Hand Posture Recognition: IR, IMU and sEMG

Richard Polfreman
Music, University of Southampton
University Road, Southampton,
SO17 1BJ, UK
r.polfreman@soton.ac.uk

ABSTRACT

Hands are important anatomical structures for musical performance, and recent developments in input device technology have allowed rather detailed capture of hand gestures using consumer-level products. While in some musical contexts, detailed hand and finger movements are required, in others it is sufficient to communicate discrete hand postures to indicate selection or other state changes. This research compared three approaches to capturing hand gestures where the shape of the hand, i.e. the relative positions and angles of finger joints, are an important part of the gesture. A number of sensor types can be used to capture information about hand posture, each of which has various practical advantages and disadvantages for music applications. This study compared three approaches, using optical, inertial and muscular information, with three sets of 5 hand postures (i.e. static gestures) and gesture recognition algorithms applied to the device data, aiming to determine which methods are most effective.

Author Keywords

Hand posture, Gesture recognition, Motion capture

CCS Concepts

• **Human-centered computing~Gestural input** • **Applied computing~Performing arts** • Social and professional topics~Hardware selection

1. INTRODUCTION

Hand gesture recognition is of interest to a number of fields such as sign-language interpretation, robotics, prosthetics, virtual reality, health applications, video games, as well as computer music. Recent devices have brought much of the necessary technology to consumer level systems, such as the Leap Motion [19] and Kinect [24] expanding the potential user base beyond scientific and research environments and into everyday use.

Real-time hand gesture recognition systems typically comprise of a hardware sensor arrangement to supply a stream of data about the user's hand, signal processing algorithms to extract features from the data and machine-learning tools to then determine the current gesture. A number of different sensor types can be used at the input stage, and in this work we wanted to compare the performance of optical (IR depth camera), surface electromyography (sEMG), and inertial measurement units (IMUs), although other sensor technologies can be used for hand tracking such as electromagnetic sensing (e.g. Polhemus) and bend sensors (e.g. Cyberglove).

1.1 Optical Systems

Optical approaches can include the use of standard video cameras, IR depth cameras and multi-camera motion capture systems, with both marker-less or marker-based techniques. Each type of system has a number of practical difficulties and advantages, but generally optical systems are subject to lighting issues, occlusion/line-of-sight problems and often quite severe spatial/orientation constraints. High-end multi-camera mo-cap systems can provide highly accurate spatial information and fast response, but can be difficult to setup in a concert environment and can be prohibitively expensive. Video cameras can be cheap and simple to set up, at the cost of spatial constraints and occlusion problems, while IR depth-camera hardware is now readily available and affordable, but again with some line-of-sight and lighting/noise problems.

Microsoft's Kinect provides full body skeletal information but currently provides little direct hand information – the second generation device providing thumb orientation and some basic hand-posture detection (*lasso*, *fist*, *open*). Others have used the Kinect or other depth cameras to track fully articulated hand gestures, including Microsoft (although not released), Intel (RealSense) and other research teams (FORTH), although often at low frame-rates and/or very high computational cost requiring GPU processing [38]. Both generations of Kinect have been used in a number of installations and music applications, including [21], [39], [33], [15]. The Leap Motion controller currently provides a detailed skeletal model of both hands and fingers and again has been explored in a number of music applications, including [11], [12], [4]. It has also been adopted commercially by some music technology companies, including Steinberg for controlling their Cubase software [37] and Fairlight (now Blackmagic Design), who embedded a Leap Motion controller into their 3D Audio Workstation for "Air panning" of sounds [2].

Camera frame-rates are an important consideration with optical systems, as this sets an upper-limit on the speed of tracking and gesture recognition. Low-cost industrial USB cameras can achieve 120fps (at the cost of resolution), high-end multi-camera mo-cap systems typically 100-500fps, while consumer depth camera systems range from ~30-100fps. In our experience, frame rates of ~50fps and higher provide systems which feel responsive in music applications, depending on the subsequent processing latency.

1.2 sEMG Systems

sEMG devices need to have skin contact near to the muscles of interest, which for hand control are in the upper forearm. They measure the electrical activity in the muscle fibres, which correlates with the exertion being made. Holding the hand in different finger postures requires different patterns of muscle activation in the forearm, and therefore the sEMG data can be interpreted to say something about the shape of the hand, although of course the sEMG data will vary with the pressure being exerted by the fingers (e.g. if a fist is being squeezed tightly or more relaxed). Indeed it is possible to tense the forearm muscles voluntarily without changing hand posture, which will affect the sEMG signals being generated.

Nymoen [29] explored a wireless sEMG device, the Myo Armband, as a digital musical instrument controller, and found issues with misattribution of hand postures, and jitter in posture detection response as limiting factors in the utility of the device in music performance. Despite this, wireless sEMG devices offer a potential solution to hand posture capture, without the lighting, spatial and line-of-sight limitations of optical systems, or the intrusion of gloves.

NIME'18, June 3-6, 2018, Blacksburg, Virginia, USA.

1.3 IMU Systems

MEMS-based IMUs are small multi-axis devices that measure the orientation and forces acting on the device. 9-axis IMUs include accelerometer, gyroscope and magnetometer (3 axes each), the signals from which can be used to track the motion of the device from a known starting point. IMUs are becoming pervasive in consumer electronic devices: fitted as standard into smart-phones, some fitness trackers and wearables, used in a number of video game controllers, such as the Nintendo wiimote and Sony DualShock 4 and VR headsets such as the Occulus Rift, and form a key part of the flight control system for drones.

For capturing hand gestures, a number of IMUs can be attached to the hand and fingers to report the relative positions of the different parts of the hand and therefore the the hand posture. For convenience the sensors can be attached to a glove, which is then worn, but this can be quite intrusive and physically cumbersome to wear while performing. IMUs do suffer from drift, but generally provide spatial freedom (if connected wirelessly), good data rates and low latency (100+fps is possible) - important factors for music applications.

"Data gloves" are a related technology, but usually involve flex sensors in the fingers rather than IMUs for the posture detection, but may have an IMU for overall hand orientation. These have typically been expensive and/or custom-made. Glove-based controllers specifically for music applications have been developed by a number of researcher/performers, with varying levels of detail in the hand capture. These include the Tod Machover's use of the *Exos Dextrous Hand Master* [31], Morita's conducting follower with data glove [27], *Lady's Glove* [3], *VAMP* [16] and *mi.mu* gloves [26], a design adopted by others such as [35] in control of physical model synthesis.

IMUs should provide reliable data for gesture recognition in this work, free from spatial, line-of-sight and finger posture confusion, but are the most intrusive in terms of disrupting the performer's natural hand state and can suffer from drift and bias issues.

2. HAND POSTURE RECOGNITION

A large number of studies over the past ~30 years have explored hand gesture recognition in a number of application areas, although most commonly in sign-language reading, using either data-gloves, sEMG or optical systems as the input device. Some studies use dynamic gestures, since often sign-language words often involve movement as well as a hand posture, while others consider static gestures, which we refer to as postures here.

2.1 Optical

Huang [13] used a 3D neural network to successfully recognise 15 sign language gestures on video with a recognition rate of 91%, but this was far from real-time at the time (10s processing time). [Chen et al 2003] used Hidden Markov Models (HMM) to recognise 20 different sign-language gestures in real-time with a 30fps camera system, with a success rate of 93% and a latency of about 1s. HMMs are often used with dynamic gestures since they consider a time series of states, rather than the static postures that we consider here.

[8] achieved an accuracy of 85-91% with a set of six hand postures (open, fist, extended thumb in N, S, E, W directions) using 320x240 pixel video data with Adaboost for hand detection, adaptive segmentation and followed by multi-scale feature extraction (blob and ridge detection to identify hand and fingers). This process was rather slow for music applications (~100ms frame processing), but would be quicker with more recent computing hardware. This required the hand to be placed in a somewhat constrained space and orientation relative to the camera to be detected.

[16] used a depth camera with Action Graphs (an HMM variant) to recognise 12 sign-language dynamic gestures, the best performing variant using hand silhouette feature extraction and achieving accuracy of ~88%.

[23] unusually used a combination of Leap Motion and Kinect devices for the detection of ten (static) sign-language gestures, finding that they could increase accurate detection from using the Leap Motion alone. For each device, features were extracted from the device output data (e.g. orthogonal fingertip distance from palm plane from the Leap) and then used to train a multiclass Support Vector Machine (SVM) classifier. With the Leap Motion alone, a combination of fingertip distances and angles produced an accuracy of 80.86%, the Kinect alone 89.28%, and the combination 91.28%. Interestingly the Kinect alone produced better results than the Leap Motion alone, while the question of whether using two Leap Motion devices at different positions/orientations could produce similarly improved results is unanswered. In our work we wanted to isolate devices and use them independently, but we included some results of combined device data for comparison.

[4] proposed a solution to the limited capture space of the Leap Motion by developing a wrist mount to hold the sensor in a fixed position relative to the arm for musical applications. They compared the performance of the wrist mounted Leap Motion with a static Leap Motion and a data glove with 12 hand postures, although not with a recognition system – they simply compared the finger flexion measurements from each set-up. The glove appeared to be most reliable, while wrist-mounting the Leap Motion aided its accuracy by keeping the hand close to the device (accuracy declines with distance [40]), limiting hand self-occlusion issues.

2.2 IMU/Data Gloves

[Dipietro et al.] provide an extensive review of glove technologies developed over many years, and outline various applications using gesture recognition, including the Data Entry Glove (1983), which mapped sensor output patterns to alphanumerical characters in hardware, while the Pinch Glove (1990s) detects postures through surface electrical contacts on the glove which when meet indicate aspects of the hand posture without any further recognition required, with up to 1000 postures theoretically possible.

[7] identify a number of researchers that have used data gloves for sign-language recognition, including: Tahakashi and Kshino (1991) who used PCA and Cluster Analysis to code static sign language gestures with a similarity function to discriminate 30 of 46 gestures tested in pseudo real-time; Murakami and Taguchi who used recurrent 3-layer neural networks with a set of 42 dynamic gestures with a 77% success rate for subjects not in the training set, 98% for those that were; Mehdi and Khan who used ANNs to achieve 88% accuracy with 24 sign-language gestures; Liang and Ouhyoung who achieved an average success rate of 80% with 250 (dynamic) sign language gestures using HMMs and dynamic programming. More recently [14] used 18 language signs with a data glove, first segmenting the input stream based on local minima in signal variation to find gestures, then centroid-based clustering of training data and a K-NN algorithm to classify gestures with an accuracy of 95% with static gestures. [22] used a Probabilistic Neural Network (PNN) with 12 static hand gestures, using a combination of clustering algorithms selected by minimum classification error to achieve a success rate of 86%.

Mitchell at al. [26] developed a glove based music performance system using flex sensors and distinguished 8 hand postures using an ANN in the form of an MLP with back propagation supervised learning. While claiming robust results, no accuracy was quoted.

2.3 sEMG

sEMG can achieve high levels of recognition, although usually requires the sensors to be adhered to the skin and require very careful placement. [17] used 2-channel sEMG forearm data (*flexor carpi radialis* and *flexor carpi radialis brevis* muscles) to identify 9 sign-language gestures with 99.7% accuracy, using a number of time and frequency domain features extracted from the sEMG data and discriminant analysis to classify the data.

In [5] a Wavelet Packet Transform (WPT) was applied with Linear Discriminant Analysis (LDA) to 4 channel sEMG data (*ex- tensor digitorum, extensor carpi radialis, palmaris longus,* and *flexor carpi ulnaris*), followed by an MLP classifier to identify 9 dynamic gestures with an accuracy of 97.4%.

[1] used artificial neural networks (ANN) with 2-channel sEMG (*brachioradiali*, *flexor carpum ulnaris* plus wrist reference) to identify 4 hand movements with 88% success, while [34] used 2-channel sEMG (*flexor capri ulnaris* and *extensor capri radialis, longus* and *brevis*) with Empirical Mode Decomposition (EMD), extracting 8 standard features from both EMD and raw sEMG, then using a linear classifier. They found that including the EMD data improved their classification by an average of 2.6%.

8-channel sEMG forearm data (8 sensors placed around the forearm), with moving-average pre-processing and Class Augmented Principal Component Analysis (CA-PCA) for feature extraction and an Extreme Learning Machine (ELM) classifier – a form of ANN for fast training – achieved 92% accuracy across 4 hand postures in [20].

[29] achieved 98% accuracy with 2-channel sEMG (*flexor carpi radialis* and *extensor carpi radialis longus*) and a set of 6 gestures using a Mean absolute Value (MAV) feature with an SVM classifier optimized via a Cuckoo Search algorithm.

3. EXPERIMENTS

3.1 Hardware Devices

The aim of the experiments was to compare consumer level devices which can easily be deployed in musical contexts, hence high-performance high-cost devices, such as a Vicon mo-cap system, were eschewed in favour of the Leap Motion for optical capture and the Myo Armband for sEMG. A Perception Neuron system was used for MEMs based mo-cap, which while more expensive than the other devices, at ~1200USD for a complete suit including two hands, still sits within range of reasonable costs in music. Noitom has announced a new consumer VR glove containing 6 9-axis IMUs per finger embedded in the glove (rather than the surface mounted Neurons) aimed at the consumer market and which appears less intrusive than their Neuron system, but the device was not available at the time of writing for testing.

3.1.1 Leap Motion controller

The Leap Motion (2013) is a small USB device using stereo infrared cameras to view a three-dimensional space above the device. Using the image data, the software generates a detailed 3D model of hand and finger joint positions and orientations for both hands at usefully high frame rates for music of the order 100 Hz (previous software versions could run at higher speeds with less hand detail). Spatial accuracy is around 0.2 mm, [38] with a precision of 0.5 mm [10] using the earlier software – the more recent software appears to have lower accuracy and precision, while providing more detailed hand skeleton data.

A number of problems emerge in practice: lighting changes, occlusion of fingers (either within hand or between hands) due to single view point, mislabelling of left and right hands, some erratic positioning and jitter, and spatial limitations (e.g. there is a limited tracking space, and users' hands can easily exit the tracking zone accidently, as the bounds are not visible to the user). [23] used the Leap Motion in their studies, but this was the earlier software version which provided less detailed information regarding the hand – chiefly fingertip and palm position and orientation, while here we used the more recent software which now provides information on all the finger bones as well as always including finger data, even if the fingers cannot be clearly seen by the device.

An important issue with the Leap is the orientation of the device relative to the hands, as this will alter the view of the hands seen by the device, and therefore the nature of the hand self-occlusion effects [11]. In this work, the orientation of the device for each posture set was chosen to optimise the Leap response for those hand shapes.

Figure 1. The Leap Motion controller

3.1.2 Perception Neuron

The Perception Neuron (2014) is a (relatively) low-cost motion capture suit using a up to 32 neurons arranged around the wearer's body, each one containing a 9-axis MEMs-based IMU. The suit can communicate wirelessly (or via USB) with software running on a PC, which converts signals from the neurons into skeletal motion data for real-time display, recording and transmission over TCP/UDP networks. Using IMUs and a wifi connection, it allows complete freedom of movement over a significant space (such as a stage).

In this work only the hand is of interest, which is captured via a glove arranged with 9 sensors (2 thumb, 2 fore-finger, 1 each on middle, ring and little fingers, back of the hand and wrist). The glove is connected to a UART hub which then either connects to a wireless hub or directly by USB to the host computer. The system operates at a default frame-rate of 120Hz. For music applications, wearing the glove is quite intrusive, while strong magnetic fields can be problematic for the hardware, which needs to be kept apart from various electrical devices that generate such fields. While calibrated at the beginning of each session, it was fairly common for the finger positions to become confused during the experiments, presumably due to interference and drift.

Figure 2. Gloves from the Perception Neuron mo-cap suit.

3.1.3 Myo Armband

Thalmic Labs Myo Armband (2014) is a bracelet format device which is worn on the forearm of the user and which uses 8 electrical sensors around the bracelet to infer information about muscle activity (EMG) in the forearm and hence the disposition of the users hand and fingers. In addition, the device contains a 9-axis IMU sensor, which is used to provide orientation data as a quaternion. The Myo software that accompanies the device recognizes 5 hand postures (see fig. 4): relaxed, fist, extended, outward wrist (i.e. wrist stretched backwards), and inward wrist (i.e. hand stretched toward the inner forearm). The device can also be calibrated for a particular user. In practical use a number of issues arise, including placement of the armband (so that it is always positioned in the same place on the arm and at the same angle around the arm), misidentification of postures (i.e. wrong posture is indicated), wearing the band for long periods can be quite tiring, while delays and jitter in posture recognition time can be challenging for musical applications and frustrating for performances.

Figure 3. Thalmic Labs' Myo Armband

3.2 Test Postures

Three sets of static hand postures formed the basis of the study in order to give a reasonable comparison between the devices and for particular musical applications. Each set was the same size (five) and was selected on the basis of being particularly suited to one of the devices or for human ease of use.

3.2.1 Myo postures

It seems a reasonable expectation that the Myo has been optimised to detect the postures it reports, and that the device in someway is suited to detect these postures in terms of different muscle activations. This set (figure 4) was used with all devices, but with the expectation that the Myo should perform optimally here: Relaxed, Fist, Spread, Wave-Out, Wave-In.

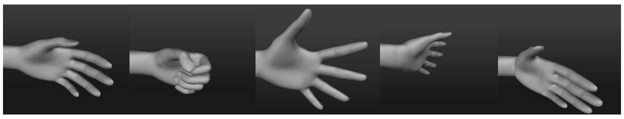

Figure 4. The Myo native set of hand postures.

3.2.2 Thumb Opposition postures

As well as being an inherently human set of postures, these can be made easily in conjunction with other hand movements that might be useful for musical control, e.g. beat/tempo indications or other parameter variations. These postures (see figure 5) activate the tendons of the forearm differently and therefore are potentially separable by the Myo, while Leap Motion and Neuron should be able to use within hand proximity values to identify the postures. It was expected that the Leap however might suffer from finger-finger and hand-finger occlusion issues in various hand orientations, and some experimentation was completed to find the best placement and orientation relative to the Leap Motion. This set was seen as the most user-friendly in terms of ease of use, minimal energy expended to maintain, and ease of combination with more global hand movement and orientation: Relaxed, Thumb-Index, Thumb-Middle, Thumb-Ring, Thumb-Little. It was also seen as most suitable for the Neuron device, since there should be a clear differentiation between postures in the mo-cap data.

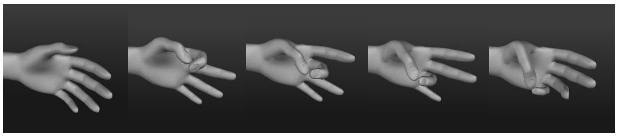

Figure 5. The Thumb Opposition set of hand postures.

3.2.3 Finger Point postures

The third set of postures were aimed as being most Leap Motion-friendly, in that this device should be able to clearly identify the postures with minimal occlusion issues. These involve extending different numbers of (thumb and) fingers, starting from a fist and ending at all fingers extended. Due the difficulty of extending the little and ring fingers independently for some people, these were either extended or retracted together, giving 5 postures altogether as with the other sets. See figure 6: Fist, Thumb, Thumb+Index, Thumb+Index+Middle, Thumb+All.

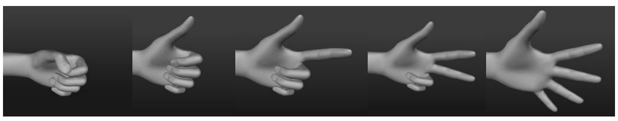

Figure 6. The finger point set of hand postures.

3.3 Software

Four software tools were initially used in the study: Max/MSP was used to interface with Myo and Leap Motion devices, record the data stream and also transmit data to/from the Gesture Recognition Toolkit (GRT) [9] for gesture identification, while Axis Neuron Pro was used to stream the Neuron data to Max for decoding and passing on to GRT. Figure 7 shows the Max Myo user interface used and similar interfaces were designed for the Leap and Neuron.

GRT is a convenient Open-Source tool for real-time gesture recognition that offers a range of machine-learning algorithms and pre-processing that can be easily tested against data sets.

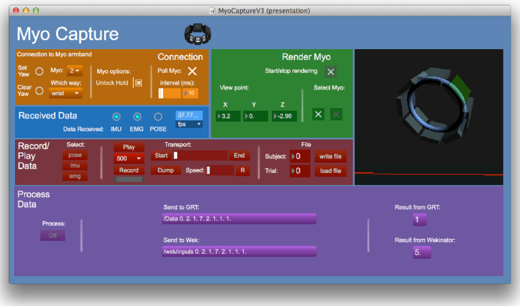

Figure 7. Myo interface in Max/MSP.

Of the three systems tested, the Leap is the only one for which line-of-sight is important, and so some consideration was given to the placement of the device relative to the hand. For the poses selected, it seemed clear that the best line of sight for the device would be to mount the Leap vertically beside the hand so that the necessary finger details would be visible to the device. The subjects' hands were placed approximately 20cm from the device, with palm centred on the Leap.

At the time of writing, the calculation data (describing the position information of the wearer's bones) transmitted by the host software (Axis Neuron Pro) did not include finger information and so BVH (BioVision Hierarchy) format data was used instead. This was decoded in a bespoke Max object (using Javascript) to provide the finger positions. The current Axis Neuron software also does not include complete information about the fingers, most notably missing adduction/abduction of the fingers, which places some limitations on gesture differences that are possible to detect.

3.4 Training and Testing

Each posture was captured for 500 samples from the device and each subject made the posture 3 times. The devices were capture successively since the Neuron glove would interfere with Leap capture quality. The samples captured variations from noise and involuntary human movement, while repeated poses captured variation in the subjects' making of the hand posture. While some approaches capture the transition from one pose to another (dynamic capture), in this first set of experiments we captured static poses only. After informal testing with a variety of classifiers in GRT, SVM was used.

For each device pre-processing was used prior to the classifier. For the Myo, rectified output from the 8 EMG sensors was smoothed with a 5-point moving average filter. For the Leap, the data processing varied according to the posture set. Given that its spatial data will depend on the size of the subject's hand, distances and position vectors were normalised by scaling by the wrist to middle finger tip distance with the hand extended – taken from the extended pose from the Myo set. For the thumb opposition postures, the distances between finger-tips and thumb-tip were used, while for the Myo and pointing postures, wrist-to-fingertip distances were used. Similar data were used from the Neuron since using hand-relative parameters would help minimise global drift problems.

In these tests the system was trained on twelve examples of each posture (30000 samples) and tested against three examples of each posture (1500 samples), not from the training set. The testing participant was selected at random from amongst the participants captured for each gesture tested. Within subject tests were not included here, although these are also of interest as training systems for specific individuals is a valuable use-case for music performance.

4. RESULTS

The initial results are summarized in tables 1-3 below. These show the relative performance of each device across the test data, which was the primary aim of the research, rather than optimizing the gesture recognition for the different hardware. Each table represents one of the sets of hand postures and shows the accuracy achieved, as well as the precision (above) and recall (below) for each of the five postures. For the Myo postures, the Myo did in fact perform very well, better than both the Leap and Neuron, although not high enough for critical music performance events. It was noted that during the testing, the Myo's own posture labeling was also below 100% - with some postures failing to be detected, and others mislabeled.

Table 1. Results for the Myo posture set

	Accuracy	P1	P2	P3	P4	P5
Leap	79.65%	1	1	1	1	0.50
		0.65	1	1	0.33	1
Myo	92.39%	0.97	0.94	0.84	0.93	0.91
		0.98	0.85	0.89	0.91	0.96
Neuron	68.81%	0.99	1	0.25	0.75	043
		0.67	1	0.11	0.99	0.67

All devices performed less well with the Oppose postures – the Leap seemed confuse the finger positions, although tracked the little finger pinch well, while the Neuron data appeared very good in some cases in terms of representing the hand shape, in other examples it clearly did not. This time the Myo performed the least well, appearing to not be able to distinguish the muscle activity effectively for these hand differences.

Table 2. Results for the Oppose posture set

	Accuracy	P1	P2	P3	P4	P5
Leap	65.59%	0.53	0.64	0.74	0.43	1
		0.38	1	0.33	0.55	1
Myo	48.13%	0.81	0.36	0.41	0.58	0.97
		0.22	0.94	0.30	0.65	0.30
Neuron	56.70%	0.46	0	0.99	0	0.5
		1	0	0.83	0	1

As expected the point posture set (table 3) seems most promising for the Leap motion sensor, although it seems to have difficulty distinguishing the fist and thumbs up postures on the basis of the fingertip to wrist distances. The Neuron should have been able to perform similarly well, but its failure to reliably maintain good position information about the fingers appears to hamper this result significantly. The Myo again performed relatively poorly, although better than for the oppose set, perhaps as here there are larger changes in hand shape.

Table 3. Results for the Point posture set

	Accuracy	P1	P2	P3	P4	P5
Leap	80.0%	1	0.5	1	1	1
		0.01	1	1	1	1
Myo	56.9%	0.45	0.62	0.25	0.59	0.76
		0.99	0.07	0.01	0.78	0.99
Neuron	67.0%	0.46	0	0.99	00	0.5
		1	0	0.83	0	1

5. DISCUSSION AND FURTHER WORK

While the devices used in these tests are attractive for musical control, there remain a number of issues which can lead to frustrating performance experiences. In our testing, the Leap would quite often mislabel the hand left rather than right, or simply lose tracking completely while the hand was stationary. In some poses, the tracking clearly became noisier as the tracking algorithm struggled maintain a lock. The Neuron suffered from frequent drift, and anomalies in the finger data – e.g. in one data capture the finger flipped periodically between two very different positions, even though the subject's hand was stationary. The Myo also sometimes had odd patterns of output (e.g. bursts of high values across all the sEMG sensors). Beyond these issues, the results indicate clearly the importance of organising posture designs and device selection to work together to give the most accurate results. Informal testing also indicated significant performance differences between classification algorithms. Although SVM appeared to give good results across devices, this warrants further investigation.

Future work will examine how to improve the quality and consistency of the raw device data (e.g. lighting and positioning for the Leap), alternative data pre-processing and testing with a range of classification algorithms. Including one of the Kinect hand-tracking systems would be desirable due to the potentially larger tracking volume available compared with the Leap.

6. REFERENCES

[1] M.R. Ahsan, M.I. Ibrahimy, O. O. Khalifa. Electromyography (EMG) Signal based Hand Gesture Recognition using Artificial Neural Network (ANN). *In Proceedings of the 2011 4th International Conference on Mechatronics (ICOM),* 17-19 May 2011, Kuala Lumpur, Malaysia.

[2] Blackmagic Design, *DaVinci Resolve with Fairlight Audio*, https://www.blackmagicdesign.com/products/davinciresolve/fairlight <accessed April 2017>.

[3] B. Bongers, Physical Interfaces in the Electronic Arts: Interaction Theory and Interfacing Techniques for Real- Time Performance. *In Trends in Gestural Control in Music, M. Wanderley and M. Battier*, IRCAM, 2000.

[4] D. Brown, N. Renney, A. Stark, C. Nash, and T. Mitchell, Leimu: gloveless music interaction using a wrist mounted leap motion. *In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME'16)*, Brisbane, Australia, 2016, pp. 300-304.

[5] F-S Chen, C-M Fu, C-L Huang, Hand gesture recognition using a real-time tracking method and hidden Markov models, *Image and Vision Computing*, 21 (2003) 745–758, Elsevier.

[6] J.-U. Chu, I. Moon, Y.-J. Lee, S.-K. Kim, and M.-S. Mun, A supervised feature-projection-based real-time EMG pat- tern recognition for multifunction myoelectric hand control. *IEEE/ASME Trans on Mechatronics*, vol. 12, no. 3, pp. 282–290, 2007.

[7] L. Dipietro, A. M. Sabatini and P. Dario, A Survey of Glove-Based Systems and Their Applications. *IEEE Trans On*

Systems, Man, And Cybernetics—Part C: Applications And Reviews, Vol. 38, No. 4, July 2008.

[8] Y. Fang, K. Wang, J. Cheng and H. Lu. A Real Time Hand Gesture Recognition Method, *In 2007 IEEE International Conference on Multimedia and Expo*. IEEE.

[9] N. Gillian, and J.A. Paradiso. The gesture recognition toolkit. *Journal of Machine Learning Research*, 15.1 (2014): 3483-87.

[10] J. Guna, J.Grega, P. Matevz, T. Saso and S. Jaka. An Analysis of the Precision and Reliability of the Leap Motion Sensor and Its Suitability for Static and Dynamic Tracking. *Sensors* 14, pp. 3702-3720. 2014.

[11] J. Han, N. Gold. Lessons Learned in Exploring the Leap Motion™ Sensor for Gesture-based Instrument Design. *In Proceedings of the International Conference on New Interfaces for Musical Expression,* London, UK, 2014.

[12] L. Hantrakul and K. Kaczmarek. Implementations of the Leap Motion in sound synthesis, effects modulation and assistive performance tools. *In Proceedings of the International Conference on New Interfaces for Musical Expression,* London, UK, 2014.

[13] C.L. Huang, W.Y. Huang, Sign language recognition using model- based tracking and a 3D Hopfield neural network, *Machine Vision and Applications* 10 (1998) 292–307.

[14] A. Ibarguren, I. Maurtua, B. Sierra, Layered architecture for real time sign recognition: Hand gesture and movement, *Engineering Applications of Artificial Intelligence*, Volume 23, Issue 7, October 2010, pp 1216-1228.

[15] A. R. Jensenius, Kinectofon: performing with shapes in planes. *In Proceedings of the international conference on new interfaces for musical expression*, Daejeon, Republic of Korea, 2013, pp. 196-197.

[16] E. Jessop. The Vocal Augmentation and Manipulation Prosthesis (VAMP): A Conducting-Based Gestural Controller for Vocal Performance. *In Proceedings of the International Conference on New Interfaces for Musical Expression,* Pittsburgh, USA, 2009.

[17] V.E. Kosmidou, L.J. Hadjileontiadis, S.M. Panas, Evaluation of surface EMG features for the recognition of American Sign Language gestures. *In Proceedings of the 28th IEEE EMBS Annual International Conference,* New York, 2006. IEEE.

[18] A. Kurakin, Z. Zhang, Z. Liu. A real time system for dynamic hand gesture recognition with a depth sensor. *In Proceedings of the 20th European Conference on Signal Processing (EUSIPCO)*, 2012. IEEE.

[19] Leap Motion Inc. *The Leap Motion*. https://www.leapmotion.com/product <accessed April 2017>.

[20] H. Lee, K. Kim, M.S. Park, J.H. Park and S.R. Oh. Verification of A Fast Training Algorithm for Multi-Channel sEMG Classification Systems to Decode Hand Configuration. *In Proceedings of the 2012 IEEE International Conference on Robotics and Automation,* Saint Paul, Minnesota, USA. IEEE.

[21] S. c S, S. W. Lee, A. Sastry, A. Daruwalla, and G. Weinberg, Crossole: a gestural interface for composition, improvisation and performance using kinect. *In Proceedings of the International Conference on New Interfaces for Musical Expression*, Ann Arbor, Michigan, 2012.

[22] O. Luzanin and M. Plancak, Hand gesture recognition using low-budget data glove and cluster-trained probabilistic neural network. *Assembly Automation*, Vol. 34 Issue: 1, pp.94-105.

[23] G. Marin, F. Dominio, P. Zanuttigh, Hand Gesture Recognition With Leap Motion and Kinect Devices. *In Proceedings of the 2014 IEEE International Conference on Image Processing*, IEEE.

[24] Microsoft Inc. *Kinect for Windows*. http://www.microsoft.com/en-us/kinectforwindows/ <accessed April 2017>.

[25] T. Mitchell and I. Heap. Soundgrasp: A gestural interface for the performance of live music. *In Proceeding of the International Conference on New Interfaces for Musical Expression,* Oslo, Norway, 2011.

[26] T. J. Mitchell, S. Madgwick, and I. Heap. Musical interaction with hand posture and orientation: A toolbox of gestural control mechanisms. *Proceedings of the International Conference on New Interfaces for Musical Expression,* Ann Arbor, Michigan, 2012.

[27] A. Mulder. Towards a choice of gestural constraints for instrumental performers. *Trends in Gestural Control of Music.* Paris, France: IRCAM, 2000.

[28] C. Dal Mutto, P. Zanuttigh, and G. M. Cortelazzo. Time-of-Flight Cameras and Microsoft Kinect, *Springer Briefs in Electrical and Computer Engineering*. Springer, 2012

[29] K. Nymoen, M.R. Haugen and A.R. Jensenius. *MuMYO* — Evaluating and Exploring the MYO Armband for Musical Interaction. *In Proceedings of Proceeding of the International Conference on New Interfaces for Musical Expression*, Baton Rouge, LA, 2015.

[30] A. Palkowski and G. Redlarski, Basic Hand Gestures Classification Based on Surface Electromyography. *Computational and Mathematical Methods in Medicine* Volume 2016.

[31] J. A. Paradiso, Electronic music: New ways to play. *IEEE Spectr*um, vol. 34, no. 12, pp. 18–30, Dec. 1997.

[32] A. Mulder, S. Fels, and K. Mase. Empty-handed gesture analysis in MAX/FTS. *In Proc. AIMI Int. Workshop Kansei—Technol. Emotion*, 1997, pp. 87–90.

[33] R. Polfreman. Multi-modal instrument: towards a platform for comparative controller evaluation. *In Proceedings of the International Computer Music Conference*, pp. 147-150, 2011, MIT Press.

[34] L. Quesada, G. López, and L.A. Guerrero L.A. Sign Language Recognition Using Leap Motion. *J. García-Chamizo, G. Fortino, S. Ochoa (eds) Ubiquitous Computing and Ambient Intelligence. Sensing, Processing, and Using Environmental Information*. Lecture Notes in Computer Science, vol 9454. Springer, Cham. 2015.

[35] C. Sapsanis, G. Georgoulas, A, Tzes, D. Lymberopoulos. Improving EMG based classification of basic hand movements using EMD. *In Proceedings of the 35th IEEE EMBC Annual International Conference*, Osaka, Japan, 2013. IEEE.

[36] S. Serafin, S. Trento, F. Grani, H. Perner-Wilson, S. Madgwick, T. Mitchell. Controlling Physically Based Virtual Musical Instruments Using The Gloves. *In Proceedings of the International Conference on New Interfaces for Musical Expression*, London, UK, 2014.

[37] Steinberg Media Technologies GmBH, iC Air for Leap Motion, https://www.steinberg.net/en/products/accessories/cubase_ic_air.html <accessed April 2017>.

[38] J. Taylor, L. Bordeaux, T. Cashman, B. Corish, C. Keskin, E. Soto, D. Sweeney, J. Valentin, B. Luff, A. Topalian, E. Wood, S. Khamis, P. Kohli, T. Sharp, S. Izadi, R. Banks, A. Fitzgibbon, J. Shotton. Efficient and Precise Interactive Hand Tracking through Joint, Continuous Optimization of Pose and Correspondences. *In Proceedings of ACM SIGGRAPH 2016*, Anaheim, CA, 2016, ACM.

[39] S. Trail, M. Dean, G. Odowichuk, T. F. Tavares, P. Driessen, A. W. Schloss, and G. Tzanetakis, Non-invasive sensing and gesture control for pitched percussion hyper-instruments using the kinect. *In Proceedings of the International Conference on New Interfaces for Musical Expression,* Ann Arbor, Michigan, 2012.

[40] F. Weichert, D. Bachmann, B. Rudak D. Fisseler, Analysis of the accuracy and robustness of the leap motion controller. *Sensors 13*, 6380–6393. 2013