‘Scrape That Barnacle’: Commanding Underwater Robot In-Contact Manipulation Tasks with Intuitive Spatial-Temporal-Force Features

Ramya Challa[1], Hunter Brown[1], Chirag Jain[1], Zachary Speiser[1], and Heather Knight[1]

Abstract— **How might a robot operator command underwater robot in-contact manipulation tasks? One of the boons of modern language models is the flexibility of what you can describe, and physical tasks have universal features, in that they occur in space and time, and that physics will predict certain outcomes (like an unstable block tower falling over). This work seeks to command spatial-temporal-force priorities and strategies between a human operator and robot, such that it can effectively perform underwater in-contact manipulation tasks, introducing a notation, study design, and ongoing theory building of effective command features. If successful, we expect such paradigms to cross apply to other out-of-water domains, such as giving an elderly person a sponge-bath, which we all know should be a *gentler* operation than hull cleaning. These kinds of adverbs are our hypothesized critical features.**

I. INTRODUCTION

All people have somatic intelligence. We use our bodies to move through the world, open doors, clean up spilled milk, or decide when to use the rough vs. soft side of a sponge. ‘Soma’ refers to body, thus somatic is people’s body-based understandings, whether trained in engineering or not. This work seeks to apply Laban Effort features – previously used in expressive robotics research [1] – to enable humans to intuitively command functional robots underwater.

Why underwater robots? People do not have gills, and many underwater environments, such as dam maintenance, currently employ both scuba divers and ROV-operators for tasks that may be high risk. Imagine tossing sand bags to fill a hole in a dam; if you put your own body in the wrong place, you may yourself be pulled into the hole (mortality rates for mistakes are high); in addition, tossing it right might let you fill it more effectively or efficiently, estimating both water flow and spatial positioning like an experienced athlete. What if humans could offer expertise with low personal risk?

While much previous work in underwater manipulation focuses on pick-and-place [2], i.e., retrieving artifacts from the ocean floor, plucking plastic bags out of the water, or pulling up a lobster cage, *in-contact manipulation is less explored.* Surface maintenance, sampling sensors, flushing out underwater communication cables of seawater, clearing biofouling (the algae growth in water with sunlight), and scraping barnacles present many opportunities for future robots if we could only communicate how to do them. Here we present our early work en route to solving this problem with body-intelligent human commanders.

II. RELATED WORK

Prior work in robot expressive motion has successfully varied robot task motions, customizing Laban features to communicate task objectives, attitudes, and inner state [3][1]. These methods have explored several ways to parameterize motion quality for mobile, flying, and upper-body humanoid robots, the most notable of which is the Laban Movement Analysis (LMA). LMA was developed for describing dance movements and defines every motion as being made up of various factors. For example, the Weight Effort defines how light or heavy a motion is. Parameterizing motion quality is broadly applicable. Claret et al. designed a robot arm controller illustrating the use of kinematic redundancy to embed motion quality as a secondary objective [4].

In-contact manipulation typically uses force controllers, such as impedance control, which attaches a virtual spring to a goal whose stiffness can be adjusted to create varied motion qualities [5], and trajectories by definition are spatial-temporal sequences of joint angles and positions. Controllers have a large impact on how a trajectory is executed. Other efforts to command underwater robots illustrate how language – such as spatial references – can be parsed to command relative robot motions [6]. *In this work, we propose to combine the human-centered motion feature descriptives of LMA with functional in-contact manipulation control, collecting trajectory examples and verbal descriptives.*

III. SPACE-TIME-FORCE NOTATION

Let us begin with a notation to define our Laban-inspired features, originally introduced in [3], focusing on three of the four Laban efforts: Space, Time, Weight. In our notation, each effort category act as instances of the Laban Effort class (using OOP metaphors). We annotate these Efforts as Ls, Lt, Lw, each of which can be set to contrasting polar values like DIP switches.

$$\begin{aligned} \text{Space, } Ls &\in \{\text{direct, indirect}\} \\ \text{Time, } Lt &\in \{\text{sudden, sustained}\} \\ \text{Weight, } Lw &\in \{\text{strong, light}\} \end{aligned} \tag{1}$$

Effort instances can also be modeled as continuous variables spanning between one pole and the other:

$$Lt \in \mathbb{R}[-1 \cdots 1], \text{e.g., sudden := -1, sustained := 1} \tag{2}$$

This feature vector f can be modified to fit a robot’s degrees of freedom. Our setting function S propagates the categories $c = \{Lt, Ls, Lw, Lf\}$ into a features vector f.

---

*This work was supported by the Office of Naval Research (ONR).

[1] Collaborative Robots and Intelligent Systems Institute, Oregon State University, Corvallis, OR, USA `challara@, brownhun@, jainch@, speiserz@ knighth@oregonstate.edu`

IV. PROPOSED SYSTEM DESIGN

Expressive motion and automated motion planning both focus on creating specific motions, however, force-based features have yet to be integrated into Laban Weight feature calculations, although touch is motion-in-collision, and therefore similar features should apply. Thus, we hope combining these two approaches together would be novel and useful because various motion qualities may be better described with force features included.

Humans are good at adapting to variation. When attempting a peg-in-hole task like plugging in a communication cable, a simple planner may work just fine to get in the vicinity of the hole. But what if the connector is delicate, or the area around the plug easy to scuff up? It would be useful to leverage human's intuition to specify how the robot should execute tasks (e.g. "carefully" or "gently"). To implement such a system will involve (1) identifying high-level goals, (2) traveling to the desired manipulation surface, and (3) manipulating the surface accordingly.

Trajectory planners view the global trajectory and handle broader elements of the overall motion quality. For example, if a robot arm is asked to perform a peg-in-hole task "carefully," this may necessitate giving larger buffers to obstacles, avoiding areas of the workspace, or limiting the robot dynamics. This trajectory planner will be designed controller-aware to maximize the utilization of the low-level controllers. Being controller aware allows the trajectory planner to fully consider the execution of a trajectory, which is particularly relevant to Laban Space features execution. At the manipulation surface, actions can additionally include force-based features.

V. COLLECTING TASK EXAMPLES AND WORDS

The hypotheses of this work: (1) people will intuitively use verbage that maps to Laban features in natural language commands instructing particular in-contact manipulation tasks – particularly when given the opportunity to describe effective and ineffective strategies, and, (2) task-relevant motion demonstrations will predict Laban features, e.g., considering the end-effector as the point from which to calculate relative angles and velocities, e.g., a plugging task would have more spatial constraints than general surface cleaning.

A user study, performed first in-air and later in-water, will validate whether participant demonstrations and natural language commands consistently select Laban-inspired features. The goal is to map varied robot control strategies to linguistic concepts like poking a surface "lightly" versus "forcefully." While our target (and past work) is the Bravo arm because it is designed for underwater use, a Franka robot arm will be used to collect demonstrations, as the former was not backdrivable, and the latter is human safe.

The above-water test environment features a built environment: tabletop, back-wall, robot, and coral and seaweed features. Note that the current emphasis for human-in-the-loop command is in-contact manipulation, due to the diversity of abilities that would enable and its novelty relative to pick-and-place. This work is approved by the ethics office of our university and the Office of Naval Research.

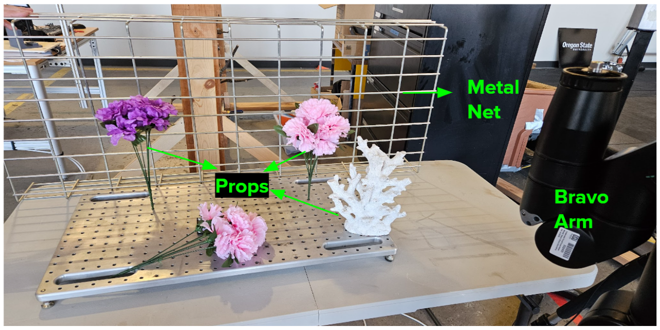

Fig. 1. Representative Setup: Backdrops will act as the Target Manipulation Surface with variable 3D Obstacles from Floor or Overhead Crossbar.

VI. DISCUSSION & CONCLUSIONS

One way to design human-centered robots is to utilize control abstractions that intuitively map to human experience and decision making. For example, young babies often hit themselves in the face with their own hand, crying because that is uncomfortable but not initially realizing that it is their own 'motor control system' (muscles). Not unlike reinforcement learning, those understandings and models improve with experience. A year later, that child may toss toys off tables to gain understanding of the physical world.

Physics is therefore a recognizable control abstraction for people seeking to instruct robots and maintain mental models of robot objectives, attitudes, and function. In fact, specialists like gymnasts, mechanical engineers, dancers, cleaners, and construction workers continue to hone and specialize their physical understandings in order to effectively perform their career. A rugby player may intentionally collide with someone to break their momentum, whereas a robot maid may gently dust off a surface with delicate glass objects.

While the current project focuses on singular overall task commands and their mappings to Laban-inspired features demonstrated in air, future underwater testing may result in distinct Laban feature mappings, as water is around 800 times denser than air and fairly incompressible, thus can add counter-forces. Future work could integrate in-the-moment corrections through speech or clarifying spatial gestures. Barnacles watch out!

REFERENCES

[1] A. Bacula and H. Knight, "Motis parameters for expressive multi-robot systems: Relative motion, timing, and spacing," *International Journal of Social Robotics*, vol. 14, pp. 1–29, 10 2022.

[2] S. Sivčev, J. Coleman, E. Omerdić, G. Dooly, and D. Toal, "Underwater manipulators: A review," *Ocean engineering*, 2018.

[3] H. Knight, "Expressive motion for low degree-of-freedom robots," Ph.D. dissertation, Carnegie Mellon University, August 2016.

[4] J.-A. Claret, G. Venture, and L. Basañez, "Exploiting the robot kinematic redundancy for emotion conveyance to humans as a lower priority task," *International Journal of Social Robotics*, vol. 9, 2017.

[5] N. Hogan and S. P. Buerger, "Impedance and interaction control," in *Robotics and automation handbook*. CRC press, 2018, pp. 375–398.

[6] I. C. Rankin, S. McCammon, and G. A. Hollinger, "Robotic information gathering using semantic language instructions," in *International Conference on Robotics and Automation*. IEEE, 2021.