Learning-Based Ergodic Control for Interactive Surface Finishing with Collaborative Robots

Stefan Schneyer
German Aerospace Center (DLR)
Institute of Robotics and Mechatronics (RM)
Münchener Str. 20, 82234 Weßling
Germany
stefan.schneyer@dlr.de

Alin Albu-Schäffer
Technical University of Munich (TUM)
School of Computation, Information and Technology
Boltzmannstr. 3, 85748 Garching
Germany
alin.albu-schaeffer@dlr.de

Freek Stulp
German Aerospace Center (DLR)
Institute of Robotics and Mechatronics (RM)
Münchener Str. 20, 82234 Weßling
Germany
freek.stulp@dlr.de

João Silvério
German Aerospace Center (DLR)
Institute of Robotics and Mechatronics (RM)
Münchener Str. 20, 82234 Weßling
Germany
joao.silverio@dlr.de

Abstract: Surface finishing is a critical yet challenging task in manufacturing, often requiring significant manual effort and expertise. This paper presents a novel approach to automate surface finishing using collaborative robots that learn from human demonstrations. By leveraging user-friendly programming methods and ergodic control techniques, our system enables real-time task execution and interaction. The proposed method maps human demonstrations onto a two-dimensional parametric space, learns the task description using a Gaussian mixture model, and employs an ergodic controller to guide the robot. Evaluation on a 7-DoF torque controlled robot demonstrates the effectiveness of our approach in generating accurate task descriptions and facilitating user interaction during execution.

Keywords: Ergodic Control, Surface Finishing, Robots, Learning, Collaboration

1 Introduction

Collaborative robots, or cobots, are designed to work alongside humans in shared environments, setting them apart from traditional robots that operate in isolation. What distinguishes cobots is their ability to interact with human partners safely and intuitively, adapting to dynamic and uncertain environments. While recent advancements in learning-based systems have significantly improved these interactions, fully integrating cobots into everyday tasks remains a substantial challenge. This paper addresses these challenges in the context of surface finishing tasks. Surface finishing operations, such as polishing, grinding, and sanding, are essential in manufacturing and construction, particularly in industries like woodworking where surface quality relies heavily on treatment methods [1, 2]. These tasks are typically repetitive, time-consuming, and hazardous to human health, so automation offers great advantages and several benefits for both workers and the economy. It reduces

8th Conference on Robot Learning (CoRL 2024), Munich, Germany.

health risks by limiting exposure to harmful dust and noise, while also increasing efficiency, cutting costs, and improving accuracy and quality in manufacturing. Robotic systems offer the flexibility needed to automate surface finishing tasks, even for complex geometries [3, 4, 5].

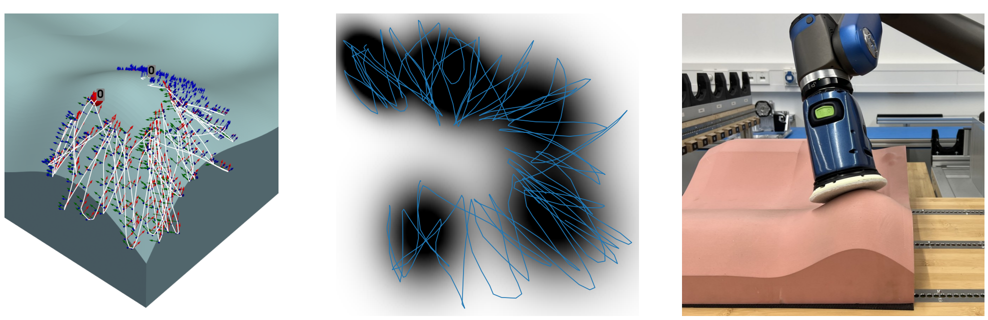
However, automating these processes is difficult due to the challenges in quantifying surface quality and predicting machining effects. As a result, surface finishing remains one of the least automated processes, with most of the work still performed manually by human workers [6, 7]. Additionally, programming robots for these tasks poses significant challenges, particularly in developing intuitive, user-friendly programming methods and a straightforward way to communicate task instructions, specifying precisely what the robot should do. These instructions must be flexible enough to accommodate changes in the environment and adapt to user feedback. Current offline programming methods often lack the adaptability required, resulting in areas being either under-processed or over-processed. Despite these obstacles, there is growing interest in advancing robotic systems for surface finishing, due to the substantial benefits they offer to both workers and industries.

Ergodic control is a method used to ensure that a robot's trajectory covers a given space in a manner that is proportional to a specified probability distribution. Unlike traditional control methods that follow a predefined path, ergodic control dynamically adjusts the robot's movements to achieve a desired spatial coverage. This is particularly useful in tasks where thorough exploration [8, 9] or coverage of an area is required, such as for cleaning tasks [10] or search and rescue operations [11].

Building on ergodic imitation methods as shown in [12], our approach infers robotic tasks using learning from demonstration without the need to code. To efficiently cover surfaces, we introduce an ergodic controller that dynamically guides the robot during task execution. The task is parameterized using probability distributions leveraging Learning from Demonstration (LfD) [13]. This simplifies programming and makes it more intuitive. Additionally, the system supports real-time user interaction, allowing users to adjust the robot's position during task execution to interrupt unwanted behavior, and direct the robot to specific areas as needed. A key feature of our approach is the use of a variable impedance scheme within a state machine, which facilitates smooth and effective interaction by adjusting the robot's stiffness based on the distance to the surface.

2 Approach

In our approach, a human operator demonstrates the surface finishing task by guiding the robot over the surface of a workpiece (e.g. shown in fig. 1a). This involves the operator manually moving the robot to perform actions such as sanding, polishing, or grinding to achieve the desired surface quality. During this process, the robot records position data, which is later used to develop a probabilistic model of the task. This model captures key aspects of the surface finishing process, including the specific areas to be processed and the required intensity of the finishing.



(a) Demonstration (b) Learned distribution (c) Robot execution

Figure 1: The figure shows the human demonstration of the surface finishing task (a), the desired distribution learned from the demonstration using a Gaussian mixture model in the two-dimensional parametric space (b), and the robot execution (c).

2.1 Learning the Task Description and Inferring the Robot Path

To facilitate learning, the complex surface of the workpiece is mapped to a two-dimensional space using an as-close-as-possible isometric surface parametrization, as shown in related work [14]. This transformation reduces the three-dimensional position data from the human demonstration to a two-dimensional space. The position data is then used to learn a probability distribution that serves as the reference distribution for the ergodic controller. This reference distribution, which we also refer to as the task description, is achieved using a Gaussian mixture model (GMM), which can be learned from the data using the expectation-maximization (EM) algorithm to find the maximum likelihood estimation for the data. The EM algorithm iteratively performs a expectation and a maximization step to update the parameters of the GMM. The maximization problem for the GMM is to find the parameters $\boldsymbol{\theta}$ that maximize the log-likelihood function as

$$\log p(\boldsymbol{X} \mid \boldsymbol{\theta}) = \sum_{i=1}^{N} \log \left(\sum_{c=1}^{C} \pi_c \mathcal{N}(\boldsymbol{x}_i | \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c) \right),$$

$$\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} \log p(\boldsymbol{X} \mid \boldsymbol{\theta}),$$

where $\boldsymbol{\theta} = \{\pi_c, \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c\}_{c=1}^{C}$ represents the parameters of the GMM, π_c are the mixing coefficients, $\mathcal{N}(\boldsymbol{x}_i | \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c)$ is the Gaussian distribution with mean $\boldsymbol{\mu}_c$ and covariance $\boldsymbol{\Sigma}_c$, and N is the number of data points. Here, $\boldsymbol{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}$ denotes the dataset consisting of the two-dimensional position data $\boldsymbol{x} \in \mathbb{R}^2$ from the human demonstration, where $\boldsymbol{x}$ represents the end-effector coordinates on the surface coordinate system. The GMM provides a probabilistic framework that allows the robot to understand and replicate the task by following the learned distribution.

For execution, an ergodic controller is used to guide the robot along the task description in real-time, optimizing ergodicity. In this work, we are using the SMC (Spectral Multiscale Coverage) [15, 16, 17] approach to guide the robot along the task description, where ergodic control aims to achieve a desired coverage, represented by a probability density function $p(\boldsymbol{x})$, which is the task description learned from human demonstrations as

$$p(\boldsymbol{x}) := p(\boldsymbol{x} \mid \boldsymbol{\theta}^*).$$

Alternative, it can also be pre-defined, as shown in [16]. The coverage at time t of a single agents is calculated as

$$c(\boldsymbol{x}, t) = \frac{1}{Mt} \sum_{i=1}^{M} \int_0^t \delta(\boldsymbol{x} - \boldsymbol{x}_i(t)) \, \mathrm{d}t,$$

where $\delta(\boldsymbol{x})$ is the Dirac delta function. This coverage function keeps track of the history of the agent across multiple interrupted paths M, indicating which areas of the state space have already been covered. The cost function, denoted as $\Phi(t)$, quantifies the deviation of the coverage $c(\boldsymbol{x}, t)$ from the target distribution $p(\boldsymbol{x})$. It is defined as

$$\Phi(t) = \sum_{k \in K} \Lambda_k (p_k - c_k(t))^2,$$

where p_k and $c_k(t)$ are the Fourier coefficients for the two-dimensional wave-number vector $k \in K$ of the desired probability density function $p(\boldsymbol{x})$ and the coverage $c(\boldsymbol{x}, t)$, respectively. For more details on the ergodic control formulation, we refer to [16]. The action $\boldsymbol{u}$ of the ergodic controller is derived from the optimal control formulation as

$$\boldsymbol{u}^* = \arg\min_{\boldsymbol{u}} \Phi(t + \mathrm{d}t) \quad \text{subject to} \quad \|\boldsymbol{u}\|_2 = u_{\max}.$$

The solution to this optimization problem is given by

$$\boldsymbol{u}^* = -u_{\max}\frac{\boldsymbol{b}(t)}{\|\boldsymbol{b}(t)\|_2},$$

where $\boldsymbol{b}(t)$ is defined as

$$\boldsymbol{b}(t) = \sum_{k\in K} \Lambda_k (p_k - c_k(t)) \boldsymbol{\nabla} f_k(\boldsymbol{x}(t)),$$

and f_k are the Fourier basis functions.

Given the control action at each time step, the robot's trajectory is generated by integrating these actions over time. The ergodic controller continuously guides the robot according to the task specification, ensuring efficient and thorough coverage of the task space in real-time. Additionally, the controller allows for dynamic adjustments, enabling the robot to replan at each time step in response to user interaction or changes in the desired coverage distribution.

2.2 State Machine and Stiffness Adaptation for Task Execution and Interaction

The robot employs impedance control [18] with a force overlay $\boldsymbol{F}_d$ to generate the desired force normal to the surface during task execution as

$$\boldsymbol{\tau} = \boldsymbol{J}^{\top}(\boldsymbol{q})[\boldsymbol{G}_P(\boldsymbol{x}_d - \boldsymbol{x}) + \boldsymbol{G}_D(\dot{\boldsymbol{x}}_d - \dot{\boldsymbol{x}}) + \boldsymbol{F}_d]$$

where $\boldsymbol{G}_P$ is the stiffness matrix and $\boldsymbol{G}_D$ the damping gains, $\boldsymbol{J}(\boldsymbol{q}) = {}^{\mathrm{d}\boldsymbol{x}}/_{\mathrm{d}\boldsymbol{q}}$ denotes the Jacobian for configuration $\boldsymbol{q}$, and $\boldsymbol{x}_d$, $\boldsymbol{x}$ the end-effector desired and current position. To facilitate seamless interaction and task execution, the robot operates using a state machine that manages different states based on the distance to the surface. This state machine allows the user to interact with the robot during task execution to change the robot's position on the surface. The stiffness of the impedance is adjusted depending on the distance to the surface. During task execution, the stiffness is set to a higher value to follow the task description, while during interaction, the stiffness is set to a lower value to make it easier for the user to move the robot. The transitions between these states are governed by specific conditions related to the height difference Δh between the tool center point and surface:

- **Idle State:** The robot lifts the tool from the surface and waits for the user input to start the task.
- **Surface Finishing State:** When the height difference is below the threshold $d_{\min}$, the robot switches to the surface finishing state, where the robot follows the task description using ergodic control. This state is always active if the robotic surface finishing tool is in contact with the surface and can be interrupted when the user lifts the robot off from the surface.
- **Interaction State:** If the height difference exceeds the threshold $d_{\min}$, the robot switches to the interaction state. In this state, the robot is not in contact with the surface, allowing the user to easily reposition it. So, the robot is in interaction state if it does not touch the surface. Depending on the distance to the surface, the robot, using impedance control, will set a lower or higher stiffness to make it easier to move. If the user releases the robot, it will slowly move back to the surface and switch back to the ergodic control state.

The stiffness of the robot's impedance control is adapted based on the height difference to ensure smooth transitions and effective task execution. The interpolation between two stiffnesses based on the height difference to the surface is calculated as

$$\boldsymbol{G}_P = \lambda \boldsymbol{G}_0 + (1-\lambda)\boldsymbol{G}_1,$$

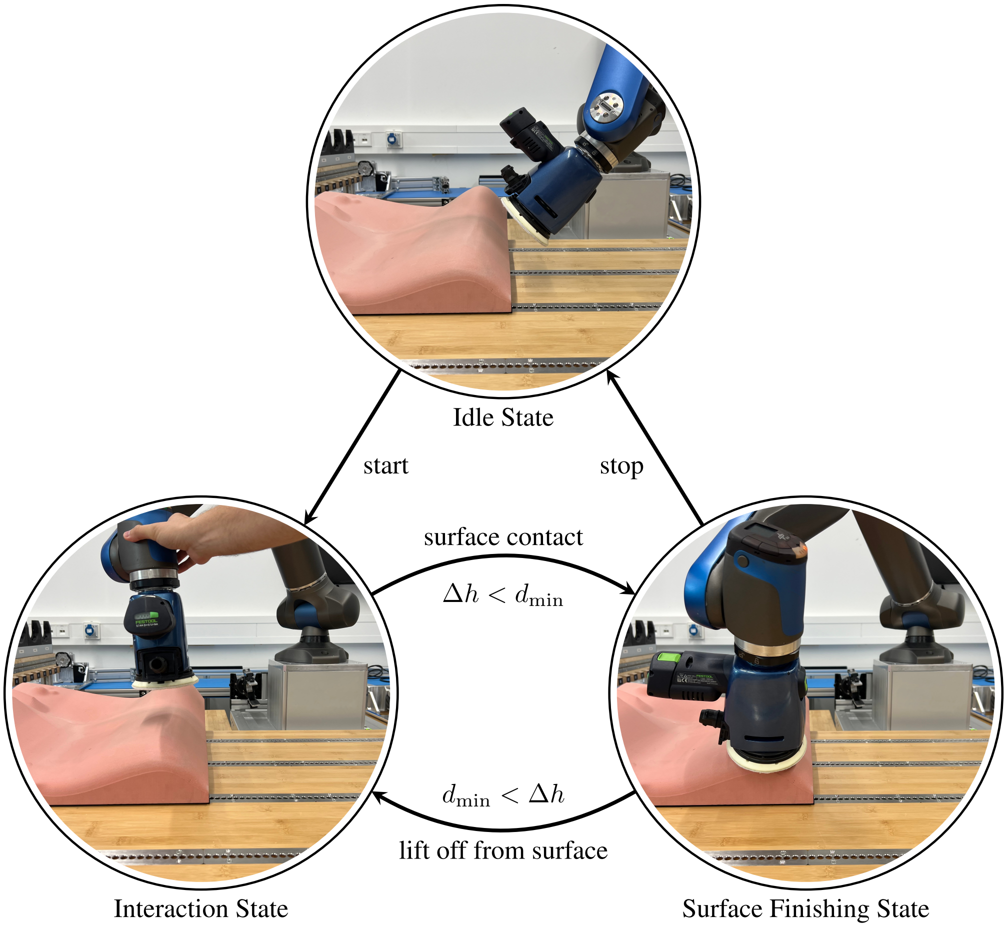


Figure 2: Robot collaboration state machine

where

$$\lambda = \begin{cases} 0 & \text{if } \Delta h \leq d_{\min} \\ \frac{\Delta h - d_{\min}}{d_{\max} - d_{\min}} & \text{if } d_{\min} < \Delta h < d_{\max} \\ 1 & \text{if } d_{\max} \leq \Delta h \end{cases} .$$

The matrix

$$\boldsymbol{G}_0 = \begin{bmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 100 \end{bmatrix} \frac{N}{m}$$

is the stiffness in the interaction state and

$$\boldsymbol{G}_1 = \begin{bmatrix} 1000 & 0 & 0 \\ 0 & 1000 & 0 \\ 0 & 0 & 800 \end{bmatrix} \frac{N}{m}$$

is the stiffness during task execution. This interpolation ensures a smooth transition between states, allowing the robot to adapt its behavior based on the distance to the surface.

3 Experimental Results

The approach is evaluated on a surface finishing task using a 7-DoF torque-controlled robot. A convex, curved surface workpiece is used to demonstrate the method. Initially, a human operator demonstrates the task to the robot, which then follows the task description generated by the Gaussian mixture model. During task execution, the user can interact with the robot to adjust its position.

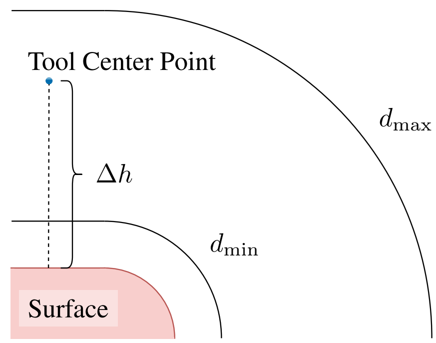


Figure 3: Interpolation between different stiffnesses and discreate states using the shortest distance to the surface

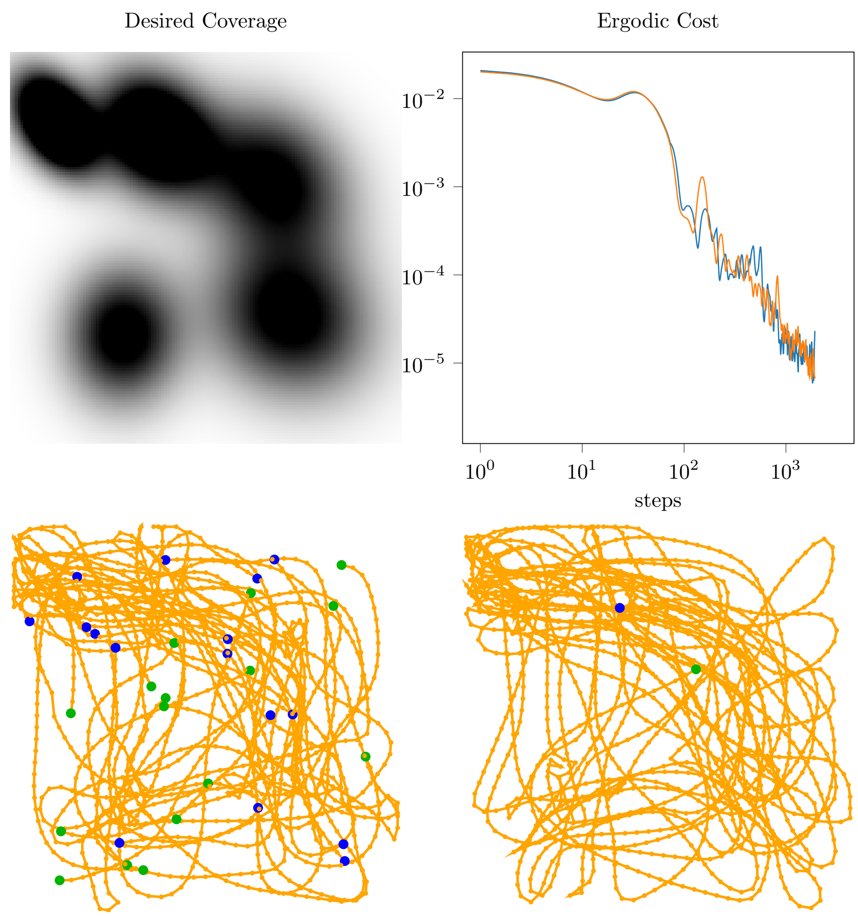


Figure 4: Comparison between execution with the proposed approach using the state machine interrupting the ergodic control path 15 times (bottom left) and execution of a single ergodic control path (bottom right) for a uniform distribution.

Fig. 4 compares two execution scenarios of the proposed approach. The image (top left) shows the Gaussian mixture model learned from human demonstrations, which is used as the desired probability density function for the ergodic controller. The plot (top right) shows the ergodic cost function for the following two scenarios over the time steps. In the first scenario (bottom left), the robot follows the task description while enabling user interaction to modify its position. This capability ensures that the robot can adapt to changes in and user inputs, providing a more flexible and intuitive surface finishing process. In this experiment, the user interrupts the ergodic control path 15 times, which causes the state machine to switch to the interaction state and allows the user to reposition the robot during task execution. In the illustration green dots mark the starting points of trajectories, and blue dots mark the end positions. The state machine effectively manages transitions between interaction and task execution states, ensuring smooth and efficient operation. In the second scenario (bottom right), the robot executes a single uninterrupted ergodic control path. The total execution time for this scenario is approximately 10 minutes. In contrast, the first scenario takes additional time due to user interactions. In both scenarios, the robot effectively follows the task description and ensures thorough coverage of the workpiece surface. The ergodic cost function converges at a similar rate in both scenarios, demonstrating that the robot can adapt to the user's interactions without significantly compromising the efficiency or accuracy of task execution. We further evaluated the approach with a uniform distribution on another surface region, as shown in the appendix A.

4 Conclusion

In this paper, we presented an approach to automate surface finishing tasks using collaborative robots that learn from human demonstrations. By leveraging ergodic control techniques and user-friendly programming methods, our system enables real-time task adaptation, where the user can interact with the robot and correct it's position, while the ergodic controller keeps track of the history already covered by the robot. The stiffness of the impedance control is adapted based on the distance to the surface to ensure smooth transitions. The approach is evaluated on a surface finishing task using a 7-DoF torque controlled robot. Experiments show that the robot adapts to the user's input without significantly affecting the execution of the tasks, with the ergodic costs decrease with the same tendency.

Future work will focus on improving the system's robustness and adaptability to handle more complex and varied surface geometries. This can be achieved by incorporating surface segmentation techniques as demonstrated in related works [19, 20, 21], which will allow the robot to handle different segments of a surface as discrete states. This will enable users to interactively switch between segments during the task. Additionally, optimizing the coverage and effectiveness of the task by considering different contact points on the surface finishing tool, as explored in [22], can be further investigated. This will allow the robot to adapt its contact strategy based on the specific requirements of the surface being finished.

A Experiment with Uniform Coverage

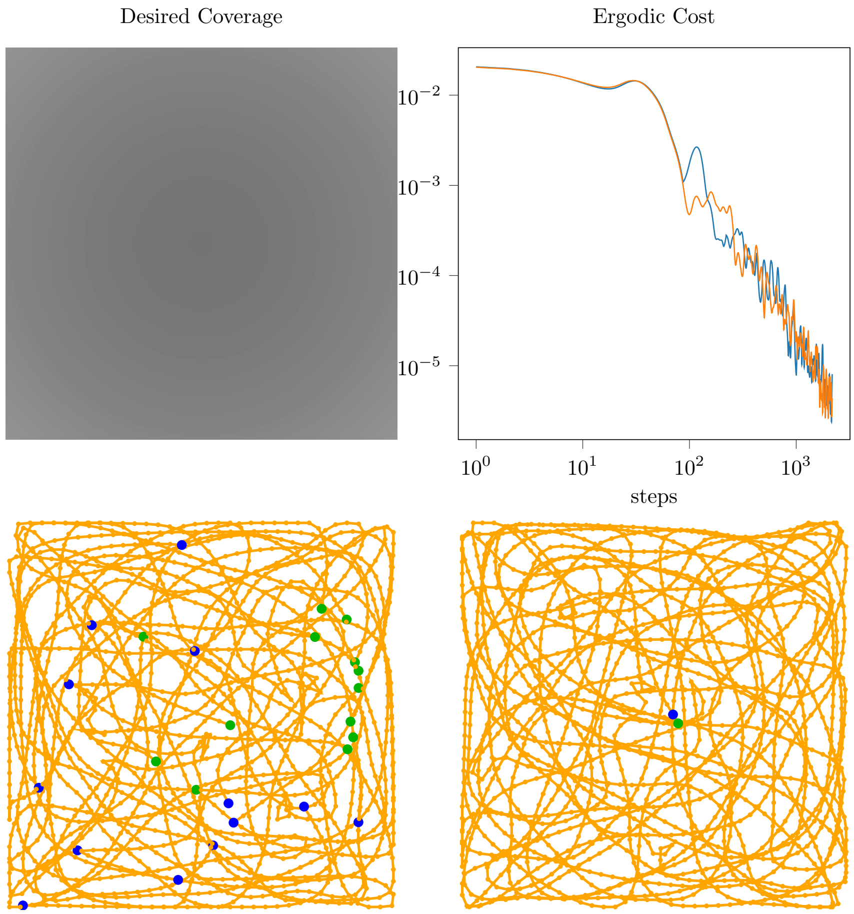


Figure 5: This figure illustrates the comparison between two execution scenarios using the proposed approach with a uniform distribution, which differs from the Gaussian mixture model used in the previous figure. The bottom left shows the robot's path when the state machine interrupts the ergodic control path 12 times, allowing for user interaction. The bottom right depicts the robot's path during a single uninterrupted ergodic control execution.

Acknowledgments

This work was supported in part by the German Federal Ministry of Education and Research (BMBF) through the "The Future of Value Creation — Research on Production, Services and Work" Program under Grant 02K20D032 and in part by the DLR internal Project "Factory of the Future Extended"

References

[1] Y. Huo, P. Li, D. Chen, Y.-H. Liu, and X. Li. Model-free adaptive impedance control for autonomous robotic sanding. In *Transactions Autom. Sci. Eng.* IEEE, 2021.

[2] Y. Zhao, J. Zhao, L. Zhang, L. Qi, and Q. Tang. Path planning for automatic robotic blade grinding. In *Proc. Int. Conf. Mechatron. Autom.*, pages 1556–1560. IEEE, 2009.

[3] 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*, 7(2):4931–4938, 2022.

[4] I. Mohsin, K. He, Z. Li, and R. Du. Path planning under force control in robotic polishing of the complex curved surfaces. *Appl. Sci.*, 9(24), 2019.

[5] M. Xiao, Y. Ding, and G. Yang. A model-based trajectory planning method for robotic polishing of complex surfaces. *IEEE Transactions Autom. Sci. Eng.*, 19(4), 2022.

[6] Y. Li, H. Chen, and N. Xi. Automatic programming for robotic grinding using real time 3d measurement. In *Int. Conf. Cyber Technol. Autom., Control Intell. Syst.*, pages 803–808. IEEE, 2017.

[7] Y. Dong, T. Ren, K. Hu, D. Wu, and K. Chen. Contact force detection and control for robotic polishing based on joint torque sensors. *Int. Journal Adv. Manuf. Technol.*, 107:2745–2756, 2020.

[8] S. Shetty, J. Silvério, and S. Calinon. Ergodic exploration using tensor train: Applications in insertion tasks. *IEEE Transactions on Robotics*, 38(2):906–921, 2022. doi:10.1109/TRO.2021.3087317.

[9] C. Bilaloglu, T. Löw, and S. Calinon. Whole-body ergodic exploration with a manipulator using diffusion. *IEEE Robotics and Automation Letters*, 8(12):8581–8587, 2023. doi:10.1109/LRA.2023.3329755.

[10] C. Bilaloglu, T. Löw, and S. Calinon. Tactile ergodic control using diffusion and geometric algebra, 2024. URL `https://arxiv.org/abs/2402.04862`.

[11] L. Lanča, K. Jakac, and S. Ivić. Model predictive altitude and velocity control in ergodic potential field directed multi-uav search, 2024. URL `https://arxiv.org/abs/2401.02899`.

[12] A. Kalinowska, A. Prabhakar, K. Fitzsimons, and T. Murphey. Ergodic imitation: Learning from what to do and what not to do. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3648–3654, 2021. doi:10.1109/ICRA48506.2021.9561746.

[13] H. Ravichandar, A. S. Polydoros, S. Chernova, and A. Billard. Recent advances in robot learning from demonstration. *Annual review of control, robotics, and autonomous systems*, 3 (1):297–330, 2020.

[14] L. Liu, L. Zhang, Y. Xu, C. Gotsman, and S. J. Gortler. A local/global approach to mesh parameterization. In *Eurographics Symp. Geom. Process.*, volume 27, 2008.

[15] G. Mathew and I. Mezić. Spectral multiscale coverage: A uniform coverage algorithm for mobile sensor networks. pages 7872 – 7877, 01 2010. doi:10.1109/CDC.2009.5400401.

[16] G. Mathew and I. Mezić. Metrics for ergodicity and design of ergodic dynamics for multi-agent systems. *Physica D: Nonlinear Phenomena*, 240(4-5):432–442, 02 2011. doi:10.1016/j.physd.2010.10.010.

[17] S. Calinon. Mixture models for the analysis, edition, and synthesis of continuous time series. 2019.

[18] N. Hogan. Impedance control: An approach to manipulation. In *American control conference*, pages 304–313. IEEE, 1984.

[19] X. Liu, Y. Li, and Q. Li. A region-based 3 + 2-axis machining toolpath generation methodfor freeform surface. *Int. Journal Adv. Manuf. Technol.*, 97:1149–1163, 2018.

[20] W. Sheng, N. Xi, M. Song, and Y. Chen. CAD-guided robot motion planning. In *Ind. Robot*, volume 28, pages 143–152, 2001.

[21] J. Liu, X. Huang, S. Fang, H. Chen, and N. Xi. Industrial robot path planning for polishing applications. In *Proc. Int. Conf. Robot. Biomimetics*, pages 1764–1769. IEEE, 2016.

[22] S. Schneyer, A. Sachtler, T. Eiband, and K. Nottensteiner. Segmentation and coverage planning of freeform geometries for robotic surface finishing. *IEEE Robotics and Automation Letters*, 8 (8):5267–5274, 2023. doi:10.1109/LRA.2023.3293309.