

# Aircraft Pilot Intention Recognition for Advanced Cockpit Assistance Systems

Stefan Suck<sup>(✉)</sup> and Florian Fortmann

OFFIS Institute for Information Technology, Escherweg 2,  
26121 Oldenburg, Germany  
{stefan.suck,florian.fortmann}@offis.de

**Abstract.** Present aircraft are highly automated systems. In general, automation improved aviation safety significantly. However, automation exhibits itself in many forms of adverse behaviors related to human factors problems. A major finding is that insufficient support of partnership between the pilot crew and the aircraft automation can result in conflicting intentions. The European project A-PiMod (Applying Pilot Models for Safer Aircraft) addresses issues of conventional automation in the aviation domain. The overall objective of the project is to foster pilot crew-automation partnership on the basis of a novel architecture for cooperative automation. An essential part of the architecture is an intention recognition module. The intention recognition module employs a Hidden Markov Model (HMM) to infer the most probable current intention of the human pilots. The HMM is trained and evaluated with data containing interactions of human pilots with the aircraft cockpit systems. The data was obtained during experiments with human pilots in a flight simulator.

**Keywords:** Aircraft crew · Intention recognition · Markov Model

## 1 Introduction

Present aircraft are highly automated systems. In general, automation improved aviation safety significantly. However, automation exhibits itself in many forms of adverse behaviors, such as automation induced complacency [11], automation bias [14], decision making errors [15], lack of system knowledge and manual control skills [12], overconfidence [20], and vigilance issues [2]. A major finding is that insufficient support of partnership between the pilot crew and the aircraft automation can result in conflicting intentions. An example is the crash of China Airlines Flight 140 [18], where conflicting intentions of the human crew and the automation ended in a tragedy.

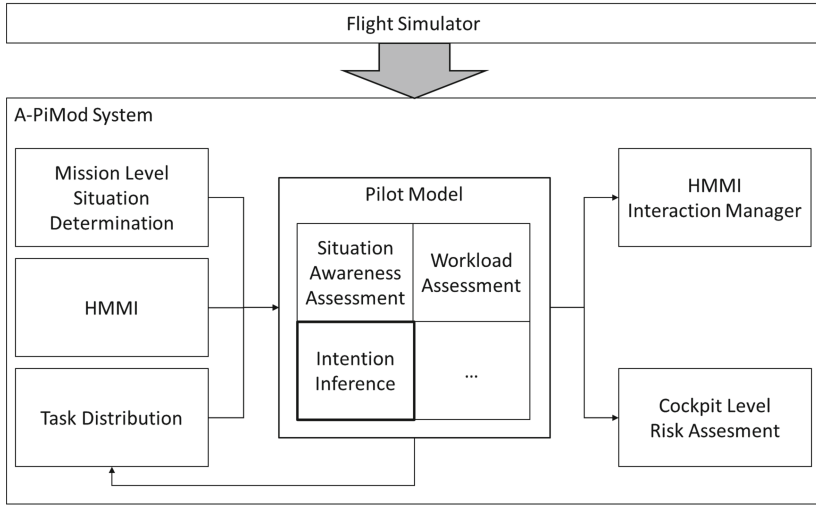
A recent research trend to overcome the problems of conventional automation is cooperative automation. The idea is that the human and the automation constitute a cooperative system in charge of the tasks of a jointly performed mission. Flemisch et al. [6] defined the term “cooperative” by key properties

such as sufficient skills to perception, cognition, action, interaction, conflict handling, internal and external compatibility, and balance between adaptivity and stability. In a cooperative automation system, humans and automation are considered as cognitive agents of a work system interacting with each other in order to achieve shared tasks. In this sense, humans and automation build a team. The task share in the human-automation team is the result of a negotiation process between the acting agents and depends on the availability and capacity utilization of resources.

The European project A-PiMod (Applying Pilot Models for Safer Aircraft) [1] addresses issues of conventional automation in the aviation domain. The overall objective of the project is to foster pilot crew-automation partnership on the basis of a novel architecture for cooperative automation. To improve the partnership between the pilots and the automation the project aims to develop a virtual crew member, which takes the position of classical aircraft automation. In the context of this project a pilot model is developed, which enables the virtual crew member to gain knowledge about the cognitive state of the human pilots. The cognitive state consists of different sub-states, which have been defined during the initial phase of the project, i.e., intentions, workload, and situation awareness. The sub-states are inferred in real-time based on data recorded during flight.

Figure 1 shows a simplified version of the A-PiMod architecture (see [8] for the holistic version). The focus of this version is on the intention inference module, and the connections of the pilot model to other selected A-PiMod modules. The Mission Level Situation Determination and the Human Machine Multimodal Interface (HMMI) deliver the real-time data from the flight simulator, data about the interactions with the cockpit systems, and the current flight phase. The Task Distribution module gives information about how pertinent flight tasks should be distributed between the human pilots and the automation in terms of risk and with regard to the abilities of the pilots and the automation. The Cockpit Level Risk Assessment continuously assesses the risk for all currently possible task distributions. The results of the intention recognition influence the Situation Awareness Assessment of the pilot model, the Cockpit Risk and eventually a new Task Distribution. The HMMI Interaction Manager adapts the output modality for certain information and the salience of certain displays and display elements in the cockpit. This adaption is also dependent on the pilot state.

In this paper, we describe the intention recognition module which is an important part of the pilot model of the A-PiMod architecture. It provides the virtual pilot with a sufficient understanding of the flight tasks the human pilots are carrying out. Understanding what the human pilots are doing allows the virtual pilot to make sense of the behaviors of the human pilots in context of the present situation, and allows to detect and to mitigate conflicting intentions. The intention recognition module employs a Hidden Markov Model (HMM) to infer the most probable current intention of the human pilots. The HMM is trained and evaluated with data containing interactions of human pilots with the aircraft cockpit systems. The data was obtained during experiments with human pilots in a flight simulator.



**Fig. 1.** Overview of the A-PiMod architecture with focus on the pilot model.

The paper is structured as follows: Sect. 2 provides an overview of aircraft pilot models and some applications of probabilistic behavior models. Section 3 explains our experimental setup and data collection approach. Section 4 describes the HMM used to recognize pilot intentions. Section 5 shows the evaluation results. Section 6 provides conclusions.

## 2 Related Work

There are already several approaches in the aviation domain to model human operators and their behavior. The Cognitive Architecture for Safety Critical Task Simulation (CASCaS) presented in [10] can be understood as a rule based approach of an operator model. CASCaS was applied in several projects among other things as aircraft pilot model which generates behavior and performs tasks similar to a human pilot. It is feasible for simulating human-machine interaction in highly dynamic environments. Contrary to our approach CASCaS is not probabilistic and generates actions and intentions of a pilot instead of monitoring the actions and inferring the associated intentions. Another approach is the adaptive pilot model of the Crew Assistant Military Aircraft (CAMA), which is described in [19]. This pilot model determines if deviations from a normative pilot model are real errors or are intended by the pilot. This approach is based on fuzzy rules and the focus is on manually flying of a military aircraft. So, interactions with the autopilot system of the aircraft are not considered. A probabilistic approach to model one part of the pilots behavior is presented in [7]. The author uses Hidden Markov Models to analyze the instrument scanning and attention switching behavior of aircraft pilots. In contrast to our approach eye-tracker data

employed whereby our approach is based on physical interactions with, e.g., buttons. Additionally our focus is not on the scanning behavior but on the flight tasks.

In other domains, there are already some similar efforts. In [4] a probabilistic model for the behavior of an Unmanned Aerial Vehicle (UAV) operator is presented. The operator is not actually flying the UAV but has to identify and tag several objects in the live stream coming from the UAV. The author trains HMMs which are then used to infer on the behavior of the UAV operator and the currently performed task by monitoring their User Interface interactions.

In the automotive domain the recognition of driver intentions or driving maneuvers is an important aspect. In [5] an overview of several approaches is given, including approaches which employ probabilistic networks. For example, the authors of [13] use a hierarchical structure of Dynamic Bayesian Networks (DBN) as driver model, to generate driving actions from driving goals. It is also possible to use this approach to monitor the driver's actions to infer on the driving goals.

In the smart home domain an important technology to monitor a person's health is to detect and to track the currently performed Activities of Daily Living of the resident. In different approaches the activities are usually inferred from the observed interactions with tagged objects or areas. For example, the approaches presented in [9, 17] employ Hidden Markov Models to recognize interleaved activities of smart home residents.

### 3 Experiment

In order to collect the necessary data to train and validate our intention inference module we performed several experiments in the Generic Experimental Cockpit (GECO) of the German Aerospace Center (DLR) in Braunschweig. The GECO is a modular cockpit simulator with interchangeable flight-mechanical models to fit the needs of different applications. It provides a generic work environment, which is able to represent any of the state-of-the-art cockpits which are currently produced by the different aircraft manufacturers. With its several hardware and software modules it forms a fix-base experimental flight simulator with many features. Figure 2 shows a view of the GECO with an A320 layout and some additional installations which were used during the A-PiMod validation experiments.

The GECO is mainly used to perform simulations with human test subjects in the loop to evaluate new display and control concepts. In contrast to simulators designed for pilot training the GECO does not strive for the highest degree of realism for one particular type of aircraft. The major objective of the GECO is to provide maximum flexibility. Thus, it can meet different requirements in the fields of cockpit research regarding new systems with human-machine interfaces and new flight procedures.

We invited six professional aircraft crews to fly scenarios in the GECO. Each crew consisted of one captain and one first officer. All these pilots were trained

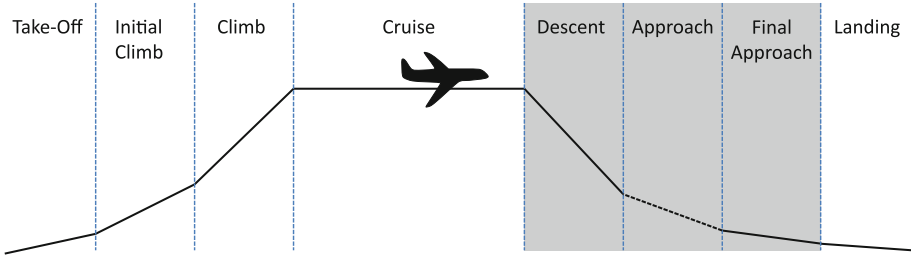


**Fig. 2.** GECCO flight simulator used for our experiments and the collection of data

and experienced in flying an Airbus A320 or a similar model. The scenarios were prepared primarily for the evaluation and the demonstration of the concepts of the A-PiMod project. For this reason the scenarios followed a “theater approach” but they did also fit our needs to gather the required data and to evaluate our intention inference module. Every scenario was constructed from a set of scenario elements and basically contained similar procedures. During each element the pilots had to perform some flight tasks. Each scenario had a duration of about 15 min and started in the descent flight phase. At this point the so called cruise phase is finished; the aircraft has left its travel altitude and is descending to approach an airport. For a better understanding, Fig. 3 illustrates the common order of the flight phases.

During one scenario the pilots first had to adapt the initially programmed flight route and after a while they had to divert to an alternative airport. Another scenario contained a missed approach. Meaning the aircraft was not in the correct configuration at a certain point of the approach to the airport. Therefore, the pilots were required to perform a Go-Around procedure to try a new approach. After the Go-Around procedure the pilots also had to adapt the programmed flight route.

While the pilots were flying the scenarios the data from the flight simulator and the different cockpit components were recorded continuously. This data contains all pilot interactions with the cockpit. Later on the data was annotated by an expert to mark the flight tasks.



**Fig. 3.** Common order of the flight phases. The flight phases of the experiment are marked with a gray background.

## 4 Modeling Approach

The intentions of a pilot usually refer to the goals he is trying to achieve [3]. To accomplish a specific goal usually a more or less complex plan exists. Complex plans can be separated into sub-plans. This way the complex plans become the goals of their sub-plans. Flight tasks which occur in an aircraft cockpit also serve the achievement of a goal. The complexity of these tasks can reach from very basic activities like changing the altitude to more complex operations like performing a go-around procedure. Complex tasks can also be separated into sub-tasks and become the goals of their sub-task. Therefore, we interpret the intentions of a human pilot as the flight tasks he is currently performing. For the execution of a task a pilot has to show a specific behavior, which means he has to perform certain interactions with the cockpit systems. Thus, there is some set of actions which is typical for this task. These actions occur in a sequence and are directly observable, while the tasks cannot be observed directly. By monitoring the interactions it is possible to infer the currently performed flight task. However, some interactions can belong to one or more flight task. Theoretically flight tasks can also occur in an interleaved manner. This means the pilot interrupts one flight task, executes another task and then resumes the previously carried out flight task. In this paper we infer 8 basic flight tasks from observable interactions of pilots with the aircraft systems.

We use a 1st order Hidden Markov Model to infer the currently performed flight task from the observed interactions with the cockpit system. HMMs are used to describe dynamic probabilistic processes [16]. The model contains a set of different states  $S = S_1, S_2, \dots, S_N$  and a set of possible emissions  $O = O_1, O_2, \dots, O_M$  which both occur as a sequence. The states are not directly observable and are therefore usually called “hidden states”. In a 1st order HMM the assumption is made that the current state at a certain point in time only depends on one previous state. So, the conditional probability of a certain state at the time  $t$  after a sequence of previous states can be simplified as shown in Eq. 1.

$$P(S_t | S_1, \dots, S_{t-1}) = P(S_t | S_{t-1}) \quad (1)$$

Each state at a point in time generates an observable emission  $O_t$ . This observation depends only on the current state  $S_t$  and not on previous observations. Thus, the probability to make a certain observation at time  $t$  after a sequence of previous states an observation can be simplified as shown in Eq. 2.

$$P(O_t|O_1, \dots, O_{t-1}, S_1, \dots, S_t) = P(O_t|S_t) \quad (2)$$

For simplification it is assumed that at the start of the process at  $t = 0$  the start state is  $S_0 = 0$ . The first transition from  $S_0$  to  $S_1$  then the first emission is  $O_1$ . The whole behavior of the HMM can be characterized with three probability matrices. The initial state probabilities  $a_{0i} = P(S_1|S_0 = 0)$ , the state transition probabilities  $a_{ij} = P(S_t = j|S_{t-1} = i)$  and the emission probabilities  $b_{jm} = P(O_t = m|S_{t-1} = j)$ . The joint distribution of a sequence of observations  $\mathbf{O} = O_1, O_2, \dots, O_T$  and a sequence of states  $\mathbf{S} = S_1, S_2, \dots, S_T$  can then be written as Eq. 3.

$$P(\mathbf{S}, \mathbf{O}) = \prod_{t=1}^T P(S_t|S_{t-1})P(O_t|S_t) \quad (3)$$

In our context the flight tasks are the hidden states of the HMM since the task are not directly observable during run-time. For the training process the states are known since we are using annotated training data. The emissions are the observed interactions of the pilot with the cockpit systems. The initial state probability matrix is generated from the frequencies of the different flight tasks in the trainings data. The state transition probability matrix is generated from the frequencies of the flight task transitions in the trainings set. The emission probability matrix is generated from the frequencies of the combinations of flight tasks and interactions in the trainings set. For the calculation of the state transition probabilities and the emission probabilities Laplace smoothing is applied. Whereby  $\#x_k$  denotes the counts of a state transition or the counts of an emission from a certain state and  $\#X = \sum_{k=1}^L \#x_k$  is the count of all state transitions or emissions from a certain state in the training data. Then the smoothed estimated probability for this state transition or emission can then be written as Eq. 4.

$$\hat{x}_k = \frac{\#x_k + 1}{\#X + L} \quad (4)$$

By doing so we avoid zero probabilities for flight task transitions which are possible but were not observed in the training data. The model would also be more robust against cases of a task execution which slightly differs from the standard procedure.

For our model we focus on 8 flight tasks. These are relevant for scenarios which were used during our simulator experiments mentioned in Sect. 3. The flight tasks are “monitor”, “arm spoiler”, “set flaps”, “set thrust”, “change speed”, “change heading” “change altitude”, and “set approach”. The task “monitor” actually means that there is none of the other defined tasks ongoing. So, we assume the pilot is just monitoring the instruments. The monitored interactions with the cockpit systems are actions like pushing a button, turning a

knob, or pulling a lever. For developing the intention recognition model, we used a knowledge engineering approach. We interviewed domain experts to gather knowledge about the relations between flight tasks and cockpit interactions, and certain context events which could be triggers for certain tasks. The parameters of the HMM are learned from manually annotated data. This data was obtained by recording the aircraft data and the interactions of professional pilots while they were flying the scenarios described in Sect. 3 in the GECO simulator with activated autopilot system.

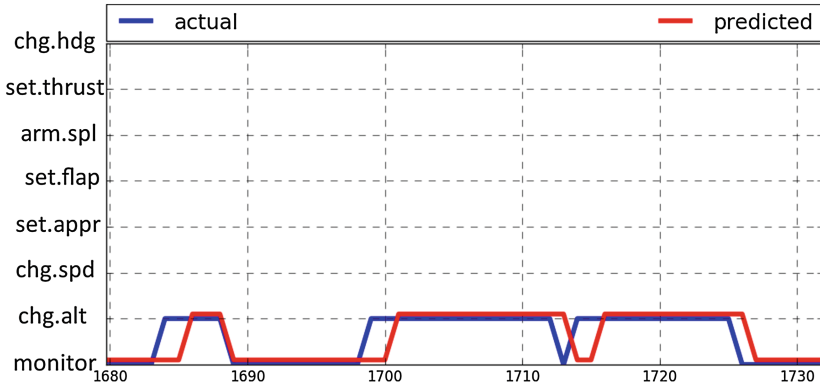
## 5 Evaluation

For the evaluation the package *HMM* for the programming language *R* was used. The flight data which was recorded during our simulator experiments was split into a training set and a test set. The test set contains two of the eight recorded scenarios which were useful. Since our data consist of very long time series it would have been impractical to perform a cross-validation with random selection of test cases. For each data point in our test set the annotated task is compared to the task with the highest probability  $P(S_t|O_{1:t})$  reported by our HMM. The model produces an error rate of 0.0069 on the test set. This means for 0.69 % of the cases in the test set the most probable tasks of the model does not match with the task given in the test set. Since our test data is strongly populated with the state ‘monitor’ we also calculated the error rate based only on the cases where the test data is not ‘monitor’. For this configuration our model produces an error rate of 0.0039. The error rates and Fig. 4 show that the model expresses the data surprisingly well. This is due to the relative high degree of determinism which is currently present in the data and in the current set of flight tasks and interactions. Figure 4 also shows that the predicted task is detected somewhat after the start of the annotated. A possible reason is that the human annotator saw more than the just the bare interaction. He perceived the context and was also able to already see the pilot reaching for certain interfaces (Table 1).

**Table 1.** Confusion matrix of task predicted by the HMM and the actual task as marked in the data.

Predicted									
		<i>monitor</i>	<i>arm.spl</i>	<i>set.flap</i>	<i>set.thrust</i>	<i>chg.spd</i>	<i>chg.hdg</i>	<i>chg.alt</i>	<i>set.appr</i>
Actual	<i>monitor</i>	21308	0	0	0	0	0	2	0
	<i>arm.spl</i>	18	12	0	0	0	0	0	0
	<i>set.flap</i>	36	0	30	0	0	0	0	0
	<i>set.thrust</i>	27	0	0	64	0	0	0	0
	<i>chg.spd</i>	53	0	0	0	123	0	0	0
	<i>chg.hdg</i>	2	0	0	0	0	3	0	0
	<i>chg.alt</i>	8	0	0	0	0	0	28	0
	<i>set.appr</i>	5	0	0	0	0	0	0	0





**Fig. 4.** Segment of the sequence of tasks detected by HMM and actual tasks as marked in the data

## 6 Conclusion

We presented a model to recognize the currently performed flight tasks of aircraft pilots by monitoring their interactions with the cockpit systems. The model uses a trained HMM which is integrated into the A-PiMod architecture as a part of the pilot model. The designated usage of the model is to receive new observations at a rate of 20 Hz during simulator flights and to infer  $P(S_t|O_{1:t})$ . The evaluation of our model showed that the model is able to detect our selected flight tasks in the flight phases ‘descent’, ‘approach’ and ‘final approach’ based on interactions with the cockpit systems. In the future we plan to detect more complex tasks with the model and eventually to add eye-tracker data and context information. This could result in a performance decrease compared to the current model but it makes the model more useful.

**Acknowledgment.** The A-PiMod project is funded by the European Commission Seventh Framework Programme (FP7/2007-2013) under contract number: 605141 Project A-PiMod. Thanks also to my colleague Mark Eilers for his advice during the creation of this work.

## References

1. Applying pilot models for safer aircraft. <http://www.apimod.eu>
2. Caldwell, J.A.: Crew schedules, sleep deprivation, and aviation performance. *Curr. Dir. Psychol. Sci.* **21**(2), 85–89 (2012)
3. Cohen, P.R., Levesque, H.J.: Intention is choice with commitment. *Artif. Intell.* **42**, 213–261 (1990)
4. Donath, D.: Verhaltensanalyse der Beanspruchung des Operators in der Multi-UAV-Führung. Dissertation, Universität der Bundeswehr München (2012)

5. Doshi, A., Trivedi, M.M.: Tactical driver behavior prediction and intent inference: a review. In: 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 1892–1897. IEEE (2011)
6. Flemisch, F., Meier, S., Neuhöfer, J., Baltzer, M., Altendorf, E., Özyurt, E.: Kognitive und kooperative systeme in der fahrzeugführung: selektiver rückblick über die letzten dekaden und spekulation über die zukunft. *Kognitive Systeme* (2013)
7. Hayashi, M.: Hidden Markov Models for analysis of pilot instrument scanning and attention switching. Ph.D. thesis, Massachusetts Institute of Technology (2004)
8. Javaux, D., Fortmann, F., Möhlenbrink, C.: Adaptive human-automation cooperation: a general architecture for the cockpit and its application in the a-pimod project. In: 7th International Conference on Advanced Cognitive Technologies and Applications (COGNITIVE 2015). International Academy, Research, and Industry Association (IARIA) (2015)
9. Landwehr, N.: Modeling interleaved hidden processes. In: Cohen, W.W., McCallum, A., Roweis, S.T. (eds.) ICML. ACM International Conference Proceeding Series, vol. 307, pp. 520–527. ACM (2008)
10. Lüdtke, A., Weber, L., Osterloh, J.-P., Wortelen, B.: Modeling pilot and driver behavior for human error simulation. In: Duffy, V.G. (ed.) Digital Human Modeling, HCII 2009. LNCS, vol. 5620, pp. 403–412. Springer, Heidelberg (2009)
11. Manzey, D., Reichenbach, J., Onnasch, L.: Human performance consequences of automated decision aids: the impact of degree of automation and system experience. *J. Cogn. Eng. Decis. Making* **6**, 57–87 (2012)
12. McGee, J.P., Mavor, A.S., Wickens, C.D., et al.: Flight to the Future: Human Factors in Air Traffic Control. National Academies Press, Washington (1997)
13. Moebus, C., Eilers, M.: Prototyping smart assistance with bayesian autonomous driver models. In: Chong, N.Y., Mastrogiovanni, F. (eds.) Handbook of Research on Ambient Intelligence and Smart Environments: Trends and Perspectives, pp. 460–512. IGI Global, May 2011
14. Mosier, K.L., Skitka, L.J., Heers, S., Burdick, M.: Automation bias: decision making and performance in high-tech cockpits. *Int. J. Aviat. Psychol.* **8**(1), 47–63 (1998)
15. Orasanu, J., Martin, L., Davison, J.: Errors in Aviation Decision Making: Bad Decisions or Bad Luck? National Aeronautics and Space Administration, Ames Research Center (1998)
16. Rabiner, L.R.: A tutorial on hidden markov models and selected applications in speech recognition. *Proc. IEEE* **77**, 257–286 (1989)
17. Singla, G., Cook, D.J.: Interleaved activity recognition for smart home residents. In: Proceedings of the 5th International Conference on Intelligent Environments, IE 2009, Barcelona, Spain, pp. 145–152 (2009)
18. Sogame, H., Ladkin, P.: Aircraft accident investigation report 96-5. Japan: Ministry of transport (1996). <http://sunnyday.mit.edu/accidents/nag-1.html>
19. Strohal, M., Onken, R.: Intent and error recognition as part of a knowledge-based cockpit assistant. In: Proceedings of the SPIE, vol. 3390, pp. 287–299 (1998)
20. Wood, S.: Flight crew reliance on automation. CAA Paper 10 (2004)