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

Quantitative Texture Analysis of MRI images for Detection of Cartilage-Related Bone Marrow Edema

Tong Kuan Chuah, Chueh Loo Poh, Kenneth Sheah

Abstract—**This paper presents a study that investigated the potential of texture analysis using Fluid Sensitive Fat Suppressed MRI images for the use in detection of bone marrow edema. A total of 168 slices of knee MRI from 10 subjects were involved. Six histogram-based textures (mean intensity, standard deviation, smoothness, third moment, uniformity and entropy) were calculated in both 2D and 3D, and were compared between healthy group and group affected by bone marrow edema. Two-sample *t*-tests were performed to assess the difference between healthy group and group affected by edema. The intensity third moment in 2D showed significant difference between the slices of healthy subjects and the slices affected by edema ($p<0.05$). Smoothness and standard deviation in 2D showed a modest significance between healthy and affected groups. No significant difference was found in the 3D textures of healthy group and group affected by edema.**

I. Introduction

Bone marrow edema, also known as bone marrow lesion (BML) is an important feature appearing in many different diseases of the knee. The diseases associated with BML include osteoarthritis (OA), arthritis, complex regional pain syndrome, osteonecrosis, osteochondritis dissecans, and bone bruise, and even tumor or transient bone marrow edema syndrome [1]. Increasing size of BML is associated with the development of knee pain [2]. Study about subchondral BML in patient with OA revealed that absence of BML is associated with decreased risk of cartilage loss whereas regions with high cartilage loss often has appearance or enlargement of BML [3].

Magnetic Resonance Imaging (MRI) has been used extensively to visualize soft tissues of the knees such as BML. This imaging modality is valuable because it has excellent contrast between hard and soft tissues. Hence, ligaments, meniscus, cartilage and bone marrow can be effectively visualized. However, much of the diagnosis and/or assessment of BML are done directly by the radiologist by looking through all the slices because quantitative tools are not yet readily available, easy to use and reliable.

Assessment of BML has been carried out both quantitatively [4-5] and semiquantitatively using scoring system [6]. BML appears as bright patch(s) in bone marrow of Fat Suppressed T2-weighted (T2w-FS) magnetic resonance (MR) image without defined boundary. Recently, Dijkstra at al. has reported an attempt to delineate the BML region from the surrounding tissues in the T2w-SPIR images and the result is encouraging [7]. The segmentation method of BML reported by Dijkstra et al. is semiautomatic, based on K-nearest neighbor classification that takes into account the intensity and spatial location of the pixels. However, in this method, the slices affected by BML needs to be identified by the operator prior to running the segmentation algorithm, which involves considerable work of perusing all the slices.

To address the issue of having to select the slices affected by BML, and having to screen patients who do not have BML from those who have, this paper presents a study that investigated parameters potentially suitable to be used for classifying image slices and/or knees (whole knee in 3D) affected by BML from those not affected by BML. Texture parameters derived from the histogram of the bone marrow of femur were analyzed. Histogram-based texture parameters are chosen because they characterize the overall intensity distribution at the macro level and BML appears at the macro level in the bone marrow.

II. Methods

A. Images

Retrospective MRI datasets were studied. These include five knees from patient without BML (healthy subjects) and five knees from patients with BML were included in the study. Five of the subjects are female. The ages of subjects included in the study range from 33 to 64 years. The weight range of the subjects is 64 to 110 kg.

Overall, there is a total of 168 slices of 2D sagittal knee MRI images (Fluid Sensitive Fat Suppressed sequences) all acquired at 3 Tesla. Fluid Sensitive Fat Suppressed images used in this study include proton density fat saturated images and T2w-FS images but they have the similar appearance. These two types of images are used practically the same way in the diagnosis process, so in this paper, no distinction is made for the two types of images. Instead, focus is directed

The work is supported by Ministry of Education, Singapore, through AcRF Tier 1 grant RG34/07. T. K. Chuah is financially supported by Nanyang President Graduate Scholarship for his current graduate study.

T. K. Chuah, and C. L. Poh are with the Division of Bioengineering, Nanyang Technological University, Singapore (corresponding author e-mail: CLPoh@ntu.edu.sg).

K. Sheah is with Radlink Diagnostic Imaging, Singapore.

to analyzing the textures of the images. The in-plane matrix size of the images is 512 × 512. The in-plane pixel spacing ranges from 0.2539 mm to 0.2930 mm. The number of slices of each subject ranges from 19 to 23 slices. Slice spacing ranges from 4 to 5 mm. All volumes were acquired with a slice thickness of 3 mm. All images were originally stored in 16-bit. They were converted to 8-bit for analysis.

B. *Analysis*

Bone marrow of the femur, which is the region of interest (ROI) in this study, was manually segmented by expert operator for each of the slices analyzed. All slices manually segmented were validated by the radiologist. Six texture parameters derived from image intensity were calculated for the ROI of each slice. The parameters are (i) Mean Intensity, (ii) Standard Deviation (abbreviated as Stdev), (iii) Smoothness, (iv) Third Moment, (v) Uniformity, and (vi) Entropy. The texture parameters are defined as follows [8]:

$$\text{Mean} = \quad m \quad = \sum_{i=0}^{L-1} z_i p(z_i)$$

$$\text{Stdev} = \quad \sigma \quad = \sqrt{\sum_{i=0}^{L-1} (z_i - m)^2 p(z_i)}$$

$$\text{Smoothness} \quad = R = 1 - \frac{1}{1 + [\sigma/(L-1)]^2}$$

$$\text{Third Moment} = \mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$$

$$\text{Uniformity} = \quad U \quad = \sum_{i=0}^{L-1} p^2(z_i)$$

$$\text{Entropy} \quad = \quad e \quad = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$$

where L is the total number of intensity levels. Its value is 255 in this study. The variable z_i is the particular intensity level at i-th count. The variable $p(z_i)$ is the probability of the particular intensity level. All texture parameters were calculated using a code written in MATLAB (The Mathworks, Inc., Natick, MA, USA).

To illustrate the reason histogram-based parameters are chosen, Fig. 1 shows a slice from healthy subject and a slice affected by BML, both with their corresponding histogram. From the MRI image of the slice affected by BML, there is clearly a brighter patch of pixels appearing at the distal part of femur, indicating BML. In contrast, in the healthy slice, no such brighter patch is visible. The brighter patch in pathological femur gives the histogram its distinct shape that differs from that of healthy femur. It can be seen that the difference in the shape of histogram could be characterized by texture parameters and the texture parameters can play an important role in distinguishing between healthy or BML effected slices.

Two types of texture analysis were done: (a) textures based on volumetric intensity in 3D and (b) textures based on 2D slices. Analyzing 2D textures enables us to distinguish slices from affected subjects and healthy subjects, whereas analyzing 3D textures enables us to distinguish between healthy and affected subjects. For the 3D texture analysis, the intensities of all pixels in the whole volume of femur bone marrow of a patient are included in calculating a single texture value. The 3D texture values from healthy subjects were then compared with the 3D texture values from the subjects affected by BML using 2-sample t-test. All t-tests were carried out using MATLAB. By carrying out t-test to test how significant the texture is different between healthy subject (Healthy Volume) and affected subject (Affected Volume), we are able to study the ability of each parameter in separating healthy subjects from affected subjects. Both versions of t-tests, equal variance and unequal variance, were performed.

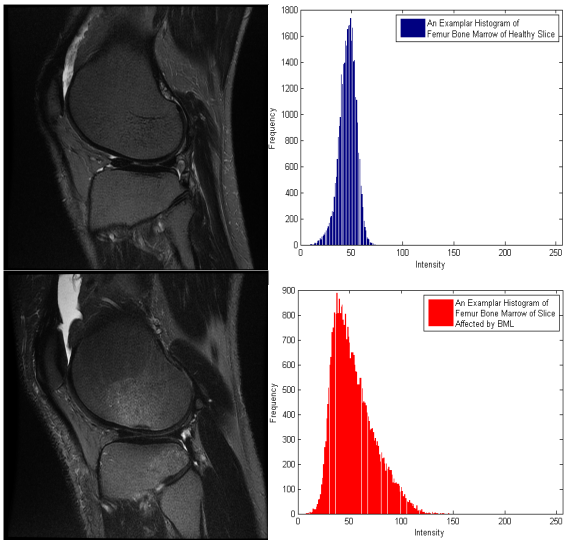

Fig. 1. Example of healthy slice with its histogram of the femur bone marrow (top); and a slice affected by BML at the femur bone marrow with its histogram (bottom). Note the bright patch at distal femur representing BML.

There is inconsistent variation of image intensity across different image slices in the same volume acquired. This variation of image intensity across slices in acquired image can influence the ability of texture parameters in characterizing difference in 3D image accurately. This variation is difficult to standardize without affecting the information in the images. In an attempt to remove this variation and to make intensities across the images more uniform, every 2D image in a volume was processed with window/level adjustment algorithm. In this algorithm, window and level are adjusted. The window is defined at the 99th percentile of the image histogram and the level is defined as the mid-point of the image. This window/level algorithm was applied in the 3D texture study.

For the 2D texture analysis, all the slices were separated into four groups: **(A)** Slices from healthy subjects, not affected by BML, this group has got 81 slices. **(B)** Slices from the subjects affected by BML. This group has got 87 slices. **(C)** Slices identified to be affected by BML, by a consultant radiologist. This group has got 15 slices. **(D)**

Healthy slices from the subjects affected by BML. This group has got 72 slices. Texture parameters were calculated for each slice, and were studied for their potential ability to separate between groups **(A) & (B)**, between **(A) & (C)**, and between **(A) & (D)**. The set relationships of the groups are shown in Fig. 2. For each pair of groups, both versions of *t*-tests including equal variance and unequal variance were performed to examine the difference in texture parameters in each pair.

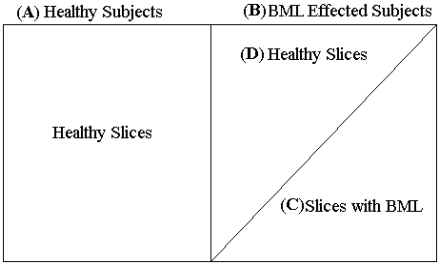

Fig. 2. Set representation of the subjects and slices studied.

III. Results

In the 3D texture analysis of this study, the *p*-values as a result of the *t*-test of 3D textures are shown in Table 1. In the analysis of this table, the texture parameters were calculated based on the 8-bit version of the acquired MRI images with and without window/level adjustment.

Table 1. Results of *t*-test of 3D texture parameters represented by the *p*-values. Textures calculated using 8-bit image.

Healthy Volume vs Affected Volume

	m	σ	R	μ_3	U	e
Equal Variance						
without Window/Level	0.919	0.060	0.056	0.386	0.114	0.105
with Window/Level	0.488	0.458	0.538	0.089	0.493	0.512
Unequal Variance						
without Window/Level	0.920	0.067	0.057	0.403	0.144	0.129
with Window/Level	0.489	0.469	0.544	0.124	0.504	0.523

In the cases without window/level adjustment, the results show no significant difference (at the level of significance 0.05) between all the 3D textures of healthy volume and affected volumes. Nevertheless, trend towards statistical significance were noted for smoothness and standard deviation. The absence of significant difference in 3D texture between healthy and affected volumes could be due to the inconsistent pixel intensities in different slices in the same volume.

Referring to the values in Table 1 that use the volume of images after window/level adjustment, similarly, the healthy subjects and subjects with BML did not show significant difference in any of the 3D texture parameters. Window/level adjustment might have undesirably reduced most of the 3D texture difference between healthy volume and affected volume.

In the 2D texture analysis, the *p*-values as a result of the *t*-test of slice-by-slice 2D textures are shown in Table 2 for unequal variance and Table 3 for equal variance.

Table 2. The *p*-values of unequal variance *t*-test of 2D texture parameters between different groups.

Unequal Variance

	m	σ	R	μ_3	U	e
(A) Healthy Subjects vs (B) Affected Subjects	0.721	0.096	0.070	0.024*	0.408	0.264
(A) Healthy Subjects (Slices) vs (C) Affected Slices	0.340	0.070	0.091	0.025*	0.085	0.068
(A) Healthy Subjects (Slices) vs (D) Healthy Slices of Affected Subjects	0.874	0.285	0.244	0.117	0.672	0.533

*Significant *p*-values at the level of 0.05.

Table 3. The *p*-values of equal variance *t*-test of 2D texture parameters between different groups.

Equal Variance

	m	σ	R	μ_3	U	e
(A) Healthy Subjects vs (B) Affected Subjects	0.719	0.098	0.073	0.028*	0.408	0.264
(A) Healthy Subjects (Slices) vs (C) Affected Slices	0.509	0.015*	0.006**	<0.001**	0.134	0.068
(A) Healthy Subjects (Slices) vs (D) Healthy Slices of Affected Subjects	0.875	0.283	0.239	0.096	0.670	0.534

*Significant *p*-values at the level of 0.05.
**Significant *p*-values at the level <0.01.

By the results of both equal variance and unequal variance *t*-tests, third moment showed significant difference between healthy slices and BML-affected slices, **(A) & (B)**; and also between healthy slices and slices of affected subjects, **(A) & (C)**. The 95% confidence interval of difference of mean between groups **(A) & (B)** is -0.029 to -0.002.

In equal variance *t*-test, groups **(A) & (C)** showed significant difference in smoothness and standard deviation. The healthy slices from affected subjects showed **(D)** no significant texture difference from the slices from healthy subjects **(A)**. This finding agrees with the notion that there is no visible difference in healthy slices regardless of whether they are from healthy subject or from a subject affected by BML.

IV. Discussion

The results showed that that third moment, which measures primarily the skewness of the histogram, is a strong parameter to be used for distinguishing a slice affected by BML from a healthy slice. The reason this parameter showed significant difference between healthy and affected slices is because healthy slices tend to have negative or zero skewness in the histograms of femur bone marrow, whereas affected slices usually have positive skewness. This explains why the confidence interval of difference in means between groups **(A) & (B)** is negative and does not include zero. The significant difference of the third moment of these two groups suggests that thresholding of this texture parameter could produce a high accuracy in screening of healthy slice from slices affected by BML.

It is not certain whether any of the parameters studied has equal variance in any two groups of interest. There could be a fair possibility that the two groups have equal variance if the sample size is large enough. Therefore in this paper both equal variance and unequal variance tests were carried out. The equal variance tests can be regarded as less stringent tests than the unequal variance tests. With assumption of equal variance, more of the parameters studied showed significant difference between healthy slices and slices affected by BML, **(A) & (C)**. Intensity smoothness and standard deviation were also showing significant difference. Standard deviation is a measure of average contrast. A significant difference in this value between the groups agrees with the notion that the contrasts are different for a healthy bone marrow and the bone marrow with BML.

Although standard deviation and smoothness did not show to be a very strong parameter to be used solely or independently to screen healthy slices from affected slices, they have good potential to be used in a multivariate classification algorithm for distinguishing the two. Moreover, these two parameters also showed trend towards statistical significance in 3D texture analysis. The study included 3D texture analysis and 2D analysis and the results suggested that 3D texture is not as robust as 2D texture in distinguishing healthy and BML effected images.

This study has explored the possibility of using texture parameters to tell slices affected by BML from healthy slices. Three histogram-based texture parameters are found to have potential. This study did not include the finding of the threshold of third moment for detecting BML, which would be a reasonable next step for testing the usefulness of the parameter. The thresholding is expected to work well and it shall address the issue of slice selection for segmentation of BML such as that in [7].

One limitation of using first order texture parameters, which are based on histogram, is that they do not characterize the spatial relationship of pixels of different intensities. Thus the parameters identified in this study do not provide distinction between different anomalies of the bone marrow that result in the same intensity distribution. For example, bone cysts also produce brighter patches on bone marrow and use of first order texture will not be effective in detecting their differences. Furthermore, in Fluid Sensitive Fat Suppressed MRI image, the bone marrow occasionally is imaged with some artifacts appeared as white stripes. This artifact also reduces the accuracy of using histogram-based texture parameters in distinguishing pathology from normality.

The investigation in this paper involves a limited sample size for 3D analysis but it has been able to demonstrate the difficulty in using 3D texture for slices acquired in 2D, with considerable inter-slice gaps. For 2D analysis, the sample size for healthy group is reasonably large; and the sample size for affected slices is of fair size.

V. CONCLUSION

This study has analyzed the histogram-based texture parameters in their ability of distinguishing healthy slices and slices affected by BML. Third moment of intensity is found to be of great potential for thresholding application because it showed significant difference between slices affected by BML and slices of healthy subjects. The significant difference is as expected because there is clear distinction in the skewness of the histograms in Fig. 1. Smoothness and standard deviation are two parameters found to be potentially useful because they showed significant difference in the case of equal variance *t*-tests. Future work will be to determine the threshold of using the texture parameter identified, and to test its accuracy of slice selection in detecting BML.

ACKNOWLEDGEMENTS

The authors would like to thank Mr. Lim Jun Hong for his assistance in manual segmentation of the femur.

REFERENCES

[1] S. Hofmann, J. Kramer, M. Breitenseher, M. Pietsch, and N. Aigner, *Bone marrow edema in the knee. Differential diagnosis and therapeutic possibilities.* Knochenmarködem im kniegelenk. Differenzialdiagnostik und therapeutische möglichkeiten, 2006. **35**(4): p. 463-477.

[2] D.T. Felson, J. Niu, A. Guermazi, F. Roemer, P. Aliabadi, M. Clancy, J. Torner, C.E. Lewis, and M.C. Nevitt, *Correlation of the development of knee pain with enlarging bone marrow lesions on magnetic resonance imaging.* Arthritis and Rheumatism, 2007. **56**(9): p. 2986-2992.

[3] F.W. Roemer, A. Guermazi, M.K. Javaid, J.A. Lynch, J. Niu, Y. Zhang, D.T. Felson, C.E. Lewis, J. Torner, and M.C. Nevitt, *Change in MRI-detected subchondral bone marrow lesions is associated with cartilage loss: The MOST Study. A longitudinal multicentre study of knee osteoarthritis.* Annals of the Rheumatic Diseases, 2009. **68**(9): p. 1461-1465.

[4] M.E. Mayerhoefer, M. Breitenseher, S. Hofmann, N. Aigner, R. Meizer, H. Siedentop, and J. Kramer, *Computer-assisted quantitative analysis of bone marrow edema of the knee: Initial experience with a new method.* American Journal of Roentgenology, 2004. **182**(6): p. 1399-1403.

[5] M.E. Mayerhoefer, M.J. Breitenseher, J. Kramer, N. Aigner, C. Norden, and S. Hofmann, *STIR vs. T1-weighted fat-suppressed gadolinium-enhanced MRI of bone marrow edema of the knee: Computer-assisted quantitative comparison and influence of injected contrast media volume and aquisition parameters.* Journal of Magnetic Resonance Imaging, 2005. **22**(6): p. 788-793.

[6] C.G. Peterfy, A. Guermazi, S. Zaim, P.F.J. Tirman, Y. Miaux, D. White, M. Kothari, Y. Lu, K. Fye, S. Zhao, and H.K. Genant, *Whole-organ magnetic resonance imaging score (WORMS) of the knee in osteoarthritis.* Osteoarthritis and Cartilage, 2004. **12**(3): p. 177-190.

[7] A.J. Dijkstra, P. Anbeek, K.G. Auw Yang, K.L. Vincken, M.A. Viergever, R.M. Castelein, and D.B.F. Saris, *Validation of a Novel Semiautomated Segmentation Method for MRI Detection of Cartilage-Related Bone Marrow Lesions.* Cartilage, 2010. **1**(4): p. 328-334.

[8] R.C. Gonzalez, R.E. Woods, and S.L. Eddins, *Digital Image Processing Using MATLAB*. p. 466.