190 (1964)
20. Watson, G.S.: Smooth regression analysis. Sankhya Ser. A **26**, 101–116 (1964)
21. Sturges, H.A.: The choice of a class interval. J. Am. Stat. Assoc. **21**, 65–66 (1926)