



Translating Expression in Taiko Performance

Shannon Cuykendall,¹ Michael Junokas,² Kyungho Lee,² Mohammad Amanzadeh²,
David Kim Tcheng,² Yawen Wang, Thecla Schiphorst,¹ Guy Garnett,² and Philippe Pasquier¹

Simon Fraser University¹ and University of Illinois²

Surrey, BC Canada¹ and Urbana-Champaign, IL United States²

scuykend@sfu.ca, junokas@illinois.edu, klee141@illinois.edu, amanzad2@illinois.edu,
dtcheng@illinois.edu, asianrainwerk@mac.com, thecla@sfu.ca, garnett@illinois.edu, pasquier@sfu.ca

Abstract

We describe our approach to collecting, analyzing and visualizing expressive movement data to support the creation of an interdisciplinary performance and installation work, *3 Movements in Translation*. We seek to understand how three perspectives (the performer, the audience, and the machine) can inform one another to create a cross-cultural performance that allows a broad audience to kinesthetically engage and empathize with expressive features of taiko performance. Taiko is a Japanese artistic practice that combines stylized movement with drumming technique. We share initial results in machine-learning analysis of taiko sound and movement. These results are used, in combination with a performer's perspective, to inform artistic visualizations of important expressive features within taiko. Through this process, we explore how multiple perspectives of taiko can inform the translation process between the performer, the audience and the machine.

Keywords

Taiko, Machine Learning, Dance, Choreography, Expressive Movement Recognition, Sound Recognition, Movement Classification, Musical Gesture, Interactive Installation, Kinesthetic Awareness, Visualization

Introduction

3 Movements in Translation is an interactive performance work that explores three perspectives of taiko performance: the performer, the audience and the machine. We seek to understand how these varying interpretations of taiko can inform one another in order to create a cross-cultural performance that allows a broad audience to kinesthetically empathize, appreciate, and understand taiko performance.

To support the creation of *3 Movements in Translation*, we formed an interdisciplinary team of researchers from Simon Fraser University, University of Illinois and an independent taiko artist. Each author focused on a different aspect of the project including machine learning, data capture, data visualization, movement analysis, choreography, and taiko drum technique. This work is still in progress, but is envisioned to be an interactive installation and performance. The installation will help expand the audience's kinesthetic awareness of taiko, while also influencing the outcome and perception of the performance.

In order to create this type of work we need to identify expressive components of taiko performance, find a method to extract these components from the data, and explore ways to create new representations of these expressive features so that they can be communicated in performance settings. According to Camurri et al (2004), "Expressive content concerns aspects related to feelings, moods, affect, intensity of emotional experience" and is not necessarily tied to a denotative meaning. This definition implies that expressive qualities can be found in any movement that contains information about the state of the mover—from everyday actions, such as walking, to actions executed in performance settings (Camurri et al. 2004).

While digital motion capture technology has helped generate an expanding corpus of visual movement features to analyze, insight into what these features mean in the context of expression has been challenging to isolate and recognize. In our research we examine how the artform of taiko is taught. Through examining taiko instruction, we can identify perceptually evident features within the technique that are important to its artistic expression and integrity.

The Japanese practice of taiko also provides an ideal form to analyze expressive components of performance because expressive movement and sound are explicitly integrated in taiko technique. This multi-modal coupling allows us to study expressive components through both sound and movement simultaneously, which can help inform the recognition process.

Overview of Taiko

Taiko, meaning "fat drum," most commonly refers to a Japanese form of ensemble drumming that became popularized as a performing art in the mid-20th century (Varian 2005). Taiko is a unique form because stylized, choreographed movement is highlighted and integrated with drumming technique. Stylistically correct sounds in taiko can only be produced with stylistically correct movement forms. Therefore, the accuracy of movement qualities can not only be *felt* and *seen*, but also *heard*.

Taiko is traditionally communicated through *kuchi-showa*, a phonetic vocal notation system that represents the timbre qualities of taiko sounds. There are six main sounds in taiko technique: *Don*, *Doro*, *Tsu*, *Tsuku*, *Ka* and *Kara*. *Don* is a loud sound, *Tsu* is a soft sound, and *Ka* is a rim sound. *Doro*, *Tsuku* and *Kara* refer to either 2 *Don*, *Tsu*, or *Ka*

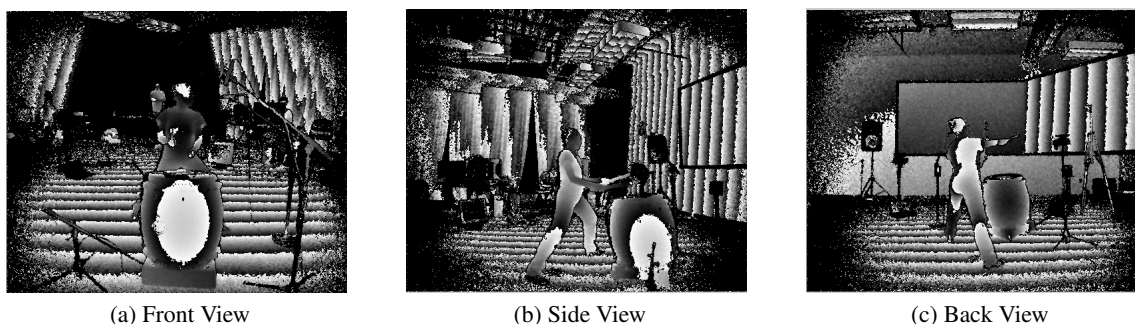


Figure 1: Kinect Sensor raw depth map data modulus 256

beats (respectively) played in succession. Jason Overy of Uzume Taiko in Vancouver British Columbia (Overy 2014) created a syllabus defining aspects of taiko that are traditionally demonstrated, but not verbally articulated in traditional taiko technique. In Overy’s syllabus all taiko sounds can be performed within five gears that refer to the body part or action initiating the sound, such as fingers, wrist, elbow, full-arm, and jumping (labeled Gear 1-5 respectively). Traditionally Don is played in high gears (e.g. Gear 4/full-arm and Gear 5/jumping) and Tsu is played in low gears (e.g. Gear 1/fingers and Gear 2/wrist). Within each gear Overy defines five levels that refer to the angle of the drumstick in relation to the drum. Level 1 is 0 degrees and level 5 is 90 degrees. This level system is inspired by techniques used in drum corps training.

Data Collection Process

Taiko Informed Data Collection

We used Overy’s syllabus to inform our data collection process. Using the gears and levels as a guide, we gathered a wide range of Don and Tsu strokes performed in similar gears and levels. We discuss initial results for distinguishing between Don and Tsu strokes performed in Gear 2 Level 4, Gear 3 Level 4, and Gear 4 Level 5. We chose to start with these particular combinations because they felt “natural” to the performer when playing both Don and Tsu sounds.

Motion Capture

In order to digitally capture movement, we used the Microsoft Kinect sensor to allow the taiko artist to perform as expressively as possible by providing an environment that limited technical intrusions into the performance. The Kinect captures movement through the generation of depth maps, utilizing a camera and an infrared sensor. From the depth maps, a variety of informative data can be extrapolated, such as a skeleton frame representing positional data of the subject. This provides subjects with a minimally invasive situation where they are not required to wear any additional devices that may inhibit their ability to replicate the physical qualities of performance. This is especially important in the case of taiko, where expressive performance demands high levels of control over a diverse range of nuanced percussion techniques. This control can be dramatically affected by extra

weight (e.g. an attached accelerometer) or atypical performance attire (e.g. a motion capture suit). The Kinect sensor is also more practical for the artwork and will allow us to present the work in a large variety of venues.

We arranged five Kinect version 2.0 sensors in a circle around the performer with one sensor directly in front of the performer, providing a panoramic view of the drummer (see Figure 1). All sensors were set at waist level except the front sensor, which was elevated to head level in order for the drumhead to be seen clearly.

Audio Capture

We captured audio from multiple perspectives in order to fully realize the expressive sounds produced by the taiko drum and performer. To accomplish this, we used an array of cardioid condenser microphones: one below the drum that captured the more resonant sounds of the drum; one above the drum that captured higher frequency attack sounds; and one off-drum to the right of the performer that captured vocalizations and the ambient room sound. The audio data captured by the array was sent through a multi-channel mixer where only minimal processing was done in the interest of maintaining sonic integrity.

Machine-Learning Analysis

To identify expressive qualities of Don and Tsu strokes in taiko, we performed analysis on the data using two different machine-learning techniques: a Naïve Bayes (NB) classifier (John and Langley 1995) and the Hidden Markov Model (HMM) (Rabiner 1989). HMM is a method for analysis of temporally sequential data and performed better than the NB classifier in distinguishing the visual data of taiko gestures. The HMM and NB classifiers performed similarly for the audio analysis. We report a summary of our results in Table 1.

Audio Analysis of Taiko Sounds

In our analysis of taiko drumming audio classes, we generated two separate sample types for NB testing: the full audio sample without normalization (Audio1) and a shorter, attack-focused sample that was normalized with a ‘room noise’ spectrogram (Audio2). For the Audio 1 test we used a window size of 512, with 256 overlap. For the Audio 2 test we

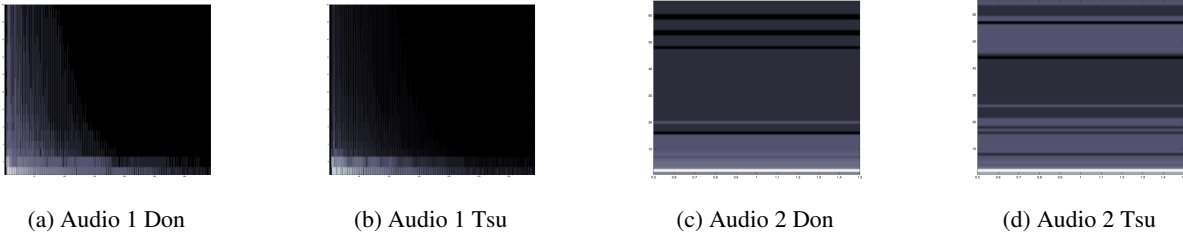


Figure 2: Plots of the square root of the magnitude of the spectrogram. Audio 1 depicts the entire duration of the sound (truncated to a range of 20hz). Audio 2 depicts a single window (only the attack).

Tests	Gear/Level	Audio1 (NB)	Audio2 (NB)	Visual (HMM)
Don/	2/4	1.00	1.00	0.97
Tsu	3/4	1.00	1.00	0.97
	4/5	1.00	1.00	1.00

Table 1: Initial results of audio and visual analysis of Don and Tsu strokes: a total of 242 samples for Audio (81 samples for 2/4, 82 samples for 3/4, and 79 samples for 4/5), and 20 samples for each gear/level combination of the visual data. Audio1 = 2s length, not normalized, Audio2 = 0.045s length, spectrally normalized, and visual samples are manually segmented (starting from raising the drumstick and ending up to 10 frames after the stroke).

used a window size of 2048 with no overlap (See Figure 2). Both audio tests used leave-one-out cross validation.

Both of our tests achieved 100 percent classification accuracy in distinguishing between Don and Tsu sounds played in different gears and levels. Furthermore, we found that even with amplitude measures minimized through normalization (Audio2), classification accuracy was not lost. This suggests that Don and Tsu sounds vary in both timbre and amplitude.

Visual Analysis of Taiko Movements

In order to analyze the style of movement when performing Don and Tsu strokes, we analyzed the average velocity of wrist movement during both the attack and recovery of the drum-strokes. Since there are only two major axes involved in performing the basic taiko drum-strokes, we analyzed the velocity in the Y-axis (vertical) and the Z-axis (camera). In the Kinect coordinate system the device is positioned at the origin, and the Z-axis represents the distance of the body from the Z-plane of the camera.

We found that the velocity of movement was the key feature in distinguishing between Don and Tsu drum-strokes. We used HMM to model and classify the wrist movement in Don and Tsu strokes. Each sample of the motion data is manually segmented, starting from raising the drumstick and ending up to 10 frames after the drum-stroke. For each gear/level combination we used 20 samples, with 70% of them for training and 30% for testing the model. We computed the accuracy through 10-fold cross validation (see Table 1) and used

various features of the motion data to train the HMMs such as velocity, acceleration, and speed. Our experiments showed excellent classification of Don and Tsu gestures.

Figure 3 compares the average velocity of the Don and Tsu strokes in Y and Z directions. On the Z-axis we observe positive average velocity when attacking the drum, and negative velocity after the drum-stroke. This suggests that both actions are performed by moving the arm towards the body. We can also observe that in comparison with the Tsu strokes, the Don strokes have lower velocity in the Y-axis and higher velocity in the Z-axis. We can interpret this as the drumstick in Don strokes tends to hit the drum tangent to the surface, while the Tsu strokes use more force in the vertical direction. These findings align with the taiko performer’s description, describing Don as having more “bounce” than Tsu and Tsu being more “controlled.” These findings also closely relate to previous work by Dahl and Altenmüller that found correlations between stroke velocity, rebound, and the perceived timbre quality of the sound produced (Dahl and Altenmüller 2008).

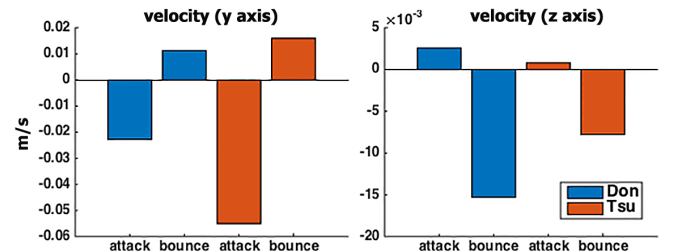


Figure 3: Comparison of the average velocity of performing the Don and Tsu sounds

Visualizing Expressive Features of Taiko

Through artistic visualizations we can mitigate the gaps between literature and artistic practices. In order to understand the expressive components of taiko from another perspective we created **visaphors**, meaning “visual metaphor” (Cox 2006).

The visaphors described below allow us to re-contextualize expressive components of taiko technique and will facilitate interactions between the audience and performer in *3 Movements in Translation*. Below we show a prototype of each visualization and describe how it was informed by var-

ious perspectives. A video of these visaphors is available at <https://vimeo.com/130595155>.

Visualizing Expressive Arm-Movements

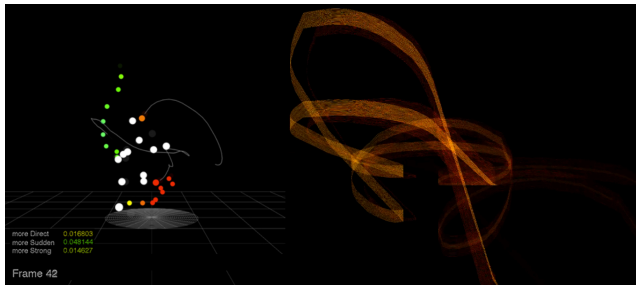


Figure 4: Visaphor 1: The trajectory of two wrists draw flaming ribbon strokes over time. ©Kyungho Lee 2015

The first *visaphor* was inspired by Kimberly Powell’s phenomenological account of practicing taiko depicted in her published work, “The Apprenticeship of Embodied Knowledge in a Taiko Drumming Ensemble.” Powell states, “I imagined colorful long ribbons trailing from my bachi [drumstick], as I snap my wrist upward and then extend my arm diagonally across my body down to an imaginary drum” (Powell 2004).

To depict this vibrant energy of taiko we used the trajectory of the performer’s right-wrist and the velocity to change the form of a flaming ribbon (See Figure 4). The flames expand the space of the arm’s trajectory, alluding to Powell’s sensation, “I felt a sensory shift in the way I perceive the boundaries of my body in relation to space. I am aware of a different sense of my body, the way it occupies positive space against negative space” (Powell 2004).

Visualizing Don and Tsu

The second *visaphor* depicts differences between Don and Tsu. As mentioned previously, Don and Tsu sounds are distinguished in taiko as loud and soft sounds respectively. However, the difference between Don and Tsu goes beyond amplitude as shown both in our machine-learning audio results and as implied by the phonetic vocalizations of taiko sounds. As indicated by the machine-learning HMM analysis, velocity also played an important role in distinguishing between Don and Tsu gestures.

We chose to visualize the difference between Don and Tsu through a spherical representation. The sphere’s surface is distorted by the momentum of each stroke. The more powerful the stroke is, the larger the displacement of the sphere and the darker the color. Figure 5 shows the Tsu stroke has less distortion than the Don stroke and is lighter in color.

Conclusion

We describe research undertaken to support the creation of our interactive performance and installation work *3 Movements in Translation*. In this work we seek to create a cross-cultural performance that translates expressive components of taiko performance into multiple forms that a broad audience can kinesthetically engage with and understand. We

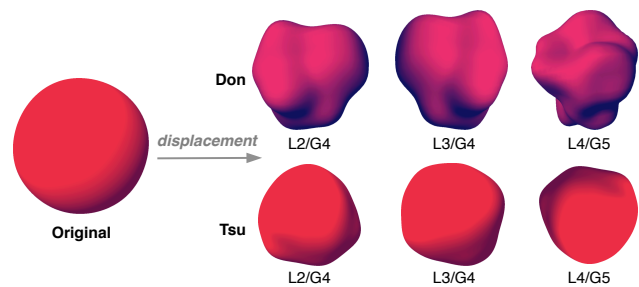


Figure 5: Visaphor 2: The original sphere’s surface is transformed by the parameters extrapolated from drumming strokes: Don and Tsu ©Kyungho Lee 2015

shared initial findings in machine-learning analysis and combined these findings with the performer’s perspective to create visaphors of important expressive features in taiko. We will use these visaphors in our interactive installation as a way to motivate audience members to partake in expressive movements that are important in taiko performance and to inform the translation process between the performer, audience and machine.

Acknowledgements

This research was supported by the Social Sciences and Humanities Research Council of Canada and the Canada Council for the Arts.

References

- Camurri, A.; Mazzarino, B.; Ricchetti, M.; Timmers, R.; and Volpe, G. 2004. Multimodal analysis of expressive gesture in music and dance performances. In *Gesture-based communication in human-computer interaction*. Springer. 20–39.
- Cox, D. 2006. Metaphoric mappings: The art of visualization. *Aesthetic computing* 89–114.
- Dahl, S., and Altenmüller, E. 2008. Motor control in drumming: Influence of movement pattern on contact force and sound characteristics. *Journal of the Acoustical Society of America* 123(5):3122.
- John, G. H., and Langley, P. 1995. Estimating continuous distributions in bayesian classifiers. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, 338–345. Morgan Kaufmann Publishers Inc.
- Overy, J. 2014. Uzume taiko. <http://www.uzume.com/>.
- Powell, K. 2004. The apprenticeship of embodied knowledge in a taiko drumming ensemble. In Bresler, L., ed., *Knowing Bodies, Moving Minds*, volume 3 of *Landscapes: the Arts, Aesthetics, and Education*. Springer Netherlands. 183–195.
- Rabiner, L. 1989. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE* 77(2):257–286.
- Varian, H. 2005. *The way of taiko*. Stone Bridge Press.