Local Track Repair for Video Tracking on Small UAVs

Stephen DelMarco[1], Mathew Antone[2], Austin Reiter[3], Todd Jenkins[4]
[1,2,3]BAE Systems Inc., 6 New England Executive Park, Burlington, MA. USA. 01803,

[4]Air Force Research Laboratory, Sensors Directorate, WPAFB, OH. USA.

ABSTRACT

Persistent aerial video surveillance from small UAV (SUAV) platforms requires accurate and robust target tracking capabilities. However, video tracks can break due to excessive camera motion, target resolution, low signal-to noise ratio, video frame dropout, and frame-to-frame registration errors. Connecting broken tracks (*video track repair*) is thus essential for maintaining high quality target tracks. In this paper we present an approach to track repair based on multi-hypothesis sequential probability ratio tests (MHSPRT) that is suitable for real-time video tracking applications. To reduce computational complexity, the approach uses a target dynamics model whose state estimation covariance matrix has an analytic eigendecomposition. Chi-square gating is used to form feasible track-to-track associations, and a set of local hypothesis tests is defined for associating new tracks with coasted tracks. Evidence is accumulated across video frames by propagating posterior probabilities associated with each track repair hypothesis in the MHSPRT framework. Global maximum likelihood and maximum *a posteriori* estimation techniques resolve conflicts between local track association hypotheses. The approach also supports fusion of appearance-based features to augment statistical distributions of the track state and enhance performance during periods of kinematic ambiguity. First, an overview of the video tracker technology is presented. Next the track repair algorithm is described. Finally, numerical results are reported demonstrating performance on real video data acquired from an SUAV.

Keywords: Video Tracking, UAV, Multi-Hypothesis Test

1. INTRODUCTION

Modern video camera systems are proliferating on small UAV (SUAV) platforms due to low cost, high image quality, and ease of data interpretation by human operators. Often in these surveillance systems, human operators visually nominate and track targets-of-interest. However, continual manual visual target tracking by an operator can reduce operator efficiency and distract from other critical activities. Automated video tracking systems can reduce operator workload by maintaining track on targets-of-interest.

The effectiveness of airborne video surveillance system depends on the quality and duration of these video tracks. A persistent problem with automated video trackers is that tracks may experience brief breaks due to excessive camera motion, short occlusions, field-of-view issues, and signal-to-noise ratio and resolution effects. **Local video track repair** (also called **track stitching**) is the operation of connecting broken tracks across these brief track drops. Local video track repair can extend mean track duration thus facilitating surveillance activities.

[1] stephen.delmarco@baesystems.com
[2] matthew.antone@baesystems.com
[3] austin.reiter@baesystems.com
[4] Todd.Jenkins@wpafb.af.mil

This work was sponsored by the Air Force Research Laboratory under Air Force Contract Number FA8650-04-C-1681.

Video tracks may be presented directly to the operator for evaluation. Alternatively, video tracks can serve as inputs to higher level trackers such as Multi-Hypothesis Trackers (MHT) [13], as in [5,6]. MHT trackers provide long term tracking, tracking in environments with high motion clutter, and tracking across extended occlusions. Local video track repair provides an MHT tracker with higher quality, coherent track data input thus leading to overall tracker performance improvement.

Simple track stitching approaches may be adequate for tracking in low target density environments. However, in complex tracking scenarios, multiple targets may gate with dropped track states, including false-alarm tracks resulting from clutter or noise detections. In these tracking environments, track stitching accuracy performance may benefit from a multi-frame approach in which track association evidence is collected across frames. Evidence accrual across video frames provides a more consistent measure of state dynamics similarity. This improved matching consistency can reduce stitching errors associated with coincidental track state alignment which may occur during a single frame. Such coincidental state alignment can occur, for example, when measurement noise obscures the separation of track states in state space. Collecting data across frames also provides the opportunity for false alarm tracks to drop, or for their state dynamics to diverge significantly from target tracks. Furthermore, processing across frames also mitigates stitching errors arising from variable track confirmation times which can cause staggered track initiation times.

For the track stitching problem, evidence accrual across frames can be cast as a sequential probability ratio test (SPRT) problem. For binary hypothesis testing problems, the SPRT has been shown to be optimal [14] in terms of providing the minimum expected number of sequential steps needed to make a decision, subject to pre-specified risk constraints. For binary hypothesis testing problems, the SPRT consists of a sequential test on the likelihood ratio. Given pre-specified risk tolerances, low and high likelihood ratio thresholds are defined, and data is collected sequentially. After each data collection step, the likelihood ratio is tested against each threshold. The null or alternative hypothesis is chosen at the first time that the likelihood ratio falls below the low threshold or rises above the high threshold, respectively.

The SPRT has been extended to the multihypothesis case thereby creating a multihypothesis sequential probability ratio test (MHSPRT). Discussion of MHSPRT development work is presented in [1]. As noted in [2], work on multiple hypothesis sequential testing has generally proceeded in two general directions. First, efforts have been made to determine optimal tests, usually under simplifying assumptions. For example work in [2] presents an optimal test under a Bayesian framework using a dynamic programming argument. Simplifying assumptions include independent and identically distributed (*i.i.d.)* random variables, constant cost per time step, and a zero-one cost function on the decision rule. However, even with simplifying assumptions, the optimal test can be complicated and difficult to implement in practice. In fact, the optimal solution to the general MHSPRT problem may not exist, or may be very difficult to determine and implement. Therefore much of the work in MHSPRT has focused on developing implementable, sub-optimal solutions, some of which have been shown to be asymptotically optimal as the risk approaches zero [1,2].

Further complications in applying MHSPRT to the track repair problem result from the variability of track stitching hypotheses across time. Probability density functions describing state dynamics evolve in time due to state prediction and covariance propagation. Assumptions such as *i.i.d.* measurements across time no longer necessarily apply. Furthermore, across frames new tracks may arise and current tracks may drop, thereby potentially changing the number of track stitching hypotheses.

In this paper, we adapt and apply the asymptotically optimal MHSPRT described in [1] to develop a local video track repair algorithm (VTR). Techniques are developed to compensate for variable hypothesis number and evolving track state probability density functions. We evaluate VTR performance on track data generated by a previously-developed video tracker operating on real video data acquired from an SUAV platform [3]. Because reducing computational complexity is critical for realtime video tracking, a kinematic-based approach using simple target dynamics is used. However, the framework supports the use of features such as target size, shape, color, and point, line, intensity, and edge features (e.g., [4,15,16,17,18]).

The rest of this paper is organized as follows. Section 2 presents the individual components of the track stitching algorithm. Section 3 integrates the components presented in section 2 with the MHSPRT to formulate the VTR. Section

4 presents numerical track repair performance results. Section 5 contains concluding remarks. Details of the SUAV video tracking system used to generate data in this report can be found in [3].

Throughout this report we use the terminology **coasted track** to refer to any track that has been terminated, but which may still correspond to a trackable object. Track states of coasted tracks are predicted forward in time using a target state dynamics model. The term **active track** refers to any track currently maintained by the tracker which represents tracking information on an object. A **stitching hypothesis** corresponds to a hypothesized **stitching association** between an active track ξ and a dropped track η and is denoted by the ordered pair (ξ, η). A **stitching solution** represents a stitching association (ξ, η) in which the two tracks are stitched. For compactness in the above terminology, the term "stitching" will often be dropped where it is understood.

For the MHSPRT formulation in [1], we assume use of the Lebesgue product measure as the dominating measure so that the Radon-Nikodym derivatives become ordinary probability density functions.

2. TRACK STITCHING COMPONENTS

2.1. Track State Prediction

Target state dynamics are modeled using a constant velocity (CV) model, defined by a partially decoupled, four-state system. Define states $x_1 = x$, $x_2 = y$, $x_3 = v_x$, and $x_4 = v_y$ where x, y represent the target x, y-direction positions, and v_x, v_y represent the target velocity component in the x, y-directions respectively, in image coordinates. Let $\overline{w}$ denote the process noise and $\overline{x}$ denote the state vector. The system model is given by $\dot{\overline{x}} = F\overline{x} + \overline{w}$ where

$$\overline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \qquad \overline{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}, \qquad F = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Because $F^2 = 0$, the state transition matrix Φ is given by $\Phi = I + F\Delta t$. The state extrapolation for the discretization of the system model is $\overline{x}_k = \Phi \overline{x}_{k-1}$, and thus the state extrapolation equations are

$$\begin{aligned} x_k &= x_{k-1} + \Delta t\, \dot{x}_{k-1} \\ y_k &= y_{k-1} + \Delta t\, \dot{y}_{k-1} \end{aligned}, \qquad \begin{aligned} \dot{x}_k &= \dot{x}_{k-1} \\ \dot{y}_k &= \dot{y}_{k-1}. \end{aligned} \tag{1}$$

Model the process noise vector as zero mean, Gaussian white noise, $\overline{w} = \mathrm{N}(0, Q)$ and model the initial state covariance as uncorrelated. The process noise covariance and initial state covariance are given respectively, by

$$Q_{ij} = \begin{cases} 0 & if\ i \neq j \\ \sigma_Q^2 & if\ i = j \end{cases}, \qquad P_{ij,0} = \begin{cases} 0 & if\ i \neq j \\ \sigma_P^2 & if\ i = j \end{cases}$$

The extrapolation of the error covariance is given by

$$P_k = \Phi P_{k-1} \Phi^T + Q. \tag{2}$$

Dropped track states are extrapolated forward in time using (1), and state error covariance is propagated using (2). Dropped track states probability distributions are modeled as Gaussian distributions centered on the extrapolated state values and with covariance from (2). It is easily proven that the covariance matrix takes the special form

$$P_k = \begin{pmatrix} \xi_k & 0 & \gamma_k & 0 \\ 0 & \xi_k & 0 & \gamma_k \\ \gamma_k & 0 & \eta_k & 0 \\ 0 & \gamma_k & 0 & \eta_k \end{pmatrix} \tag{3}$$

where

$$\begin{aligned} \xi_k &= \xi_{k-1} + 2\Delta t\gamma_{k-1} + \Delta t^2\eta_{k-1} + \sigma_Q^2, & & & \xi_0 &= \sigma_P^2 + \Delta t^2\sigma_P^2 + \sigma_Q^2 \\ \eta_k &= \eta_{k-1} + \sigma_Q^2, & & \text{for initial conditions} & \eta_0 &= \sigma_P^2 + \sigma_Q^2 \\ \gamma_k &= \gamma_{k-1} + \Delta t\eta_{k-1} & & & \gamma_0 &= \Delta t\sigma_P^2. \end{aligned}$$

2.2. Stale Track Pruning

Dropped tracks may persist for a significant time period without being stitched. Tracks that are not stitched may belong to targets that have ceased motion, traveled into occlusion zones, left the surveillance area, or are the result of a stitching error. Unstitched dropped tracks persisting for a significant time period will experience continual growth in error covariance. After sufficient time has passed without stitching, the error covariance may become prohibitively large resulting in a stale track. In this case, the dropped track should be eliminated from stitching consideration. The criterion used for testing track staleness consists of a thresholding test on the size of the eigenvalues of the covariance matrix using a variance threshold $\sigma_{STALE,i}^2$. To calculate the eigenvalues λ of the covariance matrix in (3), solve the determinant equation

$$\begin{vmatrix} \xi-\lambda & 0 & \gamma & 0 \\ 0 & \xi-\lambda & 0 & \gamma \\ \gamma & 0 & \eta-\lambda & 0 \\ 0 & \gamma & 0 & \eta-\lambda \end{vmatrix} = 0$$

to get two eigenvalues of multiplicity 2

$$\lambda_1 = \frac{(\xi+\eta) + \sqrt{\xi^2 - 2\xi\eta + (\eta^2 + 4\gamma^2)}}{2}, \qquad \lambda_2 = \frac{(\xi+\eta) - \sqrt{\xi^2 - 2\xi\eta + (\eta^2 + 4\gamma^2)}}{2}.$$

Prune dropped tracks by thresholding the eigenvalues λ_i, using the decision rule:

$$\lambda_i \underset{retain}{\overset{eliminate}{\underset{\leq}{\overset{>}{}}}} \sigma_{STALE,i}^2.$$

2.3. Homography Compensation

For video tracking applications, state information is produced relative to the coordinate system of the corresponding video frame. Before any inter-frame comparison calculations are performed, the state data must be transformed to a common coordinate system. For each dropped track, the common coordinate system to use will be the coordinate frame associated to the dropped track, immediately prior to track drop. That is, for a given dropped track η use the coordinate system F associated to the most recent set of state data measurements, prior to track drop. The motivation for using this coordinate system is to preserve the form of the covariance matrix in (3). Use of the homography, which performs the inter-frame coordinate transformation, may destroy the form of the covariance matrix thus precluding use of the closed form eigenvalue expressions (25) and (26).

Let H_k denote the homography representing the coordinate transform from frame F_{k-1} to frame F_k, i.e. $H_k : F_{k-1} \to F_k$ for time iteration k. The homography H_k expresses the coordinates of a pixel in the previous frame's coordinate system, in terms of coordinates of the pixel in the current frame's coordinate system. Homography compensation proceeds as follows. Let $F_{\eta,k}$ denote the coordinate system for dropped track η for frame number k. For each dropped track η, extrapolate all dropped track data using states expressed in coordinate system $F_{\eta,k}$. Propagate the dropped track error covariance based on state data expressed in coordinate system $F_{\eta,k}$. For track gating and stitching, transform the active track state data from the current frame to $F_{\eta,k}$.

2.4. Track Gating

Standard χ-square gating [19] is used to identify feasible stitching associations. An active track gates with a dropped track if the active track states reside within a chosen hyper-ellipsoid in state space centered at the dropped track states. In n-dimensional Euclidean space, for a multivariable Gaussian distribution with mean $\overline{m}$ and covariance matrix Σ, the equation for the state error hyper-ellipsoids is $(\overline{x} - \overline{m})^T \Sigma^{-1} (\overline{x} - \overline{m}) = \eta^2$. Substituting in the eigen-decomposition expression for the inverse of the covariance matrix, $\Sigma^{-1} = VD^{-1}V^T$, gives $\overline{e}^T D^{-1} \overline{e} = \eta^2$ where $\overline{e} = V^T(\overline{x} - \overline{m})$. Expanding this out gives:

$$\eta^2 = \sum_{i=1}^{n} \frac{e_i^2}{\lambda_i}. \tag{4}$$

where e_i denotes a component of the vector $\overline{e}$. Each e_i is Gaussian distributed with zero mean and variance λ_i, so in (4), η^2 is χ^2 distributed with n degrees of freedom. Therefore, the probability of the states residing within a hyper-ellipsoid is given by $\Pr\{\eta^2 < \tau\}$ for threshold τ. These values may be determined from statistical tables [10].

2.5. Global Hypothesis Resolution

Let C_T denote the set of coasted tracks. Let A_T denote the set of active tracks. Let $S_F \subseteq A_T \times C_T$ denote the set of all feasible stitching associations $S_F = \{(\xi,\eta)\ feasible | \xi \in A_T, \eta \in C_T\}$. A stitching association (ξ,η) is **feasible** provided three conditions are met: 1) η must not be stale, 2) the track start time for ξ must be later than the track termination time for η, 3) track ξ must gate with track η. Let G represent a real-valued gain function defined on subsets of S_F. The optimal approach to solution of the track stitching problem consists of determining the subset $S_F^{MAX} \subseteq S_F$ of specific stitching solutions that maximizes G;

$$S_F^{MAX} = \arg\max_{A \subseteq S_F} G(A).$$

Approaches to determining the optimal solution have been developed in [11]. However, determining the optimal solution can be significantly computationally expensive. Because computation time is critical in realtime video tracking applications, the VTR uses a faster suboptimal approach. The suboptimal approach consists of a two-step greedy algorithm. First, local hypothesis testing problems are defined and solved. Next, local hypothesis conflicts are resolved using global hypothesis resolution. The procedure is described as follows.

Assume ξ_i is an active track, for which there are M_i dropped tracks $\eta_1^i,\ldots,\eta_{M_i}^i$ that may be feasibly stitched to ξ_i, where $M_i > 1$. Define the **local hypothesis testing problem** associated to ξ_i:

$$\begin{aligned} H_1 &: associate\ \eta_1^i\ to\ \xi_i \\ &\vdots \\ H_{M_i} &: associate\ \eta_{M_i}^i\ to\ \xi_i \end{aligned}.$$

Each hypothesis H_i corresponds to a local stitching association (ξ_i,η_j^i). The set of hypotheses $H_i^L = \{H_1,\ldots,H_{M_i}\}$ represents a set of **local hypotheses** for stitching dropped tracks to the active track ξ_i, as illustrated in Figure 1. The sets of local hypotheses for all active tracks under consideration can be aggregated to form a set of **global hypotheses,** $H^G = \bigcup_i H_i^L$ that may require conflict resolution. Solving each local hypothesis testing problem associated to active tracks $\xi_1,\ldots\xi_N$ produces a stitching association set $A_G = \{(\xi_i,\eta_i^j)\,|\,1 \le i \le N,\ 1 \le j \le M_i\}$. For any pair $(\xi,\eta) \in A_G$ each ξ will be unique by construction. However, it may happen that η is repeated among association pairs. Local hypotheses conflict when a given dropped track associates with more than one active track, (Figure 2).

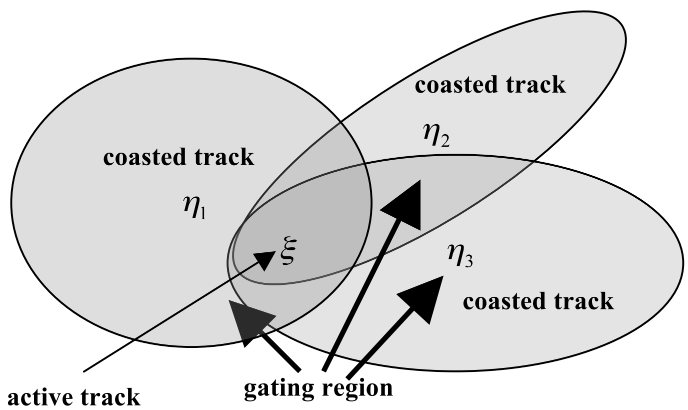


Figure 1: Local Stitching Hypothesis Set; An Active Track Gates With Multiple Coasted Tracks

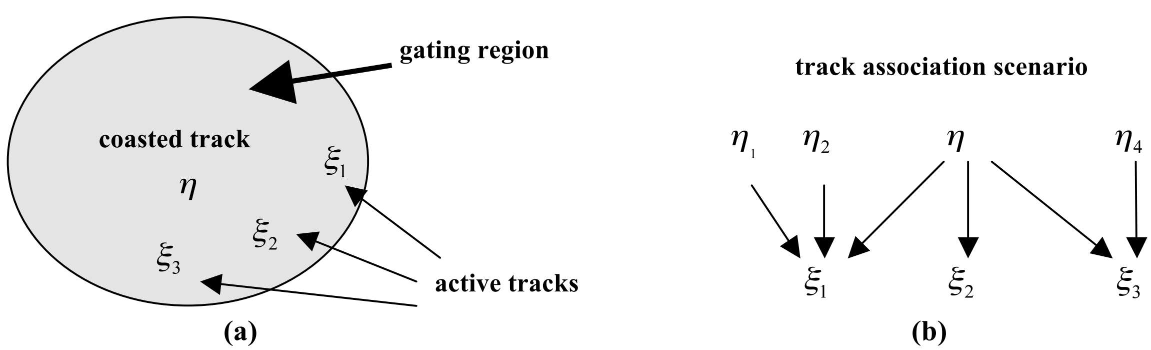


Figure 2: Global Stitching Hypothesis Set; (a) Single Coasted Track Gates with Multiple Active Tracks; (b) Single Coasted Track Associates with Multiple Active Tracks

When local hypotheses conflict, the conflict must be resolved by choosing one active track from the set of multiple active tracks associated with the given dropped track. This resolution is achieved by defining a global hypothesis testing problem. Let η denote a dropped track which has been associated with M active tracks $\xi_1,\ldots\xi_M$ through solution of the local hypothesis testing problems. Define the global hypothesis testing problem:

$$\begin{gathered} H_1 : associate\, \eta \; to\; \xi_1 \\ \vdots \\ H_M : associate\, \eta\; to\; \xi_M \end{gathered}$$

Solving each global hypothesis testing problem results in a set of de-conflicted stitching associations.

2.6. Posterior Probability Propagation

For active track ξ, assume there are $M > 1$ dropped tracks $\eta_1,\ldots,\eta_M$ feasibly stitchable to ξ. Consider the local hypothesis testing problem

$$\begin{gathered} H_1 : associate\, \eta_1 \; to\; \xi \\ \vdots \\ H_M : associate\, \eta_M\; to\; \xi \end{gathered} .$$

At time frame t_m, let $\rho(\alpha_1,\ldots,\alpha_N \mid t_m)$ denote the probability density function of the track states $\alpha_1,\ldots,\alpha_N$ for active track ξ. For dropped track η_i at time frame t_m, let $\rho_{\eta_i}(\alpha_1,\ldots,\alpha_N \mid t_m)$ denote the probability density function of the track states. Hypothesis H_i may be interpreted as $\rho(\alpha_1,\ldots,\alpha_N \mid t_m) = \rho_{\eta_i}(\alpha_1,\ldots,\alpha_N \mid t_m)$. Therefore the probability density function of the track states for ξ under hypothesis H_i is

$$\rho(\alpha_1,\ldots,\alpha_N \mid H_i, t_m) = \rho_{\eta_i}(\alpha_1,\ldots,\alpha_N \mid t_m).$$

At time frame t_m, measurements of the track states for active track ξ are obtained from the video tracker and denoted by $X_m = \{\alpha_{1,m},\ldots,\alpha_{N,m}\}$. At frame t_m define posterior probabilities $p_m^i = p(H = H_i \mid X_1,\ldots,X_m)$ where H represents the hypothesis choice, $H \in \{H_1,\ldots,H_M\}$. The hypothesis choice H is a discrete random variable so the posterior probabilities are probabilities in the probability mass function p_m^i. In the probability expressions, denote the event $H = H_i$ by H_i to simplify notation. For a mixed discrete and continuous random variable, use Bayesian update to propagate the posterior probabilities across frames:

$$p(H_i \mid X_1,\ldots,X_m,X_{m+1}) = \frac{\rho(X_{m+1} \mid H_i, X_1,\ldots,X_m)P(H_i \mid X_1,\ldots,X_m)}{\sum_{j=1}^{M} \rho(X_{m+1} \mid H_j, X_1,\ldots,X_m)P(H_j \mid X_1,\ldots,X_m)}. \tag{5}$$

The initial conditions on the prior probabilities are taken as: $P(H_i) = 1/M$ (equal priors assumption) . The probability density function $\rho(X_{m+1} \mid H_i, X_1,\ldots,X_m)$ represents the value of the probability density function of dropped track η_i at time t_{m+1}, evaluated at the latest measurement of the active track states. The conditional prior probability $P(H_i \mid X_1,\ldots,X_m)$ represents the posterior probability p_m^i calculated from the previous iteration. Note that

$$\rho(\mathrm{X}_{m+1} \mid H_i, \mathrm{X}_1, \ldots, \mathrm{X}_m) = \rho(\mathrm{X}_{m+1} \mid H_i)$$

because the track states for dropped track η_i are propagated independently of the measurements of the active track state. In fact, the probability density function $\rho_{\eta_i}(\alpha_1, \ldots, \alpha_N \mid t_m)$ depends only on the initial state conditions, the time t_m, and the assumptions in the state dynamics model. Thus $\rho(\mathrm{X}_{m+1} \mid H_i, \mathrm{X}_1, \ldots, \mathrm{X}_m) = \rho_{\eta_i}(\alpha_{1,m}, \ldots, \alpha_{N,m} \mid t_m)$.

3. TRACK STITCHING ALGORITHM

The VTR performs the following steps:

1) Initiate stitching processing on occurrence of a trigger event;
2) Determine the set of feasible track pairs for stitching (sets of local hypotheses);
3) Perform local hypothesis resolution by running an MHSPRT across video frames for each local hypothesis testing problem to choose a stopping time and hypothesis;
4) Given the resolved local hypotheses, perform global hypothesis resolution.

Track stitching is initiated whenever two conditions are satisfied: a new confirmed track is initiated, and there exists coasted tracks. Other processing triggers are also possible. If feature-aided tracking is used for example, stitching could be initiated when new features are generated.

At the stitching trigger time t_0, assume the set of feasible stitching pairs S_F consists of L active tracks, $\xi_1, \ldots, \xi_L$ with $M_i > 0$ dropped tracks $\eta_1^i, \ldots, \eta_{M_i}^i$ feasibly stitchable to active track ξ_i. To simplify notation, drop the superscript and represent $\eta_1^i, \ldots, \eta_{M_i}^i$ by $\eta_1, \ldots, \eta_{M_i}$ where it is understood that each dropped track set corresponds to an active track. Further assume that for each active track $\xi_i \in \{\xi_1, \ldots, \xi_L\}$, there exists a unique $\eta \in \{\eta_1, \ldots, \eta_{M_i}\}$ such that the stitching solution (ξ_i, η) is the correct stitching solution. The significance of this last assumption is that the stitcher assumes that gating implies the existence of a solution; the only problem becomes determining which solution is correct if multiple tracks gate. Track gates can be tightened (reduced in size) to reduce stitching errors. With small track gates, gated tracks have states with small state space deviation from the gate center. This small state space deviation makes the gated track more likely to correspond to the true track, so stitching errors are reduced. However, small track gates imply fewer gated tracks. Hence the stitcher performs stitching less frequently, but with higher probability of correct decisions.

For each $\xi_i \in \{\xi_1, \ldots, \xi_L\}$, define the local hypothesis testing problem H_i^L

$$\begin{gathered} H_1 : associate\, \eta_1 \ to\ \xi_i \\ \vdots \\ H_{M_i} : associate\, \eta_{M\,i} \ to\ \xi_i \end{gathered} .$$

If $M_i > 1$, proceed in time across video frames collecting track state measurement data for active tracks and predicting dropped track states. Across video frames, propagate the posterior probabilities using the Bayesian update in (5). During data collection across frames, new active tracks may arise, current active tracks may drop (giving rise to new dropped tracks), or dropped tracks may be eliminated, e.g. due to staleness. In the first case, each new active track requires definition and solution of an associated local hypothesis testing problem. Also, associations resulting from solution of new local hypothesis testing problems must be considered during global hypothesis resolution. In the second case, let ξ_i denote a newly dropped active track and let η denote a dropped track with $(\xi_i, \eta) \in S_F$. If (ξ_i, η) is a true stitching solution, then it is desirable to maintain the option to stitch the two dropped tracks ξ_i, η. Because ξ_i has been dropped, then measurements for the track states of ξ_i are not available. Therefore, use the coasted ξ_i track states as measurements in the Bayesian update. In the third case, assume dropped track η is removed due to staleness. For each (ξ_i, η) such that $(\xi_i, \eta) \in S_F$, the hypothesis (ξ_i, η) must be removed from the local hypothesis testing problem associated to ξ_i. If $M_i = 1$, then there are no remaining dropped tracks feasible for stitching to ξ_i. Therefore, remove ξ_i from stitching consideration, via $S_F = S_F - \{(\xi_i, \eta) \in S_F\}$. This third case is somewhat problematic for posterior probability propagation. The problem becomes one of determining the proper way to modify probabilities for current hypotheses, to account for eliminated hypotheses. The *ad hoc* approach we take is posterior probability renormalization; simply renormalize the posteriors so that they sum to one.

For each local hypothesis testing problem, the stopping time is given by

$$t_{stop} = \min\left\{m : \max_i p(H_i \mid \mathrm{X}_1, \ldots, \mathrm{X}_m) \geq p_{stop}\right\}$$

where p_{stop} represents a posterior probability threshold, and the decision rule is

$$H_j : j = \arg\max_i p(H_i \mid \mathrm{X}_1, \ldots, \mathrm{X}_k).$$

The stopping time represents the earliest time at which the maximum posterior probability surpasses the threshold. The decision corresponds to the hypothesis associated with the maximum posterior probability. When the maximum posterior probability has passed the p_{stop} threshold, then it is assumed that enough samples have been taken for a decision to be rendered in the local hypothesis testing problem. The stopping times for each local hypothesis testing problem may be different. Stitching processing terminates when all local hypothesis testing problems have been resolved. The solution to each local hypothesis testing problem produces a set of local stitching associations $S_F^{LOC} = \{(\xi_i, \eta_j)\}$. For any pair $(\xi, \eta) \in S_F^{LOC}$ each ξ will be unique by construction. However, it may happen that η is repeated among association pairs. Global hypothesis resolution must be performed to resolve the conflicting associations for repeated η.

4. NUMERICAL RESULTS

This section presents three examples of local video track repair. The video clip in this first example contains a vehicle traveling on a road, shaded by trees (Figure 3). The vehicle is tracked prior to entering the shadow occlusions. In the shadows, track is dropped. Upon exit, a new track arises. This example represents a single-target tracking case. The

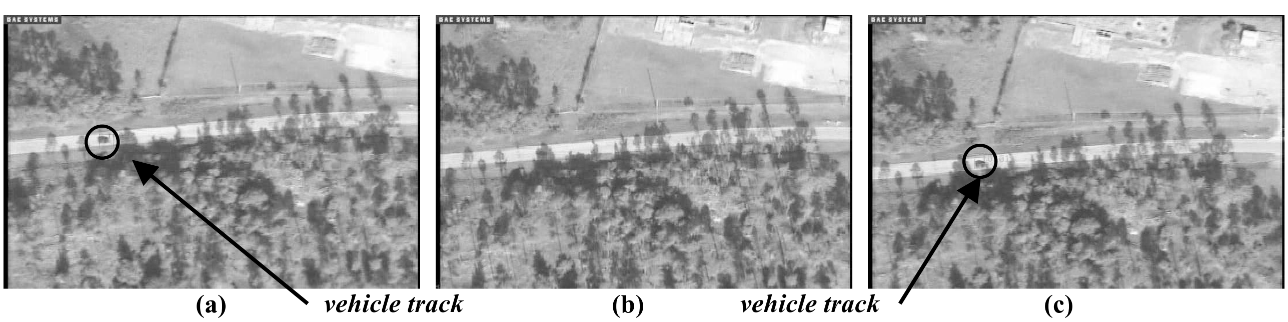
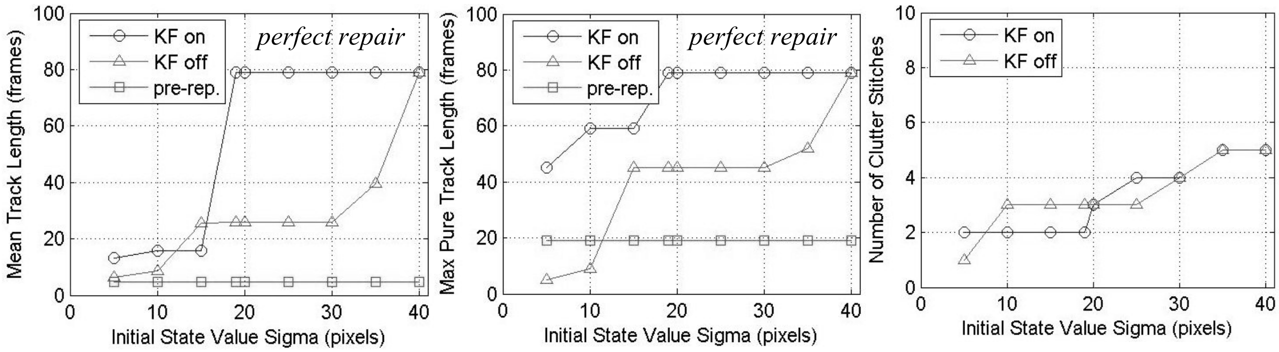


Figure 3: (a) Vehicle Tracking Prior to Entering Shadow; (b) Track Dropped in Shadow; (c) Track Initiated Upon Exit

local hypothesis testing problems in the VTR contain single hypotheses. Therefore track repair is performed simply by state extrapolation followed by track gating. Figure 4 contains track metric data resulting from performing track repair over the entire twenty two second clip, as a function of the initial state value standard deviation σ_P. The value of σ_P, together with the track staleness threshold and process noise, controls the size of the track gate. Tight gates result in fewer track repairs, but provide high probability of correct association. Looser gates encourage more frequent stitching, potentially connecting more clutter tracks. The track metrics in Figure 4 consists of mean track length (only pure tracks, stitched and un-stitched), maximum *pure* track length, and number of clutter track connections made. Track metric data is provided for before and after track repair. Due to noisy velocity measurements, σ_P is taken to be large in Figure 4.



(a) (b) (c)

Figure 4: Track Repair Metrics as Functions of Track Gate Size (a) Mean Track Length; (b) Maximum Pure Track Length; (c) Number of Clutter Track Connections

Figure 4. contains pre- and post-repair results, with track state Kalman filtering turned on and off. Results in Figure 4 indicate that track repair improves track quality by significantly increasing mean track length and maximum track length. The degree of benefit depends on track gate size; tighter track gates provide smaller improvements but also reduce the number of clutter track connections. Looser track gates provide better repair performance, at the cost of increased clutter track connection. Kalman filtering turned on provides better correct stitching performance for smaller clutter stitching.

Example 2 displays deeper VTR functionality. In this example for demonstration purposes, the track gates are greatly loosened to potentially allow clutter tracks to gate with target tracks. In this example, the local hypothesis testing problems contain multiple hypotheses, and the VTR must accumulate evidence across frames to make a decision.

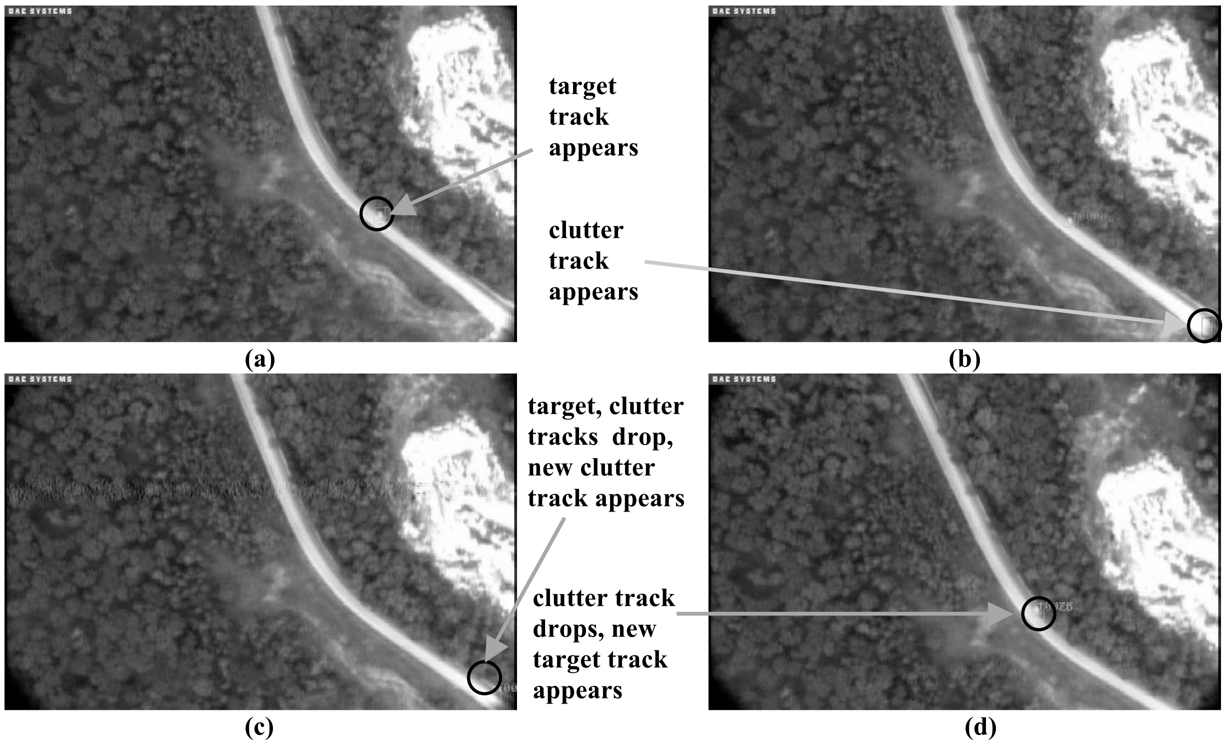


Figure 5: (a) Target Track Appears; (b) Clutter Track Appears; (c) Target and Clutter Tracks Drop; (d) New Track Arises

Figure 5 shows a sequence of frames in which target and clutter tracks arise and drop. In Figure 5 (a) a target track appears. In Figure 5(b) a clutter track appears. Next, the target and clutter tracks drop (Figure 5(c)) and a new clutter track appears. In Figure 5 (d), a new target track appears. The track repair problem consists of a local hypothesis testing problem with three hypotheses. These hypotheses correspond to new track association with one of three dropped tracks.

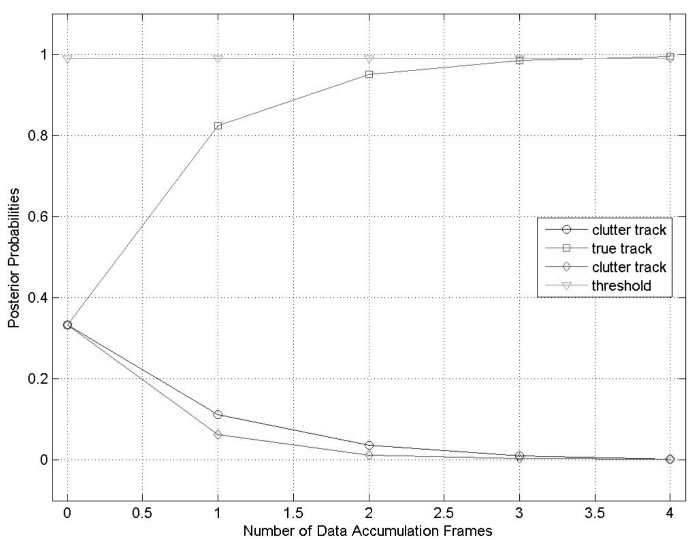


Figure 6: Propagation of Posterior Probabilities Associated to Each Repair Hypothesis

Figure 6 contains plots of the posterior probabilities corresponding to each of the three association hypotheses, as a function of number of evidence accumulation frames. Figure 6 indicates that each hypothesis begins with an equal prior. As track measurements are compared against predicted track states, the posterior probabilities evolve to reflect the comparison. Figure 6 shows that the hypothesis associated to the correct track repair association increases, while the less likely hypotheses converge to zero. After three frames, sufficient evidence has been accumulated to make a track repair decision. The correct hypothesis posterior surpasses the probability threshold, and the VTR makes the correct decision.

Example 3 demonstrates track results for parameter choices which optimize different track metrics. Probability of track purity is the percentage of correctly associated track labels for non-clutter tracks. Coverage percentage is the percentage of all connectable labels that are actually connected. Figure 7 shows track repair results for optimizing each metric, and for an intermediate case.

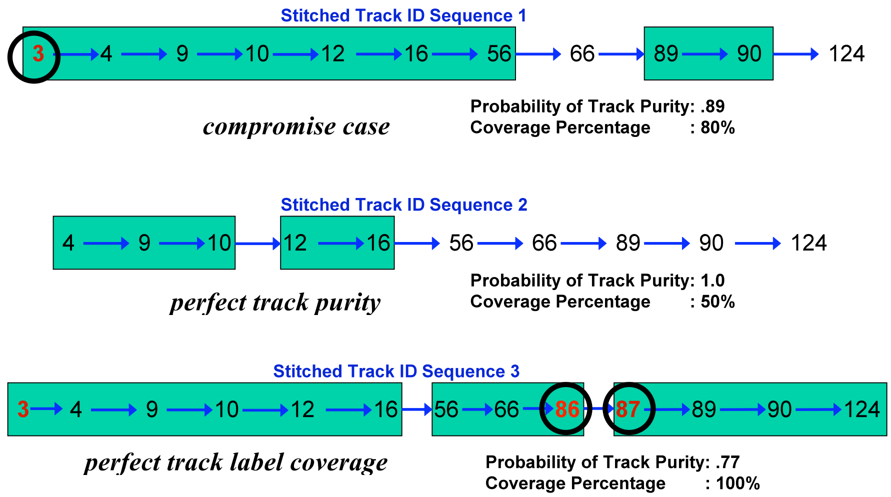


Figure 7: Track Repair Sequence for Video Clip Similar to Example 1 (Vehicle on road, Entering/Exiting Shadows)

In Figure 7 the numbers connected by arrows correspond to the true target track label sequence. The boxes show which labels are connected. Circles indicate incorrectly connected track labels. Figure 7 shows that through suitable parameter choice in the VTR algorithm, different track metrics may be optimized resulting in different track repair performance.

5. CONCLUSIONS

Performance results generated in this paper indicate that the local video track repair algorithm presented herein can provide improved tracking performance on real video data. Current and future work involves more closely examining the tradeoff between computation time and model fidelity, determining qualitative performance behavior under more general operating conditions, and adding feature-based track repair capability to improve performance.

ACKNOWLEDGEMENTS

This work was sponsored by the Air Force Research Laboratory under Air Force Contract Number FA8650-04-C-1681. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Air Force. The authors wish to thank Dr. Mark Stevens for suggesting use of evidence accumulation across frames.

REFERENCES

1. V. Dragalin, A. Tartakovsky, V. Veeravalli, "Multihypothesis Sequential Probability Ratio Tests- Part I: Asymptotic Optimality", *IEEE Trans. on Information Theory,* Vol. 45, No. 7., pp. 2448-2461, November, 1999.
2. C. Baum, V. Veeravalli, "A Sequential Procedure for Multihypothesis Testing", *IEEE Trans. on Information Theory,* Vol. 40, No. 6., pp. 1994-2007, November., 1994.
3. M. Stevens, A. Reiter, S. DelMarco, L. Vinciguerra, M. Antone, T. Jenkins, and R. Neuroth, "Video Aided Navigation for Small UAVs", *AIAA Infotech@Aerospace 2007 Conf. and Exhibit,* May 7-10, 2007, Rohnert Park, CA.
4. E. Cheng, M. Piccardi,"Track Matching by Major Color Histograms Matching and Post-Matching Integration", *Proc. ICIAP'2005*, Vol. 3617, pp. 1148-1157, 2005.

5. P. Arambel, M. Antone, M. Bosse, J. Silver, J. Krant, T. Strat, "Performance Assessment of a Video Based Air to Ground Multiple Target Tracker with Dynamic Sensor Control", *Proc. SPIE Defense and Security Symp.,* 2005.
6. P. Arambel, M. Antone, C. Rago, H. Landau, T. Strat, "A Multiple-Hypothesis Tracking of Multiple Ground Targets from Aerial Video with Dynamic Sensor Control", *Fusion'2004,* 2004
7. A. Gelb, ed., *Applied Optimal Estimation*, Cambridge: The M.I.T Press1974.
8. J. Parker, W. Blair, "Use of Target-Oriented Process Noise in Tracking Maneuvering Targets", *Technical Report NAVSWC TR 91-701*, August, 1992.
9. F.J. Breidt, A.L. Carriquiry, "Highest Density Gates for Target Tracking", *IEEE Trans. On Aerospace and Electronic Systems,* Vol. 36, No. 1, pp. 47-55, Jan., 2000.
10. D.B. Owen, Handbook of Statistical Tables, Addison-Wesley, Reading, Mass., 1962.
11. C. Stauffer, "Estimating Tracking Sources and Sinks", *Proc. IEEE Workshop on Event Mining in Video*, 2003.
12. R. Kaucic, A.G. Amitha Perera, G. Brooksby, J. Kaufhold, A. Hoogs, "A Unified Framework for Tracking Through Occlusions and Across Sensor Gaps", *Proc. CVPR'05,* Vol 1., pp. 990-997, 2005.
13. S. Blackman, "Multiple Hypothesis Tracking for Multiple Target Tracking", *IEEE Aerospace and Electronic Systems Magazine,* Vol. 19, No. 1, pp. 5-18, Jan., 2004.
14. A. Wald, and J. Wolfowitz, "Optimal Character of the Sequential Probability Ratio Test", *Ann. Math. Statist.,* vol. 19, pp. 326-339, 1948.
15. Y. Guo, H. Sawhney, R. Kumar, and S. Szu, "Vehicle Fingerprinting for Reacquisition & Tracking in Videos", *Proc. ICPR'05,* 2005.
16. B. Li, and R. Chellappa, "A Generic Approach to Simultaneous Tracking and Verification in Video", *IEEE Trans. on Image Proc.,* Vol. 11, No. 5., pp. 530-544, May., 2002.
17. P. Arambel, J. Silver, M. Antone, T. Strat, "Signature-Aided Air-to-Ground Video Tracking", *ICIF'06*, pp. 1-8, 2006.
18. J. Kang, I. Cohen, and G. Medioni, "Continuous Tracking Within and Across Camera Streams", *Proc. ICPR'03,* 2003.
19. S. Blackman, Multiple Target Tracking with Radar Applications, Artech House, Norwood, Mass., 1986.