Online Learning of Motion Constraints from Human Corrections during Ergodic Surface Finishing

Stefan Schneyer[1,2], Edoardo Fiorini[1], and João Silvério[1]

***Abstract*—Surface finishing—such as sanding, polishing, and grinding—is a time-consuming and health-critical process. Robotic automation offers a promising solution to this challenge, enabling consistent surface coverage and tool control. Ergodic control addresses this contact-rich task by covering target distributions in an online manner, accounting for coverage history. We extend our prior ergodic control framework for surface finishing [1] to enable online learning from human interactions during task execution. A human operator can physically intervene via two complementary modes: corrective guidance applied during contact update motion constraints, while repositioning actions update the desired coverage distribution by indicating where more or less attention is needed. This closes the loop between human feedback and robot learning, transforming the system from offline learning to fully interactive human-robot collaboration, where the robot continuously refines both *where* and *how* it works. A video demonstration showing online correction of motion preferences on a real workpiece is available at https://youtu.be/kXhaz11jTac.**

***Index Terms*—Ergodic Control, Human-Robot Interaction, Online Learning, Learning from Demonstration, Surface Finishing**

I. Introduction and Motivation

Surface finishing tasks such as sanding, polishing, and grinding require effective coverage of a workpiece and adherence to material- or geometry-specific motion patterns [2]–[5]. Robotic systems can automate these demanding operations while achieving effective coverage and consistent motion patterns. In our prior work [1], we applied ergodic control based on the Spectral Multiscale Coverage (SMC) method [6] to surface finishing, incorporating a tool imprint model and learning both desired coverage distributions $p(\mathbf{x})$ and motion constraints $\mathbf{\Gamma}(\mathbf{x})$ from human demonstrations, where $\mathbf{x}$ denotes the position of the robot on the surface.

However, this workflow is strictly offline: the operator provides demonstrations, $\mathbf{\Gamma}(\mathbf{x})$ and $p(\mathbf{x})$ are learned, and the robot executes without further adaptation. If the operator observes suboptimal behavior during execution — for instance, incorrect motion directions or misallocated coverage — there is currently no mechanism to update these online.

Contribution. We extend the framework from [1] with: (1) a physical human contact intention detection module that distinguishes corrective interventions and repositioning from normal execution, (2) online updating of the motion constraints ($\mathbf{\Gamma}$-map) from corrective motion data, and (3) online updating of the desired coverage distribution ($p(\mathbf{x})$) from repositioning actions, closing the loop between human guidance and robot learning.

II. Ergodic Control with Motion Preferences

The robot operates on a 3D surface, projected onto a 2D domain U via isometric surface parametrization [1]. The Spectral Multiscale Coverage (SMC) method [6] drives the robot such that its time-averaged coverage $c(\mathbf{x}, t)$ converges to a desired distribution $p(\mathbf{x})$. The ergodic cost $\phi(t) = \frac{1}{2}\sum_{\mathbf{k}} \Lambda_{\mathbf{k}}(p_{\mathbf{k}} - c_{\mathbf{k}}(t))^2$ compares the distributions via Fourier coefficients, weighted by $\Lambda_{\mathbf{k}}$. The optimal velocity is obtained by minimizing ϕ over the next time step under a speed constraint $\|\mathbf{u}\| = u_{\max}$, yielding $\mathbf{u}^* \propto -\mathbf{b}(t)$ [6], with $\mathbf{b}(t) = \sum_{\mathbf{k}} \Lambda_{\mathbf{k}}(p_{\mathbf{k}} - c_{\mathbf{k}}(t))\boldsymbol{\nabla} f_{\mathbf{k}}(\mathbf{x}(t))$ and $f_{\mathbf{k}}$ the Fourier basis functions. To encode preferred motion directions, the isotropic speed constraint is generalized to an ellipsoidal one via a positive-definite matrix $\mathbf{\Gamma}(\mathbf{x})$ [1], replacing $\mathbf{u}^* \propto -\mathbf{b}(t)$ with

$$\mathbf{u}^* = -\frac{\mathbf{\Gamma}(\mathbf{x})\,\mathbf{b}(t)}{\sqrt{\mathbf{b}(t)^{\mathsf{T}}\mathbf{\Gamma}(\mathbf{x})\,\mathbf{b}(t)}}\,. \tag{1}$$

The $\mathbf{\Gamma}$-map is learned from a demonstration dataset $\mathcal{D} = \{(\mathbf{x}^{(i)}, \mathbf{u}^{(i)})\}_{i=1}^{M}$ via kernel regression with a radial basis function kernel $k(\cdot,\cdot)$,

$$\mathbf{\Sigma}_u(\mathbf{x}) = \sum_{i=1}^{M} k(\mathbf{x}, \mathbf{x}^{(i)})\,\mathbf{u}^{(i)}\mathbf{u}^{(i)\mathsf{T}}\,, \tag{2}$$

and blended with an isotropic term to gracefully degrade in data-sparse regions,

$$\mathbf{\Gamma}(\mathbf{x}) = \lambda\,\mathbf{\Sigma}_u(\mathbf{x}) + (1-\lambda)\,\tfrac{1}{2}\mathbf{I}\,. \tag{3}$$

III. Proposed Extension: Online Learning from Human Corrections

The DLR SARA robot is equipped with joint torque sensors that enable external force estimation during task execution [7]. We extend the interaction model from [1] by distinguishing two types of interventions: *repositioning*, where the operator lifts, moves, and places the tool to indicate coverage preferences, and *corrective guidance*, where the operator applies lateral forces while in contact to guide motion direction and update the $\mathbf{\Gamma}$-map.

To distinguish corrective human forces from environmental contacts, we build on the energy-tank-based intention detection from [8]. This method tracks power $P = \mathbf{F}^{\mathsf{T}}\dot{\mathbf{x}}$ in independent energy tanks for the tangential space and normal direction of the surface at the current contact point and computes an intention ratio $h \in [0, 1]$ that differentiates human interactions. The intention ratio modulates stiffness as $K(t) = (1 - h(t))\,K_0$,

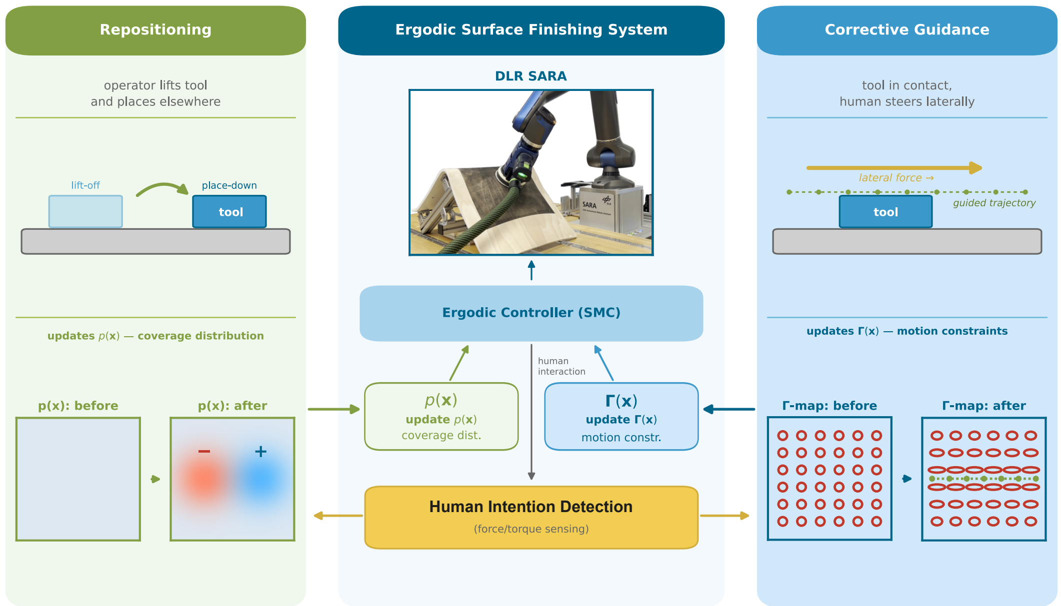

Fig. 1. The system supports two physical interaction modes during ergodic surface finishing: *repositioning* (left), where the operator lifts and places the tool to update the desired coverage distribution $p(\mathbf{x})$, and *corrective guidance* (right), where lateral forces applied while the tool is in contact update the motion constraint map $\mathbf{\Gamma}(\mathbf{x})$ via online kernel regression. Human intent is detected via energy-tank-based force/torque sensing at the DLR SARA robot.

where K_0 is the nominal stiffness in the absence of human intent, yielding compliant behavior during interaction analogous to gravity compensation mode.

When the tangential energy tank indicates human interaction while the tool maintains contact, the system enters a *correction mode* and records the guided trajectory $\{(\mathbf{x}^{(j)}_{\text{corr}}, \mathbf{u}^{(j)}_{\text{corr}})\}$ as new demonstration data. After the corrective intervention, this trajectory is appended to the demonstration dataset $\mathcal{D}$, and $\mathbf{\Gamma}(\mathbf{x})$ is re-evaluated via kernel regression. The RBF kernel ensures corrections only affect the $\mathbf{\Gamma}$-map in the vicinity of the corrected region.

When instead the operator repositions the robot, the lift-off and put-down locations carry implicit information about desired coverage allocation. We exploit this by locally decreasing $p(\mathbf{x})$ near the lift-off position, indicating the region is sufficiently covered, and increasing it near the put-down position, directing more attention there. Concretely, $p(\mathbf{x})$ is updated as:

$$p^{\text{new}}(\mathbf{x}) \propto p^{\text{old}}(\mathbf{x}) - \alpha\, k(\mathbf{x}, \mathbf{x}_{\text{lift}}) + \alpha\, k(\mathbf{x}, \mathbf{x}_{\text{place}})\,, \qquad (4)$$

where $\mathbf{x}_{\text{lift}}$ and $\mathbf{x}_{\text{place}}$ are the lift-off and put-down positions, $\alpha > 0$ controls magnitude, and the same RBF kernel ensures locality. The result is clipped to remain non-negative and renormalized. Importantly, the ergodic coverage history is preserved across all interactions, maintaining overall coverage quality.

IV. Conclusion

We proposed an extension to our ergodic surface finishing framework that enables online learning from human interactions during execution, allowing operators to refine both motion constraints ($\mathbf{\Gamma}$-map) and desired coverage distribution ($p(\mathbf{x})$) through corrective guidance and repositioning. This transforms the original offline-learning paradigm into a fully interactive human-robot collaboration system, bridging robot learning and human-robot interaction for surface finishing. A video demonstrating online motion constraint learning from human corrections is available at https://youtu.be/kXhaz11jTac.

References

[1] S. Schneyer, K. Nottensteiner, A. Albu-Schäffer, F. Stulp, and J. Silvério, "An ergodic approach to robotic surface finishing with learned motion preferences," *IEEE Transactions on Robotics*, vol. 42, pp. 342–361, 2026.

[2] H.-Y. Tam, O. Chi-hang Lui, and A. C.K. Mok, "Robotic polishing of free-form surfaces using scanning paths," *Journal Mater. Process. Technol.*, vol. 95, no. 1, pp. 191–200, 1999.

[3] M. Xiao, Y. Ding, and G. Yang, "A model-based trajectory planning method for robotic polishing of complex surfaces," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 4, pp. 2890–2903, 2022.

[4] Y. Wen, D. J. Jaeger, and P. R. Pagilla, "Uniform coverage tool path generation for robotic surface finishing of curved surfaces," *IEEE Robot. Autom. Letters*, vol. 7, no. 2, pp. 4931–4938, 2022.

[5] B. Alt, F. Stöckl, S. Müller, C. Braun, J. Raible, S. Alhasan, O. Rettig, L. Ringle, D. Katic, R. Jäkel, M. Beetz, M. Strand, and M. F. Huber, "Robogrind: Intuitive and interactive surface treatment with industrial robots," in *IEEE Int. Conf. Robot. Autom.*, 2024, pp. 2140–2146.

[6] G. Mathew and I. Mezić, "Metrics for ergodicity and design of ergodic dynamics for multi-agent systems," *Physica D: Nonlinear Phenomena*, vol. 240, no. 4-5, p. 432–442, 2011.

[7] M. Iskandar, O. Eiberger, A. Albu-Schäffer, A. De Luca, and A. Dietrich, "Collision detection, identification, and localization on the DLR SARA robot with sensing redundancy," in *IEEE Int. Conf. Robot. Autom.*, 2021, pp. 3111–3117.

[8] E. Fiorini, M. Knauer, T. Eiband, M. Iskandar, and J. Silvério, "Interactive learning via physical human feedback using uncertainty-aware energy tanks," *IEEE Robotics and Automation Letters*, 2025.