Failure Detection Methods to Predict Loss of Control Involving Human-Interface Devices Part I: Theory¹

D. W. Repperger¹, M. W. Haas¹, P. C. Schley², A. J. Koivo³

¹ - Air Force Research Laboratory
Wright-Patterson AFB, Ohio 45433

² - Veridian
Dayton, Ohio 45431

³ - Purdue University
West Lafayette, Indiana 47907

Abstract

A method is developed to detect a precipitation or loss of control in a device used as a human interface. The human-machine problem of interest involves two aircraft engaged in a position tracking scenario. The detection system developed here can be used on-line in real time or can be used in a *post hoc* sense for analysis of data already collected. This Part I paper discusses the basic theory underlying the operation of such a device.

1. Introduction

Fault detection continues to be an active area of research [1]. A fault is defined as an unpermitted deviation of at least one characteristic property of a variable from an acceptable behavior. There have been many approaches, using mathematical models to study fault detection [2-6]. This paper addresses an application of these techniques to investigate instabilities that may occur in human-machine situations which may be mitigated by the appropriate choice of an interface device.

Human-machine interactions may exhibit certain loss of control which is manifested by characteristics of the tracking error. At the brink of instability, an interaction termed a PIO (Pilot Induced Oscillation) can occur which is characterized by oscillatory behavior of an aircraft which can lead to a crash [7]. The PIO problem has been known since the Wright Brothers and their first flyer [8], and there are several theories to help explain this phenomenon. A far more serious instability that occurs in human-machine interactions can be characterized as a complete loss of control. Methodologies exist to delineate controllable and stable behavior and to distinguish it from unstable behavior [9] typical of what leads to the crash of an aircraft. This paper addresses some of the theoretical issues of the construction of a detector to give the pilot or operator of a tracking system some prior warning of the potential of losing the target. A companion paper [10] illustrates empirical validation of such a system. With a

credible prior warning, the display or controller of the human-machine system could be updated to possibly reduce the chance of a crash, or some more drastic procedures could be instituted. For example, control could be taken away from the pilot and the process could be completely automated.

2. Approach

Two approaches will be used to investigate this problem. The first method is based on fundamental properties of the closed-loop error signal and phase plane analysis. The second method utilizes Classical Control Theory to better understand how the instability arises.

2.1 Method 1 - Phase Plane Analysis and the Closed-Loop Tracking Error

Studies in human tracking [11] have demonstrated quite clearly that certain characteristics of phase planes of the closed-loop tracking error are conducive to good tracking while other attributes are indicative of poor tracking. First, in order to generate a phase plane portrait, the available error data need to be measured and estimates have to be generated on the higher-order derivatives of this signal.

2.1.1 Formulation of the Human Interface Problem - A Position Tracking Problem

Figure (1) depicts a block diagram description of the overall system under investigation. The target trajectory (f_t) represents the aircraft being followed and the pursuer aircraft is denoted by (f_p) which is under control of the operator. The tracking error $e(t)$ represents a six-dimensional pose vector (position and orientation) difference between the two aircraft. The objective of the human, who interacts with his (the pursuer's aircraft) dynamics, is to reduce the signal $e(t)$ to tolerably low levels. Only the time history $e(t)$ is available to drive a detector system.

1 - Supported, in part, by AFOSR grant, "Understanding Attentional Mechanisms Underlying Increases in Performance Associated With Multi-Sensory Displays," Task 2313CH71.

2.1.2 Estimating Higher Derivatives of $e(t)$

In Figure (1), the signal $e(t)$ is assumed to be obtained from measurements of both f_p and f_t . These data can be gleaned from sources such as radar, GPS (global positioning systems), AWACS (other aircraft), way-point display symbology, inertial guidance systems, or other sources. To develop an estimate of the higher derivatives of this signal, let $\hat{e}$ denote a low-pass filtered estimate of $e(t)=f_t-f_p$ with $\hat{E}$ and E denoting their respective Laplace transformed quantities. A transfer function is formed via:

$$\frac{\hat{E}}{E} = \frac{1}{(1 + \alpha/s)^3} \quad (1)$$

where α is the low-pass filter breakpoint. State variables are selected ($x_1=\hat{e}$, $x_2=\dot{\hat{e}}$, $x_3=\ddot{\hat{e}}$) and a state vector $x=\text{col}[x_1, x_2, x_3]$ is integrated forward using a first order Euler approximation

$$x_{t+\Delta t} \approx x_t + \dot{x}_t \Delta t \quad (2)$$

with initial conditions based on similar approximations in the error data. Figure (2) illustrates the form of the low-pass filter method to estimate the higher derivatives which can be easily implemented in either software or hardware.

2.1.3 The Detector and the Prediction of Error Divergence - Defining the Paradigm

Figure (3) illustrates how the detector works in principle. The output of Figure (2) is used to estimate the next three derivatives of the error signal utilizing the low-pass filter framework. The methodology of the detector system in the right-most block in Figure (3) is now described in detail.

The Detector Block

Since the variables ($\hat{e}$, $\dot{\hat{e}}$, $\ddot{\hat{e}}$) and also ($\ddot{\hat{e}}$) are known or estimated, the rationale for predicting error divergence can now be developed. Figures (4a-c) demonstrate the three phase planes ($\dot{\hat{e}}$ vs $\hat{e}$), ($\ddot{\hat{e}}$ vs $\dot{\hat{e}}$) and ($\ddot{\hat{e}}$ vs $\hat{e}$) if the human machine system were in a perfect oscillation. In a perfect oscillation the time spent in quadrants I and III of the phase plane is 50% of the time. It will be shown in the sequel that this type of pattern of tracking behavior is extremely destructive. This approach of finding patterns of certain types of tracking behavior is not unlike that of Ed Bristol who has successfully used such methods for adaptive control and in process control applications [12,13]. We state certain lemmas but due to space restrictions, briefly sketch out the proofs. Some definitions are introduced to better explain the lemmas.

Definitions: Let

$$r_1 = n_1/n \quad (3)$$

$$r_2 = n_2/n \quad (4)$$

$$r_3 = n_3/n \quad (5)$$

$$\text{and} \quad \gamma(t) = r_1 + r_2 + r_3 \quad (6)$$

where n is the total number of data samples under consideration and n_1 of these data samples fall in quadrants I and III of Figure 4a, n_2 fall in quadrants I and III of Figure 4b, and n_3 fall in quadrants I and III of Figure 4c. Lemma 1 addresses the approximate value of these r_i quantities when the human-machine system may be in or at a near PIO mode of operation.

Lemma (1)

If $e(t)$ is dominated by a single sinusoid, the following relationship is approximately true for a complete cycle of the sinusoid $e(t)=\sin(\omega t)$: ($r_1 \approx 0.5$, $r_1 \approx 0.5$, $r_1 \approx 0.5$ and $\gamma(t) = 1.5$).

Remark

Lemma 1 is obvious for a complete cycle of the sinusoid. This would be typical of a PIO situation. The second lemma addresses when a trajectory will show a tendency to diverge in a phase plane.

Lemma (2)

Trajectories that enter and remain in quadrants I and III of Figures 4a, b, or c lead to divergence in the sense that $\|\hat{e}_{t+\Delta t}\| > \|\hat{e}_t\|$. (A Euclidean norm is assumed).

Proof of Lemma (2)

Using a first order Taylor's series expansion of $\hat{e}_{t+\Delta t}$ yields the Euler approximation:

$$\hat{e}_{t+\Delta t} \approx \hat{e}_t + \Delta t \dot{\hat{e}}_t \quad (7)$$

Thus, from inspection of equation (7) and Figure (4a), if $\hat{e}_t$ and $\dot{\hat{e}}_t$ both have the same sign, the magnitude of $\hat{e}_{t+\Delta t}$ can only increase. By differentiating equation (7), the results then extrapolate to Figures 4b and 4c using the same arguments. An alternative way to show this result is to use a second order Taylor's series expansion which would be of the form:

$$\hat{e}_{t+\Delta t} \approx \hat{e}_t + \Delta t \dot{\hat{e}}_t + (1/2) (\Delta t)^2 \ddot{\hat{e}}_t \quad (8)$$

From observation of equation (8), one can see that if the three variables ($\hat{e}$, $\dot{\hat{e}}$, $\ddot{\hat{e}}$) have the same sign, this produces a definitive increase in the magnitude of $\hat{e}_{t+\Delta t}$. Lemma 3 addresses the converse argument and combines this measure of tracking divergence into the one variable $\gamma(t)$.

Lemma (3)

Trajectories that enter and remain in quadrants II and IV of Figures 4a, b, or c lead to convergence in the sense that $\|\hat{e}_{t+\Delta t}\| < \|\hat{e}_t\|$. A rule for determining convergence or divergence of a trajectory could be specified via: If $\gamma(t) > 2.0$, then the trajectory is diverging. If $\gamma(t) < 1$, then the trajectory is converging.

Proof of Lemma (3)

By examination of equation (8), if there is no common sign agreement among the three variables $(\hat{e}_t, \dot{\hat{e}}_t, \ddot{\hat{e}}_t)$, then the magnitude of the variable $\hat{e}_{t+\Delta t}$ can only decrease in magnitude. With this concept in mind, a detector system can now be constructed.

The Detection Algorithm

With reference to Figure (3), the detection algorithm calculates $\gamma(t)$ via:

$$\gamma(t) = r_1(t) + r_2(t) + r_3(t) \quad (9)$$

Thus the decision on whether tracking control is to be lost requires concurrence from the majority of the three phase planes as demonstrated in equation (9). To improve the quality of the detection scheme, tests were made on empirical data and [10] describes much of this detail. To summarize, two criteria were selected to help detect a loss of control or a possible PIO:

- (a) $\gamma(t)$ must be greater than some threshold value, e.g. $\gamma(t) > 2.0$ for divergence to be precipitated. (10a)
- (b) The rate of change of $\gamma(t)$ must be positive prior to the loss of control. (10b)

Further discussion and validation of this algorithm are contained in [10]. In this paper dealing with the theory of such a device, a relationship to Classical Control Theory will next be given. It is shown via classical techniques that the rationale for selecting the detection scheme (9, 10a-b) also has a sound basis from the prior work accomplished in human-machine systems. Also, when applied to data from an actual crash, the results discussed here concur with the sign relationship obtained from the discussion presented so far.

2.2 Applying Classical Control Theory to the Detector Scheme

This final section will show the concurrence of the detector scheme developed so far to well-known results from the analysis of human-machine systems. In Figure (1), the human and aircraft dynamics can be approximated by the following transfer function:

$$H(s) = F_p(s)/E(s) = (K_c/s) e^{-\tau_c s} \quad (11)$$

where K_c is the well-known "crossover frequency" for human-machine systems [14] and τ_c represents an effective time delay of the pilot-aircraft dynamics.

In Figure (1), if $H(s)$ is specified via equation (11), then the closed-loop transfer function between f_p and f_t is given by:

$$F_p/F_t = 1/(1+[se^{\tau_c s}]/K_c) \quad (12)$$

From a performance perspective, the following transfer function is of interest:

$$G(s) = E/F_t = s/(s + K_c e^{-\tau_c s}) \quad (13)$$

Performing the long division to three terms, it is seen that:

$$E/F_t = 1 - (K_c/s) e^{-\tau_c s} + (K_c^2/s^2) e^{-2\tau_c s} - \dots \quad (14)$$

and expanding $e^{-\tau_c s}$ and $e^{-2\tau_c s}$ to second order using:

$$e^{-\tau_c s} \approx 1 - \tau_c s + (1/2)(\tau_c s)^2 - \dots \quad (15)$$

$$e^{-2\tau_c s} \approx 1 - 2\tau_c s + (1/2)(2\tau_c s)^2 - \dots \quad (16)$$

provides the necessary approximations. A discussion of stability and its relationship to τ_c can now be given.

Comments on Stability

With reference to equation (13), $G(s)$ can be expanded to second order to yield:

$$G(s) = E/F_t \approx s/(s + K_c[1 - \tau_c s + 1/2(\tau_c s)^2]) \\ = s/(1/2[\tau_c^2 K_c s^2] + s[1 - K_c \tau_c] + K_c) \quad (17)$$

where typically $K_c \approx 4$ radians/second (bandwidth of the open-loop human-machine system). It is easily seen that the condition $K_c \tau_c > 1$ will produce a pole in the right half plane. If, for a given τ_c the operator were to lower the gain K_c , the bandwidth decreases resulting in the target escaping off the screen and the error trajectory diverging in the phase portrait. Hence increasing τ_c destabilizes the system as expected. Now it will be shown that the sign relationships described so far have an interesting interpretation from a time/frequency domain perspective.

Comments on Phase Portraits from A Frequency/Time Domain Perspective

We derive certain characteristics of the phase portraits and relate them to data [7] from an actual crash to show concurrence with the material presented so far.

Continuing the discussion on the transfer function E/F_t to second order, the following closed-loop characteristics can be obtained from (17) using the relationships (15-16):

$$E/F_t \approx [1 + K_c \tau_c + 2(K_c \tau_c)^2] - \\ 1/s[K_c + 2K_c^2 \tau_c] + 1/s^2[K_c^2] - s[1/2 K_c \tau_c^2] \quad (18)$$

Letting $(s=j\omega, j=\sqrt{-1})$, this yields the following real and imaginary parts of the transfer function:

$$E/F_t = a_1 + b_1 j \quad (19)$$

$$\text{where: } a_1 = 1 + K_c \tau_c + 2(K_c \tau_c)^2 - 1/\omega^2 [K_c^2] \quad (20)$$

$$\text{and } b_1 = 1/\omega [2 K_c^2 \tau_c + K_c] - \omega/2 K_c \tau_c^2 \quad (21)$$

This analysis is repeated for the transfer function of e/f_t .

As in equation (13), start first with the transfer function:

$$sE/F_t = s^2/(s + K_c e^{-\tau_c s}) \quad (22)$$

Again performing the long division to three terms, it is seen that:

$$sE/F_t = s - K_c e^{-\tau_c s} + [(K_c^2/s)] e^{-2\tau_c s} - \dots \quad (23)$$

and substituting $e^{-\tau c s}$ and $e^{-2\tau c s}$ to second order via equations (15,16) yields:

$$sE/F_t \approx [-K_c - 2(K_c \tau_c)^2] + 1/s[K_c^2] + s[1 + 2(K_c \tau_c)^2 + K_c \tau_c] + s^2[-1/2 K_c \tau_c^2] \quad (24)$$

Again letting $s = j\omega$, this can be written in a form similar to (19):

$$sE/F_t = a_2 + b_2 j \quad (25)$$

$$\text{where: } a_2 = K_c[-1 - 2 K_c \tau_c] + 1/2 [K_c^2 \omega^2] \quad (26)$$

$$\text{and: } b_2 = \omega + 2(K_c \tau_c)^2 \omega + \omega K_c \tau_c - 1/\omega [K_c]^2 \quad (27)$$

In the time domain the phase angles differ since they depend on the selected values of K_c , τ_c , and ω . To find these phase angles in the time domain, assume F_t has magnitude unity with zero phase angle. The respective time domain angles for $e(t)$ and $\dot{e}(t)$ can be determined via:

$$\angle e(t) = \tan^{-1}(b_1/a_1) \quad (28)$$

$$\text{and } \angle \dot{e}(t) = \tan^{-1}(b_2/a_2) \quad (29)$$

where $\tan^{-1}$ computes the inverse tangent based not only on the ratio of the a_i , b_i quantities, but also takes into account in which quadrants they reside. The angle difference can be expressed in the form

$$\text{Angle Difference} = \tan^{-1}(b_1/a_1) - \tan^{-1}(a_1/b_1) \quad (30)$$

3. Application to Crash Data from [7]

To show some credence to the material presented so far, the analysis will be applied to data from a known PIO incident. It was learned that the delay was $\tau_c = 0.55$ seconds and two analyses will be conducted ($K_c = 4$ rad/sec and $K_c = 6$ rad/sec). Figure (5a) illustrates the plot of the angular difference between the two variables $e(t)$ and $\dot{e}(t)$ versus τ_c [$K_c = 4$ rad/sec and $\omega \approx 2\pi/(1.1)$ rad/sec.]. Figure (5b) illustrates a similar plot for the case $K_c = 6$ rad/sec. The phase angle between e and $\dot{e}$ is seen to be very close to -1.6 radians (-91.7 degrees) for the case $K_c = 4$ rad/sec. For the case $K_c = 6$ rad/sec., the angular difference is -1.35 radians (-77.4 degrees). In either situation, the phase angle between e and $\dot{e}$ was 90 degrees or less. This verifies that the variables e and $\dot{e}$ have similar signs and their trajectories appear in quadrants I and III of the phase portrait in Figure 4a. If, however, the angle between e and $\dot{e}$ were ± 180 degrees, this would be indicative of trajectories that arise in quadrants II and IV (having opposite signs). Hence this method has shown that data from an actual crash involving a PIO would yield space trajectories that would activate the detector described in this paper.

4. Summary and Conclusions

A human tracking error detector is developed which provides prediction properties of potential error divergence. The device was derived using principles in phase plane analysis and was also examined within the framework of Classical Control Theory. Data from a

known crash was examined within the context of the analysis presented here to show concurrence of the theory to real world applications.

5. References

- [1] R. Isermann, "Supervision, Fault-Detection and Fault-Diagnosis Methods - An Introduction," Control Engineering Practice, vol. 5, no. 5, May 1997, pp. 639-652.
- [2] A. S. Willsky, "A Survey of Design Methods For Failure Detection Systems," Automatica, vol. 12, 1976, pp. 601-611.
- [3] D. M. Himmelblau, "Fault Detection and Diagnosis in Chemical and Petrochemical Processes," Elsevier: Amsterdam, pp. 343-393.
- [4] R. Isermann, "Process Fault Detection Based on Modeling and Estimation Methods - A Survey," Automatica, vol. 20, no. 4, 1984, pp. 387-404.
- [5] J. Gertler, "Survey of Model-Based Failure Detection and Isolation in Complex Plants," IEEE Control Systems Magazine, 1988.
- [6] P. M. Frank, "Fault Diagnosis in Dynamic Systems Using Analytical and Knowledge based Redundancy," Automatica, vol. 26, pp. 459-474, 1990.
- [7] M. A. Dorheim, "Report Pinpoints Factors Leading to YF-22 Crash," Aviation Week and Space Technology, Nov. 9, 1992, pp. 53-54.
- [8] R. F. Stengel, "Toward Intelligent Flight Control," IEEE Transactions on Systems, Man, and Cybernetics, vol. 23, No. 6, Nov/Dec 1993, pp. 1699-1717.
- [9] D. W. Repperger, et al., "An Algorithm to Ascertain Critical Regions of Human Tracking Ability," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-9, No. 4, April 1979, pp. 183-196.
- [10] D. W. Repperger, M. W. Haas, P. C. Schley, A. J. Koivo, "Failure Detection Methods to Predict Loss of Control Involving Human-Interface Devices- Part II Empirical Validation, Proceedings of the ACC - 1998.
- [11] D. W. Repperger, et al., "A Study on Spatially Induced 'Virtual Force' with an Information Theoretic Investigation of Human Performance," IEEE Transactions on Systems, Man, and Cybernetics, vol. 25, no. 10, October, 1995, pp. 1392-1404.
- [12] E. Bristol, "Pattern Recognition: an Alternative to Parameter Identification in Adaptive Control," Automatica, vol. 13, 1977, pp. 197-202.
- [13] E. Bristol, "The Trouble with Hairy Controls (Unpredictability)," Proceedings of The 1997 IEEE International Conference on Control Applications, Oct 1997.
- [14] D. T. McRuer and D. Graham, "Manual Control of Single Loop Systems: Part I," Journal of The Franklin Institute, vol. 283, no. 1, Jan 1967, pp. 1-29.

Figure 1 - The Human-Machine Interaction System

Figure 2 - Estimator of the Tracking Error and its Derivatives

Figure 3 - Detector System in Operation

Figure 4a - $\dot{e}$ versus e during a PIO

Figure 4b - $\ddot{e}$ versus $\dot{e}$ during a PIO

Figure 4c - $\dddot{e}$ versus $\ddot{e}$ during a PIO