34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Computational Modeling for Assessment of IBD: to be or not to be?

Franciscus M. Vos, Jeroen A.W. Tielbeek, Robiel E. Naziroglu, Zhang Li, Peter Schueffler, Dwarikanath Mahapatra, Alexander Wiebel, C. Lavini, Joachim M. Buhmann, Hans-Christian Hege, Jaap Stoker, Lucas J. van Vliet

Abstract—**The grading of inflammatory bowel disease (IBD) severity is important to determine the proper treatment strategy and to quantify the response to treatment. Traditionally, ileocolonoscopy is considered the reference standard for assessment of IBD. However, the procedure is invasive and requires extensive bowel preparation. Magnetic resonance imaging (MRI) has become an important tool for determining the presence of disease activity. Unfortunately, only moderate interobserver agreement is reported for most of the radiological severity measures. There is a clear demand for automated evaluation of MR images in Crohn's disease (CD). This paper aims to introduce a preliminary suite of fundamental tools for assessment of CD severity. It involves procedures for image analysis, classification and visualization to predict the colonoscopy disease scores.**

I. INTRODUCTION

Inflammatory bowel diseases (IBD) constitute one of the largest healthcare problems in the Western World. They affect over 1 million citizens both in Europe and in the USA, 700,000 respectively 500.000 of them suffer from Crohn's disease. Grading of Crohn's disease severity is important to determine the proper treatment strategy and to quantify the response to treatment.

Traditionally, ileocolonoscopy in combination with tissue biopsies is considered the reference standard for diagnosis and assessment of IBD. However, the procedure is invasive and requires extensive bowel preparation, which is considered very burdensome by most patients. Moreover, it only gives information on superficial abnormalities and only for the most distal part of the small bowel.

Therefore, abdominal Magnetic Resonance Imaging (MRI) is now widely used for diagnosing and grading luminal Crohn's disease (CD). It typically involves a luminal (oral) and an intravenous contrast medium in order to combine mural and extra-intestinal evaluation of disease activity. Such grading of disease activity is becoming more and more important in clinical practice given the often costly and burdensome medical treatment. Additionally, pharmaceutical trials increasingly employ MRI as outcome measure.

In a recent study Rimola et al. [1] used multivariate analyses correlating the subjective radiological features to the Crohn's Disease Endoscopic Index of Severity (CDEIS)[2]. The study confirmed the radiologic parameters that should be evaluated to that end, which were wall thickening, wall signal intensity, relative contrast enhancement, presence of edema, ulcers, enlarged lymph nodes and presence of pseudopolyps.

Unfortunately, grading the disease activity based on MRI features has intrinsic limitations related to restrictions of the MRI technique. Also, it is a subjective evaluation while varying weight is attributed to these features. For instance, a recent study by Ziech et al. [3] reported a weak to moderate interobserver agreement for most of the subjective MRI features. Based on the present methods of grading, MRI has been shown to be accurate for severe disease cases (91% accuracy), but mediocre for mild disease or remission (62% accuracy) [4].

Clearly, a system is preferred that renders a fine grading of the disease severity for accurate treatment monitoring. For an optimal evaluation of response monitoring, MRI should be a robust, objective and reproducible technique. Applying a (semi-)automated method might improve the interobserver variation and allow a finer diagnostic scale compared to the gross scale (remission – mild – severe) presently used by the radiologists. Therefore, development of computer-assisted diagnosis tools for quantitative image-based analysis of CD is pivotal.

This paper aims to introduce a preliminary suite of fundamental tools for assessment of CD severity. It involves image analysis, classification and visualization algorithms to measure disease severity from MRI features. The tools may facilitate early diagnosis and a more precise monitoring of the disease progression.

II. DATA ACQUISITION

The data employed in this paper were taken from a previous study [3]. It concerned 32 of 35 patients that consented to use of their data in the current research. MR imaging included free-breathing 3D + t DCE-MRI data acquisition on a 3.0T Philips Intera scanner by a 3D spoiled gradient echo sequence. 14 coronal slices were obtained; pixel sizes were 1.78 × 1.78 × 2.5mm, 450 of these 3D image volumes were acquired during 6.1 minutes at a rate of one volume per 0.8 seconds. Buscopan (Generic name –

* This research was funded from the European Community's Seventh Framework Programme as the VIGOR++ Project (grant nr. 270379).

F.M. Vos, R.E. Naziroglu, Z. Li and L.J. van Vliet are with Delft University of Technology, The Netherlands (phone: +31-278-7133; fax: +31-15-2786740; e-mail: f.m.vos@tudelft.nl).

F.M. Vos, J.A.W. Tielbeek and J. Stoker are with the Academic Medical Center Amsterdam, The Netherlands.

P. Schueffler and D. Mahapatra are with the Eidgenössische Technische Hochschule Zürich Zurich, Switzerland.

A. Wiebel and H.-C. Hege are with the Zuse Institute, Berlin, Germany.

Scopolamine) was administered to the subjects to minimize bowel movement. A contrast agent (Gado-vist) was injected (0.1 ml/kg) after the 10th image volume was acquired. Moreover, high resolution VIBE imaging was performed post-contrast. CRP level (a blood marker), and CDEIS (Crohn's Disease Endoscopic Index of Severity [2]) were determined serving as reference disease severity measures.

III. Image Analysis

Several image analysis procedures were devised by us for automated delineation of the colon's inner surface, measurement of the colon wall and registration of the image data.

A. Identification of the colon's inner surface.

The colonic wall thickness is typically increased with active Crohn's disease. A first step to enable automated measurement of the thickness (see below) is to identify the bowel's inner surface in the MR-images. This is a challenging problem, because the shape varies sharply, the lumen is often narrowed into stenotic parts and the contrast between wall and lumen is space variant.

A semi-automated method was devised to overcome these problems. The post-contrast MRI sequence appeared most suitable for measuring the wall thickness as it combines the best contrast with high resolution. Initially, a centerline was manually drawn by an expert through a part of the lumen suspected of Crohn's activity. Subsequently, a so-called level-set representation [5] evolved to match the bowel surface. The level-set was steered by the image information (transitions in intensity) while smoothness constraints restricted its shape. Additionally, it dealt with heterogeneities in the lumen by permitting spatial variations in intensity along the path. The method was capable to segment healthy as well as diseased bowel parts, as in Figure 1.

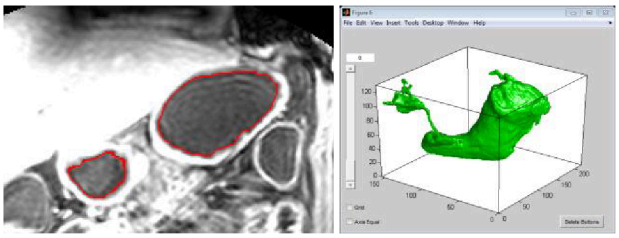

Figure 1. Example of the bowel surface segmentation: left original data, right 3D surface representation.

B. Colon wall segmentation

A method was developed for colon wall segmentation that assumes an accurately segmented colon surface (see above). Subsequently, 1D profiles of image intensities were recovered in sampled surface points by stepping in either direction perpendicular to the surface. These intensities were modeled as a sum of step functions that reflected transitions in signal from lumen to colon wall and from colon wall to surrounding tissue. Additionally, it was asserted that the step functions were convolved by Gaussians representing the directional MRI scanner's Point Spread Function. The wall thickness is simply retrieved from the parameters of the fitted model, i.e. the distance between the step functions. A representative result is shown in Figure 2.

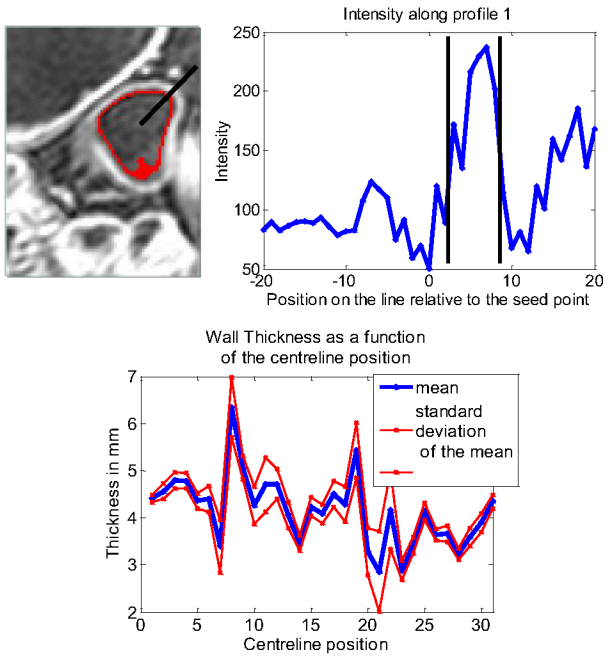

Figure 2. Part of the colon surface (top left) and profile along which intensities are sampled (top right) to retrieve the local wall thicknesss. The bottom image shows the wall thickness as a function of position along the centerline.

C. MRI data registration.

The Time Injection Curves (TICs) obtained from Dynamic Contrast Enhanced MRI (DCE-MRI), are expected to contain important information on the degree of inflammation of the bowel wall. However, respiratory and peristaltic motions complicate an easy analysis of such curves since spatial correspondence over time is lost. Therefore, a gated, 3D non-rigid motion correction method was developed that sustains robust extraction of the time intensity curves from bowel segments.

The algorithm worked in several steps: (1) images were selected from the breath-out phase since these contained the least blurring defects; (2) a non-rigid registration (alignment) procedure was adopted to compensate for misalignment, e.g. due to small peristaltic movements; such matching was performed to a position that had a minimum distance to the selected images, so that the result was unbiased; (3) the bowel wall was automatically extracted from manually indicated regions of interest.

TICs of ROIs in original, gated and registered data are compared in Figure 3. (b)-(d). The red region contains inflammatory bowel wall; the other regions are added for comparison. Notice the particularly smooth TICs of the registered data in Figure 3. (d). Clearly, the red curve supports the hypothesis for large intensity enhancement due to inflammation. A similar phenomenon was found in 12 other subjects in regions corresponding to active Crohn's disease. This might reflect that the inflammation information

could be distinguished by the TICs after our registration procedure.

The post-contrast images facilitate measurement of colonic wall thickness, which is typically increased with active Crohn's disease. As such the former images contain complementary information to DCE MRI. A non-rigid registration procedure was devised to match the DCE data to the post-contrast MR images to achieve correspondence. This algorithm proceeds from coarse to fine registration and matches based on the so-called mutual information [6]. This metric sustains a comparison of different contrasts. Typical results are shown in Figure 4.

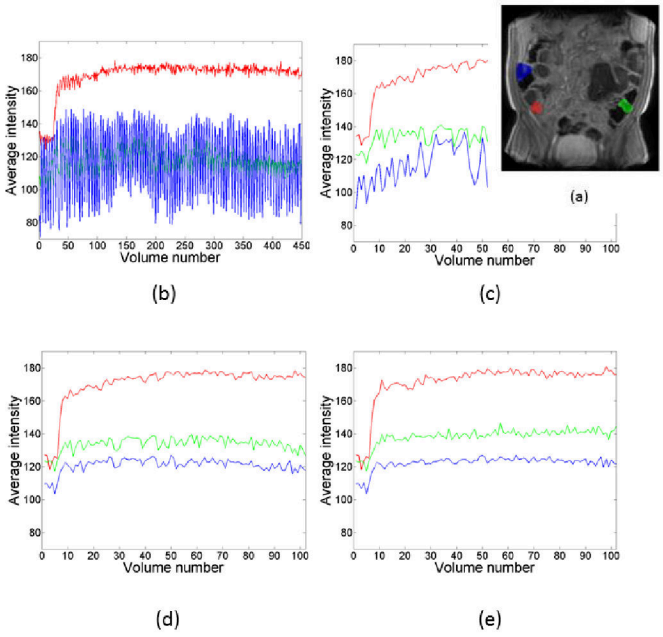

Figure 3. The TICs of a ROI. (a) is the annotated data, the red region contains a bowel segment affected by Crohn's disease. (b) - (d) show the TICs of original data, gated data and registered data (notice the reduced number of volumes in the gated dataset). (e) is a constructed TIC 'ground truth' image.

IV. Classification

We developed a framework in which a regression or classification algorithm assessed Crohn's disease severity associated with an entire MRI volume or a local region within this volume. Therefore, an expert segmented all (584) areas within our MRI data that were affected by CD. The diseased areas served as input for predicting the previously mentioned disease severity measures (see Data acquisition).

Oriented Gabor filters along six orientations (0◦; 30◦; 60◦; 90◦; 120◦; 150◦) were used to obtain texture features from all annotated regions. A 7 dimensional vector represented an MRI scan with several diseased regions containing (1) size (in voxels), (2)-(4) mean, minimum and maximum intensity, (5)-(7) mean, minimum and maximum of the texture features (see above). Each feature was calculated over all diseased regions.

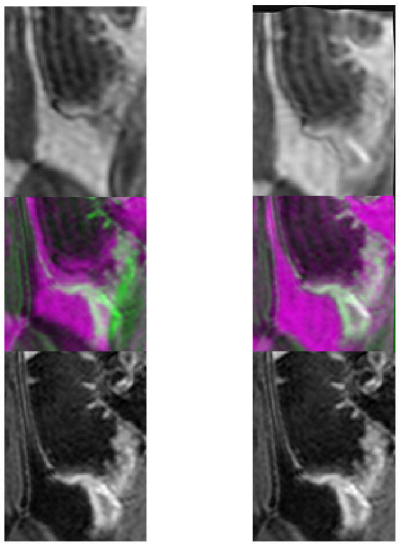

Figure 4. Example of the DCE to post-contrast image registration. The DCE image (top) is encoded in the red and blue channels (middle); the post-contrast image (bottom) in the green channel (middle). Left is prior to, right after registration.

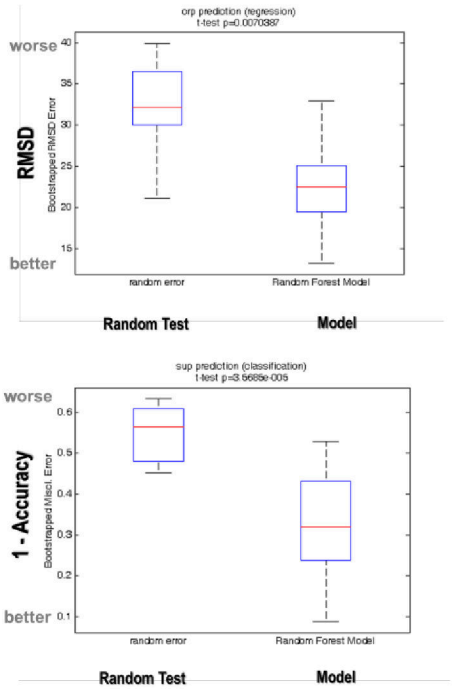

Figure 5. Top: CRP prediction of 24 patients in 10-fold bootstrap validation. Bottom: CDEIS prediction performs significantly better than random.

A Random Forest (RF) classifier with 50 trees was trained on the MRI data to predict a disease severity score[7]. For CRP and CDEIS [2], the RF was essentially a regressor, due to the continuous nature of the score. When the scores were sparse and binary (e.g. for particular CDEIS_subscores), the RF was effectively a classifier.

The RF was tested in a bootstrapped cross-validation. Repeatedly, *n* samples were drawn with replacement out of the samples in the dataset. These drawn bootstrap samples trained the RF, whereas the remaining out-of-bag samples

were used for testing. For the regression systems, the cost function to be minimized was the RMSD (root mean square deviation). For the classification systems, the cost function was the classification accuracy. This procedure was repeated 10 times.

The classification performance of our system was compared with a random severity assignment to determine if it significantly deviated from a random classification. The random model consisted of the same dataset and bootstrap cross-validation, except for that the labels of the samples were randomly permuted before analysis.

Classification/regression performance which was better than random was achieved for CRP prediction and a CDEIS sub-score, namely the amount of superficial ulcerations present (see Figure 5.). On average, the bootstrap error for the latter feature was 30 ± 10%.

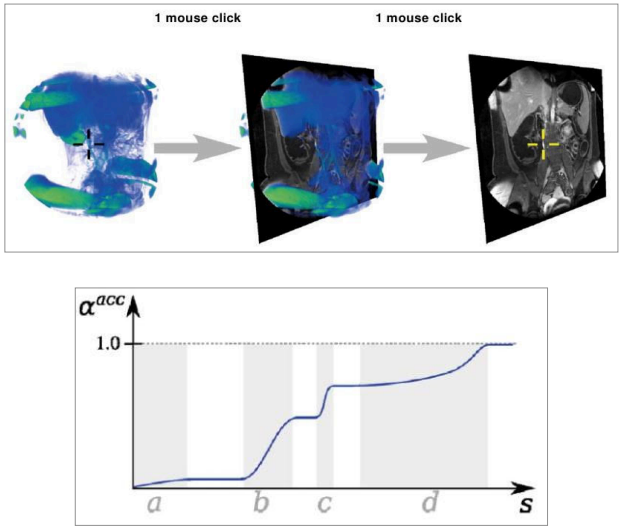

Figure 6. Top: Perception-oriented picking in volumetric rendering of bowel MRI. Bottom: Profile of accumulated opacity along the viewing ray.

V. Visualization

We developed a method for picking structures in 3D volumetric renderings by clicking in the 2D screen (see Fig. 2 left) to sustain visual inspection of the data. The picking selected the structures that were most visible to the observer, that is, it imitated perception by the human visual system [8]. The structures that contributed most to the accumulated opacity (α^{acc}) along the viewing ray were considered to be the most visible. To detect these structures we found the highest jump of α^{acc} along the ray. In the example of Figure 6. (bottom) this is the interval labeled b. To pick the actually perceived structures in the rendering is an improvement over previous methods, which either use meta data to predict the intended 3D position or can only pick structures above a certain opacity threshold. The latter is problematic for foggy renderings and very transparent structures. Automatically picking the most visible structures allowed for a very intuitive navigation through the data. So far the picking was used for easily placing cutting planes or selecting slices of the data for visualization.

The top illustration in Figure 6. shows the two-step handling for inspecting slices. The first step consisted of pointing on the interesting structure followed by clicking. This selected the slice that was then displayed together with the volumetric rendering to provide context. A second click removed the volumetric rendering to provide an unobstructed view of the slice. From the selected interval we knew the extent of the structure (along the ray). Thus, we could provide two viewing modes. One selected the front most position of the selected structure and a second one selected its center. The first was closer to what the user perceives and the latter provided a more informative view because it cut through the structure.

VI. Conclusion

There is a clear demand for semi-automated evaluation of MR images in CD. Currently, a computer-assisted diagnosis tool for automatic detection of abnormalities, ability to grade disease severity, and therewith influence clinical disease management based on MRI is missing. Development of such a system is a complex task, particularly due to the signal fluctuation inherent to MRI. Moreover, the limited thickness of the bowel wall and the presence of peristalsis further complicate the development of new techniques. A combination of (semi-)automated segmentation and different registration techniques to identify, respectively align, regions of interest in MRI images would be extremely useful. This should facilitate the measurement of descriptive properties of CD activity in the images and the application of machine learning techniques to detect and rank abnormalities. In turn, the latter would support the establishment of a combined, objective and quantitative disease severity index.

Computational modeling for assessment of IBD: to be or not to be? Well, this paper demonstrated promising preliminary results of such an assessment.

References

[1] J. Rimola, S. Rodriguez, O. García-Bosch, I. Ordás, E. Ayala, Aceituno M, et al. "Magnetic resonance for assessment of disease activity and severity in ileocolonic Crohn's disease," *Gut*, vol. 58, pp. 1113-20, 2009.

[2] J.Y. Mary and R. Modigliani, "Development and validation of an endoscopic index of the severity for Crohn's disease: a prospective multicentre study", *Gut*, vol. 30, pp. 983-989, 1989.

[3] M.L.W. Ziech, S. Bipat,J.J.T.H. Roelofs,C.Y. Nio, B. Mearadji, S. van Doorn, et al, "Retrospective comparison of magnetic resonance imaging features and histopathology in Crohn's disease patients," *Eur J Radiol*, vol. 80, pp. 299-305, 2011.

[4] K. Horsthuis, S. Bipat, P.C.F. Stokkers, J. Stoker, "Magnetic resonance imaging for evaluation of disease activity in Crohn's disease: a systematic review," *Eur Radiol*, vol. 19, pp. 1450-60, 2009.

[5] T. Brox, D. Cremers, "On local Region Models and a Statistical Interpretation of the Piecewise Smooth Mumford-Shah Functional," *Int J Comput Vision*, vol. 84, pp. 184-193, 2009.

[6] S. Klein., M. Staring., K.Murphy., M. Viergever, J. Pluim:"Elastix: a toolbox for intensity based medical image registration," *IEEE Trans. Med Imaging*, vol. 29, pp.196–205, 2010.

[7] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5-32, 2001.

[8] A. Wiebel, F.M. Vos, H.C. Hege, "Perception-Oriented Picking of Structures in Direct Volumetric Renderings," ZIB-Report 11-45, December 2011.