Leimu: Gloveless Music Interaction Using a Wrist Mounted Leap Motion

Dom Brown
University of the West of England
Bristol, UK
Dom.Brown@uwe.ac.uk

Nathan Renney
University of the West of England
Bristol, UK
nathanrenney@hotmail.com

Adam Stark
Mi.mu Limited
London, UK
adam@mimugloves.com

Chris Nash
University of the West of England
Bristol, UK
Chris.Nash@uwe.ac.uk

Tom Mitchell
University of the West of England
Bristol, UK
Tom.Mitchell@uwe.ac.uk

ABSTRACT

Camera-based motion tracking has become a popular enabling technology for gestural human-computer interaction. However, the approach suffers from several limitations, which have been shown to be particularly problematic when employed within musical contexts. This paper presents Leimu, a wrist mount that couples a Leap Motion optical sensor with an inertial measurement unit to combine the benefits of wearable and camera-based motion tracking. Leimu is designed, developed and then evaluated using discourse and statistical analysis methods. Qualitative results indicate that users consider Leimu to be an effective interface for gestural music interaction and the quantitative results demonstrate that the interface offers improved tracking precision over a Leap Motion positioned on a table top.

Author Keywords

Leap Motion, Gestural Control, Digital Musical Instruments, IMU, Wearable Technology, Motion Tracking, Data Gloves.

1. INTRODUCTION

In traditional acoustic music performance, pitch, timing and timbre are primarily controlled through the fine motor activities of the hands. To reach comparable levels of fidelity in terms of control intimacy [22] and expression [7] with digital musical instruments, the fine motions of the fingers and hands must be precisely measured with high update rates and minimal latency [32]. Consequently, an important focus for research in computer music is the conversion of hand manipulations and gestures into digital signals. When interaction is mediated through manipulation of physical objects, precise sensing of surface and tactile interactions is sufficient for many music and performance applications [6, 16, 19]. Recent advancements in technology, combined with a renewed interest in virtual reality have motivated the development of consumer motion tracking systems, designed to capture the full range of human dexterity with a particular focus on mid-air and freehand gestural interaction [1, 2, 3, 4].

This paper presents and examines a wearable optical tracking approach to music interaction named Leimu. The system integrates a wrist mounted Leap Motion, a camera-based device originally intended for tabletop use, with an Inertial Measurement Unit (IMU) for tracking fine motor activities.

NIME'16, July 11-15, 2016, Griffith University, Brisbane, Australia.

2. BACKGROUND

There is a range of hand tracking techniques available for both general and music human-computer interaction (HCI), which can be broadly divided into three categories; camera-based, wearable and combined methods.

Camera-Based Motion Tracking

Camera-based motion tracking is widely used in general and music HCI. The technique typically couples cameras or infrared sensors with machine vision and recognition algorithms to estimate a subject's pose and motion [5, 8]. Not only does this approach mirror the physiological components of human visual perception but it is advantageous as subjects are relieved of physical impediments that could constrain their free movement.

However, camera-based motion tracking presents a number of drawbacks. For instance, the approach typically relies on cameras embedded in the environment, which can limit the interaction workspace to the camera's field of view [17] and can produce tracking errors when subjects are occluded [31]. Furthermore, the temporal precision of camera-based methods are also limited by frame rates and require computationally expensive machine vision algorithms, both of which can extend action-to-response times beyond what is acceptable for time-sensitive music applications [32].

Wearable Technology

Wearable technology is a popular solution for gestural control, and is achieved by a variety of methods, such as soft sensors [30] and data gloves [14, 28]. Wearable solutions are particularly prevalent in the field of gestural music interaction, due to the ease with which sensors can be positioned across the hand to detect joint angles, orientations and translations [14, 15, 20]. However, wearable technologies such as data gloves face challenges associated with reliability and maintenance [9], and they can also be invasive [25] and cumbersome [23].

Combined Approaches

There are a number of approaches that combine camera-based tracking and wearable technology. For example, Digits [17] uses an infrared camera to capture finger motion and an IMU for tracking hand orientation, while the Lightglove [13] uses LED scanner/receive sensor arrays and a two dimensional accelerometer to enable virtual typing and pointing.

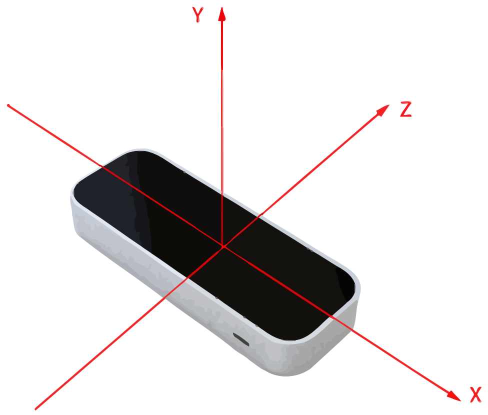

Figure 1: Cartesian Axes of the Leap Motion.

Figure 2: The Leimu wrist mount.

3. LEAP MOTION

The Leap Motion is a low-cost, consumer, camera-based gestural interaction device that has been designed for desktop use. It is able to recognise and track thin, cylindrical objects and hands by integrating the images captured by two optical infrared sensors. The device has a ~150° field of view, with an effective working distance in the region of 25–600 millimetres [1]. The device's accompanying software and development kit constructs a skeletal model of a user's hand, and exposes palm, finger and joint coordinates (Figure 1) to enable third party application development.

The Leap Motion is intended primarily for tabletop use, positioned on a flat surface, and has a limited workspace in the region above the device in which a user's hand may be tracked. Although the device and accompanying software provides highly accurate hand tracking in most circumstances, tracking is compromised when self-occlusion occurs. Furthermore, a recent study has also highlighted cases of user fatigue when operating the device for extended periods [24]. An in-depth analysis of the device's tracking accuracy has also been conducted by Guna et al [11], concluding that, while the device is able to accurately track static points, its consistency is dependant on the position of the hand, with accuracy diminishing in proportion to distance and at the periphery of the device's field of view. The analysis also found that the device produces an inconsistent update rate, with a mean of 40 Hz with significant jitter.

Several efforts have been made to evaluate the device's utility as an interface for music interaction, emulating existing musical interfaces such as a virtual keyboards and drum pads [27, 12]. Both studies highlighted usage difficulties in the absence of tactile and visual feedback when contrasted with their physical counterparts, while Silva et al noted that the device suffered unacceptable latency when used to trigger one-shot events.

4. LEIMU

Leimu has been constructed from a wrist mount comprising two 3D printed sections connected to the wearer's wrist and forearm using a pair of GoPro wrist straps (Figure 2). The connection between the two sections may be adjusted to enable the Leap Motion to be set in a range of positions with respect to the hand. Leimu may be worn with the Leap Motion positioned either above or below the hand, as the tracking algorithms are stable from either orientation.

This configuration has several advantages over a statically positioned Leap Motion. First, the issues of diminishing precision over greater distances and self-occlusion are minimised. Second, the user is not constrained to a limited desktop workspace. As the Leap Motion maintains a fixed position with respect to the wearer's hand, the palm position and orientation readings from the device are no longer meaningful. However, hand orientation and motion can be recovered by attaching a small IMU to the wrist section of the mount (Figure 2). The device used in this instance is an NGIMU - a calibrated, wireless IMU with an onboard AHRS fusion algorithm [18] providing an instantaneous estimation of orientation with respect to the earth coordinate frame. Furthermore, the high update rate of the IMU enables the accurate detection of time-sensitive musical gestures that might otherwise have been compromised by latency problems documented in prior work [27]. The Leap Motion and NGIMU have been integrated within a software interface to tools developed throughout previous studies [20, 21]. The skeletal geometry of a tracked hand is converted to 14 proximal, intermediate and distal phalange joint angles [10] for the fingers/thumb, along with orientation and gestural events derived from the IMU, and made available for analysis and arbitrary mapping to audio and music parameters.

5. EVALUATION

To assess the potential for Leimu as an interface for musical interaction the device was evaluated using a qualitative discourse analysis method and quantitative statistical analysis methods. The evaluations compared the Leimu, a table mounted Leap Motion and a data glove developed previously [20, 21]. The data glove incorporates eight resistive bend sensors; two at the proximal and intermediate joints of each finger save the thumb and little fingers, which are tracked with a single bend sensor, and is also equipped with an IMU to monitor orientation and other dynamic hand movements. The results of both studies are presented below.

5.1 Statistical Analysis

The Leimu, data glove and statically positioned Leap Motion were all evaluated to measure and compare the precision and repeatability of their joint angle measurements. An application was written to record 50 repetitions of 12 hand postures (Figure 3), 600 readings in total. The postures were displayed to the user in a randomised order. A single reading represented a vector of joint angles for each posture. The capturing exercise was completed with all three interface types and the same participant.

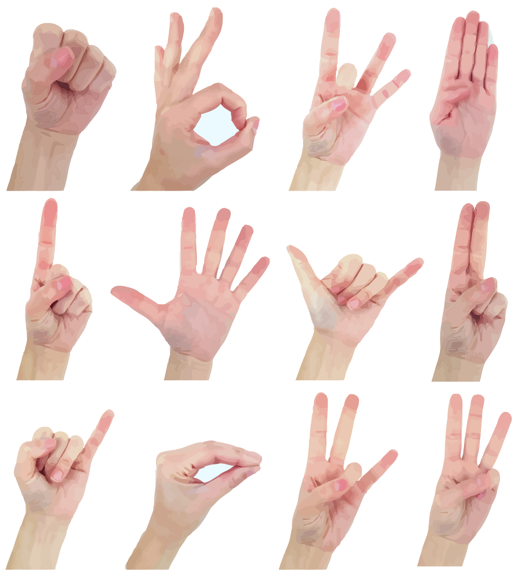

Figure 3: Posture test set

Results

Following the data collection exercise, the readings were processed to enable comparison between the three devices. The raw joint angle readings were normalised to flexion values in range of 0.0 to 1.0. The mean joint angle vector for each posture was then subtracted from each respective posture reading to centre the spread of flexion readings on zero. The results are shown in Figure 4, where each box plot represents all 600 readings for each flexion value. The labels 'P', 'I', and 'D' indicate the lower joints of the proximal, intermediate and distal phalanges respectively. The whiskers indicate 1.5× the interquartile range.

The results indicate that of the three interfaces, the data glove exhibits the smallest variance followed by the Leimu. This suggest that readings from the Leap Motion were more consistent when wrist mounted, rather than positioned on the desktop. The improvement seen in mounting the Leap in Leimu is likely to be due to the minimised possibility of self occlusion and the fixed positioning of the hand in the Leap Motion's field of view. This reflects what has been found in previous studies; that the Leap Motion's tracking accuracy wanes with distance from the sensor [11] and that it is lost entirely with self occlusion [24].

For additional indication of the relative variation between devices, Figure 5 shows the standard deviation of each joint angle for all three devices on a single plot. To give a like-for-like comparison, the distal joints of the Leap Motion readings were discarded for each finger and the mean of the two remaining joints of the little finger and thumb were taken to produce a joint angle vector that is comparable with the data glove readings. The plot also suggests that the greatest variation was measured in the intermediate joint angle readings, suggesting that some of this variance might be traced to anatomical variation in the finger positions, rather than deficiencies in the tracking apparatus.

It is important to highlight that the Leap Motion directly measures joint angle while the flex sensors in the data glove give a measure of resistance proportional to joint angle, which includes an element of non-linearity [26]. However, this becomes somewhat irrelevant, as postures where the flex sensors would be relatively straight compared to postures where the sensors are relatively bent will have the same variation in terms of joint angle.

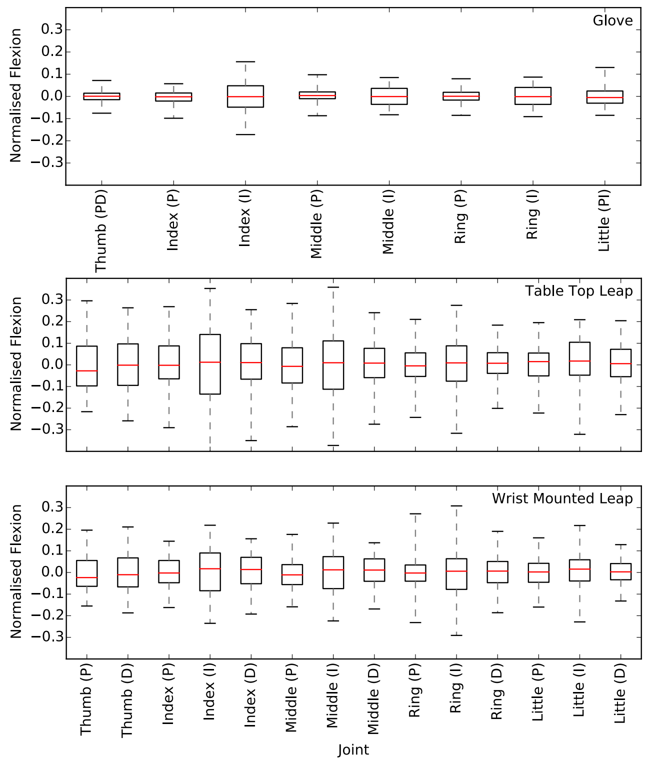

Figure 4: Normalised flexion readings for glove, Leap Motion and Leimu

5.2 Discourse Analysis

Discourse analysis is an effective qualitative approach which has been used previously by Stowell et al to evaluate digital musical instruments [29]. It uses a structured method that is able to draw more meaningful conclusions from user feedback than simple summaries. The analysis carried out invited users to contrast the Leimu in terms of musical expression with the data glove.

Method

Four users participated in the study, individually testing both the Leimu and the data glove with free and guided exploration, before engaging in a group discussion. The first interface encountered by each participant was alternated in an attempt to control any effect that ordering might have on the results. The Leimu and glove were mapped with a simple one-to-one strategy to control the parameters of the Logic Pro EFM1 synthesiser. Flexion readings controlled the FM 'amount' while roll controlled the amount of vibrato. The pitch orientation on both interfaces controlled note events.

Individual Sessions

User 1 felt that the Leimu was more suited to melodic use, while the data glove could achieve more timbral variation. They felt that the data glove was more expressive, and sensitive to their movements. While they preferred using the data glove, User 1 had positive feelings about both interfaces, and although mentioned that the Leimu's mount limited movement to a small extent, they did not feel expressively restricted.

User 2 preferred using the data glove over the Leimu. They commented on the Leimu's mount feeling insecure and expressed concern that the Leimu might be inaccurate due to the short distance between their hand and the Leap Motion. They also noted that the small amount of tactile

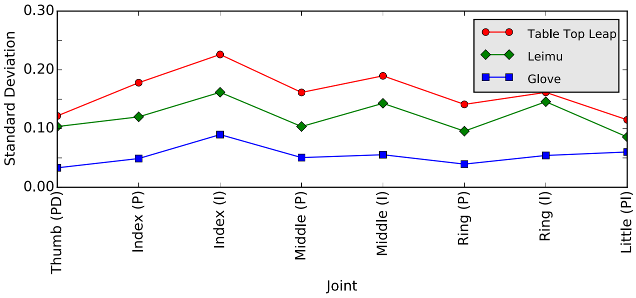

Figure 5: Flexion reading standard deviations for glove, Leap Motion and Leimu

feedback provided by the bend sensors in the data glove gave them more confidence that their actions were producing a direct response. They also thought that the greater weight of the Leimu would affect performance.

User 3 felt that the data glove gave a more fluid response than the Leimu, with finger flexion producing a smoother audio response. They also felt that the glove was more expressive due to the tactile feedback provided with the bend sensors. They thought that the Leimu occasionally "felt cumbersome", but also expressed that they felt that both interfaces would be unsuitable for "precision work" but could suit more abstract or noise based composition and performance.

User 4 immediately commented on the difference in timbral quality between the two interfaces, with the data glove "sounding better", with "more range", but they felt that the Leimu was more suited to melodic use, feeling the relationship between their hand position and note triggering was clearer and more precise. They preferred using the data glove but thought that the Leimu was "no less practical".

Group Session and Discussion

From both the solo and group sessions it was clear that the data glove was the preferred interface. However, the Leimu was also noted to be an effective controller. One issue raised in the group discussion was that the Leimu's mount was prone to slipping, which made participants more reluctant to engage in vigorous movements. The group did not perceive any delays between their input and the system's response with either interface, but all felt more confident that their movements were being tracked accurately due to the small amount of tactile feedback and resistance to motion that the bend sensors provided.

It was clear that design improvements could be made to the Leimu's mount. Future design iterations will focus on reducing the overall weight and potentially rehousing the Leap Motion's electronics within the mount itself. Stability improvements could also be made with additional wrist straps and/or supporting arms.

6. CONCLUSIONS AND FUTURE WORK

In this paper the Leimu has been presented, which is an interface that combines the benefits of camera-based and wearable motion tracking technology by wrist mounting a Leap Motion and inertial measurement unit (IMU). The aim was to address issues that have been raised in previous Leap Motion studies that limit its potential as an interface for musical interaction. By mounting the Leap Motion on the wrist, wearers are not limited to a workspace defined by its static position. Furthermore, by recovering wrist orientation and motion from a high update rate IMU, dynamic gestures can be recognised and used to trigger time sensitive musical events with acceptable latency.

The Leimu wrist mount was designed, manufactured and evaluated using both qualitative and quantitative analysis methods, and found to be an effective device for musical control as well as improving the tracking precision of the Leap Motion when compared with conventional desktop placement. The comparative study between the Leimu and the data glove raises issues that might be more widely indicative of trends in optical and wearable technologies generally, which highlights an opportunity for future research. Further work will also include improving the wrist mount design by making it lighter and more stable and releasing the software and 3D printed designs to the community, as well as undertaking more in-depth evaluations of the improved mount.

7. ACKNOWLEDGMENTS

The authors would like to thank the study participants who lent their time and to the many supporters of this work, including the University of the West of England, Innovate UK, x-io Technologies and Mi.mu Limited.

8. REFERENCES

[1] Leap motion. `leapmotion.com`, 2016. Accessed: 11th January 2016.

[2] Manus vr. `https://manus-vr.com/`, 2016. Accessed: 31st January 2016.

[3] Mi.mu gloves. `http://mimugloves.com`, 2016. Accessed: 31st January 2016.

[4] Myo. `https://www.myo.com/`, 2016. Accessed: 31st January 2016.

[5] J. K. Aggarwal and Q. Cai. Human motion analysis: A review. *Computer Vision and Image Understanding*, 73(3):428–440, 1999.

[6] F. Bevilacqua, N. Schnell, N. Rasamimanana, J. Bloit, E. Flety, B. Caramiaux, J. Françoise, and E. Boyer. De-mo: designing action-sound relationships with the mo interfaces. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, 2013.

[7] C. Dobrian and D. Koppelman. The 'e' in nime: Musical expression with new computer interfaces. In *Proc. of New Interfaces for Musical Expression (NIME)*, Paris, France, 4th–8th June 2006.

[8] D. M. Gavrila. The visual analysis of human movement: A survey. *Computer vision and image understanding*, 73(1):82–98, 1999.

[9] K. Gniotek and I. Krucinska. The basic problems of textronics. *FIBRES AND TEXTILES IN EASTERN EUROPE*, 12(1):13–16, 2004.

[10] H. Gray. *Anatomy of the human body.* Lea & Febiger, 1918.

[11] J. Guna, G. Jakus, M. Pogačnik, S. Tomažič, and J. Sodnik. An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking. *Sensors*, 14(2):3702–3720, 2014.

[12] J. Han and N. Gold. Lessons learned in exploring the leap motion ™ sensor for gesture-based instrument design. In *Proc. of New Interfaces for Musical Expression (NIME)*, pages 371–374, London, UK, 30th June–4th July 2014.

[13] B. Howard and S. Howard. Lightglove: Wrist-worn virtual typing and pointing. In *Fifth International Symposium on Wearable Computers*, pages 172–173, Zurich, Switzerland, 7th–9th October 2001. IEEE.

[14] E. Jessop. The vocal augmentation and manipulation prosthesis (vamp): A conducting-based gestural controller for vocal performance. In *Proc. of New Interfaces for Musical Expression (NIME)*, Pittsburgh, USA, 4th–6th June 2009.

[15] S. Jiang, K. Sakai, M. Yamada, J. Fujimoto, H. Hidaka, K. Okabayashi, and Y. Murase. Developing a wearable wrist glove for fieldwork support: A user activity-driven approach. In *IEEE/SICE International Symposium on System Integration (SII)*, pages 22–27, Tokyo, Japan, 13th–15th December 2014. IEEE.

[16] S. Jordà, G. Geiger, M. Alonso, and M. Kaltenbrunner. The reactable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Proc. of the 1st International Conference on Tangible and Embedded Interaction*, Baton Rouge, USA, 15th–17th June 2007.

[17] D. Kim, O. Hilliges, S. Izadi, A. D. Butler, J. Chen, I. Oikonomidis, and P. Olivier. Digits: freehand 3d interactions anywhere using a wrist-worn gloveless sensor. In *Proc. of the 25th annual ACM symposium on User Interface Software and Technology*, pages 167–176, Cambridge, USA, 7th–10th October 2012. ACM.

[18] S. O. Madgwick. An efficient orientation filter for inertial and inertial/magnetic sensor arrays. *Report x-io and University of Bristol (UK)*, 2010.

[19] A. McPherson. Buttons, handles, and keys: Advances in continuous-control keyboard instruments. *Computer Music Journal*, 39(2), 2015.

[20] T. Mitchell and I. Heap. Soundgrasp: A gestural interface for the performance of live music. In *Proc. of New Interfaces for Musical Expression (NIME)*, Oslo, Norway, 30th May–1st June 2011.

[21] T. J. Mitchell, S. Madgwick, and I. Heap. Musical interaction with hand posture and orientation: A toolbox of gestural control mechanisms. In *Proc. of New Interfaces for Musical Expression (NIME)*, Ann Arbor, USA, 21st–23rd May 2012.

[22] F. R. Moore. The dysfunctions of midi. *Computer Music Journal*, 12(1), 1988.

[23] V. Pavlovic, R. Sharma, T. S. Huang, et al. Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):677–695, 1997.

[24] L. E. Potter, J. Araullo, and L. Carter. The leap motion controller: a view on sign language. In *Proc. of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*, pages 175–178, Adelaide, Australia, 25th–29th November 2013. ACM.

[25] J. Rehg and T. Kanade. Digiteyes: vision-based hand tracking for human-computer interaction. In *Proc. of the 1994 IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, pages 16–22, Austin, USA, 11th–12th November 1994.

[26] G. Saggio, F. Giannini, M. Todisco, and G. Costantini. A data glove based sensor interface to expressively control musical processes. In *4th IEEE International Workshop on Advances in Sensors and Interfaces (IWASI)*, pages 192–195, Savelletri di Fasano, Italy, 28th–29th June 2011. IEEE.

[27] E. S. Silva, J. A. O. de Abreu, J. H. P. de Almeida, V. Teichrieb, and G. L. Ramalho. A preliminary evaluation of the leap motion sensor as controller of new digital musical instruments. 2013.

[28] L. Sonami. Lady's glove. `sonami.net/ladys-glove/`. Accessed: 20th January 2016.

[29] D. Stowell, M. D. Plumbley, and N. Bryan-Kinns. Discourse analysis evaluation method for expressive musical interfaces. In *Proc. of New Interfaces for Musical Expression (NIME)*, pages 81–86, Genova, Italy, 5th–7th June 2008.

[30] D. M. Vogt and R. J. Wood. Wrist angle measurements using soft sensors. In *SENSORS*, pages 1631–1634. IEEE, 2014.

[31] R. Y. Wang and J. Popović. Real-time hand-tracking with a color glove. *ACM Transactions on Graphics (TOG) - Proc. of ACM SIGGRAPH 2009*, 28(3), 2009.

[32] D. Wessel and M. Wright. Problems and prospects for intimate musical control of computers. *Computer Music Journal*, 26(3):11–22, 2002.