The Journal of the Society for Art and Science, Vol. 16, No. 5, pp. 138 – 153 (2017)

Motion Interpolation Using Adjectives and Motion Features

Masaki Oshita [1)] **Aoi Honda** [1)]
Maho Katsurada [1)] **Yuya Aosaki** [1)]

1) Kyushu Institute of Technology

oshita@ces.kyutech.ac.jp

Abstract
In this paper, we propose a motion interpolation method using parameters based on adjectives. We conducted a questionnaire experiment in which subjects were asked to evaluate 10 example walking motions using 41 pairs of adjectives. Based on the results, 27 pairs of adjectives that are effective for motion parameterization were selected and four primary parameters were determined by categorizing the adjectives. Our motion interpolation method allows the user to create various styles of motions using the four primary parameters and any combination of the additional 27 adjective pairs. We also conducted a questionnaire experiment on another motion set of 27 walking motions and obtained the same four primary parameters and 29 adjective pairs. These results show that our method and the four primary parameters can be generalized for walking motions. In addition, our approach was applied to motion features that are computed example motions without any questionnaire experiment and obtained the primary and additional motion feature parameters for motion interpolation. This realizes motion interpolation using parameters based on motion features. The user of our system can choose adjective or motion feature parameters to create a motion depending on the type of motion that he or she wants to create. We present the results of our experiments and demonstrate the advantages of our method.

1 Introduction

Motion interpolation is a common technique in computer animation and can be used for generating a new motion from a set of existing motions through some control parameters. Normally a small number of quantitative parameters such as the position of the end-effector, walking speed, and direction are used. However, animators often want to edit motion using adjectives. For example, they might want to create a motion that looks happy, depressed, lively, or weak.

In this paper, we propose a motion interpolation method using parameters based on adjectives. In general, there are many adjectives, and it is not easy to choose appropriate ones for motion interpolation. We hence conducted a questionnaire experiment where subjects were asked to evaluate ten example walking motions using 41 pairs of adjectives. Based on the results, 27 pairs of adjectives that are effective for motion parameterization were selected and four primary parameters were determined by categorizing the adjectives. Our motion interpolation method allows the user to create various styles of motion using the four primary parameters and any combinations of the additional 27 adjectives pairs. We also conducted a questionnaire experiment on another motion set of 27 walking motions and obtained the same four primary parameters and 29 adjective pairs. These results show that our method and the four primary parameters can be generalized for walking motions. In addition, our approach was applied to motion features that are computed example motions without any questionnaire experiment and obtained the primary and additional motion feature parameters for motion interpolation. This realizes motion interpolation using parameters based on motion features. The user of our system can choose adjective or motion feature parameters to create a motion depending on the type of motion that he or she wants to create. We present the results of our experiments and demonstrate the advantage of our method.

Our approach can be applied to any kind of motion. In this research, we chose walking motion as an example because walking is a common human behavior that can express various styles. In addition, it is a cyclic motion and can be played back repeatedly, which makes it easy to observe during experiments. We applied our method to 10 example motions, which is a relatively small number. In practice, it is not easy to create a large number of examples of a specific kind of motion with many styles. Unlike previous research that required many example motions, our method works well with this small number of example motions.

The main contributions of this paper are as follows. We propose an approach for the quantification of adjectives for motion interpolation. We also determined four primary parameters based on adjectives through our experiments. We developed a motion interpolation system with the four primary parameters and any combination of adjective pairs. Moreover, we applied our approach to motion features and determined the four primary parameters. This allows the user to chose the subjective (adjectives) or the objective (motion features) parameters to create motions. Our approach and the primary parameters should be applicable to other kinds of motion with various styles.

This paper is an extended version of our previous work [1]. We experimented our method on another motion set with a larger number of walking motions to show that our approach can be generalized (Section 3). We also applied our method to motion features (Section 5). Furthermore, we conducted an additional experiment to compare the adjective and motion feature parameters (Section 6).

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 describes the quantification and classification of adjectives based on our questionnaire experiment. Section 4 explains our implementation of motion interpolation. Section 5 explains the quantification and classification of motions features. Section 6 presents the experimental results and discussion. Section 7 concludes this paper.

2 Related Work

Motion interpolation techniques are used in computer animation [2, 3, 4, 5, 6, 7]. A motion interpolation generates a new motion from a set of example motions by blending them. A feature vector is assigned to each example motion in advance. Given a desired feature vector, the blending weights of these example motions are computed based on their feature vectors. By blending these example motions with the weights, a new motion is synthesized. Several approaches have been proposed for computing the blending weights. Rose et al. [3] combined linear approximation and non-linear adjustments with radial basis functions. This approach has been adapted by many researchers [5]. Wiley and Hahn [2] combined linear interpolations of nearby examples around the specified parameter. This approach also has been adapted by other researchers [4, 6]. However, it requires a large number of dense examples over the parameter space. To solve this problem, Kovar et al. [6] generated many examples by interpolating existing examples. Mukai and Kuriyama [7] introduced a geostatistical model for statistically estimating the correlations between feature and motion spaces. Lau et al. [8] modeled a set of example motions by their spatial temporal variations. Min et al. [9] applied principle component analysis to example motions to construct a low-dimensional statistical model for generating a motion by determining the blending weights of the principal components to satisfy the given constraints. In our research, because our primary contribution is our parameterization of adjectives and any motion interpolation method can be combined with our system, we use the standard approach [3, 5].

To apply motion interpolation, the feature space must be defined and feature vectors must be assigned to example motions. Various kinds of feature vectors have been used in previous studies. Basically, the dimension of the feature space needs to be small compared to the number of example motions. Walking speed and turning direction are used as feature vectors for walking and running motions [3, 5]. The position of the end-effector is used as a feature vector for reaching, punching, and kicking motions [6, 7]. Rose et al. [3] parameterized different styles of motions using parameters based on adverbs such as happy, sad, and angry. Their features are similar to ours. However, they used only a small number of adverbs that the authors had chosen. In contrast, our method allows the user to utilize any combination of adjectives and the primary parameters that are derived from many adjectives. Adverbs and adjectives have the same role and both can express styles of motions. In addition, our method can be applied to both. However, in this paper, we use adjectives rather than adverbs, because a mixture of adjectives and adverbs is confusing and adjectives are naturally used to express the styles of motions in Japanese, the language in which our questionnaire experiment was conducted.

Recently, Förger and Takala [10] proposed a method for motion interpolation using verbal expressions and applied it to walking motion. Unlike our method, instead of verbal expression input, they used motion features that are computed from example motions and are associated with the verbal expressions as motion interpolation parameters. They used 35 example waking motions and 13 verbal expressions such as fast, slow, aggressive, lazy, and excited in their experiment. They introduced an incremental and iterative process for motion interpolation with a pseudo-inverse matrix. Their approach requires more motions than the effective motion features and also requires more manual annotation than our method. Moreover, as the output motion depends on the series of inputs, it is difficult to reproduce the same motions.

Quantifying a subjective factor using a questionnaire experiment is a common approach, although it is not straightforward and careful design is required. Komatsu [11] quantified adjectives and onomatopoeias, which consist of consonants and vowels, using a questionnaire experiment. Although he presented an application for simple robot motion control, his quantification of adjectives was general-purpose and not intended for motion synthesis. Different methods and experiments are required for quantifying adjectives for motion interpolation.

We applied a hierarchical clustering method [12] to extract the primary parameters from the various adjectives that are used in our experiment. There are many approaches for extract-

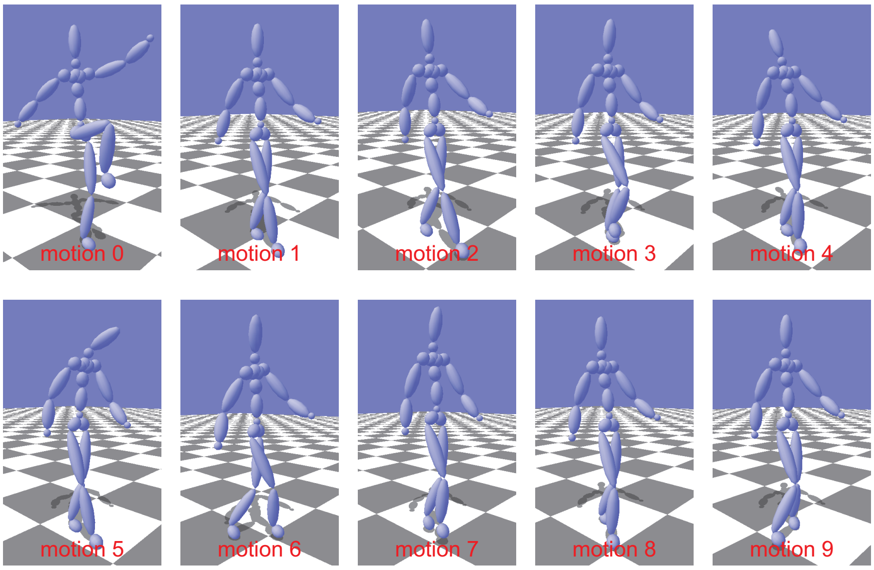


Figure 1: Example motions used in our experiment.

ing a small number of parameters from a large number of data. Principal component analysis is one popular technique. However, it does not suit our purpose, because it would extract the primary components across all the adjectives, the extracted components would become irrelevant to the adjectives, and it would be difficult to control these parameters. We rather choose to classify adjectives into a small number of groups so that the parameters based on these groups are intuitively controlled. Recently, machine learning techniques have been widely applied in many areas. For example, deep learning [13] can extract a small number of latent parameters from a large number of data by combining layers of neural networks. However, these techniques require a lot of training data, which are difficult to collect through a questionnaire-based approach and thus not applicable to our problem. However, the hierarchical clustering method [12] works well, even with a small number of data.

3 Quantification and Classification of Adjectives

We conducted a questionnaire experiment to quantify and classify adjectives. Two sets of walking motions were created and tested our method on both sets to evaluate if our method works successfully on different numbers of example motions.

3.1 Motion Data

As explained in Section 1, we chose walking motion as example for this research. Two sets of walking motions were created using the optical motion capture system OptiTrack. The motion set A contains 10 walking motions while the motion set B contains 27 walking motions. The walking motions consist of one cycle of straight walking at normal speed with various styles. A single cycle of motion can be played back in a loop to generate a continuous walking animation. These motions in the both motion sets are also presented in the accompanying video.

The motion set A contains 10 walking motions. A male subject of average body form was asked to perform various styles of walking. Because we focus on style control via adjectives, other quantitative factors such as speed and direction were fixed for all example motions. We tried to create as many variations of walking motion with distinctive styles. Finally, 10 walking motions were created. Figure 1 shows images from these example motions.

The motion set B contains 27 walking motions.

To create a large number of walking motions we introduced a systematic way. We assumed that motion style varies based on the type of emotion and speed. We defined 9 types of emotions based on Russell's model [14] that represents emotions on the circle in two dimensional space which is defined by arousal (activated - deactivated) and valence (pleasant - unpleasant) axes. This model has been widely used to classify emotions in computer graphics. The nine types of emotions are excitement, pleasure, contentment, sleepiness, depression, misery, distress, arousal and neutral. Each of nine emotions is combined with three types of speed, fast, normal and slow. As a result, 27 kinds of walking motions in total were defined. An another male subject of average body form was asked to perform these walking motions and captured them.

Our motion data contain the movements of the full body. They do not contain the movements of the fingers and face. Motion data are represented by a series of poses. Each pose is represented by the rotation of all joints and the position and orientation of the pelvis based on a hierarchical body model that is also constructed from the motion capture data. Our body model has 20 joints. The motion data have 30 frames/s.

3.2 Adjectives

We prepared 41 pairs of adjectives, as shown in Table 1. Normally, an adjective can be paired with another adjective that has the opposite meaning. Therefore, we form pairs of adjectives and treat each pair as one parameter. We selected adjectives that are commonly used for expressing styles of motions from the dictionary. With respect to the pairs of adjectives in [11], our research shares 33 pairs with the previous work and contains eight new pairs, namely, ugly-cute, heavy-light, unafraid-afraid, inelastic-elastic, thin-thick, old-young and dim-sharp. Ten pairs of adjectives in [11], for examples pesky-plain and wet-dry, were not considered applicable to motion and thus have been removed from our experiment.

Our questionnaire experiment was conducted in Japanese. All of the adjectives are in Japanese,

Table 1: List of all 41 pairs of adjectives used in our questionnaire experiment.

slow - fast	mild - violent
hard - soft	dull - quick
quiet - noisy	angular - circular
blunt - sharp	dormant - brisk
edgy - round	weak - strong
unpleasant - pleasant	inelastic - elastic
dirty - clean	sad - happy
bumpy - smooth	uncool - cool
dark - bright	thin - thick
ugly - cute	simple - gaudy
coarse - fine	old - young
heavy - light	strained - relaxed
unstable - stable	small - large
unafraid - afraid	artificial - natural
closed - open	adult - childish
vague - distinct	narrow - broad
ugly - beautiful	poor - rich
tight - free	dim - sharp
drab - clear	gloomy - cheerful
lonely - bustling	calm - excited
narrow - wide	

and the adjectives in this paper are translated versions of them. Although adjectives may not translate precisely from one language to another, in addition to the Japanese version of the interface, our system also provides English translated version of the interface.

3.3 Questionnaire Experiment

Questionnaire experiments on two motions sets were conducted separately with different subjects. Fourteen subjects who are computer engineering undergraduates and graduates participated in our questionnaire experiment. They were asked to evaluate the 41 pairs of adjectives for all the example motions on a 5-point scale (-2, -1, 0, 1 or 2). It took about 30 minutes for each subject to complete the task.

We developed a simple program to conduct our questionnaire experiment. Figure 2 shows an example screen shot. Each motion is displayed on

Figure 2: Our system for questionnaire experiment with which the subject can observe example motions and evaluate each pair of adjectives for each example motion. The green figure shows an example motion to be evaluated. The red figure shows the normal walking motion as reference.

the screen. The subject can select an adjective pair and evaluate it by clicking the corresponding button on the screen. The colors of the buttons represent the values that the subject entered. The subject can also control the camera to see the example motion from any angle. The lists of pairs of adjectives are shown on the screen.

3.4 Quantification of Adjectives

By taking the average of all answers of the subjects, we obtained the values of each pair of adjectives for each example motion. Of the 41 pairs of adjectives, the ones that are not effective for expressing the styles of the example motions were removed. If the average value of the answers from all subjects for the j-th pair of adjectives for all example motions is neutral, the pair of adjectives was considered ineffective. In addition, if the average distribution of the answers from all subjects for the j-th pair of adjectives for all example motions is high, this pair was also considered ineffective. These pairs of adjectives were removed. These conditions were evaluated by the following equations:

$$\mu_j^* = max\Big(|\mu_{ij}| : i = 0...N\Big), \tag{1}$$

$$\bar{\sigma}_j = \frac{\sum_{i=1}^{N} \sigma_{ij}}{N}, \tag{2}$$

where μ_{ij} is the average value of the answers from all subjects for the j-th pair of adjectives for the i-th example motion between -2.0 and 2.0, and σ_{ij} is its distribution. Further, N is the number of example motions. The j-th pair of adjectives was removed if the following condition was satisfied:

$$\Big(\mu_j^* < 1.0\Big) \vee \Big(\bar{\sigma}_j > 0.85\Big). \tag{3}$$

These thresholds were determined empirically. They could be adjusted based on the number of parameters to be used for motion interpolation. Through this process, 14 pairs of adjectives were removed and 27 pairs of adjectives remained, as shown in Table 2.

The average values between -2.0 and 2.0 were scaled between 0.0 and 1.0. Finally, we obtained impression matrix $\mathbf{A}$, which represents the coefficients between the 27 pairs of adjectives and 10 example motions.

3.5 Classification of Adjectives

To provide motion style control through a small number of parameters based on adjectives, we classified the pairs of adjectives into a small number of groups. We applied a popular hierarchical clustering method, Ward's method [12], which repeatedly combines the two clusters whose distance is the smallest of all pairs of clusters until a sufficient number of clusters are obtained. The distance E between two clusters is defined by the following equations:

$$\Delta E(G_i, G_j) := E(G_i \cup G_j) - E(G_i) - E(G_j), \tag{4}$$

$$E(G) = \sum_{\mathbf{a}_i \in G} d(\mathbf{a}_i, M(G)), \tag{5}$$

$$M(G) = \frac{1}{|G|} \sum_{\mathbf{a}_i \in G} \mathbf{a}_i, \tag{6}$$

where G_i and G_j are clusters, $M(G)$ is the center of a cluster and $d(\mathbf{a}_i, \mathbf{a}_j)$ is a distance function.

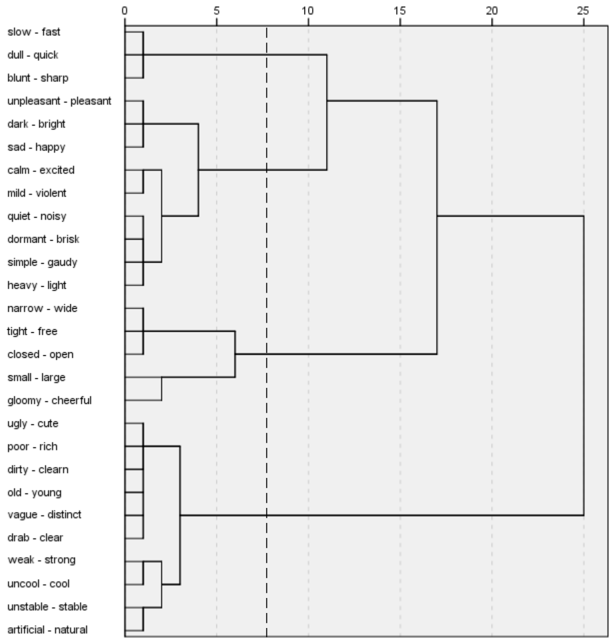


Figure 3: Constructed dendrogram representing the hierarchical clusters of 27 adjective pairs for motion set A.

We applied this method to the coefficient vectors of the pairs of adjectives from impression matrix **A**, obtained in Section 3.4. We used squared Euclidean distance, which is not a metric function but is a semi-metric function satisfying the relaxed triangle inequality, between the vectors as distance function $d(\mathbf{a}_i, \mathbf{a}_j)$.

Ward's method constructs a tree diagram called a dendrogram to illustrate the hierarchical arrangement of the clusters by repeating the process until all clusters are merged into one. Using the dendrogram, the number of clusters can be manually determined.

3.6 Results of Quantification and Classification

We have applied the above method on the collected answers from subjects for motion set A and B and constructed dendrograms in Figure 3 and 4, respectively. On both dendrograms, we chose to divide the pairs of adjectives into four clusters, as indicated by the dashed line in Figure 3 and 4, because they are clearly separated and each of them contains an adequate numbers

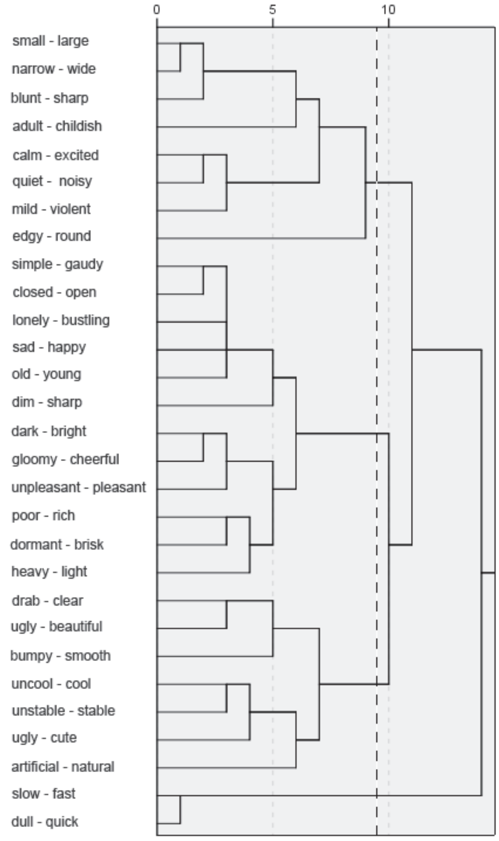


Figure 4: Constructed dendrogram representing the hierarchical clusters of 29 adjective pairs for motion set B.

of adjective pairs. Similar four clusters were obtained for two different sets of walking motions, although the remaining adjective pairs and their corresponding clusters are slightly different. This shows that the determined classification of adjective pairs can be generalized for walking motions. Table 2 and 3 shows pairs of adjectives in four clusters. We labeled these clusters "quickness," "clearness," "activeness," and "largeness."

Komatsu [11] classified pairs of adjectives into four groups labeled "sharpness," "softness," "dynamic," and "largeness" based on their factors with respect to onomatopoeias. There are some similarities between his classification and ours. However, they do not match exactly, because our classification is based on the styles of example motions, while his classification is based on onomatopoeia. Labanotation [15], which is a notation system for describing dancing movements,

Table 2: List of the selected 27 adjective pairs and their classification for motion set A.

adjectives	category
slow - fast	quickness
dull - quick	
blunt - sharp	
weak - strong	clearness
dirty - clean	
uncool - cool	
ugly-cute	
old - young	
unstable - stable	
artificial - natural	
vague - distinct	
poor - rich	
drab - clear	
calm - excited	activeness
mild - violent	
quiet - noisy	
dormant - brisk	
unpleasant - pleasant	
sad - happy	
dark - bright	
simple - gaudy	
heavy - light	
small - large	largeness
closed - open	
narrow - broad	
tight - free	
gloomy - cheerful	

Table 3: List of the selected 29 adjective pairs and their classification for motion set B.

adjectives	category
slow - fast	quickness
dull - quick	
drab - clear	clearness
ugly - beautiful	
bumpy - smooth	
uncool - cool	
unstable - stable	
ugly - cute	
artificial - natural	
simple - gaudy	activeness
closed - open	
lonely - bustling	
sad - happy	
old - young	
dim - sharp	
dark - bright	
gloomy - cheerful	
unpleasant - pleasant	
poor - rich	
dormant - brisk	
heavy - light	
small - large	largeness
narrow - wide	
blunt - sharp	
adult - childish	
calm - excited	
quiet - noisy	
mild - violent	
edgy - round	

defines four kinds of effort for the dynamic quality of movements: "space," "weight," "time," and "flow." They roughly corresponded to our factors "largeness," "activeness," "quickness," and "clearness," respectively. However, they do not match exactly either, because labanotation is specialized for dancing motions. These comparisons indicate that our four-class clustering has similarity with the classifications that have been designed in other applications and considered to be reasonable.

For motion interpolation, in addition to coefficient matrix $\mathbf{A}$ between the pairs of adjectives and example motions, coefficient matrix $\mathbf{A}'$ between the four clusters and example motions is necessary. We obtain them by taking the average of all the pairs of adjectives in each cluster.

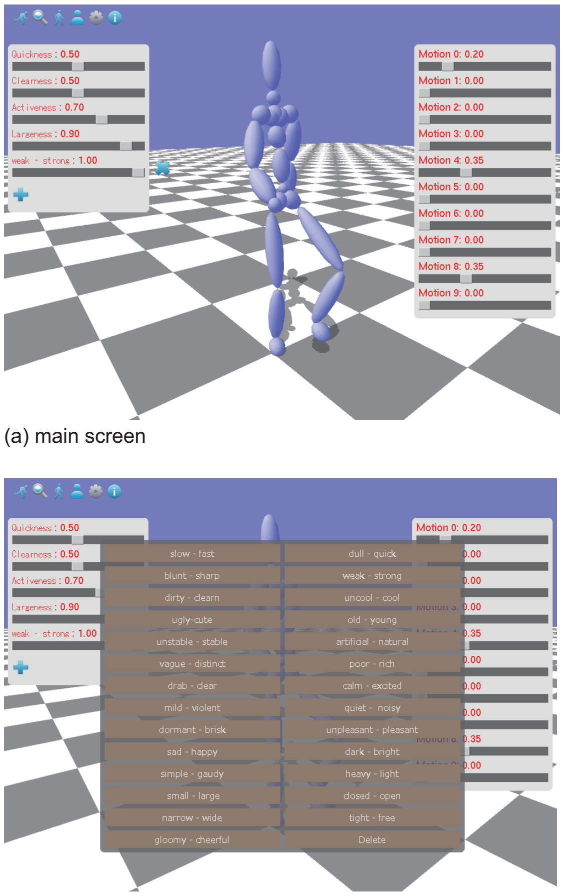
4 Motion Interpolation Using Adjectives

This section describes our method for motion interpolation using adjective-based parameters. To implement motion interpolation, we employed a common method that uses a combination of linear approximation and radial basis functions [3, 5].

4.1 User interface

A user of our method can specify values for the four primary parameters. In addition, the user can specify any additional pairs of adjectives and their values. All values are between 0.0 and 1.0. Although the quantification of adjectives is calculated based on the results of the questionnaire experiment in Japanese, our system also provides an English version of the interface. The English version provides the translated adjectives.

Figure 5(a) shows the interface of our prototype. The four primary parameters are controlled using the sliders on the left side of the screen. The user can add, delete, and alter a pair of adjectives by clicking the icons on the bottom and side of the sliders and choosing an item from the list of adjective pairs, as shown in Figure 5(b). As the user adjusts these parameters, the motion blending weights on the right side of the screen are automatically updated and the synthesized motion is also changed immediately. Alternatively, our system allows the user to control the blending weights directly by using the sliders on the right side of the screen.



(a) main screen

(b) selection of a pair of adjectives

Figure 5: User interface of our system for motion interpolation using adjectives. The user can control the adjective parameters through the sliders on the left side of the screen.

4.2 Motion Interpolation

Given the parameters $\mathbf{p}$ in M-dimensional parameter space, the blending weights $\mathbf{w}$ of N motions are computed by a combination of linear approximation and non-linear adjustments with radial basis functions. The i-th component of the weights is computed by

$$\mathbf{w}_i = \sum_{j=0}^{M} l_{ij} L_j(\mathbf{p}) + \sum_{k=1}^{N} r_{ik} R_k(\mathbf{p}), \qquad (7)$$

where $L_j(\mathbf{p})$ and l_{ij} are the linear basis and coefficients, respectively. Here, $L_j(\mathbf{p})$ is the j-th component of $\mathbf{p}$ ($j = 1...M$) and $L_0 = 1$. The linear coefficients are computed from the parameters of the example motions by solving the least squares problem with the sub-matrix of the coefficient matrix $\mathbf{A}'$ from Section 3.5. In addition, $R_k(\mathbf{p})$ and r_{ik} are respectively the non-linear radial basis and coefficients. Each $\mathbf{w}_i$ is bound between 0.0 and 1.0 and $\mathbf{w}$ is normalized. Because coefficients l_{ij} and r_{ij} depend on the parameter space, they are recomputed when the combination of adjectives is changed. For details of the algorithm, readers may refer to the previous work [3, 5].

Using the determined blending weights, the example motions are blended. To compute pose $\mathbf{q}(t)$ of the output motion at time t, the corresponding poses $\mathbf{q}_i(t_i)$ of example motion i at t_i are blended. The corresponding timing for each example motion t_i is computed by applying a time-warping based on the keytimes of example motions and keytimes of the blended motions. The keytimes of the blended motions are determined by taking a weighted average of the keytimes of

Table 4: List of 51 motion features and their classification for motion set B.

$\sigma_{torso\_left-right}$	lateral swing of the upper body
$\sigma_{head\_left-right}$	
$\sigma_{com\_left-right}$	
$\sigma_{torso\_front-back}$	
$\sigma_{div\_head_{f}ront-back}$	
$\sigma_{left-upper-arm\_front-back}$	front-back swing of the arms and legs
$\sigma_{left-forearm\_front-back}$	
$\sigma_{right-upper-arm\_front-back}$	
$\sigma_{right-forearm\_front-back}$	
$\sigma_{left-thigh\_front-back}$	
$\sigma_{right-thigh\_front-back}$	
$\sigma_{com-position\_vertical}$	
$\sigma_{com-velocity\_vertical}$	
$\sigma_{left-thigh\_left-right}$	
$\sigma_{right-thigh\_left-right}$	
$\sigma_{left-upper-arm_{l}eft-right}$	
$\sigma_{right-upper-arm_{l}eft-right}$	
$\mu_{torso\_front-back}$	bending angle of the upper body
$\mu_{head\_front-back}$	
$\mu_{left-upper-arm\_front-back}$	
$\mu_{right-upper-arm\_front-back}$	
$\mu_{left-forearm\_front-back}$	
$\mu_{right-forearm\_front-back}$	
$\mu_{left-thigh\_front-back}$	
$\mu_{right-thigh\_front-back}$	
$\mu_{torso\_left-right}$	
$\mu_{head\_left-right}$	
$\mu_{left-upper-arm\_left-right}$	
$\mu_{right-upper-arm\_left-right}$	
$\mu_{left-thigh\_left-right}$	
$\mu_{right-foot\_left-right}$	
$\mu_{com-velocity\_left-right}$	
$v_{step-length\_left-right}$	
$v_{step-length\_front-back}$	

Table 5: List of 51 motion features and their classification for motion set B (continued).

$v_{step-length\_front-back}$	step length
$\mu_{com-velocity\_front_{b}ack}$	
$\sigma_{com-velocity\_front_{b}ack}$	
$\mu_{com-position\_vertical}$	
$v_{motion-duration}$	
$\mu_{right-thigh\_left-right}$	
$\mu_{left-shin\_front-back}$	
$\mu_{right-shin\_front-back}$	
$\mu_{left-foot\_front-back}$	
$\mu_{right-foot\_front-back}$	
$\mu_{left-foot\_left-right}$	
$\sigma_{left-shin\_front-back}$	
$\sigma_{right-shin\_front-back}$	
$\sigma_{right-foot\_front-back}$	
$\sigma_{left-foot\_left-right}$	
$\sigma_{right-foot_{l}eft-right}$	
$\sigma_{left-foot_{l}eft-right}$	

example motions with the blending weights.

5 Motion Interpolation Using Motion Features

We applied our approach to motion features that are computed from example motions without any questionnaire experiment and obtained the primary and additional motion feature parameters for motion interpolation. This realizes motion interpolation using the parameters based on motion features.

The list of motion features are show in Table 4 and 5. These features are computed from each example motion which consists of a series of poses. μ_i represents the average of angle, position or velocity in a certain direction of the i body part or the center of mass (com) in the world coordinates over the example motion. The position of the center of mass is computed from the positions of all body parts and an average body model. σ_i represents the distribution of these values. v_i represents a single value that is computed for each example motion. μ_i reflects the average pose during motion, while σ_i reflects the magnitude of the movement during motion. In total, we obtain 51 motion features for each example motion.

We applied the classification method in Section 3.5 on the motion features from 27 example motions. Because each motion feature has different range unlike the adjective pairs, it is difficult to determine if the motion features is important. Therefore, we didn't apply our method for eliminating unimportant parameters in Section 3.4

and classified all 51 motion features. Using the similar approach with the adjective pairs, a dendrogram was constructed and the motion features were divide into four clusters. We labeled these clusters "lateral swing of the upper body", "front-back swing of the arms and legs", "bending angle of the upper body" and "step length" as shown in Table 4 and 5. We also obtained the coefficient matrix A' for motion interpolation.

6 Results and Discussion

To evaluate the validity and effectiveness of our method, we conducted a series of experiments. On our earlier experiment in our previous work [1], we used motion set A containing 10 example motions and compared our adjective parameters and blending weights. The subjects of our experiment were asked to create a motion that is similar to the presented target motion using each interface. The target motions were generated by determining primary adjective parameters randomly. The results showed that our interface with adjective parameters are better than using blending weights.

This paper presents the results of our second experiment where we used motion set B containing 27 example motions and compared three interfaces: adjective parameters, motion feature parameters and blending weights. Instead of generating target motions randomly, we prepared a set of fixed target motions, because we thought that appropriate interface may depend on the type of target motion. We generated many target motions randomly in the same way that is explained in [1] and selected five motions among them so that the target motions covers various kinds of walking motions. The five target motions are presented in the accompanying video.

6.1 Experimental Procedure

For each trial of the experiment, a target motion was randomly generated using our system. The subjects of our experiment were asked to create a motion that is similar to the presented target motion using our interface with adjective parameters, using our interface with motion feature parameters, and by adjusting the blending weights directly for comparison.

For each target motion, the subjects were asked to complete the following three steps, as shown in Figure 6.

1. They were asked to determine the adjective parameters, the motion feature parameters, and blending weights without seeing the synthesized motions. The distances between the target and synthesized motion for the adjective parameters, the motion feature parameters, and blending weights, D_p, D_f and D_w, respectively, were measured.

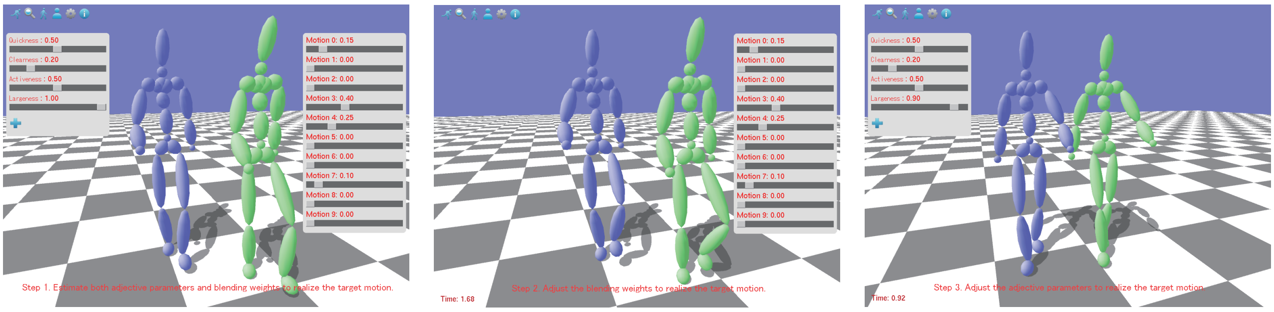
2. They were asked to create a motion that matched the target motion by adjusting the blending weights, if they could not generate it with their initial guess. The required time T_w to complete this step was measured.

3. In the same way, they were asked to create a motion that matched the target motion by adjusting the adjective parameters. The required time T_p was measured.

4. They were asked to create a motion that matched the target motion by adjusting the motion feature parameters. The required time T_f was measured.

For steps 2 and 3, if a subject could not recreate the motion in 300 seconds, the step was terminated. We introduced this time limit because subjects could not reproduce the target motion, no matter how long they tried when the target motion was difficult to create. We also collected comments from the subjects after the experiment.

Thirteen subjects who are computer engineering undergraduates and graduates participated in our experiment. None of these subjects had experience in making computer animation. Before the experiment, the subjects were told how the interfaces worked and given enough time to practice. The subjects were asked to create three random target motions to get used to the three interfaces before actual experiment. A video of all example motions was also presented to the subjects on the other screen to help them to adjust the blending weights. Each subject was asked to create five



(a) step 1: estimating both parameters (b) step 2: adjusting blending weights (c) step 3: adjusting motion parameters

Figure 6: Screen shots of our experiment. The green figure represents the target motion generated by our system. The blue figure represents the created motion.

target motions using the three interfaces. It took about 30 minutes for each subject to complete the experiment.

Steps 2 and 3 of each trial were completed when the recreated motion matched the target motion and the distance between them became lower than a threshold. Whether the two blended motions look similar cannot be simply determined based on their blending weights because different blending weights may generate similar motions. Therefore, we used the average distances of the positions of all joints over the two motions. Distance D was computed by

$$D = \sum_{t=0}^{T} \sum_{j=1}^{J} \left| \mathbf{x}_j^{target}\left(\frac{t}{T}\right) - \mathbf{x}_j\left(\frac{t}{T}\right) \right|, \qquad (8)$$

where T is the number of frames, J is the number of joints, and $\mathbf{x}_j^{target}(t')$ and $\mathbf{x}_j(t')$ are the position of the j-th joint at normalized time t' of the target and generated motions, respectively. Because the durations of the two motions are different, we used the normalized time and poses at keytimes. In our implementation, T is 10 and J is 20. The threshold was set to 0.06 m.

6.2 Experimental Results

The results are shown in Figures 7 to 11 in the same manner with our previous experiment [1]. The distances between the motion created by an initial guess using the two interfaces and target motion (D_p, D_w) are shown in Figure 7. The ratio of subjects who recreated the target motion

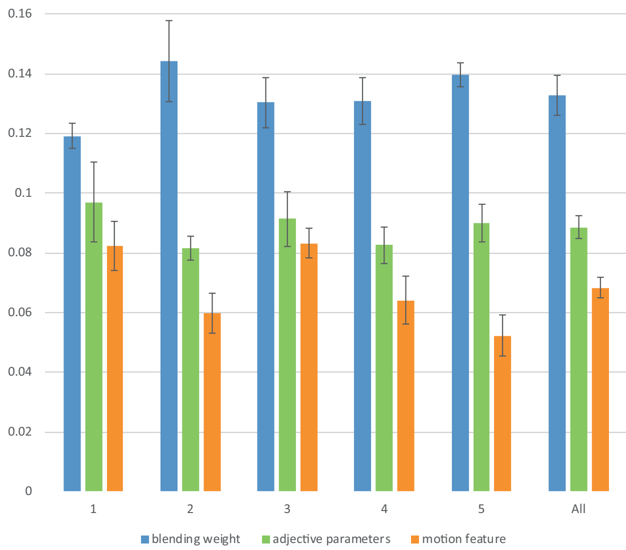


Figure 7: Average distance to the target motion from an initial guess created using both interfaces. Error bars represent one standard deviation, which is calculated under the assumption of independence between paired data.

with an initial guess is shown in Figure 8. We performed a Wilcoxon signed-rank test to compare the results of the adjective parameters and motion feature parameters. Because data were taken from the same person for the same target motion, they are paired data connected with each other. The Wilcoxon signed-rank test is a nonparametric paired difference test that is used as an alternative to the paired Student's T-test when the data cannot be assumed to be normally

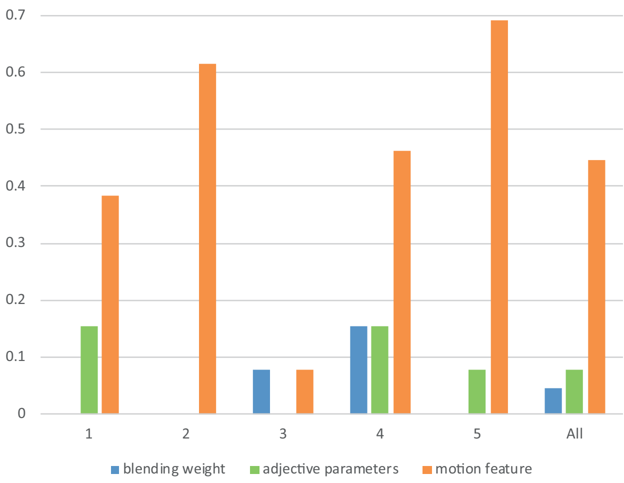


Figure 8: Ratio of subjects able to recreate the target motion with an initial guess.

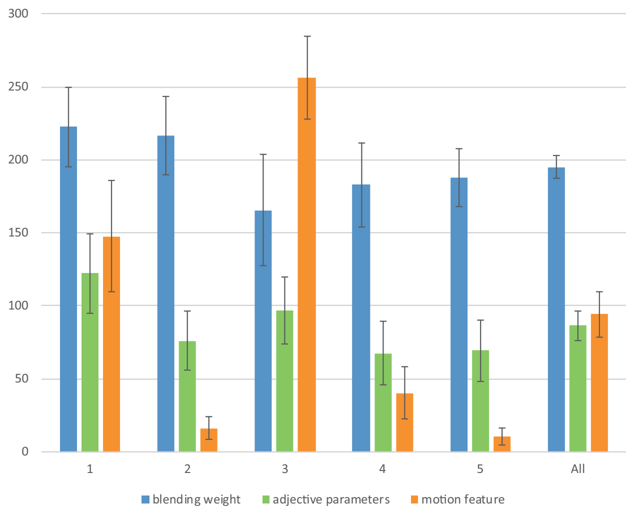


Figure 9: Average time required to create the target motion using both interfaces. Error bars represent one standard deviation, which is calculated under the assumption of independence between paired data.

distributed. The null hypothesis is that the median difference between the pairs of results of two interfaces is zero. The results of the Wilcoxon signed-rank test between the adjective parameters and the motion feature parameters were $p = 0.382, 0.011, 0.807, 0.087, 0.007$, and 0.000 for target motion 1, 2, 3, 4 and 5, and all, respectively. These results show that the motion feature parameters achieved that better results in target motion 2, 5 and all. In most cases, both interfaces were better than the blending weights, although there was no significant difference in some cases. The results of the Wilcoxon signed-rank test between the adjective parameters and the blending weights were $p = 0.133, 0.002, 0.019, 0.007, 0.004$, and 0.000 for target motion 1, 2, 3, 4 and 5, and all, respectively. The results of the Wilcoxon signed-rank test between the motion feature parameters and the blending weights were $p = 0.006, 0.002, 0.005, 0.003, 0.001$, and 0.000 for target motion 1, 2, 3, 4 and 5, and all, respectively.

The times required to create the target motion from the initial guess $(T_p$ and $T_w)$ are shown in Figure 9. Figure 10 shows the reduction in time when using our method compared with the time required when adjusting the blending weights. This is expressed as the mean of the deduction from the required time by using the blending weight and the 90% interval. The ratio of failures, which means that the subject could not make the target motion in time (within 300 seconds), are shown in Figure 11. These results show that the effective interface varies depending on the target motion. For target motion 2 and 5, the motion feature parameters achieved better results than the adjective parameters. For target motion 3, the adjective parameters achieved better results than the motion feature parameters. For target motion 1, 4 and all, there was no significant difference between the adjective parameters and the motion feature parameters. The results of the Wilcoxon signed-rank test between the adjective parameters and the motion feature parameters were $p = 0.695, 0.007, 0.002, 0.075, 0.009$, and 0.652 for target motion 1, 2, 3, 4 and 5, and all, respectively. In most cases, both interfaces were better than the blending weights, although there was no significant difference in some cases. The results of the Wilcoxon signed-rank test between the adjective parameters and the blending weights were $p = 0.034, 0.003, 0.075, 0.055, 0.08$, and 0.000 for target motion 1, 2, 3, 4 and 5, and all, respectively. The results of the Wilcoxon signed-rank test between the motion feature parameters and the blending weights were $p =$

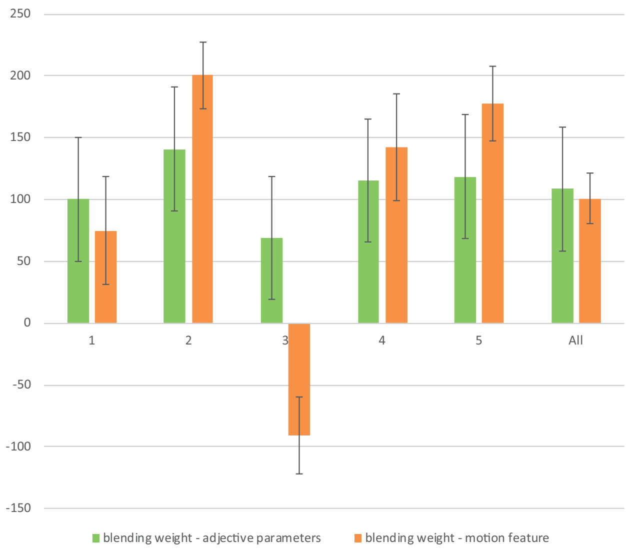


Figure 10: Average difference between the time required to use our method and that required to adjust the blending weights.

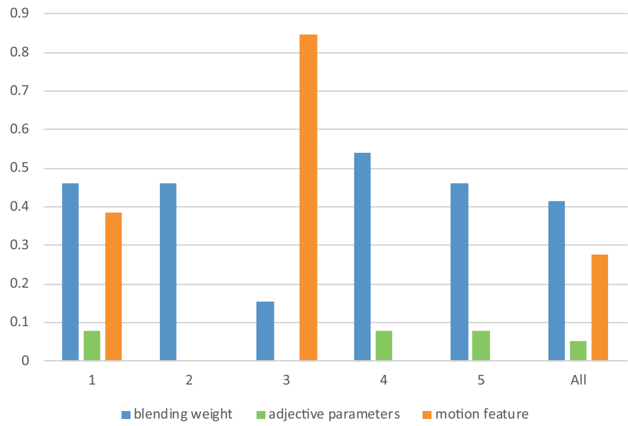


Figure 11: Ratio of failures to make the target motion.

0.201, 0.001, 0.013, 0.015, 0.001, and 0.000 for target motion 1, 2, 3, 4 and 5, and all, respectively.

These results show that the motion feature parameters are capable of creating a motion that is close to a desired motion. However, sometimes it is difficult to create the exact desired motion by using the motion feature parameters. On the other hand, the adjective parameters are capable of creating various desired motion successfully, although sometimes it requires more time than the motion features. The results show that whether the adjectives or the motion features are suitable depend on the type of target motion. The target motion with distinctive features in the average pose or body movements can be created easily by using objective motion feature parameters. On the other hand, the target motion without such distinctive features may be created easily by using subjective adjective parameters.

6.3 Discussion

In this experiment, most motions that can be created by our system are possible to create through the four primary adjective parameters without additional adjective parameters. This is probably because the degree of freedom of styles in the example walking motions was about four, that is, not particularly high. Therefore, the additional adjective parameters were not used to generate target motions in our experiment.

Our results are based on a set of example motions and a limited number of subjects. Different results may be achieved from a different group of subjects. However, the four primary adjective parameters in our results are evident, as discussed in Section 3.5. Therefore, they would be applicable to a different group of subjects.

In this research, our method is applied to walking motion. However, it could be applied to various kinds of motions. The four primary parameters are also considered suitable for other kinds of motion. The application of our method to other kinds of motion and other languages is a task for future work. When our method is applied to other kinds of motion, based on our findings, it is possible to conduct a questionnaire experiment with just the four primary parameters to reduce the effort of the subjects. However, a quantification process by human subjects is still required. It should be possible to estimate these parameters by analyzing motion data. The automatic quantification of example motions for adjective parameters is a possible direction for our future work.

The quality of synthesized motions by our method and by motion interpolation in general depends on the quality of example motions. Motion interpolation assumes that example motions are of good quality and covers the parameter space, i.e. the adjective parameter space in our method, well. In this research, we used Russell's emotion model [14] as explained in Section 3.1 to ask the actor to perform 27 example motions. There may be better way to collect s smaller number of good motions that cover the adjective parameter space well. Moreover, the best way may depend on the type of motion and the actor.

7 Conclusion

In this paper, we proposed a motion interpolation method using the parameters based on adjectives. Using our approach, various styles of motion can be controlled through intuitive adjective-based parameters from a number of precreated example motions. We applied our method on walking motions. Experimenting our method on other kinds of motions is our future work. Our method requires a questioner experiment to quantify adjectives for the target set of motions. We hope to develop a method for estimating the adjective-based features by analyzing motions. This is also our future work.

Acknowledgment

The authors would like to thank Hiroki Nakamura, Yosuke Higaki and Genki Ohishi for their help in conducting experiments. This work was supported in part by Grant-in-Aid for Scientific Research (No. 15H02704 and No. 15K05003) from the Japan Society for the Promotion of Science (JSPS).

References

[1] Masaki Oshita, Maho Katsurada, and Aoi Honda. Motion interpolation using adjectives. In *NICOGRAPH International 2016*, pages 36–43, 2016.

[2] Douglas J. Wiley and James K. Hahn. Interpolation synthesis for articulated figure motion. *IEEE Computer Graphics and Applications*, 17(6):39–45, 1997.

[3] Charles Rose, Michael F. Cohen, and Bobby Bodenheimer. Verbs and adverbs: Multidimensional motion interpolation. *IEEE Computer Graphics and Applications*, 18(5):32–40, 1998.

[4] Tadasuke Tsumura, Takeharu Yoshizuka, Takashi Nojirino, and Tsukasa Noma. T4: A motion-capture-based goal-directed realtime responsive lo-comotion engine. In *Computer Animation 2001*, pages 52–60, 2001.

[5] Sang Il Park, Hyun Joon Shin, Tae hoon Kim, and Sung Yong Shin. On-line motion blending for real-time locomotion generation. *Computer Animation and Virtual Worlds*, 15(3):125–138, 2004.

[6] Lucas Kovar and Michael Gleicher. Automated extraction and parameterization of motions in large data set. *ACM Transactions on Graphics (SIGGRAPH 2004)*, 23(3):559–568, 2004.

[7] Tomohiko Mukai and Shigeru Kuriyama. Geostatistical motion interpolation. *ACM Transactions on Graphics (SIGGRAPH 2005)*, 24(3):1062–1070, 2005.

[8] Manfred Lau, Ziv Bar-Joseph, and James Kuffner. Modeling spatial and temporal variation in motion data. *ACM Transactions on Graphics (SIGGRAPH Asia 2009)*, 28(5):Article No. 171, 2009.

[9] Jianyuan Min, Yen-Lin Chen, and Jinxiang Chai. Interactive generation of human animation with deformable motion models. *ACM Transactions on Graphics*, 29(1):Article No. 9, 2010.

[10] Klaus Förger and Tapio Takala. Animating with style: Defining expressive semantics of motion. *The Visual Computer*, 32(2):191–203, 2016.

[11] Takanori Komatsu. Quantifying Japanese onomatopoeias: Toward augmenting creative activities with onomatopoeias. In

Augmented Human International Conference 2012, page Article No. 15, 2012.

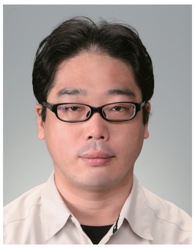
[12] Joe H. Ward. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244, 1963.

[13] Li Deng and Dong Yu. Deep learning: Methods and applications. *Foundations and Trends in Signal Processing*, 7(3-4):197–387, 2014.

[14] James A Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.

[15] Ann Hutchinson-Guest. *Labanotation: The System of Analyzing and Recording Movement.* 1954.

Masaki Oshita

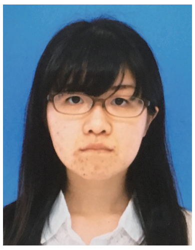


Masaki Oshita is an associate professor at Kyushu Institute of Technology. He received his BS, MS, and PhD degrees from Kyushu University in 1998, 2000, and 2003, respectively. His research interest is interactive computer animation which includes dynamic motion control of human figures, motion control interface, motion synthesis, physics-based simulation, human-computer interaction, computer vision, etc.

Aoi Hond

Aoi Hond is an associate professor at Kyushu Institute of Technology. The research interest is in the areas of nonadditive measure and integral, and its application, especially statistical data analysis.

Maho Katsurada



Maho Katsurada is a former undergraduate student at Kyushu Institute of Technology.

Yuya Aosaki

Yuya Aosaki is an undergraduate student at Kyushu Institute of Technology.