Detecting Femur Fractures by Texture Analysis of Trabeculae

Dennis Wen-Hsiang Yap, Ying Chen, Wee Kheng Leow
Dept. of Computer Science, National University of Singapore
3 Science Drive 2, Singapore 117543
chenying, leowwk@comp.nus.edu.sg

Tet Sen Howe, Meng Ai Png
Dept. of Orthopaedics / Diagnostic Radiology, Singapore General Hospital
Outram Road, Singapore 169608
tshowe@yahoo.com, gdrpma@sgh.com.sg

Abstract

30% of women and 13% of men worldwide suffer from osteoporotic bone fractures worldwide. In large hospitals, doctors need to visually inspect a large number of x-ray images to identify the fracture cases, which typically constitute about 12% of all the x-ray images examined. Automated fracture detection can help to screen for obvious cases and flag suspicious cases for closer examinations. This paper describes a method of detecting femur fractures by analyzing trabecular texture patterns. Test results show that it is more accurate than an existing method based on neck-shaft angle. Moreover, combining the methods further improve of the overall performance of fracture detection.

1. Introduction

Many people suffer from bone fractures. The International Osteoporosis Foundation [4] reported that the lifetime risk of getting osteoporotic fractures in women is 30% – 40% and 13% in men worldwide. The number of hip fractures could rise from 1.7 million worldwide in 1990 to 6.3 million by 2050. The most dramatic increase is expected to occur in Asia during the next decades. World Health Organization confirms that osteoporosis is second only to cardiovascular disease as a leading health care problem [4].

In clinical practice, doctors and radiologists in hospitals manually inspect x-ray images to determine the occurrence and the precise nature of fractures. Typically, only a small fraction of the images that the radiologists inspect contain fractures. For example, in the x-ray images that we collected from a local hospital, only about 12% of the femurs are fractured. After looking through many images containing healthy femurs, a tired radiologist has been found to miss a fractured femur among the many healthy ones. A computer vision system can assist the doctors by screening the x-ray images and flagging suspicious cases that require closer examinations. Automated screening of x-ray images can thus help to direct the doctors' attention to suspicious cases and improve the timeliness and accuracy of their diagnosis. Therefore, such as computer vision system is extremely useful for clinicians and is now feasible because all clinical radiology is going digital. Digital x-ray images are now routinely captured using digital x-ray machines. This paper presents a method of detecting fractures at the upper extremity of the femur where most fractures occur.

2. Related Work

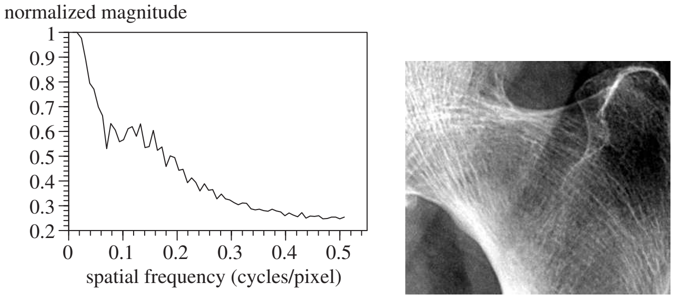
The only published work on the detection of fractures in x-ray images is that of Tian et al. [12]. Their method detects femur fractures by computing the angle between the neck axis and shaft axis, which can only detect severe fractures that cause a significant change in the neck-shaft angle.

The other related methods used non-visual techniques to detect fractures. For example, Ryder et al. analyzed acoustic pulses to determine whether a fracture has occurred [10]. Kaufman et al. analyzed mechanical vibration using a neural network model [5] whereas Singh and Chauhan measured electrical conductivity [11].

Most research effort on the analysis of orthopaedic x-ray images has been focused on the detection of osteoporosis by methods of texture analysis and fractal analysis. For example, [6, 9] computed first order statistics of texture including mean and standard deviation, and [8, 13] computed second order statistics of texture such as the co-occurrence matrix. [1, 2] derive fractal dimension and geometry for detecting osteoporosis. Caligiuri et al. [2] showed that their method appeared promising in discriminating between healthy and fractured bones. [7] applied fractal analysis on the micro x-ray images of human knees.

(a) (b)

Figure 1. (a) Severe fracture that changes the shape of the femur. (b) Fracture that does not change the femur's shape significantly.



(a) (b)

Figure 2. (a) Spatial frequency distribution of (b) trabeculae.

3. Fracture Detection by Texture Analysis

Our method of detecting femur fractures consists of three stages: (1) extraction of approximate femur contour, (2) analysis of trabecular texture, and (3) classification of femur. The extraction of femur contour in stage 1 is performed using active shape model, supplemented by active appearance models [3] at distinct feature points. Due to space limitation, this paper will focus on stages 2 and 3.

3.1. Texture Analysis of Trabecular Pattern

The trabeculae in the femur's upper extremity (Fig. 2(b)) are oriented at specific orientations to support the forces acting on the femur. A severe fracture breaks the femoral neck and causes a rotation and displacement of the femoral head. This results in a significant change in the orientations of the trabeculae at the femoral head and neck (Fig. 1(a)). On the other hand, some fractures may break the neck without rotating or displacing the femoral head significantly. They result in only local disruptions of the trabecular texture patterns (Fig. 1(b)). The method in [12] can detect the first kind of fractures but not the second kind. In contrast, our method is designed to detect both kinds of fractures.

Note that the location and size of the femur are not exactly identical in the input x-ray images. Even among healthy femurs, there are some differences in appearance and size due to differences in age and gender. One standard method of handling this problem is to normalize the location and size of the input images. However, this would either remove important texture information (if the image is shrunken) or introduce noise and artifacts (if the image is enlarged). Instead of scaling the input images, an adaptive method is used to sample the texture features so that the sampled locations in different images correspond to consistent locations in a normalized sampling grid.

Let W and H denote the width and height of the bounding box that contains the femur's upper extremity, and n_x and n_y denote the number of sampling locations along the x- and y-directions. Texture features are extracted within regions of width S_x and height S_y given by:

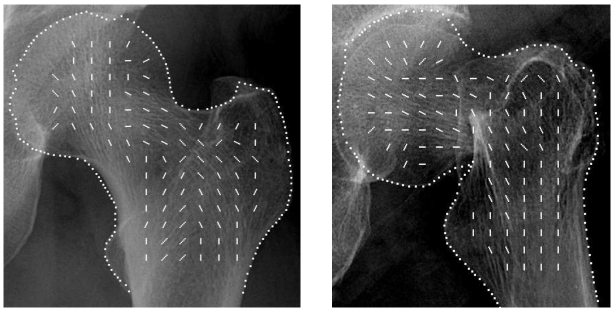
$$S_x = 2\left\lfloor \frac{W}{n_x+1} \right\rfloor \,, \quad S_y = 2\left\lfloor \frac{H}{n_y+1} \right\rfloor \,. \qquad (1)$$

In the current implementation, n_x is set to 12 and n_y to 14, and neighboring sampling regions overlap by 50% of their width and height.

A set of Gabor filters are applied to each region to extract the orientation of the texture pattern formed by the trabecular lines. A detailed study of 30 training samples shows that the frequencies of the texture patterns in different images fall within a narrow band (Fig. 2(a)). Ignoring the frequency components below 0.08 cycles/pixel, which are influenced by the DC component, the remaining components peak at around 0.13 cycles/pixel. So, a single frequency channel is sufficient. To clearly detect the main orientations of the texture patterns, 8 orientation channels are used, starting at $0°$ and incrementing at intervals of $22.5°$.

Each overlapping region of the x-ray image is convolved with the set of 8 Gabor filters. The dominant orientation of the texture in the region is identified by the orientation of the Gabor filter with the largest response. The orientation vectors outside the boundary contour of the femur's upper extremity are set to the null vector $\mathbf{0}$. Thus, an *orientation map* $\mathbf{M} = [\mathbf{u}_{ij}]$ (Fig. 3) can be obtained for each femur image, where $\mathbf{u}_{ij}$ is a unit vector that represents the orientation of the trabeculae at grid location (i, j).

To use the orientation map $\mathbf{M}$, which is a vector map, for classification, it is first converted into a scalar map by computing its difference with the *mean orientation map* $\mathcal{M}$ of healthy femurs. The mean orientation vector $\mathbf{m}_{ij}$ at po-



(a) (b)

Figure 3. Orientation maps of (a) healthy femur and (b) fractured femur. The short lines indicate trabecular orientations.

sition (i, j) in $\mathcal{M}$ is computed as:

$$\mathbf{m}_{ij} = \begin{cases} \displaystyle\sum_{k=1}^{n} \mathbf{u}_{kij} \left\| \sum_{k=1}^{n} \mathbf{u}_{kij} \right\|^{-1} & \text{if } c_{ij} > n/2 \\ \mathbf{0} & \text{otherwise} \end{cases} \quad (2)$$

where n is the number of healthy training samples, $\mathbf{u}_{kij}$ is the unit orientation vector of sample k at location (i, j), and c_{ij} is the number of samples with non-null orientation vectors at location (i, j). The difference v_{ij} at location (i, j) of the *difference map* $\mathbf{V}$ of an image is then computed as:

$$v_{ij} = \begin{cases} 0 & \text{if } \mathbf{u}_{ij} = \mathbf{m}_{ij} = \mathbf{0} \\ 1 - |\mathbf{u}_{ij} \cdot \mathbf{m}_{ij}| & \text{otherwise.} \end{cases} \quad (3)$$

For a healthy femur, the entries v_{ij} in the difference map should all be very close to 0 because its unit orientation vectors $\mathbf{u}_{ij}$ are very similar to the $\mathbf{m}_{ij}$ of the mean orientation map $\mathcal{M}$. On the other hand, for a fractured femur, many of its entries v_{ij} will have values much larger than 0 due to the disruption of texture orientations caused by the fracture.

3.2. Classification

Two classifiers are used to classify the test samples: Bayesian and Support Vector Machine (SVM).

For the first method, the sets of healthy and fracture training samples are each modeled by a multivariate Gaussian function, which are used to estimate the conditional probabilities $P(x|\,\text{healthy})$ and $P(x|\,\text{fracture})$ where x is a sample femur. We found that, in practice, the *a priori* probabilities $P(\text{healthy})$ and $P(\text{fracture})$ are not identical. Specifically, $P(\text{healthy})$ is roughly 0.88 and $P(\text{fracture})$ is 0.12. Applying Bayes' rule, we obtain

$$P(\text{class}|x) = \frac{P(x|\,\text{class})\,P(\text{class})}{P(x)} \quad (4)$$

where class is either healthy or fractured. The denominator $P(x)$ is the same for both $P(\text{healthy}|x)$ and $P(\text{fracture}|x)$ and so can be ignored. Thus, sample x can be classified as fractured if $P(\text{healthy}|x)$ is smaller than $P(\text{fracture}|x)$.

For the second method, SVM-light is applied to perform the classification.

4. Experimental Results

432 femur images were obtained from a hospital, and were divided randomly into 324 training and 108 testing images. The percentage of fractured images in the training and testing sets were kept approximately the same (12%). For the training set, 39 femurs were fractured, and for the testing set, 13 were fractured.

For the purpose of comparison, the fracture detection method of [12] based on neck-shaft angle (NSA) was also tested. Moreover, we also investigated the effectiveness of combining multiple classifiers for fracture detection. In particular, a femur was classified as fractured using one of the following three rules:

1. 1 of 2: Either one of two combined methods classified the femur as fractured. The three possible combinations are neck-shaft angle (NSA) with Bayesian, NSA with SVM, and Bayesian with SVM.
2. 1 of 3: Any one of the three methods classified the femur as fractured.
3. 2 of 3: Any two of the three methods classified the femur as fractured.

Three performance indices were computed:

- fracture detection rate: the number of correctly detected fractured femurs over the number of fractured femurs.
- false alarm rate: the number of wrongly classified healthy femurs over the number of healthy femurs.
- classification accuracy: the number of correctly classified femurs over the total number of femurs.

Table 1 illustrates the performance comparison. Both Bayesian and SVM are more accurate in classification than NSA. Bayesian has no false alarm but it is less effective in detecting fractures. SVM has the highest classification accuracy and fracture detection rate among the three methods.

Combining Bayesian with NSA or SVM improves both classification accuracy and fracture detection rate significantly. The 1-of-3 method has the highest fracture detection rate among the combined methods but a rather low classification accuracy and high false alarm rate. In contrast, the 2-of-3 method has the highest classification accuracy and no false alarm but the lowest fracture detection rate.

The remaining fractures that are not detected by the combined method are very subtle fractures that neither change

Table 1. Comparison of classification performance. NSA: neck-shaft angle method.

		texture		combined				
	NSA	Bayes	SVM	NSA or Bayes	NSA or SVM	Bayes or SVM	1 of 3	2 of 3
fracture detection rate	61.5%	46.2%	69.2%	76.9%	76.9%	76.9%	84.6%	61.5%
false alarm rate	3.2%	0.0%	3.2%	3.2%	6.3%	3.2%	6.3%	0.0%
classification accuracy	92.6%	93.5%	93.5%	94.4%	91.7%	94.4%	92.6%	95.4%

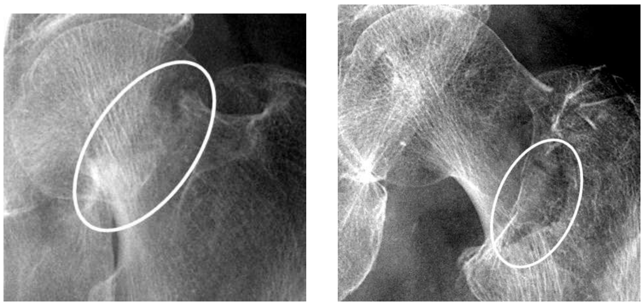


Figure 4. Subtle fractures (in white ellipse).

the femur's shape nor disrupt the trabecular pattern significantly (Fig. 4). For these fractures, other methods need to be developed to detect them. We have experimented with a method based on the shape of the 2D femur contour. But, this method does not work well because the femur's upper extremity has a complex shape. A moderate difference in the femur's orientation when it is captured in x-ray image can cause a significant change in the shape.

5. Conclusion

A method of detecting femur fractures by analyzing the trabecular texture pattern is developed. The method extracts an orientation map of the femur's upper extremity, and converts it into a scalar difference map by computing the difference between the femur's orientation map and the mean orientation map of healthy training samples. The difference map is then classified using Bayesian classifier and SVM. Experimental results show that SVM achieves higher classification accuracy and fracture detection rate than Bayesian method and neck-shaft angle (NSA) method. Combining Bayesian with SVM or NSA achieves the best overall performance in terms of high classification accuracy and fracture detection rate. Further research will be pursued to develop methods for detecting subtle fractures that are missed by the current methods.

References

[1] J. C. Buckland-Wright, J. A. Lynch, J. Rymer, and I. Fogelman. Fractal signature analysis of macroradiographs measures trabecular organization in lumbar vertebrae of postmenopausal women. *Calcif. Tissue Int.*, 54:106–112, 1994.

[2] P. Caligiuri, M. L. Giger, and M. Favus. Multifractal radiographic analysis of osteoporosis. *Medical Physics*, 21:503–508, 1994.

[3] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. In *Proc. European Conference on Computer Vision*, 1998.

[4] IOF. The facts about osteoporosis and its impact. International Osteoporosis Foundation, www.osteofound.org/press_centre/fact_sheet.html, 2002.

[5] J. J. Kaufman, A. Chiabrera, M. Hatem, N. Z. Hakim, M. Figueiredo, P. Nasser, S. Lattuga, A. A. Pilla, and R. S. Siffert. A neural network approach for bone fracture healing assessment. *IEEE Engineering in Medicine and Biology*, 9(3):23–30, 1990.

[6] T. Link, S. Majumdar, W. Konermann, N. Meier, J. Lin, D. Newitt, X. Ouyang, P. Peters, , and H. Genant. Texture analysis of direct magnification radiographs of vertebral specimens: Correlation with bone mineral density and biomechanical properties. *Acad. Radiol.*, 4:167–176, 1997.

[7] J. A. Lynch, D. J. Hawkes, and J. C. Buckland-Wright. Analysis of texture in macroradiographs of osteoarthritic knees using the fractal signature. *Phys. Med. Biol.*, 36(6):709–722, 1991.

[8] A. Materka, P. Cichy, and J. Tuliszkiewicz. Texture analysis of x-ray images for detection of changes in bone mass and structure. In M. K. Pietikäinen, editor, *Texture Analysis in Machine Vision*. World Scientific, 2000.

[9] X. Ouyang, S. Majumdar, T. M. Link, P. A. Y. Lu, J. C. Lin, D. C. Newitt, and H. K. Genant. Morphometric texture analysis of spinal trabecular bone structure assessed using orthogonal radiographic projections. *Medical Physics*, 25:2037–2045, 1998.

[10] D. M. Ryder, S. L. King, C. J. Olliff, and E. Davies. A possible method of monitoring bone fracture and bone characteristics using a non-invasive acoustic technique. In *Proc. Int. Conf. on Acoustic Sensing and Imaging*, pages 159–163, 1993.

[11] V. R. Singh and S. K. Chauhan. Early detection of fracture healing of a long bone for better mass health care. In *Proc. Annual Int. Conf. of IEEE Engineering in Medicine and Biology Society*, pages 2911–2912, 1998.

[12] T. P. Tian, Y. Chen, W. K. Leow, W. Hsu, T. S. Howe, and M. A. Png. Computing neck-shaft angle of femur for x-ray fracture detection. In *Proc. Int. Conf. on Computer Analysis of Images and Patterns (LNCS 2756)*, pages 82–89, 2003.

[13] J. Veenland, T. Link, W. Konermann, N. Meier, J. Grashuis, and E. Gelsema. Unraveling the role of structure and density in determining vertebral bone strength. *Calcif. Tissue Int.*, 61:474–479, 1997.