Local bone enhancement fuzzy clustering for segmentation of MR trabecular bone images

Jenny Folkesson^{a)} and Julio Carballido-Gamio

Department of Radiology and Biomedical Imaging, Musculoskeletal and Quantitative Imaging Research Group (MQIR), University of California, San Francisco, California 94158

Felix Eckstein

Institute of Anatomy and Musculoskeletal Research, Paracelsus Medical University, Salzburg, A5020, Austria

Thomas M. Link and Sharmila Majumdar

Department of Radiology and Biomedical Imaging, Musculoskeletal and Quantitative Imaging Research Group (MQIR), University of California, San Francisco, California 94158

(Received 22 June 2009; revised 31 August 2009; accepted for publication 27 October 2009; published 15 December 2009)

Purpose: Segmentation of trabecular bone from magnetic resonance (MR) images is a challenging task due to spatial resolution limitations, signal-to-noise ratio constraints, and signal intensity inhomogeneities. This article examines an alternative approach to trabecular bone segmentation using partial membership segmentation termed fuzzy *C*-means clustering incorporating local second order features for bone enhancement (BE-FCM) at multiple scales. This approach is meant to allow for a soft segmentation that accounts for partial volume effects while suppressing the influence of noise.

Methods: A soft segmentation method was developed and evaluated on three different sets of data; interscan reproducibility was evaluated on six test-retest *in vivo* MR scans of the proximal femur, correlation between MR and HR-pQCT measurements was evaluated on 49 *in vivo* scans from the distal tibia, and the potential for fracture discrimination was evaluated using MR scans of calcaneus specimens from 15 participants with and 15 participants without vertebral fracture. The algorithm was compared to fuzzy clustering using the intensity as the only feature (I-FCM) and a dual thresholding algorithm. The metric evaluated was bone volume over total volume (BV/TV) within user-defined regions of interest.

Results: BE-FCM had a higher interscan reproducibility (rms CV: 2.0%) compared to I-FCM (5.6%) and thresholding (4.2%), and expressed higher correlation to HR-pQCT data ($r=0.79$, $p < 10^{-11}$) compared to I-FCM ($r=0.74$, $p < 10^{-8}$) and thresholding ($r=0.70$, $p < 10^{-6}$). BE-FCM was also the method that was best able to differentiate between a control and a vertebral fracture group at a 95% significance level.

Conclusions: The results suggest that trabecular bone segmentation by BE-FCM can provide a precise BV/TV measurement that is sensitive to pathology. The segmentation method may become useful in MR imaging-based quantification of bone microarchitecture. © 2010 American Association of Physicists in Medicine. [DOI: [10.1118/1.3264615](https://doi.org/10.1118/1.3264615)]

Key words: trabecular bone, segmentation, fuzzy clustering

I. INTRODUCTION

Osteoporosis is a condition of the bone characterized by a decrease in mass, deterioration of the microstructure, and an increase in fracture incidence.¹ The current gold standard for clinical determination of osteoporosis status is bone mineral density (BMD) measurements obtained through x-ray based imaging techniques. This measurement estimates the combined trabecular and cortical bone density, where cortical bone is the outer shell of long bones, enclosing marrow and an intricate network of trabecular bone. Although the ability to predict bone strength from BMD has been established,² its capability to assess fracture risk or detect responses to therapeutic interventions is rather limited.³ Therefore, there has been an increased focus in recent osteoporosis research on

quantitative analysis of bone microstructure as a complement to BMD for a better characterization of the bone and its mechanical competence.

Magnetic resonance (MR) imaging provides the capability to quantitatively analyze bone microstructure in humans *in vivo* without the use of ionizing radiation.^{4–6} In this type of imaging data, the marrow appears as high intensity voxels and the trabecular structure is determined indirectly from the low signal content in proximity to the marrow.

The first and most fundamental step in trabecular bone structure analysis is the two-class segmentation of bone from marrow in a predefined region of interest (ROI) surrounded by the cortical shell. *In vivo* image acquisition of osteoporosis-related sites such as the proximal femur, and tibia is typically performed with phased array coils due to signal-to-noise ratio (SNR) constraints, causing intensity in-

homogeneities in the resulting images, which may be misleading the segmentation algorithm. Another challenge is the limitation in the achievable spatial resolution, which is similar to the dimension of the trabeculae (78–200 μm) in-plane, and often lower in the slice direction, which causes substantial partial volume effects. Therefore, a soft segmentation that measures the probability or partial class membership of each voxel is preferred to providing voxels with a hard label. While incorporation of labeled training data can greatly improve the segmentation performance in many medical imaging applications, incorporating such prior information is not feasible for the thin network of the trabeculae.

One commonly used segmentation method in osteoporosis studies is an empirical thresholding approach,⁵ where a global threshold is set between the intensity at the full width at half maximum of the intensity histogram of the ROI and an estimate of the bone intensity determined manually from the cortical shell. The coil-induced bias field is estimated using a low pass filtering scheme,⁷ which is sensitive to edge artifacts. Thresholding might be suitable for trabecular bone segmentation in high resolution peripheral quantitative computed tomography (HR-pQCT) data,⁸ but when designing a trabecular bone segmentation algorithm for MR images, care must be taken to counteract partial volume effects, noise, and intensity inhomogeneities.

Hwang and Wehrli⁹ used an iterative deconvolution framework to counteract noise broadening of the ROI histogram while estimating intensity inhomogeneities using local intensity averaging. Local bone volume fraction was determined from the estimated noiseless image divided by the local marrow intensity average. While the method accounts for partial volume effects, noise, and intensity inhomogeneities, it assumes that the histogram is bimodal, which is typically not the case for resolutions achievable *in vivo*. It was later found to frequently fail at identifying thin or weakly connected trabeculae.¹⁰

Vasilic and Wehrli¹⁰ later presented an alternative method based on local thresholding, in order to better determine the bone marrow signal intensity. Relying on the assumption that the image intensity and the local 2D discrete Laplacian values are correlated, the method estimated the local marrow intensities within a nearest neighbor framework. The algorithm was adapted to monomodal histograms and intensity inhomogeneities, but retained the noise in the thresholded data.

Carballido-Gamio *et al.*¹¹ used fuzzy clustering in trabecular bone segmentation. The method assigned a partial membership to each voxel based on the distance from the cluster center which accounts for partial volume effects. However, the only feature used was the intensity, which made the method sensitive to noise and intensity inhomogeneities.

In this article, we present a novel approach as an extension to the fuzzy clustering bone segmentation of Carballido-Gamio *et al.*¹¹ Besides the signal intensity, the method incorporates a local bone enhancement feature at multiple scales. The local second order structure is computed within a scale-

space framework, which suppresses noise while at the same time enhancing local relative intensity anisotropy. This permits the method to account for partial volume effects, noise, and to some extent also for signal intensity inhomogeneity. The evaluation of the method will be performed in three steps using different samples: The precision will be evaluated on test-retest *in vivo* scans of proximal femurs. The method will then be further evaluated by correlating the estimated fraction of the total ROI volume containing bone [bone volume over total volume (BV/TV)] from MR scans to corresponding measurements from HR-pQCT scans of the tibia. Finally, it will be put to test whether the method is capable of differentiating subjects with and those without vertebral fractures using MR images of the calcaneus.

II. METHODS

II.A. Coil correction

In order to correct of intensity inhomogeneities we preprocess the data with the N3 method described by Sled *et al.*¹² The underlying assumption of N3 is that the bias field blurs the intensity distribution, and in order to restore the unbiased image distribution a solution in the form of iterative deconvolution of the intensity distribution and a smoothing of the bias field estimate using *B* splines is employed. It is generic in the sense that it does not require any training data, nor prior knowledge of the number of different tissues in the image, nor the intensity distributions of those classes. We use a MATLAB (The Mathworks, Inc., Natick, MA, USA) implementation of N3.¹³

II.B. Multiscale bone enhancement

A common approach to examine the local structure of an image $I(x)$ is to consider a Taylor expansion around the point x_0 ,

$$I(x) = I(x_0) + (x - x_0)^T \nabla I(x_0) + \frac{1}{2} (x - x_0)^T \nabla^2 I(x_0) (x - x_0), \quad (1)$$

where $\nabla I(x_0)$ and $\nabla^2 I(x_0)$ denote the gradient vector and the Hessian, respectively. Hence the local second order structure of an image can be described by the intensity, the gradient, and the Hessian.

The Hessian (H) encode second order intensity information. For a 3D image the Hessian is described by

$$H(\sigma) = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix}, \quad (2)$$

where σ annotates a scale dependence. Since the Hessian is based on second order information it has a strong response at the center of the isointensity structure over the scale. For our application where the transition between the two classes is blurred due to partial volume effects the Hessian is preferable to first order measures such as the structure tensor,¹⁴ which has a strong response at structure boundaries.

FIG. 1. Trabecular bone enhancement in the tibia also seen in Fig. 3 on three scales ($\sigma=0.1, 0.15, 0.25$ mm), increasing from left to right.

We use Gaussian derivatives, defined as $I_{x_1, \dots, x_n} = I * \mathcal{D}_{x_1, \dots, x_n} G(\sigma)$, where G is a Gaussian and $\mathcal{D}$ is a differential operator,¹⁵ since the spatial averaging in scale-space causes a noise reduction that counteracts the noise amplification caused by differentiation.¹⁶

The eigenvalues of the Hessian ($\lambda_1, \lambda_2, \lambda_3$, $\lambda_1 \geq \lambda_2 \geq \lambda_3$) can characterize the resemblance of the local structure to a tube ($\lambda_1 \approx \lambda_2 \gg \lambda_3$), a sheet ($\lambda_1 \approx \lambda_2 \approx \lambda_3$), and a blob ($\lambda_1 \approx \lambda_2 \approx \lambda_3$). The eigenvalues of the Hessian have for instance been used to characterize tubelike objects for vessel segmentation,¹⁷ and sheetlike objects in sinus bone segmentation.¹⁸

The image resolution is fairly anisotropic in trabecular bone MR imaging, with an in-plane resolution typically four times higher compared to that in the slice direction. Therefore, we perform a 2D analysis of the local structure and have adopted the vessel filtering of Frangi *et al.*¹⁷ to enhance dark-appearing thin trabecular bone structure as follows. Using two geometric ratios, $\mathcal{R}_{\text{blob}} = |\lambda_2|/|\lambda_1|$ and $\mathcal{R}_{\text{noise}} = \sqrt{\lambda_1^2 + \lambda_2^2}$, a measure for bone enhancement is described by

$$\mathcal{B}(\sigma) = \begin{cases} 0 & \text{if } \lambda_1 < 0 \\ \exp\left(-\frac{\mathcal{R}_{\text{blob}}^2}{2\alpha^2}\right) \left(1 - \exp\left(-\frac{\mathcal{R}_{\text{noise}}^2}{2\beta^2}\right)\right) & \text{otherwise,} \end{cases} \quad (3)$$

where α is set to 0.5, and β is half the maximum Frobenius norm. Examples of bone enhancement on three different scales can be seen in Fig. 1.

II.C. Fuzzy C-means clustering

Voxel-based cluster analysis partitions voxels into groups according to some similarity measure. Clustering methods such as K -means assign a point to the cluster with the nearest centroid. Fuzzy C -means (FCM) clustering extends the technique by also determining a degree of cluster membership for each voxel. FCM was originally described by Dunn¹⁹ for the special case of two clusters, and later generalized by Bezdek²⁰ for multiple classes. Assuming n voxels, each represented by p features, then FCM classifies the n feature vectors ($v_1, \dots, v_n \in \mathbb{R}^p$) into one of k fuzzy clusters. The partial membership μ_{ij} for v_j of the i th cluster is required to meet the conditions $0 \leq \mu_{ij} \leq 1$, $\forall i, j$, in addition to $\sum_{i=1}^k \mu_{ij} = 1$, $\forall j$ and $0 < \sum_{j=1}^n \mu_{ij} < n$, $\forall i$. This is realized by minimizing the cost function

$$J = \sum_{j=1}^n \sum_{i=1}^k \mu_{ij}^m \|v_j - u_i\|^2, \quad (4)$$

where u_i is the centroid of the i th cluster and m is a heuristic constant controlling the membership “fuzziness.” For $m=2$, which is the value of m we use in this work, the coefficients are normalized to make their sum 1. FCM clustering implemented in MATLAB and C (Ref. 11) is applied to the trabecular bone images using a four-dimensional feature space, hence $p=4$ and $k=2$. The four features included in the bone enhancement fuzzy clustering (BE-FCM) algorithm are the intensity, and the bone enhancement feature $\mathcal{B}(\sigma)$ on three scales: $\sigma=0.1, 0.15, 0.25$ mm. The three scales are chosen to cover different sized features of trabeculae, as illustrated in Fig. 1.

II.D. Computation time

The N3 coil correction has an average run time of 2.6 min for a $512 \times 512 \times 74$ voxel volume on a Linux x86-64 2.5 GHz, 8 GB processor. On the same machine, the average run time for the FCM algorithm is less than a minute for a typical ROI size of less than 10^6 data points.

II.E. Subjects and specimens

Six healthy males were recruited for the interscan reproducibility evaluation (mean age 26 yr). As for the comparison of BV/TV measurements obtained from MR and HR-pQCT imaging, 49 postmenopausal women defined as osteopenic by WHO criteria were examined (mean age 55 yr). After informing the subjects on the nature of their respective studies, a written consent was signed in accordance with the regulations of the UCSF Committee of Human Research.

For the fracture discrimination evaluation, calcaneal from 30 formalin-fixed human cadavers (15 females, 15 males, mean age 82 ± 10 yr, Institute of Anatomy at Ludwig Maximilians University, Munich, Germany). Individuals with bone disease other than osteoporosis or osteopenia were excluded based on conventional histomorphometry of iliac crest biopsies. Specimen procedures were in accordance with local and institutional legislative guidelines.

II.F. Imaging

II.F.1. Reproducibility

For the six healthy male volunteers, MR images of the proximal femur were acquired with a 3T GE Signa scanner using a FIESTA-C [multiacquisition fully balanced steady-state free precession (bSSFP)] sequence, with a modified version of generalized autocalibrating partially parallel acquisition, with an acceleration factor of two.²¹ A four-element phased array coil was used, and scan parameters include TR/TE 11.7/4.6 ms, flip angle 60° , bandwidth ± 31.25 kHz, and resolution $0.234 \times 0.234 \times 1.0$ mm³. This sequence has been evaluated for trabecular bone imaging and shown to compare well with steady-state 3D-spin-echo and fast gradient recalled echo (FGRE) sequences in comparison

FIG. 2. Examples of the three different data sets that were used in this study. Top left: A coronal slice of a proximal femur scan, where the ROI is determined by the trochanter. Top right: The same slice preprocessed with N3 coil correction. Bottom left: An axial slice of the tibia, where the ROI is drawn within the cortical shell. Bottom right: A sagittal slice depicting the calcaneus, where the ROI is selected as a cylinder at a central location.

to HR-pQCT.²² Scans were performed twice on the same day, with a removal of the subjects from the scanning table and a repositioning of the coils between scans. Examples can be seen in Fig. 2, before and after application of coil correction.

II.F.2. MR and HR-pQCT correlation

The comparison of MR and HR-pQCT and MR imaging was evaluated in the tibia for the postmenopausal females. MR image acquisition was performed on a 3T Signa scanner using a four-element phased array coil and a bSSFP sequence with TR/TE 16.8–17.8/6.5 ms, flip angle 60°, bandwidth ± 122 Hz, and resolution $0.156 \times 0.156 \times 0.5$ mm³. An example is shown in Figs. 2 and 3. HR-pQCT images were acquired using an XtremeCT *in vivo* scanner (Scanco Medical AG, Brüttisellen, Switzerland) with source potential 60 kVP, tube current 900 μ A, integration time 100 ms, and isotropic 82 μ m resolution.²³ An example HR-pQCT slice can be seen in Fig. 4. For both MR and HR-pQCT data the ROIs were defined to be just inside the cortical bone, and for both modalities the ROI was defined to be from 22.5 to 31.5 mm proximal to the end plate in the tibia.

II.F.3. Fracture discrimination

The calcaneus specimens were imaged at 3T (Signa, GE Medical Systems, Milwaukee, WI, USA) with a two-element phased array wrist coil. The sequence was a 3D FGRE, with TR/TE 18.5/4.3 ms, 20° flip angle, 12.5 kHz BW, and resolution $156 \mu\text{m} \times 156 \mu\text{m} \times 0.5$ mm. Circular ROIs avoiding cortical bone and air artifacts were placed in the posterior part of the calcanei (see Figs. 2 and 5). Assessment of vertebral fracture status was made from radiographs of the en-

FIG. 3. Fuzzy BV/TV of one of the six healthy young males for examining reproducibility. From left to right, top to bottom: The slice from Fig. 2 magnified around the ROI, the thresholded image, intensity based fuzzy clustering, and local bone enhancement fuzzy clustering.

tire thoracic and lumbar spines specimens, with a resulting number of 17 fracture specimens. Examples of fracture and nonfracture specimens can be seen in Fig. 6. The BMD for the calcanei was determined using dual x-ray absorption using a GE/Lunar (Milwaukee, WI, USA) Prodigy scanner. The heel region was scanned with a forearm algorithm using a pencil-beam x-ray mode, and BMD (g/cm³) was measured in a circular region of interest in the posterior part of the calcanei.

FIG. 4. Fuzzy BV/TV of the tibia. From left to right, top to bottom: The slice from Fig. 2 magnified around the ROI, the thresholded image, intensity based fuzzy clustering, and local bone enhancement fuzzy clustering.

FIG. 5. Left: Slice from HR-pQCT scan of the tibia from the same subject as in Fig. 3. Right: The corresponding thresholded image.

II.G. Image analysis

The MR images are preprocessed with N3 coil correction as described in Sec. II A prior to segmentation using BE-FCM with $B(0.1, 0.15, 0.25 \text{ mm})$ and intensity I as features. In addition, the images are segmented using fuzzy clustering with the intensity as the only feature [intensity based fuzzy clustering (I-FCM)], and a dual thresholding (T) described in Ref. 5. The apparent fraction of bone per volume element (BV/TV) within the predefined ROIs is used as a metric for the segmentation; $BV/TV = 1/n \sum_j \mu_{\text{bone},j}$.

The interscan reproducibility of the apparent BV/TV in the *in vivo* proximal femur MR scans are evaluated using the root mean squared coefficient of variation (CV)

$$CV = \sqrt{\frac{1}{2n} \sum_{i=1}^n d_i^2 / \frac{1}{n} \sum_{i=1}^n \bar{x}_i} \cdot 100\%, \quad (5)$$

where n is the number of scan pairs, d_i is the difference between the measurements in the first and second visit, and $\bar{x}_i$ is the mean value for the i th scan pair.²⁴

Fracture discrimination ability is evaluated using an unpaired student's t test on a 95% confidence level for the apparent BV/TV of the calcaneus data, separated into two different groups based on vertebral fracture status.

The relation between MR and HR-pQCT images of the tibia is investigated using the Pearson's product-moment correlation coefficient (r) between the BV/TV from MR and HR-pQCT data. The significance of the correlations are established using a two-tailed student's t test on a 95% confidence level.

FIG. 6. Examples of calcaneus specimens with and without fracture. Left: A sagittal slice from an MR scan of the calcaneus from the control group. Right: A similar slice from a calcaneus associated with a vertebral fracture.

TABLE I. Evaluation of interscan reproducibility (Repro.), correlation between MR and HR-pQCT data (Corr.), and fracture discrimination (Discr.). Correlation is evaluated using r , p value, and linear regression equation (l.r.). The values are based on BV/TV obtained from the segmentation methods thresholding (T), I-FCM, and BE-FCM.

Features	Repro. (CV) (%)	Corr. r	($p <$)	l.r. ($y =$)	Discr. ($p <$)
T	4.2	0.70	10^{-6}	$0.40x - 0.05$	0.34
I-FCM	5.6	0.74	10^{-8}	$0.44x - 0.12$	0.05
BE-FCM	2.0	0.79	10^{-11}	$0.96x - 0.22$	0.02
n	6		49		30

III. RESULTS

As can be seen in Table I, the interscan reproducibility of apparent BV/TV measurements from *in vivo* proximal femur MR scans is higher when using BE-FCM (2.0%) compared to thresholding (4.2%) and I-FCM (5.6%). The higher reproducibility of BE-FCM is significant only on a 90% confidence level ($p < 0.09$) according to a paired t test, most likely due to the small sample size ($n=6$). There is no significant difference between the reproducibility using thresholding and I-FCM ($p < 0.6$).

The correlation between BV/TV measurements from HR-pQCT and MR data is higher and more significant for MR scans segmented with BE-FCM compared to both I-FCM and thresholding. A least-squares line fit of the correlated data shows that the first order coefficient is closer to unity for the values obtained with BE-FCM ($y = 0.96x - 0.22$) compared to I-FCM ($y = 0.44x - 0.12$) and thresholding ($y = 0.40x - 0.05$). In the linear regression model the MR and HR-pQCT data are on the x and y axis, respectively.

In the fracture discrimination study, the BE-FCM and I-FCM-based measurements show similar performance in ability to separate fracture versus nonfracture groups, where the discrimination is significant on a 95% confidence level. They both perform better than threshold-based apparent BV/TV ($p < 0.34$). Using the calcaneus BMD, the discrimination was significant ($p < 0.03$). The apparent BV/TV and BMD measurements are all higher in the control group compared to the fracture group. All MR-based measurements were significantly correlated with BMD, BE-FCM was most highly correlated with BMD ($r = 0.78$, $p < 10^{-6}$), closely followed by I-FCM ($r = 0.76$, $p < 10^{-6}$), followed by thresholding ($r = 0.69$, $p < 10^{-4}$).

The mean apparent BV/TV range between 0.31 and 0.34 for the three different data sets for the BE-FCM-based method. This can be compared to the I-FCM-based (0.49–0.53) and threshold-based measurements (0.24–0.42). These results are summarized in Table II. The standard deviation is lower for the BE-FCM measurements, with the highest standard deviation for the calcaneus data set, which is has the most heterogeneous population. The mean BV/TV values obtained for the HR-pQCT data of the tibia is 0.11 (± 0.02), which can be compared to the MR values for the same site which is 0.34, 0.53, and 0.42 for BE-FCM, I-FCM, and

TABLE II. Mean values (standard deviations) of MR-based apparent BV/TV measurements of the three data sets (femur, tibia, and calcaneus) obtained from the segmentation methods thresholding (*T*), I-FCM, and BE-FCM. The calcaneus specimens are divided into total, fracture group, and control group.

	Femur	Tibia	Calcaneus	Fracture	Control
<i>T</i>	0.24(0.02)	0.42(0.04)	0.38(0.05)	0.38(0.05)	0.40(0.05)
I-FCM	0.50(0.03)	0.53(0.04)	0.49(0.05)	0.48(0.05)	0.51(0.05)
BE-FCM	0.31(0.01)	0.34(0.02)	0.32(0.03)	0.31(0.03)	0.33(0.03)
<i>n</i>	6	49	30	15	15

thresholding, respectively. Example segmentations from the different data sets are demonstrated in Figs. 7, 3, and 5.

IV. DISCUSSION

In this study we evaluate apparent BV/TV measurements obtained from trabecular bone segmentation by fuzzy clustering with multiscale bone enhancement features, and compare it to fuzzy clustering using intensity as the only feature, and a conventional dual thresholding approach. The measurements are evaluated on three different data sets; the interscan reproducibility is evaluated on repeat *in vivo* MR scans of the proximal femur, correlation between MR and HR-pQCT measurements are evaluated on *in vivo* distal tibia scans, and fracture discrimination ability is evaluated on MR scans of calcaneus specimens with and without signs of vertebral fracture. For all three data sets, the BE-FCM demonstrate competitive performance. It is shown to have higher interscan reproducibility (2.0%) compared to I-FCM (5.6%) and thresholding (4.2%), and to be better correlated with HR-pQCT data ($r=0.79$, $p < 10^{-11}$) compared to I-FCM ($r=0.74$, $p < 10^{-8}$) and thresholding ($r=0.70$, $p < 10^{-6}$). BE-FCM is also method that differentiate best between control and fracture groups at a 95% significance level.

FIG. 7. Fuzzy BV/TV of the calcaneus. From left to right, top to bottom: The slice from Fig. 2 magnified around the ROI, the thresholded image, intensity based fuzzy clustering, and local bone enhancement fuzzy clustering.

The correlation between MR and HR-pQCT based BV/TV measurements correspond to those found by Kazakia *et al.*²³ for the threshold-based method ($r^2=0.5$) on the same data set. A preliminary study of correlations of trabecular bone parameters between MR and HR-pQCT data was performed by Krug *et al.*²² *In vivo* MR and HR-pQCT scans of the distal tibia of 6 healthy volunteers were compared, giving a correlation coefficient of $r=0.50$ as determined from Table V in that article. This correlation was not significant, probably due to the small number of samples. The correspondence is reasonable considering the difference in sample size.

The mean BV/TV values reported by Krug *et al.*²² are 0.41 ± 0.02 and 0.29 ± 0.07 for *in vivo* MR and HR-pQCT based measurements of the tibia, respectively. These values can be compared to those found by Kazakia *et al.*²³ who reported mean values 0.42 ± 0.04 and 0.11 ± 0.02 for MR and HR-pQCT based BV/TV measurements of the tibia, respectively. The mean BV/TV measurement from HR-pQCT determined by Krug *et al.*²² are nearly threefold that of Kazakia *et al.*,²³ where the latter can be considered more reliable due to the larger number of samples. Compared to this mean HR-pQCT based BV/TV value, the apparent BV/TV values from MR scans are threefold to fourfold higher. Reasons for this discrepancy may be related both to image acquisition differences between MR and HR-pQCT and differences in image analysis. MR images are typically acquired with a much higher in-plane resolution compared to slice thickness, and measures bone indirectly by signal voids in proximity to soft tissue such as marrow. HR-pQCT images on the other hand typically have isotropic spatial resolution and detect mineralization by x-ray attenuation. The slope of the fitted regression line for the I-FCM and thresholding derived values express some bias compared to HR-pQCT values, whereas the slope from BE-FCM-based values is close to unity (Table I). The larger variation in apparent BV/TV values from I-FCM and thresholding segmentation could be due to more global intensity variations within the region of interest, e.g., from coil shading undetected by the coil correction algorithm or different types of marrow. The bone enhancement filtering in the BE-FCM algorithm detects local intensity changes and is therefore less sensitive to more global intensity variations compared to the methods solely based on intensity magnitude.

Kazakia *et al.*²³ have compiled studies relating BV/TV measurements from μ CT, which can be considered gold standard due to the high resolution and good contrast of the

modality, demonstrating that BV/TV measurements from MR and HR-pQCT have been found to have similar agreement with μ CT with correlations of $r^2=0.87$ and $r^2=0.86$, respectively. The results also indicate that MR-based trabecular bone measurements may be overestimated (MR/ μ CT ratio = 3.1 ± 0.9), while HR-pQCT measurements may be underestimated (HR-pQCT/ μ CT ratio = 0.6). The apparent BV/TV measurements based on BE-FCM segmentation described in this work is found to have a lower mean value compared to thresholding for the tibia data set (0.34 versus 0.42). The mean apparent BV/TV value from BE-FCM is also more consistent between different imaging sites compared to thresholding.

As can be seen in Figs. 7, 3, and 5, BE-FCM captures more bone structure while being less affected by partial volume effects, noise, and intensity inhomogeneities. The bone enhancement is particularly prominent in the proximal femur, a site where fractures frequently occur but is more challenging to image compared to the extremities. Currently, MR is the only imaging modality for assessment of trabecular bone structure *in vivo* at this location.²⁵

This study demonstrates the capability of obtaining a soft trabecular bone segmentation using fuzzy clustering. Future work will involve developing analysis of other established trabecular bone parameters such as trabecular thickness, spacing, and number within a gray scale framework. The analysis is currently performed in 2D due to the difference between in-plane resolution and slice thickness for the MR scans, but the method is directly extendible to 3D.

Various extensions to the fuzzy *C*-means algorithm have been reported since the pioneering work of Dunn¹⁹ and Bezdek.²⁰ Some recent methods had modified the algorithm by incorporating neighborhood information.^{27,28} By generating hard labels, these methods have demonstrated competitive performance for unsupervised segmentation of MR brain images compared to manual segmentations of the tissues. In trabecular bone segmentation, the distribution of bone within the region of interest is irregular and inhomogeneous and the tissue is highly affected by partial volume effects. Therefore, such modifications to the FCM algorithm might be unsuitable for trabecular bone segmentation. Rather than modifying the fuzzy *C*-means algorithm, the focus of our approach is on the incorporation of bone enhancement filtering as features to improve FCM-based segmentation of trabecular bone.

Vasilic *et al.*²⁶ used eigenvalue decomposition of the tensor of inertia in a spherical neighborhood of resampled MR images to evaluate the effect of anisotropic voxel sizes on trabecular microarchitecture. With this local structure measure they show that the sensitivity to platelike and rodlike structures decrease with increasing slice distance.

While the method developed by Vasilic and Wehrli¹⁰ provides an interesting alternative for incorporating discrete local second order information, it retains noise in the soft segmentation of bone. While it shows less SNR dependence and higher bone connectivity based on Euler characteristics compared to a previous method by Hwang and Wehrli,⁹ the Euler characteristics may be affected by the noise and the evalua-

tion does not establish its ability to detect pathology or reproducibility. In this work, the linear scale-space framework ensures that derivatives are obtained in a well-posed manner, which makes the local structure analysis numerically robust.

In conclusion, we propose a trabecular bone segmentation method that is based on fuzzy clustering, using bone enhancement features. These account for partial volume effects, noise, and intensity inhomogeneities. The method is shown to be more precise (reproducible), more highly correlated with HR-pQCT and to better discriminate between participants with and without vertebral fractures than the other evaluated approaches. The method may thus have potential to become a valuable component in imaging-based studies related to the prediction of fracture risk and to the effect of treatment in osteoporosis.

ACKNOWLEDGMENTS

The authors would like to thank Dana Carpenter at UCSF for the acquisition and ROI definition of the femur MR scans, Galateia J. Kazakia, Andrew J. Burghardt, Benedict Hyun, and Roland Krug for the image acquisition and ROI definition of the tibia image data. The authors would also like to thank Eva Maria Lochmüller from the University Hospital of Gynecology, LMU, München, Germany, for her support through the German Research Foundation Grant Nos. 730/3-1 and 3-2.

^a)Electronic mail: jenny.folkesson@radiology.ucsf.edu

¹R. Baillon *et al.*, "Consensus development conference: Prophylaxis and treatment of osteoporosis," *Am. J. Med.* **90**(1), 107–110 (1991).

²S. Cummings, D. Black, M. Newitt, and W. Browner, "Bone density at various sites for prediction of hip fractures," *Lancet* **341**, 72–75 (1993).

³D. P. Delmas and E. Seeman, "Changes in bone mineral density explain little of the reduction in vertebral or nonvertebral fracture risk with anti-resorptive therapy," *Bone (N.Y.)* **34**(4), 599–604 (2004).

⁴F. W. Wehrli, "Structural and functional assessment of trabecular and cortical bone by micro magnetic resonance imaging," *J. Magn. Reson. Imaging* **25**, 390–409 (2007).

⁵S. Majumdar, H. K. Genant, S. Grampp, D. C. Newitt, V.-H. Truong, J. C. Lin, and A. Mathur, "Correlation of trabecular bone structure with age, bone mineral density, and osteoporotic status: In vivo studies in the distal radius using high resolution magnetic resonance imaging," *J. Bone Miner. Res.* **12**(1), 111–118 (1997).

⁶T. M. Link, S. Majumdar, P. Augat, J. C. Lin, D. C. Newitt, Y. Lu, N. E. Lane, and H. K. Genant, "In vivo high resolution MRI of the calcaneus: Differences in trabecular structure in osteoporosis patients," *J. Bone Miner. Res.* **13**, 1175–1182 (1998).

⁷D. C. Newitt, B. van Rietbergen, and S. Majumdar, "Processing and analysis of in vivo high-resolution MR images of trabecular bone for longitudinal studies: Reproducibility of structural measures and micro-finite element analysis derived mechanical properties," *Osteoporosis Int.* **13**, 278–287 (2002).

⁸T. L. Mueller, M. Stauber, T. Kohler, F. Eckstein, R. Müller, and G. H. van Lenthe, "Non-invasive bone competence analysis by high-resolution pQCT: An in vitro reproducibility study on structural and mechanical properties at the human radius," *Bone (N.Y.)* **44**(2), 364–371 (2009).

⁹S. N. Hwang and F. W. Wehrli, "Estimating voxel volume fractions of trabecular bone on the basis of magnetic resonance images acquired in vivo," *Int. J. Imaging Syst. Technol.* **10**, 186–198 (1999).

¹⁰B. Vasilic and F. W. Wehrli, "A novel local thresholding algorithm for trabecular bone volume fraction mapping in the limited spatial resolution regime of *in vivo* MRI," *IEEE Trans. Med. Imaging* **24**(12), 1574–1585 (2005).

¹¹J. Carballido-Gamio, C. Phan, T. M. Link, and S. Majumdar, "Characterization of trabecular bone structure from high-resolution magnetic reso-

- nance images using fuzzy logic," *Magn. Reson. Imaging* **24**, 1023–1029 (2006).
- ¹²J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Trans. Med. Imaging* **17**(1), 87–97 (1998).
 - ¹³J. Folkesson, R. Krug, J. Goldenstein, A. S. Issever, C. Fang, T. M. Link, and S. Majumdar, "Evaluation of correction methods for coil-induced intensity inhomogeneities and their influence on trabecular bone structure parameters from MR images," *Med. Phys.* **36**(4), 1267–1274 (2009).
 - ¹⁴C.-F. Westin, A. Bhalerao, H. Knutsson, and R. Kikinis, "Using local 3D structure for segmentation of bone from computer tomography images," Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR), 1997, pp. 794–800 (unpublished).
 - ¹⁵J. J. Koenderink, "The structure of images," *Biol. Cybern.* **50**, 363–370 (1984).
 - ¹⁶J. Blom, "Topological and geometrical aspects of image structure," Ph.D. thesis, Utrecht University, 1992.
 - ¹⁷A. Frangi, W. Niessen, K. L. Vincken, and M. A. Viergever, Proceedings of the International Conference on Medical Image Computing and Computer Aided Interventions (MICCAI), 1998, pp. 130–137 (unpublished).
 - ¹⁸M. Descoteaux, M. Audette, K. Chinzei, and K. Siddiqi, "Application to sinus bone segmentation and simulation of pituitary surgery," Proceedings of the International Conference on Medical Image Computing and Computer Aided Interventions (MICCAI): Bone Enhancement Filtering, 2005, pp. 9–16 (unpublished).
 - ¹⁹J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact, well separated clusters," *Journal of Cybernetics* **3**, 32–57 (1973).
 - ²⁰J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms* (Plenum, New York, 1981).
 - ²¹S. Banerjee, S. Choudhury, E. T. Han, A. C. S. Brau, C. V. Morze, D. B. Vigneron, and S. Majumdar, "Autocalibrating parallel imaging of in vivo trabecular bone microarchitecture at 3 Tesla," *Magn. Reson. Med.* **56**, 1075–1084 (2006).
 - ²²R. Krug, J. Carballido-Gamio, A. J. Burghardt, G. Kazakia, B. H. Hyun, B. Jobke, S. Banerjee, M. Huber, T. M. Link, and S. Majumdar, "Assessment of trabecular bone structure comparing magnetic resonance imaging at 3 Tesla with high-resolution peripheral quantitative computed tomography ex vivo and in vivo," *Osteoporos. Int.* **19**, 653–661 (2008).
 - ²³G. J. Kazakia, B. Hyun, A. J. Burghardt, D. C. Newitt, A. E. de Papp, T. M. Link, and S. Majumdar, "In vivo determination of bone structure in postmenopausal women: A comparison of HR-pQCT and high-field MR imaging," *J. Bone Miner. Res.* **23**(4), 463–474 (2008).
 - ²⁴C.-C. Glüer, G. Blake, Y. Lu, B. A. Blunt, M. Jergas, and H. K. Genant, "Accurate assessment of precision errors: How to measure the reproducibility of bone densitometry techniques," *Osteoporosis Int.* **5**, 262–270 (1995).
 - ²⁵R. Krug, S. Banerjee, E. T. Han, D. C. Newitt, T. M. Link, and S. Majumdar, "Feasibility of in vivo structural analysis of high-resolution magnetic resonance images of the proximal femur," *Osteoporosis Int.* **16**, 1307–1314 (2005).
 - ²⁶B. Vasilic, J. Magland, M. Wald, and F. W. Wehrli, "Advantages of isotropic voxel size for classification of trabecular bone struts and plates in micro-MR images," Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM), 2008, Vol. 16, p. 3627 (unpublished).
 - ²⁷H. Wang and B. Fei, "A modified fuzzy C-means classification method using a multiscale diffusion filtering scheme," *Med. Image Anal.* **13**, 193–202 (2009).
 - ²⁸M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty, "A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data," *IEEE Trans. Med. Imaging* **21**(3), 193–199 (2002).