VocalCords: Exploring Tactile Interaction & Performance with the Singing Voice

Max Addae[*]
MIT Media Lab
Cambridge, MA
maxaddaemusic@gmail.com

Nina Masuelli
Massachusetts Institute of Technology
Cambridge, MA
carinarm413@gmail.com

ABSTRACT

The close relationship between touch, gesture, and sound plays a critical role in expressive musical performance. Many acoustic instruments, ranging from strings to brass to percussion, involve some coupling of the "feel" of the instrument in the hands and the corresponding sound produced. The singing voice, however, is one of few musical instruments that typically does not involve touch-mediated interaction. Despite several neurological, psychological, and social connections demonstrated between the hands and voice, the coupling of touch and voice is surprisingly absent from traditional vocal performance technologies.

This provides the motivation for ***VocalCords***, which explores the design of a new digital music interface inviting tactile interaction and performance with the singing voice. The interface makes use of physical rubber cords, acting as stretch sensors, which are pulled and manipulated by the hands of the singer as they vocalize to augment and modify their voice in real-time – as if they were able to physically "touch" their own vocal cords. This approach allows for expressive, tactile control over the singing voice, which suggests a striking relationship between physical and musical tension. In this work, we explore the potential of touch-mediated vocal performance, as well as how this added tactile interaction may alter our experience with, and perception of, our singing voices.

Author Keywords

voice technology, singing, tactile interfaces, gestural control, stretch sensing

CCS Concepts

•**Applied computing** → **Sound and music computing;** *Performing arts;* •**Human-centered computing** → Interaction design theory, concepts and paradigms;

1. INTRODUCTION

Our experiences with music are, by nature, often multisensory: in addition to being heard, music is also felt, whether through rich vibration, sensorimotor synchronization, or movement along an instrument's body [16]. Touch, in particular, plays a vital role in music performance: its coupling with sound production has shown close ties to several aspects of virtuosity, including rhythmic accuracy [21, 19, 30], instrument learning [2, 24], and expressive control [28, 14].

The singing voice, however, is one of few exceptions to this phenomenon, characterized by its lack of direct tactile engagement. While this intangibility is part of what makes the voice unique, it is also somewhat surprising, given the connections between touch and voice are deeply ingrained in human experience, highlighted by the common practice of "talking with your hands", as well as the neurological proximity of hand and vocal control areas [32]. Singers, too, often incorporate physical gestures to externalize their musical expressions, which has been shown to aid singers' musical phrasing, breath control, and pitch accuracy [8, 35, 31].

Furthermore, the singing voice holds a unique position as an inherently personal, familiar, and widely accessible instrument, yet many individuals – even highly experienced vocalists – grapple with insecurities and anxieties regarding their voices. Common misconceptions about singing as an innate talent rather than a learned skill contribute to these anxieties, often leaving untrained vocalists feeling stuck with their natural voice [7]. However, advancements in voice technology offer promising avenues for reimagining our relationship with our singing voices, demonstrating potential neurological and behavioral impacts on emotion modulation and semantic content [26, 12]. This potential underscores the importance of designing vocal-driven musical experiences that encourage us to re-imagine our connection to our singing voices.

As such, this research investigates how to bring our singing voices further within our reach – that is, literally and figuratively speaking. We hypothesize that by adding a dimension of tactile control, people will feel more empowered to explore and express themselves through their voices. In other words, what if we *were* able to feel, touch, and manipulate our vocal cords as we sang?

The scope of research in this work is motivated by the following research questions (RQs):

- **RQ1:** *How can we design interactive voice technologies that invite tactile interaction with the singing voice?*
- **RQ2:** *How can the integration of touch contribute to audiences' perception of vocal expression and communication?*
- **RQ3:** *How can touch-mediated vocal performance technology alter our experience with, and perception of, our singing voices?*

---

[*] Primary Author

NIME'24, 4–6 September, Utrecht, The Netherlands.

We explore these research questions through the design, implementation, and performance of a new digital musical interface called *VocalCords*, a stretch-sensor based voice controller for tactile vocal augmentation. The interface proposes a natural, tangible means of manipulating and performing with the voice in a manner that is richly connected to the embodied practice of vocal production, unlike many traditional modes of tactile control in electroacoustic music, such as knobs, sliders, and buttons.

2. BACKGROUND & RELATED WORK

2.1 Touch and Sound

One of the most integral conceptual foundations of this research lies in the connection between touch and sound, particularly as it relates to musical learning, perception, performance, and instrument design. In pursuit of creating meaningful mappings between tactile interaction and vocal manipulation in *VocalCords*, it was important to build an understanding of how tactile experience guides musical experience.

2.1.1 Musical Learning

Recent research demonstrates a strong correspondence between our senses of touch and hearing: both senses are based on receptors that can analyze amplitude, frequency, and waveform in response to pressure stimuli (albeit with different degrees of subtlety), often within perceptual ranges that are roughly compatible [15]. As such, a great deal of multisensory integration occurs between touch and hearing when playing a musical instrument [20]. Tactile feedback from the instrument (achieved from factors such as the instrument's material, weight, arrangement of keys, strings, etc.) becomes key in allowing musicians to develop virtuosity and expressive control over their instrument [25]. As Leman describes, this feedback serves as "a multi-modal prerequisite for musical expressiveness" [28], as it gives the performer a reliable understanding of how their physical gesture translates to sound.

When it comes to the singing voice, whose sound production mechanisms are biologically hidden from the performer, the inherent lack of tactile feedback can make it challenging for inexperienced singers to develop an understanding of – and, hence, confidence in – their instrument [27, 1]. With regard to RQ3, these findings motivate *VocalCords*' investigation of how tactile connection to the singing voice may deepen performer's vocal comfort.

2.1.2 Musical Communication & Perception

While touch itself has a much lower bandwidth of transducing information for perception than vision or audition, it has a unique capacity to transmit emotional information [18]. Consequently, a performer's tactile relationship with their instrument, often represented through *musical gesture*, has been shown to play a pivotal role in conveying their musical and emotional intent. From the perspective of the listener, prior research suggests music perception fundamentally involves a "motor-mimetic" imitation of how the sound was created [13]; as such, the perceived physicality and "effort" from the performer has a significant impact on the audience's response to musical performance [11, 3]. Tactile metaphors (such as "sharp-blunt", "smooth-rough", and "warm-cold") play a significant role in a listener's processing and interpretation of musical sound. [15]. Within this framework, *VocalCords* proposes a bridge between the emotional depth of touch and our unique sensitivity to the human voice.

2.2 Tangible Musical Interfaces

Given the suggested connections between touch and sound, creative researchers have designed a variety of tangible musical interfaces. Deformable interfaces, in particular, are rapidly emerging as well suited for DMI design, as their use of flexible material can offer nuanced and responsive physical interaction with digital technologies that would not be possible with rigid interfaces or controllers. Deformable interfaces can also provide "playful, visceral, and exploratory music experiences" that allow for more inclusive musical outlets without regard for prior musical training [4, 33, 22].

Of the various modalities deformable interfaces can employ—e.g. stretching, squeezing, or bending—*VocalCords* uses stretch as its central modality, due in part to its close links with musical manipulation. In digital audio workstations (DAWs), for instance, users can often "stretch" a piece of music to increase the duration or pitch. In the physical world, the deformation of materials is often associated with sound, such as in the snapping of a stretched rubber band. Despite its rarity in digital music controllers, we demonstrate stretch offers a particularly striking metaphor for tactile voice control, as it effectively mirrors the flexibility, precision, and emotional expression commonly linked with the singing voice.

While prior works have explored stretch as a driver of sound synthesis, such as in Chang's *Zstretch* [9] and Wicaksono's *FabricKeyboard* [34], its use in parallel with a live musical source – the singing voice, in our case – has not yet been widely explored. *VocalCords*' design builds on this opportunity through its use of conductive rubber materials internally equipped with force-sensitive resistors (FSRs) to measure the material's stretch, allowing for natural sensing of the interface's deformation. This design approach pairs the expressive potential of physical stretch with the natural expressivity of the human voice, aiming to explore how this dialogue between touch and voice can enhance musical communication.

2.3 Gestural Control of the Voice

Artists and creative technologists have widely explored techniques for gestural voice control, often employing wearable sensor systems to track performers' gestures during live performances. Early examples like Michel Waisvisz's The Hands [5] utilized small keyboards, pressure sensors, and accelerometers to translate hand movements into MIDI control data, enabling manipulation of instruments and synthetic sound sources. Similarly, Elly Jessop's Vocal Augmentation and Manipulation Prosthesis (VAMP) [23] utilized a gestural vocabulary inspired by choral conducting, with intuitive hand movements correlating directly to changes in sound output. This approach became popular, as seen in interfaces like Laetitia Sonami's Lady's Glove [6] and Imogen Heap's MiMu Gloves [29], offering performers an expressive means of controlling their vocal output.

Although *VocalCords* shares similar goals of enabling physicality in the vocalist, it differs from wearable controllers in its positioning as an external physical object – necessarily bringing along its own physical language, tactile feedback, and resistance – that is in dialogue with the vocalist's hands and body. This design allows for a richly tactile vocal control mechanism, where the physical makeup of the interface can be held, deformed, and manipulated in relation to the voice's live transformation.

3. DESIGN & IMPLEMENTATION

3.1 Design Goals

In line with the aforementioned research questions, *VocalCords* was developed under the following design principles:

Expressivity

Like many musical interfaces, VocalCords aims to effectively empower and communicate musical intention. In particular, we position tension at the core of the interface's expressive language, given its dual meanings in the physical and musical world, and its fundamental role in guiding musical experience [17, 36]. By emphasizing stretch as a primary expressive modality, we create a natural connection between physical and musical tension, and present a powerful analogy of manipulating one's own "vocal cords".

Tangibility

VocalCords is designed to encourage a new kind of tactile interaction with the singing voice, aiming to push beyond the conventional "knobs-button-sliders" control paradigm present in most digital music interfaces. This principle is motivated by the fundamental role of tactile experience in musical perception and performance and, consequentially, we aim to integrate this tactile experience into the design of interactive voice technologies. While there is a specific focus on stretch as the central modality, the interface makes use of a richly varied palette of expressive, tactile gestures – including pulling, suspending, tugging, plucking, shape-forming – each with corresponding vocal processing modules evocative of the physical action.

Versatility

Finally, *VocalCords* is designed to offer control over a range of audio processing modules, allowing for performance applications in a variety of musical genres, tempi, and affective states. The system supports easily flexible gesture-sound mappings, and allows the performer to build a rich palette of rhythmic textures, harmonies, and timbral manipulations of their voice in real-time.

3.2 System Implementation

3.2.1 *Hardware & Sensor Design*

The system's hardware setup is built around two kinds of sensors: a) conductive cords made of carbon-black impregnated rubber (Figure 1) and b) a series of small 2-axis analog joysticks [1] (Figure 2), attached to both sides of the cord. The rubber cords, acting as stretch sensors, are at the core of the instrument's design, chosen both for their ease of use, as well as their potential as an intuitive abstraction of manipulating one's "vocal cords". As the string is pulled, the resistance increases as the particles get further apart, and once the force is released, the rubber slowly shrinks back to its original length and default resistance. By clipping each end of the string and connecting them into a simple voltage divider circuit, we can measure the resistance of each string as it is pulled and contracted.

To gather further gestural data while stretching the cords, we attached small joysticks to each side of the rubber cord to get approximations of the cords' axis of stretch as they are pulled or elevated. By connecting the joystick's breakout board to our Arduino microcontroller [2], we measure the X and Y movement of the joystick, providing estimations of the string's elevation and the fingers' location. By pairing the data from each joystick with the stretch sensor readings, we were able to develop a more expansive set of expressive gestures with the cords and have more flexibility with how these gestures could be creatively mapped. The joysticks also proved vital in detecting when the cords would "cross" over each other, establishing the system of cords as a connected control network. Each of the joysticks were also embedded with a button, which were used to toggle modules on/off or switch an effect's behavior.

Figure 1: Conductive Rubber Cord Stretch Sensors

Figure 2: Analog 2-axis Thumb Joystick (Adafruit Industries)

3.2.2 *Fabrication & Physical Design*

We began by creating an initial enclosure prototype for the interface out of foamcore. We chose to give the interface a "table-top"-like design, as it is a familiar and approachable design of electronic music interfaces.

To attach the cords to the joysticks, we carved out small holes in the joystick's caps, and each of the strings were then weaved through and wrapped around the center of the joystick. The backs of the joysticks and ends of the rubber cords are obscured by removable side panels, creating easy access for the cords to be tightened and replaced as they wear down over time. A sheet of self-adhesive reflective vinyl was also placed on the bottom of the enclosure underneath the cords to help accentuate and add depth to the hand movements of playing the instrument, creating an ethereal and intimate visual effect which is well suited to the instrument's sonic world.

As the foamcore prototype quickly began to wear out over time, we ultimately switched to an all-wooden enclosure for the final design, which was 3D-modeled (Figure 3) in Rhino [3] and assembled using laser-cut wood pieces.

The final physical design of the interface is shown in Figure 6.

---

[1] `https://www.adafruit.com/product/512`

[2] `https://www.arduino.cc/`

[3] `https://www.rhino3d.com/`

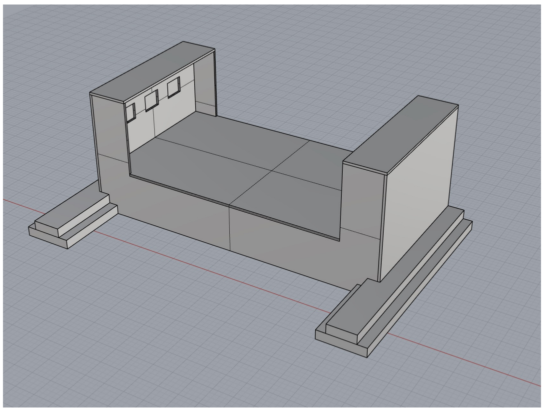


Figure 3: 3D Model of VocalCords' final wood enclosure

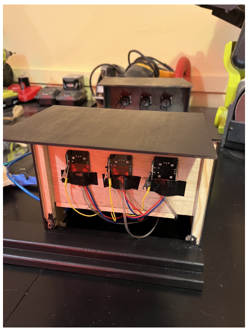


Figure 4: Wood Enclosure Assembling

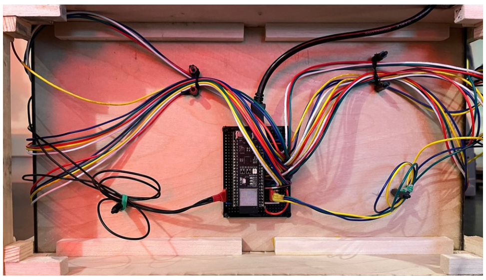


Figure 5: Circuit Board

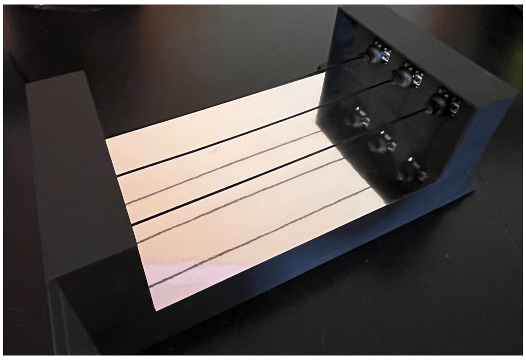
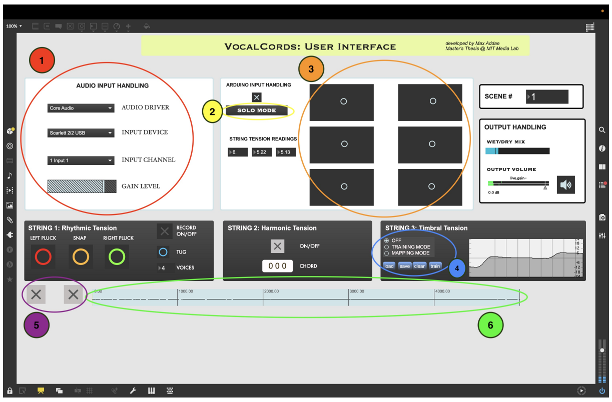


Figure 6: *VocalCords* Final Physical Design

3.2.3 *Software Design*

A small Arduino IDE program is used to measure the analog voltage for each string, and convert them back to their corresponding resistance values. These values are then packed into a list with each of the joysticks' X/Y values, and sent into Max 8 [4] through serial transmission. In Max, the sensor data is captured and linearly smoothed before routing to the respective sound processing channels. To set up, a short calibration is run to measure each string's resting resistance, stretching each string three times to measure their minimum and maximum resistance values, as they tend to vary with each use.

A simple presentation interface was created to easily visualize all of the sensor data and gesture tracking in performance settings, shown in Figure 7.



VocalCords: User Interface

developed by Max Addae
Master's Thesis @ MIT Media Lab

AUDIO INPUT HANDLING
Core Audio AUDIO DRIVER
Scarlett 2i2 USB INPUT DEVICE
1 Input 1 INPUT CHANNEL
GAIN LEVEL

ARDUINO INPUT HANDLING
SOLO MODE
STRING TENSION READINGS
6. 5.22 5.13

SCENE # 1

OUTPUT HANDLING
WET/DRY MIX
OUTPUT VOLUME
live.gain~
0.0 dB

STRING 1: Rhythmic Tension
LEFT PLUCK SNAP RIGHT PLUCK
RECORD ON/OFF
TUG
4 VOICES

STRING 2: Harmonic Tension
ON/OFF
0 0 0 CHORD

STRING 3: Timbral Tension
● OFF
○ TRAINING MODE
○ MAPPING MODE
load save clear train

Figure 7: *VocalCords* User Interface in MAX; (1) Audio Input Settings; (2) Mode Label, for detecting when the cords cross and enter new modes; (3) Visualization of the Joysticks' X/Y Positions; (4) Machine Learning Setup for Timbral String; (5) Toggles for visualizing when the rhythm string is in delay (left toggle) or granular mode (right toggle); (6) Visualization of the Recording Buffer after the rhythm string is "suspended"

3.3 Mapping Scheme

The relationship between physical and musical tension is at the core of *VocalCords*' mapping schema. In the design of the system's mapping scheme, we chose to make the strings unique in their gestural character and vocabulary with respect to their associated musical parameters, in order to clearly establish their musical and functional identities. Through this mapping approach, we aim to achieve high-level expressive connections between tactile gestures and vocal processing facilitated by rhythmic, harmonic, and timbral tension.

A diagram of the high-level mapping scheme is shown in Figure 8.

3.3.1 *Rhythmic String*

The rhythmic string's control gestures were designed for sharp, specific movements, in order to give a "rhythmic" physical behavior to the corresponding rhythmic/temporal processing. Given the variety of ways one can think about "rhythm" (e.g. in terms of repetition, tempo, texture, etc.), we chose to contextualize rhythm with respect to two temporal processes: a) delay/echo and b) granular synthesis, which allowing us to achieve each of these rhythmic "states".

---

[4] `https://cycling74.com/products/max`

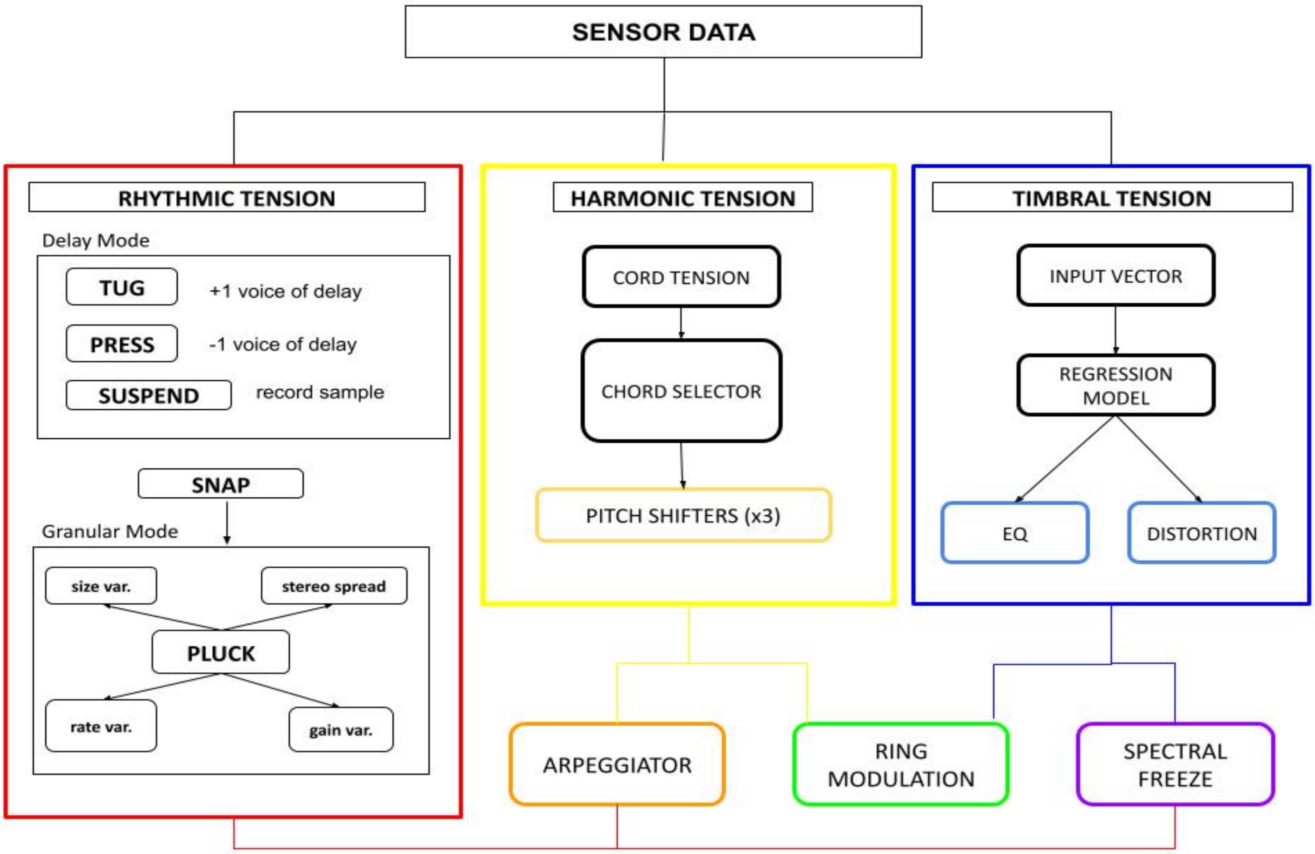


Figure 8: *VocalCords* Mapping Diagram

In the string's delay mode, an initial **tug** of the string – detected by a quick spike in the string's measured tension – initializes a set of six randomly chosen delay times ranging from 500-3500ms, and one delayed voice is turned on, as if to set an initial rhythmic pulse. With each subsequent tug, an additional delay line is toggled on, resulting in a rich layering of delayed voices and a less "secure" sense of pulse. While the delay times are typically randomized, the performer is able to set metric delay times by plucking the string four times consecutively, and the time intervals between each pluck are averaged to set the tempo, analogous to the "tap tempo" functionality present in many MIDI controllers.

Conversely, delay lines can be toggled off by **pressing** down on the string, after which a delayed voice is faded out/turned off every 1500ms for as long as the string is held down. This gesture serves as a sense of de-escalation, or a release of rhythmic tension.

When the cord is **snapped** (i.e. lifted and immediately released), a granular synthesis engine is triggered, meant to mirror the string's rapid physical oscillations following its release. Using Dan Trueman's *munger~* external in Max, the incoming audio signal is segmented into small "grains" which can vary in length, periodicity, pan spread, and amplitude[5]. The parameters are originally set to the following default values (see Table 1) to create rapid, yet regular sonic oscillations. Plucking the cord in one of four "quadrants" adds variation to the associated parameters, in order to use irregularity as a form of rhythmic tension.

Granular Synth Parameter	Default Value
Grain Size	100ms
Grain Size Variation	0
Grain Period	1ms
Grain Period Variation	0
Gain Variation (+/-)	0
Stereo Spread (ranges from 0. to 1.)	0.25

Table 1: Granular Synth Default Parameter Values

When the cord is **suspended** (i.e. lifted and held up for > 500ms), the incoming audio signal is recorded into a buffer until the cord is released. After its release, any subsequent delay/granular processing is then applied to the recorded loop instead of the live input. This allows the performer to layer and process an audio loop, as if the performer were metaphorically capturing and releasing a moment in time.

Using the embedded button of the string's left joystick, the performer can toggle between rhythmic modes (delay, granular, or both off), and the button on the string's right joystick switches the processing to be applied to either the live or recorded audio.

3.3.2 Harmonic String

In contrast to the sharp, discrete nature of the rhythmic string's control gestures, the harmonic control string was designed for more continuous gestures, allowing the chord structures to seamlessly expand and contract as the performer stretches the cord. A set of five distinct chords, represented as a series of pitch shifters applied to the audio

---

[5] `https://github.com/Cycling74/percolate/tree/master/source/projects/munger~`

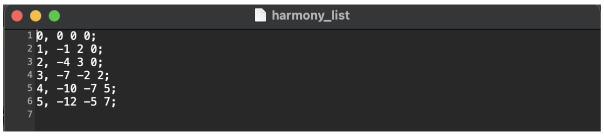
Figure 9: Performance Shot from *"In Tense Dimensions"* of the suspended string gesture (Courtesy of Jimmy Day)

input, were called upon in direct relation to the amount of tension measured from the rubber cord. We found limiting each harmonic group (or "progression") to five chords was quite effective in providing a sufficient amount of harmonic variety, while allowing each chord to have a distinct "position" within the cord to reliably call upon.

The chords were intentionally ordered such that chords with more narrow (or "closed") voicing were called upon when tension is first introduced to the string, typically beginning with a dissonant "cluster" chord consisting of the original pitch, the pitch one half step down, and the pitch two half steps up. As the string is pulled further outwards, the chords' pitches are voiced further apart resulting in chords with a wider (or "open") structure, typically ending with a chord of stacked perfect fifths as the string reaches its maximum stretch. In this way, expanding the cord becomes analogous to pulling apart the chords' voices, and contracting the cord becomes analogous to pushing the chords' voices back together. Chords within each group were chosen to loosely follow conventional voice-leading strategies of Western harmony, allowing the transitions between chords to sound smooth and natural.

By pressing on the button embedded on the string's right joystick, the performer is able to switch between chord collections, allowing for more versatility in chord progressions to use in different pieces. The stereo width of the chord voices can also be controlled with the elevation of the string, such that the widest spread occurs when the string is lifted and stretched all the way outwards.

An example of a chord progression list is shown in Figure 10.



```
0, 0 0 0;
1, -1 2 0;
2, -4 3 0;
3, -7 -2 2;
4, -10 -7 5;
5, -12 -5 7;

```

Figure 10: Example Chord Progression List, where each integer corresponds to the respective number of half steps to pitch shift on the input signal.

3.3.3 *Timbral String*

For the timbral control string, we chose to focus the control gestures around the "shape" of the string, analogous to sculpting and shaping the tone of the voice. As such, we found it particularly useful to map sensor data and sonic output using a combination of custom trained machine learning models, allowing for more flexible gesture-sound mappings. The cord's tension value, along with its joysticks' X and Y displacement positions, were combined to create an input feature vector (representing the "shape" of the cord). *ml.**[6], a machine learning toolkit for Max, is used to map the input vector to a set of timbral modification parameters using a series of multiple regression models. In particular, the following timbral control modules were called upon in relation to the string's shape:

- **EQ**: In this mode, the input feature vector is mapped to a list of biquad filter coefficients to create a filter curve roughly resembling the string's orientation. The player can then apply, for instance, a lowshelf filter to the input signal when the string is held in the corresponding graphical shape, or bring out mid-range frequency "peaks" when the middle of the string is suspended. Complex filters, by this algorithm, can then be applied when the string is held in the resemblant formations. While the mapping from orientation to filter shape is not always exact due to the limitations of the available sensor data, it suggests a striking analogy of physically shaping the applied EQ curve.

- **Distortion:** Using the *pong~* object in MAX, the input feature vector is mapped to a signal folding effect, creating a grittier, heavily distorted timbre to the voice. Depending on the shape of the string, the player can switch between several folding-modes (e.g. fold, clip, wrap) and presets, suggesting a similar analogy of distorting the voice signal in relation to the string's "distorted" orientation.

3.3.4 *Cross-String Modes*

In giving each of the strings individuality in their gestural character, it became important to consider how to still establish the cords as a connected control network. We chose to implement this by switching into new modes whenever combinations of strings would "cross" (inspired by the range of figures and expressivity of cat's cradle [7]), as if to combine their functionalities and musical controls.

Harmonic String x Rhythmic String: Crossing the harmonic string over the rhythmic string triggers an *arpeggiation* functionality, where the pitch shifting of the harmonic string is applied onto the delay/granular synthesis effects. As the harmonic string traverses over the rhythmic string from the left to right, more pitch shift values are uncovered from the list, culminating to a dense collection of pitches as the harmonic string reaches over the rhythmic string's top right corner.

Rhythmic String x Timbral String: Crossing the rhythm and timbral string over each other (creating a diamond-like formation) triggers a *spectral freeze*, which serves to capture and "freeze" a specific point in the incoming audio signal. The implementation is based on a *pfft~* subpatch in Max designed by Jean-Francois Charles [10], which uses Jitter matrices to perform a Fast Fourier Transform (FFT) on the audio signal, record eight frames of the spectral profile, and repeatedly loop/blend the spectral frames together. As such, it achieves a combined rhythmic and timbral tension by freezing temporally *and* spectrally. After freezing the signal, the timbre string can then be used to control a denoiser to further shape the tone of the drone.

Harmonic String x Timbral String: Crossing the harmonic and timbral strings controls a ring modulation effect, a form of amplitude modulation which combines the input

---

[6] `https://www.benjamindaysmith.com/ml-machine-learning-toolkit-in-max`

[7] `https://thekidshouldseethis.com/post/play-cats-cradle-string-game`

voice signal (or the modulator signal) with a selection of different carrier signals (specifically triangle and square wave), resulting in a harsher, more metallic vocal timbre. This effect also produces two output frequencies, called *sidebands*, with frequencies at the sum and difference of the two signals' frequencies. Hence, it concurrently achieves a kind of harmonic and timbral tension.

4. PERFORMANCE & DISCUSSION

VocalCords had its first public performance at the MIT Media Lab on June 30, 2023, where the first author performed an originally composed four-movement song cycle titled *In Tense Dimensions*[8] for an audience of approx. 50 people, exploring connections between physical, musical, and emotional tension, inspired by personal experiences with the strains of isolation and anxiety arising from the COVID-19 pandemic. The cycle was performed twice in full, with a brief talkback/Q&A with the audience in between performances, allowing them to first listen without preconceived notions, and listen again with more context of the tactile metaphors and emotional themes being explored. We chose to conduct the conversation in an open-ended, non-structured manner, to allow for the audience's most natural feedback without specific prompting.

During the Q&A session between performances, we were able to gain great insight into the audience's reception of *VocalCords* as a mediator of musical tension. In general, the piece and the interface both seemed to be well-received by the audience: many audience members seemed very enthusiastic to learn more about how the interface was designed, why certain mapping choices were made, and how the interface contributed to the performer's creative process.

With regards to RQ1, the audience's feedback suggested stretch as a promising expressive modality for tactile vocal manipulation. One audience member shared the following quote in discussing how the integration of tactile gesture and voice contributed to their experience: "In any vocal performance, the singer primarily emotes through their vocal choices and facial expressions. In part, they will also use their body language. However, in *In Tense Dimensions*, physical manipulation of the strings enhanced the message the singer was trying to communicate—it made me as an audience member feel the same sense of strain that the singer was communicating." Another shared: "It definitely made the performance more engaging and made it feel like you were using more than just your throat to sing."

However, the designed mappings between physical and sonic action were not always as easily interpreted as we expected; one audience member claimed: "They (the physical gestures) definitely felt pretty abstract. I got that each sensor was a different category of manipulation, but I don't think I would be able to pick up the interface and immediately know what I'm doing without further explanation on the specifics. I would see the hands manipulating the strings and I would hear things happening, but I couldn't quite make a clear correlation that I could replicate from just watching. It definitely felt expressive though and the performers' clear familiarity with the instrument made it an engaging performance." Hence, with respect to RQ2, continued refinement of the instrument's gesture-sound mapping scheme must be done in order to more reliably convey the performer's musical intentions.

Finally, with respect to RQ3, we found the integration of touch also showed promise in guiding the performer's vocal exploration and virtuosity. In the process of developing the compositions for the song cycle, the performer was able to unlock new timbres in their voice – e.g. vocalized heartbeats, pitched gasps, vocal fry – in an attempt to vocally match the intense physicality and "strain" of the tactile gestures. Additionally, while sustaining a pitch on the harmonic string, they would naturally find themself shaping their vocal dynamics in relation to the stretch of the cord, crescendo-ing as the cord was pulled further out and descrendo-ing as the tension was released. While further study would need to be conducted on a larger set of vocalists to strengthen these findings, these experiences seem to suggest tactility's potential role in guiding vocal expressivity and virtuosity.

5. CONCLUSIONS & FUTURE WORK

We have presented the design and implementation of a stretch-based musical interface inviting a new kind of real-time, tactile interaction and performance with the singing voice. As part of this, we've highlighted the importance of tactile experience in musical performance and perception, discussed the limitations of presently available interactive voice technologies, and provided motivation for designing digital voice-based musical interfaces with an embodied, tactile approach.

This work also lays the foundation for future research and development in the design of tactile live voice performance technologies. Future iterations of the interface could explore a more customizable gesture-sound mapping user interface allowing for more flexibility, or implementing *VocalCords* as a MIDI-driven controller, allowing for easy integration with a DAW's sound processing modules. Another approach could involve embedding signal processing and high-level analysis of the voice input into the system to inform *VocalCords*' sensitivity and tactile response, establishing a mutual exchange between touch and voice in the "vocal cords" abstraction, and allow for deeper investigation in the relationship between tactile and vocal gesture.

6. ACKNOWLEDGMENTS

This research was supported through the Opera of the Future group at the MIT Media Lab, as well as the Council for the Arts at MIT (CAMIT). Special thanks to Tod Machover, Joe Paradiso, and Akito van Troyer, who served as advisors for the project.

7. ETHICAL STANDARDS

This work complies with the NIME ethical standards. A formal user study was not conducted as part of this research, so no ethical issues relating to study subjects were encountered.

8. REFERENCES

[1] C. R. Abril. I have a voice but I just can't sing: a narrative investigation of singing and social anxiety. *Music Education Research*, 9(1):1–15, Mar. 2007. Publisher: Routledge _eprint: https://doi.org/10.1080/14613800601127494.

[2] M. Aho. *The tangible in music: The tactile learning of a musical instrument.* Routledge, 2016.

[3] P. Bennett, N. Ward, S. O'Modhrain, and P. Rebelo. Damper: a platform for effortful interface development. In *Proceedings of the 7th international*

---

[8] Full Performance can be viewed here: `https://www.youtube.com/watch?v=Ve1n_P_RYH8`

conference on New interfaces for musical expression, pages 273–276, 2007.

[4] A. Boem, G. M. Troiano, G. Lepri, and V. Zappi. Non-rigid musical interfaces: Exploring practices, takes, and future perspective. In *New Interfaces for Musical Expression*, 2020.

[5] A. J. Bongers. Tactual display of sound properties in electronic musical instruments. *Displays*, 18(3):129–133, May 1998.

[6] B. Bongers. Physical interfaces in the electronic arts. *Trends in gestural control of music*, pages 41–70, 2000.

[7] M. C. Brand. Male high school students' perceptions of choral singing. Master's thesis, University of Illinois at Urbana-Champaign, 2019.

[8] D. M. C. Brunkan and D. J. Bowers. Singing with Gesture: Acoustic and Perceptual Measures of Solo Singers. *Journal of Voice*, 35(2):325.e17–325.e22, Mar. 2021.

[9] A. Chang and H. Ishii. Zstretch: A stretchy fabric music controller. In *Proceedings of the 7th International Conference on New Interfaces for Musical Expression*, NIME '07, page 46–49, New York, NY, USA, 2007. Association for Computing Machinery.

[10] J.-F. Charles. A tutorial on spectral sound processing using max/msp and jitter. *Computer Music Journal*, 32(3):87–102, 2008.

[11] E. T. Cone. " musical form and musical performance" reconsidered. *Music Theory Spectrum*, 7:149–158, 1985.

[12] J. Costa, M. F. Jung, M. Czerwinski, F. Guimbretière, T. Le, and T. Choudhury. Regulating feelings during interpersonal conflicts by changing voice self-perception. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, page 1–13, New York, NY, USA, 2018. Association for Computing Machinery.

[13] A. Cox. The mimetic hypothesis and embodied musical meaning. *Musicae scientiae*, 5(2):195–212, 2001.

[14] M. Doğantan-Dack. In the beginning was gesture: Piano touch and the phenomenology of the performing body. In *New perspectives on music and gesture*, pages 243–266. Routledge, 2016.

[15] Z. Eitan and I. Rothschild. How music touches: Musical parameters and listeners' audio-tactile metaphorical mappings. *Psychology of Music*, 39(4):449–467, 2011.

[16] D. Farrell. Music Beyond Sound: Weighing the Contributions of Touch, Sight, and Balance – Frank A. Russo, Feb. 2020.

[17] W. E. Fredrickson. Perception of tension in music: Musicians versus nonmusicians. *Journal of music Therapy*, 37(1):40–50, 2000.

[18] A. Gallace and C. Spence. *In touch with the future: The sense of touch from cognitive neuroscience to virtual reality*. OUP Oxford, 2014.

[19] W. Goebl and C. Palmer. Tactile feedback and timing accuracy in piano performance. *Experimental Brain Research*, 186:471–479, 2008.

[20] J. Huang, D. Gamble, K. Sarnlertsophon, X. Wang, and S. Hsiao. Feeling music: integration of auditory and tactile inputs in musical meter perception. *PloS one*, 7(10):e48496, 2012.

[21] R. H. Jack. *Tangibility and richness in digital musical instrument design*. PhD thesis, Queen Mary, University of London, 2019.

[22] A. R. Jensenius and A. Voldsund. The music ball project: Concept, design, development, performance. In *New Interfaces for Musical Expression*, 2012.

[23] E. N. Jessop. The vocal augmentation and manipulation prosthesis (vamp): A conducting-based gestural controller for vocal performance. In *NIME*, pages 256–259, 2009.

[24] M.-L. Juntunen and L. Hyvönen. Embodiment in musical knowing: how body movement facilitates learning within dalcroze eurhythmics. *British Journal of Music Education*, 21(2):199–214, 2004.

[25] S. W. Keele. *Attention and human performance*. Goodyear Pub. Co, Pacific Palisades, Calif, Jan. 1973.

[26] R. Kleinberger. *Vocal Connection: Rethinking the Voice as a Medium for Personal, Interpersonal, and Interspecies Understanding*. PhD thesis, Massachusetts Institute of Technology, 2020.

[27] M. Latinus and P. Belin. Human voice perception. *Current Biology*, 21(4):R143–R145, 2011.

[28] M. Leman. *Embodied music cognition and mediation technology*. MIT press, 2007.

[29] T. Mitchell, S. Madgwick, and I. Heap. Musical Interaction with Hand Posture and Orientation: A Toolbox of Gestural Control Mechanisms. In *NIME*, 2012.

[30] V. Occelli, C. Spence, and M. Zampini. Audiotactile interactions in temporal perception. *Psychonomic bulletin & review*, 18:429–454, 2011.

[31] M. S. O'Modhrain. *Playing by feel: Incorporating haptic feedback into computer-based musical instruments*. Ph.D., Stanford University, United States – California, 2001. ISBN: 9780493087979.

[32] W. Penfield and T. Rasmussen. The cerebral cortex of man; a clinical study of localization of function. 1950.

[33] S. Uğur Yavuz, P. Veske, B. Scholz, M. Honauer, and K. Kuusk. Design for playfulness with interactive soft materials: Description document. In *Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '21, New York, NY, USA, 2021. Association for Computing Machinery.

[34] I. Wicaksono and J. A. Paradiso. Fabrickeyboard: multimodal textile sensate media as an expressive and deformable musical interface. In *NIME*, volume 17, pages 348–353, 2017.

[35] R. M. Wis. Physical Metaphor in the Choral Rehearsal: A Gesture-Based Approach to Developing Vocal Skill and Musical Understanding. *The Choral Journal*, 40(3):25–33, 1999. Publisher: American Choral Directors Association.

[36] L. M. Zbikowski. Metaphor and music. *The Cambridge handbook of metaphor and thought*, pages 502–524, 2008.