VOICON: An Interactive Gestural Microphone For Vocal Performance

Yongki Park
Music and Audio Research Group, GSCST
Seoul National University
Seoul, Korea
yongkipark@snu.ac.kr

Hoon Heo
Music and Audio Research Group, GSCST
Seoul National University
Seoul, Korea
cubist04@snu.ac.kr

Kyogu Lee
Music and Audio Research Group, GSCST
Seoul National University
Seoul, Korea
kglee@snu.ac.kr

ABSTRACT

This paper describes an interactive gestural microphone for vocal performance named Voicon. Voicon is a non-invasive and gesture-sensitive microphone which allows vocal performers to use natural gestures to create vocal augmentations and modifications by using embedded sensors in a microphone. Through vocal augmentation and modulation, the performers can easily generate desired amount of the vibrato and achieve wider vocal range. These vocal enhancements will deliberately enrich the vocal performance both in its expressiveness and the dynamics. Using Voicon, singers can generate additional vibrato, control the pitch and activate customizable vocal effect by simple and intuitive gestures in live and recording context.

Keywords

Gesture, Microphone, Vocal Performance, Performance Interface

1. INTRODUCTION

Vocal performance is one of the most processed music signals. In contemporary vocal performances, the use of vocal modifications and augmentations is very commonly found. However, vocal performance relies heavily on the capability of a sound engineer as the sound engineer is often in charge of audio processing control of a performance. Vocal performers are not given the freedom to use these capabilities to augment, process, and control his/her own performance. This inability of the performer to use vocal effects limits the vocal performance to be bounded within premade agreements between a sound engineer and a performer leaving no room for improvisational use of vocal modifications and augmentations in live performances. Voicon is an interactive gestural microphone for vocal performers. This gesture-sensitive microphone lets the performers to generate vibrato, control pitch and use customizable vocal effects in real-time with simple gestures and use of pre-programmed sensors, giving new possibilities of performances in live and recording context. This microphone-shaped gestural controller differs from previous gestural system as it does not require performers to wear any invasive gestural sensors.

2. RELATED WORKS

There have been previous efforts to use gesture for music control in existing literature [1][2]. This paper focuses specifically on developing a vocal interface using gesture controls. Vocal performers employ a wide range of 'gestures' during performances [3]. There are numerous wearable music controllers that use a variety of sensors to capture gestures for the enhancement of vocal performances. "The Hands" created by Michel Waisvisz incorporated small keyboards on the hands of the player. The player can manipulate force sensors using his thumb and other sensors are used to detect the tilt of the hands and distance between them. Waisvisz used this instrument to control a variety of parameters to change the sound of his voice and other sonic sources [4]. Another hand gesture based instrument is Laetitia Sonami's "Lady's Glove," developed by Sonami and Bert Bongers. This glove uses a Hall Effect sensors on the thumb, magnets on the other four fingers, flex sensors on each finger, switches on the top of the fingers, and ultrasonic receivers. During a vocal performance, data from these sensors are used to control sound, lighting and motors [5][6]. The "Bodycoder" system created by Bromwich and Wilson is another gestural controller that has been occasionally used for vocal performances. Resistive sensors on the knee and elbow joints and keypad-like switches are employed in this system. Switches are used to trigger pre-recorded samples and select particular audio and visual patches [7]. Elena Jessop's Vocal Augmentation and Manipulation Prosthesis (VAMP) allows singers to remotely control, modulate and transform their voices by capturing arm gestures with an arm sleeve embedded with a variety of sensors [8]. Donna Hewitt created the eMic [9], a microphone and a stand equipped with controls that allow a performer to filter and process his or her voice live. A major difference between Voicon and previous gestural vocal controllers is that Voicon is a standalone microphone with sensors embedded within it. Voicon does not require performers to wear or use any other unnatural external sensors keeping interaction between the performer and the device seamless and natural.

3. MAPPING

The mapping between gestures and vocal augmentation and modification consists of three parts: generating vibrato, holding note/controlling pitch and activating customizable vocal effects. The summary of the mapping between gestures and functions is shown in Table 1.

3.1 Generating Vibrato

Vocal vibrato has been recognized as a prominent characteristic of classical western singing for a long time [10]. Professional singers inherently seem to develop vibrato without

NIME'12, May 21 – 23, 2012, University of Michigan, Ann Arbor.

thinking about it and without actively striving to acquire it. It develops naturally as voice training proceeds successfully [11]. There are experimental data showing the inability of professional opera singers to change their vibrato rate, even when asked to do so [12].Using Voicon, both novice and expert vocal performers can achieve and control the desired amount of vibrato in their performance. When the performers shake the Voicon in circular motions, vibrato is generated. A circular motion is mapped to vibrato as it has features that could be easily controlled such as radius and rate of the motion. Rate of circular movement of Voicon is proportionally mapped with the rate of vibrato whereas the radius of circular motion is directly related with depth of the generated vibrato. One might argue that both vertical and horizontal motions suit better with generating vibrato. However during an active vocal performance with a lot of physical movement by the performer, vertical and horizontal movements are bound to occur more often then regular circular movements. Circular motion is chosen to eliminate false generation of vibrato. Moreover, with circular movement of Voicon, the performer has greater room to control the rate and depth of vibrato as circular motions have both vertical and horizontal movement features.

3.2 Holding Note/Controlling Pitch

The performer can hold the note/pitch he or she is singing by grasping the Voicon while placing her thumb or any part of the hands over force sensor 1 (FS1). When FS1 is pressed the audio signal that is currently going into the Pure Data patch is captured and played until the performer releases FS1. While the note is being held, the performer can control the pitch of his/her voice by tilting the Voicon up or down. The degree of tilting motion is proportional to the degree of the pitch modification. Through the use of pitch control the performer can sing higher or lower notes beyond his/her vocal range. Being able to sing in a wider range of pitch provides the performer with richer expressions. Moreover, the performer can achieve visually engaging performances while controlling pitch as he/she has the ability to control his/her pitch while leaning forward and backward. Figure 1 illustrates the motions used with Voicon.

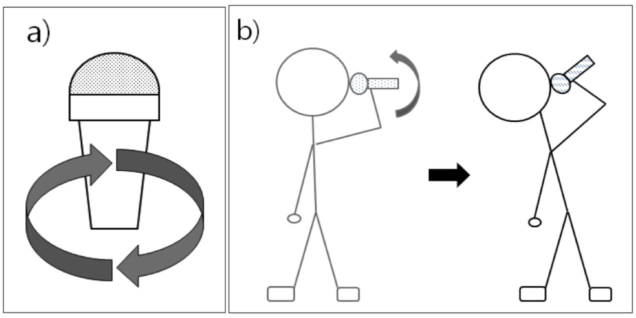

Figure 1: a)Circular motion of Voicon generates vibrato. b) Vocal performer can tilt the Voicon upward/downard to raise/lower the pitch.

3.3 Activating Customizable Vocal Effects

A variety of vocal effects are often used during live vocal performances. When using vocal effects, the timing and intensity of the effects need to be considered. Conventionally, a sound engineer and a vocal performer need to agree upon both timing and intensity of the effect. By giving the performer the ability to handle vocal effects freely, the performance will become more dynamic and expressive. A performer can activate pre-programmed vocal effects by applying pressure over force sensor 2 (FS2). The intensity of the effect is proportional to the amount of pressure applied on FS2. By establishing a co-relation between the amounts of pressure the performer applies to Voicon and intensity of vocal effect Voicon sends, controlling vocal effects through Voicon comes very natural and intuitive to its user.

Table 1: Gesture mapping to parameters of function

Gestures	Activated functions	Parameters mappings of functions
Circular motion	Generating vibrato	Radius and rate of circular motion
Tilting motion	Pitch control	Angle of inclination
Applying pressure	Activating vocal effect	Amount of pressure

4. TECHNICAL DETAIL

The Voicon system can be largely divided into microphone and PC. In the microphone, a tri-axial accelerometer(TAA) and force sensors are embedded. Obtained audio signals using the microphone are sent to a PC through Arduino Nano. Arduino Nano is a small-sized microcontroller for open-source electronics prototyping. The tri-axial accelerometer is used for measuring the acceleration of the movement in order to detect circular motions and the degree of tilting motions performed by the performer which lets the performer control vibrato and pitch, respectively. Two force sensors are used as activation switches for the holding note/controlling pitch function and the customizable vocal effect function. The force sensors were used to provide the performer room for more robust expressions of effects and natural interactions with the Voicon. The force sensors are in the form of thin film enabling it to be mounted seamlessly on the surface of Voicon. This physical feature of force sensors makes the hardware of Voicon more compact and simple. Figure 2 illustrates hardware system overview.

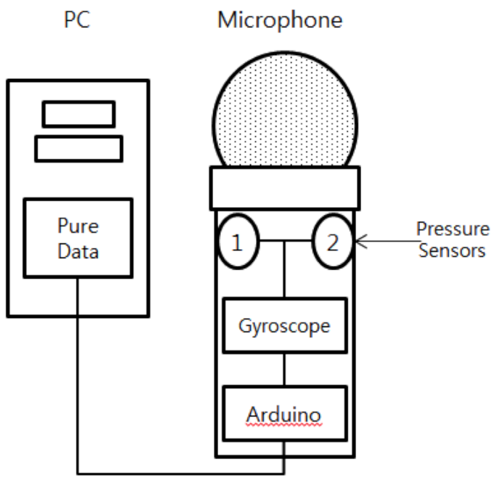

Figure 2: Hardware system overview.

4.1 Hardware System

Regarding its compactness and efficiency in real performances, our design was based on the idea that all components should be equipped on a single microphone body while not deviating so far from its original appearance. As there are enough space for the small components, installing additional components to a microphone will not affect acoustical performance of the microphone itself. Figure 3 shows the TAA mounted on Arduino Nano. Arduino Nano was selected due to its small size and it receives three kinds of acceleration data from the TAA and also transfers the data to a PC via a serial USB cable. Besides, the voice is independently sent to a PC by the original microphone cable in this prototype design. Figure 4 depicts the internal components before the final assembly.

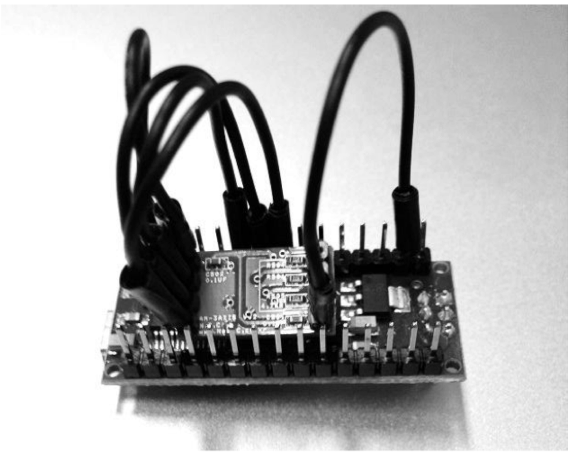

Figure 3: Arduino Nano with the tri-axial accelerometer.

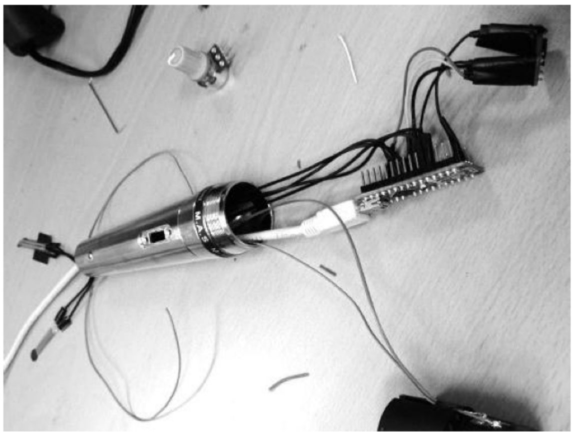

Figure 4: The internal components of Voicon.

4.2 Software System

Sensor data such as movement, tilt angle and pressure are received and processed by a program called Pure Data (PD) to perform audio signal processing. PD is an open-source real-time graphical programming environment for audio processing. PD was chosen as a sound-processing engine of the Voicon system as it is very flexible and simple, yet, it offers powerful DSP functionality. The PD patches were designed to perform each functions of Voicon. Figure 5 shows schematic diagram including the software components and the hardware components.

4.2.1 Generating Vibrato

When the performer shakes Voicon in circular motions to generate vibrato, the audio signal that is currently coming into the Pure Data patch is handled by the internal acceleration value of the TAA embedded in Voicon. These internal values such as acceleration and slope are calculated considering the initial values of sensors and then calibrated. As Voicon should not respond to the unintentional motions of the performer, a control module is designed to prevent the Voicon from responding to the unintended and irregular motions by applying a threshold value. In other words, vocal vibrato is a small oscillation in two features, which are amplitude and frequency. It is generated using a pre-defined modulator function that generates small variations in both features to express a natural vocal vibrato.

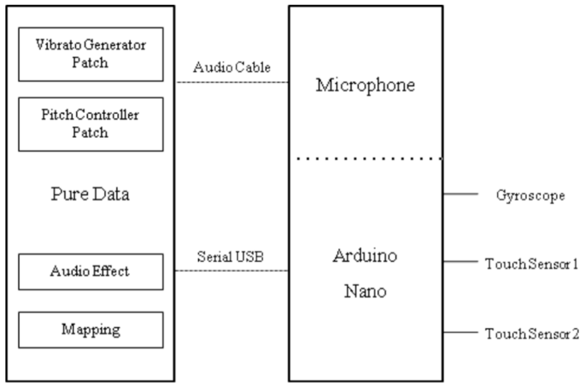

Figure 5: The internal components of Voicon.

4.2.2 Holding Note/ Controlling Pitch

At the instance of the activation of holding note/controlling pitch function, the vocal input from the previous one second is saved into a buffer of 44100 samples. This buffer is played in a loop until FS1 is released. For the pitch control, a performer can assign their own musical scales and the degree of the pitch modulation responding to degree of tilt motion from the Voicon. Pure Data patches are programmed to recognize the degree of tilting motion and increase or decrease the pitch of the performer‘s voice depending on the corresponding pitch change of 8-step musical notes. Pitch modulation and vibrato generation are processed through a short time transient process to prevent Voicon from producing unnecessary noise caused by a rapid change of audio signal. Parameter mapping between tilting angle and change in pitch is shown in Table 2.

Table 2: Tilting angle and corresponding change in pitch

Tilting Degree	Change in Pithch [semitone]	Tilting Degree	Change in Pitch [semitone]
-20	-2	30	5
-0	0	40	7
-10	2	50	9
20	4	60	11

4.2.3 Vocal Effects

The performer can also choose the amount of audio effects that he/she has assigned in advance by controlling the pressure applied on FS2. The assigned vocal effects are added to the original sound through the local buffer line. Pressure level on the corresponding sensor, which is sent to a PC, decides the amount of the effect. In a similar way to generating vibrato, the performer should put pressure on the sensor over a certain threshold level to avoid any unnecessary mishandling.

5. CONCLUSION

We have developed an interactive gestural microphone capable of generating vibrato, holding note/controlling pitch and activating vocal effects through the simple and intuitive gestures. Voicon enriches expression of voice of the performers by enabling them to manipulate their voice while maintaining the form of standalone microphone. The design of Voicon, much similar to a conventional microphone, makes it very approachable by the performers. All in all, an interactive gestural microphone was developed that allows explorable possibilities and high control level of the vocal performance.

6. FUTURE WORK

For the future work, Voicon should be developed to give tactile and visual feedback on audio augmentation and modulation it has created. Although the performer will be able to monitor the vocal effects, giving additional tactile and visual feedback apart from audio feedback will enable the performer to control vocal effects more efficiently. When applying vocal effects, tactile feedback should be designed to inform the performer using mild vibrations. The intensity of the vibrations should be mapped with the intensity of the created vocal effects. While controlling pitch, short vibrations between each change of pitch can inform the performer with the number of pitches he/she has modulated. Additionally, Voicon could be developed into the personal vocal guide using pitch detection. Voicon could alarm the performer with the tactile feedback when he/she deviates unintentionally from the appropriate melody. [7][5][4][3]

7. REFERENCES

[1] Joseph Butch Rovan, Marcelo M. Wanderley, Shlomo Dubnov, and Philippe Depalle. Instrumental gestural mapping strategies as expressivity determinants in computer music performance. In *Proceedings of KANSEI - the Technology of Emotion Workshop*, pages 3–4, 1997.

[2] Sylviane Sapir. Gestural control of digital audio environments. *Journal of New Music Research*, 31(2):119–129, 2002.

[3] C. Cadoz, A. Luciani, and J. Florens. Responsive input devices and sound synthesis by simulation of instrumental mechanisms: The cordis system. *Computer Music Journal*, 8(3):60–73, 1984. Cited By (since 1996): 17.

[4] A. J. Bongers. Tactual display of sound properties in electronic musical instruments. *Displays*, 18(3):129–133, 1998. Cited By (since 1996): 2.

[5] Bert Bongers. Physical interfaces in the electronic arts. interaction theory and interfacing techniques for real-time performance. In *Trends in Gestural Control of*, pages 41–70, 2000.

[6] L. Sonami. "lady's glve". [Web Site], 12 2008.

[7] M Bromwitch and J Wilson. A sensor suit and vocal performance mechanism for real-time performance. *Proceedings of 1998 International Computer Music Conference*, pages 292–295, 1998.

[8] Elena Jessop. The vocal augmentation and manipulation prosthesis (vamp): A conducting-based gestural controller for vocal performance. In *Proceedings of the 2009 conference on New interfaces for musical expression*, NIME '09, pages 256–259, 2009.

[9] Donna Hewitt and Ian Stevenson. E-mic: extended mic-stand interface controller. In *Proceedings of the 2003 conference on New interfaces for musical expression*, NIME '03, pages 122–128, Singapore, Singapore, 2003. National University of Singapore.

[10] C. E. Seashore. *The Vibrato*. University of Iowa, 1932.

[11] Alf Bjørklund. Analyses of soprano voices. *The Journal of the Acoustical Society of America*, 33(5):575–582, 1961.

[12] J Sundberg T Shipp, R Leanderson. *Some acousticcharacteristics of vocal vibrato*. J Res Sing, 1980.