

Musikkomposition mittels P300-basiertem BCI und einem Step-Sequencer

Bachelorarbeit

Der Technischen Fakultät der Universität Bielefeld

vorgelegt am 26. März 2013

von Benjamin Wolff

Gutachter:

Dipl.-Inform. Andrea Finke

M.Sc. Hannes Riechmann

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Bachelorarbeit selbständig verfasst und gelieferte Datensätze, Zeichnungen, Skizzen und graphische Darstellungen selbständig erstellt habe. Ich habe keine anderen Quellen als die angegebenen benutzt und habe die Stellen der Arbeit, die anderen Werken entnommen sind - einschl. verwendeter Tabellen und Abbildungen - in jedem einzelnen Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht.

Benjamin Wolff,
Bielefeld, 26. März 2013

Inhaltsverzeichnis

1	Einleitung	4
2	Theoretische Grundlagen	5
2.1	Brain-Computer Interfaces	5
2.1.1	Komponenten eines Brain-Computer Interfaces	6
2.1.2	Messverfahren	6
2.1.2.1	Nichtinvasive Verfahren	6
2.1.2.1.1	Electroenzephalografie (EEG)	7
2.1.2.1.2	Magnetenzephalografie (MEG)	9
2.1.2.1.3	Funktionelle Magnetresonanztomographie (FM-RI)	10
2.1.2.1.4	Nahinfrarotspektroskopie (NIRS)	10
2.1.2.2	Invasive Verfahren	10
2.1.2.3	Ereigniskorrelierte Potentiale (ERP)	11
2.1.2.4	Die P300-Komponente	11
2.1.3	Allgemeine Verarbeitungsmethoden und Algorithmen	15
2.1.3.1	Filter auf Fourier-Basis	15
2.1.3.2	Merkmalsextraktion durch Hauptkomponentenanalyse (Principle-Component-Analysis)	17
2.1.3.3	Fisher Diskriminanzanalyse (FLDA)	17
3	Brain-Composer: Komposition mittels P300-basiertem BCI	20
3.1	Motivation	20
3.2	Verwandte Arbeiten	20
3.3	Konzept	21
3.3.1	Ein alternativer Tonraum	21
3.3.2	Musikalische Information als Matrix	21
3.3.3	Grundlegender Aufbau von Brain-Composer	22
3.3.3.1	Grafische Oberfläche	22
3.3.3.2	Klangausgabe	23
3.3.3.3	Stimuluspräsentation	25
3.3.4	Programmablauf	27
3.4	Verwendete Technologien	27
3.4.1	Verwendete Software und Bibliotheken	28
3.4.1.1	Das UbiCi-Framework	30
3.4.2	Datenerfassung	31
3.4.3	Datenverarbeitung	31
3.4.3.1	Arbeitsschritte bei der EEG-Datenverarbeitung	32

3.4.3.1.1	Segmentierung	32
3.4.3.1.2	Balancieren	33
3.4.3.1.3	Filterung	33
3.4.3.1.4	Dimensionsreduktion	33
3.4.3.1.5	Klassifikation	34
4	Brain-Composer im Praxistest	35
4.1	Versuchsaufbau	35
4.2	Fragestellungen	36
4.3	Ergebnisse	36
4.3.1	Quantitative Ergebnisse	36
4.3.1.1	Klassifikationsgenauigkeit	36
4.3.1.2	Übertragungsgeschwindigkeit	37
4.3.1.3	Eingaben und Erzeugte Tonstrukturen	38
4.3.1.4	Fragebogen	38
4.3.2	Qualitative Ergebnisse	38
4.3.2.1	Fragebogen	38
4.3.2.2	Beobachtungen	38
4.3.3	Interpretation der Ergebnisse	39
5	Zusammenfassung	43
5.1	Ausblick	43
5.1.1	BCI-Performanz	44
5.1.2	Funktionalität und klangliche Möglichkeiten	44
5.1.3	Handhab- und Nutzbarkeit	44
A	Abkürzungsverzeichnis	46
B	Abbildungsverzeichnis	47
C	Tabellenverzeichnis	47
D	Literaturverzeichnis	49
E	Kompositionen der Probanden	52

1 Einleitung

Brain-Computer Interfaces (BCIs) bekommen eine immer größer wachsende Bedeutung. Sie erlauben dem Anwender die Steuerung von Anwendungen und Maschinen über die Gehirnaktivität. Ein Großteil der verbreiteten Eingabegeräte, wie Tastatur oder Maus, setzen die Verwendung des neuromuskulären Systems voraus. BCIs hingegen ermöglichen dem Anwender Steuersignale über einen alternativen Übertragungsweg zu senden.

Die BCI-Forschung fokussierte sich in den letzten Jahrzehnten hauptsächlich auf unterstützende Kommunikationshilfen für Menschen mit nicht intaktem motorischen Nervensystem. Es wurde jedoch wenig Forschung darüber betrieben, in welchem Rahmen ein BCI für den kreativen Ausdruck verwendet werden kann. Insbesondere die Erzeugung von Musikkompositionen über die Gehirnaktivität ist ein Terrain das noch weitgehend unbetreten geblieben ist.

Das Ziel der vorliegenden Arbeit ist die Entwicklung und Umsetzung eines Konzeptes, mit dem eine grundlegende, rudimentäre musikalische Komposition auf Basis des P300-Speller-Paradigmas ermöglicht werden soll. Ein Schwerpunkt liegt hierbei auf der gezielten Kontrolle des Klangergebnisses ohne zufällige oder nicht steuerbare Prozesse.

Die Bachelorarbeit beginnt mit einem Abschnitt über theoretische Grundlagen. Diese legt den Grundstein für das nachfolgende vorgestellte Konzept zur musikalischen Komposition mittels P300-basiertem BCI. Anschließend folgt ein Kapitel, in dem die entwickelte Software einem Praxistest unterzogen wurde. Hierbei werden zwei Benutzergruppen betrachtet: Musiker und Nicht-Musiker. Die durchgeführten Experimente werden beschrieben und die Ergebnisse zusammengefasst. In der abschließenden Zusammenfassung werden die Ergebnisse diskutiert und ein Fazit gezogen.

2 Theoretische Grundlagen

In diesem Kapitel wird der Begriff Brain-Computer Interface (BCI) eingeführt und dessen grundlegenden Komponenten beschrieben. Es folgt ein Abschnitt über gängige Messverfahren, bei dem die Funktionsweise des EEG detaillierter beschrieben und die Komponenten des EEG-Signals herausgestellt werden. Der theoretische Teil wird mit einem kurzen Abriss über gängige Verarbeitungsmethoden und Algorithmen abgeschlossen.

2.1 Brain-Computer Interfaces

Jegliche Form der Interaktion bzw. Kommunikation mit der Umwelt erfolgt über das neuromuskuläre System. Am Anfang eines Interaktionsvorganges steht die Intention des Individuums. Hierbei werden unterschiedliche Bereiche des Gehirns aktiv. Das Gehirn sendet nachfolgend Signale, die über die Efferenten Nervenfasern an die entsprechenden Motoneuronen gesendet werden, was zu den erwünschten Muskelkontraktionen führt. Unter einem BCI versteht man ein System, das eine Kommunikation mit der Umwelt ermöglicht, bei dem jedoch der neuromuskuläre Übertragungsweg verlassen wird und die Intention des Individuums über einen alternativen Signalweg übermittelt wird. Dies geschieht indem die Gehirnaktivität gemessen, somit auf die Intention geschlossen und entsprechende Signale erzeugt werden (vgl. [Graimann et al., 2010]).

Erste Forschungen begannen vor ca. 40 Jahren. Eine der Hauptmotivationen war die Entwicklung unterstützender Technologien, um Menschen, die unter einem Locked-In-Syndrom leiden (z.B. durch Amyotrophe Lateralsklerose), Kommunikation zu ermöglichen. Eine der frühesten Entwicklungen ist in diesem Zusammenhang, das "Right Justified Box"-Paradigma, das via EEG eine "Ja/Nein" - Kommunikation ermöglicht. Ein weiterer bedeutender Schritt war die erste Anwendung, die das Buchstabieren von Wörtern ermöglicht. Diese wurde 1988 erstmals mit dem P300-Speller-Paradigma beschrieben [Farwell and Donchin, 1988]. Neuere Entwicklungen, wie "Brainbrowser" ermöglichen Locked-In-Patienten das Surfen im Internet [Tomori and Moore, 2003]. Eine weitere bemerkenswerte Entwicklung, der sich verschiedenste Forschungsgruppen angenommen haben, ist ein navigierbarer Rollstuhl. Erst kürzlich stellte Toyota einen BCI-gesteuerten Rollstuhl vor, der über diskrete Parameter mit einer Genauigkeit von 95% gesteuert werden kann [Iturrate et al., 2009].

Unabhängig von den Entwicklungen, die Locked-In-Patienten eine große Hilfe bieten, zeigte sich, dass auch bei der Behandlung anderen Erkrankungen, wie z.B. Schlaganfall, Aufmerksamkeitsdefizitsyndrom (ADS), Autismus oder emotionalen Störungen der Einsatz von BCIs hilfreich sein kann [Graimann et al., 2010]. BCIs finden eine immer weitere Verbreitung und erreichen damit ein stetig größer werdendes Publikum und eine steigende Akzeptanz.

Auch die Unterhaltungsindustrie hat die Zeichen der Zeit erkannt und es werden neuerdings kostengünstige BCIs für den Konsumentenmarkt angeboten. Es ist denkbar, dass derartige Interfaces, sofern sie leistungsfähig genug sind, zu normalen Alltagsgegenständen werden könnten, wie z.B. ein Smartphone oder eine Maus zur Bedienung eines Computers.

2.1.1 Komponenten eines Brain-Computer Interfaces

Die folgenden Ausführungen stützen sich auf [Wolpaw et al., 2002].

Ein BCI lässt sich als System, dass aus mehreren Komponenten zusammengesetzt wird, beschreiben (siehe Abbildung 2.1). Als erste Komponente ist die Signalmessung zu nennen, die notwendig ist, um dem BCI Gehirnsignale zu übermitteln. Hierfür bieten sich, je nach Anwendung, unterschiedliche Verfahren an (siehe 2.1.2). Werden die Daten z.B. mittels EEG aufgezeichnet, so geschieht dies durch Elektroden, die auf der Kopfhaut platziert werden. Vor einer weiteren Verarbeitung wird das Signal verstärkt und digitalisiert. Die Signalverarbeitung kann in mehrere Teilschritte unterteilt werden. In einem Vorverarbeitungsschritt wird das Signal von Artefakten befreit z.B. durch Hoch- und Tiefpassfilter. Bei einer nachfolgenden Merkmalsextraktion wird die Datenmenge reduziert und eine Dimensionsreduktion durchgeführt. Hierbei bleiben im Idealfall jegliche Informationen, welche für die nachfolgenden Klassifikation relevant sind, enthalten. Durch die Anwendung eines Klassifikators wird aus den Daten ein Steuersignal gewonnen. Dieses Steuersignal wird von einer Anwendung empfangen um eine gewünschte Aktion auszuführen. In der Regel müssen die Parameter des Klassifikators im Vorfeld durch ein Training gewonnen werden.

2.1.2 Messverfahren

Gehirnaktivität erzeugt elektrische Signale, die mit verschiedenen Instrumenten gemessen werden können. Hierbei werden unterschiedliche Verfahren eingesetzt, die sich in Bezug auf räumlicher und zeitlicher Auflösung, sowie der Störanfälligkeit voneinander unterscheiden. Dieser Abschnitt beschreibt die gängigsten Messverfahren und ihre Funktionsweisen, mit einem Schwerpunkt auf die Electroenzephalografie. Die Ausführungen stützen sich größtenteils auf [Seifert, 2005]. Die weiteren Ausführungen basieren auf [Graimann et al., 2010], falls nicht anders gekennzeichnet.

2.1.2.1 Nichtinvasive Verfahren

Die elektrischen Signale, die bei Gehirnaktivität auftreten, sind nicht nur innerhalb des Gehirns, sondern auch außerhalb des Schädels messbar. Dies machen sich nichtinvasive Verfahren zu Nutze. Im folgenden möchte ich die verbreitesten Verfahren nennen:

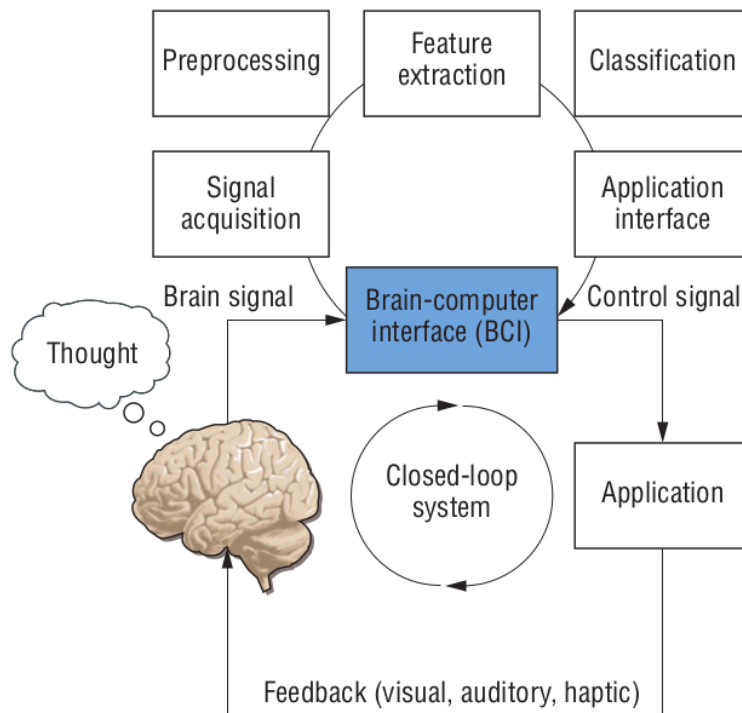


Abbildung 2.1: Die Komponenten eines Brain-Computer Interfaces [Nijholt and Tan, 2008]

Electroenzephalografie (EEG)

Die Elektroenzephalografie (EEG) misst die Hirnströme und zeichnet diese auf. Zu diagnostischen Zwecken werden die Hirnströme oft in einem Electroenzephalogramm dargestellt. Für die Ableitung der Hirnströme werden an den Schädel von außen Elektroden angebracht. Die Hirnströme resultieren aus exzitatorischen und inhibitorischen postsynaptischen Potentialen [Nijholt and Tan, 2008]. Ob ein Signal außerhalb des Schädels gemessen werden kann, hängt von vier Faktoren ab [Seifert, 2005, vgl. S.20-21]:

- Einzelne Nervenzellen sind nicht messbar, da die elektrischen Signale zu schwach sind. Es sind mindestens 10000 Nervenzellen notwendig.
- Die Zellen müssen in derselben Richtung ausgerichtet sein, damit sich die elektrischen Felder aufsummieren.
- Für das Aufsummieren ist notwendig, dass die elektrischen Felder zum gleichen Zeitpunkt entstehen. Das bedeutet, sie müssen synchron im gleichen Rhythmus feuern.
- Die elektrischen Felder dürfen nicht zu weit von der Kopfoberfläche entfernt sein. Je weiter die Quelle entfernt ist, um so niedriger wird das Signal.

Die Nervenzellen sind hierbei innerhalb des Neokortex besonders günstig gelegen, da sie nicht weit von der Kopfoberfläche entfernt sind. Zudem sind die Pyramidzellen senkrecht zur Kortexoberfläche ausgerichtet. Somit summieren sich die Potentiale mehrere Zellen auf (siehe Abbildung 2.2).

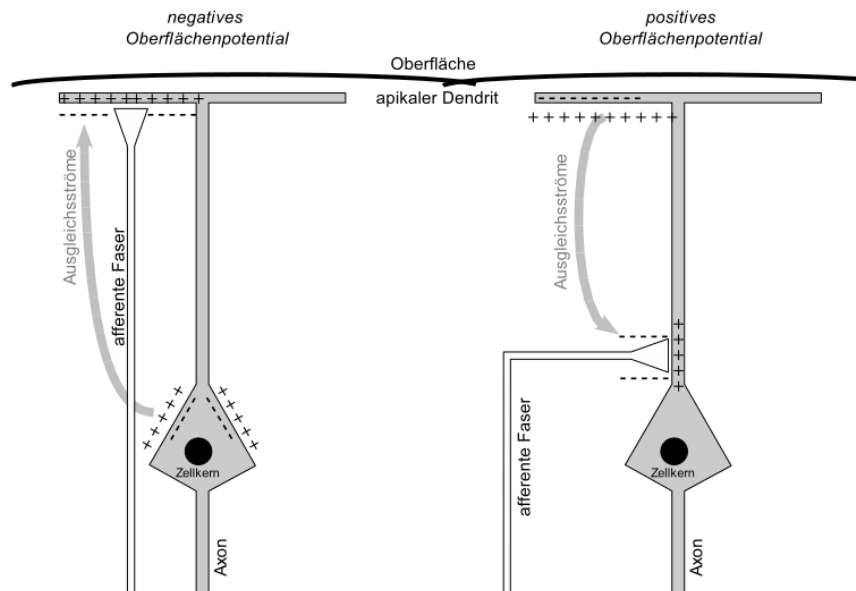
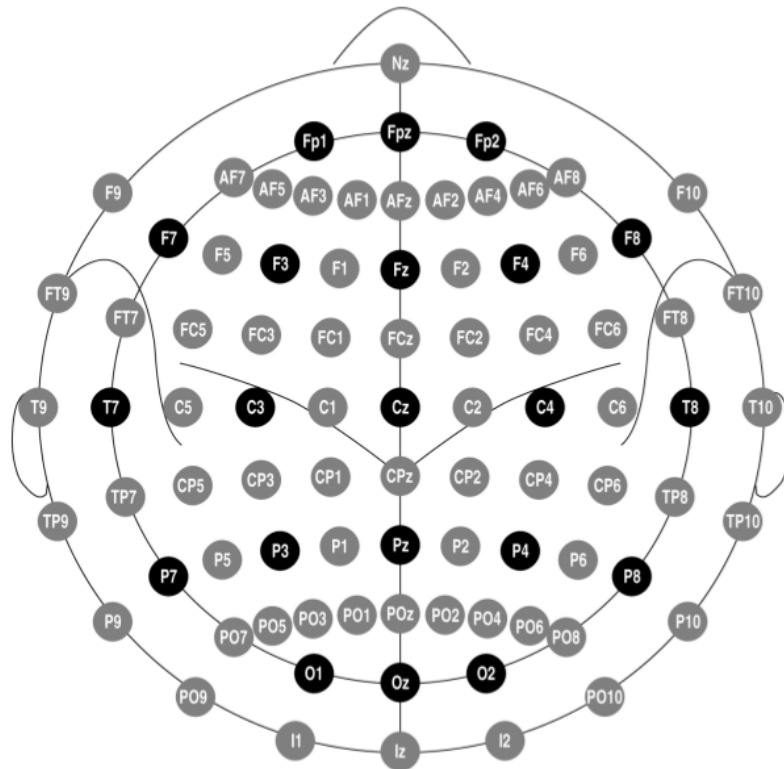


Abbildung 2.2: Die Pyramidenzellen des Neokortex sind senkrecht zur Kortex-Oberfläche ausgerichtet. Die synchrone (De-) Aktivierung ganzer Zellhaufen löst die elektrischen Ströme aus, die als EEG gemessen werden. Beim exzitatorischen postsynaptischen Potential kommt es zu einem Einstrom positiver Ladungsträger am distalen (entfernten) Dendriten. Von dort wird (etwas vereinfacht ausgedrückt) die Ladung in den Zellkern weitergeleitet, wo sie die Zelle wieder verlässt. Der extrazelluläre Strom vom Zellkern zum entfernten Dendriten schließt letztlich den Stromkreis. [Seifert, 2005]

Das EEG ist anfällig für Artefakte, die aus unterschiedlichen Quellen herrühren. Artefakte können einerseits vom Benutzer erzeugt werden, z.B. durch Muskelaktivitäten, Puls oder Schwitzen, andererseits durch externe Störeinflüsse wie elektrochemische Veränderungen an den Elektroden, durch die Netzspannung oder einen schlechten Elektrodensitz. Ein wichtiger Parameter für die Qualität des EEG-Signals wird über die Impedanz angegeben. Dazu wird jede Elektrode relativ zu einer Referenzelektrode gemessen, die üblicherweise hinter den Ohren, oder am Ohrläppchen befestigt wird. Bei einem Großteil der Experimente wird als Grenzwert für die Signalqualität 5kOhm angesetzt. Die Platzierung der Elektroden ist international mit dem 10-20 System (bzw. 10-20 extended) standardisiert (siehe Abbildung 2.3). Das EEG ist ein verbreitetes Verfahren für die Steuerung von BCIs. Die Gründe hierfür sind die hohe zeitliche Auflösung von 0.05s [Clausen, 2010, S.7], die vergleichsweise geringen Anschaffungskosten und

die hohe Portabilität. Das EEG ist das mit Abstand am meisten eingesetzte Verfahren für BCIs. Nachteilig ist die schwache räumliche Auflösung von etwa 1cm [Clausen, 2010, S.7] und ein z.T. zeitintensives Anbringen der Elektroden.



Funktionelle Magnetresonanztomographie (fMRI)

Die Funktionelle Magnetresonanztomographie (fMRI) ist ein bildgebendes Verfahren, mit dem Stoffwechselvorgänge im Gehirn sichtbar gemacht werden können. fMRI misst hierbei Veränderungen in Abhängigkeit des Sauerstoffsgehaltes im Blut (blood oxygen level dependent (BOLD)), die durch neuronale Aktivität ausgelöst werden. fMRI bietet eine hohe räumliche Auflösung, die im Millimeterbereich liegt [Yoo et al., 2004]. Patienten können die Regulierung des BOLD-Effektes in eng eingegrenzten Arealen erlernen. Somit ist dieses Verfahren auch für Echtzeit-Anwendungen geeignet [Clausen, 2010, S.7] Ein großer Nachteil ist die geringe zeitliche Auflösung. Die benötigte Ausrüstung ist nicht portabel.

Nahinfrarotspektroskopie (NIRS)

Bei der Nahinfrarotspektroskopie (NIRS) wird die Gehirnaktivität, genau wie beim fMRI, durch Veränderungen des Sauerstoffgehaltes im Blut gemessen. Es werden hierfür Lichtwellen im Bereich von 630 bis 1350nm verwendet [Nijholt and Tan, 2008]. Dieses Messverfahrens bietet einige Vorteile. Es ist kostengünstig, lässt sich ohne lange Vorbereitungen einsetzen und ist portabel. In puncto Signalqualität verbindet es die Nachteile des EEGs und des fMRIs. Die räumliche Auflösung liegt im Zentimeterbereich und die zeitliche Auflösung überschreitet mehrere Sekunden [Nijholt and Tan, 2008].

2.1.2.2 Invasive Verfahren

Bei Invasiven Messverfahren wird der Schädel durchdrungen. Dadurch lassen sich Messinstrumente gezielt platzieren, oder in das Gehirngewebe implantieren. Sie ermöglichen hierdurch eine deutlich bessere räumliche Auflösung und eine, im Gegensatz zu nichtinvasiven Verfahren, erhöhte Signalqualität. Neben der Signalqualität spielen jedoch auch Eingriffsrisiken und Langzeitfolgen, wie z.B. Gewebevernarbungen und eine erhöhte Infektionsgefahr eine Rolle. Zudem werfen Invasive Verfahren moralische und ethische Fragen auf.

Bei invasiven Verfahren werden einzelne Elektroden (Local Field Potentials (LFP)) oder Multielektrodenarrays (MEA) im Cortex platziert. Daneben gibt es mit der Elektrokortigrafie (Ecog) ein weniger tiefgreifendes invasives Verfahren. Hierbei werden Gitterelektroden direkt auf den Cortex gelegt, ohne in ihn einzudringen [Clausen, 2010].

Die Erforschung invasiver Verfahren findet größtenteils an Affen statt. Versuche am Menschen konnten jedoch zeigen, dass mit intrakortikalen Ableitungen ein Roboterarm in drei Dimensionen gezielt gesteuert werden kann [Hochberg et al., 2012]. Hiermit sind sogar komplexe Bewegungen, wie das Greifen eines Trinkgefäßes, das anschließend zum Mund geführt wird, möglich. Invasive Verfahren bieten ein hohes Potential und werden intensiv erforscht. Sie spielen jedoch für die praktische Anwendung zum derzeitigen Stand keine Rolle.

2.1.2.3 Ereigniskorrelierte Potentiale (ERP)

Bei der Verarbeitung sensorische Reize kann eine definierte Folge von negativen und positiven Auslenkungen im EEG-Signal ausgelöst werden. Hierbei spricht man von Ereigniskorrelierten Potentialen (ERP), die zu fest definierten Zeitpunkten auftreten, nachdem ein Reiz präsentiert wurde. ERPs unterscheiden sich dahingehend, mit welcher Verzögerung sie nach einem dargebotenen Reiz auftreten sowie durch ihre Topologie [Graimann et al., 2010].

ERPs haben eine geringere Auslenkung ($5\text{--}10\ \mu\text{V}$) als das Spontan-EEG (ungefähr $30\text{--}60\ \mu\text{V}$). Um die ERPs vom Spontan-EEG zu trennen, müssen daher Mittelungen über mehrere Zeitepochen durchgeführt werden um sie zuverlässig aufzuspüren. Die ERP-Komponenten werden mit einem Buchstaben und einer dazugehörigen Zahl benannt. Hier bezeichnen die Buchstaben “N” und “P” jeweils eine negative bzw. positive Amplitude und die Zahl gibt die ungefähre Latenz an, bei dem das ERP seinen Höhepunkt erreicht. [Gerloff, 2005, S.504]. Die Ausprägung und Latenz einer spezifischen ERP-Komponente kann sich bei unterschiedlichen Personen sehr stark unterscheiden. Betrachtet man jedoch mehrere gleiche ERP-Komponenten einer einzelnen Person, so sind diese Schwankungen relativ gering [Luck, 2005, S.12ff]. ERP-Komponenten können in zwei verschiedene Gruppen unterteilt werden: Exogene und endogene Komponenten.

Der Begriff ERP ist nicht scharf definiert, jedoch erfolgt in der Regel eine Abgrenzung zwischen Evozierten Potentialen Ereigniskorrelierten Potentialen. Alternativ werden hierfür auch die Begriffe exogene/endogene Komponenten verwendet. Exogene Komponenten haben eine kürzere Latenz und hängen hauptsächlich von physikalischen Reizen ab, von denen sie ausgelöst werden. Hierunter fallen u.a. somatosensorische evozierte Potentiale (SEP), visuell evozierte Potentiale (VEP) oder akustisch evozierte Potentiale (AEP). Die endogenen Komponenten hingegen haben eine höhere Latenz. Sie werden nicht zwingend durch einen Reiz ausgelöst. Endogene Komponenten basieren auf höheren kognitiven Verarbeitungsprozessen wie z.B. die Konzentration auf einen Zielreiz. [Gerloff, 2005, S.502]

Eine geeignete Komponente für BCI-Anwendungen stellt die P300 dar. Deswegen wird diese im Folgenden etwas genauer beleuchtet. Einen umfassenden Überblick über weitere Komponenten und ihre Bedeutungen findet sich in [Gerloff, 2005]

2.1.2.4 Die P300-Komponente

Nachstehender basiert im Wesentlichen auf [Gerloff, 2005], falls nicht anders gekennzeichnet.

Die P300-Komponente (auch P3 genannt) wurde erstmalig 1965 beschrieben [Sutton et al., 1965]. Grundlage war hierfür das “Zwei-Stimulusdiskriminationsparadigma” (engl. oddball-paradigma), bei dem in einer Gruppe von Hintergrundreizen an zufälligen Zeitpunkten Zielreize (engl. “oddballs”) eingestreut werden, auf die die Versuchsperson ihre Auf-

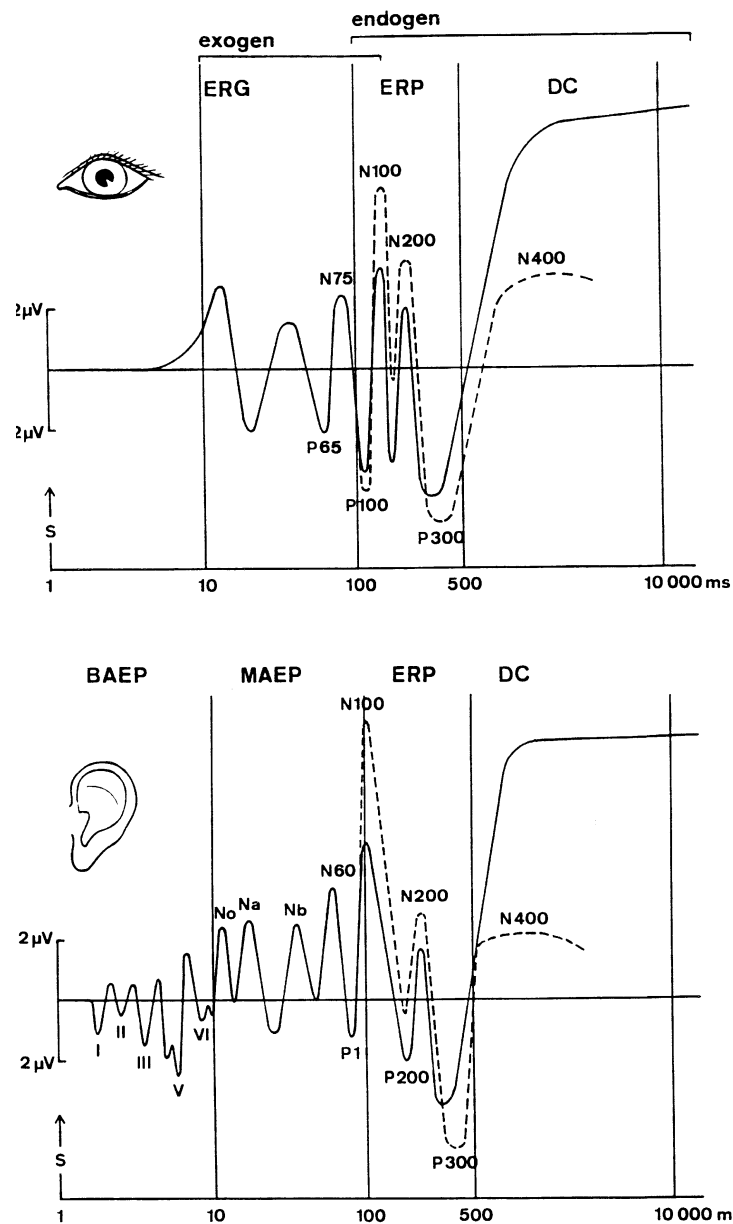


Abbildung 2.4: Schematisierte Darstellung der endogen und exogen evozierten Potenziale nach (a) visueller und (b) auditorischer Reizung auf einer logarithmischen Zeitskala. In der gestrichelten Linie sind die Potentiale bei Aufmerksamkeitszuwendung dargestellt. [Gerloff, 2005, S.503]

merksamkeit lenken soll. Bei jenem Experiment wurde im EEG-Signal ein starker positiver Ausschlag mit einer Latenz von ungefähr 350ms unmittelbar nach dem Zielreiz festgestellt.

Die P300 kann auch durch Variationen des oddball-Paradigmas ausgelöst werden. Im “Einzel-Stimulusparadigma“ (engl. single-stimulus-paradigma) wird nur der Zielreiz ohne weitere Hintergrundreize präsentiert. Dies geschieht zu unbestimmten Zeitpunkten und die Versuchsperson wird angewiesen ihre Aufmerksamkeit auf den Reiz zu lenken. Eine weitere Variation ist das “Drei-Stimulus-Paradigma“ (engl. three-stimulus-paradigma). Dem oddball-paradigma wird ein zweiter Hintergrundreiz (auch Distraktorreiz genannt) hinzugefügt, welcher jedoch wesentlich seltener als der ursprüngliche Hintergrundreiz dargeboten wird. Die Paradigmen sind in Abbildung 2.4 schematisch dargestellt.

Die psychophysiologische Bedeutung des P300 wird unterschiedlich interpretiert. Im weitesten Sinne wird eine Art Kategorisierung und Evaluation des dargebotenen Stimulusmaterials angenommen [Donchin and Coles, 1988]. Alternativ wird die Annahme vertreten, dass die P300 eine Deaktivierung nach Ende einer perzeptuellen Epoche darstellt [Deecke and Lang, 1988]. Des Weiteren gibt es die Vermutung, dass die P300 die Geschwindigkeit der kognitiven Verarbeitungsprozesse angibt [Donchin and Coles, 1988]. [Gerloff, 2005, vgl S.514]

Es werden drei verschiedene P3-Komponenten unterschieden: Die P3b, P3a und die Slow Cortical Potentials (SCP). Je nach experimentellen Bedingungen können diese gezielt hervorgerufen werden.

In der Literatur werden die Bezeichnungen P300 und P3 als Synonym für die P3b-Komponente verwendet. Sie ist die klassische P300-Komponente. Ihre Latenz liegt zwischen 340 bis 700ms mit einer maximalen Amplitude über dem zentroparietalen Kortex. Im oddball-paradigma wird sie durch Fokussierung auf einen Zielreiz, z.B. durch mitzählen, ausgelöst.

Die P3a ist eine Komponente, die besonders stark ausgeprägt wird, wenn ein Zielreiz neuartige Merkmale aufweist. Ihre Latenz liegt bei 250 bis 350ms mit einem frontozentralen Maximum. Dadurch unterscheidet sie sich deutlich von der P3b. Die P3a wird auch bei Nicht-Zielreizen ausgelöst, jedoch in geringerem Ausmaß. Nach mehreren Versuchsdurchläufen tritt eine Habituation der P3a ein.

In einem Zeitfenster zwischen 600 und 1400ms folgen die SCP. Diese tritt jedoch nur auf, falls einer der Reize für die Lösung komplexerer Aufgaben notwendig ist.

Die Amplitude der P300 ist nicht immer gleichstark ausgeprägt. In diesem Zusammenhang spielt die Auftretenswahrscheinlichkeit des Zielreizes eine große Rolle. Damit ist gemeint, wie oft im oddball-Paradigma der Zielreiz im Vergleich zu einem Hintergrundreiz auftritt. Nimmt die Auftretenswahrscheinlichkeit des Zielstimulus ab, dann steigt die Amplitude der P300 an. In der Praxis werden meist Auftretenswahrscheinlichkeiten zwischen 15 und 20% gewählt. Unterhalb von 10% verbessert sich die Qualität der P300 nicht mehr (siehe hierzu Abbildung 2.6).

Die Latenz der P300 kann variieren. Hierbei gibt es einen Zusammenhang zu der Schwierigkeit, mit der ein Zielreiz detektiert werden kann. Mit zunehmender Schwierigkeit

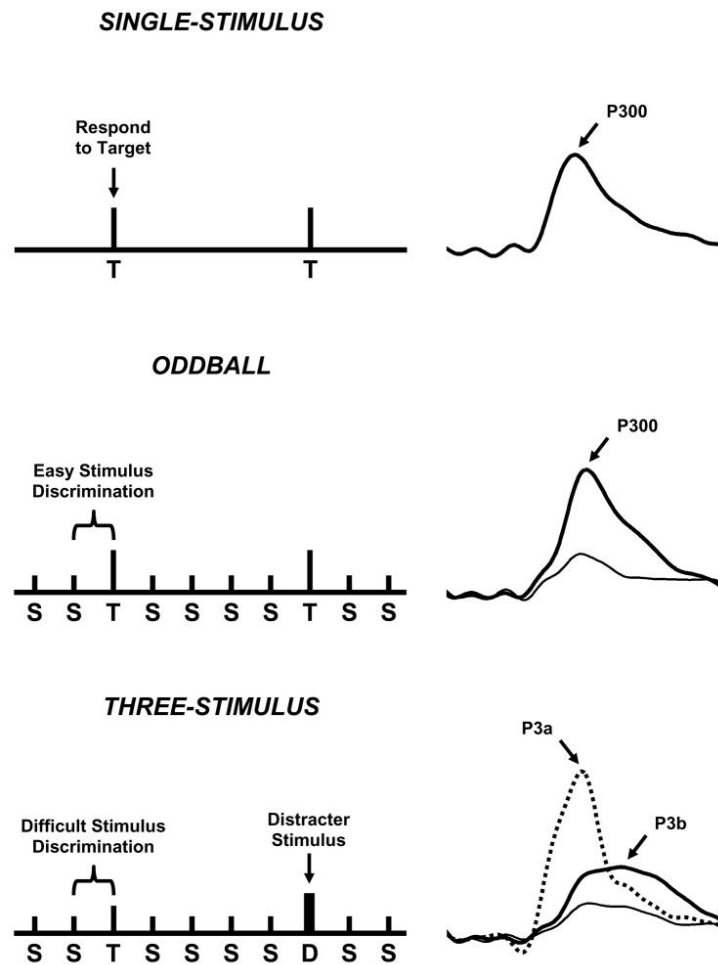


Abbildung 2.5: Schematische Skizzierung von single-stimulus (oben), oddball (mitte) und three-stimulus Paradigma. Rechtseitig wird das jeweilig ausgelöste ERP dargestellt. Im single-stimulus Paradigma wird ein Zielreiz (T) zu unbestimmten Zeitpunkten dargeboten. Das oddball-Paradigma präsentiert Hintergrundreiz (S) und Zielreiz (T) in einer zufälligen Sequenz. Das three-stimulus-Paradigma ähnelt dem oddball-Paradigma. Hier kommt ein Distraktorreiz (D), ein weiterer Hintergrundreiz hinzu, der jedoch seltener als der Hintergrundreiz auftaucht. In diesem Paradigma löst der Distraktorreiz ein P3a (gestrichelte Linie) und der Zielreiz ein P3b (durchgezogene Linie) aus. [Polich et al., 2006]

steigt die Latenz der P300 an. Dies ist beispielsweise der Fall, wenn sich Ziel- und Hintergrundreiz wenig unterscheiden. Zudem ist festzuhalten, dass sich die Latenz der P300 ab dem Erwachsenenalter kontinuierlich erhöht.

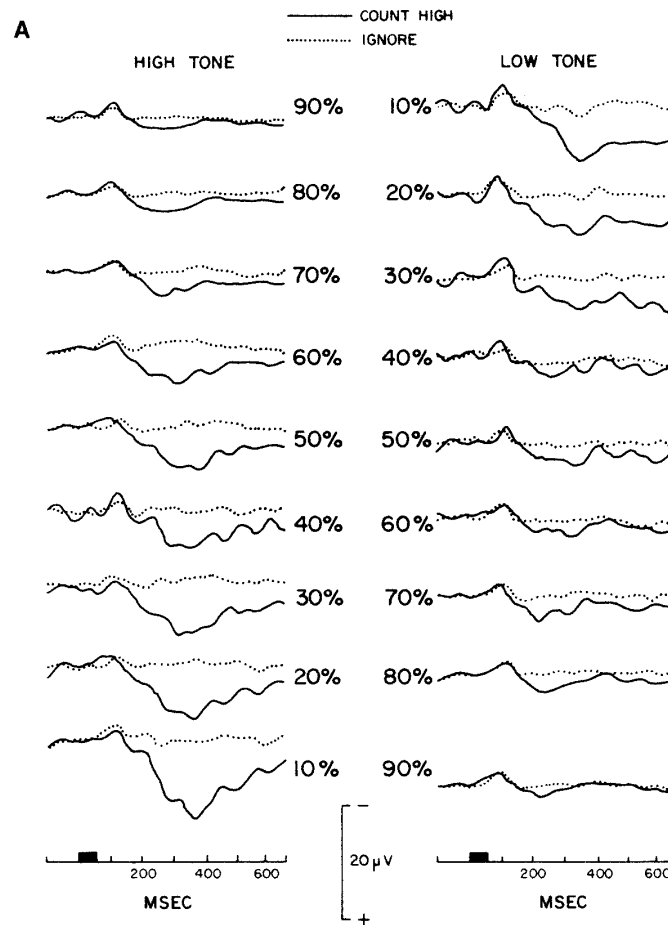


Abbildung 2.6: ERP-Plots an Elektrodenposition Pz, gemittelt über alle Versuchspersonen für hohe und tiefe Töne. Die durchgezogene Linie stellt den Verlauf dar, bei der die Versuchsperson angewiesen wurde den Zielreiz mitzuzählen. Die gestrichelte Linie den Verlauf, bei ignorierten Zielreiz. Das Auftreten des Stimulus wird durch einen schwarzen Balken auf der Zeitachse angezeigt. Bei dem Plot ist die Polarität auf den Achsen umgekehrt aufgetragen.[Duncan-Johnson and Donchin, 2007]

2.1.3 Allgemeine Verarbeitungsmethoden und Algorithmen

Im folgenden möchte ich den theoretischen Teil mit einem kurzen Abriss über gängige Verarbeitungsmethoden abschließen.

2.1.3.1 Filter auf Fourier-Basis

Jede beliebige periodische Funktion kann als Summe von Sinus- und Kosinusfunktionen ausgedrückt werden. Das geht aus der Definition der Fourier-Reihe hervor [Butz, 2005,

S.4-5].

$$f(t) = \sum_{k=0}^{\infty} (A_k \cos \omega_k t + B_k \sin \omega_k t) \quad (2.1)$$

mit

$$\omega_k = \frac{2\pi k}{T} \quad (2.2)$$

und den Filterkoeffizienten

$$A_k, B_k \quad (2.3)$$

Mittels der Fouriertransformation lässt sich ein Signal von der Zeit- in die Frequenzdomäne umwandeln. Im ersten Schritt der Fouriertransformation werden zunächst die Filterkoeffizienten gebildet. Die Fouriertransformation geht von einem stationären Signal aus, was jedoch bei EEG-Daten nicht der Fall ist. Um dieses Problem zu umgehen wird die Fast-Fourier-Transformation (FFT) angewendet. Hierbei wird das zu analysierende Signal in kleinere Segmente unterteilt. Bei der Bildung der Filterkoeffizienten müssen diese bei der FFT gefenstert werden. Die Filterkoeffizienten werden hierbei zu den Rändern hin stetig verkleinert.

Die gefensterten Filterkoeffizienten werden nun über das Signal geschoben und an jeder Position des Zeitspektrums eine Summe berechnet, die mit den Filterkoeffizienten an der entsprechenden Position multipliziert werden. [Seifert, 2005]

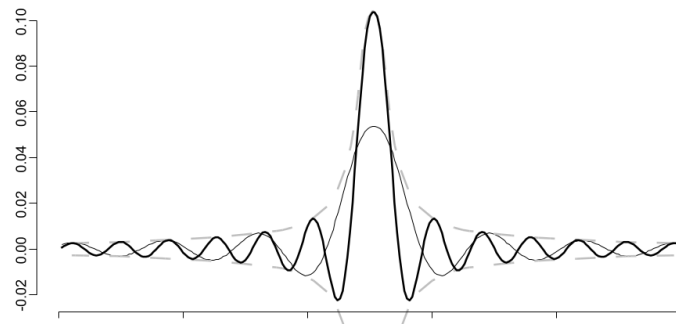


Abbildung 2.7: Zwei Mengen von Filterkoeffizienten (schwarz), die durch eine Fensterfunktion (grau gestrichelt) an den Rändern gedämpft wurden. [Seifert, 2005]

Ein derlei zerlegtes Signal lässt es sich auch in die andere Richtung, d.h. von der Frequenz- in die Zeitdomäne zurücktransformieren.

Dies können man sich bei einer einer Filterung auf Fourierbasis zu Nutze machen. Das zu filternde Signal wird in der Frequenzdomäne von unerwünschten Frequenzen befreit in dem diese gelöscht werden. Eine anschließende Rücktransformation überführt das Signal wieder in die Zeitdomäne, das nun gefiltert ist.

2.1.3.2 Merkmalsextraktion durch Hauptkomponentenanalyse (Principle-Component-Analysis)

Die Hauptkomponentenanalyse (engl. Principle Component Analysis (PCA)) ist ein Verfahren der multivariaten Statistik. Aufgabe der PCA ist es, die ursprünglichen Daten durch eine kleinere Anzahl von Variablen auszudrücken, bei gleichzeitiger Erhaltung der Varianz. Dieses Verfahren ist eine der Standardmethoden der Merkmalsextraktion auf EEG-Daten und wird vor dem Klassifikationsschritt durchgeführt. Der Grund hierfür ist, dass EEG-Daten sehr schnell hochdimensional werden. Neben den Auswirkungen auf die Performance, sind hochdimensionale Daten für Klassifikationsprobleme ineffizient. Der Grund hierfür liegt in dem sogenannten “Fluch der Dimensionalität”. Mit steigender Dimension wächst das Volumen einer Zelle im Vektorraum exponentiell. Dies führt zu großen Leeräumen zwischen Datenpunkten im Hochdimensionalen. ([Bishop, 2006, S.33-36])

Die grundlegende Idee hinter der PCA ist, durch einen Basiswechsel die Daten aus einem D -Dimensionalen in einen M -Dimensionalen Raum zu projizieren, mit $D > M$ und unter Beibehaltung der größtmöglichen Varianz. Hierfür müssen M linear unabhängige Vektoren gefunden werden, die die Basis des Vektorraums bilden, in den die Projektion erfolgen soll. [Hachmeister and Riechmann, 2010, S.51-52]

Die Basisvektoren werden aus den Eigenvektoren der Kovarianzmatrix gebildet. Die zugehörigen Eigenwerte geben an, wie groß der Anteil der Hauptkomponente an der Gesamtvarianz ist. Üblicherweise werden die Eigenvektoren absteigend nach ihrem Eigenwert angeordnet. Es können nun so viele Hauptkomponenten ausgewählt, bis ein gewünschtes Maß an Gesamtvarianz der Daten abgedeckt wird.

2.1.3.3 Fisher Diskriminanzanalyse (FLDA)

Die Fisherdiskriminanzanalyse (FLDA) ist ein lineares Klassifikationsverfahren. Die Berechnung einer Hyperebene, die zur Klassentrennung dient, ist im Mehrdimensionalen sehr komplex. Die FLDA umgeht dieses Problem durch die Projektion eines D -Dimensionalen Eingabevektors auf eine einzige Dimension. Hierbei lässt sich mittels FLDA eine optimale Projektion, mit der sich die Klassen möglichst gut trennen lassen, finden. In den folgenden Ausführungen wird ein Zwei-Klassen-Problem betrachtet. Grundlage bildet hierfür [Bishop, 2006, S.186-189].

Die FLDA erwartet einen D -Dimensionalen Eingabevektor \vec{x} und projiziert ihn auf einen skalaren Wert

$$y = \vec{w}^T \vec{x} \quad (2.4)$$

Zur Bestimmung unseres Klassifikationsergebnisses benötigen wir einen Schwellwert w_0 . Gilt für unsere projizierten Eingabedaten $y \geq w_0$, so werden diese der Klasse C_1 , andernfalls der Klasse C_2 zugewiesen. Im allgemeinen führt eine Dimensionsreduktion zu Datenverlust. Daten die sich im D -dimensionalen gut trennen lassen, können sich im

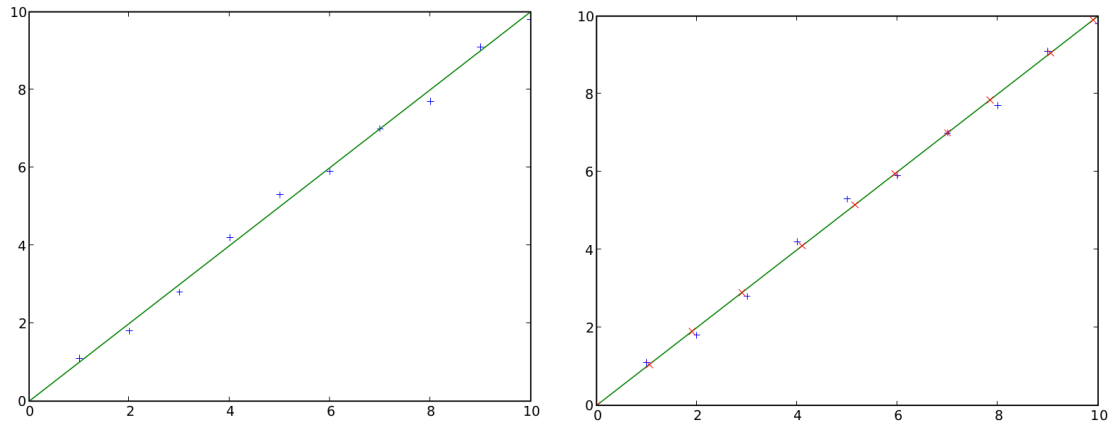


Abbildung 2.8: Beispiel für eine PCA:

Links: Der zweidimensionale Datensatz (durch blaue Kreuze dargestellt) variiert hauptsächlich in Richtung des Vektors (1,1).

Rechts: Die Daten wurden auf den Vektor (1,1) projiziert (durch rote Kreuze dargestellt). Ein Großteil der Informationen bleibt trotz Reduktion auf eine einzelne Dimension erhalten. [Hachmeister and Riechmann, 2010, S.52]

1-dimensionalen überlappen. Um ein gutes Klassifikationsergebnis zu erreichen muss dafür gesorgt werden, dass die projizierten Daten eine möglichst gute Klassentrennung erlaubt. Dies wird erreicht, in dem die Klassenmittelpunkte möglichst weit auseinandergezogen werden.

Die Mittelwerte der beiden Klassen sind definiert durch

$$\vec{m}_1 = \frac{1}{N_1} \sum_{n \in C_1} x_n \quad \vec{m}_2 = \frac{1}{N_2} \sum_{n \in C_2} x_n \quad (2.5)$$

Die Mittelwerte unserer Projektion sind gegeben durch

$$m_1 = \vec{w}^T \vec{m}_1 \quad m_2 = \vec{w}^T \vec{m}_2 \quad (2.6)$$

Nun wird die die Intervarianz maximieren, in dem die Mittelwerte der Projektion möglichst weit auseinander gezogen werden

$$\max(m_2 - m_1) \quad (2.7)$$

Dieses Kriterium ist jedoch noch nicht ausreichend. Die Intravarianz, also die Varianz innerhalb der Klassen wird hier noch nicht beachtet.

Die Intravarianz wird wie folgt definiert

$$s_k^2 = \sum_{n \in C_k} (y_n - m_k)^2 \quad (2.8)$$

Eine möglichst optimale Projektion minimiert die Intravarianz

$$\min(s_1 + s_2) \quad (2.9)$$

Mit dem Fisherkriterium lässt sich nun das Verhältnis zwischen Intervarianz und Intravarianz zu einer Kennzahl zusammenfassen. Sie gibt die Güte der Projektion für einen gegebenen Gewichtsvektor an.

$$J(w) = \frac{(m_2 - m_1)^2}{s_1^2 + s_2^2} \quad (2.10)$$

Die Trennbarkeit der Klassen wird optimal, wenn $J(w)$ maximal wird.

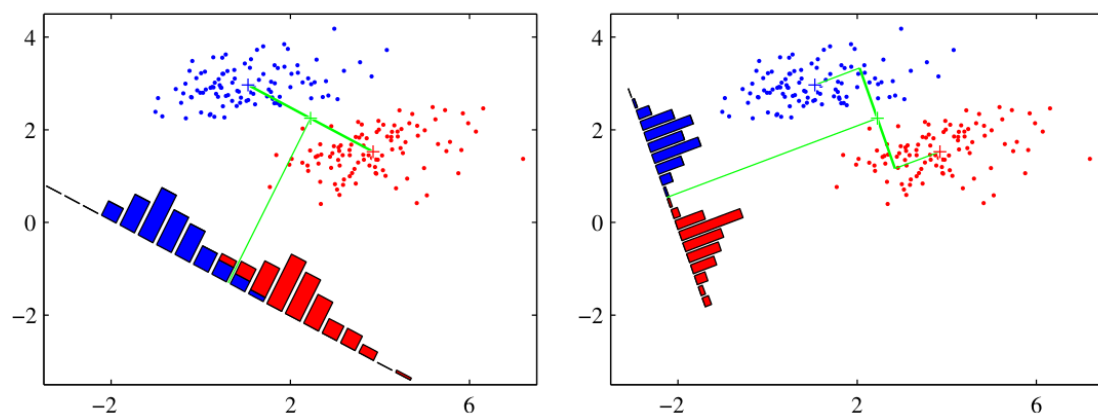


Abbildung 2.9: Projektionen von Daten aus zwei verschiedenen Klassen (in rot und blau dargestellt). In der linken Abbildung wurde die Intervarianz maxmiert. Die Klassenmittelpunkte sind auseinandergezogen, jedoch sind hier noch Überlagerungen der Klassen vorhanden.

In der rechten Abbildung wurde das Fisherkriterium verwendet. Wodurch eine bessere Klassen-separation erreicht wird.[Bishop, 2006]

3 Brain-Composer: Komposition mittels P300-basiertem BCI

3.1 Motivation

BCIs ermöglichen Kommunikation über einen alternativen Übertragungsweg und sind daher besonders interessant für Menschen, deren neuromuskuläres System beeinträchtigt ist. Daher liegt ein Schwerpunkt der BCI-Anwendungen darauf, grundlegende Kommunikationsmöglichkeiten zu schaffen. Aus diesem Grunde findet sich eine Vielzahl von BCI-Anwendungen, dessen Ziel es ist einen Text zu produzieren (z.B. um eine E-Mail zu schreiben, oder um im Internet zu surfen [Tomori and Moore, 2003]).

Auffällig ist, dass derzeit nur sehr wenige Anwendungen existieren, in denen ein BCI für kreative Tätigkeiten genutzt wird. Vorreiter ist in diesem Zusammenhang das Projekt Brainpainting [Muenssinger et al., 2010], eine BCI-Anwendung, die dem Anwender das Malen von Bildern ermöglicht. Somit wird auf Basis eines BCIs kreativer Ausdruck möglich. Brainpainting ist eines der wenigen Projekte, welches einen Schnittpunkt zwischen Kunst und BCIs schafft.

BCI-Anwendungen, die eine Komposition von Musik erlauben, sind derzeit rar gesät und bieten sehr beschränkte Möglichkeiten. Einige Anwendungen sind unter 3.2 aufgeführt.

Der Ausgangspunkt dieser Arbeit besteht darin, ein Konzept zu finden, welches musikalische Komposition über ein BCI ermöglicht. Im Gegensatz zu den bisher bekannten Arbeiten soll sich die geplante Anwendung möglichst gut mit den Eigenheiten eines BCIs vereinbaren lassen und darüber hinaus ein gewisses Maß an musikalischer Komplexität bieten.

3.2 Verwandte Arbeiten

Hier seien nur Arbeiten erwähnt, bei denen die Komposition mittels P300 basiertem BCI erfolgt.

Beim dem Projekt “Museegk“ [Chew and Caspary, 2011] wählt der Anwender über eine Spellermatrix aus einer Dur-Tonleiter (acht Töne) genau einen Ton zu einem von acht möglichen Zeitpunkten. Audiofeedback gibt es unmittelbar nach jeder Tonauswahl und die gesamte Phrase wird einmalig nach Auswahl von acht Tönen komplett abgespielt.

Eine weitere Arbeit wird in dem Paper „Brain-Computer Interface (BCI) based Musical Composition“ [Hamadicharef et al., 2010] beschrieben. Hier kann der Benutzer beliebige Viertelnoten auf ein graphisch dargestelltes Notenblatt setzen. Vorteilhaft ist, dass im Vorfeld die Tonart gewählt werden kann. Zudem können Viertelpausen gesetzt, gesetzte Noten gelöscht und die Noten jederzeit abgespielt werden. Für einen besseren Überblick

eignet sich die Videodemonstration der Software¹

Hervorzuheben ist, dass der Anwender bei beiden oben genannten Systemen über musiktheoretische Kenntnisse und musikalisches Vorstellungsvermögen verfügen muss, um die Software gezielt nutzen zu können. Dies schränkt den Raum der potentiellen Benutzer drastisch ein. Zudem ist keinerlei Polyphonie möglich ist. Die musikalischen Möglichkeiten beider Arbeiten sind zudem sehr begrenzt, da nur Viertelnoten erzeugt werden können.

3.3 Konzept

3.3.1 Ein alternativer Tonraum

Es erscheint zielführender, sich von der Notenschrift zu lösen und dem Anwender stattdessen eine andere Art der Visualisierung des Tonraumes anzubieten, die sich mit den Eigenheiten des P300 basierenden BCIs besser vereinbaren lässt. Hier ist vor allem die vergleichsweise langsame Übertragungsrate hervor zu heben.

Wird mit einer visuellen Stimuluspräsentation gearbeitet, so wird in der Regel eine Spellermatrix verwendet. Es liegt die Idee nahe, ebenfalls eine Matrix zur Repräsentation der Musik zu verwenden, um eine direkte Übersetzung zwischen dem Speller und der musikalischen Information herzustellen.

3.3.2 Musikalische Information als Matrix

Die Idee, musikalische Information in einer Matrix zu repräsentieren wurde bereits vielfach aufgegriffen. Ein derartige Organisation wird gerne bei der synthetischen Klangerzeugung verwendet. Kommt dieses Prinzip zur Anwendung spricht man von sogenannten Step-Sequencern. Musik wird erzeugt, indem der Sequencer im Vorfeld programmiert wird. Dies geschieht durch Festlegung, welche Töne in welcher Länge zu einem festgelegten Zeitpunkt abgespielt werden sollen. In der Regel wird das Klangergebnis in einer Schleife abgespielt[Lötscher, 2003].

Ein Großteil der aktuell verbreiteten Drum-Computer arbeiten nach dieser Methode. Das grundlegende Prinzip eines Step-Sequencers mit musikalischer Information als Matrix ist die Organisation der Zeit in X-Richtung und der Tonhöhe in Y-Richtung. Hierbei lässt sich jedes Feld der Matrix ein- bzw. ausschalten und somit entscheiden, welcher Ton zu welchem Zeitpunkt gespielt wird. Da dies unabhängig voneinander für jeden Ton möglich ist, wird die Erzeugung polyphoner Klänge ermöglicht. Eine Matrix stellt hierbei einen Takt dar. Die Anzahl der Elemente in X-Richtung entsprechen der Takttrasterung. Bei acht Elementen ist jede Achtel des Taktes erreichbar, bei 16 Elementen jede Sechzehntel, usw. Ein weiteres Prinzip des Matrixsynthesizers ist das kontinuier-

¹[Hamadicharef et al., 2010], <http://www.youtube.com/watch?v=t-i2ogktrj0>

liche wiederholen (bzw. loopen), einer oder mehrerer Matrizen. Dies sorgt dafür, dass der Anwender direktes Feedback über seine Eingaben bekommt.

Als prominentes Beispiel sei an dieser Stelle das Produkt Tenori-On [Nishibori and Iwai, 2006] der Firma Yamaha erwähnt, welches eine Verbindung zwischen modernem Kunstobjekt und Musikinstrument herstellt. Auch hier arbeitet im Kern ein Step-Sequencer. Weiter verfolgen die Flashsoftware ToneMatrix², das Instrument BeatBearing [Bennett and O'Modhrain, 2008] den gleichen Ansatz. Der Vorteil dieser Herangehensweise ist ein intuitiver und spielerischer Zugang zur Musikkomposition.

3.3.3 Grundlegender Aufbau von Brain-Composer

Im Verlauf der Bachelorarbeit soll zunächst ein Prototyp mit begrenzten Möglichkeiten entwickelt werden um das Potential eines BCI-gesteuerten Step-Sequencers auszuloten. Bei der Entwicklung müssen die vergleichsweise niedrigen Datenübertragungsraten, die BCI-gesteuerte Anwendungen mit sich bringen, im Auge behalten werden. Weiter soll die Anwendung eine einfache und klare Struktur besitzen, damit sich der Benutzer auf die wesentliche Aspekte konzentrieren kann.

Für einen ersten Prototyp bietet sich eine Matrix der Größe 8x5 an, um einen guten Kompromiss zwischen musikalischer Komplexität und Datenübertragungsgeschwindigkeit zu finden.

In diesem Abschnitt wird die grundlegende Funktionalität und das Verhalten der Software beschrieben. Hinweise zu den verwendeten Technologien sind in Kapitel 3.4 zu finden.

3.3.3.1 Grafische Oberfläche

Die grafische Oberfläche (Abbildung 3.1) wird dem Benutzer grundsätzlich im Vollbildmodus dargestellt. Auf dem Bildschirm befindet sich linkseitig eine Spellermatrix, wie gemeinhin aus P300-basierenden BCI-Anwendungen bekannt (siehe 3.3.3.3)

Möchte der Anwender das Klangergebnis beeinflussen, so kann er zu diesem Zweck die Matrix des Step-Sequencers manipulieren. Jedes Feld kann entweder aktiviert oder deaktiviert werden und wird entsprechend weiß oder grau dargestellt.

Braincomposer verfügt über drei verschiedene Modi: Abspiel-, Klassifikations- und Trainingsmodus.

Der Trainingsmodus wird unmittelbar zu Beginn gestartet und dient dazu, den verwendeten Klassifikator zu trainieren.

Anschließend kann der Benutzer interaktiv zwischen dem Klassifikationsmodus und dem Abspielmodus wechseln.

Im Abspielmodus fährt ein Cursor, welcher durch einen hellgrauen vertikalen Balken dargestellt wird, entlang der X-Achse. Aktivierte Elemente innerhalb des Cursor werden

²<http://tonematrix.audiotool.com>

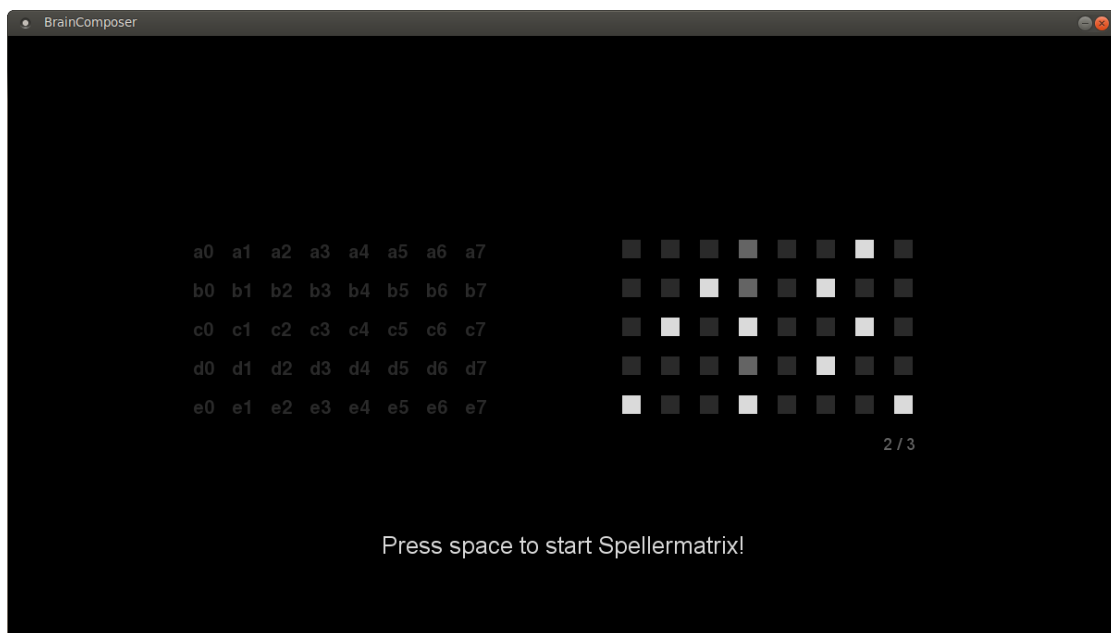


Abbildung 3.1: Braincomposer GUI

unmittelbar synthetisiert und abgespielt. Rechts unterhalb des Step-Sequencer-Matrix befindet sich das Matrixverzeichnis. Die erste Zahl gibt an, welche Matrix aktuell angezeigt wird, die zweite Zahl, wie viele Matrizen bisher angelegt wurden. Mit Hilfe der Cursortasten, kann der Anwender durch das Matrixverzeichnis blättern. Befindet sich der Anwender am Ende des Verzeichnisses und blättert weiter, so wird ein weiteres Matricelement angehängt.

Der Klassifikationsmodus wird mit der Leertaste gestartet. Hier kann der Anwender, gesteuert über das BCI, in jedem Durchlauf den Zustand genau eines Feldes der Step-Sequencer-Matrix verändern. Der Anwender lenkt hierzu seine Aufmerksamkeit auf ein Feld der Spellermatrix, welches mit dem zu setzenden Feld in der Step-Sequencer-Matrix direkt korrespondiert. Um die P300-Klassifikation nicht negativ zu beeinflussen, findet in diesem Modus keine Klangausgabe statt.

3.3.3.2 Klangausgabe

Dieser Abschnitt stütze ich mich, was die grundlegenden musiktheoretischen Grundlagen betrifft, gänzlich auf [Sikora, 2003, S.24-29].

Ein Großteil der Anwender von Braincomposer verfügt nicht über musiktheoretische Grundlagen und hat insbesondere keine Kenntnisse darüber, welche Intervalle eine tonale Dissonanz bzw. Konsonanz erzeugen.

"Dissonanzen erzeugen im Ohr ein Gefühl der Unruhe, des Konflikts, des Auflösungsbestrebens. Umgekehrt sind die Intervalle, die dem Grundton am nächsten liegen zu-

nehmend konsonanter = spannungsärmer (konsonierend = zusammenklingend) Konsonanzen vermitteln Stimmigkeit. Ruhe, Stabilität."³

Innerhalb einer Oktave können Intervalle anhand ihres Spannungsgrades angeordnet werden.

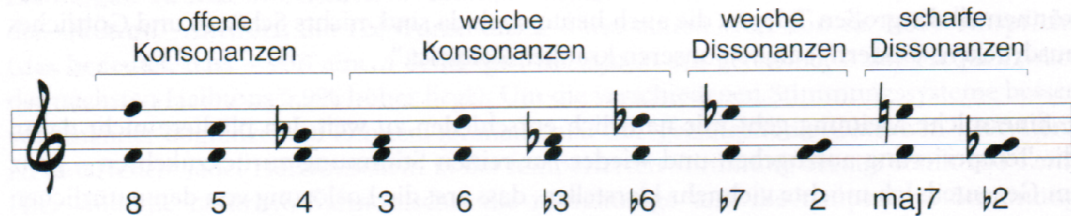


Abbildung 3.2: Anordnung der Intervalle entsprechend ihres Spannungsgrades[Sikora, 2003, S.24]

Natürlich ist eine derartige Beurteilung von den jeweiligen Hörgewohnheiten und dem kulturellen Umfeld geprägt. Die Abbildung Sikora bietet nur eine grobe Orientierung, die nicht als allgemeingültig betrachtet werden kann.

Für die Wahl eines geeigneten Tonsystems ist festzuhalten, dass die reinen Intervalle (Prim, Quarte, Quinte, Oktave) eine große Konsonanz erzeugen. Eine starke Dissonanz lösen hingegen Intervalle aus, die einen Halbton voneinander entfernt liegen. Diese Feststellung macht bereits Pythagoras in seiner Proportionstheorie: Je einfacher das Schwingungsverhältnis zweier Töne, desto konsonanter.

Im Bezug auf Braincomposer soll es dem Anwender möglich sein, sich ohne musikalische Kenntnisse dem System spielerisch zu nähern. Daher sollte ein Tonsystem zu Grunde liegen, bei dem eine starke Dissonanz der Töne ausgeschlossen wird. Hier sei angemerkt, dass Dissonanz ein wichtiges und unverzichtbares Stilmittel in der Kompositionstechnik ist. Im Rahmen der vorgestellten Software würde dies jedoch die Komplexität für den Anwender unnötig erhöhen.

Bei der Wahl unseres darunterliegenden Tonsystems haben wir festgestellt, dass sich in der vorliegenden Anwendung vor allem die reinen Intervalle eignen. Verfolgt man diesen Ansatz weiter, so gelangt man zu der Pentatonik.

"Der Begriff Pentatonik (griech. = fünftönig) subsummiert verschiedenste den Oktavraum mit fünf Stufen ausfüllende Skalen".⁴

Hierbei unterscheidet man hemitonische und anhemitonische Pentatoniken. Hemitonische Pentatoniken enthalten Halbtonschritte. Bei der anhemitonischen Pentatonik ist die kleine Sekunde das kleinste enthaltene Intervall. Die anhemitonische Pentatonik lässt

³[Sikora, 2003, S.24], S.24

⁴[Amon, 2011, S.296]

sich ausschließlich durch die Verwendung reiner Intervalle herleiten. Das Grundprinzip ist hier die Schichtung von vier Quinten, die in einem Tonraum zusammengezogen werden. [Amon, 2011]

Innerhalb von Braincomposer wird auf eine Variante der elementaren anhemitonische Pentatonik gesetzt (Abbildung 3.3 und 3.4) . Diese entsteht, indem von einem Ausgangston jeweils nach oben und nach unten zwei Quintschichtung erfolgen: $f - c \leftarrow g \rightarrow d - a \Rightarrow c - d - f - g - a$ [Amon, 2011]

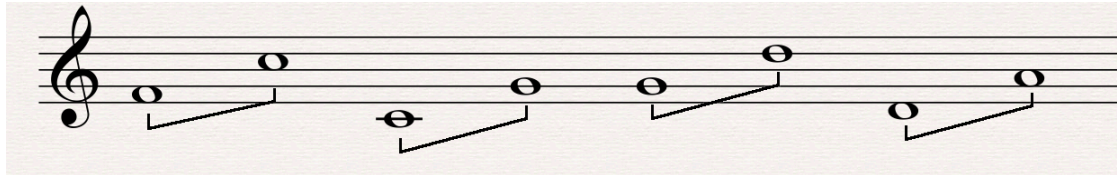


Abbildung 3.3: Herleitung der anhemitonischen Pentatonik durch Schichtung von vier Quinten



Abbildung 3.4: Zusammenziehen der vier geschichteten Quinten zu einer anhemitonische Pentatonik

Eine solche verwendete Pentatonik hat durch fehlenden kleinen Sekunden keinerlei Leittöne und wird dadurch zu einem offenem System.[Amon, 2011] Sie erzeugt einen schwebenden Charakter, der für das Ohr leicht greifbar ist. [Sikora, 2003]

3.3.3.3 Stimuluspräsentation

Braincomposer basiert auf dem P300-Speller-Paradigma, ein beliebtes und verbreitetes Verfahren um Daten über ein Brain-Computer-Interface zu übertragen. Diese Methode, aufbauend auf dem Oddball-Paradigma, ermöglicht eine zuverlässige und schnelle Datenübertragung. Erstmal wurde das P300-Speller-Paradigma 1988 in Zusammenhang mit einer Applikation, welche das Buchstabieren von Sätzen via EEG ermöglicht, eingeführt. [Farwell and Donchin, 1988]. Auch in anderen Anwendungen zeigte sich dieses Verfahren als sehr praktikabel.

Das grundlegende Prinzip eines Spellers funktioniert wie folgt: Der Anwender bekommt eine Matrix mit einer festen Anzahl von Elementen präsentiert. Nun wendet er seine Aufmerksamkeit auf genau ein Element, welches er zu auswählen beabsichtigt. Anschließend leuchten sequentiell einzelne Zeilen und Spalten in einer zufälligen Reihenfolge auf. Hierbei wird bei jedem Aufleuchten des Zielelementes ein P300 ausgelöst.

Das System mittelt die Ergebnisse über Zeilen und Spalten. Bei einer Vielzahl der eingesetzten Systeme geschieht dies über eine feste Anzahl von Durchläufen. Anschließend bestimmt ein linearer Klassifikator Zeile und Spalte mit der höchsten True-Positive Confidence.

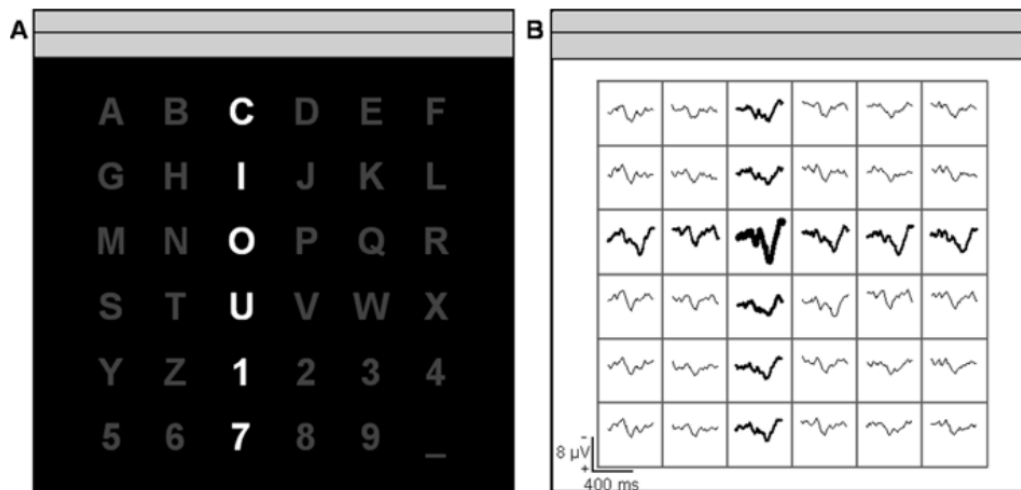


Abbildung 3.5: (A) Abbildung einer 6x6 P300 matrix. Zeilen und Spalten leuchten in zufälliger Reihenfolge auf (hier Spalte 3). (B) Gemittelte Wellenform an Elektrodenposition Pz für alle 36 Symbole der Matrix. Der Zielbuchstabe “O“ löste die stärkste P300-Antwort aus. Eine kleinere P300-Antwort ist in Spalte und Zeile 3 zu sehen. Dies rührt daher, dass diese immer gleichzeitig mit dem Zielstimuli aufleuchten. Alle anderen Wellenformen sind keine P300-Antworten. Jede Antwort ist gemittelt über 30 Stimuluspräsentationen.[Sellers et al., 2010]

Die Schnittmenge dieser beiden Zeilen und Spalten liefert das Klassifikationsergebnis (siehe Abbildung 3.5). Für die Klassifikation der Ergebnisse wird bei P300-Systemen i.d.R. eine Epoche von 800ms betrachtet.

Die verwendete Spellermatrix innerhalb von Braincomposer enthält fünf Zeilen und acht Spalten. Die Zeilen sind mit Buchstaben (a-e) und die Spalten mit Zahlen (0-7) versehen. Für jeden Klassifikationsdurchgang (Trial) werden vier Subtrials durchlaufen. Ein wichtiger Parameter der Spellermatrix ist das interstimulus-interval (ISI), der den zeitlichen Abstand zwischen zwei dargebotenen Reizen (Epochen) beschreibt. Durch ein niedriges ISI wird das System beschleunigt, wird es zu hoch gewählt, leidet die Klassifikationsgenauigkeit (Accuracy).

Braincomposer verwendet ein ISI von 140ms. In der ersten Phase wird der aktuelle Stimulus für 90ms eingeschaltet, danach für 50ms abgeschaltet.

3.3.4 Programmablauf

Die Anwendung startet zunächst im Trainingsmodus, um den verwendeten Klassifikator zu trainieren (Abbildung 3.6). In der eigentlichen Hauptanwendung kann der Anwender interaktiv zwischen Klassifikations- und Abspielmodus wechseln (Abbildung 3.7). Der gesamte Programmablauf wird detailliert in dem jeweiligen Flussdiagramm dargestellt.

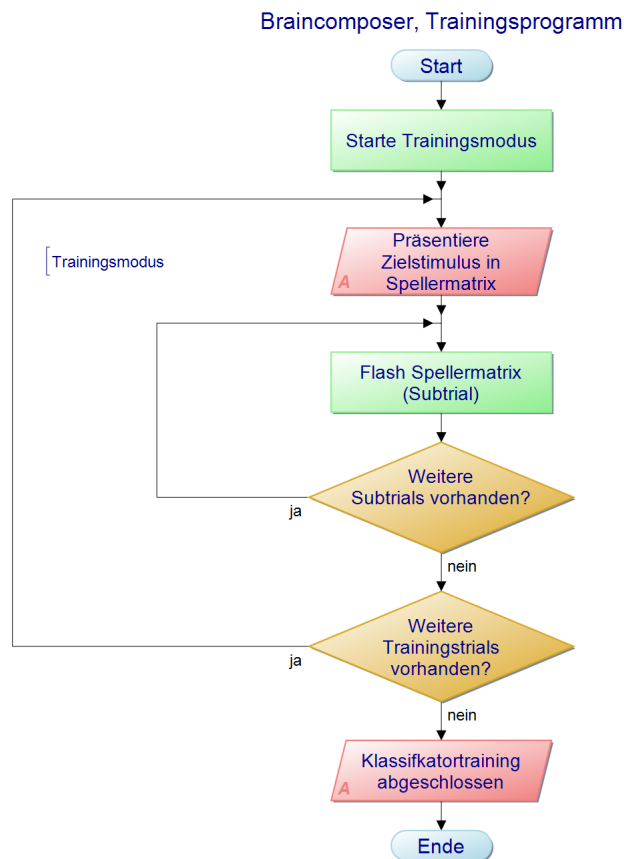


Abbildung 3.6: Braincomposer, Programmablaufplan: Training

3.4 Verwendete Technologien

In diesem Abschnitt wird beschrieben, welche Technologien verwendet wurden und wie die einzelnen Komponenten des Systems zusammenarbeiten. Hierzu bietet Abbildung 3.8 einen Gesamtüberblick.

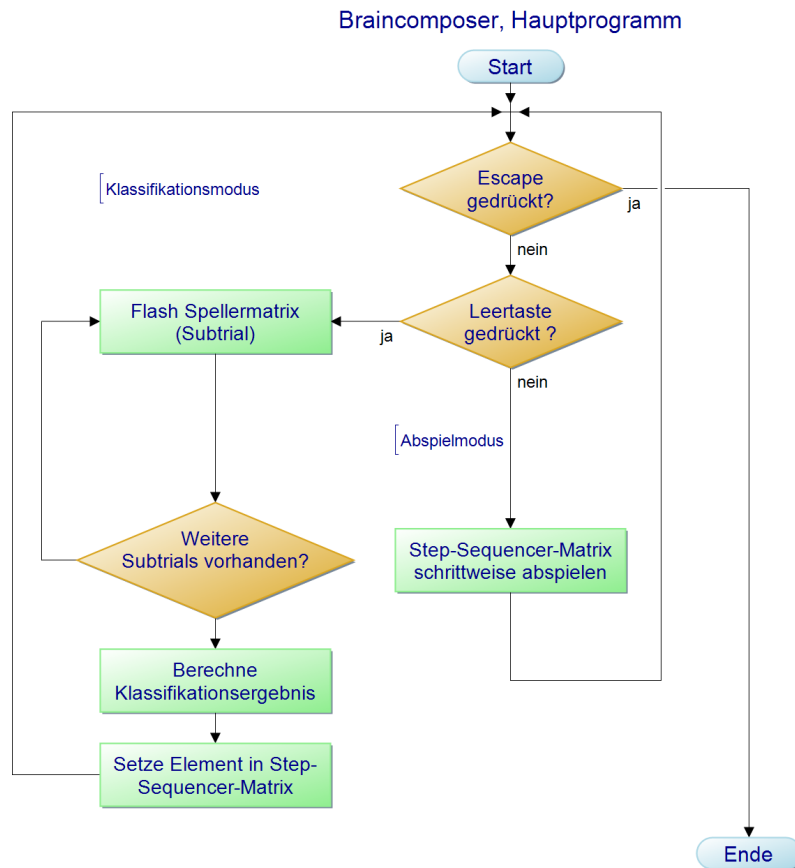


Abbildung 3.7: Braincomposer, Programmablaufplan: Hauptanwendung

3.4.1 Verwendete Software und Bibliotheken

Die Entwicklung der Anwendung Braincomposer erfolgt unter Python 2.7⁵. Diese Wahl erfolgte aus zwei Gründen:

Zum einen bietet Python ein relativ hohes Abstraktionslevel (im Vergleich zu C++ oder Java), wodurch sich vergleichsweise schnell und produktiv Programmprototypen entwickeln lassen. Zum anderen bietet das verwendete UbiCi-Framework eine Pythonschnittstelle (libubici_python). Hierfür wird eine entsprechende UbiCi-Komponente (bzw. ein Connector) erzeugt, welche den Datentransport zwischen dem UbiCi-Framework und Braincomposer regelt.

Pygame⁶ ist ein Satz von Pythonmodulen, gedacht für die Entwicklung von Computer-

⁵<http://www.python.org/>

⁶<http://www.pygame.org>

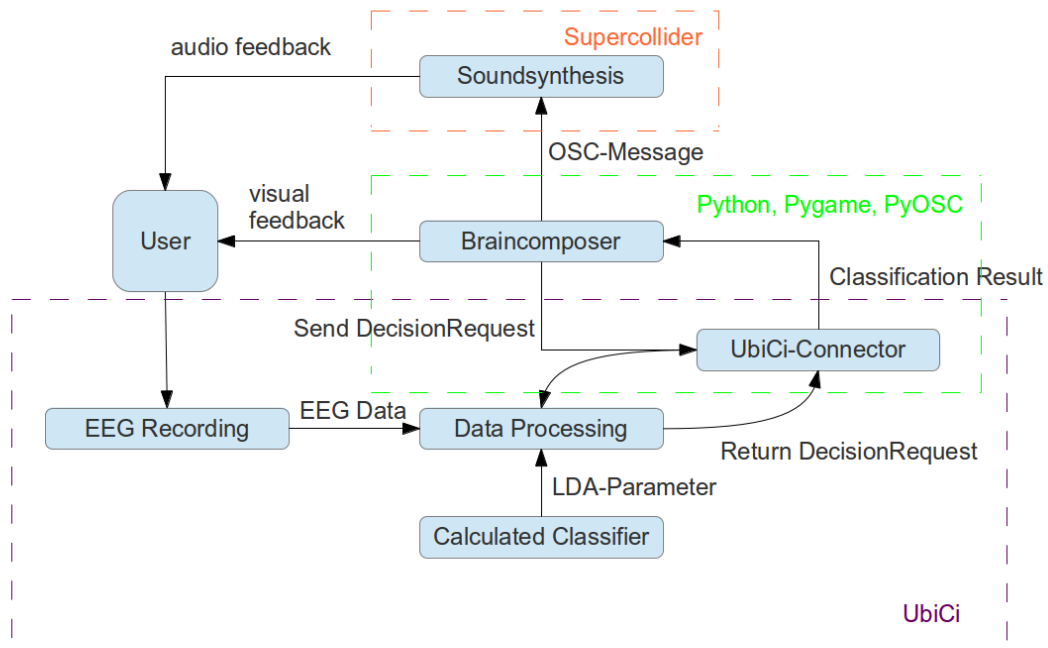


Abbildung 3.8: Braincomposer, Komponenten des Systems und deren Interaktion.

Anmerkung: Die Klassifikatorberechnung findet im Trainingsmodus statt.

spielen. Pygame baut auf der C-Bibliothek Simple Direct Media Layer (SDL)⁷ auf. Ein großer Vorteil ist, dass sich hiermit auf einem hohen Abstraktionsniveau grafische Anwendungen programmieren lassen, die portabel und performant sind. Pygame verwendet in den Kernfunktionen optimierten C und Assemblercode⁸.

Die Klänge des Step-Sequencers werden in Echtzeit synthetisiert und ausgegeben. Damit wird eine höhere Flexibilität und Erweiterbarkeit erreicht, als dies mit abgespielten Samples möglich wäre. Für die Klangsynthese wird Supercollider3.6.2 verwendet. Die Kommunikation zwischen Braincomposer und Supercollider erfolgt über das Protokoll Open-Sound-Control (OSC). Innerhalb des Pythoncodes wird das Modul PyOSC⁹ eingesetzt um OSC-Nachrichten abzusetzen.

Datenaufnahme, sowie -verarbeitung erfolgen über das UbiCi-Framework (siehe Kapitel 3.4.1.1) der Universität Bielefeld.

⁷<http://www.libsdl.org/>

⁸<http://www.pygame.org/wiki/about>

⁹<https://trac.v2.nl/wiki/pyOSC>

3.4.1.1 Das UbiCi-Framework

Das UbiCi-Framework der Uni Bielefeld wurde von Nils Hachmeister und Hannes Riechmann im Rahmen ihrer Masterarbeit entwickelt. Die Software diente hier zur Steuerung des Humanoiden Roboters iCub¹⁰, basierend auf einem hybriden Ansatz (P300 und Motor-Imagery)[Hachmeister and Riechmann, 2010]. Das Framework wurde seitdem kontinuierlich erweitert.

Der Kern des UbiCi ist die Bibliothek libubici. Hier sind die grundlegenden Komponenten zum Aufbau eines Brain-Computer-Interfaces enthalten. Sie enthält u.a. Komponenten zur Vorverarbeitung und Segmentierung von EEG-Daten, sowie Eingabe- und Ausgabekomponenten, die das Lesen und Schreiben von Dateien, sowie das Versenden und Empfangen von Daten über das Netzwerk via TCP/IP-Protokoll ermöglichen. UbiCi basiert auf QT4¹¹ und bedient sich dem “Signal and Slots-Mechanismus“. Um Matrixberechnungen zu beschleunigen, wird die Bibliothek OpenCV¹² eingesetzt.

In der UbiCi-Bibliothek “libubici_computation“ sind Komponenten zur Segmentierung, Feature-Extraktion, der Filterung und Klassifikation von EEG-Daten enthalten.

Eine weitere Bibliothek verwendete Bibliothek des UbiCi ist “libuici_python“. Diese ermöglicht durch Erzeugen einer entsprechenden Connector-Komponente, welche in Pythoncode realisiert und von einer Basiskomponente abgeleitet wird, eine bequeme Kommunikation zwischen dem UbiCi und einer Pythonanwendung. Während des Trainings und des Klassifikationsvorganges werden EEG-Daten vom UbiCi aufgezeichnet und verarbeitet. Zwischen Braincomposer und dem UbiCi besteht eine bidirektionale Verbindung. Dies ist notwendig, da zum einen die Datenverarbeitung innerhalb des UbiCi vom aktuellen Zustand des Braincomposers abhängt und zum anderen, da die Klassifikationsergebnisse zum Braincomposer zurückgeleitet werden müssen.

Wichtiger Hauptbestandteil des Frameworks ist “ubici_application“. Hier ist die Anwendung “online_bci“ enthalten, welche in Verbindung mit einer sogenannten Deployment- und einer Konfigurationsdatei aufgerufen werden. Deployments ermöglichen in einer eigenen Skriptsprache jegliche UbiCi-Komponenten flexibel zu einer Anwendung zusammenzusetzen. Diese Komponenten können detailliert in entsprechenden Konfigurationsdateien angepasst werden. Das Ergebnis ist ein leistungsfähiges und hochflexibles BCI-System.

Braincomposer benötigt mehrere (insgesamt fünf) online_bci-Programminstanzen, die z.T. miteinander über das Netzwerk kommunizieren. Der Datenaustausch zwischen den Anwendungen über das Netzwerk erfolgt via TCP/IP-Sockets.

¹⁰<http://www.icub.org/>

¹¹<http://qt.digia.com/>

¹²<http://www.opencv.org/>

3.4.2 Datenerfassung

Die Aufnahme der EEG-Daten erfolgt über den 16-kanaligen Biosignal-Verstärker G.USBamp der Firma g.Tec ¹³. Das Gerät wird hierbei über die USB-Schnittstelle mit dem Computer verbunden. Für die Aufzeichnung wird ein Linuxrechner verwendet (Distribution: Ubuntu 10.04.02)

Der Aufnahmeprozess erfolgt über das UbiCi. Das hierfür erstellte Deployment trägt die Bezeichnung **recAndSendEegData**. Der Aufnahmeprozess läuft unabhängig von Braincomposer auf einem eigenen Rechner. Auf die aufgezeichneten Daten können andere UbiCi-Prozesse über eine Netzwerkverbindung zugreifen um die Daten weiter zu verarbeiten.

Die EEG-Daten werden in dem Biosignal-Verstärker unmittelbar vor dem Versenden vorverarbeitet. Hier kommen ein Hochpassfilter mit einem low-cut von 0.1Hz und ein Notchfilter zwischen 48-52Hz zum Einsatz. Letzterer dient dazu Artefakte, die durch die Netzfrequenz im deutschen Stromnetz entstehen, zu entfernen.

Die EEG-Daten werden mit einer Samplingrate von 256Hz aufgezeichnet.

3.4.3 Datenverarbeitung

Für die weitere Verarbeitung der Daten wurden UbiCi-Deployments für die folgenden Aufgaben erzeugt:

- **trainDataReceiver:**

Empfängt Daten über das Netzwerk, wenn sich Braincomposer im Trainingsmodus befindet. Zum einen sind dies die EEG-Daten, welche "recAndSendEegData" aufzeichnet, zum anderen Informationen darüber, zu welchen Zeitpunkten welche Elemente der Spellermatrix aufgeleuchtet haben (epoch_infos), mit den entsprechenden Trainingslabel (P300, Nicht-P300). Die Daten werden gesammelt und in einer Datei zur späteren Verwendung abgespeichert.

- **calculateClassifier:**

Wurden die Daten im Trainingsmodus aufgezeichnet, werden hier die Daten segmentiert, balanciert, gefiltert, die Projektionsmatrix der PCA berechnet und die Parameter der Fisher-Diskrimante ermittelt. Die Projektionsmatrix und Eigenwerte der PCA, sowie berechneter Parameter der FDA werden abgespeichert. Diese Daten werden für die Klassifikation des Onlinesystem genutzt..

- **onlineClassifier:**

Empfängt Daten über das Netzwerk, wenn Braincomposer im Klassifikationsmo-

¹³<http://www.gtec.at/>

aus gestartet wurde. Auch hier empfangen wir wieder EEG-Daten von “recAndSendEegData“. Die Daten müssen segmentiert und gefiltert werden. Über das Deployment “braincomposerReceiverAndSender“ erfährt “onlineClassifier“, ob eine neue Klassifikation erfolgen soll (decision_request). Ist dies der Fall, so werden Informationen über den Zeitpunkt des Aufleuchtens der jeweiligen Elemente der Spellermatrix verarbeitet. Auf das entsprechende Segment wird eine Dimensionsreduktion mittels PCA durchgeführt und mit Hilfe der FDA das Klassifikationsergebnis berechnet. Das Klassifikationsergebnis wird anschließend “braincomposerReceiverAndSender“ über das Netzwerk zur Verfügung gestellt.

- **braincomposerReceiverAndSender:**

Stellt Informationen aus der Braincomposeranwendung für andere Anwendungen über das Netzwerk zur Verfügung. Hier sind Informationen darüber, zu welchen Zeitpunkten welche Elemente der Spellermatrix aufgeleuchtet haben (epoch_infos) und ob ein neuer Klassifikationsvorgang gestartet wurde (decision_request). Für einen besseren Überblick der Datenverarbeitung, sind die Deployments “calculateClassifier“ und “onlineClassifier“ in den entsprechenden Abbildungen zu sehen. (TODO)

3.4.3.1 Arbeitsschritte bei der EEG-Datenverarbeitung

In diesem Abschnitt möchte ich die wichtigsten Schritte, die bei der Verarbeitung der EEG-Daten vorgenommen werden, auflisten.

Segmentierung

Im ersten Schritt findet eine Segmentierung der Daten statt. Da die Daten auf P300-Potentiale untersucht werden, sind jene Zeitabschnitte interessant, innerhalb derer Elemente im Speller aufleuchten. Diese Informationen werden mit einem entsprechenden Zeitstempel versehen über das Netzwerk gesendet. Da in der Datenverarbeitung die EEG-Daten mit einem Zeitstempel gleichen Formats synchronisiert werden, ist nun möglich festzustellen, welche EEG-Daten mit einem bestimmten Ereignis korrespondieren. Eine gängige Segmentlänge liegt bei 800ms. Dieser Zeitraum reicht aus um die markanten Ausprägungen des P300-Potentials zu erfassen. In der vorliegenden Anwendung, mit einer Samplerate von 256Hz, werden somit 205 Samples nach Auslösen eines Ereignisses in der Spellermatrix betrachtet. Die restlichen EEG-Daten sind nicht notwendig und spielen bei der weiteren Betrachtung keine Rolle.

Da die Datenerfassung und Verarbeitung auf mehreren Computern ausgeführt wird, ist es dringend notwendig die Systemzeiten zu synchronisieren.

Balancieren

Im Trainingsmodus wird nach der Segmentierung ein balancierter Datensatz erstellt. Mit dem Balancieren der Daten wird dafür gesorgt, dass wir ein Datensatz erzeugen, der gleiche viele Datenbeispiele aus beiden verwendeten Klassen enthält. In der vorliegenden Anwendung liegt das Verhältnis von P300, zu nicht P300-Datenbeispielen bei 2:11 pro Subtrial. Für das Training des verwendeten Klassifikators würde somit bei einem nicht-balancierten Datensatz die Klasse der P300-Datenbeispiele unterrepräsentiert.

Filterung

Im folgenden Bearbeitungsschritt wird eine Datenfilterung durchgeführt. Durch die Filterung sollen ungewünschte Artefakte entfernt werden. Innerhalb der EEG-Daten haben diese unterschiedliche Entstehungsquellen. Artefakte werden u.a. verursacht durch die Netzspannung, Muskelaktivität und durch Drifts aus elektrochemischen Veränderungen an den Elektroden¹⁴

Die Filterung erfolgt auf Fourier-Basis (siehe 2.1.3.1). Das EEG-Signal wird hierdurch aus dem Zeitbereich in den Frequenzbereich transformiert. Um nun eine Hoch- bzw. Tiefpassfilterung durchzuführen, können die Frequenzkoeffizienten unter- bzw. oberhalb des entsprechenden Bereichs gelöscht werden. Es genügt diese Frequenzbereiche auf den Wert Null zu setzen. Folgt nun die Rücktransformation in den Zeitbereich, so erhalten wir ein gefiltertes Signal.

In der Datenverarbeitung wird ein FFT-Filter mit einem low-cut von 1Hz und einem high-cut von 10hz verwendet.

Dimensionsreduktion

Betrachten wir zunächst die Spaltendimension der segmentierten Daten:

Wir verwenden eine Segmentlänge von 0.8s und eine Samplingrate von 256Hz. Angenommen, wir verwenden für die Datenaufzeichnung 10 Elektroden, dann folgt:

$$\text{Spaltendimension} = 0.8s \cdot 256 \cdot 10 = 2048$$

Im nachfolgenden Klassifikationsschritt mittels FDA wird die Inverse der über beiden Klassen aufsummierten Kovarianzmatrix gebildet. Es ist sehr wahrscheinlich, dass eine derartige Matrix für höhere Dimensionen singulär und daher nicht invertierbar ist [Lenhardt et al., 2008, S.3].

Aus diesem Grund wird vor der Berechnung des Klassifikators mittels Hauptkomponentenanalyse (PCA, siehe 2.1.3.2) eine Feature-Extraction durchgeführt. Als Kriterium für die PCA wählen wir so viele Hauptkomponenten, bis 99% Gesamtvarianz abgedeckt werden. Die dazugehörige Projektionsmatrix wird auf den Trainingsdaten berechnet und abgespeichert. Im Onlinebetrieb wird sie unverändert für die Feature-Extraction genutzt.

¹⁴[Seifert, 2005], S.62ff

Klassifikation

Die Klassifikation der Segmente erfolgt anhand der Linearen Diskriminanzanalyse nach Fisher(FDA). Vor der Verwendung der FDA müssen zunächst die Parameter der FDA aus den Trainingsdaten berechnet werden. Dieser Schritt erfolgt nachdem bereits die PCA angewendet wurde.

Im Klassifikationsmodus liefert die FDA für jedes Segment ein Klassifikationsergebnis mit dazugehörigem Confidencewert. Der True-Positive-Confidence-Wert für ein P300-Potential wird für diejenigen Elemente, die zu diesem Zeitpunkt aufgeleuchtet haben, Segment für Segment aufsummiert. Am Ende eines gesamten Trials wird das Element mit dem höchsten Wert (bzw. der Gewinner), als Klassifikationsergebnis zurückgegeben.

4 Brain-Composer im Praxistest

4.1 Versuchsaufbau

Es wurden Daten von insgesamt zehn Probanden im Alter zwischen 22 und 39 Jahren erhoben. Hierbei wurden zwei Gruppen unterschieden: Musiker und Nichtmusiker. Zu jeder Gruppe zählten fünf Probanden. Zu der Gruppe der Musiker wurden in der Studie jene Probanden gezählt, die mindestens ein Jahr Erfahrung auf einem Musikinstrument besitzen und grundlegende Kenntnisse in der Harmonielehre, vorallem im Notenlesen und dem Aufbau von Dreikängen, besitzen.

Der Ablauf des Experiments bestand aus folgenden Schritten:

- Anbringung der Elektroden
- Instruieren des Probanden
- Aufnehmen der Trainingsdaten zur Klassifikatorgewinnung
- Braincomposer verwenden
- Ausfüllen eines Fragebogens

Die Messung der Hirnströme wurde via EEG über einen Biosignalverstärker der Firma g.Tec (siehe 3.4.2) auf zehn Kanälen vorgenommen. Die Aufzeichnung erfolgte in einem Raum, der weder elektromagnetisch, noch von anderen äußeren Einflüssen in besondere Weise abgeschirmt ist. Auf einer EEG-Kappe der Firma easycap¹⁵ wurden Ringelektroden angebracht. Dem 10-20-System folgend (siehe 2.1.2.1.1) wurden die Elektroden an den Positionen FZ, CZ, PZ, OZ, P3, C3, P4, C4, PO7, PO8 platziert. Um einen besseren Kontakt zwischen Elektrode und Kopfhaut herzustellen, musste die Haut mit Isopropanol vorgereinigt und anschließend mit einem abrasiven Gel behandelt werden. Bei allen Versuchspersonen lagen die Impedanzen der Elektroden unterhalb von 2k Ω . Die verwendete Samplerate des Biosignalverstärkers betrug 256Hz.

Jedee Proband absolvierte zunächst eine 14-minütige Trainingssitzung, bestehend aus 40 Trials mit je 5 Subtrials, um die Parameter für den Klassifikator zu gewinnen. Das interstimulus-interval (ISI) wurde auf 140ms gesetzt, wobei jeder Stimulus für eine Zeit von 70ms aktiv präsentiert und für 70ms abgeschaltet wurde.

Vor der Aufzeichnung der Hirnströme wurden die Probanden angewiesen, Bewegungen und Augenblinzeln während der Präsentation der Stimuli zu vermeiden um Artefakte zu reduzieren.

Für die Ausführung des Programmes Braincomposer kam ein Dell Latitude D630 (Intel Core2Duo T7500 mit 2.2Ghz, Arbeitsspeicher von 2GB) zum Einsatz. Die Präsentation

¹⁵<http://www.easycap.de>

der grafischen Oberfläche erfolgte auf einem externen TFT-Monitor (Dell 2407wfp) mit einer Diagonale von 24 und einer Auflösung von 1920x1200 Pixel.

Die Aufzeichnung und Verarbeitung der Daten im ersten Teil des Versuchs wurde auf zwei verschiedenen Workstations ausgeführt, die über das Netzwerk miteinander kommunizierten. Für die Aufzeichnung kam ein Dell Precision T3400 (Core2Duo E6750 mit 2,66 Ghz, Arbeitsspeicher 4GB), für die Speicherung und Verarbeitung ein Dell T390 (Core2 6600 mit 2,4Ghz, Arbeitsspeicher 2GB) zum Einsatz. Vor dem Start des Experiments musste dafür gesorgt werden, dass die Zeiten aller beteiligten Computer bis auf wenige Millisekunden mit dem der Zeitserver der Universität Bielefeld (stratum2-1.ntp.techfak.uni-bielefeld.de) synchronisiert wurden.

Aufgrund von externen Netzwerkproblemen, was zu einer fehlerhaften Zeitsynchronisation und unbrauchbaren Daten führte, wurden Speicherung und Verarbeitung auf den Rechner, der ursprünglich nur für die Aufzeichnung vorgesehen war, verlagert. Dies geschah in der Versuchsreihe ab Proband 8.

Die Präsentation der Audiowiedergaben erfolgte über einen einzelnen Aktivlautsprecher mit der Bezeichnung "Roland Cakewalk MA-15d". Der Lautstärkepegel wurde mit dem Probanden individuell abgestimmt.

4.2 Fragestellungen

Vor Beginn der Experimente wurden Fragestellungen formuliert, über die die Studie Aufschluss geben sollte:

- Wie gut ist das System für den Anwender nutzbar?
- Funktioniert das Konzept "Musikalische Komposition über einen P300-Speller" zufriedenstellend?
- Gibt es einen signifikanten Unterschied zwischen Musikern und Nicht-Musikern bei der Nutzung des System und in der Komplexität des musikalischen Resultates gibt ?

Der subjektive Eindruck der Probanden wurde nach dem Experiment in einem Fragebogen festgehalten (siehe Anhang).

4.3 Ergebnisse

4.3.1 Quantitative Ergebnisse

4.3.1.1 Klassifikationsgenauigkeit

Die Klassifikationsgenauigkeit des Systems kann in der Tabelle 4.1 abgelesen werden.

	\bar{x}	σ^2
Musiker	0.945	0.0072
Nicht-Musiker	0.852	0.1432
Beide Gruppen	0.899	0.016

Tabelle 4.1: Klassifikationsgenauigkeit des P3 ERP. Die Ergebnisse wurden über die einzelnen Gruppen und über alle Probanden gemittelt. Die Klassifikationsgenauigkeit per Zufall liegt bei 1/40 (0.025).

Die Ergebnisse zwischen Musikern und Nichtmusikern differieren leicht, unterscheiden sich jedoch nicht signifikant (P-Wert von 0.297).

Die Software lief verteilt auf drei Computern. Durch Netzwerkprobleme wurde das Ergebnis der Klassifikation geschmälert. Daher wurden ab VP8 nur noch zwei Computer eingesetzt, bei denen das Netzwerk fehlerfrei funktionierte. Die Klassifikationsrate verbesserte sich bei den folgenden sechs Probanden und lag im Mittel bei ca. 0.976 (beide Gruppen). Die Hälfte der Probanden erreichten sogar eine Klassifikationsgenauigkeit von 1.

4.3.1.2 Übertragungsgeschwindigkeit

Die Übertragungsgeschwindigkeit ist eine wichtige Kenngröße um die Leistungsfähigkeit eines BCIs einordnen zu können. Die Angabe erfolgt üblicherweise in *bit/s*. Für die vorliegende Anwendung muss zunächst die mittlere Datenübertragungsrate pro Trial ermittelt werden [Wolpaw et al., 2000]:

$$B = \log_2(N) + p \cdot \log_2(p) + (1 - p) \cdot \log_2\left(\frac{1 - p}{N - 1}\right) \quad (4.1)$$

Wobei N die Anzahl auswählbarer Elemente pro Trial und p die Klassifikationsgenauigkeit ausdrücken.

Die Anzahl der Elemente beträgt $N = 40$. Betrachten wir die Ergebnisse aller Versuchspersonen, ohne Trennung der beiden Gruppen ($p = 0.899$), so erhalten wir mit Formel 4.1 ca. 5.184 *bits/trial*.

Die theoretische maximale Übertragungsgeschwindigkeit in bit/s bei einem Speller berechnet sich folglich durch:

$$P = \frac{60 \cdot B}{(\text{zeilen} + \text{spalten}) \cdot \text{ISI} \cdot \text{substrials}} \quad (4.2)$$

In der vorliegenden Anwendung verwenden wir 5 Zeilen, 8 Spalten, eine ISI von 140ms und 5 substrials. Damit erhalten wir eine maximale Datenübertragungsrate von 34.18 *bit/s*. Es ist zu beachten, dass diese Zahl die maximale Datenübertragungsrate ausdrückt. In

der Praxis wird diese jedoch nicht erreicht, da der Anwender sich zwischen den Eingaben das Klangergebnis anhört, es beurteilt und sich Gedanken über die nächste Eingabe macht. Für die Leistungsfähigkeit ist in diesem Falle die Datenübertragungsrate pro Trial ein sinnvollerer Maß.

Die tatsächliche Datenübertragungsrate im Experiment über beide Gruppen lag im Mittel bei 14,69 *bit/s*.

4.3.1.3 Eingaben und Erzeugte Tonstrukturen

Um Unterschiede im Klangergebnis der zu untersuchenden Gruppen auf quantitativer Basis festzustellen, wurden die Anzahl der Eingaben und die erzeugten Tonstrukturen über beide Gruppen gemittelt. In der Tabelle 4.2 sind die Ergebnisse dargestellt.

Ein signifikanter Unterschied lässt sich hier bei der Anzahl der Eingaben feststellen (P-Wert von 0.0153). Die Musiker machten im Schnitt ungefähr 12 Eingaben weniger als die Nicht-Musiker. Trotz weniger Eingaben haben die Musiker am Ende im Schnitt drei Töne mehr in der Komposition verwendet.

4.3.1.4 Fragebogen

Die Probanden wurden gebeten nach dem Experiment einen Fragebogen auszufüllen. Die zusammengefassten Ergebnisse der geschlossenen Fragen sind in der Tabelle 4.3 dargestellt.

Bei der Beantwortung der geschlossenen Fragen konnte kein signifikanter Unterschied zwischen den Gruppen festgestellt werden.

4.3.2 Qualitative Ergebnisse

4.3.2.1 Fragebogen

Der Fragebogen umfasste neben acht geschlossenen auch zwei offene Fragen:

- Was hat dir besonders gut gefallen?
- Was hat dir weniger gut gefallen?

Positiv empfanden viele Anwender die hohe Klassifikationsgenauigkeit (4 Nennungen). Auch das Konzept, Musik über die Gedanken zu erzeugen fand positiv Erwähnung (2 Nennungen).

Auf der negativen Seite ist zu nennen, dass vielen Anwendern die lange Trainingsphase nicht gefiel (4 Nennungen), oder aber das anbringen der Elektroden (2 Nennungen).

4.3.2.2 Beobachtungen

Während die Versuchspersonen das Programm für die Komposition einsetzten, wurden sie beobachtet.

Auffällig war in diesem Zusammenhang, dass Musiker mehr Geduld mit dem System hatten. Sie hörten sich ihre Eingaben sehr genau an und planten welchen Ton sie als nächstes eingeben wollten. Hierbei konnte man oftmals beobachten wie sie das bisherige Ergebnis mehrfach in der Schleife spielten, aufmerksam zuhörten und mit einem Finger zu dem Zeitpunkt auf die Tischplatte klopfen, an dem sie den nächsten Ton planten. Es hatte den Anschein, die Musiker würden ihre Eingabe zunächst im Kopf simulieren und bewerten. Bei den Nicht-Musikern erfolgte die Eingabe eher willkürlich. Hier war auffällig, dass sich der neu eingegebene Ton teilweise gar nicht angehört wurde, bevor die nächste Eingabe vorgenommen wurde. Die Nicht-Musiker hatten eine schnellere Interaktion mit dem System, die jedoch weniger gezielt erschien.

Viele Musiker wirkten motiviert und ehrgeizig. Sie wollten ein möglichst gutes Ergebnis produzieren und dieses auch noch nach der offiziellen Zeit von 15 Minuten zu Ende bringen. Bei den Nicht-Musikern war von diesem Ehrgeiz weniger zu spüren.

4.3.3 Interpretation der Ergebnisse

Das System wurde von den Anwender sehr positiv aufgenommen. Das Komponieren bereitete allen Anwendern Spaß, den meisten sogar in hohem Maße. Dies spiegelt sich im Fragebogen in den Antworten auf die geschlossenen Fragen 05, 06, 07 und 08 wider. Hierzu steuert vor allem auch die hohe Klassifikationsgenauigkeit bei. Die Klassifikation der Eingaben funktionierte zuverlässig auf einem hohen Niveau und pendelte sich im Mittel bei 89,9% ein. Bei vier Probanden lag sie sogar bei 100%. Aus Sicht der Anwender wurde dies als höchst positiv bewertet (siehe 4.3, Frage 01). Einige Probanden waren sogar überrascht über die hohe Genauigkeit mit dem das System die Eingaben korrekt erkannte, wie aus den Antworten auf die offenen Fragen hervorgeht.

Die Gründe für die insgesamt positive Bewertung hängt möglicherweise damit zusammen, dass BCIs noch wenig verbreitet sind und es selten die Möglichkeit gibt mit einer darartigen Anwendung zu "spielen".

Die Praxistauglichkeit des grundlegenden Konzeptes, einen P300-Speller zur musikalischen Komposition mit einem Step-Sequencer zusammen zu bringen, konnte in der Studie nachgewiesen werden. Die Klänge konnten aus Sicht der Probanden deutlich nach den eigenen klanglichen Vorstellungen beeinflusst werden und waren kein Zufallsprodukt (Fragebogen, Frage 03 und 04).

Vergleicht man die Kompositionen von Musikern und Nichtmusikern auf quantitativer Basis, so lassen sich weniger signifikante Unterschiede finden als erwartet. Hier sticht jedoch ein Merkmal heraus: Nicht-Musiker senden deutlich mehr Eingaben an das System als die Musiker. Im Endeffekt erzeugen jedoch die Musiker eine Komposition die insgesamt mehr Töne besitzt (im Mittel 3 mehr). Die Nicht-Musiker korrigieren also deutlich häufiger ihre Eingaben, in dem Töne wieder gelöscht werden. Nicht-Musiker machten im Schnitt ca 10 Eingabekorrekturen, Musiker hingegen nur ca. 3. Daraus resultiert, dass die Musiker bei weniger Eingaben eine höhere Effizienz bei der Nutzung

des Systems erreichen. Dies deckt sich mit den unter 4.3.2.2 gemachten Beobachtungen. Die Musiker gehen langsamer und zielgerichteter vor und planen ihre Eingaben länger. Damit erreichen sie eine höherer Trefferquote bei der Umsetzung ihrer Klangvorstellung.

Als kritischer Punkt muss der Trainingsvorgang gesehen werden. Die Probanden verbalisierten dies in den offenen Fragen der Fragebögen. Das Training wurde als zu lang und anstrengend empfunden.

Erzeugte Klangstrukturen					
	Eingaben ins System	verwendete Töne	Polyphone Klänge	Matrizen	Töne pro Matrix
Musiker	36.8 ± 4.58	32 ± 4.86	8 ± 1.41	3.8 ± 0.748	8.56 ± 1.13
Nicht-Musiker	48.2 ± 5.85	29 ± 7.51	5.8 ± 2.93	3.6 ± 0.49	8.24 ± 2.72

Tabelle 4.2: Anzahl der erzeugten Tonstrukturen, gemittelt über beide Gruppen. Die Spalte "Eingaben ins System" gibt die Menge der tatsächlich übertragenen Töne an, die Spalte "verwendete Töne" die Anzahl der Töne die nicht wieder gelöscht wurden

Fragebogen			
	Musiker	Nicht-Musiker	beide Gruppen
01 Das System hat meine Eingaben korrekt erkannt	4.8 ± 0.4	4.4 ± 0.8	4.6 ± 0.663
02 Die Interaktion mit dem System hat mich angestrengt	2 ± 0.633	2.2 ± 0.748	2.1 ± 0.7
03 Ich konnte das Klangergebnis nach meinen Vorstellungen beeinflussen	4.8 ± 0.4	4 ± 1.095	4.4 ± 0.916
04 Das Klangergebnis war ein Zufallsprodukt	1.2 ± 0.4	1.4 ± 0.8	1.3 ± 0.64
05 Das erzeugte Klangergebnis hat mir gefallen	4.6 ± 0.49	4 ± 0.633	4.3 ± 0.64
06 Die Benutzung des Systems hat mir Spaß gemacht	5 ± 0	4.4 ± 0.49	4.7 ± 0.458
07 Das System hat mich inspiriert	4.6 ± 0.49	4 ± 0.89	4.3 ± 0.781
08 Ich könnte mir vorstellen das System zur Unterhaltung zu verwenden	4 ± 0.63	3.8 ± 1.17	3.9 ± 0.943

Tabelle 4.3: Ergebnisse des Fragebogens. Die Ergebnisse wurden über die einzelnen Gruppen und über alle Probanden gemittelt. Die Bewertungsskala reicht von 1 bis 5, mit 1 für “überhaupt nicht“ und 5 für “in hohem Maße“

5 Zusammenfassung

In der vorliegenden Arbeit wurde ein Konzept erarbeitet, dass eine musikalische Komposition auf Basis eines P300-Spellings ermöglicht. Anschließend wurde das Konzept technisch umgesetzt.

In einer Studie wurde das konstruierte System einem Praxistest unterzogen und die Ergebnisse auf verschiedenen Ebenen evaluiert. Besondere Aufmerksamkeit wurde hierbei der unterschiedlichen Herangehensweise und Verwendung des Systems von Musikern und Nicht-Musikern gewidmet.

Bei der Entwicklung des Konzeptes stand im Vordergrund, dass die erzeugte Komposition gezielt beeinflusst werden, Polyphonie ermöglicht und dem Anwender ein gewisser Rahmen an musikalischer Komplexität zur Verfügung stehen sollte. Zudem sollte, im Vergleich zu bisherigen BCI-Anwendungen mit ähnlicher Zielvorgabe, ein möglichst ansprechendes musikalisches Ergebnis erzeugt werden.

In der Benutzerstudie wurde das System von fünf Musikern und fünf Nicht-Musikern für einen Zeitraum von 15 Minuten für die Musikkomposition genutzt. Das System arbeitete zuverlässig und erkannte die Eingaben der Anwender mit einer hohen Genauigkeit.

Die Praxistauglichkeit konnte in der Studie nachgewiesen werden. Die Probanden gaben an, die Komposition gezielt nach den eigenen Vorstellungen beeinflussen zu können. Darüberhinaus machte die Verwendung des Systems allen Anwendern Spaß. Die Probanden fanden zudem gefallen an ihren Kompositionen.

Die Strukturen der erzeugten klanglichen Ergebnisse von Musikern und Nicht-Musikern wurden auf quantitativer Basis verglichen. Es ließ sich feststellen, dass Musiker sich für die Auswahl des einzugebenden Tones mehr Zeit ließen als Nicht-Musiker. Nicht-Musiker machten in dem gleichen Zeitraum mehr Eingaben, korrigierten ihre Eingaben jedoch auch deutlich häufiger, indem Töne wieder gelöscht wurden. Insgesamt erzeugten die Musiker eine Komposition mit mehr Tönen, bei weniger Eingaben.

Daraus lässt sich schlussfolgern, dass die Musiker ihre Vorstellungen bewusster und gezielter umsetzen konnten und daher weniger Korrekturen benötigten.

5.1 Ausblick

Die vorgestellte Software wurde bewusst einfach gehalten um das grundlegende Konzept zu testen. Auf Basis der vorliegenden Arbeit könnte das Konzept bzw. das System in Zukunft weiterentwickelt werden. Hierbei könnten Entwicklungen in drei verschiedene Richtungen erfolgen:

- BCI-Performanz
- Funktionalität und klangliche Möglichkeiten

- Handhab- und Nutzbarkeit

5.1.1 BCI-Performanz

Im Onlinemodus ließe sich die Datenübertragungsrate des Spellers durch Implementierung einer Dynamische-Subtrial-Limitierung erhöhen [Lenhardt et al., 2008]. Das korrekte Klassifikationsergebnis kann bereits nach wenigen Subtrials feststehen, so dass ein Durchlauf weiterer Subtrials nicht mehr notwendig ist. Für jeden Anwender wird hierfür aus den Trainingsdaten ein Schwellwert berechnet. Wird bei der Dateneingabe im Onlinemodus dieser Schwellwert erreicht, so müssen keine weiteren Subtrials durchlaufen werden.

Der eingesetzte Speller verwendet eine zufällige Reihenfolge aufleuchtender Spalten und Zeilen. Dies kann dazu führen, dass zwei Zielreize unmittelbar nacheinander aufleuchten. Beim zweiten Zielreiz führt dies zu einer P300 mit geringerer Amplitude. Daher ist die Implementierung einer Pseudorandomisierung vorzuziehen, bei der nach einem Zielreiz immer mindestens zwei Hintergrundreize folgen [Gerloff, 2005, S.516]. Hierdurch könnte die Accuracy verbessert werden.

5.1.2 Funktionalität und klangliche Möglichkeiten

Der Matrix des Step-Sequencers ließe sich unkompliziert auf eine Größe von $16 \cdot 16$ Elemente erweitern. Diese könnte in vier unterschiedliche $8 \cdot 8$ - Abschnitte (z.B. markiert durch unterschiedliche Farben) unterteilt werden. Der Anwender hätte nun die Möglichkeit eine dieser Matrizen durch eine Vorauswahl zu selektieren und kann anschließend mittels einem $8 \cdot 8$ - Speller Eingaben vornehmen. Dies würde die musikalische Komplexität stark erhöhen, bei gleichzeitig guter Handhabbarkeit.

Um das Klangergebnis feiner zu gestalten könnte der Anwender unter mehreren Instrumenten, welches für die Klanguisgabe benutzt werden soll. Eine Anpassung der Geschwindigkeit ließe sich ebenfalls unkompliziert umsetzen. Für Musiker könnte es interessant sein, auszuwählen, welches Tonsystem dem Step-Sequencer zu Grunde liegen soll.

Weiter könnte das Konzept in Richtung "parallele Matrizen" entwickelt werden. Hiermit würde sich die Möglichkeit eröffnen, eine Komposition mit mehreren parallel spielenden Instrumenten zu erzeugen.

5.1.3 Handhab- und Nutzbarkeit

Der bequeme und unkomplizierte Einsatz für den Anwender spielt eine nicht zu unterschätzende Rolle.

Die Erkennung der Eingaben war sehr zuverlässig. Dies ging jedoch mit einer Trainingsphase einher, die viele Probanden als lang und anstrengend empfanden. Die Trainingsphase bestand aus 40 Trials und 5 Subtrials, was zu einer Trainingsdauer von ca. 14 Minuten führte. Hier könnten Überlegungen angestellt werden, in welchem Umfang die Trainingsphase verkürzt werden könnte.

In der aktuellen Studie wurde mit 10 Elektroden gearbeitet, was die Probanden als wenig angenehm empfanden. Die Anbringung der Elektroden nahm ca. 25-30 Minuten in Anspruch. Es ist zu überlegen, ob die Anzahl der Elektroden reduziert werden kann, ohne dass die Klassifikationsgenauigkeit hierunter zu stark leidet.

Der Verwendung von Braincomposer mit einem EEG-Headset könnte getestet werden. Sofern hier akzeptable Klassifikationsergebnisse erreicht werden, könnte die Anwendung unkompliziert auf Veranstaltungen oder Messen von Interessenten ausprobiert werden.

A Abkürzungsverzeichnis

ADS	Aufmerksamkeitsdefizitsyndrom, page 5
AEP	akustisch evozierte Potentiale, page 11
BCI	Brain-Computer Interface, page 5
BOLD	blood oxygen level dependent, page 10
Ecog	Elektrokortigrafie, page 10
EEG	Elektroenzephalografie, page 7
ERP	Ereigniskorrelierten Potentialen, page 11
FFT	Fast-Fourier-Transformantion, page 16
FLDA	Fisherdiskrimanzanalyse, page 17
FMRI	Funktionelle Magnetresonanztomographie, page 10
ISI	interstimulus-interval, page 26
LFP	Local Field Potentials, page 10
MEA	Multielektrodenarrays, page 10
MEG	Magnetenzephalografie, page 9
NIRS	Nahinfrarotspektroskopie, page 10
PCA	Principle Component Analysis, page 17
SCP	Slow Cortical Potentials, page 13
SEP	somatosensorische evozierte Potentiale, page 11
VEP	visuell evozierte Potentiale, page 11

B Abbildungsverzeichnis

2.1	Komponenten eines BCI	7
2.2	Pyramidzellen	8
2.3	10-20 System	9
2.4	endogen/exogen evozierte Potentiale	12
2.5	P300 Paradigmen	14
2.6	ERP-Plots	15
2.7	FFT, Filterkoeffizienten	16
2.8	PCA-Beispiel	18
2.9	Visualisierung der Fisher-Diskriminante	19
3.1	Braincomposer GUI	23
3.2	Intervalle und ihr Spannungsgrad	24
3.3	Herleitung der anhemitonischen Pentatonik	25
3.4	Anhemitonische Pentatonik	25
3.5	P300-Speller	26
3.6	Braincomposer: PAP, Training	27
3.7	Braincomposer: PAP, Hauptanwendung	28
3.8	Braincomposer, Komponenten des Systems und deren Interaktion. Anmerkung: Die Klassifikatorberechnung findet im Trainingsmodus statt.	29
E.1	Komposition VP1 (Musiker)	52
E.2	Komposition VP2 (Nicht-Musiker)	53
E.3	Komposition VP3 (Musiker)	54
E.4	Komposition VP4 (Musiker)	55
E.5	Komposition VP5 (Nicht-Musiker)	56
E.6	Komposition VP7 (Nicht-Musiker)	57
E.7	Komposition VP8 (Musiker)	58
E.8	Komposition VP9 (Nicht-Musiker)	59
E.9	Komposition VP10 (Nicht-Musiker)	60
E.10	Komposition VP12 (Musiker)	61

C Tabellenverzeichnis

4.1	Braincomposer-Ergebnisse: P3-Accuracy	37
4.2	Braincomposerergebnisse: Anzahl der erzeugten Tonstrukturen	41

4.3	Braincomposerergebnisse: Fragebogen	42
-----	---	----

D Literaturverzeichnis

Literatur

- [Amon, 2011] Amon, R. (2011). *Lexikon der Harmonielehre*. Doblinger/Metzler.
- [Bennett and O’Modhrain, 2008] Bennett, P. and O’Modhrain, S. (2008). The beatbearing: a tangible rhythm sequencer. *Proc. of NordiCHI*. Vol. 2008.
- [Bishop, 2006] Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer.
- [Butz, 2005] Butz, T. (2005). *Fouriertransformation für Fußgänger*. Lehrbuch : Mathematik. Teubner.
- [Chew and Caspary, 2011] Chew, Y. C. and Caspary, E. (2011). Museegk: Design of a bcmi. *Proceedings of the 8th ACM conference on Creativity and cognition*, pages 325–326.
- [Clausen, 2010] Clausen, J. (2010). *Technik im Gehirn: Ethische, theoretische und historische Aspekte moderner Neurotechnologie*. Deutscher Ärzte-Verlag.
- [Deecke and Lang, 1988] Deecke, L. and Lang, W. (1988). P-300 as the resolution of negative cortical dc shifts. *Behav Brain Sci*, 11:379–381.
- [Donchin and Coles, 1988] Donchin, E. and Coles, M. (1988). Is the p-300 component a manifestation of contex updating? *Behav Brain Sci*, 11:357–374.
- [Duncan-Johnson and Donchin, 2007] Duncan-Johnson, C. C. and Donchin, E. (2007). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, 14(5):456–467.
- [Farwell and Donchin, 1988] Farwell, L. A. and Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 70:510–523.
- [Gerloff, 2005] Gerloff, C. (2005). Ereignis-korrelierte potenziale (ekp). *Evozierte Potenziale*, pages 499–537.
- [Graumann et al., 2010] Graumann, B., Allison, B., and Pfurtscheller, G. (2010). Brain-computer interfaces: A gentle introduction. *Brain-Computer Interfaces*.
- [Hachmeister and Riechmann, 2010] Hachmeister, N. and Riechmann, H. (2010). Implementation of a hybrid brain-robot interface for manipulating the humanoid icub robot. Master’s thesis, Universität Bielefeld.

- [Hamadicharef et al., 2010] Hamadicharef, B., Xu, M., and Aditya, S. (2010). Brain-computer interface (bci) based musical composition. *CW '10 Proceedings of the 2010 International Conference on Cyberworlds*, pages 282–286.
- [Hochberg et al., 2012] Hochberg, L. R., Bacher, D., Jarosiewicz, B., Masse, N. Y., Simeral, J. D., Vogel, J., Haddadin, S., Liu, J., Cash, S. S., van der Smagt, P., et al. (2012). Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485(7398):372–375.
- [Iturrate et al., 2009] Iturrate, I., Antelis, J., Kübler, A., and Minguez, J. (2009). Non-invasive brain-actuated wheelchair based on a p300 neurophysiological protocol and automated navigation. *IEEE Trans Robot*, 25(2):367–381.
- [Lenhardt et al., 2008] Lenhardt, A., Kaper, M., and Ritter, H. J. (2008). An adaptive p300-based online brain-computer interface. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 16:121–130.
- [Luck, 2005] Luck, S. J. (2005). *An introduction to the event-related potential technique*.
- [Lötscher, 2003] Lötscher, M. (2003). Musik am Computer, Entstehung und Bedeutung des Tonstudios im Schlafzimmer. *Diplomarbeit*, pages 6–7.
- [Muenssinger et al., 2010] Muenssinger, J. I., Halder, S., Kleih, S. C., Furdea, A., Raco, V., Hoesle, A., and Kübler, A. (2010). Brain painting: first evaluation of a new brain-computer interface application with als-patients and healthy volunteers. *Frontiers in neuroscience*, 4.
- [Nijholt and Tan, 2008] Nijholt, A. and Tan, D. (2008). Brain-computer interfacing for intelligent systems. *Intelligent Systems, IEEE*, 23(3):72–79.
- [Nishibori and Iwai, 2006] Nishibori, Y. and Iwai, T. (2006). Tenori-on. *Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NI-ME06), Paris, France*, pages 172–175.
- [Oostenveld and Praamstra, 2001] Oostenveld, R. and Praamstra, P. (2001). The five percent electrode system for high-resolution eeg and erp measurements. *Clinical Neurophysiology*, 112(4):713–719.
- [Polich et al., 2006] Polich, J., Criado, J. R., et al. (2006). Neuropsychology and neuropharmacology of p3a and p3b. *International Journal of Psychophysiology*, 60(2):172–185.
- [Seifert, 2005] Seifert, J. (2005). *Ereigniskorrelierte EEG-Aktivität*. Pabst Science Publishers.

- [Sellers et al., 2010] Sellers, E. W., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R. (2010). Bcis in the laboratory and at home: The wadsworth research program. *Brain-Computer Interfaces*, pages 97–111.
- [Sikora, 2003] Sikora, F. (2003). *Neue Jazz-Harmonielehre*. Schott Musik International, Mainz.
- [Sutton et al., 1965] Sutton, S., Braren, M., Zubin, J., John, E., et al. (1965). Evoked-potential correlates of stimulus uncertainty. *Science (New York, NY)*, 150(700):1187.
- [Tomori and Moore, 2003] Tomori, O. and Moore, M. (2003). The neurally controllable internet browser (brainbrowser). *CHI'03 extended abstracts on Human factors in computing systems*, pages 796–797.
- [Wolpaw et al., 2000] Wolpaw, J. R., Birbaumer, N., Heetderks, W. J., McFarland, D. J., Peckham, P. H., Schalk, G., Donchin, E., Quatrano, L. A., Robinson, C. J., Vaughan, T. M., et al. (2000). Brain-computer interface technology: a review of the first international meeting. *IEEE transactions on rehabilitation engineering*, 8(2):164–173.
- [Wolpaw et al., 2002] Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., Vaughan, T. M., et al. (2002). Brain-computer interfaces for communication and control. *Clinical neurophysiology*, 113, No.6:767–791.
- [Yoo et al., 2004] Yoo, S.-S., Fairney, T., Chen, N.-K., Choo, S.-E., Panych, L. P., Park, H., Lee, S.-Y., and Jolesz, F. A. (2004). Brain-computer interface using fm-ri: spatial navigation by thoughts. *Neuroreport*, 15(10):1591–1595.

E Kompositionen der Probanden

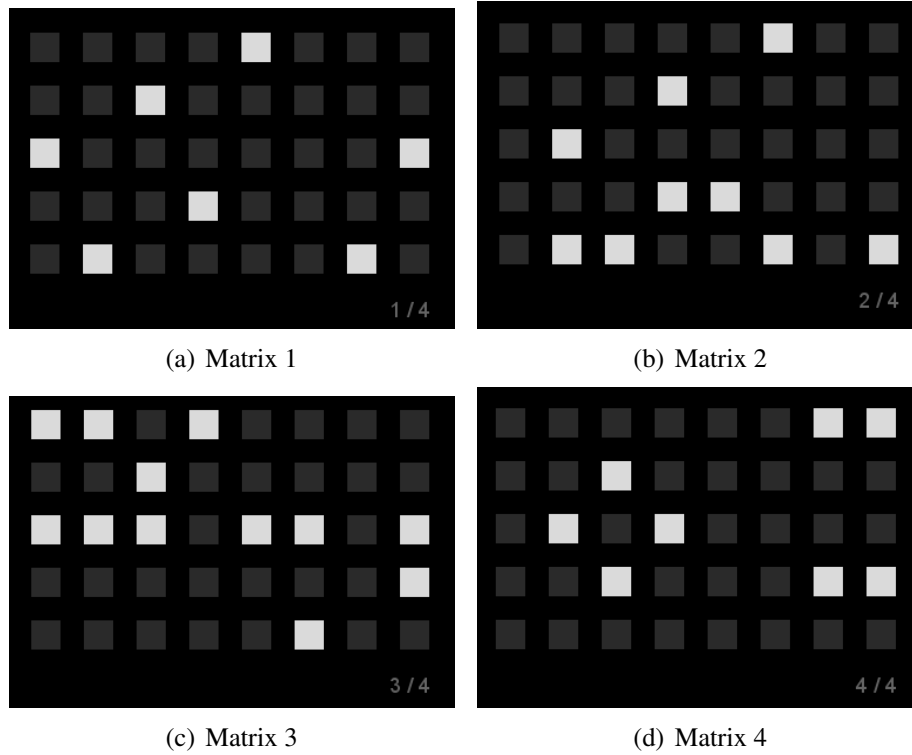


Abbildung E.1: Komposition VP1 (Musiker)

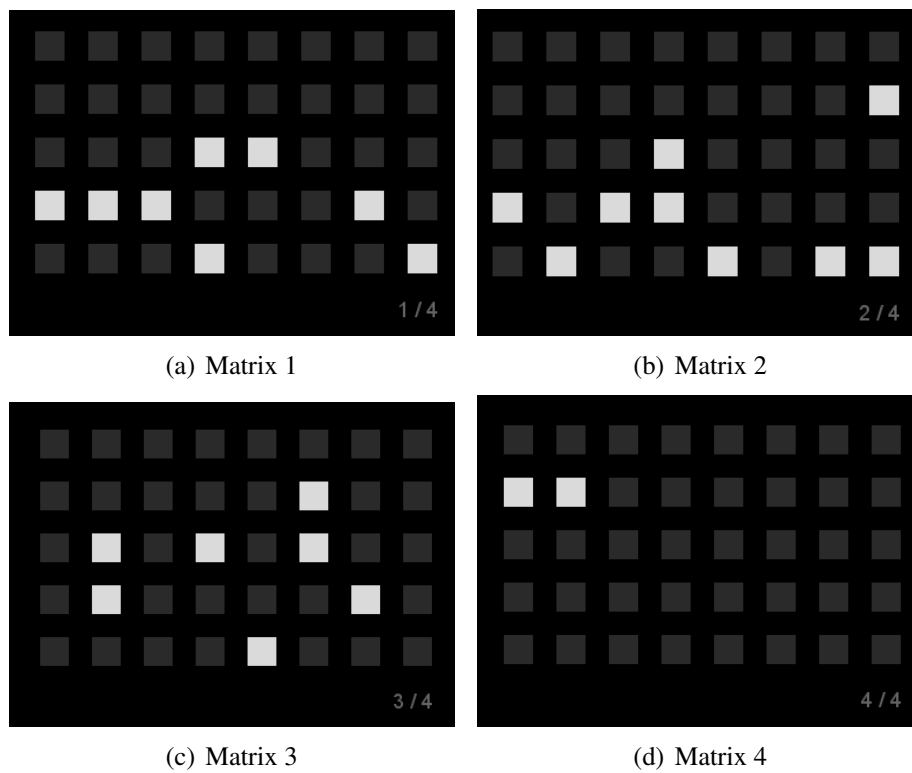
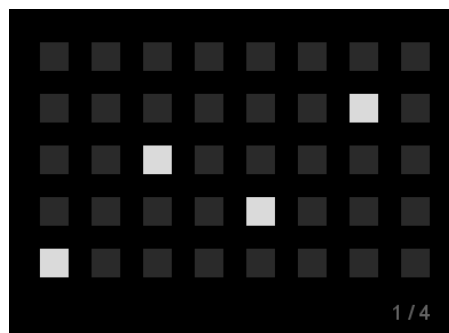
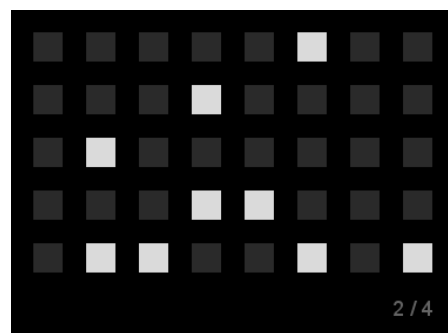


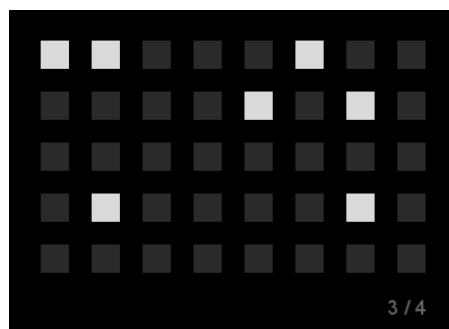
Abbildung E.2: Komposition VP2 (Nicht-Musiker)



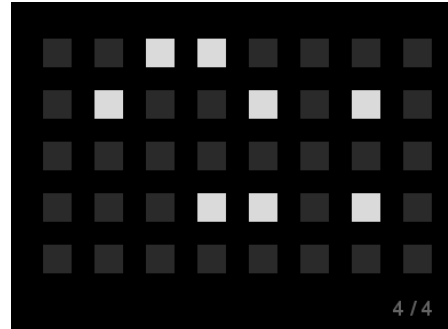
(a) Matrix 1



(b) Matrix 2



(c) Matrix 3



(d) Matrix 4

Abbildung E.3: Komposition VP3 (Musiker)

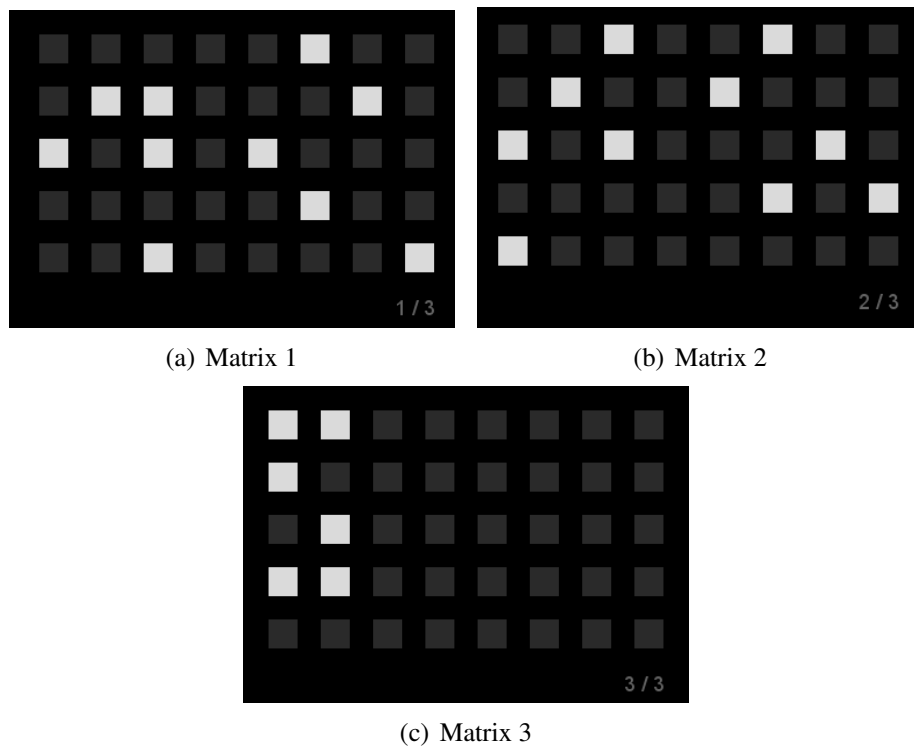


Abbildung E.4: Komposition VP4 (Musiker)

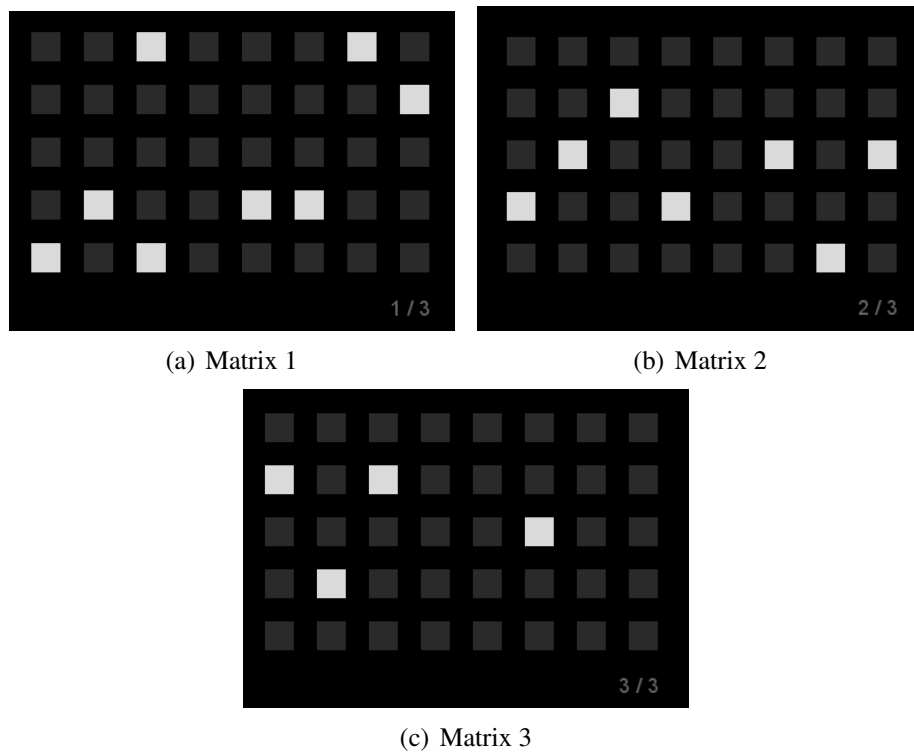


Abbildung E.5: Komposition VP5 (Nicht-Musiker)

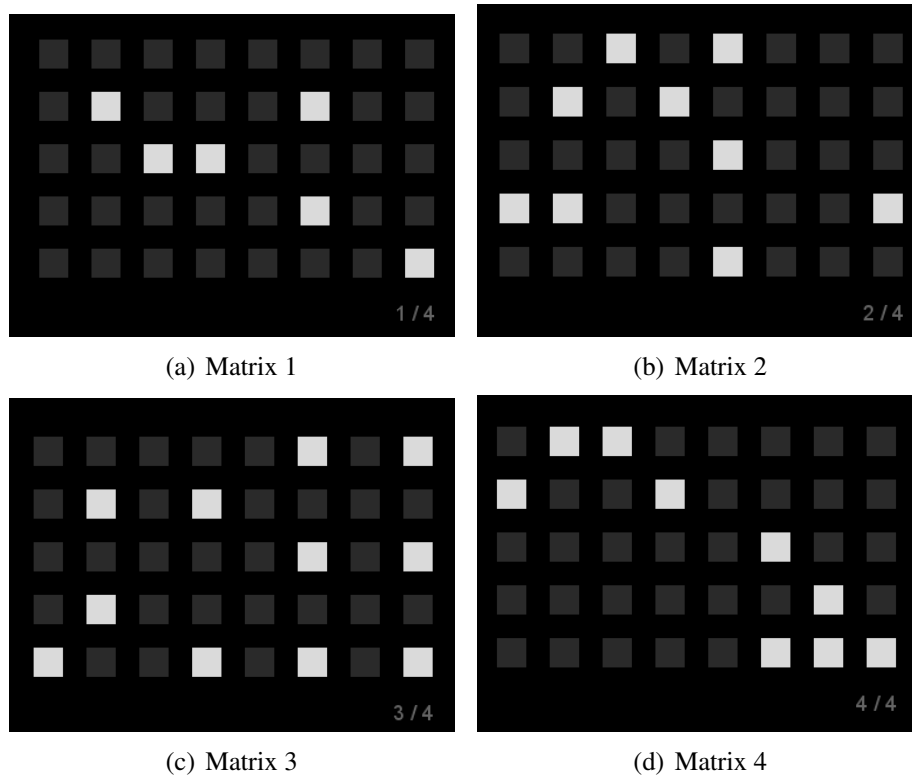


Abbildung E.6: Komposition VP7 (Nicht-Musiker)

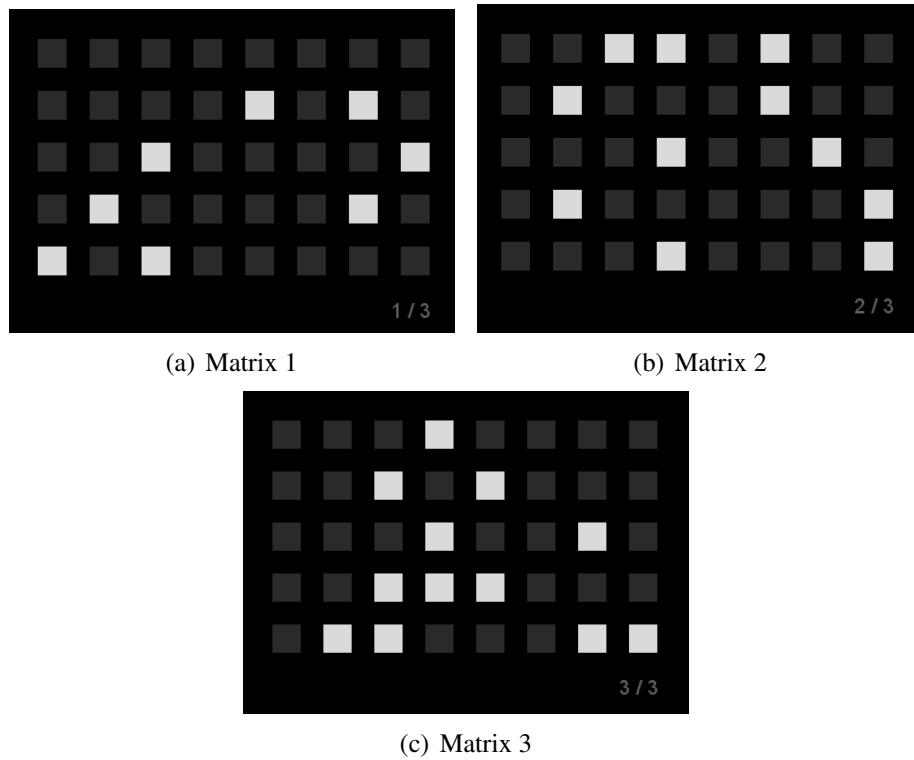


Abbildung E.7: Komposition VP8 (Musiker)

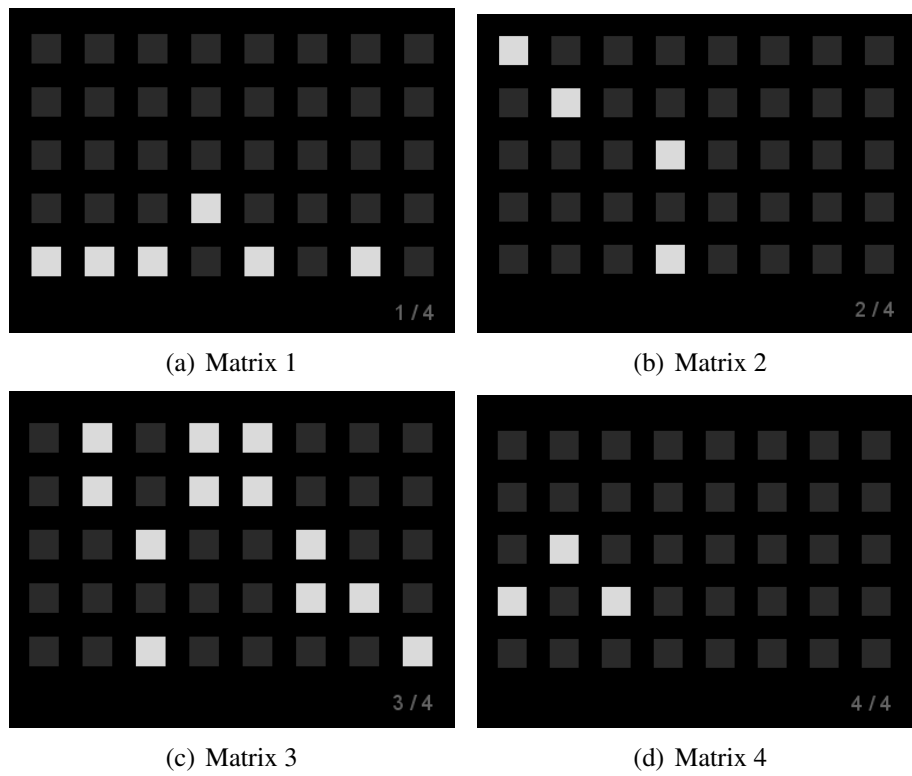


Abbildung E.8: Komposition VP9 (Nicht-Musiker)

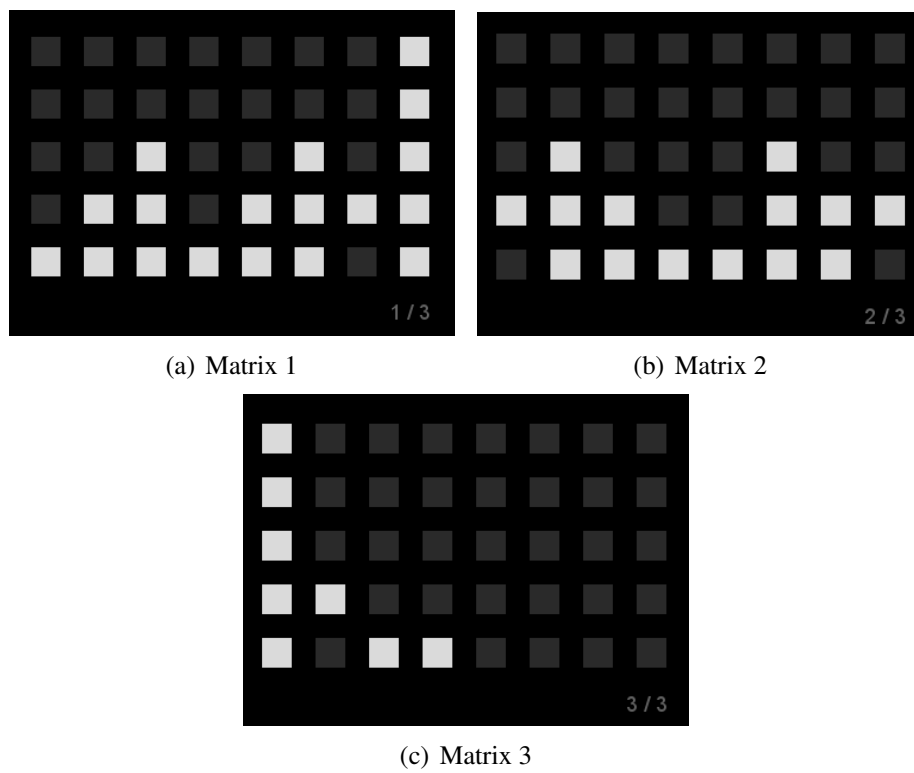
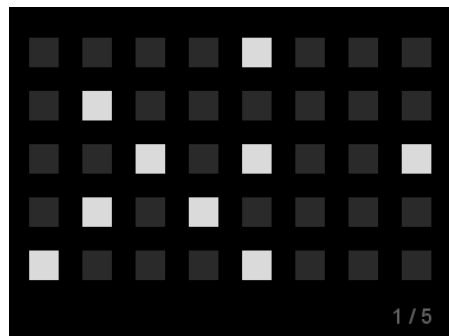
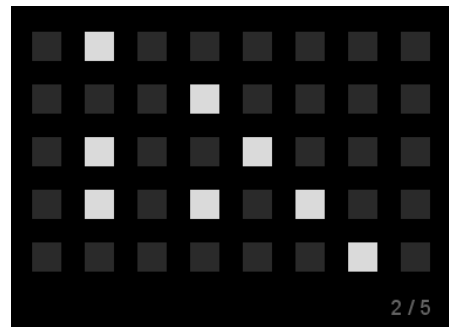


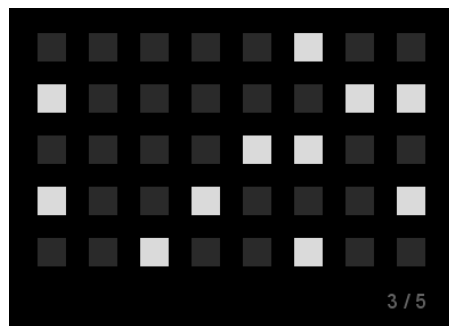
Abbildung E.9: Komposition VP10 (Nicht-Musiker)



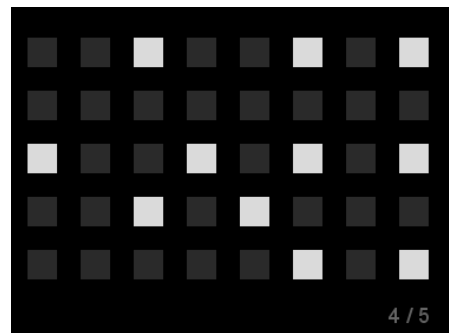
(a) Matrix 1



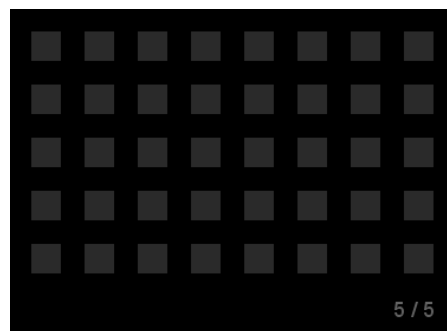
(b) Matrix 2



(c) Matrix 3



(d) Matrix 4



(e) Matrix 5

Abbildung E.10: Komposition VP12 (Musiker)