



Technische Universität Ilmenau

Fakultät für Informatik und Automatisierung

Fachgebiet Neuroinformatik und Kognitive Robotik

Klassifizierung von mentaler Beanspruchung durch Messung von Pupillendaten

Bachelorarbeit zur Erlangung des akademischen Grades Bachelor of Science

Melanie Müller

Betreuer: Dr. Matthias Nieser, Sensomotoric Instruments GmbH

Dr. Klaus Debes, FG Neuroinformatik und Kognitive Robotik

Verantwortlicher Hochschullehrer:

Prof. Dr. H.-M. Groß, FG Neuroinformatik und Kognitive Robotik

Die Bachelorarbeit wurde am 15.05.2015 bei der Fakultät für Informatik
und Automatisierung der Technischen Universität Ilmenau eingereicht.

[urn:nbn:de:gbv:ilm1-2015200147](https://nbn-resolving.org/urn:nbn:de:gbv:ilm1-2015200147)

Danksagung

Einen ganz besonders Dank möchte ich an dieser Stelle meinen beiden Betreuern Dr. Klaus Debes und Dr. Matthias Nieser aussprechen, die es mir trotz der räumlichen Entfernung möglich gemacht haben diese Bachelorarbeit anfertigen zu können. Ich möchte mich besonders für deren Geduld und Rücksicht bedanken, sowie über die kompetente Betreuung und Hilfestellung.

Außerdem bedanke ich mich bei der Firma Sensomotoric Instruments GmbH, die mir nicht nur Hard- und Software, sondern auch einen Testplatz zur Verfügung gestellt haben. Meinen Probanden danke ich für ihr Verständnis und die Bereitschaft mir und meinem Experiment ihre volle Aufmerksamkeit zu schenken.

Zuletzt möchte ich meiner Familie danken, die mich in jeder Situation unterstützt hat.

Erklärung: „Hiermit versichere ich, dass ich diese Bachelorarbeit selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel verwendet habe. Alle von mir aus anderen Veröffentlichungen übernommenen Passagen sind als solche gekennzeichnet.“

Ilmenau, 15.05.2015

.....
Melanie Müller

Kurzdarstellung

In der derzeitigen Wissenschaft wird mittels verschiedenartiger, physiologischer Messmethoden versucht, die mentale Belastung von Personen darstellbar zu machen. Dies kann mittels EKG, EEG, Hautableitwert oder Pupillometrie geschehen.

In dieser Arbeit habe wurde eine Kombination von pupillometrisch erfassten Daten, einer Vorverarbeitung dieser und einem Fuzzy-System eine Klassifizierung des aktuellen mentalen Workloads zu realisiert. Ziel dieser Arbeit ist es sein, einen automatisierten Weg zu einer solchen Klassifizierung zu finden, welche eine grobe Einschätzung über den momentanen Status der psychischen Beanspruchung zulässt.

Zur Datenaquise wurden zunächst vier verschiedene, in der Wissenschaft etablierte Experimente ausmacht und mit der Software „ExperimentCenter“ erstellt. Die Datenaufnahme wurde mit mindestens 25 Probanden durchgeführt.

Mittels linguistischer Variablen ließ sich eine Beschreibung der verschiedenen Eingangsgrößen finden, welche anschließend zur Klassifikation mittels Fuzzy-System genutzt werden konnten. Als Eingangsgrößen dieses Systems wurde der normierte Pupillendurchmesser und der normierte Anstieg der Pupille implementiert. Eine Verknüpfung dessen mittels Inferenz nach Sugeno ergab eine ermittelte Beanspruchung, die mit der Angenommenen einhergeht.

Abstract

The representation, measurement and analysis of mental workload of an individual are major subjects of scientific research. Electrocardiogram, electroencephalography, skin conductance level and pupillometric measurements are all options for obtaining values of this workload.

For this thesis, a pupillometric dataset containing test data for at least 25 individuals was obtained, in which each individual was asked to execute a series of tests with different types and intensities of psychological stress.

From the dataset, pupil size and changes in pupil size were extracted as input parameters to a software system based on linguistic variables of fuzzy-logic and Sugeno inference operators to approximate the current mental workload. This approach allowed a clear distinction between phases of low and high workload.

Inhaltsverzeichnis

1	Einleitung	3
2	Psychische Beanspruchung	5
2.1	Begriffserklärung	5
2.2	Ereignis-korrelierte Potentiale	5
2.2.1	Evozierte Potentiale	6
3	Was ist Pupillometrie?	9
3.1	Begriffserklärung	9
3.2	Psychosensorische Pupillenreakton	9
4	Eye-Tracking als Messmethode	11
4.1	Was ist Eye-Tracking?	11
4.2	Arten von Eye-Tracking Systemen	11
4.3	Vor- und Nachteile dieser Messmethode	12
5	Experiment	15
5.1	Ziel des Experiments	15
5.2	Optische Methoden in der Wissenschaft	15
5.2.1	Sternberg's Memory Scanning Paradigma	16
5.2.2	Digit span forward	18
5.2.3	Digit span backwards	19
5.2.4	Oddball paradigm	19
5.3	Umgesetzte Experimentenreihe	21

5.3.1	Anforderungen an die Probanden	21
5.3.2	Anforderungen an den Testplatz	22
5.3.3	Testablauf	24
6	Algorithmische Umsetzung	29
6.1	Eventdaten	29
6.2	Einladen der Daten in Matlab	30
6.3	Bereinigen von Ausreißern und invaliden Werten	30
6.3.1	Notwendigkeit der Bereinigung	30
6.3.2	Detektion von Ausreißern und invaliden Werten	33
6.3.3	Interpolation von Ausreißern und invaliden Werten	35
6.4	Normierung der Pupillendaten	35
6.5	Bestimmung des Anstiegs	36
6.6	Integrierung in Fuzzy-Logic-Toolbox	37
6.6.1	Eingangsdaten	38
6.6.2	Methode und Ausgabevariablen	44
6.6.3	Regeln	44
7	Vergleich Erwartungswerte mit Fuzzy-Klassifizierung	47
7.1	Digit Span-, Digit Span Backwards- und Sternberg Test	47
7.2	Oddball Test	56
7.3	Probleme der Auswertung	58
7.4	Optimierungsmöglichkeiten	59
8	Zusammenfassung und Ausblick	61
A	Ergänzende Unterlagen	63
A.1	Probandendaten	63
A.2	Software	65
	Literaturverzeichnis	68

Abkürzungsverzeichnis

AOI	Area of Interest
CR	Kornealer Reflex
EKP	Ereignis-korrelierte Potentiale
IDF	iDataFile
IR	Infrarot
LED	Light-emitting diode
POI	Point of Interest
RED	Remote Eye-Tracking Device
SMI	Sensomotoric Instruments GmbH

Kapitel 1

Einleitung

*„Das Auge ist der Punkt, in welchem Seele und Körper sich vermischen.“
Christian Friedrich Hebbel*

Dass die Augen eine besondere Rolle in der Wahrnehmung anderer Personen spielen, ist schon lange bekannt. Sie verzaubern, lassen Zorn erkennen oder wirken schon aufgrund einer vergrößerten Pupille sympathischer. Was zu Christian Friedrichs Hebbels Zeit noch undenkbar war, ist nun möglich. Kontaktlos lässt sich heute die absolute Pupillengröße mittels Eyetracking bestimmen. Dadurch lässt sich Psychische Beanspruchung, emotionale Erregung oder Anspannung greifbarer machen. Wie viele psychologische Studien gezeigt haben, besteht ein direkter Zusammenhang zwischen verschiedenen psychischen Zuständen und der Größe der Pupille. Als eines der äußerlich erkennbar am schnellst wirkenden Organe ist einer Nutzung dieser Messgröße eine hohe Wichtigkeit einzuräumen.

Ziel dieser Arbeit ist es, einen Algorithmus zu entwickeln, mit dem möglichst einfach eine Schätzung über die psychische Beanspruchung möglich ist. Dazu müssen zunächst Daten erzeugt werden, die verschiedene Ausprägungen von psychischer Beanspruchung wiedergeben. Diese Daten müssen sowohl Informationen über Stimuli als auch über die Pupillen enthalten. Tests mit verschiedenen Probanden sind somit von Nöten und unumgänglich, da keine frei zur Verfügung stehenden Messdaten existie-

ren. Nutzen könnte man diese Einschätzung in vielerlei Hinsicht. Für Usability-Tests, Lügendetektoren oder in Flugsimulatoren, um die tatsächliche Belastung von Piloten während schwierigen Flugmanövern zu erfassen.

Kapitel 2

Psychische Beanspruchung

2.1 Begriffserklärung

In der Arbeitswissenschaft wird klar zwischen psychischer Belastung und Beanspruchung unterschieden. Laut der DIN EN ISO 10075-1 [DIN, 1991] beschreibt der Begriff „psychische Belastung“ die äußeren Merkmale der Arbeitssituation. Dies können Arbeitsumfeld, Geräuschkulisse, soziale Umgebungsbedingungen oder Zeitdruck sein. „Psychische Belastung“ ist in jedem Fall neutral zu verstehen und schließt keinerlei Wertung des aktuellen Befindens einer Person mit ein.

„Psychische“ Beanspruchung hingegen ist immer wertend und beschreibt die Auswirkung der psychischen Belastung. Sie ist dabei nicht direkt abhängig von der Belastung selbst, sondern auch Summe der individuellen Eigenschaften und Fertigkeiten. Die gleiche Belastung kann also bei unterschiedlichen Personen ein anderes Ausmaß an Beanspruchung hervorrufen. So können Erwartungswerte für bestimmte Situationen hier nur als Approximation gesehen werden. (vgl. [DIN, 1991], [SCHLICK et al., 2010])

2.2 Ereignis-korrelierte Potentiale

Ereignis-korrelierte Potential (EKP) sind hirnelektrische Potentialverschiebungen, die nicht spontan, sondern durch einen Stimulus oder aufgrund eines anderen konkre-

ten Ereignisses ausgelöst werden. Diese Potentiale können durch ein Elektroenzephalogramm gemessen werden. Stimuli sind Ereignisse, dessen Zeitpunkt bekannt sind und die eindeutig wahrnehmbar für den Probanden sind. Eine Korrelation von Stimulus bzw. Ereignis und Potential muss zwingend gegeben sein. Besonders die inneren (endogenen) Ereignisse veranlassen die Bildung dieses Potentials. Unter inneren Ereignissen lassen sich die Wahrnehmung, Weiterleitung und Verabreichung des Reizes zusammenfassen. Bei verminderter Aufmerksamkeit würde der Reiz als solches nicht wahrgenommen und ein Potential nicht evoziert werden. Es gibt, wie in allen verhaltenspsychologischen Fragestellungen, geschlechtsspezifische Unterschiede. Diese bestimmen die Wahl der Stimuli, die Emotionen oder Urinstinkte ansprechen sollen. Bei Stimuli, welche neutral sind und keine wertbaren Inhalte zeigen, kann dieser Aspekt vernachlässigt werden. (vgl. [TRIMMEL, 1990], [BUCHNER, 2005], [GREHL et al., 2000], [SCHMIDT und LANG, 2013])

Die Theorie des Ressourcenmodells geht von einer Menge an begrenzten Verarbeitungsressourcen jedes Menschen aus, welche zur Informationsverarbeitung benötigt werden. Der Bedarf einer Ressource hängt von der Komplexität der zu lösenden Aufgabe ab und ist meist nur für einen bestimmten Bereich gültig. So kann ein Proband im Bereich der visuellen Verarbeitung ausgelastet sein, seine Ressourcen im auditiven Bereich müssen davon jedoch nicht beeinflusst sein. (vgl. [WICKENS, 2002], [MANZEY, 1998])

2.2.1 Evozierte Potentiale

Die evozierten Potentiale gehören zu den ereigniskorrelierten Potentialen. Sie können als Antwort des Zentralen Nervensystems auf eine Reizung von “peripheren Nerven, von sensorischen Bahnen oder Kernen” ([SCHMIDT und LANG, 2013], S.135) verstanden werden.

Es kann zwischen folgenden Arten von evozierten Potentialen unterschieden werden:

- **visuell:** visuelle Stimuli in Bild-, Text- oder Videoform.
 - **akustisch:** akustische Reize durch Tonfolgen oder aufgezeichneten Stimuli
-

- **somatosensorisch:** Reize meist durch elektrische Stimulation von Nerven (meist Nervus medianus am Handgelenk)
- **motorisch:** Magnetstimulation

(vgl. [BUCHNER, 2005], [GREHL et al., 2000])

Bei den evozierten Potentialen, die in dieser Arbeit genutzt werden, handelt es sich ausschließlich um visuelle und akustische. Diese Stimuli können am einfachsten erzeugt werden und bürden die geringsten Gefahren für die Probanden. Zudem bedarf es bei den von mir gewählten Arten der Reizbildung keiner ärztliche Aufsicht, da Risiken bei der Durchführung der Tests für die Probanden ausgeschlossen sind.

P300 Als P300 versteht man ein hirnelektrisches Signal im späten Latenzbereich, d.h. im Bereich von 300-500ms nach einem speziellen Reiz. Dieses späte Potential wird durch einen seltenen Reiz ausgelöst. Notwendig dazu ist eine Reihe von Standardreizen, die im gleichen Abstand dem Probanden vorgestellt werden. Beispielsweise das Bild eines Quadrats. In seltenen Fällen (in höchstens 5 Prozent der Reize) wird das Abbild eines Kreises als Stimulus gewählt. Da der Proband zuvor die Aufgabe erhielt bei jedem Stimulus, der einen Kreis entspricht eine Aktion auszuführen, wird dieser Aufgabe nun nachgegangen. Als Aktion kann beispielsweise ein Zählen des besonderen Reizes oder das Drücken einer Taste ausgewählt werden. Die Mitarbeit des Probanden ist bei einem solchen Test zwingend erforderlich, denn ohne eine aufgebrachte Konzentration ist ein solches EKP nicht provozierbar. Ein Beispiel für einen solchen Test ist der Oddball Test 5.2.4

Auch zur medizinischen Diagnostik kann der P300 ermittelt werden. Er gibt Rückschlüsse auf Demenz, hirnorganische Psychosyndrome und toxische Schäden. (vgl. [GREHL et al., 2000],[BUCHNER, 2014])

Kapitel 3

Was ist Pupillometrie?

3.1 Begriffserklärung

Pupillometrie beschreibt die Messung und Verwertung der absoluten Pupillengröße und dessen Veränderung über die Zeit. Die Messung erfolgt meist über ein optisches Verfahren, bei dem mit infrarotem Licht gearbeitet wird. Dieses ist für das menschliche Auge unsichtbar und blendet den Probanden somit nicht. Eine Blendung würde das Zusammenziehen der Iris und die Verengung der Pupille zur Folge haben.

3.2 Psychosensorische Pupillenreaktion

Die Iris ist eine der wenigen Erfolgsorgane, die auf beide vegetativen Nervensysteme reagiert, also den Sympathikus und den Parasympathikus. Die inneren Augenmuskeln sind für die Adaption der Pupille verantwortlich. Für die Erweiterung der Pupille ist der Sympathikus verantwortlich. Dieser innerviert den Musculus dilator pupillae. Der Parasympathikus innerviert den Musculus sphincter pupillae. Dieser besteht aus radiären Zügen glatter Muskelzellen in der Iris und erweitert die Pupille. [ULFIG, 2011], [SCHMIDT et al., 2007]

Kapitel 4

Eye-Tracking als Messmethode

4.1 Was ist Eye-Tracking?

Die Bezeichnung Eye Tracking beschreibt zunächst das Erfassen des Auges. Dies kann in Bildern oder Videos geschehen. Die häufigste Verwendung des Wortes geschieht jedoch im Zusammenhang mit der Blickrichtungsverfolgung. Die Blickrichtung, oder der Punkt des scharfen Sehens, wird weiterhin auch als Gaze oder POI (point of interest) bezeichnet und beschreibt selbiges. Das Erkennen der Blickrichtung geschieht mittels Eye Tracker.

4.2 Arten von Eye-Tracking Systemen

Es gibt grundsätzlich zwei verschiedene Arten von Eye Tracking Systemen:

- **Remote Eye Tracker:** Diese Systeme bestehen aus einer Kameraeinheit und LED-Leuchtgruppen. Die LED-Gruppen geben infrarotes Licht (IR-Licht) in Richtung des zu messenden Probanden ab. Je nach System und Hersteller variieren die Anzahl der LED-Einheiten und Kameras.
 - **Mobile Eye Tracker:** Diese Bauart ist meist in eine Aufnahmebrille implementiert und nimmt die Umgebung (Szenenkamera) und die Augen (Augenkamera)
-

auf. Eine verarbeitende Einheit, wie ein Computer oder Smartphone, ist bei beiden Arten des Eye Trackings von Nöten.

Als weitere Modellvariante kann hier auch ein Aufbau auf einem Helm oder einer anderen Kopfbefestigung erwähnt werden.

In dieser Arbeit wurde lediglich der Remote Eye Tracker „RED-oem“ des Herstellers SensoMotoric Instruments (SMI) verwandt. Dieser besteht aus einem Kameramodul und zwei LED-Baugruppen. Die IR-Licht-erzeugenden Elemente befinden sich jeweils am äußeren Rand der Platine. Sie erzeugen eine Reflektion an der Hornhaut (engl. Cornea). Diese Reflektionen werden auch als Korneale Reflexe (kurz CR) bezeichnet. Gängige Eye Tracking Methoden identifizieren das Pupillenzentrum und die CR's im Kamerabild und rekonstruieren daraus die Augenposition und die Blickrichtung.

Als Beispiel sei die P-CR Methode zu nennen, die allein auf der Messung der absoluten Differenz zwischen Pupillenzentrum und CR im Kamerabild basiert. (vgl. [HOLMQVIST et al., 2011])

4.3 Vor- und Nachteile dieser Messmethode

Im Folgenden sollen die Vor- und Nachteile der Nutzung eines Eye Tracking Systems beleuchtet werden. Da meiner Klassifikation ausschließlich die Messdaten der Pupille zugrunde liegen, befasse ich mich speziell mit den Aspekten, die für diese Art der Nutzung eines Eye Tracking Systems relevant sind.

Der größte Vorteil besteht unbestritten in der Kontaktlosigkeit dieser Messmethode. Der Proband muss keine Sensoren oder andere Messinstrumente mit sich tragen. Im Falle eines Remote Eye Trackers ist ein Monitor zwingend erforderlich, unter dem das Remote Eyetracking Gerät angebracht wird. Anderenfalls muss der Proband ein mobiles Eyetracking System tragen. Durch die Wahl des Instruments ist der Messort eingeschränkt bzw. frei wählbar.

Eine Ablenkung des zu Untersuchenden ist unwahrscheinlich, da das Gerät passiv arbeitet und es nicht zu sichtbaren Veränderungen am Gerät oder Bildschirm kommt.

Bei der Aufzeichnung der Augendaten bedarf es nach Starten der Applikation und der Kalibrierung keiner weiteren gezielten Interaktion.

Die Beeinflussung der Pupillendilatation ist nur durch Einnahme von Medikamenten oder Drogen zu erlangen, welche zum veränderten Verhalten der Pupille führen. Diese Einflussfaktoren sind messbar und somit nachweisbar. Dies kann als Vorteil gesehen werden, da bei einem Ausschluss der oben genannten Mittel von einer größtenteils Unabhängigkeit der Pupillenreaktion ausgegangen wird. Für diese Arbeit wurde auf ein Messen dieser störenden Faktoren aus Kostengründen verzichtet. Es wurde lediglich eine Befragung zur Einnahme der relevanten Stoffe durchgeführt. [HOLMQVIST et al., 2011]

Ein Trainieren oder Provozieren der Pupillenveränderung ist zudem nicht möglich.

Als Nachteil können die bisher noch hohen Kosten für das Eyetracking Gerät und die Softwarelizenzen genannt werden. Diese werden in Zukunft durch die Weiterentwicklung von Elektronik mit minimalster Größe reduziert.

Ein weiterer Nachteil sind die äußeren Gegebenheiten im Testumfeld, die eine Pupillendilatation auch aufgrund der sich ändernden Lichtverhältnisse provozieren. Da dies die Hauptaufgabe der Iris ist, lässt sie sich nicht unterdrücken. Dies muss bei der Gestaltung des Experiments und der räumlichen Umgebung beachtet werden. Ein Stimuli, welches viele Farbvariationen aufweist, führt nicht nur aufgrund emotionaler oder kognitiver Veränderungen zur Veränderung der Pupille, sondern auch durch die Akkumulation der Iris. In Kapitel 5.3.2 sind die genauen Anpassungen meiner Testsituation darauf dargelegt.

Hinzu kommt, dass das Blinzeln oder das Wegsehen vom Bildschirm ein kontinuierliches Messen nicht möglich machen. Beim Blinzeln kommt es zum vollständigen Schließen des Auges, somit ist die Messung des Pupillendurchmessers und der Blickrichtung unmöglich. Nach dem Blinzeln kann es zudem zu Ausreißern kommen, da die Pupille neu im Bild gefunden werden muss und sich der neuen Beleuchtungssituation erst anpasst. Diese Werte müssen aus der Messreihe entfernt werden, da sie zu fehlerhaften Rückschlüssen und Berechnungen führen können.

Kapitel 5

Experiment

5.1 Ziel des Experiments

Die aufgenommenen Daten sollen als Basis zur Klassifizierung unterschiedlicher mentaler Beanspruchung verwendet werden. Dazu ist es notwendig, Phasen der vermuteten endogenen Aktivität (hier mentale Beanspruchung) und belastungsfreie Phasen zu realisieren. Der Blick des Probandes sollte möglichst immer auf den Bildschirm des Testaufbaus gerichtet sein, sodass Augendaten vom Eyetracker aufgenommen werden können. Ist der Blick zu weit vom Blickerfassungssystem abgewandt oder sind die Augenlider gar geschlossen, ist eine Datenaufnahme unmöglich. Die Auswertung von einzelnen Sequenzen ist somit nicht möglich. Dies reduziert die zur Verfügung stehenden Analysedaten und beeinträchtigt unter Umständen das Ergebnis. Die Augendaten und die dargestellten Stimuli müssen mit nachvollziehbarem Zeitstempel vorliegen, sodass Augenwerte und gezeigte Testbilder/Audiospuren miteinander in Verbindung gebracht werden können.

5.2 Optische Methoden in der Wissenschaft

Im Folgenden werde ich die meist genutzten optischen Methoden zur Erforschung von psychischer Beanspruchung auführen und erklären.

5.2.1 Sternberg's Memory Scanning Paradigma

Das Sternberg Memory Scanning Paradigma ist ein Paradigma, welches von Saul Sternberg 1966 zum ersten Mal durchgeführt wurde. Der Test besteht aus 4 Stufen der Informationsverarbeitung:

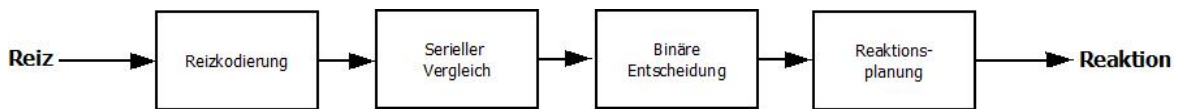


Abbildung 5.1: Vier Stufen der Informationsverarbeitung im Sternberg Task

Dieser Test spricht gleichzeitig die Speicher- und Verarbeitungseigenschaft des Arbeitsgedächtnisses an. Als Reiz wird bei diesem Test eine Zahlenfolge (Memory Set) von einer bis sechs einstelligen Zahlen auf dem Bildschirm präsentiert. Die Länge der Zahlenreihe wird von Versuch zu Versuch variiert. Jede Zahl wird einzeln für 1,2 Sekunden angezeigt. Danach folgt eine Pause von zwei Sekunden. Im Anschluss wird ein Probestimulus gezeigt, dieser besteht aus einer einzelnen Zahl. Der Proband hat nun die Aufgabe so schnell wie möglich zu entscheiden, ob die nun angezeigte Zahl Teil der präsentierten Zahlenfolge ist oder nicht. Dazu muss er einen seriellen Vergleich mit der Folge anstellen und eine binäre Entscheidung (Ja/Nein) treffen. Die Reaktion wird mental geplant und daraufhin ausgeführt. Die Wahrscheinlichkeit, dass die Probe Teil der Zahlenfolge ist, beträgt 50%.

Eine Abhängigkeit von Reaktionsgeschwindigkeit und Anzahl der Zahlen in der Folge ist zu erkennen. Die Reaktionszeit steigt linear zu der Anzahl der Zahlen im Memory Set. Der Proband versucht in einer ebenso schnellen Zeit mehr Vergleiche zu ziehen, dadurch steigt seine mentale Beanspruchung. Da die Reaktionsgeschwindigkeit allerdings auch durch andere Einflüsse - wie zum Beispiel durch Alter - verlangsamt werden kann, ist die P300 als bevorzugtes Maß zu nehmen. Die Latenz der P300 ist gleich geblieben und ist unabhängig von altersbedingten motorischen Einschränkungen, die die gemessene Zeit bis zur Resonanz verändern. (vgl. [ANDREASSI, 2000], [CORBIN

und MARQUER, 2009], [STERNBERG, 1966]) In der Auswertung dieses Tests wurden 3 mögliche Pupillenreaktionen von Beatty und Lucero-Wagoner beobachtet und durch Klinger bestätigt. [KLINGER et al., 2008]

1. Mean dilation: Dilation nach dem Stimulus in der Verarbeitungsphase
2. Peak dilation: Maximale Dilation während der Verarbeitungsphase (mean dilation)
3. Peak latency: Zeit zwischen Probestimulus und Dilatationsmaximum (peak dilation)

5.2.2 Digit span forward

Die Digit Span Methode ist ein altes und einfaches Verfahren zur Erzeugung von mentaler Beanspruchung. Sie wurde 1966 von Kahneman und Beatty zum ersten Mal in Verbindung mit der Pupillengröße veröffentlicht.

Als Stimuli werden Zahlenreihen mit der Länge von 3 bis 7 Zahlen gezeigt. Jede Zahl wird für eine Sekunde auf dem Bildschirm dargestellt. Nach der Folge schließt sich eine Pause von 2 Sekunden an, in der der Proband sich die Zahlenreihe merken muss. Anschließend wird das Gelernte in einer beliebigen Art abgefragt. Dies kann sowohl mündlich als auch schriftlich geschehen.

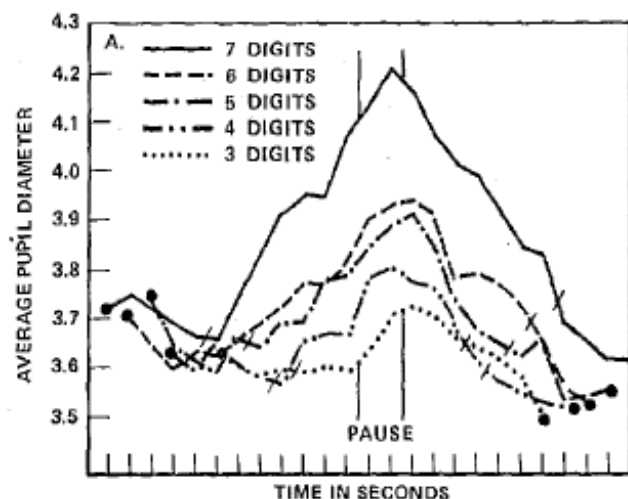


Abbildung 5.2: Pupillendurchmesser in Millimeter in Abhängigkeit der Zeit und Länge der zu merkenden Zahlenreihe während des Digit span Tests.

In dieser Graphik ist der Pupillendurchmesser in mm über die Zeit in Sekunden aufgetragen. Dieser wurde jeweils für die Zeit der Merkphase einer Zahlenreihe mit 3,4,5,6 und 7 Zahlen während des Digit span Tests gemessen. [BEATTY, 1982]

Wie in Abbildung 5.2 zu erkennen, ist ein Anstieg der Pupillengröße während der Aneinanderreihung der Zahlen zu erwarten. Eine Korrelation zwischen Anzahl der Elemente in der Folge und Pupillengröße ist deutlich zu erkennen. Während der Pausenphase geht der Proband die Zahlenreihe nochmals mental durch, was zu einem erneuten Anstieg der Pupillengröße führt. Während der Phase, in der die Reihe von

dem Probanden wiedergegeben wird, verkleinert sich die Pupille immer weiter. [BEATTY, 1982] [ANDREASSI, 2000]

5.2.3 Digit span backwards

Der Digit span backwards Test ist ähnlich aufgebaut wie der zuvor beschriebene vorwärts orientierte Digit Span Test. Eine Zahlenfolge von drei bis sechs Stellen wird visuell oder auditiv dem Probanden vorgestellt. Nach den Stimuli folgt eine kurze Phase des Merkens, welche meist 2 Sekunden andauert. Im Anschluss daran muss der Proband die Zahlenfolge rückwärts wiedergeben, beginnend mit der zuletzt dargebotenen Zahl. Dies erfordert weiterhin das Merken der gesamten Zahlenreihe und fordert mehr kognitive Leistung als die Variante, die die Zahlenreihe in der gezeigten Reihenfolge abfragt. (vgl. [BEATTY, 1982])

5.2.4 Oddball paradigm

Das Oddball Paradigma ist das am häufigsten genutzte Paradigma um ein P300-EKP (siehe 2.2.1) zu erzeugen. Dabei wird eine Reihe von schnell aufeinander folgenden Reizen dem Probanden dargeboten. Seltene Zielreize (engl. Oddball), die von den Standardreizen abweichen, werden in zufälliger Reihenfolge wiedergegeben. Standard- und Zielreiz können sowohl visueller als auch auditiver Natur sein. Beispielsweise kann ein Ton mit der Frequenz 1000Hz als Standardreiz eingesetzt werden. Der Zielreiz muss sich davon klar unterscheiden und stellt hier beispielsweise einen Ton mit der Frequenz 1250Hz dar. Es ist zu beachten, dass die Häufigkeit der Standardreize gewöhnlich 80-90% beträgt. Der Proband hat die Aufgabe, bei Identifikation eines solchen Zielreizes ein eindeutiges, sichtbares Zeichen zu geben (z.B. Klicken) oder diese Zielreize zu zählen. Zu Zwecken von psychologischen Untersuchungen (z.B. an schizophrenen Patienten) wird die Amplitude und Latenzzeit des P300 EKP aufgenommen und gemessen.

In Abbildung 5.3 ist die schematische Darstellung des akustischen Oddball tasks zu sehen. „S“ steht in diesem Falle für den Standardreiz, „Oddball“ für den seltenen Zielreiz.

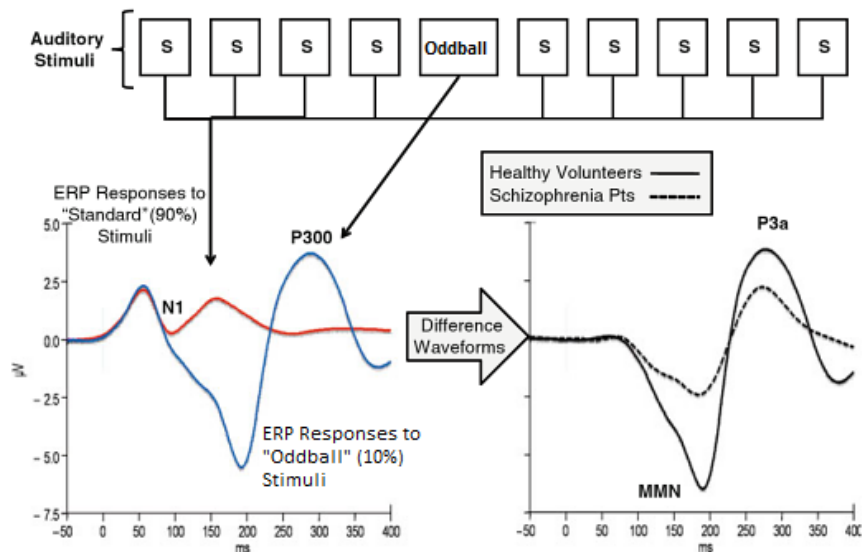


Abbildung 5.3: Oddball Test und EEG Ableitungen während Standard- (S) und Zielreize (Oddball).

Im oberen Teil der Graphik sind die Stimuli des Oddball Tasks mit entsprechenden EEG Ableitungen während der Stand- und Zielreize zu sehen. Die in Blau gezeichnete Funktion des linken Koordinatensystems entspricht der EEG Ableitung während eines Zieltons, die in Rot gezeichnete der, zum Zeitpunkt des Standardtons. Im rechten Graphen sind die EEG Ableitungen während eines Zielreizes jeweils bei einem gesunden Probanden, sowie eines Schizophrenen zu erkennen. Bei Letzterem kommt es zu einer stark verminderten Ausprägung der P300 Welle. [SWERDLOW, 2010]

Im unteren Koordinatensystemen sind die EEG-Ableitungen in mV über die Zeit aufgetragen. Beim unteren, linken Koordinatensystem ist zunächst der Unterschied von Standard- zu Zielton gegenübergestellt. Bei dem unauffälligen Standardreiz ist ein geringer Ausschlag über ca. $20\mu\text{V}$ (N1, rote Ableitung) zu erkennen. Eine große Amplitude ist hingegen bei dem Zielton zu erkennen (P300, blaue Kurve). Zunächst sinkt der Ableitstrom auf ca. -60mV und steigt in einer Zeit von 150ms auf ca. 30mV . Anschließend pendelt sich der Ableitstrom auf den Ausgangswert ein.

Im unteren, rechten Koordinatensystem sind die EEG-Ableitsignale nach dem Erscheinen des Zielreizes für gesunde sowie für schizophrene Probanden dargestellt. Es ist klar zu erkennen, dass es bei den gesunden Probanden einen deutlich größeren Ausschlag in positive wie negative Richtung auf der y-Achse gibt. Bei Schizophrenen ist eine ähnliche Signalform mit verringerter Amplitude, jedoch ohne Verzögerung, zu erkennen. Ein ähnlicher Signalverlauf ist für alkoholabhängige, demente oder depressive Patienten erkennbar. (vgl. [BLACKWOOD et al., 1987]) Die Abweichungen der Kurven sind bei schizophrenen und dementen Patienten auf neuronale Läsionen zurückzuführen, welche in der Hypocampusformation lokalisiert werden konnten. (vgl. [MAURER et al., 2005])

5.3 Umgesetzte Experimentenreihe

Bei meiner umgesetzten Experimentenreihe habe ich mich an den in der Wissenschaft etablierten Methoden orientiert. Somit kann davon ausgegangen werden, dass sich die Ergebnisse der durchgeführten Tests zu wissenschaftlich validen Zwecken nutzen lassen.

5.3.1 Anforderungen an die Probanden

Die Probanden sollten alle in einer Alterskategorie sein um vergleichbare Ergebnisse zu erhalten. Durch ein stark erhöhtes Alter sind die motorischen Fähigkeiten eingeschränkt und Aufgaben können nur mit hohem Aufwand erledigt werden. Zudem kommt hinzu, dass viele ältere Personen wenig Erfahrung mit der Benutzung eines Computers haben, was eine hohe mentale Belastung und somit keinen Ruhezustand

zur Folge hätte.

Schizophrene, Drogenabhängige und Personen mit Diabetes können zudem nicht an der Studie teilnehmen, da sie eine veränderte Resonanz zu gezeigten Stimuli zeigen würden. Auch die Pupillendilation ist bei von Drogen abhängigen oder unter Drogen stehenden Personen gestört. Meist ist eine Akkommodation bei diesen Gruppen nicht möglich.

Auch Stress, Depression oder starke Müdigkeit können die Ergebnisse verfälschen, da in den Ruhephasen nie vollkommene innere Ruhe erreicht wird und der Pupillendurchmesser nicht als wirkliche Baseline gesehen werden kann. Ein Anstieg der Pupillengröße kann so eventuell verringert sein und sich nicht klar von einem Ruhezustand unterscheiden.

5.3.2 Anforderungen an den Testplatz

Beleuchtung

Da der Pupillendurchmesser stark von der Beleuchtung abhängig ist, ist eine konstante Beleuchtung am Testplatz unbedingt erforderlich. Eine abrupte oder sich langsam verändernde Beleuchtungssituation würde den Pupillendurchmesser unabhängig von der kognitiven Beanspruchung verändern und so die Messwerte negativ beeinflussen und unbrauchbar machen. Aus diesem Grund beschränkt sich die Hintergrundfarbe des Bildschirms während der Experimente ausschließlich auf ein mittleres Grau, welches keine maximale oder minimale Pupillengröße beim Probanden hervorruft. Somit kann der äußere Einfluss auf den gemessenen Pupillenwert minimiert werden.

Eine oszillierende Beleuchtung, wie beispielsweise durch Leuchtstoffröhren, würde einen ähnlichen Effekt mit sich ziehen und gegebenenfalls zu oszillatorischen Veränderungen der Pupille führen. Ein mit LED- oder Halogenquellen beleuchteter Raum kann somit als geeignet angesehen werden. Es muss eine konstante Beleuchtungsstärke vor dem Gesicht des Probanden herrschen. Diese wird mithilfe eines Luxmeters in geringem Abstand zu den Augen des jeweiligen Versuchsteilnehmers gemessen.

Es müssen minimale und maximale Grenzwerte der Beleuchtungsstärke vorliegen, so dass es der Pupille möglich sein kann, sich zu erweitern oder ihre Größe zu reduzieren. Ist die Beleuchtungsstärke des Testplatzes beispielsweise zu gering, hat die Pupille bereits ihre maximale Ausdehnung erreicht und wird von einer veränderten kognitiven Beanspruchung nicht weiter beeinflusst. Vgl. [DIN, 1991]

Geräuschkulisse

Auch Geräusche können die Aufmerksamkeit des Probanden vom Test weg und hin zum Umgebungsgeschehen lenken. Dies könnte eine sprunghafte, kognitive Veränderung auslösen. Um dies zu vermeiden, wird die Testperson gebeten einen Gehörschutz bzw. halboffene oder geschlossene Kopfhörer während des gesamten Tests zu tragen. Die Wahrscheinlichkeit eines Konzentrationsverlustes oder einer Konzentrationsverlagerung ist so geringer. Aufgrund der Lage des Messraumes war ein völliger Ausschluss von jeglichen Geräuschen allerdings nicht möglich.

Positionierung

Der Proband wird in einem Abstand von 0,65 m zum Remote Eye-Tracking Device positioniert. Dieser Abstand wird vor Beginn jedes Experiments gemessen und die Sitzposition entsprechend angepasst. Da laut Spezifikation des Eye-Trackers eine Distanz Auge zu RED von 0,65m die Ideale ist, haben die Messwerte in diesem Bereich auch die höchste Genauigkeit.

Der Proband wurde, wie in Abbildung 5.4 zu sehen ist, vor einem separaten Monitor positioniert. Dies hat zur Folge, dass er keinen Einblick auf den Überwachungsmonitor des Testleiters hat. Dies würde zu einer großen Ablenkung führen. Der Proband verfügt zudem über eine eigene Maus, sowie Tastatur. So ist es ihm möglich, Anweisungen im Test zu folgen und während der Antwortphase jedes Trials zu interagieren.

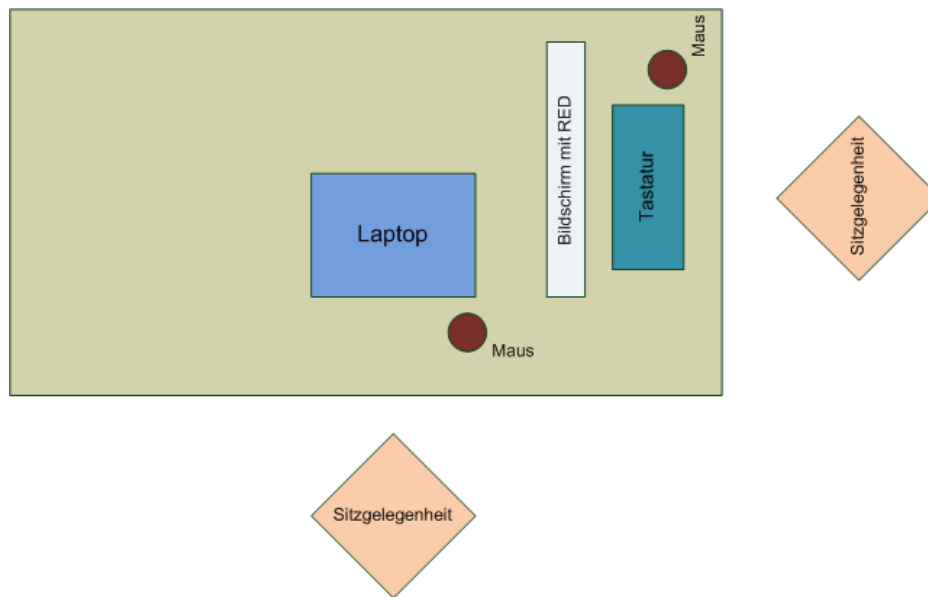


Abbildung 5.4: Aufbau und Positionierung des Testplatzes

5.3.3 Testablauf

Die Umsetzung meiner Experimentenreihe geschah mit dem Tool ExperimentSuite 3.5 der Firma Sensomotoric Instruments GmbH. Dieses ist eine Software, welche modulare Zusammenstellungen von visuellen, computerbasierten Tests ermöglicht und mit den Eyetracking Systemen, des selben Herstellers harmoniert. Text-, bild-, video- und websitesbasierte Stimuli sind mit diesem Programm aneinanderreihbar. Die Dauer der Einblendzeit ist für jedes Stimuli separat einzustellen.

Meine Experimentenreihe gliedert sich in 4 Untertests, wie in Abbildung 5.5 zu sehen ist. Diese vier Tests werden in vorher definierter, unterschiedlicher Reihenfolge durchlaufen. Eine vorige Festlegung der Reihenfolge ist nötig, um nicht beabsichtigte, subjektive Unregelmäßigkeiten zu vermeiden.

Jeder dieser 4 Tests beinhaltet einen Instruktions- und einen Aufgabenteil, wie in der Grafik 5.6 zu sehen ist. Als Trial wird in diesem Fall eine Spanne von Stimuli bezeichnet, die aus einer Pause vor der Aufgabe, einem Aufgabenteil (meist eine Zahlenfolge), einer Pause zu dem Merken der Stimuli und einer Abfrage besteht.

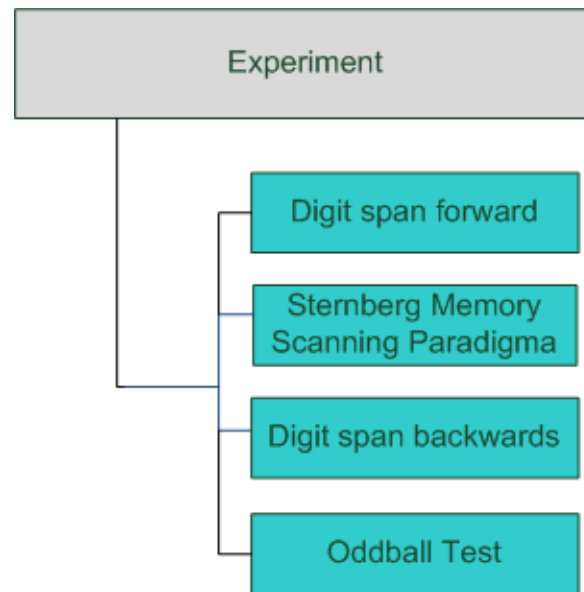


Abbildung 5.5: Möglichkeiten der Experimentenanordnung

Am Anfang des Instruktionsteils steht bei jedem der Experimente eine Kalibrierung. Diese ist je nach Trackbarkeit der Probanden eine 5-Punkt- oder eine 2-Punkt-Kalibrierung. Unter Kalibrierung versteht man hierbei das Anzeigen von 2 oder 5 Punkten auf dem Bildschirm, welche vom Probanden fixiert werden müssen. Detektiert die Software eine Fixation im Bereich des angezeigten Punktes, so wird dieser ausgeblendet und weitere Punkte werden angezeigt. Damit ist die Software anschließend individuell auf den Nutzer abgestimmt. Diese Kalibrierung wird vor jedem der 4 Experimententeile vorgenommen, damit eine Unabhängigkeit der Reihenfolge gegeben ist. Auch sind so kurze Pausen zwischen einzelnen Testsegmenten möglich.

Im Anschluss an die Kalibrierung sieht der Proband eine visuelle Erläuterung über den Ablauf des nächsten Tests. Dazu gehört die Definition der Stimuli, die zu sehen sein werden und wie der Proband darauf reagieren soll. Um möglichst wenig Fragen offen zu lassen, wird im Anschluss an die Instruktion ein Video mit einem Beispiel gezeigt. Darauf folgt eine Erklärung darüber, dass nun aufgekommene Fragen gestellt werden können, und dass das Experiment per Drücken der Leertaste gestartet werden kann. Somit kann der Proband selbst entscheiden, wann das Experiment startet und

ob er die Möglichkeit einer Nachfrage nutzen möchte. Mit dem Drücken der Leertaste wird das erste Trial des aktuellen Unterexperiments gestartet. Als Trial wird in diesem Fall eine Spanne von Stimuli bezeichnet, die aus einer Pause vor der Aufgabe, einem Aufgabenteil (meist eine Zahlenfolge), einer Pause zum Merken der Stimuli und einer Abfrage besteht.

Im Aufgabenteil werden die einzelnen Trials nacheinander als Stimuli angezeigt. Vor jedem Trial wird eine kurze Pause eingefügt, um ein Absenken der Pupillengröße zu ermöglichen. Diese Pause darf allerdings nicht beliebig lang gewählt werden, da es sonst zu Verwirrung des Probanden führen kann und dieser sich über die Situation selbst Gedanken macht, anstatt die Pause nur als Ruhezeit zu nutzen. Das eine kurze Pause dazu eher dienlich ist, kann nicht validiert werden, da spontane Gedankengänge nicht beeinflussbar oder messbar sind.

Oddball paradigma

In meiner Umsetzung des Oddball Paradigmas stellt ein 1000Hz-Sinuston den Standardstimulus dar, ein 1250Hz-Sinuston den Trigger- oder Zielreiz. Die Häufigkeit des Standardtons beträgt 90-95%. Der Proband wird angewiesen, nach dem 1250Hz-Ton so schnell wie möglich auf ein schwarzes Kästchen zu klicken, welches sich auf dem grauen Hintergrundbildschirm befindet. Dadurch wird gewährleistet, dass der Proband nicht vom Bildschirm wegsieht und somit so viele Augendaten wie möglich aufgenommen werden können.

Der Ablauf des Oddball Tests ist identisch zu dem in Abbildung 5.3 gezeigten Stimulireihe. Die Anzeigedauer beträgt für die Darbietung eines Stimulus eine Sekunde. Für die Pausen zwischen den aufeinander folgenden Stimuli habe ich eine AOI (Area of Interest)-Zone um den anzuklickenden Punkt in der Mitte des Bildschirms integriert. Wird diese Zone für mehr als 1250ms vom Probanden fixiert, ist die Pause beendet. Andernfalls dauert diese Pause 3000ms und der Proband kann daran erinnert werden, mit geöffneten Augen die Bildschirmmitte zu fixieren.

	Art der Anzeige	Dauer
Instruktionsteil	Kalibrierung	individuell
	Instruktion	individuell
	Beispiel	individuell
	Möglichkeit zu Fragen	individuell
Aufgabenteil	Pause	1.5 s
	Stimuli 1	1 s
	Stimuli 2	1 s
	Stimuli 3	1 s
	⋮	⋮
	Pause	2 s
	Abfrage	individuell
	⋮	⋮

Abbildung 5.6: Ablauf Digit Span

In dieser Abbildung wird der Ablauf des Digit Span Tests schemenhaft dargestellt. Jedes Experiment besteht aus einem Instruktionsteil (hier rot unterlegt) und einem Aufgabenteil (hier blau unterlegt). Der Instruktionsteil ist bei jedem Probanden identisch. Die Anzahl der Stimuli sind je Trial unterschiedlich lang, was mit 3 Punkten im Schaubild verdeutlicht werden soll.

Kapitel 6

Algorithmische Umsetzung

Für die Auswertung habe ich das Programm Matlab der Firma The MathWorks genutzt, da es viele Vorteile in der Handhabung großer Datenmengen bietet und Berechnungen aufgrund der Matrizenform leicht zu bewerkstelligen sind. Außerdem ist eine Reimplementierung so einfach möglich, da die Software weit verbreitet ist.

6.1 Eventdaten

Nach erfolgreichem Durchführen der Experimente mit Experiment Center, liegen die Augendaten zunächst in Form eines iDataFiles (weiter IDF genannt) vor. Diese beinhaltet alle Augen- und Kalibrierungsdaten. Das Experiment mit all seinen Ausgaben muss anschließend in das Programm “BeGaze“ geladen werden. Dies ist eine unternehmeneigene Software, welche zur Ausgabe und Veranschaulichung der gemessenen Daten herangezogen wird.

Für meine Experimente habe ich das Programm lediglich dazu verwendet, die gespeicherten Daten aus der IDF-Datei in eine lesbare Textdatei umzuwandeln, welche aus den gemessenen Daten nur die für mich relevanten Daten enthält. Diese sind, wie in Abbildung 6.1 zu sehen: Zeitstempel, Pupillendurchmesser des linken und rechten Auges, Angabe zu Fixation und Blinzeln, sowie der angezeigte Stimulus zu jedem Messzeitpunkt. Außerdem enthält diese Datei alle Daten zur Aufnahme, welche zur Reproduktion notwendig sind.

Time	Type	Trial	L Pupil Diameter [mm]	R Pupil Diameter [mm]	Frame	Aux1	L Event Info	R Event Info	Stimulus
5610556557	SMP	1	3,55	3,67	-1		Fixation	Fixation	instruction test.jpg
5610573329	SMP	1	3,54	3,67	-1		Fixation	Fixation	instruction test.jpg
5610590014	SMP	1	3,55	3,68	-1		Fixation	Fixation	instruction test.jpg
5610606728	SMP	1	3,54	3,67	-1		Fixation	Fixation	instruction test.jpg
5610623327	SMP	1	3,52	3,67	-1		Fixation	Fixation	instruction test.jpg
5610641429	SMP	1	3,51	3,67	-1		Fixation	Fixation	instruction test.jpg
5610656707	SMP	1	3,5	3,66	-1		Fixation	Saccade	instruction test.jpg
5610673321	SMP	1	3,63	3,86	-1		-	Blink	instruction test.jpg
5610690028	SMP	1	3,58	0	-1		-	Blink	instruction test.jpg
5610706719	SMP	1	0	0	-1		Blink	Blink	instruction test.jpg
5610723381	SMP	1	0	0	-1		Blink	Blink	instruction test.jpg

Abbildung 6.1: Beispiel für Augen- und Eventdaten in Textdatei

6.2 Einladen der Daten in Matlab

Das Einladen der Textdatei in Matlab erfolgt mithilfe eines bereits existierenden Matlabscriptes. Dieses wurde mir von dem Unternehmen SMI für meine Bachelorarbeit zur Verfügung gestellt. Es durchsucht die Datei nach dem Beginn der Header und lädt die Augendaten in die Matlab Workspace. Spezielle SMI-interne Informationsmeldungen, welche in der Textdatei als “Messages“ zu erkennen sind, werden dabei eliminiert, da diese nicht zur Auswertung hinzugezogen werden. An dieser Stelle beginnt mein Eigenanteil des Programmcodes, welcher alle weiteren Berechnungen der Augendaten beinhaltet.

6.3 Bereinigen von Ausreißern und invaliden Werten

6.3.1 Notwendigkeit der Bereinigung

Da es sich beim Messobjekt um ein lebendiges Individuum handelt, ist eine Unterdrückung des physiologisch notwendigen Blinzeln nicht möglich. Diese verhindern unter anderem allerdings eine kontinuierliche Messung.

Gründe für Ausreißer und Nullsamples können sein:

- *Nullsamples durch Blinzeln:* Für die Zeitspanne des geschlossenen Auges sind alle Pupillendaten gleich Null.

- *Ausreißer durch Lidschluss*: Während sich das Augenlid gerade schließt bzw. öffnet, ist nur ein Teil der Pupille sichtbar. Die Angabe der Pupillengröße durch den Eyetracker ist in dieser Zeitspanne nicht verlässlich und ergibt häufig zu kleine Werte aus. Diese peakartigen Messwerte müssen, genau wie die Werte für das geschlossene Auge, ermittelt und geglättet werden.
- *Messfehler*: Hat nur ein einzelnes Sample den Wert Null oder liegt weit außerhalb der übrigen Werte, so handelt es sich hierbei mit großer Wahrscheinlichkeit um einen Messfehler. Da der Lidschluss, das Blinzeln und das erneute Öffnen des Lids mindestens 200-300 ms dauert, ist nicht davon auszugehen, dass es durch ein solches Event hervorgerufen wurde, da dieses eine größere Reihe von Nullsamples ergeben würde. Aufgrund der Messfrequenz von 60Hz beträgt die Differenz zwischen zwei Messwerten 16ms.

Zusätzlich kann eine Fehldetektion über mehrere Samples vorliegen. Diese Werte können sowohl auf- als auch absteigend sein. In diesem Fall normalisiert sich der gemessene Wert innerhalb einiger (<10) Samples meist wieder.

Wie jedes Messgerät liefert auch ein Eyetracker nur Werte mit einer endlichen Genauigkeit. Die Messwerte der Samples unterliegen einem statischen Rauschen, das durch unvermeidbare Ungenauigkeiten bei der Datenaquise hervorgerufen wird.

Da mich in meiner Auswertung nicht nur der absoluten Pupillendurchmesser, sondern auch dessen zeitliche Ableitungen interessieren, ist eine Glättung der Werte unbedingt notwendig. Anderenfalls könnte dies beispielsweise bei sprunghaftem Anstieg zu fehlerhaften Rückschlüssen führen.

In Abbildung 6.2 sind die Pupillenrohdaten eines beliebigen Probanden während des Digit Span Tests zu sehen. Man erkennt deutlich, dass es zu einigen Blinzeln während des gezeigten Ausschnitts kam. Dies ist immer genau dann der Fall, wenn die Größe der Pupille über mehrere Messwerte hinweg gleich Null ist.

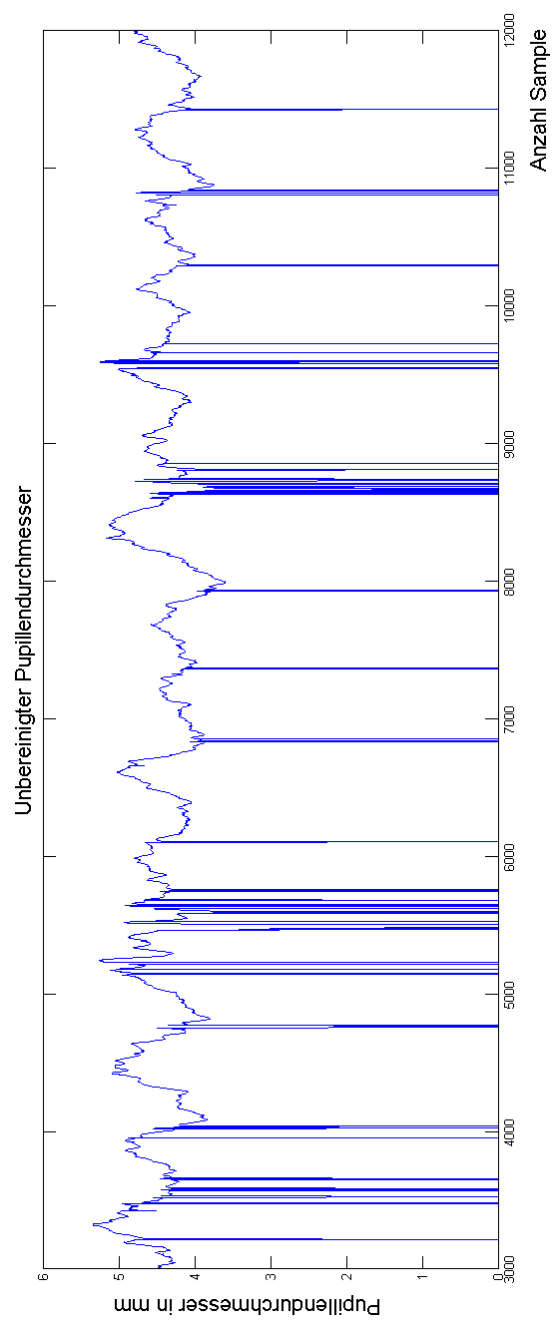


Abbildung 6.2: Unbereinigter Pupillendurchmesser über die Zeit in Sekunden

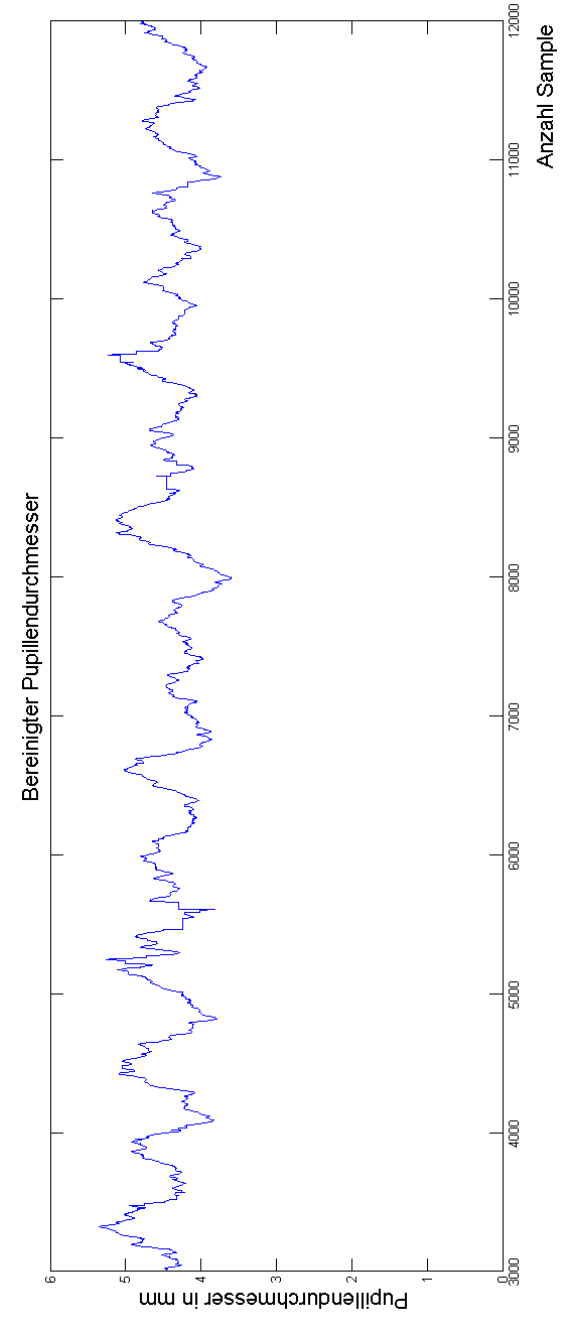


Abbildung 6.3: Bereinigter Pupillendurchmesser in Millimeter über die Zeit in Sekunden

6.3.2 Detektion von Ausreißern und invaliden Werten

Als Ausreißer verstehen sich hier diejenigen Messwerte, die sich wie im Abschnitt zuvor beschrieben durch Blinzeln, Lidschlüsse und Messfehler ergeben. Außerdem gibt es eine Reihe von Werten, die die Software auf Null setzt, da sie nach einer internen Prüfung als nicht valide gelten. Ein Beispiel für eine Messreihe mit vorhandenen Ausreißern und invaliden Werten ist in Abbildung 6.2 zu sehen.

Algorithmisch bedeutet dies für meine Umsetzung, dass zunächst die Differenz zwischen zwei Messwerten ermittelt werden muss um den Zeitpunkt großer Veränderungen zu detektieren. Diese Werte werden anschließend genullt.

Einzelne Ausreißer werden durch die absolute Differenz zum nächsten Sample ausgemacht. Ist dieser größergleich 0,2mm, ist von einem invaliden Wert auszugehen, da sich die Pupille innerhalb von 16ms (Zeitdifferenz zwischen zwei Samples) nicht in diesem Ausmaße ändern kann.

Um auch Ausreißer auszumachen, die nicht vereinzelt vorkommen, sondern an deren Wert sich stetig angenähert wird, reicht eine einfache Sample zu Sample Differenz nicht aus. Für eine solche Fehldetektion habe ich die Pupillendurchmesser von 10 aufeinanderfolgenden Werten betrachtet. Der aufsummierte Betrag aller 10 Sampledifferenzen zueinander wurden anschließend mit einem Schwellwert verglichen.

Der Schwellwert (siehe 6.4) ergibt sich aus 0,15 multipliziert mit dem Quotient aus der Anzahl von validen Werten der betrachteten Spanne und der Länge der Spanne. Sind alle Werte der Spanne ungleich Null, so wird an dieser Stelle mit 1 multipliziert. Die Bildung des Quotienten ist notwendig, da sich der Schwellwert je nach Anzahl der betrachteten Werte in einer Spanne verändert. Umso weniger Werte im vorigen Schritt summiert worden sind, umso niedriger ergibt sich der Schwellwert. Überschreitet die errechnete Summe den Schwellwert, wird der Wert des Pupillendurchmessers an der Stelle j auf Null gesetzt.

Ist der Schwellwert nicht überschritten, so behält der Pupillendurchmesser seinen ursprünglichen Wert.

Eingaben

```

1  sumDiff=0           // Summe der Differenzen von Sample zu Sample
2  diffSample=0       // Differenz zwischen einzelnen Samples
3  span = 0           // Laenge der zu überprüfenden Messwertspanne
4  countvalid = 0     // Menge der validen Samples in Spanne
5  PupilDia = {(1, y1), ..., (N, yN)} // Pupillendurchmesser in Millimeter

```

Algorithmus

Label1: // Bildung von Summe der Differenzen in Spanne

```

6  for Alle positiven PupilDia-Werte i;
7    for j = 0 bis Laenge von span;
8      ; // Summe der einzelnen Differenzen in Spanne
9      sumDiff = sumDiff + Betrag(PupilDia(i - j) - PupilDia(i - j - 1));
10     ; // Absolute Differenz von Sample zu Sample
11     diffSample = Betrag(PupilDia(i) + PupilDia(i - 1))/2;
12     countvalid = countvalid + 1; // Zählen von validen Werten in Spanne

```

Label 2: // Berechnung Schwellwert

```

13  Wenn sumDiff >= 0.15 * (countvalid/span);
14  oder;
15  Wenn diffSample >= 0.2 ;
16    Rückgabewert = invalid ;
17  sonst;
18    Rückgabewert = valid ;

```

Rückgabe

```

19  if Rückgabewert = invalid
20    PupilDia(i) = 0
21  if Rückgabewert = valid
22    PupilDia(i) = PupilDia(i)

```

Abbildung 6.4: Detektion von Ausreißern und invaliden Werten

6.3.3 Interpolation von Ausreißern und invaliden Werten

Die im vorigen Schritt genullten Ausreißer und die Blinzler, welche schon zuvor den Wert Null besessen haben, müssen nun aus den in Abschnitt 6.3.1 beschriebenen Gründen interpoliert werden.

Zur Interpolation der Daten habe ich mich zunächst für ein einfaches Ersetzen der Ausreißer- und Blinzelsamples mittels Auffüllen von Mittelwerten entschieden. Dieser Mittelwert setzt sich aus den vorangegangenen und folgenden validen Messwertes zusammen. Bei dieser Umsetzung sehe ich allerdings noch die Möglichkeit zur Verbesserung. Details könnten unter Rücksichtnahme des vorigen und sich anschließenden Anstiegs noch angepasst werden.

Als Sonderfall müssen zwei Szenarios betrachtet werden. Beginnt oder endet die Messreihe mit einem invaliden Sample, so wird nicht der Mittelwert gebildet, da der zweite Summand fehlen würde. Im Falle eines Nullwertes zu Anfang der Messreihe, werden die invaliden Messwerte mit dem nächsten validen Wert ersetzt. Bei einem Enden der Messreihe mit dem Wert Null wird er mit dem zuletzt als gültig detektierte Wert ersetzt.

6.4 Normierung der Pupillendaten

Die Standardisierung der Pupillendaten ist notwendig, um eine Vergleichbarkeit zwischen verschiedenen Probanden zu ermöglichen, da die Pupillengröße im Ruhezustand für verschiedene Probanden unterschiedlich ist. Nicht nur die höherwertigen Messwerte sind relevant, sondern auch die kleineren Messwerte. Teil der Standardisierung, welche in Gleichung 6.1 zu sehen ist, ist die Subtraktion des Mittelwertes \bar{x}_j der gesamten Messreihe vom aktuellen Wert. Dadurch wird die unterschiedliche Grundpupillengröße der einzelnen Probanden eliminiert. Eine Vergleichbarkeit wird somit erleichtert. Der zweite Teil der Standardisierung ist das Dividieren des vorigen Ergebnisses durch die Standardabweichung, welches durch den Matlab-Befehl “std(X)” errechnet wurde. X steht hierbei für die bereits bereinigten Pupillendaten eines Auges. [KESSLER, 2007]

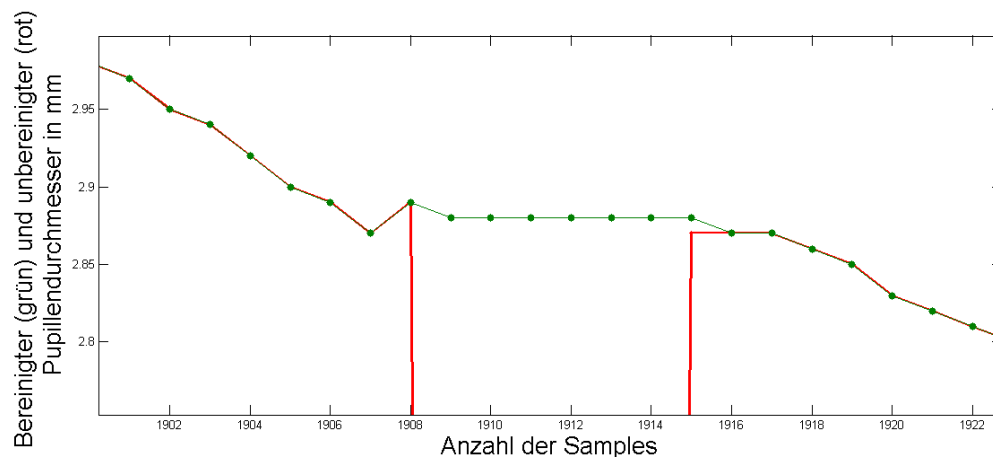


Abbildung 6.5: Beispiel für interpolierte Messwerte

Der hier rot gezeichnete Verlauf zeigt die ungefilterte Messreihe. Dabei kommt es bei 7 Messwerten zu einem Wert, der gleich Null ist. Grün sind hier die von mir interpolierten Messpunkte dargestellt. Es ist eine Verbindung mittels einer Geraden zu erkennen.

$$z_i = \frac{x_i - \bar{x}}{s_i} \quad (6.1)$$

i = 1... N Objekte

x_i = Messwert für das Objekt i

\bar{x} = Mittelwert über alle Messwerte

s_i = Standardabweichung der Pupillendurchmesser

z_i = Standardisierter Pupillenmesswert

6.5 Bestimmung des Anstiegs

Die Messungen wurden mit einer Frequenz von 60 Messwerten pro Sekunde aufgenommen. Bei der Messung treten aufgrund der physikalischen Limitierungen des Messgerätes Ungenauigkeiten auf. Die Messwerte sind sowohl zeitlich nicht exakt gleichverteilt, als auch mit einem natürlichen Rauschen behaftet.

Gerade bei der Berechnung von Ableitungen verursacht Rauchen große Fehler. Das Signal muss also vorher stark geglättet werden, um lokale Anstiegsanomalien zu unterdrücken.

Für meinen Algorithmus habe ich eine Fenstergröße von 60 Samples gewählt. So bleiben die relevanten Änderungen erhalten, die Unterschiede die nur sampleweise vorliegen sind jedoch geglättet. Um nun den Anstieg zu bestimmen wird die Punktsteigungsform 6.2 genutzt, da die Punkte der x- und y- Richtung bereits bekannt sind. Als Wert der y-Achse fungiert der gefensterte Pupillendurchmesser in mm, als Werte der x-Achse die aktuelle Anzahl der Messwerte. Dazu wurde zunächst ein Zähler initialisiert.

Der aktuelle Wert bildet in meinem Algorithmus die Mitte des betrachteten Fensters. Die Existenz der nachfolgenden Messwerte mit einer Anzahl von mindestens der halben Fenstergröße muss dazu gegeben sein. Eine leichte Verzögerung ist somit notwendig, da eine Vorlaufzeit von, in diesem Fall, 0.5 Sekunden, gegeben sein muss.

$$y_{(t-\frac{span}{2})} - y_{(t+\frac{span}{2})} = m \cdot (x_{(t-\frac{span}{2})} - x_{(t+\frac{span}{2})}) \quad (6.2)$$

$span$	=	<i>Spanne der Anstiegsbestimmung</i>
$y_{(t-\frac{span}{2})}$	=	<i>gegltteter Pupillendurchmesser zur Zeit $t_{(t-\frac{span}{2})}$</i>
$y_{(t+\frac{span}{2})}$	=	<i>gegltteter Pupillendurchmesser zur Zeit $t_{(t+\frac{span}{2})}$</i>
$x_{(t-\frac{span}{2})}$	=	<i>Anzahl der Samples zur Zeit $t_{(t-\frac{span}{2})}$</i>
$x_{(t+\frac{span}{2})}$	=	<i>Anzahl der Samples zur Zeit $t_{(t+\frac{span}{2})}$</i>
m	=	<i>Steigung</i>

Wie in zahlreichen Veröffentlichungen beschrieben, steigt der absolute Durchmesser der Pupille mit steigender mentaler Beanspruchung. Anhand des Anstiegs soll dies auch in meinen Algorithmus einfließen.

6.6 Integrierung in Fuzzy-Logic-Toolbox

Die Möglichkeit der unscharfen Beschreibung von Zuständen und Größen innerhalb von Fuzzy-Sets eignet sich aufgrund Verschiedenartigkeit der Messgrößen bei Lebewe-

sen besonders gut. Eine Beschreibung des aktuellen normierten Pupillendurchmessers anhand einer linguistischen Variable ist somit nicht zwischen zwei absoluten Werten eingeschränkt, sondern lässt sich offener erfassen.

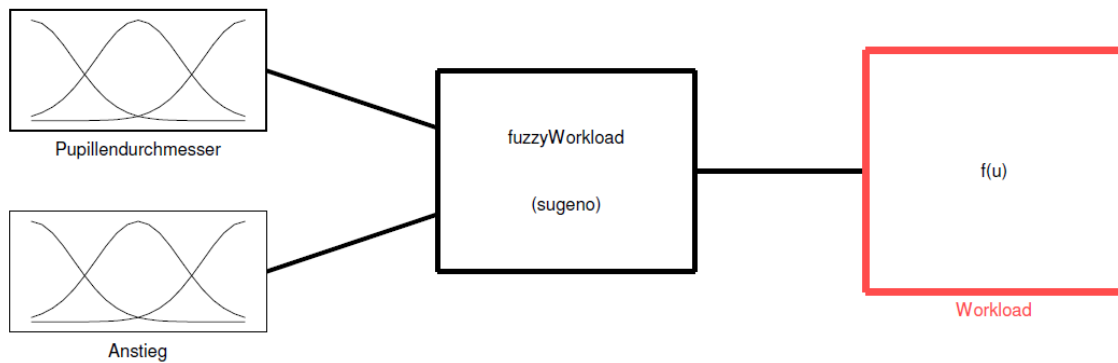


Abbildung 6.6: Übersicht über Fuzzy-System

Das Fuzzy-System besteht, wie im obigen Schema zu erkennen, aus zwei Eingangsvariablen: dem normierten Pupillendurchmesser und dem normierten Anstieg der Pupille. Als Operator wird die Inferenzbildung nach Sugeno genutzt. Dementsprechend resultiert ein einziger Ausgabewert, der den Workload beschreiben soll.

6.6.1 Eingangsdaten

Als Eingangsdaten meines Fuzzy-Klassifikators habe ich die, in Abschnitt 6.4 beschriebene, normierte Pupillengröße und den, in Abschnitt 6.5 beschriebenen normierten Anstieg der Pupillengröße verwendet.

Eine Verwendung von zwei linguistischen Variablen ist nötig, da auch bei einem geringen Pupillendurchmesser durchaus eine mentale Beanspruchung stattfinden kann, welches sich nur im Anstieg dessen zu erkennen ist. Ebenso verhält es sich mit einem mittleren positiven Pupillenwert. Dieser kann bei Vorkommen in einer Steigungsphase auf eine Erhöhung der Beanspruchung hindeuten, während eines großen Abstiegs aber auch auf das Gegenteil.

Linguistische Variablen

Als linguistische Variablen fungieren die in Abbildung 6.6 aus Input gezeigten Mengen „Pupillendurchmesser“ und „Anstieg“. Bei der Menge „Pupillendurchmesser“ handelt es sich um die in den Abschnitten 6.4 und 6.5 beschriebene Menge an Werten.

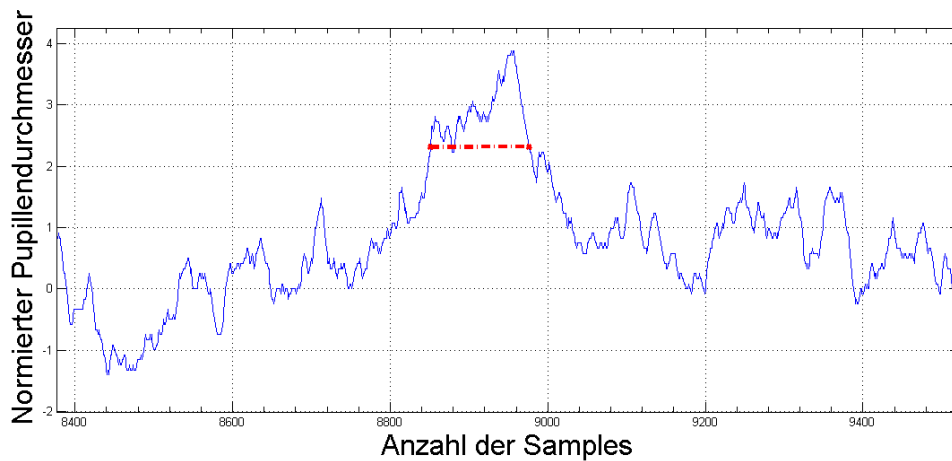


Abbildung 6.7: Beispiel für mittleren Pupillenwert in An- und Abstiegsphase

Die hier rot eingezeichnete Gerade macht sichtbar, wie sehr der Anstieg eines Punktes für die Klassifizierung seiner Zugehörigkeit ist. Der Messpunkt am linken Ende der Geraden steht stellvertretend für einen Moment, in dem sich die kognitive Belastung eventuell erhöht. Der Wert rechts der Geraden steht für einen Zeitpunkt während eines starken Abstiegs, in dem der Proband ggf. wieder weniger belastet wird.

normierter Pupillendurchmesser

Die linguistische Variable „normierter Pupillendurchmesser“ wurde in drei Terme aufgeteilt. Diese untergliedern sich wie in Abbildung 6.8 zu sehen in «Gering», «Mittel» und «Groß».

Die Zugehörigkeitsfunktionen werden durch eine Gaussfunktion 6.3 beschrieben. Diese besitzt für den Ab- bzw. Anstieg jeweils andere σ und c Parameter. Dies macht eine komfortable Anpassung der Bereiche, sowie eine Art Plateau möglich, zu der die Zugehörigkeit = 1 gehört. Die langsamere Annäherung an den Wert Null birgt hier den Vorteil, dass der Übergang zur nächsten Zugehörigkeit fließender ist. Dies macht die Zugehörigkeit „offener,, und den Übergang fließender.

Die Zugehörigkeitsfunktionen für den normierten Pupillendurchmesser (Pupil) können wie in Formel 6.3 beschrieben werden.

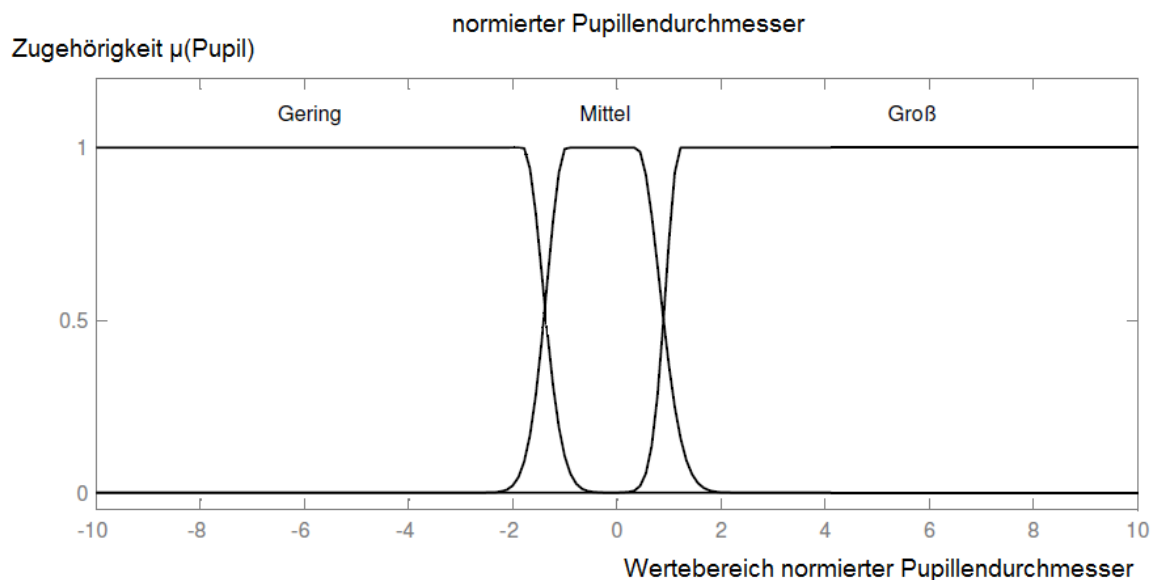


Abbildung 6.8: Zugehörigkeitsfunktion von normiertem Pupillendurchmesser

$$A(\mu_X) = e^{-\frac{(x-c)^2}{2 \cdot \sigma^2}} \quad (6.3)$$

$$\begin{aligned} \mu_X(x; \sigma, c) &= [\sigma_1 \ c_1 \ \sigma_2 \ c_2] \quad \text{für } x \in \mathbb{R} \\ \mu_{Gering}(x; \sigma, c) &= [0.03 \ -10.49 \ 0.4 \ -1.8] \quad \text{für } x < 0 \\ \mu_{Mittel}(x; \sigma, c) &= [0.3 \ -1 \ 0.4 \ 0.3] \quad \text{für } -2 < x < 2 \\ \mu_{Gross}(x; \sigma, c) &= [0.3 \ -1.2 \ 4.0 \ 11.5] \quad \text{für } x > 0 \end{aligned}$$

$$\begin{aligned} \mu_X(x) &= \text{Zugehörigkeitsgrad von Wert } x \text{ zu linguistischer} \\ &\text{Variable } X, \text{ für } \mu \in [0, 1] \\ x &= \text{aktueller Wert des normierten Pupillendurchmessers} \end{aligned} \quad (6.4)$$

Normierter Anstieg der Pupillengröße

Als Eingangsvektor fungiert hierbei der in Abschnitt 6.5 ermittelten Werte.

Die linguistische Variable „normierter Anstieg“ wurde in drei Terme aufgeteilt. Diese untergliedern sich wie in Abbildung 6.9 zu sehen in «Gering», «Mittel» und «Groß». Die Zugehörigkeitsfunktionen wurden nach ausführlicher Analyse der Messdaten ausgemacht.

Es handelt sich bei allen Termen um eine Trapezfunktion (siehe 6.5), wobei die Terme «Gering» und «Gross» nur über einen Bereich des Anstiegs verfügen. Eine Trapezfunktion wurde gewählt, da die Anstiege linear sind und die Annäherung an Null im Gegensatz zur Gausskurve steil ist. Eine striktere Abgrenzung der Zugehörigkeit war hier nach einer Datenanalyse notwendig. Alle Werte, die kleiner Zwei bzw. Größer Zwei sind, gelten als maximal anzunehmende Werte. In diesem Bereich ist der maximale An- bzw. Abstieg erreicht. Die Ausprägung dieses maximalen Ausschlags ist individuell abhängig vom Probanden.

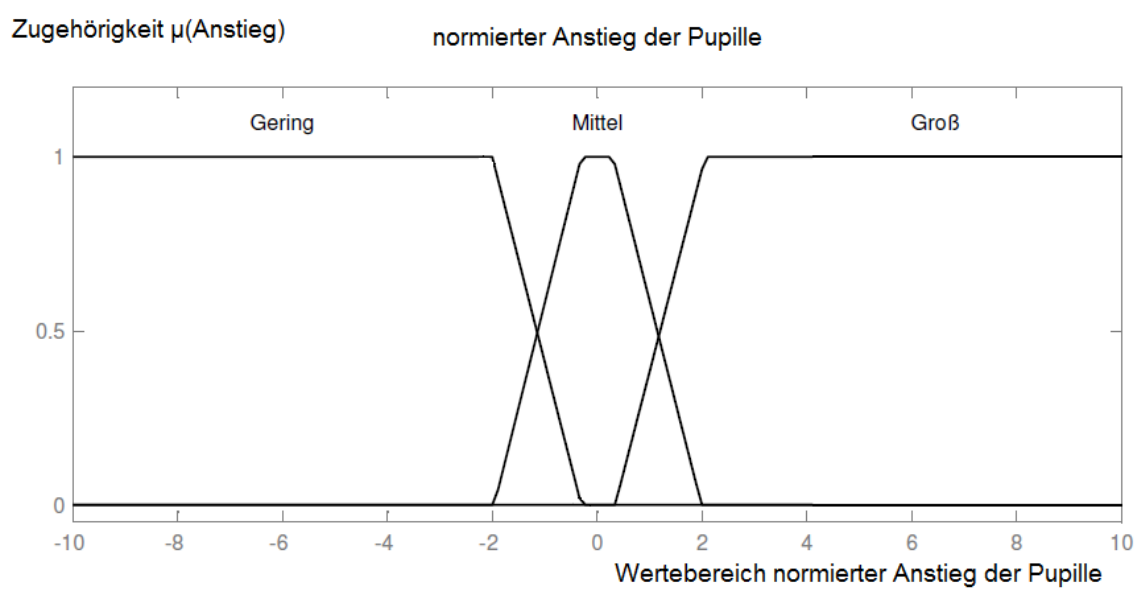


Abbildung 6.9: Zugehörigkeitsfunktion von normiertem Anstieg des Pupillendurchmessers

$$\mu_{Gering}(x) = \begin{cases} 1 & \text{für } x < -2 \\ \frac{-0.3-x}{-0.3+2} & \text{für } -2 \leq x < -0.3 \\ 0 & \text{für } x \geq -0.3 \end{cases} \quad (6.5)$$

$$\mu_{Mittel}(x) = \begin{cases} 0 & \text{für } x < -2 \\ \frac{x+2}{-0.3+2} & \text{für } -2 \leq x < -0.3 \\ 1 & \text{für } -0.3 \leq x < 0.3 \\ \frac{2-x}{2-0.3} & \text{für } 0.3 \leq x \leq 2 \\ 0 & \text{für } x > 2 \end{cases} \quad (6.6)$$

$$\mu_{Gross}(x) = \begin{cases} 0 & \text{für } x < 0.3 \\ \frac{x-0.3}{2-0.3} & \text{für } 0.3 \leq x < 2 \\ 1 & \text{für } x \leq 2 \end{cases}$$

$\mu_X(x)$ = Zugehörigkeitsgrad von Wert x zu linguistischer Menge X , für $\mu \in [0, 1]$

x = aktueller Wert des normierten Anstiegs

(6.7)

6.6.2 Methode und Ausgabevariablen

Als Operator habe ich mich für die Inferenzbildung nach Sugeno entschieden. Als Ausgabe steht hier, anders als bei der Mamdani-Implikation, keine unscharfe Menge sondern ein konkreter Wert. Dieser beschreibt den geschätzten Wert $[0, 100]$ des mentalen Workloads für einen bestimmten - z.B. visuellen - Bereich.

Die Sugeno Methode eignet sich besonders, da eine Defuzzifizierung nicht notwendig ist und somit eine hohe Rechengeschwindigkeit erlangt werden kann. Dies ist für eine Berechnung nahe des realen Messzeitpunktes wichtig.

Als Ausgabevariablen ergeben sich die in der Tabelle 6.1 gezeigten fünf zuvor definierte Stadien der mentalen Beanspruchung. Da dieser besonders sensitiv gegenüber gesteigertem Workload sein soll, ist eine differenzierte Aufteilung im Bereich über 50 vorgenommen worden.

Ausgangsvariable	dazugehöriger Wert
sehr viel	100
viel	90
etwas mehr	80
mittel	50
wenig	20

Tabelle 6.1: Ausgangsvariablen

6.6.3 Regeln

Aus den Eingangsvektoren ergeben sich nun die in Tabelle 6.2 gezeigten neun Regel für die Inferenzbildung. In der nachfolgenden Tabelle wird die linguistische Variable des normierten Pupillendurchmessers als „Pupille“ und die des normierten Pupillenanstiegs als „Anstieg“ bezeichnet. Als Verknüpfung der beiden Mengen wird die Schnittmenge der Eingänge verwandt, was dem UND-Elementaroperation entspricht. [NOLL, 2002]

Die Fuzzy-Regeln spiegeln hier das Expertenwissen im Bezug auf die, in Kapitel 2 vorstellte, psychische Beanspruchung wider.

WENN	Pupille	Gering	UND	Anstieg	Gering	DANN	Workload	Wenig
WENN	Pupille	Gering	UND	Anstieg	Mittel	DANN	Workload	Mittel
WENN	Pupille	Gering	UND	Anstieg	Gross	DANN	Workload	etwas mehr
WENN	Pupille	Mittel	UND	Anstieg	Gering	DANN	Workload	Mittel
WENN	Pupille	Mittel	UND	Anstieg	Mittel	DANN	Workload	etwas mehr
WENN	Pupille	Mittel	UND	Anstieg	Gross	DANN	Workload	viel
WENN	Pupille	Gross	UND	Anstieg	Gering	DANN	Workload	etwas mehr
WENN	Pupille	Gross	UND	Anstieg	Mittel	DANN	Workload	viel
WENN	Pupille	Gross	UND	Anstieg	Gross	DANN	Workload	sehr viel

Tabelle 6.2: Fuzzyregeln

Kapitel 7

Vergleich Erwartungswerte mit Fuzzy-Klassifizierung

Um einen Vergleich der Fuzzy-Klassifizierung anstellen zu können, benötigt man zunächst einen Referenzwert des aktuell zu vermutenden Workloads. Dieser Referenzwert muss wie die Klassifizierung für jeden Zeitpunkt t der betrachteten Messreihe vorliegen. Der Vergleich wird in zwei Gruppen aufgeteilt: Tests innerhalb einer Spanne ansteigenden Workloads und Test mit einfachem, peakartigem Anstieg des Workloads. Zusammenfassen lässt sich dies, da es im Falle des Digit Span-, Digit Span Backwards- und Sternberg Tests zu einer längeren Merkphase für eine Zahlenreihe kommt. Beim Oddball Test ist eine periodische, peakartige Erhöhung des Workloads zu erwarten, jedoch kein kontinuierlicher Anstieg über mehrere Sekunden. Zunächst soll die Variante mit einem Anstieg über eine festgelegte Spanne betrachtet werden.

7.1 Digit Span-, Digit Span Backwards- und Sternberg Test

Wie schon Kahnemann und Beatty (vgl. [BEATTY, 1982]) in frühen Studien festgestellt haben, hängt der aufzuwendende Workload zur Bearbeitung einer speziellen Aufgabe mit der Schwierigkeit dieser zusammen. Bei Tests, welche die Kurzzeitmerkfähigkeit

von Zahlenreihen beinhalten, kann von einem linearen Anstieg des Ressourceneinsatzes mit zunehmender Länge der Zahlenreihe ausgegangen werden.

Aufgrund dieses linearen Zusammenhangs der Länge der präsentierten Stimuli und der Schwierigkeit habe ich eine Gerade berechnet. Diese verbindet den Workload zu Beginn jedes Trials mit der Merkphase direkt im Anschluss an die Stimuluspräsentation. Die zeitliche Länge der Stimuli ist direkt proportional der Anzahl der gezeigten Zahlen, da jede Zahl, genau wie in Abschnitt 5 beschrieben, eine Sekunde lang angezeigt worden ist. Da die Latenzzeit der menschlicher Pupillenreaktion mindestens 100ms beträgt, wird für diese Zeit die Gerade weitergeführt. Dies entspricht genau der Hälfte der Merkphase. (vgl. [MANZEY, 1998])

Bei Beginn der Stimulipräsentation habe ich für den Workload einen Wert von 50 angenommen, da ab dem ersten Moment des Erscheinens der Zahlenfolge eine erhöhte Aufmerksamkeit gegenüber der Pause notwendig ist. Anschließend steigt die in Abbildung 7.1 gezeigte rote Gerade bis zu einem Maximum von 85 an. Das Maximum liegt genau bei 100 Samples nach der zuletzt gezeigten Stimuluszahl. Der anschließende Verlauf der Kurve (hier in blau dargestellt) ist lediglich eine quantitative Annahme und ist durch keinerlei Literatur belegt. Sie dient lediglich der besseren Unterscheidung der verschiedenen Abschnitte eines Trials.

Bestimmung des DiffSample- Wertes

Der DiffSample-Wert gilt als Maß für die Abweichung meines klassifizierten Workloads vom vorher definierten Erwartungswert.

Für diese Spannen des sicheren Erwartungswertes, habe ich nach dem Anwenden meines Fuzzy-Operators die absolute Differenz zwischen Ausgabewert des Fuzzysystems und Erwartungswert gebildet (siehe Formel 7.1). Diese Differenzen wurden weiterhin aufsummiert und anschließend durch die Anzahl der Samples während der Stimulusdarbietung dividiert. Daraus resultiert ein Faktor für jeden Probanden, der angibt um welches Maß der mit dem Fuzzy-System klassifizierte Wert von dem des Erwartungswertes abweicht. Dieser Wert wird weiterhin als DiffSample bezeichnet.

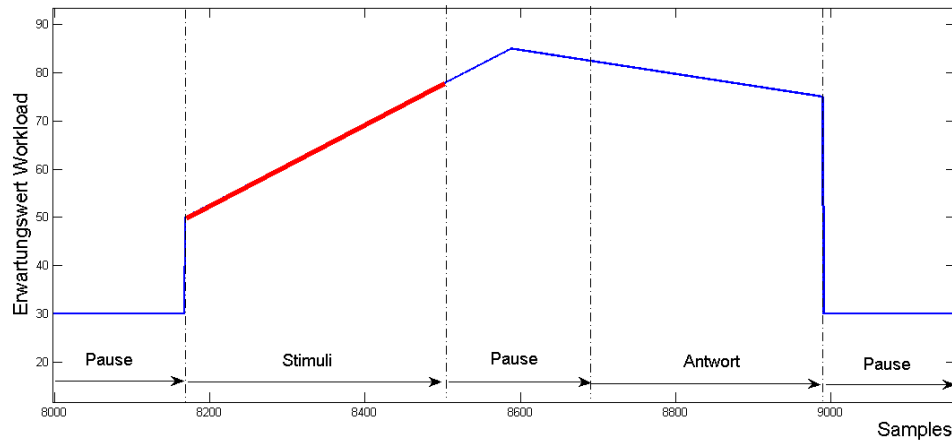


Abbildung 7.1: Erwartungswert für Stimulispanne

Im rot markierten Linienabschnitt ist mit einem linearen Anstieg der mentalen Beanspruchung zu rechnen, da in dieser Zeit eine immer größer werdende Reihe an Stimuli angezeigt werden. Mit steigender Anzahl dieser wird der Proband kognitiv höher belastet.

$$DiffSample = \frac{\sum_{i=0}^N |FuzzyOutput(i) - Erwartungswert(i)|}{N} \quad (7.1)$$

N = Anzahl der Samples während Stimulidarbietung

i = Sample während Stimulidarbietung

Die jeweiligen DiffSample Werte für die aufgenommenen Tests erstrecken sich, wie in Tabelle 7.1 zu sehen, von 9,2 bis 19,7. Der Mittelwert aller DiffSamples der Probandenversuche beträgt 14,9.

Die jeweilige Ausprägung von DiffSample variiert je Proband und Test. Es ist deshalb davon auszugehen, dass die Art des Tests eine Rolle für die Auswertung spielt. Die Ergebnisse der für jeden Probanden, sortiert nach Art des Tests ist in den Abbildungen 7.7, 7.8 und 7.9 zu sehen.

Eine gute Schätzung des Workloads ist in Abbildung 7.2 zu sehen. Während der Pausen

Test	DiffSample	DiffSample	DiffSample	Anzahl Probanden
	Min	Max	Mittelwert	
Digit Span	9,2	19,7	15,0	25
Digit Span Backwards	11,8	17,0	14,4	25
Sternberg	11,0	18,2	15,2	28

Tabelle 7.1: Vergleich des DiffSample Indexes bei verschiedenen Tests

vor dem Zeigen der zu merkenden Zahlenfolge ist der geschätzte Workload niedriger und steigt während der Zeit des angezeigten Stimuli an. Nach dem Zeitpunkt, zu dem die Antwort gegeben wurde, sinkt der errechnete Wert wieder und steigt erst mit einer erneuten Präsentation des Stimulus.

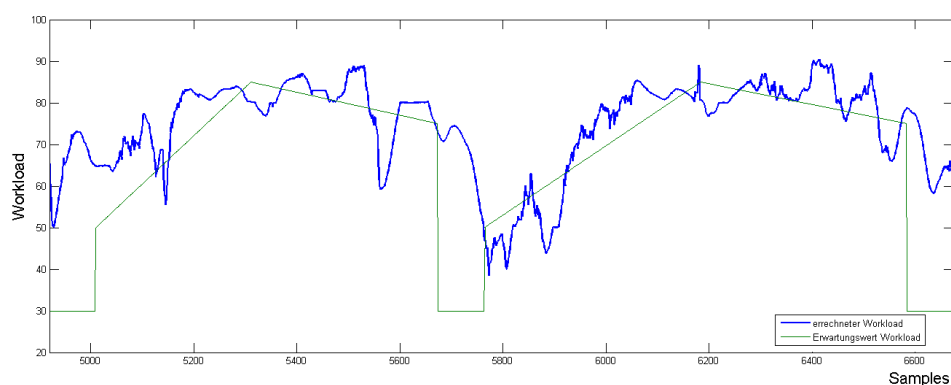


Abbildung 7.2: Positives Beispiel des errechneten Workloads im Vergleich zum Erwartungswert

Ein abrupter Abfall des Pupillendurchmessers nach dem Ausblenden der letzten Zahl in der Zahlenreihe, kann in einigen Fällen beobachtet werden. Dies lässt den geschätzten Workload während dieser Zeit stark sinken. Dieses Phänomen kann mit einer Erleichterung in Zusammenhang gebracht werden, das die Probanden verspüren, sobald sie bemerken, dass kein weiterer Parameter mehr zur Zahlenreihe hinzukommt. Da die Längen der Zahlenfolgen nach keinem vorher definierten Muster angeordnet war, hatte

der Proband darüber keine Kenntnis.

In Abbildung 7.3 ist die mit dem DiffSample Wert am schlechtesten bewertete Klassifizierung des Workloads während zwei Trials des Sternberg Tests zu sehen. Der DiffSample Wert lag für diesen Probanden bei diesem Experiment bei 18,2, welcher gleichzeitig den größtmöglichen Index aller 3 Tests darstellt. Hierbei ist leicht zu erkennen, dass es während der Pausenzeiten zu einem nur geringen Abfall der Schätzung kommt. Nach dem Zeitpunkt des Ausblendens der letzten Zahl in der Reihe kommt es jedoch zu einem sehr starken Abfall. Dies könnte probandenspezifisch sein, da dieses Phänomen, wie bereits erwähnt, jedoch öfter auftritt, ist von einer anderen Ursache auszugehen.

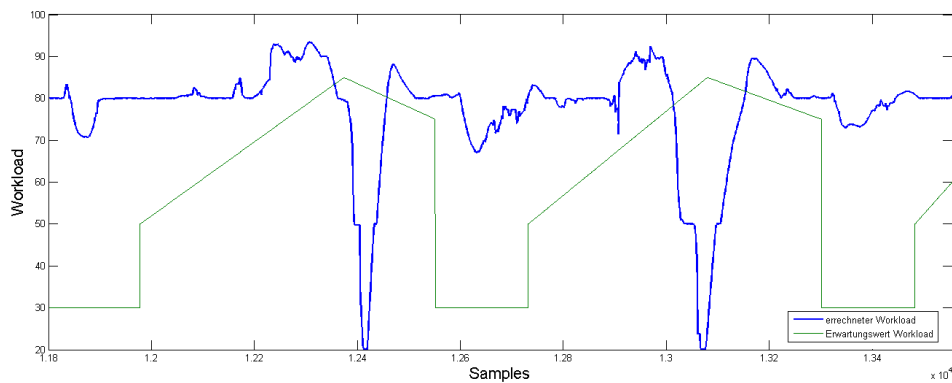


Abbildung 7.3: Negatives Beispiel des errechneten Workloads im Vergleich zum Erwartungswert

Ein Vergleich von fünf zufällig ausgewählten Probanden ist in Grafik 7.4 zu sehen. Es handelt sich hier bei allen Probanden um das selbe Trial des Digit Span Backwards Test. Der qualitative Verlauf des errechneten Workloads stellt sich zunächst an der Stelle *Samples* = 0 als ein fallender Verlauf dar. Dies rührt von der vorherigen Aktivität des Probanden. Da die Trials kontinuierlich hintereinander angeordnet sind, hat der Proband zum Zeitpunkt *Samples* = 0 soeben erst eine Antwort auf das Trial zuvor gegeben. Die Anspannung, die er dabei ausgesetzt war, sinkt nun dementsprechend. Ein Steigen der Kurve ist ab circa 90 Samples zu erkennen. Da es sich um das 5. Trial

während des Testablaufs handelt, hat der Proband bereits gelernt, dass die nächste Stimulifolge kurz bevor steht. Somit erhöht sich die Aufmerksamkeit aller Probanden bereits kurz vor dem Beginn der Stimulipräsentation. Qualitativ ist zu erkennen, dass die Anstiege des Workloads für alle betrachteten Teilnehmer während der Stimuliphase positiv sind. Lediglich bei Proband 3 und 5 kommt es zu Anfang der Stimuli kurz zu einem Abfall, welcher sich bereits nach etwa 25 Samples (etwas mehr als 4ms) wieder normalisiert hat.

Nach Beenden der auditiv dargebotenen Stimuliphase kommt es bei allen Teilnehmern zu einem Abfall mit unterschiedlich starker Ausprägung.

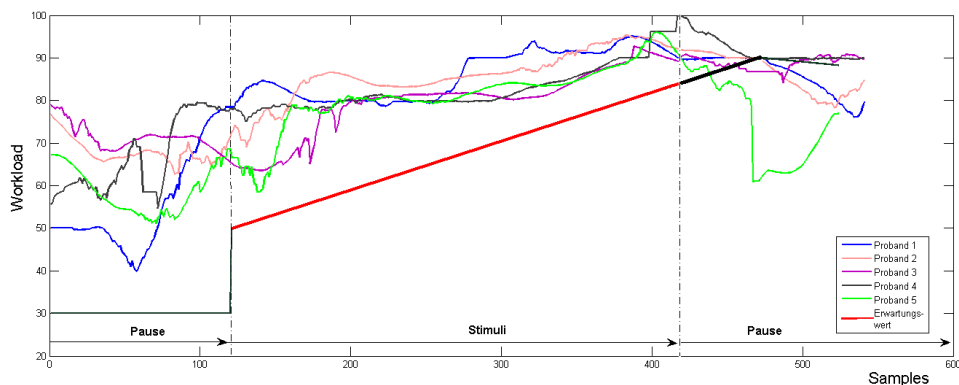


Abbildung 7.4: Vergleich des errechneten Workloads zum Erwartungswert für 5 Probanden

In der obigen Abbildung sind die errechneten Workloads pro Sample im Vergleich zum Erwartungswertes zu sehen. Ein Anstieg des Workloads zum Zeitpunkt der ersten Stimulipräsentation ist bei allen Probanden zu erkennen. Lediglich bei Probanden 2 und 5 ist ein kurzzeitiger Abfall der Kurve zu erkennen. Nach circa 25 Samples (etwas mehr als 4ms) erhöht sich dieser Wert aber stetig. Nach dem Beenden der Stimulipräsentation bei etwa 490 Samples, ist für alle Probanden ein Abfall mit verschiedenartigem Ausmaß zu erkennen.

In Abbildung 7.6 ist ein Vergleich von 5 zufällig ausgewählten Probanden während eines Trials des Sternberg Tests zu sehen. Für Probanden 1,3 und 4 ist ein Anstieg des Workloads über die Zeit der Stimulipräsentation deutlich zu erkennen. Proband 5

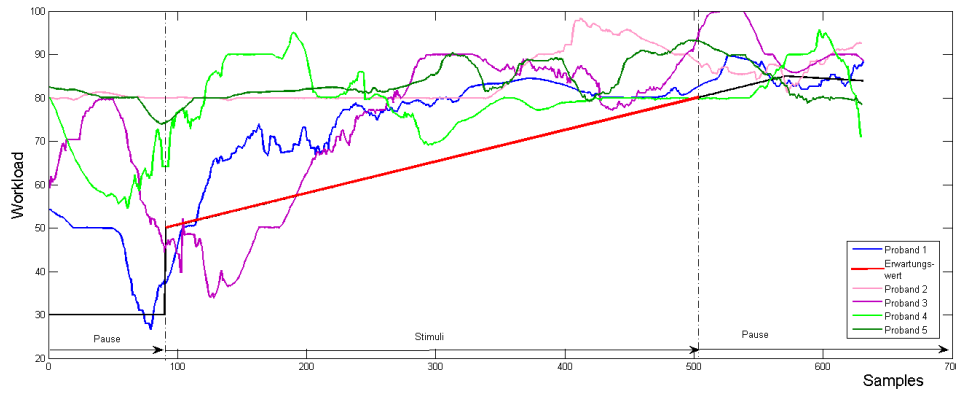


Abbildung 7.5: Vergleich des errechneten Workloads zum Erwartungswert für 5 Probanden während des Digit Span Tasks

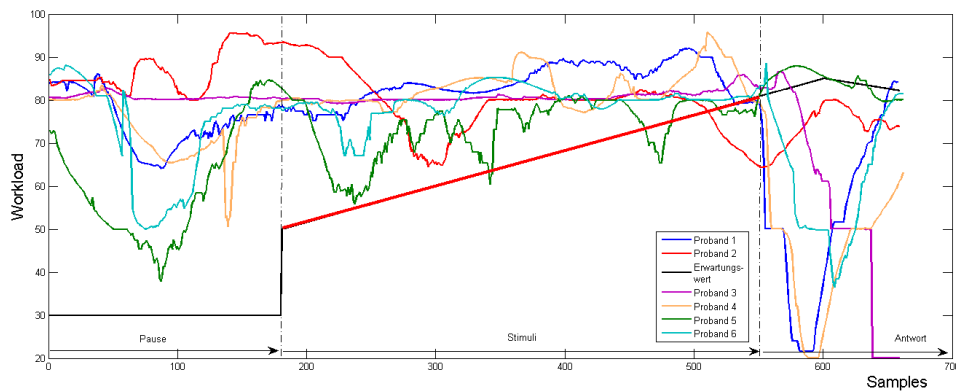


Abbildung 7.6: Vergleich des errechneten Workloads zum Erwartungswert für 6 Probanden während des Sternberg Tests

weist erst nach 300 Samples einen Anstieg auf. Zu Beginn des Trials ist der Anstieg nur sehr gering, was darauf zurückzuführen ist, dass sich der berechnete Workload während der Pause nicht stark verringert hat. Im Verlauf von Proband 2 ist zunächst keine Steigung des Workloads zu erkennen. Ab Sample 350 ist ein starker Anstieg zu verzeichnen.

Dieses Verhalten ist zunächst nicht leicht zu Interpretieren. Einerseits ist eine schlechte Klassifizierung des Workloads an dieser Stelle möglich. Zum Anderen kam es bei den Probanden 2 und 5 während der vorangehenden Pausenphase zu keiner errechneten Reduzierung des Workloads. Daraus ergibt sich das in Kapitel 7.3 beschriebene Problem. Dieses benennt die Ungewissheit des wirklich vorherrschenden kognitiven Workloads. Es ist möglich, dass die Probanden während dieser Zeit der Pause mental noch mit der zuvor gegebenen Antwort beschäftigt waren oder diese versucht haben zu verifizieren. Ein Rückgang der Beanspruchung wäre so nicht zu erwarten und die errechnete, erhöhte Beanspruchung würde an dieser Stelle mit dem realen Wert übereinstimmen. Da diese reale Beanspruchung des Probanden nicht ohne EEG messbar sind, kann hier nicht über eine Fehldetektion des Workloads entschieden werden.

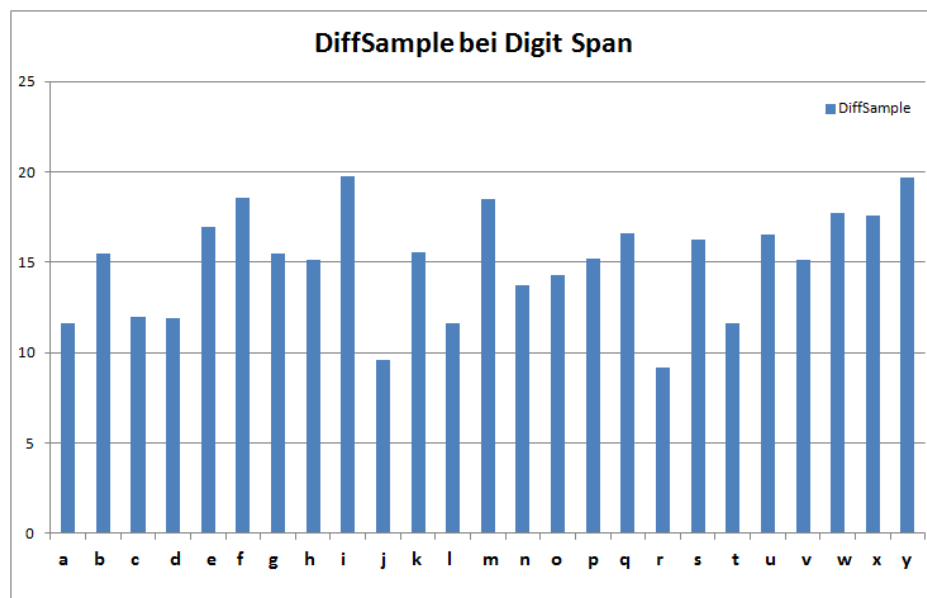


Abbildung 7.7: DiffSample aller Probanden während Digit span Test
Auf der X-Achse sind die jeweiligen Abkürzungen der Probandennamen zu sehen, über die der in 7.1 errechneten DiffSample Wert aufgetragen ist.

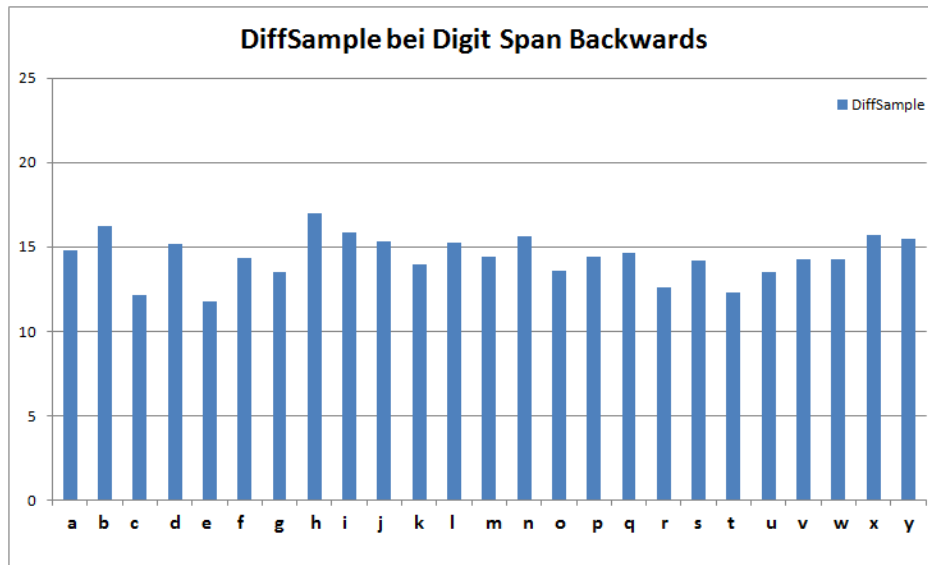


Abbildung 7.8: DiffSample aller Probanden während Digit Span Backward Test

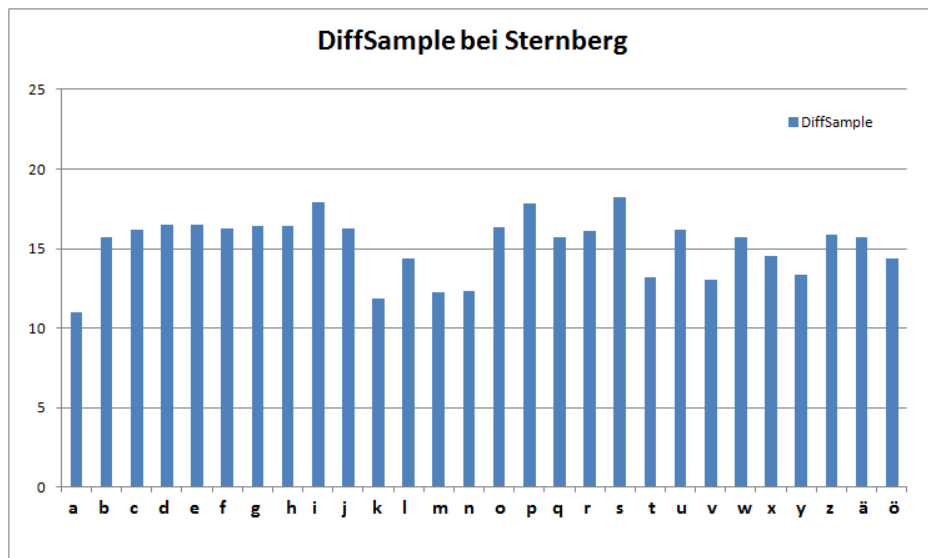


Abbildung 7.9: DiffSample aller Probanden während Sternberg Test

7.2 Oddball Test

Beim Oddball Test ist während des gesamten Tests eine erhöhte mentale Beanspruchung zu erwarten, da es in einem kontinuierlichen Ablauf zu einer schnellen Reizdarbietung kommt. Die Pausen zwischen den einzelnen Stimulitönen betragen lediglich eine Sekunde, was nicht bei den meisten Probanden nur zu einer kleinen Änderung der Pupille führte. Dies spiegelt auch das Ergebnis der Klassifikation wider. Abfälle des Workloads sind zwar fast ausschließlich in den Pausen zu beobachten, jedoch sind sie oft nicht stark ausgeprägt. Eine signifikante Steigerung während der Darbietung des seltenen Zielreizes ergibt sich nur in seltenen Fällen, da für jede Stimulidarbietung vom Probanden entschieden werden muss, ob es sich um einen Standard- oder Zielreiz handelt oder nicht. Die Wahl dieses Experimentes war folglich nicht optimal für eine Klassifizierung mittels Eyetracking. Die Phasen von erhöhtem und niedrigerer Beanspruchung sollten klarer getrennt werden können.

Test	DiffSample Min	DiffSample Max	DiffSample Mittelwert	Anzahl Probanden
Oddball	13,2	19,9	17,4	26

Tabelle 7.2: DiffSample Indexes bei Oddball Test

Die DiffSample Werte aller Probanden sind auch hier kleiner als 20. Dieser Wert wurde für alle Zeitpunkte der Stimulidarbietung berechnet.

Als Erwartungswert wurde für die Zeitspanne der Standardstimuli der Wert 80 angenommen, für die Zeitspanne der Zielreize der Wert von 90.

In Abbildung 7.10 ist gut zu erkennen, dass der Proband den Testablauf schon gut verinnerlicht hat. Bereits kurz vor der Präsentation des nächsten potentiellen Zielreizes wird ein steigender Workload erkannt.

Eine Detektion der Zielreize anhand es errechneten Workloads ist aufgrund einer hohen false-positive Rate nicht möglich, es ist jedoch eine periodische Erhöhung der Beanspruchung währen der Stimulipräsentation, wie in Abbildung 7.10 zu sehen ist, zu

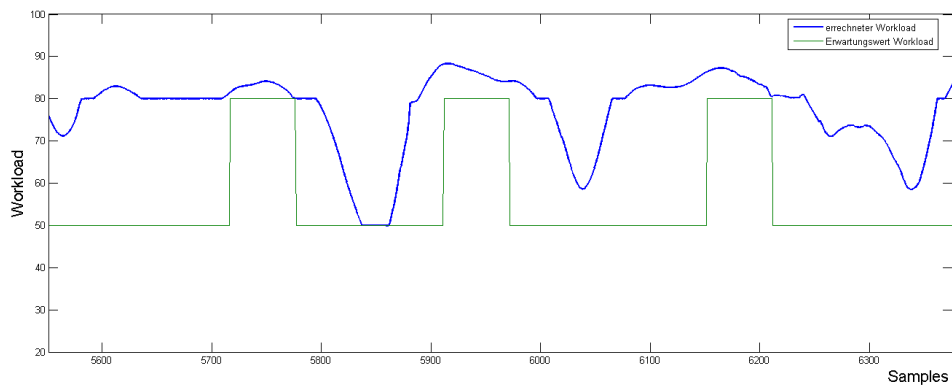


Abbildung 7.10: Beispiel des errechneten Workloads im Vergleich zum Erwartungswert

erkennen. Dieser Test ist für die eine differenziertere Unterscheidung Beanspruchungsphasen folglich nicht sensitiv genug.

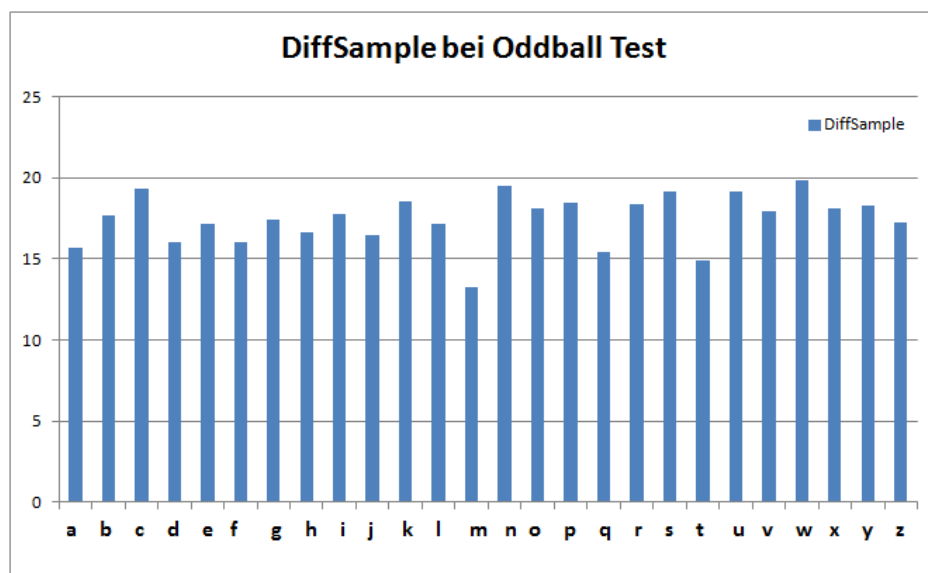


Abbildung 7.11: DiffSample aller Probanden während Oddball Test

7.3 Probleme der Auswertung

Wird der geschätzte, mentale Workload bereits zu Beginn des neuen Trials als hoch eingeschätzt, obwohl der Erwartungswert zunächst einen geringeren Wert vorgibt, muss es sich hierbei nicht um eine fehlerhafte Klassifikation handeln. Es besteht die Möglichkeit, dass der Proband sich mental noch mit der zuvor beantworteten Aufgabe beschäftigt, sein Ergebnis dabei anzweifelt oder verifiziert.

Dies würde auch nach Beantwortung der vorigen Frage einen mentalen Workload mit sich ziehen. Somit kann eine falsch positiv Identifikation des ermittelten Workloads nicht per se anhand des Erwartungswerts angenommen werden. Eine falsch positiv Einschätzung ist im Bezug auf den mentalen Workload diejenige, die eine hohe Beanspruchung annimmt obwohl diese nicht stattgefunden hat.

Nur der Vergleich mit einem EEG würde eine differenziertere Aussage zu dem aktuellen Beanspruchungsstatus zulassen. Eine solche Vergleichsstudie überschreitet an dieser Stelle aber den Umfang dieser Arbeit.

Während der Stimulipräsentation kann hingegen immer von einer, wenigstens leicht, erhöhten Beanspruchung ausgegangen werden, sobald der Proband einen Teil der gezeigten Stimuli im Anschluss wiedergeben kann. Ohne gänzliche Beanspruchung während dieser Spanne wäre ihm dies nicht möglich.

7.4 Optimierungsmöglichkeiten

Zum einen besteht die Möglichkeit die Kurve des Erwartungswertes noch besser auf die Schwierigkeit der gezeigten Stimuli anzupassen. Eine Variation des Maximalwertes abhängig von der Schwierigkeit des Stimulus wäre hier vorstellbar. Somit ließe sich die Überprüfung der Fuzzyresultate noch differenzierter betrachten.

Auch eine Modifizierung der in Abschnitt 6.6.3 definierten Regeln könnte zu einem noch aussagekräftigerem Ergebnis führen. Eventuell könnten die linguistischen Variablen in noch mehr Terme unterteilt werden. Demzufolge würden sich mehr Regeln ergeben, welche aber auch Gefahren der Überregulierung mit sich bringen würden.

Als Möglichkeiten zur Verbesserung sehe ich die weitere Differenzierung von Stadien der geringen mentalen Beanspruchung, sowie das Problem des Einbruchs der geschätzten Beanspruchung nach der Präsentation des letzten Stimuli. Aus den mir vorliegenden Daten konnte ich zu diesem Phänomen keine andere Einschätzung des Workloads ziehen. Zur genaueren Klärung dieses Phänomens bedarf es weiterer psychologischer Einschätzung und weiterführender Studien.

Eine Kompensation der beleuchtungsbedingten Pupillendilatation oder eine Erkennung dieser wäre für eine laborferne Umgebung wünschenswert. In diesem Punkt sehe ich noch weiteres Potential für die Forschung.

Zudem ist die Umsetzung der Datenverarbeitung in Matlab noch sehr langsam. Der Algorithmus könnte dahingehend noch weiter verbessert werden.

Kapitel 8

Zusammenfassung und Ausblick

In meiner Arbeit habe ich eine Fuzzy-Logik basierte Möglichkeit geschaffen, ereignis-korrelierte Potentiale mittels Auswertung von Pupillendaten zu erfassen.

Zur Datenaquise habe ich zunächst vier verschiedene, in der Wissenschaft etablierte Experimente ausmacht und mit der Software „ExperimentCenter“ erstellt. Die Datenaufnahme wurde mit mindestens 25 Probanden durchgeführt.

Nach der Datenverarbeitung und der Implementierung dieser in die Fuzzy-Logic-Toolbox ergab sich eine Approximation des Workloads in einem bestimmten Bereich der Informationsverarbeitung des Probanden. Diese Approximation wurde mit einem in der Wissenschaft beschriebenen Erwartungswert verglichen.

Als Ergebnis dieses Vergleichs lässt sich feststellen, dass der Trend der ermittelten Beanspruchung mit dem Angenommenen einhergeht. Mittels linguistischer Variablen lässt sich eine Beschreibung der verschiedenen Eingangsgrößen finden, welche anschließend zur Klassifikation genutzt werden.

Die Güte der Approximation wird mit dem *DiffSample – Wert* beschrieben (siehe 7) und gibt eine Durchschnittliche Abweichung der Approximation an. Dieser Wert ist noch verbesserungswürdig, lässt gleichzeitig aber eine Beurteilung über den ungefähren Zustand einer mentalen Beanspruchung zu. Es kann deutlich Unterschieden werden, ob sich ein Proband in einer Phase geringer oder erhöhter Beanspruchung befindet.

Wie in Abschnitt 7.4 beschrieben sehe ich noch weiterführende Möglichkeiten zur Optimierung dieses Verfahrens. Diese sind unter Anderem die bessere Differenzierung von

Phasen mit geringem Workload und die Kompensation von beleuchtungsbedingter Pupillendilatation.

Anhang A

Ergänzende Unterlagen

A.1 Probandendaten

Im folgenden sind alle Probanden, die sich zu einem Test bereit erklärt haben und diesen anschließend durchgeführt haben, aufgelistet. Des Weiteren sind Informationen zu Geschlecht, Händigkeit, Alter zum Zeitpunkt des Tests und Art der erfolgreich teilgenommenen Tests, der abgehangenen Tabelle zu entnehmen.

Name	m	w	Händig	Alter	Digit Span	Digit Span Back	Sternberg	Oddball
a	x	O	Rechts	31	x	x	x	x
b	x	O	Rechts	32	x	x	x	x
c	O	x	Rechts	31	x	x	x	x
d	O	x	Rechts	29	x	x	x	x
e	x	O	Rechts	34	x	O	x	x
f	x	O	Rechts	27	x	x	x	O
g	x	O	Rechts	28	x	x	x	O
h	x	O	Rechts	29	x	x	x	x
i	x	O	Rechts	34	x	O	x	x
j	x	O	Rechts	35	x	x	x	x
k	x	O	Rechts	23	x	x	x	x
l	x	O	Rechts	34	x	x	x	x
m	O	x	Rechts	23	x	x	x	x
n	O	x	Rechts	35	x	x	x	x
o	x	O	Rechts	31	O	O	x	x
p	x	O	Links	18	x	x	x	x
q	x	O	Rechts	24	x	x	x	x
r	x	O	Rechts	29	x	x	x	x
s	x	O	Rechts	24	x	x	x	x
t	x	O	Links	25	x	x	O	x
u	x	O	Rechts	26	O	x	x	x
v	x	O	Rechts	35	x	x	x	x
w	O	x	Rechts	25	O	x	x	O
x	x	O	Links	25	x	x	x	x
y	x	O	Rechts	35	x	x	x	x
z	x	O	Rechts	30	x	O	x	x
ä	x	O	Rechts	27	x	x	x	x
ö	x	O	Rechts	34	O	x	x	x
ü	x	O	Rechts	35	x	x	x	x

x - zutreffend

O - nicht zutreffend

A.2 Software

Experiment Suite

Für die Erstellung meiner durchgeführten Experimente habe ich die Softwareversion „Experiment Suite - 3.5.200 “ der Firma Sensomotoric Instruments GmbH genutzt. Diese beinhaltet die bereits erwähnten Softwarekomponenten „ExperimentCenter“ und „BeGaze “ Diese Software kann allerdings nur mit einem dazugehörigen Dongle der Firma Sensomotoric Instruments verwendet werden.

Der Client, welcher die Schnittstelle zwischen Eyetracking-Server und Experiment-Center bildet, habe ich in der Version 3.2.156 verwendet. Dieser kann auf der Homepage <http://www.smivision.com/en/gaze-and-eye-tracking-systems/support/software-download.html> heruntergeladen werden. Diese Software ist nur in Verbindung mit einem Remote Eyetracking System der Firma Sensomotoric Instruments GmbH zu verwenden.

Matlab

In dieser Arbeit wurde die von The Mathworks™ angebotene Matlab® Version 7.7.0.471 (R2008b) für die Auswertung der Augendaten verwendet.

Die Ausgabe der errechneten Daten erfolgt in der Workspace des Tools. Eine Übersicht der errechneten DiffSample-Werte im Vergleich wird je Kompilierung im jeweiligen Benutzerordner des PCs erstellt.

Übersicht zur Strukturierung der Matlab-Files

- **startWork.m** : Diese Datei enthält die Aufrufe aller weiteren Matlabdateien, die zur Verarbeitung der Eingangsdaten, zur Klassifizierung, zur Speicherung und zum plotten benötigt werden. Diese sind: readAllInput, readSample, applyFuzzy2,
 - **startWorkWithEyeData.m**: Diese Datei enthält ebenfalls die zuvor genannten Aufrufe der folgenden Skripte mit Ausnahme der des readAllInput- Skriptes,
-

welches beim Einladen der bereits vorhandenen Workspaces übersprungen werden kann.

- **readAllInput.m:** Liest alle zuvor definierten Daten ein
 - **readSample.m:** erzeugt zwei voneinander getrennte Pfade in der Workspace: eyeData und estimations. Der Pfad eyeData enthält alle Eingangsdaten, bearbeiteten Daten und Zwischenergebnisse. Die Daten im Pfad estimations enthalten Informationen zu den gezeigten Stimuli.
 - **applyFuzzy2.m:** Dieser Algorithmus beinhaltet die Berechnung der Erwartungswerte, den Aufruf der Fuzzy-Logic-Toolbox und die Ausgabe des errechneten Workloads in den estimation-Pfad der Workspace
 - **plotTrial.m:** Plottet den Erwartungswert und den errechneten Workload für das im Aufruf definierte Trial
-

Abbildungsverzeichnis

5.1	Vier Stufen der Informationsverarbeitung im Sternberg Task	16
5.2	Pupillendurchmesser in Millimeter in Abhängigkeit der Zeit und Länge der zu merkenden Zahlenreihe während des Digit span Tests.	18
5.3	Oddball Test und EEG Ableitungen während Standard- (S) und Ziel- reize (Oddball).	20
5.4	Aufbau und Positionierung des Testplatzes	24
5.5	Möglichkeiten der Experimentenanordnung	25
5.6	Ablauf Digit Span	27
6.1	Beispiel für Augen- und Eventdaten in Textdatei	30
6.2	Unbereiniger Pupillendurchmesser über die Zeit in Sekunden	32
6.3	Bereiniger Pupillendurchmesser in Millimeter über die Zeit in Sekunden	32
6.4	Detektion von Ausreißern und invaliden Werten	34
6.5	Beispiel für interpolierte Messwerte	36
6.6	Übersicht über Fuzzy-System	38
6.7	Beispiel für mittleren Pupillenwert in An- und Abstiegsphase	39
6.8	Zugehörigkeitsfunktion von normiertem Pupillendurchmesser	40
6.9	Zugehörigkeitsfunktion von normiertem Anstieg des Pupillendurchmessers	42
7.1	Erwartungswert für Stimuluspanne	49
7.2	Positives Beispiel des errechneten Workloads im Vergleich zum Erwar- tungswert	50
7.3	Negatives Beispiel des errechneten Workloads im Vergleich zum Erwar- tungswert	51

7.4	Vergleich des errechneten Workloads zum Erwartungswert für 5 Probanden	52
7.5	Vergleich des errechneten Workloads zum Erwartungswert für 5 Probanden während des Digit Span Tasks	53
7.6	Vergleich des errechneten Workloads zum Erwartungswert für 6 Probanden während des Sternberg Tests	53
7.7	DiffSample aller Probanden während Digit span Test	54
7.8	DiffSample aller Probanden während Digit Span Backward Test	55
7.9	DiffSample aller Probanden während Sternberg Test	55
7.10	Beispiel des errechneten Workloads im Vergleich zum Erwartungswert	57
7.11	DiffSample aller Probanden während Oddball Test	57

Literaturverzeichnis

- [DIN, 1991] (1991). *ISO 10075, Ergonomic principles related to mental work-load - General terms and definitions.*
- [ANDREASSI, 2000] ANDREASSI, J. L. (2000). *Psychophysiology - Human behavior and physiological response.* Lawrence Erlbaum Associates.
- [BEATTY, 1982] BEATTY, J. (1982). *Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources.* Psychological Bulletin, 91(2):276–292.
- [BLACKWOOD et al., 1987] BLACKWOOD, D.H., L. J. WHALLEY, J. E. CHRISTIE, I. M. BLACKBURN, D. M. CLAIR und A. MCINNES (1987). *Changes in auditory P3 event-related potential in schizophrenia and depression.* The British Journal of Psychiatry, 150:154–160.
- [BUCHNER, 2005] BUCHNER, H. (2005). *Evozierte Potentiale, neurovegetative Diagnostik, Okulographie.* Thieme.
- [BUCHNER, 2014] BUCHNER, H. (2014). *Praxisbuch evozierte Potentiale.* Thieme.
- [CORBIN und MARQUER, 2009] CORBIN, L. und J. MARQUER (2009). *Individual differences in Sternberg's memory scanning task.* Acta Psychologica, (131):153–162.
- [GREHL et al., 2000] GREHL, H., F. REINHARDT und J. NEUBERGER (2000). *Checkliste Neurologie.* Thieme.
-

- [HOLMQVIST et al., 2011] HOLMQVIST, K., N. M., R. ANDERSSON, R. DEWHURST, H. JARODZKA und J. V. D. WEIJER (2011). *Eye Tracking - A comprehensive guide to methods and measures*. Oxford University Press.
- [KESSLER, 2007] KESSLER, W. (2007). *Multivariate Datenanalyse: fuer die Pharma, Bio- und Prozessanalytik*. John Wiley and Sons.
- [KLINGER et al., 2008] KLINGER, J., R. KUMAR und P. HANRAHAN (2008). *Measuring the Task-Evoked Pupillary Response with a Remote Eye Tracker*. Association for Computing Machinery Inc.
- [MANZEY, 1998] MANZEY, D. (1998). *Enzyklopädie der Psychologie, Themenbereich C, Serie I, Band 5, Ergebnisse und Anwendungen der PSycho-physiologie*. Hogrefe Verlag für Psychologie.
- [MAURER et al., 2005] MAURER, K., N. LANG und J. ECKERT (2005). *Praxis der Evozierten Potentiale: Sep, Aep, Mep, Vep*. Steinkopff Verlag.
- [SCHLICK et al., 2010] SCHLICK, CH., R. BRUDER und H. LUCZAK (2010). *Arbeitswissenschaft*. Springer.
- [SCHMIDT und LANG, 2013] SCHMIDT, R. F. AND THEWS, G. und F. LANG (2013). *Physiologie des Menschen*. Springer.
- [SCHMIDT et al., 2007] SCHMIDT, R. F., F. LANG und M. HECKMANN (2007). *Physiologie des Menschen*. Springer.
- [STERNBERG, 1966] STERNBERG, S. (1966). *High-Speed Scanning in Human Memory*. Science, 153(3736):652–654.
- [SWERDLOW, 2010] SWERDLOW, N. R. (2010). *Behavioral Neurobiology of Schizophrenia and Its Treatments*. Springer.
- [TRIMMEL, 1990] TRIMMEL, M. (1990). *Angewandte und Experimentelle Neuropsychiologie*. Springer.
-

[ULFIG, 2011] ULFIG, NORBERT (2011). *Kurzlehrbuch Histologie*. Georg Thieme Verlag.

[WICKENS, 2002] WICKENS, C. D. (2002). *Multiple resources and performance prediction*. Taylor and Francis Ltd.

[NOLL, 2002] NOLL, P. (2009). *Statistisches Matching mit Fuzzy Logic* . Springer