

---

# Human Coded Orchestra: a System for Extemporaneous Group Singing Performance

**Yuzu Saijo**

The University of Tokyo  
3-8-1 Komaba, Meguro-ku,  
Tokyo, Japan  
yuzu0125@gmail.com

**Kenta Suzuki**

*\*Joint first authorship*  
University of Tsukuba  
1-2 Kasuga, Tsukuba, Ibaraki,  
Japan  
Ikuzus.atanek.0626@gmail.com

**Nobutaka Ito**

The University of Tokyo  
3-8-1 Komaba, Meguro-ku,  
Tokyo, Japan  
earth.nobu.light@gmail.com

**Amy Koike**

University of Tsukuba  
1-2 Kasuga, Tsukuba, Ibaraki,  
Japan  
amy23kik@gmail.com

**Yoichi Ochiai**

University of Tsukuba  
1-2 Kasuga, Tsukuba, Ibaraki,  
Japan  
wizard@slis.tsukuba.ac.jp

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).  
ACE2016, November 09-12, 2016, Osaka, Japan  
ACM 978-1-4503-4773-0/16/11.  
<http://dx.doi.org/10.1145/3001773.3001811>

**Abstract**

We propose Human Coded Orchestra, a new approach to enable a group of individuals to sing in harmony by using computed directional speakers. The possibilities of musical performance by an untrained group have been explored in the fields of science and art. However, previous work has rarely proceeded beyond simple rhythm-based music and failed to achieve musical complexity. Human Coded Orchestra employs a number of directional speakers, each set at a different pitch, enabling them to deliver different pitches to each participant to sing to, according to their positions. Experiments demonstrated that participants succeeded in singing in harmony extemporaneously, and they reported that they enjoyed both the experience of singing and the feeling that they were able to participate in an activity with others. Our system does not require preparation on the part of singers, which opens up the possibility of practical application in the area of interactive performance.

**Author Keywords**

musical performance; directional speaker; singing.

**ACM Classification Keywords**

H.5.3 [Information interfaces and presentation (e.g., HCI)]: Group and Organization Interfaces

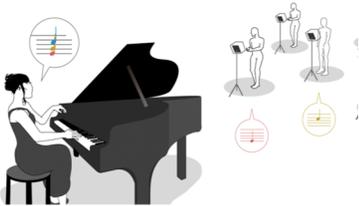


Figure 1: The ideal use of Human Coded Orchestra.

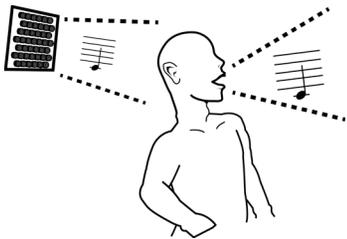


Figure 2: The core idea of Human Coded Orchestra is to place a directional speaker towards a person, and the person vocally imitates the pitch he or she hears.

## Introduction

Continuous efforts to advance ubiquitous computing, together with the development of small devices, such as it is easier than ever for audiences to become involved during stage performances. The audience no longer exists passively by simply remaining seated and watching the performance quietly. In fact, members of the audience are increasingly becoming active components of the stage.

Advanced interactions between artists and the audience can be observed especially in the music world. A British rock band, Coldplay, introduced an interactive wristband called Xylobands [12] to their concerts. Xylobands are LED-embedded wristbands synchronized with the live music performance. This enabled the band to involve their audience as a part of their stage design. A mobile app named Wham Light City [11] also exemplifies audience participation in music performance. At concerts, the app receives an ultrasonic data signal from the speakers, preprogrammed to be synchronized with live music. This app enables the audience to participate in the stage design. Considering these examples, it is clear that the boundary between musicians and audiences has become blurred, and on various occasions, has actually been remade.

The aim of the present project is to advance this boundary even further. It goes beyond merely involving the audience in the visual design of the stage; rather, it aims to draw the audience into the process of creating the musical performance, i.e., to encourage the audience to actually sing during the performance. Enabling audience members to contribute to a musical performance using their own body is likely to make them feel more connected and involved.

In this paper, we present Human Coded Orchestra, a computed system that allows a group of people to sing a complex song in harmony without much practice

beforehand (Figure 1). Its core idea is to create a controlled environment in which participants are directed to utter sounds that are produced by computed directional speakers.

The present work is inspired by the idea of human direction. Haptic Turk [5] directed users by providing them with computed visual cues, and this approach seems to be well suited to producing a complex output from untrained individuals.

Following this work, our system computes the environments surrounding participants using directional speakers instead of requiring them to learn complicated rules. Human Coded Orchestra expects participants to perform only one simple action, namely to sing to the pitch they hear. By providing only auditory cues, the system has simplified the rule (Figure 2). It is designed not only to produce a complex musical output but also to be universal and expandable.

## Related Works

Musical interfaces for people without prior in-depth training have been explored in the context of both art and science for years.

Recent studies attempted to enable multiple users in the same location to improvise together using technology. Musical interfaces for group improvisation are exemplified by interactive systems such as Jam-O-Drum [10] or Fire Flies [3], which give users control of simple rhythm, or COOL [9], which employs a single, shared musical instrument to be played by multiple users around it. More recent attempts include Polymetros [1], which also shares one musical instrument with multiple users.

Group improvisation has attracted attention in the world of art. In the performing arts, a work called The Sine Wave Orchestra [4] allows users to gather with their mobile devices to play a sine wave with which

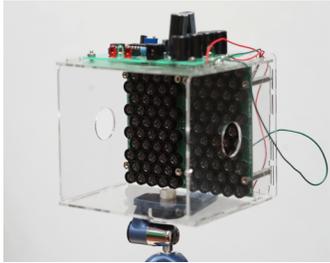


Figure 3: Directional speakers are made of 100 ultrasonic transducers, each 10 mm in a diameter. The transducers emit super directional audio signals, which are transmitted by ultrasound. Input sound signals come through a stereo cable from a computer, and get modulated in an IC in the speaker before being emitted.

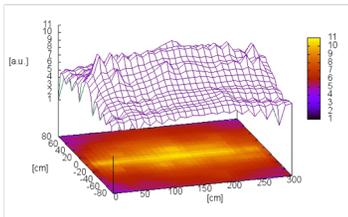


Figure 4: We measured sound pressure at each point using one speaker and a microphone. This indicates that if directional speakers are placed 1m apart from two participants of Human Coded Orchestra, the distance between the two should be about 1m so that they both hear clearly from one speaker.

they would like to harmonize. The approaches presented above can help users participate in a social activity and share the joy of creation. Group improvisation, however, definitely requires certain musical skills if the players expect to create genuine, harmonious music, rather than random noise. Thus, it would be difficult for musical beginners to improvise together even with the systems mentioned above. Furthermore, group improvisation is likely to fail to achieve both rhythm and melodic complexity and to result in a mere simple mix of sounds.

In another contribution to the performing arts, Matarisama [6], each performer holds one of a pair of percussion instruments in each hand. They sit in a circle facing the same direction and place their hands on the shoulders of the person in front of them. Each participant plays one of the musical instruments by tapping the shoulder of the person in front whenever the person behind them did the same thing. The instruments they play must be on the side corresponding to the shoulder, i.e., a player who was patted on the right shoulder is expected to hold their instrument in the right hand when tapping the next person. The performers were told to switch hands each time to ensure that the music would have variety.

As another example, in his concerts American jazz singer Bobby McFerrin implemented something closest to what Human Coded Orchestra aims to achieve [2]. He jumped to the right and left and sang different pitches according to his relative location, and asked the audience to sing along at the specified pitch. By jumping around, he provided the audience with visual cues of the pitches to sing to and succeeded in unifying the large audience. The Hub [8] also intended to encourage the audience to perform, by utilizing trackballs and MIDI keyboards. However, the scale it was able to produce was only one to six, and much practice on the side of the audience was required before an actual performance.

The work mentioned above certainly succeeded in enabling an unpracticed group of individuals to create the intended music. However, the music these works produced lacked the complexity of the output performances. In Matarisama, the musical performance resulted in monotonous rhythm-based music with only two different sounds. McFerrin's performance only allowed the audience to sing one note together at a time and did not achieve harmonious music by fully utilizing the large number of singing participants.

Human Coded Orchestra utilizes computation to solve this problem. The system uses directional speakers controlled by computers in an attempt to create music with enhanced complexity.

### Implementation

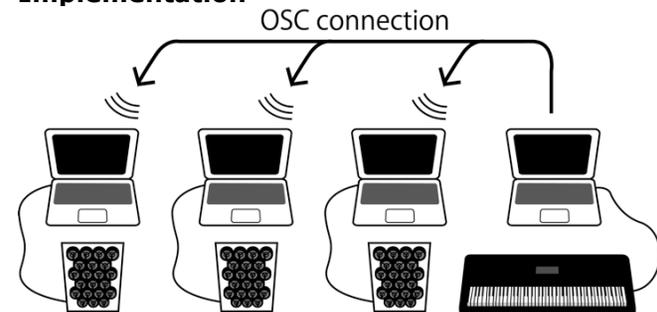


Figure 5: In Human Coded Orchestra, the main computer that is connected to a MIDI keyboard sends sound signals to other computers through OSC connection. The directional speakers emit sound based on the signal sent to the computer.

Human Coded Orchestra is composed of multiple directional speakers (Figure 3 and Figure 4), computers, and a MIDI keyboard (see Figure 5 for connections and communication among the components).

### Experiment

We conducted an experiment to verify that Human

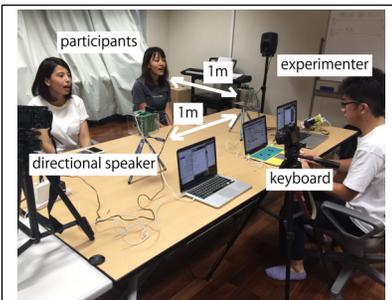


Figure 6: In the experiments, participants sat in front of the directional speakers, 1m away from them. When multiple speakers were employed, they were placed 1m away from each other.

|   | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|----|
| A | × | × | × | × | × | × | × | × | × | ×  |
| B | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ×  |
| C | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ×  |
| D | × | ○ | × | × | × | × | × | × | × | ×  |
| E | ○ | × | × | × | × | × | × | × | × | ×  |
| F | × | × | × | × | × | ○ | × | × | × | ×  |
| G | ○ | ○ | × | × | × | × | ○ | ○ | ○ | ○  |
| H | ○ | ○ | × | × | × | × | × | × | × | ×  |
| I | × | × | × | × | × | × | ○ | × | × | ○  |

Table 1: The vertical axis, A to I, represents each group of two singers. The horizontal axis, 1 to 10, stands for each trial. Successful trials are represented as ○, and unsuccessful trials as × in this table.

**Pilot Study:** In the pilot study, two participants each time were asked to sing in harmony with a normal speaker. In each trial, participants heard a two-note chord sequence. (Figure 6 and Table 1)

Coded Orchestra enables untrained participants to sing in harmony without much practice (Figure 6). Forty non-musicians (14 females) were recruited for the study. Their ages ranged from 15 to 44 ( $M=23.5$ ,  $SD=7.5$ ). Six of them had undergone formal singing training.

The experiments were conducted over two days (Figure 7). Participants were asked to enter the room ( $5m \times 10m$ ) and to randomly choose experiments in which to participate. They all signed an agreement and completed a questionnaire about their musical background. After conducting a pilot study, the experiential study was conducted with 24 people. In this study three participants were randomly selected to evaluate the musical performance using Human Coded Orchestra. Afterwards, each of the two groups (the three participants and the remaining 21 participants) was asked to fill out separate evaluation forms.

In all the experiments, all the musical parts were recorded separately using iPhones. We subjected the recorded data to Fourier analysis and visualized the results using a desktop application named Sound View. We visually inspected the analyzed data and calculated the accuracy and latency. In the pilot study, approximately 15 notes were played from each speaker, which took about 1 minute in total each time. In the evaluation process, 10-note sequences were selected from each audio recording where sound waves were clearly observed visually to enable the latency and accuracy of human voices to be analyzed. The greater importance of absolute pitch compared to relative pitch for group performance was clear, as stated in [7]. Since Human Coded Orchestra is designed for group performance, only absolute pitch errors are analyzed in this study.

### Experiential Study

The experiential study was conducted to evaluate the actual experience of singing to a song and the

performance quality using Human Coded Orchestra. Three randomly selected participants used Human Coded Orchestra to evaluate the group singing performance. They were asked to sit against the speakers so that they could only hear the participants singing, rather than the sounds from the speakers. The remaining participants were divided into three groups of seven each, and sat in front of the assigned speaker and sang to the part they heard. Before the actual performance, the singers had a two-part practice session, limited to the necessary minimum. First, each group listened to their part and practiced singing together as a group. Second, all the groups sang several passages of their parts at the same time. The duration of the practice session was roughly 1 minute.

### Results and Observations

The results obtained via the questionnaires were positive. On average, participants rated the joy of their singing experience as 7.4 on a 10-point scale ( $SD=1.8$ ), and rated the feeling of participation in the performance as 7.6 on a 10-point scale ( $SD=2.0$ ). One participant stated that he enjoyed the singing experience using Human Coded Orchestra especially when he realized that he was singing a different part from others. Another participant commented that it was surprising how she was able to sing in harmony merely by imitating the sounds from Human Coded Orchestra. The most likely reason for the lower enjoyment score was the low volume of the speakers because six participants mentioned the difficulty of hearing the sounds from the directional speakers. One participant stated that he would have experienced more joy listening to the harmony generated using Human Coded Orchestra if it had been easier to hear the sounds.

The three audience members in the experiential study, i.e., those who only listened to the performance, all rated the quality of the performance as 9 on a 10-point scale, which is clearly high. Two of the audience members commented that the harmony was clearer



Figure 7: These two photos are from the experience study where 21 participants sang in harmony using Human Coded Orchestra.

**Experience Study:** In this experience study, the participants and the authors played "Let it be" by the Beatles together. The participants sang the guitar part of the song, which was divided into three different parts, using Human Coded Orchestra. One of the authors played the MIDI keyboard to control the sounds from the directional speakers, another author played the guitar, and the other author sang the lyrics as the main vocal.

than they had expected. One member stated that considering that the singers had not practiced before, their performance using Human Coded Orchestra was admirable. In summary, the experiential study proved that the singers enjoyed the actual singing experience using Human Coded Orchestra, although the joy can be enhanced by increasing the sound emitted by the directional speakers. In addition, our work proved that untrained singers were able to sing in harmony successfully.

### Discussion

One of the main goals of this work was to attempt to invent a new musical interface for a group of individuals without extensive practice. The accuracy achieved by the experiments was promising as mentioned above.

With this positive result, it could be safely stated that Human Coded Orchestra succeeded either in generating harmonious music and creating a situation as if each participant were a musical instrument, or in orchestrating him or her. Our system also achieved the production of music of a more intricate nature than attempts by others. In the experimental study, the Human Coded Orchestra prototype was able to generate 12 different pitches to play three notes at most to create harmony. This can be seen as a considerable improvement in this field, considering that Matarisama created only two types of sounds with monotonous rhythm, and McFerrin could produce about 10 different pitches with various rhythms but failed to achieve harmony using different pitches simultaneously.

The experiments involving Human Coded Orchestra also proved that its participants had felt that they had been part of the performance, which is another important aspect of the present study. This ensures that application of this system to actual musical performances could be expected to change the relationship between the artists on stage and the audience fundamentally. It can provide the audience

with a stronger feeling that they do not merely exist passively but form an essential part of the creative process of the musical performance.

Apart from inventing a new musical interface for an untrained group, another contribution of the present work lies in determining the characteristics of group singing as a musical instrument: the extent to which the latency and accuracy change according to conditions such as the number of singers, the number of parts, and the difference in pitch. Considering these findings, the authors have succeeded in creating programmed directional speakers that function as a sequencer. The speakers play a sequence perfectly to enable latency and accuracy to be taken into account based on all the pitch differences in the sequence, the number of parts, and the number of individuals present. This would only be achievable with computational technologies and our findings from the experiments.

### Applications

A possible application of Human Coded Orchestra is interactive stage performance, as mentioned in the Introduction (Figure 8). In addition to positive ratings as to the joy of the singing experience using Human Coded Orchestra, the experiential study showed that only approximately 1 minute was required before a song could be sung in harmony. This indicates that Human Coded Orchestra would be effective if applied to live musical scenes to involve the audience in real time. For instance, musical artists would be able to play a song together with the audience, who could work as an additional musical instrument as a whole, by using Human Coded Orchestra (Figure 9). Considering the results of our experimental evaluation, an application such as this presents the possibility of enhancing interaction during a musical performance in an exciting and innovative way.



Figure 8: User study of our system. Participants sing a song successfully without much practice beforehand.

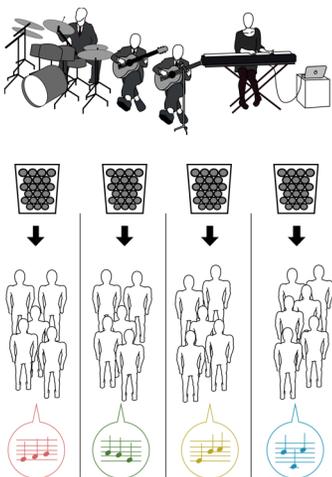


Figure 9: Human Coded Orchestra can help artists on stage to involve their audience by making them sing in harmony and participate in the live music.

## Conclusion and Future Work

We proposed Human Coded Orchestra to achieve non-improvisational musical performance by groups instantaneously without any thorough practice beforehand. The experimental results show that singing in harmony using the present system entertained the participants and made them feel part of the performance. Considering the results, the proposed system has the possibility to enhance the level of interaction in the music world. Future work aims to conduct additional experiments with a larger number of participants in an actual concert hall to envisage future applications.

## References

1. Ben Bengler, and Nick Bryan-Kinns. 2013. Designing collaborative musical experiences for broad audiences. In *Proceedings of the 9th ACM Conference on Creativity & Cognition (C&C '13)*, Ellen Yi-Luen Do, Steven Dow, Jack Ox, Steve Smith, Kazushi Nishimoto, and Chek Tien Tan (Eds.). ACM, New York, NY, USA, 234-242. DOI=<http://dx.doi.org/10.1145/2466627.2466633>
2. Bobby McFerrin. 2009. Watch me play ... the audience! Retrieved October 8, 2016 from [https://www.ted.com/talks/bobby\\_mcferrin\\_hacks\\_your\\_brain\\_with\\_music](https://www.ted.com/talks/bobby_mcferrin_hacks_your_brain_with_music)
3. Gili Weinberg, T Lackner, and Jason Jay. 2000. The Musical Fireflies-learning about mathematical patterns in music through expression and play. *Proceedings of Colloquium on Musical Informatics Laquila, Italy*, 146-149
4. Ken Furudate, Jo kazuhiro, and Daisuke Ishida. 2002. The SINE WAVE ORCHESTRA. Retrieved October 8, 2016 from <http://swo.jp/yebizo2016/>
5. Lung-Pan Cheng, Patrick Lühne, Pedro Lopes, Christoph Sterz, and Patrick Baudisch. 2014. Haptic turk: a motion platform based on people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3463-3472. DOI=<http://dx.doi.org/10.1145/2556288.2557101>
6. Masahiro Miwa. 2002. The MATARISAMA. Retrieved October 8, 2016 from <http://www.iamas.ac.jp/~mmiwa/XORensemble.html>
7. Pfordresher Peter, and Mantell James. 2009. Singing as a form of vocal imitation - Mechanisms and deficits. In *Proceedings of 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM'09)*.
8. Scot Gresham-Lancaster. *The Aesthetics and History of the Hub: The Effects of Changing Technology on Network Computer Music*. 1998. *Leonardo Music Journal*. 8(1), 39-44.
9. Tim Hankins, David Merrill, and Jocelyn Robert. 2002. Circular optical object locator. In *Proceedings of the 2002 conference on New interfaces for musical expression (NIME '02)*, Eoin Brazil (Ed.). National University of Singapore, Singapore, 1-2. 425-430
10. Tina Blaine and Tim Perkis. 2000. The Jam-O-Drum interactive music system: a study in interaction design. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques (DIS '00)*, Daniel Boyarski and Wendy A. Kellogg (Eds.). ACM, New York, NY, USA, 165-173.
11. Wham City Lights, Inc. 2014. Wham City Lights. Retrieved October 8, 2016 from <http://www.whamcitylights.com/>
12. Xylobands. 2011. Xylobands. Retrieved October 8, 2016 from <http://www.whamcitylights.com/>