# Context-Aware Learning from Demonstration: Using Camera Data to Support the Synergistic Control of a Multi-Joint Prosthetic Arm

Gautham Vasan and Patrick M. Pilarski

***Abstract*— Muscle synergies in humans are context-dependent—they are based on the integration of vision, sensorimotor information and proprioception. In particular, visual information plays a significant role in the execution of goal-directed grasping movements. Based on a desired motor task, a limb is directed to the correct spatial location and the posture of the hand reflects the size, shape and orientation of the grasped object. Such contextual synergies are largely absent from modern prosthetic robots. In this work, we therefore introduce a new algorithmic contribution to support the context-aware, synergistic control of multiple degrees-of-freedom of an upper-limb prosthesis. In our previous work, we showcased an actor-critic reinforcement learning method that allowed someone with an amputation to use their non-amputated arm to teach their prosthetic arm how to move through a range of coordinated motions and grasp patterns. We here extend this approach to include visual information that could potentially help achieve context-dependent movement. To study the integration of visual context into coordinated grasping, we recorded computer vision information, myoelectic signals, inertial measurements, and positional information during a subject's training a robotic arm. Our approach was evaluated via prediction learning, wherein our algorithm was tasked with accurately distinguishing between three different muscle synergies involving similar myoelectric signals based on visual context from a robot-mounted camera. These preliminary results suggest that even simple visual data can help a learning system disentangle synergies that would be indistinguishable based solely on motor and myoelectric signals recorded from the human user and their robotic arm. We therefore suggest that integrating learned, vision-contingent predictions about movement synergies into a prosthetic control system could potentially allow systems to better adapt to diverse situations of daily-life prosthesis use.**

## I. INTRODUCTION

Humans have an astounding ability to learn a variety of motor skills, ranging from tying shoelaces, threading a needle to shooting a basketball. A wide number of interacting elements are involved in learning such skills. In order to be proficient in multiple motor skills, humans learn to gather task-relevant sensory information efficiently and use it in decision making and the selection of control strategies.

When any muscle or muscle group is activated, the resulting movement is dependent on the context—the relationship between muscle excitation and movement is variable and this variability is known to be context conditioned [1]. It is also well known that vision provides critical sensory information in the execution of goal-directed reaching and grasping movements. The dorsal stream in our visual cortex is directly involved in estimating position, shape and orientation of target objects for reaching and grasping purposes. It provides us with the ability to interact with our environment in a quick, reliable fashion [2].

By contrast, the *myoelectric control* of artificial limbs (specifically upper-limb prostheses) relies predominantly on the activity of muscle tissue to dictate the movement of robotic actuators within the prosthetic chassis [3]. Motor neurons transmit electrical signals that cause muscles to contract. These signals, also known as electromyographic (EMG) signals, contain information about the neural signals sent from the spinal cord to control the muscles. Despite decades of research and development, myoelectric control of upper-limb prostheses still has not reached its full potential, as a large proportion of amputees stop using myoelectric prostheses due to non-intuitive control, lack of sufficient feedback, and insufficient functionality [4]. A fundamental remaining issue is that there is a significant mismatch between the number of functions available in modern powered prosthetic arms and the number of functions an amputee can control at any given moment [5].

Apart from EMG signals traditionally used for the myoelectric control, there are a number of other physiological and non-physiological signals that could be used to assist in the control of multiple prosthetic degrees-of-freedom (DOF). To use additional signals effectively involves building a processing unit which performs sensory and motor information processing to effect a more rich behavioral repertoire. Using additional sensors is important, as there are typically a limited number of viable sites on an amputated arm for recording EMG signals. This limits the possible space of movements to discrete, gross movements. We hypothesize that a prosthesis controller can emulate highly synergistic movements traditionally considered the responsibility of the user if equipped with non-physiological sensors (e.g., vision and tactile sensors). By combining sensory information from different sources, we expect a controller should be able to detect and analyze the current context, plan the movement strategy, and simultaneously and proportionally control multiple DOF available in the prosthesis.

In our previous work [6], [7], we outlined a *learning from demonstration* (LfD) approach that could potentially allow an amputee to use their non-amputated arm to teach their prosthetic arm how to move in a natural and coordinated

This research was undertaken, in part, thanks to funding from the Canada Research Chairs program, the Canada Foundation for Innovation, and was supported by the Alberta Machine Intelligence Institute (Amii), Alberta Innovates – Technology Futures (AITF), the Government of Alberta, and the Natural Sciences and Engineering Research Council (NSERC).

G. Vasan is with Kindred Inc. P. M. Pilarski is with the Department of Computing Science and the Department of Medicine, University of Alberta, Edmonton, AB T6G 2E1, Canada. G Vasan contributed during his M.Sc. at the University of Alberta. Please direct correspondence to: vasan@ualberta.ca; pilarski@ualberta.ca

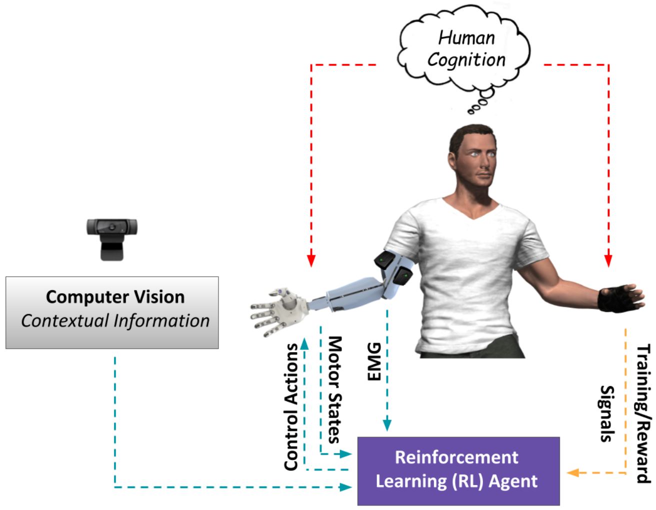


Fig. 1: Schematic showing the flow of information during the training period for a vision-enhanced LfD approach.

fashion. This could also be considered as a different type of pattern recognition method that could include most if not all aspects of traditional pattern recognition methods. But our prior approach still has an upper bound on the number of synergistic movements that could be learned by the reinforcement learning (RL) agent. For example, in the case of a transhumeral amputation (someone missing their hand, forearm, and elbow), generated EMG signals are not separable enough to identify various grasping patterns. While the LfD approach described in Vasan and Pilarski [6] learns to match a particular synergistic movement to the user's muscle signals, it cannot learn to discriminate between context-dependent grasping patterns that are characterized by highly similar EMG signals. In order to achieve situation-dependent movement based on muscle excitation, we suggest the control system should also be given relevant contextual information and meta-data about the user, the robotic limb and its environment.

In the present work, we explore how additional sensory information can be exploited to improve synergistic control of a multi-DOF prosthetic arm. More specifically, we equip the prosthetic arm with artificial vision to perceive the state of the user, prosthesis and the environment. Using this additional information, we expect the controller should be able to learn to distinguish between different grasp patterns according to the context and user's intentions (communicated using EMG signals as shown in Fig. 1). We describe an online approach based on reinforcement learning that could potentially achieve context-dependent motion if presented with contextually relevant learned or hand-designed visual representations. When integrated with our prior work, we believe such an approach can learn to use feedback from the user or the control environment to continually update or adapt its control policy. This implies that an amputee could in principle teach his/her prosthesis to *unlearn* and *relearn* unique reaching and grasping patterns for different target

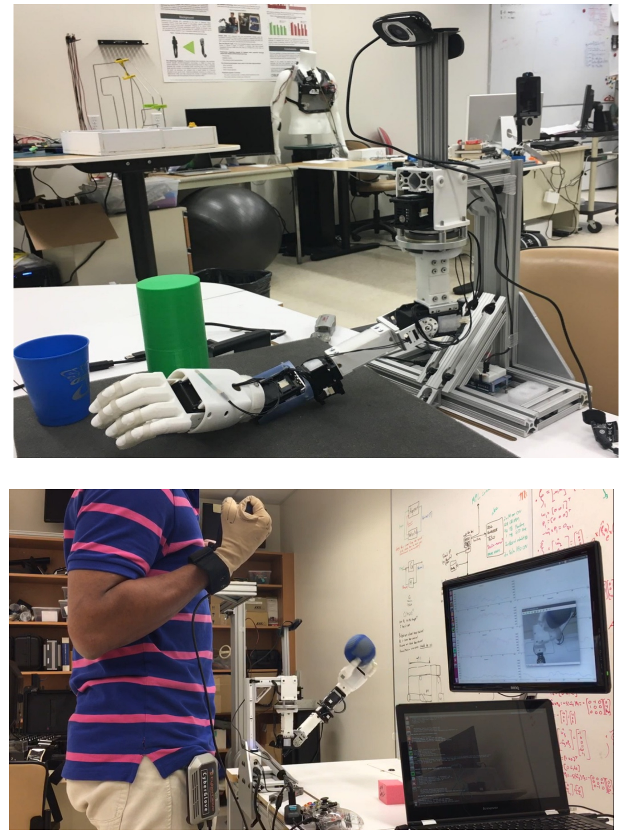

Fig. 2: Experimental setup which includes a high-definition camera, Bento Arm, Thalmic Myo Armband and CyberTouch II. The Bento Arm as used in our trials had 5 active DoFs including shoulder rotation, elbow flex/extend, wrist pronation/supination, wrist flex/extend and hand open/close.

objects or settings. In this preliminary study, for clarity we first test our hypothesis using real-time prediction learning experiments. As described by Pilarski et al. [8], such learned predictions can be integrated into a prosthetic control system in straightforward ways to enhance multi-joint prosthesis use, or act as a method by which to select the information that should be presented to a more complex control learning algorithm.

## II. Related Work

Markovic et al. [9] introduced a sensor fusion and computer vision based control approach for the context-aware control of a multi-DOF prosthesis. Their work uses a combination of sensing units, comprising myoelectric recording, computer vision, inertial measurements and embedded prosthesis sensors (position and force), to develop a controller that could allow a multi-DOF prosthesis to perform simultaneous, coordinated movements. The method relies on sensor fusion which allows for the perception of the user (proprioception), the environment (exteroception) and their interaction, leading to simultaneous, proportional control of multiple DOFs through context-dependent behavior. Ghazaei et al. also introduce a classification-based approach to grasping that leverages the representational power of deep networks [10].

## III. METHODS

Similar to the experiments described by Vasan and Pilarski [6], we assume a setting in which our target users have undergone a unilateral, transhumeral amputation. To provide a proof of concept in this setting, we first study the case where a non-amputee participant has one biological hand (right hand), and one robotic arm (left hand) that they wish to train to appropriately respond to the commands being generated by the muscle tissue in the user's upper left arm (simulating the residual limb). Compared to our prior work [6], there are two significant changes in our experimental setup—first, we mount a Logitech HD 1080p webcam on top of the robotic arm (see Fig. 2 (top)) and second, we use a Thalmic Myo armband instead of the Delsys Trigno Wireless system to obtain EMG signals.

### *A. Hardware*

*Bento Arm:* The Bento Arm is a myoelectric training tool to assess and train upper-limb amputees in how to use their muscle signals prior to being fit with myoelectric prostheses [11]. We use a 5-DOF Bento Arm for the purposes of our experiments. It is a 3D printed prototype with Dynamixel MX Series servos.

*EMG Data Acquisition:* We used a 8-Channel Thalmic Myo armband to record EMG signals from our subjects. The Myo armband is fashioned as a wearable gesture control and motion control device that can be worn around the forearm/upper arm. It also provides inertial measurements which can be used to calculate rotation and translation with respect to a fixed frame of reference.

*Motion Capture Glove:* The desired joint angle configurations for wrist flexion/extension $(\theta^*_{wf})$ and hand open/close $(\theta^*_h)$ were defined by the subject using a CyberTouch II system (CyberGlove Systems LLC) worn on the hand of their training arm.

*Computer Vision:* We used a Logitech C920 HD Pro Webcam mounted on top of the Bento Arm. The camera captured images at 50 frames-per-seconds. Examples of the images obtained from the camera are shown in Fig. 3.

The experimental task was divided for the participant into multiple segments as follows: grasp the object on the table, bring it up by flexing the elbow (as if the object is being examined closely), extend the arm and bring it down and finally drop it on the table. While the EMG control commands remained the same for flexing and extending the elbow, the corresponding wrist and hand trajectories for manipulating each object was different. The hand, wrist flexor and wrist rotator joints (denoted by $\theta_h, \theta_f$ and $\theta_r$ respectively) were correlated with the angular position of the elbow joint (denoted by $\theta_e$) such that there exists a policy that could map any given elbow position uniquely into a combination of higher dimensional joint movements. The goal for the machine learning approach was to accurately differentiate and predict the three different muscle synergies shown by the user for grasping a large blue ball, a red sponge, and a smaller yellow smiley ball as shown in Fig. 3.

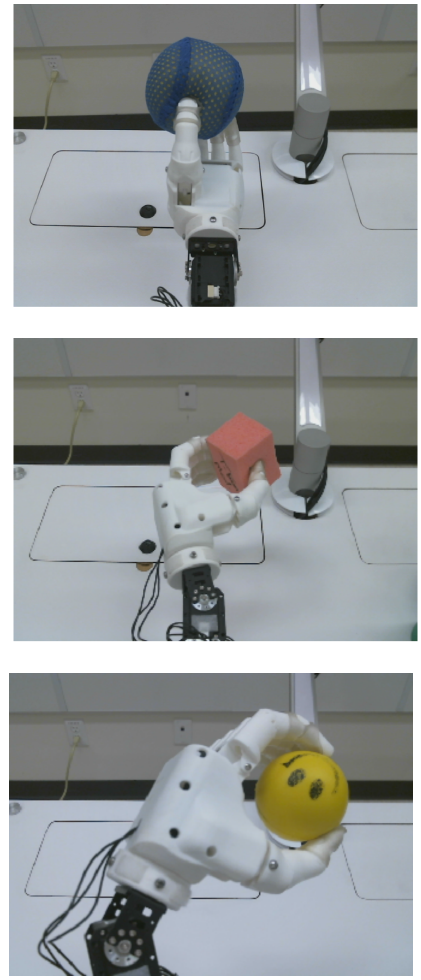

Fig. 3: Three different objects for our experiment—a large blue ball, a red sponge and a small yellow smiley ball—with different grasp patterns. Raw input images were processed and provided to the learning system as input features.

### *Phase I: Recording training data*

In this phase, the able-bodied participant was instructed to execute the task as a repetitive sequence of simple reaching and grasping movements that were mirrored by both their control and training arms (for an amputee, this would correspond to trying to perform identical movements using their non-amputated arm and the prosthetic arm). The training arm demonstrated the desired movement and grasp pattern to the prosthetic arm. During training, the elbow of the Bento Arm was actuated via proportional myoelectric control from the subject's control arm, while the wrist and hand of the Bento Arm were actuated via direct teleoperation—i.e., the Bento Arm copied the training arm's movements as reflected to the contralateral side. As shown in Fig. 1, Fig. 2 (bottom) and described above, we recorded computer vision information, EMG signals from the Myo Armband and desired angles from the subject's training arm (wrist and finger joints) using the motion capture glove and inertial measurement system. The participant demonstrated each reaching and grasping movement for $\sim 15min$. The entire trial lasted $\sim 45min$.

This task was similar to the one described in Vasan and Pilarski [6], in which we had our participants demonstrate a single, desired movement and grasp pattern in 3 DOF to an actor-critic control learner. In this experiment, our participant instead demonstrated the wrist and grasp patterns for manipulating three different objects.

### *B. Learning Predictions with Temporal-Difference Methods*

Temporal difference (TD) learning uses changes or differences in predictions over successive time steps to drive the learning process. It learns how to predict a quantity that depends on future values of a given signal [12]. TD algorithms are most commonly used in reinforcement learning to predict the expected return of a reward signal.

More recently, Sutton et al. described a generalized prediction learning approach based on temporal-difference learning [13]. They were successful in learning temporally extended predictions about non-reward sensorimotor data. It has also been demonstrated that learned temporally extended predictions can accurately forecast signals during prosthesis control by both amputees and able-bodied subjects [8], [14]–[16]. Predictions were phrased as a linear combination of a learned weight vector, here denoted $w_p$, and a state approximation $\phi_p$ of current and next states $s_p$ and $s'_p$ respectively. Predictions P for a given signal $r_p$ were then computed using $P_p = {w_p}^T \phi$, where $w_p$ was updated on each time step according to the TD error $\delta_p$ using the following incremental procedure, with parameters $\alpha_p$ (learning rate), $\gamma_p$ (discounting parameter) and $\lambda_p$ (decay rate of the eligibility trace $e_p$):

- $\delta_p \leftarrow r_p + \gamma_p {w_p}^T \phi_p(s'_p) - {w_p}^T \phi_p(s_p)$
- $e_p \leftarrow \gamma_p \lambda e_p + \phi_p(s_p)$
- $w_p \leftarrow w_p + \alpha_p \delta_p e_p$

## IV. Real-time Prediction learning using General Value Functions (GVFs)

In this experimental setting, the EMG control signals from the user are extremely (and deliberately) similar across the three context-dependent control tasks. Artificial vision is the only distinguishing input feature that could help a learning system differentiate between the tasks. Learning accurate predictions about the desired target trajectories for each object gives a clear measure by which to show a system's ability to ascertain this simple form of context. As such, we use standard temporal-difference learning [12] of General Value Functions (GVFs, [13]) general to allow the system to make predictions about the desired joint angles.

We created three GVFs for predicting the three signals of interest ${\theta^*}_j$ in the robotic system. We make temporally extended predictions at a short time scale (1.0s) about three target joint angles—wrist rotation, wrist flexion and hand open/close. The learning agent was presented with a signal space consisting of the following:

- Elbow joint angle and velocity $\langle \theta_e, \dot{\theta}_e \rangle$
- Object specific features: input *rgb* images were converted into *hsv* format and analyzed for how often certain colors appear and in what proportions they are to be found in different types of images. As a very simple contextual representation based on visual features not available to standard prosthetic hardware, the dominant range of 'r', 'g' and 'b' values of the corresponding objects was then used to classify the target object. E.g., based on the number of blue pixels within a particular threshold, target objects could be uniquely classified and used as a single integer signal to the learner.
- The EMG control signal, given as the difference between the mean-absolute-value of the EMG signals obtained from Myo sensors 3 and 8 (placed directly over the biceps and triceps respectively).

We used tile coding for linear function approximation [12]. Our state representation consisted of 32 incrementally offset tilings (width=1) for better generalization. Each tiling had a resolution level $N_R = 10$. The binary feature vector of length 5,000,000 was hashed down to a memory size of 8192 and we also added a bias unit which was always active. At every timestep, 4 continuous signals were provided to the tile coder and $m = 33$ features were active. The learning parameters were set as follows: $\alpha = 0.1/m, \gamma = 0.99, \lambda = 0.7$. Weight vectors $w, e$ were initialized to zero. Each GVF received its target joint angle $\theta^*$ as the cumulant $r_p$.

Performance of the learning system was measured based on its ability to predict desired joint angles. All learning algorithms were run on a Lenovo Y700 Laptop with Intel Core i7-6700HQ @2.60GHz x 8 and 8GB RAM. We used the Robot Operating System (ROS) Kinetic on Ubuntu 16.04 to send and receive information and commands from the Bento Arm, CyberTouch II and the Thalmic Myo armband. All sensorimotor information was communicated between different systems using ROS topics. We recorded all the sensorimotor information (from ROS topics) using rosbags. Rosbags avoid deserialization and reserialization of the messages. After recording, we can playback the data in a time synchronized fashion and simulate real-time sampling and learning conditions for different hyper-parameter choices.

## V. Results

As shown in the results, the system was able to successfully anticipate the joint trajectories initiated by the subject for all three different objects. Accurate predictions were observed after $5 - 6min$ of real-time sampling (simulated by playing back recorded data and synchronized using timestamps) and learning. Here the normalized predictions $\bar{P}_p$ at a time scale of 1.0s (colored lines) are compared against corresponding target joint angles (grey lines). The prediction $P_p$ is dependent on the timescale (i.e., time constant) of return predictions determined by $\gamma$ (i.e., the discount factor). For comparison with raw signals, normalized return predictions ($\bar{P}_p$) are therefore scaled according to the time constant, i.e., $\bar{P}_p = P_p * (1 - \gamma_p)$. Fig. 4 shows an example of wrist rotation angle prediction after two offline learning passes through $10mins$ of recorded training data from the subject demonstrations. The agent was tasked with learning three target joint angles in parallel—wrist flexion, wrist rotation and hand open/close. The training phase lasted for $\sim 90mins$

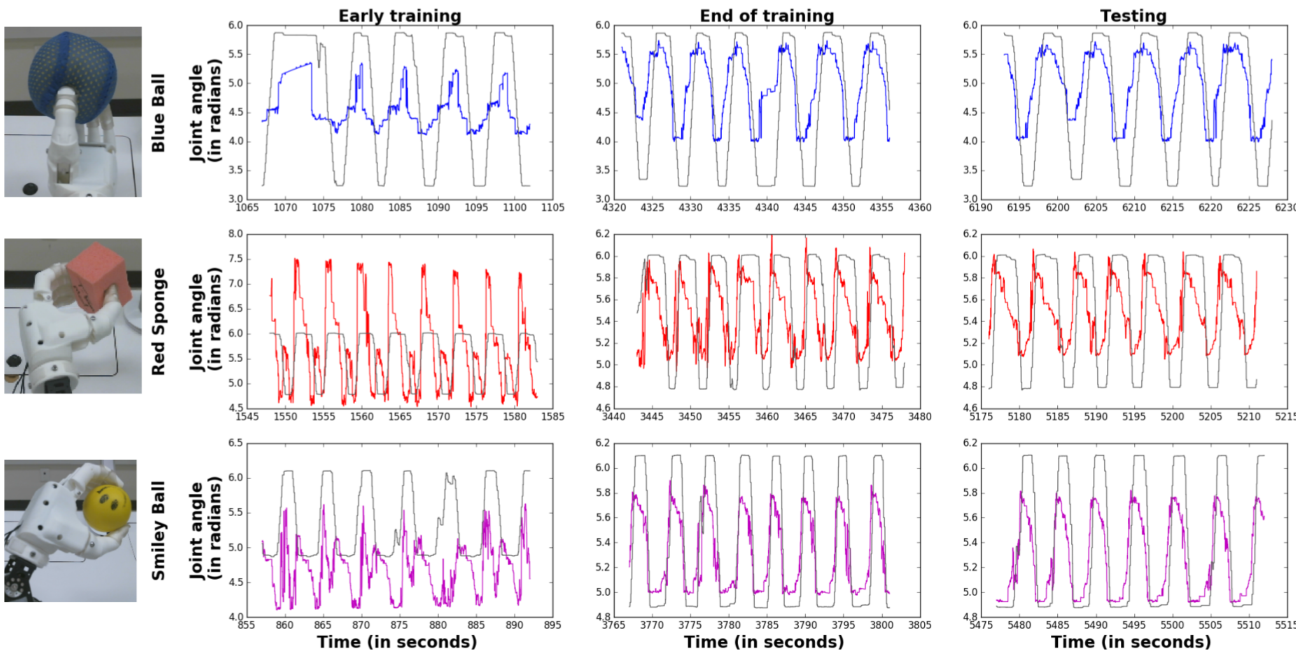


Fig. 4: Comparison of target (grey line) and predictions (colored lines) of ***wrist rotation trajectories*** over training and testing periods. This plot shows the joint angle predicted by the TD learner for the able-bodied subject during training and testing.

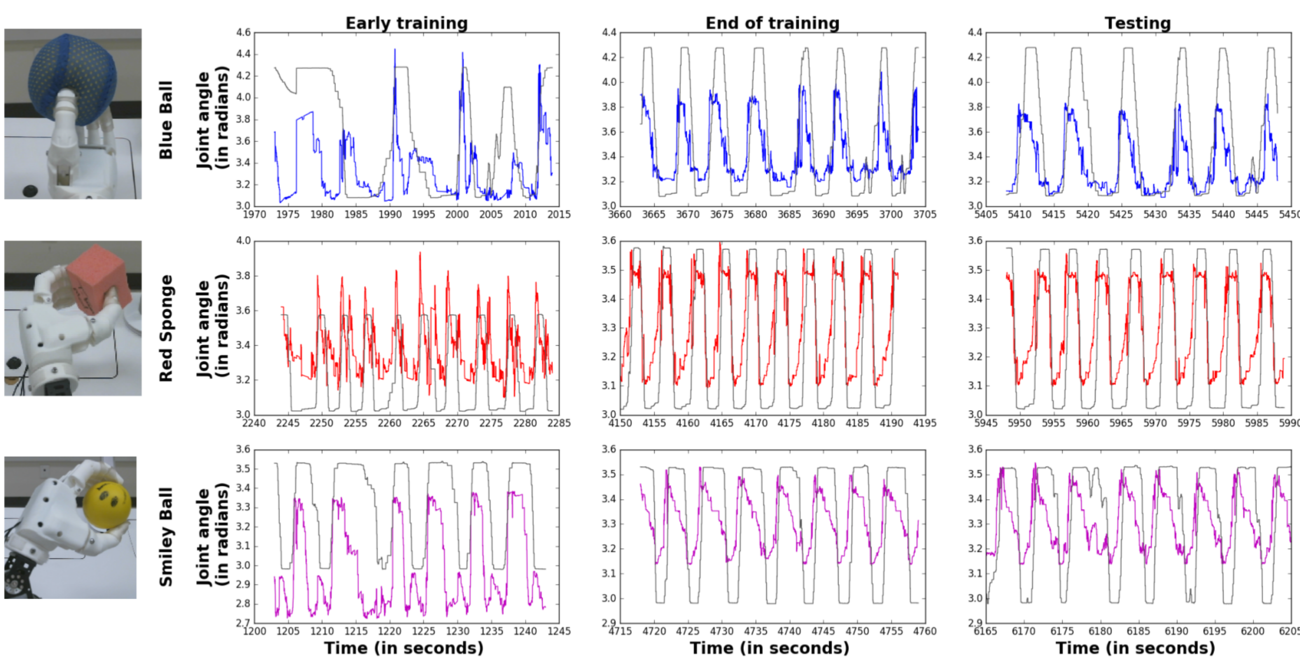


Fig. 5: Comparison of target (grey line) and predictions (colored lines) of ***wrist flexion trajectories*** over training and testing periods. This plot shows the joint angle predicted by the TD learner for the able-bodied subject during training and testing.

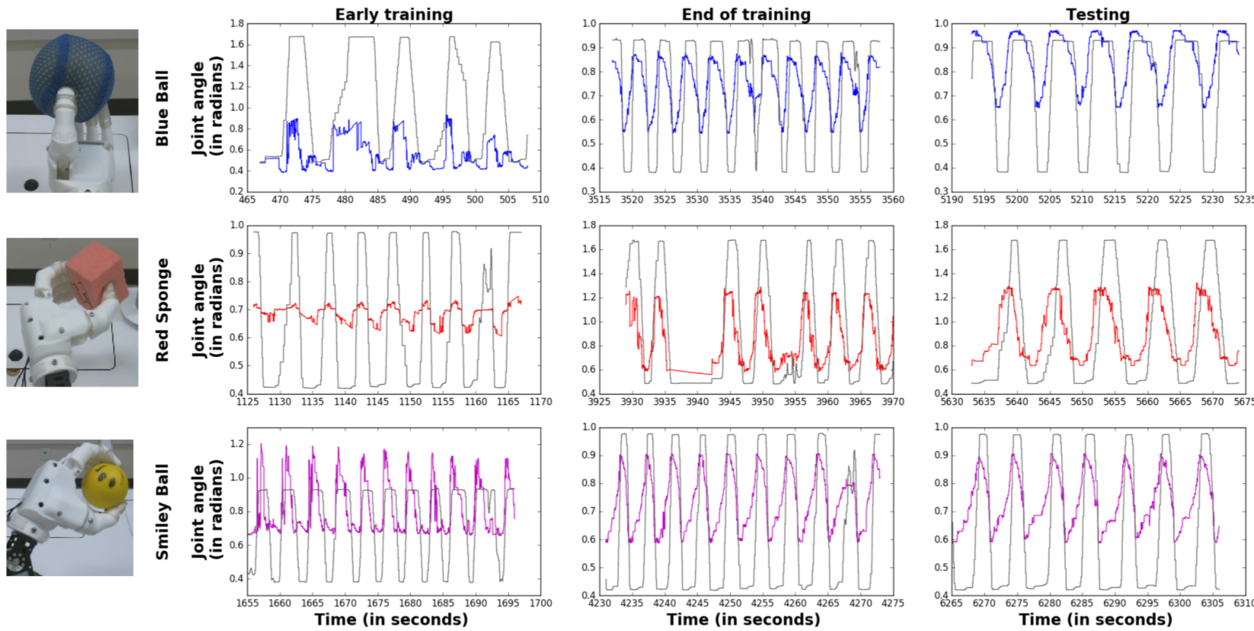


Fig. 6: Comparison of target (grey line) and predictions (colored lines) of ***gripper hand trajectories*** over training and testing periods. This plot shows the joint angle predicted by the TD learner for the able-bodied subject during training and testing.

(two offline passes through $15mins$ of recorded data for each demonstration) and the testing phase lasted for $\sim 15mins$ (one offline pass through $5mins$ of recorded data for each demonstration). Fig. 5 and 6 show examples for normalized joint angle predictions for wrist flexion and gripper hand respectively over training and testing periods.

## VI. Discussion

### A. Prediction in Adaptive Control

Highly skilled motor behavior relies on our brain learning both to control its body as well as predict the consequences of this control. Flanagan et al. studied the relation between predictions as control during motor learning [17]. They found different time scales of learning for predictions and control, with predictions being learned much faster than control. Pilarski et al. integrated learned anticipatory predictions into the control the actuators of a multi-joint prosthesis for use by amputees, especially amputees with limited signal recording sites on their amputated limbs [8]. They were able to make accurate, anticipatory predictions at different timescales about various joint angles, dynamics, and EMG signals. Their integration of real-time prediction and control learning promises to speed up control policy acquisition, allow unsupervised adaptation in myoelectric controllers, and facilitate coordinated, synergistic movements in a mutli-DOF prosthetic limbs. In this paper, we test the ability of our system to learn accurate, temporally abstracted predictions about the actuator positions of the joints controlled by the learning agent when supplemented with contextual visual information. In the future, we hope to implement a similar integration of real-time prediction learning and control to learn a larger repertoire of motor behaviors, either through these predictions being provided to a control system as state information (as per Pilarski et al. [8]), or via the identification of more complex visual features that can be provided directly to the control learning approach of Vasan and Pilarski [6].

### B. Sensor Fusion for Context-Aware Control

As discussed in Vasan and Pilarski [6], the LfD paradigm is not applicable only to prosthetic arms, but could also be extended to wearable robots such as exoskeletons, powered orthotics, supernumerary limbs, functional electrical stimulation systems, and other assistive rehabilitation technologies. While these devices face challenges in both hardware and software design, a major challenge is that the robot usually lacks the capability to adequately recognize the actions and intentions of the human user. Consequently, it cannot assist the user appropriately, a drawback that has been especially emphasized in the rehabilitation robotics domain.

In most wearable robots, many sensors are already built into the device, such as joint angle sensors, electrophysiological measurements such as electromyography (EMG) or electroencephalography (EEG), or alternatively mechanical sensors or inertial measurement units (IMUs) placed on a part of the body that is not covered by the wearable robot. We suggest that, in the same way as vision was used in the present work, we can and should combine this multi-modal information (combining different sensor types) to better learn and adapt to the needs of the user.

In the rehabilitation robotics domain, the degree of control (DOC) for a device can greatly outnumber the number of input channels the user has available. For example, in the case of an amputee user, the disparity between the DOC and the number of available input signals greatly increases as the level of amputation increases (e.g., those with transhumeral amputations can provide even fewer control signals than those with transradial amputations). Unfortunately, this mismatch makes the control of wearable robots difficult and tedious. As suggested here for the case of visual sensing during grasping, sensor fusion can potentially alleviate some of the issues associated with controlling large DOFs with a small subset of input signals.

In addition to the standard electro-physiological signals, IMUs and joint angle measurement units, it may be fruitful to add artificial vision, gaze vectors (to know where the user is looking) and tactile sensation systems (for example, a camera and capacitive touch systems respectively) to robotic devices prostheses. These systems could provide useful sensory information to the learning agent that could be used to better perceive the environment and needs of the user.

### C. Representation Learning

In this paper, we showed that it is feasible to learn to distinguish between the desired grip trajectories for three different objects by the addition of very simple vision-based features. A main contribution of this work is to highlight the utility vision-based features can have for building contextual LfD algorithms in the prosthetic setting. As would be expected, altered lighting, displacements and other factors would not be well handled by the current proof-of-concept learning representation used in the present work. While out of scope for the present comparisons, and as described below, there are numerous state-of-the art representational approaches, both learned and engineered, that could be implemented to well support prediction and control learning in prosthetic LfD.

Modern day prostheses could receive a huge density of data about the user, their physiological and psychological needs and their environment. For example, camera data or even additional sensors on the socket of a prosthesis can readily provide enough contextual information to allow an actor-critic RL system to produce varied motor synergies in response to similar EMG signals from the user—e.g., a system can use additional sensor and state information to help manage the user's degree-of-freedom problem, generating synergies that artfully align to different situations in the user's daily life. It is therefore important that efficient ways of structuring prosthetic data are developed to better represent context to a machine learning prosthetic control system without facing the curse of dimensionality. For example, representation learning methods built on Kanerva Coding could potentially be used to better handle this large number of real-world state signals [18].

While the idea of using a single state representation to better leverage the multi-modal sensory information is extremely appealing, it's been shown that different function approximators can be better at learning about different types of data. For example, convolutional neural nets (CNNs) have been widely successful in image classification and object detection datasets [19]. Similarly, motor primitives have been successful in encoding rhythmic and discrete movements. Recurrent neural networks (RNNs) have also been extremely successful in speech recognition and text translation [20]. Though with its own inherent challenges, one simple way to combine all the sensory information is to extract the features for each modality separately, then input all features into a single sensor fusion algorithm.

### D. Study Limitations and Future Developments

The experimental evaluation in this study was designed specifically to test the hypothesis that a learning system's ability to make accurate predictions about context-dependent joint trajectories in the prosthetic LfD setting will increase when supported by visual sensor information or visual scene features. Therefore, the scope of this pilot study was limited to having our participant perform a discrete set of functionally relevant movements (e.g., grasping and lifting distinctly colored, differently oriented objects) designed to approximate those facing a multi-DOF prosthesis user, while also taking into account the constraints of the setup (e.g., desk mounted Bento Arm, and vision sensor mounted on top of the arm instead of on the socket or the terminal device). We are fully aware of the importance of and state of the art in standardized evaluations of the functional effectiveness of powered prostheses, and ongoing work is exploring how our LfD approach generalizes across different tasks, objects, and settings via validated eye and motion capture metrics.

## VII. Conclusion

Our approach was able to learn contextually-accurate predictions from joint trajectories demonstrated by an able-bodied user though the integration of simple visual features into the state space of the prediction learning algorithm. These results therefore suggest important extensions to the generality of our previously published methods for LfD in the prosthetic setting, wherein a prosthetic user with only limited control channels at their disposal may still be able to effect a variety of synergistic, coordinated movements in a situation-appropriate fashion. This suggestion remains to be demonstrated in practice through precise control learning follow-up experiments, and provides a path to research into long-term control adaptation. An interesting extension to this work is to explore the use of contextual-prediction architectures along with ACRL or alternative control approaches in complex real-world activities with gold-standard outcome metrics, and to further evaluate predictions for context-aware control adaptation with a population of participants with amputations.

## Acknowledgement

We thank Craig Sherstan, Kory Mathewson, Michael Rory Dawson, Jaden Travnik, McNiel Keri and Vivek Veeriah for many helpful discussions.

## References

[1] M. T. Turvey, H. L. Fitch and B. Tuller, "The Bernstein perspective, I: The problems of degrees of freedom and context-conditioned variability," in *Human motor behaviour: An introduction,* Ed. J.A.S. Kelso. Hillsdale, NJ: Lawrence Erlbaum Associates, 1982.

[2] E. Chinellato, A.P. del Pobil, "The Neuroscience of Action and Perception." In: "The Visual Neuroscience of Robotic Grasping," *Cognitive Systems Monographs*, vol 28. Springer, Cham, 2016.

[3] C. Castellini, P. Artemiadis, M. Wininger, et al., "Proceedings of the first workshop on peripheral machine interfaces: Going beyond traditional surface electromyography," *Frontiers in Neurorobotics*, vol. 8, no. 22, Aug. 2014.

[4] B. Peerdeman, D. Boere, H. Witteveen, et al., "Myoelectric forearm prostheses: State of the art from a user-centered perspective," *J. Rehab. Res. Dev*, vol. 48, no. 6, pp 719–738, 2011.

[5] C. Sherstan, "Towards prosthetic arms as wearable intelligent robots," M.Sc Thesis, University of Alberta, Dept. Computing Science, 2015.

[6] G. Vasan and P. M. Pilarski, "Learning from demonstration: Teaching a myoelectric prosthesis with an intact limb via reinforcement learning," *Proc. IEEE Int. Conf. on Rehab. Robotics (ICORR)*, London, 2017, pp. 1457–1464.

[7] G. Vasan,"Teaching a powered prosthetic arm with an intact arm using reinforcement learning." M.Sc Thesis, University of Alberta, Dept. Computing Science, 2017.

[8] P. M. Pilarski, T. B. Dick, and R. S. Sutton, "Real-time prediction learning for the simultaneous actuation of multiple prosthetic joints," in *Proc. IEEE Int. Conf. on Rehab. Robotics (ICORR)*, Seattle, USA, June 24–26, 2013, pp. 1–8.

[9] M. Markovic, S. Dosen, D. Popovic, B. Graimann and D. Farina, "Sensor fusion and computer vision for context-aware control of a multi degree-of-freedom prosthesis," *Journal of Neural Engineering*, vol. 12, no. 6, 2015, pp. 066022.

[10] G. Ghazaei, A. Alameer, P. Degenaar, G. Morgan, and K. Nazarpour, "Deep learning-based artificial vision for grasp classification in myoelectric hands," *Journal of Neural Engineering*, vol. 14, no. 3, 2017, pp. 036025.

[11] M. R. Dawson, C. Sherstan, J. P. Carey, et al., "Development of the Bento Arm: An improved robotic arm for myoelectric training and research, in *Proc. of MEC'14: Myoelectric Controls Symposium*, Fredericton, New Brunswick, August 18–22, 2014, pp. 60–64.

[12] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction.* MIT Press Cambridge, Massachusetts; London, England, 1998.

[13] R. S. Sutton, J. Modayil, M. Delp, et al., "Horde: A Scalable Real-time Architecture for Learning Knowledge from Unsupervised Sensorimotor Interaction," *Proc. of 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS),* May 2-6, Taipei, Taiwan, 2011, pp. 761–768.

[14] P. M. Pilarski et al., "Adaptive artificial limbs: a real-time approach to prediction and anticipation," *IEEE Robotics & Automation Magazine*, vol. 20, no. 1, pp. 53–64, March 2013

[15] A. L. Edwards, J. S. Hebert, P. M. Pilarski, "Machine learning and unlearning to autonomously switch between the functions of a myoelectric arm," *6th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, June 26-29, 2016, Singapore, pp. 514-521.

[16] A. L. Edwards, M. R. Dawson, J. S. Hebert, C. Sherstan, R. S. Sutton, K. M. Chan, P. M. Pilarski, "Application of real-time machine learning to myoelectric prosthesis control: A case series in adaptive switching," *Prosthetics & Orthotics International*. vol. 40, no. 5, 573-581, 2016.

[17] J. R. Flanagan, P. Vetter, R. S. Johansson, and D. M. Wolpert, "Prediction precedes control in motor learning," *Current Biology*, vol. 13, no. 2, pp. 146–150, 2003.

[18] J. B. Travnik and P. M. Pilarski "Representing high-dimensional data to intelligent prostheses and other wearable assistive robots: A first comparison of tile coding and selective Kanerva coding", *2017 International Conference on Rehabilitation Robotics (ICORR)*, pp. 1443–1450, 2017.

[19] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet classification with deep convolutional neural networks", *Advances in Neural Information Processing Systems 25*, 2012, pp. 1097–1105.

[20] A. Graves, A.-r. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks", *2013 IEEE Int. Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 6645–6649.