Behaviour recognition of ground vehicle using airborne monitoring of unmanned aerial vehicles

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This is an Accepted Manuscript of an article published by Taylor & Francis Group in International Journal of Systems Science on 4 Mar 2013, available online at: http://www.tandfonline.com/10.1080/00207721.2013.772677

Metadata Record: https://dspace.lboro.ac.uk/2134/16068

Version: Accepted for publication

Publisher: © Taylor and Francis

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
This paper proposes a behaviour recognition methodology for ground vehicles moving within road traffic using unmanned aerial vehicles in order to identify suspicious or abnormal behaviour. With the target information acquired by unmanned aerial vehicles and estimated by filtering techniques, ground vehicle behaviour is first classified into representative driving modes, and then a string pattern matching theory is applied to detect suspicious behaviours in the driving mode history. Furthermore, a fuzzy decision making process is developed to systematically exploit all available information obtained from a complex environment and confirm the characteristic of behaviour, while considering spatiotemporal environment factors as well as several aspects of behaviours. To verify the feasibility and benefits of the proposed approach, numerical simulations on moving ground vehicles are performed using realistic car trajectory data from an off-the-shelf traffic simulation software.

**Keywords:** Behaviour Recognition; Unmanned Aerial Vehicles; Trajectory Classification; Tracking Filter; String Matching; Decision Making

1. Introduction

Recently airborne surveillance and reconnaissance systems become a challenging and emerging issue in the area of aerospace and robotics, with the rapid improvement of UAV (unmanned aerial vehicle) systems and of associated sensing technology (Gunetti et al. 2011, Sharma and Ghose 2009). For instance, Casbeer et al. (2006) analysed the feasibility of using multiple low-altitude, short endurance UAVs to cooperatively monitor and track the propagation of large forest fires. The interest for airborne surveillance was also extended to the roadway traffic monitoring (Puri 2004, Oh et al. 2012). Compared to traditional traffic surveillance sensors such as loop detectors and video cameras positioned at fixed locations, UAV aerial sensing can provide better coverage with the capability to survey large areas at a high speed without being confined to prescribed ground navigation routes (Coifman et al. 2006). Thus, this airborne monitoring enables suspicious or unusual behaviour in road traffic to be identified and investigated promptly so that operator’s situational awareness is increased in support of border patrol, law enforcement and protecting infrastructure.

In particular, UAV surveillance with an onboard MTIR (Moving Target Indicator Radar) sensor can provide accurate information of a large number of moving targets and consequently
allows rapid and consistent detection of possible threats. However, an operator still needs to analyse the resulting position and motion data and construct a picture of events, in order to detect suspicious behaviour. This usually requires several highly-skilled human operators, which is expensive and unsustainable under a deluge of data and information which could result from complex scenes. Hence, there is a strong need to develop high-level analysis algorithms to process target information and detect anomalous behaviours, to reduce the human operator’s workload.

For sensing and modelling motion dynamics of ground vehicles, a research group at the Delft Technical University (Schreuder et al. 2003, Ossen 2008) has done the research about calibration of microscopic traffic simulations. Their objective was to analyse and model driving behaviour of individual vehicles by using high-resolution digital image data collected with a camera attached to a manned helicopter. There are also several classical studies on modelling driver behaviour: drowsiness estimation based on electroencephalogram (EEG) (Lin et al. 2005); emotional behaviour by collecting driving performance data, psychophysiological responses and eye movement data (Cai et al. 2007); and driver handling behaviour in a drive-vehicle-environment system using neural network architectures (Lin et al. 2005).

For multi-sensor management and information fusion, a large number of studies have been actively done, and the related issues and approaches can be found in detail at reference (Xiong and Svensson 2002). Specifically, Deming and Perlovsky (2007) tried to develop a concurrent multi-target tracking and navigation based on a probabilistic technique by using a swarm of flying sensors. Looney and Liang (2003) developed an integrated multi-phase approach focusing on middle and high level data fusion for situation and threat assessments of ground battle spaces using a fuzzy belief network. Modi et al. (2011) explored an integrated information fusion framework in which signals received from various sensors are used by several types of experts to infer the human state. Even though this technique was proposed in the general context, it can certainly be used for specific applications like the abnormal behaviour identification.

In terms of guidance and control, Ariyur and Fregene (2008) proposed a ground vehicle tracking method by aerial vehicles using chaser and prey models but focused on the tracking control law development with a simple double integrator target dynamics. Kim and Kim (2008) proposed an optimal path planning to keep a close line of sight from a UAV to a ground target in a densely populated area. Oh et al. (2012) proposed the standoff tracking guidance for a ground target moving on the road networks using a nonlinear model predictive control along with a road-constrained filtering.

Detecting anomalous behaviour can be classified largely into two categories: The first approach codifies the behaviours using the experience and domain knowledge of experts, while in the second approach the behaviours are extracted and learned from the data (Will et al. 2011). Purely learning based approaches can provide good performance (Johansson and Falkman 2007, Du et al. 2006, Loy et al. 2009), however, they require massive data set in advance, or tend to suffer from a high computation burden for real-time applications. On the other hand, there are several algorithms which deal with behaviour or activity analysis in the context of codified (or classified) behaviour models, with the aid of the learning approach for both maritime and ground traffic surveillance scenarios.

In this regard, Srivastava et al. (2011) introduced a method to detect anomalies of the ground vehicle by observing the patterns in its velocity trajectory using a hypothetical coordinated system. The shape of velocity trajectory is used to detect anomalies by reference to paths derived from a trained Gaussian mixture model. In Li et al. (2007), moving target trajectories are expressed as discrete pattern fragments, known as motifs along with several attributes such as location, duration of motif and average speed. With this motif-based feature space representation, hierarchical rule-based classifier is used to detect abnormal behaviours. Fraile and Maybank (1998) proposed the idea of dividing the trajectories of ground vehicles into distinct driving modes, using video images obtained for ground traffic surveillance. However, this classification is limited to slow speed car manoeuvres in an urban parking area.
This paper proposes a behaviour recognition methodology against moving ground vehicles within road traffic using UAVs, in order to identify suspicious or abnormal behaviour, thereby reducing the workload of human operators. For this study, vehicular traffic modelling should be firstly addressed, which is a central problem of traffic engineering (Shvetsov 2003, Chakroborty 2006). This is generally classified into two main approaches differing in the level of detail: macroscopic and microscopic. Macroscopic models attempt to describe the traffic flow by interpreting the traffic flow as a compressible fluid without distinguishing between vehicles. In contrast to the macroscopic ones, microscopic models define the traffic behaviour by describing the behaviour of individual drivers in different situations. This study focuses on the microscopic-level differential geometric modelling since there is a definite need of tracking each individual vehicle’s trajectory for irregular behaviour recognition.

Current research using UAVs has focused on the development of separate algorithms, e.g. sensing and modelling, multi-sensor management and information fusion, or guidance and control required for the ground traffic monitoring as mentioned above. However, they have not been integrated into a system to aid situation assessment. On the other hand, this study describes a systematic and integrated development for autonomous airborne monitoring of ground traffic behaviour. It includes the following techniques: target tracking, sensor fusion and trajectory refinement, trajectory classification, and behaviour recognition. More specifically, one of contributions of this study is to propose a classification methodology for driving behaviour represented as a sequence of integers used to categorise specific differential geometric quantities. Additionally, a simple but effective detection scheme for irregular driving behaviour is proposed, by applying string matching theory (Theodoridis and Koutroumbas 2006) which has been widely used in the area of text-processing.

Although the aforementioned string matching algorithm provides a measure of suspicious behaviour, additional information needs to be considered to finally confirm the characteristic of behaviour in order to avoid frequent false alarms. For instance, a geographical database or other domain knowledge describing surrounding environments, such as traffic flow density or velocity data of normal vehicles could be of great help in efficiently identifying the intention of suspicious ground vehicles. Therefore, a fuzzy decision making process is developed to systematically exploit all available information obtained from a complex environment. The proposed fuzzy expert rule-based decision making system can concurrently accommodate several aspects of behaviour, as well as taking into account spatiotemporal environment factors, thus providing a more effective level of alert to the operator monitoring complex scenes acquired by airborne UAVs.

The overall structure of this paper is given as follow. Section 2 introduces the ground target tracking filter design, trajectory classification to model the behaviour of ground vehicles, and behaviour recognition algorithm using string matching theory. Section 3 introduces a rule-based decision making algorithm to detect suspicious or anomalous behaviour based on fuzzy logic principles. Section 4 presents numerical simulation results of behaviour monitoring for both military and civilian traffic scenario using realistic ground vehicle trajectory data. Lastly, conclusions and future work are addressed in Section 5. Figure 1 shows a flow chart of the overall algorithm for the proposed automatic behaviour monitoring scheme.

2. Behaviour Modelling and Detection

This section introduces behaviour modelling and detection algorithm of ground vehicles. The ground target tracking filter using UAVs is first explained. Trajectory classification is followed to model the behaviour of ground vehicles, and lastly behaviour recognition algorithm using string matching theory is presented.
2.1. Target tracking

This study considers acceleration dynamics to apply it to the tracking of ground vehicle. This model defines the target acceleration as a correlated process with a decaying exponential autocorrelation function, which means if there is a certain acceleration rate at a time $t$, then it is likely to be exponentially correlated at a time instant $t + \tau$ as:

$$x_k^t = F_k x_{k-1}^t + \eta_k$$

where the state vector is $x_k^t = (x_k^t, \dot{x}_k^t, \ddot{x}_k^t, y_k^t, \dot{y}_k^t, \ddot{y}_k^t)^T$, and where $\eta_k$ is a process noise which represents the acceleration characteristics of the target. The state transition matrix $F_k$ is given by:

$$F_k = \begin{bmatrix} 1 & T_s & \Phi & 0 & 0 & 0 \\ 0 & 1 & (1-e^{-\alpha T_s}) & 0 & 0 & 0 \\ 0 & 0 & e^{-\alpha T_s} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & T_s & \Phi \\ 0 & 0 & 0 & 0 & 1 & (1-e^{-\alpha T_s}) \\ 0 & 0 & 0 & 0 & 0 & e^{-\alpha T_s} \end{bmatrix}$$

where $\Phi = (e^{-\alpha T_s} + \alpha T_s - 1)/\alpha^2$, and $\alpha$ is a correlation parameter which models different classes of targets: a small $\alpha$ for targets with relatively slow manoeuvres and a high $\alpha$ for targets with fast and evasive manoeuvres. The details of the covariance matrix $Q_k$ of the process noise $\eta_k$ and other characteristics of this model can be found in (Mehrotra and Mahapatra 1997, Bar-Shalom et al. 2001).

Besides, this study assumes that the UAV is equipped with a MTIR (Moving Target Indicator Radar) sensor to localise the position of the target. Since the measurements of the MTIR are composed of range and azimuth of the target with respect to the radar location, the actual measurements are the relative range and azimuth with respect to the position of the UAV. The radar measurement $(r, \phi)^T$ can be defined as the following nonlinear relation using the target
position \((x_k^i, y_k^i)^T\) and the UAV position \((x_k, y_k)^T\) as:

\[
\mathbf{z}_k = \begin{pmatrix} r_k \\ \phi_k \end{pmatrix} = h(x_k^i) + \nu_k = \begin{pmatrix} \sqrt{(x_k^i - x_k)^2 + (y_k^i - y_k)^2} \\ \tan^{-1} \frac{y_k^i - y_k}{x_k^i - x_k} \end{pmatrix} + \nu_k
\]

(3)

where \(\nu_k\) is a measurement noise vector, and its noise covariance matrix is defined as:

\[
V[\nu_k] = R_k = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_{\phi}^2 \end{bmatrix}.
\]

(4)

Considering that the measurement equation is nonlinear, the localisation of target is designed using the EKF (Extended Kalman filter) as (Lewis 1992):

**Prediction**

\[
x_{k|k-1}^i = F_k x_{k-1|k-1}^i
\]

(5)

\[
P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k
\]

(6)

**Measurement update**

\[
K = P_{k|k-1} H_k^T \{H_k P_{k|k-1} H_k^T + R_k\}^{-1}
\]

(7)

\[
x_{k|k}^i = x_{k|k-1}^i + K \{\mathbf{z}_k - h(x_{k|k-1}^i)\}
\]

(8)

\[
P_{k|k} = (I - KH_k) P_{k|k-1}.
\]

(9)

where \(\mathbf{z}_k\) is a measurement vector. The output matrix \(H_k\) is a Jacobian of Eq. (3) with respect to the time-update state \(x_{k|k-1}^i\) as

\[
H_k = \frac{\partial h}{\partial x_k^i}|_{x_k^i=x_{k|k-1}^i}
\]

(10)

where

\[
\frac{\partial h}{\partial x_k^i} = \begin{bmatrix} \frac{x_k^i - x_k}{\sqrt{(x_k^i - x_k)^2 + (y_k^i - y_k)^2}} & 0 & 0 & \frac{y_k^i - y_k}{\sqrt{(x_k^i - x_k)^2 + (y_k^i - y_k)^2}} \\ - \cos^2 \{\tan^{-1}(\frac{y_k^i - y_k}{x_k^i - x_k})\} & \frac{x_k^i - x_k}{(x_k^i - x_k)^2} & 0 & \cos^2 \{\tan^{-1}(\frac{y_k^i - y_k}{x_k^i - x_k})\} \frac{1}{x_k^i - x_k} \\ \end{bmatrix}
\]

(11)

In addition to the EKF, assuming a pair of UAVs track the same targets, a sensor fusion technique using a Covariance Intersection (CI) algorithm (Julier and Uhlmann 1997) is applied. Sensor fusion is a well-known technique where multiple sources of data are fused together in order to get a better estimate of underlying information. In particular, the CI algorithm uses a convex combination of the means and covariances of two random variables \(a\) and \(b\) in the information space. The real statistics of these variables are assumed to be unknown due to noise; they, in turn, provide the only information of consistent estimates of the means and covariances. In order to define the consistency of these two variables, it is assumed that the means and covariances of these variables are \(\{\bar{a}, P_{aa}\}\) and \(\{\bar{b}, P_{bb}\}\). The information given by variables \(a\) and \(b\) can then be combined together in order to get a new estimate \(\{\bar{c}, P_{cc}\}\). When \(P_{cc}\) lies within the intersection of \(P_{aa}\) and \(P_{bb}\) for any possible choices of \(P_{ab}\), then an update strategy \(P_{cc}\) in the intersection region will be consistent. The tighter the updated covariance fits the region of intersection, the
more information is extracted. This intersection can be formulated by the CI algorithm as:

\[
P_{cc}^{-1} = \omega P_{aa}^{-1} + (1 - \omega) P_{bb}^{-1}
\]

\[
P_{cc}^{-1} e = \omega P_{aa}^{-1} a + (1 - \omega) P_{bb}^{-1} b
\]

where \( \omega \in [0, 1] \). Free parameter \( \omega \) manipulates the convex weights which are assigned to \( a \) and \( b \). Different choices of \( \omega \) can be used to optimise the covariance update with respect to different performance criteria, e.g. minimising the trace or the determinant of \( P_{cc} \). The cost functions are convex with respect to \( \omega \) and will have a unique optimum in the range of \( 0 \leq \omega \leq 1 \).

Lastly, as behaviour description for the ground vehicle requires the trajectory history over a specific length of time (rather than instantaneously), the recent past history of the state estimates can be used to enhance the tracking accuracy periodically. In this study, an optimal fixed-interval smoothing (Lewis 1992) is applied with the EKF. This smoothing algorithm is composed of a forward filter and a backward filter as shown in Fig. 2. The basic idea is that, if the measurements between \( t \) and \( T_f (> t) \) are available, the estimates of the forward filter at time \( t, \hat{x}_f(t) \), can be adjusted, based on the estimates of the backward filter at that time, \( \hat{x}_b(t) \). From now on, superscripts \( f \) and \( b \) designate the forward filter and the backward filter, respectively. The detail derivation of the optimal smoother can be found in (Lewis 1992, M.S.Grewal and Andrews 2008), and briefly explained as follows.

The optimal fixed interval smoothing requires three passes through the measurements and data derived therefrom, at every discrete time \( t_k \) in the entire fixed interval:

1. A complete filter pass in the forward direction (i.e., with measurement time increasing), saving the values of the \textit{a priori estimates} \( \hat{x}_f^k(+) \) and the associated covariance of estimation uncertainty \( P_{f}^k(+) \) (which are equivalent to those in Eqs. (8)~(9)).

2. A complete filter pass in the backward direction (time decreasing), saving the \textit{a priori estimates} and associated covariance of estimation uncertainties. Let us assume the following variable transformation.

\[
S_k(\pm) \triangleq P_{k}^b(\pm)^{-1}
\]

\[
\bar{y}_k(\pm) \triangleq P_{k}^b(\pm)^{-1} \hat{x}_{k}^{b}(\pm) = S_k(\pm) \hat{x}_{k}^{b}(\pm)
\]

If the last discrete sampling time is \( N \), the transformed estimates and error covariance matrix can be initialized using:

\[
\bar{y}_N(-) = 0
\]

\[
S_N(-) = 0.
\]
The backward filter performs the measurement update first:
\[
\hat{y}_k(+) = \hat{y}_k(-) + H_k^T R_k^{-1} z_k
\]  
\[
S_k(+) = S_k(-) + H_k^T R_k^{-1} H_k
\]

Then, the time update is performed using the following equations.

\[
K^b_k = S_k(+)(S_k(+)+Q_k^{-1})^{-1}
\]

\[
S_{k-1}(-) = F_k^T(I-K^b_k)S_k(+F_k
\]

\[
\hat{y}_{k-1}(-) = F_k^T(I-K^b_k)\hat{y}_k(+)
\]

3. A third, smoother pass combining the forward and backward data to obtain the smoothed estimate and the covariance of smoother uncertainty,

\[
K_k = P^f_k(+)(S_k(-)(I+P^f_k(+S_k(-))-1
\]

\[
P_k = (I-K_k)P^f_k(+
\]

\[
\hat{x}_k = (I-K_k)\hat{x}^f_k(+)+P_k\hat{y}_k(-)
\]

2.2. Trajectory classification

In order to recognise the driver’s behaviour, the trajectory is classified into driving modes. The purpose of the classification is to simplify and categorise the characteristics of manoeuvres associated with forward or lateral driving, by assigning them to driving modes. This classification approach will enable recognition of ground traffic behaviour in a computationally-efficient and flexible way using differential geometric quantities.

The assumption for this work is that the driving behaviour persists over a finite length of time rather than from sample to sample time. This implies that the vehicle will perform particular behaviours over a period of time that can be identified. For this, a moving-window-based trajectory approximation is applied using a third-order polynomial function which generates a trajectory with a virtually increased sampling time over a certain time interval. Note that, running a filtering algorithm with an increased sampling time could also be used, providing similar or possibly more accurate trajectory than that of a polynomial approximation. However, this approximation technique can be useful in case that the filter and smoothing algorithms are not working properly due to large process and sensor noises, or unavailable (e.g. only trajectory information with discrete time step is available without target and sensor model).

Let us assume a new time sequence within a moving window, \(0 < T_n < 2T_n < \ldots < (N_T - 1)cT_n = (N_T - 1)T_s\) where \(T_s\) is an original sampling time of the tracking filter, \(T_n\) is the new virtual sampling time, and \(N_T\) is the number of samplings for a moving window. In this study, it is assumed that \(N_T = 4, T_s = 0.5,\) and \(c = 5,\) and thus the new virtual sampling time is 0.1 seconds. The selection of \(N_T = 4,\) i.e. a 1.5 seconds moving window, reflects the bandwidth for lane changing of at least 1.0Hz as described in reference Lin and Ulsoy (1995). The velocity \((\hat{x}^i(i), \hat{y}^i(i))\) and acceleration \((\ddot{x}^i(i), \ddot{y}^i(i))\) histories with a new time sequence are then used to compute the minimum speed \(U,\) the rate of change of orientation \(\theta^i(i),\) and the forward acceleration \(a_f^i(i)\) of the vehicle at the current time step \(k\) for each \(i\) in a moving window (i.e.
\[ k - c(N - 1) + 1 \leq i \leq k \] as:

\[ U = \min v(i) = \min \sqrt{\dot{x}^2(i) + \dot{y}^2(i)} \quad (24) \]

\[ \theta(i) = v(i)\kappa(i) \]

\[ = \sqrt{\dot{x}^2(i) + \dot{y}^2(i)} \frac{\dot{x}^2(i)\ddot{y}(i) - \dot{y}^2(i)\dot{x}(i)}{(\dot{x}^2(i) + \dot{y}^2(i))^{3/2}} \quad (25) \]

\[ a_f(i) = \dot{x}^2(i) \cos \psi(i) + \ddot{y}(i) \sin \psi(i) \quad (26) \]

where \( \kappa \) is a curvature, and \( \psi = \tan^{-1}(\dot{x}/\dot{y}) \) is the heading angle from North. Using above equations, a selective driving mode \( m^d_k \) among the driving mode set \( M^d = \{0, \cdots, 8\} \) at time step \( k \) can be obtained for each moving window with a frequency of \( 1/T_s \) as:

- **Stopping (0)**, \( U < 1 \): Since 1 m/s equals to 3.6 km/h, it can be assumed that the car does not move or is about to stop or start moving.
- **Left turn (1)**, \( \max(\theta) \min(\theta) > 0 \) and \( \max(\theta) > \theta_{th} \): The inspection of the sign change of \( \theta \) is used to distinguish a pure turning manoeuvre from a lane change.
- **Right turn (8)** \( \max(\theta) \min(\theta) < 0 \) and \( \max(\theta) < -\theta_{th} \)
- **Left lane change (2)** \( \max(\theta) \min(\theta) < 0 \), \( \max(|\theta|) > \theta_{th} \), and \( \theta(0) > 0 \): The difference to a left turn is obtained by detecting a sign change of the rate of orientation change. As one can see in Fig. 3, the sign of the curvature transits from positive to negative for a left lane change.

- **Right lane change (7)** \( \max(\theta) \min(\theta) < 0 \), \( \max(|\theta|) > \theta_{th} \), and \( \theta(0) < 0 \): As can be seen in Fig. 3, the sign of the curvature transits from negative to positive in case of a right lane change.
- **Closing gap (6)** \( \max(a_f) \min(a_f) < 0 \), and \( a_f(0) > 0 \): Consider a linear trajectory that is sampled at several instances with a known sample period (not necessarily constant) as shown in Fig. 4. This illustrates the forward acceleration profile obtained by velocity change over each segment with a length of \( L \). Hence, velocity and acceleration at \( i \)-th instance are:

\[ v_i = \frac{L_i}{T_{i+1} - T_i} \quad (27) \]

\[ a_i = \frac{v_{i+1} - v_i}{T_{i+1} - T_i} \quad (28) \]

When the driver wants to close gap with the preceding vehicle, the sign of acceleration transits from positive to negative as shown in Fig. 4(a).
Figure 4. Acceleration history

- **Widening gap** (3) $\max(a_f) \min(a_f) < 0$, and $a_f(0) \leq 0$: Contrary to the case of gap closing, the sign of acceleration transits from negative to positive for this case.
- **Accelerating ahead** (5) $\max(a_f) \min(a_f) > 0$, and $a_f(0) > 0$: The sign of acceleration stays positive as shown in the Fig. 4(b).
- **Decelerating ahead** (4) $\max(a_f) \min(a_f) > 0$, and $a_f(0) \leq 0$: The sign of acceleration stays negative.

Figure 5 shows a full flowchart of the car trajectory classification algorithm proposed in this study. Using this technique, complicated trajectories can be simplified as string of numbers having essential characteristics of driver’s behaviours.
2.3. Behaviour detection

This section introduces behaviour detection scheme to detect suspicious behaviour using driving mode histories of ground vehicles. The key tools for our behaviour detection scheme are symbolic dynamics and string matching. The mathematical subject of symbolic dynamics originally arose in the theory of dynamical systems and was motivated by the qualitative approach to dynamics in which the character of trajectories is more important than their numerical values. String matching theory is a well-developed area of text processing. String matching consists in finding all of the occurrences of a string (called a pattern) in a text where the pattern is a string $x$ of length $m$, while the text is a string $y$ of length $n$. In this study, using the driving mode set $M^d = \{0, \cdots, 9\}$, a symbolic time series of driving modes $y^d_k = \{m^d_l \in M^d | l = 1, \ldots, N_{sm}\}$ is generated by trajectory classification for each time step $k$, where $N_{sm}$ represents a moving window length for string matching. The suspicious behaviour is also expressed as strings $x_s$ consisting of ten numbers.

The intuitive string matching method is an exact matching which detects exactly the same pattern in the driving mode history as the pre-defined suspicious string. However, to find the exact string to a certain behaviour class is not technically easy and computationally burdensome. Let us assume that we are interested in a reference pattern (driving mode sequence) ‘154058’, which represents a ground vehicle which turns left, accelerates ahead, decelerates ahead, stops, accelerates ahead, and lastly turns right. Sometimes, the test pattern appearing as the driving mode sequences, ‘154458’ or ‘154558’ cannot be ignored in the detection scheme, whose fourth element of the string could be one of the following forward driving modes: ‘3’, ‘4’, ‘5’, ‘6’, but not ‘0’. However, it is not efficient to run the string matching repeatedly using all of these possible driving-mode sequences. In this case, we need to define a cost, measuring the distance of the similarity between the reference pattern and the test patterns, which is here defined as the Edit Distance.

The Edit Distance (Theodoridis and Koutroumbas 2006) between two string pattern $S_1$ and $S_2$ is defined as the minimum total number of changes $C$, insertions $I$, and deletions $R$ required to change pattern $S_1$ into $S_2$:

$$D(S_1, S_2) = \min_j [C(j) + I(j) + R(j)]$$

(29)

where $j$ represents all possible combinations of symbol variations in order to obtain $S_2$ from $S_1$. A dynamic programming methodology is employed to compute the required minimum number of editions. For this, let us form a grid by placing the symbols of the reference pattern along the horizontal axis and the test pattern along the vertical axis, as shown in Fig. 6. For optimal path searching using dynamic programming, consider the following constraints (Theodoridis and Koutroumbas 2006).

- $D(0,0)=0$: The cost $D(0,0)$ is zero.
- A complete path is searched.
- Each node $(i, j)$ can be reached only through three allowable predecessors: $(i - 1, j), (i - 1, j - 1), (i, j - 1)$.

The cost of the three transitions can be defined for diagonal transitions and horizontal/vertical
transitions, respectively.

\[
d(i, j|i-1, j-1) = \begin{cases} 
0 & \text{if } r(i) = t(j) \\
1 & \text{if } r(i) \neq t(j)
\end{cases} \quad (30)
\]

\[
d(i, j|i-1) = d(i, j|i, j-1) = 1 \quad (31)
\]

Using these constraints and the Edit Distance as the performance index, dynamic programming can be applied to identify the minimum Edit Distance. The proposed detection scheme has a good theoretical basis on symbolic dynamics, and it is intuitive, robust within threshold bounds, computationally efficient and flexible since string patterns to search for can be easily adapted.

Note that, detecting a cyclical/circular shift of a reference string could be also of great importance for a suspicious string consisting of a set of repeated driving modes. By comparing a pre-defined suspicious string as well as its circularly shifted strings with the driving mode history of a certain length, the proposed behaviour detection scheme could cope with this cyclical/circular shift. For instance, let us assume a driving mode history \( a = [4 \ 5 \ 6 \ 1 \ 2 \ 3] \), and a suspicious string \( b = [1 \ 2 \ 3 \ 4 \ 5 \ 6] \). By shifting the string \( b \) progressively, Edit distance \( D \) will be changed as:

\[
\begin{align*}
D(a, \text{circshift}(b, 0)) &= D(a, [1 \ 2 \ 3 \ 4 \ 5 \ 6]) = 6 \\
D(a, \text{circshift}(b, 1)) &= D(a, [6 \ 1 \ 2 \ 3 \ 4 \ 5]) = 5 \\
D(a, \text{circshift}(b, 2)) &= D(a, [5 \ 6 \ 1 \ 2 \ 3 \ 4]) = 3 \\
D(a, \text{circshift}(b, 3)) &= D(a, [4 \ 5 \ 6 \ 1 \ 2 \ 3]) = 0 \\
D(a, \text{circshift}(b, 4)) &= D(a, [3 \ 4 \ 5 \ 6 \ 1 \ 2]) = 3 \\
D(a, \text{circshift}(b, 5)) &= D(a, [2 \ 3 \ 4 \ 5 \ 6 \ 1]) = 5
\end{align*}
\]

where \( \text{circshift}(b, i) \) circularly shifts the values in the string \( b \) by the shift size \( i \). Examining the minimum Edit distance \( D \) amongst all shifted strings makes it possible to detect the suspicious string having aforementioned shifting characteristic in the driving mode history.

Although Edit Distance can provide a measure of detection by computing the similarity between pre-defined suspicious strings and driving mode history, additional information needs to be considered to finally confirm characteristic of behaviour while avoiding frequent false alarms. From the following section, what types of information can be used and how to combine them will be dealt with.
3. Fuzzy Expert Rule-Based Decision Making

For airborne behaviour monitoring, this section proposes a decision making algorithm to detect suspicious or anomalous vehicle based on a fuzzy logic. To systematically exploit all available information obtained from complex environment, a fuzzy system is applied because of its ability to classify complex sources into simple and intuitive forms in the form of a rule-base. The proposed fuzzy expert rule-based decision making allows concurrent accommodation of several aspects of behaviour as well as spatiotemporal environmental factors, aided by supervision by a human operator.

A fuzzy system first classifies all available input information using membership functions (fuzzification), applies a rule-base and then produces an output result based on the rules (defuzzification). The fuzzy system used in this study consists of four fuzzy membership functions to classify the inputs and one output constructed from 36 expert inference rules, as shown in Fig. 7. Note that, the values in membership functions and inference rules used in the following are for illustration only, rather than extracted from experiments or experts.

3.1. Fuzzification

The fuzzy input for behaviour monitoring includes four parameters: location, behaviour cost, speed of the vehicle, and environmental data. The details are described as:

- **Location:** A time history of the position of the suspicious ground vehicle relative to a critical area (e.g. the centre of complex activities and the base walls of military facilities) or an index of the roadmap that the ground vehicle has moved along. Assuming that the local roadmap information is readily available in advance, the indexes of the local roads in region of interest can be annotated by a sequence of road numbers. If the vehicle travelling on one of identified roads of interest, the location is categorised as ‘Region of interest (R)’; otherwise it is categorised as ‘General (G)’, as shown in Fig. 8(a).

- **Behaviour cost:** As a key factor in behaviour monitoring, the Edit Distance \( D \) from specified test strings can be used to provide a time history of the behaviour cost. Let \( X_s = \{x_s^1, \ldots, x_s^{N_s}\} \) be the set of pre-defined suspicious behaviours. Then, the behaviour cost \( C_b^k \) with respect to current time series of driving modes \( y_d^k \) and suspicious behaviours at time step \( k \) can be defined as:

\[
C_b^k = \frac{1}{\min_{i \in X_s} D(x_s^i, y_d^k) + 1}
\]  

(32)

![Figure 7. Structure of fuzzy decision making](image)
Three fuzzy membership functions defined with linguistic variables ‘Normal (N)’, ‘Suspicious (Su)’, and ‘Worrying (W)’ are used to categorise the Behaviour Cost as shown in Fig. 8(b). Note that this Behaviour Cost can be replaced by other costs such as the anomaly score from the Gaussian process (Loy et al. 2009) instead of using Edit Distance.

- **Speed:** The speed of the vehicle with respect to its position or time also needs to be investigated since it can provide a measure of suspicious or abnormal behaviour. A membership function with linguistic variables ‘Slow (Sl)’, ‘Moderate (M)’, and ‘Fast (F)’ are used, as shown in Fig. 8(c).

- **Environment:** The last input considers an environmental condition with human interaction for the behaviour decision process. Depending on the traffic flow density, two membership functions with linguistic variables ‘Normal traffic (Nt)’ and ‘Congestion (C)’ are used, as shown in Fig. 8(d). Even though only traffic flow is used in this study, inputs such as time zone (day/night or weekday/weekend), or any other environmental parameter. Human operator input can also be used in the rule-base instead of relying on an autonomous decision process.

![Figure 8. Membership functions for fuzzy inputs](image)

The fuzzy output for behaviour monitoring rule-base is the level of alert for each ground vehicle, consisting of a membership function with four linguistic variables ‘Allow’, ‘Monitor’, ‘Investigate’, and ‘Respond’, as shown in Fig. 9.

### 3.2. Fuzzy inference

The fuzzy inference system is designed using a Mamdani model (Mamdani 1977). Expert knowledge can be expressed in a natural way using the linguistic variables previously defined to create a rule base as shown in Table. 1~2. From the table, the rules can be interpreted as:

- **Rule 1:** If Location is ‘G’ and Behaviour is ‘N’ and Speed is ‘Sl’ and Environment is ‘Nt’, then Alert is ‘Allow’.
Figure 9. Membership function for fuzzy output

Table 1. Fuzzy rule 1-18: location is ‘G’ (General road)

<table>
<thead>
<tr>
<th>Rule No.</th>
<th>Behaviour</th>
<th>Speed</th>
<th>Environment</th>
<th>Alert</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 / 2</td>
<td>N</td>
<td>Sl</td>
<td>Nt / C</td>
<td>Allow / Allow</td>
</tr>
<tr>
<td>3 / 4</td>
<td>N</td>
<td>M</td>
<td>Nt / C</td>
<td>Allow / Allow</td>
</tr>
<tr>
<td>5 / 6</td>
<td>N</td>
<td>F</td>
<td>Nt / C</td>
<td>Monitor / Investigate</td>
</tr>
<tr>
<td>7 / 8</td>
<td>Su</td>
<td>Sl</td>
<td>Nt / C</td>
<td>Monitor / Allow</td>
</tr>
<tr>
<td>9 / 10</td>
<td>Su</td>
<td>M</td>
<td>Nt / C</td>
<td>Monitor / Allow</td>
</tr>
<tr>
<td>11 / 12</td>
<td>Su</td>
<td>F</td>
<td>Nt / C</td>
<td>Investigate / Investigate</td>
</tr>
<tr>
<td>13 / 14</td>
<td>W</td>
<td>Sl</td>
<td>Nt / C</td>
<td>Investigate / Monitor</td>
</tr>
<tr>
<td>15 / 16</td>
<td>W</td>
<td>M</td>
<td>Nt / C</td>
<td>Investigate / Monitor</td>
</tr>
<tr>
<td>17 / 18</td>
<td>W</td>
<td>F</td>
<td>Nt / C</td>
<td>Respond / Respond</td>
</tr>
</tbody>
</table>

Table 2. Fuzzy rule 19-36: location is ‘R’ (Region of interest)

<table>
<thead>
<tr>
<th>Rule No.</th>
<th>Behaviour</th>
<th>Speed</th>
<th>Environment</th>
<th>Alert</th>
</tr>
</thead>
<tbody>
<tr>
<td>19 / 20</td>
<td>N</td>
<td>Sl</td>
<td>Nt / C</td>
<td>Investigate / Monitor</td>
</tr>
<tr>
<td>21 / 22</td>
<td>N</td>
<td>M</td>
<td>Nt / C</td>
<td>Allow / Allow</td>
</tr>
<tr>
<td>23 / 24</td>
<td>N</td>
<td>F</td>
<td>Nt / C</td>
<td>Investigate / Monitor</td>
</tr>
<tr>
<td>25 / 26</td>
<td>Su</td>
<td>Sl</td>
<td>Nt / C</td>
<td>Investigate / Monitor</td>
</tr>
<tr>
<td>27 / 28</td>
<td>Su</td>
<td>M</td>
<td>Nt / C</td>
<td>Monitor / Monitor</td>
</tr>
<tr>
<td>29 / 30</td>
<td>Su</td>
<td>F</td>
<td>Nt / C</td>
<td>Investigate / Investigate</td>
</tr>
<tr>
<td>31 / 32</td>
<td>W</td>
<td>Sl</td>
<td>Nt / C</td>
<td>Respond / Investigate</td>
</tr>
<tr>
<td>33 / 34</td>
<td>W</td>
<td>M</td>
<td>Nt / C</td>
<td>Investigate / Investigate</td>
</tr>
<tr>
<td>35 / 36</td>
<td>W</td>
<td>F</td>
<td>Nt / C</td>
<td>Respond / Respond</td>
</tr>
</tbody>
</table>

Note that, depending on the location and the environment, the rules are modified. For instance, if the location of the vehicle is ‘G’ (i.e. general area), the speed ‘Sl’ does not mean something significant, thus leading to an ‘Allow’ output, whereas if the location is ‘R’ (i.e. region of interest), slow speed or stopping of the vehicle can be identified as suspicious (monitoring the military base or placing of improvised explosive devices) leading to an ‘Investigate’ output, using Rule 1 and 19. However, even though the location is ‘R’ and the speed is ‘Sl’, if the environment is ‘C’ (i.e. congestion), its alert level should be alleviated using Rule 20 since slow speed is more likely to be observed in this area.
3.3. Defuzzification

Using the input variables and the defined fuzzy rules, the fuzzy outputs for all rules are then aggregated into one output fuzzy set. Finally, to obtain a crisp decision value for the level of alert, a defuzzification process needs to be performed. Even though there are several algorithms for this defuzzification, this study uses the method of taking the centre of gravity of the aggregated output fuzzy set (D. Driankov and Reinfrank 1993). Figure 10 shows the recommended level of alert using fuzzy input variables with defined inference rules. From Figs. 10(a)∼(c), it can be observed that the level of alert tends to decrease as the environment value (i.e. traffic flow density) increases when the location is ‘G’ (equivalently, 0.5). Meanwhile, in case the location is ‘R’ (equivalently, 1.0), as shown in Figs. 10(d)∼(f), representing the vehicle moving in the region of interest, slow speed causes a high alert, as shown in Fig. 10(d). Its effect decreases as traffic flow increases. Once a final decision value is made for a ground vehicle, corresponding tasks can be performed by UAVs such as following the identified target more closely for further monitoring (Kim et al. 2013).

![Figure 10](image)

Figure 10. Recommended level of alert according to fuzzy input variables (Location, Environment)

4. Numerical Simulations

This section presents numerical simulations for both civilian traffic and military scenarios using the proposed behaviour recognition framework for moving ground targets using two UAVs loitering over a certain area.

4.1. Civilian traffic scenario

The ground target trajectory is obtained using S-Paramics (Limited 2011) traffic model of Devizes in the United Kingdom, sampled at 2Hz, as shown in Fig. 11. It is used to generate MTIR
measurements composed of relative range and azimuth angle with respect to the position of UAVs. Generated MTIR measurements of a pair of UAVs were mixed with the zero-mean white Gaussian noise having the standard deviation of \((\sigma_r, \sigma_\phi) = (10\text{ m}, 3\text{ deg})\).

Figure 12(a) shows trajectory estimation result of a ground vehicle by the covariance intersection (CI) with the optimal smoother based on the EKF using two UAVs. Figure 13 describes the process of a new covariance ellipse (green one) generation by optimally combining those from two UAVs using CI algorithm at the final four time steps. In this study, in addition to given S-Paramics data, frequent lane changes are inserted artificially, in order to generate suspicious behaviour as shown in Fig. 12(b). This manoeuvre is labelled as weaving or evasive, and can be viewed as one of the most dangerous behaviours in civilian traffic. To detect this, a suspicious behaviour string \(x_s = 272727\) and ‘7272’ is selected (2: right lane change and 7: left lane change). In this scenario, every road is assumed to be a general road (i.e. location is ‘G’), and the size of driving mode history \(y^D\) is set to \(N_{sm} = 4\) which is the same as that of test sets \(x_s\).

Figure 14 shows the fuzzy rule-based decision making result including Behaviour Cost with trajectory classification and speed of the vehicle for 100 seconds. Note that even if \(y^D\) and \(x_s\) are totally different, since Edit Distance \(D\) between them is four, the lowest Behaviour Cost would be 0.2 instead of zero using Eq. (32), as shown in Fig. 14(b). In normal traffic shown by the blue line in Fig. 14(d), the level of alert has a high value when the Behaviour Cost is high, which means that evasive manoeuvre is likely over a period of 20~40 seconds, or velocity is fast at approximately 10 second. In case of congested traffic as shown by red dashed line in Fig. 14(d), although the level of alert shows the same tendency as the previous example, the effect of the Behaviour Cost is reduced as frequent lane change is more likely to happen due to traffic congestion. On the contrary, the effect of fast speed is enhanced at about 10 second since a fast moving vehicle in a congested traffic condition could be regarded as dangerous one.

Figure 11. Trajectory of a ground vehicle within the road network with GIS satellite data overlaid thanks to Google earth

4.2. Military scenario

Figure 15 shows the map description where a ground vehicle is moving around a military base. In the map, at the southern area of a river, there is a military base of strategic importance to be protected, which has a surrounding roadmap used by civilian ground vehicles, which has roads near the base wall. The ground vehicle considered in this scenario circles clockwise round the military base twice. During that time, the vehicle stops for ten seconds in the middle of road.
The trajectory classification histories show a reasonable performance, capturing the turning at the corner or stopping manoeuvre as shown in Fig. 16 in conjunction with the trajectory estimation result shown by blue lines and numbered road index history in Fig. 15. In this scenario, only road 3 and 4 are assumed to be of interest (i.e. location is ‘R’) as shown by red lines in Fig. 17(a), and suspicious behaviour string $x_s^{\prime} 4 4 0 0 0 0$ is selected (which is interpreted as deceleration and then stopping) to detect if the vehicle stops around the military base suspiciously.

Figure 17 shows the result of behaviour recognition including the time history of the road

segment 3 at approximately 420s. After that, it crosses the bridge and then travelling on the general road network. The true MTIR measurements of a pair of UAVs are mixed with the same noise characteristic as in the previous scenario.

The trajectory classification histories show a reasonable performance, capturing the turning at the corner or stopping manoeuvre as shown in Fig. 16 in conjunction with the trajectory estimation result shown by blue lines and numbered road index history in Fig. 15. In this scenario, only road 3 and 4 are assumed to be of interest (i.e. location is ‘R’) as shown by red lines in Fig. 17(a), and suspicious behaviour string $x_s^{\prime} 4 4 0 0 0 0$ is selected (which is interpreted as deceleration and then stopping) to detect if the vehicle stops around the military base suspiciously.

Figure 17 shows the result of behaviour recognition including the time history of the road
index, Behaviour Cost, speed of the vehicle, and the level of alert. In normal traffic, shown as the blue line in Fig. 17(d), the level of alert output from the fuzzy decision making process has a high value when the location is ‘R’, the Behaviour Cost is high and the speed is slow. However, if there is congestion in the traffic, the effect of the Behaviour Cost and the slow speed on the level of alert is reduced shown as the red dashed line in Fig. 17(d), since those conditions are more likely to happen due to traffic congestion.

5. Conclusions and Future work

This paper proposed an integrated behaviour recognition methodology for ground vehicles using UAVs, to identify suspicious or abnormal behaviours including the following techniques: target tracking filter, sensor fusion, trajectory classification, behaviour detection, and fuzzy decision making. Numerical simulation results using a synthetic military scenario and realistic car trajectory data from an off-the-shelf traffic simulation program demonstrated the feasibility of the proposed approach, providing the recommended level of alert for ground vehicles of interest. The proposed approach can be applied to various anomalous behaviour detection scenarios for ground and maritime traffic monitoring from the air. For future work, additional relevant aspects of behaviour will be considered as decision inputs. An example is the cultural background related to driving habits in the region where this system is to be deployed. The performance and robustness of the proposed scheme will also be investigated using both Monte Carlo simulation as well as real GMTI sensor data for different types of cars. Application of an adaptive fuzzy system would be of interest for more flexible human interaction in an uncertain and dynamic environment.
REFERENCES

Figure 15. Trajectory estimation of a ground vehicle for military scenario

Figure 16. Trajectory classification for military scenario

References


Figure 17. Fuzzy rule-based decision making results for military scenario

System, 153(1), 11.


**Hyondong Oh** received the BSc. and MSc. degrees in Aerospace Engineering from KAIST, South Korea, in 2004 and 2010, respectively. He is currently pursuing his Ph.D. degree in the Department of Engineering Physics, School of Engineering, Cranfield University, UK. His research interests include cooperative control and path planning of autonomous systems, target surveillance and tracking guidance, and information/sensor fusion.

**Seungkeun Kim** received a BSc. degree in Mechanical and Aerospace Engineering from Seoul National University, Seoul, Korea, in 2002 and then acquired a PhD on Three-dimensional Optimum Controller Design for UAV Formation Flight Using Behavioural Decentralized Approach from Seoul National University in 2008. He is currently an assistant professor at Dept of Aerospace Engineering, Chungnam National University, Daejeon, Korea. Previously he was a lecturer in Cooperative Control and Estimation at Autonomous Systems Group in Cranfield University, UK between Dec 2010 and Feb 2012 and also worked as a Research Fellow at Cranfield University between Apr 2008 and Dec 2010. He has actively published book chapters, journal and conference papers related to unmanned systems. He was also awarded Postdoctorates and Research Fellows Program from Korea Research Foundation in 2008. His research interests cover nonlinear guidance and control, estimation, sensor and information fusion, fault diagnosis, fault tolerant control, and decision making for unmanned systems.

**Hyo-Sang Shin** received his BSc from Pusan National University and gained an MSc on flight dynamics, guidance and control in Aerospace Engineering from KAIST and a PhD on cooperative missile guidance from Cranfield University. He is currently Lecturer on Guidance, Control and Navigation Systems in Autonomous and Intelligent Systems Group at Cranfield University. He has actively published book chapters, journal and conference papers and has been invited for many lectures in Universities, industries, and research institutes. His current research interests include cooperative guidance and control for multiple vehicles, coordinated health monitoring and management, and information-driven sensing.

**Antonios Tsourdos** obtained a MEng on Electronic, Control and Systems Engineering from the University of Sheffield (1995), an MSc on Systems Engineering from Cardiff University (1996) and a PhD on Nonlinear Robust Missile Autopilot Design and Analysis from Cranfield University (1999). He is a Professor of Control Engineering with Cranfield University. Appointed Head of the Autonomous Systems Group in 2007. Professor Tsourdos was member of the Team Stellar, the winning team for the UK MoD Grand Challenge (2008) and the IET Innovation Award (Category Team, 2009).
Brian A. White received the B.Sc.Eng. degree from the University of Leicester, Leicester, U.K., in 1967 and the M.Sc. degree in automatic control and the Ph.D. degree in bond graph modeling from the University of Manchester Institute of Science and Technology, Manchester, U.K., in 1971 and 1973, respectively. He is a Professor Emeritus with Cranfield University, Cranfield, U.K. He is with the Autonomous Systems Group, Cranfield Defence and Security, Cranfield University, Defence Academy of the United Kingdom, Swindon, U.K. His research interests include robust control, nonlinear control, estimation, observer applications, inertial navigation, guidance design, soft computing, and sensor and data fusion.