Motion Interpolation Using Adjectives

Masaki Oshita, Maho Katsurada and Aoi Honda
Kyushu Institute of Technology
680-4 Kawazu, Iizuka, Fukuoka, 820-8502, Japan
Email: 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, we selected 27 pairs of adjectives that are effective for motion parameterization and determined four primary parameters 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 present the results of our experiments and demonstrate the advantages of our method.**

Keywords*—*motion interpolation; adjective; questionnaire experiment

I. 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, we selected 27 pairs of adjectives that are effective for motion parameterization and determined four primary parameters 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 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. Our approach and the primary parameters should be applicable to other kinds of motion with various styles.

The remainder of this paper is organized as follows. Section II reviews related work. Section III describes the quantification and classification of adjectives based on our questionnaire experiment. Section IV explains our implementation of motion interpolation. Section V presents the experimental results and discussion. Section VI concludes this paper.

II. Related Work

Motion interpolation techniques are used in computer animation [1], [2], [3], [4], [5], [6]. 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. [2] combined linear approximation and non-linear adjustments with radial basis functions. This approach has been adapted by many researchers [4]. Wiley and Hahn [1] combined linear interpolations of nearby examples around the specified parameter. This approach also has been adapted by other researchers [3], [5]. However, it requires a large number of dense examples over the parameter space. To solve this problem, Kovar et al. [5] generated many examples by interpolating existing examples. Mukai and Kuriyama [6] introduced a geostatistical model for statistically estimating the correlations between feature and motion spaces. Lau et al. [7] modeled a set of example motions by their spatial temporal variations. Min et al. [8] 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 [2], [4].

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 [2], [4]. The position of the end-effector is used as a feature vector for reaching, punching,

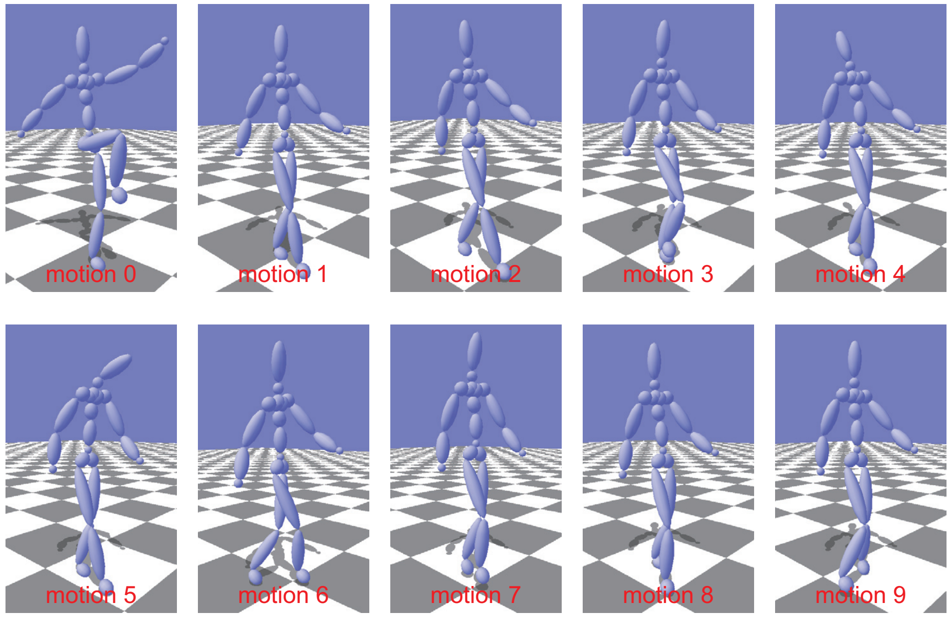


Fig. 1. Example motions used in our experiment.

and kicking motions [5], [6]. Rose et al. [2] 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 [9] 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 [10] 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 [11] to extract the primary parameters from the various adjectives that are used in our experiment. There are many approaches for extracting 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 [12] 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 [11] works well, even with a small number of data.

III. Quantification and Classification of Adjectives

We conducted a questionnaire experiment to quantify and classify adjectives.

A. Motion Data

As explained in Section I, we chose walking motion as example for this research. We created 10 walking motions using the optical motion capture system OptiTrack. The walking motions consist of one cycle of straight walking at normal speed with various styles. A male subject of average body form was asked to perform various styles of walking. Because we

TABLE I. 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

focus on style control via adjectives, other quantitative factors such as speed and direction were fixed for all example motions. Although we tried to create as many variations of walking motion as possible, we captured walking motions with distinctive styles. Finally, 10 walking motions were created. Figure 1 shows images from these example motions. A single cycle of motion can be played back in a loop to generate a continuous walking animation. These motions are also presented in the accompanying video.

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.

B. Adjectives

We prepared 41 pairs of adjectives, as shown in Table I. 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 [10], our research shares 33 pairs with the previous work and contains eight new pairs. Ten pairs of adjectives in [10] 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, and the adjectives in this paper are translated versions of them. Although adjectives may not translate precisely from one language to another, we consider that our approach and results in Japanese can be also applied to other languages as long as the adjectives are translated, because motions are basically universal. Our system provides both Japanese and English versions of the interface.

C. Questionnaire Experiment

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 the 10 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.

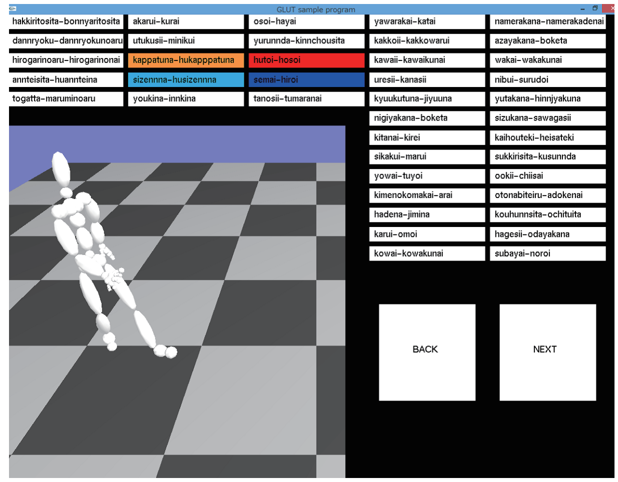


Fig. 2. Our system for questionnaire experiment with which the subject can observe example motions and evaluate each pair of adjectives for each example motion.

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

D. 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 II.

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.

E. 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 [11], 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. We applied this method to the coefficient vectors of the pairs of adjectives from impression matrix $\mathbf{A}$, obtained in Section III-D. 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. The dendrogram in Figure 3 was constructed. We chose to divide the pairs of adjectives into four clusters, as indicated by the dashed line in Figure 3, because they are clearly separated and each of them contains an adequate numbers of adjective pairs. Table II shows 27 pairs of adjectives in four clusters. We labeled these clusters "quickness," "clearness," "activeness," and "largeness."

Komatsu [10] 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 [13], which is a notation system for describing dancing movements, 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 is supported by the classifications that have been designed in other applications and is reasonable.

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

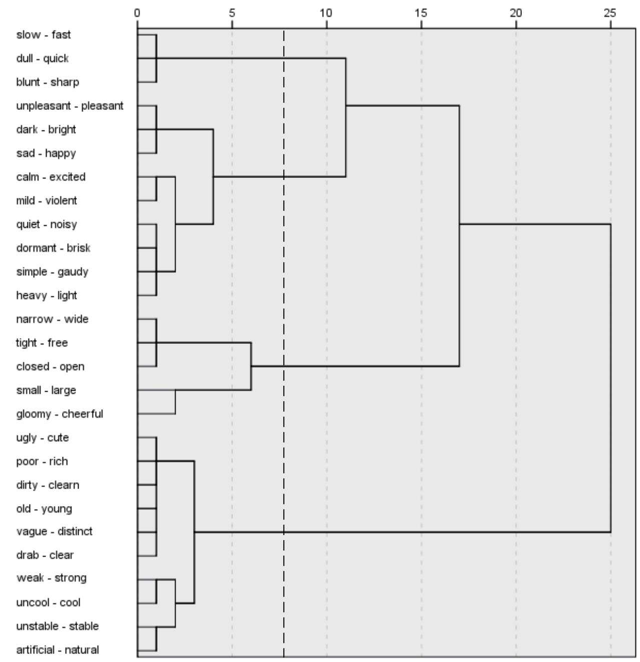


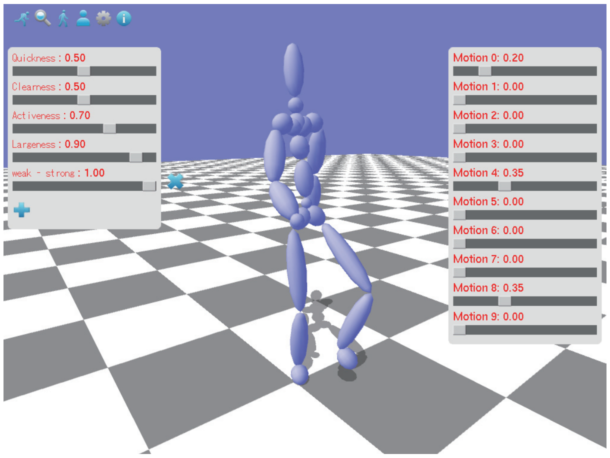
Fig. 3. Constructed dendrogram representing the hierarchical clusters of 27 adjective pairs.

TABLE II. LIST OF THE SELECTED 27 ADJECTIVE PAIRS AND THEIR CLASSIFICATION.

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	

IV. 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 [2], [4].



(a) main screen

(b) selection of a pair of adjectives

Fig. 4. 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.

A. 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 4(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 4(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.

B. 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

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

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 III-E. 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 [2], [4].

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 example motions with the blending weights.

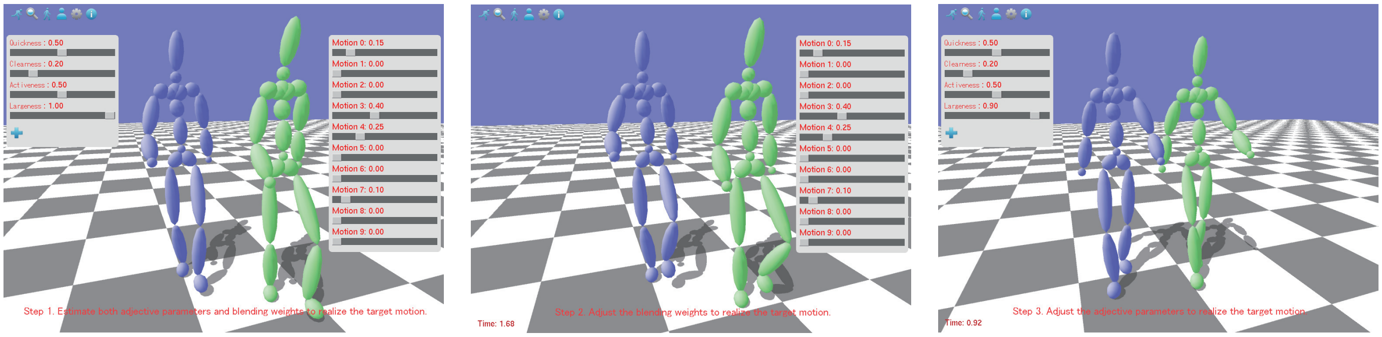
V. Results and Discussion

A. Experiment

To evaluate the validity and effectiveness of our method, we conducted an experiment. 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 and also by adjusting the blending weights directly for comparison. A video of ten example motions was also presented to the subjects on the other screen to help them to adjust the blending weights.

For each target motion, the subjects were asked to complete the following three steps, as shown in Figure 5.

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



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

Fig. 5. Screen shots of our experiment. The green figure represents the target motion generated by our system. The blue figure represents the created motion.

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.

The target motions were generated randomly as follows. Some primary adjective parameters were randomly given and other parameters were set to neutral (0.5). A target motion was then generated based on the adjective parameters. To avoid creating an easy target motion that could be recreated without changing the adjective parameters at all, if the generated motion was close to the neutral motion, which is generated by setting all parameters to neutral values (0.5), this process was repeated to create a target motion that was not close to the neutral motion. We created four kinds of target motions (Levels 1, 2, 3, and 4) depending the number of primary adjective parameters that were changed. A Level 1 target motion should be easily reproduced by adjusting one primary parameter, while a Level 4 target motion could require adjusting four primary parameters and should be difficult to reproduce. However, this algorithm was disclosed to the subjects. Therefore, they may have adjusted more parameters than necessary. Although we could generate target motions by randomly determining the blending weights, we did not take this approach, because we wanted to control the level of difficulty, and most of the motions were possible to recreate through the four primary adjective parameters. We discuss this further in Section V-C.

Nine 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. Each subject was asked to create one target motion from each level. Hence, they had four trials in total. 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

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

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.

B. Results

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 6. There was no significant distance between the results of the two interfaces. We performed a Wilcoxon signed-rank test to compare the results of the two interfaces. 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 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 were $p = 0.578, 0.499, 0.441, 0.878$, and 0.411 for Levels 1, 2, 3, and 4 and all levels, respectively. In all cases, p is larger than the significance level of 0.05, which means that the null hypothesis is not rejected and there was no significant difference between the pairs of results. In addition, the ratio of subjects who recreated the target motion with an initial guess is shown in Figure 7. The results show that the subjects reproduced the target motions by using adjective parameters in many cases compared with using the blending weights, even though there was no difference in the average distances. This is probably because using adjective parameters worked very well when the target motion matched the adjective parameters that the subject had thought of.

The times required to create the target motion from the initial guess $(T_p$ and $T_w)$ are shown in Figure 8. Figure 9 shows the reduction in time when using our method compared with the time required when adjusting the blending weights. This is

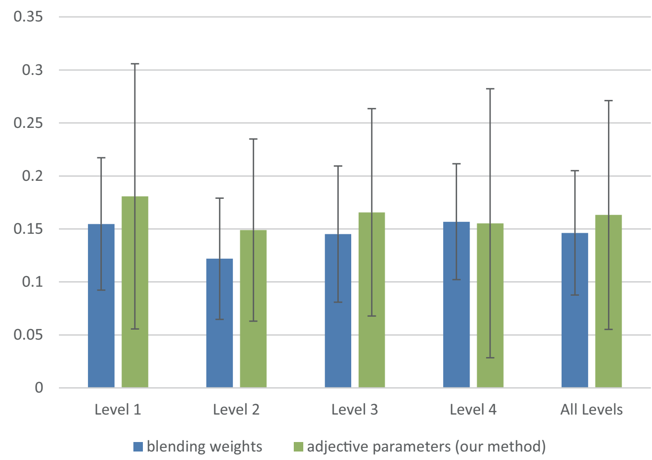


Fig. 6. Average distance to the target motion from an initial guess created using both interfaces. Error bars represent one standard deviation, which is calculated under the assumption of independence between paired data.

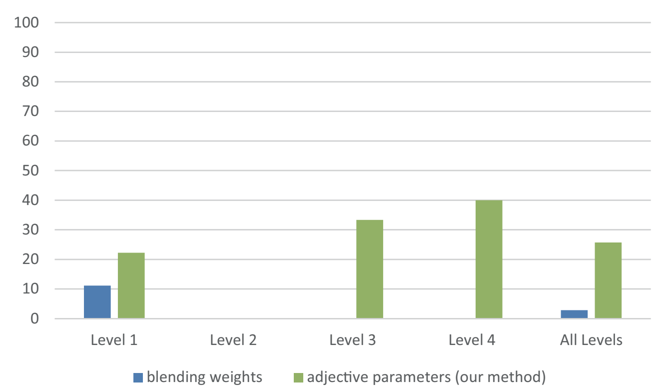


Fig. 7. Ratio of subjects able to recreate the target motion with an initial guess.

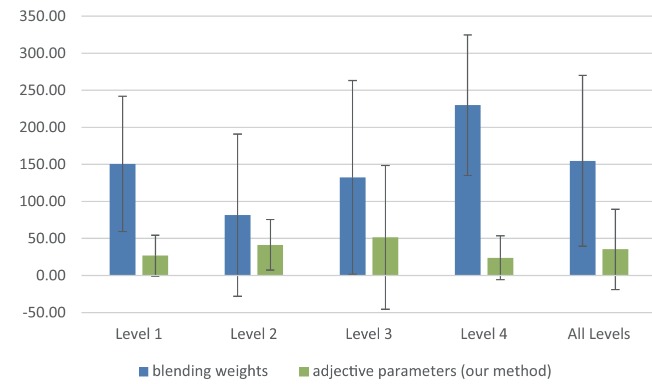


Fig. 8. Average time required to create the target motion using both interfaces. Error bars represent one standard deviation, which is calculated under the assumption of independence between paired data.

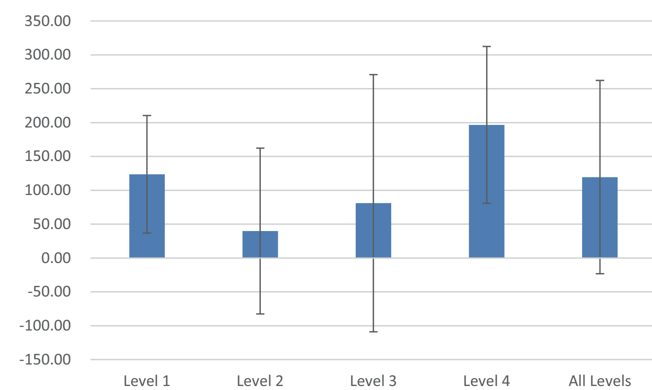


Fig. 9. Average difference between the time required to use our method and that required to adjust the blending weights.

expressed as the mean of the deduction from the required time by using the blending weight and the 90% interval. From this figure, we can deduce that the mean of the required time by using adjective parameters is less than when using blending weights at a 95% confidence level in Levels 1 and 4 and all levels. In addition, the results of the Wilcoxon signed-rank test were $p = 0.011, 0.866, 0.173, 0.007$, and 0.000 for Levels 1, 2, 3, and 4, and all levels, respectively. There were significant differences for Levels 1 and 4 and all levels. The ratio of failures, which means that the subject could not make the target motion in time (within 300 seconds), are shown in Figure 10. Clearly, the subjects could not make the target motion many times when they used the blending weights. The failure data are approximated by 300 seconds in the results in Figures 8 and 9. We remark that they should be more precisely approximated by some statistical methods, but approximation as 300 seconds does not disadvantage the method using the blending weights. Adjective parameters required less time to use than blending weights. This is probably because using adjective parameters was intuitive for the subjects and the number of parameters to be adjusted was small.

There were no significant differences between the results for all four levels. Because the algorithm for generating target motion was not explained to the subjects, they tried to adjust all four primary adjective parameters at every level. Therefore, the target motions at the lower levels were not easier to create than the target motions at any of the higher levels. Most of subjects mainly used only the four primary adjective parameters. There were some subjects who used additional adjective parameters actively.

The subjects commented that using the primary adjective parameters was intuitive and they were able to create the intended motions easily. The subjects also mentioned that using the blending weights might be effective for creating a target motion that is close to one of the example motions, but mostly it was difficult to adjust all blending weights to create a desired motion.

These results indicate that our adjective parameter-based interface is better than the interface that adjusts blending weights. In this experiment, the number of example motions was 10, which is relatively small. In practice, more example motions can be used, which makes adjusting the blending weights much harder. Our adjective parameter-based interface is effective for controlling motion style, which is difficult to

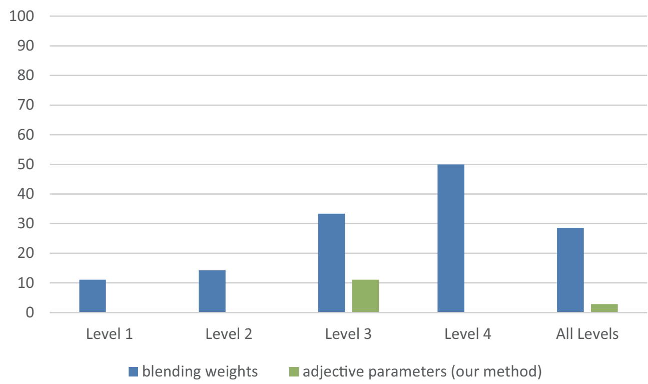


Fig. 10. Ratio of failures to make the target motion.

parameterize. As mentioned above, most subjects used only the primary adjective parameters. A few subjects commented that the primary adjective parameters were too abstract and the additional adjective parameters were more intuitive. In contrast, a few other subjects commented that there are so many additional adjective parameters that it is difficult to choose an appropriate one from them. Our interface allows any combinations of the primary and additional adjective parameters. A user of our interface can choose the approach that they prefer. This is an advantage of our interface.

C. 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 10 example 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 III-E. 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.

VI. 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 thorough 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 quantifies 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 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] D. J. Wiley and J. K. Hahn, “Interpolation synthesis for articulated figure motion,” *IEEE Computer Graphics and Applications*, vol. 17, no. 6, pp. 39–45, 1997.

[2] C. Rose, M. F. Cohen, and B. Bodenheimer, “Verbs and adverbs: Multidimensional motion interpolation,” *IEEE Computer Graphics and Applications*, vol. 18, no. 5, pp. 32–40, 1998.

[3] T. Tsumura, T. Yoshizuka, T. Nojirino, and T. Noma, “T4: A motion-capture-based goal-directed realtime responsive lo-comotion engine,” in *Computer Animation 2001*, 2001, pp. 52–60.

[4] S. I. Park, H. J. Shin, T. hoon Kim, and S. Y. Shin, “On-line motion blending for real-time locomotion generation,” *Computer Animation and Virtual Worlds*, vol. 15, no. 3, pp. 125–138, 2004.

[5] L. Kovar and M. Gleicher, “Automated extraction and parameterization of motions in large data set,” *ACM Transactions on Graphics (SIGGRAPH 2004)*, vol. 23, no. 3, pp. 559–568, 2004.

[6] T. Mukai and S. Kuriyama, “Geostatistical motion interpolation,” *ACM Transactions on Graphics (SIGGRAPH 2005)*, vol. 24, no. 3, pp. 1062–1070, 2005.

[7] M. Lau, Z. Bar-Joseph, and J. Kuffner, “Modeling spatial and temporal variation in motion data,” *ACM Transactions on Graphics (SIGGRAPH Asia 2009)*, vol. 28, no. 5, p. Article No. 171, 2009.

[8] J. Min, Y.-L. Chen, and J. Chai, “Interactive generation of human animation with deformable motion models,” *ACM Transactions on Graphics*, vol. 29, no. 1, p. Article No. 9, 2010.

[9] K. Förger and T. Takala, “Animating with style: Defining expressive semantics of motion,” *The Visual Computer*, vol. 32, no. 2, pp. 191–203, 2016.

[10] T. Komatsu, “Quantifying Japanese onomatopoeias: Toward augmenting creative activities with onomatopoeias,” in *Augmented Human International Conference 2012*, 2012, p. Article No. 15.

[11] J. H. Ward, “Hierarchical grouping to optimize an objective function,” *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236–244, 1963.

[12] L. Deng and D. Yu, “Deep learning: Methods and applications,” *Foundations and Trends in Signal Processing*, vol. 7, no. 3-4, pp. 197–387, 2014.

[13] A. Hutchinson-Guest, *Labanotation: The System of Analyzing and Recording Movement*, 1954.