IMU-sic: A Novel Gestural Controller for Live Vocal Processing

Removing the barriers to singers' creativity on stage

Sam Castle

Dyson School of Design Engineering, Imperial College London

DE4 Master's Project Final Report – Design, Build, Test

Impact Theme: AI and Data

Supervisor: Dr Lorenzo Picinali

Word Count: 5277

Written for consideration at the International Conference on New Interfaces for Musical Expression (NIME)

1 ABSTRACT

Since a paradigm shift in the late 1960s, live music today largely focusses on the recreation of meticulously sculpted studio sounds. Although many current artists' 'sound' is heavily defined by the processing applied to their vocals, most performers are content for their live vocal processing to be controlled at the mixing desk according to a prescheduled performance plan. Artists who demand more control, however, are constrained in their expression and movement by traditional musical interfaces.

This paper investigates how natural singing gestures might be used by singers to actuate and manipulate effects live on stage. A novel gestural MIDI controller was proposed, IMU-sic, and a proof-of-concept prototype was built (Figure 1). The device addresses the design limitations of current commercial effects controllers by providing full access to any combination of standard digital effects, through complex gesture recognition and hand orientation. In addition, this controller provides a scalable experience, and commercial sonic landscape, to engage all levels of singer.

A user study was piloted, demonstrating the core novel interaction as functional and compelling, both for applying effects presets, and for manipulating the parameters of those presets. The study also highlighted key development areas including integrating a more stable orientation algorithm and combatting the accidental recognition of unwanted gestures. Early user engagement with the device gives confidence that with future development, IMU-sic has the potential capability to garner widespread uptake within a famously conservative industry.

Keywords: Gesture Recognition, Digital Processing Effects, New Interface for Musical Expression, MIDI Controller, Interaction Design

N.B. All the code for this project can be found at: https://github.com/sam-castle/IMU.sic

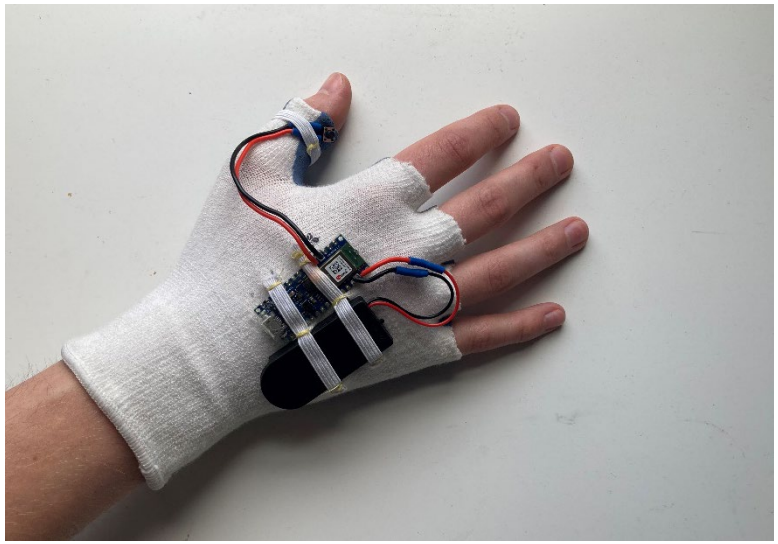

Figure 1 - IMU-sic gesture controller (proof-of-concept prototype)

CONTENTS

2 INTRODUCTION

2.1 LIVE MUSIC: AS IT STANDS

The experimental recording techniques used in the production of The Beatles' 1967 concept album 'Sergeant Pepper's Lonely Hearts Club Band' revolutionised studio recording forever. Whilst music production had previously been limited to recording live takes, in their bid to create "never-heard-before sounds" [1], the band and their sound engineer Geoff Emerick experimented with novel tape techniques and microphone arrangements. The resulting sounds heard on tracks such as 'Strawberry Fields Forever' and 'A Day in the Life', were so progressive that they heralded the arrival of a new era of music production [2]. Following the release of the album, there was a noticeable shift towards artists writing specifically for the studio instead of the stage. In turn, live performances began to attempt to recreate the sounds painstakingly sculpted in the recording studio [3].

In practice, this requires "cues" to be triggered, whereby special effects and audio presets are added to the performance. In addition, other elements can be included in the show, such as lighting and set changes. This system has become the foundation for the smooth running of the huge, mega-produced spectacles now regularly put on by the industry's biggest popstars (e.g., Figure 2). With the overall number of stage elements ever increasing, their accurate scheduling has become of paramount importance. As a consequence, backing tracks, rather than live bands, and the use of Timecode[1] are now prevalent within the industry. These techniques have led to the night-to-night performances of a tour being almost identical, and with it the loss of a considerable part of the 'magic' that unique and emotional live performances can give.

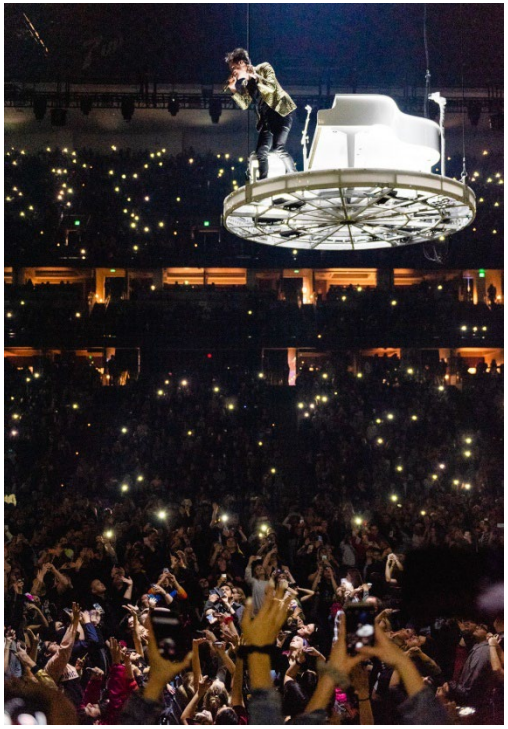

Figure 2 - Panic! at the Disco performing at the Honda Centre

[1] Timecode – a sequence of numeric codes generated at regular intervals by a timing synchronization system

2.2 Live Vocal Performance: As It Stands

The management of a performer's live vocal processing is treated somewhat similarly. Generally, predefined combinations of effects (e.g., delay and reverb) are saved as presets. These presets are used throughout the performance and are managed at the mixing desk by a sound engineer, according to the planned schedule of the set (Figure 3). If the singer would like to stray from the defined performance plan, say to request for more volume on their vocals, then they are forced to awkwardly gesture towards the control booth, or stop the performance to speak to their engineer.

The majority of artists tolerate this, however some demand more control over their vocal processing, to unlock more interesting and flexible sounds on stage. Currently, to achieve this versatility, an artist must use a laptop, mixing console or effects pedals on stage (Figure 4). The setup of these options acts to tether them to a single spot and peering into a laptop or bending to twist a knob significantly diminishes the artist's ability to express themselves creatively on stage.

The aim of this work is to explore how natural movements and gestures could be utilised to actuate effects presets, and both finely tune and experimentally manipulate those effects live on stage. Doing so in an accessible and intuitive way, the aim is to give singers the ability to physically shape the sonic profile of their voice live and remove the limitations to their creativity and expression.

Figure 3 - Live mixing desk

Figure 4 - Claudio performs at TEDxPerth

3 EXPLORATION

3.1 GESTURE CONTROLLED MUSICAL INTERFACES

Addressing the topic of gesture controlled musical interfaces, initial consultations were held with Camilla Tassi, award winning projection designer and music producer, and Bret Battey, multimedia composer and Professor of Audiovisual Composition. Following guidance from these experts, a thorough literature and product review was carried out.

Synthesising and manipulating sound with movement is not a new concept, and has been experimented with for over a hundred years. First however, the difference between a musical (MIDI[2]) controller and a musical instrument must be defined: a musical instrument is a device that produces musical sound [4]; a MIDI controller generates data, used to trigger sounds or control effects, like reverb or fuzz [5].

3.1.1 Historical Solutions

In 1921, Leon Theremin gave the first performance of a new musical instrument: 'The Theremin' (Figure 5). Played in thin air, the instrument uses the principles of capacitance and heterodyning to create sound from movement [6], and produces a tone similar to a cello. In 1929 it became the first commercially available electronic musical instrument [7].

There are several other documented pieces of the work in the field over the subsequent years, however, following the invention of MIDI in 1981 [8], there was an explosion of research into new musical interfaces. These included Bernard Szajner's 'Laser Harp' [9], Michel Waisvisz's 'The Hands' [10], Laetitia Sonami's 'Lady's Gloves' [11] and Tod Machover's conception of 'Hyperinstruments', a class of established instruments whose physical limitations are bypassed, and expressive capabilities extended, using electronic sensors [12]. These devices were mostly developed within an academic setting (although the 'Laser Harp' was used extensively in concert by Jean Michel Jarre [9]) and mostly operated within a challenging experimental soundscape with a focus on sound synthesis, rather than signal processing (Figures 6 - 9) .

Figure 5 - Leon Theremin playing his Thereminvox in 1932

Figure 6 - Jean Michel Jarre plays the Laser Harp in concert

Figure 8 - Laetitia Sonami performing with 'The Lady's Glove'

Figure 9 - Michel Waisvisz performing with 'The Hands'

Figure 7 - The Hypercello at the MIT Media Lab

[2] MIDI – musical instrument digital interface

3.1.2 State of the Art Solutions

More recently there has been an increase in the development of commercially viable products. There are several MIDI Controller Rings now on the market, 'Neova' by Enhancia [13], the Genki Wave [14], and 'Hot Hand 3' by Source Audio [15], all of which use IMU[3] data generated from hand movements to manipulate a MIDI input, e.g., from a synthesiser or within a DAW[4]. In addition, Auug's 'Motion Synth' gained some traction via a successful Kickstarter campaign [16] and investment on Shark Tank Australia [17]. This device converts a smartphone into a synth controller where notes are played on the touch screen and the sound is created and altered with motion [16].

In addition, there are several controllers specifically focussed on controlling vocal processing. OWOW's recently funded 'Wiggle Kit' (Figure 11) is an extremely accessible device which uses simple roll, pitch, and yaw rotations with a corresponding X, Y, and Z button-push to individually vary three selected effects [18].

However, the most significant and advanced controller on the market are Imogen Heap's 'Mi.Mu Gloves' (Figure 12). The gloves use an IMU to measure the orientation of the wrist, and flex sensors on the knuckles to measure the bend of the fingers, allowing both large sweeping gestures and specific hand shapes, like a fist, to be detected to control various live effects. The gloves also incorporate vibrotactile haptic feedback to increase the performers' accuracy when performing a gesture [19]. Championed by the Grammy Award winner in her own performances, the gloves have also been trialled by several notable artists, including Ariana Grande, proving the greatest mainstream penetration of any of the discussed devices. However, when comparing Grande's performances [20] with Heap's [21], it is clear that significant rehearsal with the gloves is required to deliver a compelling performance.

Figure 10 - Neova ring used with a MIDI keyboard

Figure 11 - OWOW's Wiggle Kit

Figure 12 - Imogen Heap performs with the Mi.Mu Gloves

[3] IMU – inertial measurement unit
[4] DAW – digital audio workstation, e.g., Logic

3.2 Aims of the Work

Through this literature review, it is obvious that there is still a considerable gap in the market for a device which allows singers to physically interact with their sound by giving direct access to their vocal processing whilst promoting a more emotional connection with their audience. Achieving this via a means that is not perceived as synthetic by the audience is paramount to creating a naturally expressive performance [22]. In addition, creating sound that is adequately commercial for a mainstream audience to appreciate is important for the widescale uptake of the device within the industry.

Therefore, this work aims to create a fully configurable and discrete device where programmed gestures can be used to precisely tune and experimentally manipulate any combination of vocal effects live on stage. In doing so, the objective is that performers can unlock a more expansive sonic landscape, allowing singers to express themselves beyond the limits of their own physical singing ability.

Furthermore, the device will aim to combat the interfacing issue found with many traditional musical instruments as well as many of the devices highlighted previously, whereby a significant skill level must be obtained initially in order to operate the instrument/controller effectively. For example, think how many years it takes for a beginner violinist to start making a pleasant and rewarding sound. Therefore, the device will aim to create an engaging musical experience from the outset, all the way through to a higher skill level. This can be achieved by giving a choice in the level of responsibility over the controlling of the effects, commensurate with the user's ability level.

3.3 Gesture Identification Study

In order to design a gesture controller appropriate for vocal performance, discovery into gestures and their implications, and specifically the types of gestures used by singers, was undertaken.

Gestures are movements which transmit additional information when verbally communicating [23] and are an intrinsic part of linguistics, used in every language and culture [24]. When speaking descriptively, they accompany around 90% of verbal expressions [25].

Singing, in every way, presents a heightened emotional expression of speech. As such, gestures, both conscious and subconscious, are even more prominent in vocal performance than speech. Davidson asserts that gestures used in song can be divided into five main categories based on Ekman and Friesen's gesture labels [22]. 'Adaptors' are gestures which express ones mood; 'Regulators' assist with co-ordination, perhaps with other performers; 'Emblems' are gestures with a direct verbal translation, but are not necessarily tied to the speech they accompany; 'Illustrators' also have a direct verbal translation, but are directly tied to the speech; and 'Displays' are employed purely to show-off to the audience [22] [26].

To gain a greater understanding of what singing gestures look like, how they relate to the lyrical and melodic content that they accompany, and what they might mean to the performer, a study was designed and carried out.

3.3.1 Participants

Three experienced singers (1M/2F, aged 19-24), each from a different genre of singing were selected to participate: the frontman of an indie rock band; an established, classically trained mezzo soprano; and a Musical Theatre student. All three participants were recruited from within the author's personal network.

3.3.2 Singing Video Analysis

Each singer submitted self-recorded performances of two contrasting songs, to an accompaniment or backing track. Therefore, six performances were analysed, equalling 14min 11sec of footage. The videos (Figures 13 - 15) were filmed in balanced lighting, high resolution, and against a light-coloured background. This allowed easy gesture identification and leaves the option

for more complex tracking techniques to be used in the future.

The lyrics (and translations where required) to accompany the tapes were researched, and for each line of the lyrics the singers' gestures were analysed. A description of each gesture, classified as per Davidson [22], was given and how it related to the lyrics, emotional sentiment or melodic movement was noted. In addition, each singer's general technique was recorded.

3.3.3 Semi-Structured Interviews

The results of the gesture analysis were corroborated with short semi-structured interviews with each participant. The questions centred around the identified gestures: whether they were conscious or subconscious; what their purpose or meaning was; and how they aided their performance, e.g., by helping convey emotion or physically supporting their singing technique.

The interviews were carried out in person and were recorded with the participant's signed consent. Key notes were taken from the interviews and added to the notes for each video performance.

Figure 13 - Pop/Rock vocalist

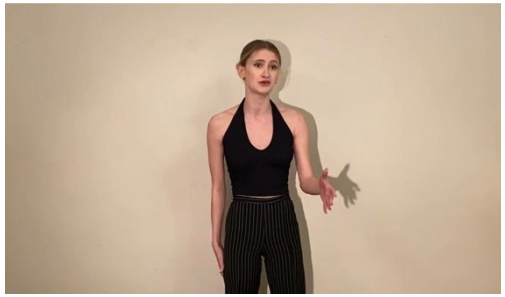

Figure 14 - Musical Theatre vocalist

Figure 15 - Classical/Operatic Vocalist

3.3.4 Taxonomy of Significant Gestures

Three main findings were evident from the study:

1. Gestures are incredibly personal, with each singer repeatedly using a small set of unique gestures;
2. Hand gestures are most prevalent, contributing 72.8% of the significant identified gestures;
3. Gestures are genre specific. In pop/rock hand gestures are used extremely frequently, 14.7/min, and are almost all subconscious. In Musical Theatre the hands are used less frequently, 6.7/min, and are generally thought-through or even choreographed to build a convincing character. Hand gestures are least prevalent in classical singing, 1.6/min, and seem mostly to respond spontaneously to the emotion of the song.

The analysis for each recording was reviewed and a selection of the most significant and meaningful gestures were summarised into a taxonomy of gestures (Figure 16). The full analysis for each performance can be found in Appendix i.

Gesture Description	Type	Spontaneous/ Choreographed	Significance
Reach up, hands clasped in 'begging' position	Adaptor	Spontaneous	Responding to the melody and lyrics – empathising the words
Raise hands and letting them drop	Adaptor	Spontaneous	Repeated natural gesture
Pointing to ground, self, and others	Illustrator/ Emblem	Both	Narrative gestures – enunciate key words and characters
Drawing the vocal line with finger	Adaptor	Spontaneous	Describing melody line / pitch visually
Marking out the rhythm of the line	Regulator	Mostly spontaneous	Moving to the music to keep the body invested. Also used to coordinate with other performers if music is slipping out of time (more calculated)
Hand tilt/twist	Adaptor	Spontaneous	Describing sliding melody line / pitch visually
Hand swipes	Adaptor	Spontaneous	Repeated natural gesture

Figure 16 - Taxonomy of significant gestures

4 DEVELOPMENT

4.1 IMU-SIC PROTOTYPE

Whilst effects have long been applied with the feet, as a result of their development for guitars, the Gesture Identification Study highlighted the hands as the source of most expressive gestures. Technically they also exhibit the most degrees of freedom and are least likely to effect good singing technique. Therefore, a proof-of-concept prototype which focusses on extracting hand gestures to control effects was developed.

4.1.1 Concept

A morphological analysis (Figure 17; Appendix ii) was used to explore the design possibilities and establish a concept to prototype. IMU sensing was chosen for gesture tracking, as this allows a standalone device to be developed and negates the need for tracking cameras around the performance space. This reduces cost and does not constrain the performer's movement to the cameras' fields of view.

The controllers' function can be defined by three main tasks:

1. The selection of effect presets: achieved by training the system to detect certain gesture profiles in the tri-axis accelerometer and gyroscope;
2. The manipulating chosen parameters of those effects: the raw IMU data is processed to calculate the orientation of the device, with the roll, pitch and yaw angles used to drive the parameters;
3. Applying effects to the audio input. This was undertaken in Max[5] for simplicity of prototyping digital sound processors, building presets, and routing audio.

Other elements of the morphological analysis, like the device's embodiment and feedback systems, are less important at this proof-of-concept stage, but certainly required for future development of the product.

Identify Body Movements	Select Effects	Control Effect Parameters	Create Effects
Infrared motion tracking systems e.g. Vicon, Optitrack	Gesture recognition, certain gesture applies to an effect	Potentiometers	Guitar pedal board, mechanically actuated
RGB colour tracking	Individual finger bent corresponds to an effect	Accelerometer/gyroscope thresholding	Custom circuitry
Pose tracking using machine learning e.g. Google's MediaPipe	Hand shapes correspond to different effects	Hand orientation tracking – roll, pitch yaw	Effects added with a DAW, e.g. Ableton
Wearable IMU sensor data	Buttons pressed correspond to effects	Pressure detection	Digital effects built in Max 8
Bend sensors used to calculate joint angles	Body part in certain 'zone' corresponds to selecting effect	Bend of fingers (or other joints) corresponds to level of an effect	Acoustic (mechanical) effects
Built-in sensors in smartphone accessed	Swipe through gesture presets sequentially	Smart textiles – stretch	
Microsoft Kinect	Proximity sensors	Incremental increase/decrease with swipe gestures	
Leap Motion hand tracking	Hall effect effects and magnet	Ultrasonic/distance measurement	
Smart textiles – measuring stretch	Capacitive touch	Touchscreen display	
		Pneumatic or hydraulic squeeze detection – reservoir in palm	

Figure 17 - Preview of morphological analysis - for full analysis see Appendix ii

[5] Max – a visual programming language for music and multimedia

4.1.2 IMU-sic Specification

With the product's primary functions defined, a full product specification was formulated (Figure 18; Appendix iii). The specification defines IMU-sic's critical criteria, providing the basis for effect user testing, imperative for proving clear success metrics.

Category	Specification Point
1) Selecting Effect Presets	1.1) Must be able to control a number of standard effects
	1.2) Must be able to create personal presets
	1.3) Must be able to choose personal selection gestures
	1.4) Must be able to choose which gesture triggers which preset
	1.5) Must be able to route effects in series and parallel
2) Controlling Effect Parameters	2.1) Must be able to choose which effect parameters are controlled live
	2.2) Must be able to precisely tune parameters
	2.3) Must be able to experimentally manipulate parameters
	2.4) Must be able to manipulate several effects at once
3) Impact on Performance	3.1) Device / use of device must not be distracting to the audience
	3.2) Device / use of device must not be distracting to the performer
4) General Usability	4.1) Must be operable with one hand
	4.2) Must not inhibit instruments to be played and handled
	4.3) Must be wireless
	4.4) Device's function must be scalable
	4.5) Must be comfortable to use
	4.6) Must be robust and durable
5) Commercial Viability	5.1) Must create a commercially viable sound
	5.2) Must be compact
	5.3) Must be lightweight
	5.4) Must be low cost (potential for)

Figure 18 - IMU-sic Specification – For full, justified specification see Appendix i

4.1.3 Initial Test

An initial test was carried out with the GyrOSC app (Figure 19). Raw IMU data, pulled from the smartphone's sensors, was sent to Max to control some simple effects. This confirmed that: a) raw IMU data is unsuitable for controlling effects, since the data describes rates of change and therefore exhibit spikes of data with motion rather than stable values; and b) whilst the availability of smartphones makes an app a more accessible product, they are large, cumbersome and indiscreet.

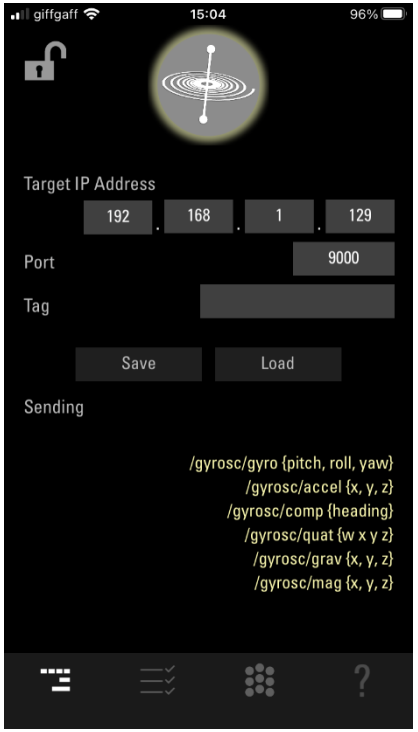

Figure 19 – the GyrOSC app was used to test raw IMU data

4.1.4 System Architecture

The IMU-sic system (Figure 20) consists of a peripheral device which recognises gestures and calculates hand orientation. This information is sent to a paired central device via BLE[6] which transfers the data to Processing[7] via serial communication. The data is converted to OSC[8] messages and sent to Max, where it is used to process the microphone input.

4.1.5 Hardware

The Arduino Nano 33 BLE Sense was chosen as the basis for the peripheral controller. The board is built around the nRF52840 microcontroller, runs on Arm Mbed OS, and was chosen for its tiny form factor and the host of inbuilt components, in particular the LSM9DS1 9-axis IMU for motion and orientation detection, and the u-blox NINA-B306 which allows for Bluetooth Low Energy connection. In addition, the compatibility with TensorFlow Lite Micro allows machine learning models to be loaded onto the board itself. This allows the device to act independently of the Internet, minimises form factor, and maximises battery life.

Further, a tactile switch drives a digital pin input, and the board is powered by two CR2032 coin cells, housed in series in a battery holder with an integrated on/off switch. A 30-minute, high load power test was performed, using an Innovateking-EU A3 USB Multimeter with a current accuracy rated at ±0.8% [27]. With the discharge rate measured at 22.2mA, the 210mAh cells can power the device continuously for 18.9hrs.

The assembly is mounted onto the back of a fingerless glove with the switch attached to the inside of the thumb, leaving the tip free to play instruments. The central device is also an Arduino Nano 33 BLE Sense, and is wired directly into the computer's USB port.

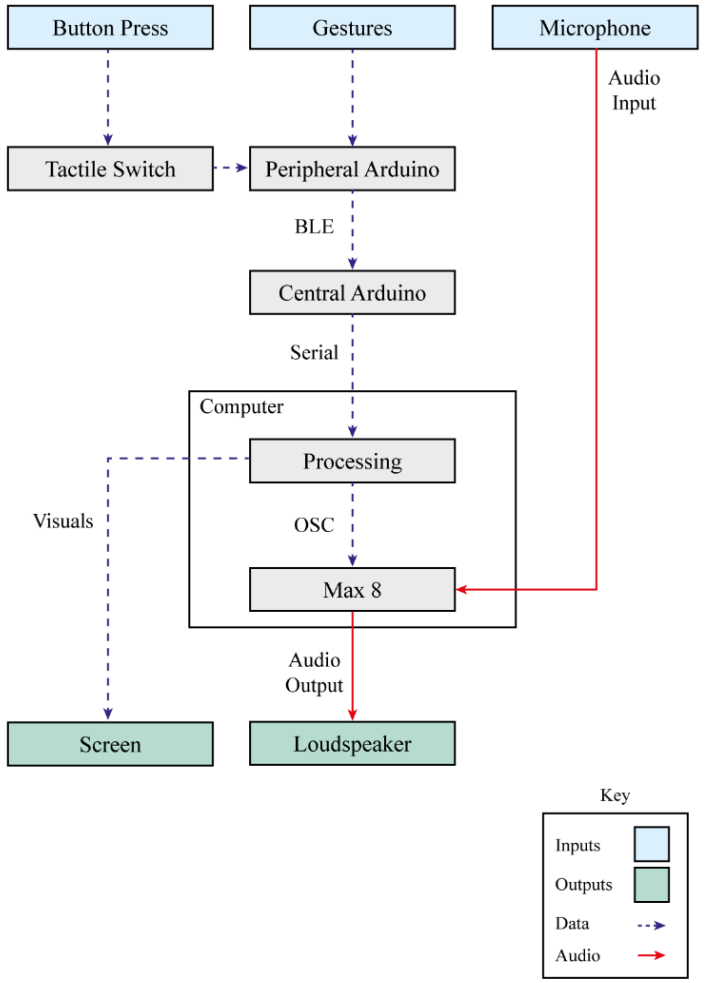

Figure 20 – IMU-sic system diagram

[6] BLE – Bluetooth Low Energy
[7] Processing – Java based integrated development environment for electronic art
[8] OSC – Open Sound Control

4.1.6 Recognising Gestures with Arduino

An Arduino sketch creates usable gesture data from the raw IMU data with an orientation algorithm and a gesture recognition machine learning model (Figure 21). A second sketch, for the central device, receives this data via BLE and writes it to the USB serial port (Figure 22).

Hand Orientation

The Madgwick IMU algorithm was used to calculate the orientation of the device by combining the data from the tri-axis accelerometer and gyroscope sensors. This optimised gradient-descent algorithm is computationally inexpensive, effective at low sampling rates, and more accurate than similar Kalman-based algorithms [28]. The algorithm was accessed with the 'MadgwickAHRS.h' library.

The hand's orientation is constantly calculated, keeping the device globally calibrated – however, the orientation data is only sent when the tactile switch is depressed. The user therefore must press and hold the button whilst orientating their hand. When a particular sound has been found, the button is released, locking the parameter values. Since the 3D axis remains fixed, theoretically, certain sounds possess fixed hand orientations within the workspace, allowing them to be repeatedly found and altered.

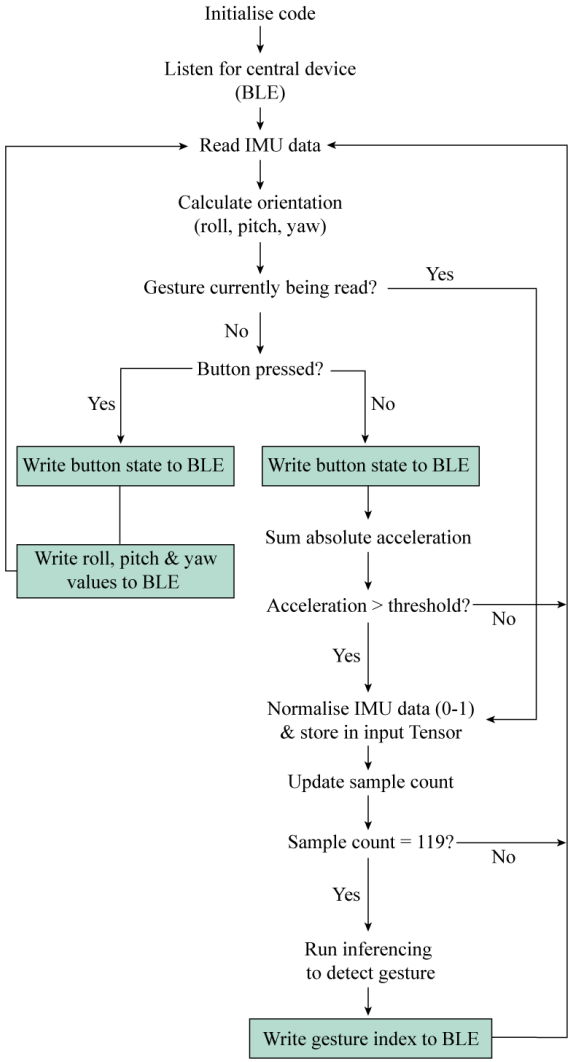

Figure 21 - Peripheral device Arduino sketch

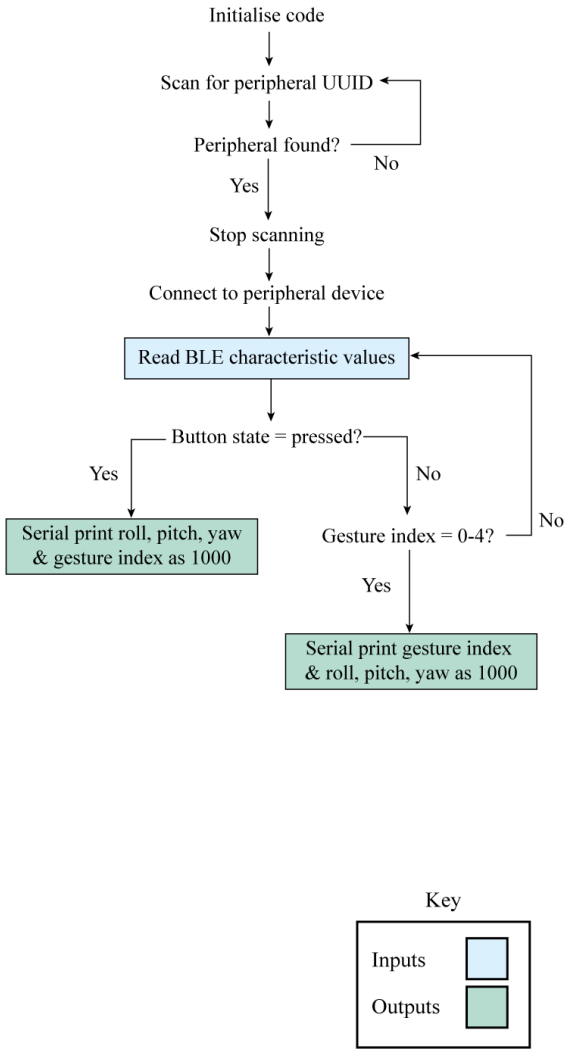

Figure 22 - Central device Arduino sketch

Gesture Training and Identification

When a gesture is performed, a pattern of sensor values in each of the IMU's 6 axes can be observed. Combining all 6 axes creates an IMU profile which corresponds to the gesture and is capable of being recognised, provided that it is distinct enough. Five gesture profiles were collected: Double Tap, Swipe Right, Swipe Left, Circle, and Twist (Figure 25). The gestures, recorded for 1 second after significant motion was detected, were each performed 20 times. The resulting datasets were used to train the model.

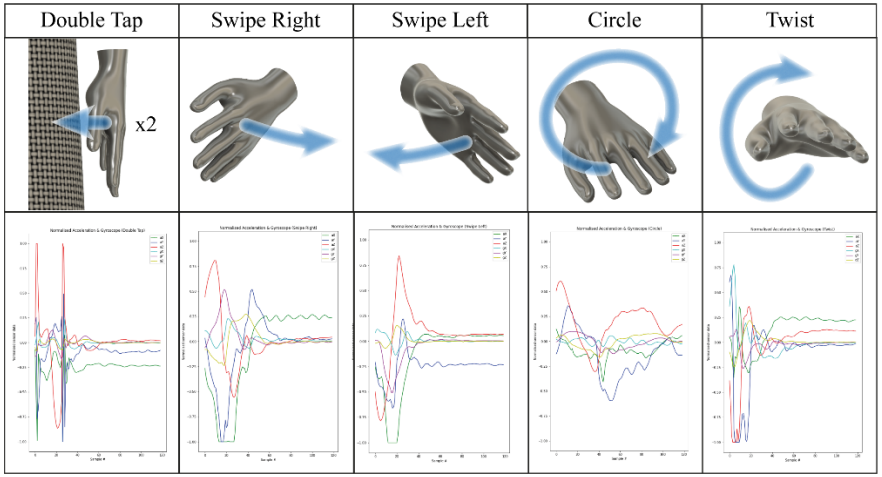

Figure 25 - Chosen gestures and normalised IMU profiles, inspired by the Gesture Identification Study

After being normalised between 0 and 1, the data was randomly split into training (60%), validation (20%), and test (20%) sets. These were used to build and train a TensorFlow model using the Keras Sequential model API. After 600 epochs, the trained model was extremely accurate (training: loss<3e-11, MAE[9]<3e-06; validation: loss<5e-08, MAE<4e-05), as is shown in Figure 23 & 24 In addition, inputting the test data and plotting the predicted values against the actual values showed extremely high accuracy (Figure 26).

The model was formatted for Arduino, so that when significant motion is detected (acceleration>2.8G), the program attempts to recognise a gesture. If a gesture is recognised at >90% accuracy, the corresponding gesture index is sent out.

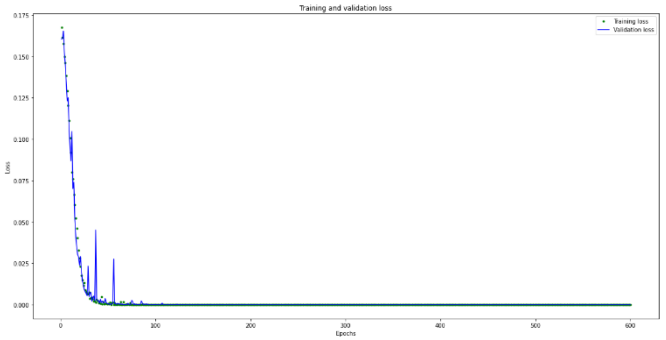

Figure 24 - Training and Validation loss

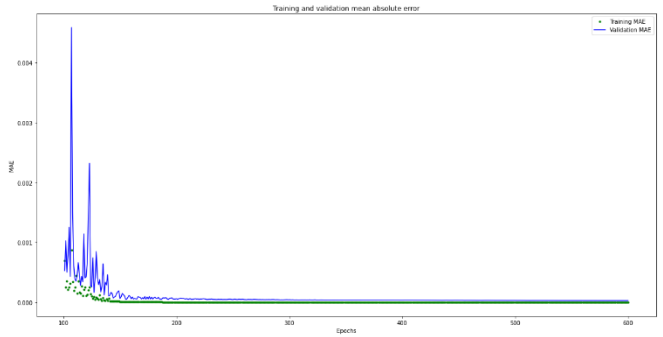

Figure 23 - Training and Validation mean absolute error

Figure 26 - Training data, predicted vs actual values

[9] MAE – Mean Absolute Error

4.1.7 Processing

The serial driver used by Mbed OS is incompatible with Max. Therefore, the serial data is read in Processing and sent as OSC messages to Max (Figure 27).

Although the built-in LEDs on the peripheral device indicate the two gestural functions, these LEDs are out of the direct eyeline of the user when using the device. Therefore, Processing was used to give visual feedback to the user, drawing the effect preset names and a simple orientation visual (Figure 28). Processing was designed for animations and generative art, so there is great potential for its usage in synchronising visuals with the audio effects.

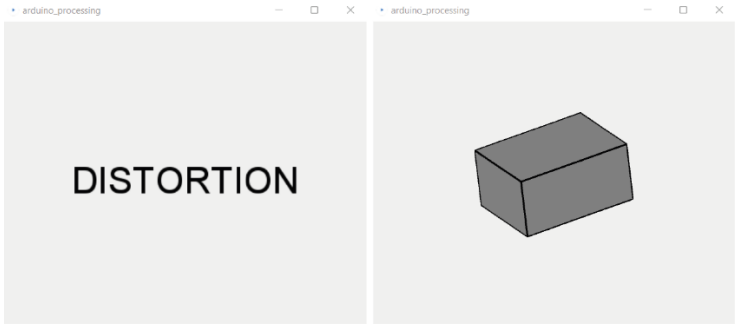

Figure 28 - Processing provides visual feedback

4.1.8 Creating Audio Effects in Max

The audio processing was handled in a custom Max Patch (Figure 29). Six conventional effects were chosen and built in Max: Volume, Pitch Shift, Flanger, Distortion, Delay, and Reverb. Effect combinations are built much like a guitar pedal board, with the choice to link them in either series or parallel, and the ability to mix the levels of the effects. The routing pathways are saved as presets, alongside initial values for effect parameters, e.g., Distortion Sample Rate. The parameters are altered live with the orientation of the hand: pitch, roll and yaw rotations correspond to different parameters for different presets.

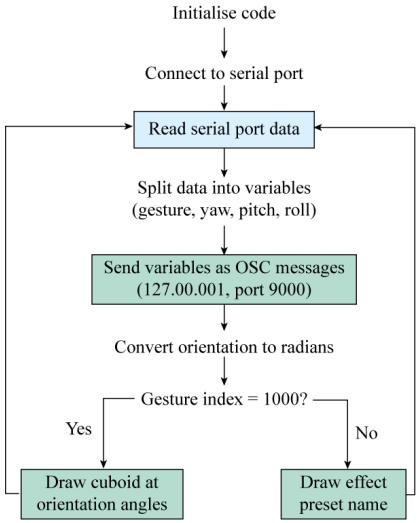

Figure 27 - Processing script overview

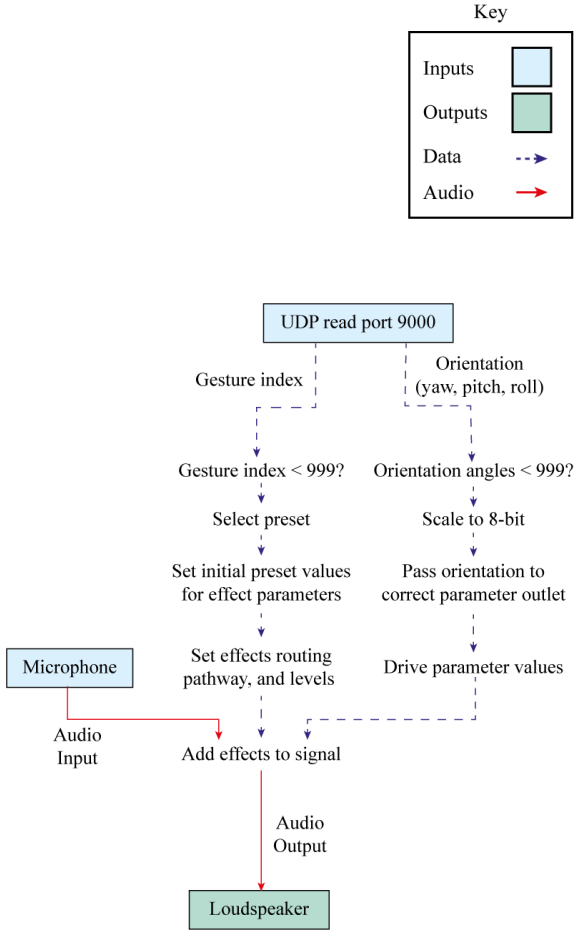

Figure 29 - Max patch overview

5 EVALUATION

5.1 PRODUCT USER TESTING

A mixed-methods user testing strategy was used to test IMU-sic's functionality with respect to the product specification.

5.1.1 Participants

The participants (n=10, 7M/3F, aged 19-24) were recruited from the London music scene, with the majority of participants attending the Royal College of Music and Rose Bruford College. The singers were selected from backgrounds that included professional pop session work, contemporary a cappella, and indie rock. Participants were not renumerated.

5.1.2 Procedure

Moderated, in-person testing sessions were administered to evaluate the product's general functionality. The participants were instructed to experiment with two user interfaces by applying all five effect presets and manipulate the relevant control parameters in a simulated vocal performance. The first interface was computer based, constituting the Max patch in 'Presentation Mode' (Figure 30) and was used as a control. The second interface was the IMU-sic controller (Figure 31). Both interfaces provided the exact same functionality, in order to investigate the success of the gestural interaction, and were introduced to the participants in a random order to remove bias.

Following a short demonstration, each interface was given to the participant to experiment with for at least 15 minutes. During the test, the users' behaviour was observed and any queries were answered. In addition, a series of unstructured questions allowed users to point out deficiencies and suggest improvements.

The participants evaluated their experience after using each system via a digital questionnaire. The questions were designed to directly address the product's specification, with some questions worded negatively to reduce bias. The users self-reported their responses on 5-point Likert scales.

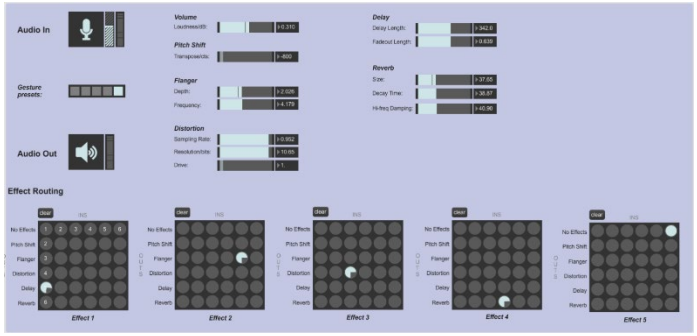

Figure 30 - Computer interface (control)

Figure 31 - IMU-sic controller

5.1.3 Results

The results of the questionnaire (Appendix iv), scored 1 (worst) to 5 (best), and the testing notes (Appendix v) are analysed in this section. Unpaired t-tests were used to establish significant differences between the control interface and the IMU-sic controller.

The participants were shown to be a highly receptive audience, with the two core principals of the device strongly validated in the results. It was strongly believed that the ability to control their vocal effects live would elevate their performance (Q1: μ=4.6, std=0.97), and 100% of participants strongly agreed that freedom of movement is important to feeling natural and confident on stage (Q2: μ=5, std=0).

Selecting Effects Presets

The users were able to effectively apply effect presets using gestures (Q3B: μ=4.3, std=0.67), and the gesture recognition was generally accurate (Q4B: μ=3.8, std=0.42), although there was some confusion between 'Swipe Right' and 'Circle' gestures {U2, U4, U6, U10}. The 'Double Tap' gesture, to clear effects, was occasionally accidently triggered when the hand was dropped too aggressively {U3, U6, U8, U9, U10} and there was some concern that dancing to the music {U5}, or playing an instrument {U7, U8}, may inadvertently actuate effects.

The visual cues for loaded presets were very effective (Q3B: μ=4.8, std=0.42), but displaying the gesture type, as well as the preset name, could aid with selection {U6}.

Choosing personalised effect presets and training one's own gestures to trigger them was strongly desired (Q6A: μ=4.9, std=0.32; Q6B: μ=4.9, std=0.32). Personalised gestures would likely improve the accuracy of the gesture selection {U2, U4} and using fewer effects would make the device simpler to operate {U2, U6, U9}.

Controlling Effects Parameters

Users felt more able to creatively manipulate their sound (Q8B: μ=4.8, std=0.42), than precisely control the effects parameters (Q7B: μ=3.4, std=0.52). Precise control was significantly (p=0.005) more achievable using the computer interface (Q7A: μ=4.4, std=0.84) than the IMU-sic controller.

Whilst it was obvious when effects parameters were being controlled (Q9B: μ=4.7, std=0.67) it was less obvious which parameters those were (Q10B: μ=3.6, std=1.07), with some participants sometimes struggling to determine which direction to rotate the hand to control certain parameters {U4, U10}. This difficulty was not seen in the control data (Q10A: μ=4.3, std=0.48). Roll rotations were more effective than pitch and yaw rotations {U4, U10}.

The ability to choose which parameters to control live as well as which hand rotations corresponded to them was strongly desired by the participants (Q11A: μ=4.5, std=0.85; Q11B: μ=4.5, std=1.08).

Impact on Performance

Performing gestures did not distract the singers significantly from their singing, (Q12B: μ=3.8, std=0.79) and all participants were confident that with practice they would become fluent at performing gestures whilst singing (Q13B: μ=4.9, std=0.32). That said, one participant found the device off-putting as it required 'too much coordination' {U2}. Nevertheless, controlling the computer interface was significantly (p=0.0023) more distracting to their performances (Q12A: μ=2.4, std=0.97).

General Usability

The participants were able to use the device easily whilst holding a microphone in their other hand (Q14B: μ=4.5, std=0.71) but were less confident about holding or playing an instrument (Q15B: μ=3.6, std=0.84). For example, an open palm is required for guitar playing to allow the strings to be damped {U7, U9}.

The IMU-sic controller allowed the singers to maintain their freedom of movement (Q16B: μ=4.2, std=0.63), and the movements required to operate it felt mostly natural (Q17B: μ=3.8, std=0.42). The computer interface inhibited movement significantly more than the controller did (Q16A: μ=2.1, std=0.99; p<0.0001), and was significantly less comfortable to use (Q19A: μ=3.1, std=1.37; Q19B: μ=4.7, std=0.67; p=0.0038).

Commercial Viability

All the participants were able to create a sound that a) they found artistically interesting, and b) was commercially viable (Q20B: μ=4.9, std=0.32; Q21B: μ=4.7, std=0.48). 90% of the participants would recommend IMU-sic to other performers looking for new ways to express themselves on stage (Q22B: μ=4.7, std=0.67). However, only 50% would recommend the computer interface (Q22A: μ=3.5, std=0.85; p=0.0025).

5.2 DISCUSSION

Overall, the testing proved substantially that the proposed gestural interaction is appropriate and dynamic for actuating and manipulating effects. With all the testable specification points met to some extent, and the IMU-sic controller outperforming the control in key areas, the concept was successfully validated and informs the scope for future development of the product. Notably, with all users capable of creating artistically interesting sounds whilst remaining sonically within a commercial space, uptake of the device within the industry appears plausible, with neither the experience of musician nor that of the audience being compromised.

5.2.1 Design Limitations

The testing process highlighted two significant functional shortcomings of the IMU-sic prototype. First was the robustness of the gesture recognition system. Some gestures exhibited similar profiles making it difficult for the model to differentiate between them. This caused some unwanted effects to be applied. In addition, any significant motion can be inadvertently interpreted by the system as a gesture attempt. Despite a high confidence threshold, the system interprets some movements as gestures e.g., dropping the hand as the 'Double Tap' gesture. Whilst a higher acceleration threshold could be implemented to ensure only purposeful gestures are recognised, this would promote aggressive, unnatural gestures. Instead, another switch could be added, which when depressed would activate the gesture recognition.

The second concern was the precision of the orientation algorithm. The device showed significant sensor drift, particularly in the yaw axis, causing global orientations to slowly move over time. The IMU algorithm implemented is an outdated version and simplified to use only accelerometer and gyroscope sensors. The most recent Madgwick AHRS[10] Fusion algorithm should be employed instead. This version incorporates magnetometer data to compensate for magnetic distortion and gyroscope bias drift [29].

5.2.2 Study Limitations

The sample size and range of demographics were too small for real statistical conclusions to be drawn. In addition, the time spent with the device (approximately 15 mins) was not sufficient to highlight all the possible implications of highly pressurised performance situations or long-term usage. Further, whilst the building of personalised presets and training of gestures were discussed, the users were not able to autonomously trial these features, due to time constraints. This restricts the analysis of the device's scalability to a theoretical level. Finally, despite several precautions built into the method, all participants were known to the author to some degree, which is likely to establish a level of positive bias.

However, overall, the study was successful in establishing significant qualitative design insights and served as a pilot for a more in-depth longitudinal investigation. By developing a higher fidelity product with several more interaction possibilities, and evaluating them comparatively over a longer usage period with a larger sample size and a greater range of targeted demographics, an effective empirical evaluation of IMU-sic's function and viability can be achieved.

5.3 FUTURE WORK

5.3.1 Design Development

There are several key areas for future development:

Firstly, the design embodiment was of secondary importance at this proof-of-concept stage, with several decisions made purely for ease of prototyping. Now, with the function of the device proven, an Industrial Design approach should be taken to optimise the touchpoint design. Similarly, a UX[11] design approach should be taken to develop a commercial control user interface to ensure creating presets and routing effects is simple and intuitive.

Secondly, whilst using BLE was useful for energy efficiency, there are limits to the frequency of data output, and device range. The latter is an issue for larger venues, and as such development should prioritise Wi-Fi data transmission instead.

Thirdly, the visual feedback of the device was only partially successful, with reliance on the display undesirable in maintaining the performer's freedom of movement. Therefore, development should investigate what and how information is displayed. Further, the incorporation of vibrotactile haptic feedback could go some way to reducing the dependence on screen displays.

Finally, minimal latency is a critical performance metric in audiovisual systems. Whilst a level of latency was expected and tolerated at this prototyping stage, future

[10] AHRS – Attitude and Heading Reference System

[11] UX – User experience

development must focus on improving the system's efficiency, both in hardware and software, to reduce latency to an imperceptible level. A more efficient system also draws less power, making battery life longer.

5.3.2 Gestures for Incremental Changes

Minimally investigated during development, gestures could instead trigger changes incrementally, e.g., increasing or decreasing parameter values in small amounts, or presets scrolled through chronologically. Whilst less powerful, optimising the device to function in this way could provide a simpler, more robust interaction. Catering to a wider, less skilled market, functioning this way would promote precision control, over experimental manipulation.

6 CONCLUSION

This work investigates how natural singing gestures might be used by singers to actuate and manipulate effects live on stage. In doing so, a novel gestural controller was proposed, built, and tested. The IMU-sic controller addresses the design issues and limitations of current commercial effects controllers by providing full access to any combination of standard digital effects, discreetly, and using only gestures. By sculpting the quality of their voice physically through hand gestures, singers are untethered from traditional, static effects interfaces, instead enabling and promoting natural, expressive, and unique performances. In addition, this controller tackles the shortcomings of other gesture-based controllers by providing a scalable experience and commercial sonic landscape, to engage all levels of performers, from beginner and recreational users through to those who are pushing the boundaries of modern music.

A preliminary user testing study demonstrated the core novel interaction as both functional and compelling, in both the application of effects presets and the manipulation of the parameters of these presets. In addition, strong qualitative insights highlighted the key areas of development and refinement required, namely integrating a more advanced and stable orientation algorithm, and combatting the accidental recognition of unwanted gestures. With the computational process and function of the device proven, a holistic design approach to developing the least invasive, and most useful touchpoint, and user interface for IMU-sic has the potential to transform this prototype into a compelling product. Early user engagement with the device gives confidence that with this future development, IMU-sic has the potential capability to garner widespread uptake within a conservative industry.

7 ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr Lorenzo Picinali, for all his guidance and support.

8 REFERENCES

[1] G. Emerick, Interviewee, *Sound engineer Geoff Emerick remembers recording The Beatles Sgt Pepper's album.* [Interview]. 25 May 2017.

[2] K. Murnane, "'Sgt. Pepper's' Was A Perfect Storm Of Musical And Recording Creativity," Forbes, 2017.

[3] B. A. P. Bloomberg, "Making Musical Magic Live: Inventing modern production technology for human-centric music performance," Massachusetts Institute of Technology, Massachusetts, 2020.

[4] J. A. Westrup and T. C. Grame, "musical instrument," Encyclopedia Britannica, 27 November 2020. [Online]. Available: https://www.britannica.com/art/musical-instrument. [Accessed 27 November 2021].

[5] Music Repo, "Music Technology Glossary of Terms," 16 November 2020. [Online]. Available: https://www.musicrepo.com/music-technology-glossary-of-terms/. [Accessed 27 November 2021].

[6] M. Kennedy and J. B. Kennedy, "Theremin," in *The Concise Oxford Dictionary of Music*, Oxford University Press, 2007.

[7] J. E. Oliver La Rosa, "Theremin in the Press: Instrument remediation and code-instrument transduction," *Organised Sound,* vol. 23, no. 3, pp. 256-269, 2018.

[8] D. Smith and C. Wood, "The 'USI', or Universal Synthesizer Interface," Audio Engineering Society, San Jose, 1981.

[9] laser-harp.com, "What is a Laser Harp," 22 August 2019. [Online]. Available: https://laser-harp.com/what-is-a-laser-harp/. [Accessed 28 November 2021].

[10] G. Torre, K. Andersen and F. Baldé, "The Hands: the making of a digital musical instrument," *Computer Music Journal,* vol. 40, no. 2, pp. 22-34, 2016.

[11] L. Sonami, "Lady's Glove," 2021. [Online]. Available: https://sonami.net/portfolio/items/ladys-glove/. [Accessed 29 November 2021].

[12] T. Machover and J. Chung, "Hyperinstruments: Musically Intelligent and Interactive Performance and Creativity Systems," *International Computer Music Association,* vol. 1989, pp. 186-190, 1989.

[13] Enhancia, "Neova," Enhancia, 2021. [Online]. Available: https://www.enhancia-music.com/product/neova/. [Accessed 28 November 2021].

[14] Genki, "Wave," Genki, 2021. [Online]. Available: https://genkiinstruments.com/products/wave. [Accessed 28 November 2021].

[15] Source Audio, "Hot Hand 3 Universal Wireless Effects Controller," Source Audio, 2015. [Online]. Available: https://www.sourceaudio.net/hot-hand-3.html. [Accessed 28 November 2021].

[16] AUUG, "The Motion Synth: Turn Movement into Music," Kickstarter, April 2014. [Online]. Available: https://www.kickstarter.com/projects/1892750571/the-motion-synth-turn-movement-into-music. [Accessed 28 May 2022].

[17] Shark Tank Australia, "Sharks Blown Away By Revolutionary Mobile App," Youtube, 25 Decemeber 2019. [Online]. Available: https://www.youtube.com/watch?v=yjyhXKOq0Jg&ab_channel=SharkTankAustralia. [Accessed 28 May 2022].

[18] OWOW, "Wiggle Kit - A New Instrument For Singers," Kickstarter, 16 February 2018. [Online]. Available:

https://www.kickstarter.com/projects/owow/wiggle-kit-a-new-instrument-for-singers. [Accessed 28 November 2021].

[19] MiMU, “Mimu Gloves,” 2021. [Online]. Available: https://mimugloves.com/gloves/. [Accessed 28 November 2021].

[20] C. Rand, “Ariana Grande (Mimu Gloves),” Youtube, 2 March 2015. [Online]. Available: https://www.youtube.com/watch?v=VcdvBGEEXcU. [Accessed 29 November 2021].

[21] NPR Music, “Imogen Heap: NPR Music Tiny Desk Concert,” Youtube, 28 June 2019. [Online]. Available: https://www.youtube.com/watch?v=3QtklTXbKUQ. [Accessed 28 November 2021].

[22] J. W. Davidson, “The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox,” *Musicae Scientiae,* vol. 5, no. 2, pp. 235-256, 2001.

[23] D. McNeill, “So You Think Gestures Are Nonverbal?,” *Psychological Review,* vol. 92, no. 3, pp. 350-371, 1985.

[24] P. Feyereisen and J.-D. de Lannoy, “Gestures and speech: Psychological investigations,” Cambridge University Press, New York, 1991.

[25] D. McNeill, “Gesture and Thought,” University of Chicago Press, Chicago, 2005.

[26] P. Ekman and W. V. Friesen, “The Repertoire Of Nonverbal Behavior Categories, Origins, Usage, and Coding,” *Semiotica,* vol. 1, no. 1, pp. 49-98, 1969.

[27] Innovateking-EU, “USB Color Tester Instructions - Model: A3/A3-B,” [Online]. Available: https://www.amazon.co.uk/Voltmeter-Bluetooth-Multimeter-Multi-Function-Detector/dp/B07DCS11GM/ref=cm_cr_arp_d_product_top?ie=UTF8&th=1. [Accessed June 2022].

[28] S. O. Madgwick, “An efficient orientation filter for inertial and inertial/magnetic sensor arrays,” IEEE, Bristol, 2010.

[29] xioTechnologies, “Fusion - AHRS Algorithm,” 30 May 2022. [Online]. Available: https://github.com/xioTechnologies/Fusion. [Accessed June 2022].

9 APPENDIX

Appendix i

Gesture Identification Study Analysis:

Pop/Rock Vocalist – Wild Blue (John Mayer)

General vocal technique:

- Feet shoulder width apart
- Weighted on right leg – tries not to do this
- Natural, wide shoulders
- Head tilted upwards slightly
- Left hand in trouser pockets for half of the performance

Labelled expressive gestures according to the lyrics of 'Wild Blue' by John Mayer:

Lyrics	Gesture descriptor
Verse 1	
Never seen the sun before	*Adaptor* – Hands cupping face
Lyin' on the ocean floor	*Adaptor* – Hands in 'praying' position
I'm walkin' through the wilderness	*Adaptor* – Shaking head in time
And livin' off the loneliness	
Chorus	
Oh wild blue, deeper than I ever knew	*Adaptor* – Gentle pointing with both hands – fingertips together *Illustrator* – Point to self on 'I' *Adaptor* – Outline stepping down melodic line with right index
Wild blue on a bed of grey	*Regulator* – Mark opening word of the line *Adaptor* – All fingertips held together

Oh baby, what a wild blue	Adaptor – Point with right index and draw the vocal line
I found myself when I lost you	*Illustrator* – Point to self on 'I' *Illustrator* – Point to audience on 'you'
SILENCE AND STILLNESS	*Adaptor* – Nodding head and tapping foot to instrumental break
Verse 2	
All the tears I meant to cry	*Regulator* – Right hand, with index and thumb touching, marks out the rhythm of the line
Dance across the evening sky	*Adaptor* – Swipes right hand, with index and thumb touching, across the body
And in my sorrow I can see	*Regulator* – Right hand, with index and thumb touching, marks out the rhythm of the line
That missin' you belongs to me	*Illustrator* – Points out on 'you' *Emblem* – Point down to floor on 'me'
Chorus	
Oh wild blue, deeper than I ever knew	*Adaptor* – Rock head back at top of melodic line *Illustrator* – Point to self on 'I'
Wild blue on a bed of grey	*Regulator* – Right hand, fingers outstretched, marks out the rhythm of the line
Oh baby, what a wild blue	*Regulator* – Right hand, fingers outstretched, marks out the rhythm of the line
I found myself when I lost you	*Regulator* – Right hand, fingers outstretched, marks out the rhythm of the line
SILENCE AND STILLNESS	*Adaptor* – Nodding head and tapping foot to guitar solo
Chorus	
Wild blue, deeper than I ever knew	*Adaptor* – Right hand wave with fingers outstretched
It's a wild blue since you went away	*Adaptor* – Point with right index and draw the vocal line

Oh baby, what a wild blue	*Regulator* – Right hand, fingers outstretched, marks out the rhythm of the line
I found myself when I lost you	*Illustrator* – Point to self on 'I'
SILENCE AND STILLNESS	
Outro And you'll never know	*Adaptor* – Point with right index and draw the vocal line
The unlikely beauty in letting you go	*Regulator* – Right hand, fingers outstretched, marks out the rhythm of the line

Pop/Rock Vocalist – Lovesick Blues (Hank Williams)

General vocal technique:

- Feet shoulder width apart
- Weighted on right leg
- Natural, wide shoulders
- Head tilted upwards slightly
- Both hands used to gesture for majority of the performance

Labelled expressive gestures according to the lyrics of 'Lovesick Blues' by Hank Williams:

Lyrics	Gesture descriptor
Chorus 1 I got a feelin' called the blues, oh Lord	*Adaptor* – Both thumbs together, fingers splayed *Adaptor* – Beat hands on 'Lord'
Since my baby said goodbye	
Lord, I don't know what I'll do	*Illustrator* – Point to self, with both hands, on 'I'

All I do is sit and sigh, oh Lord	*Adaptor* – All fingertips together, fingers splayed
That last long day, she said goodbye	*Adaptor* – Upper body lean backwards
Well Lord, I thought I would cry	*Adaptor* – Hands split, hanging by chest, fingers splayed
She'll do me, she'll do you	*Adaptor* – Beat hands on 'me' and 'you'
She's got that kind of lovin'	*Adaptor* – Left index and right thumb touch
Lord, I love to hear her when she calls me sweet da-a-addy	*Adaptor* – Left index and right thumb touch, right fingers splayed *Adaptor* – Beat hands on 'Lord' *Adaptor* – Upper body lean forwards
Such a beautiful dream	*Regulator* – Shake head in time
I hate to think it's all over	*Illustrator* – Point to self, with both hands
I've lost my heart it seems	*Regulator* – Low right hand, splayed gun shape, marks rhythm of the line
I've grown so used to you somehow	*Regulator* – Higher right hand, gun shape, marks rhythm of the line *Adaptor* – Hand falls in sync with falling melodic line
Well, I'm nobody's sugar daddy now	*Regulator* – High right hand, index and thumb touching, marks rhythm of the line
And I'm lo-o-onesome	*Regulator* – High right hand, index and thumb touching, marks rhythm of the line *Adaptor* – Hand tilts back and forwards with yodel pitch
I got the lovesick blues	*Regulator* – High right hand, index and thumb touching, marks rhythm of the line
Verse Well, I'm in love, I'm in love with a beautiful gal	*Regulator* – Relaxed right hand, marks rhythm of the line *Adaptor* – Both thumbs together, fingers splayed
That's what's the matter with me	*Regulator* – Raised left hand higher, marks rhythm of the line *Adaptor* – Hand lowers in sync with falling melodic line

Well, I'm in love, I'm in love with a beautiful gal	*Adaptor* – Both thumbs together, fingers splayed *Regulator* – Marks rhythm of 'beautiful gal'
But she don't care about me	*Illustrator* – Points to self with both hands
Lord, I tried and tried to keep her satisfied	*Adaptor* – Both thumbs together, fingers splayed
But she just wouldn't stay	*Emblem* – Point down to floor on 'she'
So now that she is leavin'	*Adaptor* – Both thumbs together, fingers splayed *Adaptor* – Right hand, thumb, index and middle finger splayed, waves to highlight yodel on 'leaving'
This is all I can say	*Adaptor* – Both thumbs together, fingers splayed
Chorus 2 I got a feelin' called the blues, oh Lord	*Adaptor* – Both thumbs together, fingers splayed, in an open pray shape
Since my baby said goodbye	*Adaptor* – Swipes right hand, with index and thumb touching, across the body
Lord, I don't know what I'll do	*Regulator* – Right hand, with index and thumb touching, marks the beat
All I do is sit and sigh, oh Lord	*Adaptor* – Throws right hand to head height on 'all' *Regulator* – Right hand, thumb, index and middle finger splayed, waves side to side to rhythm of yodel on 'sigh'
That last long day, she said goodbye	*Adaptor* – Right hand waves with thumb, index and middle finger splayed *Adaptor* – Shake head and rock forwards to falling melodic line on 'goodbye'
Well Lord, I thought I would cry	*Illustrator* – Point to self on 'I'
She'll do me, she'll do you	*Regulator* – Heavy hands at chest (both), fingers splayed, marks out the rhythm of the line
She's got that kind of lovin'	*Regulator* – Heavy hands at chest, fingers splayed, marks out the rhythm of the line

Lord, I love to hear her when she calls me sweet da-a-addy	*Adaptor* – Right hand relaxed point, waves to highlight yodel on ‘daddy’
Such a beautiful dream	*Adaptor* – Right hand relaxed point, draws small circles with index *Adaptor* – Index drops on dream
I hate to think it's all over	*Adaptor* – Right hand, touch chin with index and middle fingers Adaptor – Right hand, open, describes falling melodic line
I've lost my heart it seems	*Regulator* – Right hand, relaxed waves, marks the rhythm of the line Adaptor – Palm twists to vertical on vocal slide on ‘seems’
I've grown so used to you somehow	*Regulator* – Higher right hand, gun shape, marks rhythm of the line *Adaptor* – Hand falls in sync with falling melodic line *Adaptor* – Upper body lean forwards with melodic line
Well, I'm nobody's sugar daddy now	*Regulator* – Heavy hands at chest (both), fingers splayed, marks out the rhythm of the line
And I'm lo-o-onesome	*Adaptor* – Right hand waves on yodel on ‘lonesome’ *Adaptor* – Upper body lean forwards with melodic line
I got the lovesick blues	*Regulator* – Higher right hand, relaxed with vertical palm, marks rhythm of the line

Interview Notes

Think about it in the moment when the constraints of the music allow me to be more emotionally expressive. When not playing.

Subconsciously add to lyrically but not thinking about it but it does relate (I and you etc)

More likely to react melodic line

When accapella think about body more freedom when not playing instrument.

Regulators when live – coordinate with the other guys on stage – when things might be slipping or getting – more calculated – trying to get the band back in time. On own its more about moving to the music – keeping your body invested in the music.

Musical Theatre Vocalist – Get Out and Stay Out (Dolly Parton)

General vocal technique:

- Feet shoulder width apart
- Rooted in spot – little movement
- Natural, relaxed shoulders
- Balanced head
- Trying not to sway
- Pulling latissimus dorsi muscles down to support high notes

Labelled expressive gestures according to the lyrics of 'Get Out and Stay Out' by Dolly Parton:

Lyrics	Gesture descriptor
SILENCE AND STILLNESS	*Display* – Pause before starting to sing
Intro Well, It's funny how you waltzed in here, assuming I'd come back	*Adaptor* – Shoulder shrug *Adaptor* – Raise both hands slightly and let them drop
Well let me tell you something, you are way off track	*Emblem* – Shakes head at audience
Can't you see I'm different, or are you still that blind	*Adaptor* – Shoulder shrug *Adaptor* – Raise both hands slightly and let them drop
No - you stand right here and take it, there's no love to hide behind	*Illustrator* – Pointing down to ground with left index on 'here'
Well, I am proud to tell you I'm really feeling good	*Adaptor* – Raise both hands slightly and let them drop *Adaptor* – Squeeze back of both thighs
I'm doing so much better than you ever thought I would	*Adaptor* – Raise both hands slightly and let them drop
Got my own place, my own space, where I can dream and plan	*Adaptor* – Look to floor to right hand side
It took me this long to realize I do not need a man	*Emblem* – Smiling out wide at audience

SILENCE AND STILLNESS	*Display* – Pause before starting to sing
Verse 1 I used to need you, then I finally learned	*Illustrator* – Curved right hand touching stomach (point) on 'I' *Regulator* – Regular rocking from side to side
I used to want you, now the tables turned	*Illustrator* – Both hands curved touching stomach (point) on 'I' *Adaptor* – Raise both hands slightly and let them drop
I used to love you, now it's your time to squirm	*Illustrator* – Point to audience on 'your'
Cause I'm saying goodbye and I won't wait for your return	*Adaptor* – Raises left hand out forward then drops
Chorus 1 So get out and stay out, I've finally had enough	
Don't kiss me on your way out, it wouldn't move me much	*Adaptor* – Left hand point to audience
You used me, abused me, you cheated and you lied	
So get out and stay out, I'm taking back my life	*Adaptor* – Left hand point to audience *Emblem* – Point to ground on 'life'

Musical Theatre Vocalist – Once You Lose Your Heart (Noel Gay & Douglas Furber)

Labelled expressive gestures according to the lyrics of 'Once You Lose Your Heart' by Noel Gay & Douglas Furber:

Lyrics	Gesture descriptor
SILENCE AND STILLNESS	*Display* – Pause before starting to sing
Section 1 Once you lose your heart,	

Once somebody takes it,	*Adaptor* – Raise left hand slightly and let drop
From the place it rested in before.	
Once you lose your heart,	
Once somebody wakes it,	
There's one thing certain from the start,	*Adaptor* – Raise left hand slightly and let drop
You'll find forever,	
You've got to follow your heart.	
Section 2	
They say a girl should never be without love,	*Adaptor* – Spread both arms, fingers outstretched, and let drop
And all the joy that love alone can bring.	*Adaptor* – Raise both hands slightly and let them drop
All that I have ever learnt about love,	
Tells me it's a very funny thing.	*Adaptor* – Raise both hands slightly and let them drop *Adaptor* – Shoulder shrug
Section 3	
Once you lose your heart,	*Adaptor* – Squeeze back of both thighs
Once somebody takes it,	
There's one thing certain from the start,	*Adaptor* – Left hand relaxed point to audience, then moves to side
You've got to follow,	*Adaptor* – Both hands squeeze to fists *Adaptor* – Both hands open, slightly raised
You've got to follow your heart.	*Adaptor* – Raise left hand further, pulse slightly on long note on ‘heart’

Interview Notes

Most gestures are subconscious

Narrative based, i.e. talking about someone you might point to where they've just left, specific gestures might be choreographed – if the director wants you to acknowledge that's who your talking about

Natural movements to do with the song content – the emotion

At the beginning its very speech quality/diction based so gestures respond more directly to the lyrics

Rocking more energetically with the energy of the music increasing – responding the heightened emotion and the build in music. Towards a big blow out of emotion towards the end

So the narrative gestures like 'here', 'I', harsh verbs your thinking about enunciating

Character adaptation – you would build in ticks and movements specific for the character – if all of the given circumstances for the character show she's shy and insecure your movement would be minimal – as she's gains confidence she might move more to own the space – have to do the same every show to match lighting plots and stage elements – becomes pretty choreographed

Definitely don't use gestures to help with technique – can't rely on it in case you need to be doing some movement/dance routine whilst also singing.

Classical/Operatic Vocalist – Frühlingsglaube (Franz Schubert)

General vocal technique:

- Feet shoulder width apart
- Rooted in spot – little movement
- Natural, wide shoulders
- Balanced head

Labelled expressive gestures according to the lyrics of 'Frühlingsglaube' by Franz Schubert:

Lyrics and (*translation)*	Gesture descriptor
SILENCE AND STILLNESS	*Display* – Pause before starting to sing
Verse 1 Die linden Lüfte sind erwacht, *(Balmy breezes are awakened;)*	*Regulator* – Regular rocking from side to side
Sie säuseln und weben Tag und Nacht, *(they stir and whisper day and night,)*	*Display* – Nodding and smiling at audience

Sie schaffen an allen Enden, an allen Enden. *(everywhere creative.)*	
O frischer Duft, o neuer Klang, *(O fresh scents, O new sounds!)*	*Adaptor* – Reach up, hands clasped in 'begging' position *Adaptor* – Flare elbows at phrase top
O neuer Klang! *(O new sounds!)*	*Adaptor* – Outline melodic line with falling left hand
Nun, armes Herze, sei nicht bang! *(Now, poor heart, do not be afraid.)*	*Adaptor* – Lean into phrase top
Nun muss sich alles, alles wenden. *(Now all must change.)*	
Nun muss sich alles, alles wenden. *(Now all must change.)*	*Adaptor* – Lean into phrase top
SILENCE AND STILLNESS	*Display* – Slowly look across the audience smiling
Verse 2	
Die Welt wird schöner mit jedem Tag, *(The world grows fairer each day;)*	*Regulator* – Regular, deep rocking from side to side
Man weiss nicht, was noch werden mag, *(we cannot know what is still to come;)*	*Display* – Slowly shaking head at audience
Das Blühen will nicht enden, es will nicht enden. *(the flowering knows no end.)*	*Adaptor* – Hands clasped in 'begging' position *Regulator* – Pulsing elbows in time

Es blüht das fernste, tiefste Tal, es blüht das tiefste Tal. *(The deepest, most distant valley is in flower)*	*Adaptor* – Slowly drop hands, left hand lingers
Nun, armes Herz, vergiss der Qual! *(Now, poor heart, forget your torment.)*	*Display* – Slowly shaking head at audience
Nun muss sich alles, alles wenden. *(Now all must change.)*	*Adaptor* – sway into ornamented line
Nun muss sich alles, alles wenden. *(Now all must change.)*	
SILENCE AND STILLNESS	*Display* – Slowly look across the audience smiling

Classical/Operatic Vocalist – Zur Rosenzeit (Edvard Grieg)

General vocal technique:

- Feet shoulder width apart
- Rooted in spot – little movement
- Natural, wide shoulders
- Balanced head

Labelled expressive gestures according to the lyrics of 'Zur Rosenzeit' by Edvard Grieg:

Lyrics and (*translation)*	**Gesture Descriptor**
SILENCE AND STILLNESS	*Display* – Pause before starting to sing
Ihr verblühet, süße Rosen, *(You fade, sweet roses,)*	*Regulator* – Very slow regular rocking from side to side

Meine Liebe trug euch nicht; *(My love did not wear you;)*	*Display* – Slowly shaking head at audience
Blühet, ach! dem Hoffnungslosen, *(Ah! You bloom for one bereft of hope,)*	
Dem der Gram die Seele bricht! *(Whose soul now breaks with grief!)*	
SILENCE AND STILLNESS	*Display* – Slowly look across audience
Jener Tage denk' ich trauernd, *(Sorrowfully I think of those days,)*	*Adaptor* – Reach up, hands clasped in 'begging' position
Als ich, Engel, an dir hing, *(When I, my angel, set my heart on you,)*	*Adaptor* – Pulse 'begging' hands in circular motion
Auf das erste Knöspchen lauernd *(And waiting for the first little bud,)*	*Adaptor* – Drop right hand slowly, reach left to the side with an open palm
Früh zu meinem Garten ging; *(Went early to my garden;)*	
Alle Blüten, alle Früchte *(Laid all the blossoms, all the fruits)*	*Regulator* – Deep rock to the right on second 'alle'
Noch zu deinen Füßen trug *(At your very feet,)*	

Und vor deinem Angesichte *(With hope beating in my heart,)*	
Hoffnung in dem Herzen schlug. *(When you looked on me.)*	*Adaptor* – Nodding head on ‘Herzen’
Ihr verblühet, süße Rosen, *(You fade, sweet roses,)*	*Regulator* – Very slow regular rocking from side to side
Meine Liebe trug euch nicht; *(My love did not wear you;)*	
Blühet, ach! dem Hoffnungslosen, *(Ah! You bloom for one bereft of hope,)*	*Adaptor* – Shake head on ‘ach’
Dem der Gram die Seele bricht. *(Whose soul now breaks with grief.)*	
SILENCE AND STILLNESS	*Display* – Slowly look across the audience

Interview Notes

Rocking is subconscious, actually trying to avoid – bit too rooted trying to relax. More tense in pressurised situations – keep voice flowing

Like to look at the audience and keep them engaged. People like to feel like your singing to them – want to connect.

Reaching up with hands clasped – just happens in the moment – to empathise the words.

Not for physical act of singing.

Gestures respond to melodic and lyrical content of the song but are not choreographed – very spontaneous.

Don’t use many – unnecessary to them everywhere – too many don’t mean anything. So its to do with it being operatic –

Would use more gestures when singing pop/rock etc. and a cappella.

Appendix ii

Morphological Analysis:

Identify Body Movements	Select Effects	Control Effect Parameters	Feedback	Create Effects	Touchpoint Embodiment
Infrared motion tracking systems e.g. Vicon, Optitrack	Gesture recognition, certain gesture applies to an effect	Potentiometers	Vibrotactile, haptic motors	Guitar pedal board, mechanically actuated	No device – camera tracked
RGB colour tracking	Individual finger bent corresponds to an effect	Accelerometer/gyroscope thresholding	Material resistance e.g. in stretching	Custom circuitry	Motion capture suit
Pose tracking using machine learning e.g. Google's MediaPipe	Hand shapes correspond to different effects	Hand orientation tracking – roll, pitch yaw	LED colours show which effect is selected	Effects added with a DAW, e.g. Ableton	Wristband/armband
Wearable IMU sensor data	Buttons pressed correspond to effects	Pressure detection	Screen display	Digital effects built in Max 8	Glove
Bend sensors used to calculate joint angles	Body part in certain 'zone' corresponds to selecting effect	Bend of fingers (or other joints) corresponds to level of an effect	Audio feedback into earpiece (but not to audience)	Acoustic (mechanical) effects	Handheld remote controller
Built-in sensors in smartphone accessed	Swipe through gesture presets sequentially	Smart textiles – stretch	Animations		Ring
Microsoft Kinect	Proximity sensors	Incremental increase/decrease with swipe gestures	Projections		Stick on sensors
Leap Motion hand tracking	Hall effect effects and magnet	Ultrasonic/distance measurement			LED light suit to make colour tracking easier
Smart textiles – measuring stretch	Capacitive touch	Touchscreen display			Clip/stick to microphone – modular
		Pneumatic or hydraulic squeeze detection – reservoir in palm			

Appendix iii

Full Product Specification:

Category	Specification Point	Justification/Explanation
1) Selecting Effect Presets	1.1) Must be able to control a number of standard effects	e.g. reverb, delay, etc.
	1.2) Must be able to create personal presets	Singers often use special combinations of multiple effects to create a specific sound
	1.3) Must be able to choose personal selection gestures	Gestures are personal, as shown in the Gesture Identification Study. Singers must feel natural and comfortable when performing them, and must suit their individual performance style
	1.4) Must be able to choose which gesture triggers which preset	The meaning of a gesture is defined by the individual, in this case the meaning is the effect applied
	1.5) Must be able to route effects in series and parallel	The sound of an effect can vary massively due to the signal routing
2) Controlling Effect Parameters	2.1)Must be able to choose which effect parameters are controlled live	e.g. distortion sample rate, flange frequency, etc.
	2.2)Must be able to precisely tune parameters	Finding specific sounds requires precision tuning of effect parameters
	2.3)Must be able to experimentally manipulate parameters	The freedom to improvise and creatively experiment on stage creates ‘magic’ live moments
	2.4)Must be able to manipulate several effects at once	There should be no limits to the performer’s creatively
3) Impact on Performance	3.1)Device / use of device must not be distracting to the audience	The emotion of the performance must not seem synthetic to the audience
	3.2)Device / use of device must not be distracting to the performer	Use of the device must not inhibit the performers ability to convey their message
4) General Usability	4.1)Must be operable with one hand	To allow microphone to be held in the other
	4.2)Must not inhibit instruments to be played and handled	Many singers play instruments during their performances, the device must not stop that
	4.3)Must be wireless	The device must not inhibit the performer’s natural movements
	4.4)Device’s function must be scalable	The device must be configurable to the performer’s ability level
	4.5)Must be comfortable to use	To the 95th percentile
	4.6)Must be robust and durable	Long tours with multiple shows and travel can be taxing on kit
5) Commercial Viability	5.1)Must create a commercially viable sound	To promote widespread uptake in the industry
	5.2)Must be compact	To avoid affecting the performance
	5.3)Must be lightweight	To avoid fatigue on stage
	5.4)Must be low cost (potential for)	To promote accessibility to all performers

Appendix iv

Questionnaire Results:

	Question	Category	Question	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	Mean	Adjusted Mean	Standard Deviation
Background	1	Personal Values	Access to the control of my vocal effects live would elevate my performance	1	4	2	1	1	1	1	1	1	1	1.4	4.6	0.966091783
	2		Having freedom of movement is important to feeling natural and confident on stage	1	1	1	1	1	1	1	1	1	1	1	5	0
Computer Interface (Control)	3A	Selecting Effects / Effects Presets	I felt capable of using the controller effectively to apply effects/effects presets using the computer interface	1	2	1	1	2	4	1	1	1	2	1.6	4.4	0.966091783
	4A		I was able to apply the effect preset I wanted accurately	1	1	2	1	2	4	1	1	1	1	1.5	4.5	0.971825316
	5A		It was not obvious which effect preset I had loaded (-)	3	5	5	3	5	2	5	2	4	2	3.6	3.6	1.349897115
	6A		Choosing my own effect presets doesn't interest me (-)	5	4	5	5	5	5	5	5	5	5	4.9	4.9	0.316227766
	7A	Controlling Effect Parameters	I felt capable of precisely controlling the parameters of the selected effects	1	3	1	3	2	1	2	1	1	1	1.6	4.4	0.843274043
	8A		I felt capable of creatively experimenting with the sound I was creating	4	2	1	1	2	2	2	2	2	2	2	4	0.816496581
	9A		It was obvious when I was manipulating effects parameters	1	1	2	1	1	1	1	2	1	2	1.3	4.7	0.483045892
	10A		It was not obvious which parameters I was in control of (-)	5	4	4	4	4	5	4	4	4	5	4.3	4.3	0.483045892
	11A		Choosing which effects parameters I can control live doesn't interest me (-)	5	3	3	5	5	5	5	5	5	4	4.5	4.5	0.849836586
	12A	Impact to Performance	Adjusting parameters and selecting presets distracted me from my singing/performance (-)	1	2	3	3	4	3	3	2	2	1	2.4	2.4	0.966091783
	13A		With more practice I believe I would become fluent at selecting presets and controlling parameters whilst singing	4	3	1	2	1	1	3	1	2	3	2.1	3.9	1.100504935
	14A	General Usability	I was able to use the device whilst also handling a microphone	4	3	1	2	1	1	2	3	3	3	2.3	3.7	1.059349905
	15A		I would feel comfortable playing/holding an instrument whilst operating the system	5	4	2	3	1	3	4	4	4	5	3.5	2.5	1.269295518
	16A		When using the system my movements were constrained by it (-)	1	1	4	2	3	3	2	2	1	2	2.1	2.1	0.994428926
	17A		The movements required to use the system felt natural to me	4	2	1	2	2	2	3	4	3	3	2.6	3.4	0.966091783
	18A		The computer display restricted the spontaneity of my movement (-)	1	3	4	5	3	4	3	2	1	2	2.8	2.8	1.316561177
	19A		Operating the system whilst singing was uncomfortable (-)	3	2	5	5	5	3	2	2	2	2	3.1	3.1	1.370320319
	23		I found it difficult to understand how to route the effects (series/parallel) to my liking (-)	2	4	2	4	4	2	5	4	4	4	3.5	3.5	1.08012345
	20A	Commercial Viability	I was unable to create a sound which I found interesting artistically (-)	5	4	5	5	5	5	5	5	5	5	4.9	4.9	0.316227766
	21A		I was able to create a sound that would be commercially viable	1	3	2	1	2	2	2	1	1	1	1.6	4.4	0.699205899
	22A		I would recommend this product to other performers within the industry who are looking for new ways to express themselves on stage	3	2	2	1	2	2	4	3	3	3	2.5	3.5	0.849836586

IMU-sic Controller	3B	Selecting Effects / Effects Presets	I felt capable of using the controller effectively to apply effects/effects presets using gestures	1	2	2	3	1	2	2	1	1	2	1.7	4.3	0.674948558
	4B		The gesture recognition was accurate, I was able to apply the effect preset I wanted	2	2	3	3	2	2	2	2	2	2	2.2	3.8	0.421637021
	5B		It was not obvious which effect preset I had loaded with my gesture (-)	4	5	5	5	5	4	5	5	5	5	4.8	4.8	0.421637021
	6B		Choosing my own gestures and training the device to recognise them excites me	1	2	1	1	1	1	1	1	1	1	1.1	4.9	0.316227766
	7B	Controlling Effect Parameters	I felt capable of precisely controlling the parameters of the selected effects	3	3	2	3	2	3	2	3	2	3	2.6	3.4	0.516397779
	8B		I felt capable of creatively experimenting with the sound I was creating	1	2	2	1	1	1	1	1	1	1	1.2	4.8	0.421637021
	9B		It was obvious when I was manipulating effects parameters	1	1	3	1	1	1	1	1	1	2	1.3	4.7	0.674948558
	10B		It was not obvious which parameters I was in control of (-)	2	5	4	3	4	3	5	2	4	4	3.6	3.6	1.0749677
	11B		Choosing the axes of hand rotation to control chosen parameters interests me	1	4	3	1	1	1	1	1	1	1	1.5	4.5	1.08012345
	12B	Impact to Performance	Performing gestures distracted me from my singing/performance (-)	2	3	4	5	4	4	4	4	4	4	3.8	3.8	0.788810638
	13B		With more practice I believe I would become fluent performing the gestures whilst singing	1	2	1	1	1	1	1	1	1	1	1.1	4.9	0.316227766
	14B	General Usability	I was able to use the device whilst also handling a microphone	2	3	2	1	1	1	1	1	1	2	1.5	4.5	0.707106781
	15B		I would feel comfortable playing/holding an instrument whilst wearing the device	2	4	2	3	1	2	2	2	3	3	2.4	3.6	0.843274043
	16B		When using the controller, my movements were constrained by it (-)	4	3	5	4	4	4	4	5	5	4	4.2	4.2	0.632455532
	17B		The movements required to use the controller felt natural to me	2	3	3	2	2	2	2	2	2	2	2.2	3.8	0.421637021
	18B		The computer display restricted the spontaneity of my movement (-)	2	3	5	4	2	5	2	3	2	3	3.1	3.1	1.197219
	19B		The controller was uncomfortable to use (-)	5	5	5	3	5	5	4	5	5	5	4.7	4.7	0.674948558
	20B	Commercial Viability	I was unable to create a sound which I found interesting artistically (-)	4	5	5	5	5	5	5	5	5	5	4.9	4.9	0.316227766
	21B		I was able to create a sound that would be commercially viable	1	2	2	1	1	1	2	1	1	1	1.3	4.7	0.483045892
	22B		I would recommend this product to other performers within the industry who are looking for new ways to express themselves on stage	1	3	1	1	1	1	1	1	1	2	1.3	4.7	0.674948558

Appendix v

Testing Notes:

User	About	Notes (Author Responses)
U1	Studying at Laine Theatre Arts Musical Theatre and Pop Female, 19	Was excited to try the device, has been looking for a way to apply reverb to voice in gigs Quickly got to grips with the gesture recognition – had no real problems Effects are a bit much for what she would use them for Quite a bit of latency – more than shes used to Swipe right mistaken for circle a few times but mostly worked as intended Struggled to accurately control the effects parameters at the top and bottom of sliders – jumping back to the bottom – need to have upper and lower limits
U2	Royal College of Music A cappella and Classical Female, 21	Used the device quite tentatively More intense/out there effects weren't to her liking – own presets would be much more subtle Pitch shift effect is off putting – definitely wouldn't use that effect The device could cut out some post processing for the videos she produces – the effects added live whilst recording rather than in post Generally too many options – would prefer to just swipe through a set of presets – only two gestures this way (forward and back) rather than a complicated 'bop-it' routine Requires too much coordination Circle and sipe right getting confused Wouldn't want to spend loads of time trying to learn how to use it Probably would only use presets – not live adjustments Might only use delay / reverb Pitch shift could be used by bass singer in accapella group
U3	Rose Bruford College Actor Musician Male, 23	Pitch shift is really funny, particularly setting it low Hard to stay in tune with it on – could be useful for chorusing up or down the octave though Building more harmonies would be fun – perhaps a major triad for one gesture, minor for another – build up 'vocoder' effects without having to play the keys on the piano – I believe this is a feature on the Mi.Mu Gloves, definitely on Auug Motion Synth Continuous scale on the pitch shift is nice but might be useful to 'stick' to notes that are in-tune – less micro-tonality Dropping hand quickly clears results unwantedly

U4	Rose Bruford College Actor Musician Male, 22	Couldn't get the twist gesture to work very well – most specific movement that was trained, I had a go straight after and it worked fine so must be that he couldn't quite work the movement out properly Doesn't think he would have the same problem if he chose and trained his own gestures Swipe right and circle gestures got mistaken for each other a few times Found it a bit difficult to figure out which way to rotate the hand to control various parameters – when shown again he gained confidence Twisting in X (roll) was very effective – had good control Twisting in Y was less controlled
U5	Rose Bruford College Actor Musician Male, 21	Picked up gestures quickly Would be useful for recording – also asked to record something able to do in Max with 'Quick Record' Could be a more fluid way to find sounds with which to build tracks up from Sometimes is getting into the groove of a sound he's found and starts dancing along/doing natural gestures which trigger the controller to try and change the effect preset This then causes him to lose the sound hes just found
U6	Royal College of Music Session Singer & Classically trained Female, 23	Once gestures were demoed she could get the device to change presets fairly easily – swipe_right gesture gets confused with circle gesture – similar profiles Liked the effect parameters flicking over the top back to minimum – saves having to twist all the way around Show gesture name as well with the effects preset name to make sure the correct gesture was recognised Effects get cut when dropping hand – thinks it is the double leg tap The button keeps moving – needs to be secured more firmly Gesture presets keep jumping to maximum values not the preset – FIXED - bug in Max Would only use reverb probably so the device could afford to be simpler
U7	Rose Bruford Indie Band Member Male, 24	Dived right in – confident using gestures and hand orientation Palm of glove would need to be open for guitar to be played – piano would be fine Strumming guitar would likely set off effects without meaning to Device is only right handed – what happens if they're left handed Choice to be able to switch hands might be quite useful since you might switch hands with your microphone too Effects routing is very similar to how he puts his guitar pedalboard together so quite intuitive

U8	Rose Bruford Indie Band Member & Musical Theatre Male, 21	Wouldn't want too much to think about gestures if playing and singing already Drummer in the band and mainly on backing vocals so would probably not need effects as much and already constrained in movements Drumming whilst wearing would set off unwanted effects Would love to use for his solo music Hand dropping cuts effects sometimes Was able to use all the gestures well Particularly liked experimenting with the distortion effect
U9	Rose Bruford Indie Band Member Male, 22	Most effects work well – picked them up quickly Dropping hand triggers Double Tap Some difficultly with twist – showed again and managed to get it working Would really like to try training his own effects – would make using more natural Would like to add some crazy effects like a granulator Loops could be fun to work with, might be a little confusing though Glove would get in the way of playing bass unless completely fingerless and with an open palm
U10	Rose Bruford Indie Band Member Male, 22	Changing effects parameters is a little difficult to work out which way to twist but quite good fun exploring to find out Probably wouldn't do a lot of live changing – mainly stick to the presets Might whack reverb up on a last note or something Little confusion with button, didn't realise it needed to be held down Twisting in roll was easiest Some unwanted effects – circle and double tap getting applied by accident at times