Published in final edited form as:

Proc SPIE Int Soc Opt Eng. 2013 February 9; 8672: . doi:10.1117/12.2006943.

Stable Automated Segmentation of Liver MR Elastography Images for Clinical Stiffness Measurement

Bogdan Dzyubak¹, Sudhakar K. Venkatesh², Kevin Glaser², Meng Yin², Jayant Talwalkar³, Jun Chen², Armando Manduca², and Richard L. Ehman²

¹Mayo Graduate School, Mayo Clinic, Rochester, Minnesota, United States

²Radiology, Mayo Clinic, Rochester, Minnesota, United States

³Gastroenterology, Mayo Clinic, Rochester, Minnesota, United States

Abstract

Magnetic Resonance Elastography (MRE) is an MRI-based technique that is used for the clinical diagnosis and staging of liver fibrosis by quantitatively measuring the stiffness of the liver. Due to the complexity of the signal characteristics and the presence of artifacts both in the acquired images and in the resulting stiffness images, the selection of the ROI for the stiffness measurement is currently performed manually, which may lead to significant inter- and intrareader variability. An algorithm has been developed to fully automate this analysis for liver MRE images.

Automated segmentation of liver MRE images is challenging due to signal inhomogeneity, low contrast, and variability in patient anatomy. An initial liver contour is found by fitting Gaussian peaks to the image histogram and selecting the peak that comprises intensities in the expected range and produces a mask near the expected location of the liver. After correction to reduce intensity inhomogeneity, an active contour based on intensity, with morphology used to implicitly enforce smoothness, is used to segment liver tissue while avoiding blood vessels. The resulting mask is used to initialize another segmentation which splits the region of the elastogram belonging to the liver into homogeneous liver tissue and areas with inclusions, partial volume effects, and artifacts.

In a set of 88 cases the algorithm had a $-6.0 \pm 14.2\%$ stiffness difference from an experienced reader, which was superior to the $6.8 \pm 22.8\%$ difference between two readers. The segmentation was run on an additional 200 cases and the final ROIs were subjectively rated by a radiologist. The ROIs in 98% of cases received an average rating of “good” or “acceptable.”

Keywords

MR elastography; liver; automation; segmentation; hepatic fibrosis

1. Introduction

Hepatic fibrosis is a significant world health problem causing as many as 44,000 deaths and over 100,000 acute hospitalizations annually^[1]. If untreated, it can progress to cirrhosis, which can lead to the need for liver transplantation or the development of hepatocellular carcinoma in as many as 50-70% of cases^[2]. Fibrosis can be reversed at early stages^[3-5],

however, which makes timely diagnosis highly important. Liver biopsy, which is currently the standard method for detecting hepatic fibrosis, is a highly invasive procedure, often confounded by the small sample size and subjectivity of grading which leads to variability in the readings^[6]. For this reason, significant research has been undertaken to develop non-invasive quantitative diagnostic techniques. Most such techniques, including morphometric methods, contrast enhanced MRI and CT, and diffusion tensor imaging with MRI have only been shown to be able to distinguish healthy volunteers from cirrhotic patients^[7]. On the other hand, Magnetic Resonance Elastography (MRE) and Transient Ultrasound Elastography^[8-10] involve a specialized acquisition for the measurement of hepatic stiffness (which is associated with high collagen deposition that progressively occurs in the diseased liver) and have been shown to have a high sensitivity and specificity for diagnosing fibrosis ^[1].

MRE is used clinically to diagnose and stage liver fibrosis as well as to evaluate the response to treatment. It measures tissue stiffness by applying acoustic vibrations to the body, imaging the resulting elastic wave propagation through the tissue with MRI, and analyzing images of the wave propagation to create images of the mechanical properties of the tissue (elastograms)^[11]. To achieve the correct stiffness measurement, a careful analysis of the images and exclusion of these artifacts is required. Wave interference near edges, as well as heterogeneous structures such as blood vessels, can lead to errors in the elastograms as illustrated in Figure 1.

Due to the complexity in the selection of an artifact-free liver tissue-only ROI, the analysis is currently performed manually using the following steps:

1. Analyze the images showing the wave propagation and exclude from the ROI regions with wave interference which may give rise to stiffness reconstruction errors.
2. Analyze the elastogram, which has lower resolution than the MR magnitude image, to exclude from the ROI areas with partial volume artifacts due to the processing.
3. Threshold the ROI using the confidence map that is calculated during the stiffness reconstruction and describes the quality of the wave data.

The ROI drawing process contains a number of subjective steps, particularly steps 1 and 2. Thus, variations in technique between readers can cause an average difference in the reported stiffness as high as 23%^[12].

In a prior study^[12], we proposed a fully automated algorithm that segmented the liver, excluded stiffness outliers, and reported the final liver stiffness. This method was shown to agree with the stiffness reported by a trained reader with an agreement similar to that reported by two experienced readers. To be able to use the algorithm as a standardized processing tool or as a quality control tool for clinical MRE, it needs to perform successfully across a broad range of clinical cases. Gradient echo (GRE) MRE is the implementation used for clinical MRE at a number of institutions. Segmentation of the liver in the magnitude images from these exams is challenging for most standard approaches as these images have signal inhomogeneity due to uncorrected coil sensitivities, as well as low

contrast due to thick slices and the motion inherent in MRE. Additionally, if the algorithm is to be used in the clinical environment it must be stable with respect to changes in patient and image characteristics. This paper presents an improved algorithm for reading MRE exams that can segment an appropriate ROI and report accurate stiffness values across a broad range of patient cases, as demonstrated in a retrospective study of our MRE clinical archive.

2. Method

The algorithm was evaluated in 288 clinical cases acquired in compliance with our institutional review board. The MRE images were acquired using a 1.5 T GE Signa Excite HDx scanner, an 8-channel receive-only torso coil, and a plastic MRE driver (producing vibrations at 60 Hz) placed on the chest wall of the patient^[1]. The GRE-based acquisition had an FOV of 44 cm in most cases, with a 256×64 acquisition matrix (interpolated to 256×256), parallel imaging acceleration factor of 2, and TR/TE/flip angle/bandwidth of 50 ms/20 ms/30°/32 kHz. Four contiguous 1-cm thick slices were acquired with 1st-moment gradient nulling. The tissue motion was recorded into the phase of the MR images using 16.7-ms, 3.2 G/cm, 1st-moment nulled motion-encoding gradients (MEG) applied in the superior-inferior direction that were synchronized with the tissue vibration. Each slice, collected during separate 14-sec breath holds performed at end expiration, included 4 time offsets between the motion and the MEG which were used to reconstruct the stiffness images (i.e., elastograms).

The elastograms were reconstructed using a Direct Inversion (DI) of the Helmholtz wave equation that included directional filtering in 4 orthogonal directions^[13] to yield elastograms with pixel values representing local tissue stiffness. Inversion was performed with an 11×11 pixel kernel. A confidence map based on the quality of local fits of second-order or fourth-order polynomials to the wave images was also calculated and regions with less than a 95% confidence were masked out from the final ROI to avoid high noise and wave interference areas.

The proposed fully automated algorithm is comprised of three stages – initialization, liver tissue segmentation, and removal of non-parenchymal tissues, such as blood vessels and tumors, as well as their partial volume. The initialization relies on prior knowledge of the relative image intensities of abdominal tissues as well as the large size of the liver and its position on the left side of the image in the axial imaging plane (right side of the body). Given the T1-weighted nature of the imaging sequence (TR/TE = 50/20 ms), the liver is known to have a medium intensity that is higher than that of air but lower than that of adipose tissue. To determine the appropriate threshold for the initialization mask, the histogram of the left half of the image (containing the liver) is calculated and three Gaussian peaks, corresponding to liver tissue, lungs and adipose tissue, are fit to it. The starting intensities of the peaks are $center = \text{mean}(\text{left half of image})$, $L = \text{max}(\text{min}(\text{left half of image}), center-w)$, and $R = \text{max}(\text{min}(\text{left half of image}), center+w)$, where w is a tunable width that was set to 75 intensity levels. The starting widths and heights of the peaks were set to $w/3$ intensity levels and $1/3$ of the total number of pixels in the histogram, respectively. If present, peaks that have means outside of the [L, R] interval or have a standard deviation greater than half of the L-R interval were attributed to the baseline and

inhomogeneity artifacts and are excluded from further analysis. A typical histogram comprising the three tissue types, as well as the peaks fit to them (P_{1-3}) is shown in Figure 2.

Due to inhomogeneity of the images as well as the variability in body composition between patients, sometimes one tissue type may have only a negligible contribution to the histogram or, alternatively, may require multiple peaks to fit. Thus spatial information was used in addition to intensity information to find the initial liver mask. For every peak contained within the $[L, R]$ interval, a tissue mask was generated consisting of pixels with intensities for which the Gaussian being considered was higher than the other two combined. The mask comprising at least 500 pixels and having the smallest mean distance to the middle of the left half of the image (expected liver center) was selected for initialization.

Before performing liver tissue segmentation, a simple inhomogeneity correction was performed by normalizing the image by a version of itself smoothed with a 30×30 kernel. Active contour segmentation was then used to separate the image into liver and nonliver tissue based on intensity (Equation 1):

$$F = -2\log(\sigma_u/\sigma_v) - (I - u)^2/\sigma_u^2 + (I - v)^2/\sigma_v^2 \quad (1)$$

$$F = F * e^{-|F-C|/\alpha} \quad \text{for all points with } F \leq C \quad (2)$$

In Equation 1, I is the vector of pixel intensities along the contour, u and v are the means of the regions inside and outside the contour, σ_u and σ_v are the standard deviations inside and outside the contour and F is the force associated with every contour pixel. An additional exponential cost term was used to prevent leakage into low-contrast areas (Equation 2). The term C is the step-size for the active contour (set to 0.5) such that points with $F < C$ do not move in a given iteration. Finally, α is an arbitrary tunable parameter that determines the strictness of the leakage prevention and was set to a value of 5.

Many active contour implementations are only effective if the initialization is close to the desired boundary as they contain a smoothness term which reduces leakage but prevents the contour from bypassing internal structures, such as blood vessels, to reach distant edges of the organ. To be able to capture a larger portion of the liver while avoiding vessels, the contour was allowed to adopt arbitrary shapes for short time intervals with no explicit smoothness term used. Hole-filling and morphological opening after every 50 iterations was used to remove isolated narrow areas of leakage while preserving parts of the contour that bypassed internal structures and reconnected. The holes being filled, likely corresponding to blood vessels, were kept track of for subsequent analysis.

Since the liver is a homogeneous organ and fibrosis is a diffuse disease, it is assumed that wave interference artifacts causing stiffness reconstruction errors occur only in and around small structures and edges, as found by the liver tissue segmentation. To further refine the ROI, the segmented magnitude image, with vessels excluded, was first thresholded at the mean $\pm$ one standard deviation in terms of intensities contained within the mask, and then again in terms of the stiffness values. This process excluded parts of vessel/tumor areas

which may have been missed by the segmentation. The 95% confidence mask generated based on the polynomial fits to the waves images was then also applied to the ROI removing areas with poor wave amplitude and wave interference. The resulting ROI was used to initialize the segmentation of the elastogram, within the liver segmentation mask, which removes the partial volume effects. This segmentation was performed with the same active contour used for the magnitude image segmentation. The final ROI produced by represents high-confidence, homogeneous hepatic tissue and excludes blood vessels, tumors, low-confidence areas, and their partial volume. The slice area-weighted average stiffness was then calculated from the 4 slices of the MRE exam.

The stiffness measurement obtained using the automated algorithm was compared to the stiffness reported by the more experienced of two readers (with 4 and 6 years of experience reading hepatic MRE cases) in a study of 88 cases using the Intraclass Correlation Coefficient (ICC) and the Bland-Altman test, as per our previous study. Additionally, as a study of stability, the algorithm was used to generate ROIs in a retrospective sample of 200 clinical cases. A radiologist who uses MRE in his clinical practice was asked to inspect the results and score them on a 3-point scale: “1 – I would accept the ROI” (good), “2 – Minor issues but I would accept the ROI” (acceptable), “3 – I would redraw the ROI” (bad). The difference in the stiffness measurement attained by using all the slices and only the slices with “good” ROIs was calculated in all cases where multiple grades of ROIs were present. All differences are reported as 95% confidence intervals (mean $\pm$ 1.96* standard deviation). A paired two-tail t-test was also performed to determine if the stiffnesses measured using all slices and only the slices with “good” ROIs were different.

3. Results

The ICC between the stiffnesses measured by the algorithm and each of the two readers was greater than 0.96 for the 88 cases. Using the Bland-Altman test, the 95% confidence interval for the difference between the algorithm and the more experienced reader was $-6.0 \pm 14.2\%$ which was superior to the agreement of $-6.8 \pm 22.8\%$ between the readers.

In the 200 case dataset evaluated by the expert radiologist 98% of cases were given a slice-average rating lower than 2 (better than “acceptable”). A histogram of slice-average ratings is shown in Figure 3. In cases where both “good” and “acceptable/bad” ROIs were available ($n = 39$), the average difference between taking the stiffness measurement obtained using only the “good” slices and all of the slices was $-3 \pm 5.7\%$. Furthermore, a two-tail t-test comparing the resulting stiffness measurement yielded a p-value of 0.42. Thus, the inclusion of sub-optimal slices did not significantly affect the measured stiffness in those cases. The main issue reported by the radiologist in “acceptable” and “bad” cases was the limited extent of the ROIs. In rare cases leakage into other organs or the inclusion of areas with wave interference was reported.

Representative images of the diverse nature of MRE magnitude images and the ROIs resulting from the magnitude image segmentation are shown in Figure 4.

An example of the segmentation of the elastogram leading to effective exclusion of the partial volume region around a blood vessel and an out-of-plane hepatic fissure is shown in Figure 5.

4. Discussion

The algorithm has demonstrated high stability with respect to image and patient variations by segmenting 288 cases successfully. The discrepancy in the stiffness reported by the algorithm and an experienced reader in 88 cases was less than the difference between two readers. The relative improvement in these 88 cases compared to our previous study was small. However, the use of prior knowledge of contrast and spatial information allowed successful ROI initialization in all 288 cases, approximately 5-10% of which would not have resulted in useable initialization using our previous method.

The expert evaluation of the automatically generated ROIs was very favorable with 98% of cases receiving interslice average ratings better than “acceptable.” Furthermore, slices with “acceptable” or even “bad” ratings did not produce a statistically significant effect on the average stiffness. Thus, as long as the ROI is sufficiently large, typically over 10,000 pixels, and has sufficiently good quality, small regions of missed tissue and leakage into other organs do not affect the final measured stiffness.

The inhomogeneity correction used in this study is very simple and, in principle, is better suited towards images where intensities do not vary spatially on a large scale. Within large homogeneous organs pixels are normalized by the average of other pixels of the same tissue and inhomogeneity is removed. On the other hand, pixels near boundaries with tissues of significantly different intensities are divided by a small value if the adjacent tissue is darker and by a large value if the tissue is brighter. Thus, the edge contrast at boundaries with significantly different intensities is further increased while the boundary is made thicker. The first effect is positive in that it may reduce leakage while the second makes the segmentation unable to capture some of the tissue near the edges. Since a typical liver contains tens of thousands of pixels, and the effect described above occurs only near boundaries with high tissue contrast and extends no further than the size of the blurring kernel (30 pixels), the reduction in the sample size is unlikely to affect the stiffness measurement significantly. This simple inhomogeneity correction approach was effective given the goal of measuring an average stiffness value. However, a more universal inhomogeneity correction, such as one using an entropy-based polynomial fit to model and remove inhomogeneity, should be investigated.

The blood vessel exclusion during the segmentation step was generally effective, as shown in Figure 4. However, due to the 1-cm slice thickness and motion present in the images, edge contrast is low in these images and vessels may be missed. The approach in this study was to additionally threshold the segmentation mask for outliers in terms of intensity and stiffness which would allow the active contour running on the elastogram to sample additional regions. Due to the smoothness of the elastogram itself, small regions that were excluded by the threshold but did not have sharp boundaries with their neighbors would again be included in the mask while only significantly stiffer or softer objects would be

excluded. The use of an initialization and a magnitude image segmentation technique that is more stable with respect to low contrast would allow better exclusion of blood vessels and reduce contour leakage, likely removing the need for thresholding of the segmentation mask. Ideally, this technique would take into account discontinuities in the wave-propagation image which may have better contrast than the intensity image itself. A method for doing this is under development.

The elastogram segmentation step is useful in removing partial volume artifacts due to blood vessels and edges. Unfortunately, it further reinforces the assumption that the liver has a homogeneous stiffness. While this assumption is made of necessity in many MRE approaches, a detailed study of inhomogeneity by clinicians is warranted and the algorithm presented here may play a key role in such a study if appropriately extended.

5. Conclusions

The fully automated algorithm presented here was able to produce reliable ROIs and stiffness measurements across a wide range of clinical MRE exams, as determined by comparing the stiffness measurements it produced to those from an expert reader as well as by subjective evaluation by a radiologist. Future improvements include better inhomogeneity correction and the development of a segmentation technique that is more sensitive to edge contrast. The improvements presented here allow the algorithm to process a broad range of cases and will be important in the upcoming prospective study of the algorithm as a quality control tool.

Acknowledgments

NIH EB01981

References

1. Yin M, Talwalkar JA, Glaser KJ, et al. Assessment of hepatic fibrosis with magnetic resonance elastography. *Clin Gastroenterol Hepatol*. 2007; 5(10):1207–1213. [PubMed: 17916548]
2. Fattovich G, Stroffolini T, Zagni I, Donato F. Hepatocellular Carcinoma in cirrhosis: Incidence and risk factors. *Gastroenterol*. 2004; 127(5 Suppl 1):35–50.
3. Bataller R, Brenner DA. Liver fibrosis. *J Clin Invest*. 2005; 115:209–218. [PubMed: 15690074]
4. Tsukada S, Parsons CJ, Rippe RA. Mechanisms of liver fibrosis. *Clin Chim Acta*. 2006; 364:33–60. [PubMed: 16139830]
5. Friedman SL. Liver fibrosis -- from bench to bedside. *J Hepatol*. 2003; 38:38–53.
6. Regev A, Berho M, Jeffers LJ, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. *Am J Gastroenterol*. 2002; 97:2614–2618. [PubMed: 12385448]
7. Afdhal NH, Nunes D. Evaluation of Liver Fibrosis: A concise review. *Am J of Gastroenterol*. 2004; 99:1160–1174. [PubMed: 15180741]
8. Bonekamp S, Kamel I, Solga S. Can imaging modalities diagnose and stage hepatic fibrosis and cirrhosis accurately? *J Hepatol*. 2009; 50:17–35. [PubMed: 19022517]
9. Talwalkar JA, Yin M, Fidler JL, Sanderson SO, Kamath PS, Ehman RL. Magnetic resonance imaging of hepatic fibrosis: emerging clinical applications. *Hepatology*. 2008; 47:332–342. [PubMed: 18161879]
10. Fatemi M, Greenleaf JF. Probing the dynamics of tissue at low frequencies with the radiation force of ultrasound. *Phys Med Biol*. 2000; 45(6):1449–64. [PubMed: 10870703]

11. Muthupillai R, Lomas DJ, Rossman PJ, Greenleaf JF, Manduca A, Ehman RL. Magnetic resonance elastography by direct visualization of propagating acoustic strain waves. *Science*. 1995; 269(5232):1854–7. [PubMed: 7569924]
12. Dzyubak B, Glaser K, Yin M, et al. Automated Liver Stiffness Measurement with Magnetic Resonance Elastography. *J Magn Reson Imaging*. 2012 Epub ahead of print.
13. Manduca A, Lake DS, Kruse SA, Ehman LR. Spatio-temporal directional filtering for improved inversion of MR elastography images. *Med Im Anal*. 2003; 7(4):465–473.

Figure 1.

Magnitude (A), wave (B), and stiffness (C) images. A set of blood vessels (indicated by the arrow) has very low contrast in the magnitude image but can be observed as a discontinuity in the wave image. This disruption results in a region with apparent significantly decreased stiffness that is blurred due to the size of the processing kernel (11×11 pixels).

Figure 2.

Algorithm initialization by peak fitting to the intensity histogram. Each of the peaks, intended to fit the air in the lungs (A), liver tissue (B), and adipose tissue (C), is used to create and initialization mask. The mask with the smallest mean distance to the expected center of the liver is selected (mask B, in red).

Figure 3. Histogram of inter-slice average ROI ratings

The “good” category comprises cases with average rating below 1.25, the “good-acceptable” category comprises cases with ratings above 1.25 and below 2, while the “acceptable-bad” category contains cases with average ratings above 2.

Figure 4. Examples of clinical MRE magnitude images and their segmentations

These ROIs (the area between the internal and external green contours) comprise homogeneous liver tissue and exclude blood vessels and fissures. These ROIs would be further adjusted by using stiffness information and the confidence mask before a final measurement is reported.

Case A – Many blood vessels. The full liver was found successfully while excluding the many hepatic blood vessels.

Case B – Unusual body composition. Even though the full liver was not segmented in this case due to the relatively small liver size and unusual shape, the resulting ROI still represents homogeneous liver-only tissue.

Case C – Motion artifact and poor contrast. A homogeneous area of the liver was found and blood vessels were excluded.

Figure 5.

Segmentation of the elastogram to remove partial volume artifacts from the ROI. The magnitude image with the liver mask (green) is presented in A. In image B, the contour for the refined ROI (in red) shows the initial exclusion of blood vessels and areas with low confidence, while image C shows the final ROI with the partial volume artifacts fully excluded. The checkerboard area is the masked out region with poor wave propagation as determined by the stiffness inversion algorithm ($<95\%$ confidence).