Abstract—UAVs (unmanned air vehicles) have contributed greatly to situational awareness through surveillance missions. However, complete autonomy of a UAV has yet to be realized due to the lack of reliable onboard sensing capabilities. Current research at the Air Force pairs UAVs with Unattended Ground Sensors (UGS) to create a system that autonomously patrols, detects, and isolates intruders on a road network. During the patrol phase a UAV must visit each UGS, transitioning to the isolation phase if an intruder is detected. Optimizing the UAV flight plan during the patrol phase will lead to a faster response time and higher probability of capturing the intruder. The goal of this work is to investigate a learning-based approach that will enable more efficient and effective surveillance operations. In the proposed framework, techniques from adaptive feedforward iterative learning control and a region of attraction-based tracking approach have been used to optimize the UAV flight plan between surveillance flights. The proposed approach has resulted in the development of a novel learning control framework that leverages a region-based tracking requirement to minimize overall distance traveled, while guaranteeing convergence within the required tracking zone. Simulation results for a 1D example system demonstrate the validity of the control framework through an 8% reduction of overall distance traveled as compared to traditional surveillance strategies requiring tracking convergence to a single point. The proposed framework has the potential to significantly decrease the resources required for surveillance-based applications.

I. INTRODUCTION

UAVs (unmanned air vehicles) and other airborne or land-based surveillance systems have contributed greatly to situational awareness through surveillance missions [1]. A key limitation to these types of unmanned surveillance systems is the lack of truly autonomous UAVs due to limited reliable onboard sensing capabilities. Autonomous surveillance vehicles are often equipped with very limited sensor capabilities in order to minimize weight and energy requirements [2]. In contrast, unattended ground sensors (UGSs) have the capacity to capture a wide range of information; however, the transfer capabilities are often limited to relatively small distances. Recent surveillance strategies from the Air Force combine a UAV and multiple UGSs into a single cohesive unit in order to leverage the attributes of each system [3], [4]. Combining these two units into a single system results in a cooperative surveillance scenario in which the UAVs rely on information from the UGSs in order to detect and respond to an intruder.

This work was supported by the American Society for Engineering Education under the Air Force Summer Faculty Fellowship Program 2012. K. Barton is with University of Michigan, Ann Arbor, MI 48109, e-mail: BartonK@umich.edu. D. Kingston is with the Wright Patterson Air Force Base Vehicles Division, 2210 8th St., B20146 R301, Wright Patterson, OH 45433, e-mail: derek.kingston@wpafb.af.mil

Fig. 1. Diagram of traditional approach to UAV/UGS surveillance.

Traditional approaches for optimizing the UAV route between subsequent UGS locations have utilized a traveling salesman strategy for optimizing the order of the visits between UGSs. As a means of guaranteeing information transfer, the UAV is instructed to fly directly over the center point location of each UGS, see Figure 1. Although this path guarantees communication and therefore sensor information hand-off between the UAV and UGS, it is a conservative approach that does not take advantage of the larger communication range of the UGS. If a region of successful communication transfer around an UGS is known, then an optimized surveillance flight that only requires the UAV to fly within the region around an UGS can be determined.

The goal of this research is to develop a learning-based approach for the identification of a region-based map of the UGSs and the optimization of the UAV patrol path. This is a two-part problem that will be addressed in phases. This paper presents phase I: the development of a novel control framework for optimizing path planning that will significantly decrease the resources currently required for surveillance-based applications. Given the repetitive nature of surveillance, performance improvements can be realized by utilizing novel control approaches to leverage the information collected by the system during routine patrol missions. Specifically, the proposed framework will incorporate theory and techniques from the field of iterative learning control (ILC). ILC is a high performance control technique that updates an input signal based on information from the measured error. The novel contribution from this work is a region-to-region ILC strategy that combines learning and a region-of-attraction concept into a single framework.

This work provides a unique approach to UAV path planning that can serve as a stand-alone technique or an augmentation on an existing control strategy. This research represents the first step towards the use of a learning-based
feedforward control technique with UAVs for enhanced situational awareness.

II. UAV/UGS Scenario Description

A. Scenario Description

The limited onboard sensing capabilities of many UAVs motivates recent proposed scenarios in which a team of UAVs query a set of unattended ground sensors (UGSs) located on a road network (see Fig. 2). Many current UGS systems require line-of-sight between UGSs or a ‘gateway’ node for networking. This constrains the system by requiring careful setup of the sensor network and a constant power source. The UAV-UGS Road Surveillance (U2R2) problem combines UAVs and UGSs to relax these constraints. The combined system improves surveillance of the road network by enhancing both intruder detection and information flow about potential threats.

![Fig. 2. Visual representation of the UGS/UAV road surveillance problem (U2R2) proposed by the Air Force Research Lab (AFRL).](image)

In this scenario, UAVs are used to query the UGSs and deliver data to an operator [4]. This allows both a higher sensor density and UGS placement in areas with difficult line-of-sight between sensors. The four phases of the U2R2 problem include 1) patrol, 2) isolation, 3) delivery and classification, and 4) response. This work focuses on the patrol phase in which a UAV repetitively visits and queries each UGS to collect surveillance information.

B. Current Approach: Traveling Salesman Problem

The goal of the patrol phase is to efficiently and effectively query the network of UGSs. The key challenge in this phase consists of determining the optimized flight pattern to achieve this goal. Traditionally, optimized flight patterns have been identified by posing the patrol phase as a traveling salesman problem (TSP), in which a UAV is required to visit each UGS only once, starting from any UGS and returning to the original place of departure (denoted as Home Base in Figure 1). TSP has been studied since the 1930s and serves as a benchmark for many optimization problems [5]. Although computer technology has continued to improve, the number of locations that must be visited and additional system constraints, such as designated length of each visit, lead to increased computational complexity and run time.

Despite the computation constraints, there exist several TSP approaches that are capable of solving large optimization problems [5], [6]. These approaches have been used in the UAV/UGS scenario to provide the shortest possible route that visits each designated location exactly once and returns to the original starting point [7]. The route is comprised of the appropriate ordering, with no UAV heading or orientation information. Smooth flight patterns that result in the UAV passing directly over the center of the UGS are combined with a TSP in order to generate flight patterns for patrol (Figure 1). These patterns are inherently conservative and do not utilize the sensor communication range of the UGSs. Recent work by Obermeyer [7] introduced a polygon visiting TSP approach in which an optimized path through a set of polygons was determined. Obermeyer’s approach focused on regions associated with visual sensor data. One critical limitation of this approach is the inability of the framework to account for potential changes in the assumed regions. The proposed approach outlined here seeks to address this limitation by combining an iterative learning controller that incorporates updated sensor information from the patrols with an ordered TSP solution.

C. Proposed Approach: Region-to-Region ILC

Surveillance and information gathering requires a UAV to be within communication range of the UGS. The region for successful communication varies for each UGS based on: the orientation of the UGS, the topology of the local environment, battery life, type of sensors, etc. Because the communication range cannot be determined a priori, the current practice is to design optimized surveillance paths that require the UAVs to fly to locations directly over the center of each UGS. This approach can lead to unacceptable time requirements for time-sensitive operations.

Region-to-region ILC provides an alternative approach through a unique combination of route flexibility between the UGSs and learning. Region-to-region ILC enables the controller to optimize the path between UGSs by utilizing information from previous passes. This information can be used to extend the critical UAV fly-over locations from a central point directly above the UGS to a region of communication surrounding the UGS. In this manner, an optimized surveillance path that targets each communication region can be determined. Figure 3 presents an illustrative comparison between the current and proposed surveillance approaches. Note that the proposed region-based ILC approach results in modified surveillance patterns aimed at reducing the overall distance traveled.

III. Iterative Learning Control

As a means of leveraging the repetitive nature of surveillance, the novel control approach described in this work incorporates iterative learning control (ILC) concepts into
a modified learning framework. Importantly, the learning applies to the path planning aspect of the UAV/UGS surveillance problem rather than the tracking performance of the UAVs. ILC is a high performance control technique that seeks to improve system performance along the time and iteration domain. ILC is an anticipatory technique that provides an updated feedforward signal that enters the system at the control input or with the reference signal. Closing the loop from iteration-to-iteration, ILC acts as a feedback controller in the iteration domain.

ILC has been successfully applied to repetitive applications in robotics [8], [9], manufacturing [10]–[12] and chemical processing [13], [14]. In these applications, ILC was implemented to minimize the trajectory tracking errors of the system through iterative updates to the control signal. Additional details about ILC can be found in [15]–[17].

A. Norm Optimal Framework

This research utilizes the time-domain or norm optimal approach, in which the discrete-time behavior of the system and the relevant signals are represented by matrices and vectors containing the stacked time-domain data from one surveillance cycle. For the norm-optimal approach, the system is represented by a convolution matrix, \( P \), using impulse response data \( H_{ii,jj} \).

\[
P = \begin{bmatrix}
  H_{0,0} & 0 & \cdots & 0 \\
  \vdots & \ddots & \ddots & \vdots \\
  0 & \cdots & H_{N-1,0} & \cdots & H_{N-1,N-1}
\end{bmatrix}
\tag{1}
\]

For multi-input multi-output (MIMO) linear time-varying (LTV) systems, \( H_{ii,jj}(k) \) contains the impulse response from each of the \( q_i \) inputs to each of the \( q_o \) outputs and can be derived using time-varying matrices.

\[
H_{ii,jj} = \begin{cases} 
  D(ii), & ii = jj \\
  C(ii)A(ii-1)\ldots A(jj+1)B(jj), & ii > jj
\end{cases}
\tag{2}
\]

In Eqn. (2), \( (A(k), B(k), C(k), D(k)) \) are appropriately sized iteration-invariant real-valued matrices used to describe the dynamic behavior of the system, \( P \), and \( k = 0, 1, \ldots, N - 1 \) is the discrete time index.

B. ILC Update Law

The common time-domain update law for ILC is derived from a quadratic optimization problem in which the objective is to minimize a cost function [18],

\[
J = e_{j+1}^T Q e_{j+1} + u_{j+1}^T S u_{j+1} + (u_{j+1} - u_j)^T R (u_{j+1} - u_j).
\tag{3}
\]

In Eqn. (3), \((Q, R, S)\) are symmetric positive definite weighting used to design the tradeoff between performance, convergence rate, and robustness. The common form of these matrices is given as \((Q, R, S) \triangleq (qI, rI, sI)\).

The ILC update law is given as [18], [19],

\[
\begin{align*}
  u_{j+1} &= L_u u_j + L_e e_j 
\end{align*}
\tag{4}
\]

where

\[
\begin{align*}
  e_j &= r - y_j \\
  y_j &= Tr + Pu_j.
\end{align*}
\tag{5}
\]

In Eqs. (4)-(6), \( e_j \) is the tracking error signal between the reference \( r \) and the system output \( y_j \). \( T \) is the complementary sensitivity function, \( P \) is given in (1), and \( j = 0, 1, \ldots \) is the iteration index.

The design elements in the update law include the \( L_u \)-filter and \( L_e \)-filter which are functions of the weighting matrices and the plant dynamics. These filters are of the form,

\[
\begin{align*}
  L_u &= (P^T Q P + S + R)^{-1}(P^T Q P) \\
  L_e &= (P^T Q P + S + R)^{-1}P^T Q.
\end{align*}
\]

\( L_u \) is generally designed as a low-pass filter to limit the learning bandwidth and provide robustness to the system. \( L_e \), also known as the learning filter, is designed to maximize the learnable bandwidth and convergence rate.

C. Convergence

The goal in designing \( L_u \) and \( L_e \) is to ensure contraction mapping or monotonic convergence of the control signal. Substituting Eqs. (5) and (6) into Eqn. (4) and rearranging the terms results in the following control iteration dynamics.

\[
\begin{align*}
  u_{j+1} &= (L_u - L_e P)u_j + L_e (I - T)r 
\end{align*}
\tag{7}
\]

A contraction mapping from \( u_{j+1} \) to \( u_j \) is accomplished by ensuring,

\[
\|(L_u - L_e P)\|_2 < 1, \tag{8}
\]

where the induced two norm \( \| \cdot \|_2 \) is defined as

\[
\| \cdot \|_2 \triangleq \max |\cdot|.
\]

Satisfying Eqn. (8) guarantees monotonic convergence of the asymptotic control \( u_\infty \) and error signals \( e_\infty \), respectively.

IV. REGION-TO-REGION ITERATIVE LEARNING CONTROL

In many applications [8], [10], the desired objective consists of static predetermined trajectories defined over the entire cycle time interval. Alternatively, there exist applications in which the control objective can be simplified to minimizing the tracking error at critical points along the trajectory (e.g. pick and place robots). For these systems, the tracking between points is not critical and can vary from iteration to iteration.
Fig. 4. A) Illustration of a point-to-point tracking problem. The control objective requires precise tracking only at the critical points. Tracking between regions is subject to system constraints and competing performance objectives.

Although point-to-point ILC provides improved performance over traditional ILC, it may result in a conservative approach for those applications that require convergence within a given region, rather than a specific set point. Relaxing the tracking constraint from a set point to any point within a given critical region provides flexibility for the optimization of additional performance objectives such as speed, energy consumption, and distance traveled.

A. Region-to-Region ILC

In this section we present the novel learning framework for region-based tracking. Region-to-region ILC (r2r-ILC) is a generalized form of point-to-point ILC in which the performance objective is described in terms of a designated region, see Figure 4 B). The output points of interest are determined through a projection operator, ψr, which is used to map the full error signal into the designated region. The critical error points are defined as, 

\[ ε_{r,j+1} = ψ_r \cdot e_{j+1} \]  

where \( ε_{r,j+1} \) is the projected error signal. Equation (9) relies on the following assumption:

Assumption 1 (Boundedness of ε): The parameter vector \( ε \) belongs to a given compact convex set \( ε \in D \subseteq R^n \). In assumption 1, D is the specified region in which the system must converge to meet the performance specifications.

In order to leverage additional design flexibility, the problem can be constructed as a pareto optimization problem in which there are multiple competing objectives. The time-domain cost function for r2r-ILC can be minimized by a quadratic optimization:

\[ J = e_{j+1}^T ψ_r^T Q ψ_{e_{j+1}} + u_{j+1}^T S u_{j+1} + Δu_{j+1}^T R Δu_{j+1} + (\cdot)^T W_1(\cdot) + (\cdot)^T W_2(\cdot) + ... \]  

where the competing objectives are included in the cost function as additional weighted signals. This results in a modified update law (Eqn. (11)) with additional linear operators acting on the competing objectives. For additional details describing pareto-based ILC, see [22]. The modified update law is given as,

\[ u_{j+1} = L_u u_j + L_e e_j + \sum L(\cdot) \]  

where 

\[ L_u = Z^{-1} (P^T ψ_r^T Q ψ_r P + R + \sum P^T M_i^T W_i M_i P) \]  
\[ L_e = Z^{-1} (P^T ψ_r^T Q ψ_r) \]  
\[ L(\cdot) = Z^{-1} (P^T M_i^T W_i). \quad ∀i = 1, 2, ... \]

In the definitions of \( L_u - L(\cdot), W_i \) is a positive definite weighting matrix of appropriate dimensions used to weight the additional performance metrics. \( M_i \) is a mapping matrix used to map the output signal \( y_j \) to the \( i = 1, 2, ... n \) competing objectives.

B. Control Design Considerations

For the cost function and update law described in Eqns. (10) and (11), there are performance tradeoffs between the designs of the Q and W_i weighting matrices due to the competing objectives. For example, in the UAV/UGS scenario the primary goal is to optimize the patrol path, while guaranteeing sufficient signal strength for data transfer. In this manner, path optimization would take precedence over signal strength within an allowable range. Once the signal strength fell outside the allowable range, a penalty would be incurred by the control algorithm, leading to a more conservative surveillance path. In order to leverage the design flexibility and expand the control search for more aggressive surveillance flights, a missed signal from a single UGS would be tolerated within the design. A missed signal in one path would contribute to the overall environmental information. And, such a miss would be corrected in subsequent surveillance flights.

Through the designs of the linear operators, \( L(\cdot), L_u \) and \( L_e \), the values which fall within the regions of interest and the signals associated with the additional performance metrics are used to update the input signal, resulting in a multi-objective, region-to-region ILC design framework.

V. SIMULATION RESULTS

To validate the feasibility of the region-to-region ILC approach, consider the following simplified 1-dimensional (1D) example.

A discrete-time system is given as,

\[ G(z) = \frac{z - 0.5}{(z - 1)(z - 0.925)}. \]  

Discrete-time is assumed because ILC requires the storage of signals, which is typically done digitally. Here, G represents
Fig. 5. Simulation trajectory with 5 designated regions of interest.

a servo-positioning system with viscous friction. The system is stabilized with a proportional feedback controller,

\[ C(z) = 0.425. \]  \hspace{1cm} (13)

The single-input single-output (SISO) system was tasked with visiting five designated locations within a given time frame. To further simplify the problem, the five designated locations were assigned specific time stamps. The 1D regions were then defined as symmetric lines emanating along the 1D position axis from the center point of a designated location, corresponding to design flexibility along the position axis, not the time axis (see Figure 5).

A. Simulation Scenario

The objective of the simulation was to minimize the overall distance traveled along the position axis, while guaranteeing that the system passed through each region of interest. In this manner, the optimization aimed to guarantee that the system passed within the regions during the five specified times. To satisfy these requirements, a pareto optimization was constructed based on the total distance traveled and a weighting applied to locations outside the associated regions of interest during the designated time stamps.

The following assumptions were made to utilize the region-to-region iterative learning scheme:

A1) The center point locations of all regions of interest were known.
A2) Updated trajectory paths existed within the span of realizable trajectories.
A3) The regions of communication were 1-unit length lines emanating from the center points of the regions of interest.
A4) Identical start (home base) position from pass-to-pass.
A5) Identical ordering of designated location visits from pass-to-pass.

Assumptions A1-A3 were required to ensure realizable updated trajectory paths. Assumptions A4-A5 were required for the learning algorithm. Maintaining consistent ordering of the designated location visits, as well as the same start position, provided output data that could be directly compared to previous passes for learning.

For the research presented here, the region of interest corresponded to UGS signal strength. More specifically, signal strength was assumed to be a function of the distance between the system and the center point of the designated region (i.e., a line in the 1D case). This distance provided a measure of the relative strength of the communication signal between the two. The signal could be detected a maximum distance, \( \bar{x}_{max} \), from the center point. To simplify the derivation, the strength was assumed to be constant within a given distance epsilon from the center point. Mathematically, the signal strength can be defined as:

\[ ss_j(k) = d(k) \cdot \bar{x}_j(k) \]  \hspace{1cm} (14)

where

\[ \bar{x}_j(k) = r_{pp} - x_j(k) \]

\[ d(k) = \begin{cases} \frac{1}{\epsilon} & \bar{x} \leq \epsilon \\ \frac{1}{\bar{x}} & \epsilon < \bar{x} \leq \bar{x}_{max} \\ 0 & \bar{x} > \bar{x}_{max} \end{cases} \]

In the distance equation, \( r_{pp} \) denotes the designated region at varying locations, \( pp = 1, 2, \ldots 5 \). Initial simulation results assumed perfect signal transmission from anywhere within the linear signal transmission range. Additionally, the region of communication was assumed to be known and constant for all UGSs. Future work will focus on relaxing these assumptions to investigate a learning-based approach to mapping unknown regions of communication (e.g., region shape, as well as signal strength within the region) for a given topography and UAV/UGS scenario.

The total distance traveled denotes the minimum distance required to visit each region within the network.

\[ t_j = \sum (x_j(k) - x_j(k-1)), \ \forall k = 1 \ldots N \]  \hspace{1cm} (15)

B. Design Tradeoff: Efficient Patrol vs. Signal Strength

In this simulation example, there was a tradeoff between minimizing the total distance traveled and maximizing the signal strength. For the proposed scenario, the primary goal was to minimize the total distance traveled, while guaranteeing passage through the region of interest at the specified times. Once the signal strength fell outside of the allowable range (i.e., outside the defined region), a penalty was incurred by the control algorithm, leading to a more conservative trajectory.

C. Results

In this scenario, only movements along a single axis were considered. Five critical locations were identified at specific time and position points within the cycle time. The projection operator for this simple 1D example was a diagonal matrix with ones in the entries corresponding to the time indices for the five critical regions. The region-based error points were a function of signal strength rather than tracking position to simulate the UAV/UGS sensor transmission signal. The primary performance metric was to minimize the total distance traveled, \( t_j \).

The tuning parameters used to trade-off between performance objectives were the \( \{q, w\} \) gains associated with the \( Q, W \) weighting matrices applied to the signal strength and overall distance traveled, respectively. The traditional center-point approach focused on maximizing the signal strength.
TABLE I
WEIGHTING GAINS FOR THE LEARNING CONTROLLERS

<table>
<thead>
<tr>
<th>Controller</th>
<th>q</th>
<th>w</th>
<th>r</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center-point</td>
<td>1</td>
<td>1</td>
<td>1e-8</td>
<td>1e-8</td>
</tr>
<tr>
<td>Region-to-region</td>
<td>1</td>
<td>12</td>
<td>1e-8</td>
<td>1e-8</td>
</tr>
</tbody>
</table>

Fig. 6. Simulation results of a 1D scenario with 5 regions of interest.

Future efforts will extend the framework to a two-dimensional scenario, in which the nonlinear performance objectives pose an additional design challenge for a linear learning control framework. Additionally, phase II of this research problem will explore the effect of non-symmetric regions and present an approach for updating the regions of communication based on learned information from sensor data collected during routine surveillance missions.

REFERENCES