

## **Evidenzbasierte Pilotentätigkeitserkennung unter Berücksichtigung unterschiedlich zuverlässiger Beobachtungen**

Fabian Honecker & Axel Schulte

### **Zusammenfassung**

Workload-adaptive Assistenzsysteme sollen den Menschen in kritischen Beanspruchungssituationen möglichst aufgabenorientiert unterstützen. Die Beanspruchung wird dabei im Wesentlichen durch die Pilotentätigkeit, also denjenigen Aufgaben, mit denen sich der Pilot gegenwärtig beschäftigt, bestimmt. Um ein Workload-adaptives Assistenzsystem zu realisieren, ist es deshalb erforderlich, diese Aufgaben maschinell und in Echtzeit zu ermitteln, was den Fokus dieses Beitrags darstellt. Dies soll im Umfeld bemannt-unbemannter Hubschraubermissionen untersucht werden. Als Methode wird dabei eine vereinfachte Version der Evidenztheorie von Dempster und Shafer angewendet. Sie erlaubt eine echtzeitfähige, flexible und robuste Erkennung von Tätigkeiten auf Basis von Beobachtungen und ist dabei in der Lage, auch mit widersprüchlichen Beobachtungen umzugehen.

### **1 Einführung und Problemstellung**

Die Führung mehrerer unbemannter Luftfahrzeuge (UAVs) aus dem Cockpit eines bemannten Hubschraubers ist eine sehr komplexe Aufgabenstellung und kann die menschlichen Operateure (Piloten) je nach Situation und Grad der Automatisierung über-, aber auch streckenweise unterfordern. Der Mensch besitzt nur dann die beste Leistungsfähigkeit, wenn sich seine Beanspruchung (*Mental Workload MWL*) auf einem mittleren Niveau befindet (Veltman & Jansen, 2006; Young et al., 2015). Außerhalb dieses Bereiches kommt es zu einem Einbruch der Leistung des Operateurs und folglich auch zu einer geringeren Gesamtmissionsleistung. Bild 1 visualisiert diesen Sachverhalt.

Ein Ansatzpunkt, um diesem Problem zu begegnen, ist die Unterstützung von Operateuren (Piloten) durch Workload-adaptive kognitive Assistenzsysteme. Diese sollen den Menschen in abweichenden Beanspruchungssituationen wie Unter- oder Überbeanspruchung durch das Ergreifen entsprechender Maßnahmen (z.B. Aufmerksamkeitslenkung, Aufgabenvereinfachung, Übernahme von ganzen oder Teilaufgaben) unterstützen, um so die Beanspruchung des Menschen wieder auf ein mittleres Niveau zu bringen (Donath, 2012). Eine Abschätzung der Beanspruchung lässt sich zum Beispiel über Ressourcenmodelle unter Kenntnis

der durch den Operateur bearbeiteten Aufgaben durchführen (Maiwald, 2013). Darüber hinaus ist die Kenntnis der Pilotentätigkeit wichtig um weitere „Zustandsgrößen“ des Menschen wie Aufmerksamkeit (Rouse & Rouse, 1983), Absicht (Wittig, 1994) oder Verhalten (Ruckdeschel, 1997) abschätzen zu können und dem Piloten eine individuelle, adaptive Hilfestellung zu ermöglichen.

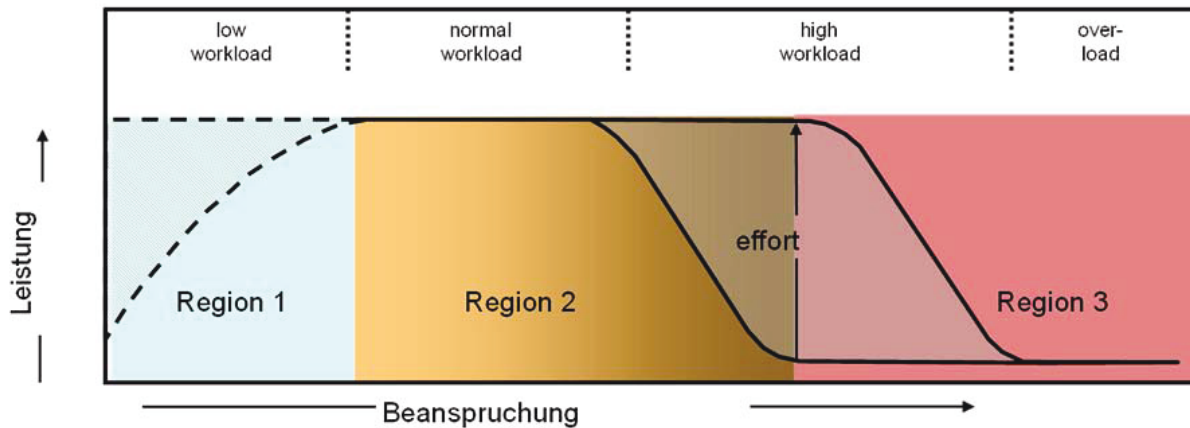


Bild 1: Zusammenhang Leistung-Beanspruchung nach Veltman & Jansen (2006)

Schulte et al. (2015) zeigen ein Konzept auf, das die Beanspruchung hinsichtlich mehrerer Gesichtspunkte operationalisiert. Ein zentraler Punkt ist dabei die Pilotentätigkeit, welche zusammen mit dem Aufgabenkontext die Workload maßgebend beeinflusst.

In dieser Arbeit werden unter der Pilotentätigkeit diejenigen Aufgaben verstanden, die der Pilot während der Missionsdurchführung zu einem bestimmten Zeitpunkt gleichzeitig bearbeitet. So kann ein Pilot beispielsweise gleichzeitig den Hubschrauber steuern und den Flugfunksprechverkehr durchführen. Diese Tätigkeiten lassen sich im Allgemeinen nicht direkt messtechnisch erfassen, sondern können nur indirekt logisch erschlossen werden.

Dazu liefert die Beobachtung der Interaktionen des Operateurs mit dem technischen System und der Umwelt (Aufgabenkontext) wichtige Evidenzen für die Pilotentätigkeit. Evidenzen sind beobachtbare Fakten, die auf die (nicht direkt beobachtbare) Bearbeitung einer Aufgabe durch einen Piloten schließen lassen. Diese werden im vorliegenden Fall über unterschiedliche verfügbare Sensoren im Cockpit wie Steuerknüppel, Schalter, Taster, Mikrofone oder ein Blickbewegungsmesssystem messtechnisch erfasst. Jeder dieser Sensoren liefert dabei Messdaten verschiedener Güte. Die aus den verschiedenen Informationsquellen gewonnen Evidenzen besitzen unterschiedliche Zuverlässigkeit und können gegebenenfalls auch widersprüchlich sein. Ein Beispiel sei die Aufgabe „Flugfunksprechverkehr durchführen“: Der Pilot drückt dabei zunächst den Funkknopf, was die Hypothese „Pilot führt Flugfunksprechverkehr durch“ unterstützt. Während der Pilot den Funkknopf drückt, spricht er in ein Mikrofon. Die Beobachtung „Pilot spricht“ wird maschinell dadurch erkannt, dass die Lautstärke einen gewissen Pegel überschreitet. Spricht der Pilot zu leise kann es passieren,

dass das System aufgrund eines Messfehlers dies nicht erkennt und geht davon aus, dass der Pilot nicht spricht. Dies lehnt die Hypothese „Pilot führt Flugfunkprechverkehr durch“ ab.

Ein maschinelles Verfahren, das auf Basis von Evidenzen, bzw. Kombination dieser Evidenzen, auf Tätigkeiten schließt, muss folglich in der Lage sein, diese Randbedingung entsprechend verarbeiten und berücksichtigen zu können. So dürfen beispielsweise Evidenzen von Quellen niedriger Zuverlässigkeit das Endergebnis nicht im Übermaß beeinflussen, da diese unter Umständen zu falschen Annahmen des Assistenzsystems über die vom Piloten ausgeführten Tätigkeiten und damit in Zusammenhang stehenden Aufgaben führen können. Dies resultiert letztendlich in unbrauchbaren Hinweisen, Handlungsempfehlungen oder ungünstiger Lenkung der Aufmerksamkeit des Piloten durch das Assistenzsystem.

Aber nicht nur die Sensoren sind mit Fehlern behaftet, auch das menschliche Verhalten selbst kann fehlerhaft sein. Der Mensch kann Aufgaben vergessen oder versehentlich falsche Interaktionen durchführen, die falsche Messwerte zur Folge haben. Trotzdem soll das Gesamtsystem robust sein und die Pilotentätigkeit noch richtig erkennen können. Wie stark eine fehlerbehaftete Beobachtung zum Gesamtergebnis beiträgt, wird in einem Evidenzmodell berücksichtigt (Abschnitt 3.1.3).

## 2 Methode

### 2.1 Methodenauswahl

Gemäß den oben genannten Randbedingungen, muss eine geeignete Methode mit Unsicherheiten umgehen können und robuste Ergebnisse liefern. Methoden, die Unsicherheiten berücksichtigen, sind beispielsweise Verfahren wie Bayes-Netze (Pearl, 1988), Sicherheitsfaktoren (Shortliffe, 1976) oder Verfahren basierend auf der Dempster-Shafer-Theorie (Dempster, 1967; Shafer, 1976). Wahrscheinlichkeitsbasierte Verfahren wie Bayes-Netze haben das Problem, dass vor dem Einsatz des Modells die Modellparameter vollständig bestimmt sein müssen. Das heißt alle a-priori und bedingte Wahrscheinlichkeiten müssen vorab aufwändig durch Häufigkeitsanalysen in Wissenserwerbsexperimenten ermittelt werden. Das Sicherheitsfaktorenmodell ist hingegen ein heuristisches Verfahren, bei dem die Modellparameter durch Expertenbefragung gewonnen werden können. Ein Anwendungsbeispiel für Pilotenassistenzsysteme findet man in Wittig (1994).

Diese Arbeit stellt ein Konzept auf Basis einer vereinfachten Version der Evidenztheorie von Dempster und Shafer vor, welches mit der intrinsischen Modellierung von Unwissen zu einem robusteren System führen kann. Es verbindet die Vorteile der Wahrscheinlichkeitstheorie mit der Dempster-Shafer-Theorie: Modellparameter können entweder durch Häufigkeitsanalysen als Wahrscheinlichkeiten bestimmt werden oder durch Expertenbefragung als

Glaubenswerte (Abschnitt 2.2). Darüber hinaus müssen nicht alle Parameter a-priori vollständig ermittelt werden.

## 2.2 Berücksichtigung von Unsicherheiten und Darstellung als Glaubensmaße gemäß der Evidenztheorie von Dempster und Shafer

Die Tätigkeitsinferenz geschieht auf Basis einer vereinfachten Version der Evidenztheorie von Dempster und Shafer, welche logisches Schließen unter Unsicherheit ermöglicht. Es wird die Nomenklatur nach Dempster (2008) verwendet.

### 2.2.1 Glaubenstriplet

Die Evidenztheorie kann als Erweiterung der Wahrscheinlichkeitstheorie aufgefasst werden. Das Intervall zwischen 0 und 1 wird nicht in die beiden Teile Wahrscheinlichkeit  $P$  und Gegenwahrscheinlichkeit  $1 - P$  aufgeteilt, sondern um einen dritten Teil, dem Unwissen, erweitert (Bild 2).

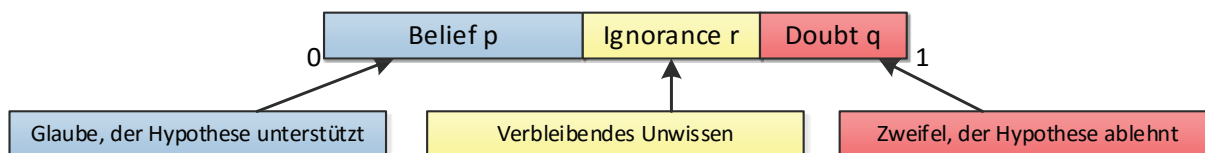


Bild 2: DST-Glaubenstriplet  $Q = (p, q, r)$ .

Anstelle von Wahrscheinlichkeiten  $P$  werden in dieser Arbeit Glaubenstriplets der Form  $Q = (p, q, r)$  analog zu Dempster (2008) verwendet. Jede Zustandsvariable  $A$  wird durch Glaubenstriplet  $Q(A)$  beschrieben. Dabei gilt die Normierungsbedingung

$$p + q + r = 1 \quad (1)$$

Werden Hypothesen durch Glaubenstriplets beschrieben, so haben die Glaubensfunktionen folgende Bedeutung (Interpretation nach Shafer, 1976):

- $p$ : *Belief*: Glaube, der eine Hypothese unterstützt.
- $q$ : *Doubt*: Zweifel, der die Hypothese ablehnt.
- $r$ : *Ignorance*: Verbleibendes Unwissen.

### 2.2.2 Bedingte Glaubenswerte

Ähnlich wie in Bayes-Netzen (Pearl, 1988) können bedingte Glaubenswerte analog zu bedingten Wahrscheinlichkeiten betrachtet werden und als Tabelle oder Matrix bedingter Glaubenswerte dargestellt werden:

$$\mathbf{M}(B|A) = \begin{pmatrix} p_t & p_f & 0 \\ q_t & q_f & 0 \\ r_t & r_f & 1 \end{pmatrix} \quad (2)$$

$A$  und  $B$  sind Zustandsvariablen.  $M(B|A)$  liefert die Glaubenswerte für Zustand  $B$ , falls der Zustand  $A$  bekannt ist. Die Komponenten haben folgende Bedeutung:

$p_t$	Glaube, dass $B = true$	falls $A = true$ .
$q_t$	Glaube, dass $B = false$	falls $A = true$ .
$r_t$	Glaube, dass $B = unknown$	falls $A = true$ .
$p_f$	Glaube, dass $B = true$	falls $A = false$ .
$q_f$	Glaube, dass $B = false$	falls $A = false$ .
$r_f$	Glaube, dass $B = unknown$	falls $A = false$ .

Es gelten die Normierungsbedingungen  $p_t + q_t + r_t = 1$  und  $p_f + q_f + r_f = 1$ .

### 2.2.3 Modellierung von Unwissen

Speziell die Zuverlässigkeit  $Z$  (hier Wert zwischen 0 und 1), beziehungsweise der relative Messfehler eines Sensorwerts, kann durch Unwissen  $r = 1 - Z$  als integriertes Maß im Modell ausgedrückt werden. Der Sensor liefert nur dann Glaubenswerte, wenn er sich sicher ist. Ein Assistenzsystem kann dieses Wissen um Nichtwissen nutzen und nur dann Hinweise geben, wenn es sich sicher ist, dass es etwas weiß. Auf diese Weise sollen Fehlentscheidungen des Piloten aufgrund falscher Assistenzsystemmeldungen minimiert werden.

Des Weiteren ist mit Hilfe der expliziten Berücksichtigung von Unwissen eine partielle Ausmodellierung des Evidenzmodell (Abschnitt 3.1.3) im Pilotenaufgabenmodell möglich, das heißt, das Modell kann bereits benutzt werden, bevor alle Modellparameter definiert wurden. Nicht ausmodellerte Parameter werden als Unwissen  $r$  dargestellt. Das Ergebnis bei der Aufgabenerkennung wird dann folgerichtig mit einem erhöhten Gesamtunwissen behaftet. Je besser das Evidenzmodell jedoch ausmodelliert wird, desto geringer wird das Unwissen.

Robustheit wird dadurch erreicht, dass nur das Wissen im Modell modelliert wird, das auch wirklich eine Aussage macht. Wenn kein Wissen hinterlegt wird, kommt anders als in einem binären Wahrscheinlichkeitsmodell Unwissen heraus. Es bleibt dann Interpretationssache des Assistenzsystems, wie es sich beim Vorliegen von Unwissen verhält.

## 3 Konzept

Die Erfassung des Pilotenzustands kann grundsätzlich als Klassifikationsproblem aufgefasst werden, bei dem für jede mögliche Aufgabe, die ein Pilot während einer typischen Mission durchführen kann, ermittelt wird, ob der Pilot diese gerade bearbeitet oder nicht. Das mögliche Aufgabenspektrum wird hierzu in einem Zustandsraummodell definiert (Abschnitt 3.1.1). Die Pilotentätigkeitserfassung geschieht durch systematische Beobachtung des Piloten und der Umwelt mit Hilfe von Messsensoren. Dabei wird berücksichtigt, dass jede Messung mit Messfehlern behaftet ist. Die Zuverlässigkeit der verschiedenen verwendeten Sensoren wird im Sensormodell (Abschnitt 3.1.2) modelliert. Bei der

Pilotenbeobachtung werden viele einzelne Evidenzen generiert, die auf die momentane Pilotentätigkeit schließen lassen. Die hierfür nötigen Modellparameter sind im Evidenzmodell hinterlegt (Abschnitt 3.1.3). Je mehr einzelne Evidenzen für eine Aufgabe kombiniert werden, desto besser können sich Messfehler ausmitteln. Dies steigert die Robustheit des Verfahrens. Die zeitliche Gültigkeit von Evidenzen wird durch ein Dynamikmodell modelliert (Abschnitt 3.1.4).

Das Zustandsraummodell, Evidenzmodell und Dynamikmodell sind Teil eines Pilotenaufgabenmodells, welches in Honecker & Schulte (2015) beschrieben wird.

### 3.1 Modelle

#### 3.1.1 Zustandsraummodell

Das Zustandsraummodell ist eine abstrakte Darstellung von Aufgaben, Evidenzen und Beobachtungen, auf dessen Grundlage mit mathematischen Methoden die Tätigkeit eines Piloten ermittelt werden kann. Es beinhaltet welche Zustandsvariablen zur Verfügung stehen und wie diese miteinander in Zusammenhang stehen.

In diesem Ansatz wird für jede Aufgabe  $X$  ein eigenes Zustandsraummodell erstellt. Jede Aufgabe kann durch mehrere Evidenzen  $E_j \in \{E_1, \dots, E_M\}$  erkannt werden, wobei jede Evidenz aus einer Beobachtung  $S_j \in \{S_1, \dots, S_M\}$  resultiert. Bild 3 stellt ein Beispiel für ein Modell mit drei Evidenzen als gerichteten Graph dar. Die Pfeile sind die Richtung der Informationsverarbeitung von den Sensoren  $S_j$  über die Evidenzen  $E_j$  bis hin zur Aufgabe  $X$  (diagnostische Richtung).

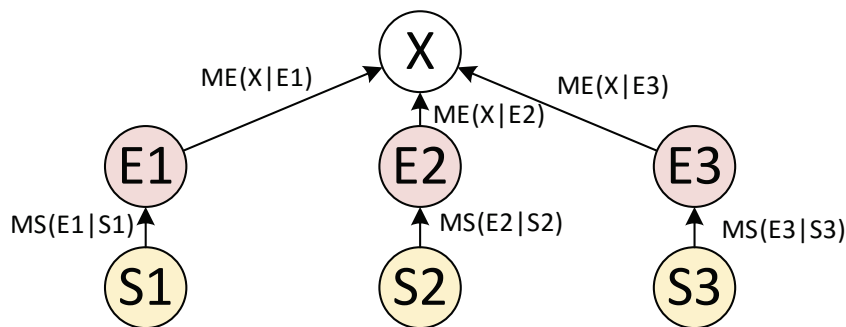


Bild 3: Zustandsraummodell für eine Aufgabe.

Jede Zustandsvariable  $X, E_j, S_j$  kann drei Zustände einnehmen (*true*, *false* oder *unknown*) und wird durch ein Triplet an Glaubenswerten (Abschnitt 2.2.1) beschrieben.  $M_E(X|E_j)$  ist das Evidenzmodell und  $M_S(E_j|S_j)$  das Sensormodell (Abschnitte 3.1.2 und 3.1.3).

#### 3.1.2 Sensormodell

Das Sensormodell ist ein Fehlermodell, das die Zuverlässigkeit des Sensors, beziehungsweise den relativen Messfehler, abschätzt. Jeder Sensor wird durch eine Tabellen bedingter Glaubenswerte (Abschnitt 2.2.2) dargestellt. Je nach

Detaillierungsgrad der Modellierung können verschiedene Sensormodelle definiert werden:

Allgemeiner Sensor:

$$\mathbf{M}_S(E_j|S_j) = \begin{pmatrix} p_t & p_f & 0 \\ q_t & q_f & 0 \\ r_t & r_f & 1 \end{pmatrix} \quad (3)$$

Perfekter Sensor:

$$\mathbf{M}_S(E_j|S_j) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4)$$

Sensor mit konstanter Zuverlässigkeit  $Z$  modelliert als Unwissen:

$$\mathbf{M}_S(E_j|S_j) = \begin{pmatrix} Z & 0 & 0 \\ 0 & Z & 0 \\ 1-Z & 1-Z & 1 \end{pmatrix} \quad (5)$$

Wahrscheinlichkeitsbasierter Sensor:

$$\mathbf{M}_S(E_j|S_j) = \begin{pmatrix} P_t & 1-P_f & 0 \\ 1-P_t & P_f & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (6)$$

Dabei ist die linke obere 2x2-Submatrix die bedingte Wahrscheinlichkeitstabelle, wie man sie in Bayes-Netzen (Pearl, 1988) für binäre Zufallsvariablen findet.

Für die Bestimmung der Sensormodellparameter gibt es verschiedene Möglichkeiten:

- Näherung als perfekte Sensoren
- Schätzung durch Wissensingenieur
- Stochastische Bestimmung aus Messdaten
- Individuelle Fehlermodelle (z.B. Blickbewegungsmessung)

### 3.1.3 Evidenzmodell

Das Evidenzmodell ist Teil eines Pilotenaufgabenmodells. Es werden Evidenzen für jede indirekt beobachtbare Aufgabe, welche während einer typischen Multi-UAV Hubschraubermission möglich ist, definiert. Dabei gibt es pro Beobachtung ein Evidenzmodell  $\mathbf{M}_E(X|E_j)$ , welches beschreibt, wie stark das Vorliegen der Evidenz  $E_j$  für die Aufgabe  $X$  spricht. Die Werte gelten unter der Annahme eines perfekten, also fehlerfreien Sensors. Das Modell einer Evidenz wird als Tabelle bedingter Glaubenswerte (Abschnitt 2.2.2) beschrieben:

$$\mathbf{M}_E(X|E_j) = \begin{pmatrix} p_t & p_f & 0 \\ q_t & q_f & 0 \\ r_t & r_f & 1 \end{pmatrix} \quad (7)$$

Einen Spezialfall für verschwindendes Unwissen ( $r_t = r_f = 0$ ) stellt das Wahrscheinlichkeitsbasierte Evidenzmodell dar:

$$\mathbf{M}_E(X|E_j) = \begin{pmatrix} P_t & 1 - P_f & 0 \\ 1 - P_t & P_f & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (8)$$

Ähnlich wie bereits die Sensormodellparameter, können auch die Daten für die Evidenzmodelle wie folgt gewonnen werden:

- Schätzung durch Wissensingenieur
- Expertenbefragung
- Stochastische Bestimmung aus Messdaten

### 3.1.4 Dynamikmodell

Das Dynamikmodell beschreibt, wie sich die Glaubensmassen der Evidenzen über die Zeit verhalten, das heißt, welche zeitliche Gültigkeit die Evidenzen in Bezug auf eine entsprechende Aufgabe besitzen.

Obwohl die Bearbeitung einer Aufgabe eine längere Zeit andauern kann, können die dadurch induzierten Beobachtungen nur sehr kurz bis fast nicht messbar auftreten. Der Pilot ist jedoch die ganze Zeit über beansprucht. So kann der Pilot beispielsweise längst eine Aufgabe begonnen haben, bevor beobachtbare Evidenzen auftreten, oder nach der Beobachtung der letzten Evidenz mental noch länger mit einer Aufgabe beschäftigt sein. Aus diesem Grund wird ein dynamisches Modell benötigt, welches das Auftreten einer Beobachtung mit der Länge der eigentlichen Ausführungsdauer einer Aufgabe korreliert. Da die Aufgabenerkennung in Echtzeit ablaufen soll, können keine Schlüsse rückwirkend gezogen werden, d.h. eine Modellierung der Glaubensmassen vor dem Eintreten einer Evidenz macht keinen Sinn.

Eine einfache Beschreibung der zeitlichen Änderung der Glaubensmasse  $dQ/dt$  nach einer Beobachtung ist ein exponentielles Abklingen der Glaubensmassen in Form eines nichtlinearen Dynamikmodells:

$$\mathbf{M}_D(X, t): Q_D(X, t) = Q_D(Q_E(X), t) = \begin{pmatrix} p_D(p_E, t) \\ q_D(q_E, t) \\ r_D(r_E, t) \end{pmatrix} \quad (9)$$

Das Dynamikmodell wird beispielhaft für  $p$  beschrieben und gilt analog auch für  $q$ . Dabei ist  $p$  durch  $q$  zu ersetzen. Das Unwissen  $r$  wird über die Normierungsbedingung (Gl. 1) bestimmt. Bild 4 stellt den zeitlichen Verlauf der Glaubensmasse  $p$  grafisch dar.



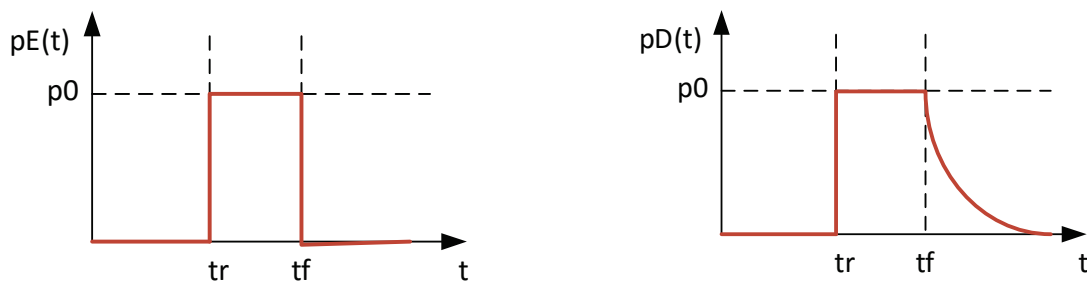


Bild 4: Dynamikmodell: Exponentielles Abklingen der Glaubensmasse (Eingangssignal links und Ausgangssignal rechts).

Das Eingangssignal  $p_E(t)$  ist die Glaubensmasse aufgrund einer Beobachtung. Sie steigt zunächst mit Beginn der Beobachtung zum Zeitpunkt  $t = t_r$  instantan auf  $p_E = p_0$  an und fällt nach der Beobachtung zum Zeitpunkt  $t = t_f$  wieder ab.

Nach der Beobachtung soll das Ausgangssignal  $p_D(t)$  jedoch nicht instantan wieder abfallen, sondern über die Zeit langsam abklingen. Dabei nehmen die Glaubenswerte ab und das Unwissen nimmt zu:

$$p_D(X, t) = \begin{cases} 0 & \text{für } t < t_r \\ p_E(X) & \text{für } t_r \leq t \leq t_f \\ p_0 \exp\left(-\frac{t-t_f}{\tau}\right) & \text{für } t > t_f \end{cases} \quad (10)$$

Dabei ist  $\tau$  die Abklingzeitkonstante und kann durch die Halbwertszeit  $T_{1/2}$ , also die Zeit, nach der sich der Glaube halbiert und das daraus resultierende Unwissen verdoppelt hat, ausgedrückt werden:

$$\tau = \frac{T_{1/2}}{\ln(2)} \quad (11)$$

Das Dynamikmodell für  $t > t_f$  kann auch in differentieller Form dargestellt werden:

$$\frac{dp}{dt} = -\frac{p_0}{\tau} \exp\left(-\frac{t-t_f}{\tau}\right) \quad (12)$$

## 3.2 Prozess der Tätigkeitsermittlung

Der Klassifikationsprozess für jede Aufgabe  $X$  verläuft in zwei Stufen: Erstens die Generierung von Evidenzen  $Q(E_j)$  aus Beobachtungen  $S_j$  (Abschnitt 3.2.1) und zweitens die Kombination der Einzelhypothesen  $Q(E_j)$  zur Gesamthypothese  $Q(X)$  (Abschnitt 3.2.2).

### 3.2.1 Sensorprozesskette: Generierung von Evidenzen aus Beobachtungen

Unter der Sensorprozesskette wird der Prozess von der Umweltbeobachtung bis hin zur Evidenzgenerierung verstanden. Die einzelnen Verarbeitungsschritte werden im Folgenden genauer beschrieben (Bild 5).

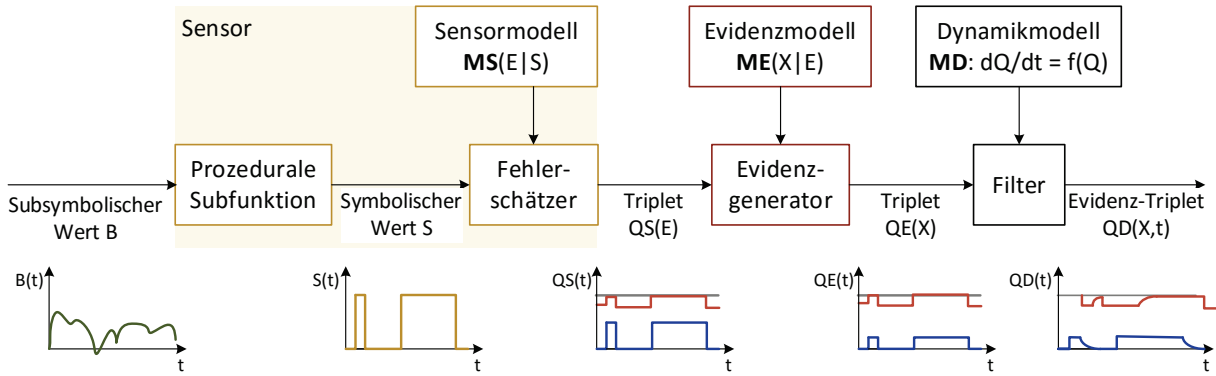


Bild 5: Sensorprozesskette für eine Evidenz  
(zur besseren Übersichtlichkeit sind keine Indizes  $j$  dargestellt).

### Prozedurale Subfunktion

Die Umwelt ist vieldimensional, nur teilweise beobachtbar, zeitkontinuierlich und hochgradig nichtlinear. Diese Beobachtung wird vom System zunächst als subsymbolischer Wert  $B$  erfasst. Über eine prozedurale Subfunktion wird daraus zunächst eine symbolische Hypothese  $S \in \{true, false, unknown\}$  generiert. Beispiel: „Pilot schaut auf Fahrtmesser“ ( $S = true$ ), „Pilot schaut nicht auf Fahrtmesser“ ( $S = false$ ) oder Sensorausfall ( $S = unknown$ ). Ein Sensorausfall kann beispielsweise dann auftreten, wenn das Blickbewegungsmesssystem defekt oder durch den Piloten teilweise verdeckt ist. Trotzdem soll die Tätigkeitserkennung noch funktionieren – dann allerdings mit einem vergrößerten Unwissen.

### Abschätzung des Sensorfehlers

Mit Hilfe eines Sensormodells (Abschnitt 3.1.2) wird die Zuverlässigkeit der Hypothese  $S$  abgeschätzt und zusammen mit der Hypothese als Glaubensmassentriplet  $Q_S(E)$  codiert. Es stellt den Glauben dar, mit welchem der Pilot zum Beispiel tatsächlich auf den Fahrtmesser schaut, wenn der Sensor „Pilot schaut auf Fahrtmesser“ als Hypothese liefert. Für die Berechnung gilt:

$$Q_S(E) = \mathbf{M}_S(E|S) Q_B(S) \quad (13)$$

$$Q_B(S) = \begin{cases} (1 & 0 & 0)^T, \text{ falls } S = true \\ (0 & 1 & 0)^T, \text{ falls } S = false \\ (0 & 0 & 1)^T, \text{ falls } S = unknown \end{cases}$$

### Generierung von Evidenzen aus fehlerbehafteten Sensorwerten

Das Sensortriplet  $Q_S(E)$  wird durch Vorwärtspropagation in Glaubensnetzen, wie in Smets (1993) oder Yaghlane & Mellouli (2006) beschrieben, mit dem Evidenzmodell kombiniert. Das Ergebnis ist erneut ein Evidenzmassentriplet  $Q_E(X)$ , das aussagt, wie stark die Messung der Beobachtung mit diesem fehlerbehafteten Sensor für oder gegen die Durchführung der Aufgabe  $X$  spricht.

Für die Vorwärtspropagation des Sensorwerts gilt:

$$\begin{aligned}
Q_E(X) &= \mathbf{M}_E(X|E)Q_S(E) \\
&= \begin{pmatrix} p_t & p_f & 0 \\ q_t & q_f & 0 \\ r_t & r_f & 1 \end{pmatrix} \begin{pmatrix} p_s \\ q_s \\ r_s \end{pmatrix} = \begin{pmatrix} p_t p_s + p_f q_s \\ q_t p_s + q_f q_s \\ r_t p_s + r_f q_s + r_s \end{pmatrix} \quad (14)
\end{aligned}$$

### Generierung des Dynamikverhaltens

$Q_D(X, t)$  erhält man durch Einsetzen von  $Q_E(X)$  in das nichtlineare Dynamikmodell (vgl. Gl. 9 und 10).

### 3.2.2 Tätigkeitsinferenz: Kombination der Evidenzen

Die Kombination der Evidenzen findet im Anschluss an die Modellierung der Dynamik statt. Im Folgenden wird zur besseren Übersichtlichkeit die Zeitabhängigkeit weggelassen.

Mit Hilfe einer Kombinationsregel können jeweils zwei Evidenztriplets  $Q_1(X)$  und  $Q_2(X)$  kombiniert werden:

$$Q_{12}(X) = Q_1(X) \oplus Q_2(X) = \begin{pmatrix} p_{12}(X) \\ q_{12}(X) \\ r_{12}(X) \end{pmatrix} \quad (15)$$

$$p_{12}(X) = p_1(X) \oplus p_2(X) = \frac{p_1 p_2 + p_1 r_2 + r_1 p_2}{1 - (p_1 q_2 + q_1 p_2)}$$

$$q_{12}(X) = q_1(X) \oplus q_2(X) = \frac{q_1 q_2 + q_1 r_2 + r_1 q_2}{1 - (p_1 q_2 + q_1 p_2)}$$

$$r_{12}(X) = 1 - p_{12}(X) - q_{12}(X)$$

Diese Formel kann aus der Kombinationsregel von Dempster für Evidenzmassen hergeleitet werden. Die Kombinationsregel von Dempster wird in Gordon & Shortliffe (1985) anschaulich beschrieben.

Durch iterative Anwendung dieser Regel können alle Einzelevidenzen, die für dieselbe Aufgabe sprechen, kombiniert werden:

$$Q_{ges}(X) = Q_1(X) \oplus Q_2(X) \oplus \dots \quad (16)$$

## 4 Technische Umsetzung

Die Funktionskette von der Erfassung der Beobachtungen über die Generierung von Evidenzen bis hin zur Tätigkeitsinferenz wurde im Hubschraubersimulator des Instituts für Flugsysteme der Universität der Bundeswehr München prototypisch implementiert. Ferner wurde eine umfassende Pilotenbeobachtung in einer interaktiven Echtzeitsimulation realisiert. Die mit Hilfe der Pilotenbeobachtung gesammelten Daten werden unter Verwendung eines echtzeitfähigen Erkennungsalgorithmus‘ kombiniert.

#### 4.1 Technische Umsetzung der Sensorprozesskette

Bisher wurden hauptsächlich perfekte Sensoren angenommen. Das vorhandene Unwissen stammt im Wesentlichen aus dem nicht vollständig ausmodellierten Evidenzmodell.

##### 4.1.1 Schalter, Taster und Drehknöpfe

Schalter können zwei oder mehrere Zustände einnehmen, in denen sie sich dauerhaft befinden. Betätigt der Pilot einen Schalter, so findet ein Zustandsübergang im Schalter statt. Ein typischer Schalter mit zwei Zuständen ist beispielsweise der Fahrwerkshebel. Drehknöpfe, wie zum Beispiel zum Einstellen einer Funkfrequenz, werden analog zu normalen Schaltern behandelt. Taster unterscheiden sich von Schalter dadurch, dass sie nur zwei Zustände besitzen, wobei sich der Taster die meiste Zeit in einem Zustand befindet und nur für eine sehr kurze Zeit in den anderen wechselt, während der Pilot ihn betätigt. Ein derartiger Taster ist beispielsweise der Funkknopf. Für die Tätigkeitserkennung ist neben dem Zustand selbst vor allem das Ereignis der Betätigung, also die zeitliche Änderung des Zustands (mathematisch: die Ableitung), relevant. Erst durch die Modellierung mittels eines Dynamikmodells (Abschnitt 3.1.4), bei dem die kurzen Ereignisse zeitlich gestreckt werden, wird die hier vorgestellte diskrete Evidenzkombination ermöglicht. Schalter, Taster und Drehknöpfe können mit guter Näherung durch das Modell eines perfekten Sensors approximiert werden.

##### 4.1.2 Steuerknüppel

Der Steuerknüppel liefert kontinuierliche Signale der Steuereingaben. Diese werden durch eine prozedurale Subfunktion in einen digitalen Wert („Pilot bewegt den Steuerknüppel“ / „Pilot bewegt den Steuerknüppel nicht“) umgewandelt. Immer wenn die Änderung der Steuergröße einen Schwellwert überschreitet, wird ein Ereignis ausgelöst. Als Fehlermodell eignet sich beispielsweise ein Sensor mit konstanter Zuverlässigkeit.

##### 4.1.3 Blickbewegungsmessung

Die Blickbewegungsmessung ist komplex und geschieht zunächst durch prozedurale Subfunktionen im Rahmen der semantischen Blickbewegungsmessung (Bild 6).

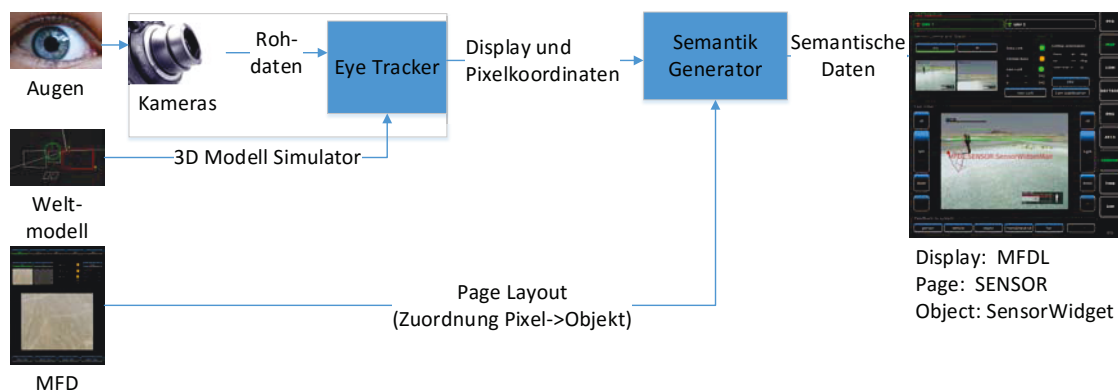


Bild 6: Semantische Blickbewegungsmessung.

Während der Pilot im Cockpit sitzt, blickt er auf ein (Multifunktions-) Display. Ein kommerzielles Blickbewegungsmesssystem (*Eye-Tracker*) bestimmt lediglich eine Pixel-Koordinate  $(x_0, y_0)$  auf diesem zuvor im Weltmodell modellierten Bildschirm. Zusammen mit dem Layout des entsprechenden Displays zu diesem Zeitpunkt wird eine semantische, objektbezogene Aussage getroffen wie zum Beispiel „Pilot schaut auf den Fahrtmesser“.

Für das Sensormodell der Zuverlässigkeit dieser Aussage wird der Blickstrahl als gaußförmige Verteilung (Normalverteilung) in zwei Raumrichtungen modelliert:

$$\rho(x, y) = \left( \frac{1}{2 \pi \sigma_x \sigma_y} \right) \exp \left( -\frac{1}{2} \left[ \left( \frac{x-x_0}{\sigma_x} \right)^2 + \left( \frac{y-y_0}{\sigma_y} \right)^2 \right] \right) \quad (17)$$

Das Modell beschreibt die Wahrscheinlichkeitsdichte, mit der der Pilot wirklich auf das auf dem Bildschirm dargestellte Objekt blickt. Dabei sind  $x_0$  und  $y_0$  der Blickpunkt auf einem Bildschirm, den das Blickbewegungsmesssystem als Rohdaten liefert. Die Strahlbreiten (Standardabweichungen) in  $x$ - und  $y$ -Richtung sind dabei  $\sigma_x$  und  $\sigma_y$ . Integriert man die Dichte  $\rho$  über ein Bildschirmobjekt, so erhält man ein Zuverlässigkeitsmaß  $Z$  und kann wie beim Sensor mit konstanter Zuverlässigkeit (Gl. 5) die Matrix bedingter Glaubensgrade erstellen.

$$Z = \iint_{\text{Objekt}} \rho(x, y) dx dy \quad (18)$$

Da das Blickbewegungsmesssystem für gewöhnlich nicht überall auf dem Bildschirm gleich gut funktioniert, sind die Standardabweichungen, also die Breite des Gauß-Strahls, selbst ortsabhängig und müssen durch eine Kalibrierung bestimmt werden (Bild 7).

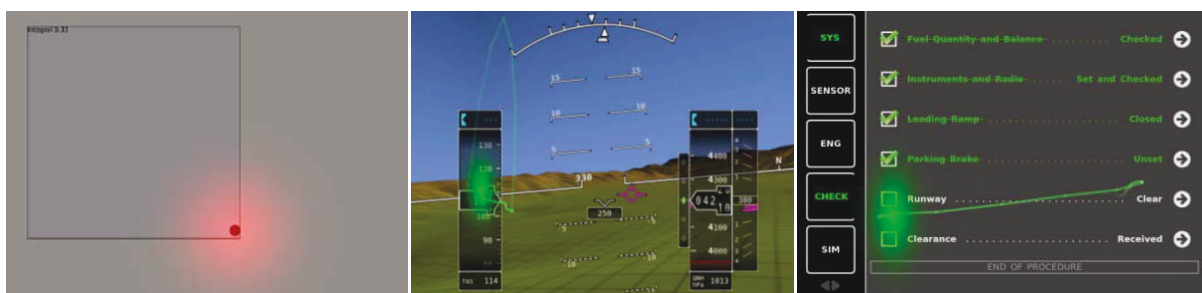


Bild 7: Modellierung der Blickbewegung als gaußförmigen Strahl.

#### 4.1.4 Touchscreen

Jedes Multifunktionsdisplay (MFD) im Simulatorcockpit verfügt über eine berührungsempfindliche Oberfläche (*Touchscreen*). Im MFD werden die rohen Pixelkoordinaten zusammen mit dem Layout direkt in eine objektbezogene symbolische Information umgewandelt. Für den Touchscreen wird ein ähnliches, gaußförmiges Fehlermodell für den Blickstrahl verwendet (Abschnitt 4.1.3), jedoch wird die Breite der Verteilung über den gesamten Bildschirmbereich als konstant angenommen. Ein menschlicher Finger besitzt eine endliche Größe und

hat Schwierigkeiten, kleine Objekte auf dem Bildschirm zu treffen. Je kleiner das Objekt, desto unzuverlässiger die Messung.

#### 4.1.5 Kommunikation

Unter der Kommunikation werden sowohl die auditive Kommunikation zwischen den beiden Piloten über das Intercom, als auch der Flugfunksprechverkehr mit einer Gegenstelle verstanden. Die Kommunikation wird über integrierte Mikrofone in den Headsets der Piloten erfasst. Die Erfassung der Kommunikation geschieht binär. Solange der Aufnahmepegel einen Schwellwert überschreitet, wird die symbolische Information „Pilot spricht“ über eine prozedurale Subfunktion generiert. Fehler in der Erkennung entstehen zum Beispiel durch Wahrnehmung von Umweltgeräuschen oder schlecht angelegte Headsets.

#### 4.1.6 Umwelt

Für Umweltinformationen werden zunächst mit Hilfe prozeduraler Subfunktionen symbolische Umweltwerte erstellt. Beispiele sind „Hubschrauber in Kontrollzone“ oder „Hubschrauber durch feindliche Kräfte gefährdet“. Auch die Ermittlung des aktuellen Flugzustands des Hubschraubers wird durch eine prozedurale Subfunktion abgebildet. Die Fehlermodelle der Umweltsensoren sind für jede Beobachtung individuell. Für die Umweltinformationen eignen sich entweder die Näherungen als perfekte Sensoren oder Sensoren mit begrenzter konstanter Zuverlässigkeit.

### 4.2 Umsetzung der Tätigkeitsinferenz: Kombination der Evidenzen

Es wurde ein echtzeitfähiger Algorithmus geschrieben, der die Kombinationsregel (15) umsetzt. Zurzeit werden insgesamt etwa 500 Evidenzen automatisiert generiert, die auf circa 50 mögliche, im Aufgabenmodell modellierte, Pilotenaufgaben schließen lassen. Zur Kontrolle wurden Diagnose-Displays entwickelt. Bild 8 zeigt einen Ausschnitt der Tätigkeitserkennung im Hubschraubersimulator während erster Engineering-Versuche.

	Task	Graphical	Belief	Plausibility	Doubt	Ignorance
1	FlyManual		0.8	1.0	0.0	0.2
2	TransitFriend		0.7	1.0	0.0	0.3
3	CrossAirspaceBorderCheckFuel		0.6	1.0	0.0	0.4
4	CheckOil		0.6	1.0	0.0	0.4
5	CheckFuel		0.6	1.0	0.0	0.4
6	CheckRPM		0.6	1.0	0.0	0.4
7	CheckAltitude		0.2	1.0	0.0	0.8

Bild 8: Ergebnisse der Tätigkeitserkennung während eines Transitflugs.

Der Pilot fliegt gerade manuell im freundlichen Gebiet kurz vor dem Überqueren einer Luftraumgrenze in potenziell bedrohtes Gebiet. Man erkennt, dass er gerade mit hohem Glaubensgrad (*Belief*) seine Instrumente kontrolliert. Gut zu erkennen ist dabei das vorhandene Unwissen (*Ignorance*). Zu diesem Zeitpunkt liegt keine Skepsis (*Doubt*) vor, was gegen die hier gezeigten Aufgaben spricht. In der

grafischen Darstellung sind links die momentanen Werte als horizontale Balken dargestellt (von links nach rechts: Glaube (blau), Unwissen (gelb), Skepsis (rot)) und rechts daneben vertikal die Glaubenswerte über den zeitlichen Verlauf mit der Gegenwart am rechten Rand. Man erkennt deutlich das exponentielle Abklingen der Glaubenswerte, welches aus dem Dynamikmodell resultiert. Die Tätigkeiten sind nach absteigender Wahrscheinlichkeit (schwarzer horizontaler Balken) sortiert, welche man erhält, wenn man das Unwissen in gleichen Teilen auf Glaube und Skepsis aufteilt (Pignistische Transformation).

## 5 Zusammenfassung und Ausblick

Erste Engineering-Versuche haben gezeigt, dass die Tätigkeitserkennung und die darauf aufbauende Workload Erkennung (Maiwald & Schulte, 2011; Schulte et al., 2015) mit dem gewählten Ansatz möglich ist. Der adaptive Loop muss aber noch geschlossen werden und die Modelle müssen auf die Anwendung angepasst werden.

Derzeit wird das Konzept stufenweise im Simulator umgesetzt. Ausgehend von perfekten Sensoren werden nach und nach die Fehlermodelle der Messsensoren, wie in den Abschnitten 3.1.2 und 4.1 beschrieben, verfeinert, um insbesondere diejenigen Situationen zu berücksichtigen, in denen das Blickbewegungsmesssystem keine Daten liefert. Zustandsraummodell, Evidenzmodell und Dynamikmodell wurden als Teil eines Pilotenaufgabenmodells implementiert. Bisher wurden im Evidenzmodell die Parameter prototypisch abgeschätzt. Schwierigkeiten bereiten dabei die Interpretation der Glaubensmaße  $p$ ,  $q$ ,  $r$  und deren Quantifizierung. Auch die Halbwertszeiten im Dynamikmodell sind ohne eine experimentelle Analyse schwer zu bestimmen. Bisher wurden im Zustandsraummodell nicht alle Aufgaben ausmodelliert und die Glaubensmaße im Evidenzmodell nur teilweise quantifiziert.

Trotz der nur partiellen Ausmodellierung ist das System lauffähig, liefert allerdings für nichtmodellerte Evidenzen, wie erwartet, Unwissen. Um das Unwissen im Pilotenaufgabenmodell zu verringern sind Wissenserwerbsexperimente mit Piloten der Bundeswehr geplant. Dabei sollen vor allem die Glaubensmaße im Evidenzmodell genauer quantifiziert werden. Ferner sind in einer weiteren Experimentalkampagne Evaluierungsexperimente des Gesamtsystems bestehend aus Workload-adaptivem Assistenzsystem mit Pilotenzustandserfassung geplant.

## Literatur

- Dempster, A.P. (1967). Upper and Lower Probabilities Induced by a Multivalued Mapping. *Annals of Mathematical Statistics*, 38(2), 325–339.
- Dempster, A.P. (2008). The Dempster–Shafer calculus for statisticians. *International Journal of Approximate Reasoning*, 48(2), 365–377.
- Donath, D. (2012). *Verhaltensanalyse der Beanspruchung des Operateurs in der Multi-UAV-Führung*. Dissertation. Universität der Bundeswehr München.

- Gordon, J. & Shortliffe, E.H. (1985). A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space. *Artificial Intelligence*, 26(3), 323–357.
- Honecker, F. & Schulte, A. (2015). Konzept für eine automatische evidenzbasierte Online-Pilotenbeobachtung in bemannt-unbemannten Hubschraubermissionen. In 4. *Interdisziplinärer Workshop Kognitive Systeme: Mensch, Teams, Systeme und Automaten – Verstehen, Beschreiben und Gestalten Kognitiver (Technischer) Systeme*. Bielefeld.
- Maiwald, F. (2013). *Maschinelle Beanspruchungsprädiktion zur ressourcengerechten Adaption eines Pilotenassistenzsystems*. Dissertation. Universität der Bundeswehr München.
- Maiwald, F. & Schulte, A. (2011). Mental resource demands prediction as a key element for future assistant systems in military helicopters. *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6781 LNAI, 585–591.
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, California: Morgan Kaufmann Publishers Inc.
- Rouse, W.B. & Rouse, S.H. (1983). *A Framework for Research on Adaptive Decision Aids*. Wright-Patterson Air Force Base, Ohio.
- Ruckdeschel, W. (1997). *Modellierung regelbasierten Pilotenverhaltens mit Petrinetzen*. VDI-Verlag, Aachen.
- Schulte, A., Donath, D. & Honecker, F. (2015). Human-System Interaction Analysis for Military Pilot Activity and Mental Workload Determination.
- Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press.
- Shortliffe, E.H. (1976). *Computer-based Medical Consultations: MYCIN (Artificial intelligence series)*. Elsevier Science Ltd.
- Smets, P.P. (1993). Belief Functions : the Disjunctive Rule of Combination and the Generalized Bayesian Theorem. *International Journal of Approximate Reasoning*, 9(1), 1–35.
- Veltman, J.A. & Jansen, C. (2006). *The Role of Operator State Assessment in Adaptive Automation* (Vol. TNO-DV3 20). Soesterberg: TNO Defence Security and Safety.
- Wittig, T. (1994). *Maschinelle Erkennung von Pilotenabsichten und Pilotenfehlern über heuristische Klassifikation*. *Fortschritt-Berichte VDI*. München: VDI-Verlag.
- Yaghlane, B.B. & Mellouli, K. (2006). Belief Function Propagation in Directed Evidential Networks. In *IPMU*.
- Young, M.S., Brookhuis, K.A., Wickens, C.D. & Hancock, P.A. (2015). State of science: mental workload in ergonomics. *Ergonomics*, 58(1), 1–17.

## Autoren

Dipl.-Phys. F. Honecker  
Prof. Dr. A. Schulte

Universität der Bundeswehr München  
Institut für Flugsysteme

Kontakt: [fabian.honecker@unibw.de](mailto:fabian.honecker@unibw.de)