

Bachelorarbeit von Tamara Flemisch

geboren am 23. November 1990 in Neu-Ulm

Entwurf und Implementierung eines parametrisierten Whole-Body-Gestensets für humanoide Roboter zur dynamischen Feedback-Gestaltung

10. September 2013

Betreuer:

Dipl.-Inf. André Viergutz
Prof. Dr.-Ing. Raimund Dachzelt

Technische Universität Dresden
Institut für Software- und Multimediatechnik
Professur für Multimedia-Technologie

Kurzfassung

Diese Arbeit untersucht auf Basis der aktuellen Forschung Möglichkeiten, die Interaktion mit humanoiden Robotern variabler und somit attraktiver zu gestalten. Hierzu werden Möglichkeiten betrachtet, Emotionen und Absichten dynamisch mittels parametrisierter Gesten auszudrücken.

Zunächst werden Gesten und mögliche Parameter analysiert, die zur Generierung eines Whole-Body-Gestensets notwendig sind, um anschließend diese auf Grund ihrer Auswirkungen auf die Gestenphasen in Intra- und Interphasenparameter zu unterteilen. Dadurch werden die innere und äußere Expressivität einer Geste definiert, die zur Bildung der Single Gesture Expression und der Multi Gesture Expression führen. Diese stellen das Steigerungspotential von Gesten beziehungsweise Absichten dar. Gesten werden anschließend nach Feedback-Arten untergliedert und in Abhängigkeit der Expressivität miteinander in Beziehung gesetzt. Somit entsteht ein Gestenset, das durch die Parameter „Feedback-Art“ und „Expressivität“ bestimmt wird.

Besagtes Gestenset wurde daraufhin zur Verwendung durch den humanoiden Roboter Nao innerhalb der Entwicklungsumgebung Choregraphe umgesetzt. Zudem wurde eine Verhaltensbibliothek entwickelt, die sämtliche, implementierte Gesten zur Wiederverwendung enthält.

Als Ergebnis entstand ein parametrisiertes Gestenset in Theorie und Praxis, das für die ausdrucksstarke Verwendung durch einen humanoiden Roboter geeignet ist und die Kommunikation und Interaktion zwischen Mensch und Roboter erweitert, verfeinert und verbessert.

Abstract

This work investigates the possibilities of developing a more variable, and thus attractive, interaction with humanoid robots based on current research. For this purpose possibilities of dynamically expressing emotions and intentions associated with parametrised gestures are contemplated.

First, gestures and possible parameters, which are necessary for the generation of a whole-body gesture set, are analysed. Afterwards, these are classified into intraphase-parameters and interphase-parameters depending on their impact on gesture phases. A gesture's inner and outer expressivity are thereby defined which leads to the establishment of the single gesture expression and the multi gesture expression. These constitute the gesture's and the intention's potential for increase. Gestures are subsequently classified by feedback types and are related depending on their expressivity. Hence, a gesture set which is defined by the parameters "feedback type" and "expressivity" is developed.

As a result, the aforementioned gesture set has been implemented with the development environment Choregraphe to be used by humanoid robots. Additionally a behaviour library is generated for reusability which contains every implemented gesture.

The result reveals a parametrised gesture set in theory and practice which is suitable to be expressively used by a humanoid robot and enhances, refines and improves the human-robot-interaction.

Erklärung

Hiermit versichere ich, Tamara Flemisch, die vorliegende Bachelorarbeit zum Thema

Entwurf und Implementierung eines parametrisierten Whole-Body-Gestensets für humanoide Roboter zur dynamischen Feedback-Gestaltung

selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht sind und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Dresden, den 10. September 2013

Tamara Flemisch

Inhaltsverzeichnis

1	Einleitung	9
2	Theoretische Hintergründe	11
2.1	Überblick über Gesten	11
2.2	Verwendung von Gesten in der Mensch-Roboter-Interaktion	13
2.2.1	Erkennung von Gesten einer Person durch einen Roboter	13
2.2.2	Verwendung von Gesten durch einen Roboter	15
2.3	Generierung und Parametrisierung von Gesten	17
2.4	Anwendungen von Gesten unter Verwendung des Nao	19
3	Entwurf eines parametrisierten Gestensets	23
3.1	Gestenanalyse	23
3.1.1	Entstehung und Entwicklung von Gesten	23
3.1.2	Analyse alltäglicher Gesten	26
3.1.3	Analyse von Gesten zur Feedback-Gestaltung	29
3.2	Parametrisierung von Gesten	33
3.2.1	Analyse und Bewertung möglicher Parameter	33
3.2.2	Auswahl und Kategorisierung der Parameter	35
3.2.3	Zuordnung der Parametern zu den Gesten	40
3.3	Kategorisierung und Auswahl relevanter Feedback-Gesten	41
3.3.1	Arten des expressiven Ausdrucks	41
3.3.2	Kategorisierung der Gesten anhand von Expressivitätsspannen	43
4	Implementierung des Gestensets für einen humanoiden Roboter	51
4.1	Technische Grundlagen	51
4.2	Grundüberlegungen zur Umsetzung	53
4.2.1	Umsetzung der Gestenerzeugung	53
4.2.2	Umsetzung der Multi Gesture Expression	54
4.3	Implementierung in Choregraphie	56
5	Abschließende Bemerkungen	61
5.1	Zusammenfassung	61
5.2	Fazit	62

5.3 Ausblick	63
Literaturverzeichnis	65
Abbildungsverzeichnis	69
Tabellenverzeichnis	71
Anhang	73
A Auswirkungen der Steigerung der Parameter auf die Gesten	73
B Anpassung der Expressivitätsspannen an den Nao und Unterteilung der Gesten in Stufen	79
C Aufbau der implementierten Gestenerzeugung in der Entwicklungsumgebung Choregraphie	83

1 | Einleitung

Roboter haben heutzutage im alltäglichen Leben der meisten Menschen noch keinen Platz. Sie rücken jedoch durch Forschungsfortschritte und -erkenntnisse weiter in den Fokus. Neben Industrierobotern, die bereits länger in vielen Produktionsstätten eingesetzt werden, werden soziale Roboter unter anderem für die Kommunikation genutzt. Vor allem humanoide Roboter werden punktuell zur Hilfe und Unterstützung von Menschen eingesetzt, ob als Therapiepartner für autistische Kinder, als Hilfe für Menschen mit motorischen Einschränkungen oder als Museumsführer. Sie können dadurch in vielfältiger Art und Weise in den Alltag der Menschen integriert werden und somit deren Leben bereichern, erleichtern und verbessern.

Da oftmals beim Einsatz von humanoiden Robotern die Kommunikation im Vordergrund steht, ist die intuitive Mensch-Roboter-Interaktion von hoher Relevanz, da laut Kim et al. [Kim12] die Möglichkeit der Kommunikation eine der wichtigsten Motivationen für Projekte mit humanoiden Robotern ist. Zusätzlich sollte eine multimodale Interaktion, die durch mehrere und vielfältige Eingabemöglichkeiten definiert wird, mit Robotern möglich sein, da sie die Erwartungen des Menschen an den Robotern unterstützt.

Die Kommunikation teilt sich generell in zwei Kanäle, den verbalen und den nonverbalen, die beide in der Interaktion mit humanoiden Robotern eingesetzt werden sollen. Vargas [Var86] stellte fest, dass in mehr als 65 % der zwischenmenschlichen Kommunikation nonverbale Signale übermittelt werden. Somit ist vor allem dieser Kanal für eine intuitive Interaktion relevant, wodurch sich herausstellt, dass nonverbale Kommunikation zusätzlich zur verbalen Kommunikation betrachtet und in die Arbeit mit humanoiden Robotern miteinbezogen werden muss. Weiterhin fanden Beck et al. [Bec10] heraus, dass Roboter, um sozial akzeptiert zu werden, Emotionen darstellen müssen. Daraus folgt, dass sie menschenähnlich wirken müssen. Zusätzlich lässt die menschliche Erscheinung des Roboters darauf schließen, dass Eingabemodalitäten wie Sprache, Mimik oder Gestik genutzt werden können. Eine der wichtigsten Methoden, Emotionen abzubilden, ist der Einsatz von Gesten. Zudem zählen Gesten zum nonverbalen Kanal und werden nahezu immer in der Kommunikation zwischen Personen verwendet. Sie verschönern nicht nur die Sprache, sondern sind Teil des Erzeugungsprozesses der Sprache, wie McNeill [McN08] feststellte. Des Weiteren existieren Kommunikationssprachen, die lediglich aus Gesten bestehen. Autonome Gesten werden eigenständig, das heißt nicht im Zusammenhang mit anderen Kommunikationskomponenten, wie auditive Äußerungen, verwendet. Auf diese Weise lassen sie sich zu einer eigenständigen Sprache zusammenfassen, welche

als Gebärdensprachen bezeichnet werden.

Zusammengefasst wird die Notwendigkeit deutlich, dass ein Roboter durch die Verwendung von Gesten von Menschen als adäquater Gesprächspartner wahrgenommen wird. Somit existiert eine berechtigte Grundlage für die Verwendung von Gesten in der Interaktion mit humanoiden Robotern.

Jedoch stellt die Erstellung von Gesten und Gestensets einen sehr hohen Aufwand dar, der nicht immer für kleinere Projekte in Kauf genommen werden kann, um die Kommunikation zu erweitern. Somit wäre es wünschenswert, wenn es ein modifizierbares Gestenset gäbe, das für viele Zwecke anpassbar ist.

In dieser Arbeit soll untersucht werden, inwiefern Gesten geeignet sind, Emotionen und Absichten eines Roboters auszudrücken und wie das Feedback eines Roboters dynamischer und anpassbarer erzeugt werden kann. Weiterhin stellt sich die Frage, ob und in welcher Art und Weise es Eigenschaften und Parameter gibt, mit deren Hilfe sich die Gesten an Situationen anpassen lassen. Die Ergebnisse dieser Untersuchungen sollen anhand einer Interaktionsmöglichkeit in einem Prototyp auf dem humanoiden Roboter Nao implementiert werden.

Zunächst wird im nächsten Kapitel ein Überblick über die nötigen Grundlagen der Mensch-Roboter-Interaktion und die Verwendung von Gesten gegeben. Anschließend werden einige verwandte Arbeiten, im Speziellen mit dem Roboter Nao, erläutert und besprochen. Das darauffolgende Kapitel handelt von dem Entwurf des Gestensets, wobei zunächst Gesten analysiert und kategorisiert werden, um zudem Parameter zu analysieren und für die ausgewählten Gesten zu bestimmen. Weiterhin werden die Ausdrücke und Absichten innerhalb einer Kommunikation weiter klassifiziert und mit Hilfe der Gesten und von sogenannten Expressivitätsspannen definiert. Im vorletzten Kapitel wird auf die Implementierung mit dem humanoiden Roboter Nao eingegangen. Sowohl technische Grundlagen als auch weitere Überlegungen zur Umsetzung und der Vorgang der tatsächlichen Implementierung werden an dieser Stelle beschrieben und erläutert. Schlussendlich wird die Arbeit zusammengefasst und ein Fazit gezogen, worauf abschließend noch ein Ausblick auf zukünftige Arbeiten folgt.

2 | Theoretische Hintergründe

Dieses Kapitel beschäftigt sich mit den theoretischen Hintergründen von Gesten und ihrer Verwendungen in der Mensch-Roboter-Interaktion. Betrachtet wird, wie Gesten generiert werden und welche Parameter sie aufweisen. Es werden hierbei verwandte Arbeiten vorgestellt, die sich im Speziellen mit dem humanoiden Roboter Nao von Aldebaran Robotics [Ald] beschäftigen.

2.1 Überblick über Gesten

Gesten stellen neben Mimik, Tonfall, Berührungen, Gerüchen und vielen weiteren Aspekten einen Kanal der nonverbalen Kommunikation dar. Dieser wird neben der verbalen Kommunikation, der Sprache, verwendet, um sich mit anderen Menschen zu verständigen. Nonverbale Signale, die vor allem unbewusst ausgesendet werden, sind ein wichtiger Teil der Kommunikation und erst durch sie entsteht Flexibilität in der zwischenmenschlichen Interaktion.

Gesten werden allgemein als bewusste oder unbewusste Bewegungen von Körperteilen, wie zum Beispiel der Arme, Hände, des Kopfes oder der Beine, bezeichnet, die eine indirekte Mitteilung oder Funktion übermitteln. Funktionen von Gesten wurden bereits von Cadoz [Cad94] definiert. Bei der „epistemic function“ werden Hände beispielsweise zur Wahrnehmung eingesetzt, um etwas zu ertasten. Die „ergotic function“ bildet die Interaktion mit der Umgebung, indem man unter anderem einen Gegenstand bewegt. Die Aufgabe, die in dieser Arbeit hauptsächlich betrachtet wird, ist die „semiotic function“. Die Funktion dieser Gesten ist es, den verbalen Kanal, die Sprache, zu begleiten. Gesten, die mitunter diese Funktion erfüllen, werden coverbale Gesten genannt, die unter anderem von McNeill [McN92] in einer Taxonomie klassifiziert wurden. Im Folgenden wird ausschließlich auf diese Taxonomie eingegangen, da sie zum einen in vielen verwandten Arbeiten verwendet wurde und zum anderen sich speziell auf coverbale Gesten fokussiert und diese klassifiziert.

Hierbei stellt McNeill vier Unterkategorien vor: metaphorische Gesten, ikonische Gesten, deiktische Gesten und Beat-Gesten. Metaphorische Gesten (Abbildung 2.1a) repräsentieren abstrakte Konzepte, wie zum Beispiel eine Wiege-Geste beim Abwägen zwischen zwei Gesichtspunkten. Ikonische Gesten (Abbildung 2.1b) bilden die Realität ab, indem sie ein Objekt, eine Handlung oder eine räumliche Situation nachahmen. Die Darstellung von verschiedenen Positionen im Raum mit den Händen zählt in dieser Taxonomie zu den ikonischen Gesten. Deiktische Gesten (Abbildung 2.1c) sind Zeigegesten,

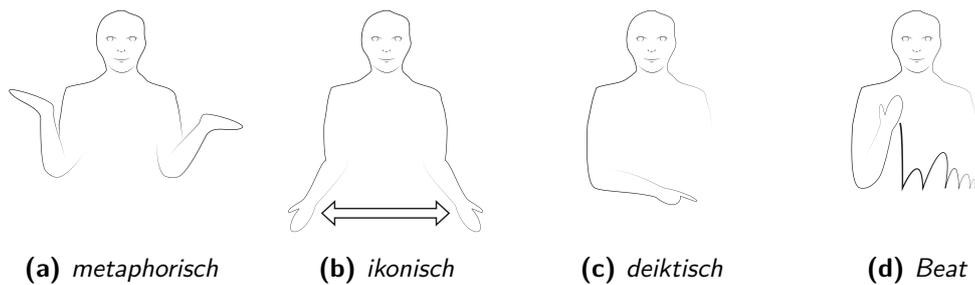


Abb. 2.1: *Gesten-Kategorien der McNeill-Taxonomie*

wie beispielsweise das Hinweisen auf ein Objekt mit einem Finger. Der Rhythmus der Sprache wird während des Sprechens durch Beat-Gesten (Abbildung 2.1d) unterstützt, die vor allem bei Reden häufig und deutlich eingesetzt werden. Diese Gesten betonen die Sprache und werden oftmals durch Auf- und Abwärtsbewegungen der Hände dargestellt.

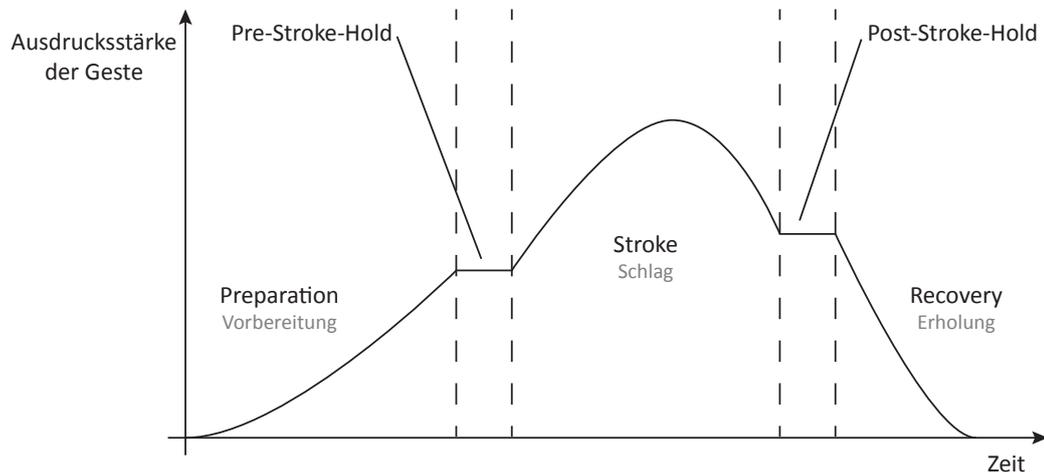


Abb. 2.2: *Schematische Darstellung einer in die drei Phasen (Preparation, Stroke, Recovery) unterteilten Geste mit den zwei Zwischenphasen (Pre-Stroke-Hold, Post-Stroke-Hold)*

Die Unterteilung einer einzelnen Geste erfolgt nach Kendon [Ken04] in drei Phasen, die in Abbildung 2.2 zu sehen sind: Preparation, Stroke und Recovery. Der Stroke ist der wichtigste Abschnitt, da in ihm die eigentliche Geste ausgeführt wird. Die Preparation leitet den Stroke ein, ist ihm vorangestellt und enthält die Anspannung des Körpers und die Bewegung in die richtige Körperhaltung für die Geste. Die Recovery oder Retraction wird nach dem Stroke ausgeführt und enthält das Lockern und Zurückziehen der verwendeten Körperteile, wie zum Beispiel der Hand. Zwischen Preparation und Stroke kann unter Umständen noch ein Pre-Stroke-Hold auftreten, in dem die Startposition des Strokes gehalten wird. Äquivalent verhält es sich mit dem Post-Stroke-Hold, der

zwischen Stroke und Recovery auftreten kann.

Um Gesten für Maschinen zugänglich und lesbar zu machen, entwickelten Forscher Gestenbeschreibungssprachen, wie die Behaviour Markup Language von Kopp et al. [Kop06], die auf XML basiert. Die Gesten werden anhand ihrer Kategorien, die an die McNeill-Taxonomie angelehnt sind, und ihrer verwendeten Verhaltenselemente, zum Beispiel Kopfbewegung, Blick und Sprache, beschrieben. Sprachen, die Gesten beschreiben, finden sowohl in der Mensch-Computer-Interaktion als auch in der Mensch-Roboter-Interaktion Anwendung.

2.2 Verwendung von Gesten in der Mensch-Roboter-Interaktion

In der Mensch-Roboter-Interaktion existieren zwei übergeordnete Verwendungsrichtungen für Gesten. Zum einen können Gesten von einem Roboter zur Kommunikation mit einem Interaktionspartner verwendet werden und zum anderen ist die gegensätzliche Richtung möglich. Hierbei werden die Gesten des Menschen, der mit dem Roboter interagiert, erkannt und ausgewertet. Beide Arten der Verwendung werden in Abbildung 2.3 dargestellt.

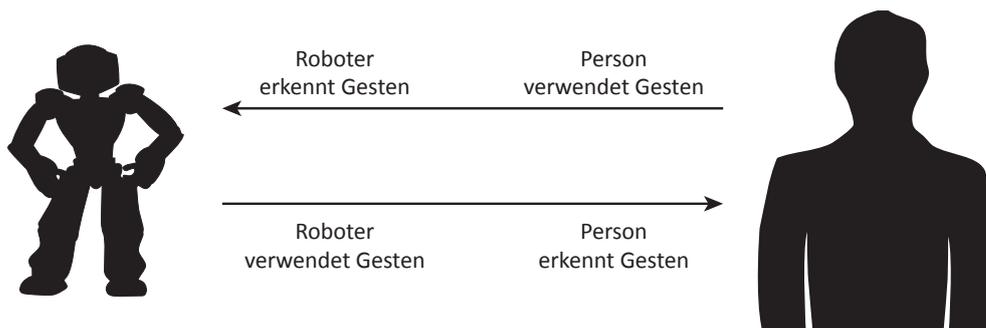


Abb. 2.3: Zwischen einem Roboter und einer Person können vor allem zwei verschiedene Arten der Kommunikation über Gesten stattfinden. Zum einen verwendet eine Person Gesten, die anschließend von einem Roboter erkannt werden. Zum anderen kann ein Roboter selbst Gesten verwenden.

2.2.1 Erkennung von Gesten einer Person durch einen Roboter

Die Erkennung von Gesten, die von einer Person verwendet werden, wird hauptsächlich eingesetzt, um die Nutzerschnittstellen natürlicher zu gestalten, so dass eine intuitive Bedienung und Kommunikation mit einem Roboter möglich ist. Hilfreich ist vor allem die Erkennung von Zeigegesten, um beispielsweise die Aufmerksamkeit des Roboters auf ein bestimmtes Objekt zu lenken oder ihn zu bestimmten Orten zu führen.

Im Folgenden wird nur ein Überblick über die Anwendung der Gestenerkennung gegeben, da der Fokus dieser Arbeit auf der Verwendung der Gesten durch einen Roboter liegt.

Nach Yang et al. [Yan07], die sich mit der Erkennung von Whole-Body-Gesten innerhalb von Bewegungsabläufen beschäftigen, wird das Problem der Gestenerkennung generell in zwei Teile untergliedert: Segmentierung und Erkennung. Die Segmentierung beschäftigt sich mit dem Finden des Anfangs und des Endes einer Geste. Anschließend wird bei der Erkennung die gefundene Geste mit einer Bibliothek aus vordefinierten Gesten verglichen, um eine eindeutige Identifizierung zu erhalten.

Eine Möglichkeit, Gestenerkennung in der Mensch-Roboter-Interaktion einzusetzen, zeigen Kobayashi et al. [Kob08] anhand eines Roboters, der in einem Museum mit den Besuchern interagieren und ihnen Informationen über die Ausstellungsstücke liefern soll. Hierbei erkennt er nonverbale Signale und sendet sie ebenso aus. Die Erkennung beschränkt sich auf Sprach- und Gesichtserkennung. Somit ist es möglich, dass der Blick des Roboters auf das Gesicht des Interaktionspartners gerichtet ist, auch wenn der Roboter seinen Kopf in eine andere Richtung dreht. Des Weiteren wird die Blickrichtung der Person berechnet, um Informationen über den Aufmerksamkeitsfokus zu erhalten, darauf reagieren zu können und gegebenenfalls die Erklärungen anzupassen.

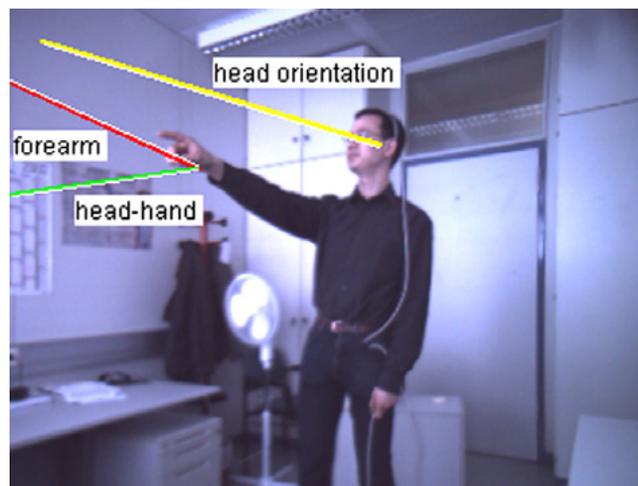


Abb. 2.4: Darstellung der Ansätze für die Richtungserkennung einer deiktischen Geste mit Hilfe der Ausrichtung des Kopfes, der Zeigerichtung des Unterarms und der weitergeführten Verbindung zwischen Kopf und Hand

Ein weiteres Projekt, das sich mit Gestenerkennung beschäftigt, ist von Waldherr et al. [Wal00]. Der Fokus liegt auf Service-Robotern und den Möglichkeiten, diesen den Weg zu weisen. Somit werden hier vor allem deiktische Gesten erkannt, um es dem Nutzer zu ermöglichen, auf natürliche Weise dem Helfer den Weg an die gewünschten Orte zu weisen. Zusätzlich kann der Roboter Personen tracken, um ihnen zu folgen, wenn gerade keine Richtungsanweisung zur Verfügung steht. Die Erkennung deiktischer Gesten wird ebenfalls bei Nickel & Stiefelhagen [Nic07] verwendet. Auf Abbildung 2.4

sieht man, dass zusätzlich zum Tracking der Hände auch die Bewegungen und Ausrichtung des Kopfes aufgezeichnet wurden, um von der Blickrichtung der Person auf die Zeigerichtung während einer Zeigegeste zu schließen. Dieses System führte zu einer erheblichen Verbesserung der Performance der Gestenerkennung und der Schätzung der Zeigerichtung einer deiktischen Geste.

Demzufolge fällt auf, dass Gestenerkennung in der Mensch-Roboter-Interaktion häufig eingesetzt wird. Jedoch wird sie ebenfalls innerhalb anderer Bereiche der Informatik oftmals und verstärkter eingesetzt.

2.2.2 Verwendung von Gesten durch einen Roboter

In der Mensch-Roboter-Interaktion bietet laut Scheutz et al. [Sch06] die Verwendung von Gesten durch einen Roboter und somit das Aussenden nonverbaler Signale, wesentlich mehr Potential als die Erkennung der nonverbalen Signale, die ein Mensch aussendet. Zusätzlich übermittelt die Verwendung nonverbaler Signale dem Nutzer ein erweitertes Verständnis für den inneren Zustand des Roboters und lässt ihn lebendiger und akzeptabler wirken. Jedoch treten hier Beschränkungen der Glaubwürdigkeit auf, vor allem wenn Gesten, die ein Roboter aussendet, zu sehr denen eines Menschen gleichen. Diese Tatsache erkannten Kim et al. [Kim12] im Rahmen einer Nutzerstudie, bei der die Nutzer sich wohler fühlten, wenn der Roboter nicht versuchte, menschliche Gesten zu exakt zu imitieren.



(a) *introvertiert*



(b) *extrovertiert*

Abb. 2.5: *Persönlichkeiten eines Roboters mit unterschiedlichem nonverbalem Verhalten: Zum einen extrovertiert mit deutlichen Augen und angehobenem Kopf und zum anderen introvertiert mit gesenktem Kopf und undeutlichen Augen*

In einer Arbeit von Aly & Tapus [Aly13] wird das nonverbale Verhalten eines Roboters in Abhängigkeit der Persönlichkeit eines Nutzers betrachtet, da diese einen großen Einfluss auf die von den Menschen eingesetzten Gesten und andere Verhaltenszüge hat.

Hierzu wird das „Big Five Framework“ verwendet, welches die Persönlichkeit in fünf Charakterzüge unterteilt: Offenheit für Erfahrungen, Gewissenhaftigkeit, Extraversion, Verträglichkeit und Neurotizismus. Die Gesten des introvertierten Roboters wurden weniger ausschweifend, langsamer und mit einer geringeren Wiederholungsrate generiert als die des extrovertierten Roboters. Auf Abbildung 2.5 wird der Unterschied der beiden generierten Persönlichkeiten deutlich. Bei einer Studie, in der Nutzer einem introvertierten Roboter und einem extrovertierten Roboter ausgesetzt wurden, stellte sich heraus, dass Nutzer einen Roboter bevorzugten, der ihrer Persönlichkeit ähnelt, wobei diese Tendenz bei introvertierten Nutzern geringer ausfiel.

Als Anwendung von Gesten dient beispielsweise der Museumsroboter, der bereits in Kapitel 2.2.1 erwähnt wurde. Er kann zusätzlich zur Erkennung auch nonverbaler Signale aussenden. Der Fokus wurde auf die wichtigsten Elemente, wie Kopfbewegungen und Zeigegesten, gelegt. Die Bewegungen des Kopfes werden zur Vermittlung von Aufmerksamkeit verwendet, wenn der Roboter gerade nicht mit einem Besucher interagiert. Sobald ein Gesicht erkannt wird, folgt er diesem mit seinen Kopfbewegungen. Während der Erklärung der Ausstellung werden deiktische Gesten verwendet, um auf besondere Ausstellungsstücke zu verweisen, über die der Roboter im Moment spricht.

Eine weitere Möglichkeit nonverbales Verhalten zu übermitteln, wurde von Wendt & Berg [Wen09] untersucht. Sie beziehen sich auf den Ausdruck von nonverbalem Humor durch Gesichtsausdrücke, Bewegungen und Gesten. Beispielsweise wird durch das Necken des Interaktionspartners Humor erzeugt. Zusätzlich werden menschliche Reaktionen auf alltägliche Reize imitiert, wie zum Beispiel Freude, die durch Klatschen ausgedrückt wird. Als weiteres Element wurde eine Tanzsequenz verwendet, wodurch die Stimmung aufgelockert wurde.



(a) während der Entwicklung



(b) vollendetes Stadium

Abb. 2.6: Zwei Stadien der Entwicklung des humanoiden Roboters Leonardo

Der Roboter „Leonardo“ zeigt eine Vielzahl nonverbaler Verhaltensweisen. Er wurde am MIT Media Lab von Breazeal et al. [Bre05] entwickelt. Als humanoider Roboter kann er sowohl visuelle als auch auditive Eingabeströme erfassen und verarbeiten. In Abbildung 2.6 ist die Entwicklung dargestellt, wodurch ein früher und der vollendete

Zustand sichtbar werden. Seine Interaktionsmöglichkeiten sind ausschließlich nonverbal. Er reagiert auf Äußerungen des Interaktionspartners mit Gesichtsausdrücken und Gesten. Beispielsweise zeigt er bei fehlgeschlagener Spracherkennung mit Hilfe von Gesten, wie Schulterzucken, dass er verwirrt ist. Somit weiß der Nutzer zu jedem Zeitpunkt, in welchem inneren Zustand sich der Roboter im Moment befindet. Zusätzlich begrüßt er den Partner mit Winken und reagiert auf Objekte in seiner Nähe mit einer Geste, die ausdrückt, dass er an dem Gegenstand interessiert ist. Augenblinzeln und spontane Veränderungen des Blicks und der Körperhaltung werden außerdem eingesetzt, um ihn lebensechter wirken zu lassen. Daraufhin wurde in einer Studie überprüft und gezeigt, dass von den Nutzern ein mit nonverbalen Signalen agierender Roboter besser angenommen wird als ein konventioneller Roboter.

Aus diesen Beispielen der Verwendung nonverbaler Signale wird deutlich, dass sie verschiedene Einsatzmöglichkeiten birgt und viele humanoide Roboter erweitert, wodurch die Kommunikation zum Menschen verbessert werden kann. Die Erkenntnis, dass die Verwendung nonverbaler Signale effektiver ist als die bloße Erkennung, die zu Beginn dieses Kapitels erwähnt wurde, erklärt sich durch diese Anwendungsfälle und wird selbstverständlich. Somit ist es essentiell, sich weiterhin mit dieser Möglichkeit, einen Roboter hinsichtlich der Kommunikation zu verbessern, zu beschäftigen.

2.3 Generierung und Parametrisierung von Gesten

Der Schwerpunkt der Arbeit liegt auf der Parametrisierung von Gesten, weswegen im Folgenden einige Modelle hierfür erläutert werden. Ein Thema, das eng mit der Parametrisierung zusammenhängt, ist die Generierung von Gesten, da diese oftmals durch parametrisierte Modelle erzeugt werden.

Ein Begriff, der die Parameter, nach denen Gesten ausgewählt und kategorisiert werden, näher beschreibt, ist die Expressivität. Es handelt sich dabei um die Ausdrucksstärke beziehungsweise die Ausdruckskraft, die Qualität einer Geste und wie diese durch die Eigenschaften der Geste zustande kommt.

Eine Möglichkeit, Gestensets zu generieren, ist mittels Text-Eingabe. Hierfür liefert das von Cassel et al. [Cas04] entwickelte BEAT-Toolkit ein gutes und umfangreiches Beispiel. Es wurde im Projekt von Aly & Tapus [Aly13] eingesetzt, das bereits in Kapitel 2.2.2 erwähnt wurde. Bei diesem Toolkit werden mit Hilfe der Sprach- und Kontextinformationen der Eingabe die Bewegungen der Hände, der Arme und des Gesichts gesteuert. Zu dem Gestenset liefert das Toolkit zusätzlich eine synchronisierte Sprachausgabe. Grundlage dafür ist die Spezifikation der Gesten durch eine Notation, die die Handformen und die Bewegungslinien beider Arme unabhängig voneinander beschreibt. Dabei sind häufige Gesten, wie eine Beat-Geste und deiktische Gesten, bereits enthalten und können auch erweitert werden. Ikonische Gesten wurden nicht vordefiniert, da sie sich je nach Person unterscheiden. Zunächst werden die einzelnen Wörter der Text-Eingabe getaggt und Kategorien zugeordnet, wodurch eine Baumstruktur entsteht, die mit Hilfe von XML gespeichert wird. Die Vorschläge für Verhaltensweisen werden mittels Durchlaufen des XML-Baums realisiert. Hierbei werden alle potentiell passenden

Verhaltensweisen generiert und anschließend gefiltert, um ein endgültiges Gestenset zu erhalten. Für die Erzeugung der Gesten werden verschiedene Generatoren verwendet, die durch Tags unterschiedliche Gesten erzeugen können. Die einzelnen Parameter wurden nicht beschrieben, da die Gesten von den Nutzern des Toolkits definiert werden und somit die Entwickler keinen Einfluss auf die verwendeten Gesten und ihre Parameter hatten.



Abb. 2.7: *Der Embodied Conversational Agent Greta*

Ein weiteres Projekt, das ähnliche Funktionalitäten wie das BEAT-Toolkit aufweist, ist die GRETA-Plattform von Niewiadomski et al. [Nie09]. Hierbei handelt es sich um ein System, das aus einer Kommunikationsabsicht und einer Emotion nonverbales Verhalten mit unterschiedlicher Expressivität generiert. Es handelt sich um einen Embodied Conversational Agent (ECA), eine virtuelle Person, die mittels Gesten, Gesichtsausdrücken und Sprache mit einem Nutzer kommunizieren kann, deren virtuelle Persönlichkeit in Abbildung 2.7 zu sehen ist. Verwendet wird unter anderem die Behaviour Markup Language [Kop06], die bereits in Kapitel 2.1 erwähnt wurde. Für dieses System entwickelten Hartmann et al. [Har06] eine Möglichkeit Gestensets zu synthetisieren. Das Ziel ihrer Arbeit war es, Eigenschaften, wie Stimmung, Persönlichkeit und Gefühle auf Low-Level Parameter, wie zum Beispiel Winkel an Gelenken abzubilden. Dies erreichen sie mitunter durch die Verwendung von Parametern für die Ausdrucksstärke einer Geste.

Beide bisher gezeigten Methoden, das BEAT-Toolkit und GRETA, basierten auf Conversational Agents. Im Gegensatz dazu steht eine Methode von Ng-Thow-Hing et al. [NTH10], bei der ein humanoider Roboter verwendet wird. Stilistische Eigenschaften des Textes werden automatisch durch Analyse extrahiert und mit Tags versehen. Anschließend wird über eine passende Geste entschieden. Weiterhin existiert ein Expressivitätsparameter, der sich je nach Kategorie der Geste unterscheidet. Beispielsweise haben ikonische Gesten generell eine tendenziell höhere Expressivität als Beat-Gesten. In diesem System werden Gesten durch „TCB“ (tension-continuity-bias) Splines beschrieben, die Bewegungslinien und deren Schlüsselpunkte eines Körperteils während der Geste festlegen. Dieser Vorgang ist in Abbildung 2.8 dargestellt. Die verwendeten Bewegungslinien

sind Funktionen, die mit Hilfe von Parametern verändert werden können, um festzulegen, wie genau eine Geste dieser vorgegebenen Linie folgt.

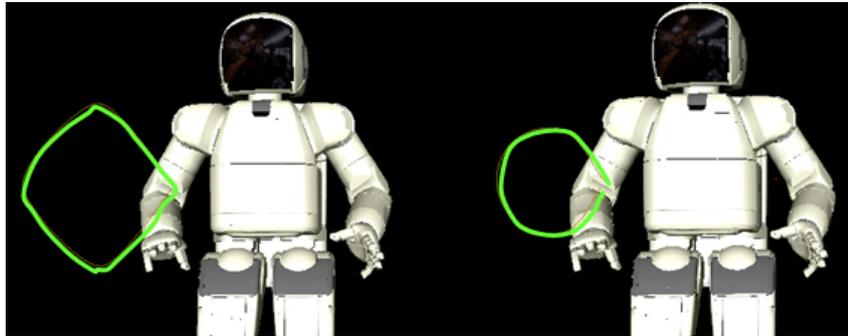


Abb. 2.8: Auf der linken Seite der Abbildung ist eine sehr expressive Geste zu sehen, die ihren Positionen sehr genau folgt. Eine weniger expressive Geste, die diesen Positionen nur grob folgt, ist auf der rechten Seite dargestellt.

Folglich gibt es mehrere Arten parametrisierter Gestensets, die sich in gewissen Parametern und Eigenschaften überschneiden, aber auch unterscheiden. Auf weitere Einzelheiten bezüglich der Konzepte zur Parametrisierung wird im weiteren Verlauf, vor allem im Kapitel 3.2.1, eingegangen.

2.4 Anwendungen von Gesten unter Verwendung des Nao

Die Implementierung von Gesten mit einem Nao ist ein weiterer Schwerpunkt der Arbeit, weswegen dazu verwandte und relevante Arbeiten betrachtet werden. Nao ist ein humanoider Roboter, der von Aldebaran Robotics [Ald] entwickelt wurde. Die äußere Erscheinung des 58 cm hohen Roboters ist auf Abbildung 2.9 dargestellt. Zusätzlich werden weitere Eigenschaften in Kapitel 4.1 erläutert.

Monceaux et al. [Mon09], Mitarbeiter von Aldebaran Robotics, entwickelten eine Gestenbibliothek, in der ungefähr 40 Gesten und emotionale Ausdrücke kategorisiert und implementiert wurden. Eingeteilt in drei Gruppen drücken sie entweder Freude, Neutralität oder Missfallen aus. Zur Kategorie „Freude“ gehören Gesten und Ausdrücke wie „interessiert“ oder „fröhlich“. „Neutralität“ enthält „warten“ oder „Konzentration“ und „Missfallen“ beinhaltet zum Beispiel „Ekel“ oder „Müdigkeit“.

In dem bereits in Kapitel 2.2.2 angesprochenen Projekt von Aly & Tapus [Aly13] wird der Nao zum einen als eher zurückhaltender und schüchterer Roboter und zum anderen als offener und extrovertierter Kommunikationspartner eingesetzt. Hier wurde zusätzlich das BEAT-Toolkit verwendet, auf das ebenfalls in Kapitel 2.3 eingegangen wurde.

Auf Grund der Erscheinung des Naos, die freundlich und kindlich wirkt, wurde er vielfach zu Therapiezwecken eingesetzt. Beispielsweise haben Shamsuddin et al. [Sha12] ihn als Kommunikations- und Therapiepartner für Kinder mit autistischem Verhalten verwendet und nonverbale Verhaltensweisen und Gesten eingesetzt um eine Kommunikation mit den Kindern aufzubauen.



Abb. 2.9: *Der Nao-Roboter von Aldebaran Robotics*

Bei einem weiteren Projekt von Ham et al. [Ham11] war das Ziel, einen Nao überzeugender wirken zu lassen. Hierfür wurden zwei Strategien, Überzeugungskraft durch Gesten und durch den Blick, entwickelt und überprüft. Er sollte innerhalb der Erzählung einer Geschichte möglichst glaubhaft wirken. Genutzt wurden Gesten, die mit Hilfe einer Videovorlage eines Schauspielers, der die gleiche Geschichte wie der Nao erzählt, nachgebildet wurden. Die Anpassung der Gesten an den Nao wird beispielhaft in Abbildung 2.10 gezeigt. Die Gesten wurden nicht automatisch generiert, sondern begleitend zur Geschichte programmiert. Verwendet wurden hierbei eine Geste für Rufen, Zeigegesten und eine Adaption des Gehens. Herausgefunden wurde, dass der Blick an sich die Überzeugungskraft einer Geschichte erhöht und zusätzliche Gesten diese noch verstärken. Jedoch sind Gesten ohne Blickkontakt nicht ausreichend, um eine gleiche Überzeugungskraft wie mit dem Blick zu erreichen.

Meena et al. [Mee12] erweiterten das Verhalten des Naos, so dass er lebensechter und ausdrucksstärker wirkt. Genutzt wurde WikiTalk, ein System mit dessen Hilfe der Nao seine Gesprächsinhalte aus Wikipedia-Artikel nehmen kann. Implementierte Gesten sind vor allem Positionen der Arme und des Kopfes, die die Sprache des Naos unterstützen. Zusätzlich wurde die Veränderung der Blickrichtung genutzt, so dass der Nao der Person mit seinem Blick folgt. Ein weiterer Schwerpunkt ist die „Navigation“ durch WikiTalk. Nao beobachtet sein Gegenüber und reagiert auf Aktionen des Nutzers, die erwartet werden, sobald der Nao ein Thema beendet hat. Die Gesten wurden durch Nutzung der Phasen der Gesten (siehe Kapitel 2.1) implementiert. Für jede Phase wurde eine Position festgelegt, wobei anschließend auf diese drei Posen ein Interpolationsalgorithmus angewendet wurde. Entstanden ist ein Gestenset und eine Gesten-Bibliothek mit parametrisierbaren Gesten, die bezüglich ihrer Dauer und ihrer Amplitude verändert werden können.

Der erwähnte Algorithmus zur Interpolation von Schlüsselpositionen verschiedener



Abb. 2.10: Bei der Erstellung der Gesten für deinen Nao, der Geschichten erzählt, wurde auf Videos eines Schauspielers zurückgegriffen. Die Adaption der Gesten für den Nao wurde anschließend durchgeführt.

Gesten des Nao wurde von Beck et al. [Bec10] entwickelt. Ziel war es den Nao trotz Mangel an Gesichtsausdrücken, mit dem Ausdruck von nonverbalem Verhalten und Gefühlen auszustatten. Es ist möglich, dass der Nao automatisch zwischen zwei Schlüsselpositionen interpoliert, dies als Bewegung generiert und diese Bewegung anschließend ausführt. Abbildung 2.11 zeigt zusätzlich die einzelnen Phasen der Interpolation.



Abb. 2.11: Interpolation über Schlüsselpositionen von 100 % Traurigkeit bis zu 100 % Angst über gemischte Positionen, die sich zu unterschiedlichen Teilen aus Traurigkeit und Angst zusammensetzen.

Basierend auf der GRETA Plattform (siehe Kapitel 2.3) und deren Verhaltens-Engine versucht eine Gruppe um Le und Pelachaud einen Nao mit expressiven Kommunikationsgesten auszustatten, damit er überzeugend eine Geschichte erzählen und ein ausdrucksstarkes Verhalten verwenden kann. Die Gesten, die durch die GRETA Plattform erzeugt werden, werden im Anschluss an die Generierung an den Nao angepasst. Die Gesten, die

ursprünglich für einen Embodied Conversational Agent (ECA) optimiert wurden, mussten für einen humanoiden Roboter adaptiert werden, der im Vergleich zum Embodied Conversational Agent über die Bewegung der Beine verfügt und von der Gravitation beeinflusst wird. Jedoch kann der Nao seine Hände nicht bewegen und es ist ihm nicht möglich Gesichtsausdrücke zu zeigen. Zudem ist der Nao langsamer in der Ausführung als der Embodied Conversational Agent, wie man in Abbildung 2.12 sieht. Die Gesten werden schlussendlich in einer Datenbank mittels Behaviour Markup Language (siehe Kapitel 2.1) gespeichert und dynamisch als Gelenkwerte erzeugt. Diese Instanziierung berücksichtigt zudem die Expressivitätsparameter von Hartmann et al. [Har06], auf die in Kapitel 3.2.1 eingegangen wird. [Le11; Le12; Pel10]

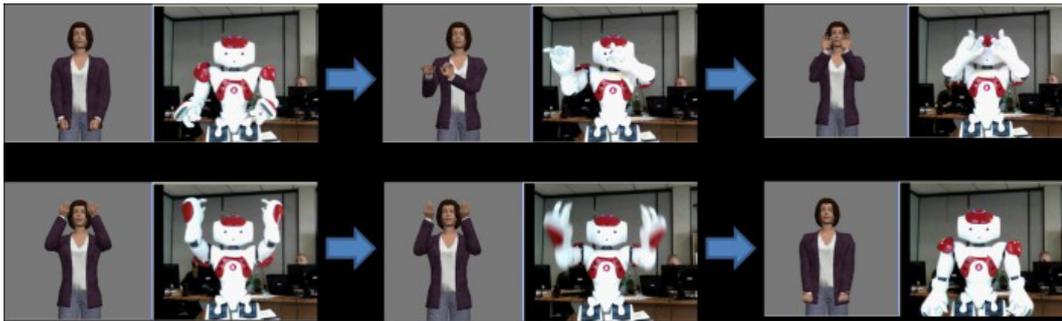


Abb. 2.12: Bei dem Vergleich der gleichen ausgeführten Geste von GRETA und einem Nao fällt vor allem auf, dass der Nao etwas langsamer in seinen Bewegungen ist. Dies wird im unteren mittleren Bild deutlich.

Mittlerweile gibt es unzählige Projekte, die sich mit dem Nao und der Verwendung von Gesten beschäftigen, so dass es den Rahmen sprengen würde, an dieser Stelle alle Arbeiten zu erwähnen. Humanoide Roboter werden in vielen Feldern der Mensch-Roboter-Interaktion eingesetzt und können vielfältige Anwendungen innehaben, wie als Hilfe für Menschen, als Gesprächspartner, als Museumsführer oder als Therapiemittel. Jedoch nutzen viele dieser Fälle die Interaktion mit dem Roboter als Schwerpunkt ihrer Verwendung, wodurch sich die Frage stellt, inwiefern diese Interaktion verbessert werden kann.

3 | Entwurf eines parametrisierten Gestensets

Um ein Gestenset zu entwerfen, müssen zunächst Gesten analysiert werden. Anschließend muss geklärt werden, welche Gesten in unserem Alltag überhaupt verwendet werden und wie und zu welchem Zweck diese Verwendung stattfindet. Weiterhin können diese Gesten in Kategorien unterteilt und beschrieben werden, um das nonverbale, menschliche Verhalten zu ordnen und in eine übersichtliche Struktur zu bringen. Unter anderem beschäftigt sich dieses Kapitel neben bereits durch andere Arbeiten vorhandenen Parametern und ihren Anwendungen auch mit deren Verbesserung und Verfeinerung.

3.1 Gestenanalyse

Um ein Gestenset zu entwickeln, müssen zunächst die Grundlagen gelegt werden, die vor allem aus Erkenntnissen über Gesten bestehen, die oftmals im Alltag verwendet werden, und deren typische Anwendungsszenarien. Im Folgenden werden aus diesem Grund Gesten und ihre Entwicklung analysiert und beschrieben.

3.1.1 Entstehung und Entwicklung von Gesten

Es stellt sich zunächst die Frage, wie Gesten entstehen und wie die Menschen sie entwickeln, um sie schlussendlich in der Kommunikation effektiv einsetzen zu können. Weiterhin wird erörtert, wie es möglich ist, passende Gesten zu menschlichen Intentionen zu finden und darzustellen. Der Vorgang, vom prozessauslösenden Reiz bis hin zur ausgeführten Geste, ist auf Abbildung 3.1 schematisch dargestellt. Die verwendeten Gesten werden in der Regel nicht bewusst, sondern reflexartig, also unbewusst, generiert und eingesetzt.

Der Reiz, den die Gestengenerierung zum Anstoß benötigt, kann unterschiedlicher Art sein. Möglich ist zum einen, dass die Person, die die Geste erzeugt, beabsichtigt etwas verbal oder nonverbal auszudrücken. Somit stammt der Reiz von der Person und ist damit intern. Beispielsweise bemerkt eine Person, dass sie unbedingt ihrem Gegenüber etwas mitteilen möchte. Zum anderen gibt es die Möglichkeit, dass der Reiz durch einen Sachverhalt in der Umwelt der Person ausgelöst wurde. Konkreter können diese vor allem optischer, akustischer oder haptischer Natur sein. Beispiele hierfür sind Dialoge, in denen eine Person auf eine Äußerung des Gegenübers reagieren möchte. Ein optischer Anreiz

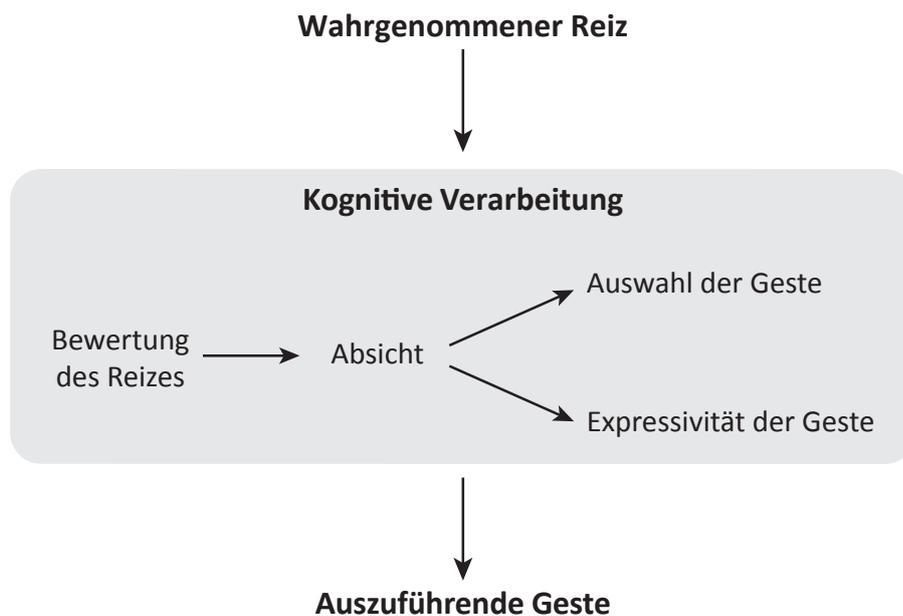


Abb. 3.1: Schematische Darstellung der Entstehung und Entwicklung einer Geste über den wahrgenommenen Reiz, die kognitive Verarbeitung durch das menschliche Gehirn, bis hin zur ausführbaren Geste.

besteht aus einer Beobachtung der Person, wodurch sie sich einer anderen Person darüber mitteilen möchte. In einem weiteren, haptischen Fall wird eine Person unerwartet von einer anderen Person am Arm berührt und reagiert auf dieses überraschende Ereignis.

Nach der Wahrnehmung des Auslösers folgt die interne beziehungsweise kognitive Verarbeitung des Reizes im Gehirn. Zunächst findet eine Bewertung des ursprünglichen Reizes statt, so dass sich eine Absicht daraus bildet. Eine Bewertung kann bei einer Berührung am Arm zum Beispiel ein überraschendes bis hin zu einem ablehnenden oder aggressiven Verhalten sein, falls es sich um eine Person handelt, zu der kein freundschaftliches Verhältnis besteht. Die Absicht kann man folglich als einen Ausdruck wahrnehmen, den man als Reaktion auf den Reiz zeigen wird.

An dieser Stelle werden zwei verschiedene Komponenten ausgewählt. Zum einen wird eine passende Geste ausgewählt, mit der man die gewünschte Absicht zum Ausdruck bringen kann. Der detaillierte Ablauf der Auswahl der Geste wird jedoch nicht betrachtet, da dies nicht Teil dieser Arbeit ist. Zudem ist laut Hartmann et al. [Har06] die genaue Ausführung der Gesten von Emotionen, der eigenen Persönlichkeit, der Kultur, der sozialen Rolle, dem Geschlecht und semantischen und pragmatischen Aspekten der Äußerung selbst abhängig.

Zum anderen entsteht eine Expressivitätsstufe, die mit der Ausdrucksstärke einer Geste gleichzusetzen ist. Beispielsweise kann die Geste „Winken“, wie alle Gesten, in ihrer Expressivität gesteigert werden, so dass sie einen stärkeren Ausdruck mit sich bringt.

Ein eher zaghaftes Winken, das nur auf Brusthöhe ausgeführt wird, hat nur eine geringe Expressivität, wohingegen ein Winken, welches über dem Kopf mit gestrecktem Arm und schnellen Bewegungen ausgeführt wird, eine sehr hohe Expressivität hat. Diese Variation in der Ausdrucksstärke wird durch Werte verschiedener Parameter definiert. Die einzelnen Parameter selbst werden in Kapitel 3.2 noch analysiert, definiert und kategorisiert.

Am Ende dieses Vorgangs steht die Geste, die sich aus den bereits besprochenen Bestandteilen zusammensetzt und dadurch von einer Person ausgeführt werden kann. Somit wird im Anschluss an den kognitiven Vorgang die gewählte Zielgeste in Gelenkpositionen und -bewegungen umgesetzt. Zusammengefasst bildet somit eine Geste eine Ausgabe, die aus äußeren Einflüssen und einer internen Verarbeitung mit diversen Parametern besteht.

Die auf Abbildung 3.2 zu sehende Hierarchie zeigt die Ebenen, die bei dem Vorgang der Gestenentstehung durchlaufen werden. Diese erstreckt sich vom allgemeinen Ausdruck, wie Wut oder Fröhlichkeit, bis hin zu einzelnen Gelenkpositionen und -bewegungen, die beispielsweise mit Winkeln oder Bewegungslinien angegeben werden können und die auf eine spezielle Bewegung zugeschnitten sind. Eine relativ allgemein gehaltene Geste ist unter anderem das „Jubeln“. Gestenprimitive können bei dieser Geste „Hände zu Fäusten ballen“ und „Arme in die Höhe heben“ sein. Diese werden daraufhin in einzelne Positionen der Gelenke übersetzt.

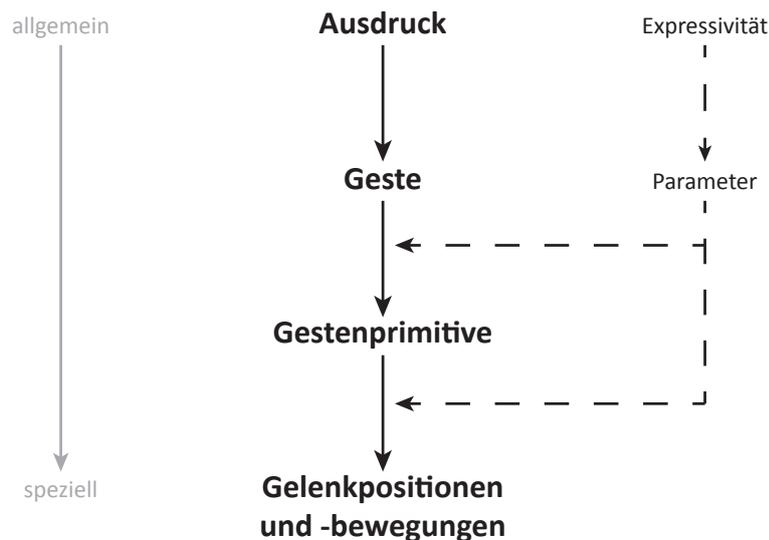


Abb. 3.2: Die Hierarchie innerhalb der Kommunikation mit Gesten unterteilt sich vom Allgemeinen zum Speziellen in vier unterschiedliche Komponenten: Ausdruck, Geste, Gestenprimitive und die einzelnen Gelenkpositionen und -bewegungen.

Selbstverständlich können sich Ausdrücke überlagern, so dass eine Person zur gleichen Zeit eine Kombination aus Wut und Traurigkeit ausdrückt. Jedoch wird dieser Punkt für die weitere Arbeit vernachlässigt, da es die Komplexität außerordentlich steigern würde.

Für eine effektive Produktion eines parametrisierten Gestensets für einen humanoiden Roboter ist die Umsetzung und Adaption dieses beschriebenen Vorgangs notwendig. Jedoch soll in dieser Arbeit nicht der Reiz oder dessen Bewertung im Fokus liegen, sondern die Verarbeitung einer Absicht zu einer möglichst aussagekräftigen Geste. Genau dieser Teil der kognitiven Verarbeitung des Reizes, der eine Geste auslöst, soll auf einem humanoiden Roboter mit Hilfe von Parametern nachgebildet werden.

3.1.2 Analyse alltäglicher Gesten

Bei der Beobachtung alltäglicher Interaktion zwischen Menschen wird deutlich, dass mehr Gesten und nonverbale Ausdrücke verwendet werden, als man annehmen würde. Gesten sind allgegenwärtig und treten in nahezu jeder Kommunikation mit anderen Menschen auf. Wenn sich die Interaktionspartner nicht unmittelbar sehen, zum Beispiel bei einem Telefonat, verwenden sie Gesten. Der Grund hierfür ist wahrscheinlich die Gewohnheit, mit der man in normalen Dialogen nonverbale Signale aussendet. Somit besitzt jeder Mensch ein Gestenset, welches essentiell für eine flüssige Kommunikation ist. Voraussetzung dafür ist das Verständnis des Gegenübers für das Gestenset.

Beispielsweise haben Xu et al. [Xu07] untersucht, inwiefern Gesten einen eigenen Kommunikationskanal darstellen. Überprüft wurde eine Kommunikation zwischen zwei Personen, die ausschließlich aus Gesten bestand, und deren Verständlichkeit, um herauszufinden, ob dies als eigener Kanal gelten kann. Das Ergebnis der Studie war positiv, so dass sich zwei Personen nur mittels Gesten verständigen können. Jedoch ergeben sich Probleme, wenn die Beteiligten komplexere Aufgaben ausführen und diese einem Partner mitteilen sollen. In diesem Fall, bei ungenügender Verständigung durch Gesten, muss die Sprache als zweiter, ergänzender und unterstützender Kanal hinzugezogen werden.

Alltägliche Gesten treten in vielen Formen auf und lassen sich untergliedern, indem man sie nach ihren zugehörigen Bedeutungen sortiert. Natürlich können manche Gesten mehrere Ausdrücke übermitteln, wobei sie in Kombination mit anderen Gesten, oftmals eine eindeutige Ausdrucksweise bilden.

Um die Gesten in ihrer alltäglichen Verwendung zu kategorisieren, werden in erster Linie die sechs Primär- oder Hauptemotionen verwendet, die von Ekman [Ekm92] definiert wurden. Hierbei handelt es sich um Fröhlichkeit, Traurigkeit, Wut, Angst, Ekel und Überraschung. Anzumerken ist, dass Emotionen nicht immer die Basis für das Erzeugen einer Geste sind und Gesten auch durch andere kognitive Prozesse hervorgerufen werden können. Des Weiteren werden zusätzlich noch die Ausdrücke „Unsicherheit“ und „Langeweile“ hinzugefügt, um die Variabilität der Kommunikation zu konkretisieren und zu erweitern. Dieser Vorgang findet bereits bei der Erstellung einer Gestendatenbank von Gunes & Piccardi [Gun06] Verwendung.

Die unterschiedlichen Ausdrücke und ihre passenden Gesten und Gestenprimitive sind in Tabelle 3.1 aufgelistet. Diese unterschiedlichen Gesten können sowohl einzeln als auch in Kombination auftreten. Einzelne Gesten können mehrdeutig sein und stellen verschiedene, teilweise widersprüchliche Ausdrücke dar. Deswegen ist es in der Kommunikation oftmals sinnvoll, nicht nur eine Geste für einen Ausdruck zu verwenden, da diese missverstanden werden kann. Personen nutzen diese Gesten unbewusst in ihren Kombinationen,

jedoch muss humanoiden Robotern diese Anwendung erst „beigebracht“ werden. Zur Veranschaulichung dient die offene Körperhaltung, die sich durch einen festen Stand, aufrechte Haltung und nicht-hängende Schultern auszeichnet. Sie kann Fröhlichkeit und Wut darstellen. Erst durch weitere, verwendete Gesten wird der Ausdruck deutlicher und differenzierter. Das Hinzunehmen einer geballten Faust wirkt weiterhin nicht deutlich, denn erst die Anwendung dieser führt zu einer eindeutigen Ausdrucksweise, da sie entweder zum Jubeln oder zum Drohen beziehungsweise zum heftigen Schlagen auf einen Tisch verwendet werden kann.

Tab. 3.1: *Ausdrücke und Absichten einer Handlung und ihre passenden Umsetzungsformen in Gestik und nonverbalem Verhalten*

Ausdruck	Entsprechende Gesten und Körperhaltungen
Neutral	kein besonderer Gesichtsausdruck Hände entspannt fester Stand
Fröhlichkeit	offene Körperhaltung angehobene Arme, die vom Körper entfernt sind Klatschen Hände zu Fäusten geballt Hüpfen von einem Bein auf das andere springen
Traurigkeit	geschlossene Körperhaltung Verlagerung des Körpergewichts langsame Bewegungen geneigter Kopf Oberkörper nach vorne gebeugt Teile des Körpers bedeckt Hände tiefer halten als normal Hände berühren Kopf/Hals Hände bedecken Gesicht Hände über dem Kopf Hände geschlossen und zusammengeführt gesenkte Schultern

Ausdruck	Entsprechende Gesten und Körperhaltungen
Wut	<ul style="list-style-type: none"> offene Körperhaltung Hände geschlossen/zur Faust geballt Hände angehoben bebende/zitternde Hände/Finger Hände an der Hüfte/Taille Handflächen nach unten zeigend mit dem Finger zeigen verschränkte Arme fester Stand mit Fuß aufstampfen mit schweren Schritten gehen
Angst	<ul style="list-style-type: none"> geschlossene Körperhaltung Zurückweichen Arme um den Körper gelegt Hände oberhalb des Kopfes Hände geschlossen/zur Faust geballt Hände bedecken Kopf/Hals/Gesicht Gesicht durch Hände bedeckt Beine eng aneinander
Ekel	<ul style="list-style-type: none"> Zurückweichen Bewegung nach links und rechts Änderung der Ausrichtung Hände in der Nähe des Körpers Hände bedecken Kopf/Hals/Mund Hände angehoben ein Bein angehoben Zeigefinger in Mund stecken (Andeutung von Erbrechen)
Überraschung	<ul style="list-style-type: none"> Zurückweichen langsames Kopfschütteln Hände fassen an Kopf/Mund/Gesicht Hände oberhalb des Kopfes Hände angehoben Hände bedecken Wangen/Mund Stolpern

Ausdruck	Entsprechende Gesten und Körperhaltungen
Langeweile	Verlagerung des Körpergewichts Bewegung nach links und rechts Änderung der Ausrichtung Hand unter dem Kinn, Ellenbogen aufgestützt Hände hinter dem Kopf überkreuzte Beine Umherlaufen
Unsicherheit	unruhige Körperhaltung geneigter Kopf am Kopf kratzen Schulterzucken beide Hände hinter dem Kopf Hände berühren Ohren/Kinn/Stirn/Hals Nägelkauen mit Fingern auf Gegenstände klopfen Hand unter dem Kinn, Ellenbogen aufgestützt nicht Stillstehen Körper verlagert sich von einem Bein auf das andere

Jedoch wurden die Gesten der Tabelle und ihre Bedeutungen nicht überprüft, sondern sind durch Allgemeinwissen über das menschliche Verhalten und durch eine bestehende Übersicht von Gunes & Piccardi [Gun06] entstanden und zusammengefasst worden. Folglich muss für weitere Arbeiten eine Übersicht mit validierten Ausdrücken und zugehörigen Gesten und Bewegungen erstellt werden, die im Anschluss evaluiert werden muss.

Eine weitere Schwierigkeit stellen die Ausdrücke Ironie und Sarkasmus durch ihre Doppeldeutigkeit der Gesten dar. Denn, wenn man erneut das vorige Beispiel verwendet, kann die Geste für „Jubeln“ (geballte Fäuste in die Höhe heben) sowohl als Jubeln selbst verwendet werden als auch für ironisches, wobei man sich nicht über eine Tatsache freut, sondern eher das Gegenteil der Fall ist.

3.1.3 Analyse von Gesten zur Feedback-Gestaltung

An die Analyse der alltäglichen Gesten und der Kategorisierung nach Emotionen beziehungsweise Ausdrücken schließen sich Gesten zur Feedback-Gestaltung an und die Frage, wie man diese unterteilen kann und wann man sie benötigt.

Bei der Arbeit mit einem Roboter kann es zu Situationen kommen, in denen ein Missverständnis zwischen Nutzer und Maschine entsteht. Dabei weiß der Mensch nicht, was im Einzelnen in dem Roboter vorgeht und warum er eine bestimmte Handlung ausführt oder auch nicht. Das Gestenset des Roboters ist folglich noch nicht ausgereift,

das heißt, dass es ist für den Menschen nicht interpretierbar ist. An dieser Stelle fehlt das Verständnis für den Zustand des Roboters, welches der Mensch benötigt. Jedoch kann der Mensch nicht die inneren Vorgänge eines Roboters erkennen. Somit muss der Roboter in diesen Situationen von selbst reagieren und seinen Zustand mitteilen, um dem Nutzer zu vermitteln, aus welchem Grund er bestimmte Aktionen ausführt oder Reaktionen zeigt.

Feedback, was wörtlich übersetzt „Rückmeldung“ bedeutet, wird benötigt und muss generiert und eingesetzt werden, um dem Menschen einen besseren Einblick in die inneren Vorgänge des Roboters zu gewähren. Anwendungsfälle für Feedback sind vor allem die Interaktionen mit einem Roboter und seine generelle Funktionsweise, da der Nutzer zum Beispiel bei Fragen ein Feedback in Form einer Antwort erhalten möchte. Weiterhin möchte der Nutzer ebenso in Kenntnis gesetzt werden, ob der Roboter eine zu erledigende Aufgabe ausführen kann und wann die Ausführung erfolgreich war.

Das Feedback kann mit verschiedenen Mitteln dargestellt werden. Zum einen existiert die Möglichkeit einer Sprach- oder einer einfachen Textausgabe. Eine weitere Art ist ein farbcodiertes Feedback, das zum Beispiel auf den Ampelfarben basiert. Zum anderen kann man Gesten nutzen, um den Nutzer über Zustände des Roboters in Kenntnis zu setzen. Darüber hinaus gibt es noch andere Arten, Feedback darzustellen, zum Beispiel Musik und Geräusche.

Arten des Feedbacks sind nicht alleinstehend, sondern bedienen sich der Ausdrücke und sind an diese angelehnt. In Tabelle 3.2 werden die äquivalenten Ausdrücke zu Feedback-Arten dargestellt und veranschaulicht. Allerdings weist zum Beispiel negatives Feedback nur eine geringe Ähnlichkeit und Überschneidung mit den Gesten, die für Ekel verwendet werden, auf. Gegensätzlich hierzu steht das Feedback für Wartend/Neutral, da diese Kategorie dem neutralen und langweiligen Ausdruck entspricht.

Tab. 3.2: *Arten des Feedbacks und die Ausdrücke, die sie beeinflussen beziehungsweise zu denen sie eine Ähnlichkeit aufweisen*

Art des Feedbacks	Positiv	Negativ	Unverständnis	Wartend/Neutral
beeinflussender Ausdruck	Fröhlichkeit	Ekel	Unsicherheit	Neutral/Langeweile

Hinzuzufügen sind der Tabelle zwei weitere Arten des Feedbacks, wenn es vor allem um die Kommunikation mit humanoiden Robotern geht. Ein „Fehler“ sollte als eigene Feedback-Art gelten, da die Stellung des Fehlers als eine besondere herausgebildet wird, obwohl es sich dabei lediglich um ein stark negatives Feedback handeln könnte. Die zweite Art, die zusätzlich eingeführt werden soll, ist „Beginn/Ende“, denn der humanoide Roboter soll einer Person vermitteln können, ab welchem Zeitpunkt er seine Aufmerksamkeit dem Gesprächspartner schenkt, das heißt wann er einsatzfähig ist. Gleiches gilt für das Ende einer Kommunikation. Jedoch entsprechen die Gesten, die für den Beginn oder das Ende einer Konversation oder Interaktion verwendet werden, keinem

der in Kapitel 3.1.2 genannten Ausdrücke. Für diese Feedback-Arten werden spezielle, einzigartige Gesten verwendet, die oftmals nur diesem Verwendungszweck dienen.

Zusammenfassend ergeben sich somit sechs verschiedene Feedback-Arten, die ein humanoider Roboter benötigt, um einem Interaktionspartner ausreichend Informationen über seinen inneren Zustand zu übermitteln:

- **Positiv** oder Akzeptanz
- **Negativ** oder Abweisung
- **Fehler**
- **Unverständnis**
- **Wartend/Neutral**
- **Beginn/Abschluss** oder Begrüßung/Verabschiedung

Diese Kategorien beziehen sich auf das extrinsische Feedback, welches eine Reaktion auf äußere Reize darstellt. Das intrinsische Feedback hingegen soll den inneren Zustand des Roboters ausdrücken, der ohne direkte, äußere Einflüsse vorhanden ist, wie zum Beispiel Emotionen. Diese Art findet jedoch an dieser Stelle jedoch weniger Betrachtung.

Bei der Zuordnung einzelner Gesten und Gestenprimitiven zu Feedback-Arten in Tabelle 3.3 fallen vor allem die Überschneidungen zwischen den Ausdrücken und den Kategorien des Feedbacks auf, die in Tabelle 3.2 beschrieben wurden. Diese Gesten können, wie erwähnt, auch kombiniert werden, um die Expressivität zu erhöhen und zu differenzieren. Mit Hilfe der verwendeten Gliedmaßen, zum Beispiel der Hände oder der Beine, kann weitere Variation in der Anwendung erzeugt werden und die Expressivität kann zusätzlich gesteigert oder abgeschwächt werden.

Tab. 3.3: Arten des Feedbacks und Möglichkeiten diese Rückmeldung mit Hilfe von Gesten auszudrücken

Art des Feedbacks	Entsprechende Gesten und Körperhaltungen
Positiv oder Akzeptanz	OK-Geste (Zusammenführung des Daumens und des Zeigefingers) High-Five Nicken Körper zum Gegenüber ausgerichtet dem Gegenüber auf die Schulter klopfen Faust mit Daumen nach oben offene Körperhaltung Klatschen Hände zu Fäusten ballen und heben (Jubeln) Hüpfen

Art des Feedbacks	Entsprechende Gesten und Körperhaltungen
Negativ oder Abweisung	Hand mit der Handfläche zum Gegenüber Kopfschütteln erhobener Zeigefinger Hände vor das Gesicht heben Faust mit Daumen nach unten mit Fuß aufstampfen Zurückweichen hängende Schultern verschränkte Arme
Unverständnis	Kopf wiegen Kopf schieflegen Schulterzucken Schultern anheben Arme angewinkelt, Handflächen nach oben zeigend Hand ausstrecken und horizontal kippen am Kopf kratzen Hand unter dem Kinn, Ellenbogen aufgestützt
Wartend/Neutral	Hände entspannt fester Stand Verlagerung des Körpergewichts Änderung der Ausrichtung Arme in die Hüfte gestemmt Hand unter dem Kinn, Ellenbogen aufgestützt Hände hinter dem Kopf umherschweifender Blick Umherlaufen mit dem Fuß klopfen
Beginn/Abschluss oder Begrü- bung/Verabschiedung	Winken Verbeugen Hand zum Gruß anheben Umarmen „Luftkuss“ Hand reichen und schütteln Zunicken Gegenüber „heranwinken“ Körper zuwenden/abwenden Aufstehen/Hinsetzen Küsschen

Schlussendlich ist ein weiteres Merkmal aller Gesten, dass sie vor allem kulturell abhängig sind. Beispielsweise gelten in Europa für bestimmte Gesten andere Bedeutungen als in anderen Kontinenten und Kulturen der Welt, wobei die unterschiedlichen Interpretationen teilweise schon zwischen Nachbarländern auftreten.

3.2 Parametrisierung von Gesten

Nachdem diverse Gesten gesammelt, analysiert und in Kategorien unterteilt wurden folgt die eigentliche Parametrisierung dieser Gesten. Zunächst findet eine Analyse statt, welche sich auf die möglichen Parameter bezieht und diese bewertet. Auf Grund der vorangegangenen Bewertung werden die relevanten Parameter ausgewählt und anhand ihrer Eigenschaften erneut unterteilt. Als letzte Stufe der Parametrisierung findet die Zuordnung der Parameter zu den Gesten statt, wobei anschließend dargelegt wird, wie eine Geste mittels Parameter in ihrer Expressivität gesteigert werden kann.

3.2.1 Analyse und Bewertung möglicher Parameter

Als Parameter sollen die Eigenschaften festgelegt werden, die die einzelnen Gesten in ihrer Ausdrucksstärke beeinflussen und diese stärken oder schwächen können. In verschiedenen Arbeiten, auf die im Kapitel 2.3 übergreifend eingegangen wurde, wurden Parameter erforscht, die diese Aufgabe übernehmen.

Insbesondere Hartmann et al. [Har06] entwickelten für die GRETA-Plattform [Nie09] Parameter, die die Ausdruckskraft einer Geste beschreiben. Sie stellen in ihrer Arbeit die folgenden sechs Parameter auf:

- **Overall Activation:** Anteil der Bewegungen in einer Konversation (passiv/statisch oder animiert/engagiert)
- **Spatial Extent:** Ausdehnung der Bewegungen (Wie viel Raum wird vom Körper eingenommen?)
- **Temporal Extent:** Dauer der Bewegungen (zum Beispiel schnelle oder längere Handlungen)
- **Fluidity:** Stetigkeit und gleichmäßiger Fluss der Bewegungen (weich oder ruckartig)
- **Power:** dynamische Eigenschaften der Bewegung beziehungsweise Beschleunigung und Abbremsen (schwach oder stark)
- **Repetition:** Tendenz zu rhythmischer Wiederholung bestimmter Bewegungen

Alle Parameter außer „Overall Activation“ können sich innerhalb $[-1, 1]$ befinden, wobei 0 für eine Handlung ohne Expressivität steht. „Overall Activation“ hat einen Wert zwischen 0 und 1. Die Null hat zur Folge, dass kein nonverbales Verhalten generiert

wird. Der Parameter „Overall Activation“ wird durch das Taggen von Wörtern der Texteingabe generiert. Die getaggten Wörter verfügen über eine Intensitätsangabe, aus der der Parameterwert berechnet wird. „Spatial Extent“, „Temporal Extent“, „Fluidity“ und „Power“ werden mit Hilfe mathematischer Formeln und Analysen der Bewegungen berechnet. Einen Eindruck der Parameter erhält man am Beispiel von „Spatial Extent“ durch Abbildung 3.3.

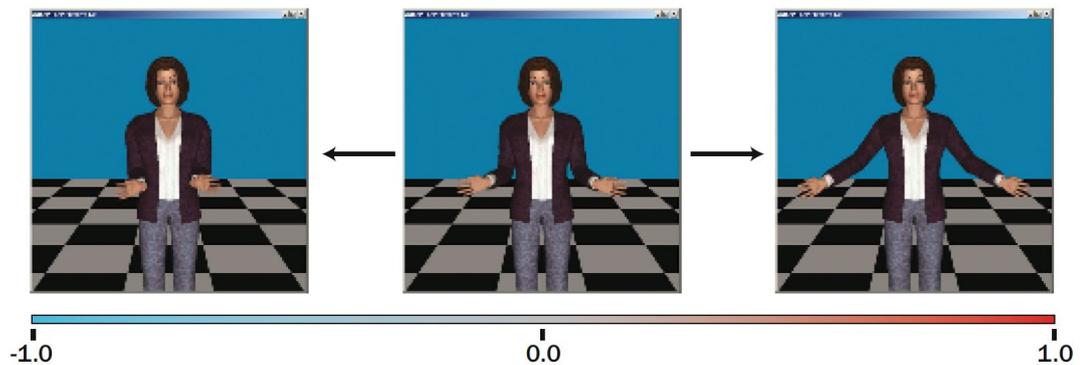


Abb. 3.3: Der Parameter „Spatial Extent“ der GRETA-Plattform kann Werte zwischen $-1,0$ und $1,0$ annehmen. Bei $-1,0$ ist die Geste wenig ausdrucksstark und nimmt sehr wenig Platz ein. Im Gegensatz dazu sieht man bei einem Wert von $1,0$, dass hier die Geste sehr deutlich gezeigt wird und somit auch viel Platz in Anspruch nimmt.

Weiterhin werden Parameter in dem in Kapitel 2.3 besprochenen Projekt von Ng-Thow-Hing et al. [NTH10], bei dem ebenso wie bei der GRETA-Plattform ein zu einem Text passendes Gestenset erstellt werden soll, verwendet. Vorgestellt werden folgende Parameter:

- **Amplitude:** vergleichbar mit „Spatial Extent“ der GRETA Plattform
- **Frequency:** vergleichbar mit „Repetition“ der GRETA Plattform
- **Tension:** vergleichbar mit „Power“ der GRETA Plattform
- **Continuity:** vergleichbar mit „Fluidity“ der GRETA Plattform
- **Bias:** räumliche Ausrichtung der Geste, wie zum Beispiel nach links oder nach rechts (kein vergleichbarer Wert bei der GRETA Plattform)

Die drei zuletzt aufgeführten Parameter werden für jeden Frame der Bewegung einzeln definiert, wobei zusätzlich die Zeit pro Frame und ein zufällig generierter Offset zur Startpunktberechnung berücksichtigt werden. Diese Parameter werden verwendet, um Eigenschaften der Gesten, wie „ruhig“, „neutral“ und „aufgeregt“ zu erzeugen, die mit Hilfe des Expressivitätsparameters noch variiert werden können. Bei einer anschließenden Evaluation des Systems wurde herausgefunden, dass schnelle Gesten eher positiv und langsame Gesten eher negativ wahrgenommen werden.

Wie man in der Auflistung der einzelnen Parameter beider Projekte erkennen kann, überschneiden sich die Forschungsergebnisse hinsichtlich ihres Inhalts und ihrer Parameter. Dass sowohl die räumliche Ausdehnung einer Geste als auch die zeitliche Ausdehnung eine Relevanz bezüglich der Ausdrucksstärke haben, ist trivial. Eine Geste, die mehr Platz einnimmt, wirkt in erster Linie in ihrem Ausdruck stärker als eine Geste, die kaum Raum in Anspruch nimmt. Hinsichtlich der Zeit gibt es keine klare Abgrenzung, ob eine schnelle oder eine langsame Geste expressiver ist. Diese Entscheidung hängt von der Geste selbst ab. Eine Zeigegeste wirkt beispielsweise deutlicher, wenn sie langsamer ausgeführt wird, da das Objekt, auf das gezeigt wird, länger im Fokus steht. Andererseits wirkt ein angedeuteter Hieb eindrucksvoller je schneller er ausgeführt wird. An dieser Stelle wird die Wertigkeit des Parameters „Power“ beziehungsweise „Tension“ deutlich. Er ist für die Beschleunigung und das Abbremsen einer Geste zuständig. Folglich sorgt bei dem Beispiel des Hiebs vor allem hier ein hoher Wert für eine starke Ausdruckskraft. Weiterhin gibt es den Parameter „Repetition“ oder „Frequency“, dessen Bedeutung ebenso klar ersichtlich ist. Eine Geste, die öfter wiederholt wird, wie das Anheben der Arme über den Kopf beim Jubeln, wird in ihrem Ausdruck stärker und dem Beobachter wird der gewünschte Ausdruck deutlicher. Ein weiterer Parameter, der nicht allgemein beschrieben werden kann, sondern für jede Geste einzeln zu definieren ist, wird „Fluidity“ oder „Continuity“ genannt. Wie stetig oder gleichmäßig eine Geste ausgeführt wird, hängt mit ihrer Expressivität zusammen. Tendenziell wird eine gleichmäßigere Bewegung eine höhere Ausdrucksstärke hervorrufen als eine ruckartige und stockende Geste. Jedoch kann diese Annahme in einzelnen Fällen abweichen. Zuletzt bleibt „Bias“, das übersetzt Ausrichtung oder Orientierung bedeutet. Eine Geste, die in Richtung des Interaktionspartners ausgeführt wird, ist in der Regel expressiver, als eine Geste, die nicht auf ihn ausgerichtet ist. Dieser Parameter verhält sich bei einigen ablehnenden Aktionen gegenteilig, da die Abwendung vom Partner bereits eine Abweisung darstellt.

Die definierten Parameter beider Forschergruppen wurden behutsam gewählt und decken das Feld der Expressivität ausreichend ab, so dass man mit Hilfe dieser Eigenschaften Variationen innerhalb der Ausdrucksstärke schaffen kann. Dennoch können Probleme durch eine ungenügende Definition der Parameter auftreten. Beispielsweise beeinflussen sich die einzelnen Parameter untereinander, wodurch sie zudem voneinander abhängig sind. Dieses Phänomen tritt hauptsächlich bei den zeitlichen Parametern, zeitliche Ausdehnung, Gleichmäßigkeit und Frequenz, auf. Wenn die Frequenz erhöht wird, steigt die generelle Dauer der Geste und mit dem Sinken der Gleichmäßigkeit, nimmt die Dauer erneut zu. Des Weiteren ist der Parameter „Bias“ ambivalent, da er zwar eine Komponente der Expressivität darstellt, aber nur für wenige und spezielle Gesten, wodurch er unter Umständen zu vernachlässigen ist, vor allem wenn es sich in einem bestimmten Anwendungsfall um lediglich zwei Interaktionspartner, den Roboter und einen Menschen, handelt.

3.2.2 Auswahl und Kategorisierung der Parameter

Um die in Kapitel 3.2.1 genannten Vorteile zu erhalten und die Probleme zu lösen, werden an dieser Stelle keine neuen Parameter entwickelt, da sie bereits in einer umfassenden

Form vorliegen, sondern sie werden erneut definiert, differenziert und anschließend in sinnvolle Kategorien unterteilt.

Trotz ihrer anscheinenden Sinnhaftigkeit werden nicht alle Parameter übernommen, da wie angesprochen vor allem der Parameter „Bias“ zu wenig Aussagekraft für den hier betrachteten Anwendungsfall besitzt. Somit wird dieser aus genannten Gründen nicht in die Betrachtung mit aufgenommen. Ein weiterer Grund dafür ist, dass man die Ausrichtung mit in die eigentliche Geste integrieren kann und sie nicht anhand ihrer Richtung parametrisieren und variieren können muss.

Ebenso kann der Parameter „Overall Activation“ vernachlässigt werden, da dieser angibt, ob und wie viel nonverbales Verhalten innerhalb des Gestensets verwendet werden soll. Diese Eigenschaft wird für die Entwicklung eines parametrisierten Gestensets nicht benötigt, da generell nonverbales Verhalten erwünscht ist und man mittels Expressivität in dessen Qualität variieren kann.

Zur weiteren Auswahl und Parametrisierung der Gesten werden somit folgende Parameter zur Betrachtung herangezogen und im weiteren Verlauf verwendet:

- **Räumliche Ausdehnung** (Spatial Extent)
- **Zeitliche Ausdehnung** (Temporal Extent)
- **Beschleunigung** (Acceleration)
- **Frequenz** (Frequency)
- **Stetigkeit** (Continuity)

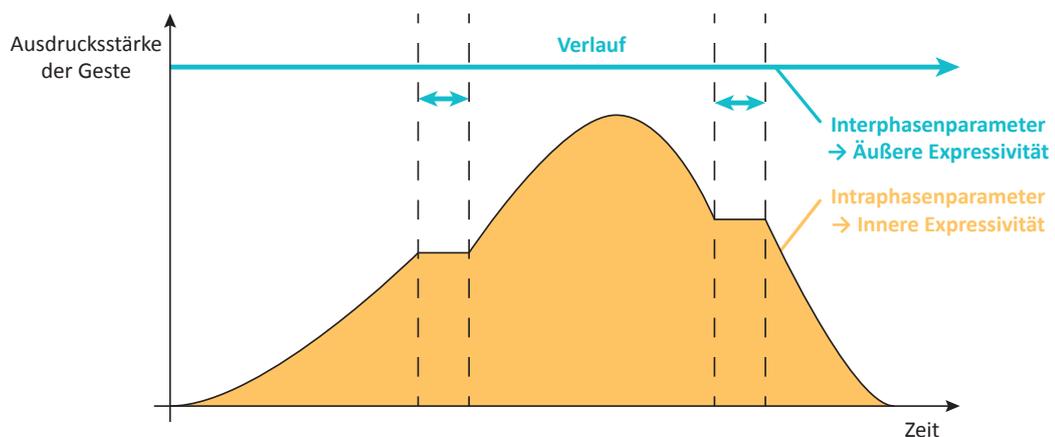


Abb. 3.4: Darstellung der Einflüsse der unterschiedlichen Parametergruppen, Intraphasenparameter und Interphasenparameter, auf die Gestenphasen und die Expressivität der Geste im Einzelnen und deren Zusammensetzung

Bevor diese Parameter näher erläutert und definiert werden, soll auf ihre allgemeine Bedeutung eingegangen werden, da sie in zwei verschiedene Gruppen unterteilt werden

können. Einerseits können räumliche Ausdehnung, zeitliche Ausdehnung und Beschleunigung als **Intraphasenparameter** bezeichnet werden, andererseits sind Stetigkeit und Frequenz als **Interphasenparameter** definiert. Diese Kategorisierung bezieht sich auf den Einflussbereich der einzelnen Parameter. Als Referenzrahmen dienen die von Kendon [Ken04] festgelegten Phasen einer Geste (siehe Kapitel 2.1). Die zuerst genannte Parameter-Gruppe bezieht sich auf die Expressivität der Geste selbst und auf die einzelnen Phasen, Preparation, Stroke und Recovery, und deren Ausführung. Dem entgegen steht die zweite Parameter-Gruppe, die vor allem für den Verlauf der Geste und die Beziehungen der einzelnen Phasen untereinander verantwortlich ist. Wenn sich die Intraphasenparameter verändern, steigt beziehungsweise sinkt dadurch die sogenannte innere Expressivität der ausgeführten Geste. Jedoch kann die gesamte Expressivität zusätzlich durch die Interphasenparameter manipuliert werden, indem man den Verlauf der Geste verändert. Beide Kategorien der Parameter und ihre jeweils durch diese bestimmten Ausdrucksstärken lassen sich zur allgemeinen Expressivität einer Geste zusammenfügen. Dieser Sachverhalt wird in Abbildung 3.4 aufgegriffen. Zudem erhalten alle Parameter Werte innerhalb des Intervalls $[0, 1]$.

Intraphasenparameter

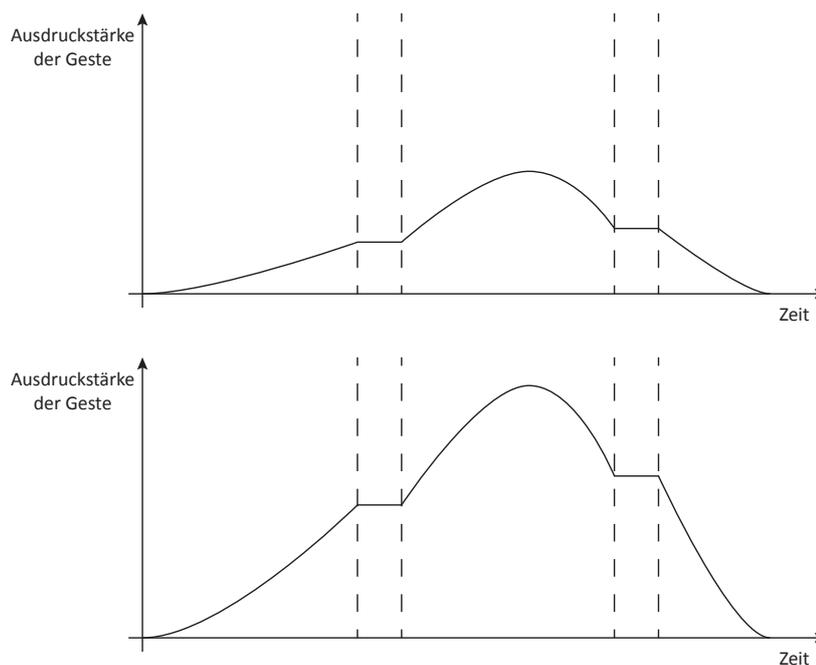


Abb. 3.5: Darstellung der Intraphasenparameter und ihres Einflusses auf die Expressivitätskurve einer Geste: Zum einen werden Parameterwerte verwendet, die eine geringe Expressivität hervorrufen, und zum anderen Werten, die für eine hohe Expressivität verantwortlich sind.

Die Intraphasenparameter beeinflussen die innere Expressivität der Geste und die Ausführung der Bewegungen, die eine Geste beinhaltet. Sie stellen die erste Stufe der Ausdrucksstärke dar. Zu der Gruppe der Parameter, die innerhalb einer Geste wirken, gehören räumliche Ausdehnung, zeitliche Ausdehnung und Beschleunigung, deren Einfluss auf die Geste zusätzlich auf Abbildung 3.5 zu sehen ist.

Räumliche Ausdehnung Die räumliche Ausdehnung beschreibt wie großflächig eine Geste ausgeführt wird. Bei dem Wert 0 wird die kleinstmögliche Ausdehnung verwendet, so dass die Geste gerade noch zu erkennen ist, und bei einem Wert von 1 wird der maximal verfügbare Raum verwendet, der durch die Umgebung oder die eigene Reichweite der Gliedmaßen bestimmt wird.

Zeitliche Ausdehnung Die Dauer einer Geste wird durch die zeitliche Ausdehnung beschrieben, wobei bei einem Wert von 0 die Geste eine möglichst geringe Dauer hat und somit schnell ausgeführt wird. Bei einem Wert von 1 ergibt sich die längste Dauer, wobei diese auf einen Maximalwert begrenzt sein muss.

Beschleunigung Wenn eine Geste deutlich und mit großer Spannung ausgeführt wird, ist dafür die Beschleunigung verantwortlich. Eine größere Beschleunigung und schnellere Abbremsung führen zu einem starken Ausdruck. Bei dem Wert 0 findet kaum Beschleunigung statt, da die Bewegung ohne große Geschwindigkeitsunterschiede abläuft. Im Gegensatz dazu wird bei einem Wert von 1 die höchste Geschwindigkeit erreicht, da die Beschleunigung hoch ist.

Interphasenparameter

Im Gegensatz zu den Intraphasenparametern sind die Interphasenparameter sowohl für den Verlauf der Bewegung beziehungsweise der Geste als auch für die Relationen der einzelnen Phasen zueinander verantwortlich. Gebildet wird hier eine zweite Stufe der Ausdrucksstärke, der äußeren Expressivität, die zusammen mit der inneren Expressivität die gesamte Ausdruckskraft einer Geste bestimmt. Zu dieser Kategorie gehören die Frequenz und die Stetigkeit einer Geste.

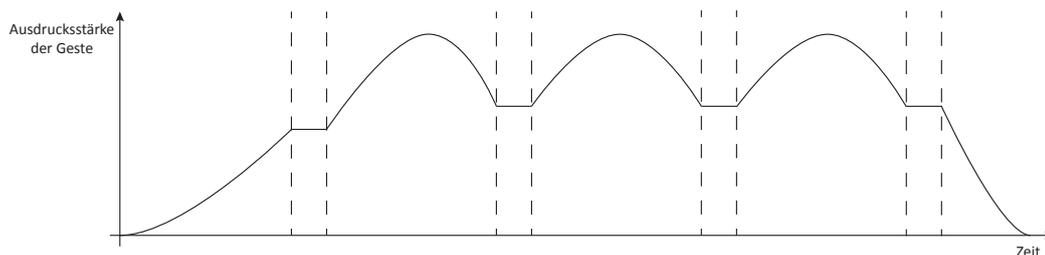


Abb. 3.6: Darstellung des Parameters Frequenz und seinen Auswirkungen auf die Expressivitätskurve durch mehrmalige Ausführung des Strokes einer Geste

Frequenz Als Frequenz wird die Anzahl der Wiederholungen der einzelnen Geste bezeichnet, um diese deutlicher wirken zu lassen. Zu beachten gilt, dass dabei der Stroke wiederholt wird, nicht die anderen Phasen. Bei einem hohen Wert der Frequenz, wie zum Beispiel 1, wird die maximale Anzahl an Wiederholungen des Strokes durchgeführt, die für jede Geste definiert werden muss. Wenn der Stroke zu häufig ausgeführt wird, wirkt die Geste leicht lächerlich oder kann einem zwanghaftem Verhalten ähneln, wodurch die Geste entwertet werden würde. Keine Wiederholungen des Strokes gibt es bei dem Wert 0. Die Auswirkungen der Frequenz auf die Expressivitätskurve sind in Abbildung 3.6 abgebildet.

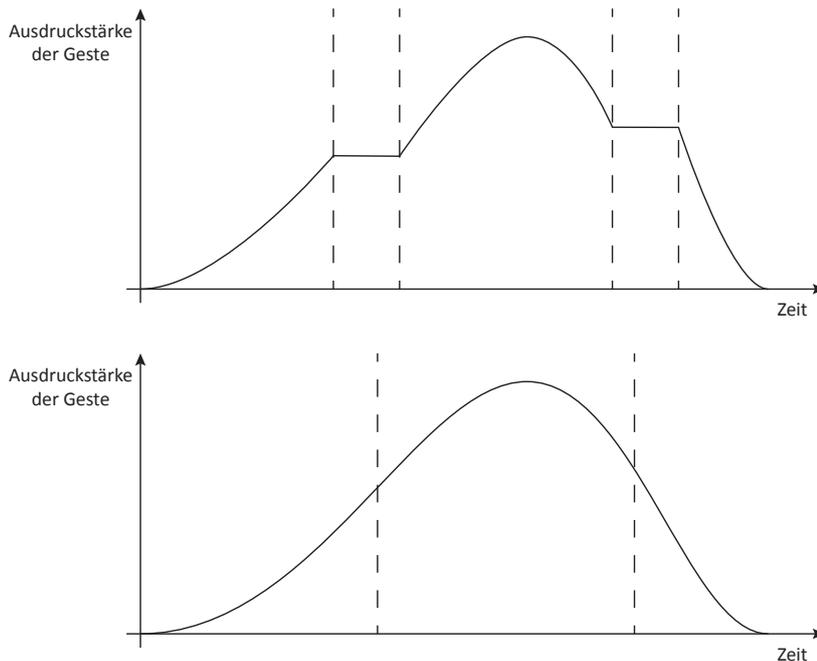


Abb. 3.7: Darstellung des Parameters *Stetigkeit* und seinen Auswirkungen auf die *Expressivitätskurve* einer Geste: Zum einen werden Parameterwerte verwendet, die eine unstete und ruckelige Kurve bilden, und zum anderen Werten, die für eine stetige und flüssige Kurve verantwortlich sind.

Stetigkeit Die *Stetigkeit* legt fest, wie gleichmäßig eine Geste ausgeführt wird. Sie beeinflusst zudem die Dauer der Holds, die zwischen den einzelnen Phasen ausgeführt werden, wie auf Abbildung 3.7 erkennbar ist. Hierzu zählen Pre-Stroke-Hold, Post-Stroke-Hold und zusätzliche Holds zwischen mehreren Strokes, die durch die Frequenz entstehen können. Bei einem Wert von 0 erscheint die Geste durch die starken Verzögerungen durch die Holds ruckartig, wobei beim Wert 1 die Geste fließend wirkt und rein theoretisch keine Holds existieren.

3.2.3 Zuordnung der Parametern zu den Gesten

Tab. 3.4: Zuordnung der Bestandteile der Geste „Winken“ zu den einzelnen Gestenphasen

Phase	Bewegung
Preparation	Anheben der Hand
Stroke	einmalige horizontale Bewegung der Hand
Recovery	Absenken der Hand

Zur Veranschaulichung kann das Beispiel „Winken“ verwendet werden. Die Bewegungen der Geste können zerlegt und den einzelnen Phasen zugeordnet werden, wie in Tabelle 3.4 dargestellt. Wenn der Parameter der räumlichen Ausdehnung erhöht wird, so wird die Bewegung aller drei Phasen deutlicher und nimmt mehr Raum ein. Bei der zeitlichen Ausdehnung wird die Bewegung der Preparation, des Strokes und der Recovery wesentlich langsamer oder schneller ausgeführt. Weiterhin kann man die Beschleunigung steigern, um die Expressivität erheblich zu steigern, da Euphorie vermittelt wird. Somit wird schneller beschleunigt und anschließend schneller abgebremst, wodurch man insgesamt eine höhere Geschwindigkeit der Bewegung erhält. Bei einer hohen Frequenz beziehungsweise bei Wiederholungen wird der Stroke erneut ausgeführt und nicht die gesamte Geste. Wenn die Stetigkeit beim Winken gering ist, haben die Zwischenphasen Pre-Stroke-Hold, Post-Stroke-Hold oder ein Hold zwischen zwei Strokes eine lange Dauer, wodurch die Geste ruckartig erscheint und sie in ihrer Expressivität mindert.

Tab. 3.5: Auswirkungen der Steigerung der Parameter auf verschiedene Gesten zur Darstellung von positivem Feedback beziehungsweise Akzeptanz (↑ = Erhöhung der Expressivität, ↓ = Verringerung der Expressivität, → = keine Veränderung der Expressivität, — = Parameter existiert nicht oder nur bedingt)
(Auszug aus Tabelle A.1)

Steigerung der	Nicken	OK-Geste	Hüpfen
räumlichen Ausdehnung	↑	—	↑
zeitlichen Ausdehnung	↓	↑	→
Beschleunigung	→	↑	↑
Frequenz	↑	↑	↑
Stetigkeit	↑	↑	↑

Im Anschluss soll eine Geste anhand ihrer Parameter in ihrer Ausdrucksstärke gesteigert werden. Beispielfhaft werden die Effekte der Parameteränderungen auf die Expres-

sivität in Tabelle 3.5 abgebildet. Diese Darstellung bezieht sich auf den Ausdruck von positivem Feedback durch Gesten, wie Nicken, der OK-Geste und Hüpfen. Deutlich wird die Beziehung zwischen der Erhöhung einer der Parameterwerte und der Expressivität. Oftmals tragen vor allem Steigerungen der räumlichen Ausdehnung und der Stetigkeit zu einer höheren Ausdruckskraft der Gesten bei.

Die Dauer einer Geste beziehungsweise ihre zeitliche Ausdehnung hat je nach Geste eine andere Bedeutung, wie man in der Tabelle erkennen kann. Beispielsweise ist ein langsames, das heißt länger dauerndes, Nicken relativ ausdruckschwach, wobei eine längere OK-Geste während dieser Zeit an Expressivität gewinnt. Die zeitliche Ausdehnung des Nickens darf nicht mit der Frequenz verwechselt werden, da bei der zeitlichen Ausdehnung ein einzelnes Nicken und dessen Dauer betrachtet wird. Die Dauer des Hüpfens ist nicht frei bestimmbar, da sie von der räumlichen Ausdehnung abhängt.

Ein weiterer Aspekt der unterschiedlichen Parameter ist die Frequenz, die in den meisten Fällen die Ausdruckskraft steigert. In Bezug auf einmalige Gesten, wie die OK-Geste oder „Daumen hoch“ wirkt die Frequenz ebenso verstärkend, jedoch kann es bei wiederholter Ausführung zu einer lächerlichen Wirkung führen.

Wie Gesten jeweils in ihrer Expressivität mit den gegebenen Parametern gesteigert werden können, kann den Tabellen A.1 bis A.5 in Anhang A entnommen werden, wobei diese Bewertungen der Gesten durch eine Evaluation verifiziert werden müssen und teilweise unterschiedlichen Vorstellungen entsprechen, wodurch Konflikte auftreten könnten.

3.3 Kategorisierung und Auswahl relevanter Feedback-Gesten

Nachdem Gesten anhand von Parametern in ihrer Ausdruckskraft verändert werden können, stellt sich die Frage, wie man Gesten zu verschiedenen Ausdrücken sinnvoll auswählt. Zunächst wird auf verschiedene Arten des Ausdrucks und der Darstellung dieser mit Hilfe der Gesten eingegangen, wodurch anschließend das Konzept der Expressivitätsspannen eingeführt und erläutert wird.

3.3.1 Arten des expressiven Ausdrucks

Wenn man verwandte Arbeiten betrachtet (siehe Kapitel 3.2.1), die ebenso Parameter verwenden, fällt auf, dass sie nur eine Geste in ihrer Expressivität mit den Parametern steigern. Jedoch bleibt unklar, wie weit diese Geste gesteigert werden kann und vor allem ob diese gesteigerten, ausdrucksstärkeren Gesten wirklich in der alltäglichen Kommunikation zwischen Menschen Verwendung finden. An dieser Stelle bekommt der Ausdruck beziehungsweise die Absicht, die ein Mensch vermitteln möchte, eine wichtige Bedeutung, da man eigentlich in der Interaktion mit anderen Personen nicht nur eine Geste an sich steigert, sondern den Ausdruck. Beispielsweise will man nicht die Geste „Klatschen“ steigern, sondern man möchte seiner Freude über etwas mehr Ausdruck verleihen, was man auf unterschiedliche Art und Weise verwirklichen kann. Nur eine davon ist die Steigerung des Klatschens.

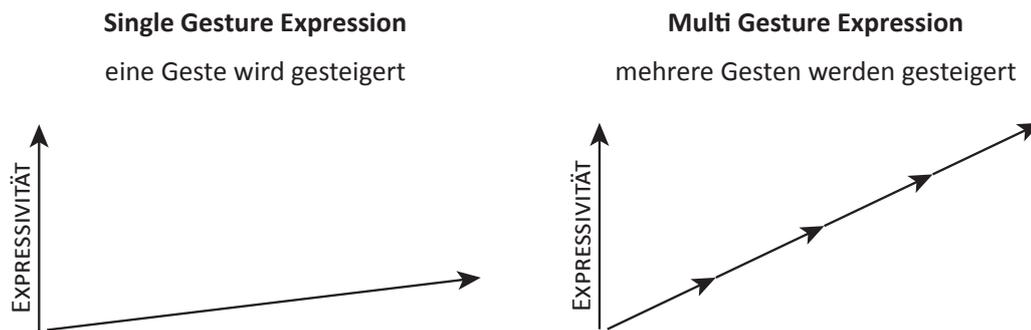


Abb. 3.8: Unterscheidung der beiden Ausdrucksformen „Single Gesture Expression“ und „Multi Gesture Expression“: Die Steigerung bei Single Gesture Expression findet lediglich über eine Geste statt, so dass die Expressivität beschränkt wird. Bei Multi Gesture Expression hingegen werden mehrere Gesten gesteigert und zwischen ihnen je nach Grundexpressivität gewechselt, um eine möglichst hohe Expressivität zu erzielen.

Diese Steigerung eines Ausdrucks kann auf zwei Arten geschehen. Zum einen existiert die Form **Single Gesture Expression**, die die Steigerung beschreibt, die von bisherigen Arbeiten verwendet wurde und die sich nur auf die Ausführung der Geste bezieht. Dabei wird die Ausdruckskraft durch Verstärkung der Geste gesteigert. Zum anderen kann die Art **Multi Gesture Expression** verwendet werden, die zwar ebenso wie die Single Gesture Expression Gesten steigert, sie aber zusätzlich noch in Abhängigkeit der Expressivität abwechselt. Wenn die Stärke des Ausdrucks erhöht werden soll, kann somit unter anderem eine andere Geste gewählt werden, um die gleiche Intention auszudrücken, da unter Umständen die vorher genutzte Geste nicht mehr ausreicht, die gewünschte Expressivität darzustellen. Beide Ausdrucksarten werden auf Abbildung 3.8 schematisch veranschaulicht, wobei man den Expressivitätsgewinn bei Verwendung von Multi Gesture Expression deutlich erkennen kann.

Betrachtet wird beispielsweise der Schlussapplaus bei einem Konzert. Hierbei wird nicht nur das Klatschen innerhalb seiner Parameter gesteigert, sondern es werden andere Ausdrucksweisen verwendet. Zunächst klatscht das Publikum, welches stärker und lauter in seiner Aktion wird. Anschließend steht das Publikum eventuell auf oder Personen heben ihre Hände, schreien oder hüpfen. Diese Variation der Ausdrucksmöglichkeiten und unterschiedlichen Expressivitätsstufen wäre mit bloßem Klatschen nicht möglich.

Die Vorteile, die die Anwendung von Multi Gesture Expression liefert, sind bereits erkennbar. Sie erweitert die Ausdrucksmöglichkeiten enorm und stellt auch eine Variabilität innerhalb eines Ausdrucks bereit, die in dieser Form bei der Verwendung der Single Gesture Expression nicht möglich wäre. Dadurch wird ein Verhalten mit deutlich höheren Abwechslungen und Überraschungen erzeugt. Weiterhin schränkt die Verwendung der Form der Single Gesture Expression die Expressivität auf die Spanne ein, die eine einzelne Geste hat, ohne dabei weitere Gesten zu nutzen, die in ihrer Ausdruckskraft wesentlich höher sind und die Bedeutung des Ausdrucks trotzdem beibehalten. Zusätzlich wird vor allem das Verhalten eines humanoiden Roboters durch die Verwendung von

Multi Gesture Expression unerwarteter und unvorhersehbarer. Somit wirkt er für einen menschlichen Kommunikationspartner wesentlich lebendiger und interessanter, da das Verhalten von Menschen ebenso oftmals unberechenbar, überraschend und selten sogar zufällig oder willkürlich ist.

3.3.2 Kategorisierung der Gesten anhand von Expressivitätsspannen

Zur Bestimmung und Realisierung der Form der Multi Gesture Expression werden Expressivitätsspannen verwendet. Jede Geste kann in ihrer Expressivität mit Hilfe ihrer Parameter und zusätzlich eingesetzter Mimik variiert und manipuliert werden. Dennoch besitzt eine Geste eine bestimmte „Grundexpressivität“, wodurch sie kein endloses Spektrum der Expressivität abdecken kann. Wenn man beispielsweise eine Person begrüßen möchte, hat man die Möglichkeit, dies mittels Winken zu tun, oder man könnte diese Person auch umarmen. Beide Gesten, das Winken und die Umarmung, können in ihrer Expressivität mit den bereits bestimmten Parametern verändert werden. Dennoch hat eine Umarmung eine wesentlich höhere Grundexpressivität im Vergleich zum Winken. Bezüglich eines Ausdrucks, zum Beispiel Fröhlichkeit, oder eines Zwecks, wie einer Begrüßung, besitzt eine Geste eine Expressivitätsspanne, für die diese Geste gilt. Diese Spanne setzt sich aus der genannten Grundexpressivität, die den Anfangspunkt der Spanne darstellt, ab dem die Geste gesteigert werden kann, und aus dem Bereich selbst, der von der Spanne eingenommen wird, zusammen. Dieser Bereich innerhalb der Expressivität kann erreicht werden, indem man die Parameter einer Geste verändert, so dass sich die Expressivität entweder erhöht oder verringert. Zusätzlicher Einsatz von Mimik kann eine Geste ebenso beeinflussen. Beispielsweise ist die Geste „Winken“ wesentlich ausdrucksstärker, wenn man diese mit einem Lächeln unterstützt.

Die Orientierung und Kategorisierung der Auswahl für die alltäglichen Gesten fand an den sechs Hauptemotionen und deren Erweiterungen, die bereits in Kapitel 3.1 verwendet wurden, statt. Für die Feedback-Gesten wurden die definierten Arten des Feedbacks aus dem gleichen Kapitel entnommen, wobei sich zusätzlich durch die Darstellungen der Expressivitätsspannen die Einflüsse der in Tabelle 3.1 aufgeführten Ausdrücke auf die einzelnen Feedback-Arten deutlich herausbilden.

Diese Expressivitätsspannen können jeden Ausdruck oder jede Absicht darstellen und als Schema abgebildet werden. Die verwendeten Gesten für jeden Ausdruck müssen erforscht, festgelegt und anschließend evaluiert werden. In dieser Arbeit würde dies jedoch den Rahmen sprengen, so dass mittels gesundem Menschenverstand und den Meinungen anderer Personen die Schemata der Expressivitätsspannen der Gesten zu den jeweiligen Feedback-Arten erstellt wurden. Jedoch kann es durch dieses Verfahren unter Umständen zu Problemen durch die nicht verifizierte Festlegung der Spannen einzelner Gesten kommen.

Auf Abbildung 3.9 ist ein Schema der Expressivitätsspannen für die Feedback-Art „Positiv“ dargestellt, dem man die Ausdrucksstärke und -spanne der Gesten entnehmen kann, die prinzipiell ein positives Feedback vermitteln können. Die exakte Ein- und Aufteilung der Gesten auf die Expressivitätsskala kann je nach subjektivem Empfinden und den eigenen Vorstellungen einer Geste variieren. Grundlegend gilt, dass lediglich einzel-

Positiv oder Akzeptanz

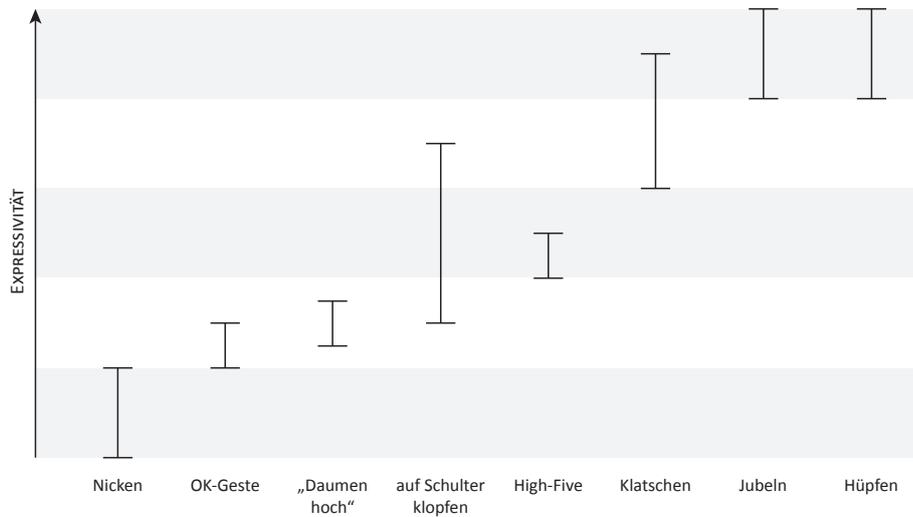


Abb. 3.9: *Expressivitätsschemata zur Darstellung verschiedener Möglichkeiten zur Erzeugung von positivem Feedback oder Akzeptanz durch Verwendung mehrerer Gesten (OK-Geste: Zusammenführen von Daumen und Zeigefinger, High-Five: gegenseitiges Aufeinanderschlagen von zwei erhobenen Händen unterschiedlicher Personen)*

Negativ oder Abweisung

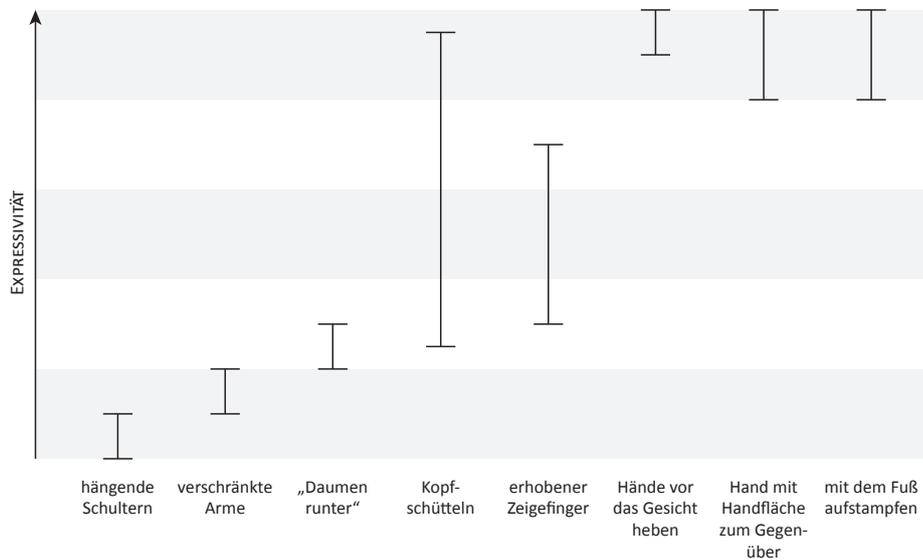


Abb. 3.10: *Expressivitätsschemata zur Darstellung verschiedener Möglichkeiten zur Erzeugung von negativem Feedback oder Abweisung durch Verwendung mehrerer Gesten*

ne Gesten betrachtet werden, welche in Kombination mit weiteren Gesten des selben Ausdrucks zu weit größerer Ausdrucksstärke führen. Eine Kombination aus mehreren Gesten ist somit in der Regel expressiver als eine einzelne Geste. Beispielhaft können die Gesten „Nicken“ und „Klatschen“ betrachtet werden. Getrennt verwendet haben sie eine durchschnittliche Expressivitätsspanne, wobei sie kombiniert unter Umständen eine höhere Expressivität als andere Gesten haben können. Ebenso verhält es sich mit Jubeln (Hände zu Fäusten ballen und heben) und Hüpfen.

Negatives Feedback wird auf Abbildung 3.10 beschrieben, bei dessen Erstellung einige Meinungsunterschiede unter den befragten Personen entstanden. Im Gegensatz zur Erzeugung von positivem Feedback, das sehr eindeutig bestimmt wurde, traten bei negativem Feedback wesentlich feinere Unterschiede innerhalb der Interpretation auf, die sich durch die teilweise sehr große Überlappung der Spannen äußern.

Ähnlichkeit zum Schema des positiven Feedbacks besitzt die Darstellung von Gesten, die Unverständnis darstellen, auf Abbildung 3.11. Vor allem in diesem Fall tritt eine erhebliche Steigerung der Ausdrucksstärke bei Kombination mehrerer Gesten auf, wie beispielsweise Schulterzucken während sich die Person am Kopf kratzt. Eine der stärksten Ergänzungen stellt die Variation mit den beiden ausdrucksstärksten Gesten dar, insofern sie vereinbar beziehungsweise gleichzeitig ausführbar sind.

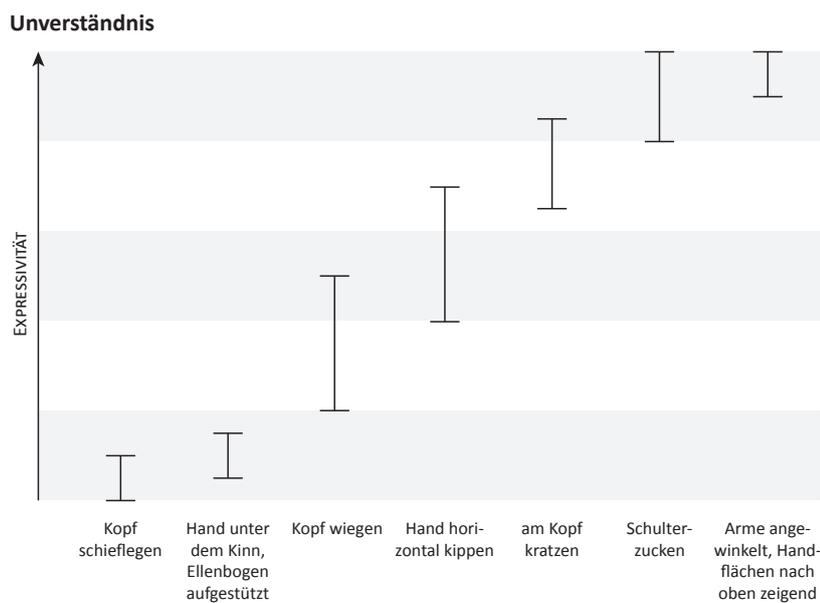


Abb. 3.11: Expressivitätsschemata zur Darstellung verschiedener Möglichkeiten zur Erzeugung von Unverständnis durch Verwendung mehrerer Gesten

Zur Erstellung eines wartenden Ausdrucks wird ein separater Vergleichswert benötigt, mit dem man zeigen kann, dass eine Person länger oder „stärker“ wartet. Dieser Wert ist die Ungeduld. Anhand dieser Eigenschaft wurden die Gesten dieses Ausdrucks in ih-

rer Expressivität bewertet. Vor allem an dieser Stelle werden viele der Gesten, die in Abbildung 3.12 dargestellt sind, gleichzeitig oder abwechselnd ausgeführt, wie beispielsweise der umherschweifende Blick mit Verlagerung des Körpergewichts und zusätzlichem Klopfen des Fußes. Anschließend kann ein Umherlaufen erfolgen, bei dem man die Arme in die Hüfte gestemmt hat.

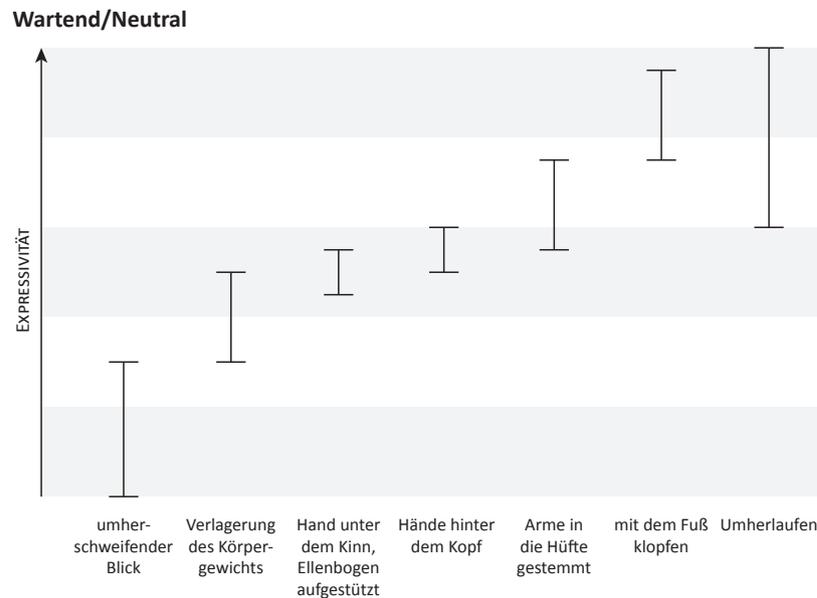


Abb. 3.12: Expressivitätsschemata zur Darstellung verschiedener Möglichkeiten zur Erzeugung eines wartenden oder neutralen Ausdrucks durch Verwendung mehrerer Gesten

Gegensätzlich verhält es sich mit Begrüßung und Verabschiedung beziehungsweise mit Beginn und Abschluss einer Interaktion. Die ausdrucksstärkeren Gesten werden fast vollkommen isoliert von anderen verwendet, um den gewünschten Ausdruck zu erzeugen. Vor allem bei diesen beiden Fällen wird deutlich, dass die verwendete Geste nicht nur von einem Ausdruck und der Expressivität abhängt. Vor allem das Küsschen sollte, zumindest in Deutschland, als Begrüßung nur bei sehr wenigen Personen und sehr vorsichtig verwendet werden. Grund dafür ist die Nähe, die diese Geste benötigt, die von eher fremden Personen als unangenehm empfunden werden kann. Dennoch wirkt sich der Körperkontakt beim Schütteln der Hände durch die starke Ritualisierung der Begrüßung und Verabschiedung nur gering auf die Expressivität aus.

Zudem existieren verschiedene Konventionen, die die Expressivitätsspannen zusätzlich beeinflussen. Eine alltägliche und bekannte Begrüßung stellt vor allem in Deutschland das Händeschütteln dar, wobei diese Geste zum Beispiel in einem Nachbarland, wie Frankreich, eher unüblich ist. Dort ist die äquivalente Geste, wie auch in einigen weiteren europäischen Ländern, das Bisou (französisch für „Küsschen“), ein Wangenkuss auf der linken und rechten Seite. Folglich sind die konventionellen Unterschiede bereits bei einer

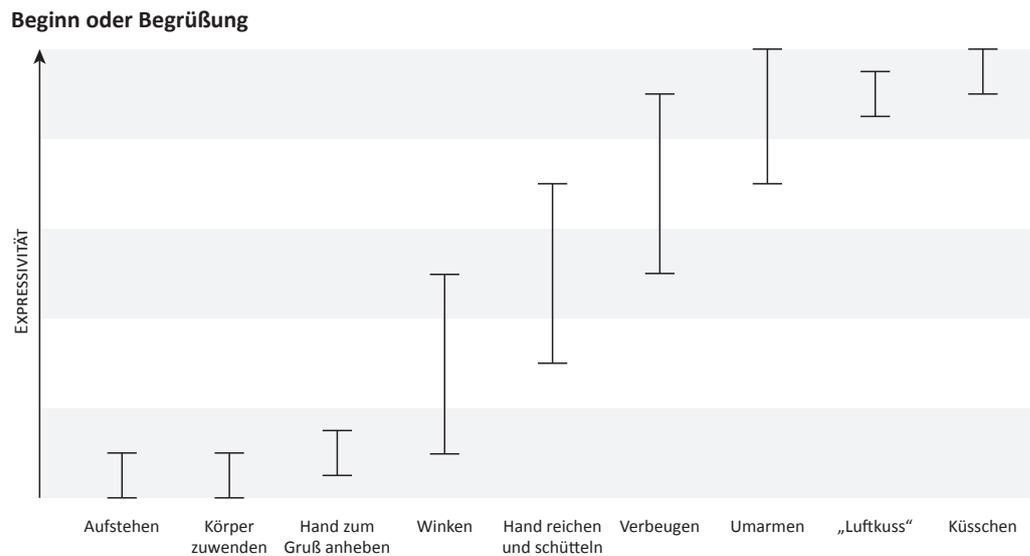


Abb. 3.13: *Expressivitätsschemata zur Darstellung verschiedener Möglichkeiten zur Erzeugung eines Beginns oder einer Begrüßung durch Verwendung mehrerer Gesten*

geringen geografischen Entfernung deutlich spürbar. Eine Verbildlichung der verwendeten Gesten zur Begrüßung und Verabschiedung sind jeweils auf Abbildung 3.13 und 3.14 zu sehen.

Die bisher nicht beschriebene Feedback-Art Fehler besitzt kein eigenes Schema, da sie prinzipiell nur ein negatives Feedback darstellt, welches aber separat gelten muss und auch eine eigene, eindeutige Geste besitzen sollte. Weiterhin stellt sich die Frage, ob ein technischer Fehler über die Kommunikation mitgeteilt werden soll oder ein technisches Feedback, wie eine LED-Anzeige, von Vorteil wäre. Wenn die Kommunikation für diesen Zweck genutzt wird, darf eine Geste, die für einen Fehler verwendet wird, zusätzlich für keinen weiteren Ausdruck verwendet werden, da ein Fehler ein übergeordnetes Feedback ist. Wenn beispielsweise ein schwerwiegender Fehler auftritt, der verhindert, dass die Interaktion fortgesetzt wird, muss ein eindeutiges Feedback generiert werden, damit der Nutzer darüber in Kenntnis gesetzt wird und diesen Fehler beheben kann. Wird dieses deutliche Feedback nicht verwendet, kennt der Nutzer den Zustand des Roboters nicht, da er nicht mehr funktioniert, womit das Problem nicht gelöst werden würde.

Allgemein ist anzumerken, dass die generellen Ausmaße der Expressivitätsspannen für jedes Feedback und für jeden Ausdruck unterschiedlich groß sein können. Zum Beispiel kann das positive und negative Feedback wesentlich deutlicher gesteigert werden und hat eine höhere Spanne der möglichen Ausdrucksstärke als beispielsweise das Feedback „Wartend/Neutral“. Trotz dieser Eigenschaft besitzen in den Schemata alle Spannen die gleiche Größe, da nicht der Vergleich der Spannen im Fokus liegt, sondern die Steigerung

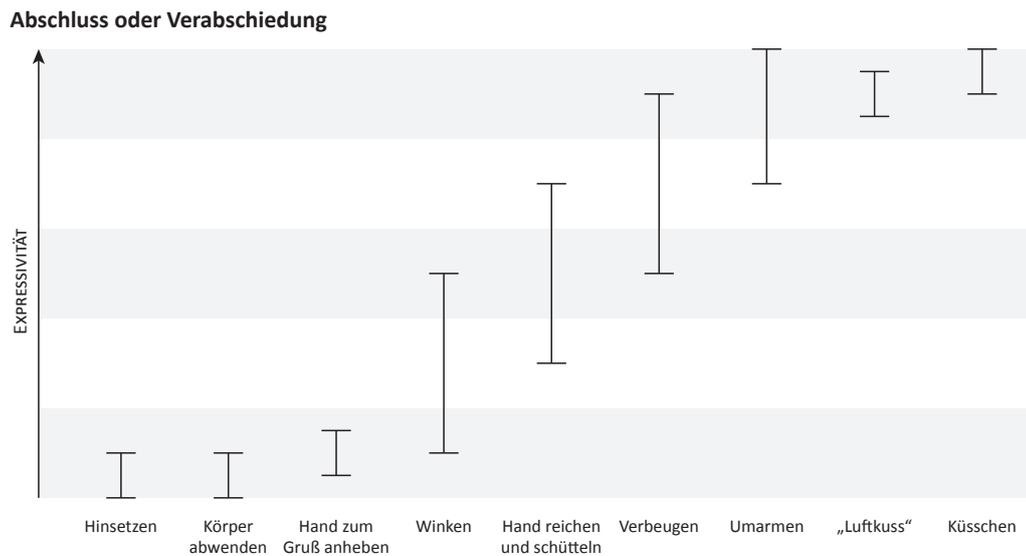


Abb. 3.14: *Expressivitätsschemata zur Darstellung verschiedener Möglichkeiten zur Erzeugung eines Abschlusses oder einer Verabschiedung durch Verwendung mehrerer Gesten*

durch Gesten innerhalb der Feedback-Arten.

Darüber hinaus besitzt auch die Größe und das Alter der ausführenden Person eine Bedeutung für den Ausdruck einer Geste. In Bezug zu humanoiden Robotern wirkt ein Küsschen oder eine Umarmung eines kleinen Roboters, zum Beispiel des Naos, tendenziell süß und lieb, da durch die Größe eine Ähnlichkeit zu einem Kind besteht. Eine Umarmung eines Roboters mit einer Größe von zwei Metern kann dagegen schnell bedrohlich und einschüchternd wirken.

In den Abbildungen 3.9 bis 3.14 wird die Anwendung der Multi Gesture Expression dargestellt. Es stehen mehrere Gesten für einen Ausdruck zur Verfügung, der in seiner Expressivität verändert werden kann. Als Generator der Multi Gesture Expression fungiert eine horizontale Linie auf Höhe der gewünschten Expressivität. Somit werden eindeutig die Gesten erkannt, die man für dieses Expressivitätsniveau verwenden kann. Jedoch müssen hier zusätzlich noch die Gesten mit Hilfe ihrer Parameter so manipuliert werden, dass sie mit einer anderen Geste und deren variierten Parameter übereinstimmen. Erst nach diesem Schritt ist eine Auswahl zwischen mehreren Gesten möglich, die sich auf der gleichen Expressivitätsstufe befinden.

Die Auswahl einer Geste aus einer Gruppe von mehreren Gesten, die den gleichen Ausdruck vermitteln und ebenso in ihrer Expressivität identisch sind, kann über verschiedene Aspekte erfolgen. Eine Möglichkeit wäre die feste Vorgabe der bevorzugten Geste. Für diesen Fall müsste man Prioritäten für Gesten einführen. Dennoch besteht hier das Problem, dass man einen großen Teil der Flexibilität der Multi Gesture Express-

sion verliert, da festgelegt wurde, welche Geste zu welchem Zeitpunkt verwendet werden sollte. Weiterhin bestünde die Möglichkeit einen humanoiden Roboter mit Hilfe von eigenen Beobachtungen und einer Tendenz zu emotionalem Verhalten selbst entscheiden zu lassen. Das Problem dabei ist der große Aufwand, der benötigt wird, und dass es bis heute nahezu unmöglich ist, einem Roboter reale Emotionen und ein Urteilsvermögen zu vermitteln. Eine Lösung dieses Problems, die auch für den Zweck genügt, ist die Entscheidung der Gestenverwendung anhand des Zufalls. Somit bleibt die Flexibilität erhalten und das Verhalten des Roboters ist nur zu einem gewissen Grad vorhersehbar.

4 | Implementierung des Gestensets für einen humanoiden Roboter

In diesem Kapitel geht es um die Umsetzung des parametrisierten Gestensets für den humanoiden Roboter Nao. Zunächst werden die Grundlagen beschrieben, die zur Verwendung des Naos essentiell sind. Die allgemeinen Gedanken zur Implementierung und die Einschränkungen durch die Verwendung des Naos werden zusätzlich thematisiert, um daraufhin den Vorgang der Implementierung näher zu erläutern.

4.1 Technische Grundlagen

Zu den technischen Grundlagen gehören zum einen die Hardware und zum anderen die Software, die zur Implementierung verwendet werden.

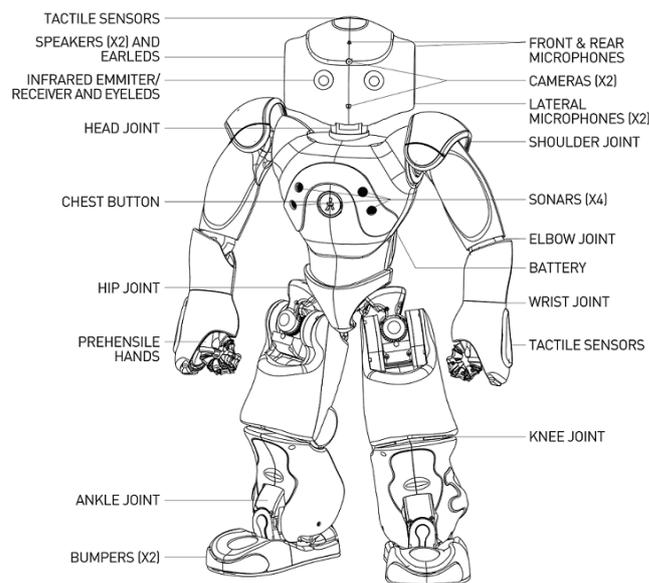


Abb. 4.1: Bestandteile und Funktionen des Naos

Da diese Arbeit sich mit der Erstellung eines Gestensets für einen humanoiden Roboter beschäftigt, wurde der Roboter Nao von Aldebaran Robotics [Ald] verwendet, der

im Jahr 2006 vorgestellt wurde und auf den Abbildungen 2.9 und 4.1 zu sehen ist. Er hat eine Größe von 58 cm, wiegt 4,5 kg, kann sich mit Hilfe elektrischer Motoren innerhalb von 25 Freiheitsgraden bewegen, besitzt zwei Kameras, mehrere Mikrofone, einen Ultraschall-Entfernungsmesser, Infrarotemitter und -empfänger, taktile Sensoren am Kopf, den Händen, den Füßen und der Brust, LEDs an den Ohren und Augen und zwei Lautsprecher. Eine vollständige Übersicht der technischen Ausstattung und der Konzeption wurde von Gouaillier et al. [Gou09] verfasst. Trotz der guten Ausstattung des Nao können manche Parameter, wie zum Beispiel die Stetigkeit, nicht reibungslos implementiert werden, da der Nao generell, wie viele Roboter, teilweise über stockende Bewegungen verfügt. Zusätzlich ist die Unterstützung der Gestik durch Mimik auf Grund der fehlenden Gesichtsmotorik des Nao weitgehend unmöglich. Zur Kompensation dieser Probleme können die LEDs der Augen und Ohren verwendet werden, um eine Art der Mimik zu erzeugen. Neben dem Nao wurde keine weitere Hardware für die Umsetzung des Gestensets verwendet.

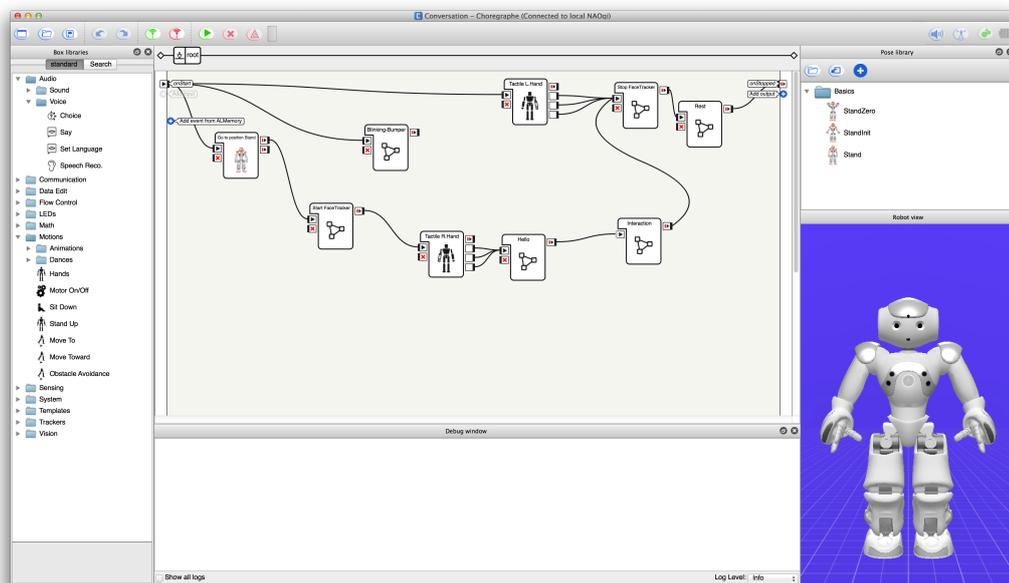


Abb. 4.2: Benutzeroberfläche der Entwicklungsumgebung Choregraphe von Aldebaran Robotics: Auf der linken Seite ist die Box Library und auf der rechten die Darstellung des Nao abgebildet. Diese Komponenten umrahmen das Flow Diagram, in dem die Boxen der Library verknüpft werden können.

Auf der Software-Seite wird zur Erzeugung von Applikationen für den Nao die Entwicklungsumgebung Choregraphe bereitgestellt, dessen grafische Oberfläche auf Abbildung 4.2 zu sehen ist. Weiterhin besitzt der Nao das Framework NaoQi, das den Zugang zu allen Funktionen des Naos gibt, die entweder mittels der Programmiersprachen C++, Python oder Urbiscript verwendet werden können. Choregraphe ist eine grafische Umgebung, die für Windows, Mac und Ubuntu verfügbar ist. Sie besteht aus drei

grundlegenden Bestandteilen. Zum einen die Box Library, die verschiedene vor- oder selbstdefinierte Verhaltensweisen enthält, und das Flow Diagram, in dem die Boxen der Bibliothek innerhalb eines Flussdiagramms angeordnet und verknüpft werden können. Zum anderen besitzt Choregraphe eine Darstellung des Naos, mit der die Gelenke des Naos direkt manipuliert werden können oder an der das erzeugte Verhalten getestet werden kann. Die angesprochenen Boxen, die die Verhaltensweisen enthalten, bestehen entweder erneut aus einem Flow Diagram, aus einer Timeline, auf der Bewegungen aufgezeichnet werden, oder aus einem Python-Skript, mit dessen Hilfe die Funktionen von NaoQi direkt angesprochen und aufgerufen werden können. Weiterhin verfügt Choregraphe über einen Animationsmodus, mit dem Gelenkpositionen des Naos mit Hilfe der Timeline gespeichert werden können. [Pot09]

4.2 Grundüberlegungen zur Umsetzung

Zur Umsetzung des entwickelten Gestensets gehören zunächst zwei Komponenten. Zum einen stellt sich die Frage wie die Gestenerzeugung stattfindet und welche Werte als Eingabe und welche als Ausgabe dienen. Zum anderen ist die Realisierung der Multi Gesture Expression und deren vielfältigen Gesten ein weiterer Punkt, der besprochen werden muss.

4.2.1 Umsetzung der Gestenerzeugung

Wie Gesten generell beim Menschen erzeugt werden, wurde in Kapitel 3.1.1 erläutert. Hierauf folgt die Frage, wie man dieses kognitive System sinnvoll mit einem Programm imitieren kann. Einen vorgangsauslösenden Reiz benötigt auch der Roboter, wobei die Verarbeitung und Klassifizierung des Reizes nicht die Aufgabe dieser Arbeit ist. An dieser Stelle wird die Gestenerzeugung bei bekannter Reaktion des Roboters betrachtet, die als Absicht bezeichnet wird. Eine praktische und sinnvolle Möglichkeit ist, dass der Nutzer Einstellungen vor Verwendung der Geste festlegen muss, die anschließend als Eingabe fungieren. Dies dient stellvertretend für die Bewertung des Reizes.

Aus einer Bewertung des Reizes folgt eine Absicht, mit der man auf das Wahrgenommene reagiert, wie zum Beispiel mit Freude oder Wut, wodurch eine Geste und eine Ausdrucksstärke generiert werden. Zur Berechnung der endgültigen Geste werden somit drei Komponenten benötigt, von denen eine, die Auswahl der Geste, durch das im Kapitel 3.3.2 entwickelte System der Parametrisierung generiert wird. Als Hilfsmittel wird zusätzlich noch eine zufällige Entscheidung verwendet, auf die in Kapitel 4.2.2 näher eingegangen wird. Die beiden weiteren Eigenschaften, die Absicht und die Expressivität der Geste, müssen vom Nutzer bestimmt werden, um eine ausführbare Geste zu erhalten, wie es zusätzlich auf Abbildung 4.3 dargestellt ist.

Die Absicht einer Geste besteht aus einer Auflistung möglicher Arten, zu denen jeweils Gesten generiert werden können. Im Gegensatz dazu wird die Expressivität durch eine Gleitkommazahl zwischen den Werten 0 und 1 angegeben, wobei 0 für keine Expressivität, folglich keine Geste, und 1 für die höchste Expressivität steht.

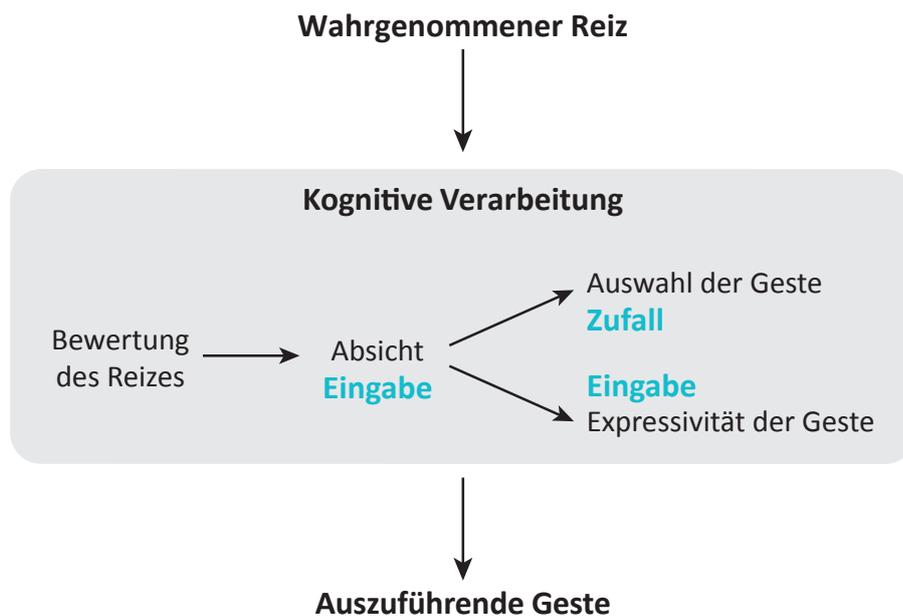


Abb. 4.3: Schematische Darstellung der Entstehung einer Geste und des Ursprungs der entsprechenden Komponenten in der Implementierung

Die genaue Angabe der Parameter, die in Kapitel 3.2.2 beschrieben wurden, ist nicht nötig, da diese aus dem Wert der Expressivität ableitbar sind. Beispielsweise müssen für eine hohe Expressivität der Winken-Geste sowohl die Frequenz als auch die räumliche Ausdehnung einen hohen Wert besitzen. Dies wird in Anhang A verdeutlicht. Folglich genügt das Wissen über die Ausdrucksstärke, um die entsprechenden Parameterwerte zu generieren.

4.2.2 Umsetzung der Multi Gesture Expression

Um das Verfahren der Multi Gesture Expression umzusetzen, müssen zunächst Anpassungen erfolgen. Hierbei stellt die Plattform, auf der die Gesten implementiert werden sollen, ein Kriterium dar, welches betrachtet werden muss. Der Roboter Nao hat eine begrenzte Anzahl an Freiheitsgraden und ist in weiteren Eigenschaften eingeschränkt. Aus diesem Grund müssen die Gesten der einzelnen Feedback-Arten auf den Nao angepasst werden, das heißt, dass die Gesten entfernt werden, die nicht ausführbar sind. Für welche Gesten dies der Fall ist und aus welchem Grund wird in Tabelle 4.1 erläutert. Durch diese Einschränkungen und die Vernachlässigung mancher Gesten müssen weiterhin die Expressivitätsspannen und -schemata, die in Kapitel 3.3.2 entwickelt wurden, für den Nao verändert werden, so dass unter Umständen Lücken geschlossen werden müssen oder die Expressivitätsskala angeglichen werden muss.

Nach Anpassung der Expressivitätsspannen wird ein Konzept benötigt, um die einzelnen Gesten anhand ihrer Parameter zu implementieren und anschließend zu variieren.

Tab. 4.1: Gesten, die vom Nao aus motorischen und technischen Gründen nicht ausgeführt werden können

Feedback-Art	Nicht ausführbare Geste	Begründung
Positiv oder Akzeptanz	OK-Geste	Keine Steuerung der Finger
	„Daumen hoch“	Keine Steuerung der Finger
	Auf Schulter klopfen	Ort des Partners unbekannt
	High-Five	Ort des Partners unbekannt
Negativ oder Abweisung	Hängende Schultern	Keine „Sprunggelenke“ vorhanden
	Verschränkte Arme	Keine Steuerung der Schultern
	„Daumen runter“	Keine Möglichkeit, die Position aussagekräftig auszuführen
Unverständnis	Erhobener Zeigefinger	Keine Steuerung der Finger
	Kopf schieflegen	Keine Steuerung der Finger
	Kopf wiegen	Keine Steuerung des Kopfes in diese Richtung
Wartend/Neutral	Schulterzucken	Keine Steuerung des Kopfs in diese Richtung
	Mit Fuß klopfen	Keine Steuerung der Schultern
		Keine Möglichkeit, die Position aussagekräftig auszuführen
Beginn/Abschluss oder Begrüßung/Verabschiedung	Körper zuwenden	Ort des Partners unbekannt
	Hand reichen und schütteln	Ort des Partners unbekannt
	Umarmen	Ort des Partners unbekannt
	„Luftkuss“	Keine Möglichkeit, die Mimik darzustellen
	Küsschen	Ort des Partners unbekannt

Eine umfangreiche Methode ist die kontinuierliche Abstimmung der Geste an die Parameter, so dass die Gelenkwerte direkt durch die Parameter beeinflusst und dynamisch generiert werden. Diese Variante ist aufwändig, weswegen eine andere, einfachere Methode verwendet wird, die dennoch ähnlich effektiv ist.

Bei Verwendung der einfacheren Methode wird die Expressivitätsspanne jeder Geste erneut in gleichgroße Stufen unterteilt. Innerhalb der Stufen wird die gleiche Geste ausgeführt. Zum Beispiel wird die Nicken-Geste in vier unterschiedliche Stufen eingeteilt, die mit Hilfe der Expressivität definiert sind. Die niedrigste Stufe beinhaltet ein sehr ausdruckschwaches Nicken, wohingegen die höchste Stufe das expressive Äquivalent enthält. Bei einer Erhöhung der Ausdrucksstärke wird ab einem bestimmten Wert die Stufe gewechselt und somit eine andere Geste verwendet. Für den Fall, dass für einen Expressivitätswert mehrere Gesten passend wären, wird eine zufällige Entscheidung zwischen den zur Wahl stehenden Möglichkeiten getroffen.

Da das eigentliche Gestenset sich der Bewegungsfähigkeiten des Rumpfes, der Extremitäten und des Kopfes bedient und dessen Mimik prinzipiell vernachlässigt, kann bei der Implementierung die geringfügige Mimik, die der Nao verwenden kann, unterstützend genutzt werden, ohne diese direkt in das Gestenset miteinzubeziehen. Der verwendete Roboter verfügt über eine geringe bis keine Mimik, da er keine Bewegung im Gesicht darstellen kann. Möglich sind lediglich Veränderungen der Augenfarbe oder Blinzeln, welches durch die LEDs der Augen imitiert wird. Bei Nutzung der Augenfarbe bietet sich beispielsweise eine Farbcodierung ähnlich einer Ampel an, wobei Grün für Positives und Rot für Negatives steht. Gelb stellt zum Beispiel Unsicherheit dar. Weiterhin ist es dem Nao möglich, Gesichter zu erkennen und diese zu tracken, das heißt ihnen mit seinem Blick beziehungsweise seiner Kopfausrichtung, zu folgen, wodurch sein Aufmerksamkeitsfokus dem Kommunikationspartner übermittelt wird.

4.3 Implementierung in Choregraphie

Die eigentliche Umsetzung des Gestensets fand in der Entwicklungsumgebung Choregraphie (siehe Kapitel 4.1) statt, wobei verschiedene Boxen zur Erzeugung der Gesten zum Einsatz kamen. Um die einzelnen Bewegungen zu erstellen, bot sich der Animationsmodus an, in dem man die aktuellen Gelenkpositionen des Naos zu Keyframes hinzufügen kann, über die anschließend interpoliert wird.

Zunächst mussten die an den Nao angepassten Expressivitätsschemata in Diagramme umgeformt werden, die die Abstufung hinsichtlich der Ausdrucksstärke als konkrete Zahlenwerte darstellen, wodurch die Grundexpressivität und die Expressivitätsspanne der einzelnen Gesten als mathematische Werte definiert werden. Weiterhin werden nahezu alle Gesten in mehrere Stufen unterteilt. In den meisten Fällen galt, dass die Anzahl der erzeugten Stufen von der Größe der Expressivitätsspanne abhängt. Demzufolge steigt die Anzahl der Stufen proportional zur Expressivitätsspanne. Jedoch variiert diese Regel an einigen Stellen. Die Expressivitätsspannen der einzelnen Gesten mussten teilweise neu angepasst werden, da beispielsweise eine Geste auf Grund von Eigenschaften des Naos nur in einer besonders ausdrucksstarken Art und Weise umsetzbar war. Das beschriebene

Vorgehen ist in den Diagrammen der Abbildungen B.1 bis B.6 in Anhang B zu sehen.

Durch die Unterteilung der Gesten entstanden eine Vielzahl an möglichen Verhaltensweisen und implementierten Gesten. Eine genauere Auflistung der Gesten, ihrer Stufen und der Anzahl der Gesten pro Feedback-Art ist in Tabelle 4.2 zu finden. Obwohl die Addition der Stufen pro Feedback-Art 62 implementierte Gesten ergibt, sind es tatsächlich 53 Gesten, da durch die mehrmalige Verwendung einiger Gesten für mehrere Feedback-Arten diese Zahl erhöht wurde. Hierzu zählen beispielsweise das Winken, Verbeugen und weitere Gesten, die sowohl für die Begrüßung als auch für die Verabschiedung verwendet werden.

Tab. 4.2: Anzahl der implementierten Stufen pro Geste und insgesamt implementierten Stufen pro Feedback-Art

Feedback-Art	Geste	Stufen pro Geste	Stufen pro Feedback-Art
Positiv oder Akzeptanz	Nicken	4	11
	Klatschen	4	
	Jubeln	3	
Negativ oder Abweisung	Kopfschütteln	4	11
	Hand mit Handfläche zum Gegenüber	2	
	Hände vor das Gesicht heben	3	
	Mit Fuß aufstampfen	2	
Unverständnis	Hand unter dem Kinn, Ellenbogen aufgestützt	1	8
	Hand horizontal kippen	3	
	Am Kopf kratzen	2	
	Arme angewinkelt, Handflächen nach oben	2	
Wartend/Neutral	Umherschweifender Blick	2	11
	Verlagerung des Körpergewichts	3	
	Hand unter dem Kinn, Ellenbogen aufgestützt	1	
	Hände hinter dem Kopf	1	
	Arme in die Hüfte gestemmt	2	
	Umherlaufen	2	
Beginn oder Begrüßung	Aufstehen	2	10
	Hand zum Gruß anheben	2	
	Winken	4	

Feedback-Art	Geste	Stufen pro Geste	Stufen pro Feedback-Art
	Verbeugen	2	
Abschluss oder Verabschiedung	Hinsetzen	2	11
	Körper abwenden	1	
	Hand zum Gruß anheben	2	
	Winken	4	
	Verbeugen	2	

Im Anschluss entstand eine Beispielanwendung, in der zunächst bei Beginn der Ausführung eine Startgeste ausgeführt wird. Bei Berührung der rechten Hand wird eine Endgeste generiert. Zusätzlich kann das Programm jederzeit mit Berührung der linken Hand beendet werden und der Nao begibt sich in eine sichere Position, in der er die Gelenkmotoren abstellen kann. Weiterhin ist es möglich mit Druck des „Bumpers“ des rechten Fußes, Blinzeln zu aktivieren, wodurch der Nao wesentlich aufmerksamer und lebendiger wirkt.

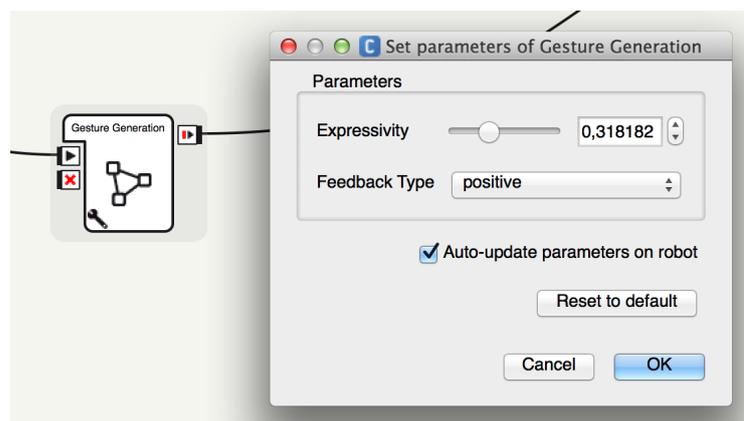


Abb. 4.4: Einstellungsmöglichkeiten der Box „Gesture Generation“: Die Expressivität kann zwischen den Werten 0 und 1 gewählt werden, wohingegen es folgende Möglichkeiten für die Feedback-Art gibt: positive, negative, incomprehension, waiting, start/greeting, end/dismissal.

Kernstück der gesamten Implementierung ist die Box zur Gestengenerierung. Bei Nutzung dieser Box kann zunächst eine Feedback-Art und eine Expressivität zwischen 0 und 1 gewählt werden, aus denen bei Ausführung des Programms die gewünschte Geste erzeugt wird. Für die Feedback-Art stehen folgende Möglichkeiten zur Auswahl: positive, negative, incomprehension, waiting, start/greeting und end/dismissal. Die Einstellungsmöglichkeiten der Parameter der Box sind auf Abbildung 4.4 zu sehen. Diese Box existiert in zwei Ausführungen. Zum einen mit Steuerung des Face Trackers und zum anderen ohne diese Möglichkeit. Grund dafür ist, dass bei jeder Bewegung des Kopfes zunächst

der Face Tracker abgestellt werden muss, um widersprüchliche Signale an den Gelenkmotor zu verhindern. Wie die Gestenerzeugungsbox im Einzelnen aufgebaut ist, ist auf den Abbildungen C.1, C.2 und C.3 in Anhang C zu sehen. Zunächst wird der Parameter „Feedback Type“ ausgewertet, wodurch anschließend die passende Box, zum Beispiel „Generate Positive Gesture“, wenn eine positive Geste erzeugt werden soll, aufgerufen wird. Innerhalb dieser Box wird mittels spezieller If-Abfragen überprüft, in welchem Bereich die Expressivität liegt. Anschließend wird, falls nötig, eine zufällige Entscheidung zwischen zwei oder mehreren Gesten getroffen und daraufhin die Stufe beziehungsweise die Expressivität neu berechnet. Dieser Vorgang ist notwendig, da die Gesten innerhalb ihrer Box erneut eine Expressivitätsspanne von 0 bis 1 haben. Dadurch sind die Boxen, die die einzelnen Gesten generieren, wiederverwendbar, so dass man diese Boxen zur Erzeugung einer separaten Geste mittels eines Werts der Ausdrucksstärke nutzen kann. Diese Anpassung wird in der Box „Level Adjustment“ vorgenommen, die folgende Berechnung enthält:

$$\text{Expressivität}_{\text{neu}} = \frac{\text{Expressivität}_{\text{alt}} - \text{Grundexpressivität}}{\text{Expressivitätsspanne}}$$

Am Beispiel einer positiven Geste mit einer Expressivität von 0,5, die der Klatschen-Geste zugeordnet wird, käme folgender Wert für die neue Expressivität zustande:

$$\text{Expressivität}_{\text{neu}} = \frac{0,5 - 0,33}{0,5} = 0,34$$

Anschließend wird in der Box der auszuführenden Geste, in besagtem Beispiel die Box für das Klatschen, erneut der Wert der Expressivität überprüft und die Einteilung in die unterteilten Stufen der Geste durchgeführt. Danach werden die Boxen aufgerufen, die die gewählte Stufe der Geste ausführen, in diesem Beispiel „Clapping: Level 2“. Zusätzlich ist noch anzumerken, dass alle Gesten hinsichtlich der Ausführung aus der stehenden Position optimiert wurden.

Unterstützt werden die Gesten der Feedback-Arten durch die LEDs der Augen. Hierbei wurde sowohl für positives Feedback als auch für den Beginn beziehungsweise Begrüßung eine grüne Farbe gewählt. Im Gegensatz dazu wird Rot für negatives Feedback und Abschluss beziehungsweise Verabschiedung verwendet. Gelb dient als Zeichen für Unverständnis. Während der Nao eine wartende Geste ausführt, wechselt die Farbe seiner Augen-LEDs von einer hohen zu einer niedrigen Intensität. Nach Abschluss der ausgeführten Geste nehmen die LEDs ihren Standardwert an.

Im Zuge der Übersichtlichkeit und der Wiederverwertbarkeit gibt es eine Box Library, die alle verwendeten Gesten und deren Boxen enthält sowie alle weiteren Boxen, die im Laufe der Implementierung erzeugt wurden. Diese Bibliothek ist wie folgt gegliedert:

- **Gesture Generation:** Boxen für die Gestenstufen der jeweiligen Feedback-Art und deren Erzeugungsbox anhand der Expressivität
- **Expressivity Related Operations:** Spezielle Operationen, die sich auf die Expressivität beziehen, wie zum Beispiel spezielle If-Abfragen oder „Level Adjustment“

- **Random Generator:** Boxen für die zufällige Entscheidung zwischen zwei, drei und vier Möglichkeiten
- **Math:** Mathematische Boxen, die nicht standardmäßig vorhanden waren
- **General Animations:** Allgemeine Bewegungen oder Verhaltensweisen, die im Laufe der Implementierung entstanden sind

Diese Bibliothek kann von jeder Person innerhalb ihres eigenen Choregraphie-Projekts verwendet werden, wodurch eine hohe Wiederverwendbarkeit garantiert ist, die zusätzlich durch die unabhängigen Expressivitätswerte zwischen Feedback-Art und eigentlicher Geste und damit deren Neuberechnung unterstützt wird.

5 | Abschließende Bemerkungen

In diesem Kapitel wird die Arbeit zusammengefasst, um die wichtigsten Punkte noch einmal zu erläutern. Anschließend folgt ein Fazit, in dem die Arbeit kritisch bewertet wird und abschließend wird ein Ausblick auf mögliche, zukünftige Arbeiten gegeben.

5.1 Zusammenfassung

In dieser Arbeit wurde die Fragestellung nach einer Parametrisierung von Gesten und die Erstellung eines Gestensets für humanoide Roboter, das dieses Vorhaben realisiert, bearbeitet. Zunächst wurde ein Überblick über vorhandene Arbeiten gegeben, die dieses Thema behandeln und die die Grundlagen für diese Arbeit darstellen. Anschließend folgte die Einführung der Hardwareplattform, auf der die parametrisierten Gesten implementiert werden sollten – dem Nao-Roboter.

Daraufhin wurden alltägliche Gesten analysiert und nach Feedback-Arten klassifiziert, auf denen anschließend das Konzept aufgebaut wurde. Somit wurden die Parameter räumliche Ausdehnung, zeitliche Ausdehnung und Beschleunigung, die zu den Intraphasenparameter zählen und für die innere Expressivität einer Geste verantwortlich sind, beschrieben. Weiterhin stellen die Parameter Frequenz und Stetigkeit die Interphasenparameter dar, die die äußere Expressivität bestimmen. Sie beeinflussen die Zusammenstellung der einzelnen Gestenphasen, wohingegen die Intraphasenparameter die Ausführung der eigentlichen Geste beeinflussen.

Anschließend wurden zwei Ausdrücke eingeführt. Zum einen Single Gesture Expression, bei der lediglich eine Geste mittels der aufgeführten Parameter gesteigert wird, und zum anderen Multi Gesture Expression, die einen Ausdruck oder eine Absicht steigert, indem mehrere Gesten gesteigert und verwendet werden. Dadurch konnten die Feedback-Arten mit zugehörigen Gesten in Expressivitätsschemata nach der zweiten Methode in ihrer Ausdrucksstärke gesteigert werden.

Dieses System bildete die Grundlage für die Implementierung. Hierbei ist es möglich mit Eingabe einer Feedback-Art und eines Expressivitätswerts, eine entsprechende Geste auf dem Nao zu generieren, die in Stufen, die die Steigerung der Parameter darstellen, unterteilt sind. Zusätzlich sind bestimmte Gesten mit Hilfe einer Angabe der Ausdrucksstärke erzeugbar.

Somit ist es möglich den Nao während einer Kommunikation Gesten ausführen zu lassen, die dem Gegenüber eine höhere Lebendigkeit und eine Bereitschaft zum Gespräch und zur Interaktion vermitteln. Gesten sind somit ein bedeutendes Mittel der Interaktion

zwischen Partnern, ob zwischen Menschen oder Mensch und Roboter. Sie unterstützen die verbale Kommunikation und machen sie verständlicher.

5.2 Fazit

Zusammenfassend lässt sich sagen, dass in dieser Arbeit ein neues Verfahren zur Expressivitätssteigerung innerhalb der Kommunikation zwischen Mensch und Roboter geschaffen wurde. Dieses beruht auf etablierten Parametern, die neu klassifiziert und spezialisiert wurden, wodurch sie zusätzlich erweitert und zu einem Parameter, der Expressivität, zusammengefasst wurden. Im Folgenden werden die Probleme beleuchtet, die es bei der theoretischen Ausarbeitung sowie der Implementierung gibt. Anschließend wird die Arbeit bewertet.

Probleme gibt es zum einen bei der vorliegenden Methode zur Parametrisierung von Gesten. Zum einen geht es hierbei um die Komplikationen, die bei einem Expressivitätswert erscheinen, der mehrere Gesten zulässt. An dieser Stelle ist die Auswahl oftmals nicht wie im Konzept an den Zufall gebunden, sondern beispielsweise an gesellschaftliche Stellungen oder die Beziehung zwischen den Kommunikationspartnern. Dies gilt außerdem für die generelle Auswahl der Geste, die oft auch kulturell beeinflusst wird.

Zum Anderen traten Probleme bei der Implementierung innerhalb der Entwicklungsumgebung Choregraphe auf. Einen Vorteil würde die zusätzliche Umsetzung mittels Python-Skripte statt der Nutzung des Animationsmodus von Choregraphe bieten, da dadurch die allgemeine Verwendung sowohl durch Choregraphe als auch durch Python-Skripte gegeben wäre. Dieses Vorhaben trat erst spät auf Grund der mangelnden Dokumentation des Programms in den Bereich des Möglichen, wodurch die kurzfristige Umsetzung nahezu unmöglich wurde. Durch die Größe und die Untergliederung der Hauptbox in Unterboxen, die in 4.3 beschrieben wurden, nimmt die einmalige Übertragung und die Erzeugung eines Programms, das diese Box einmal verwendet, bereits viel Zeit in Anspruch. Diese Zeit erhöht sich mit mehrmaliger Verwendung der gesamten Box. Eine Lösung dieses Problems stellt die Verwendung der Unterboxen statt der Hauptbox für einzelne Gesten dar, wenn beispielsweise bekannt ist, welche Geste oder zumindest welche Feedback-Art ausgedrückt werden soll. Anzumerken ist, dass während der Ausführung des Programms auf dem Nao keine Verzögerung auftritt.

Bei der Implementierung des Gestensets stellte sich heraus, dass die Gesten ausdrucksstark sind und die Absichten des Roboters effektiv darstellen. Weitere Vorteile könnten durch die zusätzliche Verwendung der Sprache erreicht werden, wodurch Gesten an Expressivität zusätzlich gewinnen würden.

Alles in allem wurde ein Gestenset entworfen und erstellt, welches seine Aufgabe der Parametrisierung erfüllt und für die Verwendung durch einen humanoiden Roboter gut geeignet ist. Durch die hohe Anpassungsfähigkeit der Gesten und der Ausdrücke ist eine Möglichkeit der Wiederverwertung innerhalb anderer Programme und Verhaltensweisen gegeben, wodurch sich diese Implementierung auszeichnet. Die Kommunikationsfähigkeit des Roboters wird durch die Verwendung des Gestensets stark erhöht, weswegen die zu Beginn gestellte Frage nach der Sinnhaftigkeit der Verwendung von Gesten durch einen

humanoiden Roboter positiv beantwortet werden kann.

5.3 Ausblick

Zukünftig gäbe es weiterführend die Möglichkeit das vorliegende System auf verschiedene Arten zu erweitern und zu verbessern, wodurch dieses besser auf humanoide Roboter abgestimmt wäre und zusätzlich weitere Komponenten der zwischenmenschlichen Kommunikation miteinbeziehen könnte.

Zunächst müssten Evaluierungen der einzelnen Gesten, deren Grundexpressivität und Expressivitätsspannen sowie der Auswirkungen der Steigerungen bestimmter Parameter und der Implementierung durchgeführt werden, da diese bisher nur geringfügig überprüft und validiert wurden.

Ein Thema, das in Zukunft betrachtet werden könnte, ist das Zustandekommen des Expressivitätswerts und der Absicht, da diese an dieser Stelle durch die Eingabe des Nutzers festgelegt werden. Hierzu zählt die Überprüfung, ob eine Geste für eine bestimmte Situation als passend eingestuft werden kann. Es existieren andere verwandte Arbeiten, wie das BEAT-Toolkit [Cas04], die diese Werte und Bewertungen beispielsweise aus dem Kontext einer Texteingabe filtern.

Außerdem gäbe es die Möglichkeit bei Erkennung von Personen durch den Roboter, diesen eine sogenannte Basissympathie zuzuweisen, die Auswirkungen auf die mögliche Expressivität, die mit den Gesten erzeugt werden kann, haben könnte. Diese Basissympathie könnte sich nach dem Bekanntheitsgrad der Person zum Roboter richten. Bei wiederholtem Aufeinandertreffen mit Personen, die der humanoide Roboter „kennt“, ist eine Erhöhung der Basissympathie möglich, da sich beide in einer vertrauteren Situation befinden als bei der ersten Begegnung. Somit würde das „Kennenlernen“ von Personen simuliert werden und auf die Beziehung zwischen Mensch und Roboter übertragen werden, wodurch Beziehungen zwischen Kommunikationspartner bewertet und imitiert werden könnten.

Die gewählten Parameter können unter Umständen zusätzlich zur Gestenerkennung genutzt werden. Je nach Detailgrad der Parameter könnten deren einzelne Werte ausreichen, um eine Geste in ihrer Form zu erkennen. Falls diese nicht ausreichen, stellt sich die Frage, welche und wie viele Informationen zusätzlich benötigt werden, um eine Geste eindeutig zu definieren. Beispielsweise könnten zu den Parameterwerten die Positionen der Extremitäten, die Gelenkwerte oder die Grundabsicht, wie Wut oder Freude, von Interesse sein.

Abschließend gibt es noch eine weitere Möglichkeit das vorgestellte System zu verbessern. Hierbei werden Gesten miteinander kombiniert, die den gleichen Ausdruck vermitteln, um eine höhere Ausdrucksstärke zu generieren. Somit würde die Multi Gesture Expression erneut verstärkt und in ihrer Expressivität erweitert werden, wodurch sich die Kommunikation zwischen Mensch und Roboter ein weiteres Mal verbessern ließe.

Die Umsetzung dieser und auch weiterer Verbesserungsmöglichkeiten würde humanoide Roboter zu lebendigeren Kommunikationspartnern machen, die vom Menschen besser akzeptiert werden würden.

Literaturverzeichnis

- [Ald] *Aldebaran Robotics*. Juli 2013. URL: <http://www.aldebaran-robotics.com/>.
- [Aly13] Amir Aly und Adriana Tapus. „A model for synthesizing a combined verbal and nonverbal behavior based on personality traits in human-robot interaction“. *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*. IEEE Press. 2013, S. 325–332.
- [Bec10] Aryel Beck, Antoine Hiolle, Alexandre Mazel und Lola Cañamero. „Interpretation of emotional body language displayed by robots“. *Proceedings of the 3rd international workshop on Affective interaction in natural environments*. ACM. 2010, S. 37–42.
- [Bre05] Cynthia Breazeal, Cory D. Kidd, Andrea Lockerd Thomaz, Guy Hoffman und Matt Berlin. „Effects of nonverbal communication on efficiency and robustness in human-robot teamwork“. *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on*. IEEE. 2005, S. 708–713.
- [Cad94] Claude Cadoz. *Les réalités virtuelles: Un exposé pour comprendre, un essai pour réfléchir*. Dominos (Paris). Flammarion, 1994.
- [Cas04] Justine Cassell, Hannes Högni Vilhjálmsón und Timothy Bickmore. „BEAT: the behavior expression animation toolkit“. *Life-Like Characters*. Springer, 2004, S. 163–185.
- [Ekm92] Paul Ekman. „An argument for basic emotions“. *Cognition & Emotion* (1992) 6.3-4, S. 169–200.
- [Gou09] David Gouaillier, Vincent Hugel, Pierre Blazevic, Chris Kilner, Jérôme Monceaux, Pascal Lafourcade, Brice Marnier, Julien Serre und Bruno Maisonier. „Mechatronic design of NAO humanoid“. *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE. 2009, S. 769–774.
- [Gun06] Hatice Gunes und Massimo Piccardi. „A bimodal face and body gesture database for automatic analysis of human nonverbal affective behavior“. *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*. Bd. 1. IEEE. 2006, S. 1148–1153.
- [Ham11] Jaap Ham, René Bokhorst, Raymond Cuijpers, David van der Pol und John-John Cabibihan. „Making robots persuasive: the influence of combining persuasive strategies (gazing and gestures) by a storytelling robot on its persuasive power“. *Social Robotics*. Springer, 2011, S. 71–83.

- [Har06] Björn Hartmann, Maurizio Mancini und Catherine Pelachaud. „Implementing expressive gesture synthesis for embodied conversational agents“. *gesture in human-Computer Interaction and Simulation*. Springer, 2006, S. 188–199.
- [Ken04] Adam Kendon. *Gesture: Visible action as utterance*. Cambridge University Press, 2004.
- [Kim12] Aelee Kim, Hyejin Kum, Ounjung Roh, Sangseok You und Sukhan Lee. „Robot gesture and user acceptance of information in human-robot interaction“. *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM. 2012, S. 279–280.
- [Kob08] Yoshinori Kobayashi, Yosuke Hoshi, Goh Hoshino, Tomoki Kasuya, Masato Fueki und Yoshinori Kuno. „Museum guide robot with three communication modes“. *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*. IEEE. 2008, S. 3224–3229.
- [Kop06] Stefan Kopp, Brigitte Krenn, Stacy Marsella, Andrew N Marshall, Catherine Pelachaud, Hannes Pirker, Kristinn R. Thórisson und Hannes Vilhjálmsson. „Towards a common framework for multimodal generation: The behavior markup language“. *Intelligent Virtual Agents*. Springer. 2006, S. 205–217.
- [Le11] Quoc Anh Le, Souheil Hanoune und Catherine Pelachaud. „Design and implementation of an expressive gesture model for a humanoid robot“. *Humanoid Robots (Humanoids), 2011 11th IEEE-RAS International Conference on*. IEEE. 2011, S. 134–140.
- [Le12] Quoc Anh Le und Catherine Pelachaud. „Generating co-speech gestures for the humanoid robot NAO through BML“. *Gesture and Sign Language in Human-Computer Interaction and Embodied Communication*. Springer, 2012, S. 228–237.
- [McN08] David McNeill. *Gesture and thought*. University of Chicago Press, 2008.
- [McN92] David McNeill. *Hand and mind: What gestures reveal about thought*. University of Chicago Press, 1992.
- [Mee12] Raveesh Meena, Kristiina Jokinen und Graham Wilcock. „Integration of gestures and speech in human-robot interaction“. *Cognitive Infocommunications (CogInfoCom), 2012 IEEE 3rd International Conference on*. IEEE. 2012, S. 673–678.
- [Mon09] Jérôme Monceaux, Joffrey Becker, Céline Boudier und Alexandre Mazel. „Demonstration: first steps in emotional expression of the humanoid robot Nao“. *Proceedings of the 2009 international conference on Multimodal interfaces*. ACM. 2009, S. 235–236.
- [NTH10] Victor Ng-Thow-Hing, Pengcheng Luo und Sandra Okita. „Synchronized gesture and speech production for humanoid robots“. *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE. 2010, S. 4617–4624.

- [Nic07] Kai Nickel und Rainer Stiefelhagen. „Visual recognition of pointing gestures for human–robot interaction“. *Image and Vision Computing* (2007) 25.12, S. 1875–1884.
- [Nie09] Radoslaw Niewiadomski, Elisabetta Bevacqua, Maurizio Mancini und Catherine Pelachaud. „Greta: an interactive expressive ECA system“. *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*. International Foundation for Autonomous Agents and Multiagent Systems. 2009, S. 1399–1400.
- [Pel10] Catherine Pelachaud, Rodolphe Gelin, Jean-Claude Martin und Quoc Anh Le. „Expressive Gestures Displayed by a Humanoid Robot during a Storytelling Application“. *New Frontiers in Human-Robot Interaction (AISB)*, Leicester, GB (2010).
- [Pot09] Emmanuel Pot, Jérôme Monceaux, Rodolphe Gelin und Bruno Maisonnier. „Choregraphe: a graphical tool for humanoid robot programming“. *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*. IEEE. 2009, S. 46–51.
- [Sch06] Matthias Scheutz, Paul Schermerhorn und James Kramer. „The utility of affect expression in natural language interactions in joint human-robot tasks“. *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM. 2006, S. 226–233.
- [Sha12] Syamimi Shamsuddin, Hanafiah Yussof, Luthffi Ismail, Fazah Akhtar Hanapiah, Salina Mohamed, Hanizah Ali Piah und Nur Ismarrubie Zahari. „Initial response of autistic children in human-robot interaction therapy with humanoid robot NAO“. *Signal Processing and its Applications (CSPA), 2012 IEEE 8th International Colloquium on*. IEEE. 2012, S. 188–193.
- [Var86] Marjorie Fink Vargas. *Louder than words: An introduction to nonverbal communication*. Iowa State University Press, 1986.
- [Wal00] Stefan Waldherr, Roseli Romero und Sebastian Thrun. „A gesture based interface for human-robot interaction“. *Autonomous Robots* (2000) 9.2, S. 151–173.
- [Wen09] Cornelia Wendt und Guy Berg. „Nonverbal humor as a new dimension of HRI“. *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*. IEEE. 2009, S. 183–188.
- [Xu07] Yong Xu, Kazuhiro Ueda, Takanori Komatsu, Takeshi Okadome, Sumi Yasuyuki und Toyooki Nishida. „Can gesture establish an independent communication channel?“ *Control, Automation and Systems, 2007. ICCAS'07. International Conference on*. IEEE. 2007, S. 727–732.
- [Yan07] Hee-Deok Yang, A-Yeon Park und Seong-Whan Lee. „Gesture spotting and recognition for human–robot interaction“. *Robotics, IEEE Transactions on* (2007) 23.2, S. 256–270.

Abbildungsverzeichnis

2.1	Kategorien der McNeill-Taxonomie nach [McN92]	12
2.2	Schematische Darstellung einer in Phasen unterteilten Geste nach [Ken04]	12
2.3	Darstellung unterschiedlicher Verwendungen von Gesten in der Mensch-Roboter-Interaktion	13
2.4	Darstellung der Ansätze für die Richtungserkennung einer deiktischen Geste [Nic07]	14
2.5	Persönlichkeiten eines Roboters mit unterschiedlichem nonverbalem Verhalten [Aly13]	15
2.6	Zwei Stadien der Entwicklung des humanoiden Roboters Leonardo http://www.slashgear.com/mits-furry-robot-is-capable-of-human-learning-235400/ (27.07.2013)	16
2.7	Der Embodied Conversational Agent Greta http://perso.telecom-paristech.fr/~pelachau/Greta/ (27.07.2013)	18
2.8	Bewegungslinien der Gesten mit unterschiedlicher Expressivität [NTH10] .	19
2.9	Der Nao-Roboter von Aldebaran Robotics [Ald]	20
2.10	Vergleich der Gesten des Schauspielers mit der Adaption des Naos [Ham11]	21
2.11	Interpolation über Schlüsselpositionen von 100% Traurigkeit bis zu 100% Angst [Bec10]	21
2.12	Vergleich der gleichen Geste von GRETA und einem Nao [Le11]	22
3.1	Schematische Darstellung der Entstehung von Gesten	24
3.2	Schematische Darstellung der Hierarchie innerhalb der Kommunikation mit Gesten	25
3.3	Darstellung des „Spatial Extent“ bei der GRETA-Plattform [Har06] . . .	34
3.4	Darstellung der Einflüsse der unterschiedlichen Parametergruppen . . .	36
3.5	Darstellung der Intraphasenparameter	37
3.6	Darstellung des Parameters Frequenz	38
3.7	Darstellung des Parameters Stetigkeit	39
3.8	Unterscheidung der beiden Ausdrucksformen „Single Gesture Expression“ und „Multi Gesture Expression“	42
3.9	Möglichkeiten zur Erzeugung von positivem Feedback oder Akzeptanz .	44
3.10	Expressivitätsschemata zur Erzeugung von negativem Feedback oder Abweisung	44

3.11	Expressivitätsschemata zur Erzeugung von Unverständnis	45
3.12	Expressivitätsschemata zur Erzeugung eines wartenden oder neutralen Ausdrucks	46
3.13	Expressivitätsschemata zur Erzeugung eines Beginns oder einer Begrüßung	47
3.14	Expressivitätsschemata zur Erzeugung eines Abschlusses oder einer Verabschiedung	48
4.1	Bestandteile und Funktionen des Naos http://robocup.bowdoin.edu/public/software/nao/NaoQi/1.10/nao-1.10/doc/site_en/greendoc/getting_started/nao_parts.html (05.09.2013)	51
4.2	Benutzeroberfläche der Entwicklungsumgebung Choregraphe von Aldebaran Robotics [Ald]	52
4.3	Schematische Darstellung der Implementierungskomponenten der Entstehung einer Geste	54
4.4	Einstellungsmöglichkeiten der Box „Gesture Generation“	58
B.1	Darstellung der Anpassung des Expressivitätsschema des positiven Feedbacks	79
B.2	Darstellung der Anpassung des Expressivitätsschema des negativen Feedbacks	80
B.3	Darstellung der Anpassung des Expressivitätsschema des Feedbacks für Unverständnis	80
B.4	Darstellung der Anpassung des Expressivitätsschema des wartenden oder neutralen Feedbacks	81
B.5	Darstellung der Anpassung des Expressivitätsschema des Beginn- oder Begrüßungsfeedbacks	81
B.6	Darstellung der Anpassung des Expressivitätsschema des Abschluss- oder Verabschiedungsfeedbacks	82
C.1	Aufbau der Box „Gesture Generation“	83
C.2	Aufbau einer Box, die eine Geste zu einer Feedback-Art erzeugt	84
C.3	Aufbau einer Box zur Generierung einer einzelnen Geste am Beispiel des Klatschens	84

Tabellenverzeichnis

3.1	Ausdrücke einer Handlung und ihre Umsetzungsformen in der Gestik nach [Gun06]	27
3.2	Arten des Feedbacks und beeinflussende Ausdrücke	30
3.3	Arten des Feedbacks und ihre Umsetzungsformen in der Gestik	31
3.4	Zuordnung der Geste „Winken“ zu den Gestenphasen	40
3.5	Auswirkung der Parametersteigerung bei positivem Feedback (Auszug aus Tabelle A.1)	40
4.1	Gesten, die vom Nao aus motorischen und technischen Gründen nicht ausgeführt werden können	55
4.2	Anzahl der implementierten Stufen pro Geste und insgesamt implementierten Stufen pro Feedback-Art	57
A.1	Auswirkung der Parametersteigerung bei positivem Feedback	74
A.2	Auswirkung der Parametersteigerung bei negativem Feedback	75
A.3	Auswirkung der Parametersteigerung bei Unverständnis	76
A.4	Auswirkung der Parametersteigerung bei wartendem Ausdruck	77
A.5	Auswirkung der Parametersteigerung bei einer Begrüßung oder Verabschiedung	78

Anhang

A Auswirkungen der Steigerung der Parameter auf die Gesten

In den folgenden Tabellen werden die Effekte der Steigerungen der einzelnen Parameter — räumliche Ausdehnung, zeitliche Ausdehnung, Beschleunigung, Frequenz und Stetigkeit — auf die Gesten zu allen Feedback-Arten betrachtet. Dabei wird dargestellt, wie sich die Expressivität verändert, wenn ein Parameter einen höheren Wert erreicht.

Die Bedeutungen der einzelnen Zelleninhalte sind folgende:

- \uparrow = Erhöhung der Expressivität
- \downarrow = Verringerung der Expressivität
- \rightarrow = keine Veränderung der Expressivität
- — = Parameter existiert nicht oder nur bedingt

Tab. A.1: Auswirkungen der Steigerung der Parameter auf verschiedene Gesten zur Darstellung von positivem Feedback beziehungsweise Akzeptanz

Steigerung der	Nicken	OK-Geste	„Daumen hoch“	auf Schulter klopfen	High-Five	Klatschen	Jubeln	Hüpfen
räumlichen Ausdehnung	↑	—	—	↑	↑	↑	↑	↑
zeitlichen Ausdehnung	↓	↑	↑	↑	↓	↓	↓	→
Beschleunigung	→	↑	↑	↑	↑	↑	↑	↑
Frequenz	↑	↑	↑	↑	↑	↑	↑	↑
Stetigkeit	↑	↑	↑	↑	↑	↑	↑	↑

Tab. A.2: Auswirkungen der Steigerung der Parameter auf verschiedene Gesten zur Darstellung von negativem Feedback beziehungsweise Abweisung

Steigerung der	hängende Schultern	ver-schränkte Arme	„Daumen runter“	Kopf-schütteln	erhobener Zeigefinger	Hand mit Handfläche zum Gegenüber	Hände vor das Gesicht heben	mit dem Fuß aufstampfen
räumlichen Ausdehnung	↑	—	—	↑	—	—	↑	↑
zeitlichen Ausdehnung	↑	↑	↑	↑	↑	↑	↑	—
Beschleunigung	↓	↑	↑	→	↑	↑	↑	↑
Frequenz	—	—	↑	↑	↑	—	—	↑
Stetigkeit	↑	↑	↑	↑	↑	↑	↑	↑

Tab. A.3: Auswirkungen der Steigerung der Parameter auf verschiedene Gesten zur Darstellung von Unverständnis

Steigerung der	Kopf schieflegen	Hand unter dem Kinn, Ellenbogen aufgestützt	Kopf wiegen	Hand horizontal kippen	am Kopf kratzen	Schulter- zucken	Arme angewinkelt, Handflächen nach oben zeigend
räumlichen Ausdehnung	↑	—	↑	↑	—	↑	↑
zeitlichen Ausdehnung	↑	↑	↑	↑	↑	↑	↑
Beschleunigung	↓	→	→	→	→	↑	→
Frequenz	—	—	↑	↑	↑	↑	—
Stetigkeit	↑	↑	↑	↑	→	↑	↑

Tab. A.4: Auswirkungen der Steigerung der Parameter auf verschiedene Gesten zur Darstellung eines wartenden oder neutralen Ausdrucks

Steigerung der	umher- schweifender Blick	Verlagerung des Körper- gewichts	Hand unter dem Kinn, Ellenbogen aufgestützt	Hände hinter dem Kopf	Arme in die Hüfte gestemmt	mit dem Fuß klopfen	Umher- laufen
räumlichen Ausdehnung	—	↑	—	—	↑	↑	↑
zeitlichen Ausdehnung	↑	↑	↑	↑	↑	↑	↑
Beschleunigung	—	↑	↑	↑	↑	↑	↓
Frequenz	↑	↑	—	—	—	↑	↑
Stetigkeit	↑	↑	↑	↑	↓	↓	↑

Tab. A.5: Auswirkungen der Steigerung der Parameter auf verschiedene Gesten zur Darstellung eines Beginns oder einer Begrüßung und eines Abschlusses oder einer Verabschiedung

Steigerung der	Aufstehen/ Hinsetzen	Körper zu- wenden/ abwenden	Hand zum Grüß anheben	Winken	Ver- beugen	Hand reichen und schütteln	Um- armen	„Luft- kuss“	Küss- chen
räumlichen Ausdehnung	—	—	↑	↑	↑	↑	↑	—	—
zeitlichen Ausdehnung	↓	↓	↑	↑	↑	↓	↑	→	↑
Beschleunigung	↑	↑	↑	↑	↓	↑	↑	—	—
Frequenz	—	—	—	↑	—	↑	—	—	—
Stetigkeit	↑	↑	↑	↑	↑	↑	↑	↑	—

B Anpassung der Expressivitätsspannen an den Nao und Unterteilung der Gesten in Stufen

Die nachfolgenden Abbildungen zeigen die Anpassungen der Expressivitätsschemata an den humanoiden Roboter Nao auf Grund der Einschränkungen durch einzelne Bewegungen, die nicht ausgeführt werden können. Weiterhin wird dargestellt, wie die Expressivitätswerte auf einer Skala von 0 bis 1 verteilt sind und welcher Wert welchen Gesten zugeordnet ist. Die Einteilung der Gesten in Stufen, die die gewünschte Feedback-Art zusätzlich verfeinern, wird zudem deutlich.

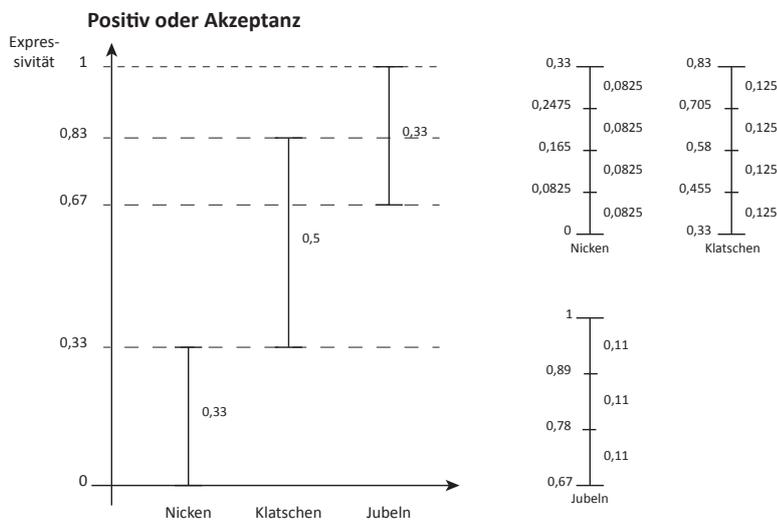


Abb. B.1: Darstellung der Anpassung des Expressivitätsschemata an den Roboter Nao und Unterteilung der Gesten des positiven Feedbacks in Unterstufen

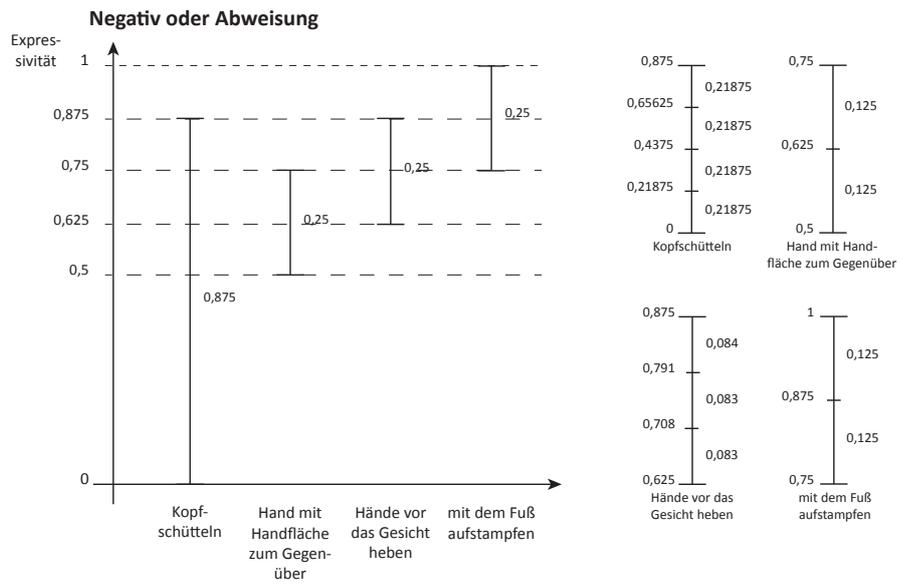


Abb. B.2: Darstellung der Anpassung des Expressivitätsschemata an den Roboter Nao und Unterteilung der Gesten des negativen Feedbacks in Unterstufen

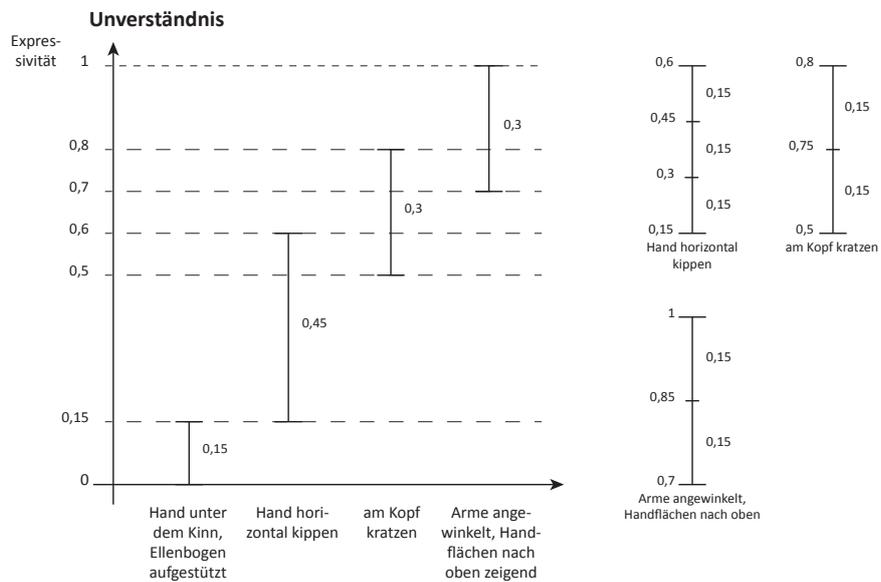


Abb. B.3: Darstellung der Anpassung des Expressivitätsschemata an den Roboter Nao und Unterteilung der Gesten des Feedbacks für Unverständnis in Unterstufen

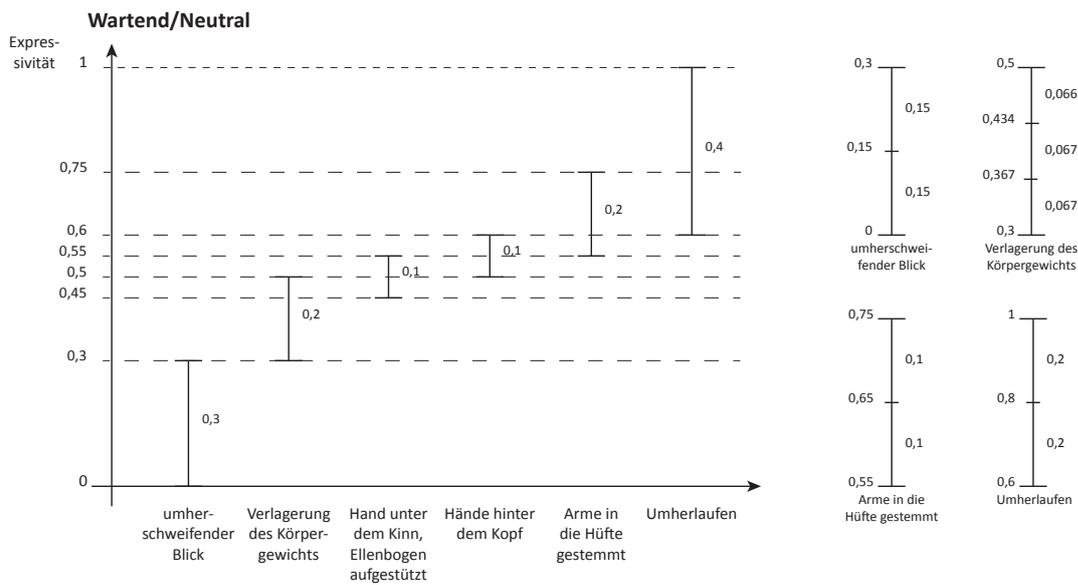


Abb. B.4: Darstellung der Anpassung des Expressivitätsschemata an den Roboter Nao und Unterteilung der Gesten des wartenden oder neutralen Feedbacks in Unterstufen

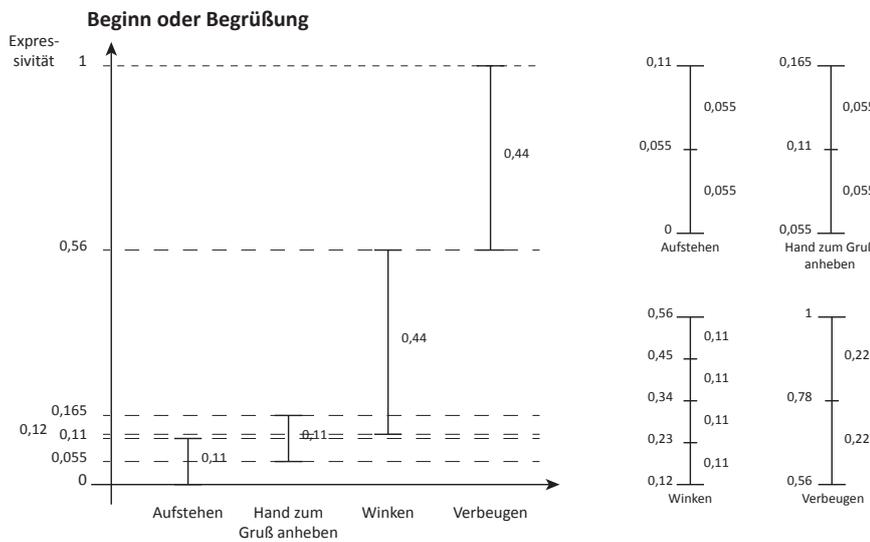


Abb. B.5: Darstellung der Anpassung des Expressivitätsschemata an den Roboter Nao und Unterteilung der Gesten des Beginn- oder Begrüßungsfeedbacks in Unterstufen

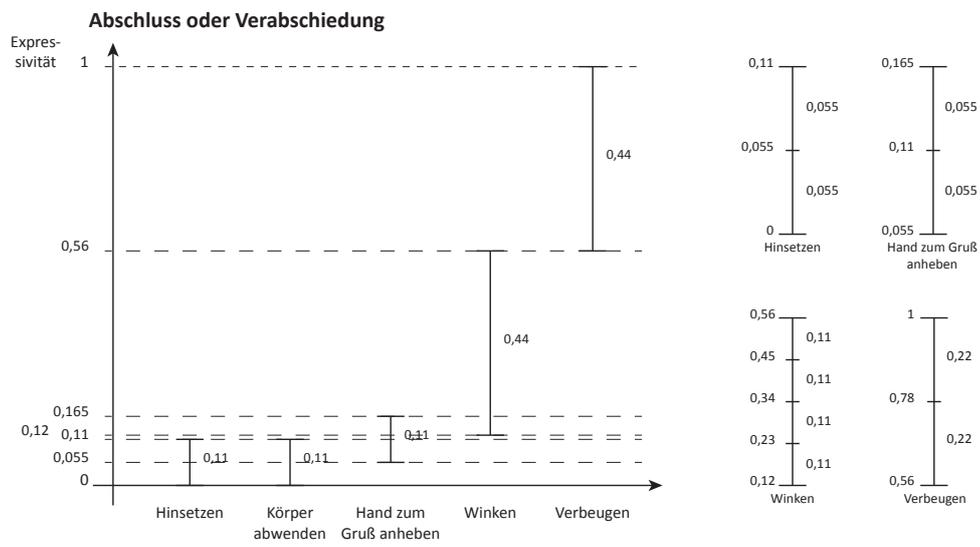


Abb. B.6: Darstellung der Anpassung des Expressivitätsschemata an den Roboter Nao und Unterteilung der Gesten des Abschluss- oder Verabschiedungsfeedbacks in Unterstufen

C Aufbau der implementierten Gestenerzeugung in der Entwicklungsumgebung Choregraphe

In den nachfolgenden Abbildungen ist die Implementierung des Gestensets innerhalb der Entwicklungsumgebung Choregraphe zu sehen. Die Abbildungen stellen die Hauptbox mit ihren Unterboxen und deren Aufbau dar, der in allen Unterboxen eine ähnliche Erscheinung hat, wodurch ein Beispiel der Box zu positivem Feedback sowie eines zur Klatschen-Geste zur Veranschaulichung genügt.

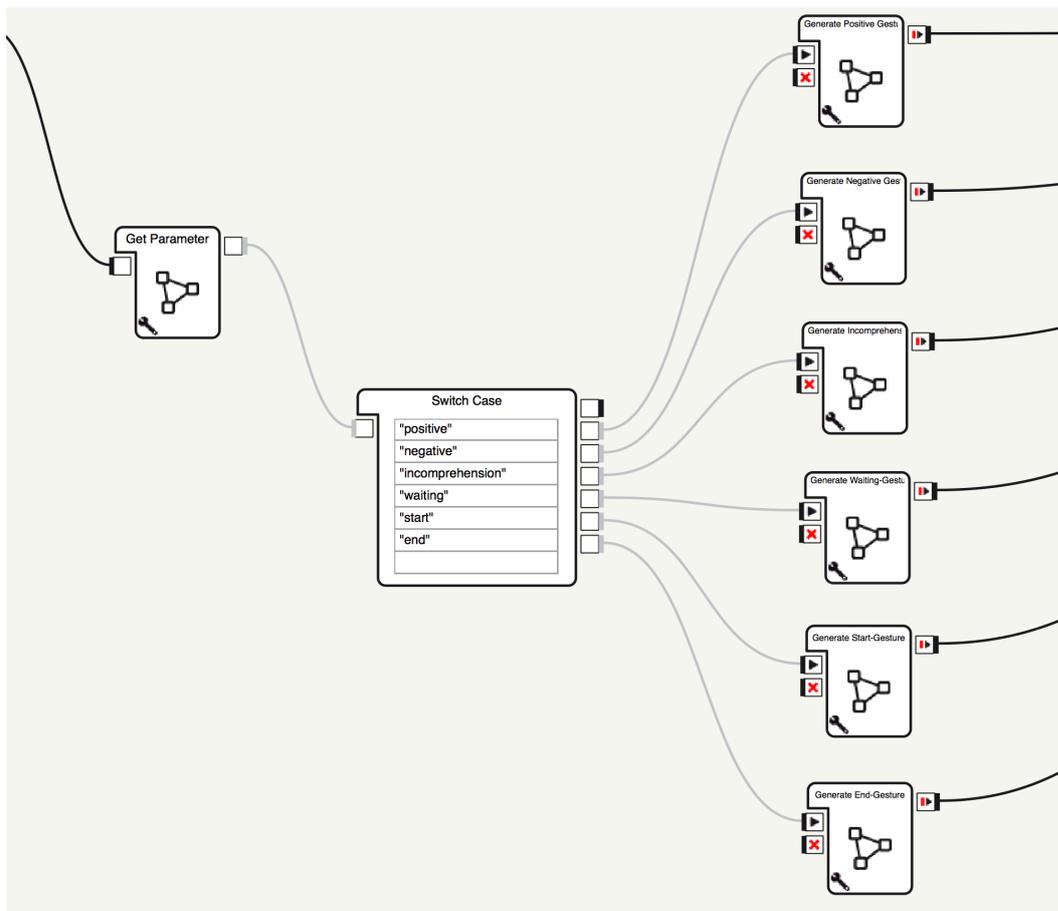


Abb. C.1: Aufbau der Box „Gesture Generation“ mit der Überprüfung der Feedback-Art und Weiterleitung an die passende Box

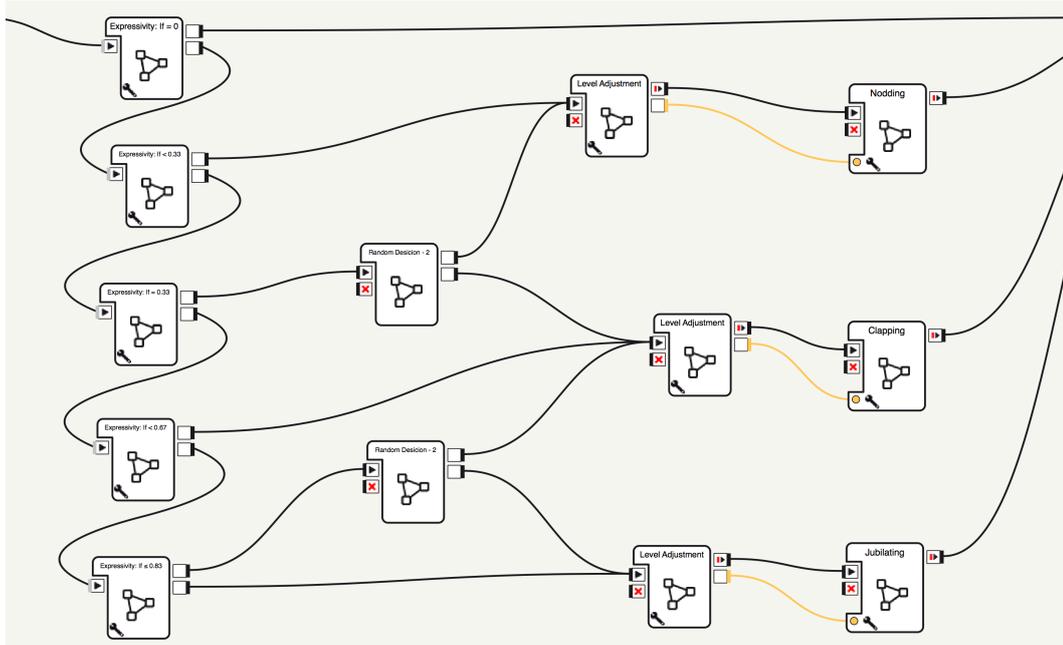


Abb. C.2: Aufbau einer Box, die eine Geste zu einer Feedback-Art erzeugt, am Beispiel des positiven Feedbacks: Durch Unterscheidung der einzelnen Pfade durch die Höhe der Expressivität und je nach der Anzahl der passenden Gesten wird eine zufällig Entscheidung getroffen.

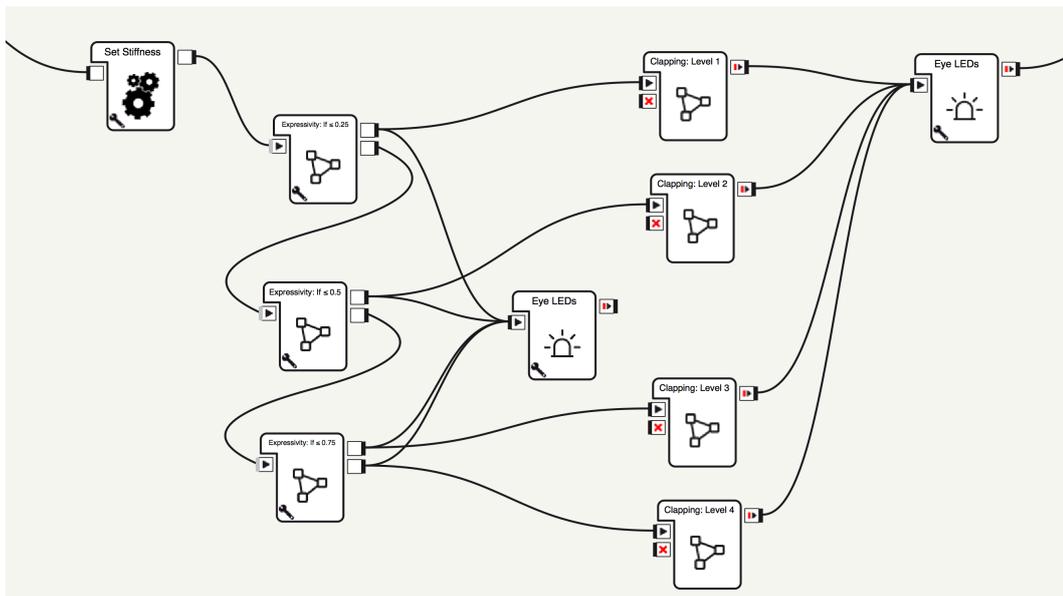


Abb. C.3: Aufbau einer Box zur Generierung einer einzelnen Geste am Beispiel des Klatschens: Die Unterteilung der Stufen nach Höhe der Expressivität und anschließende Weiterleitung zur Generierung der ausgeführten Geste.