

# Velocity-based gesture segmentation of continuous movements during virtual reality exergaming

Shanmugam Muruga Palaniappan<sup>1</sup>, Shruthi Suresh<sup>1</sup>, Bradley S. Duerstock<sup>1,2</sup>

<sup>1</sup>Weldon School of Biomedical Engineering, Purdue University, West Lafayette, Indiana, USA, <sup>2</sup>School of Industrial Engineering, Purdue University, West Lafayette, Indiana, USA

## INTRODUCTION

Gesture segmentation also referred to as gesture extraction or gesture spotting is a critical step in the process of gesture recognition. Many gesture recognition systems including those employed in Virtual Reality (VR) often assume that all input gestures are deliberate and relevant [1]. Gesture segmentation is employed to separate continuous motion data into discrete gestures [2], which can be divided into two classes - intentional gestures and unintentional movements [3]. Extraction techniques are required to determine start and end points of a deliberate motion from continuous time series data, particularly when there is a change in the gestures performed by the individual. However, gesture segmentation can often be difficult due to sensor drift, sensor inconsistencies, sensor moving independent of hand motion, inherent hand tremor, and unintentional motions [4-6].

Several gesture segmentation systems have been proposed by different researchers using velocity-based profiling depending on the nature of data stream and type of gestures expected. However, the disadvantage of the previously proposed methods in the literature is the need for training data or labelled gestures, which includes the determination of a start and endpoint. Additionally, while these systems are trained and function for able-bodied individuals, the same techniques can prove to be challenging for individuals with large variability in upper limb movements due to a motor impairment. This is largely the case in rehabilitation of individuals with upper limb mobility impairments such as stroke [7], traumatic brain injury, or spinal cord injury (SCI).

In this paper, a gesture segmentation technique was developed based on an approach that identifies the beginning and end of intentional gestures using temporal changes in continuous movements of the upper limbs. We developed a baseline VR tool, which was used by individual users to determine the threshold velocity for individuals. In addition, a commercial VR system using modified trackers for individuals with upper limb mobility impairments and wheelchairs was used for at-home rehabilitation exergaming. This velocity profile-based approach for gesture segmentation permits unsupervised gesture detection for tracking the purposeful movements of users that are not able to follow a set of predetermined gestures during rehabilitation [8].

## METHODS

### Experimental Setup

Six participants (1 female and 5 male) with cervical level (C4-C7) SCIs were recruited from the Rehabilitation Hospital of Indiana and Purdue University. The mean age of the participants was  $37.5 \pm 9.9$  and had been injured for  $15 \pm 11.2$  years. The study protocol with participant informed consent was approved by the Institutional Review Board.

The commercially available HTC Vive® platform was used for VR exergaming. The HTC Vive® comprises of two base stations, called lighthouses, which have spinning infrared (IR) lasers, head mounted display (HMD), and trackers with a constellation of IR receivers to precisely determine their position and orientation. Baseline motion data of a participant's hand was collected while performing a 3-D spatial VR test [9]. Briefly, participants were presented with six rings consisting of several virtual spheres spawned around them from above their heads to levels below their knees (Fig. 1). The participants were instructed to push as many spheres as they could as far outwards as they could in 2 minutes. The 3D coordinates representing the motion data was recorded at 90Hz. The participants were recorded with a video camera while they performed the VR game for mixed reality analysis.

### Gesture Segmentation

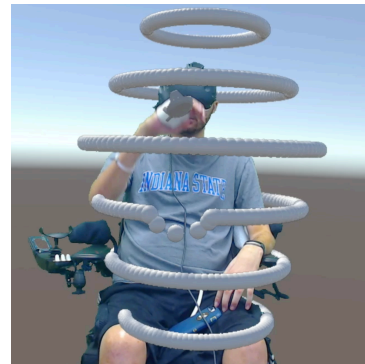


Figure 1: Participant performing 3-D VR movement test to record baseline upper limb motions by pushing away spheres in suspension at different levels around their body in order to determine threshold ranges of motion and gesture velocities.

Gesture segmentation involved a two-step process. First, a threshold velocity was measured during the participants' performance of the VR baseline movement test to determine if a deliberate gesture was being performed. Second, velocities of motion were calculated to delimit gestures. The velocity of arm motion was calculated by performing a time derivative of the 3D hand position data extracted during the baseline task. Threshold velocities were attained either at areas of rest or at intermediate poses between two intentional gestures known as a 'turning point'. Areas of rest were determined through a combination of Kernel density estimation (KDE) with a Gaussian kernel and K-means clustering algorithm. Density scores ( $d_i$ ) were calculated through equation 1 below, wherein  $i$  ranges from 1 to the length of the time series data recorded from gameplay, then used to color the 3D coordinates to plot a heat map of the gameplay.

$$d_i = \hat{f}_h(x_i, y_i, z_i) \quad (1)$$

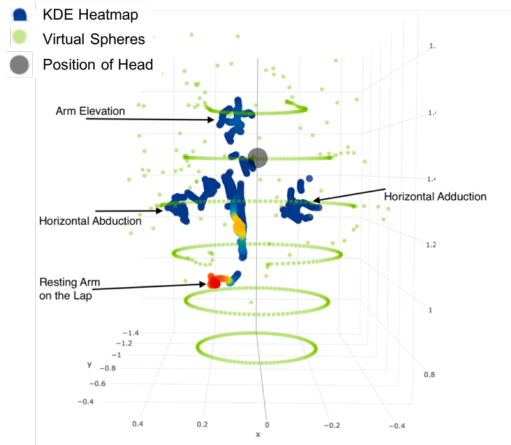


Figure 2: 3-D heatmap generated by kernel density estimation showing areas of high frequency movement of the arm. The green markers are the spheres from the baseline task. The position of the participant's head is indicated by the grey sphere.

A K-means algorithm was used to identify and separate various clusters with similar density scores but distinct (spatially separated) 3D coordinates. This allowed determination of clusters which are areas that were frequently visited by the hand [10]. Of these clusters, the cluster that was associated with the area of rest for the particular individual, defined as the cluster with the highest density and confirmed visually through mixed reality recording.

In Figure 2 is a typical heat map of areas of frequent movement from a participant generated by KDE (shown in blue). Many virtual spheres (indicated in green) can be seen to have been displaced (out of plane) while some are untouched by the participant still lie in their rings. The area of rest has the largest density as generated by the KDE (Fig. 2 shown in red).

We defined the threshold velocity as the local maxima in velocity within the area of rest. Additionally, a threshold velocity could also be reached at a turning

point which was an intermediate pose between two gestures. In this intermediate pose, the hand has a velocity of zero while the acceleration is non-zero as the direction is changing, thus a *turning point*. Acceleration of the hand is calculated through a double derivative of the position of the hand. At turning points and rest points, the velocity was either very low or close to zero.

A deliberate gesture was defined as one where the velocity of the recorded movement exceeds the determined threshold velocity. The temporal data was then sliced up using the threshold velocity as the delimiter, yielding individual gestures of variable length. All the extracted gestures exhibited a sequence wherein they start and end at a *resting position* or a *turning point*.

The resting position was used as the reference point for determining the resting velocity threshold. Through this delimitation of individual gestures based on velocity profiles, we were able to separate gestures for each participant, despite differences in arm mobility and their individual performance.

## RESULTS

We were able to delimit gestures using the velocity and acceleration of the hand positions while the

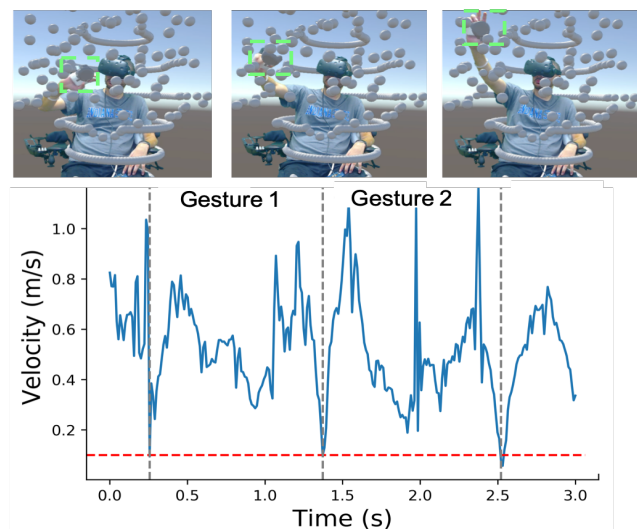
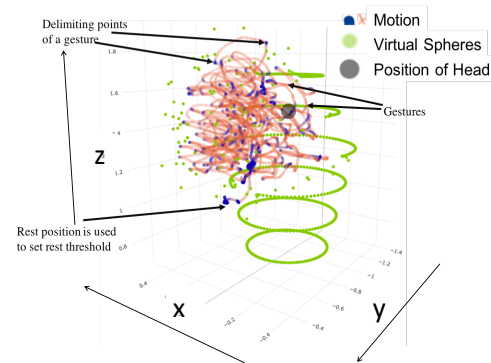


Figure 3: Representation of gesture segmentation based on the velocity profile to segment gestures with corresponding screenshots of participant performing the distinct gestures (green dashed box). Red dashed lines on the velocity profiles represent the calculated threshold velocity with the grey dashed lines representing end points of the gestures.

participants performed the various actions. Figure 3 shows a 2-D output of gesture separation over time for one of the participants. The calculated threshold velocity varied for each participant. When the velocity of the hand dropped below the threshold velocity, it demarcated the start/end points of a gesture (Fig. 3, grey lines).

The gestures extracted through the velocity delimitation were compared with videos taken of the participants during gameplay. Through manual inspection of each participant's videos, it was identified that the gestures delimited through the velocity profiles matched the change of gestures performed by each participant (Fig. 3). In addition, it was identified that through the velocity profiling, it was possible to discriminate the start and end points of gestures performed by the participants despite the variation of gestures performed.

As shown in Figure 4 the faint red points represent tracked hand motions. Blue points represent delimiting points of gestures. Green points represent virtual stimuli (spheres) presented to the participants in the VR game. We observed that end points (blue markers) occur closest to the displaced green spheres in the positive Y direction. The end points of gestures are clustered either close to the body (closest to the head) or at the bounds of motion away from the body (farthest from the head). These are the locations where the participant interacted with the virtual spheres. The gesture delimitation through velocity profiles was performed for all six participants resulting in similar observations.



**Figure 4: Gesture segmentation demonstrated through a 3D plot. The red lines are the gestures which were extracted through the algorithm and the blue are the areas where the velocity of motion is below the set threshold velocity indicating a delimiting point of a gesture.**

## DISCUSSION

Gesture segmentation plays a critical role in being able to discriminate between gestures, particularly with a continuous stream of movement data. We propose a novel technique to extract gestures based on the change in velocity of movement of the participant's hand using a commercially available VR system. Through the velocity profiles, it is possible to segment gestures and identify start and end points without requiring training data and with lesser computational power. Through this gesture extraction method, we could streamline the process of analysis and provide implementation of the calculations on a mobile device or portable computer during real-time gameplay. This information can be used to track the performance of players' movements for various exergaming applications.

## Applications in Rehabilitation

VR systems are becoming increasingly popular as a tool for rehabilitation for individuals with differing motor disorders [11,12]. Rehabilitation at home is critically needed as the durations of rehabilitation in hospitals are continually being decreased due to increasing costs [13]. Therefore, there is a need to provide portable and inexpensive rehabilitative tools for patients for use in their homes. Rehabilitation applications would benefit from the ability to discriminate between purposeful and other non-useful gestures being performed by participants in the virtual space. These purposeful gestures could be evaluated by clinicians remotely to track the rehabilitation performance of outpatients. During an example exergame application, participants targeted and popped balloons at their own pace resting when necessary. For purposes of rehabilitation, we were only interested in recognizing the balloon-popping movements and not the retraction of the arm for rest or waiting for the next balloon to pop. Using this approach, it was possible to perform gesture segmentation and real-time after VR gameplay. This allows clinicians to better understand the type of motions that are performed by the individual over time or after different gameplay scenarios [14]. The tool would allow clinicians to correct the participant if there is an erroneous gesture or a compensatory gesture being performed. Compensatory gestures are often considered a 'bad habit' during rehabilitation as it does not allow patients to recover complete usability of the targeted limb/muscle [15].

This gesture segmentation technique can be used to group similar gestures longitudinally to ensure that the performance of the patient is improving over time. For example, clinicians can track an upward arm motion relating to the performance of the deltoid muscles through comparison of metrics such as peak velocity of the gesture. Changes in this metric can be tracked over time to provide clinicians better insight on progress and make informed rehabilitation recommendations to improve motor outcomes or prescribe further corrective measures.

## Implications in Gesture Recognition in VR Systems

Many gesture recognition systems perform gesture segmentation using machine learning models based on a lexicon of pre-trained gestures. This requires training to be repeated several times for various people making it cumbersome in VR applications. Our velocity profile technique enables gesture segmentation is trained using the VR baseline movement test not from predefined gesture data. In addition, more computational methods result in slower gesture recognition preventing real-time gesture segmentation [16]. Through the use of gesture velocity profiles, it is possible to optimally perform computationally non-intensive segmentation.

For individuals with upper limb mobility impairments, it is difficult to universally characterize user-specific gestures in order to make generalizable assertions based on the idiosyncratic movements of an individual's motor capabilities [8]. The velocity profiles technique is advantageous since it does not require complex training algorithms or pre-determined gestures. This allows for greater flexibility in deployment of VR exergames for rehabilitation. In summary, through velocity profile-based gesture segmentation it is possible to extract individual gestures, which can then be used for unsupervised training for gesture recognition algorithms. It would be particularly beneficial in applications where it is not possible to follow a set of pre-determined gestures.

## ACKNOWLEDGEMENTS

Thanks to Dawn Neumann, George Hornby, and Emily Lucas from the Rehabilitation Hospital of Indiana for their help with recruitment of participants. We are grateful to the Indiana Spinal Cord and Brain Injury Research Foundation through the State Department of Health for their funding of this project. This research was also made possible through the Regenstrief Center for Healthcare Engineering at the Purdue Discovery Park.

## REFERENCES

- [1] Bhuyan MK, Ghosh D, Bora PK, Continuous Hand Gesture Segmentation and Co-articulation Detection, Springer, Berlin, Heidelberg, 2006, pp. 564–575.
- [2] Mitra S, Member S, Acharya T, Gesture Recognition: A Survey, *Appl. Rev.*, vol. 37, no. 3, 2007.
- [3] Pavlovic VI, Sharma R, Huang TS, Visual interpretation of hand gestures for human-computer interaction: A review, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 677–695, 1997.
- [4] Esculier, VJ, BP, GK, LE T, Home-based balance training program using the Wii and the Wii Fit for Parkinson's disease, *Movement Disorders*, vol. 26. pp. S125–S126, 2011.
- [5] Hofmann FG, Heyer P, Homme GI, Velocity profile based recognition of dynamic gestures with discrete hidden markov models, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1998, vol. 1371, pp. 81–95.
- [6] Kang H, Woo Lee C, Jung K, Recognition-based gesture spotting in video games, *Pattern Recognit. Lett.*, vol. 25, no. 15, pp. 1701–1714, Nov. 2004.
- [7] Hoffman M, Varcholik P, LaViola JJ, Breaking the status quo: Improving 3D gesture recognition with spatially convenient input devices, in *Proceedings - IEEE Virtual Reality*, 2010, pp. 59–66.
- [8] Jiang H, Duerstock BS, Wachs JP, A machine vision-based gestural interface for people with upper extremity physical impairments, *IEEE Trans. Syst. Man, Cybern. Syst.*, 2014.
- [9] Palaniappan SM, Duerstock BS, Developing Rehabilitation Practices Using Virtual Reality Exergaming, in *2018 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2018*, 2019, pp. 90–94.
- [10] Palaniappan SM, Zhang T, Duerstock BS, Identifying comfort areas in 3D space for persons with upper extremity mobility impairments using virtual reality, in *ASSETS 2019 - 21st International ACM SIGACCESS Conference on Computers and Accessibility*, 2019.
- [11] Rizzo AA, Schultheis M, Kerns KA, Mateer C, Analysis of assets for virtual reality applications in neuropsychology, *Neuropsychological Rehabilitation*. 2004.
- [12] Sveistrup H, Motor rehabilitation using virtual reality, *J NeuroEngineering and Rehabilitation*. 2004.
- [13] NSCIS Center, NSCISC National Spinal Cord Injury Statistical Center: 2014 Annual Report Complete Public Version: Spinal Cord Injury Model Systems., 2014.

- [14] Palaniappan SM, Suresh S, Haddad J, Duerstock BS, Adaptive Virtual Reality Exergame for Individualized Rehabilitation for Persons with Spinal Cord Injury, 16<sup>th</sup> European Conference on Computer Vision (ECCV), *In Proc. of 8<sup>th</sup> International Workshop on Assistive Computer Vision and Robotics (ACVR2020)*, Glasgow, UK, 2020.
- [15] Thielman GT, Dean CM, Gentile AM, Rehabilitation of reaching after stroke: Task-related training versus progressive resistive exercise, *Arch. Phys. Med. Rehabil.*, vol. 85, no. 10, pp. 1613–1618, Oct. 2004.
- [16] Sagayam KM, Hemanth DJ, Hand posture and gesture recognition techniques for virtual reality applications: a survey, *Virtual Real.*, vol. 21, no. 2, pp. 91–107, Jun. 2017.