

Fast Online Learning of CLiFF-maps in Changing Environments

Yufei Zhu^{1,2}, Andrey Rudenko², Luigi Palmieri², Lukas Heuer^{1,2}, Achim J. Lilienthal^{1,3}, Martin Magnusson¹

Abstract—Maps of dynamics are effective representations of motion patterns learned from prior observations, with recent research demonstrating their ability to enhance various downstream tasks such as human-aware robot navigation, long-term human motion prediction, and robot localization. Current advancements have primarily concentrated on methods for learning maps of human flow in environments where the flow is static, i.e., not assumed to change over time. In this paper we propose an online update method of the CLiFF-map (an advanced map of dynamics type that models motion patterns as velocity and orientation mixtures) to actively detect and adapt to human flow changes. As new observations are collected, our goal is to update a CLiFF-map to effectively and accurately integrate them, while retaining relevant historic motion patterns. The proposed online update method maintains a probabilistic representation in each observed location, updating parameters by continuously tracking sufficient statistics. In experiments using both synthetic and real-world datasets, we show that our method is able to maintain accurate representations of human motion dynamics, contributing to high performance flow-compliant planning downstream tasks, while being orders of magnitude faster than the comparable baselines.

I. INTRODUCTION

Safe and efficient operation in complex, dynamic and densely crowded human environments is a critical prerequisite for deploying robots in various tasks to support people in their daily activities [1, 2, 3]. Considerable efforts are dedicated to support the deployment of mobile robots by enabling them to follow social norms and achieve legible and socially compliant navigation [4].

Extending the environment model with human motion patterns using a *map of dynamics* (MoD) is one way to achieve unobtrusive navigation compliant with existing motion flows in the environment [5, 6], or avoid crowded areas [7]. MoDs are an increasingly popular tool to store spatial-temporal information about patterns of motion in an environment, such as the motions of humans [8]. Maps of dynamics extend the geometric world model, improving the temporal and contextual awareness of the robot’s own configuration space. These maps are useful for several downstream tasks, e.g. to predict long-term motion trajectories [9, 10], improve localization in dynamic environments [11] and robot navigation in crowded spaces [12], particularly planning safe and unobtrusive trajectories [5, 13] and efficiently allocating complex tasks [14]. In

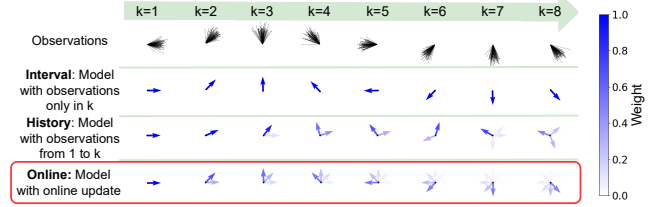


Fig. 1: Online update results of toy example data compared with using only new observations and using all observations to build the model. The **top row** shows raw observations for each of the eight directions ($0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ$), provided in each iteration k . Blue arrows depict the mean vectors of a CLiFF Gaussian mixture model, with transparency indicating the component weights. Three modeling approaches are compared: the **second row** shows models built using only observations from the current iteration k ; the **third row** shows models built with cumulative observations from iteration 1 to k , which are overgeneralized and fail to prioritize recent observations; the **fourth row** shows the proposed online update method, which incorporate new data while retaining relevant historical patterns, offering a dynamic representation of the motion pattern over time.

particular, the CLiFF-map (circular-linear flow field) [15] is a powerful and flexible representation to store velocity, orientation and turbulence of the multi-modal human motion flows in the environment, implicitly incorporating such features as common goal locations, restricted areas and other semantic attributes [16]. Practically implemented on a robot, CLiFF-map generalizes prior observations in a compact spatial representation that can be efficiently queried by downstream components.

However, prior art typically assumes that once learned, the map of dynamics remains constant [17], or follows fixed periodic patterns every day [18]. In reality, as the robot operates continuously, the motion patterns may change due to various events or alterations in the environment topology and semantics [14]. The map of dynamics should actively detect and accommodate these changes. This motivates the need for an online update method, similar to the one developed for iterative improvement of periodic maps of dynamics [19]. Furthermore, as the robot operates over extended periods, the accumulation of human observation data makes rebuilding the map from scratch increasingly costly. An online learning model becomes practical and necessary in scenarios where handling large datasets is infeasible due to storage limitations or privacy concerns. This situation motivates the need for incremental and efficient map updates, which augment previously accumulated knowledge with recent observations without retaining all past dynamics samples.

In this paper we propose a method for online updating of the CLiFF-map of dynamics for mobile robots using a variation of the stochastic expectation maximization algorithm.

¹Robot Navigation and Perception Lab, AASS Research Center, Örebro University, Sweden yufei.zhu@oru.se

²Robert Bosch GmbH, Corporate Research, Stuttgart, Germany

³Chair: Perception for Intelligent Systems, Technical University of Munich, Germany

This work has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101017274 (DARKO).

As new observations are collected, our goal is to update the existing representation to effectively and accurately integrate the new information. At the same time, the robot should not immediately dismiss previously learned patterns while avoiding the need to store the entire historical dataset. The proposed online update method maintains the probabilistic representation in each observed location, updating parameters by continuously tracking sufficient statistics. As shown in Fig. 1, our method not only ensures that the model remains adaptively accurate in reflecting the most recent human motion but also maintains consistency with historical data, thereby preserving a comprehensive understanding of the environment over time. In experiments on both a synthetic dataset and the real-world ATC [20] dataset, we show that our method quickly recognizes changes in environments with sparse and dense motion flows, while being significantly faster than baseline methods. Qualitative results further show that the maps learned with our online update method represent recent trajectories more accurately, and are therefore better suited for informing planning algorithms.

II. RELATED WORK

There are several types of Maps of Dynamics (MoDs) described in the literature, generally striving to provide an efficient tool for storing and querying information about historical or expected changes in states within the environment. Some MoDs model the expected state of occupancy or other binary signals [11, 21], or recurring discrete states of objects [22]. MoDs can be built from various sources of input, such as trajectories [23], dynamics samples, or information about the flow of continuous media (e.g., air or water) [24]. Furthermore, these models can feature diverse underlying representations, including evidence grids, histograms [25], graphs, or Gaussian mixtures. We refer the reader to [8] for a comprehensive survey. In this section, we will focus on MoDs that are designed for, or have been applied specifically to, the motion of people (trajectories, directions, velocities).

Bennewitz et al. [23] learn a collection of human motion patterns from clustered trajectory data (with k -means clustering), using Gaussian mixtures and expectation maximization (EM). A drawback of this method is that such clustering does not lend itself well to online updates with new data and relies on having complete observations of full trajectories in order to cluster the motion patterns that make up the map. Doellinger et al. [26] employ a convolutional neural network trained on simulated trajectories to predict where people are likely to walk, based on an input occupancy grid map. While it does not require full trajectories for training, it cannot easily integrate new observations to update the model. Similarly, [27] predicts a map of dynamics given a static occupancy map. However, the output map in this case is a Bayesian floor field, where each point in the map stores the probabilities of observing motion along a discrete set of directions, rather than only the probability of occupancy.

The CLiFF-map representation [15] has been used recently to facilitate efficient human-aware motion planning [5, 13, 6]

and long-term human motion prediction [28, 9, 29]. CLiFF-map represents local flow patterns as a multi-modal, continuous joint distribution of speed and orientation, as further described in Section III. A benefit of CLiFF-map is that it can be built from incomplete or spatially sparse velocity observations [29], without the need to store a long history of data or deploy advanced tracking algorithms. However, like most other types of existing MoDs, including those mentioned above, CLiFF-maps are typically built offline. A key reason for this is the high computational costs associated with the building process. This constraint limits their applicability in real environments, where patterns can change over time, e.g. due to changes in environment topology and semantics.

STeF-map [18] is another representation that builds a local motion model per cell of a grid map. STeF-map uses a discrete set of eight directions, maintaining a model of when motion can be observed at that point in that direction to learn periodic flow patterns. As a periodic spatiotemporal map of dynamics, it can be used to predict activities at specific times of the day, under the assumption that the periodic patterns exist in this environment. Most closely related to the present work is [19], which uses incremental online updates to iteratively improve the periodic patterns in the STeF-map. In contrast, we assume that the map may be non-stationary (e.g., due to changing obstacle configurations or temporary external conditions [16]), and our online update is designed to adapt to such changes quickly, while not forgetting past flow patterns.

To enable incremental online updates of the CLiFF-map, we use a stochastic expectation maximization algorithm [30] to process streaming data, as further described in Section III, and to achieve an accurate representation of dynamics, gradually balancing historical observations with more recent samples, as demonstrated in Section V.

III. ONLINE CLiFF-MAP UPDATE METHOD

A. Problem statement

A given geometric environment is denoted as $M \subseteq \mathbb{R}^2$, which is discretized to a set of locations $L = \{\text{loc}_1, \text{loc}_2, \dots, \text{loc}_{|L|}\}$. Each location $\text{loc}_l \in L$ aggregates observations collected within a radius r of loc_l . Our aim is to learn a dynamics model at each location that effectively and accurately represents non-stationary local motion patterns. As new observations become available, the model will be dynamically updated to integrate the information from them.

B. Definitions

CLiFF-map represents motion patterns using multimodal statistics to represent speed and orientation jointly [15]. In the CLiFF-map, each location is associated with a Semi-Wrapped Gaussian Mixture Model (SWGMM) [31] to capture the dependency between the speed and orientation, representing motion patterns based on local observations.

The SWGMM represents speed and direction jointly as velocity $\mathbf{V} = [\theta, \rho]^\top$ using direction θ and speed ρ , where $\rho \in \mathbb{R}^+$, $\theta \in [0, 2\pi)$. This semi-wrapped probability density function (PDF) over velocities can be visualized as a function

Algorithm 1: Online MoD update

Input: Number of observing iterations K
Output: online-CLiFF-map $\Xi_{\text{loc}_1, \dots, \text{loc}_{|L|}}$

```

1 for  $k = 1, \dots, K$  do
2   for  $\text{loc}_l = \text{loc}_1, \dots, \text{loc}_{|L|}$  do
3      $\mathbf{y}_k \leftarrow \text{getObservation}(\text{loc}_l)$ 
4     if  $\Xi_{\text{loc}_l} = \emptyset$  then
5        $\Xi_{\text{loc}_l} \leftarrow \text{buildCLiFFmap}(\mathbf{y}_k)$ 
6     else
7        $\hat{\mathbf{s}}_k \leftarrow \text{sE-step}$ 
8        $m_{1, \dots, J}, \boldsymbol{\mu}_{1, \dots, J}, \boldsymbol{\Sigma}_{1, \dots, J} \leftarrow \text{M-step}$ 
9 return  $\Xi_{\text{loc}_1, \dots, \text{loc}_{|L|}}$ 

```

on a cylinder. Direction values θ are wrapped on the unit circle and the speed ρ runs along the length of the cylinder.

Since the direction θ is a circular variable and the speed is linear, a mixture of *semi-wrapped* normal distributions (SWNDs) is used. Encoding the multimodal characteristic of human motion flow, an SWGMM is a PDF represented as a weighted sum of J SWNDs:

$$p(\mathbf{V}|\xi) = \sum_{j=1}^J m_j \mathcal{N}_{\boldsymbol{\Sigma}_j, \boldsymbol{\mu}_j}^{\text{SW}}(\mathbf{V}), \quad (1)$$

where $\xi = \{\xi_j = (m_j, \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) | j \in \mathbb{Z}^+\}$ denotes a finite set of components of the SWGMM, m_j denotes the mixing factor and satisfies $0 \leq m_j \leq 1$, and an SWND $\mathcal{N}_{\boldsymbol{\Sigma}, \boldsymbol{\mu}}^{\text{SW}}$ is formally defined as

$$\mathcal{N}_{\boldsymbol{\Sigma}, \boldsymbol{\mu}}^{\text{SW}}(\mathbf{V}) = \sum_{w \in \mathbb{Z}} \mathcal{N}_{\boldsymbol{\Sigma}, \boldsymbol{\mu}}([\theta, \rho]^\top + 2\pi[w, 0]^\top), \quad (2)$$

where $\boldsymbol{\Sigma}, \boldsymbol{\mu}$ denote the covariance matrix and mean value of the directional velocity $(\theta, \rho)^\top$, and w is a winding number. Although $w \in \mathbb{Z}$, the PDF can be approximated adequately by taking $w \in \{-1, 0, 1\}$ for practical purposes [32].

When the first batch of observations for location loc_l is available, SWGMM parameters are estimated using expectation maximization (EM) [33] along with mean shift [34] to obtain the number and initial positions of dynamics modes. When collecting new observations at the same location loc_l , SWGMM parameters are updated using the stochastic EM method (sEM [30]), which is a fast online variant of the EM algorithm. In sEM, the expectation step of the original EM algorithm is replaced by a stochastic approximation step, while the maximization step remains unchanged.

C. Online map update algorithm

The online update method is summarized in Alg. 1. During each observation iteration, the online CLiFF-map is built and updated with new observations. Let $k \geq 0$ be the iteration number for collecting observations. During iteration k , at location loc_l , new observations $\mathbf{y}_k$ are collected over the time interval $[t, t + \Delta t]$ (line 3 of Alg. 1). If no model has previously been built at loc_l , i.e., no motion has been observed at loc_l yet, an SWGMM will be initiated with $\mathbf{y}_k$ (line 5 of Alg. 1), where $\mathbf{y}_k$ denotes the set of velocity

vectors observed in iteration k of human motion. Otherwise, $\mathbf{y}_k$ will be used to update the existing model for loc_l . The variable N_k represents the number of new observations collected at loc_l during the interval. Let $\hat{\mathbf{s}}_k$ be the estimated sufficient statistics at iteration k , for each component j , where $\hat{\mathbf{s}}_j$ is composed of $(s_j^{(1)}, s_j^{(2)}, s_j^{(3)})$ and the initial value $\hat{\mathbf{s}}_0$ computed using Eq. (7) and initial observations. Sufficient statistics contain all the information needed to estimate model parameters. The sEM method tracks the sufficient statistics using a stochastic approximation procedure. The SWGMM parameters, $m, \boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are then updated using sEM, which consists of two sub-steps:

$$\textbf{sE-step: } \hat{\mathbf{s}}_k = \hat{\mathbf{s}}_{k-1} + \gamma_k (\mathbf{S}_k - \hat{\mathbf{s}}_{k-1}), \quad (3)$$

$$\textbf{M-step: } m_j = \hat{s}_j^{(1)}, \quad (4)$$

$$\boldsymbol{\mu}_j = (\hat{s}_j^{(1)})^{-1} \hat{s}_j^{(2)}, \quad (5)$$

$$\boldsymbol{\Sigma}_j = (\hat{s}_j^{(1)})^{-1} \hat{s}_j^{(3)} - \boldsymbol{\mu}_j \boldsymbol{\mu}_j^\top. \quad (6)$$

In the sE-step, the sufficient statistics of SWGMM are computed and updated with a sequence of stepsizes $(\gamma_k)_{k \geq 1}$. For iteration k and SWGMM components $j \in [1, J]$, the sufficient statistics $\mathbf{S}_k$ is composed of $(S_j^{(1)}, S_j^{(2)}, S_j^{(3)})$. For partitions of the sufficient statistics $p \in [1, 3]$, we have $S_j^{(p)} = N_k^{-1} \sum_{i=1}^{N_k} S_{i,j}^{(p)}$. For each observation i , the observed data is $\mathbf{y}_i$, $S_{i,j}$ is

$$S_{i,j}^{(1)} = \eta_{i,j}, \quad S_{i,j}^{(2)} = \eta_{i,j} \mathbf{y}_i, \quad S_{i,j}^{(3)} = \eta_{i,j} \mathbf{y}_i \mathbf{y}_i^\top, \quad (7)$$

where the responsibility η is

$$\eta_{i,j} = \frac{m_j \mathcal{N}_{\boldsymbol{\Sigma}_j, \boldsymbol{\mu}_j}^{\text{SW}}(\mathbf{y}_i)}{\sum_{j=1}^J m_j \mathcal{N}_{\boldsymbol{\Sigma}_j, \boldsymbol{\mu}_j}^{\text{SW}}(\mathbf{y}_i)}. \quad (8)$$

In each iteration k , the step size γ_k determines the impact of new observations on the model. To effectively update the model, we consider two factors: the motion intensity and the time elapsed since observation. Observations that occurred a long time ago are considered less important than more recent ones. To incorporate them, we use a dynamically decreasing total number of observations as an indicator, $N_{\text{ind}}^k = \lambda N_{\text{ind}}^{k-1} + N_k$, where $0 < \lambda < 1$ is the *decay rate*. The step size γ_k is defined as the ratio of the size of the new observations to N_{ind}^k , $\gamma_k = N_k / N_{\text{ind}}^k$. This ensures the influence of each new observation diminishes over time as more observations are accumulated, and more recent data have a proportionally higher influence than older ones. This dynamic adjustment helps the model gradually incorporate new information while reducing the weight of the older data points. During the online update, if the new observations do not fit the current model, the number of SWGMM components is increased based on the condition $\sum_{j=1}^J \sum_{i=1}^{N_k} \eta_{i,j} < \eta_{\text{thres}}$, where η_{thres} is a predefined threshold. A smaller η_{thres} makes it easier to classify new observations as not fitting well with the current model, thereby triggering an adjustment in the number of SWGMM components. The updated SWGMM then consists of both previous and new SWNDs, built using recent observations. Their weights are

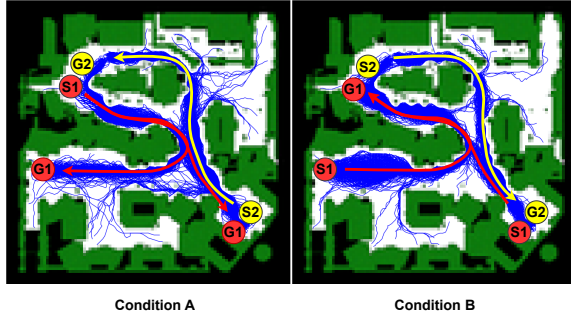


Fig. 2: The *den520d* synthetic dataset simulates trajectories between positions S1, S2, G1 and G2. The initial flow is shown in Condition A (left). To simulate the flow change, the start (S1, S2) and goal (G1, G2) positions are switched Condition B (right), reversing the dominant flow direction.

scaled by γ_k for new SWNDs and $1 - \gamma_k$ for previous ones. All weights are then normalized to integrate the historical and new components within the model.

IV. EXPERIMENTS

A. Datasets

We evaluate the online CLiFF method and a set of baselines on both real-world and synthetic datasets. To reflect the iterative nature of building the maps of dynamics from older and more recent observations, we divide each dataset into batches of trajectory data, corresponding to the input iterations as observed by the robot. The aim is to show how well the most recent map of dynamics, calculated with our method and each baseline, represents the trajectory data from the most recent iteration. The datasets used and how they were divided into batches are described below.

den520d: To evaluate the performance of the online update model in scenarios where human flow changes, we introduce a synthetic dataset *den520d*. This dataset is generated using a map from the Multi-Agent Path-Finding (MAPF) Benchmark [35] and features two distinct flow patterns: Condition A and Condition B, see Fig. 2. We simulate a change in human flow from Condition A to Condition B, where the dominant flow in Condition B is reversed compared to that in Condition A. Furthermore, the *den520d* dataset is used in the downstream robot navigation task (see Sec. IV-E) to assess its practical application in flow-aware planning. We discretize the *den520d* obstacle map with 1 m cell resolution. In this map, randomized human trajectories are simulated using stochastic optimal control path finding, based on Markov decision processes [36]. In each condition 1000 trajectories are generated. We use the condition B for evaluation in a downstream task (robot motion planning) in Fig. 7.

ATC: To evaluate the MoD online update in a real-world dataset with scenarios where human motion patterns change during the day, we use the ATC dataset [20]. Using multiple 3D range sensors, the trajectories were collected between 9 am and 9 pm on Wednesdays and Sundays, including 92 days of observations in total. We use the data from the entire first day for evaluation.

In the experiments across both datasets, the grid resolution for MoD is set to 1 m. The observation rate is downsampled from above 10 Hz to 1 Hz. In our quantitative evaluation, for

each condition in the *den520d* or each hour in the ATC, we randomly sample 10% of the data for testing and use the remaining 90% for training.

B. Baselines

We evaluate several variations of processing the data sequentially using CLiFF-map, and additionally evaluate the prior art STeF-map. The variations in applying CLiFF-map to sequential data include: our online update model (referred to as **online**), a model built from all observations in iteration 1 to k (referred to as **history**) and a model with observations only in iteration k (referred to as **interval**). When training the CLiFF-map, the convergence precision is set to $1e-5$ for both mean shift and EM algorithms, with a maximum iteration count of 100. For online update model, η_{thres} is set to 0.1.

STeF-map [18] is a spatio-temporal flow map, which models the likelihood of human motion directions on a grid-based map using harmonic functions. STeF-map captures long-term changes of crowd movements over time. Each cell in the grid map contains k_{stef} temporal models, corresponding to k_{stef} discretized orientations of people moving through the given cell over time. As suggested in [18], k_{stef} is set to 8 in the experiments and the model orders for training STeF-map, i.e. the number of periodicities, is set to 2. In our experiments, STeF-map is evaluated using the ATC dataset. Due to the characteristics of STeF-map, it can only be applied to data that exhibit meaningful *periodic* time variations and therefore lends itself better to datasets that span multiple days.

C. Metrics

To quantitatively evaluate the accuracy of modeling human motion patterns (MoD model quality), we use negative log likelihood (NLL) as the evaluation metric. The average NLL is computed for the most recent maps of dynamics over the corresponding test portion of the dataset. A lower NLL value indicates better accuracy. If no dynamics pattern is available for a given location – a situation more frequent with interval models, which may fail to capture complete motion patterns due to limited data in the recent iteration – the likelihood is set to a very low value ($1e-9$). For runtime evaluation, we report the runtime across all iterations, measured in minutes.

D. Robot deployment

To verify the method application on a real robot using live input from a human perception stack, we collected motion data of several people in a small indoor environment (40 m^2) in four periods (approx. three minutes of observation each). Each period has a unique motion pattern, e.g. in changing the direction of movement and turns.

E. Downstream task: human-aware motion planning

In order to evaluate the effects of having more accurate maps of dynamics in the downstream task, we use the *den520d* dataset to demonstrate socially-compliant flow-aware motion planning. With the *den520d* map, an A* algorithm [37] modified to use flow-aware costs [13] is used for generating socially-aware paths for agents.

Dataset	Online (ours)	Average NLL		
		Interval	History	STeF-map
ATC	1.18	2.61	1.23	4.38
<i>den520d</i>	1.41	2.36	1.65	-

TABLE I: Evaluation results on the ATC and *den520d* datasets using average negative log likelihood (NLL) as the metric. Lower values are preferable. Values in bold indicate the best performance.

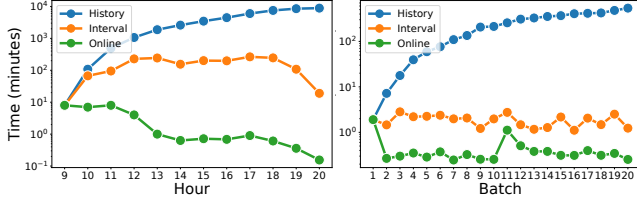


Fig. 3: Runtime of each iteration in the ATC (left) and *den520d* (right) datasets. In *den520d*, the first 10 batches are in Condition A and the second 10 batches are in Condition B. In both datasets, the online model shows significantly reduced runtime compared to the history and interval models.

V. RESULTS

In this section, we present the results obtained in the ATC and *den520d* datasets with our approach compared to baselines. The performance evaluation is conducted using both quantitative and qualitative analysis.

A. Quantitative evaluation

1) *Runtime*: The first experiment evaluates the runtime of the online method, compared to the history model and interval model, in both datasets. We use a desktop computer with an Intel i9-12900K processor running Ubuntu 20.04. As shown in Fig. 3, the online update model exhibits significantly reduced computation time compared to the interval and history models. In the initial iteration, the MoD is empty, and all three models are built from scratch using all available observations. Subsequently, the online update model computes the sufficient statistics with N_k observations in each iteration. In contrast, the interval model runs the EM algorithm to converge (to $1e-5$) using N_k observations, and the history model processes cumulative, $\sum_{k=1}^K N_k$ observations. On average, over all iterations, the online model requires 0.1% of the time used by the history model in ATC and 0.2% in *den520d* dataset.

2) *Accuracy*: The second experiment evaluates the accuracy of modeling human motion patterns. We present the average negative log likelihood (NLL) values, with lower values indicating better modeling accuracy. These values are optimized using the decay rate parameter. Table I shows how well the MoDs reflect changes in human dynamics, using trajectory data from Condition B of the *den520d* dataset and the last hour of the ATC dataset.

Compared to the STeF-map, which discretizes orientations into eight bins, the CLiFF-map provides a continuous probabilistic representation of velocities, achieving better accuracy in modeling human motion patterns. Among variations of CLiFF-map construction, the online update model achieves the best accuracy when the motion patterns change, as it accurately captures these changes while maintaining flow completeness. In contrast, the history model fails to adapt

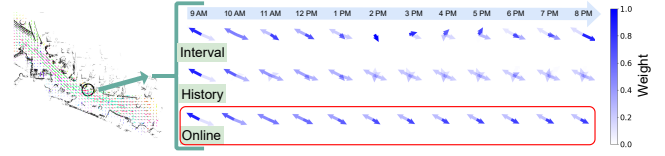


Fig. 4: An example from the east corridor in the ATC dataset. Blue arrows show the mean vectors of SWGMMs, with transparency indicating component weights. The interval model (top row) uses only the last 60 minutes of data, disregarding previously learned patterns (e.g., at 2 PM and 3 PM). The history model (second row), which treats all observations equally, tends to obscure patterns and fails to capture dominant movements effectively. Conversely, the online model (third row) adapts more effectively to changing flows and accurately represents the primary patterns.

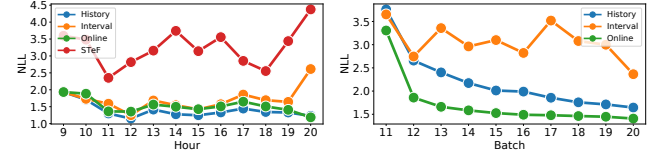


Fig. 5: **Left**: NLL evaluation results at different times of a day in the ATC dataset. The STeF-map struggles to accurately capture human motion flow, resulting in poor performance. Notably, the interval model exhibits worse NLL results mostly after 16:00. It struggles to capture the full spectrum of motion patterns and is not able to effectively manage outlier trajectories outside of the main flow. **Right**: NLL evaluation results over batches in Condition B of the *den520d* dataset. The online model quickly adapts to changes in human flow, achieving lower NLL values in fewer batches.

to changes in human flow, and the interval model suffers from insufficient data within its given interval, leading to several parts of the map having no flow representation, failing to preserve flow completeness. In the ATC dataset, the most notable human motion pattern changes occur in the east corridor. A detailed view of this area is provided in Fig. 4, where the online update model's adaptations are shown, transitioning from predominantly upward movement to primarily downward movement.

Fig. 5 shows the evaluation results for each hour in the ATC dataset and each batch in Condition B of the *den520d* dataset. In the ATC dataset, STeF-map struggles to accurately capture human motion flow, resulting in poor performance. The interval model exhibits worse NLL results, mostly after 16:00, due to its failure to capture the full spectrum of motion patterns. During the middle of the day, when hourly motion patterns are closer to the general motion pattern, the history model performs better by leveraging accumulated data over time. However, without prioritizing recent observations, the history model fails to accurately reflect the most recent motion patterns. In contrast, the online model, with its ability to effectively adapt to changes, achieves better performance in the last hour of the ATC dataset. In the *den520d* dataset, which exhibits a clear flow change from Condition A to Condition B, the history model takes more batches to reflect the flow change, while the online model quickly adapts to changes and achieves lower NLL values in fewer batches.

Parameter analysis: We evaluate the average NLL value with different decay factor values λ , varying from 0.1 to 0.9, in both ATC and *den520d* datasets. The evaluation results are shown in Fig. 6. We observe robust performance across different decay rate values.

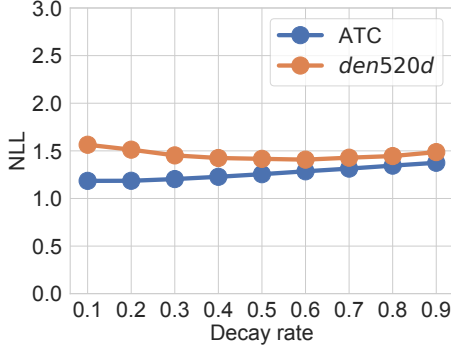


Fig. 6: Evaluation results of the average NLL for the online update models in the ATC and *den520d* datasets using various decay rates, showing robust performance.

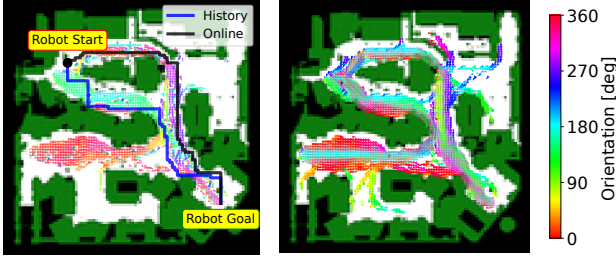


Fig. 7: Flow-aware robot navigation results (left), using the *den520d* map, where we use an A* algorithm [37] modified to use flow-aware costs [13] to generate the agent's path. The MoD of the online model is depicted as an underlying reference for flow direction. Starting from the upper-left corner, the MoD from the history model (right) fails to adapt to flow changes. Consequently, the generated path continues to follow the previous human flow from Condition A, contradicting the current flow in Condition B. In contrast, the online method, which quickly adapts to flow changes, enables the agent to navigate in alignment with the existing human flow.

B. Qualitative evaluation

Fig. 7 shows the CLiFF-maps built with online and history models in Condition B of the *den520d* dataset. When human flow shifts from Condition A to B, the online model quickly accommodates the change, while the history model preserves both flow directions and struggles to adapt. This inaccuracy in the MoD built by the history model affects the downstream task of flow-aware robot navigation; as shown in Fig. 7 (left), the path generated using the history model continues to align with the previous human flow from Condition A, which contradicts the current flow in Condition B, resulting in increased collisions. Conversely, the online method quickly adapts to flow changes, enabling the agent to navigate in a manner that aligns with the current flow pattern.

In the real-robot experiment, we use live input from a human perception stack, with each period showing a unique motion pattern. To demonstrate the runtime bottleneck of baseline methods for real-time applications, we compare runtime on the most recent batch of observations. The history and interval models require 990 s and 29 s, respectively, making them impractical for real-time deployment. In contrast, the proposed online model runs in 5 s, using only 0.5% of the history model's time, making it more suitable for real-time operation. When comparing the quality of the generated MoD, Fig. 8 shows that in the last period, while the history

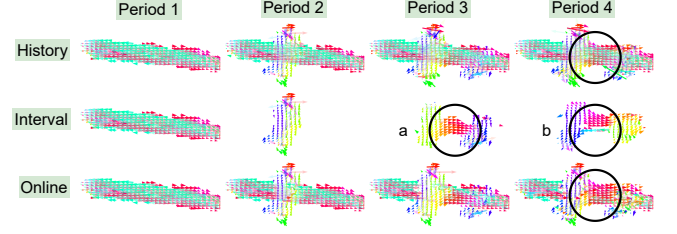


Fig. 8: Models built from real-world data collected from the robot. This experiment highlights that using all historical data typically converges to the possibility of motion in all directions, whereas the online model reflects the dominance of recent observations, while maintaining the completeness of human flow. For instance, the parts of the map influenced by Periods 3 and 4, highlighted by circles *a* and *b*, are captured more accurately by the online model (bottom row), compared to the history model (top row).

model contains motion possibilities in all directions, and the interval model completely forgets previous information and leaves parts of the map of dynamics blank, the online model effectively maintains the dominance of recent observations while keeping the completeness of human flow.

VI. CONCLUSION

In this work we propose a method for online updates of the CLiFF-map of dynamics representation, enabling it to quickly adapt to human flow changes. We use the stochastic expectation maximization algorithm to update the existing representation when new observations are collected. The method is evaluated using both synthetic and real-world datasets, and compared with three baselines. Additionally, we evaluate the method in a downstream task of socially-compliant, flow-aware motion planning. The results demonstrate that our online version can efficiently and accurately accommodate changes in human motion patterns. Furthermore, it is significantly faster than the baselines and avoids the large memory consumption associated with storing the entire dataset of historic observations.

The proposed method has limitations. As a result of our experiments with various modes of tracking the dynamics observation history, we notice that maintaining an up-to-date MoD is a challenging problem due to the need to balance the historic patterns vs. deciding when the data is outdated [38]. While the proposed online update method significantly improves runtime, making it practical for real-time robotics applications, it may not always deliver optimal performance. In situations where dynamic patterns and flow of motion in the environment are complex, historical models can achieve better accuracy as they are better equipped to handle outlier trajectories that deviate from the main flow pattern. Conversely, when the human motion flow is more laminar and there is sufficient current observational data to accurately represent human flow, the interval model may capture the most accurate patterns of human motion and can serve as a benchmark for evaluation. In these instances, the online update methods may perform less effectively compared to the interval model.

In future work we intend to use the online update method as an initial building block to enable efficient robot exploration in dynamic environments.

REFERENCES

- [1] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch. “Human-aware robot navigation: A survey”. In: *J. of Robotics and Autonomous Systems* 61.12 (2013), pp. 1726–1743.
- [2] R. Triebel et al. “Spencer: A socially aware service robot for passenger guidance and help in busy airports”. In: *Field and service robotics*. Springer. 2016, pp. 607–622.
- [3] C. Mavrogiannis, F. Baldini, A. Wang, D. Zhao, P. Trautman, A. Steinfeld, and J. Oh. “Core challenges of social robot navigation: A survey”. In: *ACM Transactions on Human-Robot Interaction* 12.3 (2023), pp. 1–39.
- [4] P. A. Lasota, T. Fong, and J. A. Shah. “A Survey of Methods for Safe Human-Robot Interaction”. In: *Foundations and Trends in Robotics* 5.4 (2017), pp. 261–349.
- [5] L. Palmieri, T. P. Kucner, M. Magnusson, A. J. Lilienthal, and K. O. Arras. “Kinodynamic motion planning on Gaussian mixture fields”. In: *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*. IEEE. 2017, pp. 6176–6181.
- [6] C. S. Swaminathan, T. P. Kucner, M. Magnusson, L. Palmieri, S. Molina, A. Mannucci, F. Pecora, and A. J. Lilienthal. “Benchmarking the utility of maps of dynamics for human-aware motion planning”. In: *Frontiers in Robotics and AI* 9 (2022). ISSN: 2296-9144.
- [7] J. Doellinger, M. Spies, and W. Burgard. “Predicting occupancy distributions of walking humans with convolutional neural networks”. In: *IEEE Robotics and Automation Letters* 3.3 (2018).
- [8] T. P. Kucner et al. “Survey of maps of dynamics for mobile robots”. In: *Int. J. of Robotics Research* (2023).
- [9] Y. Zhu, A. Rudenko, T. P. Kucner, L. Palmieri, K. O. Arras, A. J. Lilienthal, and M. Magnusson. “CLIFF-LHMP: Using Spatial Dynamics Patterns for Long-Term Human Motion Prediction”. In: *Proc. of the IEEE Int. Conf. on Intell. Robots and Syst. (IROS)*. 2023.
- [10] T. R. de Almeida et al. “THÖR-Magni: Comparative Analysis of Deep Learning Models for Role-Conditioned Human Motion Prediction”. In: *Proc. of the Int. Conf. on Comp. Vision Worksh.* 2023.
- [11] T. Krajník, J. P. Fentanes, J. M. Santos, and T. Duckett. “FreMEn: Frequency Map Enhancement for Long-Term Mobile Robot Autonomy in Changing Environments”. In: *IEEE Trans. on Robotics (TRO)* 33.4 (2017), pp. 964–977.
- [12] S. Molina et al. “The ILIAD safety stack: human-aware infrastructure-free navigation of industrial mobile robots”. In: *IEEE Robotics & Automation Magazine* (2023).
- [13] C. S. Swaminathan, T. P. Kucner, M. Magnusson, L. Palmieri, and A. J. Lilienthal. “Down The CLIFF: Flow-aware Trajectory Planning under Motion Pattern Uncertainty”. In: *Proc. of the IEEE Int. Conf. on Intell. Robots and Syst. (IROS)*. IEEE. 2018, pp. 7403–7409.
- [14] Y. Liu, L. Palmieri, I. Georgievski, and M. Aiello. “Human-flow-aware long-term mobile robot task planning based on hierarchical reinforcement learning”. In: *IEEE Robotics and Automation Letters* (2023).
- [15] T. P. Kucner, M. Magnusson, E. Schaffernicht, V. H. Bennetts, and A. J. Lilienthal. “Enabling Flow Awareness for Mobile Robots in Partially Observable Environments”. In: *IEEE Robotics and Automation Letters* 2.2 (2017), pp. 1093–1100.
- [16] T. Schreiter et al. *THÖR-MAGNI: A Large-scale Indoor Motion Capture Recording of Human Movement and Robot Interaction*. 2024. arXiv: 2403.09285 [cs.RO].
- [17] T. P. Kucner, A. J. Lilienthal, M. Magnusson, L. Palmieri, and C. S. Swaminathan. *Probabilistic mapping of spatial motion patterns for mobile robots*. Springer, 2020.
- [18] S. Molina, G. Cielniak, T. Krajník, and T. Duckett. “Modelling and Predicting Rhythmic Flow Patterns in Dynamic Environments”. In: *Annual Conf. Towards Autonom. Rob. Syst.* Springer. 2018, pp. 135–146.
- [19] S. Molina, G. Cielniak, and T. Duckett. “Robotic Exploration for Learning Human Motion Patterns”. In: *IEEE Trans. on Robotics and Automation (TRO)* (2022).
- [20] D. Bršćić, T. Kanda, T. Ikeda, and T. Miyashita. “Person tracking in large public spaces using 3-D range sensors”. In: *IEEE Trans. on Human-Machine Systems* 43.6 (2013), pp. 522–534.
- [21] J. Saarinen, H. Andreasson, and A. J. Lilienthal. “Independent Markov chain occupancy grid maps for representation of dynamic environment”. In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2012, pp. 3489–3495.
- [22] W. Burgard, C. Stachniss, and D. Hähnel. “Mobile robot map learning from range data in dynamic environments”. In: *Autonomous Navigation in Dynamic Environments*. Springer, 2007, pp. 3–28.
- [23] M. Bennewitz, W. Burgard, G. Cielniak, and S. Thrun. “Learning motion patterns of people for compliant robot motion”. In: *Int. J. of Robotics Research* 24.1 (2005), pp. 31–48.
- [24] V. H. Bennetts, T. P. Kucner, E. Schaffernicht, P. P. Neumann, H. Fan, and A. J. Lilienthal. “Probabilistic Air Flow Modelling Using Turbulent and Laminar Characteristics for Ground and Aerial Robots”. In: *IEEE Robotics and Automation Letters* 2.2 (2017), pp. 1117–1123.
- [25] Y. Zhu, H. Fan, A. Rudenko, M. Magnusson, E. Schaffernicht, and A. J. Lilienthal. “LaCE-LHMP: Airflow Modelling-Inspired Long-Term Human Motion Prediction By Enhancing Laminar Characteristics in Human Flow”. In: *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*. 2024.

- [26] J. Doellinger, V. S. Prabhakaran, L. Fu, and M. Spies. “Environment-Aware Multi-Target Tracking of Pedestrians”. In: *IEEE Robotics and Automation Letters* 4.2 (2019), pp. 1831–1837.
- [27] F. Verdoja, T. Kucner, and V. Kyrki. “Generating people flow from architecture of real unseen environments”. In: *PNARUDE Workshop: Perception and Navigation for Autonomous Robotics in Unstructured and Dynamic Environments*. 2022.
- [28] Y. Zhu, A. Rudenko, T. P. Kucner, A. J. Lilienthal, and M. Magnusson. “A Data-Efficient Approach for Long-Term Human Motion Prediction Using Maps of Dynamics”. In: *ICRA Worksh. (LHMP)*. 2023.
- [29] T. R. de Almeida, Y. Zhu, A. Rudenko, T. P. Kucner, J. A. Stork, M. Magnusson, and A. J. Lilienthal. “Trajectory Prediction for Heterogeneous Agents: A Performance Analysis on Small and Imbalanced Datasets”. In: *IEEE Robotics and Automation Letters* (2024), pp. 1–8.
- [30] O. Cappé and E. Moulines. “On-Line Expectation–Maximization Algorithm for latent Data Models”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 71.3 (2009), pp. 593–613.
- [31] A. Roy, S. K. Parui, and U. Roy. “SWGMM: a semi-wrapped Gaussian mixture model for clustering of circular-linear data”. In: *Pattern Anal. Appl.* 19.3 (2016), pp. 631–645.
- [32] K. V. Mardia and P. E. Jupp. *Directional Statistics*. Wiley, 2008.
- [33] A. P. Dempster, N. M. Laird, and D. B. Rubin. “Maximum Likelihood from Incomplete Data Via the EM Algorithm”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 39.1 (1977), pp. 1–38.
- [34] Y. Cheng. “Mean shift, mode seeking, and clustering”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1995).
- [35] R. Stern et al. “Multi-Agent Pathfinding: Definitions, Variants, and Benchmarks”. In: *Symposium on Combinatorial Search (SoCS)* (2019), pp. 151–158.
- [36] A. Rudenko, L. Palmieri, J. Doellinger, A. J. Lilienthal, and K. O. Arras. “Learning Occupancy Priors of Human Motion From Semantic Maps of Urban Environments”. In: *IEEE Robotics and Automation Letters* 6.2 (2021), pp. 3248–3255.
- [37] P. E. Hart, N. J. Nilsson, and B. Raphael. “A formal basis for the heuristic determination of minimum cost paths”. In: *IEEE transactions on Systems Science and Cybernetics* 4.2 (1968), pp. 100–107.
- [38] P. Biber and T. Duckett. “Dynamic Maps for Long-Term Operation of Mobile Service Robots”. In: *Robotics: Science and Systems*. 2005.