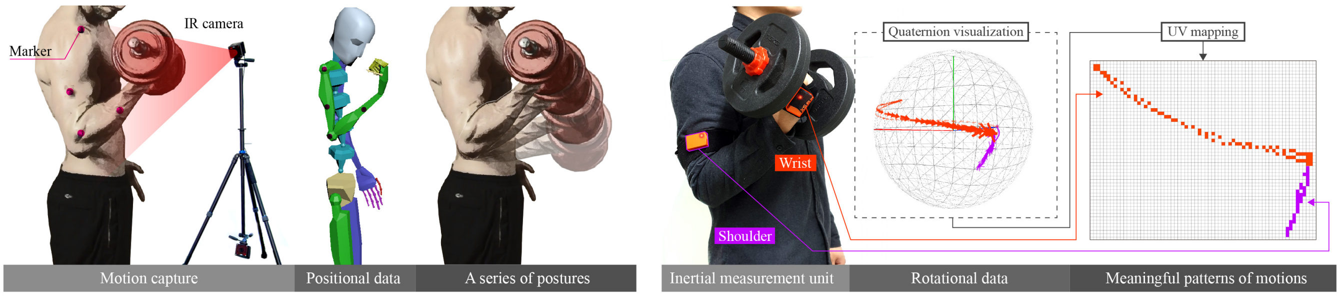
Pilot Experiment of a 2D Trajectory Representation of Quaternion-Based 3D Gesture Tracking

Anonymous Author(s)*



(a) Motion capture

(b) Gesture tracking

Figure 1: (a) Motion capture is the process of recording the movement of human body for future analysis and animation, (b) Gesture tracking is a process of tracking the trajectory of human body movement over time and space.

ABSTRACT

This paper presents ongoing research work on developing a protocol framework for human motion recognition using complex and continuous 3D motion data into more intuitive 2D trajectory representation based-on the quaternion visualization. Quaternions are very compact and free from gimbal lock for representing orientations and rotations of objects in 3D space. In this study, the focus is only on the arm orientation and not the position. In our pilot experimental evaluation, we examine our approach to visually recognize several biceps curl using quaternions data collected using wireless inertial sensors attached to the human arm. The results of the analysis indicate that the proposed framework makes it possible to represent 3D motion data in the form of a 2D trajectory for continuous motion patterns.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; **Interaction techniques**; **Gestural input**; **Visualization**; **Visualization design and evaluation methods**; • **Hardware →** *Sensor devices and platforms*; Sensor applications and deployments.

KEYWORDS

motion tracking, quaternions, wireless inertial sensors, unit sphere, trajectory, IMU, quaternion visualization, gesture, UV mapping

Valencia '19, June 18–21, 2019, Valencia, Spain

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

ACM Reference Format:
Anonymous Author(s). 2019. Pilot Experiment of a 2D Trajectory Representation of Quaternion-Based 3D Gesture Tracking. In *Proceedings of Valencia '19: ACM SIGCHI Symposium on Engineering Interactive Computing Systems (Valencia '19).* ACM, Valencia, Spain, 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In recent years, motion capture and gesture recognition technology enables many potential applications that span from health, medical, arts, and sports [14]. These technologies are gaining more interest from many researchers and engineers in the past two-three decades. Advancement in Micro-Electro-Mechanical System (MEMS) sensors with wireless communication technologies have motivated by a variety of applications, especially in medical monitoring (rehabilitation, physiotherapy), in sports for athletes performance monitoring, in computer vision for human-computer interaction and context-aware pervasive interaction system.

Human movement recognition is a very challenging problem as the human body is very flexible and has 244 kinematic degrees of freedom (DoF) [21]. The main disadvantage is that the human movement can generate different variations for the same basic movement. For instance, in physical fitness scenario, multiple biceps curl movements may lead to slightly different variations every single time. This problem increases as the human population is different in terms of their gesture, physic, and style. Therefore analysts generates a set of hypotheses to tackle the issues more tractable.

A gesture is a form of non-verbal communication or physical activity of the human body that are meaningful to convey information or to communicate with the environment [26], and automation of unmanned aerial vehicle control [29], and more. Many gesture recognition and tracking technologies study have been proposed [14]. Gestures can be hand and arm gestures, head and face gestures or full body movements. It is important to note that unlike motion capture, gesture tracking involves tracking body movement over time as depicted in figure 1. A gesture is a small movement and an activity can be considered as a series of gestures performed in sequence.

This paper presents a novel framework of on-going prototype work, implementation of an Inertial Measurement Unit (IMU) based human gesture tracking system that can visually distinguish and classify between various gestures. Current work is focusing on the development of the visual representation of the quaternion data trajectory from sensors. Among the various data given out by the IMU sensor, we use quaternion data for motion trajectory representation. In our experimental study, quaternion data was collected using an Xsens MTW IMU [27] wireless motion tracking sensors which are attached to the arms.

The paper is structured as follows: In section 2 discusses the survey of literature work on various sensing technologies and advancement in a motion capture and gesture tracking systems. Section 3 details about the motion tracking system and its data collection and orientation estimation. Using the collected data we show the experimental result of visualization on the ongoing work in section 4. Finally, we conclude with discussion of future work.

2 LITERATURE REVIEW

Figure 2 shows the time of the adoption of technology for motion capture and their impact. Some of these technologies have been adopted by the industry and has reached production maturity. Although those technologies at the bottom of the graph are not adopted widely, current research is striving for adoption within the industry. An ultra-low powered motion sensor is discussed in [6]. To reduce the power consumption of the wearable accelerometer, an electric interference device is used to compute the voltage difference between different types of human motions. Largely this includes identifying the location of the device held by the subject using static electric fields. Various studies [1, 4, 7, 8] show that electric field based motion detection. Sound and Electromagnetic sensors are used likewise for touch recognition and position estimation [22].

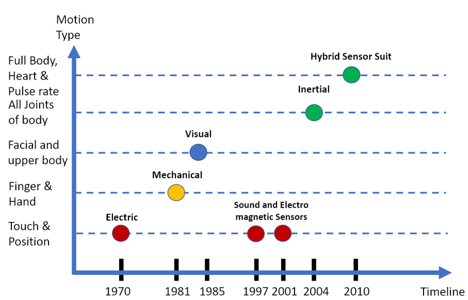


Figure 2: Overview of time of adaption and features of various motion capture and gesture tracking technologies.

Acoustic sensors are mainly used to detect and analyze the motion of the inner organs of the human body. Intravascular ultrasound is a classic example of such systems [9]. Mechanical sensors, on the other hand, use wearable strain sensors that can detect the motion of a body part [3, 23]. Although it is compact and wearable, accuracy cannot be achieved for minute movements that are targeted in the current work.

Research on motion capture and motion recognition using camera images and image processing has flooded the researching community for over 3 decades now. [12] discusses a method to interpret the 3D structure of a person based on six feature points available in the image plane. Stick figures, CAD models are used to visualize the body pose and facial expressions [5, 15]. Visual or image-based technique although are ubiquitous in pose estimation and motion detection, it requires the subject to be within a certain constrained frame. Classifying smaller movements are extremely difficult in vision-based approaches.

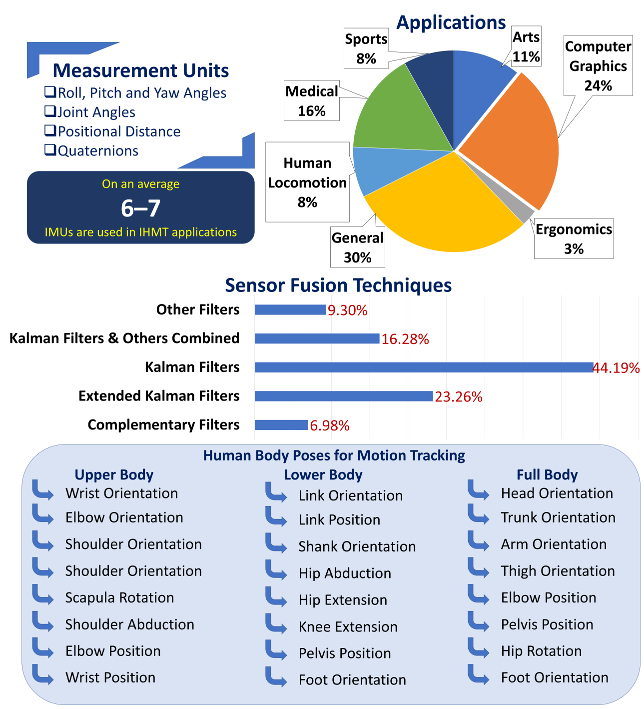


Figure 3: An info-graphics of inertial sensor-based gesture tracking survey.

Inertial sensors have matured over a period of time. Although it was majorly used in defense and navigation systems earlier[24], it is now adopted to estimate human motion to millimeter precision. Various sensor fusion techniques are used to combine multiple motions and orientations of human body [13]. [30] discusses estimation of ankle extension, knee extension, hip extension and abduction using 7 inertial sensors. Kalman Filters technique is used for sensor fusion and to estimate the position of lower limbs. Likewise, [28] proposes a 9 inertial sensor combination for orientation of the full human body in real time. This work is now developed commercially and sold under a brand called xsens[27]. figure 3 gives an info-graphics compiled from information of literature review on IMU based Human Motion Tracking (IHMT) systems. [10, 14].

Home built inertial sensors are used to represent touch and shake for expressive control of a 3D virtual avatar in a virtual environment

[18]. Gesture patterns for kicking and punching actions are mimicked in a reasonably meaningful way. [19] discusses wireless IMU (WIMU) based hand motion analysis technique for handwriting recognition in three-dimensional (3D) space. WIMU device incorporates magnetic, angular rate, and gravity sensors (MARG) and a sensor fusion algorithm to automatically distinguish segments that represent handwriting from non-handwriting data in continuous hand motion data.

Quaternion-based orientation estimation is proposed in [20] which uses 10 inertial sensors to estimate the joint orientation of 5 links. Visualizing the orientation of these joint have also been attempted with the help of a hypersphere.

In a work that is very close to current work, home built wearable inertial sensors are used for gesture recognition [2]. 6 different gestures are classified using a Support Vector Machines and Artificial Neural Networks on the data set. However, the work concerns with the gestures of which are visually different and easily classifiable. No visualization of the quaternion data of the gestures is shown as part of that work. The current research extends the idea to classify gestures which are visually very similar but different at a millimeter scale. Also, in the current work, the difference in various motion are plotted on a unit sphere and two dimensional (2D) plane as quaternions. This enables visually differentiating small movements that are otherwise unrecognizable.

3 MOTION TRACKING AND VISUALIZATION

This section discusses a motion tracking system for data collection and orientation estimation to represent 3D and 2D trajectory.

3.1 Motion Tracking System

In this work, a commercial inertial wireless motion tracking system is used. Commercial motion tracking systems are widely available in the market. For example, Perception Neuron from Noitom [17] and MTw from Xsens [27] etc. Each system has its advantages and disadvantages in providing ease of use because of interfacing suit requires complicated preparation and accurate calibration steps. The motion tracking system used in this study is MTw, which is a miniature WIMU as shown in figure 4. It is a small, lightweight and completely wireless 3D motion tracking product based on inertial sensors manufactured with MEMS technology. This sensor returns 3D orientation, acceleration, angular velocity, static pressure, and earth-magnetic field intensity. Only 3D orientation is considered for motion trajectory visualization. Basically, MTw outputs three measurements: Unit quaternion, Euler angles, and rotation matrix. We focus on quaternion representation of orientation as it is free from gimbal lock [16]. The focus is only on the arm orientation and not the position. The sensors are clipped on straps attached on the upper arm and lower arm as shown in figure 4 to collect motion tracking data.

3.2 Orientation Estimation

The ultimate objective of current experimental project is to estimate the 3D orientation of arm motion tracking and represent orientation in the form of 3D and 2D trajectory based on the quaternion visualization.

Figure 4: Overview of motion tracking system: (left-top) wireless motion tracker (MTw), (left-bottom) Awinda station and placement of two MTw sensors on arm for data collection (right).

Quaternions are a compact and complete representation of rotations in 3D space unlike Euler angles [16]. The objective of orientation estimation is to provide rotation information of tracked arm part in their local coordinate frame. Several steps are usually required in order to get an accurate orientation estimation, including transformation between local and global coordinates.

A quaternion is a four-dimensional entity that can be used to encode any rotation in a 3D coordinate system.

$$q = q_0 + q_1 i + q_2 j + q_3 k \tag{1}$$

A quaternion (q_0, q_1, q_2, q_3) has one real part (q_0) which represents the angle of rotation and three imaginary parts (q_1, q_2, q_3) representing the axis of rotation. The three imaginary dimensions, which are i, j, and k, are unit length and orthogonal to one another and it has a norm of 1 (i.e$||q|| = 1$). The conjugate of a quaternions is represented as-

$$q^* = q_0 - q_1 i - q_2 j - q_3 k \tag{2}$$

and for unit quaternions, the inverse equals the conjugate, i.e.-

$$q^{-1} = q^* \tag{3}$$

The X and Y direction of the sensor is defined in accordance with the longitudinal and lateral axis, and the Z direction is perpendicular to the plane defined by the X and Y axis. Hence, for visualization of quaternions on a 2D sphere of unit length, the starting point of the trajectory is set at (0,0,1). Every time a certain motion is plotted on a 2D sphere, the quaternions are rotated such that the initial quaternions is plotted at (0,0,1). In equations 4, the quaternions q_i is rotated in the opposite direction of first quaternions q_{first} to negate all rotations to bring the axis near the center of the sphere at (0,0,0).

$$\Delta_n = q_n q_{first}^{-1} \tag{4}$$

for $q_n = q_{first}, \Delta_n = (1, 0, 0, 0)$ To rotate the point p=(0,0,1) about Δ_n two multiplications are performed as in equation 5.

$$r_n = \Delta_n p \Delta_n^{-1} \tag{5}$$

Where, $p = (0, 0, 0, 1)$ in quaternion form and r_n is of form (0, x, y, z).

3.3 Quaternion Visualization

In computer graphics, quaternions play a key role in the representation of rotations. The idea is to transform complex, continuous motion patterns into quaternion 3D space and convert it into a more intuitive 2D space representations. Through this, we intend to propose a generalized framework for the representation of motion tracking.

In the experimental setup, three variants of a biceps curl are considered to evaluate the proposed system namely, Standard, Close and Wide bicep curl.

3.3.1 3D Representation of 3D Motion. To plot the motion trajectory on a wire-frame unit sphere, we use Unity 3D real-time engine. A total of 40x40 grids constitute a latitudinal and longitudinal axis with 20 in front and back face each. As discussed in section 3.2, the trajectory starts at prefixed point (0,0,1) on the unit sphere. As shown in figure 5, the trajectory starts with smaller arrows and grows bigger towards the end of the trajectory.

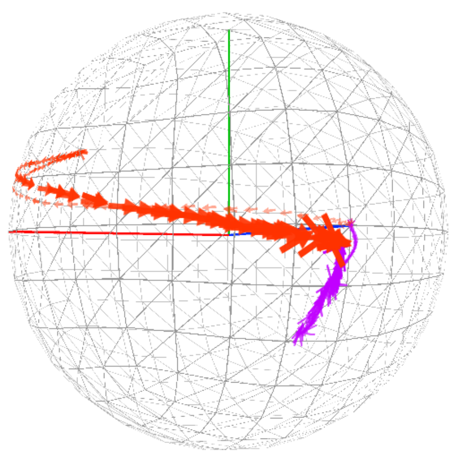


Figure 5: Standard biceps curl trajectory: Lower arm trajectory in red and upper arm trajectory in purple.

Although the representation of motion trajectory on the sphere is visually intuitive, while viewing from the front, it does not give a sense for the trajectory scale due to the curvature of the sphere. And due to the 3D nature of the unit sphere, repeated patterns of the same motion may overlap making it visually indistinguishable. To counter the scaling and overlapping issue, a 2D representation of 3D motion is proposed.

3.3.2 2D Representation of 3D Motion. In order to represent 3D motion in 2D, UV mapping technique is adopted and implemented to visualize on 2D plane (UV map) representation by using popular open source tool called Visualization Tool Kit (VTK) [11]. UV mapping is a technique used for texture mapping in 3D modeling [25].

U and V represent the axis of the 2D plane and map to a unique point on a 3D sphere. Equation 6 and 7 map points from a 3D sphere to a 2D plane.

$$u = 0.5 + \frac{arctan2(z, x)}{2\pi} \quad (6)$$

$$u = 0.5 - \frac{arcsin(y)}{\pi} \quad (7)$$

Where, (x, y, z) is the trajectory point r_i on the sphere from equation 5.

UV mapping shown in equation 6 and 7, is a one-to-one mapping of grids size 40x40 as shown in figure 6(a) and 6(b). However, it was observed that trajectory is not legible at lower girds sizes and becomes clearer visually as the grid size increases as shown in figure 6(c) and 6(d). The grids size beyond 160 causes the trajectory to break intermittently. Hence, for future 2D representation of the trajectory, the grid size of 160 is used.

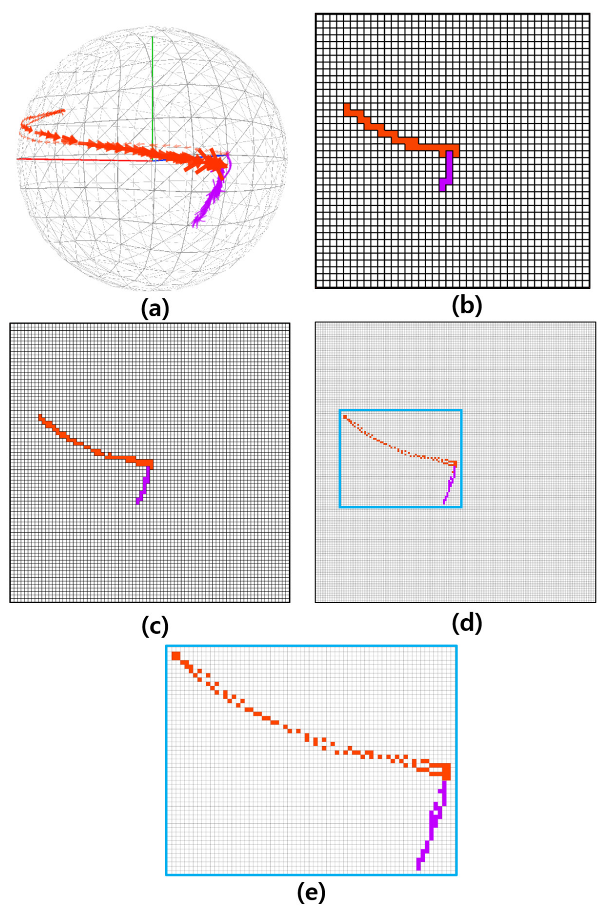


Figure 6: 2D representation of Standard biceps curl 3D motion: (a) 3D trajectory representation with 40x40 grids on a sphere, (b) 2D trajectory representation with 40x40 grids on a plane, (c) with 80x80 grids, (d) with 160x160 grids, (e) highlighted part of (d).

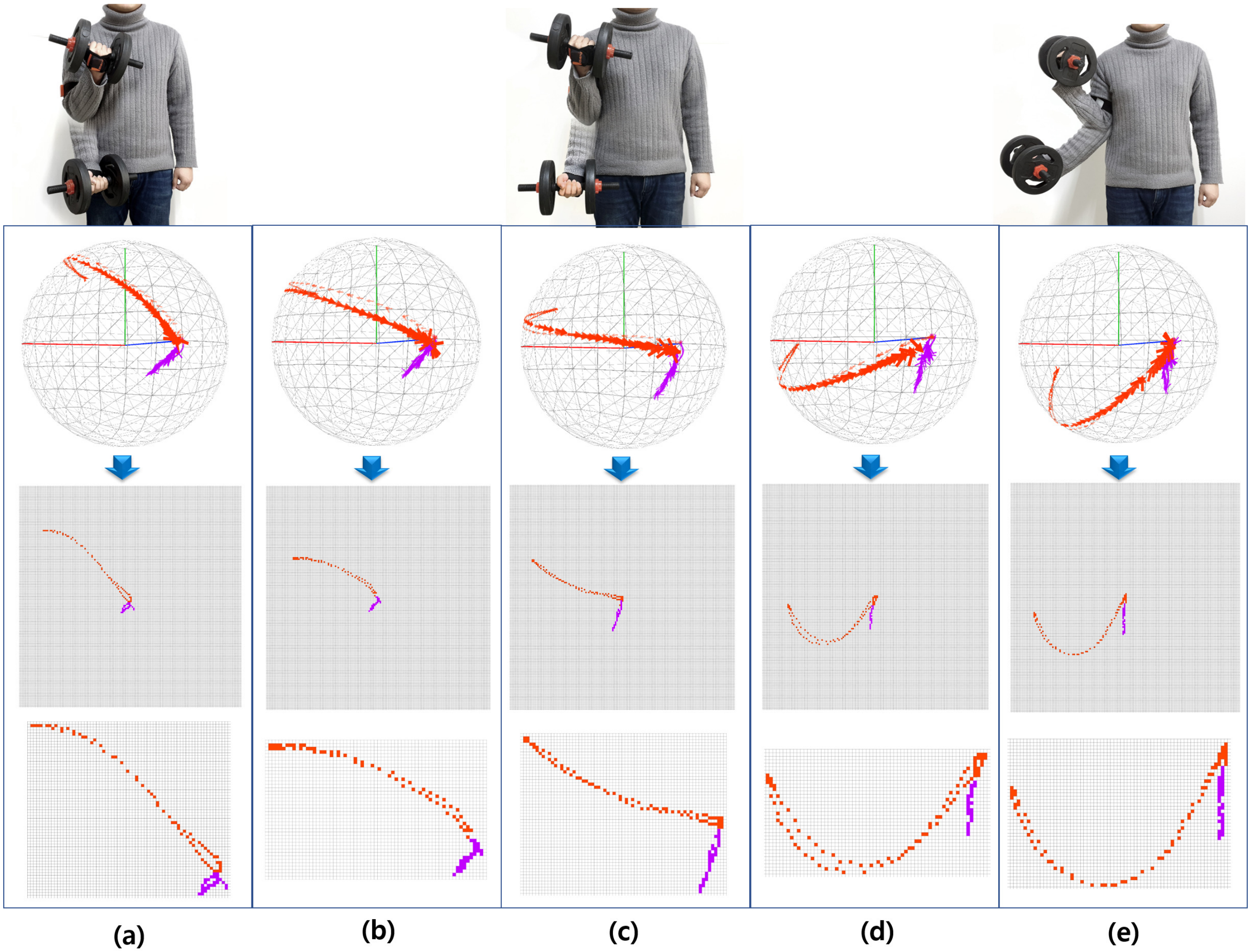


Figure 7: Three variants of a biceps curl: (a) Close bicep curl, (b) little variation in Close bicep curl, (c) Standard bicep curl, (d) little variation in Wide bicep curl, (e) Wide bicep curl.

4 EXPERIMENTAL EVALUATION

The patterns of three variant biceps curl mentioned earlier are captured. figure 7 shows the 3D motion trajectory representation on the unit sphere and 2D plane (the last row in figure 7 shows the focus on trajectory). In figure 7(a), 7(c) and 7(e), the pattern of all the three variants biceps curl are shown. Two slightly varied trajectories of close and wide bicep curl are shown in figure 7(b) and 7(d).

On the unit sphere, two different color trajectory represents the respective two IMU sensors trajectory, where the red color trajectory represents the lower arm and purple color trajectory represents the upper arm. With keen observation, the trajectory of the upper arm is shorter compared to the lower arm trajectory, since the motion data of the upper arm is less as compared to the lower arm. Similarly, with respect to the 2D representation, for upper arm trajectory data are accumulated densely, whereas for lower arm trajectory data are distributed loosely on the grid.

5 CONCLUSIONS

Quaternions are one of the most widely used methods for rotation calculation without any singular axis, unlike Euler Angle. In this pilot study, based on quaternions visualization a human arm 3D gesture tracking and its 2D trajectory representation are presented. The gesture tracking system used in this work is an Xsens wireless IMU sensor, which is fastened to the arm (upper and lower arm) and it outputs orientations in the unit quaternions form. The basic idea is to transform complex, continuous motion patterns into quaternions 3D space and convert it into a more intuitive 2D space representations. Through this, we intend to propose a generalized framework for gesture tracking representation.

In the experimental setup, three variants of a biceps curl are considered to evaluate the proposed system. The proposed framework shows that it is possible to represent a 3D motion data in the form of a 2D trajectory for continuous motion patterns. While there are still areas which could be improved, the current system as presented here is a basic functional foundation upon which to build upon in the future. For further extending the proposed method, gesture tracking will be improved and extended to gesture recognition by employing different classification methods.

REFERENCES

[1] Sandip Agrawal, Ionut Constandache, Shravan Gaonkar, Romit Roy Choudhury, Kevin Caves, and Frank DeRuyter. 2011. Using Mobile Phones to Write in Air. In *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services (MobiSys '11)*. ACM, New York, NY, USA, 15–28. https://doi.org/10.1145/1999995.1999998

[2] Shamir Alavi, Dennis Arsenault, and Anthony Whitehead. 2016. Quaternion-Based Gesture Recognition Using Wireless Wearable Motion Capture Sensors. *Sensors* 16, 5 (2016). http://www.mdpi.com/1424-8220/16/5/605

[3] Antal K Bejczy. 1980. Sensors, controls, and man-machine interface for advanced teleoperation. *Science* 208, 4450 (1980), 1327–1335.

[4] W-T Chen and H-R Chuang. 1998. Numerical computation of the EM coupling between a circular loop antenna and a full-scale human-body model. *IEEE Transactions on Microwave Theory and Techniques* 46, 10 (1998), 1516–1520.

[5] Zen Chen and H-J Lee. 1992. Knowledge-guided visual perception of 3-D human gait from a single image sequence. *IEEE transactions on Systems, Man, and Cybernetics* 22, 2 (1992), 336–342.

[6] Gabe Cohn, Sidhant Gupta, Tien-Jui Lee, Dan Morris, Joshua R. Smith, Matthew S. Reynolds, Desney S. Tan, and Shwetak N. Patel. 2012. An Ultra-low-power Human Body Motion Sensor Using Static Electric Field Sensing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. ACM, New York, NY, USA, 99–102. https://doi.org/10.1145/2370216.2370233

[7] Gabe Cohn, Daniel Morris, Shwetak Patel, and Desney Tan. 2012. Humantenna: Using the Body As an Antenna for Real-time Whole-body Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 1901–1910. https://doi.org/10.1145/2207676.2208330

[8] Gabe Cohn, Daniel Morris, Shwetak N. Patel, and Desney S. Tan. 2011. Your Noise is My Command: Sensing Gestures Using the Body As an Antenna. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 791–800. https://doi.org/10.1145/1978942.1979058

[9] Francisco Escolano, Miguel Cazorla, Domingo Gallardo, and Ramón Rizo. 1997. Deformable templates for tracking and analysis of intravascular ultrasound sequences. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*. Springer, 521–534.

[10] Alessandro Filippeschi, Norbert Schmitz, Markus Miezal, Gabriele Bleser, Emanuele Ruffaldi, and Didier Stricker. 2017. Survey of Motion Tracking Methods Based on Inertial Sensors: A Focus on Upper Limb Human Motion. *Sensors* 17, 6 (2017). https://doi.org/10.3390/s17061257

[11] KitwareâĂŹs. 2019. Visualization Toolkit. https://vtk.org/about/#overview

[12] Hsi-Jian Lee and Zen Chen. 1985. Determination of 3D human body postures from a single view. *Computer Vision, Graphics, and Image Processing* 30, 2 (1985), 148–168.

[13] Matjaž Mihelj. 2006. Inverse kinematics of human arm based on multisensor data integration. *Journal of Intelligent and Robotic Systems* 47, 2 (2006), 139–153.

[14] S. Mitra and T. Acharya. 2007. Gesture Recognition: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37, 3 (May 2007), 311–324. https://doi.org/10.1109/TSMCC.2007.893280

[15] Thomas B Moeslund and Erik Granum. 2001. A survey of computer vision-based human motion capture. *Computer vision and image understanding* 81, 3 (2001), 231–268.

[16] Ramakrishnan Mukundan. 2002. Quaternions: From classical mechanics to computer graphics, and beyond. In *Proceedings of the 7th Asian Technology conference in Mathematics*. 97–105.

[17] Noitom. 2019. Perception Neuron. https://neuronmocap.com/

[18] Shashidhar Patil, Harinadha Reddy Chintalapalli, Dubeom Kim, and Youngho Chai. 2015. Inertial Sensor-Based Touch and Shake Metaphor for Expressive Control of 3D Virtual Avatars. *Sensors* 15, 6 (2015), 14435–14457. https://doi.org/10.3390/s150614435

[19] Shashidhar Patil, Dubeom Kim, Seongsill Park, and Youngho Chai. 2016. Handwriting recognition in free space using wimu-based hand motion analysis. *Journal of Sensors* 2016 (2016).

[20] Gerard Pons-Moll, Andreas Baak, Juergen Gall, Laura Leal-Taixe, Meinard Mueller, Hans-Peter Seidel, and Bodo Rosenhahn. 2011. Outdoor human motion capture using inverse kinematics and von mises-fisher sampling. In *2011 International Conference on Computer Vision*. IEEE, 1243–1250.

[21] Boris I Prilutsky and Vladimir M Zatsiorsky. 2002. Optimization-based models of muscle coordination. *Exercise and sport sciences reviews* 30, 1 (2002), 32.

[22] Mike Topping. 2001. Handy 1, a robotic aid to independence for severely disabled people. In *7th IntâĂŹl Conf. on Rehabilitation Robotics*. 142–147.

[23] Chunya Wang, Xiang Li, Enlai Gao, Muqiang Jian, Kailun Xia, Qi Wang, Zhiping Xu, Tianling Ren, and Yingying Zhang. 2016. Carbonized silk fabric for ultra-stretchable, highly sensitive, and wearable strain sensors. *Advanced materials* 28, 31 (2016), 6640–6648.

[24] Wikimedia. 2019. Lockheed SR-71 Blackbird - Astro inertial navigation system. https://en.wikipedia.org/wiki/Lockheed_SR-71_Blackbird#Astro-inertial_navigation_system

[25] Wikipedia. 2019. UV mapping. https://en.wikipedia.org/wiki/UV_mapping

[26] Wikipidea. 2019. Gesture. https://en.wikipedia.org/wiki/Gesture

[27] Xsens. 2019. The leading innovator in 3D motion tracking technology. https://www.xsens.com

[28] Alexander David Young. 2010. Use of body model constraints to improve accuracy of inertial motion capture. In *2010 International Conference on Body Sensor Networks*. IEEE, 180–186.

[29] Y. Yu, X. Wang, Z. Zhong, and Y. Zhang. 2017. ROS-based UAV control using hand gesture recognition. In *2017 29th Chinese Control And Decision Conference (CCDC)*. 6795–6799. https://doi.org/10.1109/CCDC.2017.7978402

[30] Shaghayegh Zihajehzadeh and Edward J Park. 2017. A novel biomechanical model-aided IMU/UWB fusion for magnetometer-free lower body motion capture. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 47, 6 (2017), 927–938.