Predicting finger joint angle using ultrasound images of the forearm

Michael Gidaro[1], Linghai Wang[2], Roberta L. Klatzky[3], George Stetten[4]

CMU Robotics Institute Technical Report - CMU-RI-TR-23-30

July 6, 2023

[1] Departments of Bioengineering and Electrical Engineering, University of Pittsburgh, <mgidaro@verizon.net>
[2] Department of Bioengineering, University of Pittsburgh, <qyaoza@gmail.com>
[3] Department of Psychology, Carnegie Mellon University, <klatzky@cmu.edu>
[4] Robotics Institute, Carnegie Mellon University; Department of Bioengineering, University of Pittsburgh, <stetten@andrew.cmu.edu>

Acknowledgments

Supported by a grant from the Center for Medical Innovation at the University of Pittsburgh.

Abstract

Identifying the pose of fingers is used for basic research on the human hand and analyzing dexterous movements. Joint angles can be extracted from instrumented gloves or by tracking markers with video cameras. Such methods are cumbersome, however, and can even impede complex manipulatory tasks as in surgery. We propose a novel solution, using ultrasound images of the muscles in the forearm to determine finger pose. We calibrate its use for monitoring the contraction and relaxation of the flexor digitorum superficialis muscles in the index finger, relative to marker-based tracking, and demonstrate machine classification of finger angle from ultrasound.

Key Words: hand, finger, joint angle, manipulation, ultrasound, forearm, machine learning.

Introduction

The evolution of the human hand prepared our species for highly dexterous activities. Much research has been devoted to understanding the functional morphology of the hand, which is uniquely advantaged for use of tools through finger/thumb opposition and supporting muscles, tendons, and nerves [1-6]. The forearm muscles of humans, in particular, appear to be essential to manipulatory advantages over lower primates [2]. In addition to anatomical analyses, tracking the motion of the fingers during dexterous movement is a valuable tool for evaluating skills [7], predicting overuse in syndromes like carpal tunnel [8], or assessing recovery from adventitious events like stroke [9].

Present methods for tracking finger motion generally fall into two categories. The first depends on mechanical devices, mounted on the finger and hand, or built into a glove, that determine joint angles directly. The use of mechanical devices is inherently cumbersome, inhibiting their use for many activities. The second method depends on video surveillance of the hand, with one or more camera tracking markers mounted at key locations on the hand, or by using computer vision techniques to identify and track the un-instrumented hand. Camera tracking suffers from the requirement of a direct line of sight between the hand and the camera(s) as well as appropriate lighting. Direct line of sight and sufficient lighting are not always possible to maintain, for example, when studying the manipulation of surgical tools during an operation.

We propose a novel solution, using ultrasound images of the muscles in the forearm to determine finger pose. The flexor digitorum tendons, located in the forearm, cause the index, middle, ring, and small fingers to flex. Using ultrasound imaging, the flexor digitorum superficialis muscles, attached to the flexor digitorum superficialis tendons, can be seen contracting and thickening when the fingers are curled, potentially allowing an image recognition computer program to calculate the position of the finger based on the width of the muscle at any point in time. By placing an ultrasound scanner on a subject's forearm, it should be possible to track the fingers without impeding the hand itself with a mechanical device or depending on an external camera.

To demonstrate the basic concept, we have developed a method limited to a single finger. The method determines the flexing of the index finger from ultrasound images monitoring the contraction and relaxation of the flexor digitorum superficialis muscle. We describe our method and results below.

Methods and Results

To establish ground truth for joint angle, the index finger was tracked by means of three circular reflective markers. The markers were attached to the lateral side of the index finger at the center of each phalanx, where they could be the most easily visible without being obscured by other fingers. The positions of the markers were recorded using a digital webcam (Hamilton Buhl SuperFlix, model: WEBCAM). Ultrasound imaging of the flexor digitorum muscles was accomplished using a Terason t3000 ultrasound scanner and a Terason 12L5V ultrasound transducer probe operating at 5.98 MHz, which provided a real-time video recording of the forearm.

Two studies were conducted so that the overall feasibility of ultrasound tracking could be evaluated prior to the use of any machine learning techniques. The first of these studies used computer vision analysis to measure the width of the muscle and the angle of the finger. These time series measurements were compared using linear regression to create a direct numerical relationship between the instantaneous angle of the finger and the width of the muscle. The second study trained and tested a neural network that attempted to identify the finger angle directly from the ultrasound image, without the mediating step of computer vision. These studies are described in detail next.

Figure 1. Setup for tracking the finger using markers via a camera, while scanning the muscles of the forearm with the Terason ultrasound scanner in an axial orientation.

Study 1: Establishing the joint-angle to muscle-width relationship

In the first study, the transducer was oriented in the sagittal plane along the muscles in the forearm, as shown in Figure 1. The webcam was positioned to image the lateral side of the index finger at the same view as is seen in Figure 1. In each frame of the webcam video, the edges of the markers were isolated using the built-in functions findContours() and approxPolyDP() in the OpenCV library (https://opencv.org). The function minEnclosingCircle() was used to return the vertical and horizontal coordinates of the center of each marker. The Pythagorean theorem was used to find the distances between each marker (in pixels), allowing the marker on the medial phalanx to be identified as the point with the shortest combined distance to the other two markers, highlighted in green in Figure 2 a. Using the medial marker as the vertex, the angle between the three markers was calculated. For each ultrasound frame, the functions findContours() and approxPolyDP() were again used to isolate the edges in the image. The flexor digitorum superficialis appears as an approximately rectangular contour spanning the width of the image and positioned directly below the skin. The function boundRect() was used on the contour corresponding to the flexor digitorum superficialis to draw a bounding rectangle around it (Figure 2 b), allowing the vertical width of the rectangle to be measured in pixels.

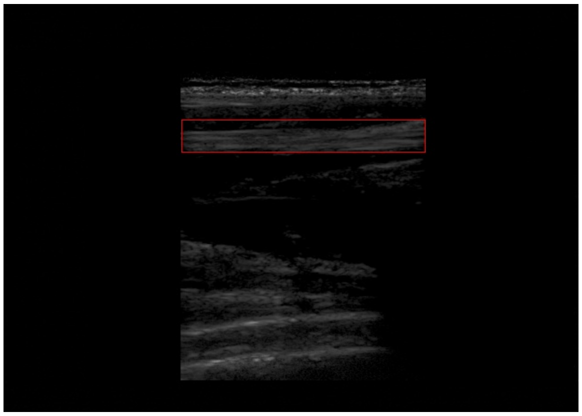

Figure 2a: Sample ultrasound image from trial 1. The flexor digitorum muscle is outlined in a red bounding box. The length of the vertical dimension of the box is recorded as the muscle width.

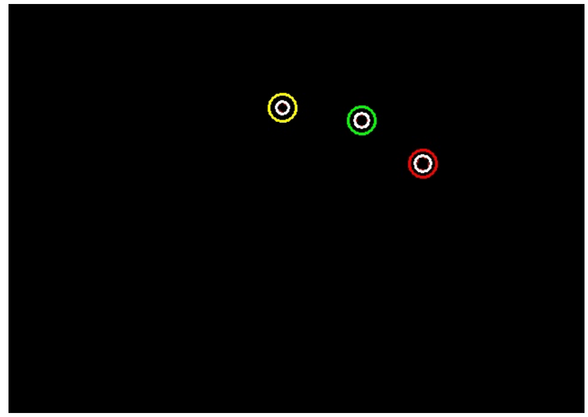

Figure 2b: Positions of the reflective markers identified by OpenCV. The medial segment, outlined in green, is used as the vertex of the calculated angle.

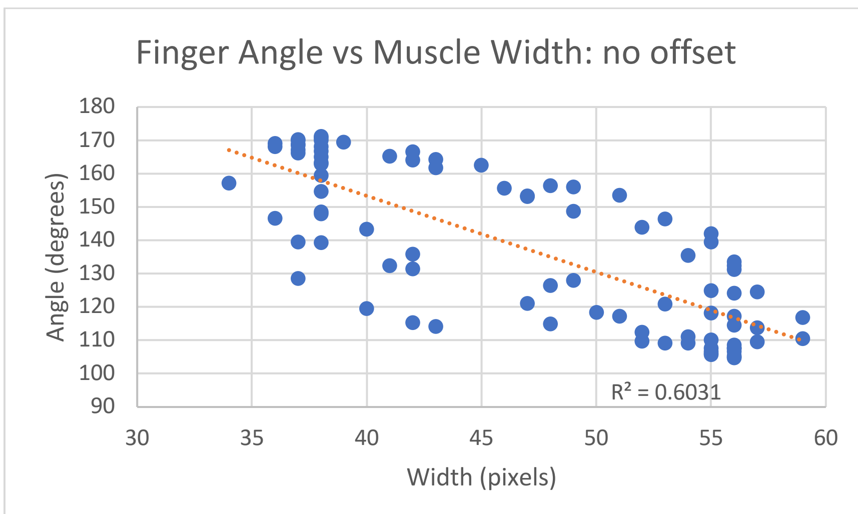

Figure 3. Finger Angle (from markers) vs Muscle Width (from ultrasound scanner) on a frame by-frame basis, with no offset.

Video frames were captured from both the ultrasound scanner and the webcam, creating two synchronous recordings in which the index finger was repeatedly flexed and extended. Both recordings used a rate of 14 fps and were 100 frames in length. The first 11 ultrasound frames were too blurry for muscle width to be reliably determined by the OpenCV algorithm and were omitted from analysis in both recordings. Linear regression was used to assess the relation on a frame-by-frame basis between the ground truth finger angle from the webcam image and muscle width determined from the ultrasound image. Initial examination of the data indicated a weak correlation of $R^2 = 0.60$ (Figure 3), suggesting a delay in ultrasound processing relative to the camera. Accordingly, the data were adjusted by shifting the ultrasound frame number forward in time relative to the webcam frame until an optimal temporal offset was identified by the maximum R^2 value in the regression relating the two variables. The optimal value was found to be 214 ms (3 frames). Linear regression analysis suggests a numerical relationship shown in equation 1, where θ_{finger} is the angle of the index finger (degrees) and w_{muscle} is the width of the flexor digitorum superficialis muscle (pixels), resulting in $R^2 = 0.91$ (Figure 4). We conclude that the joint angle is, to a significant degree, linearly related to the muscle width as measured by ultrasound.

$$\theta_{finger} = -2.764 * w_{muscle} + 266.7 \quad \text{Eq. 1}$$

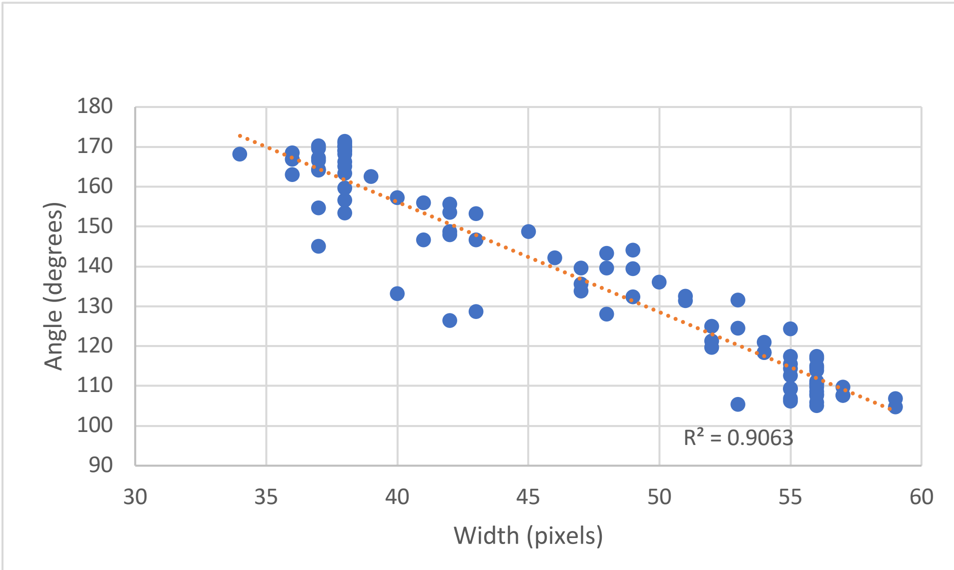

Figure 4. Finger Angle (from markers) vs Muscle Width (from ultrasound scanner) on a frame-by-frame basis, with offset of 214 ms.

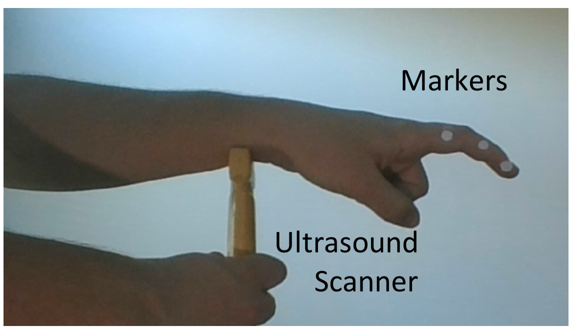

Figure 5. Setup for tracking the finger using markers via a camera, while scanning the muscles of the forearm with the Terason ultrasound scanner in a transverse orientation.

Study 2: Training a neural network to determine joint angle from the ultrasound images

A Convolutional Neural Network (CNN) was implemented with the goal of classifying images of the flexor digitorum muscles based on the measured angle of the finger. As in Study 1, recordings were made at 14 fps using a webcam of the reflective markers on the index finger along with ultrasound images, using the same 3-frame delay to ensure synchrony. However, this time the ultrasound transducer was oriented transversely across the arm (Figure 5), resulting in cross-sectional images of the flexor digitorum superficialis muscle. A total of 812 frames were recorded. The finger-angle measurement for each frame was obtained from the webcam data, as before, and each ultrasound frame was labelled with the corresponding finger-angle value. The angles were then quantized into ten 10° bins, with each bin labeled by its midpoint (e.g., angles between 90° and 100° were collected in a bin labelled 95°). Thirty-three frames (4.1%) evidenced blurring or smearing and were eliminated from analysis. Among the frames in each 10° bin, 80% were used as the training set, 10% were used as the validation set, and 10% were used as the test set.

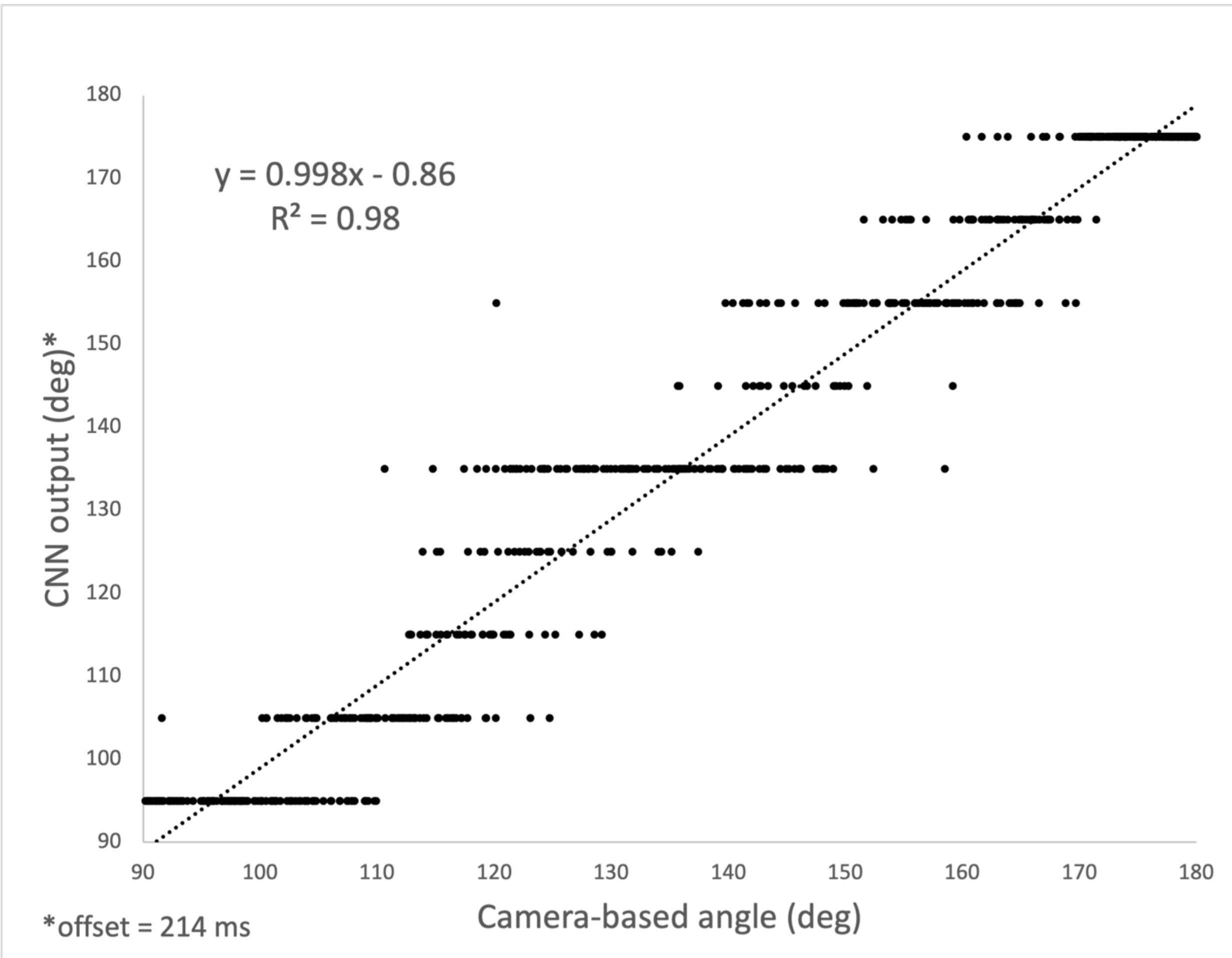

Figure 6. CNN predicted angle measure as a function of the angle derived from the webcam algorithm.

The open-source machine learning framework, PyTorch (https://pytorch.org) was used to create the CNN. The goal was to classify an input ultrasound image of the flexor digitorum muscles according to the angle of the finger as measured by the webcam. To speed up the process of training the network, VGG-11 was used for transfer learning. VGG-11 is a model containing 11 weighted layers and has been pretrained on the ImageNet dataset. In a process called "finetuning," the weights of the model were trained to classify the ultrasound images. Training utilized Cross Entropy Loss and the RMSprop optimizer, both of which are pre-built into PyTorch. A learning rate of .001 and momentum of 0.9 was used. Training was run for 20 epochs, after which the model with the highest validation accuracy was saved. The test accuracy of the selected model was 80%. The model was then implemented into software that fed the ultrasound images to the CNN in real time and reported joint angle measurements.

Figure 6 shows the CNN predicted angle measure (adjusted for lag) against the data provided by the webcam. Although the effect of binning the output of the CNN is evident, there is a strong linear relationship with slope near 1.0. Figure 7 shows the frame-by-frame correspondence between the model and data. These results indicate that by using a trained CNN, it is possible to track the position of a finger with high accuracy using only ultrasound images of the forearm.

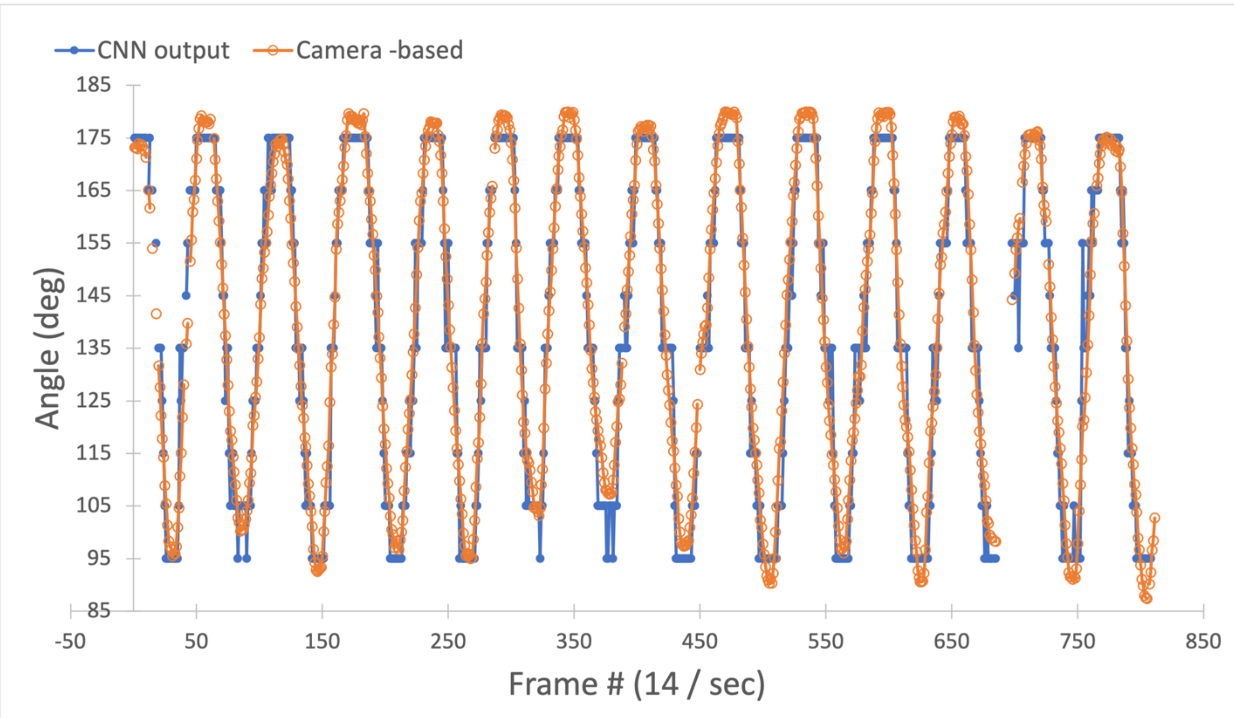

Figure 7. Frame-by-frame correspondence between the CNN output and the angle provided by the webcam. Missing data from thirty-three frames that experienced blurring or smearing are excluded.

Discussion

Our demonstration of the basic concept of extracting finger position and motion in real time from ultrasound images of the forearm is promising, but presently limited to the joint angle of one finger. It is possible that the demand for monitoring multiple joint positions might be reduced, by capitalizing on predictable synergistic relations between fingers or across multiple joints of a single digit. When required, multi-finger or multi-joint tracking with the current approach would require training of independent neural networks for each case. An alternative is to increase the available ultrasound data on the muscles, and hence the scope of the CNN, by using a real-time

3D matrix-array ultrasound scanner to gather volumetric data. Another limitation of our prototype is that only the finger's flexion and extension were predicted. Fingers also experience abduction and adduction, that is, lateral motion relative to each other. While flexion and extension are controlled by muscles in the forearm, the so-called extrinsic muscles, abduction and adduction are controlled by muscles in the hand itself, the so-called intrinsic muscles. To monitor these, an ultrasound scanner would need to scan the hand, thus becoming more cumbersome to dexterous movement. A possible solution to this may be provided by ultrasound scanners that are rapidly becoming smaller, cheaper, and designed for flexibility. One experimental transducer, for example, uses a flexible substrate for the ultrasound array to conform to the skin like a bandage [10]. Such designs could eventually be more non-obtrusive for our application than present ultrasound scanners and permit a practical alternative to present methods of determining finger position and motion.

Supplemental Video Clip

A video screen capture of the system resulting from training the CNN in Study 2 is available online at:

http://www.vialab.org/main/Images/Movies/CNN_output_unfiltered_29Hz_short.mov

The video shows 4 panes simultaneously: (Upper Right) cross-sectional ultrasound images of the flexor digitorum superficialis muscle in the forearm, (Upper Left) camera images of the 3 reflective markers on the phalanges of the index finger used to establish ground truth of joint angle, (Lower Right) printed results for each frame of the CNN output in degrees, with approximate frame rate (Hz) at that point in time, and (Lower Left) visualization of the joint angles predicted by the CNN, with proximal, middle, and distal phalanges represented by the green, red, and blue rectangles, respectively. A visual comparison of the Lower Left and Upper Left panes lends confidence to our conclusion that the prediction of the CNN does indeed follow the ground truth and determine joint angle.

References

1. Day M H, Napier J. The two heads of the flexor pollicis brevis. *J Anat* 1969; 95: 123-130.
2. Diogo R, Richmond B G, Wood B. Evolution and homologies of primate and modern human hand and forearm muscles, with notes on thumb movements and tool use. *J Human Evolution* 2012; 63:64-78. doi: 10.1016/j.jhevol.2012.04.001.
3. Nanayakkara V K, Cotugno G, Vitzilaios N, Venetsanos D, Nanayakkara T, Sahinkaya M N. The role of morphology of the thumb in anthropomorphic grasping: A review. *Front in Mech Eng* 2017; 3. doi: 10.3389/fmech.2017.00005
4. Marzke M W. Tool making, hand morphology and fossil hominins. *Phil Trans R So. B* 2013; 368(1630): 1-8. doi: 10.1098/rstb.2012.0414
5. Susman R L. Hand function and tool behavior in early hominids. *J Hum. Evol* 1998; 35: 23-46

6. Tocheri M W, Orr C M, Jacofsky M C, Marzke M W. The evolutionary history of the hominin hand since the last common ancestor of *Pan* and *Homo*. *J Anat* 2008; 212: 544-562.
7. Saleh, G M, Voyazis Y, Hance J, Ratnasothy J, Darzi, A. Evaluating surgical dexterity during corneal suturing. *Arch Ophthalmol* 2006; 124: 1263-1266. doi:10.1001/archopht.124.9.1263
8. Kozak A, Schedlbauer G, Wirth T, Euler U, Westermann C, Nienhaus A. Association between work-related biomechanical risk factors and the occurrence of carpal tunnel syndrome: An overview of systematic reviews and a meta-analysis of current research. *BMC Musculoskelet Dis*ord Sep 1 2015; 16: 231. doi: 10.1186/s12891-015-0685-0.
9. Yao S, Vargas L, Hu X, Zhu, Y. A novel finger kinematic tracking method based on skin-like wearable strain sensors. *IEEE Sensors* 2018; 18(7): 3010-3015. doi: 10.1109/JSEN.2018.2802421
10. Singh R S, Culjat M O, Vampola S P, Williams K, Grundfest W S, Lee H, Brown E R, Simulation, fabrication, and characterization of a novel flexible, conformable ultrasound transducer array. *Proc. IEEE Ultrasonics Symp* 2007; 1824-1827. doi:10.1109/ULTSYM.2007.459