

Motion Coaching with Emphatic Motions and Adverbial Expressions for Human beings by Robotic System —Method for Controlling Motions and Expressions with Sole Parameter—

Keisuke Okuno* and Tetsunari Inamura

Abstract—Whole-body gestures and verbal expressions should be bound according to given tasks and the current situation in intelligent human-robot interaction systems. Moreover, modification of expressions, such as emphasis of motions and change in verbal expressions, plays an important role for successfully completing tasks according to user reaction. Slight differences in motions can be conveyed by binding an emphasized motions and an verbal expression. In robotics, however, even though the synthesis of gestures and speech has been discussed, how to bind synthesized emphatic motions and verbal expressions from an engineering point of view has not been adequately discussed. Synthesis of motion and speech requires recognition of user reaction, we therefore should integrate 1) recognizing reaction, 2) planning to complete tasks, 3) modification of motions and speech, and 4) maintaining a bi-directional interaction loop consisting of processes 1)–3). Thus, using a phase space, we propose a method for binding emphatic motions and adverbial expressions, and for evaluating and controlling these four required processes by using a sole scalar parameter. In the phase space, variety of motion patterns and verbal expressions can be expressed as static points. To evaluate the feasibility of the proposed method, we also propose a bi-directional motion coaching system using the method. We show the feasibility and effectiveness of robotic motion coaching systems through experiments of actual sport coaching tasks for beginners. From the results of participants' improvements in motion learning, we discuss and conclude the factors affecting such motion coaching systems that realizes binding and controlling emphatic motions and adverbial expressions using a sole scalar parameter in a phase space.

I. INTRODUCTION

In developing effective and intelligent human-robot interaction systems, whole-body gestures and verbal expressions should be strongly connected according to given tasks and the current situation. Additionally, not only fixed expressions but also modification of expressions, such as emphasis of motions and change in speech words, is also an important function for successfully completing tasks according to user reaction. Analysis of the connection between gestures and speech is often discussed in psychology; however, synthesis and emphasis of gestures and speech from an engineering point of view has not been adequately discussed. Since synthesis of motion and speech also requires recognition of the current situation, such as user reaction, we therefore

should integrate 1) recognizing reaction, 2) planning to complete tasks, 3) synthesizing and emphasizing motions and speech, and 4) maintaining a bi-directional interaction loop consisting of processes 1)–3).

In this paper, we focus on a robotic system that coaches humans in performing certain motions in order to address the above issues. The robotic coaching system should include all four above-mentioned processes. 1) recognizing reaction: a robot should evaluate a human motion and analyze similarities and differences between that motion and the coached target motion. 2) planning to complete tasks: the robot should allow the human to perform a better motion based on feedback consisting of properly bound motion and speech expressions. 3) synthesizing and emphasizing both motion and speech: the robot should modify and emphasize motion demonstrations and speech expressions, based on the results of process 1) analysis. 4) The robot repeats the above three processes to maintain a continuous loop of bi-directional interaction for improving motion learning.

Fasola and Mataric developed a robot instructor for the elderly [1]. The robot establishes a loop of interaction by combining motion patterns and verbal expressions. However, it does not synthesize emphatic motions based on feedback of a person's imitations. We believe that this feedback is important because providing feedback not only with verbal expressions but also with emphatic motions would be effective in motion learning interaction for humans.

On the other hand, there have been many studies related to synthesis of motions in computer graphics [2][3][4][5]. However, how to synthesize motions are subjectively determined by designers, and how to bind motions and verbal expressions has not considered.

With regard to research on the binding of motion patterns and verbal expressions in imitation learning frameworks, the method by Iwahashi [6] realizes binding a verbal expression to several similar but different motion patterns. Similarly, stochastic binding method by Takano and Nakamura [7] realizes generation of a motion pattern from a sentence and of a sentence from several similar motion patterns. However, they did not discuss how their methods could be used for binding several different but similar motion patterns to different verbal expressions, which explain slight difference in the motions.

We believe a common problem not being considered in

Keisuke Okuno is with the Department of Informatics, The Graduate University for Advanced Studies (SOKENDAI). Hitotsubashi 2-1-2, Chiyoda-ku, Tokyo, 101-8430, Japan. k-okuno@nii.ac.jp

Tetsunari Inamura is the corresponding author and is with National Institute of Informatics (NII) / SOKENDAI. Hitotsubashi 2-1-2, Chiyoda-ku, Tokyo, 101-8430, Japan. inamura@nii.ac.jp

the above related work is that the four required processes are separated. Since each process is complex, we propose a simple framework to integrate these processes by using a sole scalar parameter to connect them. In motion coaching tasks, considerable factors, such as the similarity of performed motions, degree of emphasis of motion, and variety of verbal expression using adverbial expressions for feedback, can be evaluated using such a scalar parameter.

The objective of this paper is to propose a method for binding motion patterns and verbal expressions, and for evaluating and controlling these above-mentioned factors using a sole scalar parameter in a phase space. Furthermore, we show the feasibility and effectiveness of robotics motion coaching systems based on the proposed method through experiments of actual coaching of a tennis forehand swing for beginners.

In Section II, we explain our method using a scalar parameter. In Section III, we explain the framework of robotic motion coaching systems. In Section IV, we show experimental results. Finally, we discuss the results and conclude factors affecting motion coaching systems.

II. METHOD

In this section, we explain a method how to bind motions and expressions and how to control synthesis of emphatic motions. For this purpose, we applied the mimesis model [8][9]. It is because the mimesis model can be used for binding the motions and expressions, for controlling the motions and the expressions, and for integrating the four above-mentioned processes using a sole parameter in a phase space.

A. Motion Recognition

We applied the mimesis model to recognize motion patterns [8]. It abstracts motion patterns $\mathbf{M} = [\theta[0]\theta[1]\cdots\theta[t]]$, which is a matrix, where $\theta[t] = [\theta_1[t]\cdots\theta_i[t]\cdots\theta_n[t]]^T$ is a vector of time series joint angle, of humans and humanoid robots, where i is a index of a joint. The abstracted representations called *proto-symbols* are calculated using left-to-right continuous Hidden Markov Model (*CHMM*), which consist of a set of parameters $\lambda = \{\mathbf{Q}, \mathbf{A}, \mathbf{B}\}$, where $\mathbf{Q} = \{q_1, \dots, q_N\}$ is a finite set of states, $\mathbf{A} = \{a_{ij}\}$ is a state transition probability matrix from q_i to q_j , and $\mathbf{B} = \{b_i\}$ is a vector of output probabilities of $\mathbf{O}[t]$, at q_i , corresponding to joint angle vector $\theta[t]$ at a discrete time t . Assuming a left-to-right model, the set $\{a_{ij}, b_i\}$ determines the behavior of the stochastic process *proto-symbols*. A geometric space called *proto-symbol space (PSS)* is defined to express the relationship of the *proto-symbols* as distances among static points \mathbf{x} in the space. When a set of motor patterns \mathbf{M} is defined as $\mathbf{M} = \mathbf{O}$ and a database of motor patterns is defined as $D = \{\mathbf{O}_1, \mathbf{O}_2\}$, the *PSS* is represented by $P = F_{build}(D)$ ($P = \{\mathbf{x}_1, \mathbf{x}_2\}$, where \mathbf{x} is a coordinate of a static point in the *PSS*), where F_{build} is a construction

process of the *PSS* from D . The distances between *proto-symbols* in the space are calculated using the Bhattacharyya Distance[10] between *CHMMs* that correspond to the similarities of motor patterns.

Similarly, motion patterns are recognized by first abstracting motion patterns to the *CHMMs* using the Baum-Welch algorithm then calculating the Bhattacharyya Distance [10] between *CHMMs*. As a result, a motion pattern can be recognized as a static point \mathbf{x} in the *PSS*,

$$\mathbf{x} = F_{recog}(\mathbf{O}). \quad (1)$$

(for more detail, please refer to [8])

B. Synthesis of Emphatic Motions

For synthesizing output probability $b_i(\mathbf{O})$, we use a single Gaussian model for the output such that an intuitive synthesis of joint angle vectors can be achieved by using the mean and variance vectors of a Gaussian distribution. With the mean and variance vectors and by using the Monte Carlo method, the motion pattern is generated [8]. A function for generating sensorimotor patterns form \mathbf{x} , F_{gen} is defined as

$$\mathbf{O} = F_{gen}(\mathbf{x}) \quad (2)$$

Instead of directly interpolating/extrapolating both a_{ij} and $b_i(\mathbf{O})$ for synthesizing motion patterns, (1) the state transition probabilities and the output probabilities are separately operated upon, and (2) the state transition matrices are calculated in a different domain, i.e., the time domain (for more detail, please refer to [9]).

A function for synthesizing an internal/external dividing point \mathbf{x}_s , which corresponds to an interpolated/extrapolated novel motion pattern, from static points $\mathbf{x}_i, \mathbf{x}_j$ is represented as

$$\mathbf{x}_s = c_i\mathbf{x}_i + c_j\mathbf{x}_j, \quad (3)$$

where c_i and c_j are coefficients for internal/external dividing point. To generate synthesized emphatic motion \mathbf{O}_s , Eq. (2) was used as $\mathbf{O}_s = F_{gen}(\mathbf{x}_s)$.

C. Proposed Method, Controlling Emphatic Motions with Sole Scalar Parameter

With one parameter, α , it is possible to control both the degree of emphasis of synthesized motion and choice of adverbial expression. We defined α by applying Eq. (3),

$$\mathbf{x}_s = \mathbf{x}_p + \alpha(\mathbf{x}_t - \mathbf{x}_p), \quad (4)$$

where α is a weight coefficient for extrapolation, \mathbf{x}_t and \mathbf{x}_p are static points in the *PSS* corresponding to motions \mathbf{O}_t and \mathbf{O}_p , respectively. It is defined that $1.0 \leq \alpha \leq 2.5$, and when $\alpha = 1.0$ the synthesized motion \mathbf{x}_s is identical to the imitation target motion \mathbf{x}_t . When α is negative, missing elements are emphasized instead of complimenting the missing elements as in this paper. As depicted in Fig. 1, α corresponds to the degree of emphasis of the synthesized motion.

Example Synthesized Emphatic Motion X_s

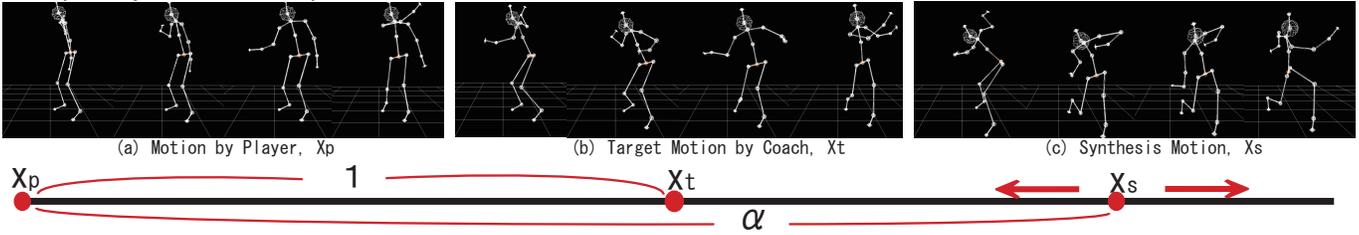


Fig. 1. Example of relationship between weight coefficient α and synthesized emphatic motion. (a): a motion imitated by a beginner player x_p , (b): the imitation target motion demonstrated by the coach x_t , (c): a synthesized emphatic motion x_s with $\alpha = 2.0$

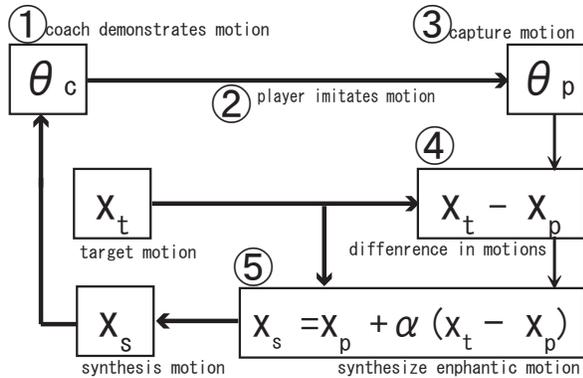


Fig. 2. Flow of motion coaching

III. FRAMEWORK OF MOTION COACHING SYSTEM

A. Coaching Flow

First, database $D = \{\theta_t, \theta_p\}$, which consists of one imitation target motion pattern θ_t and one motion pattern performed by a player θ_p , is prepared. Then, Proto-symbol Space (PSS), denoted as P , is built using D with the F_{build} process. Let us define the static point x_t in the P as that corresponding to the imitation target motion θ_t .

As depicted in Fig. 2, the motion coaching task is executed as follows.

- 1) The coach (an agent in a virtual environment in this paper) demonstrates motion pattern θ_c as the imitation target motion. Using the generation function F_{gen} (Eq. (2)), θ_c is generated from x_t in the initial trial, and from x_s synthesized in step 5), in later trials.
- 2) The human player imitates θ_c .
- 3) The coach observes the player's imitated motion pattern θ_p and converts it to a static point x_p in the P , using recognize function F_{recog} (Eq. (1)).
- 4) If x_p is not close to x_t , then it is interpreted that the player's imitated motion is imperfect. The coach calculates the missing components in the imperfect imitation of the player by $x_t - x_p$.
- 5) The coach calculates the external dividing point x_s by adding the missing components ($x_t - x_p$) to x_p , using Eq. (4). The x_s is used as a re-demonstration motion pattern θ_c for the next trial.

Repeat loop of processes 1)-5) as needed. 1 loop is considered as 1 trial in the experiment.

B. Example Results

Example outputs of the coaching system are depicted in Fig. 1-(c). Figure 1-(b) shows the imitation target motion x_t demonstrated by the coach. The objective of the player is to imitate this motion as close to the target motion as possible. Figure 1-(a) shows the observed motion pattern imitated by the player x_p . As can be seen, the imitation is not perfect. For example, the trajectory of the left arm is not the same as in x_t . Figure 1-(c) shows an emphatic motion synthesized using the proposing method. Not only the left arm has been considered as a missing elements and emphasized but also the degree at which the right knee bent and that at which the upper body bent forward were taken into consideration for emphatic motion and synthesized, using the system according to the proposed method. As a result, the synthesized motion seems to be fairly reasonable. In addition, this shows the feasibility of the mimesis model in qualitatively extrapolating/interpolating motions for real-world application, which has never been discussed before. In the next section, quantitative evaluation of the coaching system and controlling emphatic motion and adverbial expressions using a sole scalar parameter is discussed with experiments of an actual sports coaching task.

IV. EXPERIMENTS

We conducted a series of experiments in which a forehand tennis swing was coached to male beginning tennis player by a robotic system. The proposed method and the flow of the coaching system explained in the previous section were used with the following procedure.

A. Common Conditions

- The coaching agent demonstrates motion pattern x_t , shown in Fig. 1-(b), as the imitation target motion and is displayed on a wall.
- The given instruction is "please imitate this".
- The view point of the imitation target motion is fixed and always from the front.
- Motions θ consists of 17 joints, each with DoF of 3.

- The number of elements in database D , which were used for building the P is 2. In other words, only the imitation target motion θ_t and subject's imitated motion θ_p were used; $D = \{\theta_t, \theta_p\}$.
- Five swings are used to abstract each player's swing to an HMM for each trial.
- Each trial is executed at five minutes intervals.

B. Experiment 1

In experiment 1, we tested 11 participants for learning how different α would affect motion learning. The emphatic motions with different $\alpha = \{1.25, 1.50, 1.75, 2.00, 2.25, 2.50\}$ were synthesized and shown to the participants. We considered two cases, Case (a): without adverbial expression and Case (b): with adverbial expression "more".

C. Results for Experiment 1

Distance was introduced to evaluate the results.

$$d_{il} = \|x_t - x_p^{il}\|, \quad (5)$$

where i is an ID number of a participant, l is the trial number and it is $l = \{1, 2, 3, 4\}$ for this paper. Therefore, x_p^{il} corresponds to the performance of participant i in trial l . This means that the smaller the d_{il} , the better the imitation. When $d_{il} = 0$, the imitation is perfectly identical. Thus, d_{il} corresponds to an imitation error of participant ID i in trial l . To reduce possible errors in the evaluation results that could be caused by the difference in frame number of the evaluated motions, we used motion clips that had a similar frame number, at most a 20% difference. The average distance of d_{il} in trial l was also introduced as

$$\bar{d}_l = \frac{\sum_{i=1}^m d_{il}}{m}, \quad (6)$$

where m is the number of participants ($m = 11$ for experiment 1).

As shown in Fig. 3, it is reasonable to assume that $\alpha = 2.0$ would provide the best result for motion learning in both cases. Thus, in experiment 2, $\alpha = 2.0$ was used.

D. Experiment 2

In experiment 2, four different cases were tested to evaluate the proposed method.

- The number of participants was 13.
- The order of conducting each experimental cases was randomly shuffled for each participants.
- The adverbial expression, "more like this" was introduced.
- Four experimental cases were introduced to evaluate how the emphatic motion and adverbial expression contributed to motion learning. The cases are also summarized in Table I.

- Case 1:

The coaching agent repeated only demonstrating the target motion θ_t , i.e., it coached using motions with $\alpha = 1.0$ and with no adverbial expression.

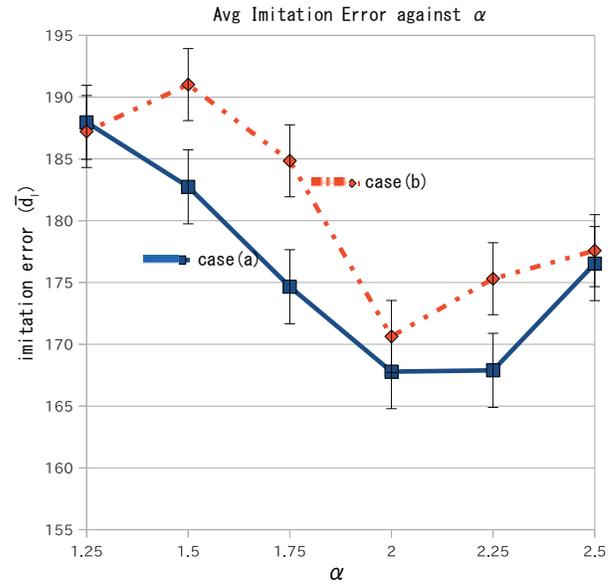


Fig. 3. Results of experiment 1: Imitation error \bar{d}_l in Eq. (6) against different values of α s (\bar{d}_l adjusted so that maximum value of each participant would be 200)

- Case 2:
The agent coached using motions with $\alpha = 1.0$ and adverbial expression "more".
- Case 3:
The agent re-demonstrated with emphatic motion patterns x_s synthesized using the proposed method, i.e., the agent coached using motions with $\alpha = 2.0$ and with no adverbial expression.
- Case 4:
The agent coached using emphatic motions with $\alpha = 2.0$ and adverbial expression "more".

TABLE I
CASES IN EXPERIMENT 2

| - | $\alpha = 1.0$ | $\alpha = 2.0$ |
|----------------|----------------|----------------|
| without "more" | Case 1 | Case 3 |
| with "more" | Case 2 | Case 4 |

E. Results of Experiment 2

To evaluate the results of experiment 2, the following measures were introduced. The average error ratio in imitation at trial l was:

$$\bar{R}_l = \frac{\sum_{i=1}^m \frac{d_{il}}{d_{i1}}}{m} \quad (7)$$

If the imitation error is smaller than the initial trial in the same case, \bar{R}_l will be less than 1.0. If the imitation error is larger than the initial trial in the same case, \bar{R}_l will be larger than 1.0. When imitation is perfect, $\bar{R}_l = 0.0$. The d_{i1} would not be zero in practice because $d_{i1} = 0$ means perfect imitation of the target motion and it would not happen.

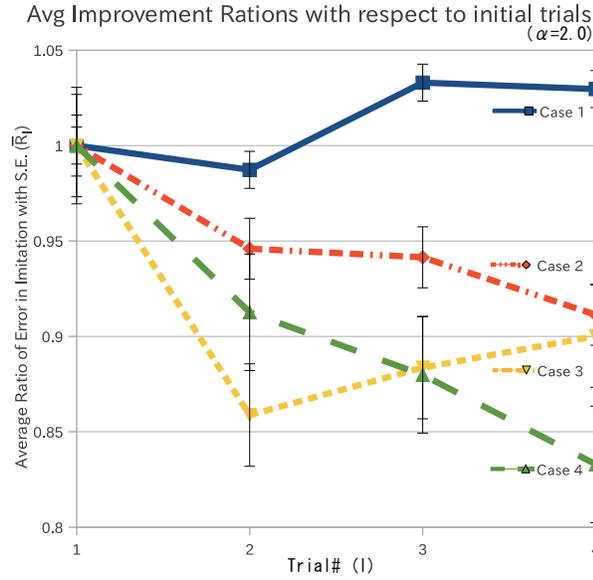


Fig. 4. Results of experiment 2: Average ratio of imitation error: \bar{R}_l in Eq. (7)

TABLE II

ONE-TAILED T-TEST WITH \bar{R}_l AT 4TH TRIAL, $l = 4$, IN EQ. (7) FOR EXPERIMENT 2. [NOTE: **($p < 0.01$), * ($p < 0.05$), + ($p < 0.10$)]

| Case 1 vs. 2 | Case 1 vs. 3 | Case 2 vs. 4 | Case 3 vs. 4 |
|--------------|--------------|--------------|--------------|
| 0.00091** | 0.0000079** | 0.061+ | 0.48 |

From Fig. 4 and Table II, it would be safe to say that emphatic motions contributed somewhat to improving the motion learning of the participants. In particular, comparison between Cases 1 and 2, and between 1 and 3 eliminates possibility of improvement by the repeated watching of several demonstrations. However, we were not able to confirm if the adverbial expression contributed to improving motion learning or not. This is because significant difference $0.05 < p < 0.10$ was found between Cases 2 and 4, but no significant difference $p > 0.10$ was found between Cases 3 and 4.

F. Experiment 3

To investigate the contribution of emphatic motions and adverbial expressions to improving motion learning, we conducted additional experiments as Cases 5 and 6 with 11 out of the 13 participants in experiment 2.

- Case 5:
The agent coached using motions with optimized α and no adverbial expression.
- Case 6:
The agent coached using motions with optimized α and adverbial expression "more".

In Case 5 and 6, optimized α was used for each participant, as summarized in Table III, because the best α may be different from the average for some participants, as shown in Fig. 5. For example, participant-A had optimized- $\alpha =$

TABLE III
CASES IN EXPERIMENT 3

| - | $\alpha = 1.0$ | optimized α |
|----------------|----------------|--------------------|
| without "more" | Case 1 | Case 5 |
| with "more" | Case 2 | Case 6 |

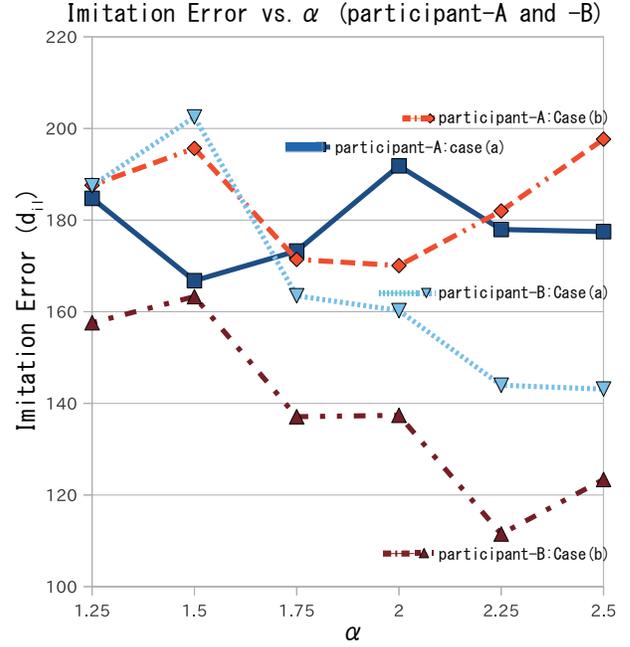


Fig. 5. example of individual imitation error against different values of α (participants-A and -B): d_{ii} in Eq. (5)

1.5 for Case (a) and optimized- $\alpha = 2.0$ for Case (b). Similarly, participant-B had optimized- $\alpha = 2.5$ for Case (a) and optimized- $\alpha = 2.25$ for Case (b). The same method was applied to determine optimized α for the rest of the participants.

G. Results of Experiment 3

From the results in Fig. 6 and Table IV, we conclude that emphatic motions contribute to motion learning because we found a significant difference with $p < 0.05$ between Case 2 and 6. However, we could not conclude whether adverbial expressions contribute to motion learning or not since there was no significant difference found, $p > 0.10$, between Case 5 and 6.

TABLE IV

ONE-TAILED T-TEST WITH \bar{R}_l (EQ. 7) IN 4TH TRIAL, $l = 4$, FOR EXPERIMENT 3. [NOTE **($p < 0.01$), * ($p < 0.05$), + ($p < 0.10$)]

| Case 1 vs. 5 | Case 2 vs. 6 | Case 5 vs. 6 |
|--------------|--------------|--------------|
| 0.0000062** | 0.033* | 0.27 |

V. CONCLUSION AND DISCUSSION

We proposed a method to realize binding motions and expressions and controlling emphatic motions and adverbial

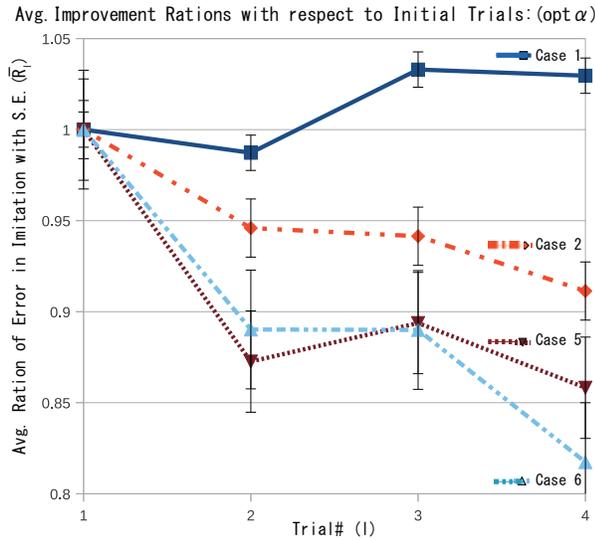


Fig. 6. Results of experiment 3: average ratio of imitation error, \bar{R}_t in Eq. (7), with optimized α for each participant

expressions using a sole scalar parameter in a phase space by applying the mimesis model. Using the proposed method, a robotic motion coaching systems was also proposed – (1) to evaluate the similarity between performed motions, (2) to control the binding method of emphatic motions and adverbial expressions, (3) to control the degree of emphasis of synthesized motions, and (4) to maintaining a continuous loop of bi-directional interaction.

We showed the feasibility and effectiveness of robotic motion coaching systems through experiments of a tennis forehand swing coaching task for beginners. From the results of players’ improvement in motion learning, we confirmed that the proposed method is beneficial for robotic systems for having effective interaction with humans by using whole-body gestures and verbal expressions, and therefore, we confirmed the feasibility of the proposed method for binding and controlling emphatic motions and adverbial expressions using a sole scalar parameter in a phase space. Specifically, we confirmed that emphatic motion, not repetition of observing similar motions, is an effective factor. However, even though we were able to confirm that adverbial expressions seemed to have a positive relation to the improvement of motion learning, we were not able to confirm whether it is an effective factor.

To investigate the effect of the adverbial expression to improving motion learning, we plan to enrich the adverbial expressions. The adverbial expression used for this paper is to show the feasibility of the proposed framework as a first step, and therefore it was fixed to a simple expression, “more like this”. For future work, we plan to conduct experiments introducing several adverbial expressions chosen according to the value of α . For example, use of the adverbial expressions “a bit more” when $1.0 \leq \alpha < 1.5$, “more” when $1.5 \leq \alpha < 2.0$, and “much more” when $2.0 \leq \alpha < 2.5$.

Other methods can synthesize emphatic motions, but there is a strong benefit with the proposed method by applying the mimesis model. With the mimesis model, for example, it would be possible to instruct by taking into account the tendencies of players, such as “Please, swing not like jumping, but more like squatting”, instead of the current simple verbal expression “more like this”. For this research, we focused on the feasibility of the proposed method; therefore, a simple verbal expression was used. The technical basis to achieve such motion synthesis by using natural expressions is that the mimesis model can convert high dimensional complex real-world property to static points in low dimensional space, *PSS*, and can convert back from the static points to the high dimensional property, such as motion patterns. If motions, such as “jump” and “squat”, can be labeled, then the natural expression could be converted into external/internal division of proto-symbols in the *PSS*; then it corresponds to proper extrapolation/interpolation of motion patterns in the high-dimensional space. This means that the *PSS* can be used as the bridge between natural expression and motion synthesis. This is the most important reason the mimesis model was used.

We only used joint angles as \mathbf{O} : however, \mathbf{O} can be extended to deal with sensory patterns $\mathbf{S} = [s[0]s[1] \cdots s[t]]$, where s are time series data. That is, $\mathbf{O} = [\mathbf{M}^T, \mathbf{S}^T]^T$. This way, we believe that systematic binding of sensorimotor patterns and more complex verbal expressions that refer to sensory patterns can be achieved.

REFERENCES

- [1] J. Fasola and M. Mataric, “Robot exercise instructor: A socially assistive robot system to monitor and encourage physical exercise for the elderly,” in *Proc. of 19th IEEE International Symposium in Robot and Human Interactive Communication*, 2010.
- [2] A. Bruderlin and L. Williams, “Motion signal processing,” in *SIGGRAPH*, 1995, pp. 97–104.
- [3] C. Rose, M. Cohen, and B. Bodenheimer, “Verbs and adverbs: multidimensional motion interpolation,” *IEEE Computer Graphics and Applications*, vol. 18, no. 5, pp. 32–40, 1998.
- [4] P. Glargon, R. Boulic, and D. Thalmann, “Pc-based walking engine using motion capture data,” in *Proc. of the Computer Graphics International*, 2004, pp. 292–298.
- [5] K. Hoshino, “Interpolation and extrapolation of repeated motions obtained with magnetic motion capture,” *IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences*, vol. E87-A, pp. 2401–2407, 2004.
- [6] N. Iwahashi, *Robots That Learn Language: A Developmental Approach to Situated Human-Robot Conversations*. I-Tech Education and Publishing, 2007, ch. Nilanjan Sankar (ed.), Human-Robot Interaction, pp. 95–118.
- [7] W. Takano and Y. Nakamura, “Integrating whole body motion primitives and natural language for humanoid robots,” in *Proc. of International Conference on Humanoid Robots*, 2008, pp. 708–713.
- [8] T. Inamura, H. Tanie, and Y. Nakamura, “From stochastic motion generation and recognition to geometric symbol development and manipulation,” in *Proc. of International Conference on Humanoid Robots*, 2003.
- [9] T. Inamura and T. Shibata, “Geometric proto-symbol manipulation towards language-based motion pattern synthesis and recognition,” in *Proc. of Int’l Conf. on Intelligent Robots and Systems*, 2008, pp. 334–339.
- [10] A. Bhattacharyya, “On a measure of divergence between two statistical populations defined by their probability distributions,” *Bulletin of the Calcutta Mathematical Society*, vol. 35, pp. 99–109, 1943.