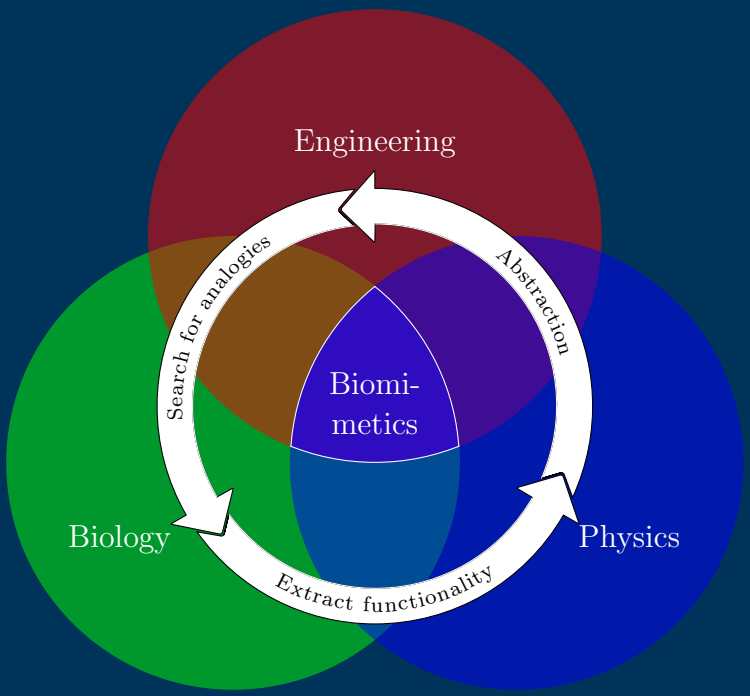
Advances in Social and Underwater Robotics through Biomimetic Design



Dissertation

Stefan Georg Sosnowski

Lehrstuhl für Steuerungs- und Regelungstechnik
Technische Universität München
Univ.-Prof. Dr.-Ing./Univ. Tokio Martin Buss

Advances in Social and Underwater Robotics through Biomimetic Design

Stefan Georg Sosnowski

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Gordon Cheng, Ph.D.

Prüfer der Dissertation:

1. Priv.-Doz. Dr.-Ing. habil. Kolja Ernst Kühnlenz
2. Univ.-Prof. Dr.-Ing. Sandra Hirche

Die Dissertation wurde am 20.06.2013 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 29.11.2013 angenommen.

Preamble

When starting my studies of electrical engineering with a specialization in high voltage in mind, I could never have imagined to end up in robotics and be captured by biomimetic systems this much. I am constantly amazed by the marvels nature evolves and by learning about the complexity and subtleties of human interaction. It is really enjoyable to take on the challenge to achieve similar (or daringly: even better) results with robots. This thesis summarizes part of the research I conducted at the Institute of Automatic Control Engineering (LSR) at the Technische Universität München since late 2007. The journey from the first ideas to this work has been an exciting one and I am glad for everything I was able to explore. I could not have undertaken it without the support of numerous people.

First of all, I would like to thank my advisor Dr. Kolja Kühnlenz. Starting from my bachelor thesis, he invoked my interest in Human-Robot-Interaction (HRI) and accompanied me along all of my academic milestones. I am very grateful for his support and guidance, while enabling me to keep an open mind and pursue my research freely. My thanks go also to my co-advisor Prof. Sandra Hirche. Without her support, it would not have been possible to bring the Snookie project to its current form. This applies also to Prof. Martin Buss, as the open work environment at his institute highly facilitates research.

Over the years, I have been involved in a number of projects and collaborations. Some of the research conducted in these can be found in this thesis. For the Human-Robot-Interaction part, I especially thank Dr. Christoph Mayer for his dedication to merge our research into several demonstrators and the insightful discussions on the countless challenges we faced. This holds true also for Barbara Kühnlenz. I am very grateful for her inspirations in our cooperations and discussions about the subtleties of psychology and social science. I would like to thank Prof. Norbert Bischof, Dr. Michael Zehetleitner and Dr. Felix Schönbrodt for introducing me and my student Isabell Borutta to the Zurich Model of Social Motivation and the fruitful collaborations. We both appreciated the dedication and support very much. Thanks to Isabell for making the model and experiments work.

The Snookie project would be inconceivable without Dr. Moritz Franosch and Andreas Vollmayr. My sincerest gratitude to Dr. Franosch for his dedication and input during the first three years of the project. Equal gratitude goes to Andreas Vollmayr, who took over in the same manner when Dr. Franosch left academia. To say that I have learnt a lot from both during our lab work and in our discussions (also on the different point of views of engineers and physicists) is an understatement. Many thanks to all my students in the project, especially David Lenz, Nader Kuhenuri-Chami, Philipp Mittendorfer, Nora Martiny, Martin Thoma and Florian Helmhold.

From the Cotesys Multi Joint Action team, my thanks go to Jürgen Blume, Alex Bannat

and Tobias Rehrl for the countless hours spent in integrating our systems into a demonstration scenario and making the lab a fun place.

One of my first encounters with a larger project was the Autonomous City Explorer (ACE) project. My thanks go to my colleagues Dr. Georgios Lidoris, Dr. Klaas Klasing, Dr. Florian Rohrmüller, Dr. Quirin Mühlbauer, Dr. Andrea Bauer, Dr. Tinging Xu and Dr. Tianguang Zhang for all the effort put into making ACE go to Marienplatz. Never again would I underestimate the power of *printf* debugging.

I appreciate that our work in ACE and my work on EDDIE has been carried on in IURO. In particular, I would like to thank Christian Landsiedel and Malte Buß for putting up with my code.

Besides research, a lot of people have made daily life at the institute enjoyable. Many thanks to my long term office mate Thomas Schauß for always having an open ear and a nice word to all sorts of topics. To Ulrich Unterhinninghofen and Nikolay Stefanov for letting me in on their electronics wizardry. My insatiable hunger for hardware for many a project was professionally taken care of by Josef Gradl, Horst Kubick, Domenik Weilbach, Robert Geng, Tobias Stoeber, Thomas Lowitz, and Wolfgang Jaschik. Thanks for working so much with me that I was considered part of the workshop "inventory". With the shift to more administrative duties during my time at LSR, I also came to know the qualifications of Dr. Dirk Wollherr, Dr. Marion Leibold, Larissa Schmid, Waltraud Werner, and Wibke Borngesser in all sorts of non-research related matters. My sincerest gratitude for all the support, both in professional and family matters, and making my life easier.

Last but not least, I would like to thank my parents and my brothers Sebastian and Fabian for believing in me and my work and all they have done so far. Especially my father has nurtured my love for technology through his skilled craftsmanship, the shared interests and providing me with all the tools and gizmos a child could wish for. Giving me my first computer at the age of seven, at a time when this was not as common as today (after I had taken apart his computer first), had a major influence on my career path. The biggest thank you has to go to my wife Lena and my son Simon. Thanks for all the love, support, patience and bearing with the sacrifices this thesis brought along for more than 2000 days. The importance of this backing cannot be overstated.

Munich, June 2013 Stefan Sosnowski

Acknowledgments

This work was supported in part within the DFG excellence initiative research cluster *Cognition for Technical Systems (CoTeSys)*, see www.cotesys.org, by the Institute for Advanced Study (IAS), Technische Universität München, see also www.tum-ias.de, within the EU-STREP project IURO (Interactive Urban Robot), supported by the 7th Framework Programme of the European Union, ICT Challenge 2 Cognitive Systems and Robotics, contract number 248317, see www.iuro-project.eu, and the Bernstein Center for Computational Neuroscience (BCCN) Munich.

Contents

Abstract

In this work, biomimetic design is applied to such distinct fields of research as social and underwater robotics. The key aspect of the first two parts of this work is the enhancement of the perception of a biomimetic robotic head as a social actor through transfer of findings from social sciences, anatomy and semiotics. In the third part, a biomimetic lateral line system enables underwater robots to utilize flow field information for object detection, reconstruction and navigation. Necessary intermediate steps include the identification of analogies, abstraction of function principles and implementation in a technical context.

In the first part, frameworks are introduced for the synthesis of implicit communication modalities with both face and neck. They enable a robot to express emotions and other non-verbal signals, including biological analogies of zoosemiotic signals. Experimental evaluations show overall good recognition rates for emotional expressions and significant impact of the zoosemiotic features and the biomimetic neck posture approach.

Based on the expression capabilities established, the second part focuses on the employment of these in social interaction. Mimicry and smiling are identified as beneficial in service encounters and other dyadic interactions and are therefore assessed for their applicability in artificial agents. A novel model for mimicry allows automatic imitation of facial and emotional expressions. Furthermore, an extended version of the system-theoretic model of smiling enables agents to generate context-sensitive smile variations as well as emotional mimicry. Both models show in experimental evaluations with the prior developed biomimetic head significant improvements in terms of subjective performance and empathy ratings.

Fish demonstrate remarkable abilities such as object avoidance, object discrimination and environment mapping by sensing the hydrodynamic image of the surrounding flow field on their body. As first step to make the hydrodynamic image usable in robotics, neuromasts, the basic components of a lateral line system, are abstracted by hot thermistor anemometry. Sensors are integrated on the robot *Snookie*, which is a specifically designed test bed for the artificial lateral line system. Measures of sensor characteristics reveal agreement with theory and compliance with the requirements derived from the findings on neuromasts. Further steps introduce two different methods of flow field information extraction. An analytical model is presented, that enables detection of walls and objects. Experimental evaluations demonstrate the applicability of the proposed method and the developed sensors. Furthermore, simulations of *Snookie* show the feasibility of obstacle avoidance with a robot solely based on flow field information. The second method is a numerical approach, implementing a novel process for flow field reconstruction. By means of this approach, the state-of-the-art is extended to object reconstruction with arbitrary static solid boundaries. Based on the reconstruction, simultaneous localization and mapping (SLAM) techniques add to the applicability of artificial lateral line systems from reactive obstacle avoidance to more sophisticated navigational tasks. Successful object reconstruction and navigation in complex environments is shown in simulation.

Zusammenfassung

In dieser Dissertation wird das Prinzip des biomimetischen Designs auf so unterschiedliche Forschungsfelder wie soziale Robotik und Unterwasserroboter angewandt. Der entscheidende Aspekt für die ersten beiden Teile dieser Arbeit ist die Steigerung der sozialen Fähigkeiten eines biomimetischen Roboterkopfs durch den Transfer von Erkenntnissen aus Sozialwissenschaften, Anatomie und der Semiotik. Wichtige Zwischenschritte beinhalten die Identifikation von Analogien zur Problemstellung, die Abstraktion von Funktionsprinzipien und die Implemetierung in einem technischen System.

Im ersten Teil werden Module zur Synthese von impliziten Kommunikationsmodalitäten mit Gesicht und Genick vorgestellt. Diese befähigen einen Roboter Emotionen und andere nicht-verbale Signale auszudrücken, einschließlich der biologischen Analogien der zoosemiotischen Signale. Experimentelle Auswertungen zeigen weitgehend gute Erkennungsraten der emotionalen Ausdrücke und einen signifikanten Einfluß sowohl der zoosemiotischen Merkmale als auch des biomimetischen Genickmodells.

Der zweite Teil der Arbeit fokussiert auf die Anwendung der eingeführten Ausdrucksmöglichkeiten in sozialen Interaktionen. Mimikry und Lächeln werden in der Literatur als vorteilhaft für Servicetätigkeiten und anderweitige Interaktionen identifiziert und daher auf Ihre Eignung in technischen Systemen überprüft. Ein neuartiges Mimikry-Modell erlaubt die automatische Imitation von Gesichts- und emotionalen Ausdrücken. Darüber hinaus erlaubt eine erweiterte Version des system-theoretischen Lächelmodells sowohl die Darstellung von kontext-sensitiven Lächelvariationen, als auch die Imitation von emotionalen Zuständen. Beide Modelle erzielen im Zusammenspiel mit dem zuvor vorgestellten biomimetischen Roboterkopf signifikante Verbesserungen in der Bewertung von subjektiver Performanz und Empathie gegenüber dem Roboter.

Fische zeigen bemerkenswerte Fähigkeiten, wie zum Beispiel Kollisionsvermeidung, Unterscheidung von Objekten und Kartographierung der Umgebung. Dies geschieht durch das Fühlen des hydrodynamischen Abbilds des sie umgebenden Strömungsfelds mittels Ihres Seitenlinienorgans. Ein erster Schritt um dieses hydrodynamische Abbild für Roboter nutzbar zu machen ist die Abstraktion der Neuromasten, der Basiselemente des Seitenlinienorgans, durch Hitz-Thermistor-Anemometrie. Diese Sensoren werden auf dem Roboter *Snookie* installiert, der als Prüfstand für das künstliche Seitenlinienorgan konzipiert ist. Messungen der Sensoreigenschaften zeigen, dass diese weitgehend der Theorie entsprechen und die aus den Erkenntnissen über Neuromasten abgeleiteten Anforderungen erfüllen. Weiterhin werden zwei verschiedene Methoden zur Extraktion von Informationen aus dem Strömungsfeld vorgestellt. Eine analytische Näherung erlaubt die Detektion der Anwesenheit einer Wand oder eines Objekts. Experimentelle Ergebnisse zeigen die Anwendbarkeit dieser Methode und der entwickelten Sensoren. Des weiteren wird eine numerische Methode vorgestellt, die eine neuartige Möglichkeit zur Rekonstruktion des Strömungsfelds eröffnet. Durch diese Methode wird der Stand der Technik auf die Rekonstruktion von beliebig geformten statischen Grenzen (z. B. Objektränder) erweitert. Basierend auf der Rekonstruktion wird der Einsatzbereich des künstlichen Seitenlinienorgans von reaktiver Objektvermeidung auf Selbstlokalisierung und Kartographierung vergrößert. Simulationsergebnisse zeigen erfolgreiche Objektrekonstruktion und Navigation in komplexen Umgebungen.

1 Introduction

"If we take nature as a guide, we will never go astray."
(Marcus Tullius Cicero)

The biodiversity of an estimated 8.7 million species [11] on earth constitutes a huge source of inspiration for innovations in technical systems. Of the already known 1.2 million species, members of flora and fauna have developed strategies to conquer habitats with such diverse environmental conditions as the darkness and depth of the deep sea, high altitudes, the cold of antarctica, the dryness and heat of the deserts or toxic environments. Besides the environmental adaption, the evolutionary race of arms and struggle for survival has lead to the development of abilities that have yet to be rivaled by technical systems (at least in the same scale): Tardigrades show their ruggedness by surviving unassisted in outer space [13]. Polychaete worms are able to live next to deep sea black smokers (up to 400°C water temperature), solely sustained by a symbiosis with chemoautotrophic bacteria feeding on methane [5]. Snailfish found in 7703 m depth seem to be highly active, despite the low energy environment they live in [6]. This depth is topped by *Abyssobrotula galathea*, which was seen in a depth of 8370 m [6]. In such depths in salt water, approximately 875 kN press on each square centimeter of its body. Also on the plant side, some bristlecone pines are as old as 5000 years [4], while few artificial structures have proven the same durability. Some animals excel in the power to weight ratio: The rhinoceros beetle can carry 850 times its own weight, while carrying the extra loads with "*remarkable economy*" [9]. The spittle bug, with only 6 mm length, can jump as high as 70 cm [15]. Bar-tailed godwit have been recorded to fly 11 500 km non-stop, without refreshment of their energy store [8], while bar-headed geese are reported to fly across Himalaya from sea level in one night, with peak heights up to about 6500 m [10]. Many species in flora and fauna sport a plethora of sensory abilities, covering the information extraction from physical phenomena such as wave propagation in solid, liquid and gaseous media (light, sound and vibrations in spectra not perceivable by humans) and electro and magnetic fields, or chemical signals. While the adaption through morphological changes and new abilities enabled the habituation of an environment in the first place, the manifestations can often be considered as being optimal, following the *Optimal Design Principle* [14], which is a consequence of the struggle for survival. With this promise of optimality, taking inspirations from nature is even more worthwhile.

1.1 Terminology

To make use of the study of biological systems for technical innovations, several disciplines have to be involved, the most common being biology, physics and engineering. The combination of these interdisciplinary fields has been given varying names. While the basic ideas remained the same, the specific terms of denoting this kind of research branched over time, but are often used synonymously.

One of the first denominations, *biophysics*, was given by Otto Schmitt, a Professor of Biophysics, Bioengineering, and Electrical Engineering at the University of Minnesota [12]. He is most famous for his work on the Schmitt Trigger, an analog comparator based on the functionality of the nerves in squids.

> *"Biophysics is not so much a subject matter as it is a point of view. It is an approach to problems of biological science utilizing the theory and technology of the physical sciences. Conversely, biophysics is also a biologist's approach to problems of physical science and engineering, although this aspect has largely been neglected."* (Otto Schmitt [12])

He later broadened *biophysics* to *biomimetics* [1]. The term biomimetics stems from the ancient greek *bios*, meaning *life* and *mimesis*, standing for *to imitate*. Another name was given by Jack E. Steele, coining *bionics* as a blend of *biology* and *electronics* on a conference in 1960 [7]. The prominence of bionics was promoted through pop culture, as it was a central theme of television series and novels. A more focused view on the knowledge transfer from biological to technical systems by Norbert Wiener and Julian Bigelow is *cybernetics* [7]. These studies of the similarities and differences in biological and artificial systems aim mainly on the principle of *homeostasis* (remaining the same). A more recent addition to the terms is made by Benyus:

> [Biomimicry is ...] *"A new science that studies nature's models and then imitates or takes inspiration from these designs and processes to solve human problems."* (Janine Benyus [3])

Nachtigall imposes some constraints on the aim of the research: Not imitation, but the transfer of the functional principles of biological to technical systems is the goal [345]. In the course of this work, Schmitt's term of *biomimetics* is used, with the additional constraint from Nachtigall. In contrast to biomimicry and cybernetics, it is the least limiting one in the field of study or application. Through the popularization of bionics by the novel *Cyborg* and the TV series *The Six Million Dollar Man*, it is often wrongly associated solely with the fusion of living organisms and machines.

1.2 Biomimetic Design Process

Biomimetics includes the whole design process from the initial research question to the final product. The description of the process of the development of biomimetic systems is two-fold, depending on the starting point: The top-down approach starts with a specific problem and looks for solutions, whereas bottom-up starts with the solution and looks for a technical problem to solve. A more detailed description is given as follows, illustrated in Figure 1.1.

Top-down Approach

Figure 1.1 depicts the top-down approach from the initial challenge to the final product on the left. In the top-down approach, a technical challenge exists, for which a solution is to be found. With a problem formulation resulting from the analysis of the challenge, biological analogies are investigated. This involves the translation of technical to biological terminology, which can result in a reformulation of the problem description. Once a possible analogy is found, the design principles of the specific feature of the biological system are analyzed. From these design principles, an abstract representation of the functionality is derived. This is a crucial point in the development of a biomimetic system, as the decoupling allows for more freedom in the manifestation of the technical solution (for example: materials, form or underlying mechanism). The process of analyzation – including design principles and abstraction – could also lead to a refined understanding of the properties of the biological system. The abstracted concept is now implemented in a technical prototype, which does not necessarily resemble the biological system. After a successful evaluation of its initially stated requirements, the technical solution is integrated in the final biomimetic product.

Bottom-up Approach

The bottom-up process, shown in Figure 1.1 on the right, bears many similarities to the top-down approach. However, the starting situation is different. In the bottom-up process, research in biology brings up an interesting property of a biological system, which is further investigated for its range of application in a technical system. The basis for this development of applications is the studying of the functional morphology, which enables certain abilities or properties of interest. As in the top-down approach, the design principles of the morphology are extracted and abstracted. Within this process, additionally collected information on the biological system can lead to an improved understanding of the previous stages of research. Without the clear goal that is given in the top-down approach, the technical implementation is of explorative nature, as is the application in a biomimetic product.

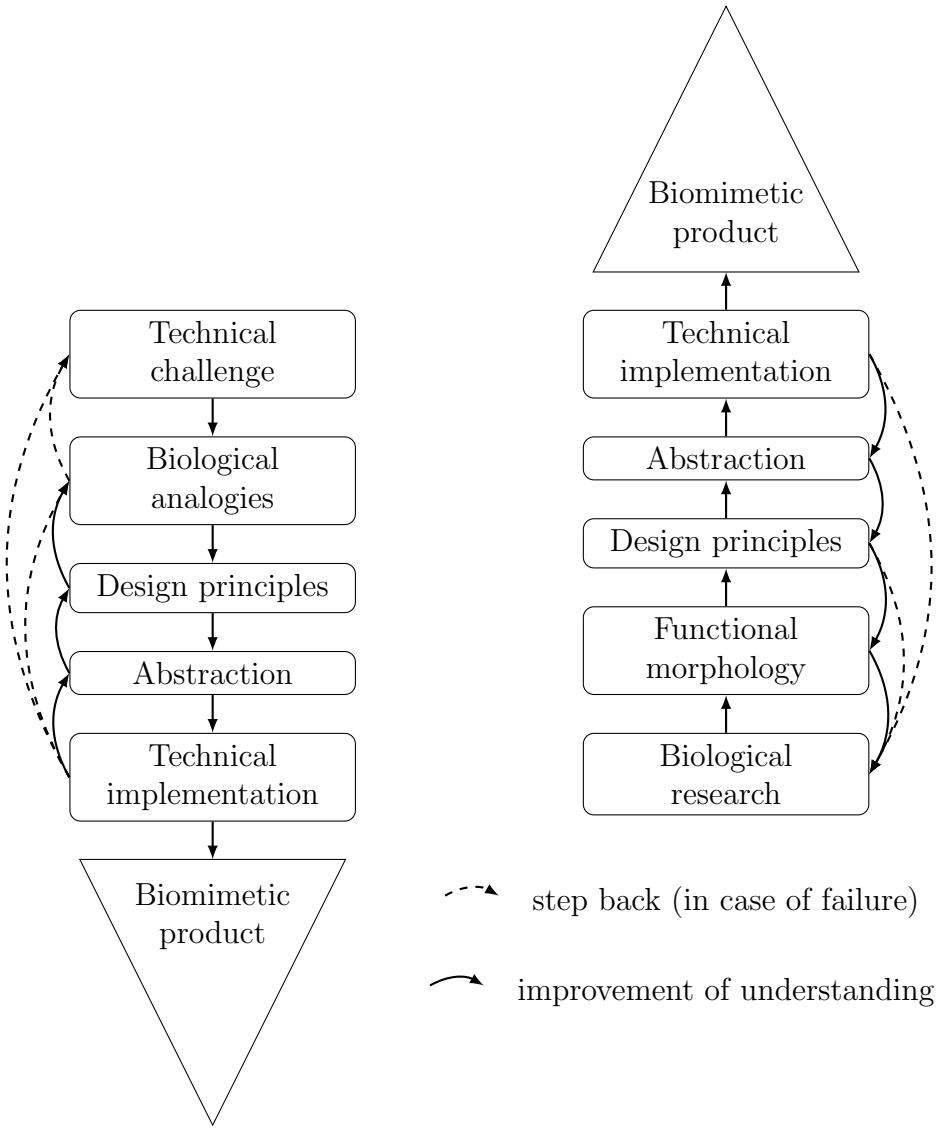


Figure 1.1: Top-down and bottom-up approach, broken down into the single steps that result in the developmental process.

1.3 Problem Statements and Challenges

In this work, challenges in Human-Robot-Interaction (HRI) are addressed with the top-down approach. Human-Human-Interaction (HHI) is regarded as an expert system that presents the solution to the challenges, which is studied extensively in psychology and social science. The realization and application of an artificial lateral line system on the other hand is taken as bottom-up approach and the challenges being of exploratory nature. Research questions are derived from the application range of the biomimetic sensory system. In the following subsections, the challenges are broken down for the specific topics.

Signal expression

The first challenge is to identify employable interaction modalities to enhance HRI, preferably ones that do not interfere with a task of an artificial agent[1]. Aspects of these modalities have been discovered for HHI. With this body of methods and evidence for the effectiveness in HHI, these methods are to be transferred to HRI and made computable. Remaining questions are about the kind of abstractions required for the transfer to technical systems and the limitations of these abstractions. Current research on artificial agents focuses on human appearance, motivated by the direct transfer of HHI capabilities. As the biomimetic top-down approach specifically includes biological analogies, it should be investigated which traits besides human ones can be applied. In previous cases of the implementation of non-human expressive features, the integration in the expression framework lacked justification of the semiotic meaning. Open questions remain on how these facial and extra features can be combined in a computable model and how to fuse them with separately generated signals. While several methods are known in literature to map emotional states of the agent to its joint space, these mappings are limited to specific agent architectures. Therefore, it is of interest to achieve more generalizable abstractions.

Use in interaction

With the expression capabilities as a basis, the second challenge is addressing the need for expressing communication signals other than explicit signals, which is not fully justified for technical systems, yet. In order to compensate the extra effort of including implicit communication signaling in technical systems, benefits for a user need to be investigated. At the same time benefits for the system to aid with the fulfillment of its tasks are of interest. Again the starting question is what findings can be transferred from HHI to HRI. With the focus on the behaviors of mimicry and smiling, which are identified as beneficial in dyadic interactions, analogies and design principles have to be determined and abstracted. Questions in the context of mimicry and smiling remain to what extent these findings can be replicated in artificial systems, with only a subset having been tested, yet. Studies conducted in this area hint at a transferability, but are subject to specific narrow implementations, with partially contradicting results. Moreover, as artificial agents

[1]The term “artificial agent” is used throughout the thesis to describe a technical avatar system. Since the methods are not bound to the implementation of the agent, it can refer both to virtual agents and robots.

commonly serve a purpose besides mere satisfaction in interaction, methods and effects of mimicry on the perception of task performance need to be investigated.

Artificial lateral line

A different challenge arises from the biomimetic bottom-up approach. Behavioral experiments with fish show their ability to detect the direction of motion, speed, shape and size of solid objects. The basic functionality and morphology of the responsible lateral line system is well known. However, the exact transfer from the hydrodynamic stimulus to the excitation of the sensor, the resulting neuronal signals and their processing is still under investigation. This means that so far, attempts to rebuild the lateral line system can only lead to an approximation of the biological source. While biomimetic cilia might come close to the morphology of the biological source of inspiration, the robustness, manufacturing complexity and signal-to-noise ratio are still challenges that prevent the application in an autonomous underwater vehicle. This requires a different abstraction of the biological source. Besides challenges in the implementation of the sensory system, open questions mostly remain in the extraction of information from the perceived flow field. Studies focused on the Mexican cave fish usually only consider the forward problem, modeling the stimulus that occurs from the hydrodynamic interference with objects on the body of fish. To utilize data gathered from the sensors on a robot and get information about the environment, the inverse problem has to be solved. Current solutions to the inverse problem require assumptions about the shape and the number of objects – typically only one – presented to the lateral line system. If the object is assumed to be a dipole, its position can be extracted from its hydrodynamic image. This is a first step towards underwater navigation, providing the information for obstacle avoidance. The application in a real world environment, however, requires the detection and reconstruction of arbitrary shaped objects without prior assumptions. For detection of static objects or mapping of the environment, more information about the object is crucial.

1.4 Contributions and Outline of this Thesis

With regards to the previously stated challenges, the following contributions are part of this thesis:

Implicit Communication Modeling in a Biomimetic Head

In Chapter 2, a framework for the synthesis of non-verbal communication signals for the head, consisting of face and neck, is introduced. It is the result of the abstraction and transfer of findings from HHI to artificial agents. Novel methods and algorithms are proposed to solve the intermediate steps of expressing a non-verbal signal, from the generation of expressions based on the semiotic meaning of the signals to the fusion of the overall expression. Furthermore, work in this chapter extends the state-of-the-art in biologically inspired robotics and expressive agents by not only taking into account the animalistic morphology, but also the semiotic meaning of non-human features. Based on the proposed methods, robotics systems are developed and evaluated. Experimental evaluations show significant

impact of the introduced framework and methods on the perception of the technical system in terms of social aspects.

Integration of Implicit Communication Signals Mimicry and Smiling

In Chapter 3, the application of the aforementioned expression capabilities in interactions are investigated. Behavioral concepts from HHI are narrowed down to mimicry and smiling, as these concepts are identified as beneficial in envisioned scenarios of robots as service agents. Two different models are developed, one based on a mimicry model in cognitive developmental psychology and one based on social motivation. Through extension of the social motivation based system-theoretic model of smiling, it is enabled to perform emotional mimicry besides generating seven smile variations, combining both concepts. The models are integrated and evaluated in an automatic mimicry system for an artificial agent. Benefits through mimicry and smiling identified in literature are verified in experimental evaluations, showing significant support of the initial hypothesis of the transferability of these concepts.

Transfer of a Biomimetic Lateral Line System to an Underwater Robot

In chapter 4, the description of a biomimetic bottom-up approach is given. The transfer is made from the biological system of the lateral line system to a new thermistor based sensor design, which incorporates the flow sensing in the tangential plane and quick response time of the lateral line neuromasts. Experimental results on the sensors show the agreement between measurements and theory. Two different approaches of modeling hydrodynamic stimuli on the artificial lateral line are given. The first is based on the concept of image charges. With the combination of this model and the neuroscientific population vector coding, obstacles can be detected theoretically and in experiments. Also a robot in simulation is able to perform obstacle avoidance solely based on input from the artificial lateral line system. The second approach is to reconstruct the surrounding flow field from the hydrodynamic image on the sensor array. This is a novel solution based on the boundary element method to the inverse problem mentioned in the problem statements. With this method, not only can the array detect obstacles, but also reconstruct their position and arbitrary shape in vicinity of it, eliminating the restrictions and assumptions on the objects necessary in previous state-of-the-art methods. Simulations of the reconstruction in 2 dimensions show, that the spatial resolution of the reconstructed objects is sufficient for SLAM with as few as 9 sensors in the array, enabling a simulated robot for the first time to navigate purely on spatial data generated from the flow field.

2 Implicit Communication Modeling in a Biomimetic Head

New fields of application in close cooperation with humans set novel requirements on the interaction capabilities of robots. In this chapter, a solution to the problem of how to enhance these interaction capabilities is provided by the biomimetic top-down approach. Given that humans are experts in HHI, they can act as the biological source for inspiration. Since a plethora of studies have been conducted in social science and psychology on HHI, a literature review of modalities humans use in communication is conducted and analyzed for the transferability to technical systems. The focus is on non-verbal communication, as this provides means to influence interaction partners besides the task related direct communication. Computable models for the expression of non-verbal social cues are introduced, split into facial expressions and neck postures, as these body parts are most influential. Evaluations with robotic systems show the mostly correct interpretability of the social cues.

2.1 Introduction

> *"Human-Robot Interaction (HRI) is a challenging research field at the intersection of psychology, cognitive science, the social sciences, artificial intelligence, computer science, robotics, engineering and human-computer interaction."* (Kerstin Dautenhahn [82])

The domain and purpose of robots is changing. While the industrial manufacturing domain is still the most important for robotic applications, robots are gradually arriving in our homes, offices, hospitals, homes for the aged, etc., taking over jobs that make interaction with humans necessary. With the paradigm shift from seeing robots as machines for repetitive, highly structured work packages that are processed isolated (or in joint work with other robots), to referring to robots as "co-workers" and service providers, these fields of application require interaction capabilities different from industrial automation. The nature of these interaction capabilities is subject to the research in HRI. The quote of Kerstin Dautenhahn's research field description highlights the interdisciplinarity, arising from the complexity of the challenges faced. Biomimetics fits in this context with providing an approach that can derive input from human-human-interaction, whose level of interaction quality reflects the goal or at least a milestone for HRI. Considering humans as experts in HHI, the biomimetic top-down approach is applicable with the intention of analyzing and transferring the crucial features, that make interaction between humans successful, to a robotic system.

Motivation

The diffusion of robots from industry, with its work environment adapted to the robotic requirements and expert users specifically trained for the handling of automation systems, to households and public areas requires a change in the way of interaction with robotic systems. Non-expert users benefit from intuitive and safe interaction methods that come close to the interaction between humans. Non-verbal communication and the use of social cues play a key-role in these scenarios, providing an intuitive human-machine-interface, that increases efficiency in working with the robot [58] and requires no user-training. Part of this is that the user is able to perceive internal states of the robot in a simple and robust way and that the robot can recognize mental and/or emotional states of a user. For example, recognition of the intention of a user by the robot benefits from this, because our decision process is heavily influenced by emotions [112, 141, 150]. On the other hand, including emotions in the robot architecture can improve the decision process. Adaptive HRI that accounts for the human emotional state becomes possible, enabling the robot to react appropriately in a socially accepted context. An example are agents that do not rely on scripted behavior, enhancing the user experience in games and simulations. A common design goal for them and the robots is to appear and behave *life-like*, which results in *"believable characters"* providing the illusion of life [143].

Another aspect is the use of robots in the care for elderly, a field that arises due to the demographic change in society towards more elderly, but less young people. Robots are expected to help both the elderly and care-takers: The care-takers by assisting with heavy lifting or repetitive tasks, easing the physical demands or leaving more time for interaction with the patients [157]. For the elderly, robots could provide more self-determination (as a tool) and be used for entertainment or even therapeutic purposes [125].

Robots can also benefit from the interaction with humans. Due to the usually unstructured environment that can be found in households or public spaces, robots will need to rely on human interaction partners as sources of information [96].

Outline

The outline of this chapter is as follows: first, in Section 2.2 a brief overview on the necessary aspects of interaction between humans is given and analyzed for the transferability to HRI. The focus is especially on the modalities a technical system can use in interaction, followed by the need to define affect and emotion for affective communication. With these basic aspects in mind, a biomimetic system is developed in Section 2.3 that is able to enrich HRI with emotional expressions and other means of nonverbal communication. Methods to generate these expressions are introduced and applied to two exemplary robotic systems. An evaluation of the performance in HRI is given and concluded with a discussion of the results. Section 2.4 continues the top-down biomimetic design approach by the analyzation of human neck movements. A model for humanoid neck movements is deduced from measurements of human neck movements and mapped to a representation with reduced degrees of freedom. This model is then applied to the problem of finding a solution for position and orientation, while fixating a point in space, for a neck with redundant configuration and evaluated in an experimental setup.

2.2 Aspects of Interaction

In general, people are social beings and enjoy interacting with each other. The quality of interaction ranges from nodding acquaintances to teaming in professional environments to strong bonds in mother-child care and many more. The ability to cope with such different circumstances in a flexible way and still rely on the same methods of interaction seem very promising to enhance the future way on how to interact with machines. So HHI can serve as a reference for the interaction that is desirable to achieve in HRI, providing inspirations on how to design aspects of robotic systems. This is congruent with the approach in top-down biomimetics to look for solutions to a given problem in the biological reference system. Thus, HHI is analyzed and - as far as applicable - transferred to HRI.

For the analysis of HHI in the context of this work, it seems best to start with the question on how humans are interacting. Thus, in the next subsection the modalities for interaction are examined.

2.2.1 Interaction Modalities

Without the aid of technology, human interaction is usually face-to-face in close vicinity, involving some sort of information exchange. The information exchange can be described as

> *"Each noticeable, conscious or unconscious, directed or undirected change of behavior (...), through which a human persuades willingly or unwillingly perception, feelings, affect and thoughts of others."* (R. A. Spitz [163])

.

Following this description, interaction is composed of an action, the change of behavior that others can notice, and a form of persuasion, that influences the interaction partner through communication encapsulated in the action.

The communication can be split into verbal communication, with the meaning conveyed by the content of a message, and non-verbal communication, which provides information - that is not necessarily congruent with the verbal message - through visual, auditorial or haptic channels. Mehrabian [56] and Cowie [71, 181] term this dichotomy *"explicit"*, with an emphasis on verbal-linguistic cues, and *"implicit"*, regarding information about the sender conveyed through subtle communication phenomena. The importance of non-verbal signals is often summarized in Mehrabian's rule[1], stating that words only account for 7% of the judgement of a statement, whereas the tone of voice accounts for 38 % and the body language for 55% of the rating [56].

A system that is designed to interact with an user to act towards a shared goal must have some means for making explicit statements to communicate. This can be achieved via the visual channel by text or signs or via speech. So the more interesting aspect is

[1]It should be noted that this rule is only valid within the context of the original experiment and the communication of feelings (i.e. like - dislike). Experiments by Ekman *et al.* [78] on the "Relative Importance of Face, Body and Speech in Judgements of Personality and Affect" could not replicate such a distinct distribution, but rather a dependency on the context of the communication.

how to enrich the communication by implicit, non-verbal, information. A categorization of implicit behavior is given by Ekman and Friesen [42], which can be used as a template for the transfer to HRI:

- **Emblem:** non-verbal acts that can be accurately translated into words (for example a handshake, frown or smile)
- **Illustrator:** part of speech to emphasize meaning (for example head and hand gestures)
- **Affect display:** functions to convey affect and emotions
- **Regulator:** speech regulation of interaction partners (for example encouragement to talk)
- **Adaptor:** satisfaction of bodily needs (for example scratching)

The concepts of emblems, illustrators, regulators and adaptors seem straight-forward for a rule-based implementation in a robot. However, the concept of affect and emotions needs to be examined more in-depth, as will be shown in the next subsection.

2.2.2 Affect and Emotions

> *"Everybody knows what an emotion is, until asked to give a definition."*
> (B. Fehr and J. A. Russell [121])

Affect and emotions are an important part of the implicit message conveyed through nonverbal interaction. Therefore, a system would benefit from understanding and utilizing these concepts. The problem is that even for people, it is hard to grasp the notion of affect and emotions. The quote from Fehr and Russell illustrates one of the underlying problems: everyone experiences emotions and affect and therefore has an idea what these things feel like. The phenomena are ubiquitous in prosaic language, but these sometimes fuzzy meanings do not comply with scientific standards. On the other hand science can not give a proper, consistent, definition as well. To agree on a terminology for this work, first a working definition of affect and emotion needs to be established. This working definition does not raise the claim to be all-embracing, but should establish grounds for the implementation in a robotic system (tailored specifically to the interactive aspects). For an in-depth coverage of the historical evolution of the definitions and the connection between definitions from different disciplines, the reader is referred to some of the extensive surveys that have been written on this subject [120, 31, 30, 135, 90, 68, 139].

Definition

Despite - or maybe because of - extensive research on emotions in several disciplines over the course of the last century, researchers tend to define emotions in a way that emphasizes the aspect of emotions they are working on [46]. This leads to specific, "partial" definitions of the overall domain and is an on-going process in a variety of scientific disciplines such

as psychology, philosophy, neuroscience, cognitive sciences and – recently – engineering. A survey by Kleinginna & Kleinginna in 1981 [139] collected and compared the definitions of emotions by the number of appearances in scientific literature, which resulted in splitting the domain of emotions in four components: A subjective, a cognitive, a physiological and a behavioral component. This shows that a concise definition is not a trivial task and that existing ones are often missing elements of what is considered an emotion in research.

One of the difficulties in defining what emotions are is that they are a hypothetical construct that can not be observed directly. Only various phenomena are observable, which are assumed to be more or less a direct expression of emotions [173]. Even the term emotion stems from the latin word *"emovere"*, which means to *move, agitate*, referring to the physiological expressions. This makes finding a general definition harder that incorporates the subjective and cognitive components as well. With the application in HRI in mind, the focus on the manifestations of emotions and affect is not an issue, as this is the information that is conveyed in the implicit part of communication. Therefore, the specifically for this case relevant definitions can be narrowed down to those describing psychophysiological changes:

- **Expression:** Phenomenon that is objective and observable [181]
- **Feeling:** Phenomenon that is subjective and part of consciousness [181]
- **Affect:** Neurophysiological state, that is consciously accessible as a simple, non-reflective feeling [84]
- **Emotion:** Referring to transitory conditions of the organism [102].
- **Emergent emotion:** Archetypal emotional phenomenon, in which the above elements of emotionality briefly come together and either dominate the way a person acts and thinks, or need to be held in check by a deliberate effort [181]. It is according to Cowie *"... a short, intensive, clearly event triggered emotional event ..."* [85].

In the course of this work emergent emotions, which lead to short, noticeable expressions, are of interest. This is due to the different time-scales associated with the above categories and the primarily brief interactions between humans and robots considered here. Thus, whenever emotions are mentioned, more precisely *emergent emotions* are meant.

Representation

In order to be able to implement affect and emotions in a technical system, a representation of the emotional state within a framework is necessary. Representations are often derived from the experiments of the observable psychophysiological manifestations of emotions, like facial expressions, or the cognitive processes that are expressed in a semantic rating of words based on their emotionality. Three approaches to represent emotions or affect are described in psychological literature [115]: appraisal-based, categorical and dimensional.

In the appraisal-based approach, the emotional state is related to a sequence of stimulus evaluation checks [86]. Both the internal state and the external influences are processed in a continuous, subjective evaluation and checked for the "*weal or woe*" [153] of the organism. The appraisals of external and internal stimuli are merged to one system state, representing a particular emotion. Several alternatives for how this process works have been proposed [68, 74, 87, 99, 116]. The merging process of the stimuli is straight-forward to obtain a specific emotional state in this representation. However, the inversion of the problem that is necessary to access the emotional state of the interaction partner is ambiguous. Thus, for the implementation in the aspired HRI system, this representation is neglected.

The categorical approach is based on the concept that a discrete set of primary (or basic) emotional states exists. This approach is convenient for a technical implementation, because the complexity of emergent emotions and expressions is reduced to a limited number of states. The number of primary states and their composition is not finally agreed on and varies depending on the roots of the respective emotion theory (psychology, philosophy, neuroscience). An exemplary overview is given in Table 2.1, showing that there is – to some degree – an intersection between the different sets of emotions.

Descartes [29]	**Watson** [171]	**Ekman** [100]	**Plutchik** [44]	**Izard** [109]
wonder		surprise	surprise	surprise
joy		joy	acceptance	enjoyment
	fear	fear	fear	fear
			panic	distress
		disgust	disgust	disgust
hate	rage	anger	rage	anger
sadness		sadness		
desire	sexual activity		desire	
love				
				interest
				contempt
				shame

Table 2.1: Exemplary listing of discrete emotional state sets

One common framework used in technical system is the set of basic emotions found by Ekman and Friesen [100] (included in Table 2.1), which are recognized in their facial expressions independent of the cultural background of an observer [441]. The aspect of universality can already be found in the work of Charles Darwin [19], stating that signs of emotions are rooted in biology and therefore universal. This claim of universality does not imply that there is no cultural influence on the *rules* of displaying emotions that are used to *"manage the appearance of particular emotions in particular situations"* [441], but that the unfiltered, raw emotions and their respective expressions are unbiased. The evidence for the cultural independency and the wide acceptance in research on the classification of emotions make this set an option for implementation in a robotic system to understand and display the cues in interaction related to emotions.

In contrast to the monopolar categorical approach, the bipolar dimensional approach assumes a dependency between the emotions within an emotional state space. Dating back to at least Wundt in 1896 [119], most theories on the state space of emotions incorporate the dimensions of *valence* (pleasant - unpleasant) and *arousal* (aroused - calm). One of the renowned two-dimensional approaches is for example Russell's circumplex model of affect, with valence as axis of abscissae and arousal as axis of ordinates [166]. The extension to three dimensions can be seen in the work of Osgood *et al.*, which describes the whole semantic space of words in the former two dimensions plus *potency* (strong - weak) as a third dimension [108]. Studies by Russell [45] and Mehrabian [102] confirm these dimensions for the rating of emotional terms, with a slight variation in the terminology: their third dimension is *dominance* (dominant - submissive), which is related to potency [166]. According to Mehrabian [102], several studies showed that "*... the same or similar sets of three factors could be used to describe emotions and social cues, including postures, body positions, facial and vocal expressions, gestures, and movements ...*" [124, 108, 64, 50]. The studies were performed with a measurement technique called semantic differential. It consists of a set of bipolar adjectives that can be rated on a seven-point Likert scale. The bipolar adjectives are assigned to one of the Pleasure Arousal Dominance (PAD) dimensions and thus personal attitude (in a self-reflective process) or subjective observations can be rated along the dimensions. A description of the dimensions is given by Mehrabian [102]:

- *"**Pleasure-displeasure** corresponded to cognitive judgments of evaluation, with higher evaluations of stimuli being associated with greater pleasure induced by the stimuli."*
- *"Judgments of high-low stimulus activity corresponded to State **Arousal-nonarousal**, defined in terms of level of mental alertness and physical activity..."*
- *"**Dominance-Submissiveness** was defined as a feeling of control and influence over one's surroundings and others versus feeling controlled or influenced by situations and others..."*

The fact that supportive evidence for this representation exists and that assessment tools for the emotional state in form of the semantic-differential are available [16] favor the application in HRI. In contrast to the categorical approach, the PAD dimensional model includes subtle emotional phenomena, which are observable in HHI besides the pronounced basic emotions. Moreover, transitions between emotional states can be mapped as trajectories within the emotional state space. However, the categorical and dimensional approach are not mutually exclusive. Studies on the *Circumplex model of affect* [166] and the PAD model by Russell and Mehrabian [45] show that emotional terms, labeling the emotional states of the primary emotions, can be assigned to coordinates within the PAD space. This combination is further used to represent emotionally expressive states of the technical system for HRI.

After this brief description of the aspects of interaction, the next step is to develop a system that is able to utilize not only the explicit form of communication, but also the implicit, non-verbal part of communication. Following Mehrabian's rule by the trend that

is implied by his experimental findings, first a form of embodiment is necessary to account for the meaning in body language.

2.3 Facial Expression Synthesis

> *A man's face as a rule says more, and more interesting things, than his mouth, for it is a compendium of everything his mouth will ever say, in that it is the monogram of all this man's thoughts and aspirations.*
>
> (Arthur Schopenhauer)

For the embodiment of a robot to interact explicitly and implicitly with humans, the face is the most important part. It is used to convey information via the visual channel in terms of social cues like emotional facial expressions, gestures like winking or smiles and visualization of syllables in the form of visemes. An argument of the importance of the face over the body is given in a neurological experiment by Meeren *et al.* [98]: If face and body convey conflicting emotional information, participants choose the depicted emotional state according to the facial expression rather than the bodily expression[2].

In this section, necessary aspects for a robot to be able to non-verbally communicate will be derived, with the human face as the template. In the biomimetic process, design options are considered, which influence the perception of the robotic face. Following the design, a method to generate expressions - with the focus on emotional expressions - is introduced.

2.3.1 Problem Statement

The kinematics of a face are highly complex and thus kinematic configurations of the biomimetic system and the facial muscular structure are not necessarily congruent. This makes an abstraction of the mapping between muscular structure and actuators necessary. In this work, the abstraction is achieved by using the Facial Action Coding System (FACS), a system developed to describe facial muscle movements objectively. A method to decouple FACS expression synthesis and the mapping to robot actuators, by means of an intermediate step of motor activations, which provides a way for more generalizable mappings, is proposed. Furthermore, a challenge in the transfer of the capabilities of the face to a robot is that the appearance has direct influence on the perception of the technical system. Thus, design considerations are highlighted and FACS is extended by including zoosemiotic features. An analysis of animalistic features and their signal effects is given and possible effects in HRI evaluated.

Once the kinematic structure of the biomimetic face is determined, the question is how to model the expression generation, in order to achieve the desired implicit and explicit non-verbal communication. In this work, the FACS based framework is used to model the display of affect, signals and emblems. With the incorporation of emblems and signals, the fusion of singular expressions to an overall facial expression must be solved, which is done by an algorithm of merging singular expressions on the action unit level.

[2] Anger: Congruent display: $83.6 \pm 2.6\%$ accuracy; incongruent display: $70.6 \pm 3.6\%$ accuracy
Fear: Congruent display: $78.8 \pm 3.3\%$ accuracy; incongruent display: $63.9 \pm 5.5\%$ accuracy

For the display of emotional states, so far, most robotic systems either still apply the limited categorical approach, or use dimensional approaches which are only applicable to the specific implementation. In this work, one way of merging the categorical and dimensional approach is introduced, which serves as a basis for the facial expression generation in the PAD dimensions. The calculation of facial expressions for the respective PAD coordinates is shown via the decomposition of FACS activations for the PAD dimension octants and the concept of motor activations provides an additional hardware abstraction layer.

2.3.2 Related Work

Since Fong's survey on socially interactive robots [154] in 2003, which comes to the conclusion that social robotics is still an immature field of research, HRI has gained traction and produced a number of robotic systems. See [168] for a more recent overview. Most robots with facial expressions can be categorized according to their level of abstraction in the biomimetic design.

A minimalistic approach is taken by robots such as Sparky [43], which is an early representative of expressive robots. Scheeff *et al.* followed traditional animation principles in the design to achieve simplistic emotional expressions. In a similar way, the LEGO robot Feelix [164] was designed. With robots such as the Character Robot Face [162] and eMuu [334], the minimalism is a design parameter for achieving a simplified user interface. This is much in the sense of the initial idea by Hara [59], who wanted to employ facial expressions as an intuitive way of accessing the overall state of a complex system, e.g. nuclear plants.

iCat [91] falls in the same category for user interfaces, but at the same time integrates zoomorphic elements. These animalistic features should free the user of expectations of humanlike capabilities or interaction style. They are also integrated to make the robot appear more as a stuffed animal, lowering inhibition thresholds for children and elderly, such as in Probo [134]. The toy like appearance should not detract from the fact that these robots can be highly complex mechatronic systems, e.g. Leonardo [242]. So far, little attention has been paid to explain the signaling effects of the individual zoomorphic implementations in an overall concept and evaluate the expected effects of the elements. To the best of the author's knowledge, no comprehensive analysis or experimental study, besides [55], on the effect zoomorphic elements exists.

A step up closer to the biological source are robots with the stature – and often behavior – based on human infants. Infants are expected to have a limited set of capabilities and features they can voluntarily control. This expectation can be exploited for reduction of the design, partially as a necessity to reduce the complexity and to consciously differ from a realistic appearance to avoid the uncanny valley. An example for this design strategy is KASPAR [36], which is "*minimally expressive*". At the same time, the specific proportions and behavioral schemes of infants elicit care taking behavior in humans. Kismet [32] is well known for its successful engagement of human care takers. Other robots such as MEXI [67] or Pearl [175] use the child like proportions for a higher user acceptance. These types of robots serve also as a test bed for the emulation of the cognitive development of children, with two representatives being iCub [89] and Infanoid [88]. They follow the paradigm that cognitive development and learning of infants is coupled to the embodiment.

There are also a number of robots with a mechatronic look that have the appearance of an adult. Robots such as WE-4RII [62], Flobi [118] or Fritz [176] are therefore admitted more competence, shifting the level of interaction.

The class of androids aims at coming as close as possible to the human appearance. While part of the robotics community tries to avoid the uncanny valley, androids are often specifically designed to overcome or disprove it. Their design involves copying the human appearance in high detail, down to artificial muscles as actuators [21] or elasticity and irregularities of the skin [174]. Recently, there has been a trend to form the robot as "clone" of a specific person, for example the Geminoid series [73], the Repliee series [18, 54] or Hanson Robotics robots [174, 75]. These robots mostly resemble mature persons, which is useful for tasks which require a high level of attributed competence, such as educational tasks. In first experiments, Saya was used for teaching children [158] and Geminoid DK for a lecture in university [110].

2.3.3 Biomimetic Design

Based on the aforementioned problem statements, the goal here is to transfer the functionality of the human face, which is able to generate up to 7000 different expressions [77], to a technical system. Facial expressions are the result of the contraction or relaxation of subcutaneous voluntary muscles, which are attached to the skull and end in the skin. A mere copying of the muscle structure and skin properties of the face to a robot is not desirable, as facial musculature is know to vary between individuals, including asymmetries [151]. Therefore, an abstraction is necessary.

Facial Action Coding System

An objective description of the changes in the facial expression can be extracted from the activation of facial action units, atomic facial activities that are formed from combinations of muscles. Based on the work by Hjortsjö [81], Ekman and Friesen developed a coding system for the observable movements in a face called the Facial Action Coding System (FACS) [38]. In FACS a set of 32 action units (AUs) involving muscles and 23 action descriptors (ADs) are defined. A list of action units and the assignment to the respective muscles based on the 2002 edition of the FACS manual is given in Table A.1. Action descriptors are not directly linked to muscle groups, but describe actions like turning and tilting the head.

The generalizable abstraction of facial movements allows to design a robotic face with the actuators resembling action units, not single muscles. This greatly reduces the kinematic complexity of the design, while capturing the essential functionality.

As FACS is defined as an objective measurement of facial movements, action unit activations can be considered as being generic. The mapping of action units to the motor or joint space of a robot τ, however, is dependent on the configuration of the robot. For simplification, an intermediate step of a set of *motor activations* M_A is introduced, that resembles action units in the joint space of the robot, see definition 2.3.2:

Definition 2.3.1. $\tau : AU \rightarrow M$

Definition 2.3.2. $M_A = \{(ma_1, \ldots, ma_n) \mid ma_j \in [-1, \ldots, 1],\ j = 1, \ldots, n\}$.

The number of motor activations ma_j is dependent on the number of robotic actuators n. Motor activations are used to compute the actual motor commands in a later stage. The advantage of motor activations over a direct mapping to motor commands is the decoupling of the action unit mapping from the actual ranges of the actuators. The range extension from $[-1, 1]$ for the motor activations in contrast to $[0, 1]$ for the action unit activations is a result of the possible combination of antagonistic action units in a single robotic actuator.

The last step is the mapping from motor activations M_A to motor commands M_C.

Definition 2.3.3. $M_C = \{(mc_1, \ldots, mc_n) \mid mc_j \in [p_{min,j}, \ldots, p_{max,j}],\ j = 1, \ldots, n\}$.

The resulting motor command mc_j that is computed from the motor activation ma_j is dependent on the motion range of the actuator, which is defined for each actuator by a maximum position $p_{max,j}$, a minimum position $p_{min,j}$ and a position which can be seen as neutral $p_{neutral,j}$:

$$mc_j = \begin{cases} (p_{max,j} - p_{neutral,j})ma_j + p_{neutral,j} & if\ ma_j \geq 0 \\ (p_{neutral,j} - p_{min,j})ma_j + p_{neutral,j} & if\ ma_j < 0 \end{cases} \tag{2.1}$$

Animalistic Features

The aforementioned Facial Action Coding System is a compilation of facial actuators of a human face. Considering the fact that a biomimetic design is not limited to humanoid-inspired sources, the communicative aspect of animalistic features should be explored. Research in zoosemiotics[3] [95] shows that animals are able to communicate interspecific, i.e. with humans, making it a valuable source for additional communicative elements in technical systems.

A catalogue of animal signals and their messages is given by Smith [177], considering the signals under their relevance in behavioral messages, modifiers and identifiers, see Table 2.2 for an excerpt and Table A.2 for the complete version.

The focus here is on the behavioral messages as, according to Snowdon [41], *"The majority of the behavioral messages listed can be inferentially linked to affective states, such as aggression, fear, affiliation, sex and ambivalence, with only a few types of messages not directly related to some affective state."* The behavioral messages in Smith's catalogue can be classified based on the three main communication modalities - visual, auditory and olfactory. The latter two are not going to be explored further in this work, as humans commonly do not sense olfactory communication signals consciously and are hardly receptive for chemical signals of other species, let alone their meaning. Concerning auditory or vocal signals, additional elements like infra- or ultrasonic sounds are not audible, thus can be neglected for HRI. Sounds within hearing spectrum, like purring for pleasure, or high pitch screams of fear, are often closely related to non-verbal humanoid sounds and thus add no benefit over using sounds users are highly familiar with. Therefore, only visual signals will be described, in the for this work relevant affective dimensions pleasure, arousal and dominance. Common animalistic elements to convey visual signals include freely movable ears,

[3]The study of communication in animals

Behavioral Messages
Interactional behavior (type of interaction unspecified)
Attack
Escape
Copulation
Affiliative
Indecisiveness
Locomotion
Site specific (staying at current location)
Seeking (attempting to perform another behavior such as affiliation or escape)
Receptive (to interaction from others)
Attentative (vigilant, monitoring)

Table 2.2: Excerpt from messages of animal signals [177], with modifications from Snowdon [41].

tails and the erection of dermal appendages [19]. While the freely movable ears are usually found in mammals, tails are shared by reptiles and mammals, and dermal appendages can be found in all three vertebrae classes. Dermal appendages in this case include movable elements connected to the epidermis, such as feathers, hairs or scales.

Pleasure display, besides playful or affiliative behavior, is not as strongly represented in zoosemiotics as the display of arousal or dominance. Affiliative behaviors such as licking or rubbing can be seen as affective gestures, establishing contact with a nursing or care-taking character. One sign of valence across mammals with the necessary range of motion in the ears can be seen on the orientation. High valence results in ears facing down, stretched to the sides, as seen in dogs and cats while caressing [180]. Low valence such as sadness results in a drop in body tension. Due to the muscle relaxation, ears tend to hang down [19]. For tails, a generalizable statement is difficult, as different species produce ambivalent signals for pleasure.

Arousal manifests in the tail in either vibrating or swinging motions, or curling in excitement [19]. In contrast to periodic wagging movements, the swinging motions have an irregular character. It is also shown by the erection of dermal appendages. The display of arousal with an erected crest is a characteristic of the cockatoo family. The magnitude of erection can be varied and is tied to the level of arousal [149].

Dominance is the most widespread display of emotional states in zoosemiotics. This could be due to the fact that manipulation of others with signals in this domain is beneficial against predators, rivals and in courtship [41], all important aspects for survival.

To show dominance, the strategy for all vertebraes is to increase the perceived volume of one's own body. Examples are manifold and mostly involve the erection of dermal appendages: Gorillas raise a crest of hair, lions their mane, birds ruffle feathers, toads and frogs inhale air and reptiles such as cobras extend their neck disc or the frill of the frill-

necked lizard. Ears and tails are straightened up to enhance the overall impression, such as seen in the aggressive stance of dogs (neck hair raised, ears and tail pointing up) [19].

Very low dominance, which is associated with fear or submissiveness, is accompanied by an inversion of the aforementioned principle of volume enlargement. Submissiveness is shown by a strive for volume decrease, to appear as small as possible. Birds for example press their feathers to the body, while dogs and cats pluck in their tails and have their ears backwards and facing downwards [19, 180].

Implication for technical systems from these zoosemiotic aspects is, that all actuated mechanisms, changing either the overall volume or the height of the system significantly could alter the perception of communicated dominance of the system. Another implication is that actuators, that are identified as tails, could influence the perceived arousal level through the amplitude and irregularity of the tail motions. This holds also for the erection and movement of head-bound elements, such as antennas. Finally, the orientation of ears is an indicator for pleasure and dominance.

With the focus on facial elements, two zoosemiotic relevant features are exemplary chosen for further inspection: the crest of a cockatoo on top of the head and multi-species inspired, ear-like features at the side of a robotic head. With these features, all three emotional dimension are covered. The crest of a cockatoo is a prominent signaling feature on the head of the bird, see Figure 2.1a. It is used in non-agonistic display involving excitement or alarm, intense social interaction and aggression. A technical implementation could thus enhance the perception of arousal and dominance.

The ear-like features take inspiration from two species. The mechanical design is based on the frill of a frill-necked lizard, see Figure 2.1c. The frill of these lizards is made from skin with spines of cartilage, enabling the lizard to spread the usually folded frill to the side of its head. It is used for intraspecific communication (agonistic and non-agonistic) and predator deterrence and is, relative to body size, *"... one of the largest and most spectacular display structures seen in any animal species"* [35]. Again, a technical implementation could enhance the perception of arousal and dominance, as the frill-related behaviors of the lizard are linked to these dimensions. To account for the valence dimension, the signaling with the frill mechanism is complemented by the signaling extracted from dog ears, see Figure 2.1b.

Childlike Characteristics

Other characteristics influencing the appearance of the head, besides the mentioned facial features, are its proportions. According to Konrad Lorenz, the appearance of infants and certain young animals triggers emotional reactions, to which an observer can not resist, as they are innate [133].

These childlike features are characterized as follows:

- The head is especially thick and large in comparison to the body
- The frontal skull is more dominant than the facial bone and the frontal skull is more convex.

- The eyes are large and are usually seated below the center of the skull
- Extremities are particularly short and thick
- The shape of the body is at large roundish
- The surface condition is soft, elastic and smooth
- The chubby cheeks are round and salient

Figure 2.2 highlights the features with the outline of an infant head in comparison to an adolescent and an adult. The characteristics evoke sympathy and a care-taking behavior [133]. Therefore, they are desirable in contemporary robotic systems, which are often error-prone and limited in their capabilities. However, the attribution of competence might be hindered, if robots resemble too much of an infant. So the application of these design characteristics need to be based on the purpose of the robot. For the communicative aspect of the robot, the elicited sympathy is beneficial, thus the design should incorporate childlike features.

These design goals seem to be contradictory to the head design guidelines by DiSalvo *et al.* [175], recommending a wider than tall head and a reduced forehead. This should preserve the *"robotness"* in the appearance. However, if the *"productness"* factor is of less concern, the waiving of a skin can increase the robotness while still eliciting the care-taking behavior achieved through childlike features [32].

2.3.4 Emotional Expressions

With the decision on the FACS based kinematic structure and the overall design, the question arises on how to generate the implicit behavior of affect display from Section 2.2.1. Two concepts for the synthesis of emotional expressions based on the discrete categorical approach and the dimensional approach are introduced. This involves, in a nutshell, the following steps: First of all, given a particular emotional state, the corresponding movement of the facial elements must be determined. The FACS provides a framework for this task, which allows an objective statement of the facial elements involved in an emotional expression. Once the setting for the AUs is determined for each emotional state, the AUs need to be matched to the actuators of the robot. In a final step, the commands for the respective actuators need to be computed to emulate the AU set.

Next, these steps are formulated in a mathematical way. In order to synthesize an expression, the mapping from an emotion domain E to the corresponding AUs domain AU to the motor domain M must happen, see Figure 2.3.

Definition 2.3.4. $\nu : E_B \rightarrow AU_B$

Definition 2.3.5. $\chi : E_{PAD} \rightarrow AU_{PAD}$

These mappings will first be shown for the categorical approach. Later the mappings for the dimensional approach will be based on the basic emotions concept.

Basic emotions

This implementation of the categorical approach is based on the basic emotions by Ekman *et al.* The basic emotions concept defines six emotional states plus neutral, which form a proper subset in the emotion domain, with emotional states $E_B = \{\text{joy, sadness, surprise, fear, anger, disgust, neutral}\} \subset E$. These states need to be mapped to the set of AU activations AU, which is specified in definition 2.3.6.

Definition 2.3.6. $AU = \{(au_1, \ldots, au_{46}) \mid au_i \in [0, \ldots, 1],\ i = 1, \ldots, 46\}$.

For the function ν the circumstance can be utilized that, in the concept of universal facial expressions, each emotional state in E_B is associated with a specific object of AU activations in the subset AU_B. For example, the activation object with $au_{B,i} = 0\ \forall i$ can be defined as the *neutral* emotional expression.

Following the AU to basic emotion mapping by Smith and Scott [167], the function ν can be described as given in Table 2.3. Action units not specified are set to $au_i = 0$. Given the fact that the mapping $AU_B = \nu(E_B)$ is static and linear, it can be written as a matrix multiplication $\boldsymbol{AU_B} = \boldsymbol{N}\boldsymbol{u}$, with the input vector $\boldsymbol{u} = [e_{joy}\ e_{sadness}\ e_{surprise}\ e_{fear}\ e_{anger}\ e_{disgust}]^T$ and each state $e \in \{0, 1\}$:

$$\begin{pmatrix} au_{B1} \\ \vdots \\ au_{B46} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} e_{joy} \\ \vdots \\ e_{disgust} \end{pmatrix} \tag{2.2}$$

$\boldsymbol{AU_B}$	Joy	Surprise	Anger	Disgust	Fear	Sadness
1	0	1	0	0	1	1
4	0	0	1	1	1	1
5	0	1	1	0	1	0
6	1	0	1	1	0	0
7	1	0	1	1	0	0
9	0	0	0	1	0	0
10	0	0	0	1	0	0
12	1	0	0	0	0	0
15	0	0	0	0	0	1
26	1	1	0	0	1	0
27	1	1	0	0	1	0

Table 2.3: Basic emotions to AUs mapping $AU_B = \nu(E_B)$, based on Smith and Scott [167].

Pleasure Arousal Dominance

Here, the transformation from PAD, as a dimensional approach, to facial expressions is described. While Ekman's basic emotion are rooted in the research on facial expressions and thus provide a direct connection, the transfer from PAD to expressions is not that obvious. However, following Russell's studies on the circumplex model of affect [166]

and the three-factor theory of emotions [45], emotional terms - which correspond to the basic emotions - can be defined within the dimensions of pleasure, arousal and dominance. Similar categorizations of emotional terms in the dimensions of valence and arousal are given by Plutchik [44] and Whissell [160]. This allows to situate the basic emotions in the PAD dimensions and use these as reference points for the transformations from a PAD state to the corresponding facial expression.

The emotion domain formed by the dimensions of pleasure, arousal and dominance E_{PAD} is a subset of the overall emotion domain E, with the basic emotions E_B being a subset of E_{PAD}, thus: $E_B \subset E_{PAD} \subset E$. The emotional state e_{PAD} is defined by $e_{PAD} = (P, A, D)$, where P is the value of pleasure or valence, A is arousal and D is dominance, with $P, A, D \in [-1, 1]$. Since E_B is a subset of E_{PAD}, the basic emotions can be represented in E_{PAD} with their respective PAD coordinates, see Table 2.4.

Dimension	Joy	Surprise	Anger	Disgust	Fear	Sadness
Pleasure	0.76	0.40	-0.51	-0.60	-0.64	-0.63
Arousal	0.48	0.67	0.59	0.35	0.6	-0.27
Dominance	0.35	-0.13	0.25	0.11	-0.43	-0.33

Table 2.4: Location of basic emotions in the PAD dimension according to Russell [166].

As every basic emotional state e_B is associated with a 46-tuple of AUs $\boldsymbol{AU_B} = [au_{B,i}]$, these points in E_{PAD} can be displayed according to the previously defined mappings $\nu(\cdot)$ and $\tau(\cdot)$. For all other points, the mapping $\nu(\cdot)$ must be extended to $\chi(\cdot)$. For this extension, an interpolation between basic emotional states for each AU activation is needed. This interpolation is done octant-wise in the PAD space, with the specification that $\boldsymbol{AU_{PAD}} = f(P, A, D)$ is a continuous function between octants for all $(au_1, \ldots, au_{46})$. The division of the PAD space in octants is in compliance with the definition of the *mood octants* by Mehrabian [102] and allow for an adjusted interpolation that takes into account the extensive change in facial expressions between positive and negative pleasure, arousal and dominance values.

The activation of the animalistic features can be derived from the behavior of the inspirational sources. For the crest (Z_1) and the ear-like features (Z_2 – Z_4), the degree of erection au_{Z1} and au_{Z2} are directly mapped to the dimensions of arousal and dominance, with $au_{Z1,2} \in [-1, 1]$ and $au_{Z1,2} = max(A, D)$. This is based on the bird and frill-necked lizard behavior. The pleasure dimension is responsible for the upwards or downwards facing of the ear-like features, so the position of the ear-like features is proportional to pleasure P, with $au_{Z3,4} \in [-1, 1]$, $au_{Z3,4} = P$. Additionally for EDDIE, the spreading of the ear-like features is proportional to arousal A, with $au_{Z3} = A$ and $au_{Z4} = -A$, $\forall A \geq 0$.

2.3.5 Signals

Besides the emotion display, the biomimetic head should be able to convey additional explicit and implicit information through the face on the visual communication channel. This additional information can aid the understanding in conversations, provide signals

for social cues and rituals and make a technical system appear more lifelike. The signals included here are visemes for speech visualization, idle motions, which are small, task-independent motions, and gestures.

Visemes

In verbal communication, the articulation of speech is accompanied by visible movements of articulators in the face, such as lips. These movements provide a secondary source of speech information [40], which can be assessed by the communication partner. Both visual and auditory information have an influence upon the speech perception and are merged for the decoding, making the decoding process more robust for congruent signals or result in the perception of altered information for conflicting signals [172]. Speech information on the auditory channel can be interpreted as a sequence of discrete components, with *phonemes* being the smallest acoustically distinguishable unit [40]. Similarly, *visemes* are the smallest distinguishable visual unit, a term coined by McGurk and MacDonald [172]. As not all phonemes are visually distinguishable, the number of visemes is smaller than the number of phonemes, so that the mapping from phonemes to visemes is many-to-one.

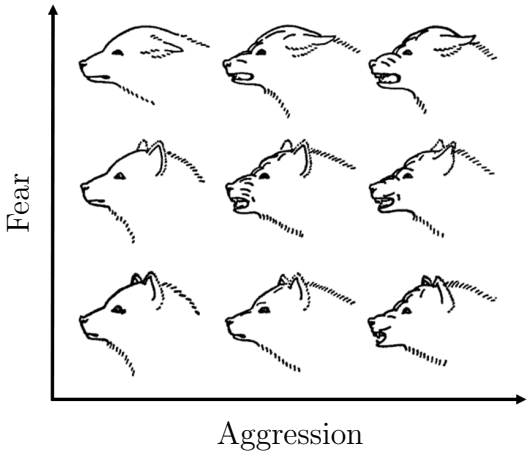
In order to visualize that the robot is talking and to benefit from the second information channel, visemes are transferred to the robotic head. Each viseme is represented with a set of action unit activations in the Facial Action Coding System, leading to a visually distinguishable movement of articulators in the face. Using the previously established action unit to motor mapping $\tau(\cdot)$, each viseme can be displayed in the motor domain.

Gestures

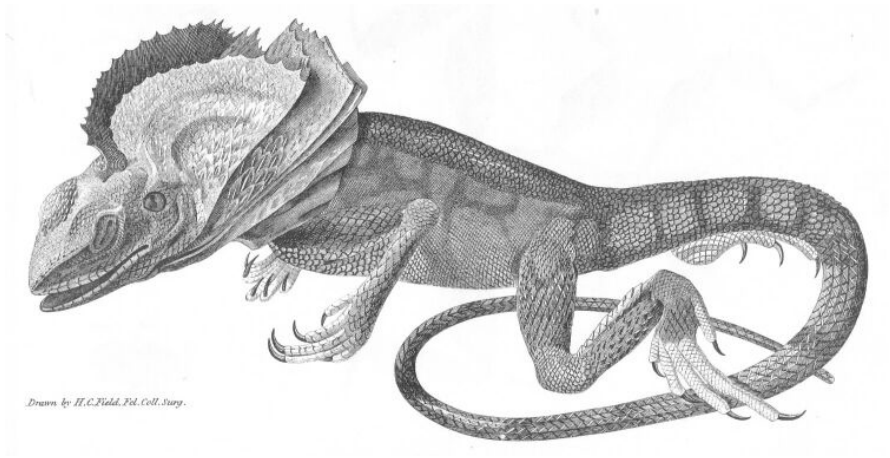
Gestures, resembling the function of emblems (see Section 2.2.1), are voluntary movements in the face to convey implicit information. They bear important functions in HHI, such as signaling turn-taking in conversations or emphasizing the meaning of speech.

The facial gestures that are of relevance here are gestures that can take the role of illustrators, regulators and punctuators during a conversation. The class of punctuators [63] is specifically tied to verbal utterances, separating sequences of words, for example with eye brow or gaze signals. A survey on facial gestures is provided by Zoric *et al.* [131], based on the gesture classes given by Pelachaud *et al.* [83]. The collection of gestures from this survey, together with a description and usage case in HRI, is given in Table 2.5. The organization of the entries in the table is as follows: The *FACS* column denotes the involved action units (AUs), movement descriptors (M) or action descriptors (ADs). *Gesture* is the action that is taken with the involved action units or action descriptors. In the *Description* column, a semantic description of the observable signals or behaviors is given. For nearly all gestures, *Attributes* can be given that can modulate the action of the specific action unit or action descriptor, and thus alter the perceived implicit message. For example, a frown with the eyebrows can either serve as a punctuation mark, if executed with a high amplitude, but at normal speed, or signal distress if done very quickly. The gestures given in this table can be represented with the activation of the corresponding action units and FACS descriptors. This allows the gestures to be included in the underlying FACS framework for the overall facial expression synthesis.

(a) Sulphur-crested Cockatoo (Cacatua sulphurea) [138]



(b) Signaling of dog ears [129]



(c) Frill-necked lizard (Chlamydosaurus kingii) [111]

Figure 2.1: Sources for the bio-inspired design of the extra facial features

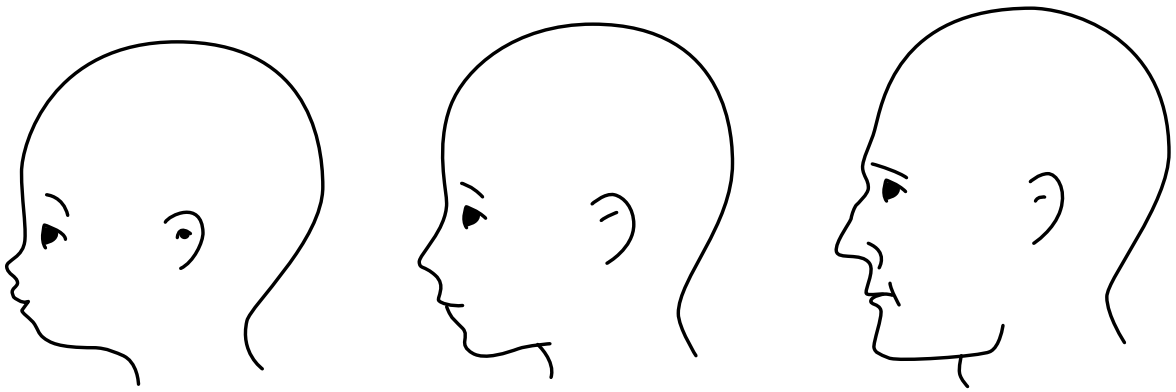


Figure 2.2: Comparison of the outline of an infant, an adolescent and an adult head [169]

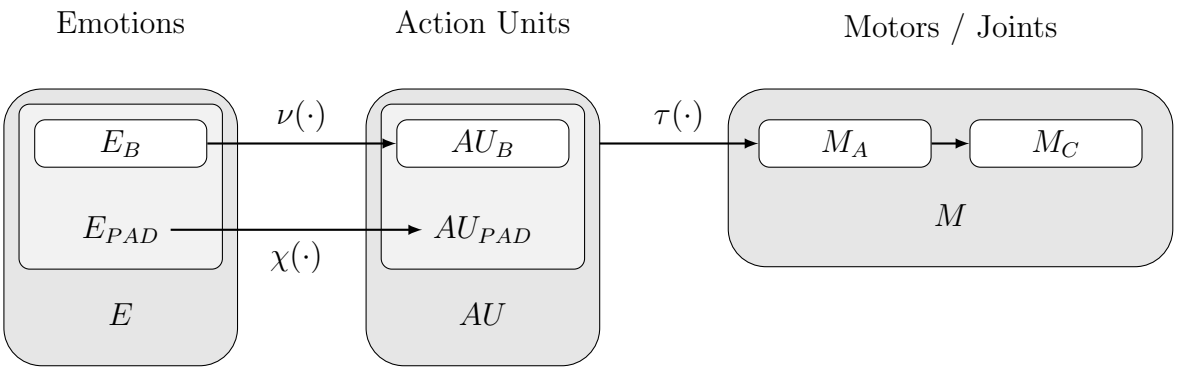


Figure 2.3: Mappings from emotion domain E to the resulting facial expression within the motor domain M, with subdomains E_{PAD} as the PAD state space and E_B the emotional state space of the basic emotions. AU_B and AU_{PAD} are the respective action unit subdomains within the overall domain AU. M_A are motor activations, which result in motor commands M_C.

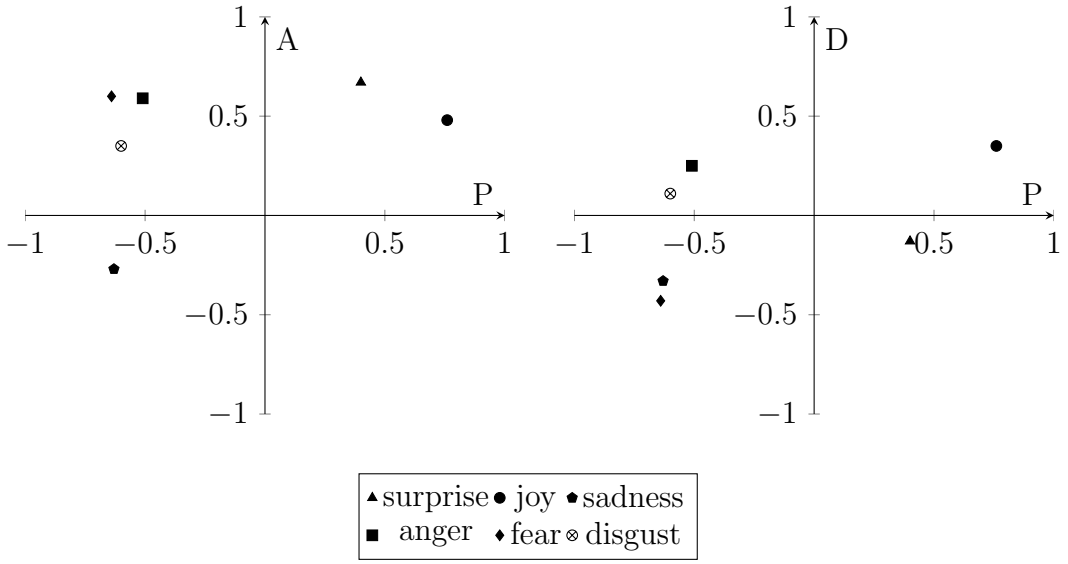


Figure 2.4: Location of the basic emotions in the PAD dimensions, based on [166].

FACS	Gesture	Description	Attributes	Function	Usage
AU1,2,4	raise	eyebrows go up or down	direction	P	when thinking
				I	show affirmation or not sure
	frown		amplitude	P	as punctuation mark
			velocity	I	distress, doubt
AU4	frown	wrinkling the brow	intensity	P	when thinking
				I	show dislike or displeasure
AU9	wrinkling	nose wrinkling		I	show dislike
AU12	smile	lip corners are pulled up	intensity	I	show liking
AU45,46	blinking	periodic or voluntary blink	velocity	A	wet the eye
			frequency	P	mark a pause
				I	emphasize speech
	winking	single-sided blink	side	I	convey a message or suggestion
M61,62, 63,64	avoidance	aversion of gaze	duration	R	when thinking
					when answering question
					hesitation about what to say
				I	while speaking
					turntaking in speech
	contact	seeking of gaze		R	pause in speech
					beginning of phrase
					asking questions
	lowered gaze	level of gaze falls		R	hesitation
				I	signal of cognitive load
	rising gaze	level of gaze rises		R	collect feedback
	saccade	rapid eye movement	direction	I	clarifies what is being said
			duration		often prior to head rotation
AD37	lip wetting	moisturizing lips	frequency	A	moisturizing lips
			velocity	P	when thinking

FACS: action unit (AU), Movement Descriptor (M), action descriptor (AD)
Function: Illustrator (I), Punctuator (P), Adaptor (A), Regulator (R)

Table 2.5: Facial gestures, adapted from Zoric *et al.* [131].

Idle Motions

Humans usually display a variety of small, involuntary movements in interaction, which serve no communicative purpose. These adaptors serve the satisfaction of bodily needs, for example shifting of the posture to avoid muscle fatigue, blinking to moisturize the eyes or scratching. While robots commonly have no such needs, the implementation of adaptors can break the static appearance of the robot. This is especially important during idle sequences between animated actions. These idle motions need to be designed in a way not to interfere with task related actions or need to be inhibited during the task. A set of idle motions is given in Table 2.5, the respective entries are marked as adaptors (A).

2.3.6 Fusion

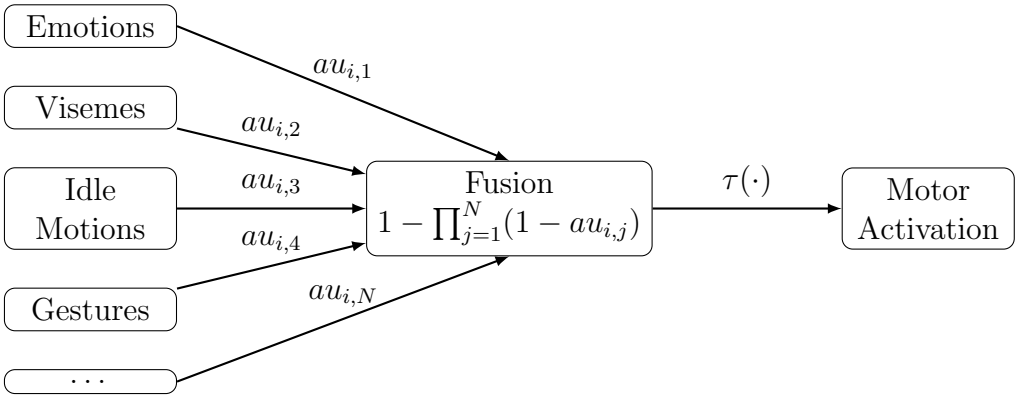


Figure 2.5: Fusion of action unit au_i accessed from several modules into an overall action unit activation.

The previously introduced components of expressions on the face manifest themselves in parallel and independent of each other. A problem arising from the parallel execution is that expressions can be overlapping in access of action units or motor activation. Thus, the contribution of several components of expressions to one specific action unit must be fused to avoid either information loss in a winner-takes-all strategy or violation of the boundaries in an additive process. The proposed fusion strategy here is to merge the contributing action unit activations $au_{i,j}$, with $i = 1, \ldots, 46$ and $j = 1, \ldots, N$, where N is the number of components accessing the action unit, see Figure 2.5, to an overall activation:

$$au_{merged\ i} = 1 - \prod_{j=1}^{N}(1 - au_{ij}), \qquad \forall\ i = 1, \ldots, 46 \tag{2.3}$$

Fusing the action unit activations with equation (2.3) results in a merged action unit activation $au_{merged\ i}$, which is bounded between $[0, 1]$. With increasing number of contributors N and $au_{ij} < 1$, the upper bound is approached asymptotically or reaches the maximum value for any $au_{ij} = 1$. Small contributions to the activation level are fused nearly additive, while activations close to the upper bound lead to a quick saturation.

2.3.7 Application

To show the application of the biomimetic design derivations and the methods of generating expressions, two robotic heads are developed that share nearly the same kinematics, but differ in the appearance. The first head is EDDIE, an "*Emotion Display with Dynamic Intuitive Interactions*" [51]. It is a mechanical looking open-frame face. The second head is the IURO head[4], built for the "*Interactive Urban Robot*" [211]. In this head the actuation and mechanisms are hidden under a cover with the shape of a stylized human face. A list of the implemented action units and the extra zoosemiotic features is given in Table 2.6. The listed action units are realizable without a flexible skin and are associated with the generation of facial expressions of basic emotions [167].

Since the configurations for EDDIE and IURO only differ in the zoosemiotic features, a common mapping τ can be given for both robot heads with respect to action units, see Table 2.7. Some of the action units specified in Table 2.3 can not be reproduced with the robot actuators, thus they are not listed. The actuator configuration, however, is different, requiring a separate motor command configuration for each head.

AU	Name	Facial muscles involved
1	Inner Brow Raiser	*Frontalis (pars medialis)*
2	Outer Brow Raiser	*Frontalis (pars lateralis)*
4	Brow Lowerer	*Corrugator supercilii*, *Depressor supercilii*
5	Upper Lid Raiser	*Levator palpebrae superioris*
7	Lid Tightener	*Orbicularis oculi (pars palpebralis)*
12	Lip Corner Puller	*Zygomaticus major*
15	Lip Corner Depressor	*Depressor anguli oris*
26	Jaw Drop	*Masseter*, relaxed *Temporalis* and internal *pterygoid*
43	Eyes Closed	Relaxation of *Levator palpebrae superioris*, *Orbicularis oculi (pars palpebralis)*
45	Blink	Relaxation of *Levator palpebrae superioris*, *Orbicularis oculi (pars palpebralis)*
46	Wink	Relaxation of *Levator palpebrae superioris*, *Orbicularis oculi (pars palpebralis)*
Z1	Crest Raiser	
Z2	Ear Turn	
Z3	Upper Ear	
Z4	Lower Ear	

Table 2.6: List of implemented action units and zoosemiotic features Z1-Z4.

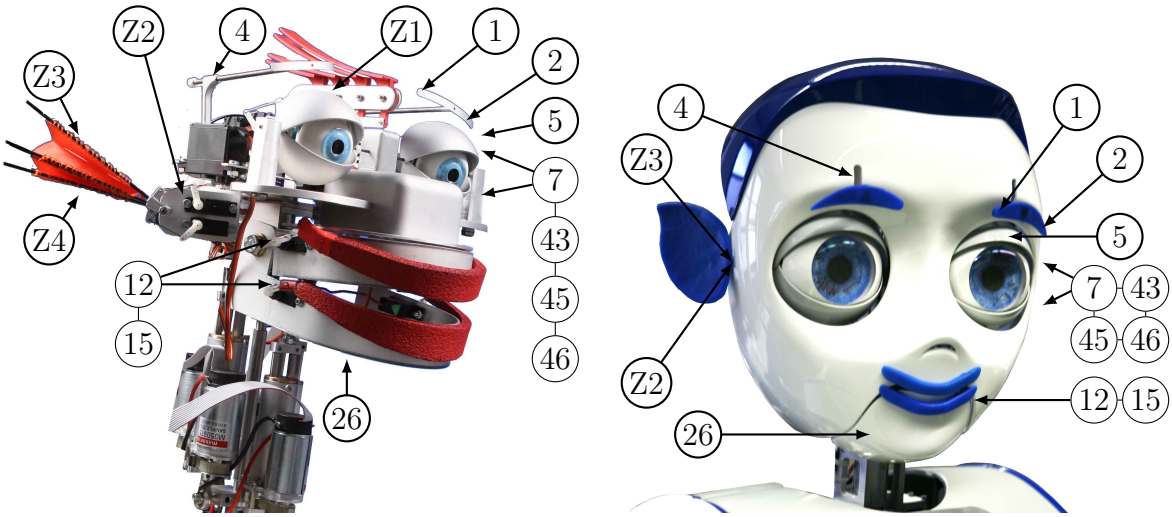
EDDIE

EDDIE is a mechatronic looking, open-frame robot head. Actuation of facial components is achieved with commercial miniature servos. Cameras are integrated into the eyes, forming

[4]Design by ACCREA Engineering

Actuator	Action Unit								
	1	2	4	5	7	12	15	26	43
Upper eyelids	0	0	0	1	-1	0	0	0	-1
Lower eyelids	0	0	0	-1	1	0	0	0	1
Upper lips	0	0	0	0	0	1	-1	0	0
Lower lips	0	0	0	0	0	1	-1	0	0
Eyebrow turn	1	-1	-1	0	0	0	0	0	0
Eyebrow raise	1	1	-1	0	0	0	0	0	0
Jaw	0	0	0	0	0	0	0	1	0

Table 2.7: action units to motor activation mapping τ.



(a) Robotic head EDDIE.

(b) Robotic head IURO.

Figure 2.6: Appearance of the robotic heads. The numbers reference the implemented action units, see Table 2.6 for the assignment.

a stereo pair. The most distinctive parts of the head are the crest, that can be erected and the ears. The ears are able to fold or unfold a membrane that is stretched across four beams. Two actuators per ear allow the independent movement of the upper and lower part of the membrane, a third actuator turns the folding mechanism. Figure 2.6a shows the robotic face EDDIE with the assignment of mechatronic actuators to the respective action units and additional zoosemiotic features.

IURO

IURO is based on the mechatronics and kinematics of EDDIE, but lacks the crest and membrane mechanisms. Instead of the folding mechanism, a rigid ear shape can be turned in two degrees of freedom (the lower ear degree of freedom is neglected). A cover encases the mechanisms and actuators with the shape of a stylized human head. Figure 2.6b shows the robotic face IURO with the action units assignment and zoosemiotic features.

2.3.8 Experimental Evaluation

The aforementioned methods and algorithms enable the synthesis of nonverbal communication with facial expressions in a HRI setup. However, in HRI the most important metric is if users can interpret the expressions in the same way the synthesis encodes the information that is intended to be conveyed. Therefore, the proposed methods need to be evaluated in setups in which users can rate or categorize the synthesized expressions on the robot heads.

Besides multiple-choice questionnaires specifically tailored for the experiment, the SAM test [53] is used in these evaluations. It is a non-verbal assessment technique using pictograms to visualize states in the tested dimensions and can be seen as a visual analogon to the semantic differential. Figure A.1 shows the test, with each row of pictograms depicting one PAD dimension. The pleasure dimension is indicated with a transition from smiling to an unhappy expression, the arousal dimension features are from sleepy to excited and the dominance dimension is characterized by a small pictogram, which denotes the feeling of being controlled and a powerful, large pictogram. Rating scores highly correlate between the SAM test and the semantic differential [28]. Therefore, it can be used to rate observations in the semantic differential-based PAD framework.

Basic Emotion Recognition

Hypothesis 2.1. *The model of displaying emotions is accurate enough, so that untrained participants can distinguish and categorize the basic emotions*

Hypothesis 2.2. *Displayed facial expressions, generated by the proposed expression synthesis, are rated correctly in the PAD dimensions to convey the intended emotion.*

The recognition rate of synthesized emotional expressions was evaluated in three experiments with the EDDIE head.

Experiment 1: In a first study [51] the six basic emotions were displayed to 24 participants. The composition of the group of participants was 8 children of ages 5 to 8 and 16 adults between 25 and 48 years. 7 of them are female and 17 are male. The robot was shown to the adult participants in groups of three, while the children had one by one sessions with the robot to avoid disturbances or influence between children. The robot would display one of the basic emotions in a randomized order and remain static for the evaluation of the expression. The task was to match the shown emotional expression to 10 given answers in a multiple-choice questionnaire. After the answer was given by the participant, the robot would show the next basic emotion without transition. Recognition rates of the basic emotions are given in Table 2.8 in the column "Recognition Rate Exp1".

Experiment 2: In a second experiment [439], EDDIE was shown to 20 participants displaying the basic emotions on video. The 20 participants are staff members of the Institute of Automatic Control Engineering (LSR), 6 of them are female and 14 male. The randomized videos showed a frontal recording of the head, with the head being nearly frame filling. In the video, the six basic emotions are shown with a change from a neutral facial expression to the respective emotional expression and back to neutral. Participants were asked to rate the emotional expression on a 5 point SAM questionnaire to capture

the perceived PAD state and also provide a free naming of the shown emotion. Results for the free naming, which was semantically grouped to match the basic emotions, is given in Table 2.8 in the column "Recognition Rate Exp2". The mean SAM ratings and standard deviations for the emotions joy, surprise and sadness are given in Table 2.9 and visualized in Figure 2.7. The figure shows the ground truth dimensional coordinates of the basic emotions, as given in Section 2.3.4, and the experimentally derived mean PAD coordinates of the perceived emotional expressions. The range of values of the SAM test is scaled to match the previously established range of values of the PAD dimensions from -1 to 1.

Emotion	**Rec. Rate Exp1** [%]	**Rec. Rate Exp2** [%]	**Ref. Germany** [%]
Joy	58	75	93
Sadness	58	90	83
Anger	54	65	71
Surprise	75	90	87
Disgust	58	20	61
Fear	42	85	86

Table 2.8: Recognition rates of basic emotions for EDDIE in experiment 1 [51] , experiment 2 [439] and the reference values for the recognition rates of human expressions in Germany [441]. Highlights show best recognition rate.

Emotion	**Pleasure**	**Arousal**	**Dominance**
Joy	0.55 (0.19)	-0.15 (0.2)	-0.05 (0.21)
Sadness	-0.9 (0.07)	-0.2 (0.24)	0.6 (0.2)
Surprise	0.7 (0.18)	0.65 (0.2)	-0.15 (0.19)

Table 2.9: Estimated mean values of PAD with standard deviations (in brackets) for three basic emotions shown with EDDIE [192]

Experiment 3: In a third study [16, 55], as a preparation for determining the influence of the zoosemiotic features of EDDIE, the robot head was tested with a reduced set of actuators. EDDIE was stripped of all zoosemiotic features (in particular: crest and ear mechanisms), leaving only those actuators relevant to FACS mounted on the head. The robot was then shown to 30 participants, displaying the six basic emotions to each participant separately and in random order. The group of participants consisted of 15 female and 15 male students and researchers (Ludwigs-Maximilian-Universität München) with an age-mean of 25 years. For the evaluation of the facial expressions, a German translation of the semantic differential was used. The results are given in Figure 2.8, showing the mean PAD ratings for each basic emotion and the respective ground truths.

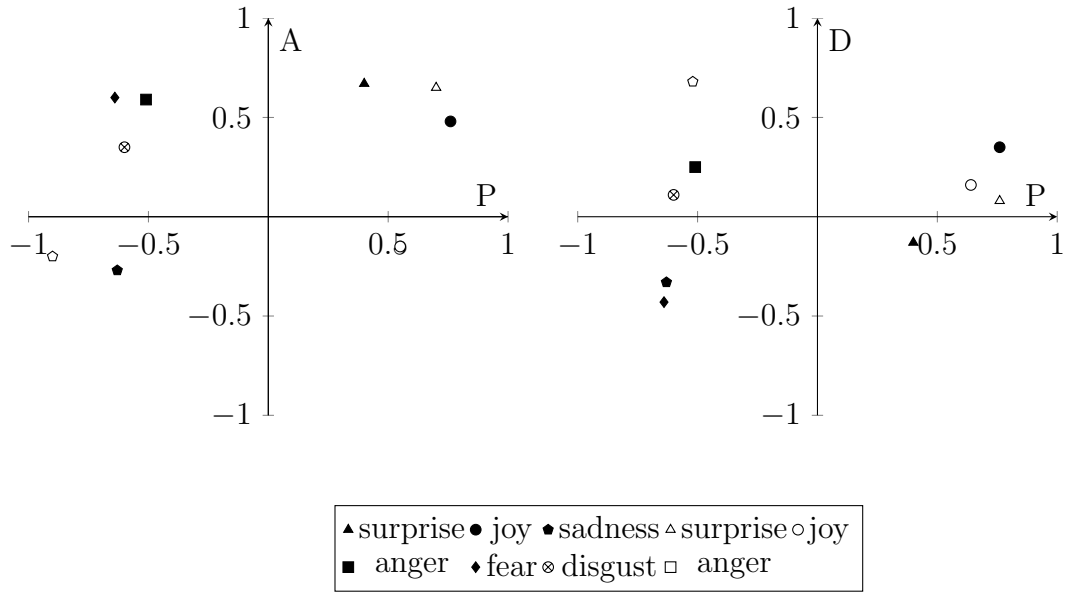


Figure 2.7: SAM ratings (hollow marks) and ground truth (filled marks) of joy, surprise and sadness for the emotion recognition experiment 2 in the PAD dimensions.

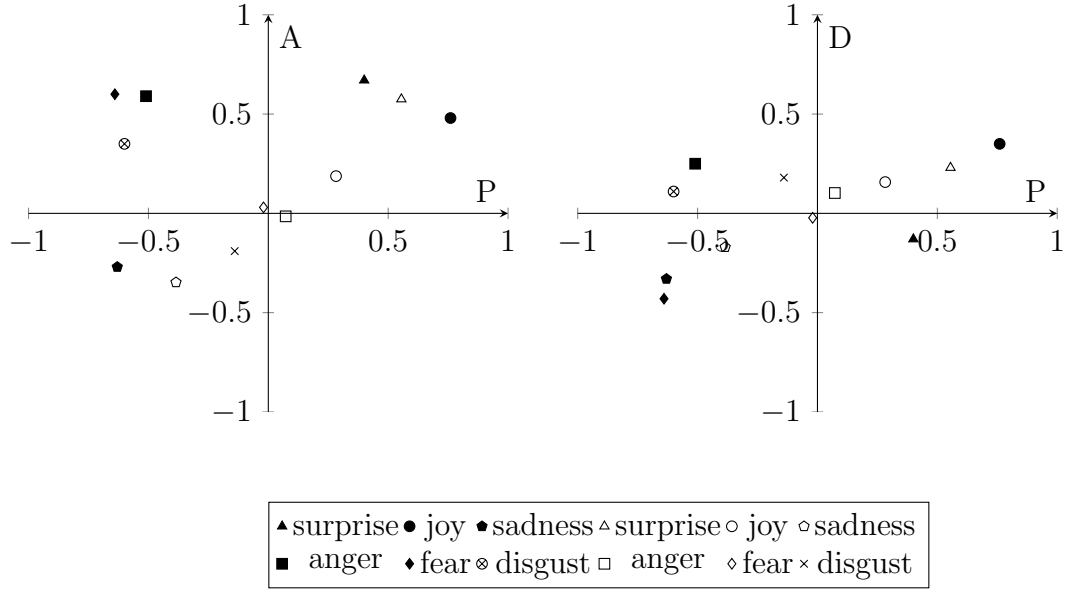


Figure 2.8: SAM ratings (hollow marks) and ground truth (filled marks) of the basic emotions for the emotion recognition experiment 3 in the PAD dimensions.

Discussion

The recognition rates of the basic emotions for the first study are in total of average quality, but significantly above chance level (16.$\bar{6}$%). The recognition rates for experiment 2 are in general higher, except for the display of disgust. The quality of the recognition rates needs to be seen in comparison to the recognition rates of human emotional facial expressions, which serve as a reference for the achievable quality of expressions. A study by Ekman *et al.* [441], performed with participants in several countries and cultures, gives a measure of recognition rates of human facial expressions of the basic emotions in Germany. 67 participants were shown three pictures of each of the six basic emotions, depicting only the face. The photographs were projected in random order and participants were asked to label the expression with one of the following terms: anger, disgust, fear, happiness, sadness and surprise, plus contempt. The methodology of the experiment makes it comparable to the studies performed with the expressive robot. The result for Germany in Table 2.8 bears qualitative similarity with the results from experiment 2. A notable similarity is in the drop of recognition rates of anger and disgust, compared to the other expressions, which can also be seen in the results from experiment 2 and, although not as distinctive, in experiment 1. It should be also noted that these reference values show that a 100% recognition rate is neither realistic nor necessary. The bottom line of recognition rates among the groups in Ekman's experiment for instance is: happiness: 69%, surprise: 78%, sadness: 76%, fear: 65%. disgust: 60% and anger: 67% [441].

The results confirm hypothesis 2.1, but there is room for improvement. Reasons for the loss of recognition rate of the robot could be found in the mechatronic structure. First of all, the head features only a reduction of implemented AUs versus the full FACS specification. This means, that not all AUs that are part of an expression are integrated and thus can contribute to the expression. An example for this is AU 9, the nose wrinkler. It is a very important AU for the display of disgust and sets this emotional expression clearly apart from other basic expressions. Therefore, only the second characteristic - the asymmetry - indicates the expression of disgust. This could explain the recognition rates of disgust. Another issue in this direction is the missing skin and thus the skin deformation. However, the skinless appearance was chosen based on design considerations of the uncanny valley [52, 51].

A remaining question is why the recognition rates between experiment 1 and 2 are so different. An explanation could be in the way of presentation of the robot. While in the first experiment the robot was immediately switched from one expression to the next, remaining static for the duration of the judging, in the second experiment the robot started from neutral, then showed the expression and was then set back to neutral. The better recognizability could result from the extra temporal dynamics of the facial expressions. Pantic *et al.* [142] divide the dynamics of emotional facial expressions into onset, apex and offset. In their context, emotional expressions are part of rapid facial signals, not static. Also, changes in the face could be more observable, emphasizing the contrast to the neutral expression.

A more detailed view on the shortcomings of the system to generate recognizable expressions is provided by experiment 2 and 3, changing the evaluation method from the categorical to the dimensional approach. This enables the decomposition of the problem

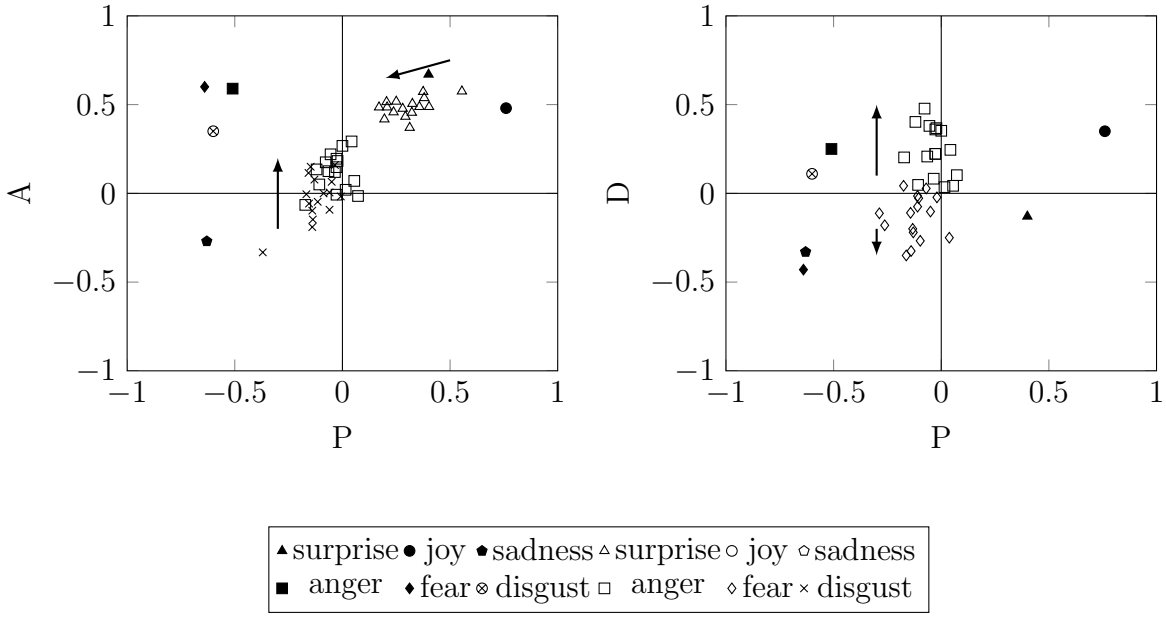


Figure 2.9: SAM ratings (hollow marks) and ground truth (filled marks) of the basic emotions, with variations in the manifestation of the zoosemiotic elements. The arrows indicate the shift in ratings for increased angles of ears and crown.

of misidentification for each specific emotion in the respective dimensions pleasure, arousal and dominance. Analyzing the PAD spread of experiment 2 in Figure 2.7, the measurements of the three prototypical emotions joy, surprise and sadness are in at least two dimensions close to the ground truth values. The arousal level of joy is only 57% of the ground truth and the dominance level of sadness is 238% of the ground truth. As can be seen from the categorization results in Table 2.9, this does not negatively influence recognition rates, as for both emotions the high pleasure/displeasure is the relevant categorization criterion. In comparison, the location of the measurement points in Figure 2.8 shows a shift towards the neutral center. This could be explained with the different scales used. Experiment 2 uses a 5 point scale, whereas in experiment 3 a 9 point scale is used. With the change from the relatively coarse 5 point scale to a finer granularity in the test, more differentiated ratings are expected, leading to less extreme overall ratings. While the overall tendency is consistent with the results from experiment 2, the high dominance of sadness could not be replicated. In experiment 3, the dominance of sadness is within the negative scale. All emotions have a shift along the pleasure dimension. A possible explanation is given by Kühnlenz [55], stating that, due to their design, the lips of EDDIE are perceived as smiling, regardless of the displayed emotional state. Hypothesis 2.2 is only partially confirmed. The displayed facial expressions are rated along the correct dimensions and are (with few exceptions) located in the correct octant of the PAD state space. However, the magnitude of emotions in the respective dimensions is not pronounced enough and the positive shift is apparent. These disturbances seem to be hardware related and not accountable to the underlying model.

Zoosemiotic Features

Hypothesis 2.3. *Zoosemiotic features aid the recognition of basic emotions*

Hypothesis 2.4. *Based on the zoosemiotic meaning of the respective signals in animals, the crown and ears influence the arousal dimension significantly*

Hypothesis 2.5. *Based on the zoosemiotic meaning of the respective signals in animals, the crown and ears influence the dominance dimension significantly*

A follow up study on experiment 3 was conducted with 30 participants to determine the influence of the zoosemiotic features on the perception of the basic emotion expressions [16, 55]. Of the 30 participants , 15 are male and 15 female. The design of the study is a 2×2 analysis of variance (ANOVA) with repeated measures, with the crest as the first factor and the ears as the second. For each factor, four conditions are defined: fully extended, semi extended, retracted and dismantled. The combination of both ears and crest dismantled resembles the conditions of the previously described experiment 3.

Each participant was asked to rate 32 combinations on the semantic differential scale. The 32 combinations are one third of all 96 possible combinations, resulting from the concatenation of the 2 factors with 4 conditions each, together with 6 basic emotions. The relationship between pairs of groups is equal, as the Mauchly test of sphericity is non-significant for all data (> 0.1). Changes to the data set were made through the linear interpolation of missing data and the exclusion of data resulting from incorrect answering behavior. As pointed out by Kühnlenz [16], this leads to the impracticability of calculating the F-test for joy and surprise in the dominance dimension and for fear and sadness in both dimensions arousal and dominance. Significant results of the described ANOVA are given in Table 2.10 and visualized in Figure 2.10 and Figure 2.11.

Emotion	Dimension	Factor	F-value	p-value
joy, surprise	P	crest	F(3,12)=4.013	0.034
joy, surprise	P	crest · ears	F(9,36)=3.631	0.003
joy, surprise	A	crest · ears	F(9,54)=3.258	0.003
fear, sadness	P	crest · ears	F(9,18)=5.843	0.001
anger, disgust	P	ears	F(3,6)=4.835	0.048
anger, disgust	P	crest · ears	F(9,18)=4.132	0.005
anger, disgust	A	ears	F(3,6)=67.582	0.000
anger, disgust	A	crest	F(3,6)=11.987	0.006
anger, disgust	D	ears	F(3,62)=46.724	0.000
anger, disgust	D	crest · ears	F(9,18)=9.463	0.000

Dimension: Pleasure Arousal Dominance (PAD)

Table 2.10: Significant results of the 2×2 ANOVA with repeated measures on the influence of additional zoosemiotic features [16].

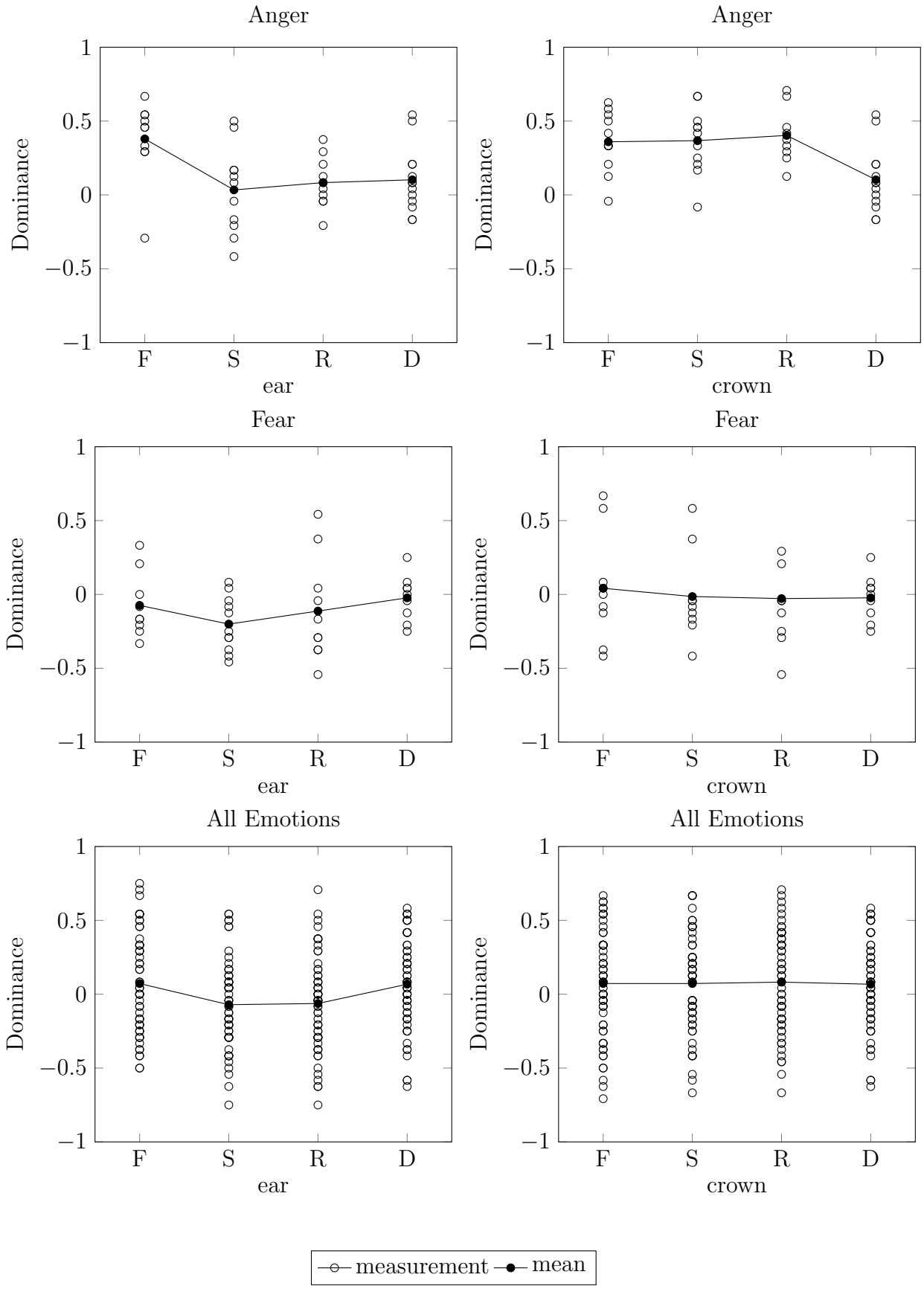


Figure 2.10: Shift in dominance means of anger, fear and overall emotion ratings for the following ear/crown conditions: F: fully extended, S: semi extended, R: retracted, D: dismantled

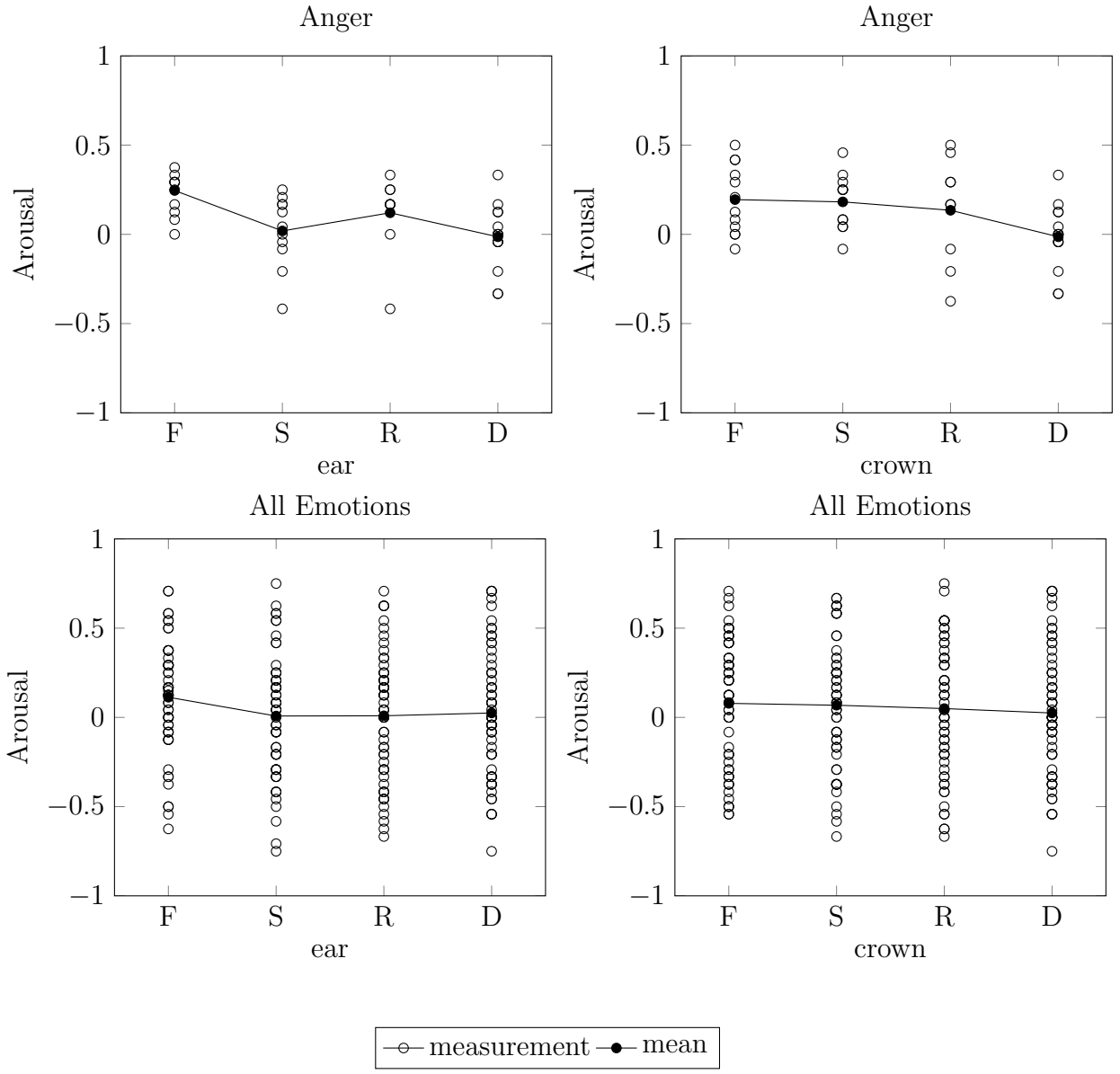


Figure 2.11: Shift in arousal means of anger and overall emotion ratings for the following ear/crown conditions: F: fully extended, S: semi extended, R: retracted, D: dismantled

Discussion

With the key question for this experiment being whether additional actuated facial elements, that are not part of human anatomy, have an influence on the perception of emotional facial expressions or not, it can be stated that there is significant impact [103]. In the case of this experiment, for the varying factors (ears, crest, ears·crest), a significant influence ($p < 0.05$) for the emotional states joy, surprise, anger and disgust could be shown. In a more detailed inspection, it is apparent that the influence of the zoosemiotic features is not of the same kind for each emotional expression and zoosemiotic feature activation. For the different emotional expressions, the activation of zoosemiotic features resulted in different shifts in the dimensions of the perceived emotional state within the PAD state space. Anger and disgust are significantly influenced on all affective dimensions, whereas joy and surprise are affected in pleasure and arousal and fear and sadness only in pleasure. This supports hypothesis 2.3, best seen in the shift towards ground truth for surprise. Also anger shows a shift in the critical dimensions arousal and dominance towards or even above ground truth levels. This is in accordance with the zoosemiotic signaling of the ears and crown as described in Section 2.3.3. In support of hypothesis 2.4 and 2.5, the effect on dominance can be best shown with anger and fear. They are roughly on the same pleasure and arousal level and mostly differ by dominance. The effect of the crown for fear on the mean rating with different positions is about the same as without crown. For anger, the mere presence of the crown has a significant result in dominance enhancement, while in average over all emotions there is no influence on the dominance perception. For the ears, there is a trend that retracted and semi-retracted ears lower the perceived dominance. The lower values for the semi-retracted in comparison to the retracted condition could be explained by the fact that the ears are partially hidden by the head structure in the retracted condition. It is therefore visually closer to the dismantled than to the semi-retracted condition. The fully extended condition however raises dominance significantly for anger.

2.3.9 Summary

Starting from studies on HHI, an expression synthesis framework was introduced for utilizing non-verbal communication in HRI. In this top-down approach, the analysis of the anatomic structure of the face leads to a technical kinematic structure based on the Facial Action Coding System. For the biomimetic design process, it should be noted that the proposed system is only a weak abstraction of the biological design, since appearance and functionality are tightly coupled in this case. FACS not only proves to be helpful for the replication of facial elements, but is also used for unifying the display of emblems, illustrators, affective display, regulators and adaptors, see Section 2.2.1. For the unification, the various non-verbal signals, that manifest themselves potentially simultaneously on the face, are merged on the action unit level. The proposed method takes into account the respective influence of various signals on one AU, while maintaining the activation limits. Definitions for the mappings involved in the generation of an emotional facial expression from the PAD state are given. For these mappings, the merging of the categorical and dimensional emotion representations allows to transfer Ekman's research on emotional facial

expressions to the PAD dimensions. This enables the use of the extensive corpus of facial expression analysis for the basic emotions within the PAD state space, which is proven to be universally understandable and independent of a specific technical implementation. The concept of motor activations and commands generalizes and eases the transfer of the action unit based expression synthesis between technical systems, decoupling the general AU computation from the kinematic structure of the technical system with the motor activations and the actual actuator implementation reacting on the motor commands. In the course of this section, possible ways of incorporating not only human facial elements, but also zoosemiotic elements, were investigated. The zoosemiotic elements are categorized based on their signaling effect in the PAD dimensions, thus being able to integrate them with the dimensional approach of emotion representation.

In order to show the applicability of the introduced methods, two robotic heads were developed and driven by the expression synthesis framework. Experimental evaluations show that the proposed expression synthesis is able to produce emotional facial expressions that can be categorized according to the basic emotions and that are qualitatively rated along the correct PAD dimensions. The introduced animalistic facial elements change the perception of some of the basic emotions significantly. They alter the dimensional rating for anger and surprise along the dominance and arousal, according to the zoosemiotic signals that are the source of inspiration.

2.4 Neck Posture Synthesis

Additionally to expressions of the face covered in the previous section, the neck aids the overall expressiveness of the head. The main task of the neck is to bring the head into a specific posture and orientation. The orientation and posture are chosen in a way that sensory input (for example for visual, auditory or olfactory stimuli) or directed actions (for example manipulations with lips and teeth or directed speech) are enabled or enhanced. To meet the performance requirements, the neck is highly agile and possesses several degrees of freedom (DoF), which are redundant if only the orientation towards an object is accounted for. Moreover, besides mere orientation, the motions of the neck and the posture convey signals, that are interpreted by interaction partners. These signals can have a standalone effect functioning as illustrators, regulators, adaptors or emblems, or be complementary to facial expressions, supporting the respective affect display.

In this section, a neck is modeled to incorporate the orientation and posture task, as well as the display of social signals. In the biomimetic design process, a simplified model of the human neck is derived from the biomechanical structure and brought into a form that is compatible with current formulations of robotic manipulators. Following the design process, possible gestures and postures are examined for their expressivity and realization with the aforementioned model. Based on the model and the gesture/posture formulations, a control strategy is described that can account for the fixation and tracking of objects, while maintaining the ability to display interaction signals.

2.4.1 Problem Statement

Most robot necks are designed with the minimal DoF necessary to accomplish the orientation task. This is sufficient if no extended workspace is needed and social interaction is of no concern. Artificial necks with redundant configurations extend the workspace of the head, but neglect to exploit the redundancy to convey additional information besides the current focus of attention. With the human neck in mind, redundancy in an artificial neck can be used to mimic the posture or movements in a more humanlike way, without interfering with the primary task of fixating a coordinate in space.

The proposed solution to define a suitable model of a humanlike head posture is to take a look at the biomechanical properties and kinematics of a human neck. From these properties, a simplified model of the kinematics, that is compatible to current rigid link robotic implementations, is to be generated. For this humanoid model, a set of model parameters is derived from experiments, such as the coupling between joints and a humanlike default posture. Furthermore, gestures and postures are assessed whether they inflict on the main task of fixating a point in space or not. A control strategy is described that is able to generate the inverse kinematics of the artificial neck while incorporating the gesture movements in the null space of the joint velocities.

2.4.2 Related Work

Numerical models of the human neck have been used in injury assessment and clinical diagnostics since the 1980's. In the two joint model [79, 107], the neck is approximated via a rigid link. The neck is connected by two pivot, universal or spherical joints to head and neck, depending on the respective model and whether the dimension of the simulation is 2-dimensional (2D) or 3-dimensional (3D). One such simulation is the SOM-LA/SOM-TA program [146], modeling human occupants in aircraft seats with a total of 29 DoF. The neck is modeled through a rigid link with two 3 DoF (spherical) joints. Despite the simplification, two joint models are suitable to describe the global motion of the head and neck in relation to the torso. However, they lack richness of detail in vertebral kinematics and cervical soft tissue deformation [61]. This is incorporated in multi-body or finite element models. They provide a more detailed analysis of not only the total head - neck complex, but the individual vertebrae, intervertebral discs, ligaments and muscles. See Dimnet [132] and more recently Chen *et al.* [61] for an overview on clinical neck models.

Similar categories can be found in robotics, with single joint neck models besides two joint and multi-body models. In robotics, the neck is a common part of the vision system for position or orientation of cameras. For cameras that are fixed on the neck, single joint model pan-tilt platforms (2 DoF) are the minimal solution for orienting the camera at any point in 3D space. Today, a wide variety of commercial platforms are available. With non-fixed camera systems, the integration of oculomotor control in the overall control framework for both neck and eyes leads to kinematically redundant systems. For the basic problem of solving the redundancy of a manipulator, an overview on standard techniques is given in [94]. A number of robotic systems are designed to solve this redundancy in a bio-inspired fashion. Kim *et al.* [76] reduce the degrees of freedom by coupling the motion of both eyes via vergence control. Two proportional-derivative (PD) controller in

a cascade, one for the eyes and one for the neck, center a target in the camera images. Breazeal *et al.* [32] decouple the eye and neck movements. The 2 DoF neck positions a head-fixed wide angle camera. The 2 DoF eyes with narrow-angle cameras, simulating the *fovea centralis*, are then positioned within wide angle picture. An overview on multi-focal systems and their control is given in [57]. The redundancy in the neck configuration itself is introduced with roll as a third DoF, so that the single joint models form a spherical joint. The additional axis is integrated to closer resemble the functionality of the upper neck joint, for example in iCub [89] in combination with 3 DoF eyes or ROMAN [117]. Their control strategy to resolve the added redundancy is to keep the head horizontally to aid vision tasks. Other control approaches include the third DoF in the overall posture generation, e.g. the overt visual attention system [148]. Single joint models are popular due to their simplicity, as it is the minimal solution for tracking objects in 3D, and in the ideal case without restrictions on the task. Two joint models in robotics are less common, but are often employed to extend the motion range of the head. Examples are the ISHA robot [37] with a 3 DoF neck and 4 DoF eyes, the Twente humanoid head [147, 161] with a 4 DoF neck and 3 DoF eyes or the Karlsruhe head with a 4 DoF neck and 3 DoF eyes [33]. The control frameworks are similar to the redundant systems with single joint models and extra DoF in the eyes, but have to account for the postural shift of the head. Both single joint and two joint models have the advantage that they resemble rigid structures, which is the dominant kinematic structure in robotics today. The stiffness of the links and possible high precision joint angle measurements benefit the primary vision task. The structure itself, however, is only an abstraction and approximation of the real kinematic structure of the neck.

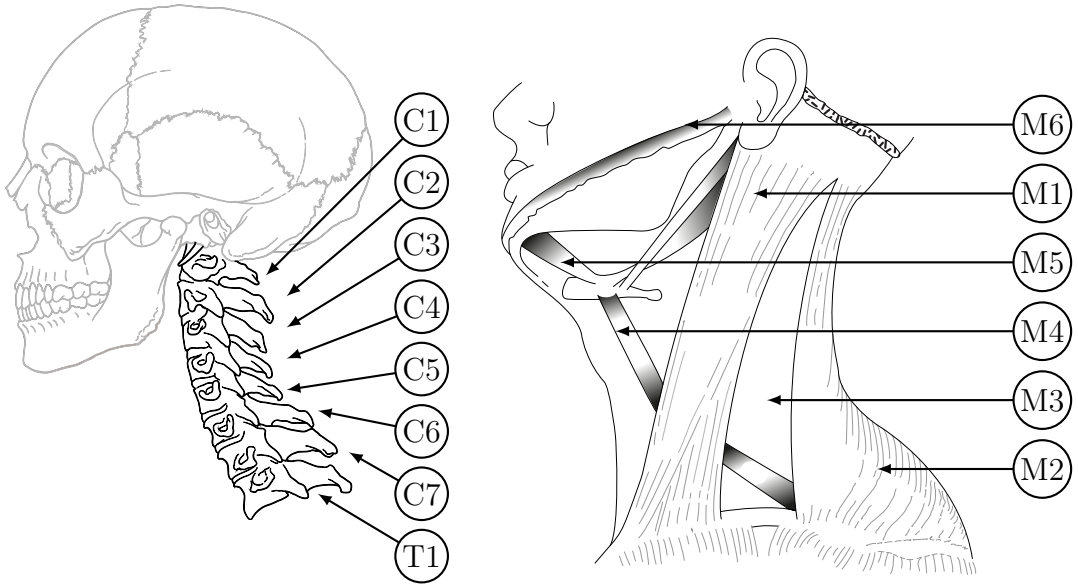
Similar to the efforts in rebuilding the human face in androids, some robotic projects focus on copying the biological design. This results in neck models similar to the clinical multi-body models. They are either built by forming a kinematic chain of several distinct segments, which are linked together, or deflecting a spring. Roos *et al.* [155] developed a three stage humanoid spine for the robot Roberta, with the links formed by cardan joints. A more advanced humanoid spine was integrated in the robot Kenshiro [26] (formerly: Kojiro). Izawa *et al.* [22] showed a neck that implements variable stiffness. Coil springs as neck basis are used by Noori *et al.* [140] in a 3DoF neck, with antagonistic tendons deflect the spring. The inverse kinematics of the neck can be solved with a model of the spring and resolved rate motion control. Another coil spring neck was developed by Hashimoto *et al.* [21], with McKibben pneumatic actuators as artificial muscles. So far, robotic necks treat all vertebraes as identical, neglecting the differences in range and distribution of motion. The advantage of close resemblance of the neck models to the biological original leads also to a number of challenges in robotics. The overall structure is more complex than the single and two joint models. Most control strategies include solutions for the forward, but rarely solutions for the inverse kinematics are given. Also the control frameworks have to compensate the elasticities introduced by intervertebral discs and tendons or pneumatic muscles. Position deviations due to unattended flexibilities lead to errors in the camera frame to world mapping, which is important for precise object localization.

Few systems and approaches explicitly incorporate non-verbal communication signals.

Breazeal integrated neck movements accompanying emotional expressions in Kismet on a single joint neck [32]. Both Reilink *et al.* [161] and Omrčen *et al.* [33] claim that gestures or implicit communication signals could be performed with their approach, utilizing the extra DoF in their respective 4 DoF necks. However, they neglect to show the integration of implicit communication signals in their methods. Takanishi *et al.* [170] described a method to achieve pursuing motion in depth with a 4 DoF neck. In the Waseda WE-4RII [62], additional DoF in the neck are used for enhancement of emotional expressions. However, the posture generation for each expression is done by hand, based on an artist's conception of the respective emotion. Studies on the social impact of neck gestures have been performed by Kose-Bagci *et al.* and Lee *et al.*. Kose-Bagci *et al.* [136] showed that simple neck gestures (nodding, side to side) in a drumming game with children results in higher enjoyment and influence on perceived intelligence of the KASPAR robot. Lee *et al.* [39] found in their study with the Mobile Dextrous and Social robot, that neck movements in addition to gestures elicit a higher number of words by participants and more intimacy in interaction, thus stimulating more engagement.

2.4.3 Biomimetic design

In this biomimetic design process, the goal is to transfer the functionality of the human neck to a technical system. While the biomechanical structure of the neck is quite complex – compared to current robotic implementations–, an abstraction of the kinematics leads to a simplified model, which can approximate the movements and postures that are relevant for non-verbal communication.



(a) Cervical vertebrae **(b)** Cervical musculature

Figure 2.12: Cervical vertebrae and simplified scheme of musculature of the human neck [72].

Bone Structure and Muscular System

The neck is part of the human spine, connecting the head with the torso in a flexible way. The neck of an adult is typically 11 to 14 cm long. It is referred to as the *cervical spine*, consisting of seven *cervical vertebrae*, which are often labeled from C1 to C7, see Figure 2.12a. The vertebrae C1 to C6 form a convex curvature, whereas C7 is already part of the concave curvature of the *thoracic spine*. Based on its structure, the cervical spine can be viewed as two-part. The upper part consists of C1 (*Atlas*) and C2 (*Axis*), with the *Os occipitale* - the docking point of the skull - referred to as C0. The connection between C0 and C1, the *atlantooccipital* joint, allows flexion of the head and leaning sideways. The *atlantoaxial* joint, connecting C1 and C2, enables turning motions and to some extent flexion and sideways motions. The lower part of the cervical spine, formed by C3 to C6, consists of vertebrae with simpler and more homogeneous shapes than Atlas and Axis. Every joint formed between vertebrae next to each other allows flexion, sideways leaning and rotation, assisted through an intervertebral disc between bones.

The positioning and orientation of the head attached to these vertebrae is done with a complex set of musculature, see Figure 2.12b for a simplified sketch. The overall setup of the cervical musculature is in an antagonistic way, but due to the kinematics and the possibility to activate larger muscle groups only one-sided, muscles can serve several functions [104]. The *Sternocleidomastoidei* (M1), for example, shift the head forward. The activation of a single *Sternocleidomastoideus* together with *Trapezius* (M2) also turns the head. This turning leads to simultaneous sideways leaning, if not counteracted by *Splenius capitis*, *Longissimus capitis* and *Semispinalis capitis* (M3). However, *Splenius capitis* is involved in turning the head and leaning sideways, while *Semispinalis capitis* streches the neck. The antagonists for the *Sternocleidomastoidei*, *Splenius* and *Levator scapulae*, shift the head backwards. But if the shoulder is in front, *Levator scapulae* leans the head sideways. Flexion is achieved by the *infra-* (M4) and *suprahyoid* (M5) muscles, which are attached to the tongue bone (*Os hyoideum*). They flex the head if the jaw muscles (M6) block the opening of the jaw.

The multifunctionality of the muscles together with the vertebrae, forming a kinematic serial chain with high degrees of freedom, require a complex interaction of the neck elements to orient and position the head. With the aim of transferring the functionality to a technical system, a simplification of the kinematic structure with the same abilities in positioning and orienting a head-like object with a comparable workspace is desirable.

The workspace of an adult neck is limited by the bone structure of the vertebrae and the elongation properties of the respective muscles and tendons. It also varies between subjects and is dependent on age. If only the active agility is considered [104], which describes the ranges achievable through muscle activity, the flexion and elongation is up to 125°. The atlantooccipital and the atlantoaxial joints (C0 to C1 to C2) contribute up to 30°, while the lower part of the neck (C4 to C7) flexes between 90° and 100°. The possible sideways leaning is up to 45° per side. The turning of the head can be up to 90° per side, of which 25° to 30° are contributed by the lower neck part.

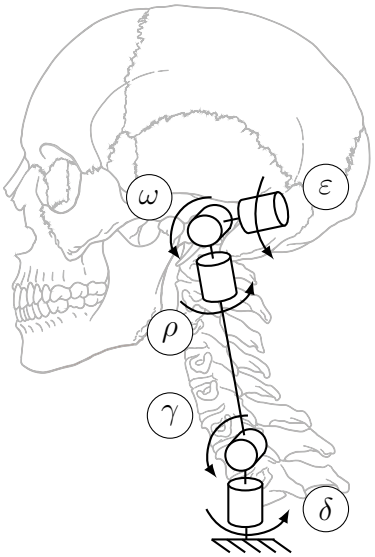


Figure 2.13: Simplified model of the human neck joint configuration.

Model

The human neck has a total of 20 joints. Most of them are coupled due to the layout of the muscles and their interaction with the vertebrae in a kinematic serial chain. Thus, the positioning and orientation of the head can be approximated with a simpler model. If considering the two-part structure of the neck, with the upper joint formed by C0 to C2 and the lower joint formed by C4 to C7, the neck can be treated as a regional structure [94], with the decoupling of positioning and orientation. The upper joint, which can be approximated by a spherical joint or three revolute joints with concurrent axes and angles ρ, ω and ε, performs the orientation of the head. The lower joint, approximated by an universal joint or two revolute joints with orthogonal axes and angles γ and δ, positions the head. With a fixed length between the upper and lower joints, the reachable positions form a sphere. With these approximations, the reduced model has only 5 DoF. In the decoupled regional structure, the 3rd DoF in the lower joint, formed by the third DoF of the vertebrae C4 to C7, has no influence on the positioning of the head and the orientation is covered by the upper joint. If no joint limits exist, it can be neglected without reducing the achievable turning of the head. Therefore, the simplified model can approximate all movements of the full neck. With the multifunctionality of several of the neck muscles, the coupling of joints needs to be modeled as well. In an experimental verification and quantification, which is described in Section A.6, the following relation of joint angles could be derived:

$$\gamma = b\omega + c + \Delta\gamma \tag{2.4}$$

$$\Delta\gamma = a|\rho| \tag{2.5}$$

$$\delta = d\rho \tag{2.6}$$

$$\varepsilon = e\rho \tag{2.7}$$

with $a = 0.068$, $b = 0.384$, $c = 0.281$, $d = 0.888$ and $e = -0.112$.

Equation (2.4) gives a relationship between the flexion of the upper and lower joint. With $b < 1$, most of the flexion is done in the atlantooccipital joint. Parameter c is a general offset in γ, that takes into account the natural deflection of the neck due to its geometry. Equation (2.6) describes the coupling of the upper and lower vertebrae while turning the head. The coefficient $d < 1$ means more turning is done in the atlantoaxial joint to reach a certain orientation. The last equation (2.7) relates the sideways roll of the head to the turning. This is not obvious at first, but the double role of *Sternocleidomastoideus* and *Splenius capitis* need to be taken into account. Both muscles turn the head and lean it sideways. The more the head is turned, the less counteraction can be performed by the antagonistic muscles. This leads to a roll motion of the head. The found relationships are in good agreement with the expected relations due to the musculature and bone layout.

2.4.4 Signals

In addition to the task of orienting and positioning, the head should be able to perform gestures and take expressive postures via the neck. These signals convey additional explicit and implicit information on the visual communication channel. This often includes gaze shifts and other gaze related behaviors, but the focus here is on the neck-induced movements. Eye movements are only considered as part of the oculo-vestibular reflex, compensating head motion while fixating a target.

Gestures

Just like the face, the head in combination with the neck can display implicit communication cues that fit the categories *adaptor*, *illustrator*, *regulator*, *emblem* and *affect display* (see Section 2.2.1).

Adaptors are in this case posture shifts of the head to avoid fatigue in the neck musculature. They occur especially for the relaxation of muscles strained by gravitational load on the head for off-center positions.

The second class of gesture types are illustrators during speech. These contain pointing gestures, such as pointing with the head towards an object, possibly accompanied by a gaze shift. They can also follow a head tilt backwards, indicating the noticing of something [23].

Gaze shifts and turning of the head are also used as regulators during conversations, signaling turn taking and information status. An example is given by Zoric *et al.*: "*The beginning of themes (already introduced utterance information) is frequently synchronized by a gaze-away from a listener, and the beginning of rhemes (new utterance information) is frequently synchronized by a gaze-toward a listener.*" Besides synchronization, involvement can be signalized via small one-way nods [144] or sudden cut off, in which the head fully turns away from the speaker [312].

The most prominent type of neck gestures are emblems such as head shake or nod. The head shake commonly is a gesture to show disapproval [126, 23], but could also be used during a question or to accompany laughter [126, 65]. The indication of agreement is signalized by the head nod [19, 127, 312, 23]. Further emblems are gestures of being taken aback, the need for closer look [285] or the head roll, which is expressing doubt through repeatedly tilting the head left and right [127].

It should be noted that the meaning of these gestures is based on studies conducted in the western civilization. Although the meaning of agreement is widely spread, head nods could also be negative [144]. Furthermore, gestures can vary between different cultures, such as the substitution of the head nod with the head wobble in Southern India [114], a cyclic tilting movement of the head. An overview on the relevant head gestures is given in Table 2.12.

Emotional Expressions (Affect Display)

Gaze and head posture are important for displaying emotions [27, 152]. Various studies have shown the share of head postures in basic emotions and the PAD dimensions (see Section 2.2.2).

For the basic emotions, the following postures add to the display of emotions besides the facial expressions: Wallbott [145] describes the basic emotions of *joy* with a raised head. In *fear*, the whole head is moved back, with the chin tucked in slightly [25]. *Sadness* is frequently associated with a bowed, hanging head [25, 19, 68]. For *anger*, the potential for aggression is displayed with a slight forward leaning [25]. The bowed head is also part of *disgust* [145]. The cited studies lack a description for the basic emotion of *surprise*. However, Hess *et al.* found that expressions of fear, surprise, and sadness are perceived as being submissive [69]. Therefore, it can be assumed that surprise also shows the submissive sign of a lowered head.

Besides the basic emotions by Ekman, a number of other categorical emotions have been described with the respective head postures: *Respect* is related with a bowed head [179], as well as *shame* [19, 137, 130, 145], *humiliation* [130] and *embarrassment* [320] or the special form of an *embarrassment smile* [285]. *Pride* is linked with both a horizontal [19] or a raised head posture [145]. Other associations with a raised head are *contempt* [137, 179] and *boredom* [145].

The categorical emotion descriptions are also considered for the dimensional emotion aspects. According to Tiedens [17] guilt, sadness, and gratitude are displays of submission, while pride and anger are related to dominance. Keltner *et al.* also see embarrassment as a display of submission [320]. Mignault *et al.* [49] show that "*... a bowed head connotes submission, inferiority emotions, sadness, joy in women, and an illusory smile. A raised head connotes dominance (especially in women), superiority emotions, happiness (i.e., joy and contentment), and an illusory downward contraction of the corners of the mouth.*"

What can be seen from these findings is that head movements are mainly a signal of the dominance dimension, whether the head is raised (dominant) or bowed (submission). Persons are rated more dominant by coders with an upright head than one "*slouching forwards*" [49]. For the other dimensions of arousal and pleasure, velocity of the movements is linked to arousal [101] and head [49] and body [66] posture can be a sign of pleasure. Kleinsmith *et al.* [128] specify the relationship with the PAD dimensions for the distinction between high and low manifestations of the respective dimension. The head is important for all three dimensions to distinguish between high and low forms and high and neutral forms (except arousal). Their guidelines for the display of PAD via the head and neck and the insights from the dimensional relations of the basic emotion are summarized in Table 2.11.

Dimension	Value	Posture	Attributes
P	high	raised	-
	low	lowered	-
A	high	forward	fast movements
	low	backward	slow movements
D	high	raised	facing interaction partner
	low	lowered	turning away

Dimension: Pleasure Arousal Dominance (PAD)

Table 2.11: Affect display via head posture

Joints	Gesture	Description	Attributes	Function	Usage
all	postural shift	overall shift of the posture	amplitude	R A	beginning of sentence strain relief
all	pointing	pointing towards an object or away from it	direction	I R	clarification of speech turn taking
all	approach avoidance	motion towards an object or away from it	direction	E E	need for closer look taken aback
ρ, δ	shake	oscillating movement around atlantoaxial joint	frequency	E I I	disapproval during a question accompanying laughter
ω, γ	nod	swing towards body and back around atlantooccipital joint	frequency	E P R	approval punctuation mark involvement
ε	roll	tilting of head repetitive tilting	amplitude frequency	E	expression of doubt

Function: Emblem (E) Illustrator (I), Punctuator (P), Adaptor (A), Regulator (R)

Table 2.12: Head gestures

In the next section, a control strategy for the neck model is described that incorporates both the biomechanic insights and non-verbal communication signals.

2.4.5 Control

The control task for the neck can be described as facing a specific point in cartesian space or looking at it. Here one assumption is that looking at an object from a specific orientation is neglected, such that the task can be described as

$$\boldsymbol{x} = [x, y, z]^T, \quad \text{with } \boldsymbol{x} \in \mathbb{R}^3. \tag{2.8}$$

The task description can also be given in the 2D image plane projection of the point in a camera system, reducing the task description to $\boldsymbol{x} \in \mathbb{R}^2$. Here the general case is considered, in which a head does not necessarily get the target information from a head mounted camera. Thus, the line of sight is represented by a virtual prismatic joint with joint variable σ, extending from either the center of the face along the surface normal or the eyes. The joint variables for the neck model described in Section 2.4.3 are then extended to

$$\boldsymbol{q} = [\varepsilon, \omega, \rho, \gamma, \delta, \sigma]^T, \quad \text{with } \boldsymbol{q} \in \mathbb{R}^6. \tag{2.9}$$

One can see that, with the DoF $n = 6$ being strictly larger than the dimension of the task description $m = 3$, the mapping from task to joint space $\boldsymbol{x} = f(\boldsymbol{q})$ needs to provide a way to deal with the kinematic redundancy of the neck. Therefore, a solution to the inverse kinematics problem for a redundant neck configuration needs to be determined. Additionally to the main task, the redundancy allows to specify the subtasks that the neck posture should appear humanlike and that gestures described in the previous section can be displayed along the main task.

Decoupled architecture A first way to solve the inverse kinematics is to consider the neck as a decoupled robotic architecture, which splits the neck movements into positioning and orientating. The position of the head with respect to the neck base is denoted by

$$\{\boldsymbol{s} = [s_x, s_y, s_z]^T \in \mathbb{R}^3 | (s_x - x_0)^2 + (s_y - y_0)^2 + (s_z - z_0)^2 = l^2\}, \tag{2.10}$$

where $\boldsymbol{s}$ is in neck base frame coordinates, which are given by $x_0 = y_0 = z_0 = 0$, and l is the length of the neck. It is a function of the lower neck joints $\boldsymbol{s} = f(\gamma, \delta)$. The orientation of the head is then described as a function of ε, ω and ρ. Using spherical coordinates with the origin $\boldsymbol{O}$ located at $\boldsymbol{s}$, the point to look at is then defined by $\boldsymbol{x} = f(\omega, \rho, \sigma)$, with the length of virtual σ link determined by the Euclidian norm $\sigma = \|\boldsymbol{x} - \boldsymbol{s}\|$. With neglecting the rotation of the line of sight, ε remains a degree of freedom, as $n = 4 > m = 3$. Introducing additional constraints for ε, such as $\varepsilon = 0$ or $\varepsilon = e\rho$ (see (2.7)), and $\sigma \geq 0$ leads to a unique solution for the inverse kinematics problem.

This approach provides the benefit that an analytic solution to the inverse kinematics problem can be found. While the main task is fulfilled, gestures can be included via the position $\boldsymbol{s}$ and velocity $\dot{\boldsymbol{s}}$. However, the resulting posture is solely dependent on the

relationship of $\boldsymbol{x}$ and $\boldsymbol{s}$, thus the subtask of a humanlike configuration, as in the way defined in Section 2.4.3, is neglected.

Coupled joints A second way to get an analytical solution for the joint configuration is to employ the joint coupling described in Section 2.4.3. Utilizing the equations (2.4) to (2.7) reduces the DoF to $\boldsymbol{x} = f(\omega, \rho, \sigma)$, so that $m = n = 3$. One can see that the result, also in the dependencies, is basically the same as in the decoupled architecture approach. The difference is that the head position $\boldsymbol{s}$ is no more freely placeable. This disables the possibility to use the head position as an expression of gestures. On the other hand, due to the joint dependencies, the solution that is obtained for the neck configuration (if a valid solution can be found) reflects the humanlike posture demanded in the second subtask.

Liégeois pseudo-inverse method The approaches mentioned before can satisfy the main task and either of the two subtasks, but not both. Especially the joint coupling enforces a strict relationship of the upper and lower joint angles, which is not as strict in the biomechanical system. The structural analysis and experimental results suggest that the neck movements are carried out according to these joint couplings, but a selective activation of muscles allows the deviation from this coupling. Therefore, a more flexible control strategy is needed that supports the joint relationships, but allows deviation from it.

The approach chosen is to get a unique inverse map through the Liégeois pseudo-inverse method [92]. In this method, the inverse solution to the task space – joint space mapping $\dot{\boldsymbol{x}} = \boldsymbol{J}(\boldsymbol{q})\dot{\boldsymbol{q}}$ can be stated as

$$\dot{\boldsymbol{q}} = \boldsymbol{J}^{\#}\dot{\boldsymbol{x}} + \boldsymbol{N}\dot{\boldsymbol{z}}_{null} \ , \tag{2.11}$$

with the weighted generalized inverse of the Jacobian matrix $\boldsymbol{J}$

$$\boldsymbol{J}^{\#} = \boldsymbol{W}^{-1}\boldsymbol{J}^{T}(\boldsymbol{J}\boldsymbol{W}^{-1}\boldsymbol{J}^{T})^{-1} \tag{2.12}$$

and the projection into null space of $\boldsymbol{J}$

$$\boldsymbol{N} = \boldsymbol{I} - \boldsymbol{J}\boldsymbol{J}^{\#} \ . \tag{2.13}$$

$\boldsymbol{W}$ is a weighting matrix as proposed by Whitney [156] (see also Hollerbach and Suh [34] for more information). Due the minimization of the weighted joint velocities, joints with higher weights are moved less. Therefore, the relation of joint angles defined in equations (2.4) to (2.7) can be expressed in the weight relations.

The null space vector $\dot{\boldsymbol{z}}_{null}$ projects the gradient of a performance criterion p onto the joint motion, so that p is reduced through subsequent motion [34]. Here the performance criterion is chosen such that the neck should keep a default posture $\boldsymbol{q}_0$, with

$$p = \frac{1}{2}\sum_{i}(q_{0,i} - q_i)^2. \tag{2.14}$$

This leads to the null space vector

$$\dot{z}_{null,i} = w_{null,i} \frac{\partial p}{\partial q_i} = w_{null,i}(q_{0,i} - q_i) \tag{2.15}$$

$$\dot{\boldsymbol{z}}_{null} = \boldsymbol{W}_{null}(\boldsymbol{q}_0 - \boldsymbol{q}) \tag{2.16}$$

$\boldsymbol{W}_{null} = diag(w_{null,i})$ is a weighting matrix, whose entries are the gain for the proportional control of the deviation from the default posture $\boldsymbol{q}_0$, with $i = 1, \ldots, n$. The default posture is relevant for achieving the subtask of a humanlike posture. While $\boldsymbol{q}_0 = \boldsymbol{0}$ seems preferable if the range of motion for each joint is symmetric [33, 161], the default vector proposed here is determined from the biomechanic properties given in Section 2.4.3, with the natural deflection in γ taken into consideration. This leads to $\boldsymbol{q}_0 = [0, -c, 0, c, 0, \sigma_0]^T$, with c as the offset parameter specified in equation (2.4) and $\sigma_0 = 1.2\,\mathrm{m}$. The choice of σ_0 is motivated from the primary use in social interaction: 1.2 m is the border between friend and stranger personal space [70]. The nonuniform shape of the personal space with narrower sides does not matter in this case, as the main task demands the facing of the interaction partner.

Similar approaches have been used to realize head-eye coordination in other systems, enabling the transfer of the following gesture synthesis to those systems. Eye movements can be consistently integrated by expanding $\boldsymbol{q}$ with the DoF for one eye and mirroring the movement on the other eye, so that for commonly used pan-tilt cameras $\boldsymbol{q}$ expands to $\boldsymbol{q} \in \mathbb{R}^8$. Another way is described by Omrčen *et al.* [33], controlling the cameras independently in a tree structure.

Implicit communication expression

The previously considered control strategies vary in the degree of support of the subtask of displaying gestures. While the joint coupling offers no way of adding the motion of gestures without breaking the coupling or neglecting the other tasks, the decoupled architecture at least can include an appetence/avoidance reaction, as the head position and thus the distance to an object is free to choose. More consistently, gestures, postures and affect display can be modeled as null space motion in Liégeois pseudo-inverse method. Therefore, possible null space vector modifications are described, which model the desired implicit communication expressions, listed in Table 2.12.

For the postural shift, the adaptor shift is automatically fulfilled, as the optimization criterion p of the null space vector leads to a return to the default posture, if no other tasks need to be satisfied. The regulator postural shift can be displayed by a small random offset to $\boldsymbol{q_0}$, so that the default posture the configuration is striving for is varied. Head shaking is a null space motion in the upper joint ρ, coupled with the lower joint δ,

$$\dot{z}_{null,\rho} = w_{null,\rho}(q_{0,\rho} + a\sin(2\pi f t) - q_\rho) \tag{2.17}$$

$$\dot{z}_{null,\delta} = w_{null,\delta}(q_{0,\delta} + da\sin(2\pi f t) - q_\delta) \tag{2.18}$$

with a denoting the amplitude and direction of the gesture movement, f the frequency of the sinusoidal wave and d the turning rate between the upper and lower joints, see (2.6).

Nodding is a null space motion in the upper joint ω and to a lesser extent γ, as again the motion is coupled due to the muscular structure, such that

$$\dot{z}_{null,\omega} = w_{null,\omega}(q_{0,\omega} + a\cos(2\pi ft) - a - q_\omega) \tag{2.19}$$

$$\dot{z}_{null,\gamma} = w_{null,\gamma}(q_{0,\gamma} + ba\cos(2\pi ft) - ba - q_\gamma) \tag{2.20}$$

with b as the flexion rate between the upper and lower joints, see (2.4). Experiments by Hashimoto *et al.* [21] support the relation of $b \ll 1$. For both nodding and head shaking, the extension of the head motion system with eyes is necessary, so that the virtual link can stay on the fixation point during the gesture. If the respective DoF in the eyes are missing (pan for shake, tilt for nod), depending on the weighting, the lower joints would achieve target fixation with counteracting the oscillation of the upper joints, leading to large amplitude movements of the head. However, one of the key characteristics of both shake and nod is that the position of the head is more or less stable, while the head is rotating around the respective upper joint. With the position stable and a forced rotation in the upper joint, the only way to ensure both target tracking and gesture execution is to counter the oscillation with the oculo-vestibular reflex. The head roll for signaling doubt can likewise be modeled as

$$\dot{z}_{null,\varepsilon} = w_{null,\varepsilon}(q_{0,\varepsilon} + a\sin(2\pi ft) - q_\varepsilon)\ , \tag{2.21}$$

A way to implement a pursuing motion has been described by Takanishi *et al.* [170], in which a head follows a target within an *"area to reaction"*. It is based on the direct shift of the position of the head, as described earlier for the decoupled architecture. The *"area to reaction"* resembles a personal space zone in its function, but is not linked to the concept of personal space. The method here is based on the default virtual link length σ_0, modifying the distance to the target in null space

$$\dot{z}_{null,\sigma} = w_{null,\sigma}(q_{0,\sigma} - a\sigma_0 - q_\sigma)\ , \tag{2.22}$$

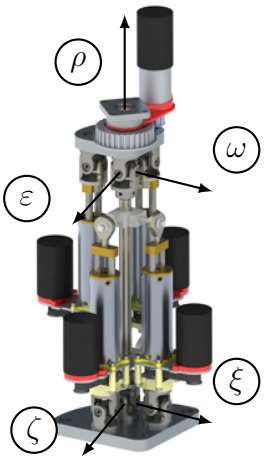
with $a \in [-1, 1]$, determining whether an approach or avoidance is performed. Simultaneously, the weight $w_{null,\sigma}$ has to be changed. During normal operation, the distance to the target is least weighted, as the it is a by-product of the other tasks. To achieve an approach or avoidance behavior, the weight needs to be temporarily raised, enhancing the gain of the proportional null space controller for $\dot{z}_{null,\sigma}$. Affect display takes a special role among the implicit communication signals. While for the categorical emotions a lookup table would be sufficient to assign the characteristics of a specific emotion to the posture of the neck, a more consistent approach is to model them through their respective PAD dimensions. The characteristics of the pose change connected to the PAD dimensions in Table 2.11 only affect the joints ω and γ, leading to a raised, bowed, forward or backward pose depending on the combination.

$$\dot{z}_{null,\omega} = w_{null,\omega}(q_{0,\omega} - K_{P,\omega}P - K_{A,\omega}A - K_{D,\omega}D - q_\omega) \tag{2.23}$$

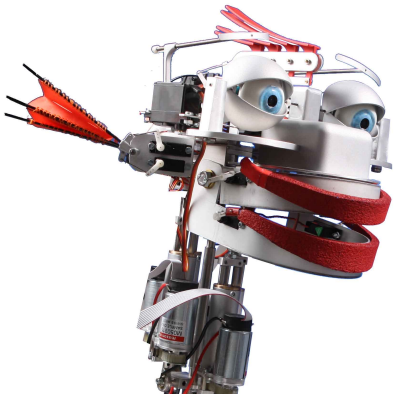
$$\dot{z}_{null,\gamma} = w_{null,\gamma}(q_{0,\gamma} - K_{P,\gamma}P + K_{A,\gamma}A - K_{D,\gamma}D - q_\gamma) \tag{2.24}$$

with gains $K_{P,i} \geq 0$, $K_{A,i} \geq 0$, $K_{D,i} \geq 0$ and PAD values $P, A, D \in [-1, 1]$.

2.4.6 Application



(a) CAD rendering of 5 DoF robotic neck



(b) Neck in combination with EDDIE

Figure 2.14: Robotic 5 DoF neck that incorporates the simplified model of the human neck as proposed in Section 2.4.3.

The applicability of the bioinspired model is shown with a robotic neck[5] that has 5 DoF, see Figure 2.14. The kinematics resemble those from the proposed simplified model, with slight differences in the placement of the joints. The universal joint at the bottom is turned by 90°, so that both rotational axes are in the xy-plane. The model joint angles δ and γ can be constructed from the rotated universal joint angles ζ and ξ. Also the order of the joints in the upper spherical joint is different, with the turning angle of the head ρ being the last joint in the serial link setup. Since all joint axes intersect in one point, the order of joints can be neglected.

The first four joints $(\delta, \gamma, \omega, \varepsilon)$ are composed of linear actuators with cardan joints, fixed to a common pole supporting the neck. A change in actuator length leads to a tilt of the pole to the side of the actuator fixture. The turning of the head is achieved with a motor turning a belt drive. The low level control for each joint is described by Troll [122]. The artificial neck is with 19.4 cm slightly longer than an average adult neck (14 cm), but is in good proportion with the robotic heads described in Section 2.3.7. The elongation is due to the travel range of the linear actuators, which is necessary to emulate the active agility of the human neck. The current version of the robotic neck thus achieves angular ranges that come close to those of an average human neck, see Table 2.13.

[5]Design by ACCREA Engineering

Movement	Robotic neck	Human neck
Flexion	$+95°$,$-80°$	$\pm 125°$
Sideways roll	$\pm 30°$	$\pm 45°$
Turning	$\pm 80°$	$\pm 90°$

Table 2.13: Comparison of movement ranges of the 5 DoF robotic neck and an average human neck.

2.4.7 Experimental Evaluation

Hypothesis 2.6. *The proposed model of humanoid neck motions is perceived as more humanlike than posture generation without the proposed default posture and joint angle dependencies.*

For the following experimental evaluation, the robotic neck has been used in combination with the EDDIE head. In this evaluation, the posture generation of the robotic neck was performed according to two conditions:

- **Min-norm**: The configuration of the robot neck is determined by a standard min-norm solution
- **Biomimetic**: The configuration of the robot results from the biomimetic approach

The min-norm solution $\boldsymbol{q} = \boldsymbol{J}^{\dagger}\boldsymbol{x}$, with $J^{\dagger} = \boldsymbol{J}^T(\boldsymbol{J}\boldsymbol{J}^T)^{-1}$ as the pseudo-inverse of $\boldsymbol{J}$, and $\boldsymbol{q}_0 = \boldsymbol{0}$ is used to generate the comparison postures. It is a standard optimization approach to get a unique solution for the underdetermined equation system. The biomimetic solution is proposed in Section 2.4.5, with default posture $\boldsymbol{q}_0 = [0, -c, 0, c, 0, \sigma_0]^T$ and the proposed joint dependencies.

For this online study 26 participants, 20 male and 6 female, with a mean age of 27.77 years (standard deviation: 1.82) were shown videos of EDDIE with the 5 DoF neck in a side view. In these videos, the robot displays 8 postures: a start posture with the respective default configuration, 3 postures with target points to the side and above or below of the robot, a fallback to the initial configuration and below in front of robot, a large sidewards motion and back to the start posture.

The online study consisted of three parts: First, participants were greeted and informed about the number of videos they would be shown. They were asked to provide demographic data (age, gender) and given instructions how to rate the videos. In the second part, a video of either condition **min-norm** or **biomimetic** was shown. Participants could replay videos as often as they liked. The *"godspeed I"* and *"godspeed II"* questionnaires [331] for anthropomorphism and animacy were provided to evaluate the video. With these ten questions, subjects would rate the robot on a five point scale along semantic differentials such as *"fake – natural"* or *"stagnant – lively"*. See Figure 2.15 and Table 2.14 for all differentials. After completion of the first questionnaire part, participants then saw the alternate condition and rated again on the godspeed I and II constructs. In the third part, both of the previous videos were shown simultaneously in a side-by-side view. Participants were asked to choose which of the videos looked more humanlike and lively, representing the overall constructs of godspeed I and II.

Construct	t	p
fake - natural	2.687	0.013
machinelike - humanlike	-3.176	0.004
unconscious - conscious	-1.614	0.119
artificial - lifelike	1.955	0.062
moving rigidly- moving elegantly	-2.23	0.035
dead - alive	5.13	0.000
stagnant - lively	7.298	0.000
mechanical - organic	-1.735	0.095
inert - (inter)active	-2.44	0.022
apathetic - responsive	3.638	0.001

Table 2.14: Results for paired t-tests between conditions, with significant results highlighted.

Significance level for all performed tests was set to $\alpha = .05$. Paired t-tests revealed several significant differences for several constructs between conditions, which are given in Table 2.14.

In the direct comparison, 20 of 26 participants (76.9%) rated the biomimetic condition as more humanlike. Furthermore, 20 participants (76.9%) rated the biomimetic model as more lively. The ratings are not identical, as 4 persons rated the **biomimetic** condition as more humanlike, but less lively, 4 rated the **min-norm** condition as more humanlike, but less lively and 2 participants chose condition **min-norm** in both categories. Binomial tests showed significant results of the **biomimetic** condition for both humanlikeness ($p = 0.011$) and liveliness ($p = 0.011$).

Discussion

The direct comparison shows a clear preference for the biomimetic posture generation. Also in the concept reflecting the hypothesis, *machinelike* versus *humanlike*, results for the biomimetic condition are significantly better. These results lead to the assumption that hypothesis 2.6 is verified. However, it should be noted that the robot is also perceived as more fake, artificial and dead in the biomimetic condition. This could be explained with the concept of the uncanny valley [52]. The robot head has been developed with a machinelike appearance, which is partially contradicted by the posture generation aiming at being more humanlike. It should be further investigated if the application of the biomimetic approach shifts the robot into the uncanny valley, leading to the discrepancy of being seen as more humanlike, but at the same time more fake.

2.4.8 Summary

In this section a method for generating humanlike head and neck postures has been introduced. The starting point has been again the anatomy of the human, which served as a blueprint for the biomimetic analysis and abstraction process. It has been abstracted in a simplified 5 DoF model, which extends the state-of-the-art models to cover more of the

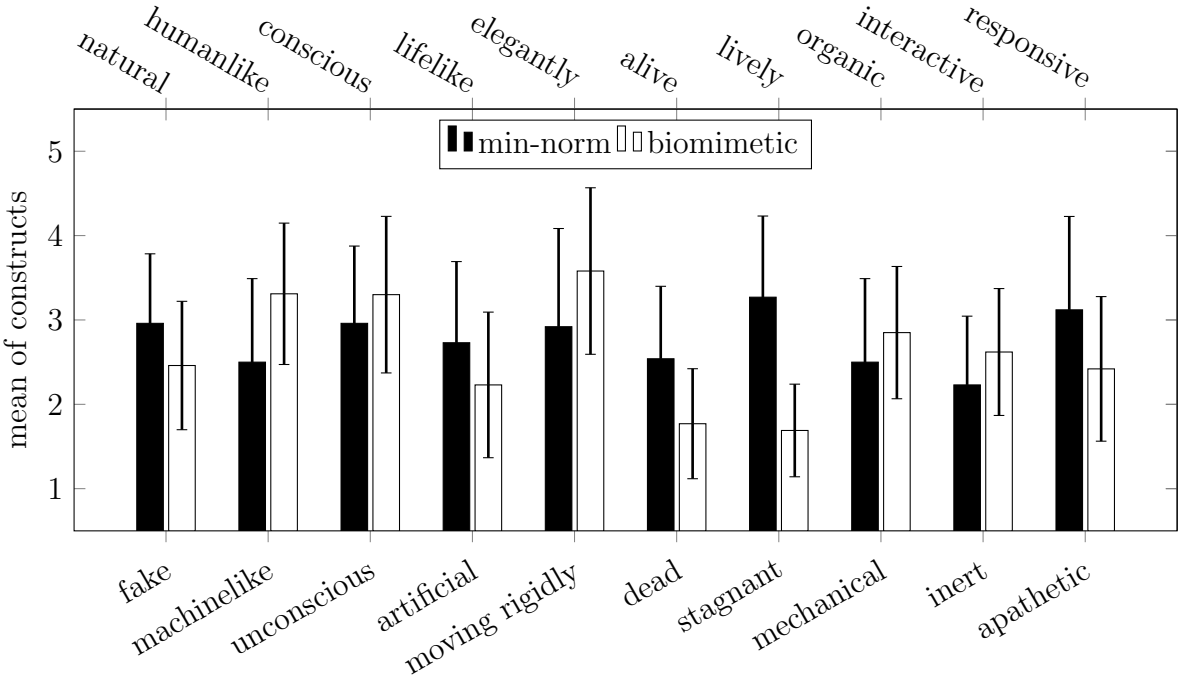


Figure 2.15: Godspeed I and II: Mean ratings (rated on Likert scales from 1 to 5 on the semantic differential) with standard deviations of each construct

possible human neck motions, but at the same time is compatible with current rigid link robotic necks. In an experiment with human subjects, coefficients for the biomechanical coupling of joints and muscles for the 5 DoF were derived and a default posture determined.

These characteristics have been included in a control method based on the Liégeois pseudo-inverse method, with the joint dependencies modeled in the weighting matrix and the default posture in the performance criterion of the null space vector. This control method allows for the execution of the primary task, i.e. fixating a target coordinate, while including the humanlike postures as subtasks. Experimental results have shown that indeed the proposed model is perceived as more humanlike than a standard optimization approach. It should be noted that, while the generated end postures are humanlike according to the model, the motion dynamics and trajectories are not necessarily computed and perceived as such. This requires an extension of the model to specifically take them into account and possibly a reevaluation of the experimental results with varying dynamics and/or trajectories in future work.

Furthermore, within the null space of the task space, gestures can be performed that do not interfere with the primary task. An extensive aggregation of implicit communication signals, such as gestures, have been described and formulated in the control method. For the inclusion of head postures in the emotional framework introduced in the previous section, posture aspects have been dimensioned along the PAD representation.

3 Integration of Implicit Communication Signals Mimicry and Smiling

In this chapter, the biomimetic top-down approach is pursued further. The key aspect is how to generate the information that is conveyed through the implicit communication channels. Again, social science serves as a template for behavior models and interaction effects for transfer to HRI. From the manifold of possible behaviors, here two exemplary behaviors are chosen that have a strategic benefit for commonly envisioned application scenarios of robots, i.e. as a interactive service provider. The first behavior is mimicry of the interaction partner. After analyzation of the various types of mimicry, a biomimetic model is built, based on the AIM model by Meltzoff and Moore [233], and evaluated according to resemblance of the original expression and effectiveness in terms of user acceptance. In the second part, a social motivation model is employed and modified as follow-up on emotional mimicry. This model, the system-theoretic model of smiling (SMS) based on the Zurich Model of Social Motivation (ZM) by Bischof [188], is also capable of generating smile variations as the second strategic behavior chosen. Seven different types of smiles are described and evaluated. A combined evaluation of facial mimicry and the modified SMS in terms of the godspeed questionnaires [331] and user acceptance [291] revealed that mimicry of the interaction partner made the robot significantly more likable and that application of the SMS enhanced empathy and subjective performance ratings of the robot.

3.1 Introduction

> *"What is surprising is that the largely subconscious social signaling that occurs at the start of the interaction appears to be more predictive[1] than either the contextual facts (attractiveness and experience) or the linguistic structure (strategy chosen, arguments employed, and so on)."*
>
> (Alex (Sandy) Pentland [213])

The paradigm shift mentioned in Section 2.1, from seeing robots as machines for repetitive, highly structured work packages that are processed isolated (or in joint work with other robots), to referring to robots as "co-workers" and service providers, requires not only new interaction capabilities in form of a new set of interaction modalities, but also the abilities of how and when to make use of these modalities. Recalling the description of information exchange (Section 2.2.1) – that interaction is composed of an action, the change of behavior that others can notice, and a form of persuasion, that influences the interaction partner through communication encapsulated in the action – in this chapter

[1] of the result of finding a mate, getting a job, negotiating a salary, and finding a place in a social network

the action selection and form of persuasion is in focus. Action selection and persuasion have to follow social rules and dynamics, transforming the robots to social actors [298]. A default strategy for humans without knowledge about social protocols is imitation.

Human social actors use social dynamics and the resulting socialization process to strengthen group cohesion, establish hierarchies [221] and manipulate relationships [197]. In this socialization process, implicit communication channels are used to convey information such as emotional contagion (observed as early as in infants) to bond within the group. Dysfunction of the processing of social communication leads to socialization difficulties, as for example observed in autism or the Asperger Syndrome. The repertoire for implicit communication signals, an integral part of the social cues involved, is given in Section 2.2.1. This leaves to investigate in which contexts to apply these signals.

A robot can take the role of a social actor as a service provider (server [230], tour guide [314], assistant [193], etc.), whenever the interaction and information exchange is the key component of the actual service to provide. It is important to note that interaction should be a necessary part of the task. There is no benefit in integrating social communication in tasks that are focused on efficiency and require no social experience [322]. However, if the social experience is a key component of the service, both user and system could benefit from the robot following social rules and/or influencing the interaction in a persuasive way. For the user, including these elements could lead to a more satisfactory interaction experience. For the robot, since it is usually (at the current state of robotics) the interaction partner lower in hierarchy, social elements in the communication allow for subtle ways of influence, without challenging the hierarchy explicitly or interfering with the task at hand. This is additionally of help, if the robot is not merely recipient of orders, but actively requires action of a human, persuading users for its own good. Example are the Interactive Urban Robot (IURO) robot [211], requiring help from users not directly benefiting from the task of the robot, or Autom, a robotic weight loss coach [228].

Motivation

Picard, one of the pioneers in attributing social abilities to technical systems, states that affective computing might lead to increased performance of the computer [206]. Here the focus is on two social behaviors, that have been proven to be beneficial for service providers: mimicking and smiling. Mimicry is a persuasive element in communication, which is low in effort, cost and risk, is unlikely to be met with new or continued rejection [293], works even if the interaction partner is aware of the persuasive intent [190] and is best for the interaction partner lower in hierarchy [196]. It is notable that according to Bailenson and Yee [227], the effect persists even when the person being mimicked is fully aware that the mimicker is an artificial agent, which indicates that this might also be applicable to the proposed system with a robot.Therefore, it is worthwhile to investigate how to perform and analyze mimicry with an artificial agent. Smiling on the other hand is known in management manuals to be a good server attitude [210]. Since there are various variants of smiles and smiling is dependent on context to be effective, a model needs to incorporate the social dynamics and contexts. Equipping robots with these interaction capabilities could transfer the benefits from HHI to HRI in a measurable and significant way.

Outline

In this chapter, systems for automatic mimicking and smiling will be modeled, applied to a technical system and evaluated. First, in the biomimetic design phase in Section 3.2.3, the various types of mimicry are described and their impact on social interaction collected. With the basic structures of mimicry identified, a model for automatic facial mimicry is defined in Section 3.2.4, applied to a robotic system and evaluated. Second, a socially motivated model that is capable of providing facial expressions and behavior based on the social context is introduced in Section 3.3. A specialized version of this model is then used to generate context-sensitive smile variations and is extended to perform mimicry as well, see Section 3.3.4. Finally, evaluations of the smile variants and a comparison of the social effects of facial mimicry and the socially motivated mimicry are performed.

3.2 Mimicry

> *"Each of us is in fact what he is almost exclusively by virtue of his imitativeness"* (William James [250])

Mimicry is unconscious [186] and unintentional [292]. It happens between interaction partners with a positive or neutral relationship [243]. The attitude towards the partner determines the degree of imitation: a positive attitude fosters mimicry, a negative attitude attenuates mimicry or leads to incongruent reactions [323], whereas in competitive interaction conditions even counter-mimicry, i.e. showing contrasting expressions, appears [294]. This attitude can be based on the in-group versus out-group affiliation of the persons: more mimicry happens between in-group personae [294]. The in-group affiliation can be seen as similarity between the subjects, with similarity leading to more mimicry [237].

In social interaction, gestures, postures, facial expressions, behaviors and emotions of others are mimicked [292, 255, 186, 254, 294, 208]. Mimicry can be observed both in children [269, 234, 256, 205] and adults [182, 287, 327, 265, 216, 286, 289]. An interesting notion is, that the history of mimicry behavior of a person can predict future behavior [222]. It is therefore not only a strong instrument of persuasion, but could also be used for a better understanding and analysis of social dynamics in technical systems.

The versatility and ubiquity in interactions make it a worthwhile effect to investigate in this section. First, the various types of mimicry are categorized. Then the influence of mimicry on measures of social dynamics such as trust and empathy is given. A subset of the types of mimicry in form of facial and emotional mimicry is then realized in a robotic application. Effects of the realized mimicry setup are then experimentally evaluated.

3.2.1 Problem Statement

Mimicry of behavior as a form of social interaction is known in literature since Adam Smith (1759) [220] and Charles Darwin (1872) [19], but the exact underlying mechanisms are still subject to research. Effects of mimicry have been investigated in a number of HHI experiments, providing evidence that mimicry has various influences on the interpersonal relations of people. Few transfers of these findings to artificial systems, being it virtual

agents or robots, have been undertaken so far. Questions in this context remain to what extent these findings can be replicated in artificial systems, with only a subset having been tested, yet. Moreover, as robots commonly serve a purpose besides mere satisfaction in interaction, methods and effects of mimicry on the perception of task performance need to be investigated. The proposed solution in this section extends the state-of-the-art by explicitly evaluating facial expression mirroring in contrast to head, arm or body gestures. Additionally, the facial expression mirroring happens automatically and online during the HRI. Experimental setups not only evaluate the effects of mimicry on the perception of the system as a social actor, but also if artificial agents could derive benefits in task related performance ratings. This would give arguments for expression capabilities in systems that are not primarily built for providing human-like interaction.

3.2.2 Related Work

The influence of behavioral mimicry has been subject to studies in the field of human-human-, human-agent- and human-robot-interaction. Previous work has already shown the transferability of inter-human-findings to virtual agents and social robots.

One work for example is a prototypical demonstrator by Bartlett *et al.* [264], which can classify seven emotions from facial expressions and display the emotional state on a virtual agent. The work of Bailenson and Yee [227] on *"digital chameleons"* concludes that embodied virtual agents mimicking head movements are viewed as more persuasive and likable compared to agents with prerecorded movements. Gratch [278] and later on Huang *et al.* [310] report on *"virtual rapport"* with virtual agents, showing benefits of mirroring head movement and posture shifts through increased speaker engagement and improvements on the behavioral and interactional level compared to unresponsive agents. Besides the reported successful mimicry of agents and positive effects on dyadic conversations, mimicry does not necessarily occur in interactions. Results by Simons *et al.* [267] showed no tendency of participants to mimic the behavior of the artificial agent. This could be due to the reduction of mimicry to eyebrow movements and self-touching gestures in their experiments.

In the field of social robotics, Kanda *et al.* [223] could improve route guidance interactions with a robot by incorporating cooperative body movements (e.g. synchronization of arm movements), enhancing both reliability and sympathy. Hegel *et al.* [185] developed a system that mirrors the emotions happiness, fear and neutral as recognized from the speech signal by facial expressions. Experimental results show that the mimicking robot is perceived as reacting more adequately to the emotional context of a situation and to recognize emotions better. Breazeal mentions observations of dynamic affective interaction of users with the robot Kismet [302]. Similar observations are reported for Sparky by Scheeff *et al.* [43]. With the robot being unaware of facial expressions of the human, the affective mirroring is more on the human side. Tscherepanow *et al.* [332] developed a system for direct imitation of human facial expressions, with a direct mapping between movements of facial elements and the joint space of the robot. Riek *et al.* [217] studied the effect of automatic head gesture mimicking with a chimpanzee robot in a Wizard-of-Oz (WOz) setup. The robot would listen to participants while either mimicking all head gestures, only nodding or no mimicking, resulting in different levels of interaction

satisfaction.

Technical methods to detect mimicry are also applied to other fields of applications. A system for airport screening by Meservy *et al.* [187] attempts to automatically detect behavioral patterns such as mimicry, that indicate deception from nonverbal behavioral cues and classifies deception and truth. Keller *et al.* [318] utilize Motion Energy Analysis [326] on the synchrony between the movements of the participants in a dyadic interaction. Pentland [213] measures mimicry in conversational audio patterns, by using auditory backchannels and short words, for the prediction of social interactions.

3.2.3 Biomimetic Design

So far, mimicry has been mainly a subject of study in HHI. For the design of an artificial mimicry system, these social science studies can be used as a blueprint to identify and abstract the essentials of the biological system. The first step is to categorize mimicry according to the various types found in literature. Commonalities of mimicry types in the characteristics, the manifestation and the effects on interaction partners aid in the selection of relevant aspects. This defines the goals of the reproduction in an artificial system, the base line that should be achieved. With these goals in mind, mechanisms to achieve them need to be designed. An inspirational source for the structure of such mechanisms can be found in neuroscience, which as a discipline is trying to uncover the fundamental neurological rules to behavior such as mimicry.

Types of Mimicry

The categorization here is based on the types of mimicry commonly found in recent literature [323, 277, 293, 208, 307, 294, 254, 292, 186, 309, 196]. Often the categorizations provided only differentiate between emotional versus non-emotional mimicry, e.g. given in Likowski *et al.* [323]. This is appropriate when the mimicked contextual *act* [309] is of interest and not the lower level motor mimicry. It should be also noted that, if the mimicry abilities of humans are the goal, in early infant-hood mimicry starts out as imitation of muscle movements and then with age (around 1 year) and experience shifts to mimicry of *acts* [309]. Therefore, with the purpose of building a model including both motor level and higher level mimicry, non-emotional mimicry is differentiated based on the involved implicit communication channels. The order of the description of the types is based on the level of abstraction involved. Facial mimicry is direct mimicking of the perceived muscle movements, whereas emotional mimicry on the other end requires the mimicking of the *concept* of emotions, not only the manifestation.

Facial Mimicry describes the mimicry of facial expressions or distinctive facial muscle movements. The imitation is specific [309], i.e. the reaction is with congruent facial expressions [330]. Facial mimicry has been extensively studied in neonates and young infants [233, 271, 194], but has been also found in adults [182, 287, 327, 265, 216, 286].

While the initial experiments of Meltzoff and Moore [233] received criticism [235], they formulated in follow-up studies a theoretical model that incorporates at least three of the 10 major aspects of infant mimicry [309]. The active intermodal mapping (AIM)

model introduced by Meltzoff and Moore forms the biological basis for the later proposed biomimetic model of facial mimicry.

Prosody and rhythm mimicry involves implicit communication channels that are mimicked to synchronize speech and movements during interaction [235, 259, 282, 316, 184, 219]. With prosody, the intonation patterns are imitated. Rhythm denotes the synchronization of speech rhythms, or timing of postural change, gesture and turn-taking.

Behavioral Mimicry, also named *"behavior matching"* [255] or *"posture sharing"* [240] is focused on the adoption of behaviors such as body configurations, gestures, manners [304, 240, 272, 186, 293, 190] or laughter [257]. Due to the involved timing, it is closely related to rhythm mimicry [292].

Emotional Mimicry involves - in contrast to the previous types - not imitation of the lower level social signals, but *"contextualized emotions"*. Hess *et al.* [208] define *contextualized* as the *"in its broadest sense ... inclusion of social information"*, which can range from explicit information, e.g. the situation in which the interaction occurs or the social status of the interaction partner (group membership), to implicit context, such as signaling of the emotional state or affiliative intentions [290]. Hess *et al.* further reason that, as the relationship between interaction partners is taken into account, emotional mimicry acts as a *"social regulator"*. The social regulation function is a matter of empathy, that is the comprehension of the emotional status of the other person through simulation [299, 325], and the motivation to control the social distance: mimicking assists in an attempt to bond with the interaction partner, but can also increase the social distance and "coolness" if subdued.

Emotional mimicry is backed up through findings by Termine and Izard [236], who found nine month old infants to be mimicking emotional expressions such as joy, sadness and anger. This is in line with the findings by Meltzoff and Moore, that approximately one year old infants shift to the mirroring of concepts [309]. It is also supported by functional magnetic resonance imaging (fMRI) experiments, showing increased neuronal activity of subjects in brain regions responsible for the facial expression of emotions, when asked to observe or mimic emotional facial expressions [198]. However, there are contradicting findings in literature about the mimicry of fear and disgust. While Fischer *et al.* [253] find no clear evidence for the mimicry of these two emotions, fMRI studies [274] show at least a connection between the observation and the feeling of disgust in the the *anterior insula* and *anterior cingulate cortex*.

Influence

In order to use mimicry goal-oriented, it is necessary to know the social parameters that are influenced by it. Furthermore, these parameters need to be evaluated for their purpose in HRI.

Affiliation is a major component of social dynamics. Several studies focus on the effect of mimicry on *"achieving social connectedness"* [270], see for instance [282, 184, 277, 303, 195], and the connection to achieve social goals via this affiliation [306, 254, 226]. In this process, encouraged by the desire to belong to a specific group [270], excluded persons mimic in-group members more than in-group members do between each other [293]. In the current state of society, robots and other artificial agents are not very likely to be seen as member of a social group. Mimicry as part of the affiliation process could thus help to make the robot appear less alien.

Trust has been shown to be raised through mimicry in sales talk. It is significantly induced in the seller, if the buyer mimics him/her. This seller trust is assumed to be the reason for the connection of mimicry and the positive outcome of negotiations [197]. Maddux *et al.* trace this effect in negotiations back to *"putting the opponent at ease"* and therefore being able to elicit information from the mimicked interaction partner [197]. A different interpretation is the enhanced predictability of the interaction partner through mimicry, as trust in early relationships, e.g. transient sales interactions, is based on the predictability of the partner [191]. An analysis of trust in machines by Muir [321] suggests similar trust-building mechanisms towards machines, as the same influential factors play a role. Again, growth of trust is dependent on the perceived predictability of the machine.

Rapport between interaction partners is fostered through behavioral mimicry [254]. It has been one of the effects linked quite early to mimicry [229, 201, 304, 240, 272, 190]. Huang *et al.* [310] report several benefits for a virtual agent through rapport, such as less tension [313], less embarrassment [252], and more trustworthiness [252], which would concur with the raise in trust.

Empathy is assumed to be generated through facial and motor mimicry. Hatfield *et al.* [263] suggest that the perception of emotions in others leads to automatic mimicry of the emotion, which in turn – in tradition of James's theory of emotion [250] – leads to feeling the emotion through the bodily feedback [300, 183, 266]. This phenomenon is referred to as *"primitive emotional contagion"* and has been connected to the mirror neuron system [202, 238, 307].

Liking of another person is evidently increased via behavioral mimicry [284, 262, 254, 270]. However, the amount of mimicry is crucial to have a positive effect on liking; experiments by Leander *et al.* [296] found that participants are *"feeling colder"* if not mimicked during interaction, whereas *"nonstandard amounts of mimicry"* are perceived as counter-stereotypic behavior and thus elicit threat responses [335]. The reaction to the nonstandard amounts is not necessarily due to an exaggerated amount of mimicry, but due to violating expected standards of social interaction [296]. While the previous social parameters would suggest a – *the more mimicry, the better* – strategy in HRI, the effects on liking call for a more differentiated view on the application of mimicry. It is an open question what the standard amount of mimicry and the expected standards of social interaction are for artificial agents in order to not negate the liking effects.

3.2.4 Mimicry Model

The proposed automatic mimicry model is based on the AIM model by Meltzoff and Moore [233, 309], see Figure B.1 in Appendix B for the schematic of the original model. In the AIM model, the infants' motor expression is compared to the perceived expression of the care taker. The comparison is achieved by encoding the perceived and the produced motor acts in a *"common (supramodal) framework"* [309], allowing the detection of equivalences. Mismatches in the comparison lead to a new goal of choosing the appropriate expression from a set of previously determined motor actions.

The AIM model works on the motor level. For the artificial system, see Figure 3.1, the abstraction of expressions in the supramodal framework can therefore be represented with the Facial Action Coding System (FACS), see Section 2.3.3. The exteroceptive object field is formed through a vision system, passing images to a FACS based expression analysis. The comparison can then either be on the FACS features, resulting in a system for facial mimicry, or emotional states extracted from the facial features for emotional mimicry. The same holds for the synthesis side in the artificial agent. The learning part that is denoted *"body babbling"* by Meltzoff is not established as learning in the artificial system, but as the action units to motor mapping and emotion to action unit mapping from Section 2.3.

3.2.5 Application

The mimicry model is applied in a demonstration setup for both facial and emotional mimicry [288] [315]. The synthesis of the expressions is performed with the expression framework established in the previous chapter. The synthesized expression can be displayed using the developed robotic heads, see Figure 3.2 for the setup with EDDIE. Images of the interaction partner are captured via the in-head camera system to analyze the facial expression.

The overall system consists of several modules. A visualization of the connected modules is given in Figure B.2 in the appendix. Additionally to expression analysis and synthesis, the robot head turns the neck to focus on the user's face. Text-to-speech is integrated to vocalize output to the user. The robot head parses text to generate adequate lip movements via visemes. A speech recognition module determines the human's verbal utterances to the robot. The modules are interconnected with a communication backbone based on the Real-time Database (RTDB) introduced by Goebl and Färber [283]. It provides a shared-memory implementation with integrated data storing and is able to handle large amounts of data in real-time, which is required for instance by the vision-based components.

Expression Analysis

The facial expression analysis is provided by Dr. Christoph Mayer[2]. Only a brief overview of the functionality is given here, for further information the reader is referred to [279] .

A model-based technique is used to determine the exact location of facial components such as eyes or eye brows in an image. Geometric models form an abstraction of real-world objects and contain knowledge about their properties, such as position, shape or texture

[2]Intelligent Autonomous Systems Group, Department of Computer Science, TU München

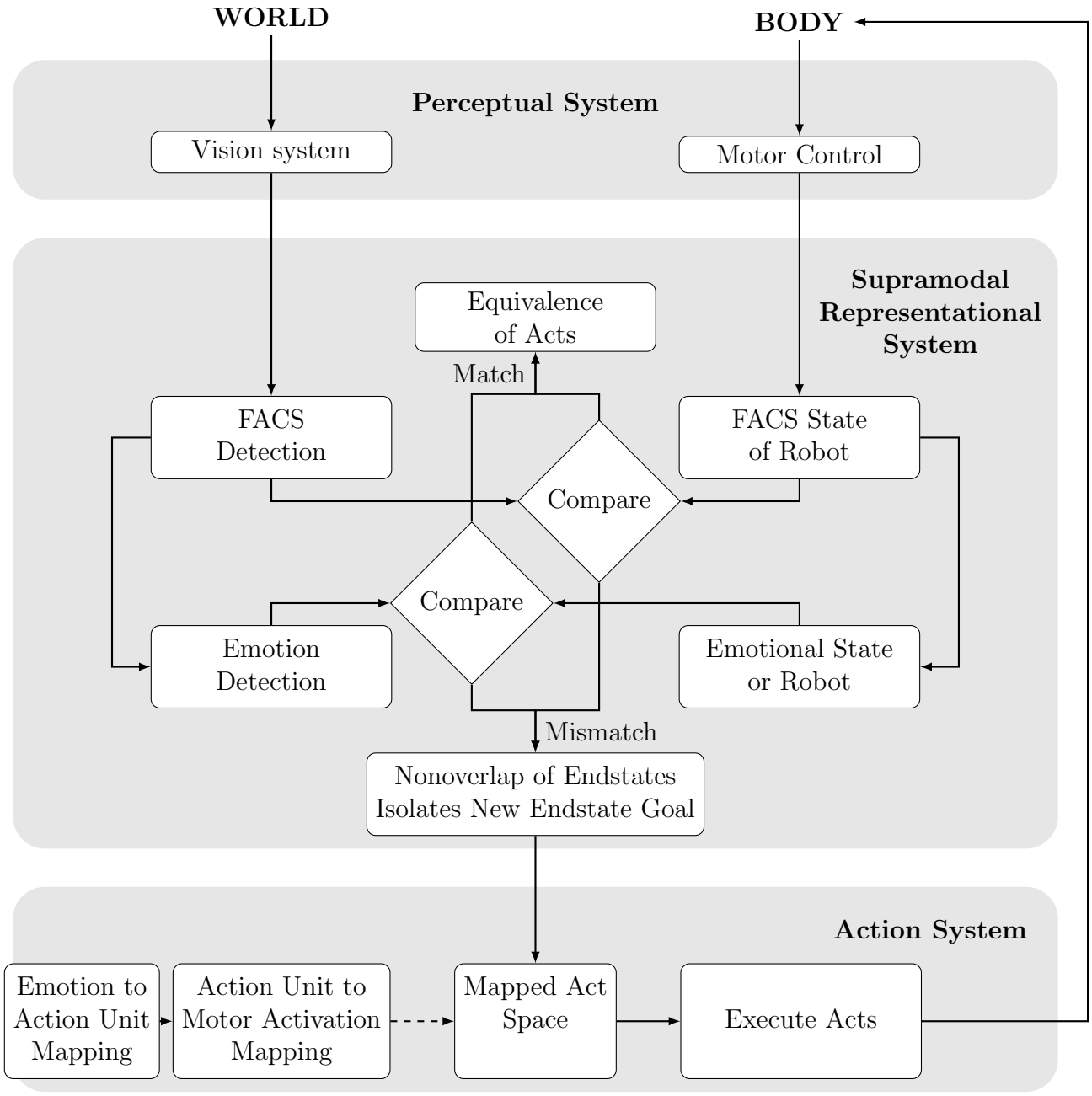


Figure 3.1: Mimicry model, based on the general structure of the AIM model [309]. Both facial expressions and emotional states can be evaluated and mirrored.

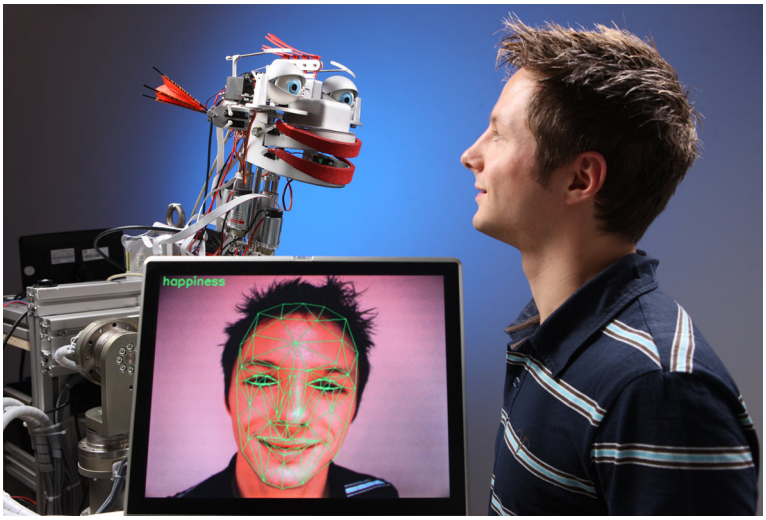


Figure 3.2: Mimicry demonstration setup
PHOTOGRAPH: KURT FUCHS

in a model parameter vector $\boldsymbol{p}$. Since no previous information is available about the image content (except for the fact that a face is visible in it) or about the person in front of the robot head, model fitting is applied to determine reference model parameters $\boldsymbol{p}_0$. These reference parameters are specific for this person and are recalculated when the system is confronted with a new person.

The Candide-3 face model used as the geometric reference is a wireframe model consisting of n = 116 anatomical landmarks [268]. It has been specifically designed for facial expression recognition [218]. The $3n$ dimensional vector $\boldsymbol{g}$ contains the vertex x-, y- and z-coordinates. The model shape is controlled by applying the shape deformation $\boldsymbol{g}_{shape} = S\boldsymbol{\sigma} + A\boldsymbol{\alpha}$ to the basic model structure $\boldsymbol{g}_{basic}$, where $\boldsymbol{\sigma}$ and $\boldsymbol{\alpha}$ contain the parameter values and the columns of A and S vertex movements for a specific parameter. The difference between the parameters in $\boldsymbol{\sigma}$ and $\boldsymbol{\alpha}$ is that A contains motion of facial components that may appear due to facial expressions and S contains motion of model vertexes that are applied to adapt the general face structure to the face structure of a specific person. To control the movement and rotation of the model in 3D-space, a rotation matrix $\boldsymbol{R}$ for the rotation r_x, r_y, r_z, a scaling factor c and a translation $\boldsymbol{t} = (t_x, t_y, t_z)$ are applied, so the model vertex coordinates are computed by $\boldsymbol{g} = c\boldsymbol{R}(\boldsymbol{g}_{basic} + \boldsymbol{g}_{shape}) + \boldsymbol{t}$ from the parameters $\boldsymbol{p} = [t_x t_y t_z c r_x r_y r_z \boldsymbol{\sigma}^T, \boldsymbol{\alpha}^T]^T$. See Figure B.3 in the appendix for the Candide-3 model with different exemplary parameterizations.

Both, the model tracking and the FACS analysis, rely on a neutral reference image of the user. A neutral reference image I_0 of the human interaction partner is obtained and corresponding model parameter values for this image are calculated. Note, that in contrast to tracking algorithms, no prior knowledge of the image content is available here. A solution to this problem is provided by Wimmer *et al.* [328]. To track the model through a sequence of images, its model points $\boldsymbol{g}$ are projected onto the image plane using perspective projection $\boldsymbol{f}$ to obtain their corresponding pixel coordinates $\hat{\boldsymbol{g}} = \boldsymbol{f}(\boldsymbol{g})$ in the camera image, obtaining $\hat{\boldsymbol{g}}_0$ for the reference image. Applying an optical flow based method, the

corresponding points $\hat{\boldsymbol{g}}_{-t}$ from I_0 and $\hat{\boldsymbol{g}}_0$ in the image captured at time step t are calculated. Afterwards, model parameters are approximated to calculate model parameters $\boldsymbol{p}_t$ that minimize the error between $\hat{\boldsymbol{g}}_t - \hat{\boldsymbol{g}}_{-t}$. A subset of model parameters $\boldsymbol{\alpha}$ refers to FACS, which is used to calculate the degree of activation of certain AUs. By selecting model parameters that refer to AUs that can be synthesized by the robotic head, the feature vector $\boldsymbol{x}_t$ from $\boldsymbol{p}_t - \boldsymbol{p}_0$ is extracted. The AUs recognized by the analysis components and synthesized by the robot are AU2 (outer brow raiser), AU4 (brow lowerer), AU5 (upper lid raiser), AU7 (lid tightener), AU13 (lip corner depressor), AU26 (yaw drop) and AU42 (eyes closed).

3.2.6 Experimental Evaluation

The previously described mimicry model and application setup enable a robotic system to mirror an interaction partner. However, for this system to work in the intended way and to provide the benefits to the artificial agent, two major assumptions are made that need to be verified. First of all, the process of mimicry requires the matching of input and output expressions. Modifications of the expression in the transformation path from perception, analyzation according to the abstraction in AUs and the synthesis due to sampling, quantification or mapping errors need to be evaluated whether they distort the perception of correspondence of faces. Second, the transferability of the effects of mimicry on social dynamics from HHI to HRI needs to be tested. The second part is evaluated in conjunction with the effects of smiling and corresponding socially motivated expressions in Section 3.3.6.

Mimicry perception

Hypothesis 3.1. *The mimicry setup creates robotic facial expressions that are perceived close to the corresponding mimicked human facial expression.*

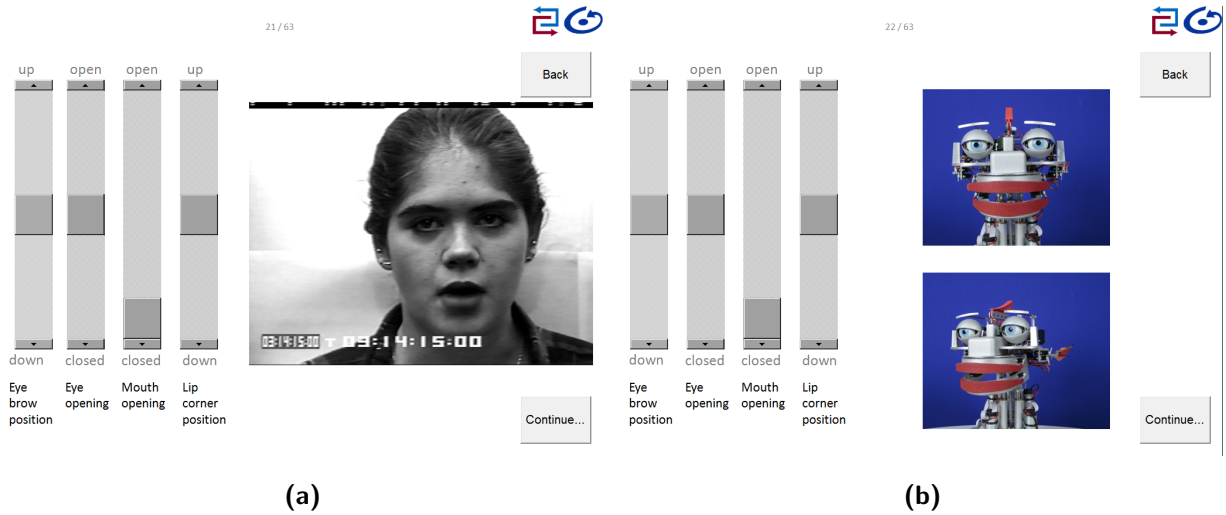
An evaluation is performed if humans perceive the generated robotic facial mimicry expression close to a corresponding human facial expression. Therefore, images are extracted from the CMU Cohn-Kanade Facial Expression Database [214]. The face model is fitted to the database pictures and in a second step the feature vector $\boldsymbol{x}_t$ is extracted. This vector is provided to the facial expression synthesis component to have the robotic head show the facial expression. It is a property of the Cohn-Kanade database that the first image of the image sequences depicts a neutral face. This neutral face serves as the reference image I_0 for the calculation of $\boldsymbol{p}_0$ to support the calculation of $\boldsymbol{x}_t$. In total, 21 pictures are taken from the image database to determine the activation of the AUs with the facial analysis module. From this procedure, pairs of images are gained, with one image depicting a human face and the corresponding second image depicting the robotic head mimicking the human face, see Figure 3.3 for an example. Since the data is automatically extracted and displayed by the system, as it would be in the mirror setup, this evaluation can also be seen as a benchmark of the mirroring. This is done the same way as an actual video stream in the live mirroring setup is processed. Evaluations of the modules as stand-alone systems are presented in Section 2.3.8 for the synthesis and by Mayer [279] for the analysis.

For the evaluation a set of powerpoint slides with automatic data logging was created. Participants got verbal instructions beforehand to follow the instructions on the screen and that they could work without a time limit. Seven persons, four male and three female, were asked to contribute to the evaluation. Since none of the persons is specifically trained on facial expression recognition or FACS coding, it was decided against asking them to rate activations of specific action unit. Instead, they were asked to rate human faces in four categories ($EyeBrows$, $EyeLids$, Jaw, $LipCorners$) and in five intensities. Example annotations were shown to the participants to prevent wrong labeling due to misunderstanding of the instructions, see Figure B.4a in the appendix for the $EyeLids$ rating instruction. For the categories $EyeBrows$, $EyeLids$ and $LipCorners$, a low intensity represents lowered eye brows, closed eyes or depressed lip corners respectively. By analogy, a high intensity reflects raised eye brows, wide opened eyes or raised lip corners. For the Jaw category, a low intensity refers to a closed mouth and a high intensity to a wide open mouth.

In the first evaluation phase, the participants were presented eight images depicting human faces in a predefined order. These images did not have corresponding images with the robotic head. The first phase served two reasons: First, to ensure that the participants had correctly understood the task, second, to have a reference of the users' ability to rate facial expressions. In the second evaluation phase, 21 images from the image database and 21 corresponding images of the robotic head were presented to the participants in random order. It should be noted that the participants were not informed that the image data included matching human-robotic head pairs. Figure 3.4 depicts an example of a human-robotic head pair without rating, i.e. all sliders are in initial state. Similar to the first phase, participants were informed on an introductory slide that now additionally a robotic head would be depicted and example ratings were given, see Figure B.4b. There was no difference in the rating mechanisms for human faces and robotic heads, except for the fact that two images (front view and slightly turned) were presented for the robotic head. Participants were allowed to navigate freely through the test with continue and back buttons.

As mentioned beforehand, the participants were presented eight training images with four sliders each, resulting in 32 slider values. From these training sets, the variance of each slider value is computed as a measurement of similarity in the participants' rating. The variance values in the training set range from 0.0 to 1.47 with a mean variance of 0.41.

A similar procedure is applied to inspect the similarity between the rating of a human face and a corresponding robotic face. The rating of the human face is denoted in one of the $L = 21$ image pairs by one of the $N = 7$ participants with $h_c^{l,n}$ with $1 \leq l \leq L$, $1 \leq n \leq N$ and $c \in \{EyeBrows, EyeLids, Jaw, LipCorners\}$. Similarly, the rating of the robotic face is denoted with $r_c^{l,n}$. To inspect the similarity between human and robotic, a rating $e_c^{l,k} = h_c^{l,n} - r_c^{l,n}$ is calculated. Per participant $21 \times 4 \times 2 = 168$ values for all image pairs l, and all categories c are calculated. Furthermore, to group the inspection by category, data vectors $\boldsymbol{e}_X$ are created that contain all values $e_{c=X}^{l,k}$. A histogram of $\boldsymbol{e}_{EyeBrows}$,$\boldsymbol{e}_{EyeLids}$,$\boldsymbol{e}_{Jaw}$ and $\boldsymbol{e}_{EyeBrows}$ is computed to obtain an intuition of the rating discrepancy distribution, see Figure 3.4. Mean and variance of $\boldsymbol{e}_{EyeBrows}$,$\boldsymbol{e}_{EyeLids}$,$\boldsymbol{e}_{Jaw}$ and $\boldsymbol{e}_{EyeBrows}$ are given in Table 3.1 and are visualized in Figure 3.5.



(a) **(b)**

Figure 3.3: A set of generated image pairs with one image depicting a human and a second image depicting the robot head mimicking the human. The participants were not aware that they were presented pairs of images.

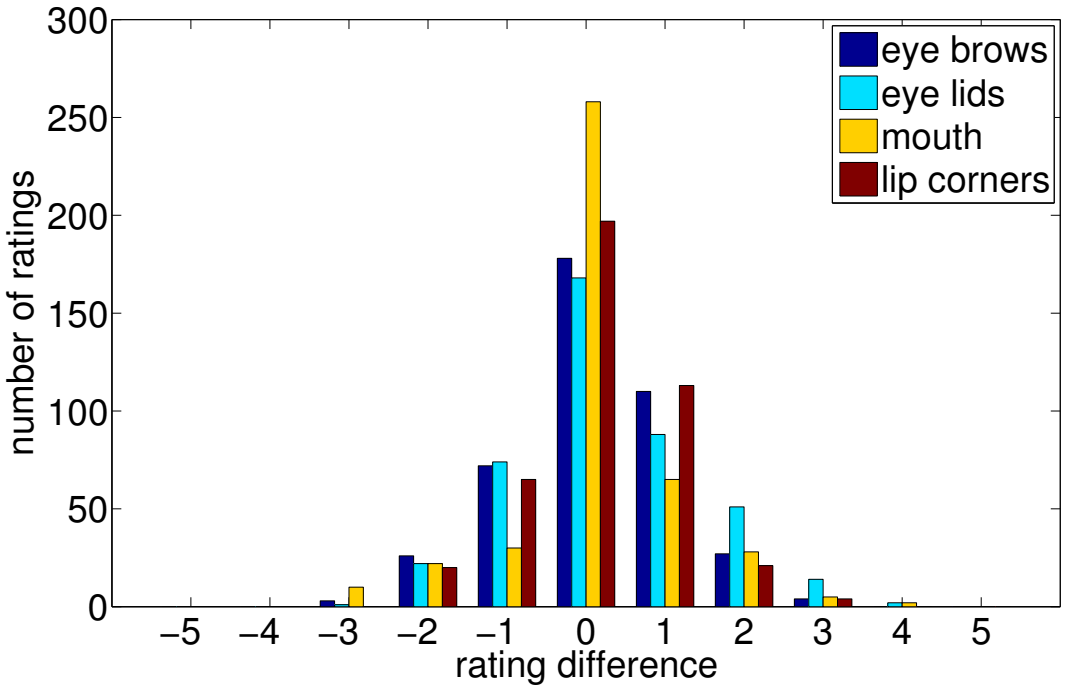


Figure 3.4: Only a small fraction of rating differences between a human face and a corresponding robotic head are larger than one slider unit.

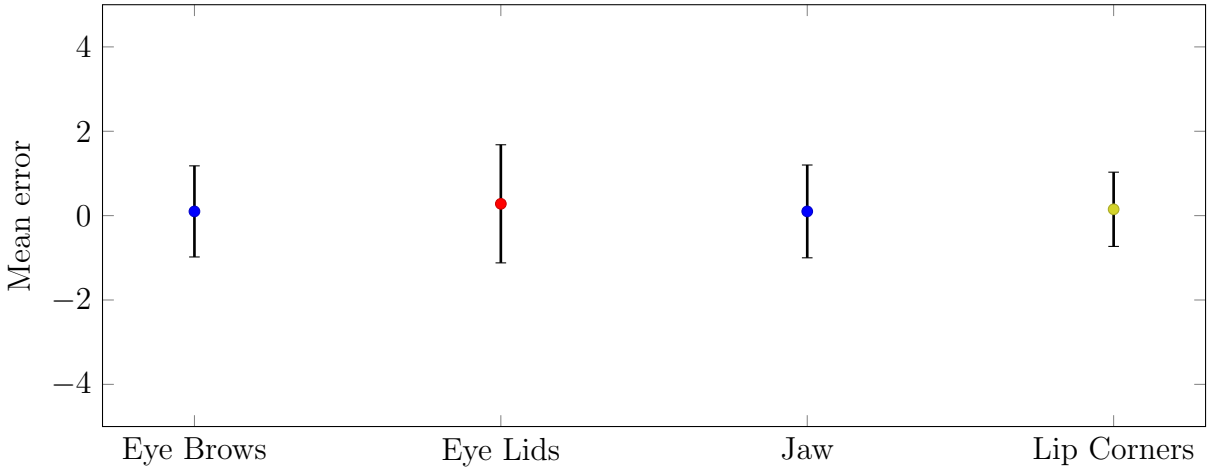


Figure 3.5: Mean errors in rating of human versus mirrored robotic expressions grouped by action units

Discussion

Following hypothesis 3.1, the evaluation demonstrates that images of human faces and images of the corresponding robotic head were rated similar by the participants. Since the participants were not aware of the fact that they were presented human/robot image pairs, this demonstrates that the robotic facial expressions are perceived close to human facial expressions. Similarity in rating demonstrates similarity in perception, rendering the ratings of different participants comparable.

The small mean variance demonstrates that participants rated the training images very similar and therefore their rating is comparable. If different participants applied a different rating scheme, their rating would not be comparable, although they might actually have the same perception of the activation of certain AUs. As can be seen in the histogram, for all categories the most frequent value of $\boldsymbol{e}_c$ is 0, which indicates that participants rated the robotic head and the human face equally. Furthermore, only a very small fraction of values of $e_c^{l,k}$ has $e_c^{l,k} < -1$ or $e_c^{l,k} > 1$ which leads to the conclusion that ratings of the human face and the robotic head only rarely differ more than one slider unit.

Inspecting the mean values confirms the findings, since they are closely to 0.0. Larger values would imply a general shift between the human and robotic head. For instance, a high mean value at *Jaw* would indicate that the robotic head's yaw drop is always

category	mean	variance
Eye Brows	0.10	1.08
Eye Lids	0.28	1.40
Jaw	0.10	1.10
Lip Corners	0.15	0.88
overall	**0.16**	**1.11**

Table 3.1: Mean and variance in the rating difference of the training set with human face and corresponding robotic head.

perceived smaller than the human one, even if both are at maximum extend. Furthermore, the variances for *EyeBrows*, *Jaw* and *LipCorners* are all close to 1.0, which further strengthens the findings that ratings of the human face and the robotic head only rarely differ more than one slider unit in these categories. *EyeLids*, however, shows a larger variance which indicates, that this has been the most difficult category to rate by the participants. However,the variance is still less than 2.0 which in turn indicates that the overall status (eyes opened/eyes closed) still has been recognized correctly in general.

3.2.7 Summary

In this section, a method of transfer of mimicry behavior to artificial systems has been introduced. Starting with the benefits that a robot could gain from integrating this behavior, the various types of mimicry described in HHI literature have been analyzed based on the implicit communication channel that is mimicked. Furthermore, the influences of mimicry on parameters of social dynamics have been given.

Mimicry behavior has been modeled for facial and emotional mimicry. It is grounded in the AIM model, providing a structure for the mimicry process that is derived from research of these processes in infants. This model has found application in a demonstration setup, forming a system to automatically analyze facial expressions from a video stream or pictures and synthesize them on a robot based on the activation of facial action units.

Two experimental evaluations have been providing insights in the quality of the mirroring and the perceived effects on social parameters in HRI. In the first experiment, 20 participants have taken part in a user study. They rated the perceived activation of four groups of action units on a human face and the robotic face on a five point scale. Results have shown that the initial and the derived facial expressions are rated very similar, with a mean error between matching faces of 0.16 and a mean variation of 1.11. This implies that the transformation from the human facial expression to the robotic one achieves the desired effect of being perceived the same way, if the expression analysis and synthesis are combined in the proposed model. Results of the second experiment, described and analyzed in Section 3.3.6, support the hypothesis that the robot mimicry behavior during interaction influences the human partner in similar ways as literature describes in HHI.

3.3 System-theoretic Model of Smiling

> "*What sunshine is to flowers, smiles are to humanity. These are but trifles, to be sure; but scattered along life's pathway, the good they do is inconceivable.*"
>
> (Joseph Addison)

The smile is a special form of emotional expression. As humans we exercise smiling in many variations everyday, for example to display felt emotions, affection or due to cultural reasons [441]. Smiling is an important form of non-verbal communication in social interactions, especially for service providers. Studies show that the task performance of service providers is rated higher, if the provider displays friendliness through smiling. The rating is even higher if the smile appears to be authentic [301]. This seems to be embedded

in the expected work role of staff, which requires friendliness and smiling as expected behavior in service occupations, eventually leading to organizational and individual salient outcomes (i.e. benefits) [261]. Smiling subjects are also attributed greater degrees of sincerity, sociability, and competence [210]. Incorporating smile reactions in robotic or virtual agent designs could thus lead to a better performance in HRI. A key aspect for integration is to determine when to smile and how to attain some degree of authenticity. This means in other words to employ the concept of display rules, that is norms prescribing which expressions should be shown under specific social circumstances [215, 231, 441, 244]. Kraut and Johnston found closer association of smiling with social interaction than with positive emotional states [273]. This association is supported by Fridlund, explaining smiles in terms of the social motives involved rather than happiness [224].

In reference to the previous section on mimicry, smiling can also be regarded as part of emotional mimicry, which is dependent on social context [208]. It is therefore desirable to treat smiles not as an universally working display of happiness and professional friendliness, but within a model that takes into account the social context. The Zurich Model of Social Motivation (ZM) offers the embedding of a system-theoretic model of smiling (SMS) within a model of social motivation, connecting the smile display with the social context and motivations of the artificial agent.

In this section the psychological system-theoretic approach to artificially generate various types of psychologically plausible smiles in a virtual agent is introduced. The basis for this system-theoretic model is the ZM, which will be described first. It is followed by the derivation of the smiling model from the general model of social motivation. All seven computable types of smiles are described in detail and finally an experimental evaluation of the hypotheses, whether the presented smiles are distinguishable and can be correctly identified, is given. As a follow-up to the previous section on mimicry, the system-theoretic model is extended to react on emotional facial expressions. The extension is resulting in according emotional reactions of the artificial agent, forming a context dependent emotional mimicry system. The influence of the mimicked expressions on interaction with human partners is experimentally evaluated.

3.3.1 Problem Statement

In this section, two challenges have to be addressed: How to generate smiles and how to model context dependent mimicry, that resembles emotional mimicry mentioned in the previous section. For the emotional mimicry part, the same problem statement as in the section beforehand applies. For smiling, few works outside of social sciences have focused on its variations and their effect on a dyadic interaction, mainly with application to virtual avatars. Even if based on psychological models, they are based on empirical categorization models, treat smiles as isolated events or require the judgement of a human operator in a WOz setup. In comparison, the ZM with the SMS specialization features a fully computable cybernetic model, with a sound psychological base. Improvement of the state-of-the-art in this section is performed via refinement and extension of the SMS to generate seven smile variants and enable mimicry. Furthermore, the smile variants and the enhancement of the pure direct mimicking in form of a social motivation model, that incorporates a temporal and modulating influence on the mirroring, are evaluated.

3.3.2 Related Work

Smiles are common emblems in embodied conversational agents (ECAs) and social robotics as a sign of happiness. A number of works on virtual avatars has aimed on extending the underlying smile models to incorporate more variety and subtleties in smiling and the impact of it on an interaction. Tanguy [329] equipped a virtual agent with an amused smile, which resembles the smiling associated with happiness and a "fake" smile, which is an asymmetric smile superimposed on a sad facial expression of the upper facial region. Experimental results attribute virtual actors displaying the fake condition as more insincere. Ochs *et al.* [311] let users create smiling expressions on a virtual agent in a web application, which were concatenated in an empirical model for amused, polite and embarrassed smiles. The implementation of this model allows a virtual agent to display said smiles. Heylen [285] mentions an embarrassed and flirting smile in context of the gaze direction and its importance on the conversational structure. Rehm *et al.* [203] created two categories of a smile on a virtual avatar. One that expresses happiness and the other a non-happy basic emotion expression, masked by a fake smile. Experimental results show a perception in the difference of the smiles, which could not be explained by participants. The happy smile was also seen as more trustable, reliable, convincing, credible and with more certainty as compared to the "fake" condition. Another method for masking facial expressions was proposed by Niewiadomski and Pelachaud [280]. Eight facial segments can express different emotional states for the overall composition of "*complex facial expressions*". While smiles are not explicitly mentioned, the combination of a positive mouth expression with variations of the other segments leads to genuine or fake expressions of joy. A different approach in varying smiles is taken by Krumhuber *et al.* [258], adjusting the dynamic characteristic of the smile and not its composition. Experiments have been conducted with job interviews of participants with a virtual face. Results show that variation of the duration of the onset and offset of the smile at a constant overall duration of 4 seconds were perceived as more authentic for longer onsets and offsets. The perceived authenticity of the smile has also an impact on the social interaction, with the job appearing more positive and suitable for authentic smiles. The smile variations in these projects are either handcrafted to resemble the expected expression or the result of a masking process, with the masking area differing between implementations. In this thesis, the superimposition of the respective emotional expressions and the dynamics are embedded in the psychological model. Also, the occurrence of smiles is explained and determined by the basic psychological assumptions of the model.

While the smile as a signal can be found in many artificial agents, mimicking of smiles has only recently come to attention. A study on the mimicry behavior of the user, not the agent, has been conducted by Krämer *et al.* [239]. In this experiment, the artificial agent would either smile frequently, occasionally or not at all during interaction, independent of the user's behavior. Results show more smiles of the user for the smiling conditions of the agent. More smiling was also rated as less introverted. Bevacqua *et al.* [225] performed a WOz study with an ECA as a listener in a dyadic interaction. During listening to a story told by the participant, the agent would perform mimicking of smiles, show random smiles or none. Results showed significant differences for the feelings of engagement and frustration, whether the agent mimicked smiles during interaction or smiled not at all. In

both approaches, the artificial agent was not able to mimic the smile on its own. Related work in the previous Section 3.2 includes automatic mimicry, but either as part of the emotional state happiness [185] or as movement of facial elements related to smiling [332].

Mimicry of smiling can be regarded as the mimicry of an *act* that has an emotional context. In literature, emotional mimicry is often linked to empathy and similarity. A number of studies have already been conducted which employ empathy and similarity as factors in human-robot or human-computer interaction to manipulate the user's attitude towards an artificial agent. They can be categorized whether the artificial agents are used to express empathy [204, 281, 245, 247, 305, 249, 232] or induce it in the user [333, 217, 247].

Empathetic expressions by the agents are mostly utilized to enhance the user experience and thus provide a benefit to the user. Depending on the correct situation awareness and choice of expression, the empathetic reactions can be comforting to the user [204], build trust [249], enhance the system perception by the user [245, 247], enhance the subjective task performance [281] and meet user expectations [305]. The expression of empathy in a particular situation is either based on empirical data [245], a theoretical model [281] or both [247]. Visual [281], auditory [281, 232] or physiological [204] cues or training data from observations of human-human interaction [245] are used to evaluate the situation of the user and express an emotion that is similar to the user's estimated emotional state. Another approach is to induce empathy in the user via similarity of the agent. This can, for example, be achieved via facial mimicry [217] or character appearance [333].

3.3.3 The Zurich Model of Social Motivation

In this section, an overview on the Zurich Model of Social Motivation (ZM) and the necessary concepts to understand the derivation of the SMS are presented. For more details, the reader is referred to publications by Bischof [189, 295, 209, 188, 319] and Gubler [275, 324].

The ZM by Norbert Bischof has its roots in Bowlby's attachment theory [297], extending it to "*describe dynamic motivated behavior from a developmental, evolutionary and systems perspective*" [207]. Furthermore, the scope of the theory is widened from infants to adolescents and adults and from security seeking only to aspects of social motivation. Social motivation according to Gubler [246] *"refers to the fact that attachment is not an isolated system. It is inseparably intertwined with a larger body of motives controlling intraspecific transactions"*. Social motives besides security include social fear and social exploration (regarding the distance and contact with strangers), sexuality and the claim for autonomy and competence.

The ZM proposes three motivational subsystems, regulating security, arousal, and autonomy in homeostatic feedback control loops, as shown in Figure 3.6. It assumes that there are set points for security, arousal, and autonomy, from which follows, that for example both too little security as well as too much security are considered an undesirable state by the system and lead to behavior aiming at the regulation of the respective level. Emotions are considered as manifestations of the motives in the respective system to achieve homeostasis. Facial expressions for the extrema of the security, arousal and autonomy dimensions are given in Figure 3.7.

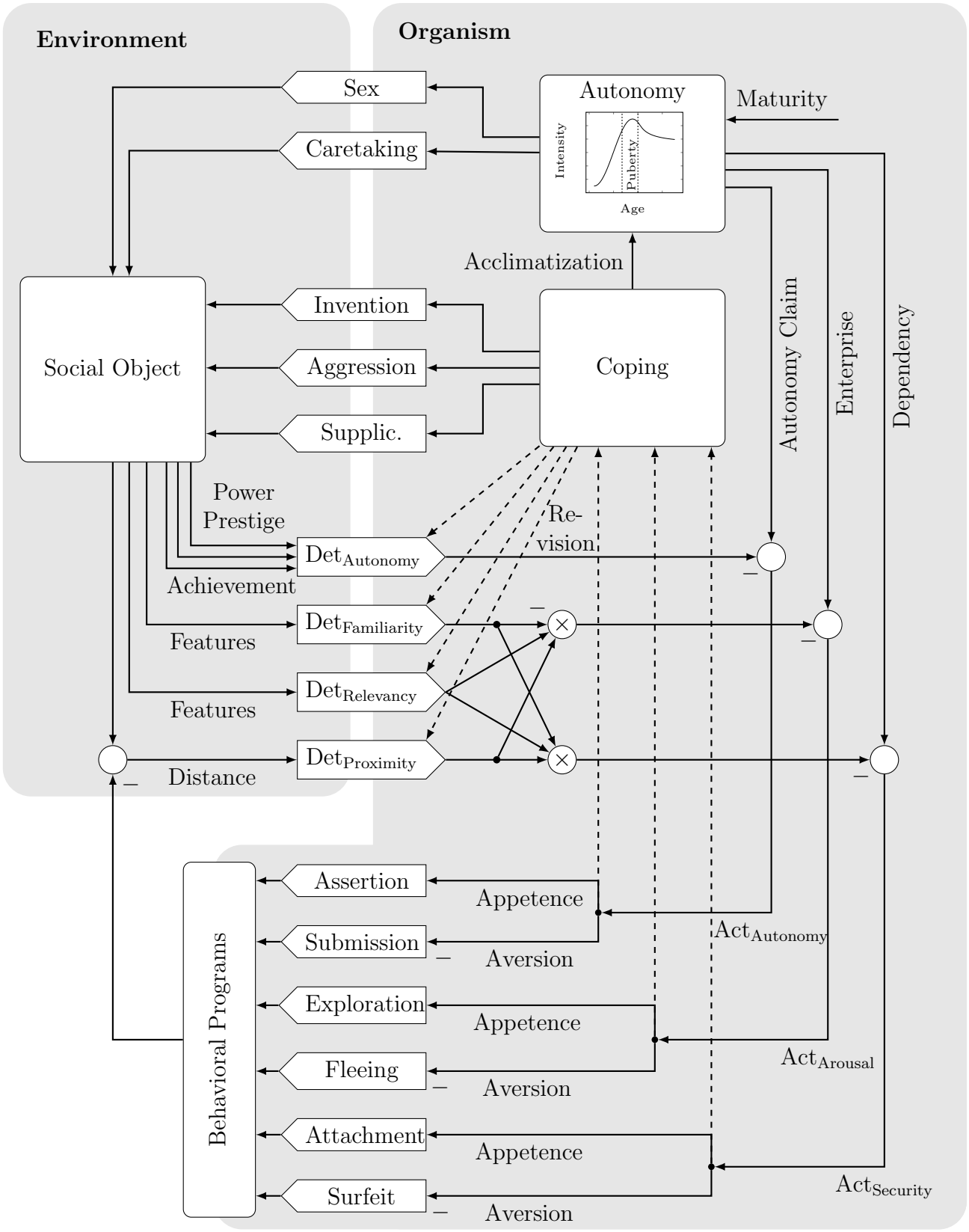


Figure 3.6: The Zurich Model of Social Motivation (ZM) in the version described in [189], with abbreviations Det = Detector, Act = Activation and Supplic. = Supplication.

(a) Grief **(b)** Curiosity **(c)** Aggressiveness

(d) Surfeit **(e)** Fear **(f)** Submission

Figure 3.7: Facial expressions for the extrema of the security, arousal and autonomy dimensions. The upper row (a-c) displays appetence, the lower row (d-f) aversion reactions.

Security is a formalism of the attachment theory, ensuring the contact to the care giver for infants [207]. It controls the behavior towards social objects in terms of familiarity. Familiarity of a "*conspecific*" [246] is a key concept for security, as it promises prosocial support, leading to an emotional response of feeling secure. As the capability of prosocial support is varying among group members (for instance due to maturity), relevancy is defined as a measure, scoring highest for adult, high-ranking conspecifics. In contrast, submissive behavior, younger age or replacement by transitional objects like a teddy bear decrease relevancy. Furthermore, security is distance dependent, with a higher distance reducing the degree of security. Bischof therefore proposes the existence of detectors capable of perceiving familiarity, distance and relevancy. Security, as well as arousal and autonomy, are homeostatic systems. The reference value for security is the degree of *dependency*. The more dependency is felt in the organism, the higher is the need for security and thus the proximity to familiar and relevant care takers. Dependency is a matter of maturity and decreases with age.

In case of the *need* for security, further called *appetence*, proximity seeking and attachment behavior is elicited to compensate the need. For the opposite case of *aversion*, surfeit leads to avoidance of the familiar social structures (often observable during puberty). Appetence and aversion are one-dimensional scalar quantities, with *tension* denoting the absolute magnitude. They result in "*motivational momentum*" [246], a locomotion motivated by the respective motive to cope for the need or surfeit, with direction and velocity determined by the *"incentive component"*. The incentive component is a vector pointing to sources or sinks in the social field of security, with the choice for sources or sinks governed by appetence or aversion.

Arousal is linked with security, as all three detectors of security are also used for arousal. The main difference is that the inverse of familiarity raises arousal, i.e. more arousal is elicited for less familiar stimuli. Familiar objects therefore can only increase arousal through unpredictable behavior, decreasing the familiarity through novelty, as in the exemplary game of peek-a-boo with infants. The reference value for the homeostasis of arousal is *enterprise*, which increases with age and maturity. Appetence of arousal leads to exploratory behavior, with an attraction to less familiar objects. Aversion on the other side leads to fearful behavior, if arousal is exceeding enterprise.

Autonomy is "*divided into three phylogenetically distinguishable motives*" [207]: power, prestige and achievement. Power describes the motive to assert oneself over others in a social group. This is mostly attributed to physical strength. Prestige on the other hand has to be earned through prosocial support for the group. The third motive describes the self-esteem, which is based on the judgements of one's own achievements. All three motives contribute to the reference value for autonomy homeostasis, denoted as the autonomy claim, which can be regarded as the amount of control over one's own life or the aspired success. The autonomy system takes a key position in the ZM, because its reference value influences the reference values of the security and the arousal system. Dependency and enterprise are connected via the autonomy claim, with enterprise related directly and dependency to the inverse of the autonomy claim. A special feature of the ZM is a process

called *"acute acclimatization"*, in which the autonomy claim – featuring as a reference value for the subsystems – is temporarily adapted to encounter the external disturbance of the homeostasis of one or more subsystems. This acclimatization plays a major role in situations, during which a barrier (in the sense of Lewin [248]) prevents the regulation of the disturbance via behavior. The autonomy claim is rising until early adolescence, with a high during puberty and a slight reduction afterwards. Bischof explains this timeline with separation desires from the family and a shift in sexuality, which is inhibited in the family. Later on autonomy is reduced to deal with demands in partnerships. The autonomy claim plays a crucial role in rank-order determination. Dynamics of hierarchical encounters, i.e. challenging the autonomy of a rival, is modeled through a cusp catastrophe. In this autonomy regulation, the challenge leads to a reciprocal increase in the autonomy claim. Depending on the respective levels of helplessness and reactance, both explained below, one opponent collapses in the autonomy claim due to the cusp catastrophe, leading to submission. Thus autonomy appetence produces *assertive* and aversion *submissive* behavior. More details on the cusp catastrophe can be found in [209] and [276].

There are further interconnections between the systems and a coping mechanism. Coping is a mechanism that gets in charge if motivational activation, e.g. security appetence, does not get reduced for a longer time [246]. The organism then applies a coping strategy, selecting an action to overcome the blocking influence. The action selection is dependent on temperament and the learning history, the repertoire can be grouped in aggression, invention and supplication strategies.

3.3.4 System-theoretic Model of Smiling

The system-theoretic account of smiling [319], is based on the reduced version of the ZM and can describe the effect of smiling based on the motivational and emotional state of a human or agent.

The core assumption of [319] in addition to [209] is that smile reactions are coupled to the first derivative of the set point of the autonomy system: whenever the autonomy claim is temporarily reduced, a smiling signal is issued. The different types of smiles relate to different situations which cause the autonomy claim to decrease. All these situations are based on the process of the previously described acute acclimatization. Therefore, according to Bischof [319] the smile reaction is coupled to the acute acclimatization of the autonomy claim, specifically to a reduction of the reference value w (see Figure 3.8 for the block diagram). The system-inputs z_{aut}, z_{ar} and z_{sec} can be modeled as external disturbances of the inner system state. Decreasing the distance of an organism to familiar persons, objects, or regions increases security z_{sec}, whereas decreasing the distance to strange, ambiguous, or discrepant objects or situations increases arousal z_{ar}. Both, security and arousal are also increased by the relevancy of the person, object or situation. They are fed into a negative feedback loop, resulting in the corresponding a_{aut}, a_{ar} and a_{sec} states. Initially the model parameters a_{aut}, a_{ar}, a_{sec} and s only describe basic emotional parameters. The motor implementation through the facial muscles is a delayed process, compared to the changes of the internal state. That is why the signals a^* and s^* are generated by first order delay terms:

$$s^* = \frac{K_0 s(t) + K_1(s(t) - s(t-1)) + T_s s^*(t-1)}{T_s + 1} \tag{3.1}$$

$$a^*(t) = \frac{K_a a(t) + T_a a^*(t-1)}{T_a + 1} \tag{3.2}$$

The non-amplified states are combined to a single offset a and dampened, as well as delayed, by the second order delay block with proportional gain. The combination with the raw value of the autonomy claim W results in the reference value of the autonomy claim w, which is fed back to the controlled values x_{aut}, x_{ar} and x_{sec}. Due to this feedback, the reference value adjusts to the actual value and not explicitly vice versa. Using (3.3), the smile reaction s can be derived from the change of the reference input w.

$$s = \begin{cases} 0 & \forall \Delta w \in]0, \infty[\\ -\Delta w & \forall \Delta w \in]-\infty, 0] \end{cases} \tag{3.3}$$

The upper part of the block diagram shows how the autonomy system can be regulated. Taking x_{aut} as a negative input, two variables *reactance* and *helplessness* can be generated, which influence the cusp control parameters λ and β in opposite directions. These parameters control a hysteresis, which is generated by a cusp catastrophe, and which results in the raw value of the autonomy claim W. Reactance denotes the behavior of increasing the autonomy claim during a challenge. Helplessness is a trait that is also increased during a challenge. The parametrization of reactance and helplessness determine the winner of a conflict in this model, with typically a reactant, less helpless individual triumphing over a more helpless or less reactant one. The concept of reactance and helplessness is later also described in the context of the superior and inferior smile, see Section 3.3.4. The outcome of a challenge is simulated via the cusp catastrophe, which can be described through the potential U. It is a polynomial of fourth order:

$$U = W^4 - 2\beta W^2 - \lambda W \tag{3.4}$$

Through an integrative feedback loop, the state W will tend towards a global or local minimum while $\nabla U(W) < 0$, with

$$\nabla U = -4W^3 + 2\beta W + \lambda\,. \tag{3.5}$$

The specific characteristic of a cusp catastrophe is the sudden change in W. For $\beta > 0$ and varying λ, the solution for W jumps from the stable solution to an alternate solution at the edges of the projection of the bifurcation surface on the $\beta\lambda$ plane. The resulting bifurcation surface and a 3D representation of the reaction surface are shown in Figure B.5 in Appendix B. Cusp parameters β and λ are both dependent on reactance and helplessness. While reactance and helplessness are both ≥ 0 and added up for β, helplessness is deducted for λ. λ has an additional base part λ_0, which is a function of a basic autonomy claim ω_0, varying with age and maturity.

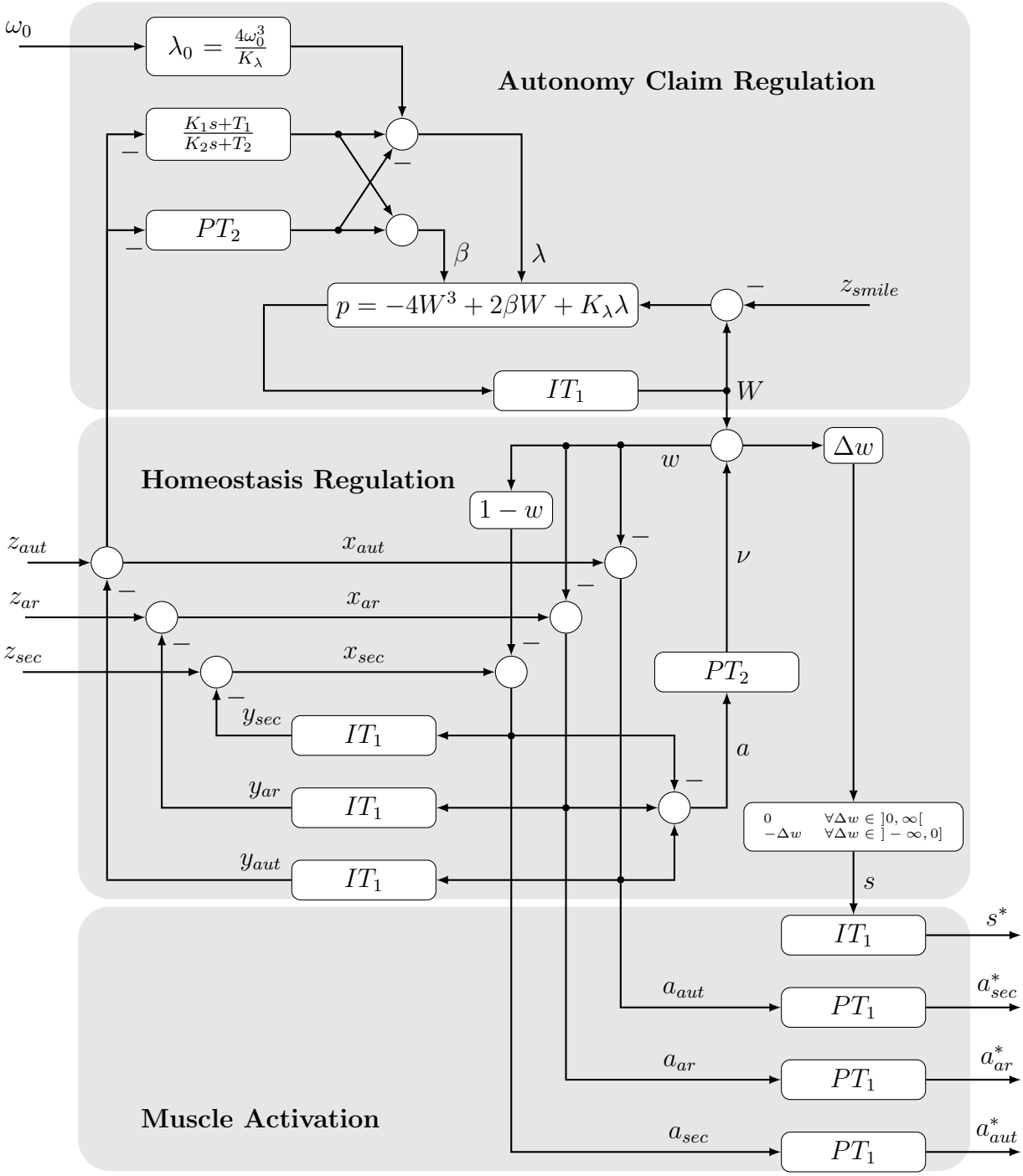


Figure 3.8: Block diagram of the SMS with aut=Autonomy, ar=Arousal and sec=Security.

Seven types of smiles

In the following section seven different types as described in [319] are introduced. Each type of smile is primarily based on one of the three state dimensions (security, arousal, or autonomy). In general, all seven types of smiling occur in contexts, which cause the setpoint value of autonomy to momentarily decrease and which have psychological semantic descriptions. To specifically show the output of the security and arousal system, the autonomy system is deactivated for all types of smiles except the superior or inferior smile.

Trustful smile The trustful smile originates in the security system and is based on security-appetence. The underlying mood is often interpreted as happiness. In humans this can be observed at the age of two to three months for the first time. During this period of life the reference value of the security system is on its ontogenetic maximum, resulting in a maximum dependency of the child. Therefore the autonomy claim is set low. Additionally the ability of acute acclimatization is very high. An exemplary situation for a trustful smile could look like this: we are observing a baby, who is not yet capable of distinguishing between strange or familiar, the ambience is still feeling safe. In this situation, the mother approaches the child. Before the mother is detectable, the ambience is of neutral security. By approaching the child the security is intensified and the baby's need for security decreases (a_{sec}), as well as the reference value w. Following (3.3), this results in a smile. As the mother now disappears again, the need for security drops slightly below a neutral mean and the reference value w rises. Being not capable of moving, y_{sec} is constantly zero. W is zero as well, as the baby has no autonomy claim, yet. In adults, corresponding situations occur, when familiar persons, such as friends or spouse are approaching. The progress of the situation is represented in Figure 3.9a.

Smile of relief The smile of relief originates also in the context of the security system. The child is now capable of moving, but still very dependent on the mother. At the beginning of the exemplary situation, mother and child are close-by, they are having the optimal distance (given through w). The mother departs and the child tries to follow, but is constrained by a barrier, which takes some time to be overtaken. When the barrier is overtaken, the child reaches it's mother and the reference value w drops, while the dependency rises. A smile reaction results out of the decrease of w. With the departure of the mother a need for security is building up, the child is security-appetent, which is neutralized with the arrival at the mother's position. The neutralization of a previous appetency of security triggered the smile. In adults, a corresponding situation would be a person in an unfamiliar situation, such as a job interview or being in a foreign country for the first time. Whenever the security rises due to approaching a familiar person or object, the security-appetence is reduced, resulting in a smile. The smile of relief situation is depicted in Figure 3.9b.

Embarrassed smile The last type of smiles that is based on security is the embarrassed smile. In contrast to the former smiles it originates in a security-aversion, that is the system encounters more security, than is necessary by the setpoint value. A descriptive situation, which leads to the embarrassed smile could be an aunt visiting an adolescent:

The visiting aunt hugs her nephew, who already has a strong autonomy claim. Thereby, the reference value w decreases, as the nephew is forced into a dependency, and smiles. The smile reactions results from a feeling of embarrassment, a certain helplessness during the shortfall of the optimal distance. The hugging can be seen as a barrier, immobilizing the nephew. When the aunt releases the child/teenager and leaves the intimate zone, the independency of the nephew rises to a normal level, the reference value acclimates. A corresponding situation for adults could be a drunken friend, invading the personal space and immobilizing a person by hugging him/her. Since the friend is no threat, the person bears the hug, smiling embarrassed. The course of the embarrassed smile is depicted in Figure 3.9c.

Anxious and surprised smile The fearful and surprised smiles are triggered by the arousal system. The actual arousal level influences the autonomy claim, when the arousal level reaches the set point, the acute acclimatization declines. The smile happens during acclimatization or the declination of the acclimatization, depending on whether the autonomy claim was boosted or reduced. Both types of smiles can be exemplified in one fictive situation: as the doorbell rings, a young boy runs to the door and opens. This new unexpected event gives raise to his initiative y_{ar} and autonomy claim. But the visitor looks unfamiliar and threatening, lowering the initiative, attenuating the tension and relaxing the acclimatization. In combination with the arousal-aversion this leads to a shy, insecure smile. Now the stranger reveals himself to be the beloved uncle, who has been traveling for a while. The boy's mood switches from arousal-aversion to appetence. He smiles due to surprise and relief. Over time the state acclimates and the smile disappears. Both anxious and surprised smile are combined in Figure 3.9d.

Superior and inferior smile These types of smiles occur during direct regulation of autonomy (compared to indirect regulation of autonomy in the previous cases). If two individuals meet in a conflicting situation, the dynamics differ from acute acclimatization, because none of the opponents should subdue sooner than necessary. The winner does not need to lose ground at all, in case of the other's surrender. The effect is even boosted if it is possible to raise the autonomy claim further and thus strengthen the tension during the conflict. This behavior is called reactance. The opposing behavior is known as helplessness. It is important to take the temporal sequence into account: first a challenge is answered with reactance, followed by helplessness if the reactive behavior was not successful.

For simulating the superior smile, an individual with a high reactance was modeled. It has a high influence on the outcome of the challenge, due to the fact that the agent with higher reactance usually dominates the conflict. When the opponent surrenders, noticeable in a sign change of a^*, the smile starts. The superior smile is first shaped by the expression of autonomy-appetence, the show-off behavior, and then is slightly accompanied by autonomy-aversion, just so that it does not appear as being submissive. It is given in Figure 3.9e. For the inferior smile, the agent tries to withstand the dominant agent in the beginning, but has a lower autonomy claim. Therefore, the agent's W value drops rapidly, the opponent wins. Due to the rapid drop in the autonomy claim, a significant smile reaction is shown. Additionally, the gaze and head are lowered in submission.

Mimicry

The ZM and the SMS do not explicitly take into account facial expressions of the interaction partner. However, the detectors of the ZM, see Figure 3.6, can be extended to perceive the disturbing influence of these expressions on the internal state. The previously employed facial expression detection for mimicry is based on the PAD dimensions. In reference to Section 2.2.2, arousal in the PAD framework is defined in terms of the level of mental alertness and physical activity or the judgement of high-low stimulus activity. This definition matches the description of the ZM arousal system. Therefore, the detected PAD arousal stimuli $ar_{PAD} \in [-1, 1]$ are linked to the arousal system:

$$z_{ar} = (1 - \underbrace{\prod_{i=1}^{N}(1 - p_i)}_{joint\ potency})(1 - \underbrace{(\frac{\sum_{i=1}^{N} f_i p_i}{\sum_{i=1}^{N} p_i})}_{joint\ familiarity}) + \sum_{i=1}^{N} ar_{PAD,i} p_i \,, \qquad z_{ar} \in [0, 1] \tag{3.6}$$

for a number of N social objects with familiarity f_i and potency p_i, dependent on the respective relevancy r_i of the social object and the distance d_i, with

$$p_i = r_i \, \frac{-1}{R} e^{\frac{-1}{R} d_i} \,. \tag{3.7}$$

The expression (or lack of) arousal thus leads to an increase or decrease of arousal in the artificial agent, dependent on the familiarity, relevancy and distance of the displaying person. The display of dominance in the PAD system is the result of "*a feeling of control and influence over one's surroundings and others versus feeling controlled or influenced by situations and others*" [102]. This is in line with the power motive of the ZM in the autonomy system. A high dominance in the PAD dimensions of the interaction partner, such as in the state of anger, can therefore be regarded as a challenge to the autonomy system. The detected power motive is seen as direct disturbance z_{aut} to the system, thus a positive dominance value $dom_{PAD} \geq 0$ is directly added to z_{aut}. A special case is the detection of smiles. It is not treated as a sign of pleasure, but follows the paradigm of being more of a social signal than related to the emotional state. It is therefore considered as an appeasing signal, lowering the raw value of the autonomy claim W, see Figure 3.8. The subtraction of the perceived smile $z_{smile} \in [0, 1]$ from W leads to two effects: First, the lowered autonomy claim during a challenging situation leads to an easing of tension in the conflict. Second, the lowering of the autonomy claim produces a smile reaction of the system itself, answering a smile with an (attenuated) smile.

The mimic reaction of the artificial agent, that occurs from the regulation of a disturbance in the related dimension, results in a mimicry effect. It can be categorized as emotional mimicry, as the abstraction of the facial display is evaluated according to the emotional dimensions and is basis of the mirroring in the respective ZM subsystem. Arousal detection in expressions such as surprise leads to increased arousal in the agent, with the resulting facial expressions showing a degree of curiosity. A dominance challenge is answered with according dominance / autonomy claim gestures and expressions. Last, a smile leads through the decrease of the autonomy claim directly to a smile response.

3.3.5 Application

For the application, the SMS is the integral part of an online simulation of the internal state of an artificial agent. In this simulation, the agent is placed in a virtual 2D world. The artificial agent is simulated along with conspecifics, i.e. other agents that represent social objects. For interaction with humans, the human partner takes the place of one of the social objects in the virtual world. The parameters and behavior of the social object then reflect the perceived parameters and actions of the human. The model allows for motion of the agent in the world as a coping mechanism. For use in interaction with static setups, the simulated agent is immobilized, so that distance regulation can only be performed by the interaction partner.

Expression Synthesis

The expressions generated by the system-theoretic model can be visualized with the previously introduced robot heads, see Section 2.3.7, or a virtual avatar. For both cases the outputs a^* and s^* of the model need to be mapped to the Facial Action Coding System, with an intensity mapping given in Table B.3. These AU activations are further mapped to the respective motor activations of the robot, allowing the display of the ZM facial expressions on the robots.

Additionally, a virtual avatar can be used to display the generated expressions [241]. Animation of the avatar is performed with the Open Source software *Xface*[3], which displays facial avatars based on the Motion Picture Experts Group (MPEG)-4 standard [308, 200]. MPEG-4 standardizes a parametric model of the face for animation, with feature points specifying the shape of a neutral face and Facial Animation Points (FAP) parameterizing the deformation. With the established a^* and s^* mapping to FACS, instead of a mapping to motor activations a mapping to FAP is needed. A solution is presented in Section B.12. Facial avatars with varying appearance are generated with the software *FaceGen*[4]. The use of a virtual avatar provides a higher degree of realism than the robotic heads, which is assumed to be beneficial for the classification of the different types of smiles. This is due to the higher number of actuated facial elements (64 controllable elements, including head posture), a more fine grained resolution in simulated muscle movements, a reciprocal effect on the skin for muscle movements and a better resemblance of the original human face. Figure 3.7 shows the maxima of the ZM based facial expressions animated with the avatar.

Expression Analysis

The facial expression analysis software that is used to generate the facial analysis input to the system is again provided by Dr. Christoph Mayer. The PAD dimensions used for the system inputs are derived from evaluation of the detected FACS activations, the reader is referred to [279] for details. The analysis software also provides the distance of the user to the agent or rather camera, which is used for the potency determination. In

[3] http://xface.fbk.eu
[4] Singular Inversions Inc.

this implementation, the determination of the degree of familiarity is not performed, but assumed to be 0.8 for all users across experiments. This familiarity indicates a well known person, but not as familiar and close as the mother, as it would be expected for a personal agent in a long-term service relation. The choice of familiarity level ensures a friendly and open attitude. An automatic detection of familiarity in future systems could be realized through facial recognition software. Accordingly, relevancy is fixed to the maximum of 1.0, as the user is the most relevant person in a service relationship. It should be noted that these settings are only fixed for the interaction with users and not for the simulation of the specific smile variations in Section 3.3.4.

3.3.6 Experimental Evaluation

With the implementation of the ZM, artificial agents are now able to generate smile variations based on the social context. Even though the model has a sound psychological basis and supportive empirical evidence, the quality of smile variations has to be evaluated in terms of being distinguishable and classifiable. This is addressed in the first of the following experiments. The second experiment deals with the question, whether the application of mimicry has an effect on interaction or not. The modified system-theoretic model of smiling and the extension to mimicry take the interaction capabilities of virtual agents and robots one step further in comparison to the previously described facial mimicry, if the effects mentioned in literature persist. Therefore, an experiment is designed to measure the effects of both approaches on a dyadic interaction.

Perception of smiles

Hypothesis 3.2. *All seven smile variations generated with the system-theoretic model are distinguishable and can be connected to the correct categories.*

In this experiment 126 videos of the virtual avatar were shown to each participant, with 18 videos showing one type of smile respectively. The 18 videos consisted of 18 different faces with the same type of smile. Three male and three female faces were used, which were varied by age (young, adult, old), so that the setup was factorized by 2x3x3. The faces were generated using the commercial software FaceGen. Each video had a length of 79 to 80 frames and was played with a speed of 15 frames per second. All stimuli were shown in a powerpoint presentation on a gray background. 20 subjects took part in the experiment, 10 of them being male and 10 female, all of them students.

In the first part of the experiment, starting with an instructional slide, the subjects were asked to answer the following question for each video: "*What happens in this situation and how does the person feel?*". A text-field for the answers was provided on each slide. Male and female faces were shown alternating, while half of the participants started with a male and half of them with a female face. Every type of smile was shown once.

In the second part, starting again with an instructional slide, the subjects had to answer this question: "*Which kind of smile is shown?*" The participants could replay the video as often as they wanted to and had to select one of the given answers: "*embarrassed*", "*superior*", "*fearful*", "*relieved*", "*inferior*", "*trustful*" and "*surprised*" (to be precise:

their german analog). In this part 126 videos had to be matched to the respective answers. For each participant the order of the videos was randomized, the answers and the number of replays for the respective videos were recorded. Mean time for the duration of the experiment was approximately 50 minutes.

Context Description Due to the less restrictive task in this part of the experiment, the descriptions of the shown situation had a great variation. Some participants described extensively their observations or what could have happened, but their descriptions lacked the emotional state of the agent. Others only evaluated part of the shown situation.

The results show that the fearful and the embarrassed smile are not seen as such. About a quarter of the subjects described a situation that includes relief for the shown smile of relief. In case of the inferior smile, answers were given with an inferior context, but subjects often mixed it up with dominance. Best results were achieved for the trustful and surprised smiles. About half of the subjects described a situation in which the agent meets a familiar person, matching the exemplary situation. Nearly all participants identified a surprising situation.

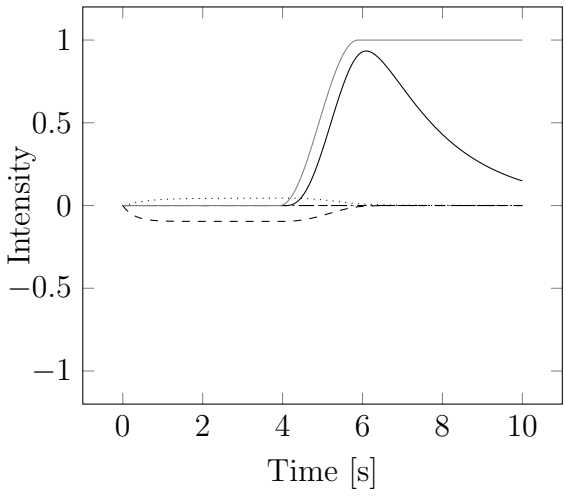
Category Matching The given answers were evaluated by the relative frequency of answers in the form of a confusion-matrix. The number of the respective answers was divided by the total number of shown stimuli of the respective type of smile. For each participant a confusion-matrix was computed and a mean over all participants calculated, see Table 3.2. The columns show the real category and the rows the experimental results. The last column includes the deviation in a category determined by the subjects, the last row describes the error rate of stimuli belonging to that category. The division of the correctly identified stimuli by the total number of stimuli results in the accuracy of classification. In this experiment the accuracy of classification is 36.8% and thus highly significant against the guessing rate of 14.29% ($t(19) = 10.6504, p = 0.00$). It should be noted that there were great variations in classification between the participants. While some of them had an accuracy of classification of only 19.84%, others managed to classify up to 50% correctly.

One can see that the embarrassed smile was often not classified as such. Mostly it was identified as a superior smile or even fearful. However, the superior smile was classified as inferior, fearful, and relieved, with superior on fourth place. The identification of the fearful smile worked better due to the fact that fearful was the most common answer. Nevertheless, it was confused with relieved, embarrassed and surprised. The smile of relief was mixed up frequently, especially with the fearful smile or the smile of relief. The answers embarrassed and surprised were given often, too. The inferior smile was classified better, only being confused with the also autonomy-based superior smile. With an accuracy of classification of 51.12%, the trustful smile was identified clearly. The best result was achieved with the surprised smile, having a classification rate of 69.83%.

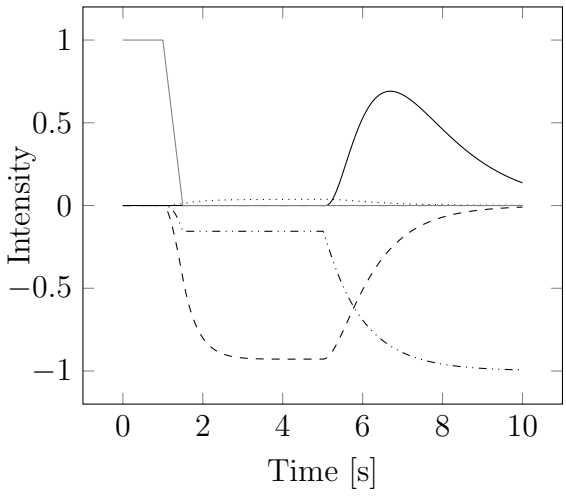
Discussion

The results show that for some types the classification works well, specifically for the trustful, the surprised and the inferior smile. In contrast, the embarrassed and the superior smile were classified worst. In the latter case, some participants never gave the correct

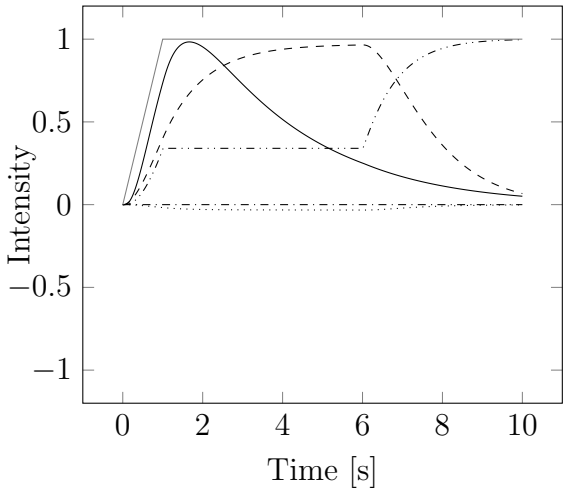
answer, whereas others classified these types 78% correctly. The anxious and the trustful smiles are in the midrange, with similar frequencies of responses, leading to the conclusion that they have been mixed up frequently. Almost all subjects matched particular smiling sequences to specific answers. Although these answers were not always the expected response, this matching shows that the different types are distinguishable. Therefore, hypothesis 3.2 is confirmed in terms of the smiles being distinguishable, but only partially in terms of the correct categorization. The errors in classification could result from different reasons: Smiling is quite often connected to positive emotions, possibly leading to a worse classification rate for types based on aversive states. The provided contexts are indicators, as for anxious types often a kind of relief was mentioned. A second reason is that the alternatives for answering were not described in detail. The meaning of the answers was subject to each participants interpretation. This could have led to different interpretations than expected, resulting in confusions. Some of the misclassifications are due to weaknesses in the animation. Although the MPEG-4 standard provides a variety of animation options, some of the needed AUs could not be implemented, due to missing corresponding FAP. This decreases the quality of animation, as well as the lack of wrinkle simulation. Moreover the focus of the animated head does not stay on the viewer if raising or lowering the head, giving them the impression that the agent looks away. Some participants stated that this behavior seems to be some kind of uncertainty, and thus is interpreted as embarrassment. Therefore, the effect of superiority was counteracted. Further, correct identification is made difficult by the fact that all situations are displayed with a duration of 5 seconds. Especially conflicts about autonomy or the visiting aunt scenario take longer. However, this was necessary to ensure consistent conditions throughout the experiment.



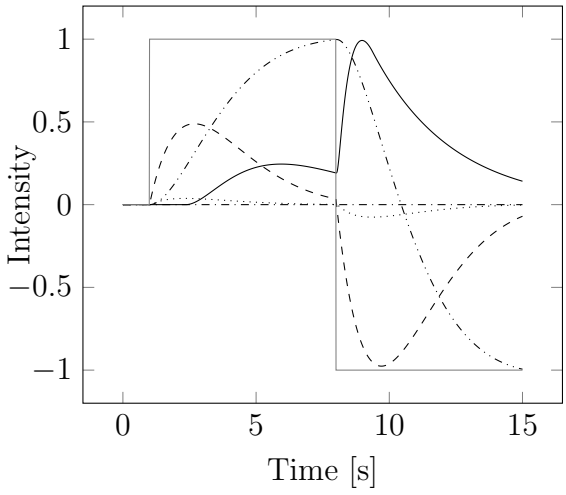
(a) Trustful smile



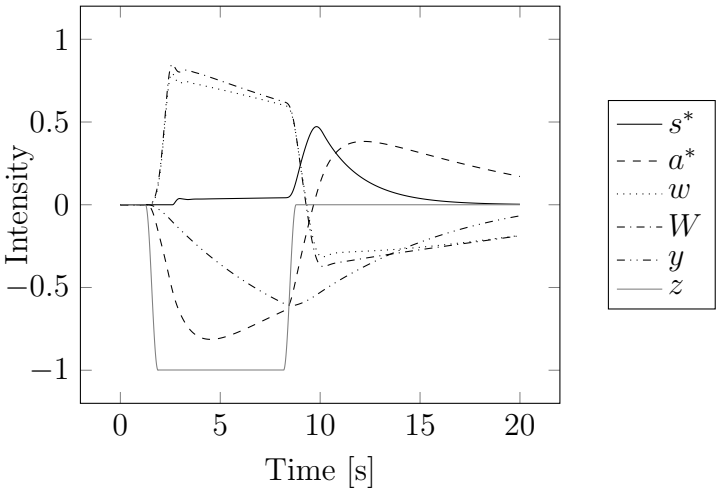
(b) Smile of relief



(c) Embarrassed smile



(d) Anxious and surprised smile



(e) Superior smile

Figure 3.9: Progress of the states of the SMS for the smile variations. Semantic descriptions of the respective situations are given in Section 3.3.4.

Real \ Exp.	embarr.	superior	anxious	relieved	inferior	trustful	surprised	Error
embarrassed	0.2479	0.0978	0.1737	0.2151	0.1056	0.0419	0.0196	72.45%
superior	0.2730	0.1453	0.0448	0.0251	0.1722	0.2291	0.0168	84.00%
anxious	0.1448	0.2430	0.3501	0.2765	0.0528	0.0698	0.0391	70.31%
relieved	0.0836	0.1760	0.2241	0.2374	0.2000	0.0587	0.0810	77.63%
inferior	0.1170	0.3045	0.0280	0.0363	0.3861	0.0894	0.0419	61.39%
trustful	0.1086	0.0251	0.0224	0.0475	0.0833	0.5112	0.1034	43.34%
surprised	0.0251	0.0084	0.1569	0.1620	0.0556	0.0556	0.6983	33.51%
Error	75.21%	85.47%	64.99%	80.45%	61.39%	48.88%	30.17%	**Real** \ **Category**

Table 3.2: Confusion-matrix of the relative frequency of answers. Highlights show the most often selected answers per category.

Social Effects of Mimicry

Hypothesis 3.3. *Mimicry influences the social aspects as described in Section 3.2.3. The influence is measurable in the five key concepts of HRI and the user acceptance concepts, with addition of empathy and subjective performance metrics.*

Hypothesis 3.4. *Mimicry of purely facial movements and mimicry with social context have a different influence on the five key concepts and the user acceptance, including empathy and subjective performance.*

In order to evaluate whether facial mimicry improves HRI or not, a second experiment is designed. The previously introduced robotic system EDDIE is set up to a communicative task, playing Akinator with a participant and guessing the thought-of person. To create a dialog with the robot head, text-to-speech is used to present questions of Akinator acoustically to the participant and speech recognition is utilized to determine the participant's answers. For further information on the experimental setup, Akinator and the dialog modules, see Sections B.7, B.5 and B.6 in the appendix.

During the interaction, EDDIE acts according one of the three experiment conditions:

- **Neutral**: EDDIE displays no facial expressions
- **Mirror**: EDDIE displays the subject's facial expressions
- **SMS**: EDDIE displays facial expressions according to its internal system-theoretic model of smiling, indirectly mirroring the subject's expression.

Subjects are divided in three groups depending on the applied condition. After the interaction, each subject fills in a computer-randomized questionnaire, which consists of two different parts that can be analyzed independently.

The first part consists of five selected constructs based on a *"limited model for studies on social abilities or social presence"* out of a toolkit for measuring user acceptance of social robots [291]. These constructs are adapted to the requirements of experimental setting and kept constant with regard to a consistent number of items, i.e. four questions for each construct. Additionally, these five constructs are enhanced by two more constructs developed by Barbara Kühnlenz (Gonsior)[5], which are proposed to measure the induced scope of empathy towards a robot, and the subjective system-performance perceived by the user [260]. These additional constructs are to reveal supposed interrelations to the other constructs on user acceptance and are therefore proposed to enhance this existing toolkit.

The second part of the applied questionnaire consists of the *"godspeed"* questionnaires [331] to evaluate the *"five key concepts of HRI"*: anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety.

Hence, the questionnaire evaluates the interaction on the dimensions of *empathy* and *subjective performance* as proposed extension of *user acceptance*, and the *key concepts* of the godspeed questionnaires.

[5]Institute of Automatic Control Engineering (LSR), Department of Electrical Engineering and Information Technology, Technische Universität München

User acceptance Heerink *et al.* [291] extended the Unified Theory of Acceptance and Use of Technology (UTAUT) model [317] by several constructs in order to adapt this model to the specific requirements of evaluating social robots. Given experimentally validated interrelations between several constructs, the five selected constructs include:

- *Trust*: The belief that the system performs with personal integrity and reliability.
- *Perceived Sociability*: The perceived ability of the system to perform sociable behavior.
- *Social Presence*: The experience of sensing a social entity when interacting with the system.
- *Perceived Enjoyment*: Feelings of joy or pleasure associated by the user with the use of the system.
- *Intention to Use*: The outspoken intention to use the system over a longer period in time.

The questionnaire evaluates each construct by four different statements, as presented in Table B.2. In order to reduce acquiescence bias, some items are negated and thus invert the scale. Subjects rate the randomized statements on five-item Likert scales from one (strongly disagree) to five (strongly agree). As the statements for user acceptance and their constructs are independent from the system performance this questionnaire is not divided into different paths if EDDIE was successful (a) or not (b) in guessing the thought-of person.

Empathy and Subjective Performance For measuring both constructs, the scope of induced empathy on the one hand, and subjective performance on the other hand, this extension to the UTAUT questionnaire is divided into two different paths depending on objective task performance, i.e. if EDDIE was successful (a) or not (b) in guessing the thought-of person. Thus, subjective performance can be compared to objective performance in order to draw conclusions on possible interrelations due to the scope of induced empathy. Therefore, subjects are asked to respond to different statements including positive, negative or inverted formulations for sharing happiness or sadness with EDDIE corresponding to the task-success or -failure of of it as shown in Table B.2. Users can reply to these statements consistent to the UTAUT questions on a five-item Likert scale from one (strongly disagree) to five (strongly agree), with inverted scales for negatively formulated questionnaire items.

Godspeed key concepts *"A series of questionnaires to measure the user's perception of robots"*[6] combines five consistent and validated questionnaires based on 5-point semantic differential scales as a standardized metric for the *"five key concepts in HRI"* [331]:

- *Anthropomorphism*: the attribution of human form, characteristics or behavior is rated on five semantic differentials.

[6]Open source version, see http://www.bartneck.de/2008/03/11/the-godspeed-questionnaire-series

- *Animacy*: the user's impression of lifelikeness and intentional behavior is rated on six semantic differentials.
- *Likability*: positive impressions of the robot are rated on five semantic differentials.
- *Perceived Intelligence*: users rate the perceived competence and intelligence of the robot on five semantic differentials.
- *Perceived Safety*: the safety of the interaction is assessed on three semantic differentials.

As recommended, the items are randomized to hide the different concepts and mask the intention. An exception are the three questions of *Perceived Safety* constantly set up the beginning of the overall questionnaire. They are measuring the emotional state directly after the interaction with EDDIE to avoid changes of the subject's emotional state while filling the questionnaire.

The study aims to unveil if mirroring improves HRI regarding the five key concepts anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety [331] on the one hand, and how mirroring influences user acceptance [291] on the other hand. Accordingly, the assumed interrelations with and between empathy and subjective performance of the system are investigated.

A key assumption for this experiment is, that the facial expressions of the robot are interpreted correctly by the human and perceived as matching expressions for the mirroring. This assumption is supported by the findings given in the previous experiment description, see Section 3.2.6.

Results can be deduced from the experimental evaluation including 55 subjects (40 male and 15 female, between 21 to 60 years with an average age of 28.8). The distribution of the subjects over experimental conditions was 13 for **Neutral**, 17 for **SMS** and 25 for **Mimicry**.

Regarding reliability, coefficients of internal consistency are calculated with Cronbach's α for the items of the novel constructs on *Empathy* and *Subjective Performance*. As a solid construct should create an Cronbach's $\alpha > .70$ all items of both novel constructs showed good reliability with Cronbach's $\alpha = .82$ for *Empathy*, and Cronbach's $\alpha > .86$ for *Subjective Performance* [260]. Since the selected constructs for user acceptance and of the Godspeed questionnaires were previously evaluated [291, 331] reliability and internal consistency are assumed.

Significance level for all performed tests was set to $\alpha = .05$. According to the results of Kolmogorov-Smirnov tests, normal distribution could be accepted for the total scores of all constructs, except *Perceived Enjoyment*. Thus, this construct has to be analyzed non-parametrically. Parametric comparisons and correlations are performed for all other constructs [260].

An ANOVA revealed significant differences between the conditions for *Empathy* ($F = 5.35, p = .008$), *Subjective Performance* ($F = 6.48, p = .003$), *Trust* ($F = 4.47, p = .016$), and *Likeability* ($F = 3.73, p = .031$). Thus, a post-hoc analysis could be conducted between the conditions. Accordingly, the assumed significance level was divided by three und thus adjusted to $\alpha = .016$. Paired t-tests revealed significant differences between

Neutral and **SMS** conditions for *Empathy* ($t = -3.01, p = .007$), *Subjective Performance* ($t = -3.51, p = .002$), and *Trust* ($t = -3.30, p = .003$). Paired t-tests revealed one significant difference between the **Neutral**- and **Mimicry** condition for the godspeed construct *Likeability* ($t = -2.03, p = .062$) and no significant differences were found between the conditions of **SMS** and **Mimicry** due to the α-value adjustment. Means, total scores and standard deviations of the five constructs on user acceptance by Heerink [291], and the two additionally introduced constructs on *Empathy* and *Subjective Performance* are displayed in Table 3.3.

Construct	**Condition**		
	Neutral	Mimicry	SMS
Empathy	3.1 (1.3)	3.7 (1.1)	4.4 (0.8)
Subjective Performance	2.8 (1.2)	3.4 (1.0)	4.1 (0.9)
Trust	3.0 (0.6)	3.3 (0.8)	3.7 (0.5)
Perceived Sociability	3.2 (1.0)	3.6 (1.0)	3.9 (0.7)
Social Presence	2.8 (0.6)	2.8 (0.7)	2.9 (0.7)
Perceived Enjoyment	2.8 (1.4)	3.9 (1.2)	4.2 (0.7)
Intention to Use	3.0 (1.3)	3.5 (1.0)	3.9 (1.0)
Total Score	2.9 (1.1)	3.5 (1.0)	3.9 (0.8)

Table 3.3: User Acceptance: Mean ratings (rated on Likert scales from 1 = strongly disagree to 5 = strongly agree) with standard deviations (in brackets) of each construct and total scores within conditions. Highlights show best results for the respective construct.

Mean values and total scores for the five key concepts in HRI, as derived from the godspeed questionnaires, are depicted in Table 3.4.

Construct	**Condition**		
	Neutral	Mimicry	SMS
Perceived Safety	3.9 (0.8)	3.6 (0.6)	3.7 (0.5)
Anthropomorphism	2.6 (0.6)	2.8 (0.5)	2.8 (0.7)
Animacy	3.1 (0.7)	3.3 (0.4)	3.3 (0.7)
Likeability	3.5 (1.1)	4.1 (0.5)	4.1 (0.7)
Perceived Intelligence	3.5 (0.8)	3.8 (0.5)	3.9 (0.5)
Total Score	1.1 (0.7)	3.5 (0.5)	3.6 (0.6)

Table 3.4: Key Concepts (Godspeed): Mean ratings (rated on Likert scales from 1 = strongly disagree to 5 = strongly agree) with standard deviations (in brackets) of each construct and total scores within conditions. Highlights show best results for the respective construct.

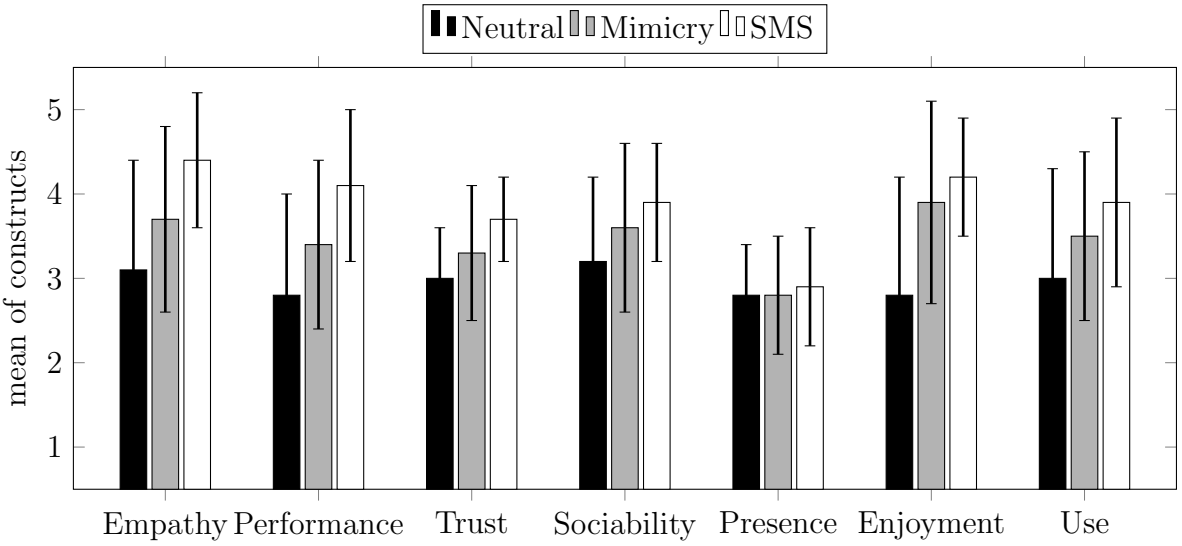


Figure 3.10: User Acceptance: Mean ratings (rated on Likert scales from 1 = strongly disagree to 5 = strongly agree) with standard deviations of each construct

Correlation analysis focused on the five selected constructs on user acceptance, along with the added constructs on *Empathy* and *Subjective Performance*. Correlation coefficients led to the finding that all constructs show significant correlations to each other ($p < .001$), except for *Social Presence* which only correlates significantly to *Trust* ($r = .36, p = .007$).

Discussion

In general, results support hypothesis 3.3 by showing a trend towards a better rating of the mimicry condition compared to the neutral condition. Results support also hypothesis 3.4, with the SMS being rated even better than mimicry in most instances. This underlines the importance of social factors to be considered for further refinement of how mirroring should be performed. Of the initially described influences mentioned in literature, direct measures for trust, empathy and liking are included as constructs in the questionnaires. The results for trust are ambivalent: The user acceptance measure for trust shows better results for both the mimicry and SMS condition, as expected. However, the related perceived safety, leading to trust in machines according to Muir [321], decreases in both conditions. Liking or likability is higher through mimicry and thus indicates that the amount and way of mimicking did not violate social conventions. Although the type of mimicry was facial and not emotional mimicry, users felt more empathy in the mimicry condition and rated for a better subjective performance. Again, this is in line with effects described in literature, with better performance ratings of human mimickers. Overall results in the user acceptance questionnaire were better for the mimicry condition than neutral, showing a positive influence on the interaction. But results for the SMS condition in all constructs of the user acceptance questionnaire showed even higher ratings than the other conditions, indicating the importance of the social context information. No direct measurements of rapport and affiliation were performed. Gains in perceived enjoyment and intention to use

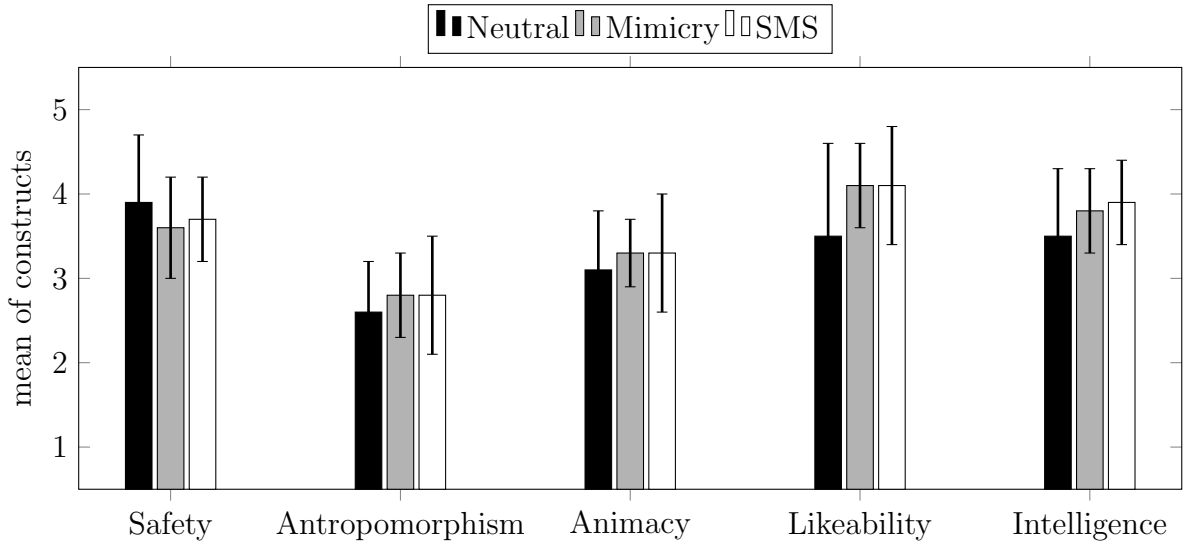


Figure 3.11: Key Concepts (Godspeed): Mean ratings (rated on Likert scales from 1 = strongly disagree to 5 = strongly agree) with standard deviations of each construct

through mimicry and SMS hint at positive effects on rapport. The constructs of sociability, social presence and anthropomorphism, which relate to the perception of the robot as social actor and "group member", indicate little effect of mimicry and only a slightly better effect of SMS on affiliation for this robot. Since EDDIE has a very machine-like appearance, it is possible that this may have dominating effects on the construct of social presence: Mean values show no noticeable increase within this construct, and no correlations could be found besides the construct on trust. However, results indicate that social presence, which is very much bound to being humanlike, is not crucial in order to induce empathy [260].

3.3.7 Summary

In this section the system-theoretic approach to artificially generate various types of psychologically plausible smiles in an artificial agent has been introduced. The basis for this system-theoretic model is the ZM, which has been described, specifically in its three motivational dimensions security, arousal and autonomy. It has been followed by the derivation of the smiling model from the general model of social motivation, with focus on the description of the smile reaction and cusp catastrophe in the autonomy system. All seven types of smiles that can be computed with the SMS have been described in detail and finally an experimental evaluation of the hypotheses, whether the presented smiles are distinguishable and can be correctly identified, has been given. Results have shown that the smiles have been distinguishable, but only partially categorized correctly.

As a follow-up to the previous section on mimicry, the SMS has been extended to react on emotional facial expressions. The extension has been resulting in according emotional reactions of the artificial agent, forming a context dependent emotional mimicry system, that is also able to display context sensitive smiles.

The influence of the extended SMS has been evaluated in terms of user acceptance, with special attributes empathy and subjective performance, and the key concepts of HRI. It has been compared to the facial mimicry system of the previous section and a neutral, non-emotional response of the robot. In general, results have been showing a trend towards a better rating of the facial mimicry condition compared to the neutral condition. Results have also been indicating better ratings of the SMS compared to facial mimicry in most instances. This underlines the importance of social factors to be considered for further refinement of how mirroring should be performed.

4 Transfer of a Biomimetic Lateral Line System to an Underwater Robot

While in the previous chapters 2 and 3 the task-dependent requirements on the robotic system were the starting point to investigate biological systems for a solution, this approach is based on the research on the biological system - in this case the lateral line system of fish - and applies the discovered capabilities and functionalities to a robot, but without an initial problem statement on the technical side. This is called the bottom-up strategy [339] and resembles the bionic process described by Nachtigall [345]. In this process a biomimetic sensor, based on the lateral line organ of fish, is developed and integrated in an autonomous underwater robot specifically designed for that purpose. Methods for modeling the sensor behavior and environmental stimuli are introduced, as well as algorithms to detect obstacles or reconstruct the environment from these stimuli. Simulated and experimental evaluations show the functionality and capabilities of the new artificial lateral line system and the robot equipped with it.

4.1 Introduction

Fish are able to locate obstacles and avoid them under poor visual conditions, or even if completely blind [346, 423]. The objects are perceived with the lateral line organ, which is distributed along the fish's body, responding to the movement of the water relative to the fish's skin [369, 376]. The presence of objects leads to an alteration of the flow field around the fish, which creates a *"hydrodynamic image"* [389] of the surroundings on the fish's body. From behavioral experiments some of the tasks the lateral system is involved in and some of the features of stimuli that are reconstructed by the lateral line system have been identified. Schooling for example can be done solely by perception of the flow fields of neighboring fish [362]. Mottled sculpin respond to the presentation of an oscillating sphere – the lowest order representation of the flow field of prey – with hunting behavior and a strike towards the dipole source [391, 413, 337, 375]. Studies on the blind cave form of *Astyanax mexicanus* and the closely related *Astyanax jordani* (previously known as *Anoptichthys jordani*) show that these fish are able to detect, avoid and also discriminate solid objects, if gliding past or towards them at close distance, [399, 355, 357]. Similar abilities have been observed in experiments carried out with goldfish [336]. For the blind Mexican cave fish, concluding from behavioral experiments [424, 383, 355], there is no doubt about its elaborate capabilities sensing its environment by usage of the lateral line system. Although it is not yet fully established what the capabilities of a lateral line system are and what tasks it can be used for, the example of the blind Mexican cave fish shows that it is obviously possible to make vital decisions solely based on information conveyed by the surrounding fluid motion.

Any object in the near surrounding disturbs the flow field on the surface of the fish compared to open water, the *hydrodynamic image*. The properties of a hydrodynamic image of a moving body mapped through an incompressible inviscid irrotational fluid are discussed in Sichert *et al.* [406] by performing a multipole expansion of the flow field of varying shapes. The flow field is measured by a transparent artificial lateral line, meaning that the presence of the artificial lateral line does not disturb the flow field of the moving body. Then from the estimated multipole moments basic information about the shape of moving body is extracted. A similar analysis has been carried out by Bouffanais *et al.* [427] in 2D for pressure sensing. The essential results are that, given a realistic resolution of the lateral line sensors, the upper bounds for the range of localization and shape reconstruction are the size of the lateral line system and the size of the moving object. The hydrodynamic image therefore only provides information about the environment in a very close range.

From these constraints, strict requirements follow for the implementation of an artificial lateral line system on an moving robot. First, the lateral line sensors must be capable of detecting small, slowly varying [416] changes in the comparably strong flow field around the moving robot. Secondly, the information processing must be very fast to enable the robot to react on detected changes of the immediate environment. And thirdly, the robot must be highly maneuverable in order to change the state of motion appropriately within this narrow range.

Motivation

This work aims at the transfer of the sensor capabilities to a technical system, more specifically an underwater robot. Transferring the mentioned capabilities of fish with their lateral line organ to a robotic system would be beneficial in a number of ways. First, it would complement existing established sensor technology. For instance sonar sensors have a minimum distance at which to measure, with a blind zone within that distance, whereas camera- or laser-based systems are dependent on visual conditions. Second, its function is purely passive and uses information that is present due to the physics of flow fields. Third, the formation control of several platforms equipped with the same sensors is possible without sensor interference or the need for data exchange. Also reflections, like with sonar systems in narrow spaces, do not interfere with the measurement. Moreover, mapping of the hydrodynamic properties of the environment is enabled. As will turn out in the course of this work, reconstruction of the environment is achievable.

Outline

This chapter is organized as follows: Section 4.2 describes the biomimetic realization of the artificial lateral line system, challenges in the design and sensor characteristics. This section also introduces the robotic submarine Snookie as a carrier for the sensor system. Based on the properties of the robot, two different methods for modeling the hydrodynamic stimuli appearing at the sensor system are explained. In Section 4.3, an analytical solution to modeling the flow field on the sensors is given. Population-vector coding is introduced as a method to detect walls and objects. Verification of the model and the detection ability with the sensor system are evaluated in experiments. Due to the limitations of the analytical

Figure 4.1: *Astyanax mexicanus*, Characidae, Blind Cave Tetra [359]
Photograph: H. Zell

model, flow field reconstruction methods in Section 4.4 utilizes the modeled stimuli for the detection of the presence of objects and their reconstruction. The reconstructed objects are then used in Section 4.4.6 to generate a map of the environment, which can be used for navigation. Finally, the methods and results are summarized in Section 4.4.8.

4.2 An Artificial Lateral Line System

The lateral line system is a mechanosensory system, that can be found in fish, tadpoles and amphibians. It is a group of sensory units, called neuromasts, that are distributed along the animal's body. These neuromasts consist of hair cells, which are bundled and encapsulated by gelatinous flags or cupulae. Each hair cell has an asymmetric structure, with one kinocilium and several stereocilia. The asymmetry leads to a directional sensitivity of the cell, with each bundle containing antagonistically oriented hair cells (for further information, the reader is referred to work in the biology domain [415, 438]). Fish possess two types of neuromasts, surface neuromasts and canal neuromasts. While the basic configuration is the same, the placement is different, see Figure 4.2. Surface neuromasts are located on the surface of the skin, with the cupulae protruding into the water. They are orientated either parallel or orthogonal to the body axis. They primarily respond to local water velocity, causing shear forces due to viscosity and pressure. Canal neuromasts are seated in sub-dermal fluid-filled canals. The canals are punctuated by pores, connecting them to the surrounding water. Between two pores a neuromast is embedded, sensing the local velocity in the canal. Since the pressure difference between two adjacent pores is accelerating the canal fluid [351], canal neuromasts are sensitive to pressure differences, which are, in turn, proportional to fluid acceleration [401]. The sensitivity and dynamics of the canal neuromasts are dependent on the physical and hydrodynamical properties of both canal and cupula [438]. The theoretical fluid flow velocity detection threshold is in the

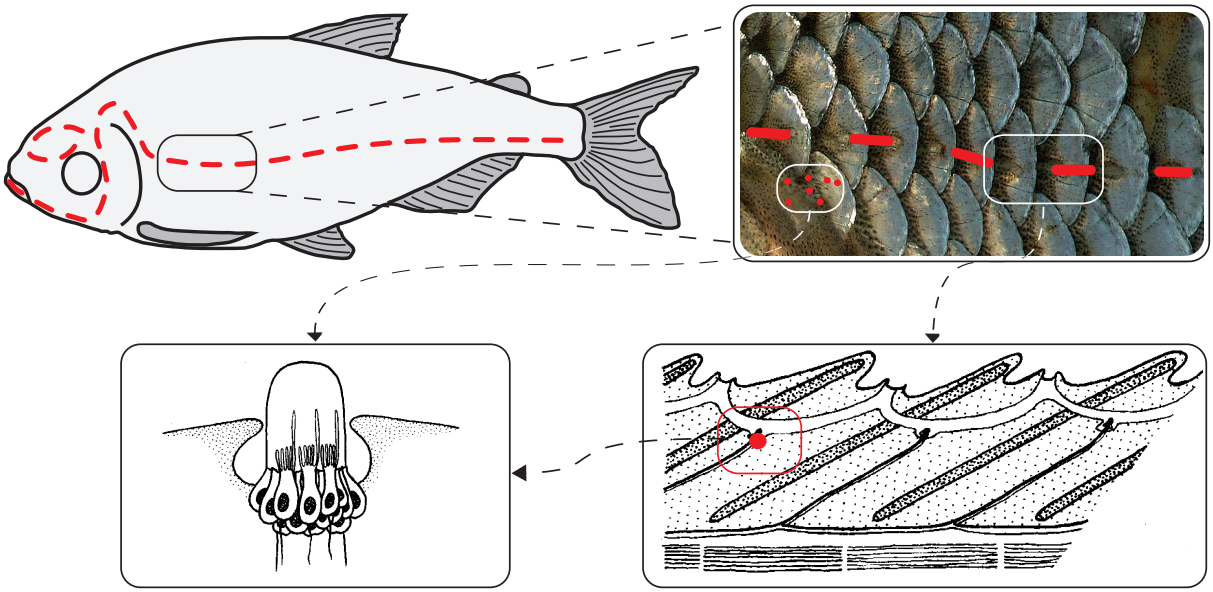


Figure 4.2: The lateral line system of *Astyanax mexicanus*. Dashed line: lateral line, indicating the placement of canal neuromasts (close up / cross section: lower right side). Dots: exemplary placement of surface neuromasts scattered all over the body (close up: lower left side). Neuromasts inside the canals are similar in structure to surface neuromasts. Adapted from Campbell [398].

range of $10\,\mu m\,s^{-1}$, according to the model by van Netten [438]. This is in good accordance with experiments on superficial neuromasts, with a measured threshold of $25\,\mu m\,s^{-1}$ [368]. The cupular impulse response has mean measured value of 4.4 ms for a ruffe [387], with a theoretical value of 3.8 ms. A notable effect of the placement of neuromasts in a canal is that the cut-off frequency due to the combined filtering properties of the canal and the cupula is $\approx 100\,Hz$ [438]. These properties of the lateral line system give a first guideline on the desired properties for the artificial later line system introduced in the next section.

4.2.1 Problem Statement

Sensors for the integration into an artificial lateral line systems require a high sensitivity to changes in the flow velocity and a short time constant, while at the same time being compact enough to build an array. Several groups have already prototyped different types of sensor concepts to realize an artificial lateral line system. At present none of the sensors is commercially available, yet. The flow sensors promising acceptable accuracy and stability available on the market can hardly be integrated to an artificial lateral line system and mounted on a robot due to size and power restrictions, e.g. Laser Doppler Anemometer (LDA). For the artificial lateral line system of *Snookie* for these reasons a conservative design decision in favor of hot thermistor anemometry was made.

4.2.2 Related Work

The basic functionality and morphology of the lateral line system is well known [396, 436, 409]. However, the exact transfer from the hydrodynamic stimulus to the excitation of the sensor [411, 387, 396, 438, 401, 434] the resulting neuronal signals [353, 371, 385] and their processing is still under investigation [420, 435, 402, 436, 360, 372].

This means that so far, attempts to rebuild the lateral line system can only lead to an approximation or abstraction of the biological source. For air, building biomimetic flow sensors is significantly simpler due to the properties of the medium, especially the viscosity and conductivity. Research in biomimetic flow sensing is driven by the upcoming interest in insect-like microflight. A review on different technologies in this sector is given by Motamed and Yan [386], highlighting sensor design and experiments. The focus is on the determination of forces acting on the microrobot as a feedback for control. One step further in terms of object/stimulus localization are projects utilizing arrays of biomimetic hair cells (*cilia*) as sensors. Work by Izadi *et al.* [338] shows the localization of a dipole source – a vibrating sphere – by measuring the deflection of artificial hair sensors. The deflection of the hair induces a capacitive change in the hair root on the substrate, which can be related to the flow velocity. Other artificial cilia are based on the piezoelectric effect, for example with polyvinylidene fluoride fibers (PVDF) [382]. Biomimetic cilia designed for underwater use are based on similar micro-electro-mechanical systems (MEMS) techniques, with a variation in the materials used to generate the piezoelectric effect. The sensors are either used as surface neuromasts [358, 428, 404], or integrated in a canal [344]. Both approaches can be used for dipole localization [350, 344]. An extension of the cilia approach is encapsulating them with a hydrogel cupula [374]. While biomimetic cilia might come close to the biological source of inspiration, the robustness, manufacturing complexity and signal-to-noise ratio are still challenges that prevent the application in an autonomous underwater vehicle. A different approach for underwater sensing is by using thermal transfer as a means for detecting the flow velocity. Hot-wire anemometers have been used for measuring flow velocities in gases and fluids [430] for a long time, but advances in the miniaturization make them applicable to artificial lateral line systems. First trial runs were done by Coombs [392], as a means of *"measuring water motions used in stimulating the mechanosensory lateral line system of a teleost fish"*. Micromachined arrays of hot-wire elements show the ability of localizing dipole sources as good as biomimetic cilia [367, 384, 363, 349].

4.2.3 Biomimetic Design

The technical realization of the sensor is an abstraction of the biological system, transferring the functionality and not the actual design. This approach is following Nachtigall's recommendation to extract the key principles [345] and apply them to a technical system. While fish and other bionic projects determine the flow via the deflection of sensors or shear forces on them, in this case the ability to sense the flow at a point on a body's surface is realized by the technique of constant temperature anemometry. Depending on the sub- or super-dermal placement of the anemometers, surface and canal neuromasts can be realized.

There is plenty of theory and experience with a very similar sensor concept, the hot wire. Hot wires were shown to in principle provide the necessary accuracy and temporal resolution [392, 343]. The energy dissipation of the smallest commercially available thermistors allows high integration densities and low energy consumption. A thermistor promises a better signal to noise ratio for small relative signal changes due to its steeper resistance curve compared to a hot wire. And finally, a small thermistor can be embedded in solid material providing the robustness necessary for operation on a moving robot. In this section, the underlying principle, as well as the sensor design are explained.

Measurement principle

The temperature of the a heated element is given by $T = T_\infty + T_\theta$ with T_∞ being the ambient temperature and T_θ the over-temperature. The heat dissipation of the element in a fluid is a function of the fluid's relative velocity v,

$$P \approx (A + Bv^n)\, T_\theta\,, \tag{4.1}$$

where $n \approx 0.5$ and the constants A and B depend on the geometry and the fluid [430, 361, 388, 414, 390, 364, 431, 400, 419]. For a sphere with diameter d [414, 400] equation (4.1) can be approximated by

$$P = \left[2 + 0.55 \left(\frac{\nu c_p \rho}{k}\right)^{0.33} \left(\frac{vd}{\nu}\right)^{0.5}\right] 4\pi \left(\frac{d}{2}\right)^2 \frac{k}{d} T_\theta\,. \tag{4.2}$$

with the specific heat capacity c_p, the heat conductivity k and the kinematic viscosity ν. Constant temperature anemometer sense the velocity of a fluid or gas by measuring the power P needed to keep a heated element on an over-temperature T_θ.

Disturbance sources

The sensors are subject to noise and other disturbances that can occur during operation, influencing the signal. Some effects are explained in the following, which lead, together with the previously mentioned requirements, to the sensor design in its current form.

Noise Since the sensor is heated, *thermal noise* should be considered. An estimation of the root mean square (RMS) of the noise voltage U_n, is given by

$$U_n = \sqrt{4 k_B T R \Delta f} \tag{4.3}$$

In this case, with a bandwidth of Δf up to 100 Hz, $T = T_\theta = 353\,\mathrm{K}$ being the temperature and $R = R_\vartheta = 240\,\Omega$ the resistance of the heated thermistor, the RMS noise voltage is approximately $22 \times 10^{-9}\,\mathrm{V}$.

Excess noise or *flicker noise* is not related to the heating process, but occurs in all electronic devices. It has a pink power spectrum and is therefore called pink noise or $1/f$ noise. According to the measurements of Keplinger *et al.* [377], the $1/f$ noise for thermistors is in the range of $10 \times 10^{-6}\,\mathrm{V}$ for 0.1 Hz or below.

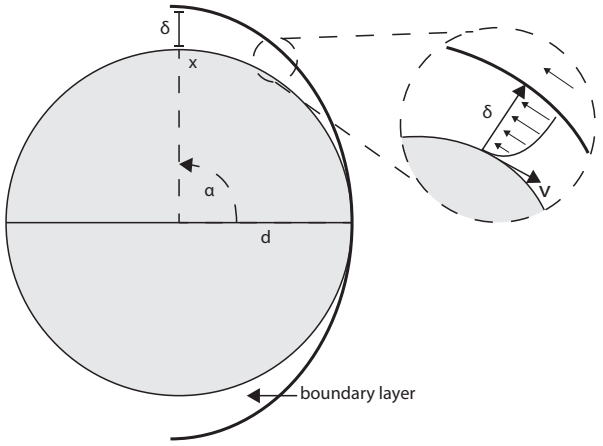


Figure 4.3: Boundary layer around a cylinder in a viscous fluid. The detailed section shows the velocity profile in the boundary layer.

Not only the sensor itself is introducing noise into the measurement system, but also the amplifiers in the controller board and electromagnetic interference in cables and boards. The amplifiers are low noise INA103 amplifiers with $U_n = 1 \times 10^{-9}\,\mathrm{V}/\sqrt{\Delta f}$ [405]. The electromagnetic interference is dependent on the environmental electromagnetic fields and therefor not quantifiable. The effect however can be minimized through shielding of the cables and boards.

The combination of these noise effects can be summed up to approximately $10 \times 10^{-6}\,\mathrm{V}$. For fluid velocities up to $v = 0.1\,\mathrm{m\,s^{-1}}$, which are relevant to this application, a signal amplitude of $\approx 100 \times 10^{-3}\,\mathrm{V}$ can be observed in experiments, as shown later. This leads to a signal-to-noise ratio (SNR) of

$$SNR = \frac{P}{P_n} = \left(\frac{U}{U_n}\right)^2 = 10^8. \tag{4.4}$$

For lower fluid velocities, for example $v = 1 \times 10^{-3}\,\mathrm{m\,s^{-1}}$, the SNR is still 10^6.

This examination on noise shows that noise effects within the system are of no significance to the measurements. Shielding on the other hand is important, since the effects can not be predicted and the magnitude of noise in a laboratory can be within the range of the signal.

Boundary layer If a fluid is in relative motion to a surface, the mechanical properties of the flow can be characterized by the dimensionless Reynolds number, which is defined as

$$\mathrm{Re} = \frac{\rho v L}{\mu} = \frac{v L}{\nu}\,, \tag{4.5}$$

with L being the characteristic length of the surface, v the free stream fluid velocity, ρ the density and μ the viscosity of the fluid. If two flows have the same Reynolds number, they can be seen as mechanically similar [425].

The boundary layer is a concept introduced by Prandtl [407]. In this concept, a fluid flowing past an object with a high Reynolds number can be divided into two unequally large

regions: a bulk region, where the viscosity can be neglected, and a very thin boundary layer close to the wall, where viscosity needs to be taken into account [425, 408], see Fig. 4.3. The boundary layer thickness δ is dependent on the Reynolds number and decreases with higher numbers. It is an artificial concept, as the transition from boundary layer to inviscid flow is continuous. The no-slip condition states that at the solid boundary of the object, the fluid will have zero velocity relative to the object boundary. With increasing distance to the object surface, the flow velocity increases nearly exponential up to the free stream velocity. In practice the boundary can be defined as the point where the flow velocity reaches a certain percentage of the outer flow, for example 99% [342, 425].

Within the boundary layer, flow can be laminar or turbulent. The partition of flow characteristics in laminar or turbulent can be based on the Reynolds number. Below the critical Reynolds number, laminar flow consists of layers of fluid with different velocities, without much exchange of fluid between layers perpendicular to the flow direction. At or above the critical Reynolds number, turbulence occurs, which is characterized by erratic, random fluctuations of the fluid.

According to Schlichting [425], the thickness of the laminar boundary layer can be approximated by

$$\frac{\delta}{x} \approx \frac{5}{\sqrt{Re}} \,. \tag{4.6}$$

It should be noted that other approximations exist, which estimate the layer to be thinner [408]:

$$\frac{\delta}{x} \approx \frac{1}{\sqrt{Re}} \,. \tag{4.7}$$

Here the more conservative estimate should be taken, but even that is only a rough estimate. For the sake of simplicity, the hydrodynamic properties of the mounting region for the artificial lateral line system are the same as for a sphere (in 3D), or a infinitely long circular cylinder (in 2D considerations), with a diameter of $d = 0.25\,\mathrm{m}$. Following equation (4.5) leads to a Reynolds number of $Re = 3 \times 10^4$, with ν taken from Table C.2, and v being the targeted exploration velocity of about $0.1\,\mathrm{m\,s^{-1}}$.

It should be noted that due to pressure gradients around the object and the loss of kinetic energy of water molecules moving in the viscous boundary layer, the boundary layer separates from the surface of the object at a certain region. For a sphere or circular cylinder, this happens approximately in the equatorial region [425]. Under the assumption that $x = \frac{d}{2}\alpha$, with $\alpha = \frac{\pi}{2}$, an upper bound for the thickness of the laminar boundary layer can be roughly estimated to $\delta \approx 6 \times 10^{-3}\,\mathrm{m}$.

This shows that surface sensors placed on sockets of about 6mm are with considerable certainty elevated out of the boundary layer. This is not relevant for canal sensors, since pressure is transferred nearly undisturbed through the layer [373, 412, 426].

Turbulence Throughout this project, it is assumed that all flow is laminar, without turbulence. Thus, turbulence must be reduced or its impact minimized.

If assumed that the fluid is initially at rest and that no external sources of fluctuations

in the fluid exist, turbulence appears only due to the motion of an object in the fluid. As previously described, the viscous friction between the object surface and the fluid creates a boundary layer in which flow can be either laminar or turbulent. The transition from laminar to turbulent flow is dependent on the Reynolds number:

$$\mathrm{Re}_{\mathrm{x_{crit}}} = (\frac{vx}{\nu})_{crit} \tag{4.8}$$

with x_{crit} marking the transition point. For a circular cylinder in 2D $\mathrm{Re}_{\mathrm{x_{crit}}} = 5 \times 10^5$ and for a sphere in the 3D case $\mathrm{Re}_{\mathrm{x_{crit}}} = 3 \times 10^5$ [425].

Turbulence also occurs due to the pressure gradients the moving object creates in the fluid, together with the separation of the boundary layer. The area between the separated layers behind the object becomes turbulent and can create a Kármán-vortex-street [429].

Thermal shielding Since the measurement is based on heat transfer, thermal shielding results in a loss of sensitivity. The thermal conductivity between the sensor and its environment is both important for the signal to noise ratio and the time constant.

For good thermal conductivity, the sensor needs direct contact with water. Air bubbles on the probe decrease the effective thermal conductivity and in worst case shield the sensor completely from the flow. Bubbles can result from the emission of dissolved gases in the water or from electrolysis. The lowest threshold for electrolysis to occur is the standard potential of the water electrolysis cell with $-1.23\,\mathrm{V}$, according to the Nernst equation, with the actual start usually given at about $1.5\,\mathrm{V}$ [430]. The voltage on the sensor is up to $10\,\mathrm{V}$, so electrical insulation is required to prevent electrolysis.

4.2.4 Application

For the artificial lateral line sensors of Snookie, glass-coated thermistors with a diameter of $0.36\,\mathrm{mm}$ from the Honeywell 111 series as heated elements are used. Thermistors are semiconductors with a non-linear negative dependency of electrical resistance upon the temperature. The resistance $R_\vartheta \approx 300\,\Omega$ of the thermistor at working temperature $T \approx 80\,^\circ\mathrm{C}$ with an over-temperature T_θ of approximately $60\,^\circ\mathrm{C}$ is, given the resistance $1400\,\Omega < R_0 < 2.4\,\mathrm{k}\Omega$ at room temperature T_0,

$$R_\vartheta = R_0 e^{\beta_\vartheta (1/T_0 - 1/T)}\,, \tag{4.9}$$

where the constant β_ϑ lies between 2000 and $5000\,\mathrm{K}$.

To sustain the thermistor at a constant temperature, the supplied electrical power P_{el} must equal the dissipated energy P, if it is assumed that all energy is converted to heat and no leakage currents, e.g. due to deficient isolation, appear.

$$P = P_{el} = UI = \frac{U^2}{R_\vartheta}\,. \tag{4.10}$$

The following rough estimates show that it is entirely legitimate to treat the thermistor adiabatically in the sense that it immediately adapts its temperature and thereby its resistance to changes in the transport of heat from it as it has been implicitly done in the

thermistor model so far. The voltage necessary to maintain a stable resistance of about $240\,\Omega$ of the thermistor in water a rest is around $1.5\,\mathrm{V}$ depending on the individual sensor. This results in a dissipated power of $0.75\,\mathrm{W}$. Power dissipation of these thermistors, when mounted on a small PCB board, fits well the power law (4.11) with $n = 0.34$, $A = 1.03\,\mathrm{mW\,K^{-1}}$, and $B = 0.74\,\mathrm{mW\,[K\,(m/s)^n]^{-1}}$ [343]. For comparison, the total heat stored in a sphere of the diameter of the size of the thermistor with over-temperature $60\,^{\circ}\mathrm{C}$ made of silicon or glass is at least one order of magnitude smaller. A change of heat transport due to changing flow conditions must therefore be immediately compensated by a change in the voltage supplied to the thermistor to hold a constant temperature. The voltage over the sensor is an adiabatic measure for the fluid velocity,

$$U^2 \approx R_\vartheta (A + Bv^n) T_\theta \,. \tag{4.11}$$

R_θ can be defined as the resistance that the sensor should have during operation at temperature T, if it is in still water. As the sensitivity of the sensor is dependent on the change in its resistance, see equation (4.11), and the resistance is directly linked to overtemperature (4.9), R_θ should provide good linearity around the setpoint and a steep gradient in the resistance-temperature curve. In case of the Honeywell 111 series, choosing $R_\theta = 240\,\Omega$ results in an operating temperature of approximately $353\,\mathrm{K}$. In a laboratory setting with controlled room temperature $T_0 \approx 290\,\mathrm{K}$, this leads to an overtemperature of $T_\theta \approx 60\,\mathrm{K}$.

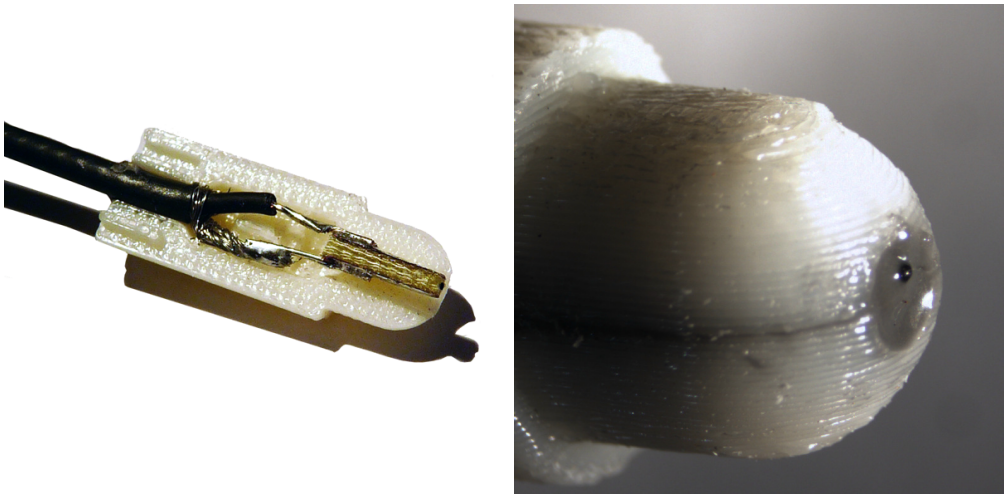
The overtemperature on the sensor is sustained by a custom analog control board. The thermistor is kept on a constant temperature with a Wheatstone bridge, in which the sensor is integrated in one of the branches. The branches are tuned to have the same resistance, a cool down of the sensor results in a resistance difference and thus a voltage across the bridge. This voltage is both the measurement signal and feedback to heat the thermistor to its set overtemperature. Since the bridge voltage is only in the range of a few millivolts, a three stage amplification system is used.

To counteract the aforementioned disturbances, the thermistor needs to be packaged, see Section C.2 for details. The shape of the sensor is a bullet form, see Figure 4.4, which resembles a $5mm$ high cupula if mounted in place. The thermistor sits at the top of the cupula.

The Submarine Snookie

The submarine *Snookie* is an autonomous underwater vehicle (AUV) specifically designed as a test bed for the artificial lateral line system. Here the concept of the robot is briefly presented.

The robot is designed and adapted in a way to provide good conditions for the new sensor system. It should also be highly maneuverable and independent in terms of computational power and energy supply. This leads to a number of requirements that need to be addressed in the design considerations. In contrast to tethered remotely operated vehicles (ROVs), AUVs need to integrated their own energy supply and computation units. This means on the one hand that operation time is limited by the amount of energy stored and the overall efficiency of the robotic system, on the other hand the operation range and



(a) Sensor structure **(b)** Embedded thermistor

Figure 4.4: Artificial lateral line sensor.

movement within the underwater environment is not limited by the tether. Since all data has to be processed onboard, the choice of the computation units is a trade-off between efficiency (and thus power consumption), size and the processing power, which is needed to accomplish all necessary tasks. In the current implementation, the computational requirements are split up into low-level control, with all the actions and computations necessary to stabilize the robot underwater and provide propulsion, and high-level control running the Robot Operating System (ROS) for more complex tasks and the analysis of the sensor data. An overview on the modules of the robot is given in Section C.3. The robot consists of a cylindrical watertight main compartment, in which all of the electronics is encapsulated, two half-spheres at the end of the cylinder and six thrusters. It has a length of 74 cm and a diameter of 25 cm. The overall mass is 32 kg, which can be fine tuned to match the water displacement of the robot to achieve neutral buoyancy. This is important for energy management and maneuverability [379].

The most important design consideration in case of this project is the shape of the robot. Not only must the components of the robot fit inside, but the outline has direct influence on the functionality of the artificial lateral line system. The sensor system is dependent on the hydrodynamic properties of the shape of the robot, which in combination with the speed accounts for the Reynolds number, see equation (4.5). As shown previously, the Reynolds number gives a measure for the estimated state of the boundary layer, being it laminar or turbulent, and if it separates from the body surface. Figure 4.5 shows the 2D simulation[1] of the velocity distribution of water around a fish body approximated by the shape of airfoil, which is a common and mostly valid approximation [347, 357]. Apart from a small area in the back of the body the flow is laminar around the fish, with the best quality around the head. This is in accordance with the distribution of the superficial neuromasts in fish over the whole body, with a concentration in the head region. This makes also sense

[1] SolidWorks Flow Simulation package, Navier-Stokes solver.

in the notion that the area in front of the fish where it is heading needs to be monitored best. Looking at Figure 4.5, one can see that in contrast to the fish shape, the flow is only laminar around the front half-sphere. This means that the only reasonable area for the velocity sensitive superficial sensors is on the front half-sphere. Later it will be shown that this helps modeling the flow at the sensors, because the complex shape of the robot can be reduced to a sphere, for which an analytical solution of the flow field exists. Also the placement of the sensors in a cross on the front sphere accounts for simulations in a 2D environment, where only the horizontal line array of sensors is considered. Additionally, the stretched cylindrical shape gives enough distance between the sensors and thrusters, which would heavily distort the flow field in the sensor area.

4.2.5 Experimental Evaluation

Several experiments were carried out to evaluate the properties and measurements of the sensors and have a comparison to results expected from theory.

Energy Dissipation in Dependence upon Overtemperature

Energy dissipation P is measured in dependence upon thermistor overtemperature T_θ by applying different constant currents (power source Toellner TOE 8733, ampere meter Fluke 45) and recording the voltages across the thermistor with the oscilloscope. According to (4.10), there should be a linear relation between power P and over-temperature T_θ which is indeed the case, see Figure 4.7a.

Energy Dissipation in Dependence upon Water Velocity

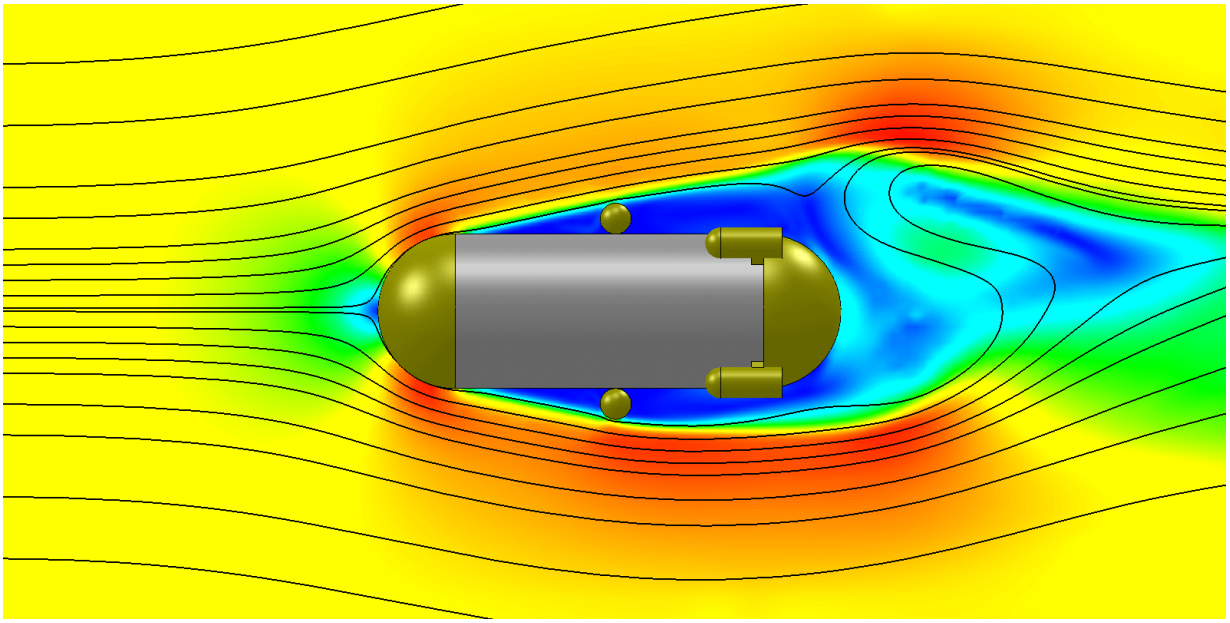
To measure energy dissipation P in relation to water velocity v, the thermistor was driven with a constant current (power source Toellner TOE 8733, ampere meter Fluke 45). A linear axis (Copley Controls Corp, velocity controlled by a Matlab/Simulink program under RTAI real-time Linux) dragged the thermistor sensor through a water basin at constant speed v.

The time-dependent voltage was measured with an oscilloscope (Tektronix TPS 2024) for different speeds v. Using voltage and current, thermal power P was calculated. Measurement data fits well to the power law (4.10); see Figure 4.8a. However, the fitted parameter B in (4.10) was much smaller than predicted by theory (4.2) and by simulation.

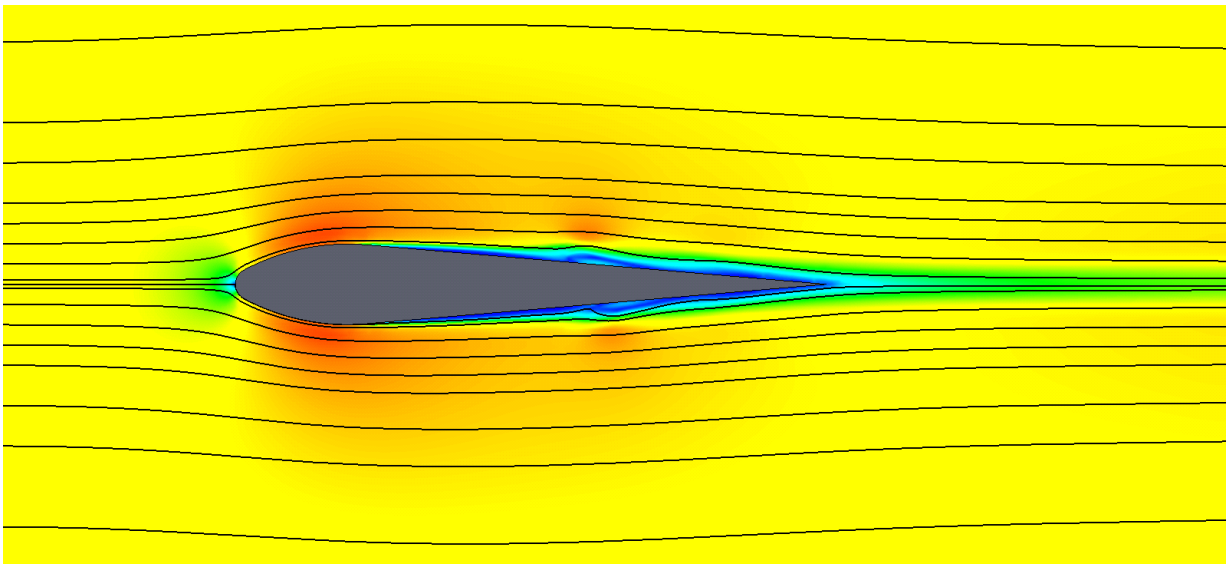
Using the power law (4.10) and linearizing the temperature-dependent resistance of the thermistor (4.9) around the resistance R_0 for velocity $v = 0$ results in

$$\Delta U \sim B v^n \tag{4.12}$$

for small velocities v, i.e. a power law with the same coefficient as in (4.10). Figure 4.8b shows that measurement data fits well to theory.



(a) The robot Snookie



(b) Approximation of the blind cave fish form with an airfoil shape, compare to [347, 357]

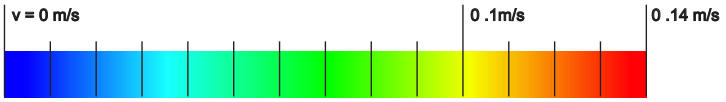


Figure 4.5: Streamlines and fluid velocity around the robot and a fish-like shape in a homogeneous flow field with initial velocity $v = 0.1\,\mathrm{m\,s^{-1}}$

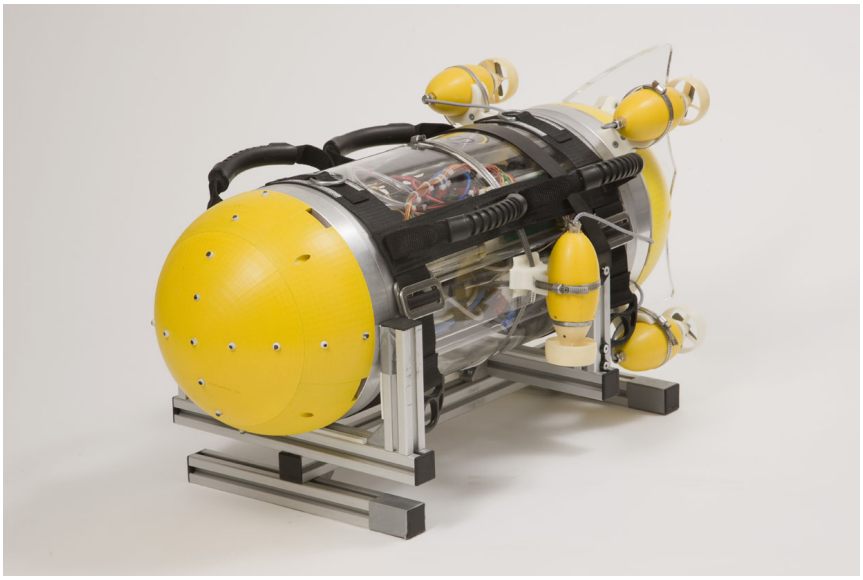


Figure 4.6: The robot Snookie
Photograph: Kurt Fuchs

Impulse Response

Hypothesis 4.1. *The time constant of the developed sensor is similar to the time constant of neuromasts.*

Figure 4.7b shows the impulse response of the sensor. The sensor was driven with a constant current of 18mA(power source Toellner TOE 8733), kicked by hand much shorter than the typical time constant and the time-dependent voltage was recorded with an oscilloscope (Tektronix TPS 2024).

Discussion

The recording shows that the time constant of sensor response is about 10 ms from response maximum to half maximum. Comparing the experimental result of a time constant to the neuromast time constant of 4.4 ms for a ruffe [387], it is clear that the sensor time constant is by a factor of 2 higher. However, the sensor behavior matches the cut off frequency of 100 Hz [438] and should therefore be suitable to achieve bioinspired flow sensing. For the purposed application, the time constant is small enough as to detect water velocity changes in time.

4.2.6 Summary

In this section a biomimetic sensor design, suitable for fitting on a robot, was presented. Following the requirements outlined in Section 4.2.1, an artificial lateral line system based on hot thermistor anemometry is developed. The flow sensors demonstrate the targeted sensing characteristics, while being small enough to fit an array of 17 sensors on a compact AUV and requiring only about 13 mW per sensor. Measures for the application underwater, preventing electrolysis and embedding the thermistor in a cupula, show no negative effects

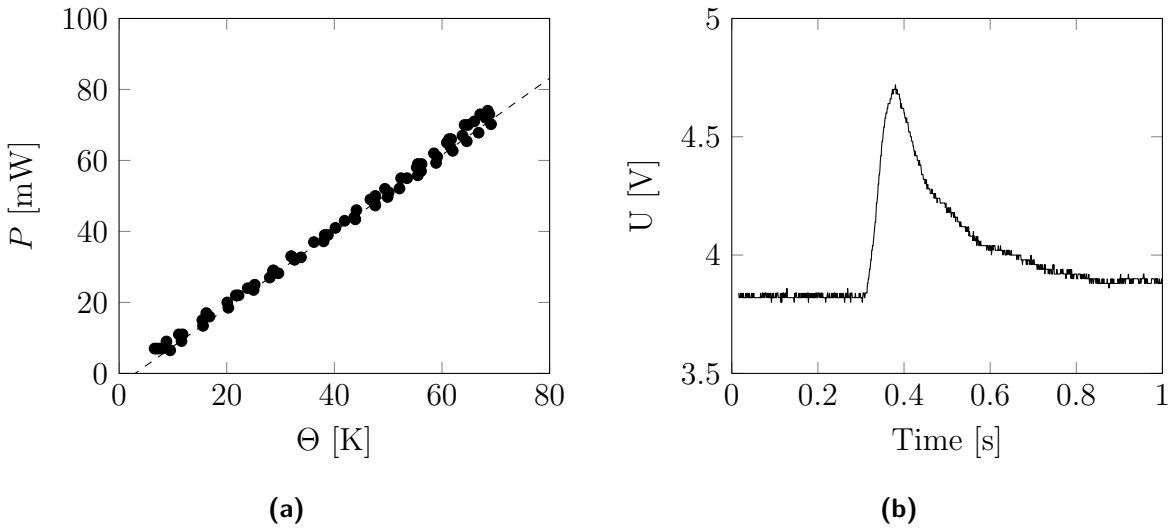


Figure 4.7: (a) Voltage impulse response of the sensor.
(b) Black dots: Measurements of energy dissipation P of a thermistor ($R_0 = 1523\,\Omega$ at $T_0 \approx 293\,\mathrm{K}$) in water mounted on a PCB board and coated, at different over-temperatures T_θ, see (4.10). Dashed line: Linear fit to the data. The relation between energy dissipation and over-temperature is perfectly linear (1.8 mW/K), as the theory (4.10) predicts.

on the performance of the sensor. Experimental results indicate that the sensors properties are close to the theoretical predictions. Additionally, the underwater robot Snookie, that was specifically designed with the implementation of the lateral line system in mind, was introduced. The frontal half-sphere section in particular is designed to establish laminar flow for the aimed at velocity range of the vehicle. As will be seen in the next sections, the lateral line system is a close range sensor system. The multi-thruster layout allows for high maneuverability that is required for avoiding obstacles at close distances.

Two constraints of the artificial lateral line in contrast to the biological source arise from the thermistor design. First of all, the heat transfer principle requires a precise knowledge of the temperature of the surrounding fluid. This is incorporated in equations (4.9) and (4.11) in the overtemperature T_θ. Integrating a thermistor as temperature measurement probe in a flooded – but shielded from flow – area can be used to update temperature reference on the fly. The neuromasts in a lateral line do not face such dependencies due to their mechanosensory nature. This is related to the second constraint of directional sensing. Neuromasts have the ability to sense flow direction due to the coupling of several directional cilia in a cupula. The thermistor design allows only to measure the magnitude of the flow at the tip of the artificial cupula. Due to the isotropic nature of the heat transfer, the only directional information extractable is that the flow is within the tangential plane of the cupula tip. This is sufficient information, as the following methods to extract information from the flow measurements are independent of the flow direction at each sensor. Furthermore, the fluid is assumed to be initially at rest and the flow field is a result of the robot moving, thus the flow direction is related to the known motion.

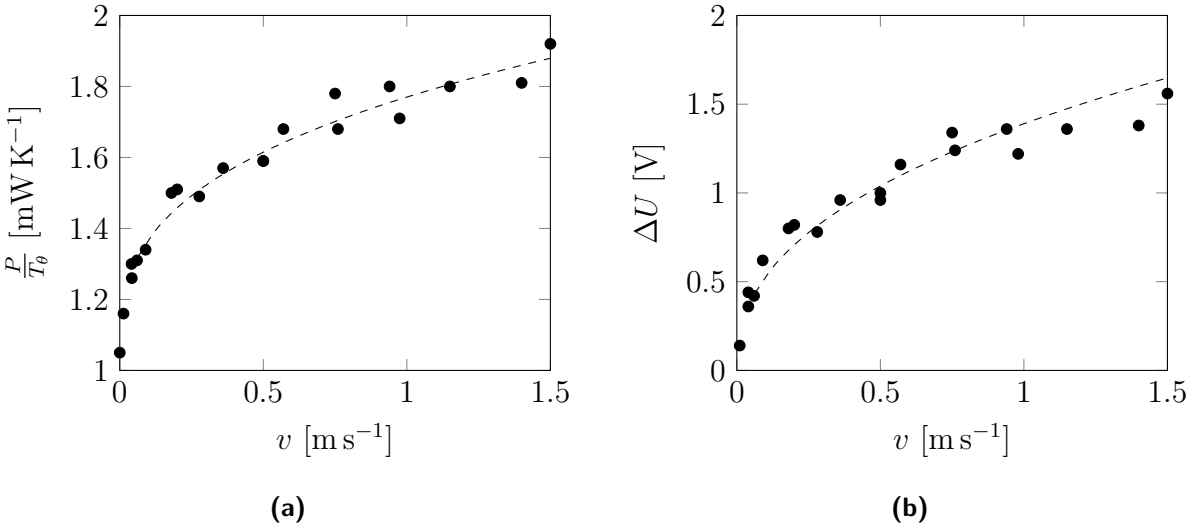


Figure 4.8: (a) Black dots: Measurements of the energy dissipation $\frac{P}{T_\theta}$ per over-temperature of the thermistor sensor from Figure 4.8b. Dashed line: Fit to the data according to the power law $\frac{P}{T_\theta} = A + B \cdot [v/(\mathrm{m/s})]^n$ with $n = 0.34$, $A = 1.03\,\mathrm{mW\,K^{-1}}$, and $B = 0.74\,\mathrm{mW\,K^{-1}}$. The data fits well to the power law of (4.10), but parameter B is much smaller than $B \approx 13.5\ \mathrm{mW\,(K\sqrt{m/s})^{-1}}$ predicted by theory (4.2). (b) Black dots: Measurements of the voltage change at a thermistor sensor driven with constant current of $18.1\,\mathrm{mA}$. The voltage across the thermistor was $3.88\,\mathrm{V}$ for water velocity $v = 0$. Dashed line: Fit to the data according to the power law $\Delta U/\mathrm{V} = B \cdot [v/(\mathrm{m/s})]^n$ with $n = 0.42$, $B = 1.39$. The data fits well to a power law and thus to the theory (4.12).

To summarize, the developed artificial lateral line system has been shown to be a useful abstraction of surface neuromasts. In the next sections, methods to extract information about the surroundings from sensor measurements are introduced. The special characteristic of the spherical front section of Snookie allows to formulate the methods as general solutions for a sphere in 3D or a circular shape in 2D.

4.3 Analytical Stimulus Modeling for Object Detection

This section describes how the hydrodynamic stimuli for the artificial lateral line system can be modeled analytically. The proposed method of image charges gives an analytical solution, but requires an analytical description of the problem, which is only feasible for basic geometries like a sphere approaching a plane or one other object. With the frontal area of Snookie being a half-sphere, this method can be used to model the hydrodynamic image on a spherical approximation of the robot when approaching a wall or similar planar object. Modeling the hydrodynamic image to project the flow field disturbance caused by an object on the surface of the robot can be used to simulate the expected sensor stimuli. Vice versa, the information about the flow field alteration of objects on the sensor readings allows the detection of obstacles in the vicinity of the robot. To avoid these obstacles,

using population-vector coding provides a way to detect the presence of an object and estimate its direction.

It should be noted that in all following sections the fluid (without the influence of an object) is assumed to be at rest.

4.3.1 Problem Statement

Blind cave fish are able to perceive objects and avoid boundaries in their environment with the lateral line. For a robot to do so, first a computational model of the hydrodynamic image on the robot surface, which is equipped with an artificial lateral line system, is needed. As a first step, an analytical solution for the hydrodynamic model of a moving body close to walls is pursued. With the focus on fish, previous work has given solutions for slender bodies. Snookie, however, is a blunt body for which a solution needs to be found. Once the relation between flow field alterations on the robot and obstacles is established, challenges remain how to interpret the sensor readings. For the boundary avoidance behavior, a method to extract the presence and direction of the boundary, typically a wall, is required.

4.3.2 Related Work

Few works so far have employed computational models for analyzation of the hydrodynamic image on fish or submarines. An analytical solution is given by Hassan [433, 397, 370, 418]. It is based on the determination of "*the potential flow around a slender body with a circular cross-section situated in a uniform stream parallel to its axis*" [397] by Handelsman and Keller [341]. Movements of a fish along a plane is modeled by the introduction of a second virtual fish mirrored at the plane [370]. An experimental evaluation is presented by Hsieh *et al.*, describing the implementation of PVDF sensors on a robotic fish. In their work, the robot is supposed to sense pressure deviations due to the presence of a wall [428]. The modeling of the wall presence is done with an image charge method that is similar to the method proposed in [380].

4.3.3 Image Charges

The method of image charges originates in electrostatics, where imaginary charges are introduced to satisfy the boundary conditions. If the flow field around the robot is described as a potential field, this method can also be applied to satisfy the boundary conditions of the hydrodynamic stimulus modeling on the surface of the robot.

Under the assumption that only the flow around the nose of Snookie is of interest, where the water velocity sensors are located, and that modeling the nose is sufficient, the whole vehicle is approximated by a sphere with radius $a = 12.5\,\mathrm{cm}$. At the relevant Reynolds numbers ($\mathrm{Re} = 2aV/\nu \approx 26000$ for velocity $V = 0.1\,\mathrm{m/s}$ with kinematic viscosity ν of water) the boundary layer on the nose of Snookie is less than $B \approx a/\sqrt{\mathrm{Re}} \approx 6\,\mathrm{mm}$ [426], see Section 4.2.3. Thus sensors are placed in the Euler flow regime outside the sphere and viscosity can be neglected there. A sphere with radius a moving with velocity $\boldsymbol{V}$ in

a non-viscous fluid generates a dipole velocity field. If the sphere is at the origin of the coordinate system, the water velocity $\boldsymbol{v}$ at position $\boldsymbol{r}$ generated by the sphere is [412]

$$\boldsymbol{v}(\boldsymbol{r}, \boldsymbol{V}) = \frac{a^3}{2|\boldsymbol{r}|^5}[3(\boldsymbol{V} \cdot \boldsymbol{r})\boldsymbol{r} - |\boldsymbol{r}|^2\boldsymbol{V}] .$$

This velocity field fulfills the continuity equation

$$\partial_x v_x + \partial_y v_y + \partial_z v_z = 0$$

as well as the no penetration boundary condition, i.e. that no flow exists through the boundary.

In case of the robot approaching a wall, the model has to be extended. An infinitely extended wall introduces another boundary condition, viz., that the fluid at the wall cannot move perpendicularly to the wall. As described in [395], this additional boundary condition can be satisfied approximately by introducing another *"mirror"* sphere, see Figure 4.9. The mirror sphere is created by mirroring the original's sphere position and velocity by using the wall as a mirror. Because of mirror symmetry, the velocity field generated by both spheres at the wall is then parallel to the wall, as required by the Euler boundary conditions.

This approach is an approximation insofar as the boundary condition at the surface of the first sphere is disturbed by the presence of the second sphere and thus boundary conditions on the spheres are only satisfied approximately. The approximation is exact for the limit case $a \to 0$ or $|\boldsymbol{D}| \to \infty$. The approximation becomes an exact solution by using an infinite series of mirror spheres [412].

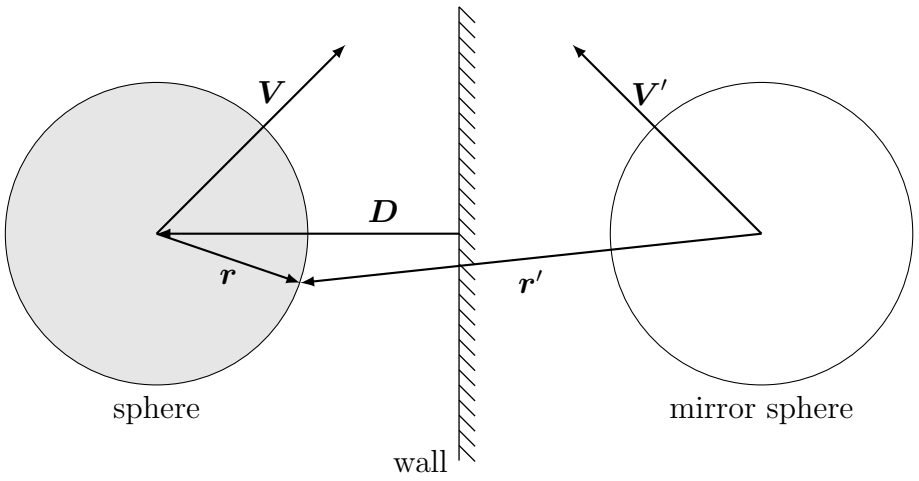


Figure 4.9: Snookie, here simplified to be a sphere, moves with velocity $\boldsymbol{V}$ towards a wall at distance $|\boldsymbol{D}|$. The mirror sphere moving with velocity $\boldsymbol{V}'$ serves to fulfill the boundary condition at the wall. A water velocity sensor sits at position $\boldsymbol{r}$ on the surface of the sphere.

Let the original sphere be at the origin of the coordinate system and let the wall be at distance $|\boldsymbol{D}|$, with the vector $\boldsymbol{D}$ pointing from the wall to the center of the sphere, perpendicularly to the wall. If the velocity of the sphere is $\boldsymbol{V}$, then the velocity of the mirror sphere is $\boldsymbol{V}' = \boldsymbol{V} - 2\boldsymbol{D}(\boldsymbol{D} \cdot \boldsymbol{V})/|\boldsymbol{D}|^2$. The center of the mirror sphere is at $-2\boldsymbol{D}$.

The velocities caused by the moving sphere and by its mirror sphere add up linearly so that the overall water velocity at position $\boldsymbol{r}$, in the laboratory coordinate system, is

$$\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{r}, \boldsymbol{V}) + \boldsymbol{v}(2\boldsymbol{D} + \boldsymbol{r}, \boldsymbol{V}') \,. \tag{4.13}$$

As $\boldsymbol{v}$ is not necessarily tangential to the surface, this velocity has still to be projected onto the surface to get an estimate of the velocity at the surface. In addition, the sphere itself is moving with velocity $\boldsymbol{V}$. Thus the water velocity $\|\boldsymbol{v}_s\|$ measured by a sensor on the surface of the vehicle with its tangential plane described by the orthonormal vectors $\boldsymbol{s}_u$,$\boldsymbol{s}_v$ is

$$\boldsymbol{v}_s = ((\boldsymbol{v} - \boldsymbol{V}) \cdot \boldsymbol{s}_u)\boldsymbol{s}_u + ((\boldsymbol{v} - \boldsymbol{V}) \cdot \boldsymbol{s}_v)\boldsymbol{s}_v \,. \tag{4.14}$$

4.3.4 Population-vector Coding

Population vector coding is a concept from neuroscience. Localization is encoded in groups of neurons forming a map, a neuronal representation of the *"outside sensory world"* [366]. Single neurons represent a preferred direction, with the firing rate depending on the proximity of the target to this direction. Calculating the vector sum of all neurons results in a vector pointing to the target.

According to van Hemmen [366], the population vector $\boldsymbol{n}$ for an assembly of motor neurons ($1 \leq i \leq N$) with firing rate ν_i encodes the direction $\boldsymbol{e}$ of movement as well as the speed of the drawing motion ν via the length of the vector:

$$\boldsymbol{n} := \nu\boldsymbol{e} = \sum_{i=1}^{N} \nu_i \boldsymbol{e}_i \tag{4.15}$$

with $\boldsymbol{e}_i$ being the preferred direction of a neuron.

For the estimation of the direction of an object with the artificial lateral line system the population vector can be reformulated:

$$\boldsymbol{D}_{\mathbf{est}} = \sum_{i=1}^{8} \bar{U}_i \boldsymbol{r}_i \,.$$

with the estimated object direction $\boldsymbol{D}_{\mathbf{est}}$, the positions $\boldsymbol{r}_i$ of the sensors relative to the center of the nose, weighted by the high-pass filtered voltage $\bar{U}_i$ of each sensor. It should be noted that the object direction is only valid for an isotropic placement of sensors on a sphere. Other sensor distributions or object forms distort the direction estimation. If only the direction is of interest, neglecting the distance information from ν, this method can be applied even if no mapping exists from sensor readings to fluid velocities. It is also robust against uniform disturbances that affect all sensors.

4.3.5 Experimental Evaluation

In this section the image charge method is evaluated for its applicability to modeling wall and object detection for an artificial lateral line system.

Wall Approaching

To verify whether the theoretical model for wall-detection is implementable in a robot, several experiments have been made with different distances of the sphere to the walls of an aquarium. The dimensions of the aquarium are $L \times W \times H = 100\,\text{cm} \times 50\,\text{cm} \times 50\,\text{cm}$. For the experiment, the voltage U_m across the bridge was captured and the corresponding water velocity has been calculated using the image charge method described above. Two sensors were placed on the sphere with radius $a = 7.5\,\text{cm}$ at positions $\boldsymbol{r}_0 = (a, \frac{\pi}{2}, 0)$ close to the wall and $\boldsymbol{r}_\pi = (a, \frac{\pi}{2}, \pi)$ at the opposite side. Figure 4.10 shows the relative velocity difference Δ, with

$$\Delta := \frac{|\boldsymbol{v}_s(\boldsymbol{r}_0)| - |\boldsymbol{v}_s(\boldsymbol{r}_\pi)|}{|\boldsymbol{v}_s(\boldsymbol{r}_0)| + |\boldsymbol{v}_s(\boldsymbol{r}_\pi)|}\,. \tag{4.16}$$

in dependence on the relative distance $\delta = \frac{|D|}{a}$ to the wall. The relative velocity difference is only dependent on the ratio δ of the distance to the wall $|\boldsymbol{D}|$ and the radius a of the sphere, specifically

$$\Delta = \frac{\frac{1}{(2\delta-1)^3} - \frac{1}{(2\delta+1)^3}}{6 + \frac{1}{(2\delta-1)^3} + \frac{1}{(2\delta+1)^3}}\,. \tag{4.17}$$

Discussion

The velocity difference increases as the distance to the wall decreases and the water velocity at the sensor next to the wall is higher, in accordance with theory. The effect is about the same magnitude as theory predicts, but sets in a bit earlier, i.e. already at larger distances from the wall in contrast to theory. This may be due to the fact that the sensor was mounted about 1 cm away from the surface of the sphere.

Object Approximation

Hypothesis 4.2. *Objects can be perceived by the lateral line when passed by.*

In the case when Snookie is passing a cylindrical object, the surface of the object is approximated by a wall tangential to the cylinder surface. This approximation is good if the radius of the cylinder is large enough. For approximations that are accurate even for small obstacles and account for fish bodies other than spheres, see [433]. Moreover, for the case of a rotationally symmetric fish-like body approaching a wall and gliding alongside a wall, approximations for the resulting flow exist [397, 370]. In this simulation, however, the image charge approximations described above are used as they describe the general case of arbitrary attack angles between Snookie and the wall. Moreover, in this case with relatively high Reynolds numbers $\text{Re} \approx 26000$, one can expect turbulence approximately where the half-spherical nose is attached to the cylindrical body of Snookie [426]. Hence,

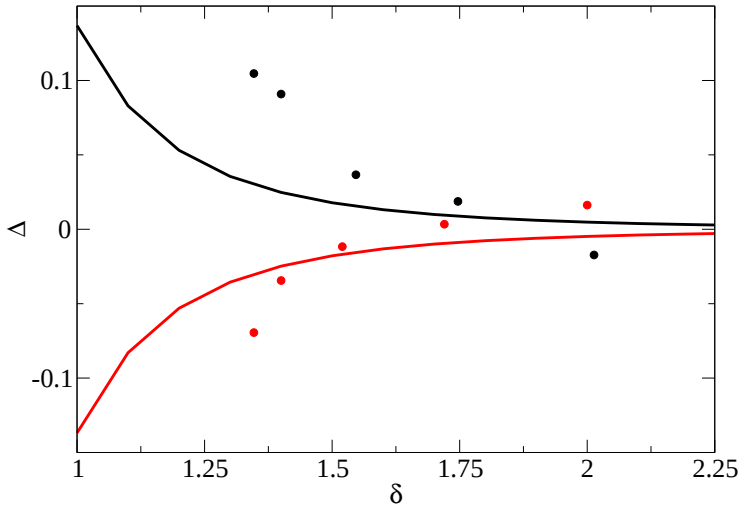


Figure 4.10: Relative velocity differences Δ from measurements (dots) compared to the theory in equation (4.17) (solid lines). If the velocity difference between left and right sensor was negative, Δ got a negative sign (red data) and a positive sign otherwise (black data). Thus the experiments resulting in the red dots were made with the sphere moving closer to the left wall and the black dots were measured with the sphere moving closer to the right wall. Qualitatively, the data matches with the theoretical calculations.

approximations that take into account body shapes other than spherical by simultaneously assuming irrotational Euler flow like [397, 370] would be inadequate in this special case.

Experiments have been done with a sphere with a thermistor sensor mounted on it was dragged past a cylindrical obstacle as in Figure 4.11. The voltage drop over the thermistor sensor is recorded [343]. Figure 4.12 shows simulated voltage responses of the sensors when Snookie passes the cylindrical object at different distances. The simulation started with the same parameters as in the experiment that lead to a voltage response shown in Figure 4.13.

Discussion

The sensors show the essential properties as desired and in good agreement with theory. Present measurements prove that this concept is technically feasible and that the sensitivity required can be calculated by the theoretical model, although the theory neglects viscosity causing boundary layer effects and turbulence. The constant B (4.11), however, has been found to be much smaller than predicted by both simulations and theory. Object detection trials, as in Figure 4.11 and Figure 4.13, show that even only one of the present water velocity sensors can already detect objects. The overall voltage is significantly higher than the measured voltage stemming from inaccuracies in the model equations (4.1), (4.13), and (4.14). The simulated *change* of voltage turned out to be much lower than the change of voltage actually measured when passing a cylindrical obstacle. This suggests that the actual water velocity increase due to presence of objects is even higher than assumed in the simulation, which would make obstacle avoidance easier. Sensitivity strongly depends on

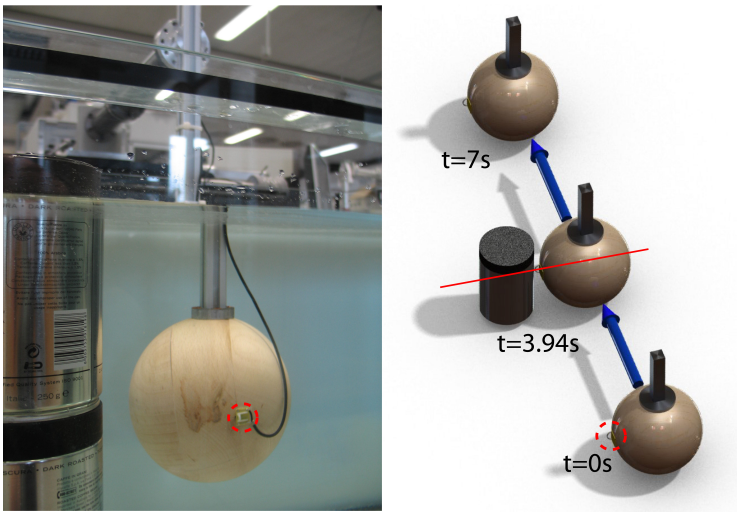


Figure 4.11: Experimental setup used for object detection. A sphere (diameter 15 cm, velocity 10 cm/s) with the thermistor sensor (dashed red circle) is dragged along a linear axis past a cylindrical object (diameter 9 cm). The minimal distance between sensor and cylinder is 1.5 cm and is reached at the position indicated by the red line.

signal-to-noise ratio. Measurement noise in the current setup is dominated by the analog to digital converter card which causes noise with a root mean square of about 1 mV. In addition, the motor of the linear axis is responsible for much of the noise that occurs during motion of the sphere. Other sources of disturbance are the limited size of the aquarium, altering the boundary conditions due to the presence of the other walls, and surface waves.

Simulated Pop-vector Coded Wall Avoidance

Hypothesis 4.3. *The dynamic properties of the robot Snookie and the sensory characteristics of the artificial lateral line allow in combination the avoidance of walls.*

A simulation of the artificial lateral line system mounted on Snookie has been conducted to test the detection and avoidance of walls. When Snookie approaches a wall, it has to first detect the presence of a wall, then estimate where the wall is, and finally start its avoidance maneuver. During the avoidance maneuver, self-movement disturbs velocity measurement, thus no update of wall position data is possible. Therefore, Snookie has to remember the wall position after it has detected a wall and then perform the wall avoidance maneuver without feedback. In the current simulations, Snookie uses eight sensors positioned as indicated in Figure 4.14a. The robot is simulated to approximate the dynamics and thrust characteristics as the real robot [380], allowing a judgement of the feasibility of wall avoidance with Snookie.

The signal to the sensors is exponentially high-pass filtered with a time constant of 2 s since slow drifts in the sensor signals probably do not indicate the presence of a wall. A double criterion has been used to detect a wall. First, the voltage difference between any two sensors must exceed 1 mV. Second, the square root of the sum of squares of the high-

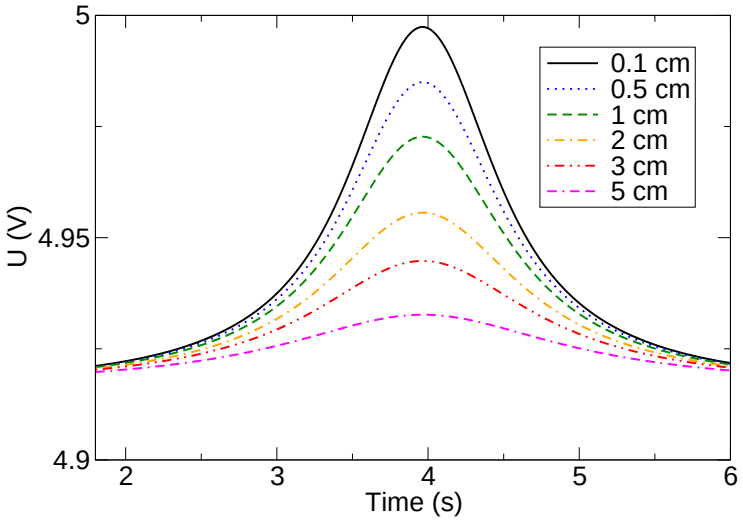


Figure 4.12: Simulated voltage responses U of Snookie's left and right sensor when Snookie is passing by a cylindrical obstacle as in Figure 4.11 in different distances. The minimal distance of the sensor to the surface of the cylindrical obstacle is indicated in the legend. Snookie itself is simulated by a sphere with the same radius as Snookie's nose (12.5 cm).

pass filtered voltages $\bar{U}_i$, $1 \leq i \leq 8$, of the sensors must exceed 1 mV. As soon as both criteria are fulfilled, the algorithm estimates the direction of the wall by the population vector code in Section 4.3.4. The resulting vector $\boldsymbol{D}_{\mathbf{est}}$ roughly indicates the direction to the wall. New desired velocity is computed by mirroring the current velocity at the estimated surface of the wall that is perpendicular to $\boldsymbol{D}_{\mathbf{est}}$, see Figure 4.14a.

$$\boldsymbol{V}_{new} = \frac{\boldsymbol{V} - 2\boldsymbol{D}_{\mathbf{est}}(\boldsymbol{D}_{\mathbf{est}} \cdot \boldsymbol{V})}{\|\boldsymbol{D}_{\mathbf{est}}\|^2} \tag{4.18}$$

Figure 4.14b shows a simulation run, where Snookie is heading with $\boldsymbol{V} = 0.1\,\mathrm{m\,s^{-1}}$ towards a wall with an incidence angle of 45°. The robot starts at (0,0), detects the wall, comes to a full stop after 0.34 m, drives backwards and turns to avoid the wall.

Discussion

A worst case scenario in terms of the robot dynamics, where the robot has to come to a full stop to avoid the wall, was chosen to test the avoidance abilities. The simulation run shows the feasibility of wall avoidance with the artificial lateral line system on the robot. However, it should be noted that the simulated sensor stimuli were free of noise. Although the voltage threshold in the algorithm was chosen to be above the typical noise level of 1 mV of the sensors, additional low pass filtering would be needed to suppress false positives. Taking into account the previous experimental results, comparing the predicted stimulus and the measurements, the simulation underestimates the sensor stimuli. Therefore it is assumed that non-simulated performance degradations of robot or sensor system are compensated by the higher stimulus level, leaving enough time to stop in front of the wall.

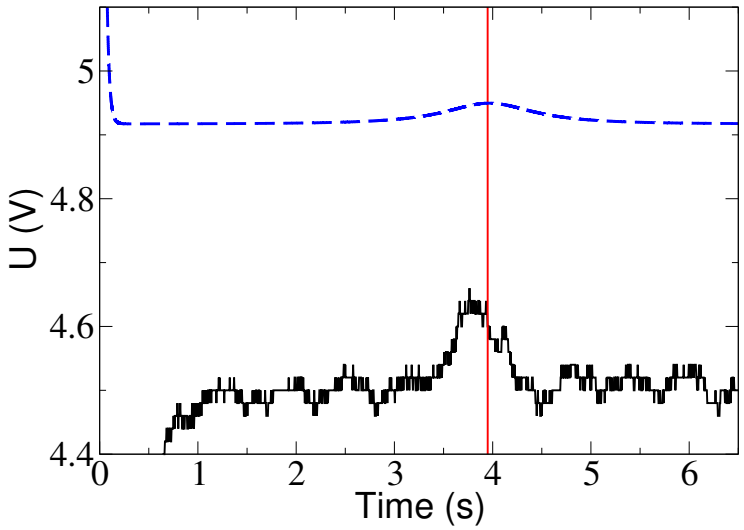
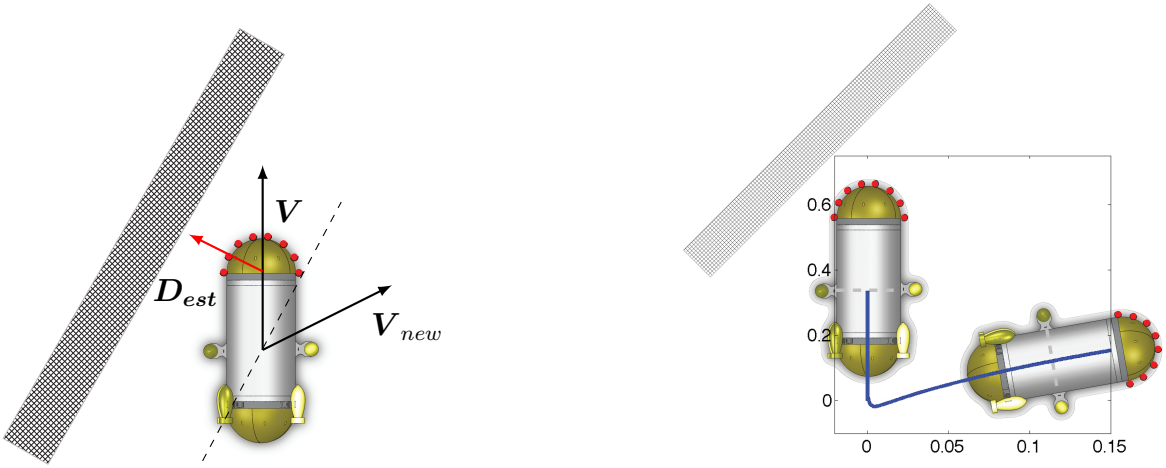


Figure 4.13: Solid black line: Measured voltage response U of the thermistor sensor of Figure 4.6 with the experimental setup of Figure 4.11. The minimal distance between sensor and cylinder is reached at the time indicated by the red line. The body is clearly detectable by the present apparatus (voltage increases in the vicinity of the red line) although there is still some noise mainly due to the power source and a non-constant velocity of the linear axis. The sensor has a resistance $R_0 = 1790\,\Omega$ and was heated by a constant current of 18 mA. Dashed blue line: Voltage response U of the simulated sensor as given by (4.12) with the parameters matching the experiment.

4.3.6 Summary

In this section, an analytical modeling approach for the hydrodynamic image was presented. Approximating Snookie as a blunt object with the hydrodynamic properties of a moving sphere, the image charge method provides a measure to calculate flow velocity changes at any point of the robot front half-sphere due to the presence of an object. Verification of the theory has been performed by experiments with the lateral line sensors on a sphere moving parallel a wall or passing a cylindrical object. Results clearly show the detectability of objects and walls with the artificial lateral line system. While the experimental results qualitatively agree with the theory, they indicate inaccuracies in the chosen modeling approach. However, experiments show higher stimulus measurements than predicted, supporting the applicability of the developed sensor system to object detection.

Pop-vector coding was introduced as a neuroscientific approach to model directional sensing with an array of sensors. It is a robust method against uniform disturbances affecting all sensors. Neglecting distance information $\|\boldsymbol{D}_{\mathbf{est}}\|$, this approach is also not dependent on determining flow velocities, but can be directly applied to raw sensor voltages. While being robust and simple to apply to the problem, no information regarding size and shape of the detected object and for few sensors only a rough direction of the nearest object can be given. This information is sufficient to reproduce observed wall avoidance behavior of blind Mexican Cave Fish with a robot in simulation. The simulation in Section 4.3.5



(a) Direction estimation

(b) Object avoidance

Figure 4.14: (a) When Snookie detects a wall, it estimates the direction to the object (red arrow) by a population vector depending on the voltages at its sensors (red dots).
(b) Illustration of object avoidance. The blue curve indicates Snookie's path during a simulation run (ideal sensory data, no environmental noise).

shows the feasibility of the chosen approach. On the other hand, the available environment information from this method is insufficient to explain more complex behaviors of blind cave fish such as object discrimination or mapping of its habitat.

4.4 Numerical Stimulus Modeling for Object Recognition

As stated in the last section, the image charge method requires an analytical description of the problem, which is only feasible for basic geometries. The proposed reconstruction method based on the boundary-element method (BEM) can find a solution numerically for arbitrary forms of the object and environment. Urban [347] shows, that for simple geometries the difference between the analytical and numerical solution is negligible. This means that it can act as a replacement for the image charge method, with the extension to complex shapes. Flow field reconstruction results in a more detailed knowledge about the surroundings.

4.4.1 Problem Statement

To be able to mimic advanced abilities of fish, such as object discrimination and habitat mapping, better models for the flow around the sensor system are needed than the analytic method described in the previous section. Modeling flow around an object is a well-known problem in computational fluid dynamics (CFD). Solver for numerical approximations of flow and pressure fields are part of most computer aided drawing (CAD) packages. For a given volumetric representation of the object and its surroundings and a defined inflow

and outflow of the fluid, the flow field around the object can be simulated. For the case of this work, however, the challenge is the inversion of the problem, i.e. the estimation of the surrounding flow field from measurements on the surface of the robot. This is typically not accounted for, therefore the number of reconstruction methods is negligible compared to solutions for the forward problem. Previous flow field reconstruction methods [406, 427] require strict assumptions about the shape and the number of objects – typically only one – presented to the lateral line system. These limitations are only compatible with simplified laboratory settings. For real world applications, these restrictions need to be reduced, especially the prior knowledge of the encountered shapes.

4.4.2 Related Work

Overviews on the state of CFD are given by Ferzinger *et al.* [365] or Anderson [378], and a more recent one by Versteeg and Malalasekera [354]. CFD is applied to modeling the stimulus that occurs from the hydrodynamic interference with objects on the body of a fish, see for example Windsor *et al.* [352, 340]. Windsor *et al.* [352] compute the pressure distribution along the fish body given the flow field experimentally determined by particle image velocimetry (PIV). The measured flow and pressure field information is used to explain observed behavior of blind cave fish. To utilize data gathered from the sensors on a robot and get information about the environment, the inverse problem has to be solved. Only recently, groups have worked on using artificial lateral line systems not only to detect, but reconstruct and recognize objects.

Fernandez *et al.* place a pressure sensor array along a submarine dummy [356, 394, 381]. Using principal component analysis, two different cross section shapes (round and square) can be classified if the object is moved along the dummy. This is a first step towards underwater navigation, providing the information for obstacle avoidance and a two class identification. For mapping of the environment, more information about the object is crucial. Bouffanais *et al.* [427] describe a method for reconstructing the shape of an object based on the pressure distribution in vicinity of the body. The shape is encoded in a parameter set of a Laurent-series, limiting the approach to the reconstruction of a single object with quasi-2-dimensional geometry perceived by a transparent sensory system. Given the same setting, the task is solved in 3D with a similar method by Sichert *et al.* [406] measuring fluid velocity. In Section 4.4 a more general method is presented that can deal with arbitrary solid stationary boundaries. The method is applicable to quasi-2-dimensional incompressible, inviscid and irrotational flow around the lateral line system.

4.4.3 Boundary Element Method

The BEM is a numerical method of solving linear partial differential equations. Boundary value or initial-value problems are formulated by boundary integral equations [393]. Discretization only takes place on the boundaries of a problem domain to create a surface (3D) or curve (2D) mesh, reducing the dimension of the problem by one and making it easier to create a mesh [393]. This is the major difference to the related finite volume method (FVM), for which the whole problem domain must be discretized in a volume

mesh, see Urban [347] for a discussion of the advantages of BEM over FVM. In this application an implementation of the BEM is used to compute the velocity potential as a function of sources and sinks on the boundaries of the robot and the environment.

Before the boundary element method can be applied here, some assumptions on the properties of the fluid and objects need to be stated:

- Objects have infinite length in z direction. This reduces the 3D problem to 2D.
- The fluid is ideal and irrotational. This means it is inviscid, incompressible and has no circulation flows.
- Flow is laminar without perturbations.

These assumptions are a non-negligible simplification of the real problems. However, the initial problem is defined in a way to come close to the assumptions. The combination of the array placement of the sensors with the horizontal movement of the robot resembles the 2D assumption. The water should be initially at rest and the placement of the sensors put them in the inviscid flow outside the boundary layer. The velocity of the robot and shape are designed in a way that the Reynolds numbers allow to assume laminar flow.

With these assumptions, the flow can be described with a potential function Φ and the fluid velocity u:

$$u = \nabla\Phi \tag{4.19}$$

For the description of the boundaries, three areas are defined with respect to the fluid and a number of objects i:

- fluid filled area in which the objects can move D
- stationary objects embedded in the fluid, which are bounded by ∂D
- moving objects with velocity $\boldsymbol{v}_i$ and occupied area D_i

According to Lamb [417], the flow field for an inviscid, incompressible and irrotational fluid can be described by the equations:

$$\begin{aligned} \triangle\Phi|_D &= 0 && (4.20)\\ \boldsymbol{n}^T \cdot \nabla\Phi|_{\partial D} &= 0 && (4.21)\\ \boldsymbol{n}^T \cdot \nabla\Phi|_{\partial D_i} &= \boldsymbol{n}^T \cdot \boldsymbol{v}_i|_{\partial D_i} && (4.22) \end{aligned}$$

Equation (4.20) implies, that there are no sources in the fluid, as it is the same as $\nabla \cdot \nabla\Phi = \nabla \cdot \boldsymbol{u}$, which is the divergence of the velocity field. (4.22) is the constraint, that the fluid normal to the surface of the moving object can only move with the same velocity of the object projected normal to the surface. (4.21) is only a special case of (4.22) for stationary objects, i.e. objects with $\boldsymbol{v}_i = 0$.

This leads to the boundary conditions that need to be satisfied:

$$\boldsymbol{u}(\boldsymbol{x}, t) \cdot \boldsymbol{n} = 0 \quad \boldsymbol{x} \in \partial D \ , \tag{4.23}$$

meaning that no fluid can flow into or out of an object, with $\boldsymbol{n}$ being the normal vector on surface of the object.

$$\boldsymbol{u} \cdot \boldsymbol{n} = \boldsymbol{v} \cdot \boldsymbol{n} \tag{4.24}$$

The fluid on the boundary of a rigid body moves with the velocity of the body. Also the Laplace equation (4.20) on D and equations (4.21) and (4.22) need to be satisfied on ∂D and ∂D_i

Following Lamb, the velocity potential Φ in D can be represented by a distribution of sources and sinks on the boundary ∂D. This transforms the initial problem of computing the velocity potential for a point $\boldsymbol{r}$ in D into calculating the strength of the sources $f(\boldsymbol{r})$ on the boundary ∂D [347]. The source density is continuous over the smooth surfaces ∂D and ∂D_i. It needs to be discretized. Here the ∂D boundary is approximated by line elements s_i with a fixed line width l. Using a constant [347] for the line source density s_i over the line elements i, the overall velocity field can be written as

$$\boldsymbol{u}(\boldsymbol{r}) = \sum_i s_i \boldsymbol{u}_i(\boldsymbol{r}) \tag{4.25}$$

In order to solve the velocity distribution, the squared error $E(\boldsymbol{s})$ for the quality of the boundary conditions needs to be minimized, resulting in an optimal solution for $\boldsymbol{s}^* = (s_0\ s_1 \dots s_N)^T$

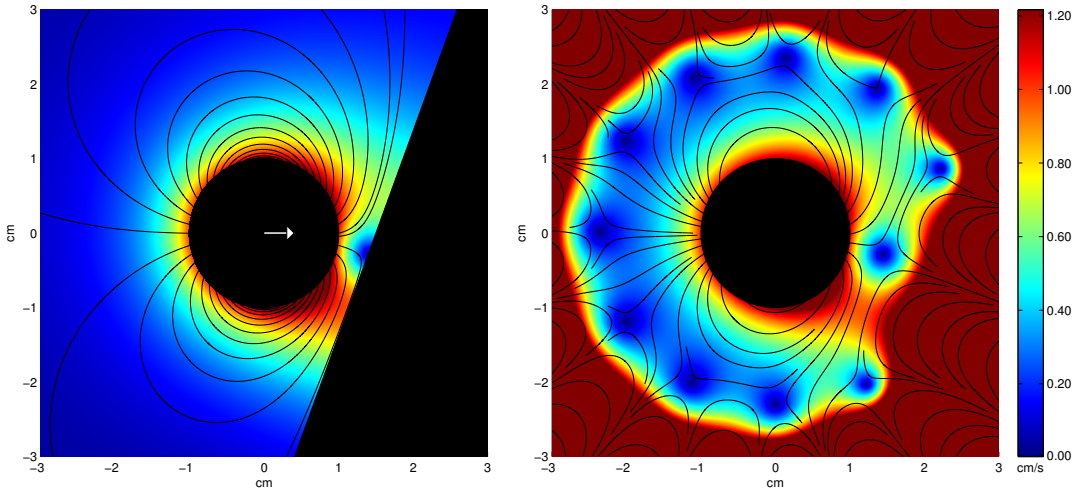
$$\boldsymbol{s}^* = \text{argmin}_{\boldsymbol{s} \in \mathbb{R}^N} E(\boldsymbol{s}) \tag{4.26}$$

with the squared error between flow for a given $\boldsymbol{s}$ and the boundary conditions

$$E(\boldsymbol{s}) = \sum_{j=1}^{N} l \left[\sum_{i=1}^{N} s_i \boldsymbol{n} \boldsymbol{u}_i(\boldsymbol{p}_j^c) - \boldsymbol{n} \boldsymbol{U}(\boldsymbol{p}_j^c) \right]^2 \tag{4.27}$$

4.4.4 Flow Field Reconstruction

The image charge method described in the previous section allows the detection of the presence of objects and walls with pop-vector coding or heuristic rules. However, the most universal solution to the problem of extracting information from the flow field is the inversion of the hydrodynamic image. An example of the reconstruction of the flow field from the hydrodynamic image is given in Figure 4.15. Based on the the velocity potential Φ, one can describe the projection of structures or objects in respect to the expected flow velocities on the surface of the robot. Considering that object geometries altering the flow have a kind of diffusive effect on the flow field [347], it is appropriate to represent the velocity potential Φ as a Fourier series [417]. This takes into account that structural information on the object is dampened frequency- and distance-dependent in the flow field. With the inversion of this projection, the coefficients of the Fourier series can be calculated and the flow field reconstructed, due to the fact that velocities on the robot surface are known from the sensors. For the transformation of the absolute flow field into a relative flow field in robot coordinates (r, φ), refer to Appendix C.4.1. The flow field



(a) Hydrodynamic image **(b)** Reconstruction

Figure 4.15: Reconstruction of the flow field around a circle. (a) A circle with 1000 sensors equally distributed on its surface moves towards a wall at a 20° angle. The flow field is plotted with respect to the FOR. (b) In the reconstructed flow field the wall can be deduced from the parallel streamlines in front of the circle.

reconstruction assumes the same fluid properties as already formulated in the boundary element Section 4.4.3, i.e. that it is inviscid, incompressible and irrotational, and that objects have infinite length in z direction. For further simplification, the following analysis is restricted to 2D.

$$\Phi(r,\varphi) = \sum_{\alpha} \left(A_\alpha \frac{r^\alpha}{r_0^{\alpha-1}} + B_\alpha \frac{r^{-\alpha}}{r_0^{-\alpha-1}} \right) e^{i\alpha\varphi} \tag{4.28}$$

Equation (4.28) is the general solution of $\triangle\Phi$ with constants A_α, B_α, r_0 and the polar variables for radius r and angle φ. r_0 is the radius of the frontal half sphere and introduced for a normalization of r in respect to the diameter of the robot. The spatial frequency information about the object geometry is encoded in $\alpha = -\infty \ldots \infty$. The Fourier equation allows an exact representation and thus reconstruction of structures in the flow field, if all spatial frequencies are present without any errors. Noise in the coefficients gets amplified exponentially with the frequency number α. To reduce the noise impact and computational complexity, spatial frequencies above α_m are cut off. The reduced frequency number α_m is termed *"reconstruction order"*.

For reconstruction, the flow velocity description at the sensor positions is needed. In polar coordinates the velocity vector can be decomposed into the radial velocity vector $u_\perp \boldsymbol{e}_r$ and the angular vector $u_\parallel \boldsymbol{e}_\varphi$:

$$\boldsymbol{u}(r,\varphi) = u_\perp \boldsymbol{e}_r + u_\parallel \boldsymbol{e}_\varphi \,. \tag{4.29}$$

The partial velocities can be derived from the flow potential:

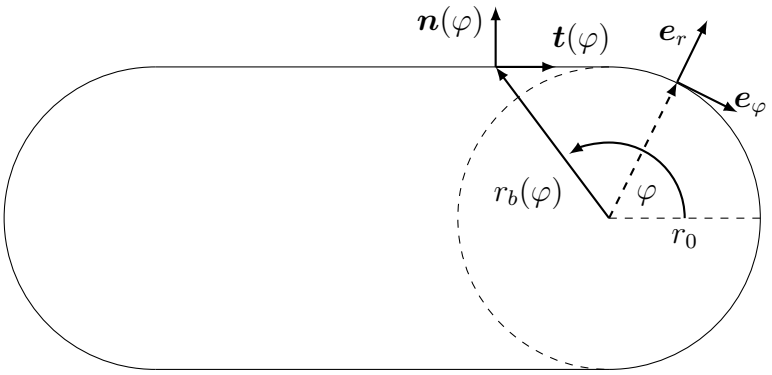


Figure 4.16: Schematic of the robotic body and the circular approximation

$$\begin{aligned} u_\perp &= \frac{\partial \Phi}{\partial r} = \sum_{\alpha=-\alpha_m}^{\alpha_m} \alpha \left(A_\alpha \left(\frac{r}{r_0}\right)^{\alpha-1} - B_\alpha \left(\frac{r}{r_0}\right)^{-\alpha-1} \right) e^{i\alpha\varphi} \\ &= \boldsymbol{g}^+(r,\varphi)\boldsymbol{A} - \boldsymbol{g}^-(r,\varphi)\boldsymbol{B} \\ &= \left[\begin{array}{cc} \boldsymbol{g}^+(r,\varphi) & -\boldsymbol{g}^-(r,\varphi) \end{array} \right] \left[\begin{array}{c} \boldsymbol{A} \\ \boldsymbol{B} \end{array} \right] \end{aligned} \tag{4.30}$$

with

$$\begin{aligned} \boldsymbol{A} &= [A_{-\alpha_m} \cdots A_{\alpha_m}]^T \\ \boldsymbol{B} &= [B_{-\alpha_m} \cdots B_{\alpha_m}]^T \\ \boldsymbol{g}^+(r,\varphi) &= \left[\begin{array}{ccc} -\alpha_m \left(\frac{r}{r_0}\right)^{-\alpha_m-1} e^{i-\alpha_m\varphi} & \cdots & \alpha_m \left(\frac{r}{r_0}\right)^{\alpha_m-1} e^{i\alpha_m\varphi} \end{array} \right] \\ \boldsymbol{g}^-(r,\varphi) &= \left[\begin{array}{ccc} -\alpha_m \left(\frac{r}{r_0}\right)^{-\alpha_m-1} e^{-i\alpha_m\varphi} & \cdots & \alpha_m \left(\frac{r}{r_0}\right)^{-\alpha_m-1} e^{i\alpha_m\varphi} \end{array} \right] \end{aligned} \tag{4.31}$$

$$\begin{aligned} u_\parallel &= \frac{1}{r}\frac{\partial \Phi}{\partial \varphi} = \sum_{\alpha} i\alpha \left(A_\alpha \left(\frac{r}{r_0}\right)^{\alpha-1} + B_\alpha \left(\frac{r}{r_0}\right)^{-\alpha-1} \right) e^{i\alpha\varphi} \\ &= i \left[\begin{array}{cc} \boldsymbol{g}^+(r,\varphi) & \boldsymbol{g}^-(r,\varphi) \end{array} \right] \left[\begin{array}{c} \boldsymbol{A} \\ \boldsymbol{B} \end{array} \right] \end{aligned} \tag{4.32}$$

At the sensors, due to the no-penetration condition, only flow parallel to the surface of the robot is measured. The scalar product of the tangential vector $\boldsymbol{t}(\varphi_i)$ for a sensor $i = 1 \ldots M$ at the sensor position r_i with the flow velocity gives the sensor reading $u_{m,i}$:

$$u_{m,i} = \boldsymbol{u}(r_i, \varphi_i) \cdot \boldsymbol{t}(\varphi_i) \qquad \forall\, , i = 1 \ldots M. \tag{4.33}$$

With the decomposition (4.29) and equations (4.30),(4.32):

$$\begin{bmatrix} \boldsymbol{g}^+(r_i,\varphi_i) & -\boldsymbol{g}^-(r_i,\varphi_i) \end{bmatrix} \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{bmatrix} t_{r,i} + i \begin{bmatrix} \boldsymbol{g}^+(r_i,\varphi_i) & \boldsymbol{g}^-(r_i,\varphi_i) \end{bmatrix} \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{bmatrix} t_{\varphi,i} = u_{m,i}$$

$$\begin{bmatrix} (t_{r,i}+it_{\varphi,i})\boldsymbol{g}^+(r_i,\varphi_i) & (it_{\varphi,i}-t_{r,i})\boldsymbol{g}^-(r_i,\varphi_i) \end{bmatrix} \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{bmatrix} = u_{m,i} \tag{4.34}$$

with $t_{r,i} = \boldsymbol{e}_r \cdot \boldsymbol{t}(\varphi_i)$ and $t_{\varphi,i} = \boldsymbol{e}_\varphi \cdot \boldsymbol{t}(\varphi_i)$.
The combination of all $u_{m,i}$ with $i = 1 \dots M$ leads to a linear equation system:

$$\underbrace{\begin{bmatrix} (t_{r,1}+it_{\varphi,i})\boldsymbol{g}^+(r_1,\varphi_1) & (it_{\varphi,1}-t_{r,1})\boldsymbol{g}^-(r_1,\varphi_1) \\ \vdots & \vdots \\ (t_{r,i}+it_{\varphi,i})\boldsymbol{g}^+(r_i,\varphi_i) & (it_{\varphi,i}-t_{r,i})\boldsymbol{g}^-(r_i,\varphi_i) \\ \vdots & \vdots \\ (t_{r,M}+it_{\varphi,M})\boldsymbol{g}^+(r_M,\varphi_M) & (it_{\varphi,M}-t_{r,M})\boldsymbol{g}^-(r_M,\varphi_M) \end{bmatrix}}_{\boldsymbol{G}_m} \underbrace{\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{bmatrix}}_{\boldsymbol{c}} = \underbrace{\begin{bmatrix} u_{m,1} \\ \vdots \\ u_{m,i} \\ \vdots \\ u_{m,M} \end{bmatrix}}_{\boldsymbol{u}_m}$$

$$\boldsymbol{G}_m \boldsymbol{c} = \boldsymbol{u}_m \tag{4.35}$$

The solution of this linear equation system satisfies the tangential flow condition at the sensor positions. A second condition comes from the boundary condition that no flow can penetrate the surface of a solid object. If the surface is described by the points $(r_b(\varphi),\varphi)$ and the normal vector $\boldsymbol{n}(\varphi)$, the boundary condition can be stated as:

$$\boldsymbol{u}(r_b(\varphi),\varphi) \cdot \boldsymbol{n}(\varphi) = 0 \qquad \forall\varphi. \tag{4.36}$$

The discretization of the boundary, as seen in the boundary element method described in Section 4.4.3, results in a second linear equation system:

$$\boldsymbol{G}_b \boldsymbol{c} = \boldsymbol{0} \tag{4.37}$$

The combination of both conditions leads to a overdetermined system

$$\underbrace{\begin{bmatrix} G_m \\ G_b \end{bmatrix}}_{\boldsymbol{G}} \boldsymbol{c} = \underbrace{\begin{bmatrix} u_{m,1} \\ \vdots \\ u_{m,M} \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\boldsymbol{d}} \tag{4.38}$$

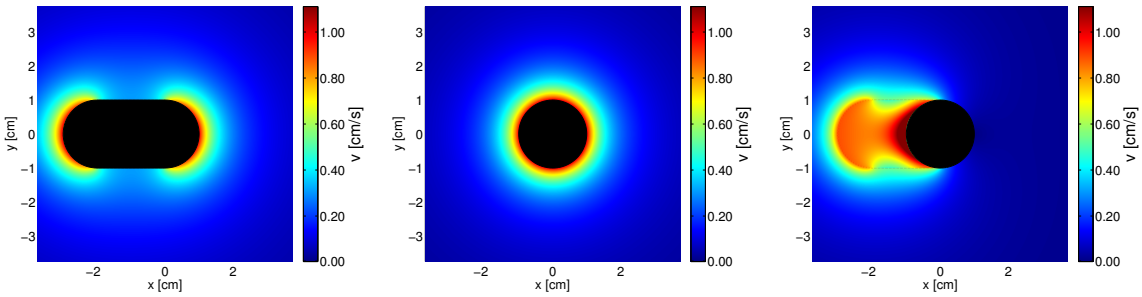
$$\boldsymbol{G}\boldsymbol{c} = \boldsymbol{d}\,, \tag{4.39}$$

with the least squares solution

$$\boldsymbol{c} = (\boldsymbol{G}^T\boldsymbol{G})^{-1}\boldsymbol{G}^T\boldsymbol{d} \tag{4.40}$$

The reconstruction approach has several advantages. First of all, the reconstruction only consists of solving a linear equation system. Here, the pseudo inverse of $\boldsymbol{G}$ is only dependent on the geometry of the problem. Thus it solely has to be computed once. Consequently, at each time step one matrix-vector multiplication is necessary. Furthermore, it is also possible to evaluate the streamline function in a simple fashion, reducing the computational effort and thus latency, which is critical in obstacle avoidance. Second, the least squares solution is expected to reject noise, especially when reconstructing with a low order of α_m. The fact, that each coefficient influences the whole flow field and is not spatially bounded is a clear disadvantage. Because of that, a reconstruction-error from a false sensor reading will affect the whole flow field.

Circular shape



(a) Absolute flow field Snookie **(b)** Absolute flow field sphere **(c)** Difference field

Figure 4.17: Comparison between robot and circular shape [348].

If the submarine body is simplified to a circle, which still leads to a decent representation of the relevant laminar flow around the sensors (see Section 4.2 for turbulence and boundary layer separation), the derived equations for the flow reconstruction are applicable. A comparison of the calculated flow fields for the robot shape and a circle is shown in Figure 4.17. It can be seen that at least for the frontal part the simplification by a circle represents the flow field with negligible deviations. For the circle, the congruency of the normal vector $\boldsymbol{n}(\varphi)$ and the basis vector $\boldsymbol{e}_r$, as well as the congruency of the tangential vector $\boldsymbol{t}(\varphi_i)$ with $\boldsymbol{e}_\varphi$ for every point r_0 on the circle, can be used to rewrite the above equations:

$$u_\perp(r_0) = 0, \tag{4.41}$$

which leads to

$$\sum_{\alpha=-\alpha_m}^{\alpha_m} \alpha \left(A_\alpha - B_\alpha\right) e^{i\alpha\varphi} = 0. \tag{4.42}$$

For $\alpha \neq 0$ the equation is only satisfied for

$$A_\alpha = B_\alpha. \tag{4.43}$$

Considering the congruency of the vectors in the equations with $t_{r,i} = 0$ and $t_{\varphi,i} = 1$, equation (4.34) simplifies to:

$$\left[i\boldsymbol{g}^+(r_i, \varphi_i) + i\boldsymbol{g}^-(r_i, \varphi_i)\right] \boldsymbol{A} = u_{m,i} \tag{4.44}$$

Similar to equation (4.39), the measurements can be combined, taking advantage of (4.43):

$$\boldsymbol{A} = (\boldsymbol{G}_m^T \boldsymbol{G}_m)^{-1} \boldsymbol{G}_m^T \boldsymbol{u}_m \tag{4.45}$$

The representation as a circle has the advantage of reducing the complexity for the reconstruction of the flow field around the robot. A lower reconstruction order can already satisfy the boundary conditions, which leads to a better representation of the surrounding flow field. A comparison between the robot shape and circle can be seen in Section C.4.2. This comparison, as well as following comparisons, utilize a flow error field [348] as quality criterion. For each discretized section of the simulated area, the norm of the error velocity $\|\boldsymbol{e}\|$ is calculated. The error velocity $\boldsymbol{e} = \boldsymbol{u} - \hat{\boldsymbol{u}}$ is the difference of the fluid velocity $\boldsymbol{u}$ computed by the boundary element method (reference) and the estimated reconstructed velocity $\hat{\boldsymbol{u}}$. For an ideal reconstruction, both velocities would be identical. The region in which the error velocity $\boldsymbol{e}$ is less than 10% is a bounded area close to the sensors. Only within this area the reconstruction is a valid representation of the flow field and is further called field of view (FOV). Information outside this FOV is discarded. The FOV is dependent on the number of sensors M, the reconstruction order α_m and, since the Fourier series includes $\frac{r^\alpha}{r_0^{\alpha-1}}$, also on the diameter of the robot. Figure 4.18 describes the FOV over the reconstruction order for a number of sensors.

Influence of noise

It is obvious that the quality of the reconstruction is subject to the level of noise in the sensor readings as well as the uncertainties in the robot's velocities.

To evaluate the influence on the sensors, an artificial error on the sensor measurement $\boldsymbol{u}_m$ is added. The noise is a zero-mean independent normal distribution with standard deviation σ (units in $[\mathrm{cm\,s^{-1}}]$). As a measure of the reconstruction quality, the previously defined FOV is taken, with a threshold $< 10\%$ of the error velocity $\boldsymbol{e}$. The noise introduces a random distortion of the reconstructed flow field, usually in the outer perimeter of the FOV. To estimate the non-corrupted area, 25 trials with noisy sensor data are computed and an average FOV area is calculated, see Figure 4.19.

The second influence is the noise in the robot velocity measurements (v_x, v_y), as the transformation in the absolute flow field relies on the relative fluid velocity measurements

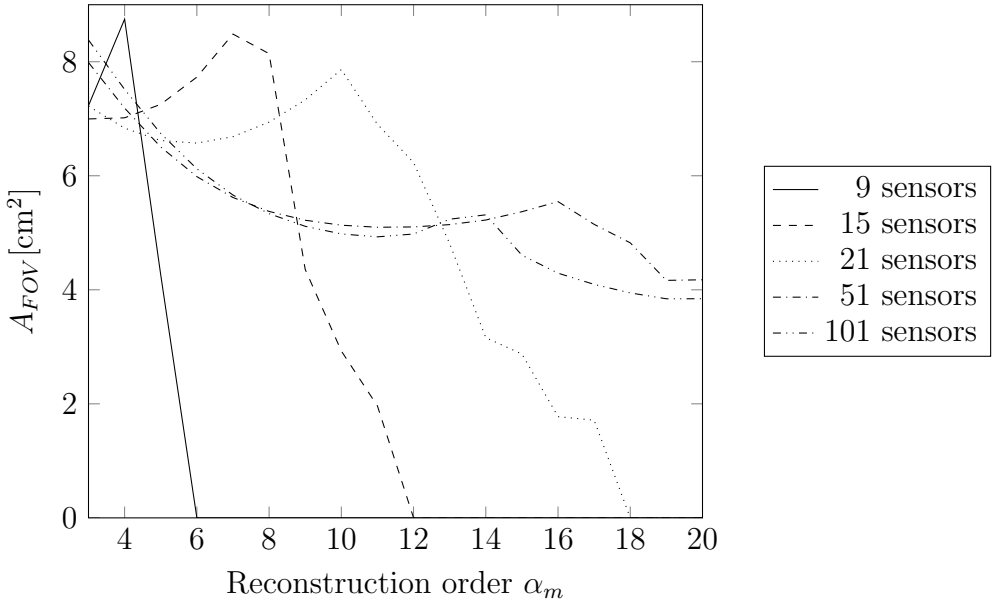


Figure 4.18: Comparison of the area of the FOV over the reconstruction order for different numbers of sensors.

and the estimated ego-motion of the robot. Figure 4.20 shows distortions in the reconstructed flow field for uncertainties c in the velocity $v_x' = v_x + c$, with $v_y = 0$. The distortions affect the whole area due to the overall shift in the transformation.

4.4.5 Object Reconstruction

After reconstructing a flow field around the front of the robot, see Figure 4.21 for an example, the next step is to find static obstacles in the flow field. One way to achieve this in 2D is to utilize the properties of streamlines. Streamlines in fluid dynamics are the curves at each instant of time at which the flow velocities are tangential. Thus, because flow is always tangential to obstacles, there is a streamline that represents the wall form. Streamlines do not intersect except at stagnation points. Therefore, it is especially not possible that a streamline intersects the surface of an obstacle. According to Lamb [417], if the flow can be described by a potential function Φ, then there always exists the dual function Ψ that is called the stream function and is defined over the relationship

$$u_x = \frac{\partial \Psi}{\partial y}, \qquad u_y = -\frac{\partial \Psi}{\partial x}. \tag{4.46}$$

with $\boldsymbol{u} = (u_x \; u_y)$ being the velocity vector of the fluid in 2D cartesian coordinates. The fact, that the contours $\Psi = c$ with some constant c represent the streamlines give the function its name. The streamline function can be stated as:

$$\Psi(r, \varphi) = -i \sum_{\alpha} \left(A_\alpha \frac{r^\alpha}{r_0^{\alpha-1}} + B_\alpha \frac{r^{-\alpha}}{r_0^{-\alpha-1}} \right) e^{i\alpha\varphi}, \tag{4.47}$$

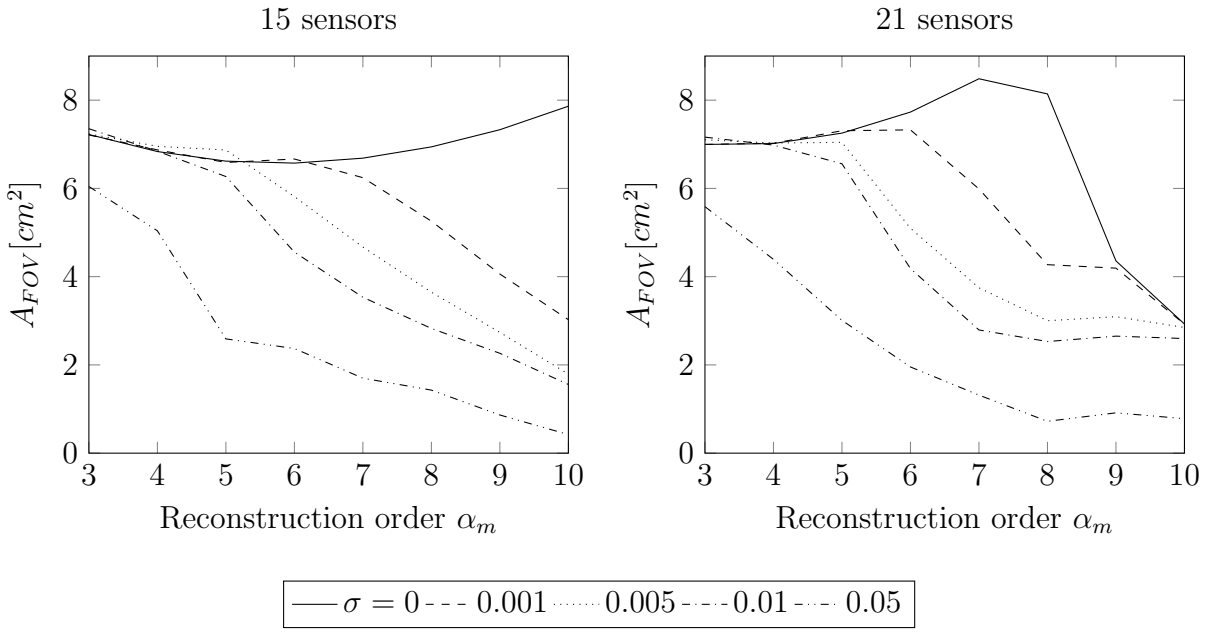
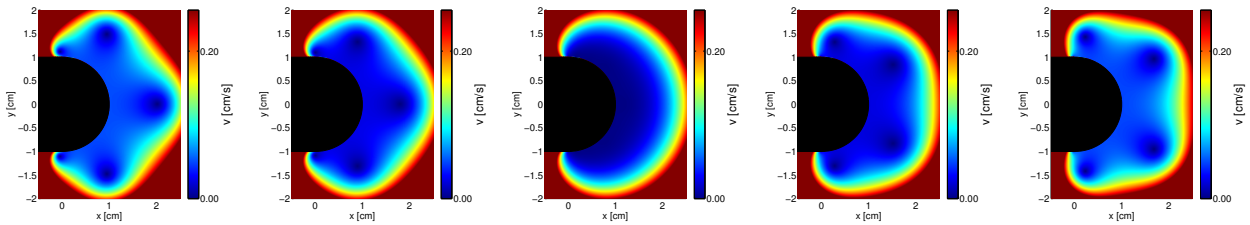


Figure 4.19: Comparison between noise influence on 15 and 21 sensors for different levels of noise with standard deviation σ.



(a) $-0.05\frac{\text{cm}}{\text{s}}$ **(b)** $-0.025\frac{\text{cm}}{\text{s}}$ **(c)** $0\frac{\text{cm}}{\text{s}}$ **(d)** $0.025\frac{\text{cm}}{\text{s}}$ **(e)** $0.05\frac{\text{cm}}{\text{s}}$

Figure 4.20: Reconstruction error flow fields for different variations c of the robot velocity v_x

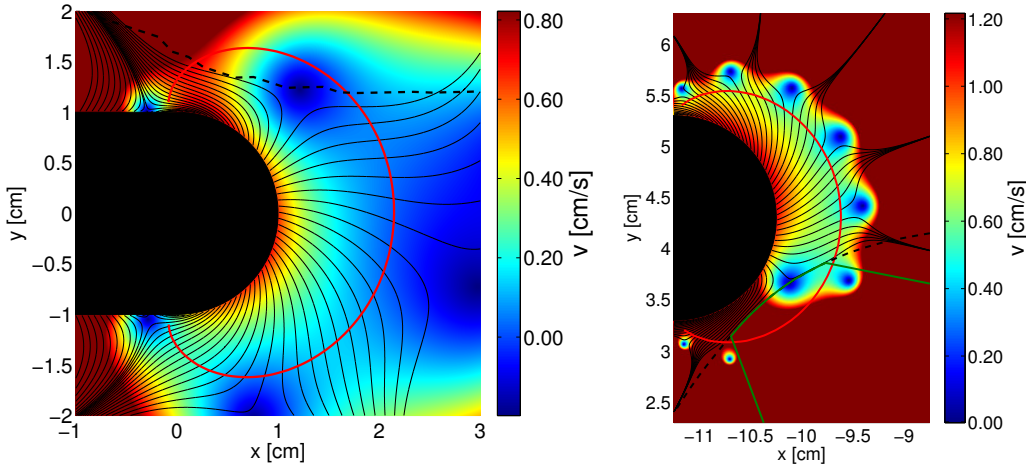
which is only different from the potential function $\Phi(r, \varphi)$ by a factor of $-i$.

In order to extract the boundaries of a solid object, the properties of the potential flow fields are utilized:

- Streamlines cannot pass through walls, i.e. they are parallel to walls
- Streamlines start and end on moving objects
- No extrema in the flow potential inside the fluid exist

Applied to this case, an algorithm can be stated as follows:

1. Find the set of extrema $\boldsymbol{P}_e$ where $u_\perp(\boldsymbol{P}_e) = u_\parallel(\boldsymbol{P}_e) = 0$. As there is no easy to find explicit solution for this set of equations, this has to be done numerically. See Figure 4.22a.



(a) Obstacle above in FOV perimeter

(b) Obstacle below in direct vicinity

Figure 4.21: Reconstructed flow field examples around the robot (black solid) with an obstacle (black dashed line). The first example shows an obstacle above of the robot within the assumed field of view of the lateral line system (red line). The second example is of an object in immediate vicinity below the robot. Simulation with $n = 21$ sensors and reconstruction order $\alpha = 5$

2. Evaluate the streamline function at each extremum i to get $s_i = \Psi(P_{e,i})$

3. Get the streamlines in the vicinity δs of the extrema. Therefore, find the contours $\Psi = s_i \pm \delta s$. See Figure 4.22b. The parameter δs has to be chosen manually, a typical value that provides good results is around 0.1.

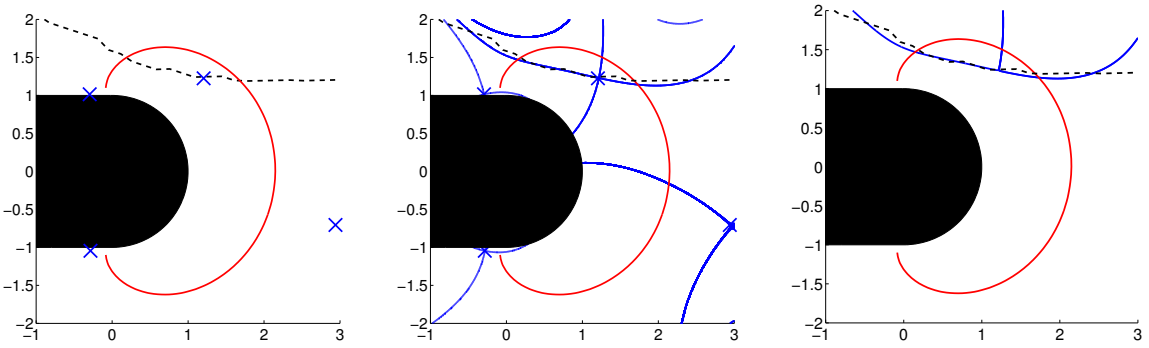
4. Remove all streamlines that go towards the robot, i.e. all streamlines that includes points where $r < r_0$. Furthermore, remove all streamlines that have no point in the FOV. See Figure 4.22c.

5. As a result, the streamlines that are left and are in the valid field of view represent walls.

Automatic Determination of Reconstruction Order

As the reconstruction fails for lower reconstruction orders, for example in the presence of a very close obstacle or misses important details for more complex obstacles, it may be preferable to choose the order automatically. Therefore a way to evaluate the current reconstruction quality is necessary. One possible way is to have a look at the residual from the least squares solution for the flow field in (4.45). After calculating the solution for $\boldsymbol{A}$, the residual vector is defined as:

$$\boldsymbol{r} = \boldsymbol{G}_m \boldsymbol{A} - \boldsymbol{u}_m. \tag{4.48}$$



(a) Extrema **(b)** Streamlines **(c)** Reduced Streamlines

Figure 4.22: Algorithm of static obstacle detection:
Figure 4.22a Find set of extrema in flow potential.
Figure 4.22b Find streamlines in their vicinity.
Figure 4.22c Reduce streamlines.

With the least square estimate, the norm of this vector $\|\boldsymbol{r}\|$ is minimized. Now if the order is not sufficient to fit the potential flow model to the sensor readings, then this error norm increases. Thus, this value can be used to determine whether the current order is sufficient or if the order has to be increased. To underline that, Figure 4.23 shows the residual norm over the reconstruction order of the two examples in the last section. One can see, that the residual error is high for low orders, but then it decreases monotonically.

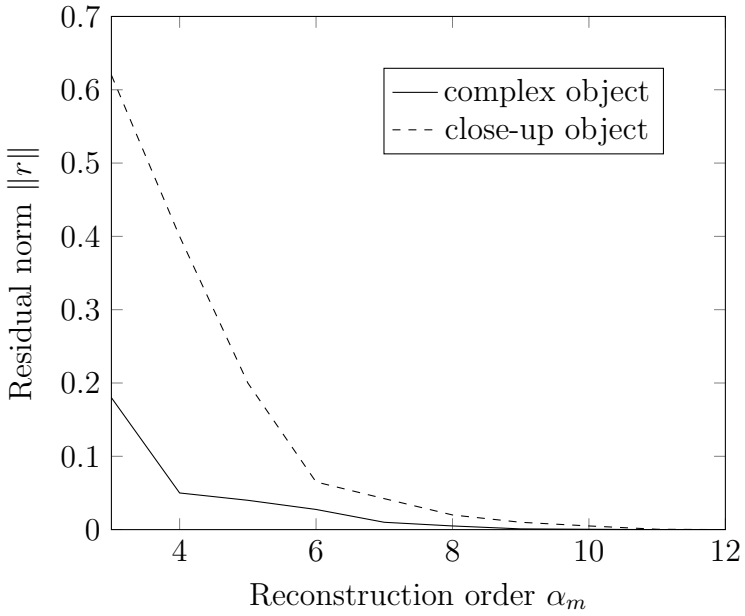


Figure 4.23: Norm of the residual over the reconstruction order for two examples

This leads to an automatic reconstruction order determination :

foreach $k = \alpha_{min} \rightarrow \alpha_{max}$ **do**
  $\boldsymbol{A}(k) = (\boldsymbol{G}_m(k)^T \boldsymbol{G}_m(k))^{-1} \boldsymbol{G}_m(k)^T \boldsymbol{u}_m$;
  $\boldsymbol{r}(k) = \boldsymbol{G}_m(k)\boldsymbol{A}(k) - \boldsymbol{u}_m$;
  if $\|\boldsymbol{r}(k)\| < \delta$ **then**
    $k, \boldsymbol{A}(k)$
  end
end

This leaves only the minimal and maximal reconstruction order α_{min} and α_{max} respectively and the threshold value δ to be defined. For a δ of 0.001, the algorithm gives an order of 9 and an order of 10 for the complex-form and the close-up example respectively. When choosing the threshold value, one has to keep in mind the level of noise that is present. If δ is chosen too small, then always a too high order is calculated. Although this algorithm might lead to an overhead in calculation, most of the time α_{min} will be sufficient. In other situations, the reconstruction will have to be calculated for more orders but as mentioned before, each reconstruction is just a matrix-vector multiplication with a constant matrix for each order. Furthermore, the obstacle detection only has to be calculated once, since it is applied after the flow reconstruction.

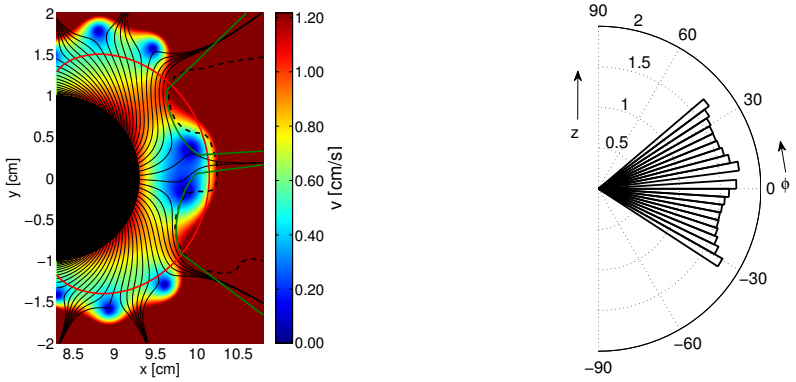
4.4.6 Mapping

Having shown that it is possible to reconstruct the local surroundings with the artificial lateral line system, the next step is to utilize the reconstruction to build a map of the environment. This again seemingly mimics the behavior of the blind cave fish, when brought into an unfamiliar environment, and shows the capabilities of the sensory system.

Sensor Virtualization

To further use the output of this sensor model, the detected obstacle is discretized in polar coordinates with respect to the robot. This is done in a fashion, that measurement pairs of (z_j, ϕ_j) for each of the virtual sensors j are generated. This representation has the advantage, that it is the same as for range finders like light detection and ranging (LIDAR) and sound navigation and ranging (SONAR) which are commonly used in robotics. This is exploited in this section to apply a mapping algorithm to the artificial lateral line system.

Simultaneous localization and mapping The problem of mapping describes the ability of a robot to build up a (metric) map of its surroundings, which enables the robot to safely plan a path and navigate without collision. This work follows the mapping algorithm described by Thrun *et al.* [403]. It is assumed that the robot movement is known exactly. At each time-step the sensor input is incorporated in a map. Therefore the space is discretized in x and y direction and for each cell m_i a probability of this cell being occupied is associated $p(m_i = occupied)$.



(a) Sensor Model Output(green boundaries)

(b) Discretized over angle

Figure 4.24: Discretization of the sensor model output in the angular space, with 50 equidistant *virtual sensors*

To update the probability, the previously described discretized sensor model is used. Each of the virtual sensors have a measurement cone with a measured obstacle distance z (If there is no measurement for one virtual sensor, then a maximum value of z_{max} that represents the maximum measuring distance is assumed). For all cells that are in the cone, but are closer than z to the robot, the occupancy probability is decreased. For all cells in the cone that have a distance of $z \pm \beta$ to the robot, the occupancy probability is increased. All other cells are kept constant. This procedure is repeated for all the virtual sensors and is depicted in Figure 4.25 for a simple example.

In general there will be no prior knowledge of the robot's trajectory to build up a correct map and also no prior map available for localization. Therefore, both tasks have to be performed simultaneously, which in literature is referred to as the simultaneous localization and mapping problem. Mathematically speaking, the problem is to estimate the probability distribution

$$p(x_t, m | z_{1:t}, u_{1:t}) \tag{4.49}$$

of the current robot pose x_t and the built-up map m given all prior measurements $z_{1:t}$ and robot controls $u_{1:t}$.

One way to represent this distribution is to use a set of particles. Each particle occupies a part of the state space (x_t, m_t) and the density of the particles in this space represents the probability. As the dimensionality of the space of all possible maps m is infinitely large, it is infeasible for a straightforward implementation of this. Therefore, the distribution is factorized in Rao-Blackwellized particle filters:

$$p(x_t, m | z_{1:t}, u_{1:t}) = p(x_t | z_{1:t}, u_{1:t}, m) p(m | z_{1:t}, u_{1:t}, x_{1:t}). \tag{4.50}$$

With this factorization, the problem can be decoupled in a localization step to determine

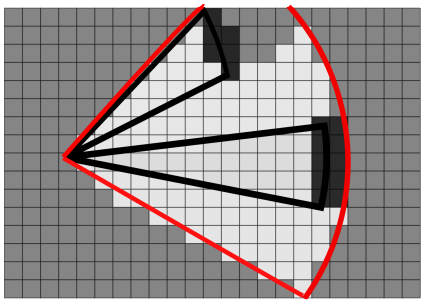


Figure 4.25: Simple mapping example for two measurement cones in the robots field of view (red) in an occupancy grid. All cells in a direction which are closer than the measurements (or where are no measurements) are weighted free space (white). All cells near to the measurement are weighted as occupied (black). All other cells are unchanged (gray).

$p(x_t|z_{1:t}, u_{1:t}, m)$ and a subsequent mapping step to update m. Each particle is building up its own map while assuming that its trajectory is correct. With each new measurement, the agreement with each particles map is calculated with a measurement model. This agreement is denoted importance weight $w^{[i]}$ for each particle i. Up to now, a particle set where each particle has a associated importance weight has been generated. The next important step is resampling of the particle set. In order to do this, each particle gets a index associated with it. Now random numbers are generated from a discrete probability distribution, where each index i gets drawn according to the importance weight $w^{[i]}$. If the index of a particle is drawn, then it is added to a updated particle set. This is repeated until a desired number of particles is reached. This step is necessary to eliminate unlikely particles in a probabilistic fashion. Note that it is possible (and in fact frequent) that a particle gets drawn more than once in the new set.

In a nutshell, the SLAM algorithm at each time-step can be summarized as following:

1. Predict motion of each particle with a motion model
2. For each particle, calculate the importance weight from measurement and map
3. Update all maps
4. Resample particle set

Motion Model In order to deal with uncertainty in the robots motion, a probabilistic motion model that includes these uncertainties is necessary. This model can be described by a probability distribution from the last state x_{t-1} to the current state x_t given the desired robot motion u_t.

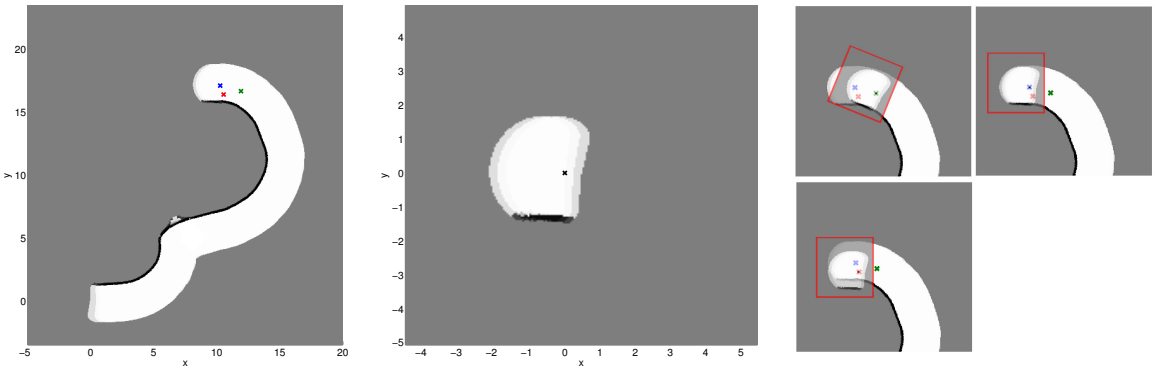
$$p(x_t|x_{t-1}, u_t) \tag{4.51}$$

For the particle filter approach it is furthermore necessary to generate particles form this distribution. With increasing time, the estimate of the robot location gets more and more uncertain which is indicated by a broadened probability distribution. Here, a simple odometry motion model described in [403] is used. This model splits the motion from

the last time step into three basic motions: rotation, translation and again a rotation. Each of these motions is faced with noise with a zero-mean normal Gaussian distribution. The standard deviation of these distributions describe the level of uncertainty in the robot motion.

Measurement Model For the measurement model, the idea of map matching [403] is pursued. Thus, a local map is generated over a short period of time in order to keep the mapping error due to uncertainty in motion small. This local map is subsequently used to compare to the global map in order to get an estimate on how well both maps coincide. This can be interpreted as

$$p(z_t|x_t, m) = p(m_{local}|x_t, m_{global}) \tag{4.52}$$



(a) Global map **(b)** Local map **(c)** Map matching

Figure 4.26: Map matching between a global (a) and a local (b) map. (c) shows the overlay of local and global map for the three different particles. Both top images show a good match between both maps, but the top right has the best match. The bottom image shows no match of both maps.

To get an estimation of $p(m_{local}|x_t, m_{global})$ a measure of the similarity between two maps is the map correlation value. Therefore the local map has to be transformed in a global coordinate system for each particle. Then the map correlation can be calculated with [403]

$$\rho_{\boldsymbol{m},\hat{\boldsymbol{m}}} = \frac{\sum_{x,y}(\hat{m}_{x,y} - \bar{m}) \cdot (m_{x,y} - \bar{m})}{\sqrt{\sum_{x,y}(\hat{m}_{x,y} - \bar{m})^2 \sum_{x,y}(m_{x,y} - \bar{m})^2}} \tag{4.53}$$

where

$$\bar{m} = \frac{1}{2N} \sum_{x,y} (\hat{m}_{x,i} + m_{x,i}). \tag{4.54}$$

Here, $\hat{m}_{x,y}$ and $m_{x,y}$ denote cells of the local and global map respectively containing position (x, y). $\bar{m}$ is the mean map value over both maps to account for the fact that the majority of cells are free space.

The variation of correlation $\rho_{\boldsymbol{m},\hat{\boldsymbol{m}}}$ between $[-1, 1]$ results in a choice of

$$p(m_{local}|x_t, m_{global}) = \max\{\rho_{\boldsymbol{m},\hat{\boldsymbol{m}}}, 0\}. \tag{4.55}$$

Figure 4.26 shows an example of global and local map and how the similarity can be used to weigh particles. The advantage of this approach is, that it is sensor independent and the only requirement is that it is possible to build a correct map with it. Furthermore, it integrates several measurements to get one larger, more reliable measurement with more included overall information. This overcomes problems in the particle filter of SLAM with overconfidence.

4.4.7 Experimental Evaluation

The following simulations provide insights on the reconstruction capabilities of an artificial lateral line system with the numerical stimulus and reconstruction method. In contrast to the analytical solution of the previous section, the detection and reconstruction of complex arbitrary shapes is enabled. Furthermore, the reconstruction capabilities are also employed in simulating robot navigation in a map with two large obstacles.

Reconstruction of Obstacle

An example of the reconstruction of a complex form can be seen in Figure 4.27 for different numbers of the reconstruction order.

Close-up Obstacle Reconstruction

A simulation of an obstacle close to the robot can be seen in Figure 4.28.

Discussion

As can be seen, it is possible to reconstruct the wall geometry, but wall-details get lost in the process, depending on the reconstruction order. With a higher reconstruction order, there are more terms in the sums for the water velocities $u_\perp$ and $u_\parallel$. Thus, there are more points possible were $u_\perp = u_\parallel = 0$ and there are more extrema in the potential function Φ. As extrema, that lie in the FOV lie on a wall as shown before, there can be more extrema on the wall and therefore there is more information implied. In Figure 4.27d this effect is shown clearly, as the bay of the wall can only be separated with two extrema. The reconstruction of the bay of the obstacle demonstrates how the reconstruction order affects the results. For orders of 3 and 5, the whole complex form is reconstructed as one big obstacle, the bay gets *smoothed* out. For an order of 7 one can start to see the two forms and for an order of 9, the two parts of the form are split up, as they are not connected in the field of view. Thus, for lower orders one gets smoother results, but for higher orders one can infer details of the forms. Limitations of a fixed reconstruction order can be seen in the close obstacle case. Although, for lower orders, the obstacle can be detected and up to a certain point reconstructed, there are artifacts in the rest of the prior calculated

FOV, which are not result of real obstacles. This can be addressed with either a higher reconstruction order or with an artificially decreased FOV.

Simulation of Mapping

After implementing the algorithms briefly discussed in the last sections, several simulations in 2D were performed to evaluate the feasibility of SLAM with an artificial lateral line system. Therefore, the odometry as well as the sensor measurements were falsified with zero-mean Gaussian distributions. An example of a simulation with 200 particles where a robot is moving two times around an obstacle can be seen in Figure 4.29. Another example of a more general trajectory is shown in Figure 4.30. This is different to the prior simulation insofar as now also times without feedback of an obstacle are included. Also, there are two different times where a loop closure happens.

Discussion

This simulation was a first attempt to use an artificial lateral line system for mapping of the environment. It is shown that the reconstruction capabilities of the simulated sensor system are sufficient to build a detailed map of the surroundings of the robot. When comparing the pure-odometry mapping with the SLAM mapping in Figure 4.29, one can see that only the SLAM algorithm is able to build up a consistent map. However, because of the small overlap of local and global map, the particle distribution is broadening over time until a loop closure, i.e. a revisit of an already known place is performed. This makes it impractical for mapping of large environments without regular loop closures.

The second scenario deals with the short range of the sensory system. Before the first loop closure, the robot travels a distance without wall feedback, where the position distribution broadens strongly. After loop closure, the robot is again able to correctly localize itself. Following self-localization, again a phase without feedback and a phase with feedback of an unknown obstacle lead to another loop closure. As now the time between loop closures was much higher, leading to a higher accumulated error, the SLAM algorithm is not able to perform a successful loop closure anymore. This could possibly be overcome with usage of more particles, such that the probability that the correct particle survives gets higher, but at the drawback of higher computational and memory costs.

4.4.8 Summary

As soon as speed and shape of surrounding objects are available, the lateral line can contribute to more complex tasks involving more than just detecting the pure presence and classifying an object, e.g. self localization and map formation. For the biological counterparts, some behavioral experiments are already available [399, 421, 437, 422]. Using a novel flow reconstruction just described, similar capabilities can be implemented on underwater robots such as *Snookie*. The main challenge in this section was the inversion of the flow field computation, i.e. the estimation of the surrounding flow field from measurements on the surface of the robot. Previous work on reconstruction methods required strict assumptions about the shape and the number of objects to be successfully reconstructed.

The presented method allows the reconstruction of arbitrary solid stationary boundaries. The reconstruction is also used to extend the state-of-the-art to the ability of creating a map of the environment with a lateral line system. Mapping and map correlation allow the application of the SLAM approach, providing a widespread toolset in robotics to the challenge of navigating solely based on flow field and odometry information. Verification of the simulated results in experiments is pending.

Although restrictions on the number and shape of objects are lifted, several limitations apply to the proposed method. First of all, the fluid is assumed to be ideal and irrotational. The resulting properties of the fluid that it is inviscid, incompressible and has no circulation flows can be approximately attributed to water (with the sensors placed in the Euler flow regime). A more critical assumption stemming from the fluid properties is, that flow is laminar and without perturbations. This is a condition that is hardly given if the fluid is initially not at rest, i.e. stagnant water, and limits the application range for underwater robotics. It should be noted that this condition can be found in the natural habitat of Mexican Cave fish. Experiments by Windsor *et al.* [357] indicate similar limitations for the blind Mexican Cave fish, as tail-beating of the fish close to a wall, disturbing the flow field, leads to a collision probability of 73%. Collision avoidance works best if the fish is in gliding phase, reducing the collision probability to 11%.

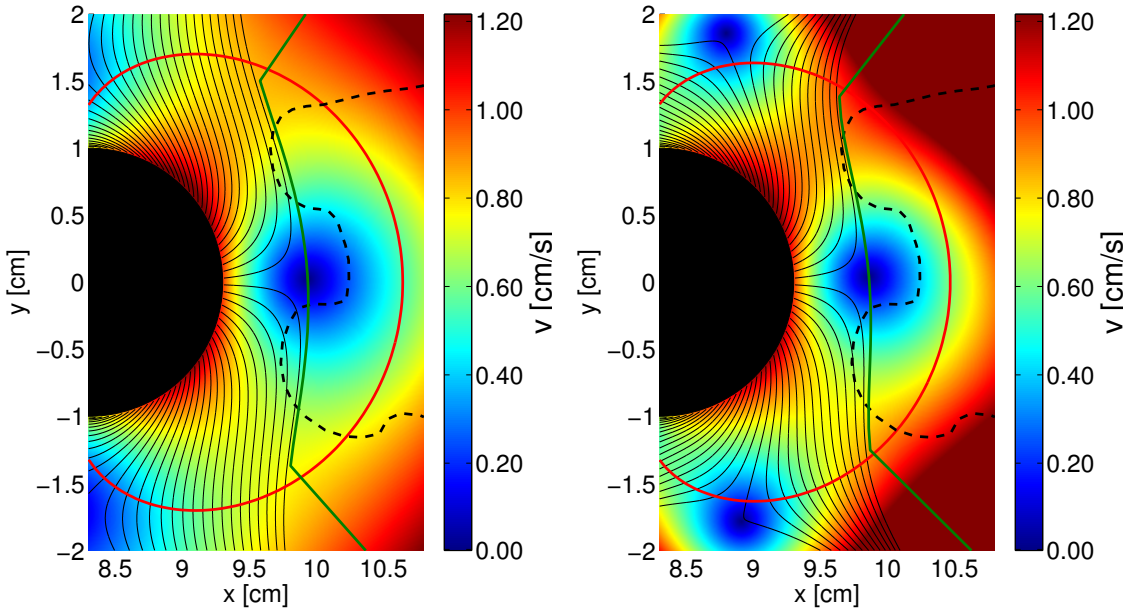
A second limitation is that each coefficient in the reconstruction affects the whole flow field and is not spatially bounded. Falsified readings of a single sensor therefore distort the whole reconstructed flow field. This requires a correction of the flow field over subsequent calculations. Using Kernels as in [347] could counteract these distortions.

Third, as has been shown in the SLAM simulations, for a successful localization the presence of obstacles is crucial. As this is true for all robot localization tasks, here the dampening characteristics of the medium requires close distance to objects for gathering spatial information. In water, the structural information on the object is dampened frequency- and distance dependent in the flow field, leading to significant information loss outside the field of view established in Section 4.4.4. So in order to apply this in a real world situation it may be necessary for the robot to pursue an active localization. This means, that the robot should preferably move along walls to ensure feedback of the environment. On a side note, this wall following behavior can be observed in blind and sighted (in dark conditions) Mexican tetra during exploration of new environments [437].

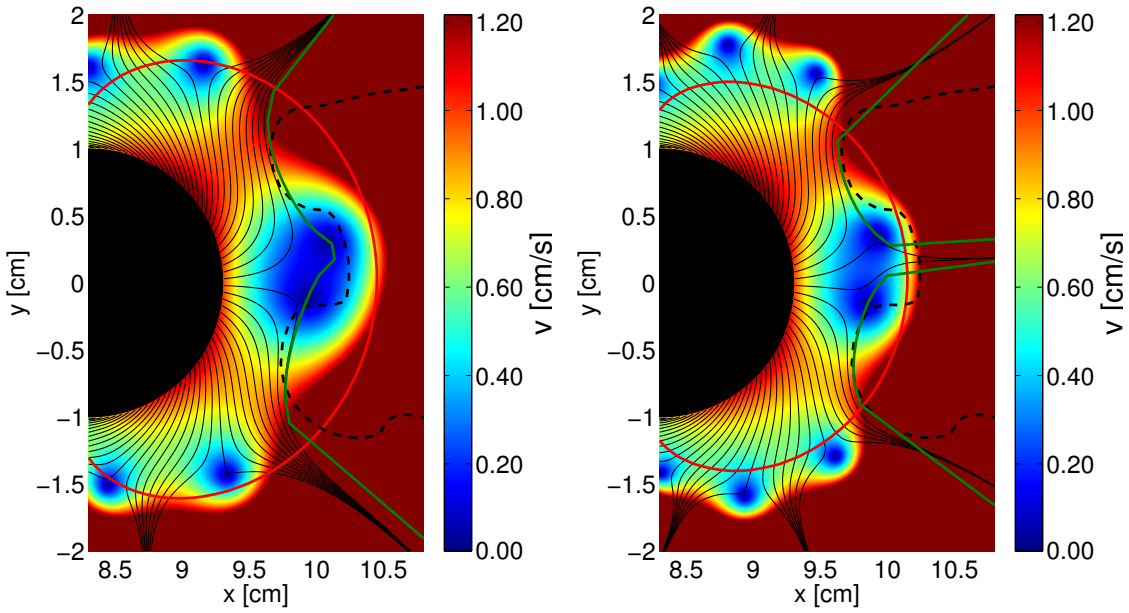
Last, the introduced approach is limited to a quasi-2-dimensional method. While it provides a significant reduction in the computational complexity, future work aims at an implementation in 3D.

The results of this section built upon the previous sections. The initial challenge of transferring an artificial lateral line sensor system to underwater robotics was approached in three steps, namely sensor development, object detection and object reconstruction, following the biomimetic bottom-up approach. Results from the sensor characteristic determination and object detection experiments showed the feasibility of the transfer and suitability of the chosen approach. While the flow field reconstruction provides a general solution to compute the flow field from measurements on the boundary of an object, the special design of *Snookie* lead to a circular simplification. Using the sensor properties and shape of *Snookie*, arbitrary shaped obstacles were detected and reconstructed in

simulations, showing novel information extraction possibilities from the *hydrodynamic image*. Simulations in this work also demonstrate successful mapping and self-localization of an unknown environment with a lateral line sensor system, significantly expanding the state-of-the-art in biomimetic underwater robotics. Results of these experiments also allow feedback for the underlying biological studies, giving an assessment of the perceptional capabilities of a lateral line system and analogies in behavior.

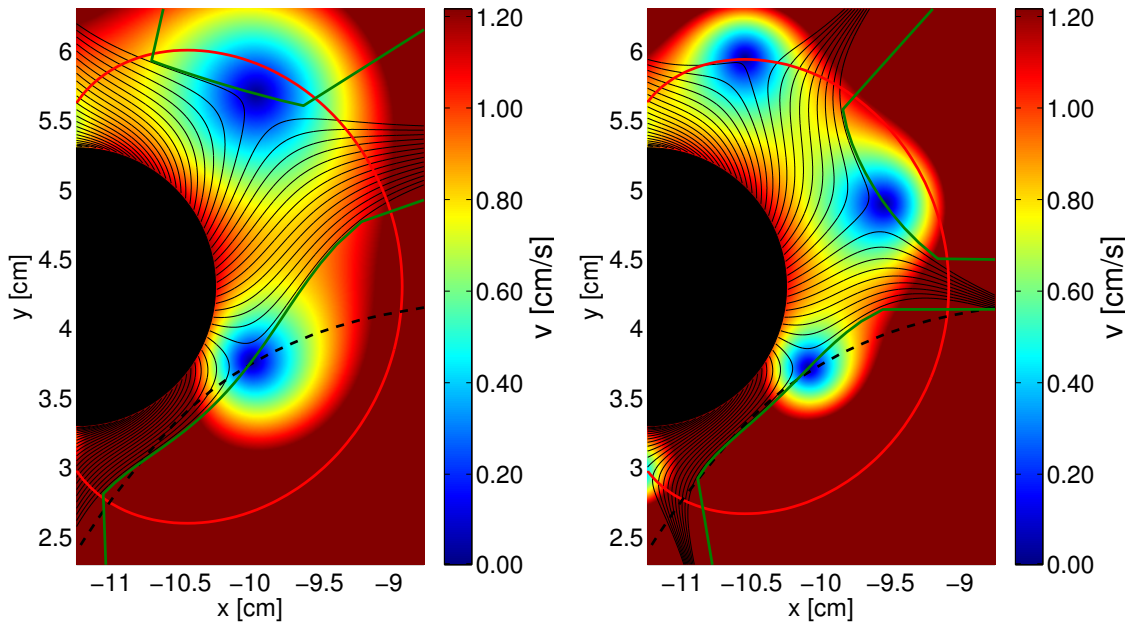


(a) $\alpha_m = 3$ **(b)** $\alpha_m = 5$

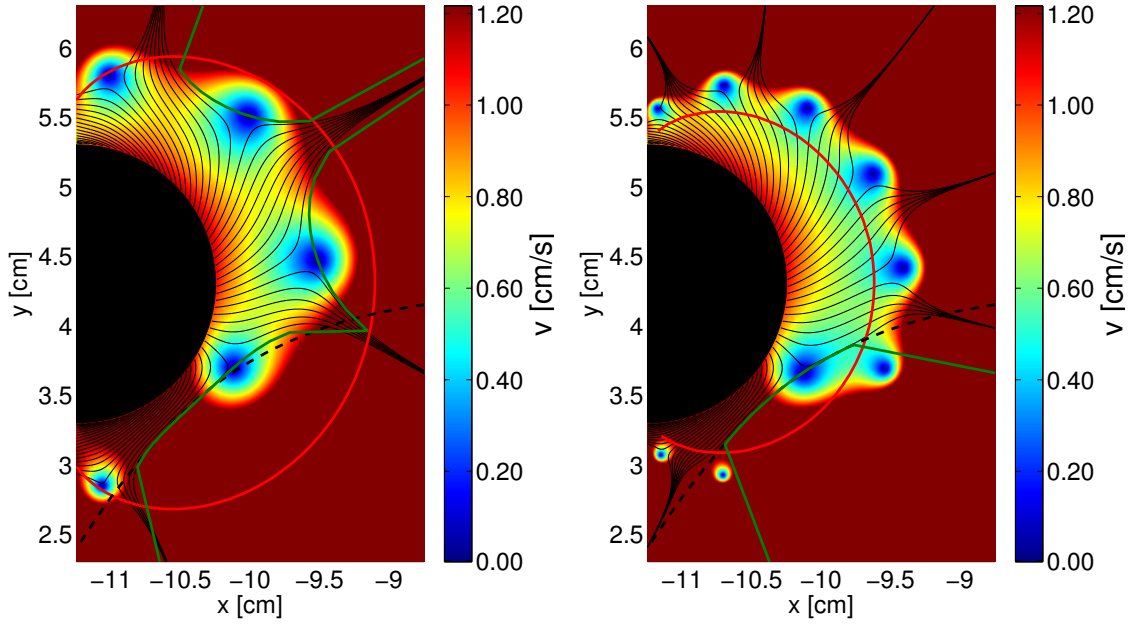


(c) $\alpha_m = 7$ **(d)** $\alpha_m = 9$

Figure 4.27: Reconstruction example with a complex form with 21 sensors and different reconstruction orders. Dashed black lines indicate the obstacle, green line indicates the reconstructed obstacle.

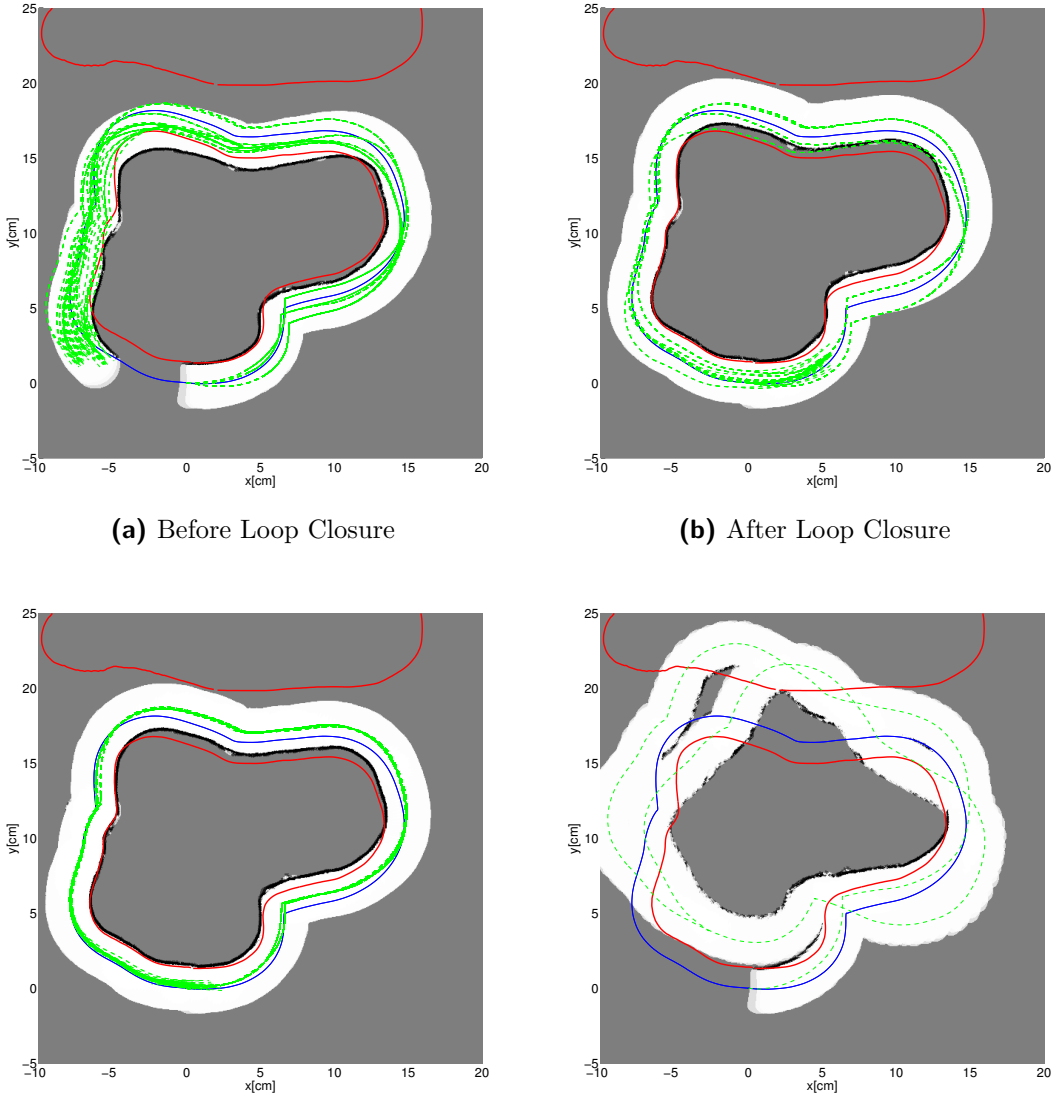


(a) $\alpha_{max} = 3$ **(b)** $\alpha_{max} = 5$



(c) $\alpha_{max} = 7$ **(d)** $\alpha_{max} = 10$

Figure 4.28: Closeup - reconstruction example with 21 sensors and different reconstruction orders. Dashed black lines indicate the obstacle, green line indicates the reconstructed obstacle, red lines the field of view.



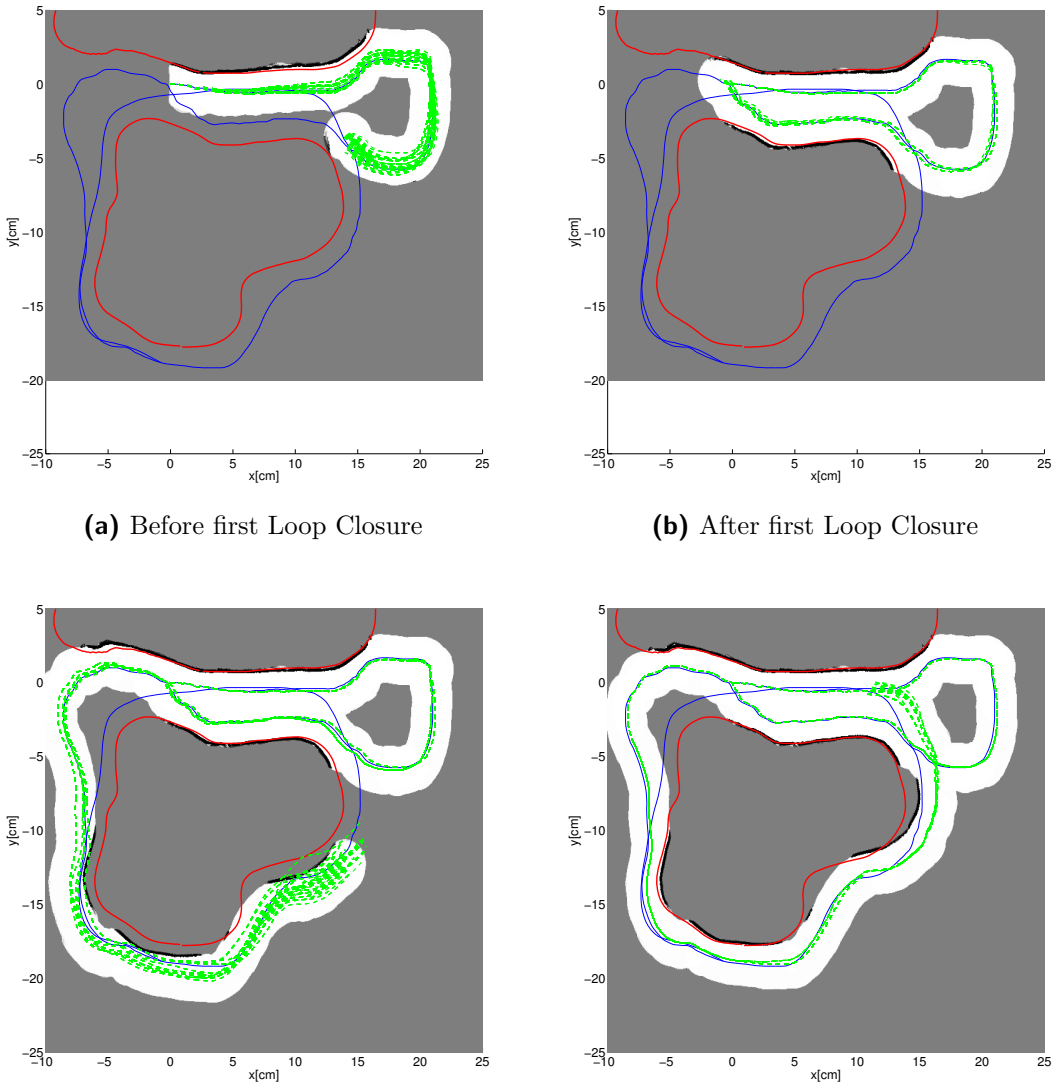
(a) Before Loop Closure

(b) After Loop Closure

(c) Complete

(d) Pure Odometry

Figure 4.29: Simulation results. In each map, the red line denotes real obstacles, black cells are occupied, white cells are free space, gray cells are unknown. Green are the paths of each particle, blue is the ground-truth robot trajectory (2x around object). (a) shows the distribution and potential map before loop closure. (b) shows the distribution shortly after loop closure. (c) shows the final result after circling the object two times. (d) shows the mapping result based on pure odometry data.



(a) Before first Loop Closure

(b) After first Loop Closure

(c) Before first Loop Closure

(d) After first Loop Closure

Figure 4.30: Simulation results. In each map, the red line denotes real obstacles, black cells are occupied, white cells are free space, gray cells are unknown. Green are the paths of each particle, blue is the ground-truth robot trajectory. (a) shows the distribution and potential map before first loop closure. (b) shows the distribution shortly after first loop closure. (c) shows the distribution and potential map before second loop closure. (d) shows the distribution shortly after second loop closure. The accumulated error between first and second loop closure could not be fully compensated.

5 Conclusions

The process of evolution has driven man and animal to develop solutions that are sufficient or even optimal for challenges arising in their natural habitat or social structures. The key element in this work is to transfer said solutions for particular challenges related to robotics – stated in Section 1.3 – to technical systems via a biomimetic design process. Biomimetic design provides a structured approach for deriving a technical solution to a problem from an expert biological system (top-down) or for developing novel techniques from a biological inspirational source (bottom-up). In this work, both methods were used to tackle particular challenges in robotics, which are addressed in three separate chapters. The main solutions and findings are chapter specific and were discussed in the respective sections. This chapter provides a summary of the main contributions in this work. Additionally, possible directions of future research work are outlined.

5.1 Summary of Contributions

This section briefly summarizes the contributions of this thesis. While the presented ideas and methods have been shown to provide solutions to the specific challenges in this work, they are not restricted to these particular implementations.

In Chapters 2 and 3, the top-down approach was applied to the field of HRI. The key aspect of this work was the enhancement of the perception of a biomimetic robotic head as a social actor. Humans are experts in interaction, therefore the goal was to transfer findings from social sciences, anatomy and semiotics to social robots. Necessary intermediate steps included the identification of analogies, abstraction of function principles and implementation in a technical context.

In Chapter 2, frameworks were introduced for the synthesis of implicit communication modalities with both face and neck, enabling a robot to express emotions and other non-verbal signals. The FACS and motor activation based expression synthesis in Section 2.3 is independent of the actual display, being it robot or virtual agent, and its configuration, as long as a mapping to the joint space can be established. With biological analogies for communication available in animals, the framework was extended to include zoosemiotic signals. They are generalizable to any feature that can be associated with a zoomorphic feature and are not necessarily limited to the face. Furthermore, the set of implicit communication signals for a neck given in Section 2.4, while formulated specifically for the two-link 5 DoF model, can be transferred at least in a subset to any neck model. Experimental evaluations showed overall good recognition rates for emotional expressions and significant impact of the zoosemiotic features and the biomimetic neck posture approach.

Based on the expression capabilities established, Chapter 3 focused on the employment of these in social interaction. Challenges addressed were when and how to apply implicit communication signals and their potential value for users and/or agents. Mimicry and

smiling were identified as beneficial in service encounters and other dyadic interactions and therefore assessed for their applicability in artificial agents. A novel technical model for mimicry, based on an abstraction of models from developmental child psychology, allowed automatic mimicry of facial and emotional expressions. Furthermore, an extended version of the system-theoretic model of smiling enabled agents to generate context-sensitive smile variations as well as emotional mimicry. Both models showed in combination with the prior developed biomimetic head and its expression capabilities significant improvements of dyadic interactions in terms of subjective performance and empathy ratings in experimental evaluations. Thus, enhancement of the perception of a robot as social actor has been demonstrated by application of a biomimetic top-down design process.

In Chapter 4, the bottom-up approach was covered. Fish, especially the blind Mexican Cave fish, demonstrate remarkable abilities such as object avoidance, object discrimination and environment mapping by sensing the hydrodynamic image of the surrounding flow field on their body. Key challenges in this part were how to make the hydrodynamic image usable in robotics and the achievable abilities. As first step in Section 4.2 neuromasts, the basic components of a lateral line system, were abstracted by hot thermistor anemometry. Sensors were developed and integrated on the robot *Snookie*, which is a specifically designed test bed for the artificial lateral line system. Determination of sensor characteristics revealed agreement with theoretical predictions and compliance with the requirements derived from the findings on neuromasts in literature. Further steps worked on information extraction about the environment from the sensor readings, introducing two different methods. An analytical stimulus model was presented in Section 4.3, enabling detection of walls and objects that can be approximated by wall-like planes. Experimental evaluations demonstrated wall and object detection with the proposed method and the applicability of the developed sensors. Furthermore, simulations of *Snookie* and the artificial lateral line showed the feasibility of obstacle avoidance solely based on flow field information. The second method was a numerical approach given in Section 4.4, implementing a novel process for flow field reconstruction. Using this process, the hydrodynamic image on the surface of the robot can be inverted, giving an approximation of the surrounding flow field based on sensor measurements. By means of this flow field reconstruction, the state-of-the-art was extended to object reconstruction with arbitrary static solid boundaries. The given detection and reconstruction methods are general solutions and therefore independent of the flow perception technique and medium (e.g. air, water or oil). With the ability to reconstruct arbitrary shapes, simultaneous localization and mapping (SLAM) techniques added to the applicability of artificial lateral line systems. The scope of application was extended from reactive obstacle avoidance to more sophisticated navigational tasks. Successful object reconstruction and navigation in complex environments based only on information from the artificial lateral line system were shown in simulation. Therefore, the development of novel flow sensing abilities mimicking those of fish was demonstrated via application of a biomimetic bottom-up approach.

5.2 Future Directions

The work and results presented in this thesis form the basis for future developments. This section outlines possible improvements as directions for future research.

Dynamics: Synthesis of non-verbal communication signals is a key aspect of social robotics. The models introduced in this thesis allow for smooth transitions between expressions and the associated internal states. While the Zurich Model of Social Motivation (ZM) explicitly incorporates the dynamics of state changes in the model, the presented PAD representation and many other models for expression generation do not take them into account. With literature on HHI suggesting the importance of timing, this could be pursued further.

Non-human features: Semiotics of non-human features in communication receive little attention so far. Besides zoosemiotic features, other features are of interest that could be perceived as communication modalities. Steps in this direction, for example, have been taken by Mirnig *et al.* [440] by investigating multi-modal cues including a head-mounted pointing device.

Cultural dependency: Studies in psychology, the most prominent by Ekman *et al.*, provide evidence for the cultural dependency of expression synthesis and perception [441]. Current frameworks in robotics and virtual agents, however, are aimed at being universally understandable. Including characteristics of the targeted society, leading to cultural diversity in social robotics with specifically tailored properties, could be worthwhile to investigate.

Social science to social robotics: In recent years the field of HRI has gained traction and the application possibilities for social robots grow. The advance in human sciences offers a tremendous pool of knowledge that could benefit social robotics as blueprints. Utilizing the design process of identifying function principles and implementing an abstraction opens up many exciting future directions. Extension of this work has already shown the successful transfer of the concept of eliciting altruistic behavior via empathy and similarity [439].

Hydrodynamic image: Perception of the *hydrodynamic image* and information extraction from it introduce a novel sensing modality in robots. The methods and algorithms presented in this work provide a significant step towards the applicability of such a sensory system to autonomous robots. However, further steps have to be taken to be able to cope fully with unconstrained real world environment: First of all, the current implementation is 2D, assuming a constant depth of the robot. An extension to 3D is planned to account for the underwater workspace. Second, necessary assumptions for the reconstruction are that the fluid is ideal and irrotational and that flow is laminar. These assumptions normally do not hold for fields of application of underwater robots or natural habitats of fish, e.g. trouts living in highly turbulent streams. Therefore, future directions could investigate how fish cope with turbulent flow. Another direction for the investigation of turbulence is

the possibility of wake or vortex tracking with Snookie, see work by Franosch *et al.* [410] or Akanyeti *et al.* [432] for feasibility.

Multi-modal fusion The blind Mexican Cave fish is a popular object of study for research on the lateral line system, as it (in all probability) only relies on this sensor system for perception of the environment. However, this is a special case and the lateral line is also a sensory system in seeing fish. Research on the fusion of lateral line and visual stimuli and a transfer to technical systems could not only allow drawing conclusions on the stimuli processing in fish, but also combine the strength of both sensor types.

Interdisciplinary convergence: Biomimetics is providing the interface between several research disciplines. For this interface to achieve its full potential, an efficient exchange of data, models and concepts is necessary. Future directions in "soft" (social science) and "intermediate" (biology) sciences towards more quantifiable and computable models would benefit the transfer to technical systems. Vice versa, technical implementations in engineering would be viable for the verification of concepts derived from other disciplines. Therefore, the interdisciplinarity of biomimetics fosters a tighter coupling of the research areas to make use of synergies and cross-checking of theories.

5.3 Concluding Remarks

In this thesis, biomimetics was used as a valuable source of inspiration and tool kit for robotics. The structured process of the top-down approach has led to the transfer of knowledge from social science, anatomy and zoology to HRI, providing design aspects for social robots, methods to generate non-verbal communication signals and utilize their influence to the benefit of both users and robot. In the same way, the bottom-up approach was consulted to transfer the flow field perception of fish to underwater robotics, aiding the development of artificial lateral line sensors and methods to extract information about the environment from the hydrodynamic image imprinted on the robot surface. In conclusion, the concepts and methods presented in this thesis significantly advance the state-of-the-art in social and underwater robotics, unifying such distinct fields via a common biomimetic design approach.

A Appendix to Chapter 2

A.1 Facial Action Coding System (FACS)

The FACS defines 46 action units, which are mapped to a set of facial muscles according to Table A.1.

AU	Name	Facial muscles involved
1	Inner Brow Raiser	*Frontalis (pars medialis)*
2	Outer Brow Raiser	*Frontalis (pars lateralis)*
4	Brow Lowerer	*Corrugator supercilii*, *Depressor supercilii*
5	Upper Lid Raiser	*Levator palpebrae superioris*
6	Cheek Raiser	*Orbicularis oculi (pars orbitalis)*
7	Lid Tightener	*Orbicularis oculi (pars palpebralis)*
9	Nose Wrinkler	*Levator labii superioris alaeque nasi*
10	Upper Lip Raiser	*Levator labii superioris*
11	Nasolabial Deepener	*Zygomaticus minor*
12	Lip Corner Puller	*Zygomaticus major*
13	Cheek Puffer	*Levator anguli oris*
14	Dimpler	*Buccinator*
15	Lip Corner Depressor	*Depressor anguli oris*
16	Lower Lip Depressor	*Depressor labii inferioris*
17	Chin Raiser	*Mentalis*
18	Lip Puckerer	*Incisivii labii superioris* and *Incisivii labii inferioris*
20	Lip stretcher	*Risorius* and *platysma*
21	Neck Tightener	
22	Lip Funneler	*Orbicularis oris*
23	Lip Tightener	*Orbicularis oris*
24	Lip Pressor	*Orbicularis oris*
25	Lips part	*Depressor labii inferioris* or relaxation of *Mentalis*, or *Orbicularis oris*
26	Jaw Drop	*Masseter*, relaxed *Temporalis* and internal *pterygoid*
27	Mouth Stretch	*Pterygoids* and *Digastric*
28	Lip Suck	*Orbicularis oris*
31	Jaw Clencher	
38	Nostril Dilator	
39	Nostril Compressor	
43	Eyes Closed	Relaxation of *Levator palpebrae superioris*, *Orbicularis oculi (pars palpebralis)*
45	Blink	Relaxation of *Levator palpebrae superioris*, *Orbicularis oculi (pars palpebralis)*
46	Wink	Relaxation of *Levator palpebrae superioris*, *Orbicularis oculi (pars palpebralis)*

Table A.1: List of Action Units and their corresponding facial muscles, based on [93].

A.2 Messages of Animal Signals

Table A.2 shows the catalogue of different messages of animal signals by Smith. This is the complete table, from which an excerpt was taken in Section 2.3.3.

Behavioral Messages
Interactional behavior (type of interaction unspecified)
Attack
Escape
Copulation
Affiliative
Indecisiveness
Locomotion
Site specific (staying at current location)
Seeking (attempting to perform another behavior such as affiliation or escape)
Receptive (to interaction from others)
Attentative (vigilant, monitoring)
Modifiers
Probability
Intensity
Stability
Direction
Identifiers
Population (individual, group, species)
Physiological state (maturity, sex, estrous)
Relationship (pair bond, family, parent-infant)
Location

Table A.2: Messages of animal signals [177], with modifications from Snowdon [41].

A.3 Idle Motions

The current implementation of idle motions in the system is in such a way that every 4 seconds a pseudo-random variable $x \in [0, 100]$ is generated. It is compared to the range assigned to each idle motion r, see Table A.3. If $x \geq r$, the animation is displayed. For inhibition, the evaluation process is suppressed.

Motion	Involved AUs	Range r
wiggle ears	Z2	90
blink	45	70
raise eyebrows	1, 2	50
open mouth	26	30
raise crest	Z1	10

Table A.3: List of implemented idle motions

A.4 Self-Assessment Manikin (SAM)

The Self-Assessment Manikin (SAM) test [53] is shown in Figure A.1.

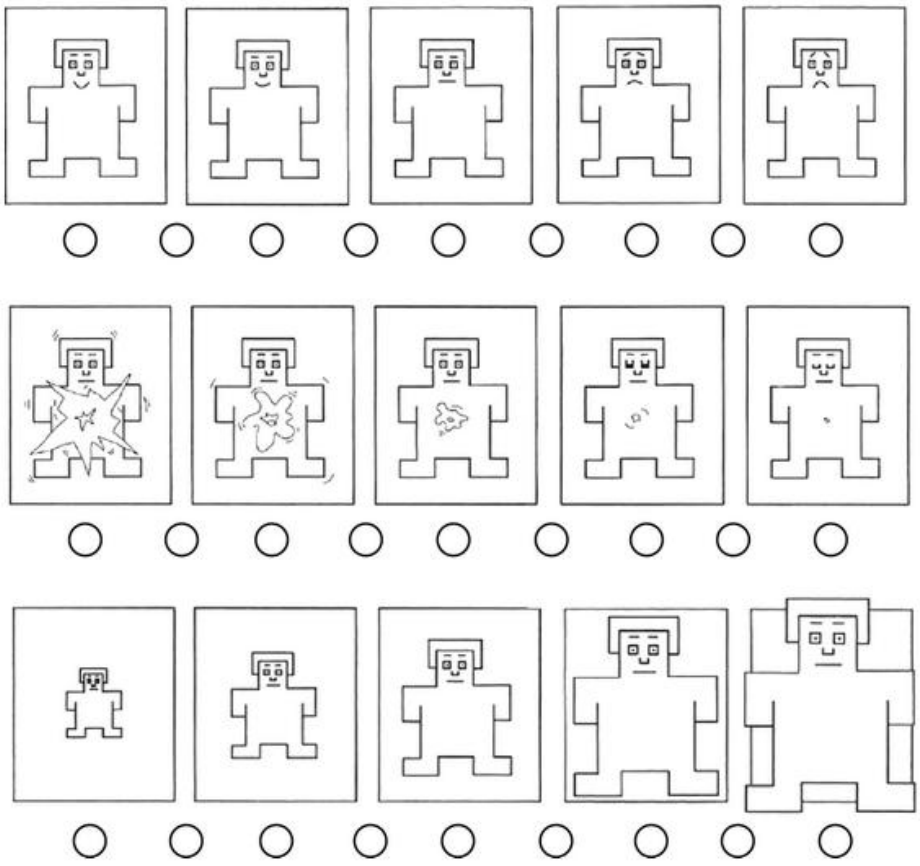


Figure A.1: Self-assessment manikin test with a 9-point scale. Top row depicts pleasure, middle row arousal and bottom row dominance [53].

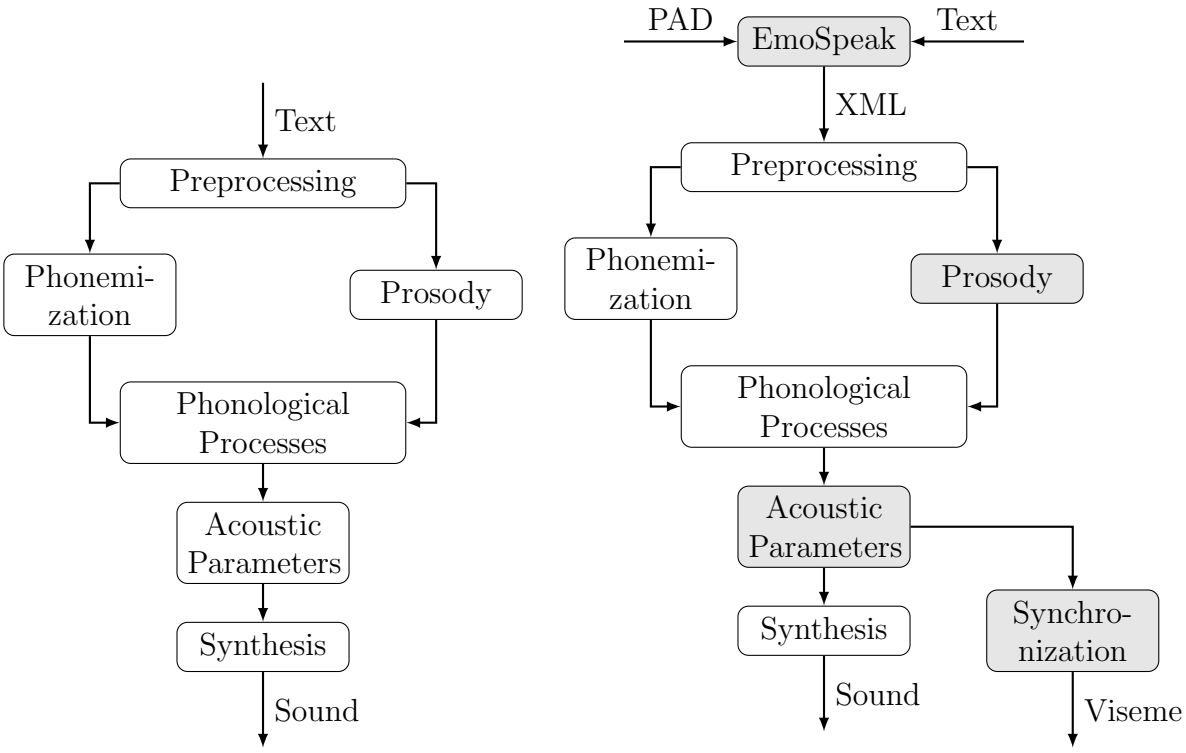
A.5 Expressive Voice

A communication channel to be considered for HRI is the voice on the auditory channel. While most information is conveyed explicitly via the semantic meaning of the spoken

words, implicit information can be given by the change in prosody. In this section, a method to generate an artificial voice with the ability to modulate prosody is introduced.

A.5.1 Voice Synthesis

An implementation of the set of rules to generate speech is incorporated in the MARY Text-to-Speech System [97] from DFKI[1], which is an open-source, multilingual Text-to-Speech synthesis platform. It is used in this system to generate verbal expressions, as the open structure allows an adaptation to specific system requirements, such the change in prosody.



(a) Simplified internal MARY structure

(b) Extended structure for emotional voice generation and viseme synchronization

Figure A.2: Modular structure of MARY TTS. Filled boxes indicate modified or additional modules in comparison to the original implementation by Schröder [60].

MARY Text-to-Speech

MARY[2] provides a modular system for the process of generating speech output from text input, with access to each intermediate step. The modular architecture is shown in a simplified version in Figure A.2a. A detailed description of the functionality is described

[1] Deutsches Forschungszentrum für Künstliche Intelligenz
[2] Modular Architecture for Research on Speech Synthesis

by Schröder [60]. Here, only a brief overview based on this description is given to explain the verbal expression generation:

In a first step, plain text or Extensible Markup Language (XML) encoded input is parsed. Within the *preprocessing*, a tokenizer separates text into tokens, which represent words, numbers special characters and punctuation marks. Rule based information about punctuation marks is added according to the surrounding context. Tokens with differences in the spoken form compared to the written form are replaced with the respective pronunciation, for example numbers, whose pronunciation depends on the context, or abbreviations, which can be either spelled out or expanded. The last step in preprocessing is a chunk parser, determining the boundaries of noun phrases, prepositional phrases and adjective phrases.

In the *phonemization* phase, words are transcribed to phonemes. Phonemic transcription of known words is lexicon based, while unknown words are processed by a letter-to-sound conversion algorithm. Words replaced during the tokenization such as abbreviations are modified with the appropriate inflection endings.

The *prosody* module assigns tones and break indices. Intonation, accents of prominent words and boundary tones associated with the end of a phrase are encoded with symbolic indices, with the actual tones assigned according to the type of sentence (declarative, interrogative or exclamative).

The combination of prosody rules and the standard phonemic string resulting from phonemization is restructured in the *phonological process*. The restructuring is based on rules concerning the phonological context information such as pitch accent or word stress.

So far, the representation of the expression to be generated was on the symbolic level. With the computation of the *acoustic parameters*, the representation changes to the parametrical domain. Symbolic tones and break indices assigned during the prosody phase are translated into fundamental frequency (F0) targets and pause durations. The output of the module is a list containing the syllables together with durations and F0 targets.

The final step is the *synthesis* of the sound output from the list that is provided by the acoustic parameter module. It is processed in the MBROLA diphone synthesizer [80] to generate the actual sounds.

Emotional Expressions

In order to generate emotional verbal expressions, prosody and acoustic parameters are influenced based on the emotional state of the system. This method is adapted from Schröder [60]. The terms evaluation, activation and power used in his work (based on Cowie *et al.* [181]) correspond directly to pleasure, arousal and dominance. Thus, the emotional component of MARY can be used within the PAD based emotional framework introduced earlier, see Figure A.2b for an overview on the modifications to the original MARY structure.

An emotional sentence is first passed from a dialog system, which generates the text to be spoken dependent on the current task, to a preprocessor module called *emoSpeak* that is upstream of the MARY architecture. This module generates the XML structure for MARY based on the current PAD state, altering a set of acoustic parameters to achieve a change in prosody:

The parameter set was selected by Schroeder for being manipulable within MARY. Table A.4 contains the maximum values for all acoustic parameters, as well as the influence of the different PAD-values P pleasure, A arousal and D dominance. Each parameter is computed by

$$\beta = 1.0 + f_P \text{ P} + f_A \text{ A} + f_D \text{ D} \tag{A.1}$$

$$p_{acoustic} = \beta v_{base} \tag{A.2}$$

The PAD-values as well as the acoustic parameter-dependent factors f_P, f_A, f_D are in the range of [-1,1]. The base value v_{base} is the value for each acoustic parameter $p_{acoustic}$ that would be used to synthesize the voice in a neutral, non-emotional way. The composition of β in equation (A.1) is based on the assumption that a linear correlation between the PAD dimensions and the acoustic parameters exists, neglecting a presumably more complex interrelation, but providing satisfying results in a perception test [60]. The values of the factors f_P, f_A, f_D originate from a combination of corpus analysis, literature review and heuristics [60].

Acoustic parameter	**Variation range** min[%]	max[%]	f_P	f_A	f_D
Pitch	-50	+30	0.27	0.27	0.09
Range	-80	+80	0	1.60	0
Pitch dynamics	-400	+400	0	2.00	2.00
Range dynamics	-400	+400	0	3.00	1.00
Rate	-70	+10	0.20	0.50	0
Accent Prominence	-100	+100	0.50	-0.50	0
Accent slope	-150	+150	1.00	-0.50	0
Number of pauses	-40	+40	0	0.40	0
Duration of pauses	-20	+20	0	-0.20	0
Vowel/nasal/liquid duration	-70	+70	0.40	0	0.30
Plosive/fricative duration	-90	+90	-0.40	0.50	0
Volume	-66	+66	0	0.66	0

Table A.4: Changes to the acoustic base parameters by the emotional speech module

The presented values in Table A.4 are a modification of the parameter set described by Schröder. Subjective tests show that high changes in pitch, range, rate and number/duration of pauses might lead an unnatural sounding voice or reduce understandability in the present setup. To prevent users from focusing on major discrepancies between the sound of the robotic voice and a human voice, a maximum range of variation for the prosodic parameters is introduced for a saturation of the impact of the parameters on the speech generation. The variation range for each parameter is heuristically tuned to achieve understandability within its limits.

Moreover, changes are applied to the f_P, f_A and f_D factors of all acoustic parameters except rate, accent prominence/slope and duration of pauses. These factors are increased to make the prosodic changes due to transitions in the emotional state more audible and distinctive (except f_P for pitch, which is slightly decreased).

This is especially important due to the continuous input provided by the Akinator game, with small alterations in the mood of the robot needing to be perceived distinguishably. The emotional states determined by the game focus on three emotions: happiness/self-assurement if the game is going well for the robot, sadness, if the game does not work out the way it should for the robot, and surprise for sudden gain or loss in confidence during the game. As a result, the acoustic parameter-dependent factors are heuristically tuned specifically for these three emotions.

A.5.2 Synchronization

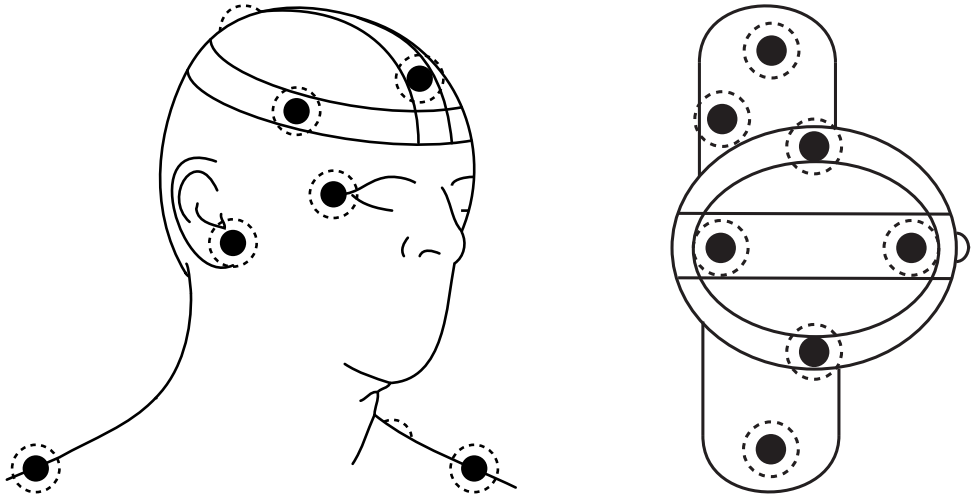
In humans, the articular process of speech generation automatically couples viseme and phoneme expression. For synthesized speech, that only involves computational models of the articulation process, the output of phonemes must be synchronized with the display of visemes presented in Section 2.3.5. In the presented system, this synchronization is achieved with the output of the acoustic parameters module. The module output is a list of syllables with durations and F0 targets. Parsing the syllables with the respective durations results in a list of visemes with durations matching the synthesizer output.

A.6 Parameter Determination of Joint Coupling

The biomimetic model of the human neck suggests dependencies in the movements of joints. These dependencies are confirmed experimentally and estimated quantitatively. Movements of the head with respect to the shoulders are tracked with a Visualeyez VZ 4000 tracking system from Phoenix Technologies Inc. with active LED markers. Figure A.3 shows the placement of the markers, split up in 6 markers at the head and 3 markers at the upper body and shoulders. The third body marker and the fourth head marker on top are for redundancy to counteract occlusions in the from above measurement. Additional eye and ear markers in reference to the top markers, together with a target marker (approximately 2.5m away) form a line of sight reference. All markers are tracked with 70 Hz.

To obtain the default posture $\boldsymbol{q}_0$, participants are asked to look straight ahead, without a target specification. The simplified model of neck joints, see Section 2.4.3, is then fitted in between the filtered marker measurements with a least squares fit.

Measurements of the joint dependencies are gathered from the participants focusing on a target grid in front of them. The target grid is partitioned in 10° steps in horizontal ($-100°$;100°) and vertical($-70°$;100°) orientation, plus diagonals.Participants are asked to look from the origin to the respective specified point, rest for a few seconds and reset back to the origin. From the maintained posture during target fixation, the angles of the simplified neck joints can be calculated using the fitted model.



(a) Side view **(b)** Top view

Figure A.3: Marker positions for the head motion measurements. Markers at the top of the head measure the inclination and position relative to the shoulder markers. Markers at the ear and eye are for reference of the line of sight. The fourth top marker and the third shoulder marker provide redundancy in case of occlusion.

B Appendix to Chapter 3

B.1 AIM Model

Figure B.1 shows the original AIM model by Meltzoff and Moore [309]. The model describes the internal structure of facial (and at a later age emotional) mimicry in infants.

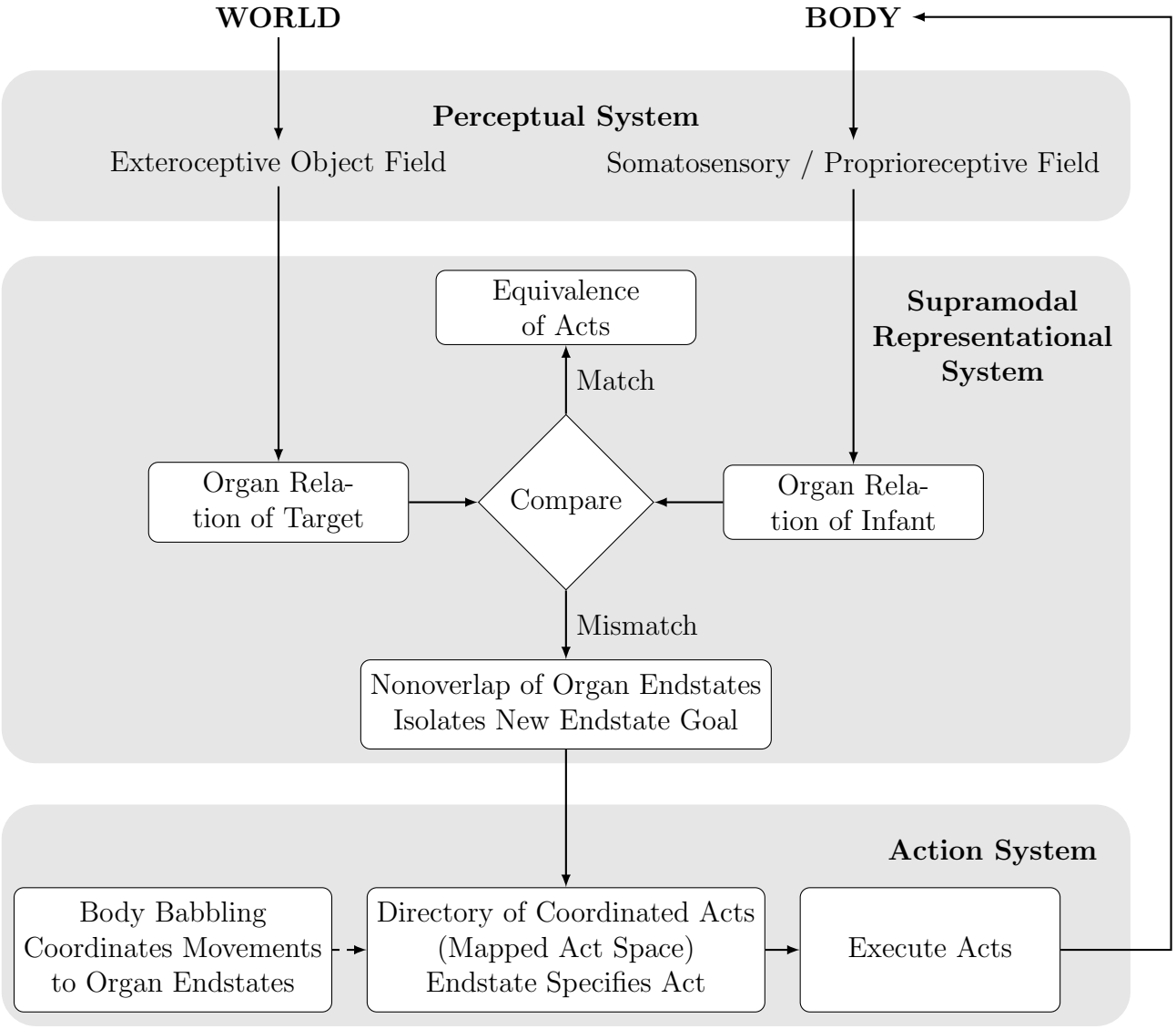


Figure B.1: AIM model according to Meltzoff and Moore [309].

B.2 System Overview

Figure B.2 provides an overview of the modules connected for the experiments in Chapter 3.

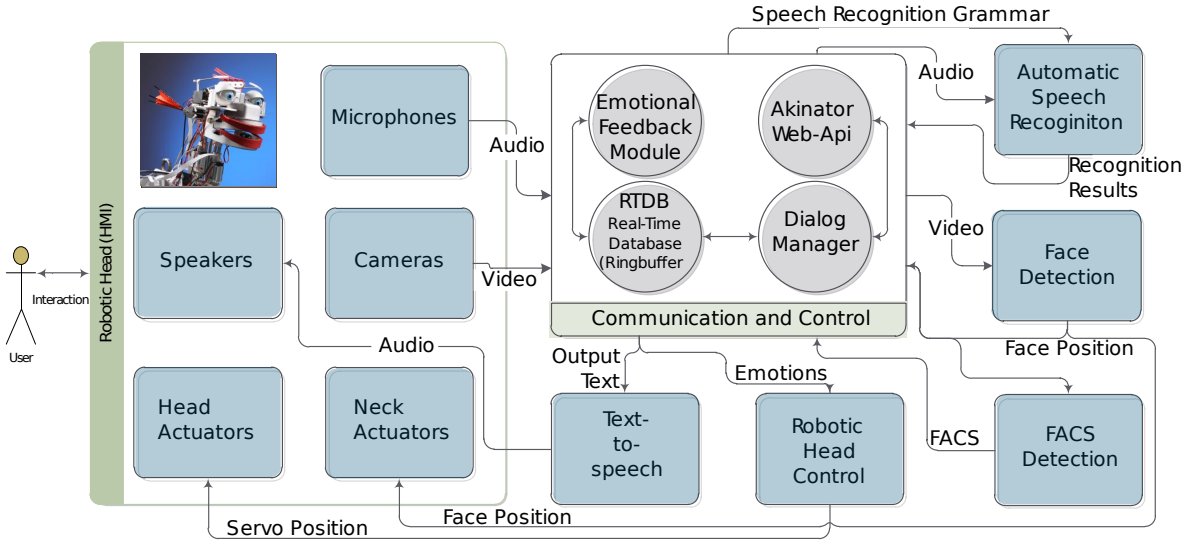


Figure B.2: System Overview of the modules connected for the experiments in Chapter 3.

B.3 Candide-3 model

Figure B.3 shows the Candide-3 model [268] with varying parameters for the 116 landmarks specified in the model.

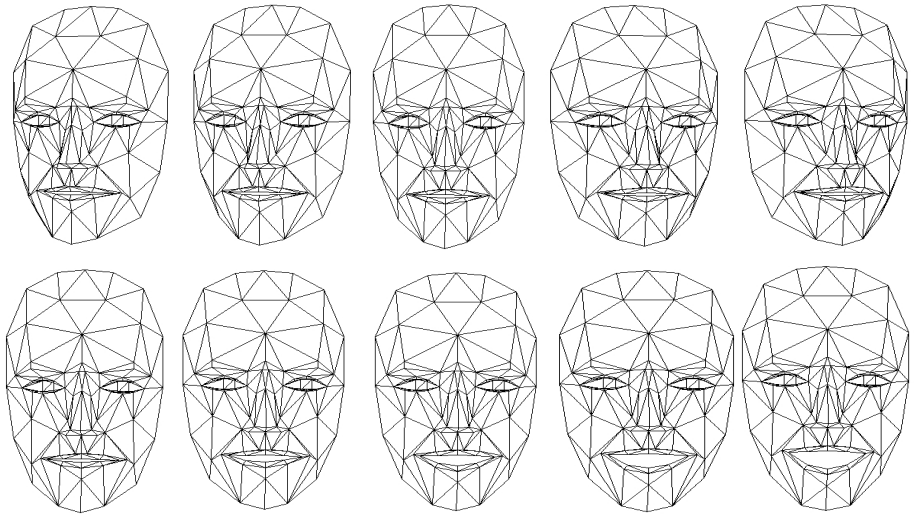
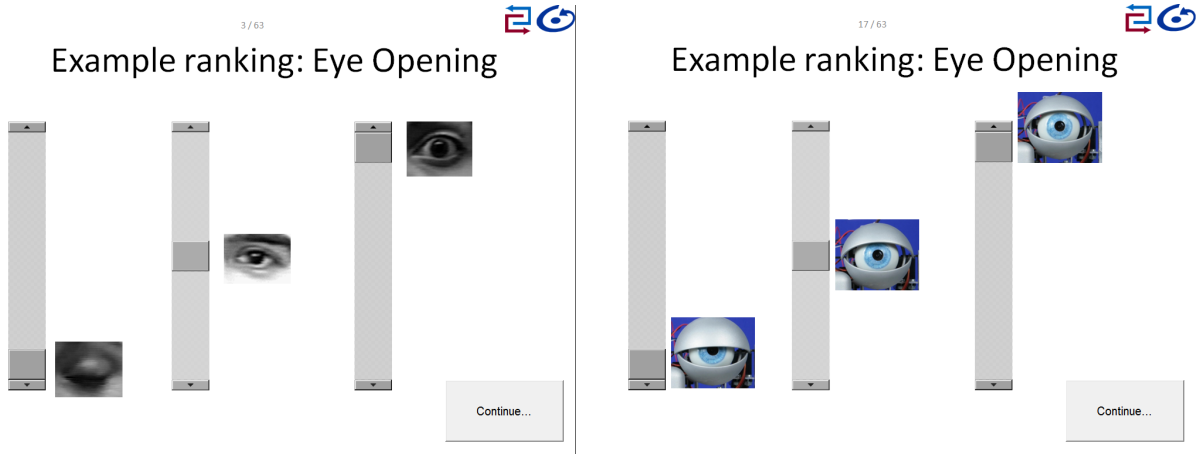


Figure B.3: Candide-3 model with different parametrization [288]



(a) (b)

Figure B.4: Participants were asked to rate human and robotic faces in four categories and five intensities.

B.4 Example Slides for Mimicry Perception Experiment

Figure B.4 shows exemplary ratings for human and robot facial expressions used in the mimicry perception experiment in Section 3.2.6.

B.5 Akinator

In order to provide structure and context to a gaming dialogue, an interface for the robot software to the *Akinator*[1], is integrated, using the application programming interface (API) provided by Elokence.com and Jürgen Blume[2]. To participate in this application, first, the user is asked to choose a person. Then, the computer tries to guess this person by asking several questions. The person may be a real or fictional person, currently living or historical, taken from literature, the media or public live. To answer Akinator's questions, a set of fixed answers is presented by the system. The set of answers is the same for every question and consists of: "Yes", "Probably" / "Partially", "I don't know", "Probably not"/ "Not really", and "No". Example questions asked by the Akinator are: "Is your character a girl?", "Does your character live in America" or "Does your character really exist?".

B.6 Dialog Manager

A dialog manager provided by Jürgen Blume keeps track of the ongoing communication to estimate when a response of the human user or the machine is expected by the dialog

[1] www.akinator.com

[2] Institute for Human-Machine Communication, Department of Electrical Engineering and Information Technology, Technische Universität München

partners. The complete dialog structure is implemented in a first-order logic representation. Tasks to be solved are represented by predicates with variables. These variables represent information which is to be determined during the dialog. Equivalence rules on these predicates are specified to navigate through the dialog by splitting a task into several subtasks.

B.7 Experimental setup

For the experimental setup a quiet room with controlled lighting conditions was chosen. The robotic head was placed on a table to be at approximately eye-level with the participants. Participants were seated in front of the robot, with a microphone placed in front of them on the table to ensure a low error rate in speech recognition. Since the task rating and enjoyment would depend on the ability of the robot to correctly understand the answers, the external microphone was preferred over the internal, which would have added to the illusion of speaking to the robot directly. The instructor greeted the person and gave a short introduction on the task and how to interact with the robot. To begin the experiment, the instructor asked the participant to think of a person of his/her own choice and give a start signal, when done. From this point, the robot started the Akinator game, speaking the questions provided by the Akinator API and listening for the answers. A sample round of Akinator can be seen in Table B.1.

After the game was finished by either the robot guessing the correct person or giving up after to many trials (dependent on the Akinator API, having a threshold influenced by the confidence and the number of trials), the subjects were asked to fill in a computer based questionnaire.

Question	**Answer**	
	given	expected
Is your character a male?	No	No
Is your character a singer?	No	No
Does your character really exist?	No	No
Does your character fight?	Not really	No
Is your character from an anime?	No	No
Does your character live in America?	No	No
Is your character a human being?	No	No
Is your character an animal?	No	No
Does your character have hair?	No	No
Is your character visible?	Yes	Yes
Is your character a robot?	Yes	Yes
Has your character played in Star Wars?	Yes	Yes
Is your character yellow?	No	No
I guess you were thinking of: R2D2		

Table B.1: Sample dialogue of a game of Akinator, looking for R2D2

B.8 Questionnaire

Table B.2 shows the questionnaire for user acceptance and the additional constructs of empathy and subjective performance developed by Barbara Kühnlenz (Gonsior).

	Empathy
1a)	I am happy that Eddie guessed my person.
1b)	It's a shame Eddie didn't guess my person.
2a)	I would have been proud if Eddie hadn't guessed my person. (inverted)
2b)	I'm proud Eddie didn't guess my person.
3a)	It would have been a pity if Eddie didn't guess my person.
3b)	It would have been nice if Eddie had guessed my person.
4a and b)	I would feel sorry for Eddie if someone tried to destroy it at that moment, thus I would try to prevent it.
	Subjective Performance
1a)	I was impressed by how fast Eddie has guessed my person.
1b)	I had the feeling that Eddie nearly guessed my person.
2)	Eddie has shown a good performance.
3)	I think that Eddie has worked efficiently.
4a)	It took Eddie long to guess my person. (negated)
4b)	It took Eddie too long to guess my person. (negated)
	Trust
1)	I would believe Eddie if he gave me advice.
2)	Eddie is inspiring confidence.
3)	I feel that I can trust Eddie.
4)	I do not trust Eddie's statements.
	Perceived Sociability
1)	I like Eddie.
2)	Eddies mimic and verbal statements fit together well.
3)	Eddie was good conversation partner.
4)	Eddie's behavior was inappropriate.
	Social Presence
1)	I had the feeling that Eddie really looked at me.
2)	I could imagine Eddie as a living being.
3)	Sometimes it felt like Eddie had real feelings.
4)	Eddies behavior was not humanlike.
	Perceived Enjoyment
1)	It was fun to interact with Eddie.
2)	The conversation with Eddie was fascinating.
3)	I consider Eddie to be entertaining.
4)	It's boring when Eddie interacts with me.
	Intention to Use
1)	I would like to interact with Eddie more often.
2)	I would take Eddie home with me.
3)	I would like to play again with Eddie within the next few days.
4)	I could imagine interacting with Eddie over an extended period of time.

Table B.2: User acceptance questionnaire for empathy and subjective performance

B.9 Cusp Catastrophe

Figure B.5 shows the bifurcation and the fold of the reaction surface of the cusp catastrophe.

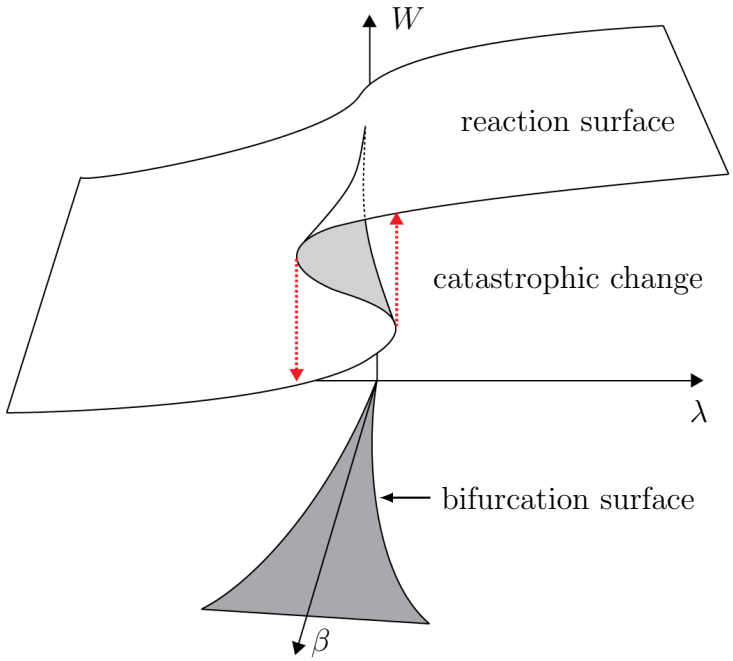


Figure B.5: Visualization of the fold of the reaction surface in 3D of the cusp catastrophe.

B.10 Appetence and Aversion Mapping to FACS

Table B.3 gives the mapping of smile s^* and expression a^* activations to facial muscles and the respective AUs. Columns 3 to 8 list the activation of the respective AU on a scale from 0 to 1.

Appetence and aversion of all three dimensions can manifest themselves in parallel and independent of each other. A problem arising from the parallel execution is that expressions can be overlapping in access of action units. Thus, the contribution of several components of expressions to one specific action unit must be fused to avoid either information loss in a winner-takes-all strategy or violation of the boundaries in an additive process. The applied fusion strategy is the same as in Section 2.3.6, merging the contributing action unit activations $au_{i,j}$, with $i = 1, \ldots, 46$ and $j = 1, \ldots, N$, where N is the number of dimensions accessing the action unit to an overall activation:

$$au_{merged\ i} = 1 - \prod_{j=1}^{N}(1 - au_{ij}),\ \forall\ i = 1, \ldots, 46$$

Muscle	AU	s^*	$a^*_{Security}$		$a^*_{Arousal}$		$a^*_{Autonomy}$	
			Ap	Av	Ap	Av	Ap	Av
Frontalis (pars mendialis)	1		0.9			1		
Frontalis (pars lateralis)	2				1	1		
Corrugator	4					1	1	
Levator palpebrae superioris	5				0.4	1		
Levator palpebrae superioris	43							0.4
Orbicularis oculi (pars orbitalis)	6	1	0.8	0.5			0.3	
Orbicularis oculi (pars palpebralis)	7		0.5					
Orbicularis oris	14						1	
Levator lab. sup. aleaeque nasi	9			1				
Levator lab. sup. caput. infraorb.	10			0.9			1	
Zygomatic major	12	1						
Triangularis	15						0.4	
Depressor labii inferioris	16		0.3					
Mentalis	17		0.1				0.3	
Risorius	20					0.5		
Masseter	26		0.3		1	0.1		0.4

Table B.3: Appetence (Ap) and aversion (Av) mapping to FACS for muscle activations of security, arousal and autonomy [199], based on [319].

Fusing the action unit activations results in a merged action unit activation $au_{merged\ i}$, which is bounded between $[0, 1]$.

B.11 Smile Smoothing

The smile detection of the facial expression analysis has a lower sampling rate and granularity than the smiling model. To counteract resulting jumps in the smile input, a double exponential smoothing filter is applied, with $\alpha = \gamma = 0.05$:

$$\begin{aligned} z^*_{smile}(t) &= \alpha z_{smile}(t) + (1-\alpha)(z_{smile}(t-1) + \beta(t)) \\ \beta(t) &= \gamma(z_{smile}(t) - z_{smile}(t-1)) + (1-\gamma)\beta(t-1) \end{aligned} \tag{B.1}$$

B.12 FACS to FAP mapping

The outputs of the system-theoretic model of smiling a^* and s^* are mapped via Table B.3 to FACS. The MPEG-4 standard, which is used for the virtual avatar, however, specifies FAP as atomic facial actuation elements. This FACS to FAP mapping is based on [251], but has been reproduced and modified. Single AUs have been animated with the avatar and matched to photographs[3] of single AUs activations of FACS coders. The lack of wrinkle

[3]photographs copyright by Medien- und Organisationspsychologie, Universität des Saarlandes http://www.uni-saarland.de/fak5/orga/Kurs/home.htm

generation and odds in muscle movement in the avatar face accounts for differences in the perception of the rendered face in comparison to the photographs. Therefore, AUs dependent on wrinkles, such as the nose wrinkler (AU9), are perceived weaker on the avatar. Also AU5 can not be adequately displayed on the avatar, as the lifting of the upper eyelid does only result in a lifting of the lid fold, but not in an opening of the eye. The resulting mapping is given in Table B.4, with one AU activating several FAP.

AU	FAP	
1	31 raise-l-i-eyebrow	32 raise-r-i-eyebrow
2	33 raise-l-m-eyebrow	34 raise-r-m-eyebrow
	35 raise-l-o-eyebrow	36 raise-r-o-eyebrow
4	31 raise-l-i-eyebrow	32 raise-r-i-eyebrow
	33 raise-l-m-eyebrow	34 raise-r-m-eyebrow
	35 raise-l-o-eyebrow	36 raise-r-o-eyebrow
	37 squeeze-l-eyebrow	38 squeeze-r-eyebrow
5	19 close-t-l-eyelid	20 close-t-r-eyelid
6	41 lift-l-cheek	42 lift-r-cheek
7	19 close-t-l-eyelid	20 close-t-r-eyelid
	21 close-b-l-eyelid	22 close-b-r-eyelid
9	21 close-b-l-eyelid	22 close-b-r-eyelid
	31 raise-l-i-eyebrow	32 raise-r-i-eyebrow
	33 raise-l-m-eyebrow	34 raise-r-m-eyebrow
	63 raise-nose	
10	4 lower-t-midlip	8 lower-t-lip-lm
	9 lower-t-lip-rm	
12	6 stretch-l-cornerlip	7 stretch-r-cornerlip
	10 raise-b-lip-lm	11 raise-b-lip-rm
	12 raise-l-cornerlip	13 raise-r-cornerlip
15	6 stretch-l-cornerlip	7 stretch-r-cornerlip
	10 raise-b-lip-lm	11 raise-b-lip-rm
	12 raise-l-cornerlip	13 raise-r-cornerlip
16	5 raise-b-midlip	10 raise-b-lip-lm
	11 raise-b-lip-rm	
17	5 raise-b-midlip	10 raise-b-lip-lm
	11 raise-b-lip-rm	12 raise-l-cornerlip
	13 raise-r-cornerlip	18 depress-chin
20	6 stretch-l-cornerlip	7 stretch-r-cornerlip
24	10 raise-b-lip-lm	11 raise-b-lip-rm
	16 push-b-lip	17 push-t-lip
26	3 open-jaw	5 raise-b-midlip
	10 raise-b-lip-lm	11 raise-b-lip-rm
	12 raise-l-cornerlip	13 raise-r-cornerlip
43	19 close-t-l-eyelid	20 close-t-r-eyelid
	21 close-b-l-eyelid	22 close-b-r-eyelid

Table B.4: Specification of the AUs to FAP mapping.

C Appendix to Chapter 4

C.1 Constants

Description	Symbol	Value	Unit
Boltzmann constant	k_B	$1.3806488 \cdot 10^{-23}$	J/K

Table C.1: Physical constants

Description	Symbol	Air	Water	Unit
Specific heat capacity	c_p	1006	4182	J/(kg K)
Density	ϱ	1.204	998	kg/m^3
Heat conductivity	k	0.0256	0.604	W/Km
Kinematic viscosity	ν	$1.511 \cdot 10^{-5}$	$0.979 \cdot 10^{-6}$	m^2/s
Compressibility	κ	$9.9 \cdot 10^{-6}$	$0.4587 \cdot 10^{-9}$	1/Pa
Prandtl number	$Pr = \frac{\nu c_p \varrho}{k}$	0.72	6.8	
Speed of sound	$c_{air/water}$	343	1484	m/s

Table C.2: Hydrodynamic constants of water and air at 20° C.

Description	Symbol	Value	Unit
Dissipation constant	K_D	1.4e-3	
Initial resistance	R_0	1790	Ω
BETA	β	-3090	
Initial temperature	T_0	293.15	K
Heat capacity	C_ϑ	9.0e-5	J/K
A-V-THERMISTOR	A	1.03e-3	W/K
B-V-THERMISTOR	B	0.74e-3	W/K
N-V-THERMISTOR	n	0.34	

Table C.3: Thermistor specific parameters in simulation

C.2 Sensor Package

The package design eases handling, provides insulation, enables mounting on the sphere and gives distance to the boundary layer. The sensor is soldered to a $1x1x5mm$ PCB stick for electrical contact and stability. This stick is placed in a custom prototyped plastic hull and contacted with a coaxial shielded cable. The connector is either a BNC RF plug or a smaller SMB RF connector. In previous designs, the coating was a weak spot, resulting in two issues: The application of the used printed circuit board (PCB) lacquer was prone to either cover the thermistor, and thus isolate thermally, or leave parts of the wires uncovered. Second, the lacquer showed failure symptoms after an extended period of use due to chemical and/or heat influence. Therefore the sensor is embedded in a silicone coating in the hull, covering the connections but leaving the thermistor glass bead free at the tip of the sensor. The used silicon is ELASTOSIL RT 628 from Wacker Chemie, which is an addition curing, two component silicone rubber.

C.3 Robot Design

Figure C.1 gives an overview on the modular structure of the robot. Each module is also represented in software to account for simulation with or without hardware-in-the-loop. Below, the modules will be briefly described and their respective models for simulation given. The motion models, control strategy and hydrodynamic stimulus modeling will be subject of the next subsections.

a) Sensors

Sensors are the artificial lateral line system, that was introduced in section 4.2. Input to the sensors is the hydrodynamic stimulus that is created either by real-world objects in the vicinity of the sensor or by simulation of the hydrodynamic environment. The sensors are arranged in an equidistant array with a cross shape on the frontal half-sphere, see Figure C.2. Currently 17 sensors are fitted to the front. The cross shape allows to treat the array either as a 3D structure or decoupled in two orthogonal planes, with each plane including 9 sensors. If assumed that the robot moves at a constant depth and that objects can be projected on the horizontal plane, the problem statement can be simplified to a 2D problem, which is covered in the simulations. In section 4.4.3 it is shown that at least in simulations the 9 sensors are sufficient for a reconstruction of the environment in the horizontal plane.

The implementation in simulation places the sensors in a robot-fixed spherical coordinate system, which corresponds to the real placement. The effect of the boundary layer and other viscosity effects are neglected in simulation, since the sensors are placed in the Euler flow regime. In the thermistor simulation a constant current of $I = 19\,\mathrm{mA}$ heats the thermistor, which causes a temperature of about $70\,^{\circ}\mathrm{C}$ in experiments [343] as well as in simulations for water velocity v = 0. For each time step, given the current temperature T of the thermistor, the simulation first calculates the resistance R_ϑ of the thermistor, the electrical power P_{el} and the dissipated thermal power P_ϑ and then integrates by applying

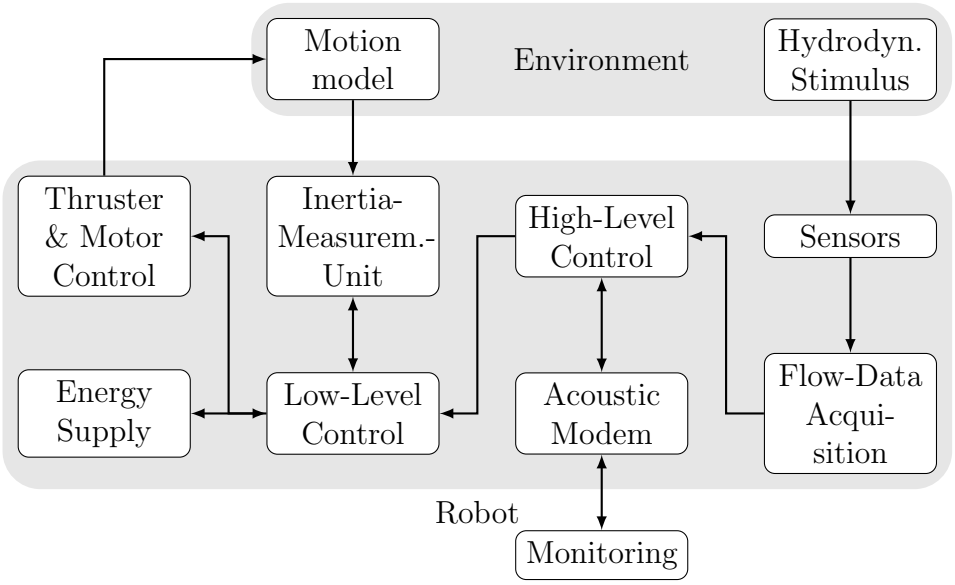


Figure C.1: Overview on the modules of Snookie

one Euler integration step with a time step of 0.01 ms. A thermal noise term $z_{thermal}$ is added to the sensor output, with rng being a value in the range of $[0;1]$ generated by a random number generator with a normal distribution.

Input: Velocity vector at sensor
Output: Voltage signal of the sensor
foreach *time-step* Δt **do**
  calculate dissipation constant: $K_D = (A_v + B_v)||\boldsymbol{v}||^{n_v}$;
  update thermistor resistance: $R_\vartheta = R_0 \cdot e^{\beta_\vartheta \cdot (1/T_0 - 1/T)}$;
  heating/cooling of the sensor: $\Delta T = \frac{I^2 R_\vartheta}{C_\vartheta} - \frac{K_D}{C_\vartheta} \cdot (T - T_0)$;
  incremental temperature: $T+ = \Delta T \cdot \Delta t$;
  thermal noise: $z_{thermal} = \sqrt{\frac{4 \cdot k_B \cdot T \cdot R_\vartheta}{\Delta t}} \cdot rng$;
  resulting sensor signal: $U = (I \cdot R_\vartheta) + z_{thermal}$;
end

b) Sensor Data Acquisition

For the acquisition and conversion from the analog sensor signal to digital, a National Instruments USB Data Acquisition Card NI USB-6218 is used. It provides 32 single-ended channels at 16Bit resolution with a sensitivity of up to 91.6 µV. The card allows to capture 250 kilo samples per second (kS/s) overall, which in the current setup would lead to a maximum of $\frac{250000}{17} = 14705$ samples per second and channel. The actually used sample rate is 10 kS/s and channel, which is then averaged over the last 10 samples to a 1 kHz sample rate. This is still by a factor of 100 faster than the expected time constant of the signal change due to environmental changes and thus fulfills the Whittaker-Nyquist-

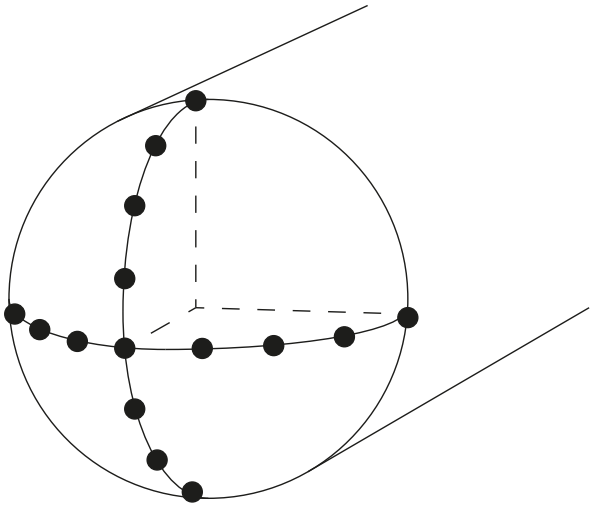


Figure C.2: Sensor array on front half-sphere. A total of 17 sensors is placed equidistantly.

Kotelnikov-Shannon sampling theorem of $f_{sampling} > 2f_{max}$. Data readings are buffered and polled from the connected high-level control PC over USB 2.0. A custom ROS node logs the raw data and pipes it to filters and/or the object detection algorithms.

In simulation, the data acquisition card is treated as an ideal analog-to-digital converter, with the sample rate of the simulation timing.

c) High-level Control

High-level control is a standard personal computer in a small form factor that is included in the robot. It provides the ROS infrastructure for the processing of the sensor data, object avoidance and recognition, data logging and interfacing to command and control. Also the online parameter estimation for the adaptive controller is computed by this PC. The current model has a Core-i3 processor, with 4GB of RAM and 500GB storage.

d) Low-level Control

The low-level control unit is based on an autopilot board by Ascending Technologies. It is the central hub for the embedded systems and controls the 3D orientation and motion underwater. It consists of two 60MHz ARM7 RISC processors, of which one is freely programmable and the other combines and preprocesses sensor data from three MEMS gyroscopes, a three-axis acceleration sensor, a three-axis magnetometer and the pressure sensor to an inertia-measurement-unit. The command unit can utilize this angular and translational data over a direct onboard link.

In simulation, virtual inertia-measurement-unit data can be generated and used in a reimplementation of the controller. Emulating the interfaces and code wrappers allow to embed the code running on the freely programmable processor directly in the simulation and thus test its functionality in a safe environment.

e) Motor Control and Thruster

Single propeller propulsion is doing well if the overall travel speed of the robot is in a speed range with good efficiency of the rudders needed to steer. However, rudders are less effective for small velocities, which will be needed to explore the environment safely. They also allow no on the spot turning, which could be helpful in narrow environments. Therefore, a helicopter-like multi-propeller propulsion system is adapted from the AMOUR V robot [379]. The basic layout incorporates four thrusters aligned horizontally in a cross shape. This allows direct control over the forward/backward movement along the robot's main axis, the pitch angle and the yaw angle. All four motors work in combination for acceleration/deceleration. Additionally two vertically mounted thrusters control depth and the roll angle.

Since motors are a source of vibrations and the propellers induce vorticity into the water due to the rotation, the thrusters have to be mounted as far away as possible from the sensors. This results in a layout with the thrusters as far back at the robot as possible, away from the front half-sphere.

f) Energy Supply

Snookie is powered by three lithium-polymer battery packs. Each pack can store 20 A h at 7 V, so the total capacity is 20 A h at 21 V in a serial connection. The current charge/discharge rate and voltage levels are reported to the low-level controller. This enables the robot to initiate a safe surfacing routine if a battery is drained.

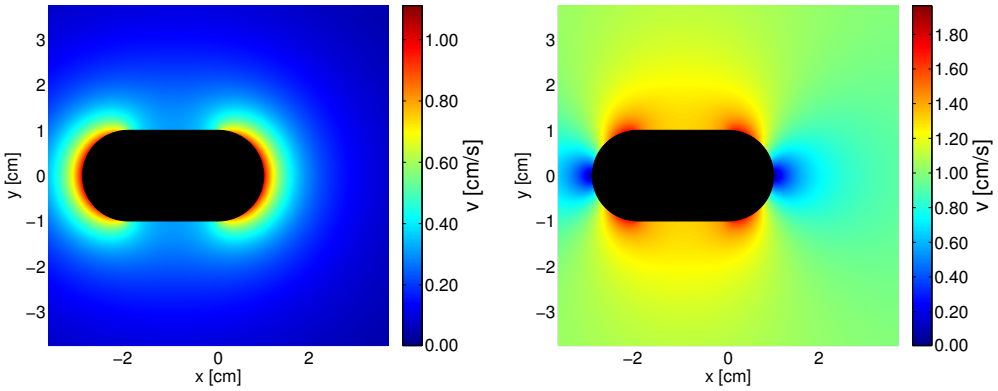
g) Command and Control

A land-based station can be used to monitor the status of the robot and/or give new commands. Direct control of the robot's movement is also possible via either a wiiMote, Joystick or keyboard. The robot can operate tethered via a Cat5 Ethernet cable for a high bandwidth communication.

For untethered operation, the link between command and control and the robot can be established via an acoustic modem by Tritech. The Micron Data Modem sends and received with 40 bits per second in simplex mode over a specified range up to 500 m. If considering the working range and the speed of sound in water, the delay between sending and receiving can be up to:

$$t = \frac{x}{c_{water}} = \frac{500\,\mathrm{m}}{1484\frac{\mathrm{m}}{\mathrm{s}}} \approx 337\,\mathrm{ms} \tag{C.1}$$

To simulate the bandwidth limitation and delays the modem introduces into the communication, a ROS node is implemented that can be inserted in the communication path between the robot and command and control. It emulates the modem as a first-in first-out buffer with limited buffer size, variable delay and throughput.



(a) Absolute flow field

(b) Relative flow field

Figure C.3: Comparison between absolute and relative flow field [348].

C.4 Flow Reconstruction

C.4.1 Relative and Absolute Flow Field

Figure C.3 shows a comparison between absolute and relative flow field.

For reconstruction of the absolute flow field (the flow field in world coordinates), a transformation from the relative flow field (in the robot's coordinate frame) is required. The relative flow field is determined by the measured relative water velocity. Given an absolute vehicle motion (v_x, v_y), the transformations are:

$$\Phi_{abs}(r,\varphi) = \Phi_{rel}(r,\varphi) + v_x r\cos(\varphi) + v_y r\sin(\varphi) \tag{C.2}$$

$$\Psi_{abs}(r,\varphi) = \Psi_{rel}(r,\varphi) + v_x r\sin(\varphi) - v_y r\cos(\varphi) \tag{C.3}$$

$$u_{abs,\perp}(r,\varphi) = u_{rel,\perp}(r,\varphi) + v_x\cos(\varphi) + v_y\sin(\varphi) \tag{C.4}$$

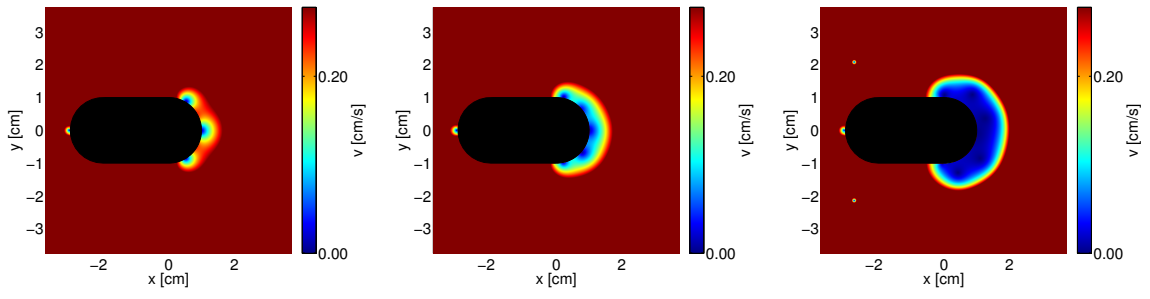
$$u_{abs,\parallel}(r,\varphi) = u_{rel,\parallel}(r,\varphi) - v_x\sin(\varphi) + v_y\cos(\varphi) \tag{C.5}$$

C.4.2 Comparison of Circle Approximation and Whole Body

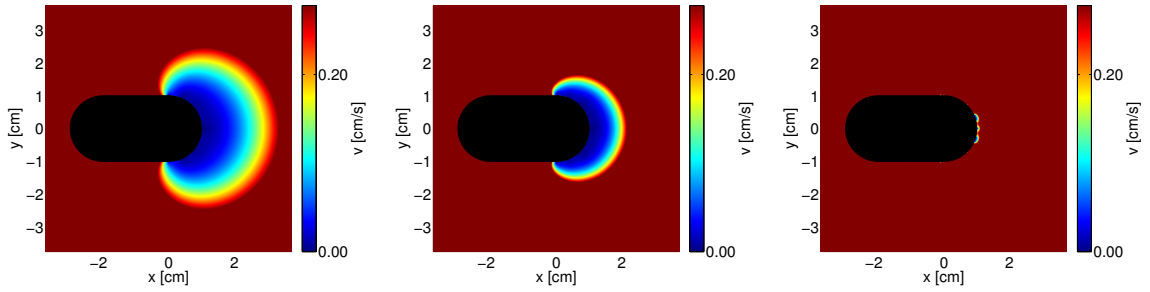
Figure C.4 shows a comparison between the reconstructed flow field for the circle approximation and a full robot body reconstruction.

C.4.3 Error Flow Fields

Figure C.5 shows error flow fields for variations of sensor number M and reconstruction order α_m. The error flow field is defined as the difference between simulated (forward) and reconstructed (inverted) flow field.

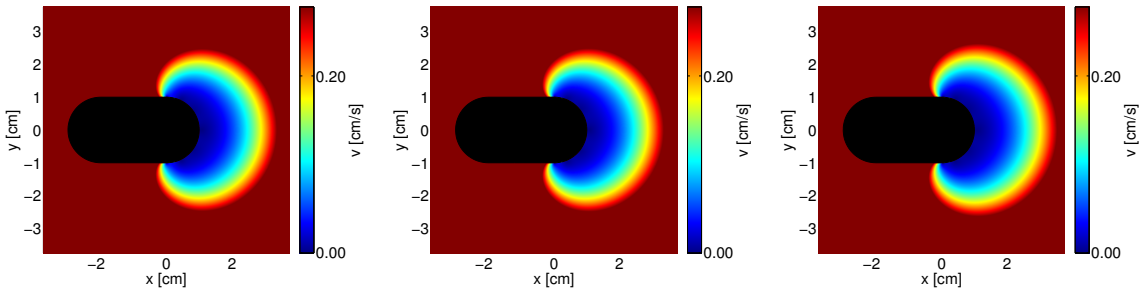


(a) Whole-Body, $\alpha_m = 3$ **(b)** Whole-Body, $\alpha_m = 5$ **(c)** Whole-Body, $\alpha_m = 10$

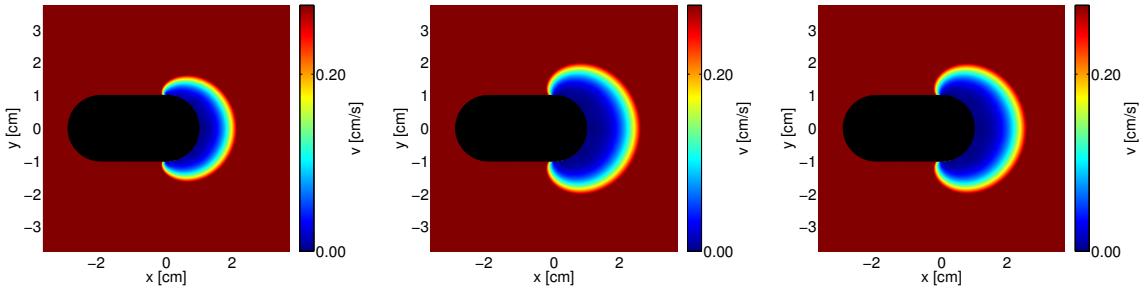


(d) Circle, $\alpha_m = 3$ **(e)** Circle, $\alpha_m = 5$ **(f)** Circle, $\alpha_m = 10$

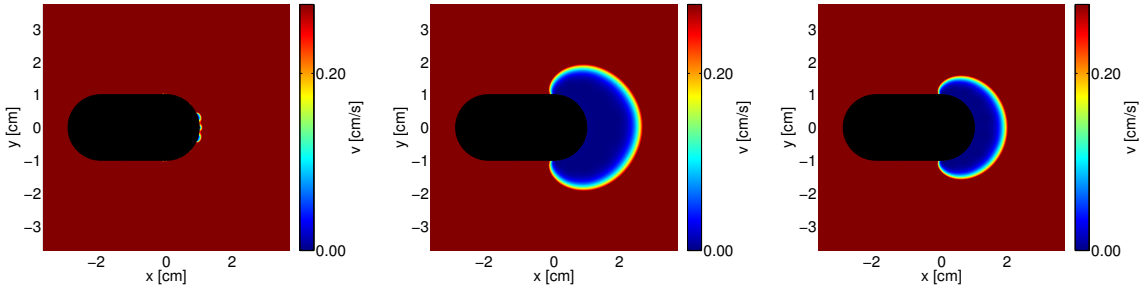
Figure C.4: Error flow field comparison between reconstruction with circle-approximation and with whole body reconstruction for different values of α_m [348].



(a) $\alpha_m = 3$ and $M = 9$ **(b)** $\alpha_m = 3$ and $M = 21$ **(c)** $\alpha_m = 3$ and $M = 51$



(d) $\alpha_m = 5$ and $M = 9$ **(e)** $\alpha_m = 5$ and $M = 21$ **(f)** $\alpha_m = 5$ and $M = 51$



(g) $\alpha_m = 10$ and $M = 9$ **(h)** $\alpha_m = 10$ and $M = 21$ **(i)** $\alpha_m = 10$ and $M = 51$

Figure C.5: Error flow field comparison with different values for M and for α_m [348].

Notations

Acronyms

2D 2-dimensional

3D 3-dimensional

ACE Autonomous City Explorer

AD action descriptor

AIM active intermodal mapping

ANOVA analysis of variance

API application programming interface

AU action unit

AUV autonomous underwater vehicle

BEM boundary-element method

BNC Bayonet Neill Concelman (connector)

CAD computer aided drawing

CFD computational fluid dynamics

CMU Carnegie Mellon University

DoF degree of freedom

ECA embodied conversational agent

EDDIE Emotion Display with Dynamic Intuitive Expressions

FACS Facial Action Coding System

FAP Facial Animation Points

fMRI functional magnetic resonance imaging

FOR frame of reference

FOV field of view

FVM finite volume method

HHI Human-Human-Interaction

HRI Human-Robot-Interaction

IURO Interactive Urban Robot

kS/s kilo samples per second

LDA Laser Doppler Anemometer

LIDAR light detection and ranging

MEMS micro-electro-mechanical systems

MPEG Motion Picture Experts Group

PAD Pleasure Arousal Dominance

PCB printed circuit board

PD proportional-derivative

PIV particle image velocimetry

PVDF polyvinylidene fluoride fibers

RF radio frequency

ROS Robot Operating System

ROV remotely operated vehicle

SAM Self-Assessment Manikin

SMB SubMiniature version B (connector)

SMS system-theoretic model of smiling

SLAM simultaneous localization and mapping

SOM-LA Seat/Occupant Model - Light Aircraft

SOM-TA Seat/Occupant Model - Transport Aircraft

SONAR sound navigation and ranging

UTAUT Unified Theory of Acceptance and Use of Technology

WOz Wizard-of-Oz

XML Extensible Markup Language

ZM Zurich Model of Social Motivation

Bibliography

C.5 Chapter 1

[1] O. H. Schmitt, "Some interesting and useful biomimetic transforms," in *Proceedings of the Third International Biophysics Congress*, p. 297, 1969. cited on page 2.

[2] W. Nachtigall, *Bionik als Wissenschaft: Erkennen - Abstrahieren - Umsetzen (German Edition)*. Springer, 1st ed., 2010. cited on pages 2, 98, and 102.

[3] J. M. Benyus, *Biomimicry: Innovation Inspired by Nature*. Morrow, William, 1st ed., 1997. cited on page 2.

[4] V. C. LaMarche Jr., "Environment in Relation to Age of Bristlecone Pines," *Ecology*, vol. 50, pp. 53–59, Jan. 1969. cited on page 1.

[5] R. J. Pledger and J. A. Baross, "Characterization of an Extremely Thermophilic Archaebacterium Isolated from a Black Smoker Polychaete (Paralvinella sp.) at the Juan de Fuca Ridge," *Systematic and Applied Microbiology*, vol. 12, pp. 249–256, Dec. 1989. cited on page 1.

[6] T. Fujii, A. J. Jamieson, M. Solan, P. M. Bagley, and I. G. Priede, "A Large Aggregation of Liparids at 7703 meters and a Reappraisal of the Abundance and Diversity of Hadal Fish," *BioScience*, vol. 60, pp. 506–515, July 2010. cited on page 1.

[7] A. G. Meyer and J.-A. Emanuel, *How to Catch a Robot Rat: When Biology Inspires Innovation*. The MIT Press, Sept. 2010. cited on page 2.

[8] P. F. Battley, N. Warnock, T. L. Tibbitts, R. E. Gill, T. Piersma, C. J. Hassell, D. C. Douglas, D. M. Mulcahy, B. D. Gartrell, R. Schuckard, D. S. Melville, and A. C. Riegen, "Contrasting extreme long-distance migration patterns in bar-tailed godwits Limosa lapponica," *Journal of Avian Biology*, vol. 43, pp. 21–32, Jan. 2012. cited on page 1.

[9] R. Kram, "Inexpensive load carrying by rhinoceros beetles," *The Journal of Experimental Biology*, vol. 199, pp. 609–12, Jan. 1996. cited on page 1.

[10] L. A. Hawkes, S. Balachandran, N. Batbayar, P. J. Butler, P. B. Frappell, W. K. Milsom, N. Tseveenmyadag, S. H. Newman, G. R. Scott, P. Sathiyaselvam, J. Y. Takekawa, M. Wikelski, and C. M. Bishop, "The trans-Himalayan flights of bar-headed geese (Anser indicus)," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, pp. 9516–9, June 2011. cited on page 1.

[11] C. Mora, D. P. Tittensor, S. Adl, A. G. B. Simpson, and B. Worm, "How many species are there on Earth and in the ocean?," *PLOS Biology*, vol. 9, p. e1001127, Aug. 2011. cited on page 1.

[12] J. M. Harkness, "In Appreciation A Lifetime of Connections: Otto Herbert Schmitt, 1913 - 1998," *Physics in Perspective (PIP)*, vol. 4, pp. 456–490, Dec. 2002. cited on page 2.

[13] N. Mø bjerg, K. A. Halberg, A. Jø rgensen, D. Persson, M. Bjø rn, H. Ramlø v, and R. M. Kristensen, "Survival in extreme environments - on the current knowledge of adaptations in tardigrades.," *Acta physiologica*, vol. 202, pp. 409–20, July 2011. cited on page 1.

[14] A. Popescu, "BIONICS, BIOLOGICAL SYSTEMS AND THE PRINCIPLE OF OPTIMAL DESIGN," *Acta biotheoretica*, vol. 46, no. 4, pp. 299–310, 1998. cited on page 1.

[15] M. Burrows, "Biomechanics: froghopper insects leap to new heights.," *Nature*, vol. 424, p. 509, July 2003. cited on page 1.

C.6 Chapter 2

[16] K. Kühnlenz, S. Sosnowski, and M. Buss, "Evaluating Emotion Expressing Robots in Affective Space," in *Human Robot Interaction* (N. Sarkar, ed.), ch. 12, pp. 1–13, I-Tech Education and Publishing, Sept. 2007. cited on pages 14, 32, and 36.

[17] L. Z. Tiedens, "Powerful emotions: The vicious cycle of social status positions and emotions.," in *Emotions in the workplace: Research, theory, and practice* (N. M. Ashkanasy, C. E. Härtel, and W. J. Zerbe, eds.), pp. 72–81, Westport, CT, US: Quorum Books/Greenwood Publishing Group, 2000. cited on page 47.

[18] T. Minato, M. Shimada, H. Ishiguro, and S. Itakura, "Development of an Android Robot for Studying Human-Robot Interaction," in *Innovations in Applied Artificial Intelligence SE - 44* (B. Orchard, C. Yang, and M. Ali, eds.), vol. 3029 of *Lecture Notes in Computer Science*, pp. 424–434, Springer Berlin Heidelberg, 2004. cited on page 17.

[19] C. Darwin, P. Ekman, and P. Prodger, *The Expression of the Emotions in Man and Animals*. Oxford, New York: Oxford University Press, 3rd ed., 2002. cited on pages 13, 19, 20, 46, 47, and 60.

[20] T. O. Meservy, M. L. Jensen, J. Kruse, J. K. Burgoon, J. F. Nunamaker, D. P. Twitchell, G. Tsechpenakis, and D. N. Metaxas, "Deception Detection through Automatic, Unobtrusive Analysis of Nonverbal Behavior," *IEEE Intelligent Systems*, vol. 20, no. 5, pp. 36–43, 2005. cited on page 62.

[21] T. Hashimoto, S. Hiramatsu, and H. Kobayashi, "Development of Face Robot for Emotional Communication between Human and Robot," in *Proceedings of the International Conference on Mechatronics and Automation*, pp. 25–30, IEEE, June 2006. cited on pages 17, 42, and 53.

[22] T. Izawa, Y. Nakanishi, and N. Ito, "Development of stiffness changeable multijoint cervical structure with soft sensor flesh for musculo-skeletal humanoids," in *Proceedings of the IEEE/RAS International Conference on Humanoid Robots*, pp. 665–670, IEEE, Dec. 2010. cited on page 42.

[23] D. Matsumoto and H. C. Hwang, "Cultural Similarities and Differences in Emblematic Gestures," *Journal of Nonverbal Behavior*, vol. 37, pp. 1–27, Oct. 2012. cited on page 46.

[24] M. Buß, *Transferability of Social-Psychological Theories to HRI : Development and Evaluation of an Emotional Communication Module*. Bachelorthesis, Technische Universität München, 2011. cited on page 32.

[25] S. Duclos and J. Laird, "Emotion-specific effects of facial expressions and postures on emotional experience.," *Journal of Personality and Social Psychology*, vol. 57, no. 1, pp. 100–108, 1989. cited on page 47.

[26] I. Mizuuchi, Y. Nakanishi, Y. Sodeyama, Y. Namiki, T. Nishino, J. Urata, K. Hongo, T. Yoshikai, and M. Inaba, "An advanced musculoskeletal humanoid Kojiro," in *Proceedings of the 7th IEEE/RAS International Conference on Humanoid Robots*, pp. 294–299, IEEE, Nov. 2007. cited on page 42.

[27] M. Argyle and M. Cook, *Gaze and mutual gaze.* Oxford, England: Cambridge University Press, 1976. cited on page 47.

[28] J. D. Morris, "Observations: SAM: The Self-Assessment Manikin; An Efficient Cross-Cultural Measurement of Emotional Response," *Journal of Advertising Research*, vol. 35, no. 8, pp. 38–63, 1995. cited on page 31.

[29] J. Panksepp, "Toward a general psychobiological theory of emotions.," *Behavioral and Brain Sciences*, vol. 5, no. 3, pp. 407–467, 1982. cited on page 13.

[30] M. M. Bradley and P. Lang, "Emotion and Motivation," in *Handbook of Psychophysiology* (J. Cacioppo, L. Tassinary, and G. Berntson, eds.), vol. Second Edi, ch. 22, pp. 602 – 642, 2nd ed., 2007. cited on page 11.

[31] R. Cowie, N. Sussman, and A. Ben-Ze'ev, "Emotion: Concepts and Definitions," in *Emotion-Oriented Systems* (R. Cowie, C. Pelachaud, and P. Petta, eds.), Cognitive Technologies, pp. 9–30, Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. cited on page 11.

[32] C. L. Breazeal, *Designing Sociable Robots*. The MIT Press, May 2002. cited on pages 16, 21, 42, and 43.

[33] D. Omrčen and A. Ude, "Redundancy Control of a Humanoid Head for Foveation and Three-Dimensional Object Tracking: A Virtual Mechanism Approach," *Advanced Robotics*, vol. 24, pp. 2171–2197, Jan. 2010. cited on pages 42, 43, and 52.

[34] J. Hollerbach and K. Suh, "Redundancy resolution of manipulators through torque optimization," *IEEE Journal of Robotics and Automation*, vol. 3, no. 4, pp. 308–316, 1987. cited on page 51.

[35] R. Shine, "Function and evolution of the frill of the frillneck lizard, Chlamydosaurus kingii (Sauria: Agamidae)," *Biological Journal of the Linnean Society*, vol. 40, pp. 11–20, May 1990. cited on page 20.

[36] K. Dautenhahn, C. L. Nehaniv, M. L. Walters, B. Robins, H. Kose-Bagci, N. A. Mirza, and M. Blow, "KASPAR - a minimally expressive humanoid robot for human-robot interaction research," *Applied Bionics and Biomechanics*, vol. 6, pp. 369–397, Nov. 2009. cited on page 16.

[37] K. Suzuki, R. Hikiji, and S. Hashimoto, "Development of an Autonomous Humanoid Robot, iSHA, for Harmonized Human-Machine Environment," *Journal of Robotics and Mechatronics*, vol. 14, no. 5, pp. 324–332, 2002. cited on page 42.

[38] P. Ekman, "Facial Expressions," in *The Handbook of Cognition and Emotion* (T. Dalgleish and T. Power, eds.), pp. 301–320, Sussex, U.K.: John Wiley & Sons, Ltd., 1999. cited on page 17.

[39] J. K. Lee and C. L. Breazeal, "Human social response toward humanoid robot's head and facial features," in *Proceedings of the 28th International Conference on Human Factors in Computing Systems (CHI EA)*, (New York, New York, USA), pp. 37–42, ACM Press, 2010. cited on page 43.

[40] B. P. Yuhas, M. H. J. Goldstein, and T. J. Sejnowski, "Integration of Acoustic and Visual Speech Signals Using Neural Networks," *IEEE Communications Magazine*, no. November, pp. 65–71, 1989. cited on page 24.

[41] C. Snowdon, "Expression of emotion in nonhuman animals," in *Handbook of Affective Sciences* (R. J. Davidson, K. R. Scherer, and H. H. Goldsmith, eds.), pp. 457–480, Oxford, New York: Oxford University Press, 2003. cited on pages 18, 19, and 153.

[42] P. Ekman and W. V. Friesen, "The repertoire of nonverbal behavior: Categories, origins, usage, and coding," *Semiotica*, vol. 1, pp. 49–98, 1969. cited on page 11.

[43] M. Scheeff, J. Pinto, K. Rahardja, S. Snibbe, and R. Tow, "Experiences with Sparky, a Social Robot," in *Socially Intelligent Agents SE - 21* (K. Dautenhahn, A. Bond, L. Cañamero, and B. Edmonds, eds.), vol. 3 of *Multiagent Systems, Artificial Societies, and Simulated Organizations*, pp. 173–180, Springer US, 2002. cited on pages 16 and 61.

[44] R. Plutchik, "A general psychoevolutionary theory of emotion," in *Emotion: Theory, research, and experience* (R. Plutchik and H. Kellerman, eds.), pp. 3–31, New York: Academic Press, 1st ed., 1980. cited on pages 13 and 23.

[45] J. A. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *Journal of Research in Personality*, vol. 11, pp. 273–294, Sept. 1977. cited on pages 14 and 23.

[46] J. G. Carlson and E. Hatfield, *Psychology of Emotion*. Wadsworth Publishing Co Inc, 1992. cited on page 11.

[47] M. Buss, D. Carton, B. Gonsior, K. Kühnlenz, C. Landsiedel, N. Mitsou, R. de Nijs, J. Zlotowski, S. Sosnowski, E. Strasser, M. Tscheligi, A. Weiss, and D. Wollherr, "Towards Proactive Human-Robot Interaction in Human Environments," in *Proceedings of the 2nd International Conference on Cognitive Infocommunications (CogInfoCom)*, (Budapest), pp. 1–6, IEEE, 2011. cited on pages 29 and 59.

[48] B. Kühnlenz, S. Sosnowski, M. Buß, D. Wollherr, K. Kühnlenz, and M. Buss, "Increasing Helpfulness towards a Robot by Emotional Adaption to the User," *International Journal of Social Robotics*, Mar. 2013. cited on pages 31, 32, and 149.

[49] A. Mignault and A. Chaudhuri, "The many faces of a neutral face: Head tilt and perception of dominance and emotion," *Journal of Nonverbal Behavior*, vol. 27, no. 2, 2003. cited on page 47.

[50] A. Mehrabian and S. Ksionzky, *A theory of affiliation*. Lexington, Massachussetts: Lexington Books, 1974. cited on page 14.

[51] S. Sosnowski, K. Kühnlenz, and M. Buss, "EDDIE - An Emotion Display with Dynamic Intuitive Expressions," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (K. Kuehnlenz, ed.), pp. 2–2, IEEE, Oct. 2006. cited on pages 29, 31, 32, and 34.

[52] M. Mori, "The Uncanny Valley," *Energy*, vol. 7, no. 4, pp. 33–35, 1970. cited on pages 34 and 56.

[53] P. Lang, "Behavioral treatment and bio-behavioral assessment: computer applications," in *Technology in Mental Health Care Delivery Systems* (J. H. Sidowski, J. H. Johnson, and T. A. Williams, eds.), pp. 119–137, Norwood, Nj: Ablex, 1980. cited on pages 31 and 154.

[54] T. Minato, M. Shimada, S. Itakura, and H. Ishiguro, "Does Gaze Reveal the Human Likeness of an Android?," in *Proceedings of the 4th International Conference on Development and Learning*, pp. 106–111, IEEE, 2005. cited on page 17.

[55] K. Kühnlenz, S. Sosnowski, and M. Buss, "Impact of Animal-Like Features on Emotion Expression of Robot Head EDDIE," *Advanced Robotics*, vol. 24, pp. 1239–1255, Jan. 2010. cited on pages 16, 32, 35, and 36.

[56] A. Mehrabian, *NonVerbal Communication*. Aldine Transaction, 2007. cited on page 10.

[57] K. Kühnlenz, *Aspects of Multi-Focal Vision*. Dissertation, Technische Universität München, 2007. cited on page 42.

[58] C. L. Breazeal, "Affective interaction between humans and robots," in *Advances in Artificial Life*, 2001. cited on page 9.

[59] F. Hara, "Artificial emotion of face robot through learning in communicative interactions with human," in *Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN)*, pp. 7–15, IEEE, 2004. cited on page 16.

[60] M. Schröder, "Dimensional emotion representation as a basis for speech synthesis with non-extreme emotions," in *Proceedings of the Workshop on Affective Dialogue Systems*, (Kloster Irsee, Germany), pp. 209–220, 2004. cited on pages 155, 156, and 157.

[61] H. Chen, L. Zhang, Z. Wang, K. H. Yang, and A. I. King, "Biomechanics of the Neck," in *Theoretical Biomechanics*, 2011. cited on page 41.

[62] M. Zecca, S. Roccella, M. Carrozza, H. Miwa, K. Itoh, G. Cappiello, J.-J. Cabibihan, M. Matsumoto, H. Takanobu, P. Dario, and A. Takanishi, "On the development of the emotion expression humanoid robot WE-4RII with RCH-1," in *Proceedings of the 4th IEEE/RAS International Conference on Humanoid Robots* (S. Roccella, ed.), vol. 1, pp. 235–252, IEEE, 2004. cited on pages 17 and 43.

[63] G. Collier and D. DiCarlo, *Emotional expression*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc., 1985. cited on page 24.

[64] L. E. Bush, "Individual differences multidimensional scaling of adjectives denoting feelings.," *Journal of Personality and Social Psychology*, vol. 25, no. 1, pp. 50–57, 1973. cited on page 14.

[65] U. Hadar, T. J. Steiner, and F. Clifford Rose, "Head movement during listening turns in conversation," *Journal of Nonverbal Behavior*, vol. 9, no. 4, pp. 214–228, 1985. cited on page 46.

[66] M. Coulson, "Attributing Emotion to Static Body Postures: Recognition Accuracy, Confusions, and Viewpoint Dependence," *Journal of Nonverbal Behavior*, vol. 28, no. 2, pp. 117–139, 2004. cited on page 47.

[67] N. Esau, B. Kleinjohann, L. Kleinjohann, and D. Stichling, "MEXI: Machine with Emotionally eXtended Intelligence - A Software Architecture for Behavior Based Handling of Emotions and Drive," in *Proceedings of the International Conference on Hybrid and Intelligent Systems (HIS)*, 2003. cited on page 16.

[68] N. H. Frijda, *The emotions: Studies in emotion and social interaction*. Cambridge University Press, 1986. cited on pages 11, 13, and 47.

[69] U. Hess, S. Blairy, and R. Kleck, "The Influence of Facial Emotion Displays, Gender, and Ethnicity on Judgments of Dominance and Affiliation," *Journal of Nonverbal Behavior*, vol. 24, no. 4, pp. 265–283, 2000. cited on page 47.

[70] E. T. Hall, *The Hidden Dimension*. Garden City, New York: Anchor, 1990. cited on page 52.

[71] R. Cowie and E. Douglas-Cowie, "Speakers and hearers are people: Reflections on speech deterioration as a consequence of acquired deafness," *Profound Deafness and Speech Communication*, pp. 510–527, 1995. cited on page 10.

[72] H. Gray, "Anatomy of the Human Body," 1918. cited on page 43.

[73] C. Becker-Asano and H. Ishiguro, "Evaluating facial displays of emotion for the android robot Geminoid F," in *Proceedings of the IEEE Workshop on Affective Computational Intelligence (WACI)*, pp. 1–8, IEEE, Apr. 2011. cited on page 17.

[74] R. S. Lazarus, *Emotion and Adaptation*. Oxford University Press, 1991. cited on page 13.

[75] J.-h. Oh, D. Hanson, W.-s. Kim, Y. Han, J.-y. Kim, and I.-w. Park, "Design of Android type Humanoid Robot Albert HUBO," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1428–1433, IEEE, Oct. 2006. cited on page 17.

[76] H. Kim, B. Lau, and J. Triesch, "Adaptive object tracking with an anthropomorphic robot head," in *From Animals to Animats 8: Proceedings of the 8th International Conference on Simulation of Adaptive Behavior* (S. Schaal, A. Ijspeert, A. G. Billard, S. Vijayakumar, J. Hallam, and J.-A. Meyer, eds.), no. July, pp. 13–17, MIT Press, 2004. cited on page 41.

[77] B. Bates and J. Cleese, *The Human Face*. BBC Books, 2001. cited on page 17.

[78] P. Ekman, W. V. Friesen, M. O'Sullivan, and K. R. Scherer, "Relative importance of face, body, and speech in judgments of personality and affect.," *Journal of Personality and Social Psychology*, vol. 38, no. 2, pp. 270–277, 1980. cited on page 10.

[79] B. Bowman and D. Robbins, "Parameter study of biomechanical quantities in analytical neck models," *Proceedings of the Stapp Car Crash Conference*, vol. 16, 1972. cited on page 41.

[80] T. Dutoit, V. Pagel, N. Pierret, F. Bataille, and O. van der Vrecken, "The MBROLA project: towards a set of high quality speech synthesizers free of use for non commercial purposes," in *Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP)*, vol. 3, pp. 1393–1396, IEEE, 1996. cited on page 156.

[81] C.-H. Hjortsjö, *Människans ansikte och mimiska spraket*. Malmö, Sweden: Studentlitteratur, 1969. cited on page 17.

[82] K. Dautenhahn, "Methodology & Themes of Human-Robot Interaction: A Growing Research Field," *International Journal of Advanced Robotic Systems*, vol. 4, no. 1, p. 1, 2007. cited on page 8.

[83] C. Pelachaud, N. I. Badler, and M. Steedman, "Generating Facial Expressions for Speech," *Cognitive Science*, vol. 20, no. 1, pp. 1–46, 1996. cited on page 24.

[84] J. A. Russell, "Core Affect and the Psychological Construction of Emotion," *Psychological Review*, vol. 110, no. 1, pp. 145–172, 2003. cited on page 12.

[85] R. Cowie, C. Pelachaud, and P. Petta, eds., *Emotion-Oriented Systems - The Humaine Handbook*. Cognitive Technologies, Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. cited on page 12.

[86] D. Sander, D. Grandjean, and K. R. Scherer, "A systems approach to appraisal mechanisms in emotion.," *Neural networks : the Official Journal of the International Neural Network Society*, vol. 18, pp. 317–52, May 2005. cited on page 13.

[87] K. R. Scherer, "Studying the emotion-antecedent appraisal process: the expert system approach," *Cognition and Emotion*, vol. 7, pp. 325–355, 1993. cited on page 13.

[88] H. Kozima, "Infanoid," in *Socially Intelligent Agents SE - 19* (K. Dautenhahn, A. Bond, L. Cañamero, and B. Edmonds, eds.), vol. 3 of *Multiagent Systems, Artificial Societies, and Simulated Organizations*, pp. 157–164, Springer US, 2002. cited on page 16.

[89] R. Beira, M. Lopes, M. Praca, J. Santos-Victor, A. Bernardino, G. Metta, F. Becchi, and R. J. Saltaren, "Design of the robot-cub (iCub) head," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (M. Lopes, ed.), pp. 94–100, IEEE, 2006. cited on pages 16 and 42.

[90] J. Panksepp, "At the interface of the affective, behavioral, and cognitive neurosciences: Decoding the emotional feelings of the brain," *Brain and Cognition*, vol. 52, pp. 4–14, June 2003. cited on page 11.

[91] A. van Breemen, X. Yan, and B. Meerbeek, "iCat: an animated user-interface robot with personality," in *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, (New York, New York, USA), p. 143, ACM Press, 2005. cited on page 16.

[92] A. Liégeois, "Automatic Supervisory Control of the Configuration and Behavior of Multibody Mechanisms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 7, no. 12, pp. 868–871, 1977. cited on page 51.

[93] P. Ekman, W. V. Friesen, and J. C. Hager, *Facial Action Coding System*. Salt Lake City, Utah, USA: A Human Face, 2002. cited on page 152.

[94] K. Waldron and J. Schmiedeler, "Kinematics," in *Springer Handbook of Robotics* (B. Siciliano and O. Khatib, eds.), ch. 1, pp. 9–33, Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. cited on pages 41 and 45.

[95] D. Martinelli, *A Critical Companion to Zoosemiotics: People, Paths, Ideas*, vol. 5 of *Biosemiotics*. Dordrecht: Springer Science+Business Media B.V. Netherlands, 2010. cited on page 18.

[96] A. Bauer, K. Klasing, G. Lidoris, Q. Mühlbauer, F. Rohrmüller, S. Sosnowski, T. Xu, K. Kühnlenz, D. Wollherr, and M. Buss, "The Autonomous City Explorer: Towards Natural Human-Robot Interaction in Urban Environments," *International Journal of Social Robotics*, vol. 1, pp. 127–140, Feb. 2009. cited on page 9.

[97] M. Schröder and J. Trouvain, "The German text-to-speech synthesis system MARY: A tool for research, development and teaching," *International Journal of Speech Technology*, vol. 6, pp. 365–377, 2003. cited on page 155.

[98] H. K. M. Meeren, C. C. R. J. van Heijnsbergen, and B. de Gelder, "Rapid perceptual integration of facial expression and emotional body language.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, pp. 16518–23, Nov. 2005. cited on page 15.

[99] A. Ortony, G. Clore, and A. Collins, *Cognitive Structure of Emotions*. Cambridge University Press, 1988. cited on page 13.

[100] P. Ekman, "An argument for basic emotions," *Cognition and Emotion*, vol. 6, pp. 169–200, 1992. cited on page 13.

[101] H. Patterson, F. E. Pollick, and A. Sanford, "The role of velocity in affect discrimination," in *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, 2001. cited on page 47.

[102] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament," *Current Psychology*, vol. 14, pp. 261–292, Dec. 1996. cited on pages 12, 14, 23, and 84.

[103] A. Bittermann, K. Kühnlenz, and M. Buss, "On the Evaluation of Emotion Expressing Robots," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2138–2143, IEEE, 2007. cited on page 39.

[104] A. Waldeyer and U. Waldeyer, *Waldeyer - Anatomie des Menschen*. Berlin, New York: Walter de Gruyter, 12th ed., 1975. cited on page 44.

[105] A. N. Meltzoff and M. K. Moore, "Newborn infants imitate adult facial gestures," *Child Development*, vol. 54, no. 3, pp. 702–709, 1983. cited on page 62.

[106] P. Ekman, W. V. Friesen, M. O'Sullivan, A. Chan, I. Diacoyanni-Tarlatzis, K. Heider, R. Krause, W. A. LeCompte, T. Pitcairn, and P. E. Ricci-Bitti, "Universals and cultural differences in the judgments of facial expressions of emotion.," *Journal of Personality and Social Psychology*, vol. 53, pp. 712–7, Oct. 1987. cited on pages 13, 32, 34, 72, 73, and 149.

[107] J. Wismans, M. Philippens, E. van Oorschot, D. Kallieris, and R. Mattern, "Comparison of Human Volunteer and Cadaver Head-Neck Response in Frontal Flexion," *SAE Technical Paper 872194*, 1987. cited on page 41.

[108] C. E. Osgood, "DIMENSIONALITY OF THE SEMANTIC SPACE FOR COMMUNICATION VIA FACIAL EXPRESSIONS," *Scandinavian Journal of Psychology*, vol. 7, pp. 1–30, Mar. 1966. cited on page 14.

[109] C. E. Izard, J. Kagan, and R. B. Zsjonc, *Emotion-Cognition Relationships and Human Development, Emotion, Cognition and Behavior*. Cambridge University Press, 1984. cited on page 13.

[110] J. Abildgaard and H. Scharfe, "A Geminoid as Lecturer," in *Social Robotics SE - 41* (S. S. Ge, O. Khatib, J.-J. Cabibihan, R. Simmons, and M.-A. Williams, eds.), vol. 7621 of *Lecture Notes in Computer Science*, pp. 408–417, Springer Berlin Heidelberg, 2012. cited on page 17.

[111] M. Curtis, "Engraving of Frill-necked Lizard (Chlamydosaurus kingii)," 1827. cited on page 25.

[112] G. H. Bower, "Emotional mood and memory," *American Psychologist*, vol. 31, pp. 129–149, 1991. cited on page 9.

[113] C. Mayer, *Facial Expression Recognition With A Three-Dimensional Face Model*. Dissertation, Technische Universität München, 2012. cited on pages 65, 68, and 85.

[114] C. Storti, *Speaking of India: Bridging the Communication Gap When Working With Indians*. Nicholas Brealey Publishing, 2007. cited on page 47.

[115] D. Grandjean, D. Sander, and K. R. Scherer, "Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization.," *Consciousness and Cognition*, vol. 17, pp. 484–95, June 2008. cited on page 12.

[116] C. Smith and L. Kirby, "Consequences require antecedents: toward a process model of emotion elicitation," *Feeling and Thinking: The Role of Affect in Social Cognition*, 2000. cited on page 13.

[117] K. Berns, C. Hillenbrand, and K. Mianowski, "The mechatonic design of a human-like robot head," in *Proceedings of the CISM-IFToMM Symposium on Robot Design, Dynamics, and Control (ROMANSY)*, pp. 263–270, 2006. cited on page 42.

[118] I. Lütkebohle, F. Hegel, S. Schulz, M. Hackel, B. Wrede, S. Wachsmuth, and G. Sagerer, "The bielefeld anthropomorphic robot head "Flobi"," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3384–3391, IEEE, May 2010. cited on page 17.

[119] W. Wundt, *Grundriss der Psychologie*. Leipzig: Alfred Körner Verlag, 13 ed., 1918. cited on page 14.

[120] H. Gunes, B. Schuller, and M. Pantic, "Emotion representation, analysis and synthesis in continuous space: A survey," in *Proceedings of the Automatic Face & Gesture Recognition and Workshops (FG 2011)*, pp. 827–834, IEEE, Mar. 2011. cited on page 11.

[121] B. Fehr and J. A. Russell, "Concept of emotion viewed from from a prototype perspective," *Journal of Experimental Psychology*, vol. 113, pp. 464–486, 1984. cited on page 11.

[122] J. A. Troll, *Regelung der Orientierung für ein 5DOF Genick*. Bachelorthesis, Technische Universität München, 2010. cited on page 54.

[123] D. Heylen, "Head gestures, gaze and the principles of conversational structure," *International Journal of Humanoid Robotics*, vol. 3, no. 3, pp. 1–27, 2006. cited on pages 46, 47, and 74.

[124] F. Williams and B. Sundene, "Dimensions of recognition: Visual vs. Vocal expression of emotion," *Educational Technology Research and Development*, vol. 13, no. 1, pp. 44–52, 1965. cited on page 14.

[125] T. Shibata, T. Mitsui, K. Wada, and K. Tanie, "Subjective Evaluation of Seal Robot: Paro -Tabulation and Analysis of Questionnaire Results," *Journal of Robotics and Mechatronics*, vol. 14, no. 1, pp. 13–19, 2002. cited on page 9.

[126] M. Argyle, *Bodily Communication*. London, UK: Methuen, 1975. cited on page 46.

[127] D. Morris, *Bodytalk: A world guide to gestures*. London, UK: Cape, 1994. cited on page 46.

[128] A. Kleinsmith and N. Bianchi-Berthouze, "Recognizing affective dimensions from body posture," *Affective Computing and Intelligent Interaction SE - 2*, pp. 1–12, 2007. cited on page 47.

[129] Lorenz, "À partir d'Eibl-Eibesfeldt 1970, tiré originellement de Lorenz, 1952.," 1952. cited on page 25.

[130] S. S. Tomkins and R. McCarter, "WHAT AND WHERE ARE THE PRIMARY AFFECTS? SOME EVIDENCE FOR A THEORY," *Perceptual and Motor Skills*, vol. 18, pp. 119–158, Feb. 1964. cited on page 47.

[131] G. Zoric, K. Smid, and I. S. Pandzic, "Facial Gestures : Taxonomy and Application of Embodied Conversational Agents," in *Conversational Informatics: An Engineering Approach* (T. Nishida, ed.), John Wiley & Sons, Ltd, 2007. cited on pages 24 and 27.

[132] J. Dimnet, "Biomechanical Models of the Head-Neck System," in *The Head-Neck Sensory Motor System* (A. Berthoz, W. Graf, and P. P. Vidal, eds.), Oxford University Press, Apr. 1992. cited on page 41.

[133] I. Eibl-Eibesfeldt, *Grundriss der vergleichenden Verhaltensforschung*. BuchVertrieb Blank GmbH, 8th ed., 2004. cited on pages 20 and 21.

[134] J. Saldien, K. Goris, S. Yilmazyildiz, W. Verhelst, and D. Lefeber, "On the Design of the Huggable Robot Probo," *Journal of Physical Agents*, vol. 2, no. 2, pp. 3–11, 2008. cited on page 16.

[135] A. Bartsch and S. Hübner, *Emotionale Kommunikation - ein integratives Modell*. Dissertation, Martin-Luther-Universität Halle-Wittenberg, 2004. cited on page 11.

[136] H. Kose-Bagci, E. Ferrari, K. Dautenhahn, D. S. Syrdal, and C. L. Nehaniv, "Effects of Embodiment and Gestures on Social Interaction in Drumming Games with a Humanoid Robot," *Advanced Robotics*, vol. 23, pp. 1951–1996, Jan. 2009. cited on page 43.

[137] C. E. Izard, *The Psychology of Emotions (Emotions, Personality, and Psychotherapy)*. Springer, 1 ed., 2004. cited on page 47.

[138] D. Janson, "Cacatua sulphurea citrinocristata," 2005. cited on page 25.

[139] P. R. Kleinginna and A. M. Kleinginna, "A categorized list of emotion definitions, with suggestions for a consensual definition," *Motivation and Emotion*, vol. 5, pp. 345–379, Dec. 1981. cited on pages 11 and 12.

[140] F. Nori, L. Jamone, G. Sandini, and G. Metta, "Accurate control of a human-like tendon-driven neck," in *Proceedings of the 7th IEEE/RAS International Conference on Humanoid Robots*, pp. 371–378, IEEE, 2007. cited on page 42.

[141] J. S. Lerner and D. Keltner, "Beyond valence: Toward a model of emotion-specific influences on judgement and choice," *Cognition and Emotion*, vol. 14, pp. 473–493, 2000. cited on page 9.

[142] M. Pantic and M. S. Bartlett, "Machine analysis of facial expressions," in *Face Recognition* (K. Delac and M. Grgic, eds.), pp. 377–416, Vienna, Austria: I-Tech Education and Publishing, 2007. cited on page 34.

[143] J. Bates, "The Role of Emotion in Believable Agents," in *Communications of the ACM*, vol. 37, pp. 122–125, Association for Computing Machinery, July 1994. cited on page 9.

[144] H. M. Rosenfeld and Hancks, "The Relationship of Verbal and Nonverbal Communication," in *The Nonverbal Context of Verbal Listener Responses* (M. R. Key, ed.), pp. 193–206, The Hague, The Netherlands: Walter de Gruyter, 1980. cited on pages 46 and 47.

[145] H. Wallbott, "Bodily expression of emotion," *European Journal of Social Psychology*, vol. 28, pp. 879–896, 1998. cited on page 47.

[146] D. H. Laananen, "Computer Simulation of an Aircraft Seat and Occupant(s) in a Crash Environment - Program SOM-LA/SOM-TA User Manual," tech. rep., Arizona State University Departmnent of Mechanical and Aerospace Engineering, Tempe, Arizona, 1991. cited on page 41.

[147] D. M. Brouwer, J. Bennik, J. Leideman, H. M. J. R. Soemers, and S. Stramigioli, "Mechatronic design of a fast and long range 4 degrees of freedom humanoid neck," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 574–579, IEEE, May 2009. cited on page 42.

[148] S. Vijayakumar, J. Conradt, T. Shibata, and S. Schaal, "Overt visual attention for a humanoid robot," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 4, pp. 2332–2337, IEEE, 2001. cited on page 42.

[149] J. W. Pepper, *The behavioral ecology of the glossy black-cockatoo Calyptorhynchus lathami halmaturinus*. Dissertation, The University of Michigan Ann Arbor, 1996. cited on page 19.

[150] H. Bless, N. Schwarz, and M. Kemmelmeier, "Mood and stereotyping: The impact of moods on the use of general knowledge structures," *European Review of Social Psychology*, vol. 7, pp. 63–93, 1996. cited on page 9.

[151] B. M. Waller, J. J. Cray, and A. M. Burrows, "Selection for universal facial emotion.," *Emotion*, vol. 8, pp. 435–9, June 2008. cited on page 17.

[152] C. L. Kleinke, "Gaze and eye contact: A research review.," *Psychological Bulletin*, vol. 100, no. 1, pp. 78–100, 1986. cited on page 47.

[153] M. B. Arnold, *Emotion and Personality. Vol. I. Psychological aspects.* Columbia University Press, 1960. cited on page 13.

[154] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and Autonomous Systems*, vol. 42, pp. 143–166, Mar. 2003. cited on page 16.

[155] L. Roos, F. Guenter, A. Guignard, and A. G. Billard, "Design of a Biomimetic Spine for the Humanoid Robot Robota," in *Proceedings of the First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)* (F. Guenter, ed.), pp. 329–334, IEEE, 2006. cited on page 42.

[156] D. E. Whitney, "Resolved Motion Rate Control of Manipulators and Human Prostheses," *IEEE Transactions on Man-Machine Systems*, vol. 10, no. 2, pp. 47–53, 1969. cited on page 51.

[157] M. Hans, B. Graf, and R. Schraft, "Robotic home assistant Care-O-bot: past-present-future," in *Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication*, pp. 380–385, IEEE, 2002. cited on page 9.

[158] T. Hashimoto, H. Kobayashi, and N. Kato, "Educational system with the android robot SAYA and field trial," in *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 766–771, IEEE, June 2011. cited on page 17.

[159] K. Bousmalis, M. Mehu, and M. Pantic, "Spotting agreement and disagreement: A survey of nonverbal audiovisual cues and tools," in *Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops (ACII)*, (Amsterdam), pp. 1–9, IEEE, Sept. 2009. cited on page 46.

[160] C. M. Whissell, "The dictionary of affect in language," in *Emotion: Theory, Research and Experience* (R. Plutchik and H. Kellerman, eds.), vol. 4, pp. 113–131, New York, NY, USA: Academic Press, 1989. cited on page 23.

[161] R. Reilink, L. C. Visser, J. Bennik, R. Carloni, D. M. Brouwer, and S. Stramigioli, "The Twente humanoid head," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1593–1594, IEEE, May 2009. cited on pages 42, 43, and 52.

[162] T. Fukuda, M. Nakashima, F. Arai, and Y. Hasegawa, "Facial expressive robotic head system for human-robot communication and its application in home environment," *Proceedings of the IEEE*, vol. 92, pp. 1851–1865, Nov. 2004. cited on page 16.

[163] R. A. Spitz, *No and Yes. On the genesis of human communication.* New York, NY, USA: International Universities Press, 1957. cited on page 10.

[164] L. D. Cañamero, "Playing the Emotion Game with Feelix," in *Socially Intelligent Agents SE - 8* (K. Dautenhahn, A. Bond, L. Cañamero, and B. Edmonds, eds.), vol. 3 of *Multiagent Systems, Artificial Societies, and Simulated Organizations*, pp. 69–76, Springer US, 2002. cited on page 16.

[165] D. Keltner and C. Anderson, "Saving Face for Darwin : The Functions and Uses of Embarrassment," *Current Directions in Psychological Science*, vol. 9, no. 6, pp. 187–192, 2000. cited on page 47.

[166] J. A. Russell, "A Circumplex Model of Affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980. cited on pages 14, 22, 23, and 26.

[167] C. A. Smith and H. S. Scott, "A Componential Approach to the meaning of facial expressions," in *The Psychology of Facial Expression* (J. A. Russell and J. M. Fernández-Dols, eds.), ch. 10, pp. 229–254, Cambridge University Press, 1997. cited on pages 22 and 29.

[168] C. L. Breazeal, A. Takanishi, and T. Kobayashi, "Social Robots that Interact with People," in *Springer Handbook of Robotics SE - 59* (B. Siciliano and O. Khatib, eds.), pp. 1349–1369, Springer Berlin Heidelberg, 2008. cited on page 16.

[169] "Childlike characteristics," 2009. cited on page 25.

[170] A. Takanishi, T. Matsuno, and I. Kato, "Development of an anthropomorphic head-eye robot with two eyes-coordinated head-eye motion and pursuing motion in the depth direction," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robot and Systems (IROS)*, vol. 2, pp. 799–804, IEEE, 1997. cited on pages 43 and 53.

[171] J. B. Watson, *Behaviorism; with a new introduction by Gregory A. Kimble*. New Brunswick, New Jersey: Transaction Publishers, 7th ed., 2009. cited on page 13.

[172] H. McGurk and J. MacDonald, "Hearing lips and seeing voices," *Nature*, vol. 264, pp. 746–748, Dec. 1976. cited on page 24.

[173] L. Schmidt-Atzert, *Lehrbuch der Emotionspsychologie*. Kohlhammer, 1996. cited on page 12.

[174] D. Hanson, A. Olney, I. A. Pereira, and M. Zielke, "Upending the Uncanny Valley," in *Proceedings of the 20th National Conference on Artificial Intelligence - Volume 4*, no. July, (Pittsburgh, PA.), pp. 1728–1729, AAAI Press, 2005. cited on page 17.

[175] C. F. DiSalvo, F. Gemperle, J. Forlizzi, and S. Kiesler, "All robots are not created equal," in *Proceedings of the Conference on Designing Interactive Systems Processes, Practices, Methods, and Techniques (DIS)*, (New York, New York, USA), p. 321, ACM Press, June 2002. cited on pages 16 and 21.

[176] M. Bennewitz, F. Faber, D. Joho, and S. Behnke, "Fritz - A Humanoid Communication Robot," in *Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (F. Faber, ed.), pp. 1072–1077, IEEE, 2007. cited on page 17.

[177] W. J. Smith, *The Behavior of Communicating: An Ethological Approach*. Cambridge, Massachusetts: Harvard University Press, 1977. cited on pages 18, 19, and 153.

[178] C. Bartneck and M. Okada, "Robotic user interfaces," in *Proceedings of the Human and Computer Conference*, pp. 130–140, 2001. cited on page 16.

[179] D. Morris, *Manwatching: A field guide to human behavior*. New York, New York, USA: H. N. Abrams, 1977. cited on page 47.

[180] E. Zimen, *Der Hund: Abstammung - Verhalten - Mensch und Hund*. Goldmann, 1992. cited on pages 19 and 20.

[181] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *IEEE Signal Processing*, no. January, pp. 32–80, 2001. cited on pages 10, 12, and 156.

C.7 Chapter 3

[182] U. Dimberg, "For Distinguished Early Career Contribution to Psychophysiology: Award Address, 1988," *Psychophysiology*, vol. 27, pp. 481–494, Sept. 1990. cited on pages 60 and 62.

[183] S. Stepper and F. Strack, "Proprioceptive determinants of emotional and nonemotional feelings," *Journal of Personality and Social Psychology*, vol. 64, pp. 211–220, 1993. cited on page 64.

[184] A. Kendon, "Movement coordination in social interaction: Some examples described," *Acta Psychologica*, vol. 32, pp. 101–125, Jan. 1970. cited on pages 63 and 64.

[185] F. Hegel, T. Spexard, B. Wrede, G. Horstmann, and T. Vogt, "Playing a different imitation game: Interaction with an Empathic Android Robot," in *Proceedings of the 6th IEEE/RAS International Conference on Humanoid Robots*, pp. 56–61, IEEE, Dec. 2006. cited on pages 61 and 75.

[186] T. L. Chartrand and J. A. Bargh, "The chameleon effect: the perception-behavior link and social interaction.," *Journal of Personality and Social Psychology*, vol. 76, pp. 893–910, June 1999. cited on pages 60, 62, and 63.

[187] T. O. Meservy, M. L. Jensen, J. Kruse, J. K. Burgoon, J. F. Nunamaker, D. P. Twitchell, G. Tsechpenakis, and D. N. Metaxas, "Deception Detection through Automatic, Unobtrusive Analysis of Nonverbal Behavior," *IEEE Intelligent Systems*, vol. 20, no. 5, pp. 36–43, 2005. cited on page 62.

[188] N. Bischof, "Zur Anwendung der Sytemtheorie in den Verhaltenswissenschaften," *Leopoldina. Mitteilungen der Deutschen Akademie der Naturforscher Leopoldina*, no. 31, pp. 129–131, 1985. cited on pages 58 and 75.

[189] N. Bischof, *Psychologie: Ein Grundkurs für Anspruchsvolle*. Stuttgart: Kohlhammer, 1 ed., 2008. cited on pages 75 and 76.

[190] R. Tanner and T. L. Chartrand, "The Convincing Chameleon: The Impact of Mimicry on Persuasion," *Journal of Consumer Research*, pp. 1–26, 2008. cited on pages 59, 63, and 64.

[191] J. K. Rempel and J. G. Homes, "How do I trust thee?," *Psychology Today*, vol. 20, no. 2, pp. 28–34, 1986. cited on page 64.

[192] M. Buß, *Transferability of Social-Psychological Theories to HRI : Development and Evaluation of an Emotional Communication Module*. Bachelorthesis, Technische Universität München, 2011. cited on page 32.

[193] D. Feil-Seifer and M. J. Mataric, "Socially Assistive Robotics," in *Proceedings of the 9th International Conference on Rehabilitation Robotics (ICORR)*, pp. 465–468, IEEE, 2005. cited on page 59.

[194] M. Kaitz, O. Meschulach-Sarfaty, J. Auerbach, and A. Eidelman, "A reexamination of newborns' ability to imitate facial expressions.," *Developmental Psychology*, vol. 24, no. 1, pp. 3–7, 1988. cited on page 62.

[195] R. B. van Baaren, R. W. Holland, B. Steenaert, and A. van Knippenberg, "Mimicry for money: Behavioral consequences of imitation," *Journal of Experimental Social Psychology*, vol. 39, pp. 393–398, July 2003. cited on page 64.

[196] A. Pentland, "Social dynamics: Signals and behavior," *International Conference on Developmental Learning*, 2004. cited on pages 59 and 62.

[197] W. W. Maddux, E. Mullen, and A. D. Galinsky, "Chameleons bake bigger pies and take bigger pieces: Strategic behavioral mimicry facilitates negotiation outcomes," *Journal of Experimental Social Psychology*, vol. 44, pp. 461–468, Mar. 2008. cited on pages 59 and 64.

[198] L. Carr, M. Iacoboni, M.-C. Dubeau, J. C. Mazziotta, and G. L. Lenzi, "Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, pp. 5497–502, Apr. 2003. cited on page 63.

[199] I. Borutta, *Ein Experiment zur Evaluierung des Zürcher Lächelmodells*. Diplomathesis, Technische Universität München, 2009. cited on page 167.

[200] K. Balci, E. Not, M. Zancanaro, and F. Pianesi, "Xface open source project and smil-agent scripting language for creating and animating embodied conversational agents," in *Proceedings of the 15th International Conference on Multimedia (MULTIMEDIA)*, (New York, New York, USA), p. 1013, ACM Press, 2007. cited on page 85.

[201] J. M. Dabbs Jr., "Similarity of gestures and interpersonal influence.," *Proceedings of the Annual Convention of the American Psychological Association*, vol. 4, no. 1, pp. 337–338, 1969. cited on page 64.

[202] M. Dapretto, M. S. Davies, J. H. Pfeifer, A. A. Scott, M. Sigman, S. Y. Bookheimer, and M. Iacoboni, "Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders," *Nature Neuroscience*, vol. 9, pp. 28–30, Jan. 2006. cited on page 64.

[203] M. Rehm and E. André, "Catch me if you can - Exploring Lying Agents in Social Settings," in *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, (New York, New York, USA), p. 937, ACM Press, 2005. cited on page 74.

[204] H. Prendinger and M. Ishizuka, "THE EMPATHIC COMPANION: A CHARACTER-BASED INTERFACE THAT ADDRESSES USERS' AFFECTIVE STATES," *Applied Artificial Intelligence*, vol. 19, pp. 267–285, Mar. 2005. cited on page 75.

[205] N. Reissland and P. Harris, "Children's use of display rules in pride-eliciting situations," *British Journal of Developmental Psychology*, vol. 9, pp. 431–435, Sept. 1991. cited on page 60.

[206] R. W. Picard, "Affective Computing," tech. rep., Perceptual Computing Section, Media Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1995. cited on page 59.

[207] F. D. Schönbrodt and J. B. Asendorpf, "The Challenge of Constructing Psychologically Believable Agents," *Journal of Media Psychology: Theories, Methods, and Applications*, vol. 23, pp. 100–107, Jan. 2011. cited on pages 75 and 78.

[208] U. Hess and A. H. Fischer, "Emotional Mimicry as Social Regulation.," *Personality and Social Psychology Review : an Official Journal of the Society for Personality and Social Psychology, Inc*, Jan. 2013. cited on pages 60, 62, 63, and 73.

[209] N. Bischof, *Das Rätsel Ödipus*. München: Piper, 1985. cited on pages 75 and 79.

[210] H. T. Reis, I. M. Wilson, C. Monestere, S. Bernstein, K. Clark, E. Seidl, M. Franco, E. Gioioso, L. Freeman, and K. Radoane, "What is smiling is beautiful and good," *European Journal of Social Psychology*, vol. 20, pp. 259–267, May 1990. cited on pages 59 and 73.

[211] M. Buss, D. Carton, B. Gonsior, K. Kühnlenz, C. Landsiedel, N. Mitsou, R. de Nijs, J. Zlotowski, S. Sosnowski, E. Strasser, M. Tscheligi, A. Weiss, and D. Wollherr, "Towards Proactive Human-Robot Interaction in Human Environments," in *Proceedings of the 2nd International Conference on Cognitive Infocommunications (CogInfoCom)*, (Budapest), pp. 1–6, IEEE, 2011. cited on pages 29 and 59.

[212] B. Kühnlenz, S. Sosnowski, M. Buß, D. Wollherr, K. Kühnlenz, and M. Buss, "Increasing Helpfulness towards a Robot by Emotional Adaption to the User," *International Journal of Social Robotics*, Mar. 2013. cited on pages 31, 32, and 149.

[213] A. Pentland, "Socially aware, computation and communication," *IEEE Computer*, vol. 38, no. 3, pp. 33–40, 2005. cited on pages 58 and 62.

[214] T. Kanade and J. Cohn, "Comprehensive database for facial expression analysis," in *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 46–53, IEEE Comput. Soc, 2000. cited on page 68.

[215] P. Ekman, "Universals and cultural differences in facial expressions of emotion," in *Symposium on Motivation* (J. Cole, ed.), vol. 19, (University of Nebraska), pp. 207–283, 1971. cited on page 73.

[216] U. Dimberg and A. Öhman, "Behold the wrath: Psychophysiological responses to facial stimuli," *Motivation and Emotion*, vol. 20, no. 2, pp. 149–182, 1996. cited on pages 60 and 62.

[217] L. D. Riek, P. C. Paul, and P. Robinson, "When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry," *Journal on Multimodal User Interfaces*, vol. 3, no. 1-2, pp. 99–108, 2010. cited on pages 61 and 75.

[218] I. Kotsia and I. Pitas, "Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 172–187, 2007. cited on page 67.

[219] J. J. McDowall, "Interactional synchrony: A reappraisal.," *Journal of Personality and Social Psychology*, vol. 36, no. 9, pp. 963–975, 1978. cited on page 63.

[220] A. Smith, *The Theory of Moral Sentiments*. 2010 ed., 1759. cited on page 60.

[221] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis.," *Psychological Bulletin*, vol. 111, no. 2, pp. 256–274, 1992. cited on page 59.

[222] I. Ajzen, "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes*, vol. 50, pp. 179–211, 1991. cited on page 60.

[223] T. Kanda, M. Kamasima, M. Imai, T. Ono, D. Sakamoto, H. Ishiguro, and Y. Anzai, "A humanoid robot that pretends to listen to route guidance from a human," *Autonomous Robots*, vol. 22, no. 1, pp. 87–100, 2007. cited on page 61.

[224] A. J. Fridlund, "Sociality of solitary smiling: Potentiation by an implicit audience," *Journal of Personality and Social Psychology*, vol. 60, no. 2, pp. 229–240, 1991. cited on page 73.

[225] E. Bevacqua, S. J. Hyniewska, and C. Pelachaud, "Positive influence of smile backchannels in ECAs," in *International Workshop on Interacting with ECAs as Virtual Characters*, pp. 13–20, 2010. cited on page 74.

[226] T. L. Chartrand, W. W. Maddux, and J. L. Lakin, "Beyond the Perception-Behavior Link: The Ubiquitous Utility and Motivational Moderators of Nonconscious Mimicry," in *The New Unconscious* (R. Hassin, J. Uleman, and J. Bargh, eds.), pp. 334–361, 2005. cited on page 64.

[227] J. N. Bailenson and N. Yee, "Digital chameleons: automatic assimilation of nonverbal gestures in immersive virtual environments.," *Psychological Science*, vol. 16, pp. 814–9, Oct. 2005. cited on pages 59 and 61.

[228] C. D. Kidd and C. L. Breazeal, "A robotic weight loss coach," in *Proceedings of the 22nd Conference on Artificial Intelligence*, pp. 1985–1986, AAAI Press, 2007. cited on page 59.

[229] E. J. Charny, "Psychosomatic Manifestations of Rapport in Psychotherapy," *Psychosomatic Medicine*, vol. 28, pp. 305–315, July 1966. cited on page 64.

[230] A. Bannat, J. Blume, J. T. Geiger, T. Rehrl, F. Wallhoff, C. Mayer, B. Radig, S. Sosnowski, and K. Kühnlenz, "A Multimodal Human-Robot-Dialog Applying Emotional Feedbacks," in *Social Robotics SE - 1* (S. Ge, H. Li, J.-J. Cabibihan, and Y. Tan, eds.), vol. 6414 of *Lecture Notes in Computer Science*, pp. 1–10, Springer Berlin Heidelberg, 2010. cited on page 59.

[231] P. Ekman, W. V. Friesen, and P. Ellsworth, *Gesichtssprache : Wege zur Objektivierung menschlicher Emotionen*. Wien: Hermann Böhlau Verlag, 1974. cited on page 73.

[232] J. M. V. D. Zwaan, "A BDI Dialogue Agent for Social Support : Specification of Verbal Support Types," in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)* (Conitzer, Winikoff, Padgham, and van der Hoek, eds.), (Valencia, Spain), 2012. cited on page 75.

[233] A. N. Meltzoff and M. K. Moore, "Imitation of facial and manual gestures by human neonates," *Science*, vol. 198, no. 4312, pp. 75–78, 1977. cited on pages 58, 62, and 65.

[234] J. M. Haviland and M. Lelwica, "The induced affect response: 10-week-old infants' responses to three emotion expressions.," *Developmental Psychology*, vol. 23, no. 1, pp. 97–104, 1987. cited on page 60.

[235] M. Anisfeld, J. C. Masters, S. W. Jacobson, J. Kagan, N. Andrew, M. K. Moore, A. B. Harris, and J. E. Vaughn, "Interpreting "Imitative" Responses in Early Infancy," *Science*, vol. 205, no. 4402, pp. 214–219, 1979. cited on pages 62 and 63.

[236] N. T. Termine and C. E. Izard, "Infants' responses to their mothers' expressions of joy and sadness," *Developmental Psychology*, vol. 24, no. 2, pp. 223–229, 1988. cited on page 63.

[237] N. Guéguen and A. Martin, "Incidental Similarity Facilitates Behavioral Mimicry," *Social Psychology*, vol. 40, pp. 88–92, Jan. 2009. cited on page 60.

[238] V. Gallese, "The shared manifold hypothesis. From mirror neurons to empathy," *Journal of Consciousness Studies*, vol. 8, no. 5-7, pp. 33–50, 2001. cited on page 64.

[239] N. C. Krämer, S. Kopp, C. Becker-Asano, and N. Sommer, "Smile and the world will smile with you—The effects of a virtual agent's smile on users' evaluation and behavior," *International Journal of Human-Computer Studies*, vol. 71, pp. 335–349, Mar. 2013. cited on page 74.

[240] M. LaFrance, "Nonverbal Synchrony and Rapport: Analysis by the Cross-Lag Panel Technique," *Social Psychology Quarterly*, vol. 42, no. 1, pp. 66–70, 1979. cited on pages 63 and 64.

[241] I. Borutta, S. Sosnowski, M. Zehetleitner, N. Bischof, and K. Kühnlenz, "Generating artificial smile variations based on a psychological system-theoretic approach," in *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 245–250, IEEE, Sept. 2009. cited on page 85.

[242] C. L. Breazeal, D. Buchsbaum, and J. Gray, "Learning From and About Others : Towards Using Imitation to Bootstrap the Social Understanding of Others by Robots," *Artificial Life*, vol. 11, no. 1-2, pp. 31–62, 2005. cited on page 16.

[243] U. Hess, "The communication of emotion," *Emotions, Qualia and Consciousness*, no. 1872, pp. 397–409, 2001. cited on page 60.

[244] E. Jakobs, a. S. R. Manstead, and a. H. Fischer, "Social Motives and Emotional Feelings as Determinants of Facial Displays: The Case of Smiling," *Personality and Social Psychology Bulletin*, vol. 25, pp. 424–435, Apr. 1999. cited on page 73.

[245] S. W. McQuiggan and J. C. Lester, "Modeling and evaluating empathy in embodied companion agents," *International Journal of Human-Computer Studies*, vol. 65, pp. 348–360, Apr. 2007. cited on page 75.

[246] H. Gubler and N. Bischof, "Computerspiele als Werkzeug der motivationspsychologischen Grundlagenforschung," *Zeitschrift für Psychologie*, no. 201, pp. 287–315, 1993. cited on pages 75, 78, and 79.

[247] M. Ochs, C. Pelachaud, and D. Sadek, "An empathic virtual dialog agent to improve human-machine interaction," in *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems (AA- MAS 2008)* (Padgham, Parkes, Müller, and Parsons, eds.), (Estoril, Portugal), pp. 89–96, 2008. cited on page 75.

[248] K. Lewin, "Behavior and development as a function of the total situation," in *Manual of Child Psychology*, pp. 791–844, Hoboken, New Jersey: John Wiley & Sons, 8 ed., 1946. cited on page 79.

[249] H. Cramer, J. Goddijn, B. Wielinga, and V. Evers, "Effects of (in)accurate empathy and situational valence on attitudes towards robots," in *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, (New York, New York, USA), p. 141, ACM Press, 2010. cited on page 75.

[250] W. James, *The principles of psychology.* NY, US: Henry Holt and Company, 1890. cited on pages 60 and 64.

[251] A. Raouzaiou, N. Tsapatsoulis, K. Karpouzis, and S. Kollias, "Parameterized Facial Expression Synthesis Based on MPEG-4," *EURASIP Journal on Advances in Signal Processing*, vol. 2002, pp. 1021–1038, Jan. 2002. cited on page 167.

[252] S.-H. Kang, J. Gratch, and J. H. Watt, "The effect of affective iconic realism on anonymous interactants' self-disclosure," in *Extended Abstracts on Human Factors in Computing Systems (CHI)*, (New York, NY, USA), pp. 4021–4026, ACM, 2009. cited on page 64.

[253] A. H. Fischer, D. Becker, and L. Veenstra, "Emotional mimicry in social context: the case of disgust and pride.," *Frontiers in Psychology*, vol. 3, p. 475, Jan. 2012. cited on page 63.

[254] T. L. Chartrand and V. E. Jefferis, "Consequences of automatic goal pursuit and the case of nonconscious mimicry," in *Responding to the social world: Implicit and explicit processes in social judgments and decisions* (J. P. Forgas, K. D. Williams, and W. von Hippel, eds.), pp. 1–26, Philadelphia: Psychology Press, 2003. cited on pages 60, 62, and 64.

[255] F. J. Bernieri, "Coordinated movement and rapport in teacher-student interactions," *Journal of Nonverbal Behavior*, vol. 12, no. 2, pp. 120–138, 1988. cited on pages 60 and 63.

[256] J. Kagan, D. Arcus, and N. Snidman, "Reactivity in infants: A cross-national comparison," *Developmental Psychology*, vol. 30, no. 3, pp. 342–345, 1994. cited on page 60.

[257] R. D. Young and M. Frye, "SOME ARE LAUGHING; SOME ARE NOT - WHY?," *Psychological Reports*, vol. 18, pp. 747–754, 1966. cited on page 63.

[258] E. Krumhuber, A. S. R. Manstead, D. Cosker, D. Marshall, and P. L. Rosin, "Effects of Dynamic Attributes of Smiles in Human and Synthetic Faces: A Simulated Job Interview Setting," *Journal of Nonverbal Behavior*, vol. 33, pp. 1–15, Nov. 2008. cited on page 74.

[259] W. S. Condon and W. D. Ogston, "Sound Film Analysis of Normal and Pathological Behavior Patterns," *The Journal of Nervous and Mental Disease*, vol. 143, no. 4, 1966. cited on page 63.

[260] B. Gonsior, S. Sosnowski, C. Mayer, J. Blume, B. Radig, D. Wollherr, and K. Kühnlenz, "Improving aspects of empathy and subjective performance for HRI through mirroring facial expressions," in *Proceedings of the 20st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 350–356, IEEE, July 2011. cited on pages 91, 93, and 96.

[261] A. Rafaeli and R. I. Sutton, "Expression of Emotion as Part of the Work Role," *The Academy of Management Review*, vol. 12, p. 23, Jan. 1987. cited on page 73.

[262] J. van der Schalk, A. H. Fischer, B. Doosje, D. Wigboldus, S. Hawk, M. Rotteveel, and U. Hess, "Convergent and divergent responses to emotional displays of ingroup and outgroup.," *Emotion*, vol. 11, pp. 286–98, Apr. 2011. cited on page 64.

[263] E. Hatfield, J. T. Cacioppo, and R. L. Rapson, *Emotional Contagion*, vol. 2. Cambridge: Cambridge University Press, 1994. cited on page 64.

[264] M. S. Bartlett, G. Littlewort, I. Fasel, and J. R. Movellan, "Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction.," in *Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop*, pp. 53–53, IEEE, June 2003. cited on page 61.

[265] U. Dimberg, "Facial Reactions to Facial Expressions," *Psychophysiology*, vol. 19, pp. 643–647, Nov. 1982. cited on pages 60 and 62.

[266] E. J. Moody, D. N. McIntosh, L. J. Mann, and K. R. Weisser, "More than mere mimicry? The influence of emotion on rapid facial reactions to faces.," *Emotion*, vol. 7, pp. 447–57, May 2007. cited on page 64.

[267] N. Simons, N. C. Krämer, and S. Kopp, "The effects of an embodied conversationational agent's subtle nonverbal behavior on user's evaluation and behavioral mimicry," in *Proceedings of the 57th Annual Conference of the International Communication Association*, 2007. cited on page 61.

[268] J. Ahlberg, "CANDIDE-3 - An Updated Parameterised Face," tech. rep., Linköping University, Sweden, 2001. cited on pages 67 and 161.

[269] K. Chisholm and J. Strayer, "Verbal and facial measures of children's emotion and empathy.," *Journal of Experimental Child Psychology*, vol. 59, pp. 299–316, Apr. 1995. cited on page 60.

[270] Y. Yabar, L. Johnston, L. Miles, and V. Peace, "Implicit Behavioral Mimicry: Investigating the Impact of Group Membership," *Journal of Nonverbal Behavior*, vol. 30, pp. 97–113, July 2006. cited on page 64.

[271] A. N. Meltzoff and M. K. Moore, "Newborn infants imitate adult facial gestures," *Child Development*, vol. 54, no. 3, pp. 702–709, 1983. cited on page 62.

[272] M. LaFrance and W. Ickes, "Posture mirroring and interactional involvement: Sex and sex typing effects," *Journal of Nonverbal Behavior*, vol. 5, no. 3, pp. 139–154, 1981. cited on pages 63 and 64.

[273] R. E. Kraut and R. E. Johnston, "Social and emotional messages of smiling: An ethological approach.," *Journal of Personality and Social Psychology*, vol. 37, no. 9, pp. 1539–1553, 1979. cited on page 73.

[274] J. Decety and P. Jackson, "A Social-Neuroscience Perspective on Empathy," *Current Directions in Psychological Science*, vol. 15, no. 2, pp. 54–58, 2006. cited on page 63.

[275] H. Gubler, M. Paffrath, and N. Bischof, "Untersuchungen zur Systemanalyse der sozialen Motivation III: Eine Ästimationsstudie zur Sicherheits- und Erregungsregulation während der Adoleszenz," *Zeitschrift für Psychologie*, vol. 202, no. 1, pp. 95–132, 1994. cited on page 75.

[276] P. T. Saunders, *An introduction to catastrophe theory*. Cambridge: Cambridge University Press, 1980. cited on page 79.

[277] J. L. Lakin, V. E. Jefferis, C. M. Cheng, and T. L. Chartrand, "The Chameleon Effect as Social Glue: Evidence for the Evolutionary Significance of Nonconscious Mimicry," *Journal of Nonverbal Behavior*, vol. 27, no. 3, pp. 145–162, 2003. cited on pages 62 and 64.

[278] J. Gratch, N. Wang, J. Gerten, E. Fast, and R. Duffy, "Creating Rapport with Virtual Agents," in *Intelligent Virtual Agents SE - 12* (C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, and D. Pelé, eds.), vol. 4722 of *Lecture Notes in Computer Science*, pp. 125–138, Springer Berlin Heidelberg, 2007. cited on page 61.

[279] C. Mayer, *Facial Expression Recognition With A Three-Dimensional Face Model*. Dissertation, Technische Universität München, 2012. cited on pages 65, 68, and 85.

[280] R. Niewiadomski and C. Pelachaud, "Model of Facial Expressions Management for an Embodied Conversational Agent," in *Affective Computing and Intelligent Interaction SE - 2* (A. Paiva, R. Prada, and R. Picard, eds.), vol. 4738 of *Lecture Notes in Computer Science*, pp. 12–23, Springer Berlin Heidelberg, 2007. cited on page 74.

[281] A. Tapus and M. J. Mataric, "Emulating Empathy in Socially Assistive Robotics Empathy in Socially Assistive Robotics," in *Proceedings of the AAAI Spring Symposium on Multidisciplinary Collaboration for Socially Assistive Robotics*, (Palo Alto, Stanford, U.S.A.), pp. 1–4, AAAI Press, 2007. cited on page 75.

[282] W. S. Condon and L. W. Sander, "Synchrony Demonstrated between Movements of the Neonate and Adult Speech," *Child Development*, vol. 45, no. 2, pp. 456–462, 1974. cited on pages 63 and 64.

[283] M. Goebl and G. Farber, "A Real-Time-capable Hard-and Software Architecture for Joint Image and Knowledge Processing in Cognitive Automobiles," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, pp. 734–740, IEEE Press, 2007. cited on page 65.

[284] U. Hess, P. Philippot, and S. Blairy, "Mimicry: facts and fiction," in *The Social Context of Nonverbal Behavior* (P. Philippot, R. Feldman, and E. Coats, eds.), pp. 213–241, New York, NY, USA: Oxford University Press, 1999. cited on page 64.

[285] D. Heylen, "Head gestures, gaze and the principles of conversational structure," *International Journal of Humanoid Robotics*, vol. 3, no. 3, pp. 1–27, 2006. cited on pages 46, 47, and 74.

[286] U. Dimberg, M. Thunberg, and K. Elmehed, "Unconscious Facial Reactions to Emotional Facial Expressions," *Psychological Science*, vol. 11, pp. 86–89, Jan. 2000. cited on pages 60 and 62.

[287] U. Dimberg, "Facial expressions as excitatory and inhibitory stimuli for conditioned autonomic responses," *Biological Psychology*, vol. 22, pp. 37–57, Feb. 1986. cited on pages 60 and 62.

[288] C. Mayer, S. Sosnowski, K. Kühnlenz, and B. Radig, "Towards robotic facial mimicry: System development and evaluation," in *Proceedings of the 19th International Symposium in Robot and Human Interactive Communication*, pp. 198–203, IEEE, Sept. 2010. cited on pages 65 and 161.

[289] L.-O. Lundqvist and U. Dimberg, "Facial expressions are contagious," *Journal of Psychophysiology*, vol. 9, no. 3, pp. 203–211, 1995. cited on page 60.

[290] U. Hess, S. Blairy, and R. Kleck, "The Influence of Facial Emotion Displays, Gender, and Ethnicity on Judgments of Dominance and Affiliation," *Journal of Nonverbal Behavior*, vol. 24, no. 4, pp. 265–283, 2000. cited on page 63.

[291] M. Heerink, B. Krose, V. Evers, and B. Wielinga, "Measuring acceptance of an assistive social robot: a suggested toolkit," in *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 528–533, IEEE, Sept. 2009. cited on pages 58, 91, 92, 93, and 94.

[292] J. B. Bavelas, A. Black, C. R. Lemery, and J. Mullett, ""I show how you feel": Motor mimicry as a communicative act," *Journal of Personality and Social Psychology*, vol. 50, no. 2, pp. 322–329, 1986. cited on pages 60, 62, and 63.

[293] J. L. Lakin, T. L. Chartrand, and R. M. Arkin, "I Am Too Just Like You Nonconscious Mimicry as an Automatic Behavioral Response to Social Exclusion," *Psychological Science*, vol. 19, no. 8, pp. 816–822, 2008. cited on pages 59, 62, 63, and 64.

[294] P. Bourgeois and U. Hess, "The impact of social context on mimicry.," *Biological Psychology*, vol. 77, pp. 343–52, Mar. 2008. cited on pages 60 and 62.

[295] N. Bischof, "Untersuchungen zur Systemanalyse der sozialen Motivation I : Die Regulation der sozialen Distanz - Von der Feldtheorie zur Systemtheorie," *Zeitschrift für Psychologie*, vol. 201, no. 1, pp. 5–43, 1993. cited on page 75.

[296] N. P. Leander, T. L. Chartrand, and J. A. Bargh, "You give me the chills: embodied reactions to inappropriate amounts of behavioral mimicry.," *Psychological Science*, vol. 23, pp. 772–9, July 2012. cited on page 64.

[297] J. Bowlby, *Attachment and loss*. New York, NY, USA: Basic Books (AZ), 1982. cited on page 75.

[298] B. Fogg, "Persuasive computers: perspectives and research directions," in *Proceedings of the Conference on Human Factors in Computing Systems (CHI)* (M. E. Atwood, C.-M. Karat, A. M. Lund, J. Coutaz, and J. Karat, eds.), (Los Angeles, California, USA), pp. 225–232, ACM, 1998. cited on page 59.

[299] L. W. Barsalou, "Perceptual symbol systems," *Behavioral and Brain Sciences*, vol. 22, no. 04, pp. 577–660, 1999. cited on page 63.

[300] W. Flack, "Peripheral feedback effects of facial expressions, bodily postures, and vocal expressions on emotional feelings," *Cognition & Emotion*, vol. 20, pp. 177–195, Feb. 2006. cited on page 64.

[301] A. a. Grandey, G. M. Fisk, A. S. Mattila, K. J. Jansen, and L. a. Sideman, "Is "service with a smile" enough? Authenticity of positive displays during service encounters," *Organizational Behavior and Human Decision Processes*, vol. 96, pp. 38–55, Jan. 2005. cited on page 72.

[302] C. L. Breazeal, C. D. Kidd, A. L. Thomaz, G. Hoffman, and M. Berlin, "Effects of nonverbal communication on efficiency and robustness in human-robot teamwork," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 708–713, IEEE, 2005. cited on page 61.

[303] C. M. Cheng and T. L. Chartrand, "Self-monitoring without awareness: using mimicry as a nonconscious affiliation strategy.," *Journal of Personality and Social Psychology*, vol. 85, pp. 1170–9, Dec. 2003. cited on page 64.

[304] M. Lafrance and M. Broadbent, "Group Rapport: Posture Sharing as a Nonverbal Indicator," *Group & Organization Management*, vol. 1, pp. 328–333, Sept. 1976. cited on pages 63 and 64.

[305] R. Niewiadomski, M. Ochs, and C. Pelachaud, "Expressions of empathy in ECAs," in *Intelligent Virtual Agents* (C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, and D. Pelé, eds.), pp. 37–44, Springer Berlin / Heidelberg, 2008. cited on page 75.

[306] J. C. Karremans and T. Verwijmeren, "Mimicking attractive opposite-sex others: the role of romantic relationship status.," *Personality and Social Psychology Bulletin*, vol. 34, pp. 939–50, July 2008. cited on page 64.

[307] U. Hess and S. Blairy, "Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy," *International Journal of Psychophysiology*, vol. 40, pp. 129–141, Mar. 2001. cited on pages 62 and 64.

[308] K. Balci, "Xface," in *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI)*, (New York, New York, USA), p. 399, ACM Press, 2004. cited on page 85.

[309] A. N. Meltzoff and M. K. Moore, "Explaining Facial Imitation : A Theoretical Model," *Early Development and Parenting*, vol. 6, no. June, pp. 179–192, 1997. cited on pages 62, 63, 65, 66, and 160.

[310] L. Huang, L.-P. Morency, and J. Gratch, "Virtual Rapport 2.0," in *Intelligent Virtual Agents SE - 8* (H. Vilhjálmsson, S. Kopp, S. Marsella, and K. Thórisson, eds.), vol. 6895 of *Lecture Notes in Computer Science*, pp. 68–79, Springer Berlin Heidelberg, 2011. cited on pages 61 and 64.

[311] M. Ochs, R. Niewiadomski, and C. Pelachaud, "How a Virtual Agent Should Smile?," in *Intelligent Virtual Agents SE - 47* (J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, and A. Safonova, eds.), vol. 6356 of *Lecture Notes in Computer Science*, pp. 427–440, Springer Berlin Heidelberg, 2010. cited on page 74.

[312] K. Bousmalis, M. Mehu, and M. Pantic, "Spotting agreement and disagreement: A survey of nonverbal audiovisual cues and tools," in *Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops (ACII)*, (Amsterdam), pp. 1–9, IEEE, Sept. 2009. cited on page 46.

[313] N. Wang and J. Gratch, "Don't just stare at me!," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, (New York, NY, USA), pp. 1241–1250, ACM Press, 2010. cited on page 64.

[314] S. Thrun, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hahnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz, "MINERVA: a second-generation museum tour-guide robot," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, vol. 3, (Detroit, Michigan), pp. 1999–2005, IEEE, 1999. cited on page 59.

[315] S. Sosnowski, C. Mayer, K. Kühnlenz, and B. Radig, "Mirror my emotions ! Combining facial expression analysis and synthesis on a robot," in *Proceedings of the 36th Annual Convention of the Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB)*, 2010. cited on page 65.

[316] A. T. Dittmann and L. G. Llewellyn, "Relationship between vocalizations and head nods as listener responses.," *Journal of Personality and Social Psychology*, vol. 9, no. 1, pp. 79–84, 1968. cited on page 63.

[317] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly*, vol. 27, no. 3, pp. 425–478, 2003. cited on page 92.

[318] E. Keller and W. Tschacher, "Prosodic and Gestural Expression of Interactional Agreement," in *Verbal and Nonverbal Communication Behaviours SE - 8* (A. Esposito, M. Faundez-Zanuy, E. Keller, and M. Marinaro, eds.), vol. 4775 of *Lecture Notes in Computer Science*, pp. 85–98, Springer Berlin Heidelberg, 2007. cited on page 62.

[319] N. Bischof, "Untersuchungen zur Systemanalyse der Sozialen Motivation IV: Die Spielarten des Lächelns und das Problem der motivationalen Sollwertanpassung.," *Zeitschrift für Psychologie mit Zeitschrift für angewandte Psychologie*, vol. 204, no. 1, pp. 1–40, 1996. cited on pages 75, 79, 82, and 167.

[320] D. Keltner and C. Anderson, "Saving Face for Darwin : The Functions and Uses of Embarrassment," *Current Directions in Psychological Science*, vol. 9, no. 6, pp. 187–192, 2000. cited on page 47.

[321] B. M. Muir, "Trust between humans and machines, and the design of decision aids," *International Journal of Man-Machine Studies*, vol. 27, pp. 527–539, Nov. 1987. cited on pages 64 and 95.

[322] B. J. Fogg, "Computers as Persuasive Social Actors," in *Persuasive Technology*, ch. 5, pp. 89–120, San Francisco, CA: Morgan Kaufmann Publishers, 2003. cited on page 59.

[323] K. U. Likowski, A. Mühlberger, B. Seibt, P. Pauli, and P. Weyers, "Modulation of facial mimicry by attitudes," *Journal of Experimental Social Psychology*, vol. 44, pp. 1065–1072, July 2008. cited on pages 60 and 62.

[324] H. Gubler and N. Bischof, "A Systems Theory Perspective," in *Infant Development: Perspectives from German-Speaking Countries* (M. E. Lamb and H. Keller, eds.), pp. 35–66, Hillsdale, New Jersey Hove: LAWRENCE ERLBAUM ASSOCIATES, 1991. cited on page 75.

[325] P. M. Niedenthal, "Embodying emotion.," *Science*, vol. 316, pp. 1002–5, May 2007. cited on page 63.

[326] A. Bobick and J. Davis, "The recognition of human movement using temporal templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 257–267, Mar. 2001. cited on page 62.

[327] U. Dimberg, "Facial reactions, autonomic activity and experienced emotion: A three component model of emotional conditioning," *Biological Psychology*, vol. 24, pp. 105–122, Apr. 1987. cited on pages 60 and 62.

[328] M. Wimmer, F. Stulp, S. Pietzsch, and B. Radig, "Learning local objective functions for robust face model fitting.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, pp. 1357–70, Aug. 2008. cited on page 67.

[329] E. A. R. Tanguy, *Emotions: the Art of Communication Applied to Virtual Actors*. Dissertation, University of Bath, 2006. cited on page 74.

[330] C. K. Hsee, E. Hatfield, J. G. Carlson, and C. Chemtob, "The effect of power on susceptibility to emotional contagion," *Cognition & Emotion*, vol. 4, pp. 327–340, Oct. 1990. cited on page 62.

[331] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots," *International Journal of Social Robotics*, vol. 1, pp. 71–81, Nov. 2008. cited on pages 55, 58, 91, 92, and 93.

[332] M. Tscherepanow, M. Hillebrand, F. Hegel, B. Wrede, and F. Kummert, "Direct imitation of human facial expressions by a user-interface robot," in *Proceedings of the 9th IEEE/RAS International Conference on Humanoid Robots*, pp. 154–160, IEEE, Dec. 2009. cited on pages 61 and 75.

[333] A. Paiva, J. a. Dias, D. Sobral, R. Aylett, S. Woods, L. Hall, and C. Zoll, "Learning By Feeling: Evoking Empathy With Synthetic Characters," *Applied Artificial Intelligence*, vol. 19, pp. 235–266, Mar. 2005. cited on page 75.

[334] C. Bartneck and M. Okada, "Robotic user interfaces," in *Proceedings of the Human and Computer Conference*, pp. 130–140, 2001. cited on page 16.

[335] W. B. Mendes, J. Blascovich, S. B. Hunter, B. Lickel, and J. T. Jost, "Threatened by the unexpected: physiological responses during social interactions with expectancy-violating partners.," *Journal of Personality and Social Psychology*, vol. 92, pp. 698–716, Apr. 2007. cited on page 64.

C.8 Chapter 4

[336] D. Vogel and H. Bleckmann, "Behavioral discrimination of water motions caused by moving objects," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 186, pp. 1107–1117, Feb. 2001. cited on page 98.

[337] S. Coombs, C. B. Braun, and B. Donovan, "The orienting response of Lake Michigan mottled sculpin is mediated by canal neuromasts," *Journal of Experimental Biology*, vol. 204, pp. 337–348, Jan. 2001. cited on page 98.

[338] N. Izadi, M. J. de Boer, J. W. Berenschot, and G. J. M. Krijnen, "Fabrication of superficial neuromast inspired capacitive flow sensors," *Journal of Micromechanics and Microengineering*, vol. 20, no. 8, p. 85041, 2010. cited on page 102.

[339] R. Bannasch, "Bionics," in *Technology Guide* (H.-J. Bullinger, ed.), pp. 178–183, Springer Berlin Heidelberg, 2009. cited on page 98.

[340] S. P. Windsor, S. E. Norris, S. M. Cameron, G. D. Mallinson, and J. C. Montgomery, "The flow fields involved in hydrodynamic imaging by blind Mexican cave fish (Astyanax fasciatus). Part II: gliding parallel to a wall," *Journal of Experimental Biology*, vol. 213, pp. 3832–3842, Nov. 2010. cited on page 123.

[341] R. A. Handelsman and J. B. Keller, "Axially symmetric potential flow around a slender body," *Journal of Fluid Mechanics*, vol. 28, pp. 131–147, Apr. 1967. cited on page 114.

[342] T. Papanastasiou, G. Georgiou, and A. N. Alexandrou, *Viscous Fluid Flow*. CRC Press, 1st ed., 1999. cited on page 105.

[343] J.-M. P. Franosch, S. Sosnowski, N. K. Chami, K. Kühnlenz, S. Hirche, and J. L. van Hemmen, "Biomimetic lateral-line system for underwater vehicles," in *Proceedings of the IEEE Sensors*, pp. 2212–2217, IEEE, 2010. cited on pages 103, 107, 118, and 171.

[344] Y. Yang, A. Klein, H. Bleckmann, and C. Liu, "Artificial lateral line canal for hydrodynamic detection," *Applied Physics Letters*, vol. 99, p. 023701, July 2011. cited on page 102.

[345] W. Nachtigall, *Bionik als Wissenschaft: Erkennen - Abstrahieren - Umsetzen (German Edition)*. Springer, 1st ed., 2010. cited on pages 2, 98, and 102.

[346] S. Dykgraaf, "Untersuchungen über die Funktion der Seitenorgane an Fischen," *Zeitschrift für Vergleichende Physiologie*, vol. 20, no. 1-2, pp. 162–214, 1933. cited on page 98.

[347] S. Urban, "Rekonstruktion stationärer Objekte durch den blinden Höhlenfisch Mexican tetra," Master's thesis, Technische Universität München, 2011. cited on pages 108, 110, 122, 124, 125, and 141.

[348] D. Lenz, "Sensory Model for Obstacle Detection with an Artificial Lateral Line System," diplomathesis, Technische Universität München, 2011. cited on pages 129, 130, 175, 176, and 177.

[349] Y. Yang, N. Nguyen, N. Chen, M. Lockwood, C. Tucker, H. Hu, H. Bleckmann, C. Liu, and D. L. Jones, "Artificial lateral line with biomimetic neuromasts to emulate fish sensing.," *Bioinspiration & Biomimetics*, vol. 5, no. 1, p. 16001, 2010. cited on page 102.

[350] N. Nguyen, D. L. Jones, Y. Yang, and C. Liu, "Flow Vision for Autonomous Underwater Vehicles via an Artificial Lateral Line," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, p. 806406, Jan. 2011. cited on page 102.

[351] E. J. Denton and J. A. B. Gray, "Some Observations on the Forces Acting on Neuromasts in Fish Lateral Line Canals," in *The Mechanosensory Lateral Line: Neurobiology and Evolution* (S. Coombs, P. Görner, and H. Münz, eds.), pp. 229–246, New York: Springer, 1989. cited on page 100.

[352] S. P. Windsor, S. E. Norris, S. M. Cameron, G. D. Mallinson, and J. C. Montgomery, "The flow fields involved in hydrodynamic imaging by blind Mexican cave fish (Astyanax fasciatus). Part I: open water and heading towards a wall," *Journal of Experimental Biology*, vol. 213, pp. 3819–3831, Nov. 2010. cited on page 123.

[353] S. Coombs and R. A. Conley, "Dipole source localization by the mottled sculpin II. The role of lateral line excitation patterns," *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, vol. 180, pp. 401–415, Mar. 1997. cited on page 102.

[354] H. K. Versteeg and W. Malalasekera, *An Introduction to Computational Fluid Dynamics: the Finite Volume Method*. Prentice Hall, 1st ed., 2007. cited on page 123.

[355] E.-S. Hassan, "On the discrimination of spatial intervals by the blind cave fish (Anoptichthys jordani)," *Journal of Comparative Physiology A: Neuroethology Sensory Neural And Behavioral Physiology*, vol. 159, pp. 701–710, Sept. 1986. cited on page 98.

[356] V. I. Fernandez, S. M. Hou, F. S. Hover, J. H. Lang, and M. S. Triantafyllou, "Lateral-Line Inspired MEMS-Array Pressure Sensing for Passive Underwater Navigation," tech. rep., MIT Sea Grant, 2007. cited on page 123.

[357] S. P. Windsor, D. Tan, and J. C. Montgomery, "Swimming kinematics and hydrodynamic imaging in the blind Mexican cave fish (Astyanax fasciatus)," *Journal of Experimental Biology*, vol. 211, pp. 2950–2959, Sept. 2008. cited on pages 98, 108, 110, and 141.

[358] C. Liu, "Micromachined biomimetic artificial haircell sensors.," *Bioinspiration & Biomimetics*, vol. 2, no. 4, pp. S162–S169, 2007. cited on page 102.

[359] H. Zell, "Astyanax Mexicanus." Wikimedia Commons, May 2011. cited on page 100.

[360] S. Künzel, H. Bleckmann, and J. Mogdans, "Responses of brainstem lateral line units to different stimulus source locations and vibration directions.," *Journal of Comparative Physiology A: Neuroethology Sensory Neural And Behavioral Physiology*, vol. 197, pp. 773–87, July 2011. cited on page 102.

[361] W. Felix, "Strömungsmessung mit Thermistoren," *Naunyn-Schmiedebergs Archiv für experimentelle Pathologie und Pharmakologie*, vol. 244, no. 3, pp. 254–269, 1962. cited on page 103.

[362] B. L. Partridge and T. J. Pitcher, "The sensory basis of fish schools: Relative roles of lateral line and vision," *Journal Of Comparative Physiology A: Neuroethology Sensory Neural And Behavioral Physiology*, vol. 135, no. 4, pp. 315–325, 1980. cited on page 98.

[363] S. Pandya, Y. Yang, C. Liu, and D. L. Jones, "Biomimetic Imaging of Flow Phenomena," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 933–936, IEEE, 2007. cited on page 102.

[364] I. C. Itsweire and K. N. Helland, "A high-performance low-cost constant-temperature hot-wire anemometer," *Journal of Physics E-Scientific Instruments*, vol. 16, no. 6, pp. 549–553, 1983. cited on page 103.

[365] J. H. Ferziger and M. Peric, *Computational Methods for Fluid Dynamics*. Berlin: Springer, 3rd ed., 1996. cited on page 123.

[366] J. L. van Hemmen and A. B. Schwartz, "Population vector code: a geometric universal as actuator," *Biological Cybernetics*, vol. 98, pp. 509–518, 2008. cited on page 116.

[367] J. Chen, J. Engel, N. Chen, S. D. Pandya, S. Coombs, and C. Liu, "Artificial Lateral Line And Hydrodynamic Object Tracking," in *Proceedings of the 19th IEEE International Conference on Micro Electro Mechanical Systems*, pp. 694–697, IEEE, 2006. cited on page 102.

[368] P. Göner, "Untersuchungen zur Morphologie und Elektrophysiologie des Seitenlinienorgans vom Krallenfrosch (Xenopus laevis Daudin)," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 47, no. 3, pp. 316–338, 1963. cited on page 101.

[369] W. J. Van Trump and M. J. McHenry, "The morphology and mechanical sensitivity of lateral line receptors in zebrafish larvae (Danio rerio).," *The Journal of Experimental Biology*, vol. 211, pp. 2105–15, July 2008. cited on page 98.

[370] E.-S. Hassan, "Mathematical description of the stimuli to the lateral line system of fish derived from a three-dimensional flow field analysis. II. The case of gliding alongside or above a plane surface," *Biological Cybernetics*, vol. 66, pp. 443–452, Mar. 1992. cited on pages 114, 117, and 118.

[371] J. Engelmann, W. Hanke, and H. Bleckmann, "Lateral line reception in still- and running water.," *Journal of Comparative Physiology A: Neuroethology Sensory Neural And Behavioral Physiology*, vol. 188, pp. 513–26, Aug. 2002. cited on page 102.

[372] G. Meyer, A. Klein, J. Mogdans, and H. Bleckmann, "Toral lateral line units of goldfish, Carassius auratus, are sensitive to the position and vibration direction of a vibrating sphere.," *Journal of Comparative Physiology A: Neuroethology, sensory, neural, and behavioral physiology*, June 2012. cited on page 102.

[373] H. Lamb, *Lehrbuch der Hydrodynamik*. Leipzig, Berlin: Teubner, 1907. cited on page 105.

[374] S. Peleshanko, M. D. Julian, M. Ornatska, M. E. McConney, M. C. LeMieux, N. Chen, C. Tucker, Y. Yang, C. Liu, J. A. C. Humphrey, and V. V. Tsukruk, "Hydrogel-Encapsulated Microfabricated Haircells Mimicking Fish Cupula Neuromast," *Advanced Materials*, vol. 19, pp. 2903–2909, Oct. 2007. cited on page 102.

[375] S. Coombs and P. Patton, "Lateral line stimulation patterns and prey orienting behavior in the Lake Michigan mottled sculpin (Cottus bairdi).," *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology*, vol. 195, pp. 279–97, Mar. 2009. cited on page 98.

[376] J. Mogdans and H. Bleckmann, "Coping with flow: behavior, neurophysiology and modeling of the fish lateral line system.," *Biological Cybernetics*, vol. 106, pp. 627–42, Dec. 2012. cited on page 98.

[377] F. Keplinger and F. Kohl, "Direct 1/f Noise Measurement on Micro-Thermistors," *Sensors, Actuators and Micro-Technology*, pp. 151–152, 2000. cited on page 103.

[378] J. D. Anderson, *Computational Fluid Dynamics*. New York, New York, USA: McGraw-Hill Book Company, Inc., 1995. cited on page 123.

[379] I. Vasilescu, C. Detweiler, M. Doniec, D. Gurdan, S. Sosnowski, J. Stumpf, and D. Rus, "AMOUR V: A Hovering Energy Efficient Underwater Robot Capable of Dynamic Payloads," *The International Journal of Robotics Research*, vol. 29, pp. 547–570, Jan. 2010. cited on pages 108 and 174.

[380] S. Sosnowski, J.-M. P. Franosch, L. Zhang, Y. Nie, S. Hirche, and J. L. van Hemmen, "Simulation of the Underwater Vehicle "Snookie": Navigating like a Fish," in *Proceedings of the 1st International Conference on Applied Bionics and Biomechanics (ICABB)*, (Venice, Italy), 2010. cited on pages 114 and 119.

[381] V. I. Fernandez, A. Maertens, F. M. Yaul, J. Dahl, J. H. Lang, and M. S. Triantafyllou, "Lateral-Line-Inspired Sensor Arrays for Navigation and Object Identification," *Marine Technology Society Journal*, vol. 45, no. 4, pp. 130–146, 2011. cited on page 123.

[382] F. Li, W. Liu, C. Stefanini, X. Fu, and P. Dario, "A Novel Bioinspired PVDF Micro/Nano Hair Receptor for a Robot Sensing System," *Sensors*, vol. 10, no. 1, pp. 994–1011, 2010. cited on page 102.

[383] R. Weissert and C. von Campenhausen, "Discrimination between stationary objects by the blind cave fish Anoptichthys jordani (Characidae)," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 143, no. 3, pp. 375–381, 1981. cited on page 98.

[384] Y. Yang, J. Chen, J. Engel, S. Pandya, N. Chen, C. Tucker, S. Coombs, D. L. Jones, and C. Liu, "Distant touch hydrodynamic imaging with an artificial lateral line," *Proceedings of the National Academy of Sciences*, vol. 103, no. 50, pp. 18891–18895, 2006. cited on page 102.

[385] B. P. Chagnaud, H. Bleckmann, and J. Engelmann, "Neural responses of goldfish lateral line afferents to vortex motions.," *The Journal of Experimental Biology*, vol. 209, pp. 327–42, Jan. 2006. cited on page 102.

[386] M. Motamed and J. Yan, "A Review of Biological, Biomimetic and Miniature Force Sensing for Microflight," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2630–2637, IEEE, 2005. cited on page 102.

[387] B. Ćurçić Blake and S. van Netten, "Rapid responses of the cupula in the lateral line of ruffe (Gymnocephalus cernuus)," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 191, no. 4, pp. 393–401, 2005. cited on pages 101, 102, and 111.

[388] H. Strickert, *Hitzdraht-und Hitzfilmanemometrie*. Berlin: VEB Verlag Technik, 1974. cited on page 103.

[389] E.-S. Hassan, "Hydrodynamic Imaging of the Surroundings by the Lateral Line of the Blind Cave Fish Anoptichthys jordani," in *The Mechanosensory Lateral Line*, pp. 217–227, 1989. cited on page 98.

[390] A. E. Perry, *Hot-wire anemometry*. Oxford: Clarendon Pr., 1982. cited on page 103.

[391] S. Coombs and R. A. Conley, "Dipole source localization by mottled sculpin. I. Approach strategies," *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, vol. 180, pp. 387–399, Mar. 1997. cited on page 98.

[392] S. Coombs, R. R. Fay, and J. Janssen, "Hot-film anemometry for measuring lateral line stimuli.," *The Journal of the Acoustical Society of America*, vol. 85, pp. 2185–93, May 1989. cited on pages 102 and 103.

[393] Y. Liu, *Fast Multipole Boundary Element Method: Theory and Applications in Engineering*. Cambridge University Press, 1 ed., 2009. cited on page 123.

[394] V. I. Fernandez, S. M. Hou, and F. S. Hover, "Development and application of distributed MEMS pressure sensor array for AUV object avoidance," tech. rep., MIT Sea Grant, 2009. cited on page 123.

[395] M. Nora, S. Sosnowski, K. Kühnlenz, S. Hirche, Y. Nie, J.-M. P. Franosch, and J. L. van Hemmen, "Design of a Lateral-Line Sensor for an Autonomous Underwater Vehicle," in *Proceedings of the 8th IFAC International Conference on Maneuvering and Control of Marine Craft (MCMC)* (D. Decio, ed.), (Sao Paulo,Brazil), pp. 292–297, Sept. 2009. cited on page 115.

[396] S. Coombs and S. M. van Netten, "The Hydrodynamics and Structural Mechanics of the Lateral Line System," *Fish Physiology*, vol. 23, pp. 103–139, 2005. cited on page 102.

[397] E.-S. Hassan, "Mathematical description of the stimuli to the lateral line system of fish derived from a three-dimensional flow field analysis. I. The cases of moving in open water and of gliding towards a plane surface," *Biological Cybernetics*, vol. 66, pp. 453–461, Mar. 1992. cited on pages 114, 117, and 118.

[398] N. A. Campbell, *Biology, 6th Edition*. Benjamin Cummings, 6th ed., 2002. cited on page 101.

[399] T. Teyke, "Collision with and avoidance of obstacles by blind cave fish Anoptichthys jordani (Characidae)," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 157, no. 6, pp. 837–843, 1985. cited on pages 98 and 140.

[400] U. Eser, *Thermisches Anemometer mit Kugelsonde zur Bestimmung kleiner Geschwindigkeitsvektoren*. PhD thesis, Universität Essen, Essen, 1990. cited on page 103.

[401] J. Goulet, J. Engelmann, B. P. Chagnaud, J.-M. P. Franosch, M. D. Suttner, and J. L. van Hemmen, "Object localization through the lateral-line system of fish: Theory and experiment," *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, vol. 194, no. 1, pp. 1–17, 2008. cited on pages 100 and 102.

[402] J. Engelmann and H. Bleckmann, "Coding of lateral line stimuli in the goldfish midbrain in still and running water.," *Zoology*, vol. 107, pp. 135–51, Jan. 2004. cited on page 102.

[403] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT Press, 2005. cited on pages 135, 137, and 138.

[404] A. Qualtieri, F. Rizzi, M. T. Todaro, A. Passaseo, R. Cingolani, and M. De Vittorio, "Stress-driven AlN cantilever-based flow sensor for fish lateral line system," *Microelectronic Engineering*, vol. 88, pp. 2376–2378, Aug. 2011. cited on page 102.

[405] Burr-Brown Corporation, *INA103 Low Noise, Low Distortion INSTRUMENTATION AMPLIFIER*, 1998. cited on page 104.

[406] A. Sichert, R. Bamler, and J. L. van Hemmen, "Hydrodynamic Object Recognition: When Multipoles Count," *Physical Review Letters*, vol. 102, Feb. 2009. cited on pages 99 and 123.

[407] L. Prandtl and A. Betz, *VIER ABHANDLUNGEN ZUR HYDRODYNAMIK UND AERODYNAMIK*. Kaiser Wilhelm - Institut für Strömungsforschung, 1927. cited on page 104.

[408] H. Oertel and K. Mayes, eds., *Prandtl's Essentials of Fluid Mechanics (Applied Mathematical Sciences)*. Springer, 2nd ed., 2004. cited on page 105.

[409] H. Bleckmann and R. Zelick, "Lateral line system of fish.," *Integrative Zoology*, vol. 4, no. 1, pp. 13–25, 2009. cited on page 102.

[410] J.-M. Franosch, H. Hagedorn, J. Goulet, J. Engelmann, and J. L. van Hemmen, "Wake Tracking and the Detection of Vortex Rings by the Canal Lateral Line of Fish," *Physical Review Letters*, vol. 103, Aug. 2009. cited on page 150.

[411] S. Coombs, M. Hastings, and J. Finneran, "Modeling and measuring lateral line excitation patterns to changing dipole source locations," *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, vol. 178, Mar. 1996. cited on page 102.

[412] H. Lamb, *Hydrodynamics*. Cambridge: Cambridge University Press, 6 ed., 1932. cited on pages 105 and 115.

[413] R. A. Conley and S. Coombs, "Dipole source localization by mottled sculpin. III. Orientation after site-specific, unilateral denervation of the lateral line system," *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, vol. 183, pp. 335–344, Sept. 1998. cited on page 98.

[414] S. Emsmann and A. Lehmann, "Entwicklung eines Thermistoranemometers zur Messung instationärer Wassergeschwindigkeiten," *Fortschrittsberichte der VDI Zeitschriften*, vol. 18, no. 8, 1975. cited on page 103.

[415] H. Bleckmann, *Reception of Hydrodynamic Stimuli in Aquatic and Semiaquatic Animals*. Stuttgart: Fischer, 1994. cited on page 100.

[416] H. Bleckmann, T. Breithaupt, R. Blickhan, and J. Tautz, "The time course and frequency content of hydrodynamic events caused by moving fish, frogs, and crustaceans," *Journal of Comparative Physiology A: Neuroethology Sensory Neural And Behavioral Physiology*, vol. 168, June 1991. cited on page 99.

[417] H. Lamb, *Hydrodynamics*. Dover publications, 6th ed., 1945. cited on pages 124, 125, and 131.

[418] E.-S. Hassan, "Mathematical description of the stimuli to the lateral line system of fish, derived from a three-dimensional flow field analysis. III. The case of an oscillating sphere near the fish," *Biological Cybernetics*, vol. 69, pp. 525–538, Sept. 1993. cited on page 114.

[419] H. H. Bruun, *Hot wire anemometry*. Oxford: Oxford Univ. Press, 1996. cited on page 103.

[420] S. Kröther, J. Mogdans, and H. Bleckmann, "Brainstem lateral line responses to sinusoidal wave stimuli in still and running water.," *The Journal of Experimental Biology*, vol. 205, pp. 1471–84, May 2002. cited on page 102.

[421] T. Burt de Perera, "Spatial parameters encoded in the spatial map of the blind Mexican cave fish, Astyanax fasciatus," *Animal Behaviour*, vol. 68, pp. 291–295, Aug. 2004. cited on page 140.

[422] P. Patton, S. P. Windsor, and S. Coombs, "Active wall following by Mexican blind cavefish (Astyanax mexicanus)," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 196, no. 11, pp. 853–867, 2010. cited on page 140.

[423] S. Dijkgraaf, "THE FUNCTIONING AND SIGNIFICANCE OF THE LATERAL-LINE ORGANS," *Biological Reviews*, vol. 38, pp. 51–105, Feb. 1963. cited on page 98.

[424] C. Campenhausen, I. Riess, and R. Weissert, "Detection of stationary objects by the blind Cave Fish Anoptichthys jordani (Characidae)," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 143, no. 3, pp. 369–374, 1981. cited on page 98.

[425] H. Schlichting and K. Gersten, *Boundary-Layer Theory*. Springer, 2006. cited on pages 104, 105, and 106.

[426] R. L. Panton, *Incompressible Flow*. Wiley, 3rd ed., 2005. cited on pages 105, 114, and 117.

[427] R. Bouffanais, G. D. Weymouth, and D. K. P. Yue, "Hydrodynamic object recognition using pressure sensing," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 467, pp. 19–38, June 2011. cited on pages 99 and 123.

[428] T.-Y. Hsieh, S.-W. Huang, L.-J. Mu, E. Chen, and J. Guo, "Artificial lateral line design for robotic fish," in *Proceedings of the IEEE Symposium on Underwater Technology (UT) and Workshop on Scientific Use of Submarine Cables and Related Technologies (SSC)*, pp. 1–6, IEEE, Apr. 2011. cited on pages 102 and 114.

[429] T. von Kármán, *Aerodynamics*. McGraw-Hill, 1963. cited on page 106.

[430] G. B. Middlebrook and E. L. Piret, "Hot wire anemometry-solution of some difficulties in measurement of low water velocities," *Industrial and Engineering Chemistry*, vol. 42, no. 8, 1950. cited on pages 102, 103, and 106.

[431] C. G. Lomas, *Fundamentals of hot wire anemometry*. Cambridge: Cambridge Univ. Pr., 1986. cited on page 103.

[432] O. Akanyeti, R. Venturelli, F. Visentin, L. Chambers, W. M. Megill, and P. Fiorini, "What information do Kármán streets offer to flow sensing?," *Bioinspiration & Biomimetics*, vol. 6, p. 036001, Sept. 2011. cited on page 150.

[433] E.-S. Hassan, "Mathematical analysis of the stimulus for the lateral line organ," *Biological Cybernetics*, vol. 52, no. 1, pp. 23–36, 1985. cited on pages 114 and 117.

[434] M. J. McHenry, J. A. Strother, and S. M. van Netten, "Mechanical filtering by the boundary layer and fluid-structure interaction in the superficial neuromast of the fish lateral line system.," *Journal Of Comparative Physiology A: Neuroethology Sensory Neural And Behavioral Physiology*, vol. 194, pp. 795–810, Sept. 2008. cited on page 102.

[435] D. T. T. Plachta, "A hydrodynamic topographic map in the midbrain of goldfish Carassius auratus," *Journal of Experimental Biology*, vol. 206, pp. 3479–3486, Oct. 2003. cited on page 102.

[436] H. Bleckmann, "Peripheral and central processing of lateral line information," *Journal Of Comparative Physiology A: Neuroethology Sensory Neural And Behavioral Physiology*, vol. 194, no. 2, pp. 145–58, 2008. cited on page 102.

[437] S. Sharma, S. Coombs, P. Patton, and T. Burt de Perera, "The function of wall-following behaviors in the Mexican blind cavefish and a sighted relative, the Mexican tetra (Astyanax).," *Journal Of Comparative Physiology A: Neuroethology Sensory Neural And Behavioral Physiology*, vol. 195, pp. 225–40, Mar. 2009. cited on pages 140 and 141.

[438] S. van Netten, "Hydrodynamic detection by cupulae in a lateral line canal: functional relations between physics and physiology," *Biological Cybernetics*, vol. 94, no. 1, pp. 67–85, 2006. cited on pages 100, 101, 102, and 111.

C.9 Chapter 5

[439] B. Kühnlenz, S. Sosnowski, M. Buß, D. Wollherr, K. Kühnlenz, and M. Buss, "Increasing Helpfulness towards a Robot by Emotional Adaption to the User," *International Journal of Social Robotics*, Mar. 2013. cited on pages 31, 32, and 149.

[440] N. Mirnig, B. Gonsior, S. Sosnowski, C. Landsiedel, D. Wollherr, A. Weiss, and M. Tscheligi, "Feedback guidelines for multimodal human-robot interaction: How should a robot give feedback when asking for directions?," in *Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 533–538, IEEE, Sept. 2012. cited on page 149.

[441] P. Ekman, W. V. Friesen, M. O'Sullivan, A. Chan, I. Diacoyanni-Tarlatzis, K. Heider, R. Krause, W. A. LeCompte, T. Pitcairn, and P. E. Ricci-Bitti, "Universals and cultural differences in the judgments of facial expressions of emotion.," *Journal of Personality and Social Psychology*, vol. 53, pp. 712–7, Oct. 1987. cited on pages 13, 32, 34, 72, 73, and 149.