# Composing an Assistive Control Strategy Based on Linear Bellman Combination From Estimated User's Motor Goal

Jun-ichiro Furukawa ⬤ , *Member, IEEE*, and Jun Morimoto, *Member, IEEE*

*Abstract*—In assistive control strategies, we must estimate the user's movement intentions. In previous studies, such intended motions were inferred by linearly converting muscle activities to the joint torques of an assistive robot or classifying muscle activities to identify the most likely movement from pre-designed robot motion classes. However, the assistive performances of these approaches are limited in terms of accuracy and flexibility. In this study, we propose an optimal assistive control strategy that uses estimated user movement intentions as the terminal cost function not only for generating movements for different task goals but to precisely enhance the motion with an exoskeleton robot. The optimal assistive policy is derived by blending the pre-computed optimal control laws based on the linear Bellman combination method. Coefficients that determine how to blend the control laws are derived based on low-dimensional feature values that represent the user's movement intention. To validate our proposed method, we conducted an assisted basketball-throwing task and showed that the performances of our subjects significantly improved.

*Index Terms*—Human performance augmentation, physically assistive devices, optimization and optimal control.

## I. INTRODUCTION

**D**UE to the recent progress in robotics technologies, such wearable robots as exoskeleton robots are expected to physically contact and assist humans in a myriad of activities. As a proof-of-principle, a hand exoskeleton [1], [2] and upper- and lower-body exoskeleton robots [3]–[5] have been studied. For these applications, surface electromyography (EMG) is a candidate approach to intuitively control the robots by estimating the user movement intentions [6], [7].

Using a linearly scaled EMG to derive the joint torques of an assistive device is a standard approach for exoskeleton robot control. However, to accomplish a given task based on this simple EMG-based control strategy, human users must generate highly tuned EMG signal patterns by themselves. Therefore, this standard approach imposes the computational burden on user's nervous system [8]. Consequently, the motion precision depends on the user's ability. Another common approach to estimate human movement intention is applying a classification method that estimates associated movements for the observed EMG signals. In this case, EMG is only used for initiating pre-designed movements. Therefore, users do not need to carefully control exoskeleton robots by themselves. However, we can only use this approach for a limited number of discrete task goals since all the assistive controllers, which are associated with those goals, need to be pre-computed.

In this study, we develop an assistive control method that exploits the above two standard approaches. In our proposed method, we estimate movement intentions by observing the initiation of the motions with muscle activities for a very short period of time. We assume that a discrete number of optimal control laws, which are associated to different task goals, are pre-computed. Then we derive the optimal assistive policy for the estimated movement intention by blending the pre-computed optimal control laws based on the linear Bellman combination method (Fig. 1). To validate our proposed method, we conducted an assisted basketball-throwing task with using an upper-limb exoskeleton robot. In these experiments, subjects shot a basketball at the center of a target on a board at different distances. The distances from the throwing point to the current task goal were estimated as the user's movement intention and the optimal control policies were derived using our proposed method. Compared with a baseline method that used a classification method, our results showed that the shooting accuracy was significantly improved.

The following are the contributions of our study:
1) Task goals were estimated from EMG signals monitored within very limited amount of time period through finding low-dimensional task-related feature space.
2) Optimal assistive policy to cope with continuous change of the task goal was derived by blending pre-computed optimal control laws based on linear Bellman combination method.
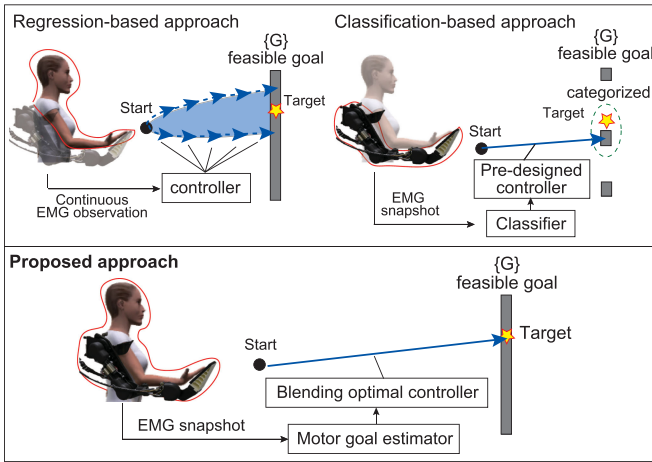
Fig. 1. Schematic illustration of conventional approaches and our approach: Motion precision depends on user's ability in regression approach, and classification approach can only be used for limited number of discrete task goals. In our approach, optimal assistive policy to cope with continuous change of task goal is derived by blending pre-computed optimal control laws based on the linear Bellman combination method. How to blend control laws is determined from estimated user's movement intentions.

3) The derived policy based on the linear Bellman combination were successfully applied to an upper-limb exoskeleton robot to accomplish an assisted basketball-throwing task.

The rest of this letter is organized as follows. In Section II, we introduce related EMG-based control strategies for the exoskeleton robot. In Section III, we explain our user's goal estimation-based optimal control approach. Section IV describes our experimental setups. Section V describes our experimental results. Finally, the conclusion is provided in Section VI.

## II. RELATED WORKS

To intuitively control such assistive devices as exoskeleton robots, regression approaches are often used. In these approaches, robot motions are generated by using identified relationships between EMG signals and joint torques [9]–[11]. In this method, the robot is continuously controlled by a user through measured muscle activities. However, since fixed parameters are used in these regression methods, human users need to carefully generate their own EMG signal patterns to cope with changing task goals. The control performances of assisted movements greatly depend on the skills of the users. Another frequently adopted approach for estimating user's movement intentions is a classification method. For example, support vector machine (SVM) [12]–[15] or linear discriminant analysis (LDA) [16]–[19] have been utilized to classify EMG signal patterns detected from human users. The classification results were used to select pre-designed control output associated with the estimated motion labels. Since this approach can only generate a limited number of pre-designed control output, these classification approaches are unsuitable for generating movements for novel task goals.

In this study, we take the advantages of these two approaches. We first pre-compute optimal assistive control laws for multiple task goals, then combine these control laws using the linear Bellman combination method [20] to generate optimal control output for a novel task goal that is estimated from an observed user's state. Therefore, human users do not need to carefully generate their own EMG signals since our method derives an optimal policy for the estimated task goal. In addition, the number of motions is not limited to the number of pre-computed control laws since the proposed method can cope with novel task goals by combining these pre-computed controllers. Furthermore, the proposed approach is computationally light and can easily derive the controller in real-time because it does not optimize a controller but combine pre-optimized controllers.

## III. METHODS

This section introduces our approach for determining the control policy to assist the user's intended movements. In Section III-A, we introduce how user's motor goal is estimated from measured sensor signals from a subject and an assistive robot. Then, in Section III-B, the linear Bellman combination method is introduced to combine pre-optimized controllers based on the estimated user's motor goal.

### A. User's Motor Goal Estimation

To cope with the noisy inputs of the user's EMG signals, we first find low-dimensional feature space for robust extraction of user's motor goal. In this study, we used the partial least squares (PLS) algorithm [21], [22] to find the task related low-dimensional feature space. Since, by using PLS, the subspace is extracted in consideration with the covariance with the motion labels, the low-dimensional features reflecting the label information can be obtained.

We first monitor the user state $\boldsymbol{\psi} \in \mathcal{R}^m$ for movement initiation period, then project the averaged user state $\bar{\boldsymbol{\psi}}$ over the movement period onto the feature space:

$$\boldsymbol{\mu} = \mathbf{W}^\top \bar{\boldsymbol{\psi}}, \tag{1}$$

where $\boldsymbol{\mu} \in \mathcal{R}^l$ $(l < m)$ is the extracted feature vector and the projection matrix $\mathbf{W} \in \mathcal{R}^{m \times l}$ is derived before the actual experiments with using pre-collected training data. We then estimate the weight parameter $\omega_i$ of the task goal for the $i$-th pre-computed optimal policy by a regression method.

### B. Optimal Control Based on Motion Intention

By combining the weighted task goals, we can derive a novel optimal controller based on the pre-computed optimal policies with the linear Bellman combination method. In this method, the novel controller is represented as the combination of the pre-computed optimal policies for pre-designed tasks:

$$\pi^* = \alpha_1(\mathbf{x}, k)\pi_1(\mathbf{x}, k) + \cdots + \alpha_n(\mathbf{x}, k)\pi_n(\mathbf{x}, k), \tag{2}$$

where, $\mathbf{x}$ is the robot state, $\alpha_i(\mathbf{x}, k)$ is the mixing coefficient, and $\pi_i$ indicates the optimal control policy for the $i$-th motor goal
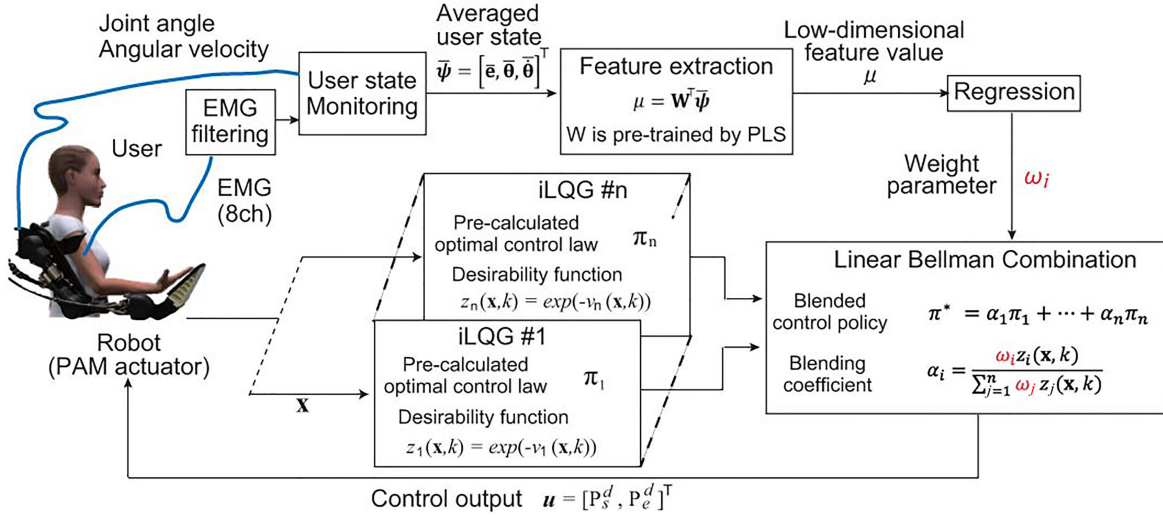
Fig. 2. Proposed method: Assistive control policy is derived by blending optimal control laws, where each controller was pre-computed using iterative LQG and computed policies were combined based on linear Bellman combination method. Mixing coefficients that determine how to blend the control laws were derived based on low-dimensional feature values extracted from measured user's EMG signals, joint angles and angular velocities.

task obtained by minimizing the value function $v_i(\cdot)$ as follows:

$$v_i^\pi(\mathbf{x}, t) = g_i(\mathbf{x}(T)) + \sum_{k}^{T-1} l(\mathbf{x}, \pi, k) \quad (3)$$

$$\pi_i \leftarrow \operatorname*{argmin}_\pi v_i^\pi \quad (4)$$

where, $g(\cdot)$ and $l(\cdot)$ indicate the terminal cost and instantaneous cost, respectively.

Assuming the multiple optimal control problems that share all setting except the terminal cost $g_i(\mathbf{x}(T))$ of (3), the linear Bellman combination approach [20], [23] can be used, and the mixing coefficient can be written as follows:

$$\alpha_i(\mathbf{x}, k) = \frac{\omega_i z_i(\mathbf{x}, k)}{\sum_{j=1}^n \omega_j z_j(\mathbf{x}, k)}. \quad (5)$$

where, $z_i(\mathbf{x}, k) = exp(-v_i(\mathbf{x}, k))$ indicates the feasibility of the policy for the current state. The value function for the $i$-th optimal control problem $v_i(\mathbf{x}, k) = v^{\pi_i}(\mathbf{x}, k)$ is derived by using the iterative linear-quadratic-Gaussian (iLQG) [24].

By using iLQG method, a quadratic approximation of value function for $i$-th movement trajectory can be computed:

$$v_i(\mathbf{x}, k) = v_i(k) + \Delta\mathbf{x}^\top(k)\mathbf{s}_i(k)$$
$$+ \frac{1}{2}\Delta\mathbf{x}^\top(k)\mathbf{S}_i(k)\Delta\mathbf{x}(k), \quad (6)$$

where $\mathbf{s}_i(k)$ and $\mathbf{S}_i(k)$ are the parameter of the quadratic value function model. These parameters can be acquired in the calculation process of iLQG. The actual control law can also be acquired as follows:

$$\pi_i(\mathbf{x}, k) = \mathbf{u}_i^*(k) + L_i(k)\Delta\mathbf{x}_i(k), \quad (7)$$

where $\Delta\mathbf{x}_i(k) = \mathbf{x}(k) - \mathbf{x}_i^*(k)$. $\mathbf{x}_i^*(k)$ and $\mathbf{u}_i^*(k)$ are the optimized open loop state and the control output of the derived $i$-th

optimal controller, respectively. Time-dependent local feedback gain $L_i(k)$ is also acquired.

Figure 2 and Algorithm below summarize our proposed method.

---

**Algorithm**

Require: User state $\psi$, Projection matrix $\mathbf{W}$,
Pre-computed value functions $v_i(\cdot)$,
Pre-computed policies $\pi_i(\cdot)$, $(i = 1, \ldots, n)$,
Task horizon length $T$
$\boldsymbol{\mu} = \mathbf{W}^\top\overline{\psi}$
$\omega_i \leftarrow$ TaskGoalEstimation $(\boldsymbol{\mu})$
**for** $k = 1 : T$ **do**
  **for** $i = 1 : n$ **do**
    $z_i(\mathbf{x}, k) = exp(-v_i(\mathbf{x}, k))$
    $\alpha_i(\mathbf{x}, k) = \frac{\omega_i z_i(\mathbf{x}, k)}{\sum_{j=1}^n \omega_j z_j(\mathbf{x}, k)}$
  **endfor**
  $\pi^*(\mathbf{x}, k) = \alpha_1(\mathbf{x}, k)\pi_1(\mathbf{x}, k) + \cdots + \alpha_n(x, k)\pi_n(\mathbf{x}, k)$
**endfor**

---

## IV. EXPERIMENTAL SETUPS

In this study, as depicted in Fig. 3, we conducted the basketball throw task to evaluate our assistive control strategy. We use our four-joint upper limb exoskeleton robot (Fig. 4) [25] for this assisted basketball-throwing task. We compared online assist control performances among a baseline classification-based method, no assist and the proposed method, where the details of the baseline method is introduced in Appendix.

### A. Upper Limb Exoskeleton Robot Control

In this study, the optimal policies were derived to control shoulder extension/flexion and elbow extension/flexion joints.
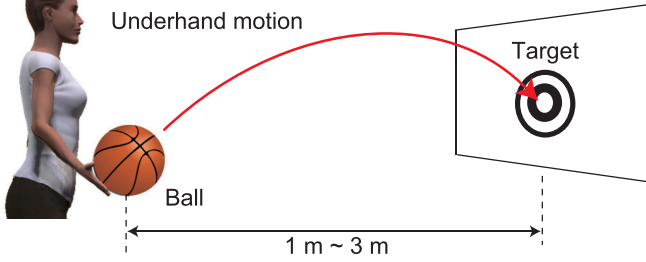
Fig. 3. Basketball throwing task. Subjects throw basketball with underhand posture towards center of target on board. Distance to target was changed as 1 m, 1.5 m, 2 m, 2.5 m and 3 m.
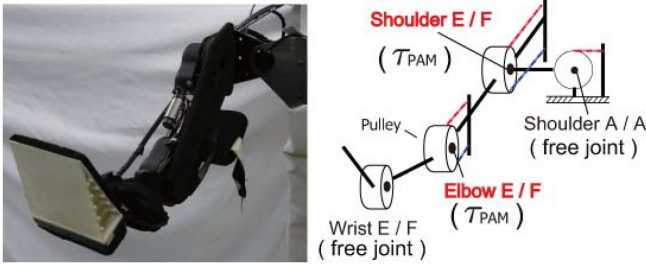


Fig. 4. Upper-limb exoskeleton robot. Shoulder extension/flexion and elbow extension/flexion are actuated by pneumatic artificial muscles (PAMs). In our experiment, shoulder abduction/adduction and wrist extension/flexion joints were not actuated and behaved as free joints.



Fig. 5. EMG channel location. We used eight EMG channels to estimate user's motor goal.

Each joint is actuated by a pneumatic artificial muscle (PAM). The joint torques are generated by a PAM as follows:

$$\tau_{pam} = r f_{pam} \tag{8}$$

where, $r$ is the pulley radius and $f_{pam}$ is the PAM force generated by the path contraction of the spiral fibers embedded in a pneumatic bladder. The details of the mechanical design and the pressure-force model were introduced in our previous studies [10], [25], [26].

In this study, we used two-link robot model as the physical constraint for iLQG calculations. The link parameters were determined with considering the CAD data of the robot and estimated body parameters of each subject [27]. The state of the robot was defined by the joint angle $\theta$, angular velocity $\dot{\theta}$, and measured inner pressure of PAM $P$ as $\mathbf{x} = [\theta_s, \theta_e, \dot{\theta}_s, \dot{\theta}_e, P_s, P_e]^\top$, where the subscripts $s$ and $e$ represent the shoulder and elbow joints, respectively. The control output $\mathbf{u}$ was the desired pressure inputs for pneumatic actuators: $\mathbf{u} = [P_s^d, P_e^d]^\top$. We modeled the dynamics of the inner pressure state of PAM as $\dot{P} = \frac{1}{t_c}(P^d - P)$, where $t_c$ is the time constant of the PAM [26].

### B. Motion Task

In the assisted basketball-throwing task, the subjects threw a basketball with underhand posture towards the center of a target on a board while sitting on a chair. Distance to the target was changed as 1 m, 1.5 m, 2 m, 2.5 m and 3 m. The subjects threw the ball five times at each distance in test experiments. The trial order
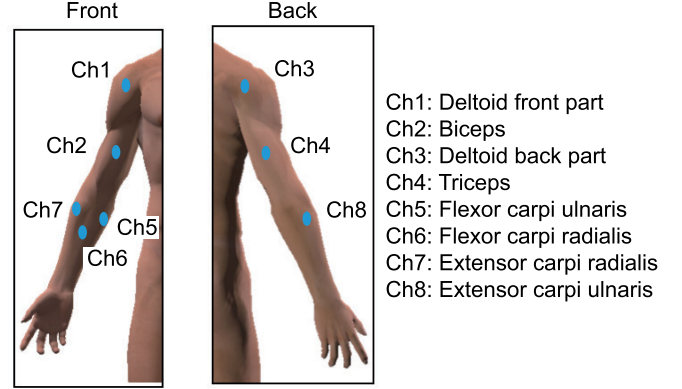
of the proposed, baseline, and no assist conditions was randomly changed for each subject. The subjects were given enough time to rest between the trials. To measure release timing of the ball, a tactile sensor was attached to the tip of the subject's middle finger.

The terminal cost $g_i(\cdot)$ for minimization was set as follows:

$$E_i^a(T) = \left\| \boldsymbol{\theta}(T) - \boldsymbol{\theta}_i^{target}(T) \right\|^2,$$
$$E_i^v(T) = \left\| \dot{\boldsymbol{\theta}}(T) - \dot{\boldsymbol{\theta}}_i^{target}(T) \right\|^2, \tag{9}$$
$$g_i(\mathbf{x}(T)) = C_a E_i^a + C_v E_i^v,$$

where, $T$ is the terminal time, $\theta_i^{target}$ is the target angles, and $\dot{\theta}_i^{target}$ is the target angular velocities at the point of release of the ball for $i$-th task goal. These desired states were calculated according to the $i$-th target distance, taking into account the dynamics of the ball movement. In this study, we pre-computed the optimal policies for the distances of 1 m and 3 m by defining the terminal cost $g_1(\cdot)$ and $g_2(\cdot)$, and tested the final control policies for five basketball shooting distances by blending them.

The instantaneous cost $l(\cdot)$ was set as

$$l(\mathbf{x}, \pi, k) = C_p \|\mathbf{u}\|^2 + C_{pd} \|\dot{\mathbf{u}}\|^2. \tag{10}$$

where, $C_a, C_v, C_p, C_{pd}$ were manually selected as in most optimal control studies.

We conducted the experiments with ten healthy right handed subjects (seven females with 21–38 and three males with age 23–39), after obtaining informed consent from them. The human research ethics committee of Advanced Telecommunications Research Institute International approved the experiment.

### C. Feature Extraction

In order to estimate the user's motor goal from an initial motion, the rectified and filtered EMG signals: $e$, angles and angular velocities of the shoulder and the elbow joints were used as the user state: $\psi = [e_1, \ldots, e_8, \theta_s, \theta_e, \dot{\theta}_s, \dot{\theta}_e]^\top \in R^{12}$. The EMG signals were measured from the right arms. Figure 5 shows the EMG channel locations probed to measure the muscle activities using eight sensor channels. We used Ag/AgCl bipolar
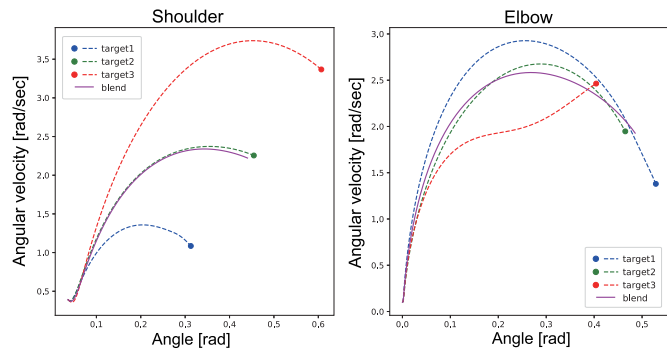
Fig. 6. Simulation results. Optimized trajectories for 1 m, 2 m and 3 m are plotted with dashed lines with blue, green, and red colors, respectively. Purple solid line shows generated trajectory by blending two optimal policies for distances of 1 m and 3 m with coefficients of $\omega_1 = 0.5$ and $\omega_2 = 0.5$. Results show that similar movement trajectory can be generated by blending pre-computed policies for novel task goal with target at 2 m.
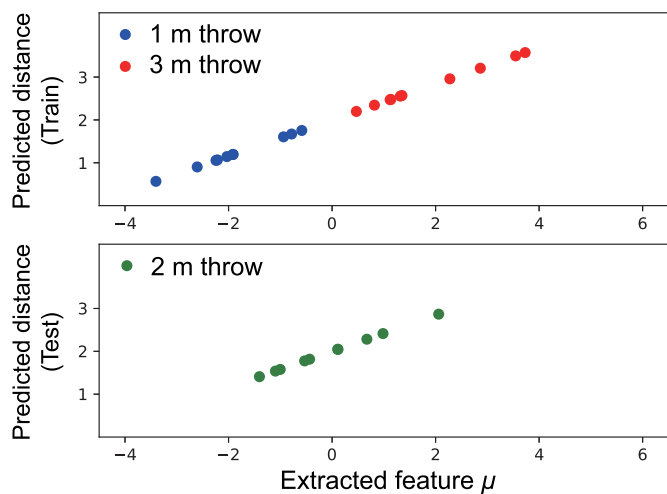


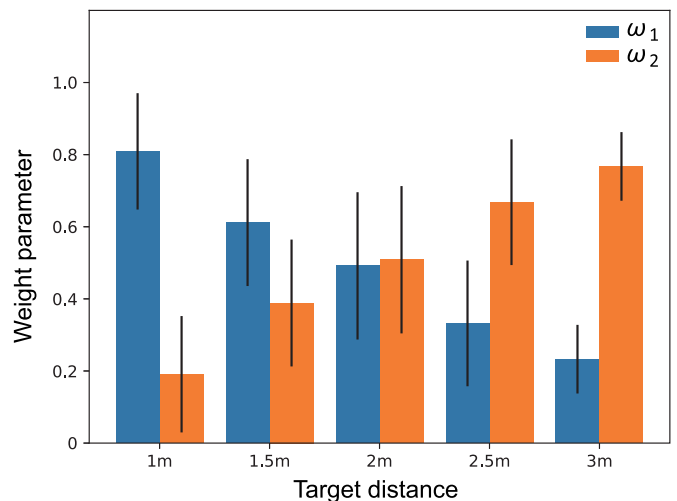Fig. 7. Relationships between feature values obtained by PLS and corresponding user's motor goals.



Fig. 8. Average of the ten user's motor goal estimation results. Weight parameter of one target was obtained by averaging five times throwing, and we measured the weights in each distance for each subject.
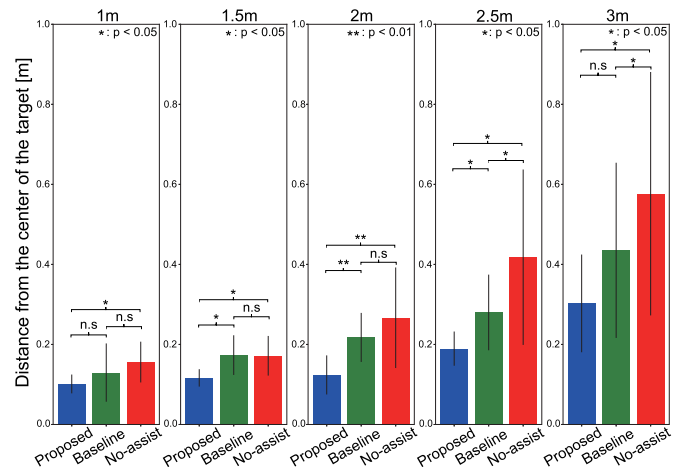


Fig. 9. Average of ball throwing accuracy for all subjects in each distance. Data for each subject in this each distance condition was obtained by averaging mean of the accuracy of five throws. Benjamini-Hochberg adjusted Wilcoxon signed-rank test was applied to performances of three methods in each distance condition.

surface EMG electrodes. Using the encoders of the upper-limb exoskeleton robot system, we simultaneously obtained the angles of the shoulder and the elbow joints. The angular velocities were derived by numerical differentiation of the measured joint angles.

During the basketball throwing motion, the point when the angular velocity of the shoulder joint exceeds a threshold was defined as the actuation starting point. Through our preliminary trials, the threshold was set to 0.2 rad/sec. We then averaged the user state $\psi$ for tens of milliseconds before the detected actuation stating point to derive the averaged state $\overline{\psi}$.

In this study, the training data were acquired when the exoskeleton robot assisted the shooting motion towards the center of the target at distances of 1 m and 3 m. Concretely, subjects threw the ball ten times at each distance condition in this training data acquisition phase before the actual test experiments. We collected user's EMG signals, joint angles, and angular velocities data for each distance and stored with the output values

of 1 and 3. Then, a projection matrix $\mathbf{W}$ for calculating the one-dimensional feature $\mu \in \mathcal{R}$ in (1) was derived by PLS.

We used a logistic regression to estimate the weight parameter for $i$-th task goal:

$$\omega_i = \frac{1}{1 + \exp(-a_i \mu - b_i)}, \tag{11}$$

where $a_i$ and $b_i$ are parameters for the $i$-th task goal and determined by using the training data.

## V. RESULTS

In this section, we first present the simulation results to show how the linear Bellman combination strategy works to derive throwing movements. We then show the comparisons of online

a) Proposed method at 2-m distance

b) Baseline method at 2-m distance

Scaled Time

Fig. 10. Control performance of ball throwing experiment with proposed and baseline methods.
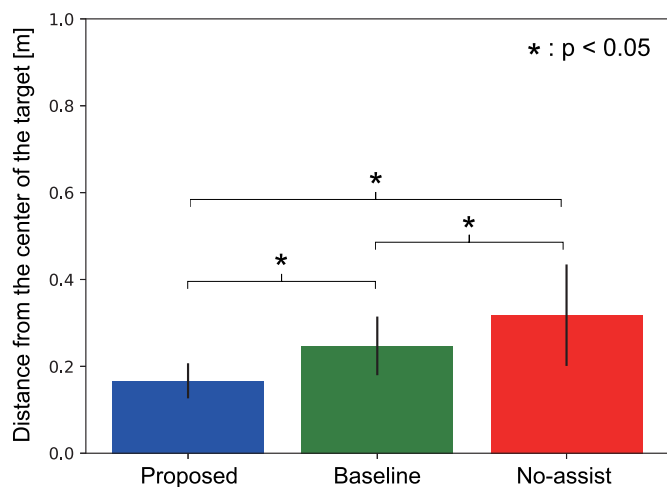


Fig. 11. Average of ball throwing accuracy for all subjects. Data for each subject in this figure was obtained by averaging mean of the accuracy of five throws at each distance over entire distance. Benjamini-Hochberg adjusted Wilcoxon signed-rank test was applied to three accuracy.

assist control performances among a baseline method, no assist and the proposed method.

### A. Simulation Results

Figure 6 shows the simulation results. Optimized trajectories for 1 m, 2 m and 3 m are plotted with dashed lines with blue, green, and red colors, respectively. These dashed lines in the figure show how the pre-computed optimal policies work. The purple solid line shows generated trajectory by blending the two optimal policies for the distances of 1 m and 3 m with weight parameters of $\omega_1 = 0.5$ and $\omega_2 = 0.5$. These weight parameters in this simulation results were fixed without being estimated from the EMGs. Results show that the similar movement trajectory to the trajectory generated by the optimized policy for the distance at 2 m can be generated by blending pre-computed policies. Note that since the corresponding terminal costs for the 2-m policy and the blended policy are different, trajectories generated by two

policies are naturally different but very close. Here, with these results, we tried to demonstrate the capability of the blending approach to generate a movement trajectory for a novel task goal.

### B. Feature Extraction

Here, we validated the relationships between feature values obtained by PLS and corresponding user's motor goals. The *user's motor goal* represents the aiming distance at which the subjects tried to throw the ball. Figure 7 shows the relationships. The parameters for PLS were trained by using the 1-m and 3-m target shot data, and tested with the 2-m target shot data. In this study, based on these results, we set the weight parameter for the 1-m target shot as $\omega_1 = 1 - \omega_2$ while $\omega_2$ for the 3-m target shot was derived as in (11). Figure 8 shows the average of ten user's estimated motor goals when throwing the ball with our proposed method at each target distance in online control test. Note that this weight (estimated user's motor goal) is one of the parameter to determine the mixing coefficient in equation (5). When the target distance was close to 1 m, the weight parameter of $\omega_1$ increased and the weight parameter of $\omega_2$ decreased. Conversely, when the target distance was close to 3 m, the weight parameter of $\omega_1$ became smaller and the weight parameter of $\omega_2$ became larger. In particular, the weight parameters at the intermediate distance (2 m) are almost the same for $\omega_1$ and $\omega_2$, and are similar to the weight parameters of the simulation result in Fig. 6. These results show that the motor goals were properly estimated according to each target distance.

### C. Task Performances

Here, we show the experimental results of the assisted basketball-throwing task with our proposed method, baseline method and no-assist condition. Detailed implementation of the baseline method is provided in Appendix. We show the evaluation results of the control policies derived from our proposed approach in terms of motion precision.

Figure 9 shows the ball throwing accuracy for each distance. The accuracy was obtained by measuring the distance from the

center of the board of the target to that of the ball when the ball hit the board. The data for each subject in this each distance condition was obtained by averaging the mean of the accuracy of five throws. We applied Benjamini-Hochberg adjusted Wilcoxon signed-rank test to three accuracy in each distance condition. In every distance condition, significant differences were found between our proposed approach and no-assist condition: 1-m distance condition ($p = 0.041$), 1.5-m distance condition ($p = 0.015$), 2-m distance condition ($p = 0.0066$), 2.5-m distance condition ($p = 0.015$), 3-m distance condition ($p = 0.043$). On the other hand, no-significant differences were found between the baseline method and no-assist condition except for 2.5-m distance and 3-m distance conditions: 2.5-m distance ($p = 0.037$), 3-m distance ($p = 0.043$). This was considered that the baseline method was also showing the assist effect under the conditions where the distance was long and the task difficulty rose. Next, we compared our proposed method with the baseline method. Although the proposed method had the highest accuracy, no-significant differences were found between the accuracy under the proposed and baseline methods when throwing the ball at a distance of 1 m and 3 m. However, under these conditions, there is no need to blend, which was a reasonable result. On the other hand, at a distance between 1 m and 3 m, significant differences were found between our proposed and baseline methods: 1.5-m distance ($p = 0.019$), 2-m distance ($p = 0.0048$), 2.5-m distance ($p = 0.019$). From these results, the effect of the control policies derived by blending was clearly seen at distance conditions between 1 m and 3 m (see also Fig. 10).

Here we also show the average of ball throwing accuracy for entire distance in Fig. 11. The data for each subject in this figure was obtained by averaging the mean of the accuracy of five throws at each distance over the entire distance. We applied the Benjamini-Hochberg adjusted Wilcoxon signed-rank test to performances of three methods, and significant differences are found between our proposed and baseline methods ($p = 0.010$), between our proposed method and no-assist condition ($p = 0.010$), and between baseline method and no-assist condition ($p = 0.022$). From these results, it was clearly shown that our proposed approach is effective in assisting the control for dynamic motion that requires accuracy.

## VI. Conclusion

Obtained results showed that the proposed approach for controlling the assistive exoskeleton robot was successful in the actual basketball throwing experiments. Furthermore, the successful real-time online control performances clearly showed the computational efficiency of our proposed approach. The ball throwing accuracy with the proposed method was the highest among all the conditions. These results showed that our proposed approach effectively improved robot-assisted task performances.

To further improve the throwing accuracy with our proposed approach, it can be beneficial to explicitly take the control of other joints, e.g., shoulder abduction/adduction, wrist joint, into account.

TABLE I
SELECTION RATES OF EACH PRE-COMPUTED POLICY BY BASELINE METHOD

| Ball shooting distance | Average | |
|---|---|---|
| | $\pi_1$ | $\pi_2$ |
| 1 m | 74% | 26% |
| 1.5 m | 72% | 28% |
| 2 m | 40% | 60% |
| 2.5 m | 26% | 74% |
| 3 m | 18% | 82% |

In this study, control policies were blended by using a linear Bellman combination approach, where control policies for the intermediate distance condition between 1 m and 3 m were derived based on the optimal control policies for the 1 m and 3 m conditions. In many tasks, adaptive change of the control output is required to cope with the continuous change of the task goal even in the same task context. In this study, we addressed this issue rather than dealing with totally different tasks. Therefore, the range of the task need to be predetermined. Our findings suggest that this approach can be useful for the coordination of new tasks within the expected range in assistive control contexts, although further studies are required to generate wider variety of movements with blending more pre-optimized polices. In addition, using other user's data for deriving the feature space to reduce the data acquisition burden would be an interesting topic as a future study [28], [29].

## Appendix

Here, we introduce the baseline method. In the baseline classification method, a standard soft margin SVM model was derived and evaluated with using the same training and test data as we used for our proposed method. The averaged user state $\overline{\psi}$ was input and the output for the 1-m target shot data was labeled as $y = -1$ and $y = +1$ for the 3-m. Hyper parameters of the SVM model were searched with using 5-fold cross validation. Then, the control policy $\pi^b$ of the baseline method for the test trials was derived by selecting one of the pre-computed optimal policies as the output of the SVM model:

$$\pi^b \leftarrow \begin{cases} \pi_1 & \text{if } y = -1 \\ \pi_2 & \text{if } y = +1 \end{cases} \qquad (12)$$

Table I shows the selection rates of each pre-computed policy averaged over the ten subjects in the ball-throwing trials. When the target distance was close to 1 m, the selection rate of using $\pi_1$ increased and the rate of using $\pi_2$ decreased. Conversely, when the target distance was close to 3 m, the selection rate of using $\pi_1$ became smaller and that of using $\pi_2$ became larger. These results show that the SVM model was properly trained and the baseline method reasonably selected pre-computed policies.

## References

[1] J. Ngeo *et al.*, "Control of an optimal finger exoskeleton based on continuous joint angle estimation from EMG signals," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Jul. 2013, pp. 338–341.

[2] C. J. Gearhart, B. Varone, M. H. Stella, and B. F. BuSha, "An effective 3-fingered augmenting exoskeleton for the human hand," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2016, pp. 590–593.

[3] J. Furukawa, T. Noda, T. Teramae, and J. Morimoto, "Estimating joint movements from observed EMG signals with multiple electrodes under sensor failure situations toward safe assistive robot control," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2015, pp. 4985–4991.

[4] A. J. Young and D. P. Ferris, "State of the Art and Future Directions for Lower Limb Robotic Exoskeletons," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 2, pp. 171–182, Feb. 2017.

[5] K. Gui, H. Liu, and D. Zhang, "A generalized framework to achieve coordinated admittance control for multi-joint lower limb robotic exoskeleton," in *Proc. Int. Conf. Rehabil. Robot.*, Jul. 2017, pp. 228–233.

[6] K. Nagata and K. Magatani, "Basic study on combined motion estimation using multichannel surface EMG signals," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2011, pp. 7865–7868.

[7] P. K. Artemiads and K. J. Kyriakopoulos, "An EMG-Based Robot Control Scheme Robust to Time-Varying EMG Signal Features," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 582–588, May 2010.

[8] C. Castellini *et al.*, "Proceedings of the first workshop on Peripheral Machine Interfaces: going beyond traditional surface electromyography," *Front. Neurorobot.*, vol. 8, 2014, Art. no. 22.

[9] D. Ao, R. Song, and J. Gao, "Movement Performance of Human–Robot Cooperation Control Based on EMG-Driven Hill-Type and Proportional Models for an Ankle Power-Assist Exoskeleton Robot," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 8, pp. 1125–1134, Aug. 2017.

[10] J. Furukawa, T. Noda, T. Teramae, and J. Morimoto, "Human Movement Modeling to Detect Biosignal Sensor Failures for Myoelectric Assistive Robot Control," *IEEE Trans. Robot.*, vol. 33, no. 4, pp. 846–857, Aug. 2017.

[11] C. Fleischer and G. Hommel, "A Human–Exoskeleton Interface Utilizing Electromyography," *IEEE Trans. Robot.*, vol. 24, no. 4, pp. 872–882, Aug. 2008.

[12] M. A. Oskoei and H. Hu, "Support Vector Machine-Based Classification Scheme for Myoelectric Control Applied to Upper Limb," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 8, pp. 1956–1965, Aug. 2008.

[13] Q. She, Z. Luo, M. Meng, and P. Xu, "Multiple Kernel learning SVM-based EMG pattern classification for lower limb control," in *Proc. 11th Int. Conf. Control Automat. Robot. Vis.*, 2010, pp. 2109–2113.

[14] F. Bian, R. Li, and P. Liang, "SVM based simultaneous hand movements classification using sEMG signals," in *IEEE Int. Conf. Mechatronics Automat.*, 2017, pp. 427–432.

[15] G. Purushothaman and R. Vikas, "Identification of a feature selection based pattern recognition scheme for finger movement recognition from multichannel EMG signals," *Australas Phys. Eng. Sci. Med.*, vol. 41, no. 2, pp. 549–559, Jun. 2018.

[16] L. J. Hargrove, E. J. Scheme, K. B. Englehart, and B. S. Hudgins, "Multiple Binary Classifications via Linear Discriminant Analysis for Improved Controllability of a Powered Prosthesis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 1, pp. 49–57, Feb. 2010.

[17] D. Zhang, X. Zhao, J. Han, and Y. Zhao, "A comparative study on PCA and LDA based EMG pattern recognition for anthropomorphic robotic hand," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2014, pp. 4850–4855.

[18] Y. Yu, X. Sheng, W. Guo, and X. Zhu, "Attenuating the impact of limb position on surface EMG pattern recognition using a mixed-LDA classifier," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, 2017, pp. 1497–1502.

[19] E. Campbell, A. Phinyomark, and E. Scheme, "Linear Discriminant Analysis with Bayesian Risk Parameters for Myoelectric Control," in *Proc. IEEE Glob. Conf. Signal Inf. Process.*, 2019, pp. 1–5.

[20] M. da Silva, F. Durand, and J. Popović, "Linear Bellman Combination for Control of Character Animation," *ACM Trans. Graph.*, vol. 28, no. 3, 2009, Art. no. 82.

[21] H. Wold, "Soft Modelling by Latent Variables: The Non-Linear Iterative Partial Least Squares (NIPALS) Approach," *J. Appl. Probability*, vol. 12, no. S1, pp. 117–142, 1975.

[22] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: a basic tool of chemometrics," *Chemometrics Intell. Lab. Syst.*, vol. 58, no. 2, pp. 109–130, 2001.

[23] E. Todorov, "Compositionality of optimal control laws," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 1856–1864.

[24] E. Todorov and W. Li, "A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems," in *Proc. Amer. Control Conf.*, 2005, pp. 300–306.

[25] T. Noda, T. Teramae, B. Ugurlu, and J. Morimoto, "Development of an upper limb exoskeleton powered via pneumatic electric hybrid actuators with bowden cable," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2014, pp. 3573–3578.

[26] T. Teramae, T. Noda, and J. Morimoto, "Optimal control approach for pneumatic artificial muscle with using pressure-force conversion model," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2014, pp. 4792–4797.

[27] D. A. Winter, *Biomechanics and Motor Control of Human Movement*, Hoboken, NJ, USA: Wiley, 2009.

[28] T. Matsubara and J. Morimoto, "Bilinear Modeling of EMG Signals to Extract User-Independent Features for Multiuser Myoelectric Interface," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 8, pp. 2205–2213, Aug. 2013.

[29] J. Furukawa, A. Takai, and J. Morimoto, "Database-driven approach for biosignal-based robot control with collaborative filtering," in *Proc. IEEE-RAS 17th Int. Conf. Humanoid Robot. (Humanoids)*, 2017, pp. 606–611.