Motor synergy generalization framework for new targets in multi-planar and multi-directional reaching task

Kyo Kutsuzawa and Mitsuhiro Hayashibe

Department of Robotics, Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan
KK, 0000-0002-5326-7847

Humans can rapidly adapt to new situations, even though they have redundant degrees of freedom (d.f.). Previous studies in neuroscience revealed that human movements could be accounted for by low-dimensional control signals, known as motor synergies. Many studies have suggested that humans use the same repertories of motor synergies among similar tasks. However, it has not yet been confirmed whether the combinations of motor synergy repertories can be re-used for new targets in a systematic way. Here we show that the combination of motor synergies can be generalized to new targets that each repertory cannot handle. We use the multi-directional reaching task as an example. We first trained multiple policies with limited ranges of targets by reinforcement learning and extracted sets of motor synergies. Finally, we optimized the activation patterns of sets of motor synergies and demonstrated that combined motor synergy repertories were able to reach new targets that were not achieved with either original policies or single repertories of motor synergies. We believe this is the first study that has succeeded in motor synergy generalization for new targets in new planes, using a full 7-d.f. arm model, which is a realistic mechanical environment for general reaching tasks.

1. Introduction

Humans can generate complex movements (such as object manipulation and locomotion) in diverse situations, using a number of muscles and joints in a coordinated manner. Previous studies have revealed that human movements can be accounted for by a small number of primitive components, known as motor synergies [1–3]. The central nervous system is considered to simplify the control problem of complex and redundant musculoskeletal systems by linearly combining a small number of
motor synergies [4,5]. Another feature of motor synergies is that the same synergy structures are often observed from various movements performed for different targets [6–9]. Such motor synergies are referred to as shared synergies, and they refer to the re-usability of motor synergy structures. It has also been reported that animals, including humans, explore combination patterns of previously acquired motor synergies during motor skill learning [10,11], suggesting that motor synergies can be combined and re-used to deal with new targets. Computationally, seeking motor synergy combinations can reduce the search space in motor learning.

There are previous studies that address the motor synergy structure computationally. The computational approach facilitates the analysis and evaluation of the role of each component using simple models that extract important factors of the actual system. For example, some studies have reported that motor synergy structures can be observed from trajectories acquired by machine learning and optimization: zero-moment point-based locomotion [12], feedback torque integration [13], optimal control [14,15], trajectory optimization [16], reinforcement learning for gait [17] and reaching [18]. These studies provide a methodology for approaching motor synergies without assuming how motor synergies are implemented in animals. However, the aspect of task generalization through the obtained synergy has not been fully discussed. In addition, Rückert & D’Avella proposed a framework that explicitly implemented a shared-synergy structure based on dynamic movement primitives [19,20], movement representations similar to motor synergies [21]. There are also some studies that examined the generalization ability of single repertoires of motor synergies. Al Borno et al. evaluated reaching tasks using a musculoskeletal model and reported that a single repertoire of motor synergies obtained by trajectory optimization could be generalized to different initial poses and targets [22]. Hagio & Kouzaki demonstrated that a hard-wired motor synergy structure could accelerate motor learning against changes in musculoskeletal structures [23]. These studies reveal the generalization ability of single repertoires of motor synergies; however, determining whether combinations of multiple motor synergy repertoires can be generalized to new targets remains unclear. In this study, to address this question, we use multiple motor synergy repertoires extracted from motor control agents (i.e. policies) that are individually trained with different targets, and we investigate the systematic generalization framework of motor synergies.

This study demonstrates that the combinations of motor synergies extracted from multiple policies can be re-used for generalization of new targets in multi-directional three-dimensional (3D) reaching tasks. We employed several types of target positions in each of which the target positions were located on different planes. The main contribution of this study is to demonstrate the generalization ability of motor synergies extracted from individually learned motor skills. Alternatively, this study provides a framework for explaining and reproducing human’s ability to generalize to new situations by combining previously acquired skills, which will be beneficial to neuroscience, machine learning and robotics.

2. Material and methods

2.1. Overview

This study demonstrates that the ability of a single repertoire of motor synergies can be extended by combining another repertoire of motor synergies. In this study, we use spatio-temporal synergies to represent the motor synergies as follows:

\[ x(t) \approx \sum_{l=0}^{L-1} h_l w_l(t), \]  

(2.1)

where \( x(t), w_l(t) \) and \( h_l \) indicate the movement, spatio-temporal synergy and activity amplitude, respectively. Considering this equation, it is expected that more diverse movements can be expressed by increasing the number of motor synergies. We demonstrate a more special case in this study. We evaluate how representation ability can be improved by combining motor synergies extracted from different policies. To accomplish that, we construct a method that consists of (i) motor learning to acquire multiple policies, (ii) motor synergy extraction from individual policies, and (iii) generalization to new targets with combined motor synergy repertoires.

Figure 1 illustrates an overview of the method. We used tasks as illustrated in figure 1a,b. The method consists of motor learning, motor synergy extraction, combination, and optimization for new targets (figure 1c). First, multiple policies are trained for different types of targets using a deep reinforcement
Thereafter, motor synergies are extracted from the time series of the joint torque performed by each policy. Finally, combining the repertoires of motor synergies linearly, we applied an optimization method for adjusting motor synergy activities to accomplish other types of targets. These are newly provided to investigate task generalization (figure 1). 

2.2. Task description

We used a seven-degrees-of-freedom (d.f.) arm such that its kinematics resembled that of human arms. The overview of the arm is illustrated in figure 1a. The shoulder, elbow and wrist in this arm have three, two and two rotational joints, respectively. The joints are controlled by torque commands. We used MuJoCo [24] for physics simulation. The joint specifications are described in table 1.

This study handles a multi-directional reciprocating reaching task; the arm first moves to a target point at a specific time and returns to the initial position at a predetermined finish time. The same type of task is often employed in studies of human arm reaching [3,25]. An observation variable of learning method (figure 1d). Thereafter, motor synergies are extracted from the time series of the joint torque performed by each policy. Finally, combining the repertoires of motor synergies linearly, we applied an optimization method for adjusting motor synergy activities to accomplish other types of targets. These are newly provided to investigate task generalization (figure 1e).
the task at time $t$, $s(t)$, consists of seven-dimensional joint angles $\theta(t)$ (rad), seven-dimensional joint angular velocities $\dot{\theta}(t)$ (rad s$^{-1}$), three-dimensional target position $q_{\text{target}}$ (m), and normalized time $\phi(t)$. Here, $q_{\text{target}}$ is fixed for each episode. The normalized time is represented as follows:

$$\phi(t) = \frac{t}{t_{\text{finish}}},$$

where $t_{\text{finish}}$ indicates the time when the task should have been terminated. Action $a(t)$ consists of a seven-dimensional joint torque $\tau(t)$ (Nm). An immediate reward $r(t)$ is expressed as follows:

$$r(t) = \begin{cases} 
\|q_{\text{target}} - q(t)\|^2 + k_1\|\dot{q}(t)\|^2 + k_2\|a(t)\|^2\Delta t, & t = t_{\text{target}}, \\
\|q_{\text{initial}} - q(t)\|^2 + k_1\|\dot{q}(t)\|^2 + k_2\|a(t)\|^2\Delta t, & t = t_{\text{finish}}, \\
k_2\|a(t)\|^2\Delta t, & \text{otherwise}. 
\end{cases}$$

This reward function forces the arm tip to reach $q_{\text{target}}$ at $t_{\text{target}}$ and returns to $q_{\text{initial}}$ at $t_{\text{finish}}$ while attempting to reduce the joint torques during the task. Here, $q_{\text{initial}} = [0.24\,\text{m}, 0.0\,\text{m}, 0.21\,\text{m}]^T$. $q_{\text{target}}$ indicates the target position to be reached. $q(t)$ and $\dot{q}(t)$ indicate the position and velocity at the centre of gravity of the fingers, respectively. The coefficient values were set to $k_1 = 0.002$ and $k_2 = 0.2$. The sampling interval of the policy was set to $\Delta t = 10$ ms. Each episode consists of 100 samples of interaction, (i.e. 1 s). In addition, $t_{\text{target}}$ and $t_{\text{finish}}$ were set to 50 and 100 (i.e. 0.5 and 1 s), respectively.

We employed four types of target positions: cases in which the target points were on a horizontal, sagittal, frontal and upper-horizontal plane. We refer to them as horizontal, sagittal, frontal and upper targets. The target positions in each case are defined as follows:

$$q_{\text{target}} = q_{\text{initial}} + [-l \sin \theta_l, -l \cos \theta_l, 0]^T,$$

$$q_{\text{target}} = q_{\text{initial}} + [-l \sin \theta_l, 0, -l \cos \theta_l]^T,$$

$$q_{\text{target}} = q_{\text{initial}} + [0, -l \sin \theta_l, -l \cos \theta_l]^T,$$

$$q_{\text{target}} = q_{\text{initial}} + [-l \sin \theta_l, -l \cos \theta_l, z]^T,$$

where

$$l = 0.15\,\text{m},$$

$$\theta_l \in \left\{ \frac{2\pi}{8} m = 0, 1, \ldots, 7 \right\}$$

and

$$z = 0.05\,\text{m}. $$

The target positions are illustrated in figure 1b. As explained below, we use only the horizontal and sagittal targets for the initial policy acquisition to acquire each repertory motor synergy, and the other targets are used for synergy-based optimization.

### 2.3. Motor synergy acquisition through reinforcement learning

We first trained policies using proximal policy optimization (PPO) [26], a type of reinforcement learning method. Each policy consists of a neural network with two hidden layers as illustrated in figure 1d. Each
hidden layer has 256 units with tanh activation. During the training, critic networks with the same hidden layer architecture as the policy networks were used. Another set-up for PPO is detailed in table 2.

After the training of the policies with PPO, motor synergies are extracted. We used non-negative matrix factorization (NMF) to extract spatiotemporal synergies.

Spatio-temporal synergies indicate primitive patterns of multi-d.f. movements linked both temporally and spatially. It can be extracted with NMF from a matrix of movements in multiple trials, $X$, as follows:

$$X \approx WH,$$

where $W$ and $H$ indicate the motor synergies and their activity amplitudes in individual trials, respectively. Among the various definitions of motor synergies [28], spatio-temporal synergies can compress both spatial and temporal patterns of the time series of actions.

Before extracting motor synergies, it is necessary to pre-process the sequences of $M$ d.f. actions that can take both positive and negative values because NMF can only handle non-negative values. First, an action at time $t$ in the $n$th trial, $a_n(t)$, is converted to two non-negative values, $a_n^+(t)$ and $a_n^-(t)$, as follows:

$$a_n^+(t) = \max(a_n(t), 0)$$

and

$$a_n^-(t) = -\min(a_n(t), 0),$$

where $a_n^+(t)$ and $a_n^-(t)$ indicate the positive and negative parts of $a_n(t)$, respectively. By definition, $a_n^+(t) - a_n^-(t) = a_n(t)$ holds. Thereafter, an action matrix $X$ is constructed as follows:

$$X = \begin{bmatrix}
a_0^+(0) & \cdots & a_N^-(0) \\
\vdots & \ddots & \vdots \\
a_0^+(T-1) & \cdots & a_N^-(T-1) \\
a_0^0(0) & \cdots & a_N^0(N-1) \\
\vdots & \ddots & \vdots \\
a_0^0(T-1) & \cdots & a_N^0(N-1)
\end{bmatrix},$$

where $T$ and $N$ indicate the number of actions in a trial and the number of trials, respectively. Because $a_n^+(t)$ and $a_n^-(t)$ can be regarded as $M \times 1$ matrices, $X$ has $2MT \times N$ elements.

Finally, $X$ can be decomposed into spatio-temporal synergies and their activities with NMF as in (2.11). Here, $W \in \mathbb{R}^{2MT \times L}$ and $H \in \mathbb{R}^{L \times N}$, where $L$ is the number of motor synergies. In this study, $T = 100$ and $M = 7$ are used.

2.4. Generalization to new targets

By training multiple policies considering the targets, we can obtain multiple repertoires of motor synergies. Supposing we have two sets of the motor synergies for the horizontal and sagittal reaching targets, let the numbers of motor synergies be $L_h$ and $L_s$, respectively. Combining them, we can finally use $L' \equiv L_h + L_s$ motor synergies to acquire new movements.

| item                        | value |
|-----------------------------|-------|
| no. training epochs         | 3000  |
| steps per epoch             | 10 000|
| discount factor              | 0.99  |
| clipping ratio              | 0.2   |
| $\lambda$ for generalized advantage estimator [27] | 0.97  |
| learning rate for the policy network | 0.0003|
| learning rate for the critic network | 0.001 |
| number of gradient descent steps per epoch | 80    |
Using these motor synergies, we finally optimized the magnitude of their activities (i.e. $h_l$ in equation (2.1)) for new targets. In summary, the optimization was performed using a numerical optimization method that explores solutions in a way of evolutionary computation.

We optimized the magnitude of the $L'$ motor synergy activities, denoted as $h' \in \mathbb{R}^{L'}$, using the covariance matrix adaptation evolution strategy (CMA-ES) [29,30]. An overview is illustrated in figure 1e. The covariance matrix adaptation evolution strategy is a stochastic optimization method that uses a Gaussian distribution to search for a solution. It searches an $L'$-dimensional space in which the coordinates correspond to the activities of motor synergies, and it maximizes the following accumulated reward $G$:

$$G = \sum_{t=0}^{t_{\text{finish}}} r(t). \quad (2.15)$$

Considering each optimization step, $(4 + \lfloor 3 \log L' \rfloor)$ candidates were sampled from a Gaussian distribution, and the parameters of the Gaussian distribution (i.e. the mean vector and covariance matrix) were updated through the evaluation of the sampled candidates. Here, $\lfloor \cdot \rfloor$ indicates the floor function. The initial parameters of the Gaussian distribution were set to $\mu = 0$ and $\sigma = 10^{-3}$. The CMA-ES explores the entire $L'$-dimensional space containing negative values, whereas the motor synergy activity should be positive. To limit the solutions to positive, a sampled candidate $x \in \mathbb{R}^{L'}$ was converted to the motor synergy activity $h'$ using the element-wise absolute value function, $|\cdot|$, as follows:

$$h' = |x|. \quad (2.16)$$

The motor synergy activity, $h'$, was finally decoded to a time series of actions $\hat{a}(0), \ldots, \hat{a}(T-1)$ as follows:

$$\begin{bmatrix}
\hat{a}_+(0) \\
\vdots \\
\hat{a}_+(T-1) \\
\hat{a}_-(0) \\
\vdots \\
\hat{a}_-(T-1)
\end{bmatrix} = [W_h \ W_s] h' \quad (2.17)$$

and

$$\hat{a}(t) = \hat{a}_+(t) - \hat{a}_-(t), \quad t = 0, \ldots, T-1, \quad (2.18)$$

where $W_h$ and $W_s$ indicate the motor synergies in the horizontal and sagittal targets, respectively.

Here we used CMA-ES instead of PPO because of the difference in learning targets. As the motor synergy activities correspond to trajectories one by one, we did not need to optimize sequential decision making but only optimize single variables. Also, we did not use neural networks but directly optimized motor synergy activities. Thus, it was suitable to use a numeric optimization method.

3. Results

3.1. Motor synergy emergence by deep reinforcement learning

Two policies were trained with PPO for the horizontal and sagittal targets (figure 1b). We refer to these policies as the horizontal and sagittal policies, respectively. Figure 2ae shows the snapshots of the horizontal and sagittal policies. The movements are also available in electronic supplementary material, movie S1. Also, figure 2hf shows the trajectories performed by the horizontal and sagittal policies. In addition, the trajectories during the training are shown in figure 2cg. The errors at $t = t_{\text{target}}$, $t_{\text{finish}}$ in the eight learned targets are summarized in table 3.

To further confirm the ability of the policies, we evaluate other target positions using different target angles $\theta$ as follows:

$$\theta \in \left\{ \frac{2\pi n}{32} \mid n = 0, 1, \ldots, 31 \right\}, \quad (3.1)$$

in which 24 targets have not been given during training. Figure 2dh shows that 32 trajectories performed horizontal and sagittal policies.
We extracted synergy structures from the actions (i.e. joint torques) performed by the horizontal and sagittal policies. Considering the eight trials for each policy, we evaluated the $R^2$-values to quantify the reconstruction performance for various motor synergies. The $R^2$-values are computed as follows:

$$R^2 = 1 - \frac{\sum_{t=0}^{T-1} ||a(t) - \hat{a}(t)||^2}{\sum_{t=0}^{T-1} ||a(t) - \bar{a}||^2},$$

where $a(t)$ and $\hat{a}(t)$ indicate the original and reconstructed actions at time $t$, respectively, and $\bar{a}$ indicates the mean values of the actions over the trial. Figure 3a,d shows $R^2$-values of the motor synergies extracted from the horizontal and sagittal policies. Four or more motor synergies achieved $R^2 > 0.95$ in both the horizontal and sagittal policies.

Figure 3b,e shows the waveforms of the four motor synergies extracted from eight reaching movements shown in figure 2b,f. Regarding the figures, the red positive values correspond to the components of $\hat{a}_x(t)$, whereas the purple-coloured negative values indicate the components of $\hat{a}_y(t)$. Moreover, figure 3c,f shows the synergy activities for the target position’s directions $\theta$. Considering the figures, the marked points correspond to the learned targets shown in figure 2b,f, whereas the curved lines correspond to the 32 targets shown in figure 2d,h.
3.2. Generalization to new target positions with motor synergies

We optimize the magnitude of the motor synergy activities $h'$ using the method illustrated in figure 1e to obtain new movements. First, we applied the method of the reaching tasks to the horizontal and sagittal
used, the cases with \( L_h = 5 \) and \( L_s = 0 \) are even different. In human studies, measured data usually contain movements of multiple subjects who

targets to investigate the ability of the motor synergies. Figure 4 shows the results in the horizontal and sagittal targets, which were used in reinforcement learning. Figure 4b,e shows the errors in \( q_{\text{target}} \) with varying numbers of motor synergies. The trajectories in the horizontal targets in \( L_h = 5 \) and \( L_s = 0 \) are shown in figure 4c. In addition, the trajectories in the sagittal targets in \( L_h = 0 \) and \( L_s = 5 \) are shown in figure 4f.

Subsequently, we evaluated the method in the frontal targets, which can be regarded as new targets in a new plane. Figure 5 shows the results. Figure 5a shows the reaching errors in \( q_{\text{target}} \). Changing the number of motor synergies, \( L_h \) and \( L_s \). Considering the results, the case with \( L_h = 5 \) and \( L_s = 8 \) resulted in the smallest error in \( q_{\text{target}} \). Figure 5c shows the trajectories in \( L_h = 5 \) and \( L_s = 4 \), where the average error is 0.621 cm; This was not the smallest error, whereas the difference from the smallest error was 0.013 cm, which was approximately 4% of the standard deviation. The movements in \( L_h = 5 \) and \( L_s = 4 \) are available in electronic supplementary material, movie S1. Figure 5d shows the trajectories during the optimization. Furthermore, considering the results when only single motor synergy repertoires are used, the cases with \( L_h = 5 \) and \( L_s = 8 \) resulted in the smallest errors. The trajectories are shown in figure 5e,f. Comparatively, the trajectories obtained by the policies are shown in figure 5g,h. The errors at \( t = t_{\text{target}}, \ t_{\text{finish}} \) are summarized in table 3.

Moreover, the same evaluation is conducted for the upper targets, which can be regarded as new targets in a new plane by sliding the horizontal plane. Figure 6a shows the reaching errors in \( q_{\text{target}} \). Changing the number of motor synergies, \( L_h \) and \( L_s \). Figure 6c shows the trajectories in \( L_h = 5 \) and \( L_s = 5 \), where the average error is 0.753 cm, and figure 6d shows the trajectories during the optimization. The movements in \( L_h = 5 \) and \( L_s = 5 \) are also available in electronic supplementary material, movie S1. Furthermore, considering the results of when only single motor synergy repertoires are used, the cases with \( L_h = 8 \) and \( L_s = 6 \) resulted in the smallest errors. The trajectories are shown in figure 6c,f. Comparatively, the trajectories obtained by the policies are shown in figure 6g,h. The errors at \( t = t_{\text{target}}, \ t_{\text{finish}} \) are summarized in table 3.

### 3.3. Transferability to different arms

We conducted an additional simulation to evaluate if the results so far are consistent when arm models are even different. In human studies, measured data usually contain movements of multiple subjects who
three arm models with different link lengths: the original model, a shorter-arm model and a longer-arm model. The link lengths were used: the original model, a shorter-arm model and a longer-arm model. The link lengths varied depending on the arm lengths. We then extracted motor synergies from the trajectories of each arm model and optimized their activities with CMA-ES for the original arm model.

Figure 7 shows the results in the frontal targets. Compared with figure 5a, the motor synergies by the shorter-arm model brought about slightly smaller errors, whereas the motor synergies by the longer-arm model brought about larger errors. When the motor synergies from the shorter-arm model were used, the error was 0.577 ± 0.27 cm in $L_h = L_s = 4$. In contrast, when the motor synergies from the longer-arm model were used, the error was 2.192 ± 1.08 cm in $L_h = L_s = 4$; this was approximately three times larger than that observed in figure 5a.

| Table 4. Link lengths of arm models. |
|-------------------------------------|
| model name             | shoulder-to-elbow (m) | elbow-to-wrist (m) |
| original model         | 0.36                  | 0.27               |
| shorter-arm model      | 0.31                  | 0.22               |
| longer-arm model       | 0.41                  | 0.32               |

have different kinematic parameters such as link lengths. Here, three arm models with different link lengths were used: the original model, a shorter-arm model and a longer-arm model. The link parameters are listed in table 4. Firstly, policies were trained for the shorter- and longer-arm models by using PPO with the same hyper-parameters as in table 2. In policy learning, target positions were calculated with equations (2.4) and (2.5), whereas $q$ initial varied depending on the arm lengths. We then extracted motor synergies from the trajectories of each arm model and optimized their activities with CMA-ES for the original arm model.

Figure 5. Optimization results in the frontal targets (new targets). (a) Error map in $q_{\text{target}}$ (cm). (b) Frontal targets. (c) Trajectories when $L_h = 5$ and $L_s = 4$. (d) Learning progress. Trajectories at the 0, 10, 20, 30, 40, 50, 100, 150, 200, 300, 400 and 500th generations are shown. (e) Trajectories when $L_h = 5$ and $L_s = 0$. (f) Trajectories when $L_h = 0$ and $L_s = 8$. (g) Trajectories by the horizontal policy. (h) Trajectories by the sagittal policy.
generations are shown. (b) Trajectories when \( L_h = 5 \) and \( L_s = 5 \). (c) Learning progress. Trajectories at the 0, 10, 20, 30, 40, 50, 100, 150, 200, 300, 400 and 500th generations are shown. (e) Trajectories when \( L_h = 8 \) and \( L_s = 6 \). (f) Trajectories when \( L_h = 0 \) and \( L_s = 6 \). (g) Trajectories by the horizontal policy. (h) Trajectories by the sagittal policy.

Figure 6. Optimization results in the upper targets (new targets). (a) Error map in \( q_{\text{target}} \) (cm). (b) Frontal targets. (c) Trajectories when \( L_h = 5 \) and \( L_s = 5 \). (d) Learning progress. Trajectories at the 0, 10, 20, 30, 40, 50, 100, 150, 200, 300, 400 and 500th generations are shown. (e) Trajectories when \( L_h = 8 \) and \( L_s = 6 \). (f) Trajectories when \( L_h = 0 \) and \( L_s = 6 \). (g) Trajectories by the horizontal policy. (h) Trajectories by the sagittal policy.

4. Discussion

4.1. Performance of the learned movements

The horizontal and sagittal policies trained with reinforcement learning demonstrated high performance and generalization ability; nonetheless, there were limitations. It is be observed that these policies have learned the reaching task and have demonstrated the generalization ability for the intermediate target positions between the trained ones (figure 2 and table 3). By contrast, when these policies are applied without a combination of policies, the performance is greatly degraded in the frontal and upper targets as shown in figure 5g, h, 6g, h and table 3. The difference in the performance arises from whether the new targets are the interpolation or extrapolation of learned targets. Considering the same target types (i.e. where only the angle \( \theta \) is changed), targets will be located at approximately the intermediate positions between the trained ones. Regarding such cases, the policies can easily succeed in the reaching tasks by interpolating learned behaviours. By contrast, when the targets were changed to the frontal targets, most of the targets were no longer at the intermediate positions between the targets in the horizontal or sagittal targets. Therefore, policies have to extrapolate from learned behaviours, which is generally difficult. It can be observed that the arm was moved to the points where the target is projected onto the horizontal/sagittal plane. The policies attempted to reach the frontal and upper targets by interpolating the learned behaviours.

The endpoint trajectories show a characteristic feature as observed in human reaching movements. The trajectories were almost along straight lines to the targets. Although the trajectories were slightly drooping in the z-direction owing to gravity, the policy compensated most of the gravity effect. The
Motor synergies are able to reconstruct the original actions (figure 3) and achieve the reaching task (figure 4). Only four motor synergies resulted in $R^2 > 0.95$ in both the horizontal and sagittal targets. Moreover, the average target errors could be reduced to approximately 1 cm in both the horizontal and sagittal targets with only five motor synergies. It can also be observed that the trajectories of motor synergies (figure 4) resemble those of the policies (figure 2). Because a single movement
consists of 100 actions of seven joints, the motor synergies significantly and efficiently reduce the dimensionality of the movements.

We observe the relationship between the reconstruction accuracy and performance. The selection of motor synergies with high reconstruction accuracy resulted in high performance, suggesting that the reconstruction accuracy could be used to determine the appropriate number of motor synergies for motor control. Previous studies that applied motor synergies to robot control employed reconstruction accuracy to determine the number of motor synergies to use [32–35]. However, the performance and $R^2$-values did not completely correspond to each other. For example, the cases of three and four motor synergies showed almost double the difference in the reaching errors, whereas there were no such remarkable changes in $R^2$-values, especially in the sagittal targets (figures 3d and 4e). A possible reason is the gap between signal and performance reconstructions as pointed out in [36].

Each repertory of the extracted motor synergies is considered to be specialized for the corresponding targets. Regarding the horizontal targets, the number of motor synergies from the horizontal policy remarkably affected the performance, whereas those from the sagittal policy hardly changed the performance (figure 4b). Considering the sagittal targets, the motor synergies from the sagittal policy affected the performance, whereas those from the horizontal policy did not (figure 4e). Moreover, neither was the repertory able to improve the performance in the frontal and upper targets by itself (figures 5 and 6). It is notable that the motor synergy repertoires from the horizontal targets themselves were not able to reach the upper targets, where only the height was different. Such a difference in the contribution of the motor synergies between the targets suggests that motor synergies are related to specific motor skills.

Focusing on each of the extracted motor synergies, we can observe that they encode task information efficiently. The waveforms of the motor synergies shown in figure 3b,e show the trend changes in $t_{\text{target}} = 0.5\,\text{s}$. For example, in figure 3b, the joint torque four in synergy zero changes from negative to positive values at $t_{\text{target}}$. Furthermore, considering figure 3c, joint torque two in synergy three rapidly decreases at $t_{\text{target}}$. These trend changes show that the movement directions change at $t_{\text{target}}$. Therefore, the motor synergies efficiently encoded the temporal information about the task, which were the changes in the reaching direction over time. This is in contrast to the other motor representations proposed in the robotics field, such as dynamic movement primitives (DMP) [19,20], which do not depend on specific tasks. Because the motor synergies encode specific skills of experts (e.g. humans and trained policies), they (the motor synergies) are considered suitable for handling a combination of multiple skills.

The motor synergies themselves and their activation patterns also seem to encode the task information. The amplitude of the motor synergy activities change in the cosine-like form for the angles of $q_{\text{target}}$, according to figure 3c,f. Each motor synergy strongly activates at a particular target direction, and movements toward the intermediate directions between them can be represented by combinations of these motor synergies. Previous studies report similar activation patterns in human reaching movements [3,25,37] although the types of motor synergies are different. There are regular spatial patterns in motor synergy activities, which can improve the generalization ability of motor command generators by explicitly incorporating it into the design.

### 4.3. Generalization by combination of motor synergy repertoires

It is demonstrated that the combination of multiple motor synergy repertoires can be generalized to new targets, even in the new planes. Considering figures 5 and 6, the combination of four to five motor synergies from each of the horizontal and sagittal policies results in comparable performances to those of reinforcement learning (table 3). In addition, the trajectories are approximately straight lines, similar to those obtained by reinforcement learning (figures 2, 5 and 6). This suggests that the skills of the original policies, such as gravity compensation, are also reproduced. The same is suggested by the fact that a straight trajectory started to appear early in the optimization process (figures 5d and 6d), contrary to reinforcement learning, where the complex curved trajectories are generated at the beginning of motor learning (figure 2c,g).

The performance was remarkably improved only when both repertoires of the motor synergies from the horizontal/sagittal policies are combined (figures 5 and 6). Considering the frontal targets, when only single repertoires were used (i.e. $L_h = 0$ or $L_s = 0$) the minimum error was 6.49 cm at $L_h = 5$. The trajectories do not seem to have reached the targets successfully in most cases (figure 5c,f). Considering the case where $L_h = L_s = 3$, the error was reduced to 4.24 cm although the number of motor synergies were approximately the same as in the above case. The error was further reduced by adding one motor synergy from each of the two repertoires, whereas it hardly changed when two motor synergies were added from single repertoires. A similar tendency was also observed in the
upper targets. Regarding these results, the essential factor for generalization is not to simply increase the motor synergies; nonetheless, it is to combine multiple repertoires of the motor synergies.

Motor synergies, however, were not always effective for different arm models. The motor synergies extracted from the longer-arm model were only able to show lower performance than that of the original arm model (figures 5a and 7d). This would be because the motor synergies depend on the inverse kinematics of the arm, which varies for different arm models. By contrast, the motor synergies from the shorter-arm model resulted in a similar performance to the original arm model. Thus, motor synergies may be transferable for different bodies in some cases and enable the new bodies to perform the same task. Whether a repertory of motor synergies is transferable for another body is not simple and remains as future work. Besides, in both cases, we can yet observe a similar tendency to the previous subsection that the performance was only improved when both repertoires of the motor synergies were combined.

Considering the combination of multiple motor synergy repertoires, linearity plays an important role. Linearity makes it possible to extend the current repertory of motor synergies by adding other motor synergies. This study directly demonstrates this using the optimization of the frontal and upper targets (figures 5c and 6c). The motor synergy repertoires extracted from the horizontal and sagittal policies enable x–y and x–z directional reaching, respectively. Each repertory itself cannot realize y–z directional reaching movements; however, combining them, a new skill that can perform y–z directional reaching can be realized. Whereas there exists a study that addresses the extension of motor synergy repertoires [38], our study extends the repertoires with motor synergies that are extracted from individually learned movements.

These findings can also contribute to machine learning and robotics. Our results suggest the possibility that a robot can construct new movements by re-using previously acquired skills through motor synergies. Even if motor synergies may not be able to work in other robots that have totally different kinematics, that fact can be beneficial toward the scalability of intelligent robots. Therefore, it will improve the efficiency of motor learning, although there is a labour of motor synergy extraction.

4.4. Comparison with the findings in neuroscience

Although it is not our goal to describe the actual adaptation process in humans, it is possible to compare our results with past findings in neuroscience. Previous studies have suggested that humans employ the superposition of motor synergies to generate new movements. Ivanenko et al. and Clark et al. reported that approximately five motor synergies were able to represent human locomotion [39,40]. Moreover, Ivanenko et al. reported that additional motor synergies make it possible to represent other types of movements such as straddling-obstacle movements [41]. Similar findings are reported in multidirectional reaching tasks [4]. Although our study did not consider the timing of motor synergy activation, as in the above studies, we computationally reproduced a similar phenomenon in which new movements were generated by extending motor synergy repertoires.

Our findings are also closely related to the shared synergies and motor synergies commonly observed in movements for different tasks [6,9]. Scano et al. showed that similar motor synergy structures were observed between reaching tasks with various target positions [9]. Isaely et al. demonstrated that a set of motor synergies in a reaching task can account for other directional movements [8]. Those studies have shown the existence of the shared synergies; by contrast, our study demonstrates a process in which motor synergies are re-used for new targets to function as shared synergies. In this study, the motor synergies are not intended to be shared with different types of targets. This is in contrast to other studies that directly designed shared-synergy structures [21,42].

However, there are differences between our method and the human adaptation process. One difference is that we only treated joint-level synergies, whereas human studies usually consider muscle-level synergies. As the numbers of muscles are usually more than those of joints, higher dimensionality needs to be treated than that of ours in muscle-level synergies. Muscle-level synergies also consider redundancy of antagonistic muscles, which do not appear in joint-level synergies. On the other hand, the tasks in this study also have redundancy and nonlinearity due to the inverse kinematics. Moreover, previous studies pointed out that joint-level synergies are highly correlated with muscle-level synergies [43,44]. These suggest that our results can also be applied to muscle-level synergies of more complex musculoskeletal systems.

Another difference is that in humans, the motor synergies themselves can change while learning new tasks. Motor synergy structures have been reported to gradually change during development or motor skill learning [45–47]. It has also been reported that the number of motor synergies increases in
experts who have experienced long-term motor training in specific tasks [48]. By contrast, our method does not have a function that makes it possible to modify the motor synergies after extraction. Although the modification of motor synergies is slow and small [11,49], fine-tuning the motor synergy structure may improve the final performance in more complex tasks. In addition, we used spatiotemporal motor synergies, in which motor activities were feed-forwardly determined, and sensory feedback played an important role in the motor pattern generation in the central nervous system [14,50–53]. Although this study addresses a situation where feed-forward motor commands are sufficient, the extension to sensorimotor feedback remains as future study.

It should be noted that the reaching task in this study is complex even compared with related works that used musculoskeletal systems because we treated a three-dimensional task. Although recent studies have dealt with simulated musculoskeletal systems with many muscles [22,54,55], most studies only handled planar physics, in which the rotational directions of joints are always constant (i.e. perpendicular to the plane). By contrast, in three-dimensional reaching, the rotational axes of joints vary depending on other joint angles; for example, the rotation direction of the elbow joint changes depending on shoulder-joint angles. Owing to such difficulty, there are still only a few studies that dealt with three-dimensional tasks with motor synergies [56].

5. Conclusion

This study reveals that a linear combination of motor synergies extracted from multiple policies can be generalized to new targets that cannot be achieved with either the original policies or single motor synergy repertories. We have demonstrated a new framework that consists of (i) motor learning to acquire individual skills, (ii) motor synergy extraction from the skills, and (iii) generalization to new targets by combining the motor synergies and optimizing their activities. We also found that the motor synergies efficiently encoded the movements achieving the tasks, and the representation ability of the motor synergies could be easily extended by superposition owing to the linearity. New targets in new planes could be realized with errors within 1 cm, even for the new plane that was perpendicular to the learned planes. These computational results suggest the advantages of motor synergies for the generalization of motor skills.

We believe this is the first study that has succeeded in motor synergy generalization for new targets in multi-directional reaching with a full seven-d.f. arm under gravity, a realistic mechanical environment for the motor task. It was previously difficult to be verified with the motion analysis performed by human subjects who had already obtained motor synergies for various motor tasks. Moreover, our results provide important contributions to machine learning and robotics field. Our finding highlights that the generalization ability of motor synergies supports the employment of motor synergy representations in the computational learning process.

Data accessibility. The codes to run simulations are available at https://doi.org/10.5281/zenodo.6395483.

Authors’ contributions. K.K.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, validation, visualization, writing—original draft, writing—review and editing; M.H.: conceptualization, funding acquisition, project administration, supervision, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein. Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was supported by JST, ACT-X grant no. JPMJAX20A8, Japan; and the JSPS Grant-in-Aid for Scientific Research on Innovative Areas (20H05458) Hyper-Adaptability project.

Acknowledgements. We would like to thank Editage (http://www.editage.com) for editing and reviewing this manuscript for English language.

References

1. Lee WA. 1984 Neuromotor synergies as a basis for coordinated intentional action. J. Mot. Behav. 16, 135–170. (doi:10.1080/00222895.1984.10735316)
2. Santello M, Flanders M, Soechting JF. 1998 Postural hand synergies for tool use. J. Neurosci. 18, 10 105–10 115. (doi:10.1523/JNEUROSCI.18-23-10105.1998)
3. D’Avella A, Portone A, Fernandez L, Lacquaniti F. 2006 Control of fast-reaching movements by muscle synergy combinations. J. Neurosci. 26, 7791–7810. (doi:10.1523/JNEUROSCI.0830-06. 2006)
4. D’Avella A, Lacquaniti F. 2013 Control of reaching movements by muscle synergy combinations. Front. Comput. Neurosci. 7, 42. (doi:10.3389/fncom.2013.00042)
5. Takei T, Confais J, Tomatsu S, Oya T, Seki K. 2017 Neural basis for hand muscle synergies in the primate spinal cord. Proc. Natl Acad. Sci. USA 114, 8643–8648. (doi:10.1073/pnas.1704381114)
6. D'Avelia A, Bizzi E. 2005 Shared and specific muscle synergies in natural motor behaviors. Proc. Natl Acad. Sci. USA 102, 3076–3081. (doi:10.1073/pnas.0500199102)

7. Chvatal SA, Ting LH. 2013 Common muscle synergies for balance and walking. Front. Comput. Neurosci. 7, 48. (doi:10.3389/fncom.2013.00048)

8. Israely S, Leisman G, Machtal C, Shnitzer T, Carmeli E. 2017 Direction modulation of muscle synergies in a hand-reaching task. IEEE Trans. Neural Syst. Rehabil. Eng. 25, 2427–2440. (doi:10.1109/TNSRE.2017.2769659)

9. Scano A, Dardan L, Moltet F, Giberti H, Tosatti LM, D’Avella A. 2019 A comprehensive mapping of muscle synergies in highly variable upper-limb movements of healthy subjects. Front. Physiol. 10, 125. (doi:10.3389/fphys.2019.01231)

10. Kargo WL, Nitz DA. 2003 Early skill learning is expressed through selection and tuning of cortically represented muscle synergies. J. Neurosci. 23, 11255–11269. (doi:10.1523/NEUR023-13-11255.2003)

11. Berger DJ, Gentner R, Edmunds T, Pai DK, Avella A. 2019 A comprehensive spatial mapping of muscle synergies in natural motor behaviors. In Motor Control: From Cope to Coop. Stockholm, Sweden 16–21 May, pp. 1776–1781. IEEE. (doi:10.1109/ICRA.2016.7487322)

12. Gunia T, Vieira PM, Costa K, Santos CP. 2016 Looking for motor synergies in Darwin-OP biped robot. In 2016 IEEE Int. Conf. on Robotics and Automation, Stockholm, Sweden, 16–21 May, pp. 776–781. IEEE. (doi:10.1109/ICRA.2016.7487322)

13. Hagio S, Kozakai M. 2018 Modularity speeds up motor learning by overcoming mechanical bias in musculoskeletal geometry. J. R. Soc. Interface 15, 20180249. (doi:10.1098/rsif.2018.0249)

14. Todorov E, Erez T, Tassa Y. 2012 MuJoCo: a physics engine for model-based control. In 2012 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Vilamoura-Algarve, Portugal, 7–12 October, pp. 9026–9033. IEEE. (doi:10.1109/IROS.2012.6386109)

15. Schaal S, Portone A, Laquaquini F. 2011 Superposition and modulation of muscle synergies for reaching in response to a change in target location. J. Neurophysiol. 106, 2796–2812. (doi:10.1152/jn.00675.2010)

16. Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O. 2017 Proximal policy optimization algorithms. (https://arxiv.org/abs/1707.06347).

17. Schulman J, Moritz P, Levine S, Jordan MI, Abbeel P. 2016 High-dimensional continuous control using generalized advantage estimation. In 4th Int. Conf. on Learning Representations, San Juan, Puerto Rico, 24–26 May. See https://dblp.org/rec/journals/corr/SchulmanMILAJ15.html

18. Chihab M, Jacobs RA. 2006 Properties of synergies arising from a theory of optimal motor behavior. Neural Comput. 18, 2320–2342. (doi:10.1162/neco.2006.18.10.2320)

19. Geotter FD, Jonkers I, Duyens J. 2014 Task constraints and minimization of muscle effort result in a small number of muscle synergies during gait. Front. Comput. Neurosci. 8, 697. (doi:10.3389/fncom.2014.00115)

20. Chai J, Hayashi M. 2020 Motor synergy development in high-performing deep reinforcement learning algorithms. IEEE Robot. Autom. Lett. 5, 1271–1278. (doi:10.1109/LRA.2020.2968067)

21. Hansen N, Ostermeier A. 2001 Completely Derandomized Self-Adaptation in Evolution Strategies. Evol. Comput. 9, 159–195. (doi:10.1162/10636560175019398)

22. Hansen N, Auger A. 2014 Principled design of continuous stochastic search: from theory to practice. In Theory and principled methods for the design of metaheuristics (eds Y Borenstein, A Moraglio), pp. 145–180. Berlin, Germany: Springer. (doi:10.1007/978-3-642-33093-3_4)

23. Saad S. 2006 Dynamic movement primitives: a framework for motor control in humans and humanoid robots. In Adaptive motion of animals and machines, pp. 261–280. Tokyo, Japan: Springer.

24. Rickert E, D’Avella A. 2013 Learned parameterized dynamic movement primitives with shared synergies for controlling robotic and musculoskeletal systems. Front. Comput. Neurosci. 7, 138. (doi:10.3389/fncom.2013.00138)

25. D’Avella A, Portone A, Laquaquini F. 2011 Superposition and modulation of muscle synergies for the solution of reaching and speed. J. Neurophysiol. 100, 1453–1454. (doi:10.1152/jn.01377.2007)

26. Alessandro C, Carbajal JP, D’Avella A. 2012Synthesis and adaptation of effective motor synergies for the solution of reaching tasks. In From animals to animats, vol. 12 (eds T Ziemke, C Balkenius, J Hallam), pp. 33–43. Berlin, Germany: Springer. (doi:10.1007/978-3-642-33093-3_4)

27. Abbéel P. 2015 Five basic muscle activation patterns account for muscle activity during human locomotion. J. Physiol. 556, 267–282. (doi:10.1113/jphysiol.2003.057174)

28. Clark DJ, Ting LH, Zajac FE, Neptune RR, Kautz SA. 2020 Merging of healthy motor modules predicts reduced locomotor performance and muscle coordination complexity post-stroke. J. Neurophysiol. 103, 844–857. (doi:10.1152/jn.00675.2020.00625.2019)

29. Ruffini Y, Pappele RE, Laquaquini F. 2006 Muscle control as a theory of motor coordination. J. Neurophysiol. 95, 844–857. (doi:10.1152/jn.00675.2020.00625.2019)

30. Neumann G, Daniel C, Paraschos A, Kupcsik A, Peters J. 2014 Learning modular policies for robotics. Front. Comput. Neurosci. 8, 55. (doi:10.3389/fncom.2014.00062)

31. Tagliaiublu M, Ciancio AL, Brochi T, Eskelinen S, Maier MA. 2015 Differences between kinematic synergies and muscle synergies during digit grasping. Front. Hum. Neurosci. 9, 1–17. (doi:10.3389/fnhum.2015.00165)

32. Esmaeili S, Karami H, Baniasad M, Shojaeefard M, Farahmand F. 2022 The association between muscle force and joint invariance. J. Neurophysiol. 121, 339–348. (doi:10.1177/107385842287987)

33. Gentner R, Gorges S, Weise D, Kupcsik A, Buttmann M, Classen J. 2010 Encoding of motor skill in the corticomuscular system of musicians. Curr. Biol. 20, 1869–1874. (doi:10.1016/j.cub.2010.09.043)
46. Dominici N et al. 2011 Locomotor primitives in newborn babies and their development. Science 334, 997–999. (doi:10.1126/science.1210617)

47. Cheung VCK, Cheung BMF, Zhang JH, Chan ZYS, Ha SCW, Chen CY, Cheung RTH. 2020 Plasticity of muscle synergies through fractionation and merging during development and training of human runners. Nat. Commun. 11, 4356. (doi:10.1038/s41467-020-18210-4)

48. Sawers A, Allen JL, Ting LH. 2015 Long-term training modifies the modular structure and organization of walking balance control. J. Neurophysiol. 114, 3359–3373. (doi:10.1152/jn.00758.2015)

49. Yang Q, Logan D, Giszter SF. 2019 Motor primitives are determined in early development and are then robustly conserved into adulthood. Proc. Natl Acad. Sci. USA 116, 12 025–12 034. (doi:10.1073/pnas.1821455116)

50. Latack ML. 2000 The organization of quick corrections within a two-joint synergy in conditions of unexpected blocking and release of a fast movement. Clin. Neurophysiol. 111, 975–987. (doi:10.1016/S1388-2457(00)00263-7)

51. Chvatal SA, Ting LH. 2012 Voluntary and reactive recruitment of locomotor muscle synergies during perturbed walking. J. Neurosci. 32, 12 237–12 250. (doi:10.1523/JNEUROSCI.6344-11.2012)

52. Safavynia SA, Ting LH. 2013 Long-latency muscle activity reflects continuous, delayed sensorimotor feedback of task-level and not joint-level error. J. Neurophysiol. 110, 1278–1290. (doi:10.1152/jn.00609.2012)

53. Nazifi MM, Yoon HU, Beschomer K, Har P. 2017 Shared and task-specific muscle synergies during normal walking and slipping. Front. Hum. Neurosci. 11, 40. (doi:10.3389/fncom.2017.00040)

54. Chen J, Qiao H. 2020 Muscle-synergies-based neuromuscular control for motion learning and generalization of a musculoskeletal system. IEEE Trans. Syst. Man Cybern. Syst. 51, 3993–4006. (doi:10.1109/TSMC.2020.2966818)

55. Kambara H, Takagi A, Shimizu H, Kawase T, Yoshimura N, Schweighofer N, Koike Y. 2021 Computational reproductions of external force field adaption without assuming desired trajectories. Neural Netw. 139, 179–198. (doi:10.1016/j.neunet.2021.01.030)

56. Chen J, Zhong S, Kang E, Qiao H. 2019 Realizing human-like manipulation with a musculoskeletal system and biologically inspired control scheme. Neurocomputing 339, 116–129. (doi:10.1016/j.neucom.2018.12.069)