Int J CARS (2011) 6:13–20
DOI 10.1007/s11548-010-0493-9

ORIGINAL ARTICLE

Liver segmentation for contrast-enhanced MR images using partitioned probabilistic model

László Ruskó · György Bekes

Received: 8 January 2010 / Accepted: 18 May 2010 / Published online: 11 June 2010

Abstract
Purpose Liver volume segmentation is important in computer assisted diagnosis and therapy planning of liver tumors. Manual segmentation is time-consuming, tedious and error prone, so automated methods are needed. Automatic segmentation of MR images is more challenging than for CT images, so a robust system was developed.
Methods An intensity-based segmentation method that uses probabilistic model to increase the precision of the segmentation was developed. The model was build based on 60 manually contoured liver CT exams and partitioned into 8 parts according to the (Couinaud) segmental anatomy of the liver. The partitioning allows using different intensity statistics in different parts of the organ, which makes it insensitive to local intensity differences from MR artifacts or pathology. The method employs a modality independent model with registration that exploits some LAVA image characteristics. This dependence can be eliminated to adapt the segmentation method for a wide range of MR images.
Results The method was evaluated using eight representative, manually segmented MR LAVA exams. The results show that the method can accurately segment the liver volume despite various MR artifacts and pathology. The evaluation shows that the proposed method provides more precise segmentation (6% average absolute relative volume error) compared with global intensity statistics for the whole organ (20% average absolute relative volume error). The compute time of the method was 30 s in average, which is acceptable for wide range of clinical applications.
Conclusion An automatic method that can segment the liver in contrast-enhanced MR LAVA images was developed and tested. The results demonstrate that the method is feasible, efficient and robust to artifacts and pathology.

Keywords Segmentation · Probabilistic model · Liver · MR

L. Ruskó (✉) · G. Bekes
GE Hungary, Budapest, Hungary
e-mail: laszlo.rusko@ge.com

Purpose

Segmentation of anatomical structures on 3D medical images is very important for several clinical applications. It can facilitate diagnosis, therapy planning and monitoring. Liver-related therapy planning is based on the liver volume, the different vessel trees, the anatomical segments and the relation of lesions to these structures. To get this information, 3D imaging modalities are used such as computed tomography (CT) or magnetic resonance imaging (MRI). MRI as well as CT images are usually enhanced with contrast agent to better visualize lesions and vasculature. The automated segmentation of contrast-enhanced MR images is very challenging for the following reasons. First, the liver is surrounded by organs, whose image signal intensity can be similar to that of the liver. Secondly, the liver parenchyma is heterogeneous due to various contrast uptake of different lesions and the artifacts of the MR acquisition. Our goal was to develop a segmentation method for liver MR images that is both quick and precise, so it can be used in wide range of clinical applications.

Methods

Several methods have been published for liver segmentation on CT images. These methods are some variants of level-set [1], region-growing [2], or deformable shape [3] algorithms, which can be extended with shape [4], probabilistic [5], or

intensity [6] model. Some methods [3,4,6] use shape models, which are extended with local intensity or shape characteristics. These methods can handle intensity variations of the liver in efficient way but still rely on CT characteristics (liver intensity is always in a well-defined range), which makes them difficult to adapt to MR modality. Moreover, using complex shape models increases the running time significantly. In order to segment liver volume, Furukawa et al. [1] and Rikxoort et al. [5] apply probabilistic model that is used to facilitate a level-set and a voxel classification method, respectively. Although the applied probabilistic model is modality independent, these methods cannot efficiently handle large intensity variations characteristic for MR liver exams. A low-computation demanding method is presented in [2]. This method uses region-growing and incorporates anatomical knowledge to eliminate under- and over-segmentation. Although the algorithm is fast, it can under-segment large lesions in the liver. It uses global intensity statistics for the entire organ, therefore its concept cannot be applied to MR images without any change. Li et al. [7] describe an automatic approach, which combines level-set and watershed method. According to their results, the algorithm provides good segmentation, but there are no statistics presented about precision and running time. Farraher et al. [8] segment the liver on MR images by clustering. This technique is semi-automatic (the user is required to inspect the segmentation results and modify parameters if needed) and time-consuming (the running time is above 5 min). Hermoye et al. [9] compare manual contouring with 2D semiautomatic deformable model technique. The semi-automatic segmentation is performed slice-by-slice started from a user-positioned circle, which is deformed to the organ's boundary. It takes about 5 min and multiple user interactions, which introduces significant inter-operator variability.

From the results cited above, one can see that there is still need for an automated method that can quickly and precisely segment the liver volume on MR images. This fact motivated us to develop a method that integrates the advantages of the intensity-based, model-based and local statistics-based methods.

Model construction

In this subsection, we present the construction of the probabilistic liver model that incorporates information about the location of the anatomical segments, hence allows segmenting different parts of the liver using different intensity statistics. The model consists of two main parts: the probability map of liver voxels and the anatomical segmentation of the liver.

The probabilistic model was built from 60 manually contoured liver volumes. The manual segmentation was performed on CT exams by radiologist. As a normalizing step, the liver volumes were resampled to the same isotropic voxel size (2.5 mm). Then, they were registered to a previously selected reference liver using 3D similarity transformation. This transform allows translation, scaling, and rotation. Due to the volume-based registration method the liver, whose volume (1,520 ml) was the closest to the average volume (1,525 ml) of 60 cases, was selected as reference. In the registration process, the mean square difference of the reference and the moving image was minimized using gradient descent optimizer (implemented in ITK). After registration, the resulted transform was applied to each liver volume and the probability map was defined by the sum of all transformed liver volumes. Finally, the values of the probability map were normalized to the range [0,1].

The key point of the model construction is to extend the probabilistic model with anatomical segments. The Couinaud segmentation [10] divides the liver into 8 independent segments. Each segment has its own vascular inflow, outflow, and biliary drainage. This partition makes possible to divide the liver into anatomically meaningful parts and use different statistics in different parts of the liver. A simplified definition of Couinaud segments uses 5 planes to separate the 8 segments. These planes are fitted to the 3 main branches of the hepatic vein and to the 2 main branches of the portal vein. In our model construction, the segment separating planes were defined for the reference volume incorporating the underlying CT image (to visualize the vessel structures). Using these planes, the segment separation was applied to the average liver shape. In order to get the average liver shape, the probability map was thresholded at 30%. The threshold level was defined such that the volume of the average liver shape is the closest to 1,525 ml.

In summary, the model construction process results in a 3-dimensional volume that assigns 2 values to each voxel: the probability of the voxel to belong to the liver and the label of the segment that involves the voxel. Figure 1 shows the partitioned average liver shape, an axial and a coronal slice of the partitioned liver probability map.

Segmentation

The following part describes how the liver model is registered to the input image and how the image is segmented incorporating the registered model. In this work the portal-venous phase of the gadolinium-enhanced MR LAVA acquisition was used as input. LAVA is a 3D FastSPGR plus fat suppression acquisition technique. This multi-phase protocol was developed for abdominal imaging and acquires only the part of the abdomen, where the liver is located (not involving chest or pelvis).

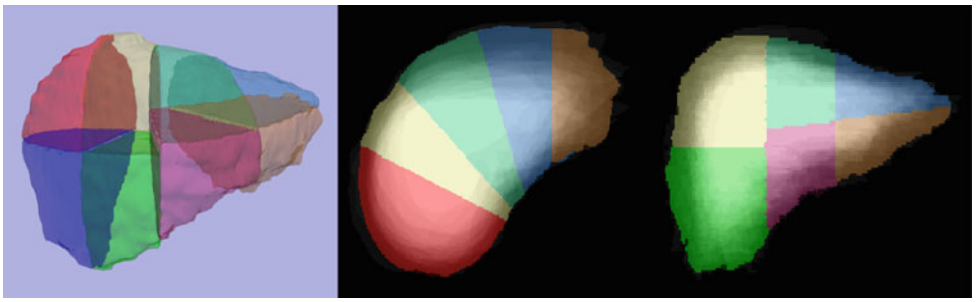

Fig. 1 Segmented liver model. The liver probability map was thresholded at 30% to get the average liver shape (*left*) that was partitioned using the separating planes defined for the reference liver. An axial (*center*) and a coronal (*right*) slice of the liver model that assigns two values to each voxel: the liver probability (*brightness*) and the segment label (*colors*)

1. The intensity range of liver tissue (referred as $[G_{\min}, G_{\max}]$) varies significantly among contrast-enhanced LAVA images (for our evaluation data set, the mean liver intensity varied between 729 and 2,954). In contrast to the CT modality, this range shall be determined in each case separately. In order to do that, the histogram of the whole image is computed. Due to the contrast enhancement and the localized acquisition, the histogram always has a well separable maximum at the mean liver intensity.

In order to find the liver maximum, the environment of all local maxima is analyzed on the smoothed histogram (h). For each local maximum position m, the nearest positions l_m and r_m in the left and the right direction are computed such that:

- $\mathrm{h}(l_m) < 0.1 \cdot \mathrm{h}(m)$ or l_m is a local minimum, and
- $\mathrm{h}(r_m) < 0.1 \cdot \mathrm{h}(m)$ or r_m is a local minimum.

After computing l_m and r_m, the area (A_m) under the histogram belonging to the range $[l_m, r_m]$ is computed (see Fig. 2). When A_m is available for each maximum position, the largest m, for which $0.05 < A_m$ is selected as mean liver intensity, while $G_{\min}$ and $G_{\max}$ are set to the values l_m and r_m, respectively. The 5% limit prevents the method to select an intensity range belonging to a large hyperdense liver lesion or any other organ.

2. In order to register the probability model to the image to be segmented we do the following. In the first step, the input image is resampled using 2.5 mm isotropic voxel size. Then, the input image is thresholded using $[G_{\min}, G_{\max}]$ and the left-posterior quarter of the image (where the spleen is located) is erased (Fig. 3). For all nonzero voxels of the threshold image, the 3D distance from the nearest contour is computed, which results in a 3-dimensional distance map (Fig. 3). Since the liver has the largest compact volume in the abdominal image, the largest values on the distance map represent the central part of the liver. Note that the probability map has the same characteristics, which makes it is possible to register the liver model with the distance map (Fig. 3). The registration computes a 3D similarity transform by minimizing the square difference between the probabilistic model and the distance map, which is also normalized to the range [0, 1]. The registration transform is initialized with scaling equal to 1, with rotation equal to the identity and with translation equal to the difference of the weighted center of gravity belonging to the distance map and the probability map. After registration the average liver shape as well as the segment information can be placed onto the image. In order to have segment information for all possible liver voxels (compensate for registration errors), 3-dimensional dilation is applied to the segments (Fig. 3).

3. Region-growing method is used to segment the liver parenchyma. In order to initialize this method, we need a seed region that does not involve any vessel or lesion. This region is extracted form the threshold image (created in the previous step) by performing 3D erosion. Since the liver has the largest volume in the abdomen, a sphere with large radius (15 mm) can be used for erosion. The largest 3D connected region (after erosion) is selected as seed region (Fig. 4).

The next step computes intensity statistics for the region-growing. We perform it for the seed region and for each segment separately. The intensity range of the seed region $[L_{\min}, L_{\max}]$ is computed based on the histogram of the seed region in the following way. Let L_{mod}, $L_{\mathrm{std(l)}}$ and $L_{\mathrm{std(r)}}$ denote the histogram modus, the left, and the right standard deviation of intensities, respectively. $L_{\min}$ is defined as $L_{\mathrm{mod}} - c \cdot L_{\mathrm{std}(l)}$ and $L_{\max}$ is defined as $L_{\mathrm{mod}} + c \cdot L_{\mathrm{std}(l)}$, where $c = 3$ is an empirical constant. This intensity range is narrower than $[G_{\min}, G_{\max}]$ because it does not involve vessels and hypo- or hyperdense lesions.

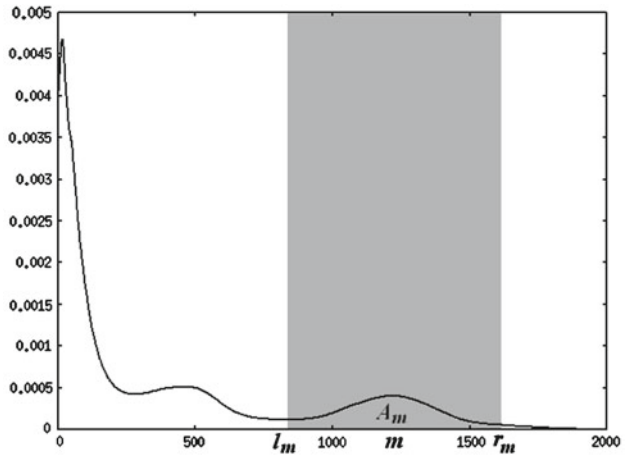

Fig. 2 Smoothed histogram of an input image. The intensity range of the liver (*gray interval*) is computed for all exams separately by analyzing each local maximum (m) of the histogram

Using the original (not-dilated) segment information, we compute the intensity range for all 8 segments in the same way as in case of the seed region. Let $[S^i_{\min},\ S^i_{\max}]$ denote the intensity range of segment $i \in [1, 8]$. Since the partitioned model is registered to the image, voxels belonging to each segment can be determined. Thus, the value of $S^i_{\min}$ and $S^i_{\max}$ can be computed based on the corresponding set of voxels. In summary, we use the following intensity statistics during the region-growing:

- Liver statistics $[L_{\min},\ L_{\max}]$: subset of the global range, represents liver parenchyma without vessels and lesions.
- Segment statistics $[S^i_{\min},\ S^i_{\max}]$: represent local intensity of segment i, including liver parenchyma, vessels and lesions (can be either narrower than $[L_{\min},\ L_{\max}]$ or wider than $[G_{\min},\ G_{\max}]$ according to the local artifacts and pathology.

The region-growing starts from the seed region. In each iteration, one voxel (j) is examined. The voxel j is added to the result, if the intensity of all voxels in its (5 mm) 3D environment is in the range $[V_{\min(j)},\ V_{\max(j)}]$ that is defined as follows:

$$V_{\min}(j) = S^{i(j)}_{\min} + (L_{\min} - S^{i(j)}_{\min}) \cdot (1 - p(j)) \quad \text{and}$$
$$V_{\max}(j) = S^{i(j)}_{\max} + (L_{\max} - S^{i(j)}_{\max}) \cdot (1 - p(j)),$$

where *i(j)* is the index of the segment containing (or located the nearest to) voxel j, and *p(j)* is the value of the probability map at voxel j. If voxel j is added to the result its 6 neighbors are also examined. This definition makes possible to balance between local (computed from segments) and global (computed from seed region) intensity ranges. The higher the *p(j)* is, the accepted intensity range is more similar to the local intensity range. This approach allows segmenting regions, where the probability of liver is high, but the intensity is very different from normal liver parenchyma (e.g. lesions). Using the model, large over-segmentation can be also eliminated because the result involves voxels with low probability only

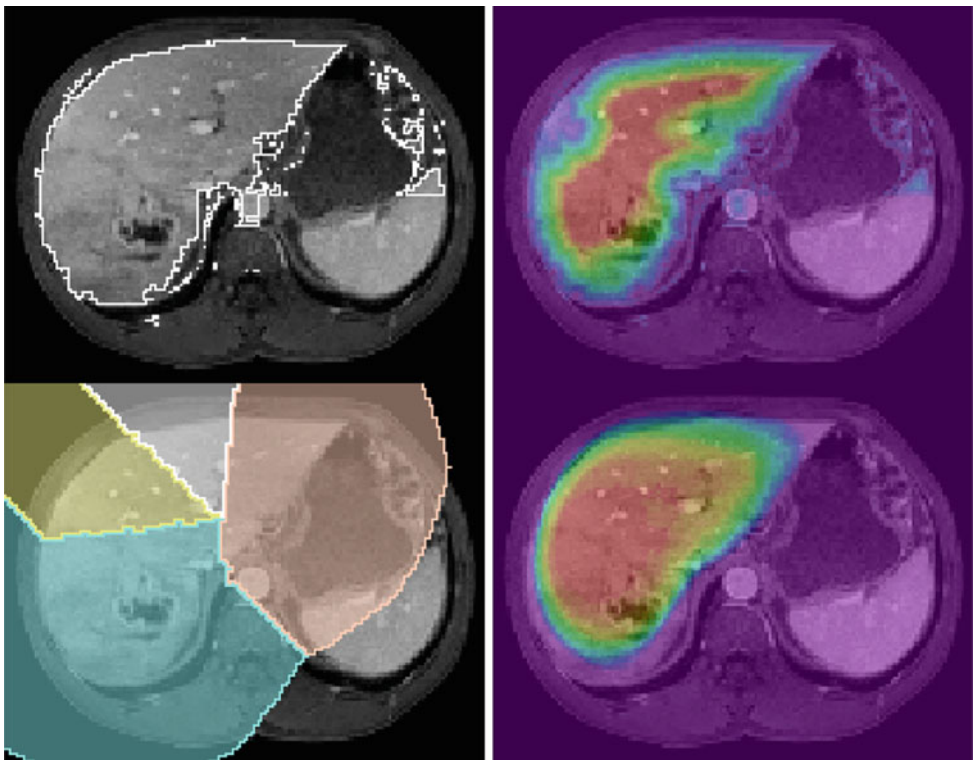

Fig. 3 Main steps of the model registration. Image is thresholded using the global liver intensity range (*top left*), distance map is computed from the thresholded image (*top right*), probability map after registering to the distance map (*bottom right*), and the dilated segment information (*bottom left*)

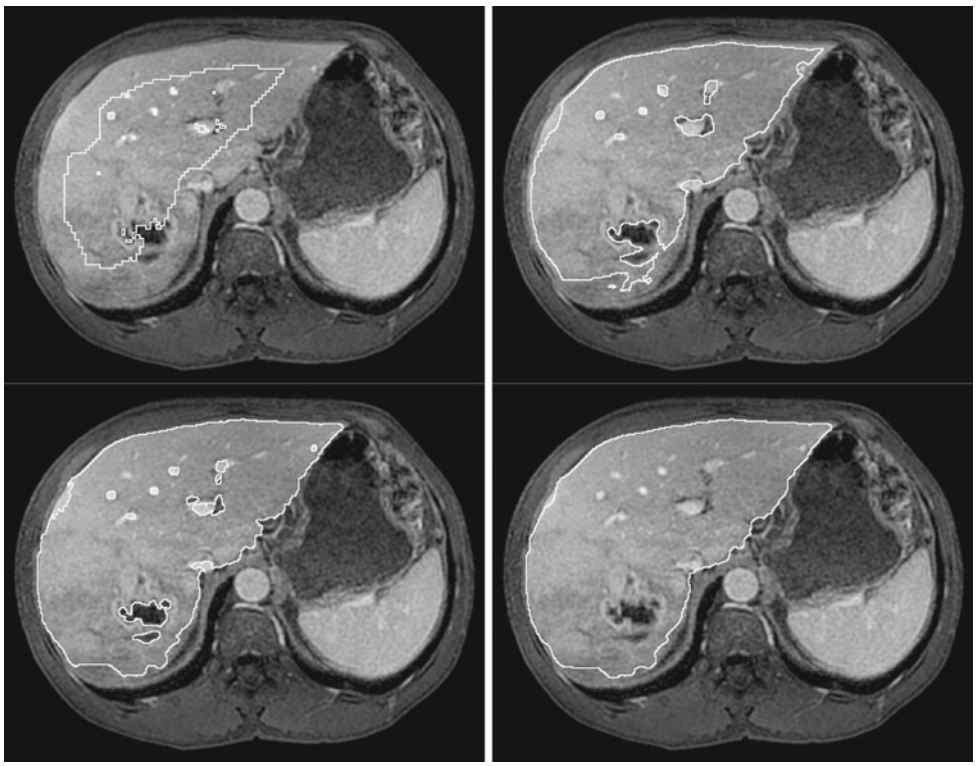

Fig. 4 Main steps of the segmentation. Initial region (*top left*), the result of the region-growing (*top right*), the result of dilation (*bottom left*), the result of the cavity filling (*bottom right*)

if their intensity is very close to the normal liver parenchyma (Fig. 4).

Since large environment (5 mm radius) is used in each iteration of the region growing, the result is under-segmented. This problem can be eliminated performing a dilation using the same kernel (Fig. 4). In addition, a 2-dimensional cavity filling is performed for each axial slice, so that vessels or lesions are included in the final result (Fig. 4).

Results

We performed the evaluation involving 8 portal-venous contrast-enhanced MR LAVA exams. The slice resolution was 512 × 512 for all cases, the slice number was 134 in average (min 91, max 176). The average pixel size was $0.82 \times 0.82\,\text{mm}^2$ ($\min 0.7 \times 0.7\,\text{mm}^2$, $\max 0.86 \times 0.86\,\text{mm}^2$), and the average slice thickness was 1.7 mm (min 1.3 mm, max 2.2 mm). The exams were selected by radiologist, so they represent typical clinical cases. The test exams (see Fig. 5) involve: (A) healthy liver with less acquisition artifacts, (B) healthy liver with acquisition artifacts at its left lobe and inferior part, (C) liver with large hypodense lesion, (D) liver with large hyperdense lesion in the posterior part of the right lobe, (E) liver with a lot of hypodense cysts, (F) liver with some hyperdense lesions in the left lobe, (G) liver with hyperdense lesions and acquisition artifacts (the border of the organ is blurred near the kidney and the stomach), (H) liver with large heterogeneous diffuse disease.

The result of the automated segmentation was compared with manual segmentation. Figure 5 shows the results of the automatic segmentation using segment specific intensity statistics. According to the images, the majority of the liver parenchyma is well segmented even if the image is affected by artifacts (A,F) or hypodense lesions (C,D,E). One can see that the results are rarely over-segmented except for (F) at the heart or (A,G) at the stomach, and most of the hypodense lesions are not under-segmented. In some cases, however, the liver is under-segmented near the boundary of the organ (G), especially in the left lobe (A,B,C), where the probability model has lower values.

In order to perform quantitative evaluation, the segmentation results were compared with manually delineated contour using the following metrics:

- True-positive rate (TPR): the true-positive volume divided by the reference volume, TPR = 100% means no under-segmentation.
- False-positive rate (FPR): the false-positive volume divided by the reference volume, FPR = 0% means no over-segmentation.
- Dice similarity coefficient (DSC): overlap of the segmented and the reference volume divided by the average of the segmented and the reference volume, DSC = 100% means no under- and over-segmentation.

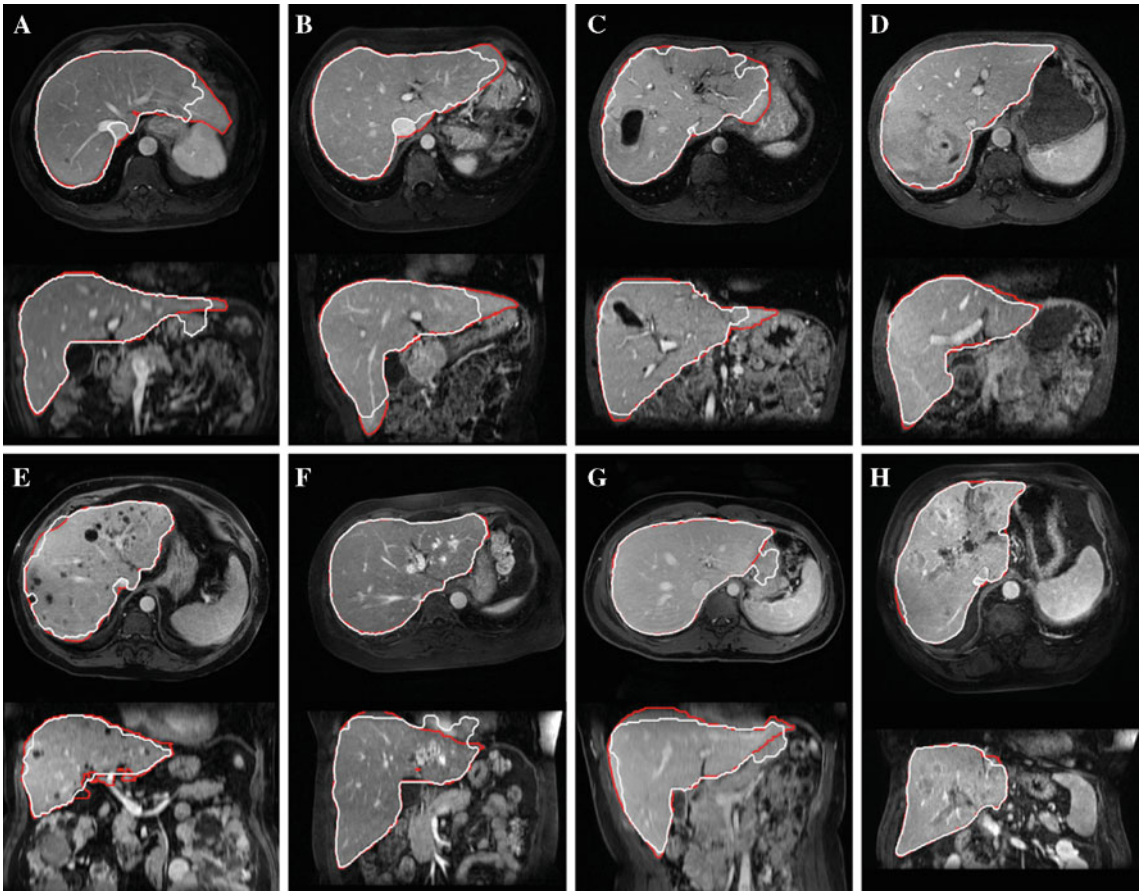

Fig. 5 Axial and coronal slices of the segmentation result for 8 test cases using local intensity statistics. *Red* contour represents the manual segmentation, *white* contour represents the result of the automatic segmentation

- Volumetric overlap error (VOE): the intersection of the segmented and the reference volume divided by the union of the segmented and the reference volume subtracted from 1, VOE = 0% means no under- and over-segmentation.
- Relative volume difference (RVD): the difference of segmented volume and the reference volume divided by the reference volume, negative RVD value means under-positive RVD value means over-segmentation.
- Average symmetric surface distance (ASSD): for each voxel along the segmented/reference border, the closest voxel along the reference/segmented border is determined, and the average of all distances is computed, ASSD = 0 mm means perfect result.
- Maximal symmetric surface distance (MSSD): for each voxel along the segmented/reference border, the closest voxel along the reference/segmented border is determined, and the maximum of all distances is computed, MSSD = 0 mm means perfect result.

The quantitative evaluation was performed for 2 different methods: the first uses local intensity statistics (as proposed), the second uses global intensity statistics (computed based on the seed region only). Tables 1 and 2 show the evaluation metrics for both methods. In case of using local statistics, the TPR and FPR values show no large over- or under-segmented regions. The average FPR is 2.6% and the average TPR is 91.2%. DSC is above 90% in all cases and the average VOE is 11.2%. The negative RVD indicates that the results are under-segmented, especially in cases B and E. The average ASSD = 2.2 mm and the average MSSD = 34 mm indicate large surface distance at some locations, as seen in cases A,B,F.

In case of using global statistics, the metrics are slightly different. VOE and RVD are higher, DSC and TPR are lower, while FPR is nearly the same in average. According to the surface metrics the average ASSD as well as the average MSSD are higher. The difference (that is statistically significant for 5 of the 7 metrics) between the average metrics show that better segmentation quality can be achieved by using local intensity statistics.

Figure 6 demonstrates two cases, where some parts of the liver are affected by typical MR artifact (signal inhomogeneity). These regions are very under-segmented, when global intensity statistics are used.

Table 1 Error metrics for the segmentation method using local (segment specific) intensity statistics

Exam	TPR	FPR	DSC	VOE	RVD	ASSD	MSSD
A	93.9	1.9	95.9	7.9	−4.2	1.6	47.6
B	86.5	0.6	92.5	14.0	−12.9	2.6	46.4
C	90.1	1.0	94.3	10.9	−8.9	2.1	26.1
D	95.3	3.7	95.7	8.2	−1.0	1.5	27.8
E	85.0	3.4	90.2	17.8	−11.6	3.1	24.6
F	94.1	4.7	94.7	10.1	−1.1	2.8	43.0
G	89.2	4.6	92.0	14.7	−6.2	3.2	32.7
H	95.3	1.1	97.1	5.7	−3.6	0.9	23.9
AVG	91.2	2.6	94.1	11.2	−6.2	2.2	34.0
SD	4.1	1.7	2.3	4.1	4.6	0.8	10.1

Table 2 Error metrics for the segmentation method using global intensity statistics

Exam	TPR	FPR	DSC	VOE	RVD	ASSD	MSSD
A	84.6	3.7	89.9	18.4	−11.7	3.0	33.1
B	66.6	0.1	79.9	33.4	−33.3	5.9	45.9
C	67.1	0.2	80.2	33.1	−32.7	6.5	48.8
D	88.8	2.2	93.0	13.1	−9.0	2.5	39.7
E	77.2	1.2	86.6	23.7	−21.6	4.1	27.6
F	82.3	0.6	90.0	18.1	−17.1	3.8	40.3
G	72.7	2.7	82.9	29.2	−24.6	6.1	30.7
H	87.7	0.5	93.2	12.7	−11.9	2.1	29.2
AVG	78.4	1.4	87.0	22.7	−20.2	4.2	36.9
SD	8.9	1.3	5.4	8.4	9.5	1.7	7.9
Δ	−12.8	−1.2	−7.1	11.6	−14.1	2.0	2.9
p	0.0007	0.0795	0.0024	0.0015	0.0004	0.0032	0.4793

Last two rows: average difference compared to method using local statistics (Δ), *P*-value of the Student's *t*-test (*P*)

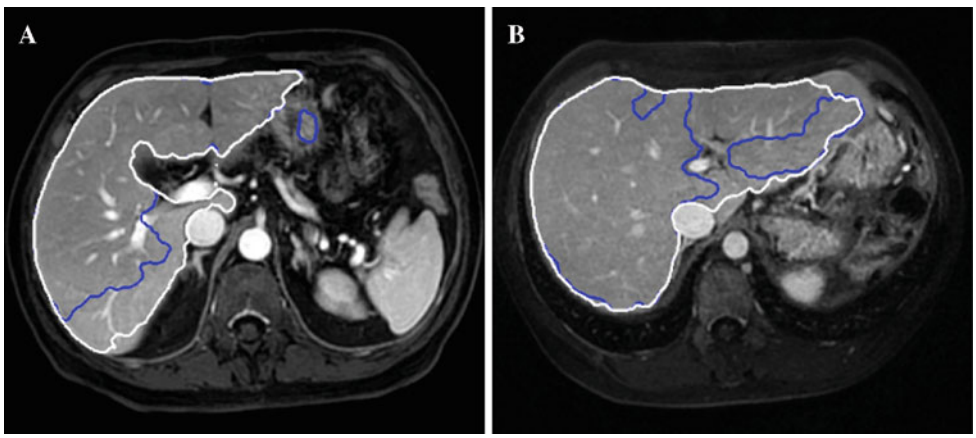

Fig. 6 Result of the segmentation for exams **a** (*left*) and **b** (*right*) with method using local (*white contour*) and global (*blue contour*) statistics

Conclusion

Our tests confirmed that the model-based liver segmentation is more precise if segment specific statistics are used. In case of using 8 segments the result involves more heterogeneous regions like lesions or artifacts. The optimal number of segments is not necessary equals to 8, which can be investigated in future work. We also note, that the segmentation can be improved if the statistics were continuously defined between neighboring segments.

We presented a segmentation method that uses probability information based on 60 manually contoured liver exams. It is important to note that under-segmentation can be decreased by incorporating the shape information of the

model. Although the presented model is modality independent, the model fitting exploits some characteristics of LAVA images, which shall be eliminated to adapt the segmentation method for wide range of MR images.

The segmentation precision was measured using various standard metrics, which allows comparing our results with those of the state of the art method, CT. Heimann et al. [11] presented a comparison of automated liver segmentation methods for CT images. According this paper, the average VOE was 11.3%, the average absolute RVD was 4.5%, the average ASSD was 2.5 mm, and the average MSSD was 35.5 mm for the 10 presented methods. Comparing these values with those of the proposed method (11.2%, 6.2%, 2.2 mm, 34.0 mm) we developed a competitive method for MR modality.

The running time of both methods was the same for all test exams. The average time was 30 s $\pm$ 8.7 using Intel Core2Duo 2.1 GHz, and 2 GB RAM, which is acceptable for wide range of clinical applications. In summary, we developed a method that can automatically segment the liver on contrast-enhanced MR LAVA images, but its extensive clinical usability shall be proved using larger dataset.

References

1. Furukawa D, Shimizu A, Kobatake H (2007) Segmentation method based on maximum a posterior probability estimation and level set method. In: MICCAI 2007 workshop proceedings: 3D segmentation in the clinic. 117–124
2. Rusko L, Bekes G, Fidrich M (2009) Automatic segmentation of the liver from multi- and single-phase contrast-enhanced CT images. Med Image Anal 13(6):871–882
3. Heimann T, Meinzer HP (2009) Statistical shape models for 3D medical image segmentation: A review. Med Image Anal 13(4):543–563
4. Seghers D, Loeckx D, Maes F, Vandermeulen D, Suetens P (2007) Minimal shape and intensity cost path segmentation. IEEE Trans Med Imaging 26(8):1115–1129
5. van Rikxoort E, Arzhaeva Y, van Ginneken B (2007) Automatic segmentation of the liver in computed tomography scans with voxel classification and atlas matching. In: MICCAI 2007 workshop proceedings: 3D segmentation in the clinic. 101–108
6. Kainmuller D, Lange T, Lamecker H (2007) Shape constrained automatic segmentation of the liver based on a heuristic intensity model. In: MICCAI 2007 workshop proceedings: 3D segmentation in the clinic. 117–124
7. Li H, Elmoataz A, Fadili J, Ruan S (2003) An improved image segmentation approach based on level set and mathematical morphology. In: Proceedings of SPIE vol 5286. 851–854
8. Farraher SW, Jara H, Chang KK, Hou A, Soto JA (2005) Liver and spleen volumetry with quantitative mr imaging and dual-space clustering segmentation. Radiology 237(1):322–328
9. Hermoye L, Laamari-Azjal I, Cao Z, Annet L, Lerut J, Dawant BM, Beers BEV (2005) Liver segmentation in living liver transplant donors: Comparison of semiautomatic and manual methods. Radiology 234:171–178
10. Bismuth H (1982) Surgical anatomy and anatomical surgery of the liver. World J Surg 6(1):3–9
11. 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 P, Chi Y, Cordova A, Dawant BM, Fidrich M, Furst JD, Furukawa D, Grenacher L, Hornegger J, Kainmuller D, Kitney RI, Kobatake H, Lamecker H, Lange T, Lee J, Lennon B, Li R, Li S, Meinzer H-P, Nemeth G, Raicu DS, Rau A-M, 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 (2009) Comparison and Evaluation of Methods for Liver Segmentation From CT Datasets. IEEE Trans Med Imaging 28(8):1251–1265