

VISION-BASED BONE IMAGE RECOGNITION USING GEOMETRIC PROPERTIES*

A. CHALECHALE^{1**}, A. BAHARI² AND M. VATANCHIAN³

¹Dept. of Computer, Faculty of Eng. Razi University, Kermanshah, I. R. of Iran
Email: chalechale@razi.ac.ir

²Dept. of Veterinary Medicine, Faculty of Paraveterinary Sciences, University of Bu Ali Sina, I. R. of Iran

³Dept. of Biology, Faculty of Sciences, University of Bu Ali Sina, I. R. of Iran

Abstract– An effective approach for bone image analysis, including segmentation and recognition is presented in this paper. Bone segmentation is carried out by K-means clustering of the bone image, whereas the recognition phase is based on feature extraction and two-level statistical classification. The proposed approach has applications in medicine and veterinary anatomy studies, orthopedics, paleontology and archaeology. Several image features, including geometric and moment invariants (regular and Zernike), are derived for recognition. The first-level classification is used to distinguish different kinds of bone and the second-level to recognize the right animal to which the bone belongs. Two-dimensional structures, called cluster-property and cluster-features matrices, have been employed to evaluate different bone characteristics. Experimental results for the first-level recognition exhibit better performance of the geometric features compared to moment invariants and Zernike moments. On the other hand, Zernike moments showed supremacy in differential diagnosis at the second level to recognize animals.

Keywords– Visual properties, bone recognition, geometric features

1. INTRODUCTION

Several aspects of bone image recognition using computerized methods and machine vision have been studied in the literature; however bone recognition is still a common problem [1-4]. This is due to the significant challenges in correctness, reliability and response time. Bone image analysis methods rely greatly on *pattern recognition*. Pattern recognition has found wide applications in many fields ranging from text processing and character recognition to speech, image and video analysis. In pattern recognition, it is often necessary to deal with problems to classify a transformed pattern. A correct classification/description of input data (i.e. raw or transformed patterns), based on either a priori knowledge or on statistical information, is the final goal of pattern recognition algorithms. In addition to correctness and effectiveness, efficiency (i.e. swiftness or being able to do the task in the requested time) is also an important aspect of pattern recognition strategies.

Most bone image analysis methods can be categorized into artificial intelligence, fuzzy and neural network-based approaches, knowledge-based approaches, and statistical classification techniques. The main aim of the approaches is to recognize human or animal bones properly, which is an essential process in anatomy and forensic studies, osteology, orthopedics, archaeology and paleontology. In recent years, several researches have been conducted to develop and describe reliable schemes for bone recognition and to diagnose the related diseases using medical imaging techniques and image processing algorithms [1, 2, 5, 6].

*Received by the editors September 3, 2009; Accepted November 20, 2010.

**Corresponding author

Neural network approach has been widely used for bone recognition. Suzuki et al. [2] have proposed an approach to suppress the contrast of ribs and clavicles in chest images by using a massive training artificial neural network (MTANN). Here, a multi-resolution MTANN is employed for three different images. The approach is based on various spatial frequencies in the ribs to detect lung nodules. A bone-image likeness, similar to the training bone images is produced first and subtracted from the corresponding chest to provide a soft-tissue image. This last image is exploited to recognize nodules.

Bone features are used to establish a correlation between bone structure and age. These features also provide information about age-related bone diseases. A fuzzy-based image processing algorithm is proposed in [1] for analysis of cross sections of human bones. The approach uses an adaptive neighborhood algorithm in conjunction with clustering and local covariance measures. Various bone features are extracted in this study and used to quantitatively correlate with age and, possibly, with age related bone diseases. Specifically, osteoporosis, an age related bone disorder, is considered to be recognized using the proposed method. Bone mass measurements techniques are used in most of the current approaches for monitoring bone conditions. However, these techniques do not completely describe the mechanisms to distinguish between osteoporotic and normal subjects. Structural parameters such as trabecular connectivity have been proposed in [6] as features for assessing bone conditions.

Geometric invariant moments have been widely used for shape recognition [7, 8]. Invariant moments and statistical classification have been used for gesture detection. This approach is employed for human-machine interface using visual-based communication [9]. Ng et al. [10] have proposed a system for automatic detection and recognition of human head movements. Invariant moments and hidden Markov model (HMM) are combined for feature extraction and recognition tasks, respectively. The main advantage of this approach is that it can operate in a relatively complex background. However, the computational requirements arising from the invariant moments extraction and HMM's application render the approach inappropriate for real-time applications where several images are involved. Zernike moments, on the other hand, using orthogonal basis functions, are less sensitive to noise than geometric moments and are more powerful in discriminating objects with n-fold symmetries. They are employed for building a region-based shape descriptor in the MPEG-7 standard [7].

An image processing system for bone image segmentation has been proposed by Liu et al. [5]. The system is based on using micro-structural and relational knowledge presented in the bone cross sections. It automates the bone image analysis and is able to produce results based on quantitative measurements from a large number of bone images. The existing correlation between bone structural features and age-related bone developments is realized using a dataset of bone images.

In this paper a new discipline on animal bone image recognition based on feature extraction and statistical classification is proposed. First, in the preprocessing step a fast and straightforward segmentation using K-means clustering is applied to the bone image. The result is then fed as the input to the feature extraction step. A two-level statistical classification is employed for recognition of the general structure of the bone image that is the first level recognition, and to recognize the animal which the bone image belongs to (the second level recognition). Experiments have been conducted to compare the recognition powers of three different sets of features, including regular and Zernike invariant moments, and the new proposed single and multi-valued geometric features.

Based on the background of the mathematical definition of the geometric and Zernike moments, the rest of the paper is organized as follows: the next section explains our approach in detail and Section 3 presents comparative results and discussion. Finally, Section 4 concludes the paper and poses some new guide lines.

2. COMPUTER-BASED RECOGNITION OF BONE IMAGES

One of the major areas in medical image processing is the analysis of bone images for a wide range of applications. General characteristics of a bone can be used, for instance, to infer the part of the body the bone belongs to. Moreover, in veterinary anatomy the details of the bone shape can be analyzed for differential diagnosis to determine the right animal that the bone is taken from. This section presents a novel two-level approach on bone image recognition. The first level is to recognize part of the body from the bone shape, and the second is to distinguish the animal based on anatomic characteristics of the bone (i.e. differential diagnosis).

The approach is based on feature extraction and statistical classification, where a bone image is first passed through a preprocessing step for bone segmentation. The segmented bone is then feature extracted using three different methods (i.e. geometric, regular and Zernike invariant moments). This is to determine the best method among the approaches with respect to their recognition power. The proposed paradigm for bone recognition is explained in detail as follows. It is important to note that the proposed procedure can be adopted for any other set of bone shapes without the need to change its general structure.

Initially, a collection of 1184 bone images is created and used in this study. The collection includes 6 different animal bone classes that are: scapula and femur of horse, ox, and sheep. Fig. 1 depicts some examples of the images in the collection. Full color bone images are next converted to gray intensity images by eliminating the hue and saturation components, while retaining the luminance. For further image processing steps the images need to be converted to binary. Due to varying lighting conditions of the images, using a unique threshold to binarize images is inadequate. Fig. 2 shows instances where a unique threshold causes inappropriate segmentation of the bone image. To avoid this, K-means clustering [11] is employed for binarization in the preprocessing stage. This successfully segments bone images from the background.

Size normalization using nearest neighbor interpolation is applied next. This is to achieve the scale invariance property, which allows varying-size bones to have similar features. The bounding box of the region of interest is found first and then normalized to $w \times h$ pixels (256×256 pixels in our experiments).

Afterwards, for each segmented-normalized bone image b belonging to bone class $B_i, i = 1 \dots I; J$ shape properties $P_j, j = 1 \dots J$ are extracted. Currently, for the bone collection, I is 6 and J is chosen to be 12 corresponding to 6 bone clusters and 12 predominant bone properties, respectively. The properties include five geometric and seven invariant moment-based functions. Geometric properties are: area (ar), perimeter (pr), major axis length (mj), minor axis length (mi), and eccentricity (ec).

Obviously, different species have different shapes in their bones. To overcome this problem we need to, firstly, define different clusters for different species, secondly, provide training data, and then set up the classifiers accordingly. In other words, we need to take several photos for those species we are interested in, then employ them in the training phase and find appropriate parameters applicable for recognition.

To determine the recognition power of each property, a classification scheme using the properties P_j is exploited. Here, classification of 300 randomly selected bone images (50 of each group) into the associated groups is attempted. Recognition rates R_{ij} for $i = 1 \dots I$ and $j = 1 \dots J$ are obtained and saved in appropriate entries in a *cluster property* matrix. The classification is based on Bayesian rule assuming Gaussian distribution for the bone patterns. To extract a decision function for our classifier, J number of 1D probability density functions are obtained. Each function involves I pattern groups governed by Gaussian densities, with means m_{ij} and standard deviation σ_{ij} . Therefore, the Bayes decision function has the following form:

$$d_{ij}(b) = p(b|B_i)P(B_i) \quad (1)$$

that is identical as

$$d_{ij}(b) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} e^{-\frac{(b-m_{ij})^2}{2\sigma_{ij}^2}} P(B_i) \tag{2}$$

for $i = 1 \dots I$ and $j = 1 \dots J$, where $p(b|B_i)$ is the probability density function of the bone pattern b from cluster B_i and $P(B_i)$ is the probability of occurrence of the corresponding cluster.



Fig. 1. Bone image examples

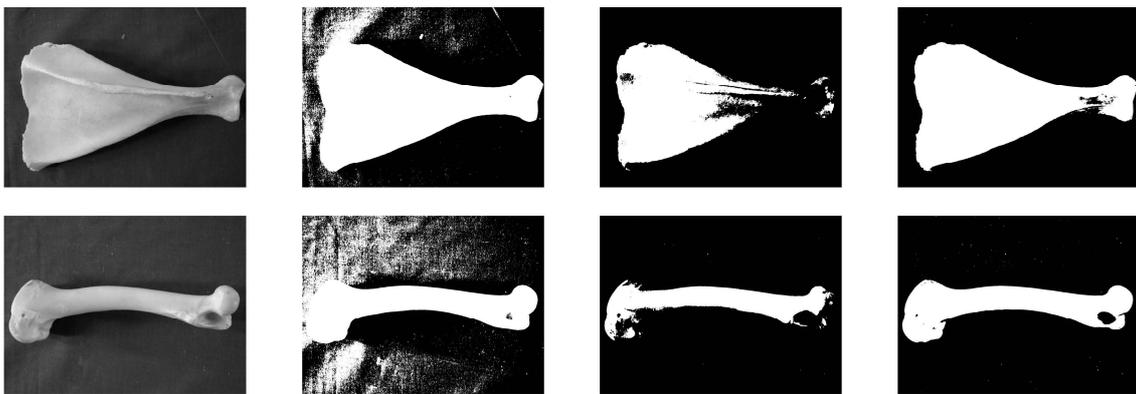


Fig. 2. Instances of gray scale images (first column) where lower thresholds (second column) make many unwanted noisy regions and higher thresholds (third column) destroy the bone region, while K-means clustering (last column) segments the bone region properly

Assuming an equally likely occurrence of all classes (*i.e.* $P(B_1) = P(B_2) \dots = P(B_i) \dots = P(B_I) = 1/I$), and because of the exponential form of the Gaussian density, which persuades the use of a natural logarithm, and since

the logarithm is a monotonically increasing function, the decision function in Eq. (2) can be modified to a more convenient form. In other words, based on the aforementioned assumption and facts, the following decision function, which is less computationally expensive and much faster for the classification of bone shapes, is being used:

$$d_{ij}(b) = \ln[p(b|B_i)P(B_i)] = \ln p(b|B_i) + \ln P(B_i) \quad (3)$$

Considering Eq. (2) and the mathematical process described in [9], an expeditious decision function is obtained as:

$$d_{ij}(b) = -\ln \sigma_{ij} - \frac{(b - m_{ij})^2}{2\sigma_{ij}^2} \quad (4)$$

for $i = 1 \dots I$ and $j = 1 \dots J$, where m_{ij} and σ_{ij} are the mean and standard deviation of bone group B_i using property P_j , and b is the corresponding scalar property of an unknown bone.

Utilizing the above classification approach we calculate recognition rates R_{ij} for each single-valued property P_j and for each bone cluster B_i and save them in the crossing cells of the corresponding rows and columns of the cluster-property matrix.

Next, to appraise a combinatory analysis and depict an efficient feature vector to be used for bone recognition, a set of $K=30$ different combinations of the geometric properties, invariant moment-based functions, and a Zernike feature vector (as described in [7]) is generated and recognition rates are obtained. Here, since the properties are multi-valued, the decision function for the classification is obtained differently. In the multi-valued case, the Gaussian density of the vectors in the i th bone class has the form

$$p(\xi|B_i) = \frac{1}{2\pi^{n/2} |C_{ik}|^{1/2}} e^{\left[-\frac{1}{2} (\xi - m_{ik})^T C_{ik}^{-1} (\xi - m_{ik}) \right]} \quad (5)$$

for $k = 1, 2, \dots, K$, where ξ is the extracted feature vector of an unknown bone, n is the dimensionality of the feature vectors and $||$ indicates matrix determinant.

Using a similar approach described in [9] a simple and expeditious decision function is obtained:

$$d_{ik}(\xi) = -\ln |C_{ik}| - (\xi - m_{ik})^T C_{ik}^{-1} (\xi - m_{ik}) \quad (6)$$

for $i = 1 \dots I$ and $k = 1 \dots K$. Note that C_{ik} values are independent of the input ξ , which means they can be calculated off-line and saved in a look-up table. They are fetched from the look-up table at the on-line stage to better accelerate the decision making process.

The recognition rates R_{ik} for $i = 1 \dots I$ and $k = 1 \dots K$ are calculated utilizing Eq. (6) and saved in appropriate entries in another structure called *cluster-features* matrix. This clearly represents not only the distinguishability of a bone image, but also the recognition power of different sets of features to describe bone shapes.

The general paradigm explained above provides a straightforward method for feature evaluation, which is used for bone recognition. It has been shown to be effective in the experimental results as well (next section). More importantly, column-wise summations in the *cluster-property* and *cluster-features* matrices indicate the recognition power of the simple properties and complex features, respectively. On the other hand, row-wise summations exhibit the discrimination power of each bone class to determine the animal that is useful in differential diagnosis.

3. EXPERIMENTAL RESULTS

The bone collection, which currently includes 1184 bone images, is used for the experiments. The collection includes 6 sets of bone clusters having a number of members from 64 to 288. In the training stage the statistical model parameters are obtained. These include means and standard deviations (scalars) for the individual properties and means (vectors) and covariance matrices for the combined features. In the recognition stage 300 different bone images (50 in each of 6 groups) were applied and classified using the approach explained in Section 2.

To more precisely investigate scale and rotation invariance of the proposed approaches, 16 resized and rotated variations of the original photographs (74 images) are included in the collection. Nearest neighbor interpolation has been used to create new variations. The images are color, in JPEG format photographed by a Casio QV-R61 camera.

For the first-level recognition we attempted to recognize the scapula and femur of all animals involved in the study (i.e. horse, ox and sheep). For this, we placed all rows of the cluster-property matrix into two rows for scapula and femur. Table 1 exhibits recognition powers of the geometric properties (i.e.; *ar*, *pr*, *mj*, *mi*, *ec*) using 1110 images for the training step. Here, 74 bone images (38 scapula bones and 36 femur bones) were applied as the test data. As indicated in the table, recognition rates using *mj* and *ec* are promising (93.2% and 96.8%, respectively). This is due to the considerable differences between the clusters.

For the differential diagnosis (i.e. second-level recognition) test images were classified using 30 combinatory feature sets. The recognition rates were obtained using the decision function in Eq. 6 and the results were saved in the *cluster-features* matrix, which currently, in our experiments has 6 rows and 30 columns. The rows correspond to bone clusters and the columns correspond to a various combinations of features including geometric properties, conventional moment invariants, and Zernike moments. For conventional moments, ϕ functions [7] were used as the image features. For Zernike moments, the MPEG-7's procedure for shape description [7] has been followed. The number of entries in the feature vectors vary from 2 to 36. There are a massive number of different combinations, but only those properties which were shown to have better discriminating power were selected. These properties have tentatively been chosen based on their independent characteristics using covariance matrix. As the 2D structures (matrices) used for the experiments are relatively large, they are not presented here. Results obtained from the experiments are summarized in Table 2.

As can be seen from Table 2, the recognition rate extracted from the moment-invariant functions showed a lack of efficacy, while the rate of geometric properties is slightly better (73.4% and 81.9%, respectively). In this test the Zernike moments exhibited superiority over the others (86.8%). Differences can be explained by 1) the complexity of the feature sets, and 2) the ability of the approach to capture more details of the bone image. For instance, Zernike moments uses a feature vector of size 36, which is the most complex set of features used in the experiments. Therefore, the best recognition rate belongs to the Zernike approach. On the other hand, the geometric approach (a set of five features) shows better recognition ability than the regular moments (a set of seven features), which arises from the fact that the selected geometric properties better distinguish similarities and differences that exist in the animal's bone.

Table 1. Recognition rates for scapula and femur of the animals

Features	<i>ar</i>	<i>pr</i>	<i>mj</i>	<i>mi</i>	<i>ec</i>
Recognition rate %	75.3	79.2	93.2	89.7	96.8

Table 2. Differential diagnosis using three different methods

Method	Regular moments	Geometric properties	Zernike moments
Recognition rate %	73.4	81.9	86.8

4. CONCLUSION AND FURTHER WORK

We have developed an efficient algorithm for bone recognition based on visual properties. It has applications in forensic and anatomy studies (human or animal), orthopedics, archaeology and paleontology. The algorithm is applied and tested for animal bone recognition. The proposed method can be easily expanded for other disciplines including human bone recognition, and for diagnosis of bone diseases.

The approach described in the paper is based on statistical classification. Here, a collection of images is used for setting up the classifier (1110 bone images in this study). Consequently, parameters of the classifier are obtained in this step. This stage can be considered as the training phase, where a large amount of input data is fed to prepare the system. Afterwards, this educated system is used to recognize an unknown bone. 74 images were applied to test the ability of the system, and the recognition rates are derived.

For the first-level recognition, i.e. to determine part of the body, the results obtained from the geometric features are highly promising and provide a source for hope. On the other hand, the second-level recognition, which is to automate differential diagnosis in order to explicitly identify the true animal needs more research. This can be accomplished by employing more sophisticated methods working on local features instead of global features. Consequently, bone image decomposition, correlating image features and anatomic characteristics are new research directions arisen from the study. As a new research direction (recommended by one of the reviewers of the paper), a system can be developed for human bone recognition utilizing the approach proposed in this paper. Feature extraction, and finding applicable properties would follow a similar discipline to that employed for animal bone recognition.

REFERENCES

1. Liu, Z. Q., Austin, T. J. & Moore, D. (1995). Image processing techniques for bone image analysis. *Proc. IEEE Intr. Conf. Image Processing*, Vol. 1, pp. 458–461.
2. Suzuki, K., Abe, H., MacMahon, H. & Doi, K. (2006). Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (mtann). *IEEE Trans. Medical Imaging*, Vol. 25, No. 4, pp. 406–416.
3. He, Y. L., Tian, L. F., Zhu, C. M., Chen, P., Li, B. and Mao, Z. Y. (2005). Development of intelligent diagnosis and report system based on whole body bone spect image. in *Proc. IEEE Intr. Conf. Machine Learning and Cybernetics*, Vol. 9, pp. 5437–5441.
4. Mahmoodi, S., Sharif, B. S., Chester, E. G., Owen, J. P. & Lee, R. E. J. (1997). Automated vision system for skeletal age assessment using knowledge based techniques. *Proc. IEEE Sixth Intr. Conf. Image Processing and its Applications*, Vol. 2, pp. 809–813.
5. Liu, Z. Q., Liew, H. L., Clement, J. G. & Thomas, C. D. L. (1999). Bone image segmentation. *IEEE Trans. Biomedical Engineering*, Vol. 46, No. 5, pp. 565–573.
6. Swarnakar, V., Acharya, R. S., Le Blanc, A., Evans, H., Chen, L., Hausman, E. & Schakelford, L. (1996). Fractal analysis of bone images. *Proc. IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, pp. 74–82.
7. ISO/IEC JTC1/SC29/WG11/N3321, (2000). MPEG-7 visual part of experimentation model version 5. Nordinwijkerhout.
8. Del Bimbo, A. (1999). *Visual Information retrieval*, Morgan Kaufmann Publishers.

9. Chalechale, A. & Safaei, F. (2008). Visual-based interface using hand gesture recognition and object tracking. *Iranian Journal of Science & Technology, Transaction B: Engineering*, Vol. 32, No. B3, pp. 279–293.
10. Ng, P. C. & De Silva, L. C. (2001). Head gestures recognition. *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Vol. 3, pp. 266–269.
11. Thornton, J., Orengo, C. & Jones, D. (2003). *Bioinformatics: genes, proteins and computers*. BIOS.