3D Shape Description of the Bicipital Groove: Correlation to Pathology

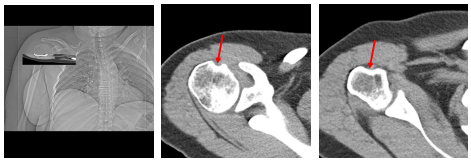
Aaron D. Ward[1], Ghassan Hamarneh[1], and Mark E. Schweitzer[2]

[1] Medical Image Analysis Lab, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada
{award, hamarneh}@cs.sfu.ca, http://mial.fas.sfu.ca/
[2] Department of Radiology and Orthopedic Surgery, NYU School of Medicine, Hospital for Joint Diseases, New York, NY, 10016, U.S.A.
{mark.schweitzer}@nyumc.org

Abstract. The bicipital groove (BG) of the proximal humerus retains the tendon of the long head of the biceps. It is understood that the shape of the BG is related to the probability of injury to the long biceps tendon (LBT). Measurements taken of the BG in previous studies from dry bones and radiographs (henceforth *classical measurements*) are of single cross sections of the humerus, and may therefore overlook important BG shape characteristics. In this study, we test the hypothesis that a novel, medial axis-based 3D shape descriptor captures all relevant features measured in previous work, plus more. To this end, we review previous studies wherein classical measurements have been taken on large numbers of BGs, yielding a distribution that reveals the nature of a normal BG. We develop an automated approach to replicating those measurements on MRI to determine, for each of our data sets, the deviation from the mean of all the classical measurements. We train a classifier by pairing our 3D representations with these deviations to evaluate the potential for computer aided diagnosis of BG pathology based on our 3D shape descriptor.

1 Introduction

The bicipital groove (BG) of the proximal humerus is located on the humeral head, and is formed by the greater and lesser tuberosities (figure 1). The long biceps tendon (LBT) is retained by the BG as the arm moves. Abnormal shape of the BG can induce injury of the LBT. Relevant shape measurements taken in previous work (henceforth *classical measurements*) include BG depth, width, and medial wall angle (figure 2) [1–5]. A deep, narrow BG can irritate the LBT, causing tenosynovitis. A shallow, wide BG can favour dislocation of the LBT. Also, the presence of the supratubercular ridge of Meyer (henceforth *the ridge*) is understood to greatly favour dislocation [1, 2]. In previous work, measurements are taken at a single 2D cross section of the humeral head. Due to large intra-subject variation in BG shape (figure 1 (b),(c)), such measurements are prone to overlooking important shape features, motivating the need for a 3D shape descriptor that captures information along the entire length of the BG.



(a) (b) (c)

Fig. 1. (a) A radiograph intended to show the location of the BG within the body (indicated by the small intersecting axial cross section). (b) A proximal axial cross section showing the shape of the BG, indicated by an arrow (CT scan shown for clarity of illustration). (c) A distal axial cross section from the same patient; note the large difference in BG shape within a single patient; this is typical.

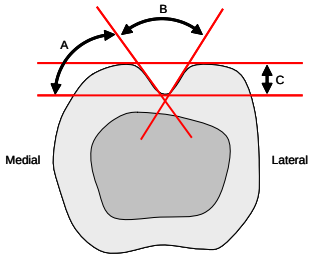


Fig. 2. BG measurements taken from a single cross section in previous literature. A: Medial opening angle. B: Total opening angle (capturing width). C: Depth. (Adapted from [5].)

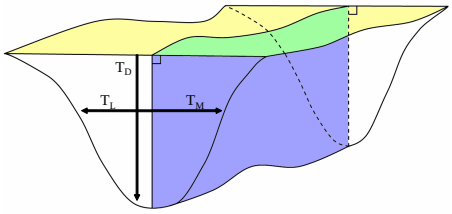


Fig. 3. Depiction of the shape descriptor used in this study [6]. An intertubercular sheet (yellow) is computed to join the tuberosities. A medial sheet (blue) is orthogonal to the intertubercular sheet and intersects as near to the deepest BG point as possible while remaining smooth. Magnitudes of vectors (T_M, T_L, T_D) emanating from sampled points on these sheets and terminating at the BG surface form several 2D thickness fields (medial wall, lateral wall, depth, and width) capturing BG shape.

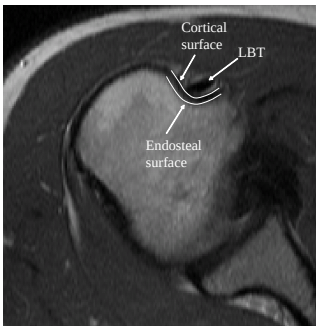


Fig. 4. Appearance of bone on T2-weighted MRI. The inner surface where bone meets bone marrow is the endosteal surface. The outer surface where bone meets surrounding tissue is the cortical surface. We must distinguish between these during surface extraction, because previous population studies of BG shape consist of measurements taken on dry bones and radiographs. Thus, our classical measurements should measure the cortical surface.

In previous work [6], we demonstrated that a 3D, medial axis-based shape descriptor captures medically relevant shape information as 2D thickness fields computed relative to a medial sheet positioned to approximately bisect the BG (figure 3). In the current paper, we investigate the performance of machine learning classifiers in determining normal vs. pathological BG shapes based on this descriptor. That is, the input is a set of thickness fields representing a BG shape, and the expected output is a proper classification of BG shape (e.g. "normal", "abnormal medial wall angle", "presence of the ridge"). The hypothesis is that the 3D shape descriptor is sufficient for automated determination of the deviation of BG shape from normal. This is a first step in answering the subsequent question, to be addressed in future work, of whether the 3D shape descriptor is sufficient for automated determination of the probability of injury to the LBT. To test the hypothesis, we examine previous studies reporting classical measurements taken on large numbers of BGs (dry bones and radiographs). This yields a distribution that allows, for a set of classical measurements taken from a BG, the determination of the deviation from the mean of the distribution. Our automated approach to classical measurement on MRI allows the determination of where each of our BGs fits within the distribution given by previous work, yielding the inputs to our tested classifiers. Our main motivation for pursuing a machine learning/classification approach to this problem is that identification of osseous spurs, the ridge, and the angle of the medial BG wall is difficult due to lack of precise definitions of these structures. This motivates the need for a 3D shape descriptor inherently capturing these features inherently which can be used in classifier training for identification of pathologies in these structures.

In the course of this work, we face the question: why not directly perform a study correlating BG shape to incidence of LBT injury? The reason is that there is an indirect relationship between the anatomical structure whose shape we are computing (BG) and that whose probability of injury we are trying to estimate (LBT). It is understood that an abnormal BG shape *may predispose* an individual to LBT injury. For example, an osseous spur (bony abnormality) inside of the BG can cause the LBT to fray and tear [2]. However, it is entirely possible that at the time of the MRI scan, a patient with an osseous spur in his BG has a perfectly healthy LBT; his LBT injury, if it is going to happen, has not happened *yet*. This presents a problem when training a classifier to learn the probability of LBT injury based on BG shape: with a practically feasible sample size, such cases (i.e. healthy LBT in abnormally-shaped BG and vice versa) can confuse the classifier. Given a sufficiently large sample, one expects the effect of this confusion to be minimized, but collecting such a sample is cost prohibitive.

The indirect relationship between the BG shape and incidence of LBT injury is one of this study's primary motivations. By computing the classical measurements of BG shape and placing each data set in context of the distribution of measurements from previous population studies, we can determine the deviation of each BG shape from normal. This allows us to identify and handle cases where the patient has an injured LBT in a normally-shaped BG and vice-versa. The other primary motivation of this study is to verify that the shape descriptor suf-

ficiently captures, at minimum, the classical measurements of BG shape. If this shape descriptor can be used to effectively train a classifier to perform BG shape diagnosis, then we can proceed with confidence in a future study establishing a relationship between this shape descriptor and the probability of LBT injury.

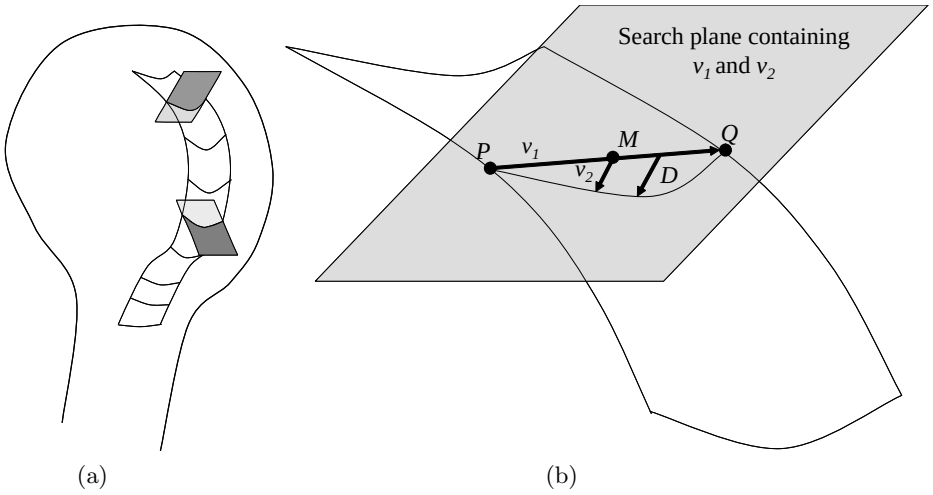


Fig. 5. Because the BG curves with the contour of the humeral head, it is important in measuring depth to ensure that depth is measured along vectors orthogonal to the intertubercular sheet, lying in search planes exemplified in (a). (b) is an enlarged view of a single search plane from (a) showing the details of the definition of the search plane at a given slice of the BG. Segment PQ is defined to join the endpoints of the intertubercular sheet, and vector v_1 is defined from P to Q. Vector v_2 is defined to originate from the midpoint M of PQ and be *normal to the intertubercular sheet* at M. The search plane is defined by v_1 and v_2. Rays are cast from points sampled uniformly along and orthogonal to PQ, constrained to lie in the search plane, and terminating at the BG surface. The length of the longest such ray (indicated by D) is determined to be the depth of the BG on this slice.

The remainder of this paper is organized as follows. In section 2 our data sets and experimental approach are described. In section 3 we give our results and a discussion, and in section 4 we give some concluding remarks.

2 Material and Methods

This study is based on 32 T2-weighted MRIs of the shoulder taken at 1.5T. 10 of these data sets correspond to patients diagnosed with a normal LBT; 22 are diagnosed as abnormal (subluxation, dislocation, or tear). Our approach is as follows (see figure 6).

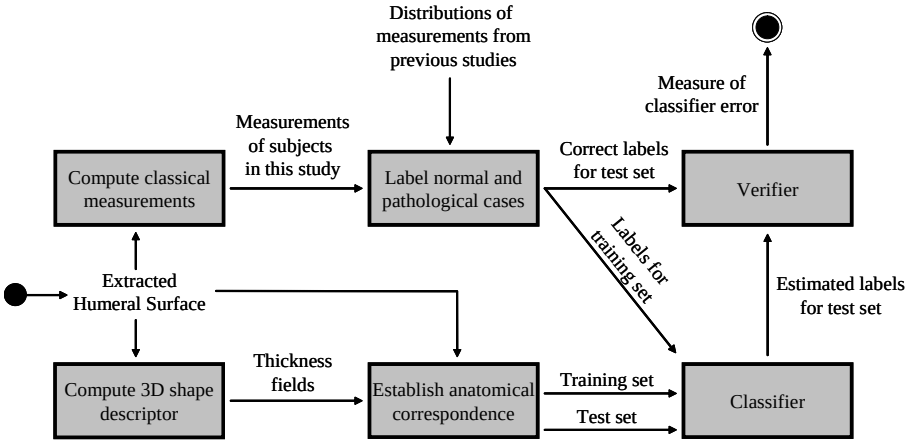


Fig. 6. Procedure followed in this investigation. Extracted BG surface points are used to compute a 3D shape descriptor (figure 3), and also used to compute "classical" measurements of BG shape used in previous work (figure 2). Based on distributions of classical measurements determined in previous studies, a diagnosis of normal versus pathological shape is made for each data set based on its classical measurements. These diagnoses form labels for a set of training shapes given to a classifier, which attempts to correctly label BG shapes in the test set. We evaluate the accuracy of the classifier based on a comparison of these labels to known labels for the test set.

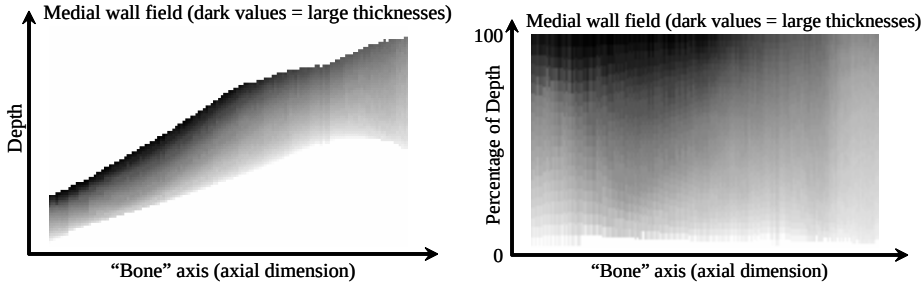


Fig. 7. (a) A medial wall field rendered as a 2D image, with one axis along the bone (axial) dimension, and the other along the depth of the BG. (b) The same field after rescaling along the depth dimension to transform it into a percentage of depth dimension for alignment.

1. **Compute classical measurements.**[1] Given a set of points sampled from a BG surface, we automatically compute the width, depth, and medial wall angle (figure 2) of each data set. Care must be taken to ensure that the

[1] The authors wish to acknowledge and thank Eli Gibson for his efforts in discussion and implementation of this approach to automated classical measurement.

cortical surface of the bone is extracted, to ensure consistency with previous population studies of dry bones and radiographs (figure 4). Segmentation is not the focus of our study and is done manually. To take the measurements, the BG surfaces are first rotated so that slicing axially yields cross sections orthogonal to the humerus. An intertubercular sheet is fitted to the tuberosities (figure 3). At each axial slice S, the intertubercular sheet appears as a line segment PQ with endpoints P and Q touching the tuberosities. The depth D of the BG on S is determined as shown in figure 5. The width of the BG at S is the length of PQ. Starting from the deepest point on the BG surface on S, points sampled along the medial wall define endpoints of line segments approximating the wall. Angles of these line segments with respect to the segment PQ are recorded for S. So, for each BG, we have a set of depth and width values and a set of medial wall angles. Since measurements taken in previous work are of a single slice, and previous authors are not specific in describing how the slice is chosen [7, 2, 4], we aggregate all of our measurements by taking the mean for a single data set. This yields a single depth, width, and medial wall angle for each data set for comparison purposes. This single-slice limitation further reinforces the need for a 3D shape descriptor.

2. **Label normal and pathological cases.** Previous studies took measurements of depth, width, and medial wall angle on dry bones and radiographs [7, 2, 4]; 130 patients in total. The results are as follows. Medial wall angle mean: 60.02°, standard deviation (SD): 15.32°. Depth mean: 4.19mm, SD: 0.96mm. Width mean: 7.9mm, SD: 1.42mm. To provide a binary classification of each BG to the classifier for training, we must specify a standard deviation cutoff defining normal vs. abnormal. We choose a threshold (1.5 SD for our data, for all measurement types) that results in half of the data being normal, presenting the greatest challenge to the classifier. Thus, attempts to classify at random result in the poorest possible performance (e.g. if 90% of the data sets were normal, a classifier could achieve 90% accuracy by simply classifying all test sets as normal). Each data set is also labeled according to expert observation of the presence of the ridge.
3. **Computation of 3D shape descriptor.** An intertubercular sheet is fitted to close the BG, and a medial sheet is computed orthogonal to the intertubercular sheet on each slice. Medial and lateral wall, depth and width fields are computed relative to the sheets (figure 3). Our previous publication [6] gives further details.
4. **Anatomical correspondence.** To prepare the thickness fields for machine learning, we establish anatomically meaningful correspondence between elements. Point (i, j) in any thickness field should correspond anatomically with points (i, j) in the thickness fields of all other data sets. Establishing correspondence is challenging for the BG as it lacks meaningful anatomical landmarks. Due to the large slice thickness, the proximal end of the BG is not reliably determined, and a method for determining the distal end is debated [4]. Our approach is indirect: since the BG is formed by the tuberosities of the humerus we align the *humeri*, consequently aligning the BGs. Our

shape descriptor is invariant to rigid transformations except axial translation, and it is not invariant to changes in scale. To establish correspondence, we find the parameters of the best-fit sphere to the humeral head using the Hough transform. We align all thickness fields such that the axial coordinate of these sphere centers are the same, thus aligning the bones. We scale thickness fields to normalize for humeral head size, according to the spheres' radii, resulting in a set of thickness fields (e.g. figure 7(a)), normalized for scale and aligned along the bone (axial) dimension. Finally, we rescale the field along the depth dimension to make the fields rectangular by resampling thickness values from 0 to 100% along the depth of the BG (figure 7(b)).

Classical Measurement	*Classifier*	*Error*	*No. Principal components*
Width	Quadratic Bayes	0.1875	6
Depth	Min. LS Linear	0.1875	6
Medial wall angle	Quadratic Bayes	0.3750	15
Supratubercular ridge	Min. LS Linear	0.1250	6

Table 1. Results of testing of classification, showing the methods that gave the best performance against our data. The error indicates the proportion of data sets that were mis-classified.

5. **Classification.** We performed dimensionality reduction using PCA on the 1000D vectors formed by the thickness fields. We then trained several classifiers against the dimensions of the thickness field data capturing 95% of the variation. We also trained classifiers to recognize the presence of the supratubercular ridge of Meyer [1] from the depth fields. Testing was performed in a leave-one-out fashion, with classification errors averaged over all rounds.

3 Results

Table 1 shows the results of classification. Accuracy was over 80% for most classifications; different types of classifiers performed best for different tasks. Using classifiers in the *PRTools v.4* package [2], we obtained best results with the minimum least square linear classifier and quadratic Bayes normal classifier. No classifier performed adequately in diagnosis of medial wall angle abnormalities. Vagueness regarding the slice locations of BG measurements taken in previous studies may also adversely affect classification. Considering these obstacles, the results are encouraging; they suggest that the majority of important BG shape features are captured by our representation. Especially encouraging is the classification performance for the ridge, which can be difficult for the human expert to identify.

[2] PRTools v.4, Delft U. of Technology

4 Conclusion

In this work, we investigated the ability of a 3D shape descriptor for the BG to capture aspects of shape known to be related to LBT injury. We showed this by demonstrating that classification algorithms can be trained, using our shape descriptor, to perform accurate diagnosis of BG shape abnormality. The outcome of this investigation is that classification performance using this shape descriptor is acceptable, given the practical obstacles of small sample size and lack of precise literature specifying how some classical measurements were taken in previous studies. The auxiliary aim of this work is to illustrate practical considerations that need to be addressed in a computational study of musculoskeletal disorders on real data, such as development of strategies for handling small sample sizes and anatomical alignment of structures that may lack clear anatomical landmarks. Future work includes establishing the relationship between the 3D shape of the BG and the incidence of injury to the LBT.

References

1. Meyer, A.W.: Spontaneous dislocation and destruction of tendon of long head of biceps brachii. Fifty-nine instances. Archives of Surgery **17** (1928) 493–506
2. Hitchcock, H.H., Bechtol, C.O.: Painful shoulder. Observations on the role of the tendon of the biceps brachii in its causation. Journal of Bone and Joint Surgery, American Volume **30A** (1948) 263–273
3. O'Donoghue, D.: Subluxing biceps tendon in the athlete. American Journal of Sports Medicine **1**(3) (1973) 20–29
4. Ueberham, K., Le Floch-Prigent, P.: Intertubercular sulcus of the humerus: Biometry and morphology of 100 dry bones. Surg. and Radiol. Anat. **20**(5) (1998) 351–354
5. Pfahler, M., Branner, S., Refior, H.J.: The role of the bicipital groove in tendopathy of the long biceps tendon. J. Shoulder and Elbow Surgery **8**(5) (1999) 419–424
6. Ward, A.D., Schweitzer, M.E., Hamarneh, G.: 3D shape description of the bicipital groove of the proximal humerus. Proceedings of SPIE Medical Imaging: Physiology, Function, and Structure from Medical Images **6143** (2006) 1–9
7. Ahovuo, J.: Radiographic anatomy of the intertubercular groove of the humerus. European Journal of Radiology **5**(2) (1985) 83–86