Paper:

Editing Robot Motion Using Phonemic Feature of Onomatopoeias

Junki Ito[*1], **Masayoshi Kanoh**[*2], **Tsuyoshi Nakamura**[*3], **and Takanori Komatsu**[*4]

[*1]Graduate School of Computer and Cognitive Sciences, Chukyo University
101 Tokodachi, Kaizu-cho, Toyota, Aichi 470-0393, Japan
E-mail: h11202m@st.chukyo-u.ac.jp
[*2]School of Information Science and Technology, Chukyo University
101 Tokodachi, Kaizu-cho, Toyota, Aichi 470-0393, Japan
E-mail: mkanoh@sist.chukyo-u.ac.jp
[*3]Graduate School of Engineering, Nagoya Institute of Technology
Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan
E-mail: tnaka@nitech.ac.jp
[*4]Faculty of Textile Science and Technology, Shinshu University
3-15-1 Tokida, Ueda, Nagano 386-8567, Japan
E-mail: tkomat@shinshu-u.ac.jp
[Received November 15, 2012; accepted January 30, 2013]

Onomatopoeias are words that represent the sound, appearance, or voice of things, thus making it possible to create expressions that bring a scene to life in a subtle fashion. Onomatopoeias can be used to make the process of robot motion generation more easily and intuitively. In previous studies, subjective quantified values of onomatopoeias have been used as indices of robot motion, but the generality of the motion has not been evaluated. In this study, we propose a method for generating robot motion using the objective quantified values of onomatopoeias. We experimentally verified that the proposed method generated more suitable motion than did the previous methods.

Keywords: onomatopoeia, P-type Fourier descriptor, auto-associative neural network

1. Introduction

In recent times, biped hobby robots for public use have become increasingly popular. To fully exploit the capabilities of such robots, it is essential to be able to edit their motions. Although most such robots are available with editing software, actually using this software to edit robot motions requires considerable knowledge about robots and experience in editing robot motions. In addition, the need to properly control more than one joint makes it difficult to intuitively imagine the edited robot motions. This also makes it difficult to accurately reflect the editor's images through robot motions, which imposes upon the editor the heavy burden of trials and errors. To overcome these problems, we have previously proposed a method to intuitively edit humanoid robot motions using onomatopoeias, which are sensuous expressions for the sounds or states of an object [1]. However, the authors [1, 2] used their subjective impressions to quantify the attribute values of such onomatopoeias or to establish a correspondence between the attribute values and the robot motions, and therefore, the images of such onomatopoeias could not be conveyed properly. While onomatopoeias are primarily used by those who observe the conditions of robot motions or states to subjectively grasp and represent such conditions, they can also be used effectively to convey some universal images through their acoustic characteristics. It would then be possible to properly convey an editor's images to a third party by using the acoustic characteristics of onomatopoeias. In this study, we propose an operation plane system to edit robot motions. In this system, we use four-dimensional attribute vectors as an objective index based on the acoustic characteristics of onomatopoeias [3]. Section 2 of this paper describes what onomatopoeias are. Section 3 presents an overview of the proposed system, and it discusses a method to quantify onomatopoeias, a method to generate the waveforms for robot motions that represent onomatopoeias, a method to extract the waveform characteristics from the waveforms for robot motions, a method to construct an operation plane, and a method to generate robot motions from the operation plane. Section 4 presents a comparison of the proposed method for generating waveforms for robot motions with the conventional method, and it demonstrates that the proposed method can represent the images of onomatopoeias more accurately than the conventional method. The accuracy of the operation plane that has been constructed with the waveforms for robot motions generated by the proposed method is also verified.

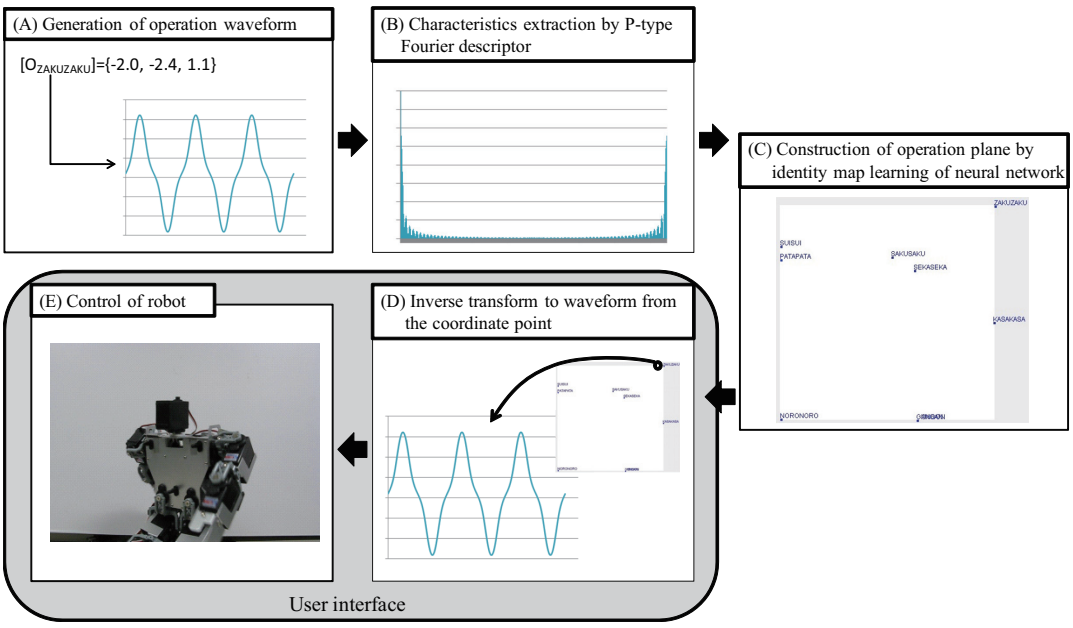


Fig. 1. Proposed system.

2. What Are Onomatopoeias?

Onomatopoeias are a collective term for "onomatopoeic words," "mimetic words," and "imitative words." As they are sensuous expressions for the states or motions of an object, they can characteristically represent things with more sense of presence and in a more subtle way than ordinary words. According to some ideas, the images felt from onomatopoeias are derived from their acoustic characteristics, independent of linguistic meanings. Such acoustic characteristics, called phonemic features, are said to give far more universal impressions than any words.

On the other hand, the impressions given by onomatopoeias include "vague impressions that cannot be well verbalized" [2]. If we can accurately grasp such "vague impressions" and appropriately provide them to a proper subject, it may assist us with our sensuous understanding and reduce our cognitive burden. Several studies in which onomatopoeias are being used as inputs from the abovementioned viewpoints are now underway. "Onomatopen" developed by Kanbara et al. [4] uses the onomatopoeias expressed by users as inputs to serve as an auxiliary function for painting tools, thus providing users with an intuitive operational environment. Onomatopen, however, can use only a limited number of onomatopoeias. To cope with this problem, Terashima et al. developed an image-editing tool called "MOYA-MOYA Drawing" [5]. MOYA-MOYA Drawing assists users in their expressive activities by quantifying input onomatopoeias and assigning corresponding effects to the subject images. Onomatopen's problem in that it can use only a limited number of onomatopoeias is solved by quantifying such onomatopoeic characteristics. In consideration of the fact that in medical settings, patients often use onomatopoeias to express their pain in both quantitative and qualitative ways, Ueda et al. developed a system to support medical interviews with onomatopoeias [6]. This system inputs patients' onomatopoeias to quantitatively estimate and output the degrees or intensities of their pain from the onomatopoeic phonemic features. This system enables medical doctors to properly grasp patients' conditions in a short time. Noting that Japanese customers often use onomatopoeias to represent dishes or tastes, Lertsumruaypun et al. developed "Onomatoperori" [7], a recipe recommendation system using onomatopoeias. This system aims at more accurate searches by computing the co-occurrence rates between cooking-related words dealing with ingredients or cooking methods and onomatopoeias. Clearly, support tools based on onomatopoeias are being developed in many areas.

3. Proposed System

3.1. Overview

Figure 1 shows an overview of the proposed system. First, the basic waveform for robot motions is transformed based on the attribute vectors of phonemic features [2] of onomatopoeias (A). Then, the waveform characteristics of the transformed waveform are extracted using P-type Fourier descriptors [8–10] (B). After that, the similarity relationships among onomatopoeias is learned through auto-associative learning by applying a neural network to the extracted waveform characteristics to construct an operation plane (C). The neural network extracts the waveform characteristics using the P-type Fourier descriptor when a user selects a certain coordinate point on the operation plane (D). Inverse Fourier transform of the waveform characteristics extracted using the P-type Fourier descriptor will restore the waveforms for robot motions to make them move (E).

Table 1. Four-dimensional attribute vectors set for four elements [3].

		Quickness	Softness	Dynamic	Width
Vowels	A	0.05	0.29	0.83	1.38
	I	0.71	−0.88	0.15	−1.23
	U	−0.73	0.73	−0.95	−0.02
	E	0.29	−0.45	−0.08	−0.61
	O	−0.69	0.55	−0.15	1.83
Consonants	K	2.05	−2.43	1.54	−0.46
	S	1.67	−0.92	1.15	−1.55
	T	1.20	−1.51	1.13	0.08
	N	−1.26	0.94	−1.55	0.04
	H	−0.01	0.45	0.17	−0.26
	M	−1.42	1.31	−1.36	0.82
	Y	−0.75	0.74	−0.47	−1.43
	R	0.10	0.31	0.67	−0.37
	W	−0.43	0.78	0.65	1.51
Others	Dull Sounds	−0.07	−1.57	−0.28	0.87
	Semi-Dull Sounds	0.36	0.76	0.88	−0.83

The most significant feature of the proposed system lies in using the operation plane as an input interface. In the proposed system, the degrees of similarity among onomatopoeias are visually represented to create a more intuitive environment for inputs than texting inputs. The proposed system, using which the operation plane is constructed through auto-associative learning, can even generate the waveforms of onomatopoeias that do not exist on the operation plane.

3.2. Method to Objectively Quantify Onomatopoeias

This study discusses a method to generate robot motions that conform to $XYXY$-type onomatopoeias such as "SAKUSAKU" and "TOBOTOBO," where the first two sounds are repeated twice. In Japanese, $XYXY$-type onomatopoeias are used most often as adverbs for adjective verbs. In English, "SAKUSAKU aruku" may be translated as *quickly* walking (or *trotting*) and "TOBOTOBO aruku," as *ploddingly* walking (or *plodding*). This study discusses how to deal with onomatopoeias that correspond to *quickly* and *ploddingly* as described above.

The attribute values of eight-dimensional attribute vectors were subjectively quantified by the authors but were found to lack objectivity [1, 2]. However, a method to quantify onomatopoeias in a more objective manner was also proposed [3]. This study discusses how to generate robot motions by using the latter method to quantify onomatopoeias.

First, all of the vowels and consonants that constitute onomatopoeias are provided using four attribute values – "quickness," "softness," "dynamic," and "width" – as shown in **Table 1**. Next, for onomatopoeia O of $XYXY$-type, four attribute values O_i ($i \in q =$ quickness, $d =$ dynamic, $s =$ softness, $w =$ width) are calculated from the attribute values of the component vowels $X^{(v)}$, $Y^{(v)}$ and consonants $X^{(c)}$, $Y^{(c)}$.

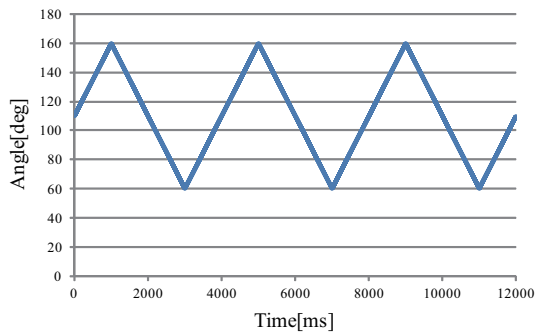


Fig. 2. basic waveform.

$$\left.\begin{aligned} O_q &= 0.60X_q^{(c)} + 0.52Y_q^{(c)} \\ O_d &= 0.59X_d^{(c)} + 0.4Y_d^{(c)} \\ O_s &= 0.56X_s^{(c)} + 0.46Y_s^{(c)} + 0.22Y_s^{(v)} \end{aligned}\right\} \quad . \quad . \quad (1)$$

The abovementioned attribute values have been objectively obtained; however, the attribute value O_w for width has not been obtained. In this study, therefore, we use O_q, O_d, and O_s to generate the waveforms for robot motions.

3.3. Method to Generate Waveforms for Robot Motions

We now propose the method to generate the waveforms for robot motions by using the attribute values O_i obtained using the above-described method. Specifically, the motion waveforms for onomatopoeia O are generated by transforming the amplitude, period, and shape of the basic waveform (**Fig. 2**).

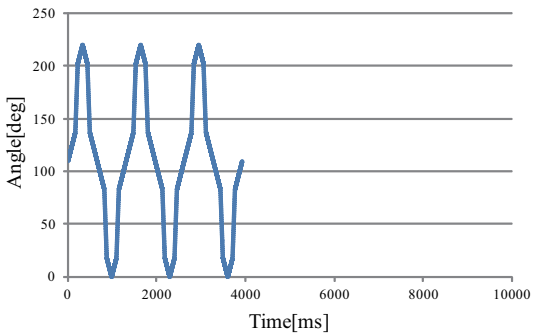


Fig. 3. Waveform for "GASIGASI."

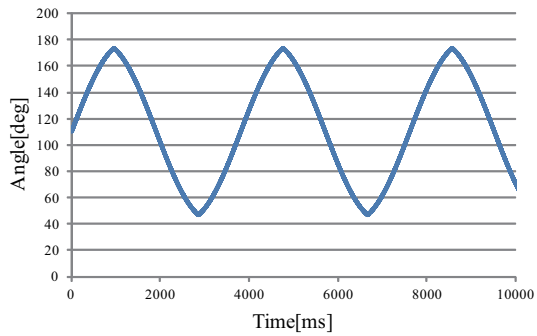


Fig. 4. Waveform for "UROURO."

3.3.1. Adjustment of Periods by Quickness

Assuming that the attribute values for quickness affect the speeds of robot motions, the period T_O of the motion waveforms can be given by the following equation:

$$T_O = \begin{cases} T_B(|O_q|+1.0) & (O_q < 0) \\ \dfrac{T_B}{|O_q|+1.0} & (O_q \geq 0) \end{cases} \quad \ldots\ldots \quad (2)$$

where T_B denotes the period of the basic waveform.

3.3.2. Adjustment of Amplitudes by Dynamic

Assuming that the attribute values for the dynamics affect the scale of robot motions, we have determined the amplitude A_O of the motion waveforms using the following equation:

$$A_O = \begin{cases} |O_d|+1.0 & (O_d \geq 0) \\ \dfrac{1}{|O_d|+1.0} & (O_d < 0) \end{cases} \quad \ldots\ldots \quad (3)$$

3.3.3. Waveform Transformation by Softness

Assuming that the attribute values for softness affect the trajectory of robot motions, we decided to transform the basic waveform.

First, the attribute value O_s is normalized within $[-1,1]$. The normalized O_s is denoted as O_s^n. Then, the motion waveforms $f_O(n)$ $(n = 0,\ldots,T_O)$ are generated as follows, using the abovementioned T_O and A_O.

$$f_O(n) = A_O\left(f_B\left(\frac{T_B}{T_O}n\right) + |O_s^n|\left(A_B g(n) - \left(f_B\left(\frac{T_B}{T_O}n\right) - f_B(0)\right)\right)\right) \quad (4)$$

where $f_B(n)$ denotes the basic waveform, A_B denotes the amplitude of the basic waveform, and $g(n)$ is defined as follows:

If $O_s \geq 0$,

$$g(n) = \sin\left(\frac{2\pi n}{T_O}\right). \quad \ldots\ldots\ldots\ldots \quad (5)$$

If $O_s < 0$,

$$g(n) = \begin{cases} 0 & \left(0 \leq n < \frac{3}{24}T_O, \frac{9}{24}T_O \leq n < \frac{15}{24}T_O, \frac{21}{24}T_O \leq n \leq T_O\right) \\ 1 & \left(\frac{4}{24}T_O \leq n < \frac{8}{24}T_O\right) \\ -1 & \left(\frac{16}{24}T_O \leq n < \frac{20}{24}T_O\right) \\ \frac{24}{T_O}n - 3 & \left(\frac{3}{24}T_O \leq n < \frac{4}{24}T_O\right) \\ \frac{-24}{T_O}n + 9 & \left(\frac{8}{24}T_O \leq n < \frac{9}{24}T_O\right) \\ \frac{-24}{T_O}n + 15 & \left(\frac{15}{24}T_O \leq n < \frac{16}{24}T_O\right) \\ \frac{24}{T_O}n - 21 & \left(\frac{20}{24}T_O \leq n < \frac{21}{24}T_O\right) \end{cases} \quad \ldots\ldots\ldots\ldots \quad (6)$$

The abovementioned waveform transformation produces more sinusoidal and trapezoidal waves in the case of larger and smaller attribute values for softness, respectively. The waveforms generated for "GASIGASI" and "UROURO" by the abovementioned method are shown in **Figs. 3** and **4**, respectively. "GASIGASI" is an onomatopoeia that represents solid and strong robot motions and "UROURO," repetition of purposeless to and fro motions.

3.4. Extraction of Waveform Characteristics

We extract the waveform characteristics of robot motions, using P-type Fourier descriptors [8–10]. A P-type Fourier descriptor is used to expand complex functions with an integral curvature function $\theta(i)$ in the exponential part. We present a brief overview of P-type Fourier descriptors below.

Line graph C (**Fig. 5**) is a polygonal diagram that connects two neighboring points $z(j)$ and $z(j-1)$; the dis-

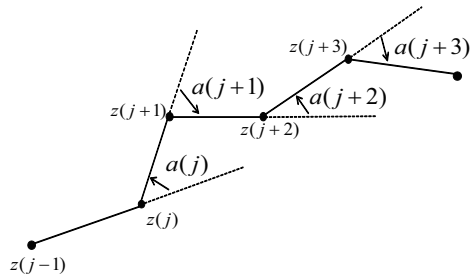


Fig. 5. Line graph C.

tance δ between the two points is assumed to be constant.

$$\delta = |z(j) - z(j-1)| \quad (j = 1, \ldots, n) \quad . \; . \; . \; . \quad (7)$$

The angle between two vectors $z(j) - z(j-1)$ and $z(j+1) - z(j)$ on line graph C is positive in the anticlockwise direction and is denoted by $a(j)$ $(-\pi \leq a(j) < \pi)$. Next, the integral curvature function $\theta(j)$ on Line Graph C is defined using the angular bend function $a(j)$ as follows:

$$\begin{cases} \theta(0) = a(0) \\ \theta(j) = \theta(j-1) + a(j) \quad (j = 1, \ldots, n-1) \end{cases} \quad (8)$$

where $a(0)$ denotes the angles between vector $z(1) - z(0)$ and the x-axis. The complex function $\omega(j)$ is defined with $\theta(j)$ as follows:

$$\omega(j) = \exp(i\theta(j)) \quad (j = 1, \ldots, n-1) \quad . \; . \; . \quad (9)$$

where i denotes the imaginary unit. The discrete Fourier transform of $\omega(j)$ is expressed by the following equation:

$$c(k) = \frac{1}{n} \sum_{j=0}^{n-1} \omega(j) \exp\left(-2\pi i \frac{jk}{n}\right). \quad . \; . \; . \; . \quad (10)$$

The use of the P-type Fourier descriptor affords the following advantages: (1) the end points of the original curve and the reproduced curve always come at the same position; (2) nothing needs to be changed for parallel translation, enlargement, or reduction of the curves; (3) the reproduced curve visually comes closer to the original curve as the degree of reproduction increases; and (4) as data on the pattern class of the original curve are contained in a very limited number of low-region components, it can serve as an excellent feature parameter. These waveform characteristics enable us to utilize such a limited number of low regions as feature parameters.

Now, we describe a method to reproduce the waveforms for robot motions by using the inverse Fourier transform. The following equation is obtained by solving Eq. (10) in terms of $\omega(j)$:

$$\omega(j) = \frac{1}{n} \sum_{j=0}^{n-1} c(k) \exp\left(-2\pi i \frac{jk}{n}\right) \quad . \; . \; . \; . \; . \quad (11)$$

where $\hat{c}(k)$ is defined as follows:

$$\hat{c}(k) = \begin{cases} c(k) & \left(k = 0, 1, \ldots, \frac{n}{2}\right) \\ c(n+k) & \left(k = -\frac{n}{2} + 1, \ldots, -1\right) \end{cases} \quad . \; (12)$$

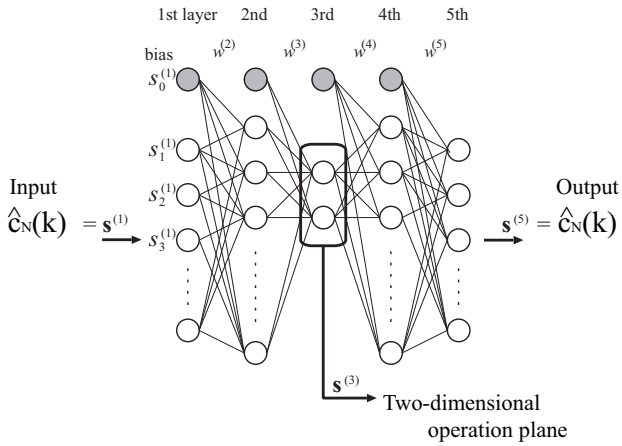


Fig. 6. Auto-associative neural network.

$\omega_N(j)$ using only the lower-period region, $|k| \leq N$, of $c(k)$ denotes an N-th-degree P-representation of line graph C and is expressed by the following equation:

$$\omega_N(j) = \sum_{k=-N}^{N} \hat{c}_N(k) \exp\left(-2\pi i \frac{jk}{n}\right). \quad . \; . \quad (13)$$

The curve obtained from Eq. (13) is called the reproduced curve of the N-th-degree.

$$z_N(j) = z_N(0) + \delta \sum_{r=1}^{j-1} \omega_N(r) \quad . \; . \; . \; . \; . \; . \; . \; . \quad (14)$$

In this study, we use $\hat{c_N}(k)$ as the waveform characteristics.

3.5. Construction of Operation Plane

Although we have extracted the waveform characteristics by using P-type Fourier descriptors, we still have to grasp the similarity and positional relationships among onomatopoeias. Therefore, we make use of auto-associative learning using the neural network to compress the waveform characteristics into two dimensions so that the relationships among onomatopoeias can be made visible.

The auto-associative neural network in the shape of a sand hourglass with a squeezed middle layer is a learning network with the same learning data input in the input and the output layers [11]. The internal structures contained in the learning data are acquired in the middle layer through auto-associative learning [12]. The generalization capability of the neural network makes it possible to generate intermediate onomatopoeias by learning basic data only. Auto-associative networks have been used in various systems to analyze or synthesize the facial expressions of humans or robots [12–15].

In this study, we use a five-layer auto-associative network, as shown in **Fig. 6**. This network is constructed such that the third layer has fewer units than the input/output units. This five-layer network can provide better nonlinear mapping performance than a three-layer net-

work. In this five-layer auto-associative network, information about the characteristics of input data is extracted in the third layer. In this study, we set the number of units extracted in the third layer to two, so that such an extracted characteristic space can be utilized as an operation plane.

The following is the learning process of the auto-associative neural network.

$\mathbf{s}^{(1)}$ denotes the waveform characteristics $\hat{c_N}(k)$ to be input in the neural network.

Output $s_j^{(k)}$ of the j-th unit in the k-layer is given by the following equation:

$$s_j^{(k)} = f\left(u_j^{(k)}\right) \quad \ldots\ldots\ldots\ldots \quad (15)$$

where $f(x)$ denotes a sigmoid function and $u_j^{(k)}$ is given by the following equation:

$$u_j^{(k)} = \sum_i w_{ij}^{(k)} s_i^{(k-1)} \quad \ldots\ldots\ldots\ldots \quad (16)$$

where $w_{ij}^{(k)}$ denotes the coupling load and $u_0^{(k)}$, the bias unit.

The errors between the input and the output are given by the following equation:

$$E = \frac{1}{2}\sum_i \left(s_i^{(1)} - s_i^{(5)}\right)^2. \quad \ldots\ldots\ldots \quad (17)$$

Learning is performed by minimizing E by using the error back-propagation algorithm.

$$w_{ij}^{(k)}(t+1) = w_{ij}^{(k)}(t) + \Delta w_{ij}^{(k)}(t) \quad \ldots\ldots \quad (18)$$

where

$$\Delta w_{ij}^{(k)}(t) = \varepsilon d_j^{(k)} s_i^{(k-1)} + \eta \Delta w_{ij}^{(k)}(t-1), \quad \ldots \quad (19)$$

$$d_j^{(k)} = \begin{cases} f'\left(u_j^{(k)}\right) \sum_l w_{jl}^{(k+1)}(t) d_l^{(k+1)} & (k \neq 5) \\ f'\left(u_j^{(k)}\right)\left(s_j^{(1)} - s_j^{(k)}\right) & (k = 5). \end{cases} \quad \ldots\ldots\ldots\ldots \quad (20)$$

where ε denotes the learning rates and η, the momenta.

Through the abovementioned learning process, information about the characteristics of the input data is extracted in the third layer. In this study, we utilize the characteristic space extracted in the third layer as an operation plane.

3.6. Generation of Robot Motions

Robot motions are generated by extracting motion waveforms from the neural network. Specifically, as shown in **Fig. 7**, a certain coordinate point on the operation plane is first input into the third layer of the neural network, and then, waveform characteristics $\hat{c_N}(x)$ are extracted from the fifth layer for inverse Fourier transform so that a reproduced curve can be obtained for generating robot motions.

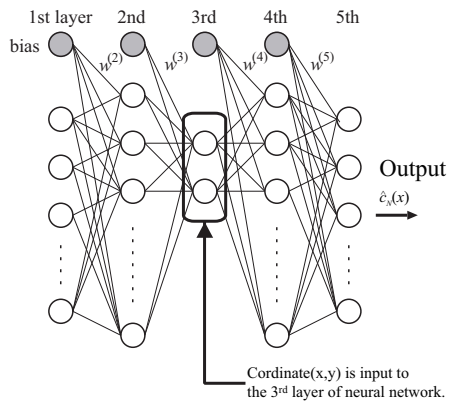


Fig. 7. Generation of motions.

4. Experiments

In this study, we evaluated the input interface by the operation plane based on the following two conditions: (1) the proposed method to generate the waveforms for robot motions can properly convey the editor's images to a third party, and (2) the neural network can reproduce high-accuracy waveforms for robot motions.

4.1. Evaluations of Waveforms Generated for Robot Motions

4.1.1. Experimental Settings

To verify the effectiveness of the method for generating the waveforms for robot motions, as proposed in Subsection 3.3, we examined whether the generated waveforms properly represent the images of the onomatopoeias provided for a walking robot to swing the arms back and forth. First, we selected twenty types of onomatopoeias that can be associated with such arm-swinging motions. Next, using both the proposed method and a comparative approach, we generated the waveforms for robot motions that will correspond to these twenty onomatopoeias. We showed videos of a robot that makes some motions following such generated waveforms to 36 human subjects to evaluate, on a 5-point Likert scale, how closely they feel the robot's motions correspond to their own images of the onomatopoeias. In the experiments, we used silent videos to avoid any disturbances such as the noises made by the robot while moving. The robot used for the experiments is a KHR-2HV biped walking robot made by Kondo Kagaku Co., Ltd. (**Fig. 8**).

4.1.2. Comparative Approach

Now, we describe the eight-dimensional attribute model [1] (conventional method) that we have adopted as a comparative approach.

First, we assign the following eight attributes to the vowels and consonants that constitute Japanese words: "strength," "hardness," "wetness," "smoothness," "roundness," "elasticity," "quickness," and "warmth." The attribute values are set within $[-2, 2]$. When phonemes

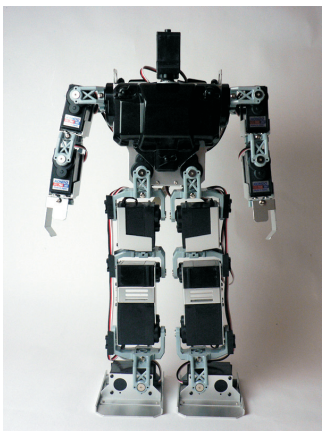


Fig. 8. KHR-2HV.

in an onomatopoeia $XYXY$ are denoted by $X^{(v)}$, $Y^{(v)}$ for vowels and by $X^{(c)}$, $Y^{(c)}$ for consonants, the i-th attribute value O_i for an onomatopoeia O is calculated using the following equation:

$$O_i = 2X_i^{(c)} + X_i^{(v)} + \frac{\left(2Y_i^{(c)} + Y_i^{(v)}\right)}{2}. \quad . \; . \; . \; (21)$$

$Z_i^{(s)}$ denotes the attribute values when $Z \in \{X, Y\}$, $s \in \{v =$vowel, $c =$consonant$\}$, $i \in \{$strength, hardness, wetness, smoothness, roundness, elasticity, quickness, warmth$\}$. The waveforms for robot motions that correspond to the onomatopoeias are generated by modifying the amplitude and period of the basic waveform according to the eight-dimensional vector values to be obtained from Eq. (21). The waveforms for "GASIGASI" and "UROURO" are illustrated in **Figs. 9** and **10**, respectively.

The details of the waveform transformation can be referred to from a previous study [1].

4.1.3. Results/Consideration

The findings of the abovementioned questionnaire are shown in **Fig. 11**.

These findings indicate that the proposed method was evaluated to be equal to or better than the conventional method for several onomatopoeias. To verify the evaluations, we conducted a test on the scores received by both the proposed method and the conventional method. As a result, some significant differences ($p = 0.000$) were recognized between the two methods (**Fig. 12**). This indicates that the images of onomatopoeias can be more accurately conveyed to users if the proposed method is used or that the proposed method is more effective as an objective index than the conventional method.

The proposed method and the conventional method adopt different approaches to generate the waveforms for robot motions or to quantify the onomatopoeias. For the generation of waveforms for robot motions, although both methods adopt the same approach to calculate the amplitude and period of the waveforms, the methods for waveform transformation differ considerably: in the proposed method, the waveforms for robot motions present smaller

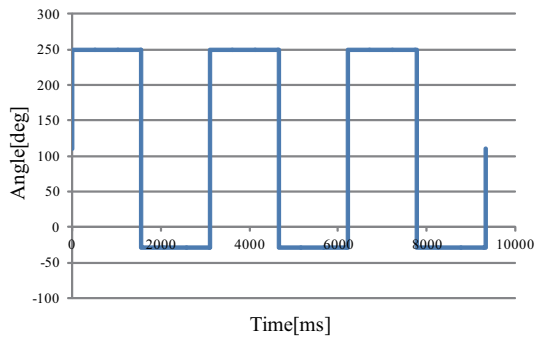


Fig. 9. Waveform for "GASIGASI."

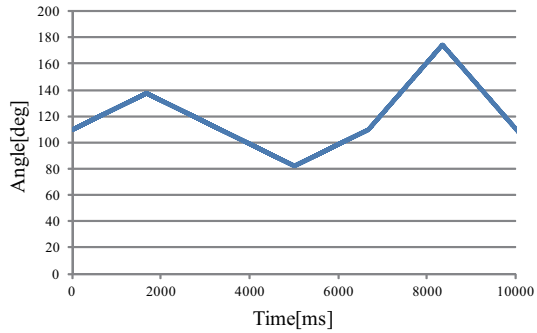


Fig. 10. Waveform for "UROURO."

and more continuous changes in shape than in the conventional method. Minute changes in shape from basic waveforms such as a trapezoidal or a sinusoidal wave are difficult to appreciate intuitively [2]. This suggests that human subjects can hardly appreciate minute changes in the shape of the waveforms generated by the proposed method. Human subjects seem more likely to focus on the amplitude and period of the arm-swinging motion. Accordingly, we consider that any improvements in the evaluation results can be attributed to the change in the methods used for quantifying onomatopoeias (introduction of objective numerical values).

By using the proposed method, the evaluations for three types of onomatopoeias – "MOTAMOTA," "BURUBURU," and "UROURO" – are lower than those for the conventional method. First, we illustrate the waveform for "MOTAMOTA" motion in **Fig. 13**. By using the proposed method, the waveform for "MOTAMOTA" motion has a shorter period than in the conventional method, indicating that the proposed method seems to fail to well represent the "sluggish motion with things falling into arrears" that could be associated with the "MOTAMOTA" motion. Next, the waveforms for "BURUBURU" motion and "UROURO" motion are shown in **Figs. 14** and **15**, respectively. "BURUBURU" is an onomatopoeia that represents a heavily vibrating motion and "UROURO," one that represents a motion for moving around back and forth. For both "BURUBURU" and "UROURO," by using the conventional method, the amplitudes of the waveforms vary with different periods, whereas the proposed method presents the same amplitudes for any period. This

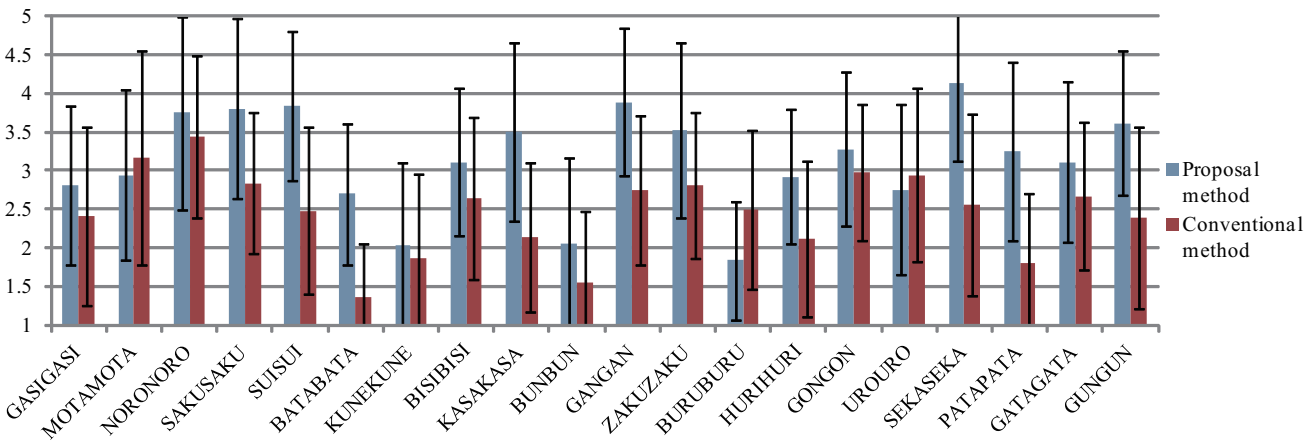


Fig. 11. Evaluations for twenty types of onomatopoeias.

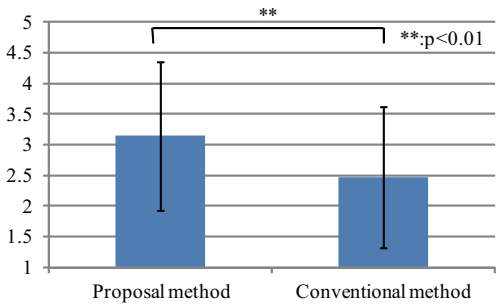


Fig. 12. Experimental results.

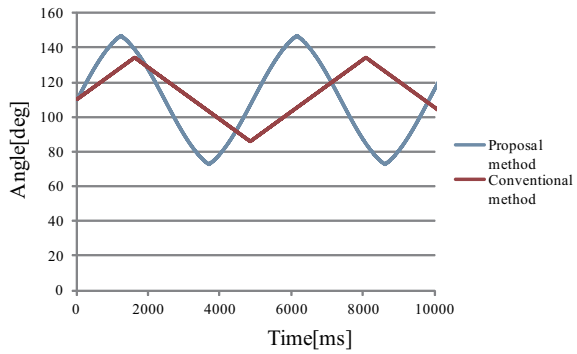


Fig. 13. Waveform for "MOTAMOTA."

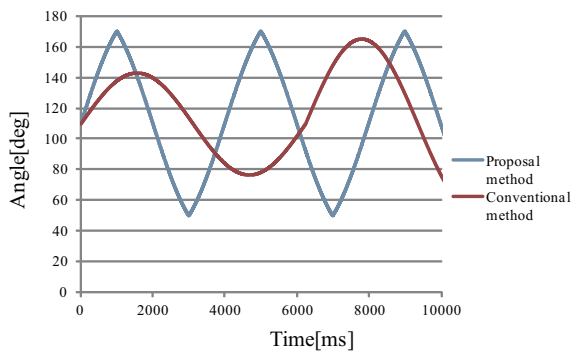


Fig. 14. Waveform for "BURUBURU."

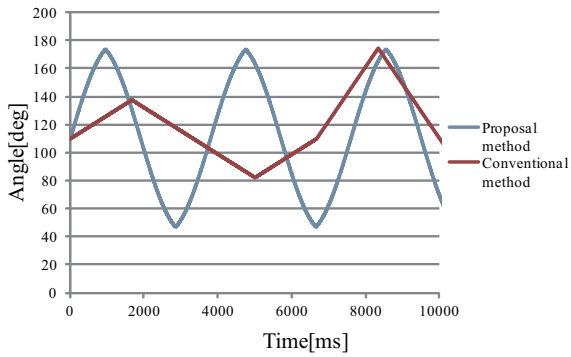


Fig. 15. Waveform for "UROURO."

explains why the proposed method fails to well represent the irregular images that are closely associated with "BURUBURU" and "UROURO," thus lowering the evaluations for these onomatopoeias. In other words, for the representation of the images for onomatopoeias, it may be useful to change the amplitudes of the waveforms according to the periods.

4.2. Evaluation of Accuracy of Operation Plane

To evaluate the accuracy of the operation plane, we have constructed an operation plane using the top ten types of onomatopoeias that have gained high evaluations in the experiments described in Sub-section 4.1. First, we seek the P-representation $O_i(\hat{c_N}(k))$ for onomatopoeia i. Next, the specters are normalized by the following equation:

$$normalize(O_i(\hat{c}_N(k))) = \frac{O_i(\hat{c_N}(k)) - \min_j O_j(\hat{c_N}(k))}{\max_j O_j(\hat{c_N}(k)) - \min_j O_j(\hat{c_N}(k))}. \quad . \quad . \quad (22)$$

The operation plane is constructed through auto-associative learning with such normalized specters used for input and teacher signals ($N = 16$).

The constructed operation plane is shown in **Fig. 16**. **Fig. 16** also shows the curve reproduced from the out-

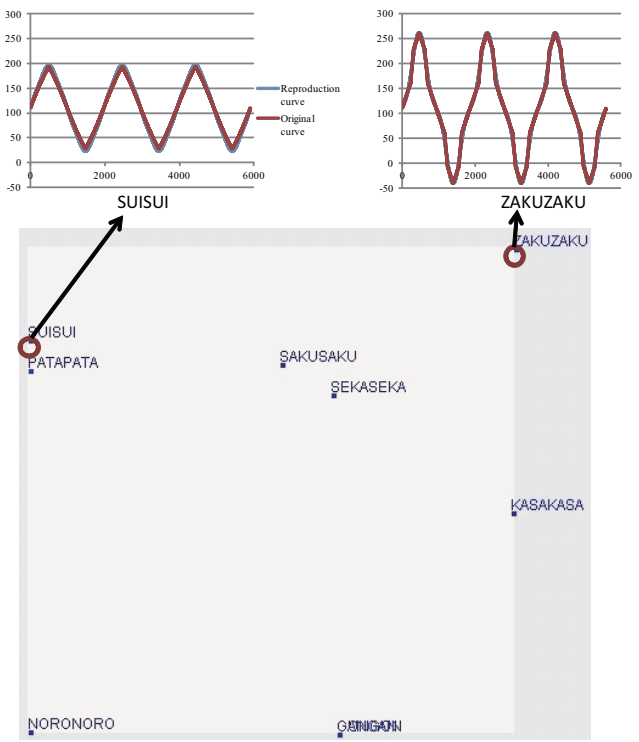


Fig. 16. Operation plane.

put in the fifth layer (output curve) and the original curve. For the ten types of onomatopoeias subjected to auto-associative learning, we have confirmed that their mean squared errors at the control points on the output curve and the original curve are $6.5^\circ \times 10^{-4}$, being nearly equal to each other. This proves that proper learning of the motion parameters has been accomplished by auto-associative learning.

5. Conclusion

In this study, we have proposed an operation plane system to edit robot motions using four-dimensional attribute vectors as an objective index based on the phonemic features of onomatopoeias. Through evaluation experiments, we have verified that the proposed system can generate robot motions that will conform to the phonemic features of onomatopoeias. The proposed system can contribute considerably toward reducing the burden on editors who are engaged in editing robot motions.

In this study, we have attempted, using the proposed system, to represent the onomatopoeias for simple motions after providing some phonemic features to a one-DOF motion, and as a result, we have found that some onomatopoeias failed to gain high evaluations. This may be attributed to the limits of the phonemic features of the onomatopoeias that can be represented by such one-DOF motion. In the future, we will need to further evaluate the proposed system for phonemic features for moving more than one joint.

References:

[1] J. Ito, R. Arisawa, M. Kanoh, T. Nakamura, and T. Komatsu, "Operation Generation of the Intuitive Robot by the Operation Plane Using the Neural Network," Annual Conf. of the Japanese Society for Artificial Intelligence, 2012 (in Japanese).

[2] T. Komatsu and H. Akiyama, "Expression System of Onomatopoeias for Assisting Users' Intuitive Expressions," The Trans. of the Institute of Electronics, Information and Communication Engineers, Vol.J92-A, No.11, pp. 752-763, 2009 (in Japanese).

[3] T. Komatsu, "Quantifying Japanese Onomatopoeias: Toward Augmenting Creative Activities with Onomatopoeias," Augmented Human Int. Conf., 2012.

[4] K. Kanbara and K. Tsukada, "Onomatopen: Painting Using Onomatopoeia," Int. Conf. on Entertainment Computing, pp. 43-54, 2010.

[5] H. Terashima and T. Komatsu, "MOYA-MOYA Drawing, Development of A Drawing Tool That Can Utilize Users' Expressed Onomatopoeias as An Drawing Effect," Annual Conf. of the Japanese Society for Artificial Intelligence, 2012 (in Japanese).

[6] Y. Ueda, Y. Shimizu, and M. Sakamoto, "System Construction Supporting Medical Interviews with Foreign Doctors Using Onomatopoeia Expressing Pains," Int. Workshop on Modern Science and Technology, pp. 132-136, 2012.

[7] K. Lertsumruaypun, C. Watanabe, and S. Nakamura, "Onomatoperori: Recipe Recommendation System Using Onomatopoeic Words," IEIEC 2nd technical committee document, 2010.

[8] Y. Uesaka, "Spectral Analysis and Complexity of Form," Symposium on Applied Functional Analysis, pp. 18-29, 1985.

[9] Y. Uesaka, "Spectral Analysis of Form Based on Fourier Descriptors," Int. Symposium for Science on Form, pp. 405-412, 1986.

[10] M. Kanoh, S. Kato, and H. Itoh, "Efficient Joint Detection Considering Complexity of Contours," Lecture Notes in Artificial Intelligence, Vol.1886, pp. 588-598, 2000.

[11] C. M. Bishop, "Neural Networks for Pattern Recognition," Oxford University Press, 1995.

[12] F. Kawakami, S. Morishima, H. Yamada, and H. Harashima, "Construction and Psychological Evaluation of 3-D Emotion Space," J. of Biomedical Fuzzy Systems Association Biomedical Fuzzy and Human Sciences, Vol.1, No.1, pp. 33-42, 1995.

[13] M. Kanoh, S. Iwata, S. Kato, and H. Itoh, "Emotive Facial Expressions of Sensitivity Communication Robot "Ifbot"," Kansei Engineering Int., Vol.5, No.3, pp. 35-42, 2005.

[14] M. Kanoh, T. Nakamura, S. Kato, and H. Itoh, "Affective Facial Expressions Using Auto-associative Neural Network in Kansei Robot "Ifbot"," Y. Dai et al. (Eds.), Kansei Engineering and Soft Computing, Theory and Practice, pp. 215-236, IGI Global, 2010.

[15] N. Ueki, S. Morishima, H. Yamada, and H. Harashima, "Expression Analysis/Synthesis System Based on Emotion Space Constructed by Multilayered Neural Network," Systems and Computers in Japan, Vol.25, No.13, pp. 95-107, 1995.

Name:
Junki Ito

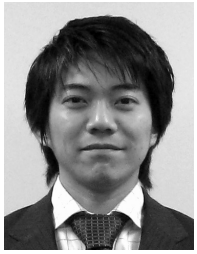
Affiliation:
Graduate School of Computer and Cognitive Sciences, Chukyo University

Address:
101 Tokodachi, Kaizu-cho, Toyota, Aichi 470-0393, Japan

Brief Biographical History:
2012 Received B.S. degree from Chukyo University
2012- Graduate School of Computer and Cognitive Sciences, Chukyo University

Main Works:
- "An Operation Plane Using a Neural Network for Intuitive Generation of Robot Motion," Int. Conf. on Soft Computing and Intelligent Systems, 2012.

Membership in Academic Societies:
- Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT)
- The Robotics Society of Japan (RSJ)



Name:
Masayoshi Kanoh

Affiliation:
Associate Professor, Department of Mechanics and Information Technology, School of Information Science and Technology, Chukyo University

Address:
101 Tokodachi, Kaizu-cho, Toyota, Aichi 470-0393, Japan

Brief Biographical History:
2004 Received Ph.D. degree from Nagoya Institute of Technology
2004-2010 Assistant Professor, Chukyo University
2010- Associate Professor, Chukyo University

Main Works:
• "Evolutionary Multi-valued Decision Diagrams for Obtaining Motion Representation of Humanoid Robots," IEEE Trans. on Systems, Man and Cybernetics, Part C, Vol.42, No.5, pp. 653-663, 2012.

Membership in Academic Societies:
• The Institute of Electrical and Electronics Engineers (IEEE)
• The Robotics Society of Japan (RSJ)
• Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT)

Name:
Tsuyoshi Nakamura

Affiliation:
Associate Professor, Department of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

Address:
Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan

Brief Biographical History:
1998 Ph.D., Nagoya Institute of Technology
1998- Research Associate, Nagoya Institute of Technology
2003- Associate Professor, Nagoya Institute of Technology

Main Works:
• "Estimating Subjective Assessments using a Simple Biosignal Senso," World Congress on Computational Intelligence (WCCI 2012, FUZZ-IEEE), pp. 1248-1253, 2012.

Membership in Academic Societies:
• The Institute of Electrical and Electronics Engineering (IEEE)
• Association for Computing Machinery (ACM)
• The Institute of Electronics, Information and Communication Engineers (IEICE)
• Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT)

Name:
Takanori Komatsu

Affiliation:
Associate Professor, Department of Kansei Engineering, Faculty of Textile Science and Technology, Shinshu University

Address:
Tokida 3-15-1, Ueda, Nagano 386-8567, Japan

Brief Biographical History:
2003 Ph.D., The University of Tokyo
2003- Research Associate, Future University Hakodate
2007- Assistant Professor, Shinshu University
2012- Associate Professor, Shinshu University

Main Works:
• "Artificial Subtle Expressions: Intuitive Notification Methodology of Artifacts," In Proc. of the 28th ACM Conf. on Human Factors in Computing Systems (CHI2010), pp. 1941-1944, 2010.

Membership in Academic Societies:
• Association for Computing Machinery (ACM)
• Information Processing Society of Japan (IPSJ)
• The Japanese Society for Artificial Intelligence (JSAI)
• Japanese Cognitive Science Society (JCSS)
• Human Interface Society (HIS)