Developing a Feedback System to Augment Monitoring Performance of Aircraft Pilots

Florian Fortmann, Stefan Suck
Human Centered Design
OFFIS – Institute for Information Technology
Oldenburg, Germany
{florian.fortmann, stefan.suck}@offis.de

Denis Javaux
Symbio Concepts and Products
Bassenge, Belgium
denis.javaux@symbio.pro

Joan Cahill, Tiziana C. Callari
Centre for Innovative Human Systems
Trinity College Dublin
Dublin, Ireland
{cahilljo, callartc}@tcd.ie

Andreas Hasselberg
Institute of Flight Guidance
German Aerospace Center (DLR)
Braunschweig, Germany
andreas.hasselberg@dlr.de

Abstract—As a supervisor of a highly automated technical system, a human operator is the “ultima ratio” in abnormal situations that cannot be handled by the automation. Therefore, a human operator must adequately monitor the automated system throughout the entire operation. Unfortunately, humans tend to fall prey to a couple of demons leading to inadequate monitoring behavior, such as boredom, attentional tunneling or perservation. As a consequence, the human might lose situation awareness and be unable to detect and handle an abnormal situation in the given amount time. In this paper, we present a feedback system to augment monitoring performance on the part of an aircraft pilot. The feedback system has been developed under the umbrella of the European research and development project A-PiMod, where novel team-centered concepts for pilot-automation interaction are investigated.

Keywords—feedback systems; monitoring behavior; human supervisory control; eye tracking

I. INTRODUCTION

Human factors research has shown that there are several issues associated with automation [1] which are related to the fact that automation does not detach human operators from control but pushes them into the role of supervisors. The human-machine interaction paradigm related to this role is commonly known as supervisory control. Duties of human operators in charge of supervisory control include planning system actions, setting system goals and constraints, diagnosing system failures, intervening in system control as necessary, and learning from experience [2]. The core function performed by human operators in charge of supervisory control is the continuous monitoring of the status of the automated technical system and how it performs the delegated tasks [3]. Hereby, human operators typically have to manage the distribution of visual attention to information provided by a set of visual displays. Given the fact that human operators cannot supply visual attention to each visual display at the same time, they must periodically shift their visual attention between different information sources. While doing so, human operators might fall prey to a couple of demons distracting them from performing adequate monitoring behavior, such as boredom [4], attentional tunneling [5] or perservation [6].

Adequate monitoring behavior is determined by the continuous, efficient, and effective performance of visual sampling actions towards information sources demanding attention at any point in time. Depending on the complexity of a present situation, the pertinent tasks, and the size of the information space that needs to be supervised, performing adequate monitoring behavior can be very demanding for a human, as he or she must permanently decide which information source should be sampled, in which order, how often, and how long. Because visual attention is a limited resource, a human supervising a highly automated technical system need to apply a function that allows him or her to continuously perform monitoring actions associated with low performance costs, but high performance benefits [3]. The performance costs of a certain monitoring action depend, e.g., on whether an information source is easy or hard to reach (physical effort of head and eye movements), and whether the provided information is easy or hard to understand (cognitive effort for interpreting information). Both, physical and cognitive efforts consume time, which is an additional cost factor. The performance benefit of a certain monitoring action depends, e.g., on whether the sampled information source contains new and relevant information, or not.

Monitoring performance directly drives the generation of situation awareness [7]. The consequence of inadequate monitoring behavior can be insufficient situation awareness,
which has been identified as the major cause of human error in context of supervisory control [8]. Finally, the consequence of human error can be dangerous incidents threatening human life and the environment.

A classical example, which sadly proves the possible consequences of inadequate monitoring behavior, is the crash of Eastern Airlines Flight 401 [9]. According to the official accident report from the National Transportation Safety Board [10], the crash occurred as a result of the flight crew becoming preoccupied with a burnt-out landing gear indicator light, failing to notice that the autopilot had inadvertently been disconnected from aircraft control. As a consequence the aircraft gradually lost altitude and crashed in the Everglades although the Altitude Alarter was issued. During this crash 101 of the 176 people on board lost their lives.

In this paper, we present a feedback system to augment monitoring performance of an aircraft pilot. The feedback system has been developed under the umbrella of the European research and development project A-PiMod\(^1\). The A-PiMod project investigates novel team-centered concepts for pilot-automation interaction. These team-centered concepts are built around the vision that automation should be considered and designed as a supporting team member in the cockpit, which does not only observe and assess the flight situation, but also observes and assesses the mental states of the pilots, and understands how pilot behaviors fit to the demands of the pertinent tasks and mission objectives.

II. RELATED WORK

The work described in this paper focuses on the assessment and augmentation of the monitoring performance of human operators. Therefore, this section provides a review of related research in these two areas.

A. Assessment of Monitoring Performance

A function for the optimization of monitoring performance was described by Onken and Schulte [11] who recommend focusing on the information sources associated with the most urgent task. However, this exclusive-or function fails to attend to the fact that information sources associated with other less urgent tasks, also need to be sampled periodically to avoid loss of situation awareness [7].

Wickens\(^2\) [12] SEEV model describes monitoring behavior as a selective process. The SEEV model allows to estimate how likely it is that a human operator shifts his or her visual focus from one information source \(i_s\) to another information source \(i_v\), based on the four eponymous parameters: Salience, i.e., the salience of \(i_v\); Effort, i.e., the effort required to redirect a human operator’s visual focus from \(i_s\) to \(i_v\); Expectancy, i.e., a human operator’s expectancy that \(i_v\) contains new information; Value, i.e., the value of \(i_v\) in context of the tasks that the human operator has to perform. The parameters Salience and Effort are determined by the environment and referred to as bottom-up parameters. The parameters Expectancy and Value are knowledge-driven and referred to as top-down parameters. Wickens argues that adequate monitoring behavior is driven only by the top-down parameters.

Wortelen and Lüdtke [13] developed the AIE model, which describes the effect of information frequency on monitoring behavior. They argue that an information source with a high information frequency needs more frequent sampling actions than an information source with a low information frequency.

Hooey et al. [14] integrated the SEEV model into the Man-Machine Integration Design and Analysis System (MIDAS) in order to assess SA of a virtual pilot. SA has been quantified using values reaching from 0 (= bad SA) to 1 (= good SA). Their approach uses a function to assess how often information sources are sampled. As this perspective on SA is directly linked to monitoring behavior, we consider the function’s output primarily as an estimation of monitoring behavior adequacy. As will be described later, our work is highly related to this approach.

B. Augmentation of Monitoring Performance

There are basically two directions of solutions to augment monitoring performance. One the one hand, there are static solutions, such as display design guidelines. Display design guidelines give designers advice on how to design user interfaces that help human operators to find the information they need at the right time [15]. These solutions are applied at the design stage and therefore limited as they cannot deal with dynamic situations and circumstances. On the other hand, there are dynamic solutions, such as feedback systems. Feedback systems are dynamic and therefore far more powerful as they adapt to the specific demands of the situations and circumstances that human operators are typically faced with. In the following, two examples are introduced.

Graham [16] investigated the effect of an auditory feedback display on human operators supervising unmanned aerial vehicles (UAV). In a study, participants had to maintain awareness of UAV course deviations. Continuous auditory cues were issued to help the participants to improve their awareness of current and future UAV course deviations.

Fortmann et al. [17] investigated the effect of an ambient visual feedback display on human operators supervising multiple UAV. In a study, participants monitored highly automated UAV performing a firefighting mission. Monitoring behavior was recorded using an eye tracker and the performance was assessed. Ambient visual cues were continuously issued to provide human operators feedback on their monitoring performance. Study results indicated that ambient feedback on monitoring performance can augment monitoring performance. The work described in this paper transfers technology used in this research from a UAV control station to an aircraft cockpit.

\(^1\) Details can be found on the website http://www.apimod.eu
III. DESIGN

In this section, we present the design of the feedback system. As shown in Fig. 1, the feedback system closes a continuous loop on the monitoring performance of the pilot. The loop includes three basic functions: (1) observation of the monitoring performance of the pilot; (2) assessment of the adequacy of the pilot’s monitoring performance in context of the current flight phase; and (3) augmentation of the pilot’s monitoring performance to achieve a better allocation of visual attention towards the relevant information sources.

Each function is carried out by a dedicated component of the feedback system. The first component is an eye tracker, which is used to observe the monitoring performance of the pilot during flight. The second component is the MBA-Tracer software, which assesses the pilot’s monitoring performance. The third component is the Mission and Cockpit Management Display (MCMD), which provides the pilot – besides to a couple of other information – feedback on his or her monitoring performance. The communication between the components is realized via a network. In general, the feedback system is intended to support the entire cockpit crew. Due to limited access to eye tracking systems, we actually developed our prototype for only one pilot. In the following, we provide further details on each component of the feedback system.

A. Eye Tracker

In the last years, eye tracking has become a very popular technology for the observation of monitoring behavior. Reasons include the increased usability and accuracy, as well as reduced cost of eye tracking systems [18]. Eye tracking measures the orientation of a human’s eye ball. Based on knowledge it is possible to determine a human’s line of sight, where he or she looks at a certain point in time, how long he or she looks at a target area, and the path his or her eyes follow. At the moment, we use the Dikablis eye tracking system, which consists of a head unit and a laptop with installed software to detect when the eyes of the pilot are directed towards information sources, such as cockpit displays or the outside view. However, the feedback system would also work with any other eye tracking system, which allows to determine the target areas of eye movements.

B. MBA-Tracer

The MBA-Tracer is a software tool, which facilitates the analysis of eye tracking data to assess the monitoring behavior adequacy (MBA) of the pilot. The analysis focuses on the detection of dwells on information sources. A dwell is an uninterrupted sequence of fixation and saccade time towards an information source of at least 200ms [19]. The implemented dwell detection approach works as follows: once a dwell is detected, the information source is marked with a flag. The information source stays marked as long as the pilot’s eyes stay there. The dwell last from the point in time where the human operator started to look at the information source until the eyes move away from it. If the eyes shift to another information source, the flag is immediately removed.

The assessment is based on the dynamic Attention Supply Demand Model (ASDM). In the ASDM, a human operator is considered as a supplier of visual attention, and each information source as a demander of visual attention. The assumption underlying the ASDM is the eye-mind assumption: by sampling information sources a human operator supplies visual attention to information sources. Therefore, he or she satisfies their demand for attention (DFA). The ASDM integrates a formal construct for the identification of situations, pertinent tasks, and relevant information. Moreover, the ASDM defines functions to calculate DFA-scores for each information source, and MBA scores for the overall monitoring performance of the pilot. We apply a simplified version of the ASDM, which does not perform a hierarchical break-down of information sources into rich information objects and atomic information elements.

During flight, the MBA-Tracer analyzes the current situation, i.e., the aircraft and environment states, determines the flight phase, and assigns to each information source a relevance level and an attendedness level. For each flight phase-information source pair a relevance level needs to be defined a priori. The assignation of the current relevance level of an information source in context of a flight phase is performed dynamically. A relevance level is a value in the range [0, 1], where 0 indicated a low relevance and 1 indicates a high relevance of an information source. The attendedness level of an information source is the ratio between the actual neglect time and the a priori defined maximum neglect time of the information source. The assignation of the current attendedness level of an information object is performed dynamically. The attendedness level is a value in the range [0, 1], where 0 indicates a low attendedness and 1 indicates a high attendedness of an information source.

![Fig. 1. The feedback system closes a continuous loop on the monitoring performance of a pilot, including the functions observe, assess, and augment. Each function is carried out by a dedicated component of the feedback system.](image-url)
The DFA-score of an information source derives from the trade-off between the relevance level and the attendedness level. The calculation of the DFA-score of an information source is shown in (1):

\[ DFA_i = (a \ast rl_i) \ast (1 - b \ast al_i), \]

where \( i \) is an information source, \( rl_i \) is the relevance level of \( i \), and \( al_i \) is the attendedness level of \( i \). Further, the parameters \( a \) and \( b \) are weights, where \( a \) is a value in the range \([0,1]\) and \( b\) is a value in the range \([0,1]\). These parameters can be used to alter the influence of \( rl_i \) and \( al_i \) on the result \( DFA_i \), which is also a value in the range \([0,1]\), where 0 represents a low DFA-score and 1 represents a high DFA-score. \( DFA_i \) can only be equals 1 if \( rl_i = 1 \) and \( al_i = 0 \). Further, \( DFA_i \) can only be equals 0 if \( rl_i = 0 \) or \( al_i = 1 \). The DFA-matrix in Fig. 2 visualizes possible results with \( a = 1 \) and \( b = 1 \).

The ASDM also defines functions to assess the demand for attention of the overall information space and the overall monitoring behavior adequacy of a human operator. However, these functions have not been used for the development of the feedback system described in this paper. Therefore, no further descriptions are given in this paper. Details on these functions can be found in [13].

For our prototype, we consider the flight phases take-off, climb, cruise, decent, approach and landing. Further, we apply the MBA-Tracer to the following set of aircraft cockpit information sources: Primary Flight Display (PFD), Navigation Display (ND), Electronic Centralized Aircraft Monitor (ECAM), Mission and Cockpit Management Display (MCMD), Flight Control Unit (FCU), Multifunctional Control and Display Unit (MCDU), and the outside view (OV). We asked a pilot to determine for each flight phase-information source pair a relevance value and a maximum neglect time in seconds. Therefore, we applied a structured stepwise procedure. We first asked the pilot to order the information sources for each flight phase from “not relevant” to “very relevant”, and from “low maximum neglect” time to “high maximum neglect time”. Then, we asked the pilot to assign relevance levels and maximum neglect times to the information sources. The results of this procedure are shown in Table I and Table II. These look-up tables are the knowledge base of our feedback system. The ASDM uses this knowledge to calculate during flight the DFA-scores of each information source. Because the knowledge base is generated on the basis of subjective assessments the feedback of only one pilot is of course not very reliable. In order to increase the reliability of the knowledge base further pilots need to be involved in the future.

C. Mission and Cockpit Management Display

The DFA-scores, which are provided by the MBA-Tracer, are handed over to the Mission and Cockpit Management Display (MCMD). The MCMD is a central component of the A-

<table>
<thead>
<tr>
<th>TABLE I. DETERMINATION OF RELEVANCE LEVELS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PFD</strong></td>
</tr>
<tr>
<td>Take-off</td>
</tr>
<tr>
<td>Climb</td>
</tr>
<tr>
<td>Cruise</td>
</tr>
<tr>
<td>Decent</td>
</tr>
<tr>
<td>Approach</td>
</tr>
<tr>
<td>Landing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II. DETERMINATION OF MAXIMUM NEGLECT TIMES (S)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PFD</strong></td>
</tr>
<tr>
<td>Take-off</td>
</tr>
<tr>
<td>Climb</td>
</tr>
<tr>
<td>Cruise</td>
</tr>
<tr>
<td>Decent</td>
</tr>
<tr>
<td>Approach</td>
</tr>
<tr>
<td>Landing</td>
</tr>
</tbody>
</table>

Fig. 2. DFA-matrix for an information source \( i \) with \( a = 1 \) and \( b = 1 \). The colorization highlights low (green) and high (red) DFA-scores.
PiMod project, as provides access to several functions of the investigated team-centered concepts for pilot-automation interaction. The MCMD runs on a Microsoft Surface tablet. Fig. 3 shows the MCMD.

The MCMD consists of a dedicated mission level view (top) and a cockpit level view (bottom). In the mission level view, pilots can monitor the mission status in terms of risks and activate alternative missions if the risk for a mission turns out to high. For example, the mission “Land at EDDM” could be threatened by an increased possibility for loss of control during the approach. As shown in Fig. 3, loss of control can be the consequence of a thunderstorm in the proximity of the destination airport EDDM. In this case the MCMD proposes possible alternative missions, e.g., “Hold at EDDM” or “Divert to EDDF” (just for clarification: EDDM and EDDF are airport identification codes given by the International Civil Aviation Organization). In the current implementation, five different risk states are possible. Similar to the risk states in the mission level view, the escalation states are color-coded from green (non-critical) to red (critical). The displayed color represents the escalation state of the information source with the currently highest DFA-score. On the other hand, a pilot can always enter a view showing the current DFA-scores for each information source. The view can be entered by pressing the button labelled “>” next to the task label named “Monitor”. The feedback view allows the pilot to get feedback on demand, e.g., in situations where there is less to do in the cockpit. The view shows a bar for each information source. The length and the color of a bar represents the DFA-score of the associated information source.

The MCMD implements two strategies to augment a pilot’s monitoring performance during flight. Both strategies are depicted in Fig. 4. On the one hand, a notification is issued by the MCMD to prompt an inattentive pilot to check certain information sources if DFA-scores fall below a defined threshold. In the current implementation, five different escalation states are possible. Similar to the risk states in the mission level view, the escalation states are color-coded from green (non-critical) to red (critical). The displayed color represents the escalation state of the information source with the currently highest DFA-score. On the other hand, a pilot can always enter a view showing the current DFA-scores for each information source. The view can be entered by pressing the button labelled “>” next to the task label named “Monitor”. The feedback view allows the pilot to get feedback on demand, e.g., in situations where there is less to do in the cockpit. The view shows a bar for each information source. The length and the color of a bar represent the DFA-score of the associated information source.

**IV. TARGET PLATFORM**

The target platform is the Generic Experimental Cockpit (GECO) simulator of the Institute for Flight Guidance of the German Aerospace Center in Braunschweig, Germany. The GECO simulator is used to conduct studies with pilots in order to evaluate new display and control concepts. Research also focuses on the development and evaluation of innovative operational procedures that can be applied in the future using new technologies. Examples include systems for satellite-based
determination of position, maneuvering area lighting, air-ground communication, collision detection and avoidance as well as new sensors. The outside view is simulated using three high-resolution projectors that project an image on a mirror system with a 6-meter diameter. This allows an area of 180° by 40° to be displayed, giving a realistic perception of depth. Fig. 5 shows a panoramic view of the cockpit.

As the MCMD runs on a tablet, it will not be a physically integrated display of the cockpit environment. In order to prevent that the tablet on which the MCMD flies around in the cockpit during a turbulent flight, the MCMD is mounted on a rack in a central position of the cockpit. The position provides both pilots access to the display. The rack consists of a plate to which the MCMD is attached, and an arm with multiple movable joints, so that a pilot can easily pull the MCMD towards him or her for interaction. Fig. 6 shows the rack and the MCMD mounted in a mobile cockpit during an exhibition at the AeroDays 2015 in London, where intermediate results of the A-PiMod project were presented to and discussed with participants from academia and industry. The final integration into the GECO simulator is currently running.

V. Evaluation Approach

In the A-PiMod project, we need to evaluate the adaptive automation and multimodal cockpit concepts — including the feedback system described in this paper — from (1) a cockpit perspective (i.e. usability, functionality and integration of technology), and (2) a broader operational/safety perspective. The evaluation methodology combines formal and informal human-machine interaction design and evaluation approaches, along with an integrated stakeholder approach to evaluation [21]. The validation with internal (i.e. representatives of each project partner) and external (i.e. representatives of the aviation-related industry and flight operational system) stakeholders is referred to as the A-PiMod Community of Practice [22]. Currently, the A-PiMod Community of Practice comprises 16 members. Both sets of stakeholders are involved in the specification and evaluation of the emerging concepts and technologies. This spans several activities pertaining to the specification and evaluation of user and technical requirements and user interface design prototypes. This includes desktop evaluation, interviews and workshops.

So far, twenty two pilots have been involved in validation activities. In the first validation cycle, ten commercial pilots were recruited and involved both in desktop and simulator validation exercises; in the second validation cycle (which will happen in November 2015 and elapse over two weeks), twelve commercial pilots will be involved in simulator validation exercises and collaborative evaluation sessions. A series of scenarios have been selected to evaluate the adaptive automation and multimodal cockpit concepts and to demonstrate the achievement of the project impact statements and selected key performance indicators. These scenarios involve a “theatre approach” and will combine distinctive scenario elements. One of these scenarios focuses on pilot monitoring performance. Based on what happened in the accident of Eastern Airlines Flight 401 [11, 12], the pilots will be instructed to artificially neglect a pre-determined set of relevant displays. As described in this paper, deviations from adequate monitoring will be detected and notifications will be displayed in the MCMD. Overall, the assessment of operational impact for the feedback system will focus on benefits in relation to pilot situation awareness, error detection and error management.

VI. Discussion

The A-PiMod project investigates novel team-centered concepts for pilot-automation interaction. These team-centered concepts are built around the vision that automation should be considered and designed as a supporting team member in the cockpit, which does not only observe and assess the flight situation, but also observes and assesses the mental states of the pilots, and understands how pilot behaviors fit to the demands of the pertinent tasks and mission objectives. In general, our results of previous evaluation research on technology developed in the A-PiMod project indicate that pilots are open to novel technology and concepts, which can improve flight safety.
Careful attention needs to be given to the technology and measures used to form a picture of pilot state and augment pilot performance. This refers, e.g., to the reliability and robustness of recording and assessment technology. Eye tracking is actually a very interesting technology to observe pilot monitoring performance. However, we found that the technology needs improvements to be reliable and robust inside an aircraft cockpit. An important constraint is the fact that an aircraft cockpit is often a dark environment, which can affect the eye tracking measure. Incorrect measurement can of course lead to incorrect assessments and this may cause unappreciated notifications on the MCMD. Further, the function used to assess pilot monitoring performance described in this paper depends on what we know about the pilot actual monitoring performance, i.e., if they are or are not looking at the right information sources. Hereby, the presented approach does not consider individual differences, e.g., pilot experience. It may not be a problem if an experienced pilot is not monitoring the navigation display in the defined frequency, e.g., because the pilot is very familiar with the route. Further, we have to think about how the technology presented in this paper can be applied on a two pilot crew. Given the case that two pilots can be observed and assessed, the question arises how individual feedback on monitoring performance can be provided to each crew member. At the moment, notifications are issued on the MCMD, which is a display shared by the pilot crew.

VII. CONCLUSIONS

The observation, assessment and comprehension of the mental states of the pilots, and how they behave in relation to pertinent tasks and mission objectives allows to design team-centered solutions for pilot-automation interaction, where the automation is considered as a supporting team member in the cockpit. The feedback system described in this paper is one contribution to this overall vision. It will, e.g., detect deviations from adequate monitoring and notify pilots – just like a good crew member would do – that the pilot should be more attentive. The evaluation of the overall concept developed in the A-PiMod project, and as part of this also the described feedback system, will be performed during the second validation cycle happening in November 2015.

ACKNOWLEDGMENT

The work described in this paper was carried out under the umbrella of the A-PiMod project. A-PiMod is funded by the European Commission Seventh Framework Programme under contract number: 605141 Project A-PiMod.

REFERENCES