33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Automated Scoring of Liver Fibrosis through Combined Features from Different Collagen Groups

Shuoyu Xu, Dean Tai, Aileen Wee, Roy Welsch, Peter So, Hanry Yu, Jagath Rajapakse

***Abstract*— Liver biopsy remains the gold standard in monitoring progression of liver fibrosis associated with an abnormal increase in collagen. Descriptive scoring systems are still being widely used to grade biopsy samples. In this study, we propose a new set of features by clustering collagen fibers into three groups first based on their localization and connectivity properties, and then by extracting morphological features of collagen fibers. The new feature set is compared to the earlier features used in classification of liver fibrosis, which were based on the total amount of collagen fibers. Our results show that new features lead to more accurate grading of liver fibrosis.**

I. Introduction

Liver fibrosis is a result of a wound healing process that accumulates extracellular matrix (ECM) proteins such as collagen [1]. Liver biopsy remains the gold standard in monitoring the progression of fibrosis in which a small sample of tissue is removed with a needle, stained, and examined under a microscope, and then graded based on descriptive or semi-quantitative scores by trained pathologists. However, pathological features used in these systems do not have clear definitions and are somehow ambiguous, which makes the grading and staging scores subjective. Therefore, inter- and intra- observer variations can be as high as 35% and it is difficult to obtain highly reproducible results from these systems [2].

To exclude observer discrepancies, several studies have reported building automated image analyses systems to grade liver fibrosis by using the features extracted from images [3]. Most of these systems measure the amount of collagen as the only measurement to grade and stage fibrosis. Our previous work showed how better accuracy of grading is achieved by using both morphological features of the collagen fibers and the amount of collagen [4]. However, different pathologies can result in different morphologies of collagen distribution as liver damage progresses. Collagen deposition often starts around the portal region in early stages and eventually connects to form nodules in advanced fibrosis, resulting in liver cirrhosis. As a result, the changes of collagen amount and morphologies are not the same for all the collagen fibers in the sample along the fibrosis progression.

Other pathological features such as collagen architecture change play a more important role in grading and staging fibrosis [5], which leads to the need for computer-aided systems with more measurements than collagen amount only. To address this problem, we clustered collagen fibers into three groups based on their connectivity properties and location before extracting features of collagen fibers in order to represent collagen morphology changes more accurately. The performance of grading with the new features is compared with those achieved with features extracted from total collagen fibers and previously used features. The results suggest that combined features from different collagen groups can lead to more accurate scoring of liver fibrosis.

II. Methods

The proposed automated scoring system consists of three major steps as shown in Fig 1. First, collagen fibers are clustered into three groups according to their location and connectivity properties. Second, the major axis of each collagen fiber is identified and the morphological features of collagen fibers are quantified for each collagen group. Finally, the best feature set is selected from all the features and is used to train a classifier for scoring.

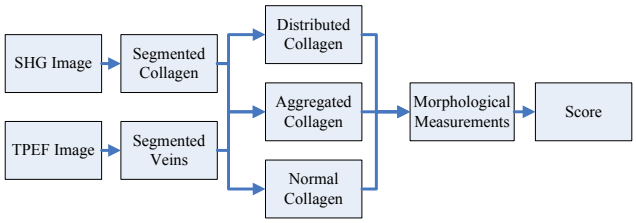

Fig. 1 Illustration of the proposed computer-aided scoring system for liver fibrosis.

A. Collagen Segmentation

Liver slices were scanned by second harmonic generation (SHG) microscopy, which is a non-linear optical process

Manuscript received April 7, 2011. This work was supported in part by Singapore-MIT Alliance Research and Technology (SMART) Center.

S. X. is with Computation and System Biology Program, Singapore-MIT Alliance. (phone: 65-9729-4710; e-mail: s070066@e.ntu.edu.edu).

D. T. is with Histoindex Pte.Ltd. (e-mail: dean.tai@histoindex.com).

A. W. is Professor and Senior Consultant at the Department of Pathology, Yong Loo Lin School of Medicine, National University of Singapore, National University Hospital.(e-mail: aileen_wee@nuhs.edu.sg).

R.W. is Professor of Sloan School of Management, Massachusetts Institute of Technology. (e-mail: rwelsch@mit.edu).

P.S. is Professor of Mechanical Engineering and Biological Engineering, Massachusetts Institute of Technology. (e-mail: ptso@mit.edu).

H.Y. is Professor in Physiology, Yong Loo Lin School of Medicine, National University of Singapore. (e-mail: hanry_yu@nuhs.edu.sg).

J.R. is Professor and Director of Bioinformatics Research Center, School of Computer Engineering, Nanyang Technological University. (e-mail: asjagath@ntu.edu.sg).

specific for imaging fibrillar collagens without the need of staining [6]. SHG images were first segmented into collagen and background by using a segmentation algorithm based on Gaussian mixture models [7]. It was assumed that the intensity of pixels in the SHG image can be modeled as the mixture of two Gaussian distributions, one representing collagen area with strong SHG signals and the other representing the background. The segmentation results were compared with those generated from other segmentation methods such as global thresholding and clustering methods in our previous study, which illustrated a more accurate quantification of fibrosis progression by using a segmentation algorithm based on Gaussian mixture models.

B. *Identification of collagen around veins*

Collagen disperses differently in different locations by starting around veins and tracts and eventually connecting together. Therefore, it is important to identify those collagen fibers at different locations such as those around the veins and tracts.

The veins were first segmented from an image generated by two-photon excitation microscopy (TPEF) for the same region of SHG image. Since the TPEF image records the auto-fluorescence signals of the tissue, the vein areas were empty spaces with low auto-fluorescence signal. The pixels were classified into three groups based on intensity by K-means unsupervised clustering. The group of pixels with the lowest average pixel intensity was recognized as the veins and tracts area, while the other two groups refer to damaged cells area and healthy hepatocytes area.

Let $f:\Omega \to \{0,1\}$ denotes the collagen image after segmentation where $f(z)=1$ when pixel $z=(x,y)\in\Omega$ is a collagen pixel, and $f(z)=0$ otherwise. C denotes the set of collagen pixels where $f(z)=1, z\in C$.

For each collagen pixel, we proposed a normalized distance γ to reflect how far it is from its nearest vein space:

$$\gamma = \frac{\min_{b\in B} d(z,b)}{d(z,c)}, z\in C$$

where $d(z,c)$ is the Euclidean distance between the collagen pixel z and the center of mass c of the nearest vein space, and $d(z,b)$ is the Euclidean distance between the collagen pixel z and the boundary point b which belongs to the set of all boundary points B of the nearest vein space, see Fig 2(a). The normalized distance γ of a collagen pixel which is close to a vein will be close to zero whereas it will be near to one for a collagen pixel far from the veins. Fig 2(b) illustrates the histogram of γ from Fig 2(a), and two peaks can be observed. Collagen pixels are then divided into two groups according to γ by the Otsu method [8]. The group of collagen pixels with value of γ close to zero is identified as collagen around portal tracts and central veins. The rest of the collagen pixels are recognized as collagen between portal tracts and central veins.

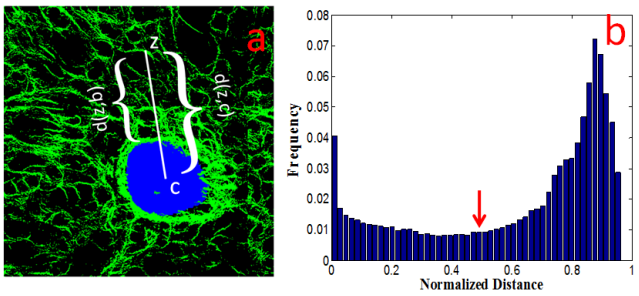

Fig 2. The parameters used to calculate normalized distance are defined in (a), where z is a collagen pixel and c is the center of mass of the nearest vein space. Collagen pixels are shown in green and vein space is shown in blue. The histogram of normalized distance from (a) is presented in (b), and two peaks can be observed. The red arrow points to the threshold generated by the Otsu method.

C. *Separation of aggregated and distributed collagen*

Collagens between portal tracts and central veins are further grouped into aggregated collagen and distributed collagen. If a collagen fiber links to other fibers, it has more than one cross-link and is classified as aggregated collagen. Distributed collagens are those collagen fibers with no cross-links with others.

To identify the cross-links of each fiber and to further extract morphological features, a fiber network extraction algorithm [9] was performed on the binary image of the segmented collagen to track the skeleton of each collagen fiber. The principle of the algorithm is to calculate the minimal distance from a collagen pixel to a background pixel and to trace along the maximal ridges of the distance function. The algorithm is briefly described here.

For a collagen pixel $z\in C$, we define its neighbors $N_r(z)$ in a searching box with radius r and $B_r(z)$ on the boundary of a searching box with radius r as below:

$$N_r(z) = \{v : \|z-v\| < r; v\in\Omega\},$$
$$B_r(z) = \{v : \|z-v\| = r; v\in\Omega\}.$$

Firstly, nucleation points N of each fiber are located as the starting points for fiber extending and tracking in the following steps. To identify nucleation points, the distance transform is first computed for all collagen pixels given by:

$$d_{\min}(z) = \min_{v\in\bar{C}} \|z-v\|, z\in C.$$

The nucleation points N are then defined as the global maximum points in a neighborhood of the distance function:

$$\mathrm{N} = \{z : d_{\min}(z) \geq \max_{v\in N_r(z)} \{\alpha_1, d_{\min}(v)\}, z\in C\}$$

where α_1 is a threshold parameter and is set to 1.5 in the experiment.

Secondly, starting from each nucleation point, fibers were traced through a set of local maxima points until the end of

the fiber or another nucleation point was reached. Given a nucleation point $z \in \mathrm{N}$, its local maxima points are given by:

$$LMP = \{v : v \in B_{d_{\min}(z)}(z), d_{\min}(v) \geq \max_{b \in B_1(v)} \{\alpha_2, d_{\min}(b)\}, z \in \mathrm{N}\}$$

where α_2 is a threshold parameter and is set to 0.2.

From each nucleation point, the fiber is extended by adding its local maxima points to the list. For each of these local maxima points, its local maxima points are calculated by the same equation described and added to the current fiber so that the fiber is extended further. The extension stops when there is no local maxima point found or another nucleation is found in the neighborhood of current point.

Finally, those fibers shorter than a certain threshold are recognized as danglers and will be removed. Moreover, if the distance between two fiber ends are smaller than a threshold and the difference of the orientation of these two fibers are smaller than a threshold as well, these two fibers are linked to form a single fiber.

Cross-link points are defined as those points belonging to more than one fiber. Those collagen fibers having no cross-link points were classified as distributed collagen, which refers to fine collagen fibers distributed in sinusoidal regions. The rest of the collagen fibers containing one or more cross-link points were aggregated collagen.

Table 1. Connectivity Property and Location of Three Collagen Groups.

Feature Set	Symbol	Cross-links	Location
Total Collagen	F_T		
Vein Collagen	F_N	>=1	Around veins
Aggregated Collagen	F_A	>=1	Between veins
Distributed Collagen	F_D	=0	Everywhere

D. Collagen classification and feature extraction

The collagens are classified into vein collagen, aggregated and distributed collagen according to the location identified in section *B* and number of cross-links extracted as described in section *C*. The connectivity property and location of each collagen group are summarized in Table 1. The clustering results for a sample image are shown in Fig 3, whereas vein collagen, aggregated collagen and distributed collagen are coded in blue, green and red colors.

Eight morphological features were extracted for total collagen fibers as well as for vein collagen and aggregated collagen: collagen amount, fiber number per mm^2, average fiber length, average fiber width, average curvature, average coordination number, average cross-link density, and average cross-link space, based on collagen mask after segmentation and major axis extracted from each collagen fiber. Since distributed collagen fibers have no cross-link points, only six features were extracted by excluding cross-link density and cross-link space.

III. Experimental Results

All the liver tissue samples in this study were extracted from bile duct ligated rats. Bile duct ligation (BDL) of rats will generate a wound in the liver and then lead to fibrosis. A total of 15 rats were ligated and sacrificed at intervals of 2, 4 and 6 weeks (n = 5 per week). 5 control rats were also sacrificed at week 0. A tissue slice with 50um thickness is sliced from each liver and a total of four images (4068 x 4095 pixels, ~ 4.1 x 4.1 mm) were scanned for each slice by SHG microscopy. Another tissue specimen with 4um thickness is sliced from each liver, stained, imaged by light microscope, and scored by a pathologist. The score is then used as the ground-truth of fibrosis stage of each liver tissue to train and test the performance of the scoring system with proposed feature sets.

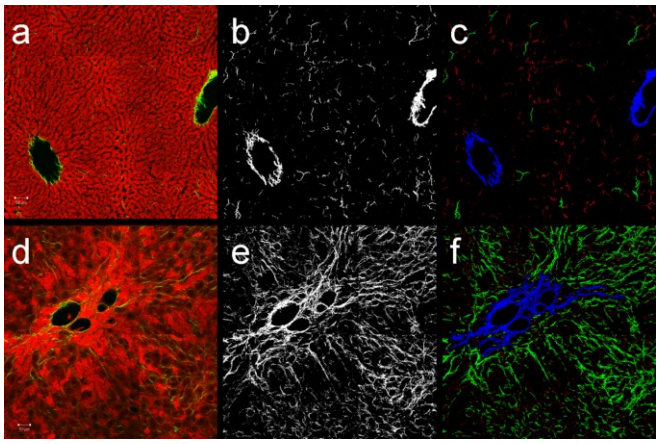

Fig. 3 Clustering of collagen from one healthy and one late stage fibrosis liver sample. (a) SHG/TPEF image of a healthy sample. (b) Segmented collagen fibers of (a). (c) Collagen clusters of (b), normal, aggregated and distributed collagens are presented in blue, green and red respectively. (d) SHG/TPEF image of a late stage sample. (e) Segmented collagen of (d). (e) Collagen clusters of (e).

The bootstrap method was used to determine significance of the performance difference between the feature sets. The dataset used in the experiment contained 129 images. One thousand bootstrap samples were generated by sampling the images with replacement, which means that each new set of sampled images contained the same number of images as the original set. For each set, feature selection based on minimum redundancy and maximum relevancy (MRMR) criteria [10] was performed to rank the features and 5-fold cross-validation using a one-against-one multi-class support vector machine classifier was performed to decide the number of selected features according to weighted cost which is the weighted summation of wrongly predicted samples.

The classification results indicated that, by combining features from vein collagen, aggregated collagen and distributed collagen with features from the total amount of collagen, better scoring performance can be achieved. The

feature set F_T contains eight features extracted from total amount of collagen whereas the best performance of the system is accomplished when the top five ranked features are selected. By adding features from vein collagen, aggregated collagen and distributed collagen to F_T, the new feature set consists of thirty features and is able to improve the system performance significantly since the average weighted cost decreased about 25% from 20.45 to 14.9 and the average scoring accuracy increased by 5%, while top eight ranked features are selected and used, see Table 2.

Table 2. Feature selection results of the scoring system using the combination of features.

Feature Set	No. of Features	No. of Selected Features	Average Weighted Cost	Average Accuracy
F_T	8	5	20.45±6.67	85.39±0.04%
$F_T+F_N+F_A+F_D$	30	8	14.96±4.62	89.26±0.03%

With the combination of features from different collagen groups, the sensitivity and specificity of the scoring system are listed in Table 3. Comparing to the sensitivity and specificity using the total collagen feature set only, the improvements of sensitivity and specificity are observed for most of the stages except stage 1. Because the number of images of stage 1 group is limited and is much less than the numbers of other groups, this unbalanced dataset could result in the big variance of the results of stage 1. Overall, while good sensitivity and specificity are achieved for early and late stages, the system also performs well for separating mid stages using the combination of features from different collagen groups which again indicates the importance of introducing the new features from different collagen groups.

Table 3. Sensitivity and specificity of the scoring system using the combination of features from different collagen groups, and using total collagen feature set only.

Stage	0	1	2	3	4	Average
F_T						
Sensitivity	93.7± 5.2%	86.4± 17.2%	74.3± 10.4%	85.2± 7.4%	94.5± 5.3%	86.8±4.8%
Specificity	98.4± 1.4%	98.7± 1.2%	92.6± 3.3%	93.6± 3.0%	99.1± 0.9%	96.5±1.1%
$F_{T+N+A+D}$						
Sensitivity	96.1± 4.6%	81.5± 20.1%	83.9± 8.2%	91.5± 5.1%	97.8± 3.9%	90.2±4.4%
Specificity	98.9± 1.1%	98.5± 1.2%	95.7± 2.3%	97.1± 1.9%	99.1± 0.9%	97.9±0.7%

IV. Conclusion

In this study, we have demonstrated a new set of collagen features for automated scoring of liver fibrosis. The collagen fibers are grouped into three classes based on their connecting properties and location first and morphological features are extracted for different collagen groups. We suggested that the combination of collagen features of different groups with total collagen features can lead to more accurate scoring of liver fibrosis. This is motivated by the pathological knowledge that different pathologies can result in different morphologies of collagen distribution as liver damage progresses, which makes the usage of total collagen features only not accurate enough to monitor fibrosis progression. Our suggestion is supported by the statistical analysis results. The combination feature sets with collagen features of all three clusters performed significantly better than the total collagen feature set based on the comparisons of classification accuracy and sensitivity using a multi-class SVM. This is due to the more accurate representation of collagen morphology changes by grouping collagen fibers into different classes. Further study on different animal models and human samples will be conducted to demonstrate the value of using feature sets of different collagen clusters. Moreover, although images used in our study are from SHG microscopy, all the image analysis and feature extraction procedures are applied on binary images after collagen segmentation. So these methods can be applied to stained samples directly, while the only difference is that the collagen segmentation algorithm should be altered for stained samples.

References

[1] Bataller, R., D.A. Brenner: Liver fibrosis. *J. Clin. Invest.* 115, 209-218 (2005)

[2] Perrault, J., Mcgill, D.B., Ott, B.J., Taylor, WF: Liver biopsy: complications in 1000 inpatients and outpatients. *Gastroenterology* 74(1), 103-106(1978)

[3] Germani, G., Burroughs, A.K., Dhillon, A.P.: The relationship between liver disease stage and liver fiborsis: a tangled web. *Histopathology*. 57, 773-784(2010)

[4] Xu, S., Tai, D.C.S., So, P.T.C., Yu, H., Rajapakse, J.C., Automated Scoring System for Liver Fibrosis Diagnosis with Second Harmonic Generation Microscopy, *ICONIP 2010*, 22-25 Nov, Sydney

[5] D.Goodman, Z., Grading and staging systems for inflammation and fibrosis in chronic liver diseases. *Journal of Hepatology*, 2007. **47**: p. 598-607.

[6] Tai, D., Tan, N., Xu, S., Kang, C.H., Chia, S.M., Cheng, C.L., Wee, A., Wei, C.L., Raja, A.M., Xiao, G., Chang, S., Rajapakse, J.C., So, P.T., Tang, H.H., Chen, C.S., Yu, H.: Fibro-C-Index: comprehensive, morphology-based quantification of liver fibrosis using second harmonic generation and two-photon microscopy . *J Biomed Opt* 14(4), 044013(2009)

[7] Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum Likelihood from Incomplete Data via the EM algorithm. *J. Roy. Statistical Society Series B* 39(1), 1-38(1977)

[8] Otsu, N.: A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics* 9(1), 62-66(1979)

[9] Stein, A.M., Vader, D.A., Jawerth, L.M., Weitz, D.A., Sander, L.M.: An algorithm for extracting the network geometry of three-dimensional collagen gels. *J Microsc* 232(3), 463-475(2008)

[10] Peng, HC., Long, FH., Ding, C., Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy**,** *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 27, No. 8, pp.1226-1238, 2005.