The Journal of the Society for Art and Science, Vol. 24, No. 2, pp. 7:1-7:10 (2025)

Detecting Hodo and Zu in Bunraku Nonverbal Communication using Hilbert-Huang Transform

Xunan LIU[1)] **Ran DONG**[2)] (Member) **Dongsheng CAI** [3)]

1) Graduate School of Science and Technology, University of Tsukuba

2) School of Engineering, Chukyo University

3) Faculty of Management, Nagoya University of Commerce and Business

liuxunan@cavelab.cs.tsukuba.ac.jp

Abstract

Nonverbal communication relies on human actions to convey information through signals. In Bunraku puppetry performance, a UNESCO Intangible Cultural Heritage, three puppeteers manipulate a single puppet using the rhythmic structure of "Hodo" and "Zu". The puppeteer master embeds cues, "Hodo" and "Zu", into the puppet's movements, guiding coordination. However, the mechanics of these cues remain unclear. Despite this, shared gestures and rhythms enable seamless collaboration. This paper proposes a signal analysis framework using Principal Component Analysis and Hilbert-Huang Transform to investigate how puppeteers achieve nonverbal communication, synchronizing movements and rhythmic variations to manipulate the puppet with natural fluidity. The correlation analysis using Procrustes Analysis, Mutual Information, and Canonical Correlation Analysis, along with the examples identified through this method, all meet statistical significance across these three validation approaches, confirming the effectiveness of our method.

Keywords: non-verbal cooperation, Bunraku, Principal Component Analysis, Hilbert-Huang Transform

1. Introduction

Communication is the act of conveying information through signals, cues, and other means. Non-verbal communication, by coordinating cues, suggests a structural hierarchy and priority among the participants in the communication, facilitates the flow of the interaction, and provides communication and meta-feedback [1].

Nonverbal communication is also utilized in various artistic performances, such as theater and dance. This paper focuses on nonverbal communication in Bunraku (文楽), a traditional Japanese performing art. Bunraku, also known as Ningyo Joruri (人形浄瑠璃), is one of the world's most beautiful forms of puppetry and has been inscribed on UNESCO's Intangible Cultural Heritage list. In Bunraku, three puppeteers manipulate a puppet using special cues called "Hodo" and "Zu". These cues serve as a nonverbal means of communication that enables the three puppeteers to control the puppet's movements, yet many aspects of their function remain unclear [2].

There are numerous studies on Bunraku. For example, [3] and [4] investigate the relationship between synchronization and asynchronization of breathing and movement in Bunraku puppeteers. [5] examines the coordination mechanisms among Bunraku puppeteers. [6], [7], and [8] focus on the cooperative movements of Bunraku puppets. [9], [10], and [11] analyze the relationship between the phase and amplitude of the motion axes of Bunraku puppets. [12] and [13] study the mechanism of "Jo-Ha-Kyu (序破急)" in Bunraku. [14], [15], and [16] explore the application of the "Jo-Ha-Kyu" mechanism in robotics.

However, none of these studies have sufficiently analyzed the nonverbal communication among the three Bunraku puppeteers. In particular, there has been no analysis of "Hodo" and "Zu" in the nonlinear frequency domain of puppet motion. Additionally, "Hodo" and "Zu" are embedded within the improvisational rhythmic variations of "Jo-Ha-Kyu", causing the puppet motion that follows these rhythm changes to become a nonlinear signal. To analyze these nonlinear movements of "Hodo" and "Zu", this paper employs Principal Component Analysis and Hilbert-Huang Transform.

In this paper, we analyze nonverbal communication in Bunraku puppet manipulation by applying Principal Component Analysis and the Hilbert-Huang Transform to motion-captured puppet motion data. This method extracts features in the nonlinear frequency domain of the puppet's motions. By comparing the extracted features with the rhythms of the shamisen (三味線) and Gidayu (義太夫), we discuss insights and new findings regarding "Hodo" and "Zu".

1.1 The structure of Puppet

Much of the theatrical expression in Bunraku, a traditional Japanese puppet theater, is one of the typical forms of nonverbal communication [2]. Bunraku consists of three elements: the shamisen, the narrator (Gidayu), and the puppet itself.

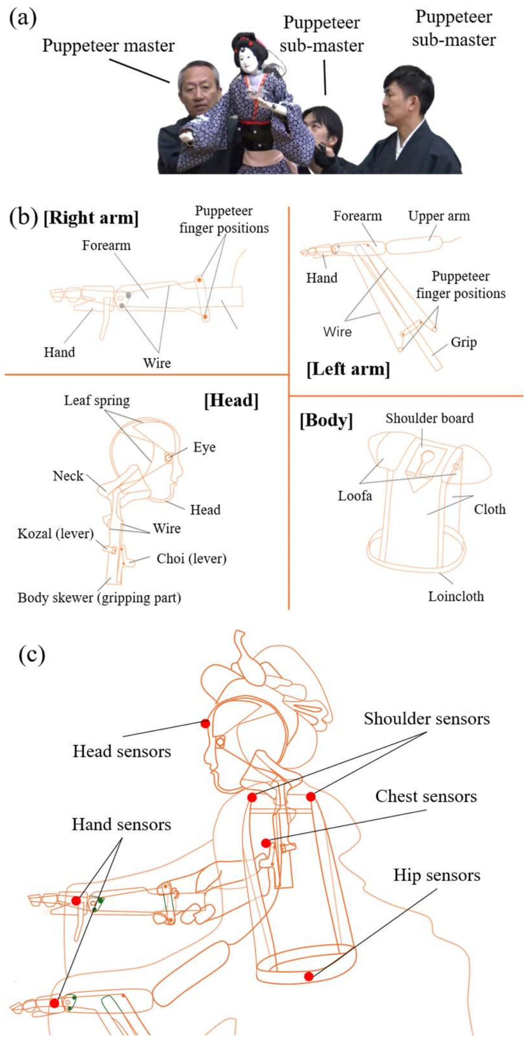


Figure 1. (a) One bunraku puppet. One puppet is operated by one puppeteer master, one left puppeteer, and one foot puppeteer[1]. (b) Structure of a bunraku puppet. The puppet consists of four parts: a right arm, a left arm, a head, and a body. (c) Sensor positions of one bunraku puppet in this experiment (head, left shoulder, right shoulder, left arm, right arm, and chest) [2].

The performance is carried out by a shamisen player who plays the music, a narrator who provides the narration, and three puppeteers who manipulate the special puppets. As shown in Figure 1(a), a single puppet is operated by a main puppeteer, a left-hand puppeteer, and a foot puppeteer. As shown in Figures 1(b) and (c), the puppeteer master controls the head and the right hand, the left-hand puppeteer controls the left hand, and the foot puppeteer controls the feet. The puppeteers grasp the handles of the puppet and manipulate its various parts.

1.2 "Hodo" and "Zu"

The following are the basic movements of "Hodo" and "Zu", which have been partially clarified in studies of Bunraku puppet theater by Ueda and others [2]:

The Shamisen and Gidayu control the puppeteer's movements by giving a slow and rapid rhythm.

Each of the three puppeteers manipulates a different part of the puppet by using the handle of the puppet. While one puppeteer manipulates the movement of one puppet, the other two puppeteers are aware of the nonverbal cues that are given in response to the movements of the main puppeteer. These cues are "Hodo" and "Zu".

The overall movements and actions of the puppets can be categorized into approximately 140 different "kata (型)", or short coordinated movements. The puppeteer master uses "Hodo" to signal the start of coordinated movements, and "Zu" to direct the movements of the left puppeteer and foot puppeteer.

The term "Hodo" refers to the subtle signaling motion made by the master puppeteer using their left hand to guide the puppet movement (back and forth, left and right, or up and down), typically timed with the backbeat of the Gidayu. While the puppet position remains fixed in front of the audience, this signal initiates a directional shift. The puppet then follows the indicated direction with a neck rotation, known as "Zu".

However, some principal puppeteers, especially those who are called "masters", give instructions to the left puppeteer only with "Hodo". Here, "Hodo" is defined as a signal by the principal puppeteer to instruct the left puppeteer to start moving the left hand by shifting the puppet back and forth, left and right, up and down with the left arm.

1.3 Puppet Motion Capture and Retargeting

The puppet used for the motion capture in this paper was not a standard Bunraku puppet without joints, but a special Bunraku puppet with joints, as shown in Figure 2, prepared by Living National Treasure Kanjuro Kiritake (桐竹勘十郎). This puppet features a cotton-stuffed torso and limbs with joints. The scene in Figure 2 captures the motion of the protagonist, Omiwa (お三輪), from the fourth act, "Sugisakaya (杉酒屋)", of the Bunraku play "Imoseyama Onna Teikin (妹背山婦女庭訓)".

As shown in Figure 2(a), optical and magnetic motion capture systems (optical motion capture: MAC3D System, magnetic motion capture: Polhemus G4) were used to capture the movements of the bunraku puppets' movements. The sampling rate was 120 Hz. The sensors are attached to the puppet by both sticking and sewing to the surface of the puppet's clothing with thread. Because the sensors used are very light and there are no extra wires to interfere with the operation of the puppet, there will be no burden that affects the performance, as confirmed by the puppeteer masters. Regarding the noise, the data was captured and processed by Nac Image Technology Inc. to ensure that the data accuracy can meet the analysis requirements of this paper.

[1] Credit: BUNRAKU-KYOKAI (The Bunraku Association).

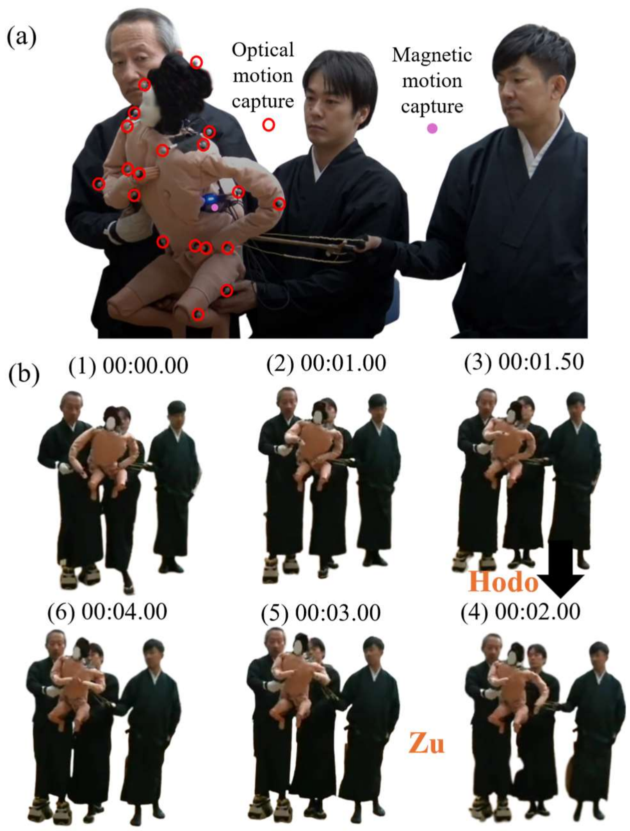


Figure 2. (a) The positions of the sensors on the puppet. (b) One of the representative "kata (型)" motion, "Joining the hands"[2].

Figure 1(c) shows the conceptual arrangement of the motion capture sensors attached to the puppet. The sensor positions are shown in Figure 1(c). However, as shown in Figure 2(a), the actual Bunraku puppet used was one provided by Living National Treasure Kanjuro Kiritake, which has joints in the limbs and torso. The sensors were placed at positions where the puppet's joints were easiest to fix. Apart from the presence of joints and torso, the Bunraku puppet prepared by Kiritake is identical to the one shown in Figure 1.

As a result, the position data is obtained by fixing the optical maker to the puppet body as in Figure 2(a). Then, to perform motion analyses from a kinematic view, we reconstructed the puppet's skeleton and applied inverse kinematics to estimate the joint angles in each frame using MotionBuilder. Given the structural complexity of the Bunraku puppet, we carefully designed the joint setup by focusing on the key rotation centers that are most actively manipulated during performance. The resulting motion data were exported in BVH format [17], which represents motion using a hierarchical skeletal structure. All subsequent analysis in this study was conducted based on this skeletal model.

1.4 The "Hodo" and "Zu" captured in the motion capture experiment

In this study, the focus will be on the coordinated communication between the puppeteer master and the left-hand puppeteer, while the nonverbal communication with the foot puppeteer will not be discussed.

The puppeteer sub-master synchronizes the left hand with the head's movement and begins to move it in the direction of the puppeteer master. Eventually, the left and right hands of the puppet are brought together at the center of the puppet's body, near the chest.

The synchronization signal from the puppeteer master to the puppeteer sub-master is the "Zu". These two synchronized movements are fully integrated into the puppet's actions, which are synchronized with the rhythm of the Gidayu narration.

The motion of the Bunraku is complex and the signals are superimposed, so they have to be viewed in a decomposed way. Though there are other ways to decompose the action signals as well, such as Short-time Fourier Transform (STFT) [18] and Continuous Wavelet Transform (CWT) [19]. However, according to previous research [20-22], HHT is more suitable than both STFT and CWT in motion decomposition. Thus, we use HHT to analyze Bunraku's motion in this study.

1.5 Hilbert-Huang Transform

1.5.1 Definition

The Hilbert-Huang Transform (HHT) consists of Empirical Mode Decomposition (EMD) [23] and Hilbert Spectrum Analysis (HSA) [24]. HHT primarily uses EMD to decompose a signal into Intrinsic Mode Functions (IMFs) and then applies HSA to each IMF to obtain its instantaneous frequency spectrum [25]. The characteristic of HHT is its ability to represent nonstationary and nonlinear signals as instantaneous frequencies.

1.5.2 Hilbert Transform

For any monochromatic signal $c_i(t)$, Hilbert Transform (HT) $d_i(t)$ can be calculated [26]. Here, the monochromatic time signal $c_i(t)$ is treated as a pseudo-monochromatic wave IMF. However, i represents the index of the IMF.

$$d_i(t) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{c_i(\tau)}{t-\tau} d\tau \tag{1}$$

P is the Cauchy principal value of the singular integral, and the monochromatic analytic signal $z_i(t)$, its real part $c_i(t)$, imaginary part $d_i(t)$, instantaneous amplitude $a_i(t)$, and instantaneous phase angle $\theta_i(t)$ are expressed as follows.

$$z_i(t) = c_i(t) + j d_i(t) = a_i(t) e^{j\theta_i(t)} \tag{2}$$

$$a_i(t) = \sqrt{c_i^2(t) + d_i^2(t)} \tag{3}$$

$$\theta_i(t) = \tan^{-1} \frac{d_i(t)}{c_i(t)} \tag{4}$$

Here, j is the imaginary unit, and the instantaneous frequency $\omega_i(t)$ is

$$\omega_i(t) = \frac{d\theta_i(t)}{dt} \tag{5}$$

The nonmonochromatic original signal $x(t)$ can be decomposed into

[2] Credit: BUNRAKU-KYOKAI (The Bunraku Association).

each IMF, which is assumed to be monochromatic, using EMD, and the original signal can be expressed by the following equation:

$$x(t) = \mathrm{Re}\sum_{i=1}^{n} a_i(t)e^{j\theta_i(t)}$$
$$= \mathrm{Re}\sum_{i=1}^{n} a_i(t)e^{j\int \omega_i(t)dt} \qquad (6)$$

Here, n represents the number of decomposed IMFs, and Re denotes the real part of the signal.

In contrast, traditional methods such as the Fourier Transform decompose a signal into components with fixed frequency and amplitude, expressed as $x(t) = \mathrm{Re}\sum_{k=1}^{\infty} a_k e^{j\int \omega_k dt}$. When applied to non-stationary and nonlinear signals, such as human or puppet motion, a large number of components k is required, resulting in limited physical interpretability of the decomposition. Given that the "Hodo" and "Zu" motions are subtle and complicated, often characterized by localized and temporally varying movement, HHT provides superior interpretability and resolution for identifying their frequency characteristics compared to fixed-basis methods. This theoretical advantage is one of the key motivations for selecting HHT in our analysis.

2. Data analyzed by Principal Component Analysis (PCA) and HHT

The motion capture data of the puppet is retargeted and converted into BVH data [17]. BVH data consists of Euler angles on local coordinates fixed to each joint, and Euler angles are introduced by Leonhard Euler to describe the rotation directions of a rigid body relative to a fixed coordinate system.

Euler angles can be defined by a combination of basic geometric shapes or rotations [27]. The Euler angles are local coordinate angles fixed to the joints, and they do not necessarily correspond to the movement angles of the puppet's motions. To discover the primary angular directions aligned with the puppet's actions, this paper applies Principal Component Analysis (PCA) [28] to the θ_x, θ_y, θ_z angles of each joint, where they are, respectively, the rotation angles of x, y and z axes.

On the one hand, since each joint in the motion data has three degrees of freedom corresponding to 3D rotational values, we need to reduce the dimensionality before analyzing the phase of each joint. On the other hand, PCA has been proven to be very effective in reducing the dimension of motion data [29-31]. Then, we adopt PCA to project the 3D rotational data of each joint onto a one-dimensional principal component. Specifically, Multivariate EMD (MEMD) [32] is performed on the principal component rotation angle θ_p, which is the first principal component of each joint's rotation motion, to analyze the "Hodo" and "Zu" actions performed by the puppeteer master. Dimensionality reduction allows us to analyze the synchronization between the movements of puppeteer masters (Spine and Head joints) and puppeteer sub-masters (Left forearm), thereby confirming the temporal correspondence between the detected "Hodo" and "Zu".

3. Results

In this study, we analyze the relationship between the puppeteer master and puppeteer sub-master in a scene from the fourth act "Sugisakaya (杉酒屋)", of the Bunraku play Imoseyama Onna Teikin (妹背山婦女庭訓), where one of the protagonists, Omiwa (お三輪), puts the hands together. This analysis is conducted using PCA, the Hilbert spectrum obtained through the Hilbert-Huang Transform, and the Hilbert instantaneous phase (Equation (4)).

In this experiment, about 11 cases of motion data in which the left hand was used were collected. The same results were obtained in all cases. Among them, the results of the putting the hands together scene, in which cooperative motion is particularly important, are shown below.

Since "Joining the hands" operation in Figure 2 is a rather complex task, Figure 3 shows a time series of nonverbal communication as a conclusion. Figure 3 is used as a reference to understand the analysis in section 3.1 and below.

In Figure 3, the puppeteer master begins with "Hodo" at ❶ (Figure 5(a, b), IMF5). Before finishing the "Hodo", the puppeteer master starts the "Zu" at ❷ (Figure 6(a), 7(a), IMF6), and 0.09 seconds later, the puppeteer master finishes the "Hodo" ❸ (Figure 5(a, b), IMF5).

From the start of the "Zu" in ❷, after 0.19 seconds, in 4 the puppeteer sub-master begins to follow the "Zu" (Figure 6(b), 7(b), IMF6). In❺, the puppeteer master finishes the "Zu" (Figure 6(a), 7(a), IMF6), and after 0.25 seconds, the hands come together. In 6 the puppeteer sub-master finishes following the "Zu" (Figure 6(b), 7(b), IMF6). During about 2 seconds, a dense nonverbal communication takes place between the shamisen player, the Gidayu, the puppeteer master, and the puppeteer sub-master.

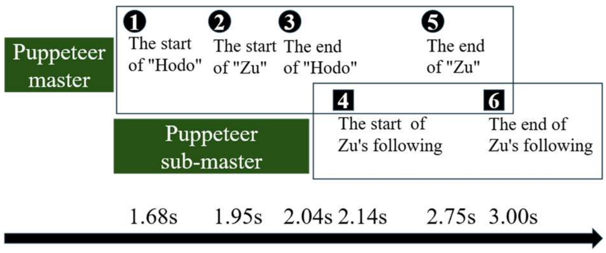


Figure 3. Time series of Bunraku nonverbal communication.

In particular, during the period between❶ and ❸, while the "Hodo" signal is being given, it is clear that the "Zu" signal begins at ❷. This indicates that for the puppeteer master, "Hodo" and "Zu" are not separate, distinct signals, but unified signals. Additionally, from the timing of the signals given by the narrator and the puppeteer master, it is revealed through Empirical Mode Decomposition that "Hodo" and "Zu" are each decomposed into IMF 5 and 6, respectively.

The following sections follow the details of the time series data, focusing on Sections 3.1 ,3.2 and 3.3.

3.1 Comparison of results for STFT, CWT and HHT

Figure 4 shows the comparison of the original signals of spine and head in PC1 direction after PCA processing and the PC1 results after

STFT, CWT and HHT processing. As shown in Figure 4, only the HHT method can analyze the non-linear motions like "Hodo" and "Zu" in the instantaneous frequency domain, which cannot be observed by other methods that pointed out by the red circles. For example, STFT and CWT can capture the approximate timing of the "Zu" motion, as shown in Figures 4(f) and 4(g). In contrast, HHT (Figure 4(h)) reveals more detailed features of "Zu", including a 1 Hz IMF that corresponds to a 1-second motion segment, which was confirmed by the video analysis. This is also consistent with previous research [20-22]. Therefore, in the present study, we use HHT to detect and analyze the "Hodo" and "Zu".

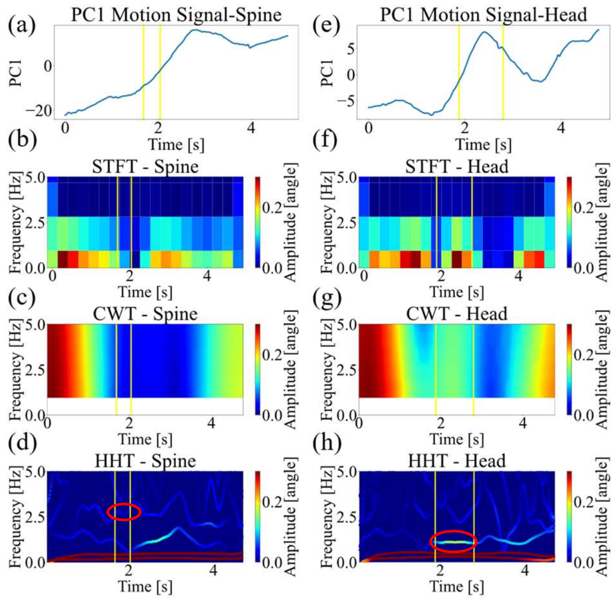


Figure 4. Analytical results of STFT, CWT, and HHT in the time-frequency domain. (a)-(d) are the results of the spine's motion on PC1, STFT, CWT and HHT, respectively. (e)-(h) are the results of the head's motion on PC1, STFT, CWT and HHT, respectively.

3.2 Analysis Using Instantaneous Phase and Instantaneous Frequency

3.2.1 Analysis Procedure

In the following analysis, we identify the IMF signals that capture the cueing behavior emitting "Hodo" and "Zu". That is, based on the Hilbert spectrum intensity or the magnitude of each IMF signal amplitude in Figure 6, the IMF signal of the desired action is identified as IMF5 for "Hodo" and IMF6 for "Zu". The event occurrence and termination times are then estimated from the phase angle data.

Based on the amplitude, IMF5 and IMF6 were identified as the estimated signals. To estimate the timing of the puppeteer master's signal, "Hodo" and "Zu", and the subsequent movements of the left forearm following these cues, the phase angles of IMF5 ("Hodo") and IMF6 ("Zu"), obtained through intrinsic mode decomposition in the Hilbert-Huang transform, are used.

3.2.2 Analysis of "Hodo"

Next, since the "Hodo" here is a simple and small movement in which the puppeteer master simply shifts the puppet to the left, Figure 5(a) shows the Hilbert spectrum obtained by applying PCA and HHT to the BVH data of the Euler angle of the marker on the back of the puppet body. The x-axis is time, the y-axis is instantaneous frequency, and each curve is the spectrum of the decomposed IMF. The color bar on the right indicates the instantaneous amplitude of each spectrum.

Figure 5(b) shows the Hilbert instantaneous phase angles $\theta(t)$ of the Intrinsic Mode Function (IMFs) 5 and 6 defined by Equation (4) for the back of the puppet (Spine). The x-axis represents time, and the y-axis represents the Hilbert instantaneous phase angles $\theta(t)$ of IMFs 5 and 6. The range of $\theta(t)$ in the figure is $-\pi \leq \theta(t) \leq \pi$.

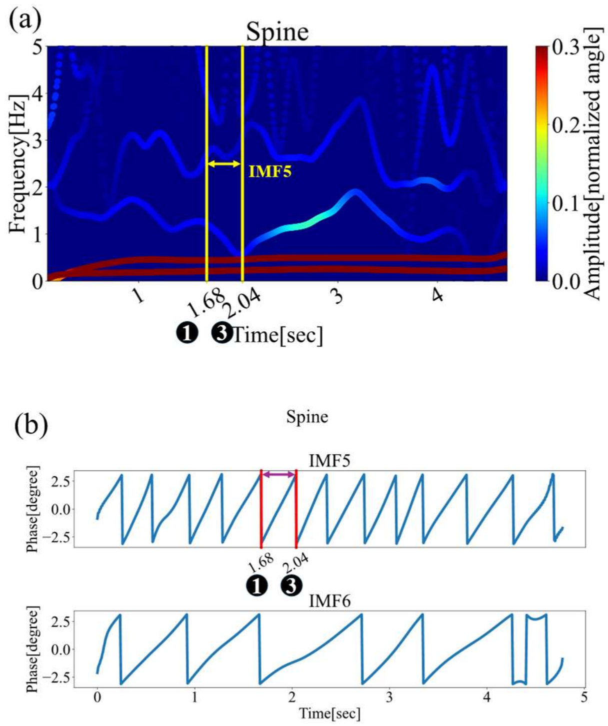


Figure 5. (a) Hilbert Spectrum of the spine. (b) Hilbert Instantaneous Phase of the spine.

In Figure 5, at approximately 1.68 seconds, the puppeteer master shifts the puppet to the right, signaling "Hodo". This is confirmed by IMF5 in Figure 5(b). The puppeteer manipulates the puppet in accordance with the rhythm, and each cycle is one movement. The red line indicates the Hilbert instantaneous phase of "Hodo" (hereafter called instantaneous phase). As shown in Figure 5, about 0.4 seconds from about 1.68 seconds to 2.04 seconds of IMF5 is the "Hodo", and the "Hodo" signal is detected by the instantaneous phase.

Here, the puppeteer master's "Hodo" movement is a small movement that merely shifts the puppet to the left. Although no strong spectrum can be read from the IMF5 spectrum in Figure 5(a), it is found of the IMF5 instantaneous phase in (b).

3.2.3 Analysis of "Zu"

Figure 6 shows the Hilbert spectrum of the instantaneous frequency of the first principal component joint angle after applying principal value analysis to each Euler angle of the head and left forearm joint. The x-axis indicates time, the y-axis indicates instantaneous angular frequency, each curve indicates the decomposed IMF spectrum, and the color bar on the right indicates the instantaneous vibration amplitude.

Figures 6(a) and (b) show the Hilbert spectrum of the instantaneous angular frequencies of the head and left forearm joint, respectively. IMF6 in Figure 6(a) is considered to indicate the "Zu" movement of the head. In Figure 6(a), the yellow vertical line (start time of "Zu") is approximately 1.95 seconds, and the next vertical line (end time of "Zu") is approximately 2.75 seconds.

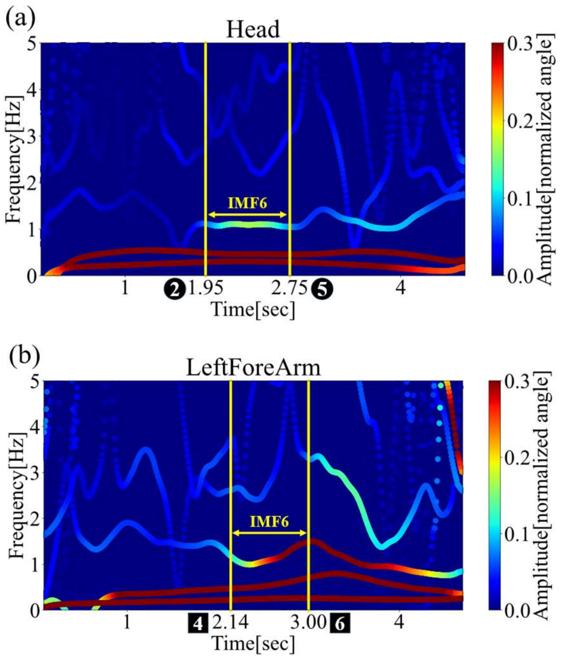


Figure 6. Hilbert Spectrum of Joint Angles. (a) and (b) represent the Hilbert spectrum for the head and left forearm, respectively.

Figure 6(b) shows the Hilbert spectrum of the left forearm joint. The first yellow vertical line represents approximately 2.14 seconds, and the next yellow vertical line represents approximately 3.0 seconds. IMF6 represents the Hilbert spectrum of the left forearm's movement, where the puppeteer sub-master moves the left hand to clasp it in front, following the puppeteer master's "Zu" cue.

Figure 7 shows the instantaneous phase of each head and left forearm joint angle IMF5 and 6, respectively, in Figure 7 x-axis represents time and y-axis represents the instantaneous phase angle $\theta(t)$. In Figure 7(a), the IMF6 spectrum from about 1.95 seconds to 2.75 seconds, which is surrounded by the red line, corresponds to the "Zu" produced by the puppeteer master. Figure 7(b) shows the instantaneous phase angles of IMF5 and IMF6 in the left forearm. The period from about 2.14 seconds to 3.0 seconds shows the left-hand movement of the left forearm following the "Zu" and clasping the hands together. The left forearm responds to the "Zu" cue of the puppeteer master with a delay of about 0.2 second and moves for about 0.2 second longer.

"Hodo" is the signal for the start of "Zu". In other words, the puppeteer master uses the direction, speed, and magnitude of the "Hodo" movement, to send information to the puppeteer sub-master. With this information, the puppeteer sub-master can smoothly follow the "Zu" and bring out the left hand of the puppet.

In this sense, the puppeteer master gives the next size, speed, direction, etc. of the next action as a signal of "Hodo". Thus, the speed and direction of the puppets' motion are improvised by the puppeteer master on the spot during the performance and are communicated to the other Bunraku or Joruri performers by the "Hodo".

The role of "Hodo" is important in this way, as the puppeteer master and puppeteer sub-master need to synchronize each movement of the left hand with the right hand to move it accurately. In this manner, the puppeteer sub-master, in coordination with the puppeteer master, uses "Hodo" and "Zu" to achieve collaborative puppet manipulation. The "Hodo" and "Zu" are embedded in the puppet performance and remain unnoticed by the audience.

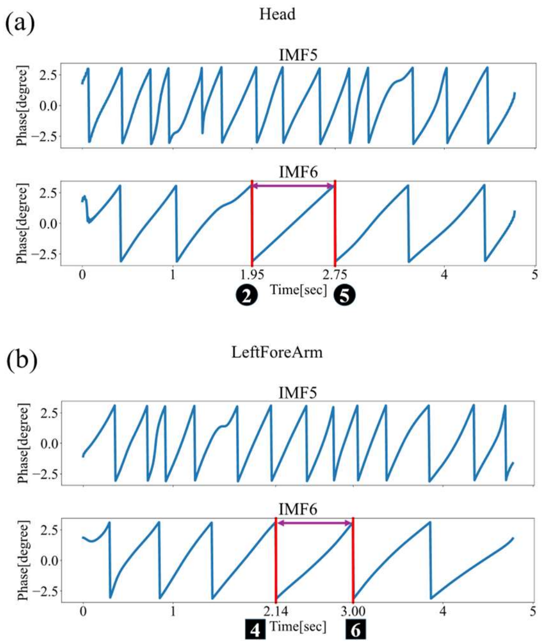


Figure 7. (a) and (b) Hilbert phase of IMF6 corresponding to "Zu" for the head and left forearm.

3.3 Time series similarity analysis of "Hodo" and "Zu"

As mentioned above, in Section 3.1, we calculated the occurrence times of "Hodo" and "Zu". Next, we will analyze the similarity of their motion trajectories in terms of velocity vectors and direction vectors. To quantify the relationship between "Hodo" and "Zu", we employ Procrustes analysis (PA) [33], Mutual Information (MI) [34], and Canonical Correlation Analysis (CCA) [35], which are commonly adopted for calculating the similarity of motions [36-38], to measure the similarities between both velocity and direction vectors of Hodo trajectories that are "spine" ones and Zu trajectories that are "head" or "left forearm" trajectories.

PA, MI, and CCA were selected for their wide applicability and minimal assumptions. On the one hand, the motion capture data used in this study, consisting of time-series 3D velocity and directional vectors, do not strictly follow a normal distribution. On the other hand, PA and MI are non-parametric methods and do not require normality. While CCA assumes linear relationships and benefits from normally distributed data for statistical testing, the canonical correlation scores obtained in our analysis exhibit an approximately symmetric distribution across samples. Although other methods could be considered, the three selected similarity analysis techniques are sufficient for the sensitivity analysis conducted in this study. Therefore, we used them as the basis for the t-tests in our evaluation of detecting "Hodo" and "Zu" using HHT.

Specifically, we will prepare two sets of paired 3-dimensional vectors to verify our detection method: one representing "Hodo" and "Zu" and another consisting of unrelated "Hodo" and "Zu" vectors. The first set includes 11 paired vectors, where the spine vector for "Hodo" has dimensions $3 \times F_1$; and the head vector for "Zu" has dimensions $3 \times F_2$. Since the frame lengths F_1 and F_2 differ, interpolation is applied to normalize their lengths. The second set consists of 11 unrelated paired vectors, where both the spine and head vectors have dimensions $3 \times F_3$. To ensure a fair comparison, the frame length of each pair was determined based on the duration of the corresponding "Zu" segments. Since the original "Hodo-Zu" pairs are not aligned on a shared timeline, the unrelated pairs were randomly selected as the arbitrary motion segments from other sequences while matching the same frame length, thus preserving the similar temporal structure as much as possible for a fair comparison.

For each set, we will compute the similarity of speed vectors and direction vectors using Procrustes Analysis, MI, and CCA. This results in 11 similarity scores for each measure in both sets. Finally, a t-test will be conducted to compare the mean similarity scores between the two sets and determine whether the difference is statistically significant.

Table 1 presents the similarity analysis of speed vectors between the spine and head using Procrustes Analysis, Mutual Information (MI), and Canonical Correlation Analysis (CCA). Table 2 provides similar analysis of direction vectors between the spine and head using Procrustes Analysis, MI, and CCA. Table 3 presents the similarity analysis of speed vectors between the spine and left forearm using

Table 1. Speed vector (spine and head)

	PA		**MI**		**CCA**	
	Hodo-Zu	Arbitrary	Hodo-Zu	Arbitrary	Hodo-Zu	Arbitrary
sample1	**0.3888**	0.1021	**0.8652**	0.5093	**0.9686**	0.7423
sample2	**0.0471**	0.0411	**0.3505**	0.2652	**0.7473**	0.4367
sample3	**0.6954**	0.2601	**1.5821**	0.2736	**0.9806**	0.9327
Sample4	**0.3172**	0.2608	**0.9393**	0.4048	**0.9804**	0.8898
sample5	0.1253	**0.2858**	**0.8849**	0.4907	**0.9837**	0.8396
sample6	**0.2631**	0.0316	**0.9883**	0.0047	**0.8998**	0.5207
Sample7	**0.4285**	0.1758	**0.9855**	0.4962	**0.9561**	0.8204
Sample8	**0.3990**	0.1273	**0.8787**	0.6044	**0.9878**	0.4635
Sample9	**0.4414**	0.1307	**0.7746**	0.1691	**0.9564**	0.8456
Sample10	**0.6055**	0.1750	**1.8578**	0.3373	**0.9924**	0.7632
Sample11	**0.5983**	0.4907	**1.0123**	0.8524	**0.8946**	0.8674
Sample (mean)	**0.3918**	0.1892	**1.0108**	0.4007	**0.9407**	0.7384
T-test (p-value)	**0.00425939**		**0.00147688**		**0.00130710**	

Table 2. Direction vector (spine and head)

	PA		**MI**		**CCA**	
	Hodo-Zu	Arbitrary	Hodo-Zu	Arbitrary	Hodo-Zu	Arbitrary
sample1	**0.6599**	0.2822	**1.2563**	0.5941	**0.9553**	0.8272
sample2	**0.0507**	0.0409	0.2350	**0.4219**	0.4247	**0.5779**
sample3	0.1786	**0.3799**	**0.8789**	0.5733	0.9109	**0.9261**
Sample4	**0.6510**	0.2026	**0.7429**	0.4266	**0.9717**	0.8290
sample5	0.0774	**0.3082**	**0.6822**	0.5042	**0.9469**	0.8138
sample6	**0.6312**	0.4469	**1.0524**	0.4923	**0.9686**	0.8841
Sample7	**0.5178**	0.2313	**0.8956**	0.8667	**0.9820**	0.7951
Sample8	**0.4045**	0.1230	**0.5203**	0.1867	**0.9796**	0.8199
Sample9	**0.4972**	0.1358	**0.6298**	0.4316	**0.9038**	0.7193
Sample10	**0.4259**	0.0896	**1.2789**	0.6250	**0.9801**	0.5338
Sample11	**0.3696**	0.2940	**0.8528**	0.6829	0.8708	**0.8919**
Sample (mean)	**0.4058**	0.2304	**0.8205**	0.5278	**0.8995**	0.7835
T-test (p-value)	**0.03175573**		**0.00391436**		**0.02988056**	

Table 3. Speed vector (spine and left forearm)

	PA		**MI**		**CCA**	
	Hodo-Zu	Arbitrary	Hodo-Zu	Arbitrary	Hodo-Zu	Arbitrary
sample1	**0.3996**	0.2257	**0.7850**	0.2467	**0.9799**	0.7180
sample2	0.0132	**0.0314**	**0.4862**	0.2665	**0.7574**	0.6467
sample3	**0.2928**	0.1924	**0.8352**	0.1517	**0.8958**	0.7429
Sample4	**0.3799**	0.1259	**0.9564**	0.4968	**0.9481**	0.6041
sample5	0.0547	**0.1840**	**0.4040**	0.3691	**0.9187**	0.8077
sample6	**0.4132**	0.0406	**1.0699**	0.2633	**0.9899**	0.7133
Sample7	**0.5702**	0.1649	**0.7347**	0.3487	**0.9797**	0.8454
Sample8	**0.3198**	0.0756	**1.0145**	0.1009	**0.9930**	0.6992
Sample9	**0.6654**	0.1201	**0.9362**	0.5114	**0.9454**	0.7669
Sample10	**0.3442**	0.0497	**1.6244**	0.5529	**0.9031**	0.7614
Sample11	**0.5480**	0.1044	**0.9046**	0.3069	**0.9589**	0.8936
Sample (mean)	**0.3637**	0.1195	**0.8865**	0.3286	**0.9336**	0.7454
T-test (p-value)	**0.00250580**		**0.00011303**		**0.00004249**	

Table 4. Direction vector (spine and left forearm)

	PA		**MI**		**CCA**	
	Hodo-Zu	Arbitrary	Hodo-Zu	Arbitrary	Hodo-Zu	Arbitrary
sample1	**0.4390**	0.1180	**1.1459**	0.4408	**0.9650**	0.6911
sample2	**0.1447**	0.0479	**0.4520**	0.3613	**0.6247**	0.4514
sample3	0.0618	**0.2016**	**0.4058**	0.2931	**0.7744**	0.7221
Sample4	**0.4041**	0.0920	**0.6878**	0.3289	**0.9589**	0.5048
sample5	0.0215	**0.0815**	**0.2525**	0.0650	**0.5964**	0.5810
sample6	**0.5903**	0.3749	**1.0489**	0.5338	**0.9871**	0.9146
Sample7	**0.4707**	0.2471	**0.6274**	0.3956	**0.9835**	0.8846
Sample8	**0.2144**	0.0977	**0.6662**	0.1485	**0.9726**	0.6557
Sample9	**0.7664**	0.2213	**0.8838**	0.1890	**0.9795**	0.7892
Sample10	**0.3387**	0.1138	**1.2767**	0.6441	**0.8914**	0.7803
Sample11	**0.4656**	0.0348	**0.8793**	0.3639	**0.9795**	0.9426
Sample (mean)	**0.3561**	0.1482	**0.7569**	0.3422	**0.8830**	0.7197
T-test (p-value)	**0.00629343**		**0.00013190**		**0.00266519**	

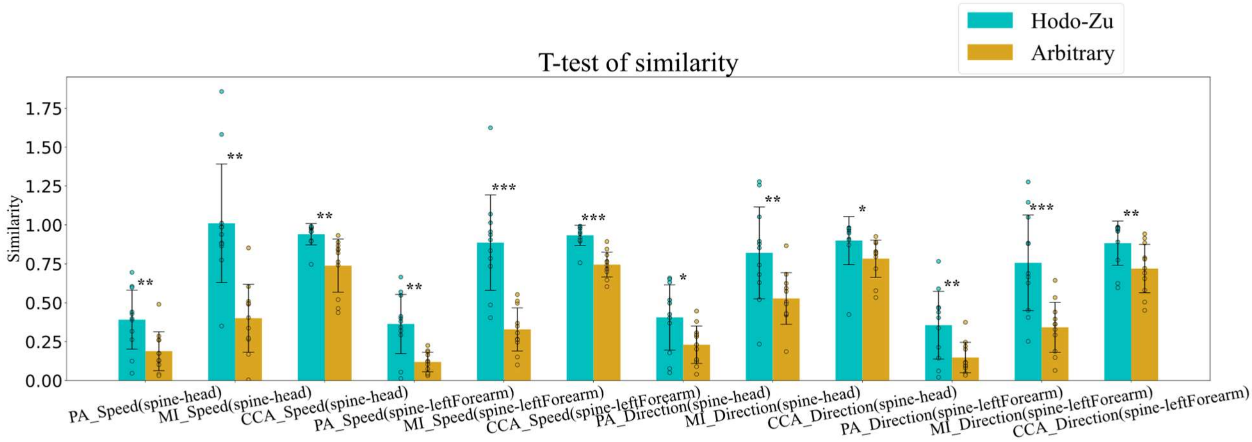


Figure 8. The mean values of PA, MI and CCA under the "Hodo-Zu" and non-"Hodo-Zu" conditions.

Procrustes Analysis, MI, and CCA. Table 4 provides the similarity analysis of direction vectors between the spine and left forearm using Procrustes Analysis, MI, and CCA. The bold values indicate the best results between the two groups for each analysis method.

At the end of Tables 1, 2, 3, and 4, a t-test is conducted for Procrustes Analysis (PA), Mutual Information (MI), and Canonical Correlation Analysis (CCA) to compare "Hodo-Zu" and non-"Hodo-Zu" cases, with the results presented as p-values.

Figure 8 illustrates the mean values of Procrustes Analysis (PA), Mutual Information (MI), and Canonical Correlation Analysis (CCA) across four categories: Speed Vector (spine and head), Direction Vector (spine and head), Speed Vector (spine and left forearm), and Direction Vector (spine and left forearm), under both "Hodo-Zu" and non-"Hodo-Zu" conditions. The figure also includes significance markers from the t-test results. The x-axis represents the categories, while the y-axis indicates the mean values. Green bars represent the "Hodo-Zu" condition, and yellow bars represent the non-"Hodo-Zu" condition. Statistical significance from the t-test is, respectively, denoted by asterisks: * for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$.

As shown in Figure 8, the correlation of both speed vectors and direction vectors is significantly higher in the "Hodo-Zu" condition compared to the non-"Hodo-Zu" condition. Additionally, all p-values are below 0.05, indicating strong statistical significance.

These results show that the paired vectors formed by "Hodo" and "Zu" exhibit significantly higher similarity in both speed and direction vectors compared to randomly selected paired vectors (arbitrary motions). This indicates that the "Hodo" motion, initiated by the puppeteer master through the movement of the spine, successfully drives the subsequent "Zu" motion, involving the head (puppeteer master) and the left forearm (puppeteer sub-master). Furthermore, the three different similarity measures all consistently highlight the same pattern, reinforcing the reliability of the findings. These results validate the effectiveness of our proposed Hilbert-Huang Transform-based detection method in analyzing the interaction between "Hodo" and "Zu" in Bunraku nonverbal communication.

4. Conclusion

Bunraku puppet manipulation relies heavily on nonverbal cues like "Hodo" and "Zu", which have been passed down through hands-on training and remain largely undocumented. Despite recent research using motion capture, analyzing these nonlinear and improvisational movements remains challenging. To address this, we propose a novel motion capture and analysis framework, enabling the extraction of "Hodo" and "Zu" as nonlinear signals (IMF 5, 6) for the first time. Our analysis, employing retargeting, Principal Component Analysis, and Hilbert-Huang Transform, proves effective for complex motion studies beyond Bunraku. Moreover, the 11 examples identified using this method, all of which meet statistical significance across these three validation methods, confirm the effectiveness of our approach, though further applications remain for future research. This study can be summarized as follows:

(1) In this study, motion sensors were attached to the joints of the upper body of a Bunraku puppet with the joint structure shown in Figure 2, which was prepared by the Living National Treasure, Kiritake Kanjuro's group (桐竹勘十郎一座). The motion of one of the characters, Omiwa (お三輪), in Bunraku performances such as Sugisakaya, Yondan, Imoseyama Onna Teikin, (妹背山婦女庭訓四段杉酒屋) was captured.

(2) The captured motion data was retargeted and corrected to match the standard human body structure, and was primarily processed as BVH data as follows.

(3) PCA was applied to the motion data of each joint, represented by θ_x, θ_y, θ_z. For the first principal component of each joint, which is the rotational motion angle (i.e., the main rotational motion angle θ_p), EMD was performed to obtain the Hilbert spectrum and instantaneous phase. Then, the "Hodo" and "Zu" signals were extracted.

(4) In this experiment, the following findings were discovered and confirmed regarding the operation of Bunraku puppets:

① The puppeteer master uses a signal called "Hodo", which involves gently shaking the puppet's torso up and down or side to side, to notify the other puppeteers to begin the coordinated movement.

② The puppeteer master moves the puppet's head and uses the signal "Zu" to guide the puppeteer sub-master's left-hand movements. The puppeteer sub-master moves the left hand in response to the head's "Zu" movement.

③ It was confirmed that the "Hodo" movement corresponds to the phase of IMF5 in Figure 5(b), and the "Zu" movement corresponds to the phase of IMF6 in Figure 7. The phase period of "Hodo" is about half of that of "Zu". In other words, the puppeteer master instructs the puppeteer sub-master on the speed and magnitude of the next movement ("Zu").

(5) Through Procrustes Analysis and correlation analysis using MI and CCA, we observe that the correlation in the "Hodo-Zu " condition is significantly higher than in the non-"Hodo-Zu " condition. This suggests that in Bunraku, puppeteers exhibit a high degree of coordination during "Hodo-Zu".

(6) "Hodo" plays an important role in puppet movements, and it is assumed that the puppeteer master uses "Hodo" to convey various types of leading information. However, this remains for future research to explore.

For future work, we will explore a broader range of statistical methods to analyze and improve the detection algorithm of "Hodo" and "Zu" with higher precision, ultimately facilitating the development of AI-driven interactive systems based on our findings of the Bunraku nonverbal communication mechanism.

Acknowledgement

We would like to express our sincere gratitude to the Bunraku performers Tsukoma Takemoto (Tayu), Sosuke Takezawa (Shamisen), and the puppeteers team led by Kanjuro Kiritake, along with two additional members, for their generous cooperation in providing performance data. We also gratefully acknowledge the support of the BUNRAKU-KYOKAI (The Bunraku Association) for making this research possible.

This study was partially supported by JSPS KAKENHI Grant Numbers JP24K02970, JP25K21325 and JST SPRING.

Reference

[1] Duncan Jr S. Nonverbal communication[J]. Psychological Bulletin, 1969, 72(2): 118.

[2] 植田一博，櫻哲郎，成田友紀，等. 文楽人形遣いの協調のメカニズム[C]//人工知能学会全国大会論文集 第 27 回（2013）. 一般社団法人人工知能学会, 2013: 1H4OS02b5-1H4OS02b5.

[3] 渋谷友紀，森田ゆい，福田玄明，等. 文楽人形遣いにおける呼吸と動作の非同期的関係：日本の古典芸能における「息づかい」の特殊性[J]. 認知科学, 2012, 19(3): 337-364.

[4] 渋谷友紀，櫻哲郎，佐々木正人，等. 文楽における浄瑠璃と人形遣いの呼吸の同期[J]. 認知科学, 2017, 24(4): 518-539.

[5] 董然，蔡東生，中川志信，等. 文楽における序破急のメカニズム解析[C]//IEICE Conferences Archives. The Institute of Electronics, Information and Communication Engineers, 2019.

[6] 櫻哲郎，成田友紀，沢井和也，等. 文楽人形遣いの協調操作を実現する非言語情報通信[J]. 電子情報通信学会論文誌 D, 2013, 96(1): 195-208.

[7] 櫻哲郎，森田寿郎，植田一博. 文楽人形の協調操作における動作特徴の抽出[J]. ヒューマンインタフェース学会論文誌, 2009, 11(3): 255-264.

[8] 櫻哲郎，森田寿郎，植田一博. 文楽人形の動作表現技術の解明‐協調操作を実現する動作要素“ほど”の抽出[C]//日本デザイン学会研究発表大会概要集 日本デザイン学会 第 56 回研究発表大会. 一般社団法人 日本デザイン学会, 2009: A16-A16.

[9] Hattori M, Furuta S, Nishizawa S, et al. An analysis of the amplitude factors of the Bunraku puppet's motion axis-for the description, analysis, and generation of humanoids' motions[C]//Smc 2000 conference proceedings. 2000 ieee international conference on systems, man and cybernetics.'cybernetics evolving to systems, humans, organizations, and their complex interactions'(cat. no. 0. IEEE, 2000, 2: 900-905.

[10] Hattori I, Furuta S, Nishizawa S, et al. A method to generate humanoid's actions based on the phase and amplitude factors of its motion axes-an analysis of the Bunraku puppet's motion axis[C]//Proceedings 9th IEEE International Workshop on Robot and Human Interactive Communication. IEEE RO-MAN 2000 (Cat. No. 00TH8499). IEEE, 2000: 388-393.

[11] Hattori H, Nakabo Y, Tadokoro S, et al. An analysis of the Bunraku puppet's motions based on the phase correspondence of the puppet's motions axis-for the generation of humanoid robots motions with fertile emotions[C]//IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 99CH37028). IEEE, 1999, 2: 1041-1046.

[12] Dong R, Cai D, Hayano S, et al. Investigating the Effect of Jo-Ha-Kyū on Music Tempos and Kinematics across Cultures: Animation Design for 3D Characters Using Japanese Bunraku Theater[J]. Leonardo, 2022, 55(5): 468-474.

[13] Dong R, Ni S, Wu B, et al. Analyzing Jo-Ha-Kyu Mechanism in Japanese Traditional Performing Art Ningyo Joruri[C]//2022 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech). IEEE, 2022: 1-5.

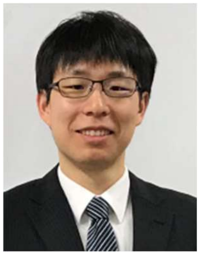
[14] Dong R, Chen Y, Cai D, et al. Robot motion design using bunraku emotional expressions focusing on Jo-Ha-Kyū in sounds and movements[J]. Advanced Robotics, 2020, 34(5): 299-312.

[15] Dong R, Chang Q, Er M J, et al. Motion Capture-Based Robotic Imitation: A Keyframeless Implementation Method Using Multivariate Empirical Mode Decomposition[J]. IEEE/ASME Transactions on Mechatronics, 2024.

[16] Dong R, He Y, Cai D, et al. Interacting with Humanoid Robots: Affective Robot Motion Design with 3D Squash and Stretch Using Japanese Jo-ha-kyu Principles in Bunraku[M]//ACM SIGGRAPH 2021 Talks. 2021: 1-2.

[17] Meredith M, Maddock S. Motion capture file formats explained[J]. Department of Computer Science, University of Sheffield, 2001, 211: 241-244.
[18] Griffin D, Lim J. Signal estimation from modified short-time Fourier transform[J]. IEEE Transactions on acoustics, speech, and signal processing, 1984, 32(2): 236-243.
[19] Rioul O, Duhamel P. Fast algorithms for discrete and continuous wavelet transforms[J]. IEEE transactions on information theory, 1992, 38(2): 569-586.
[20] Dong R, Cai D, Asai N. Nonlinear dance motion analysis and motion editing using Hilbert-Huang transform[C]//Proceedings of the computer graphics international conference. 2017: 1-6.
[21] Dong R, Cai D, Ikuno S. Motion capture data analysis in the instantaneous frequency-domain using hilbert-huang transform[J]. Sensors, 2020, 20(22): 6534.
[22] Dong R, Lin Y, Chang Q, et al. Motion Feature Extraction and Stylization for Character Animation using Hilbert-Huang Transform[C]//Proceedings of the 2021 ACM International Conference on Intelligent Computing and its Emerging Applications. 2021: 16-21.
[23] Huang N E. Hilbert-Huang transform and its applications[M]. World Scientific, 2014.
[24] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences, 1998, 454(1971): 903-995.
[25] Wang G, Chen X Y, Qiao F L, et al. On intrinsic mode function[J]. Advances in Adaptive Data Analysis, 2010, 2(03): 277-293.
[26] Fried B D, Conte S D. The plasma dispersion function: the Hilbert transform of the Gaussian[M]. Academic press, 2015.
[27] Diebel J. Representing attitude: Euler angles, unit quaternions, and rotation vectors[J]. Matrix, 2006, 58(15-16): 1-35.
[28] Abdi H, Williams L J. Principal component analysis[J]. Wiley interdisciplinary reviews: computational statistics, 2010, 2(4): 433-459.
[29] Glardon P, Boulic R, Thalmann D. PCA-based walking engine using motion capture data[C]//Proceedings Computer Graphics International, 2004. IEEE, 2004: 292-298.
[30] Federolf P, Reid R, Gilgien M, et al. The application of principal component analysis to quantify technique in sports[J]. Scandinavian journal of medicine & science in sports, 2014, 24(3): 491-499.
[31] Warmenhoven J, Cobley S, Draper C, et al. Considerations for the use of functional principal components analysis in sports biomechanics: examples from on-water rowing[J]. Sports Biomechanics, 2019, 18(3): 317-341.
[32] Rehman N, Mandic D P. Multivariate empirical mode decomposition[J]. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2010, 466(2117): 1291-1302.
[33] Gower J C. Generalized procrustes analysis[J]. Psychometrika, 1975, 40: 33-51.
[34] Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information[J]. Physical Review E—Statistical, Nonlinear, and Soft Matter Physics, 2004, 69(6): 066138.
[35] Hardoon D R, Szedmak S, Shawe-Taylor J. Canonical correlation analysis: An overview with application to learning methods[J]. Neural computation, 2004, 16(12): 2639-2664.
[36] Adams DC, Cerney MM. Quantifying biomechanical motion using Procrustes motion analysis. Journal of Biomechanics. 2007 Jan 1;40(2):437-44.
[37] Gazula H, Chang CC, Lu ML, Hsiang SM. Using mutual information to capture major concerns of postural control in a tossing activity. Journal of biomechanics. 2015 Apr 13;48(6):1105-11.
[38] Bashir K, Xiang T, Gong S. Cross view gait recognition using correlation strength[C]//Bmvc. 2010: 1-11.

Xunan Liu is currently enrolled in the Degree Programs in Systems and Information Engineering, University of Tsukuba, and engaged in research on computer graphics.



Ran Dong received his PhD degree from the Graduate School of Systems and Information Engineering, University of Tsukuba. He is currently a Lecturer with the School of Engineering, Chukyo University, His research has focused on human–machine interaction and artificial intelligence.

Dongsheng Cai received his Ph.D. in Engineering from the Graduate School of Engineering, University of Tokyo, in 1989. He is currently a Professor at the Faculty of Management, Nagoya University of Commerce and Business. He was formerly an Associate Professor at the Institute of Systems and Information Engineering, University of Tsukuba. His research focuses on scientific computing and visualization.