Recognizing Behavior and Strategy for Enjoyment and Affection in Touch-based Play with a Humanoid Robot

A dissertation submitted to

THE GRADUATE SCHOOL OF ENGINEERING SCIENCE

OSAKA UNIVERSITY

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN ENGINEERING

By

Cooney, Martin Daniel

March 2014

Abstract

Well-being—related to happiness, life satisfaction, and quality of life—is being increasingly afforded attention as an indicator of the success of human social systems, in part due to its being linked with healthier and longer life. We believe that (1) “social well-being”, one aspect of well-being due to social interactions, can be facilitated via playing with a companion robot; (2) this long-term phenomenon can be affected by providing short-term perceptions of enjoyment and affection; and (3) to evoke such perceptions a robot should recognize human behavior and employ an appropriate behavior strategy. In the current dissertation, we focused on the fundamental case of a dyadic touch-based interaction with a humanoid robot, because touch serves an important role in playing for enjoyment and communicating affection, and the humanoid form is a familiar interface which offers potential for rich interactions. The foremost challenge faced was that human social perceptions toward a robot—of how to behave toward a robot and how to interpret a robot’s behavior—are highly complex and not well understood. Our approach involved observing interactions with robotic prototypes specifically designed to address key issues and querying participants to gain insight into how interactions were perceived. We conducted our investigation in four steps. First we built a system to recognize how people play for enjoyment by moving the body of a small held humanoid robot. Second, we investigated what kind of behavior strategy would allow a robot to use this recognition capability to provide enjoyment. Third, we expanded this initial scenario to also address recognizing how people use touch to express affection (liking or disliking) toward a robot. Fourth, we proposed an affectionate strategy which would allow a robot to approach people to initiate interactions. The knowledge gained provides a fundamental indication of how a humanoid companion robot can recognize a person’s behavior motivated by enjoyment or affection and use such recognition results to structure its behavior, toward facilitating perceived social well-being and improving lives of interacting persons.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

"They're my friends. I made them."

J.F. Sebastian, in the movie Bladerunner, 1982 describing some robots.

This dissertation reports on our work on designing interactions with a humanoid companion robot conducted in the belief that social well-being in interacting individuals can be facilitated [1-6]. We begin by discussing the wider context of human-robot interaction (HRI).

People expect robots to perform various useful tasks. These include both troublesome tasks which may be dirty, dull, or dangerous, as well as difficult tasks requiring excellent memory and calculation skills [7]. Some examples of well-known robots which perform tasks are Roomba which cleans rooms, TALON which is used in security applications, and AIBO which provides companionship [8]. Although these robots differ greatly, one important similarity is that people seek to interact with them almost as if the robots were human; for example, naming, talking to, presenting awards to, and even holding funeral services for robots such as TALON [9]. This suggests the potential usefulness of a robot's interactive capabilities.

Some ambiguity exists in what is required to develop good interactions. A number of writers and researchers have set forth opinions. For example, Asimov proposed the importance of safety in interactions [10]. Picard predicted that emotional computing would be desirable [11]. Scheutz *et al*. hypothesized that naturalness was an important criterion [12]. Thus, many opinions were set forth, but it was not clear what the fundamental goal might be. For example, safety is important but a form of minimal requirement, in the sense that not interacting with a robot at all would be safest; something more is needed. Similarly, it could possibly be useful for a robot to express emotions such as disgust or surprise, but this may also not be our deepest concern. Furthermore, we may not always want robots to behave in a natural or human-like manner; for example, seeing a robot pull out a pen and paper to perform difficult calculations, or express uncertainty with its answer, could be disconcerting during a financial transaction.

1.1 Our Focus: Definitions and Motivation

We think the deepest, most important question in HRI is how to provide social well-being; i.e., how to make people feel good via interactions. We define our terminology formally in several steps as follows. Well-being in general is “a good or satisfactory condition of existence; a state characterized by health, happiness, and prosperity” [13]. This term has been used interchangeably with “happiness” and “life satisfaction” in the past [14]. Well-being is typically considered to be a long-term quality influenced not only by predisposition but also by recent and past events over a period of time [15]. As well, we are concerned not with an objective status, but rather a subjective feeling [16]. “Social well-being” here means well-being supported by social interactions [15]. To bring together all of the above points, we define social well-being as “a subjectively perceived long-term state of existence characterized by happiness, life satisfaction, health, and other prosperity afforded by social interactions”. This, and some other definitions of important terms which appear in this work, are summarized in Table 1.

The importance of social well-being has been suggested by a number of famous and influential thinkers including Aristotle, John Stuart Mill, and Seligman, although the idea of “being well” has been interpreted as either “feeling well” or “behaving well”. In Ancient Greece, Aristippus seems to have first advocated a hedonic notion of well-being (that feeling well is important), hypothesizing that the goal of life was to seek pleasure. Alternatively, Aristotle, one of the “greatest philosophers of all time”, felt that eudaimonic well-being (behaving well) was the ultimate goal of life [17]. Augustine affirmed that “all men agree in desiring . . . happiness” [18], which Aquinas stated could be attained in an imperfect form in this life [19]. The utalitarians including Mill, “the most influential English-speaking philosopher of the nineteenth century” [20], taught that maximizing

Table 1. Some important definitions for providing social well-being via robotic interactions

Term	**Our definition**
Social well-being	A subjectively perceived long-term state of existence characterized by happiness, life satisfaction, health, and other prosperity which is afforded by social interactions, leads to longer, healthier life, and may result from perceived enjoyment and affection
Enjoyment	A feeling of joy, pleasure, satisfaction, and gratification caused by some entertaining activity or stimulus
Affection	Fond attachment, devotion, love, liking, positive feeling, and gentleness directed toward someone
Play	An enjoyable, free, absorbing activity performed by people of all ages often involving touching and moving an artifact

hedonic well-being over the greatest number of individuals was of the utmost importance. Maslow, Rogers, Erikson, and Allport, considered respectively the 10th, 6th, 12th, and 11th most eminent psychologists of the 20th Century [21], also postulated ideas with regard to well-being. Recently, Helliwell and Putnam described well-being as possibly “the ultimate ‘dependent variable’ in social science” [14]. Seligman, “the 13th most frequently cited psychologist in introductory psychology textbooks throughout the century, as well as the 31st most eminent overall” [21] described well-being as making life more fulfilling [22], and played a role in inspiring Cameron, the Prime Minister of the U.K., to also consider this important criterion for assessing national prosperity [17]: Cameron professed that “improving our society's sense of well-being is . . . the central political challenge of our times” [24].

The importance of well-being is also suggested by some objective data. Davidson and Begley found that well-being leads to improved health [25]. Also, Frey observed that well-being relates to extended lifespan [26]. Thus the importance of well-being has been described by both subjective and objective sources.

One problem is that how social well-being can be provided and adequately measured is unclear. Well-being is perceived over the long term as a complex result of numerous events and causes [15], making it difficult to evaluate how it is affected by short experiments in a laboratory, and directly asking respondents has been described as unwieldy [27]. For example, asking about a participant’s well-being after an interaction with a robot could yield answers which are more influenced by other recent experiences and would not clarify how the interaction affected them. Likewise, an attempt to start with a long-term study could also fail or yield few results over much time, because research on this topic in robotics is relatively new and our knowledge is incomplete, as will be described in Chapter 1.3. This suggested the usefulness of finding some likely determinants of well-being which could be affected and measured over the short term.

For the current work, we hypothesized that two such factors, "enjoyment" and "affection", would be important for experiencing social well-being. First we define these terms, and later describe why we believe these two factors are important. Enjoyment has been described in general as a pleasant feeling caused by some entertaining activity or stimulus [28], characterized by joy, gratification, and satisfaction [29]; as such it relates to usability (ISO 9241-11). We use this word in its broad sense, without distinguishing pleasure from gratification [30]. Factors supporting enjoyment include perception of sufficient challenge, reward, control over what occurs, and unambiguous objectives [31-32]. Enjoyable activities could include playing, drinking a favorite beverage, or taking a warm bath. Unenjoyable activities could include performing a boring or stressful task, eating a disliked food, or being soaked by an icy rain shower.

Affection connotes "fond attachment, devotion, or love" [33]; it relates to positive attitude, gentleness, niceness, and regard toward someone [34-37]. We caution that "affection" and "affect" are two different words, the latter being used as a synonym for emotion in general. Affectionate experiences could include being patted on the back, receiving a smile, or being congratulated. Unaffectionate experiences could include being shoved, observing someone scowling at one's self, or being insulted.

We believe that enjoyment and affection are closely related to social well-being, based on six sources of evidence: (1) they are logically correlated, (2) recent results of studies by leading experts and surveys indicate a correlation (3) they also form part of most if not all past models of well-being which have been proposed by some of the most famous and influential thinkers, (4) although some caveats have been raised, to our knowledge no one considers them to be detrimental, (5) objective evidence exists that hormones associated with well-being which have been released due to enjoyable and affection stimuli. We describe these points in detail below.

(1) First, we feel it is logically intuitive that good feeling toward something or someone in the short-term (enjoyment and affection) may contribute to good feeling in the long-term (well-being), because a long-term period can be described in terms of a sequence of short-term periods.

(2) This intuition was supported by the most recent model of well-being to our knowledge proposed by an expert in the field, which includes these two factors. Seligman, a founder of the new field of positive psychology [22], set forth five factors which he felt were most central to well-being: enjoyment (positive emotion, pleasures and gratification), affection (positive relationships), accomplishment, meaning, and engagement [30, 38]. Five factors, although not so many for a model in psychology prescribing how people should live, seemed quite daunting in our context of robotic interactions; therefore, we wished to know which factors were most fundamentally associated with well-being.

We obtained some insight from a recent survey performed by the American Association of Retired Persons (AARP) Research & Strategic Analysis, which asked 4397 respondents to describe what they perceived to be the most important causes of well-being. The result was that the most important factors perceived to be associated with well-being other than health were enjoyment (pleasure) and affectionate relationships, which were rated higher than the other factors (accomplishment, meaning, and engagement). In particular, the strongest contributors to well-being were felt to include enjoyable experiences of something positive, funny, or unexpected, as well as affectionate experiences of hugging or kissing a loved one, and being with loved ones such as family, friends, or pets [39].

But these expectations of an expert and regular people alone might not be valid. Some extra evidence is offered by Kahneman and Krueger who measured the amount of time people spend in an unpleasant state (based on participants' self-evaluations using the Experience Sampling Method and Day Reconstruction Method) to get insight into well-being at the level of feelings; they found a high correlation between well-being and enjoyment of 0.73, stating that this is intuitive because "it would be odd to feel happy but not feel enjoyment" [40]. Furthermore, a study by Helliwell and Putnam also provided evidence that well-being was closely correlated with affectionate interactions with family and friends, based on data from two sources, the Social Capital Benchmark Survey in the US with 29000 respondents and the Canadian Equality, Security and Community survey (ESC) with 7500 respondents [14]. Thus, these recent studies suggested that enjoyment and affection are expected to be particularly important factors for experiencing well-being.

(3) This supposition was also supported by the earlier literature, in which much discussion can be found of enjoyment and affection in conjunction with well-being. In Ancient Greece, Aristippus proposed the importance of interpreting circumstances in such a way as to provide enjoyment [41]. Plato presented some debate questioning whether or not enjoyment is the only good [42]. Plato's student, Aristotle, thought that while enjoyment would be experienced in a happy life its pursuit alone would be insufficient to achieve well-being; he counseled the importance of actively and genuinely behaving with affection toward others, as part of living a virtuous and excellent life [17]. Agreeing that people should live rightfully, Epicurus described the importance of both enjoyment, which included avoiding fear and pain, and affection, in the form of friendships [43].

Enjoyment and affection continued to be discussed in conjunction with well-being in the modern time. Similar to Hutcheson [44], Bentham claimed the importance of enjoyment in the form of his quantitatively hedonic principle of "greatest happiness", stating that, "Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do, as well as to determine what we shall do" [45]. Mill adopted a slightly different qualitatively hedonic position, arguing the existence of "higher" and "lower" forms of enjoyment [46].

In the twentieth century, Maslow proposed an hierarchy of requirements for well-being, including a basic need for affection which he thought would result in deficiency if unfulfilled [34]. Extending work by Freud, Erikson expressed a developmental conceptualization of well-being, which he believed would stem from successful mastering of some typical psychosocial crises; in this process, enjoyment perceived from interacting could be hampered by guilt over one's desires, and affection could be hindered by fear of intimacy or self-centeredness [47]. Rogers described a "good life" in which fully functioning persons were expected to more intensely experience enjoyment and pain, affection and loss; like Maslow, he believed that people feel a basic need for affection (positive regard) [48]. Allport used the term "psychological maturity" to refer to a state of well-being afforded by a capability to find enjoyment in circumstances via humor and sustain affectionate (warm) relationships; enjoyment, in providing intrinsic motivation, was considered to be an important driving force in human behavior [49]. Based on such previous work by Maslow, Erikson, Rogers and Allport, Ryff compiled a model of what she called "psychological well-being"; there she also described the importance of affection in terms of positive relations with others [50]. Thus, much discussion has revolved around enjoyment and affection in the context of proposals for well-being.

We note that some other possible correlates of well-being were proposed by only one or a small number of authors. For example, Plato's debate in Protagoras mentioned five virtues: wisdom, temperance, courage, justice and holiness [42]. In modern times; Rogers also highlighted the importance of openness and creativity [48], Allport furthermore discussed social involvement, emotional security, and empathy [27]; and Ryff additionally described self-acceptance, autonomy, environmental mastery, purpose, and growth [28]. In contrast, enjoyment and affection appear to form a fundamental part of most, if not all, earlier models of well-being. Perhaps this is because, as Epicurus noted, such eudaimonic factors are not independent but may themselves yield enjoyment [43]. For example, behaving wisely, exerting our creativity, interacting socially, and achieving mastery over our environment can feel good. Thus, the literature suggested to us that enjoyment and affection were important and should be investigated.

(4) Likewise, we could not find any sources which claim that enjoyment and affection are detrimental or do not lead to well-being. The most salient caveats we encountered are as follows. A consideration about providing enjoyment is that momentary pleasure does not necessarily lead to future pleasure. For example, if we succumb to the enjoyment of watching television instead of doing some

necessary work, we might regret our choice the next day. However, it does not entail that enjoyment is necessarily bad for people. As above, in the sense described by Epicurus, we can think of "work" as also providing enjoyment in the future if is done [43]. Thus, we believe that a robot should not force people to play with it. Also, although we focus on short-term enjoyment in our current work, it will also be important in later work to consider what impact a robot's behavior can have on a person's enjoyment not only in the present, but in the future as well.

Similarly, we did not find objections to the general idea of providing affection. Only in the context of interactions with relational artifacts, some ethical concerns have been voiced in regard to the possibility that such interactions could potentially be damaging, in the following ways: (a) by promising emotional connections which are not "real" and prevent and take the place of real human relationships, (b) by teaching that relationships can be controlled just as we wish, and (c) by inadvertently hurting people's feelings (e.g., a robot which breaks down could seem to be ignoring a person out of dislike) [51]. We offer some of our own thoughts on these topics. (a) People can feel one-sided affection toward characters in literature, songs, or movies, as well as some pets, without sacrificing their relationships to other people; rather, such affection could even facilitate interactions by yielding topics to discuss with others and motivating people, through the empathy they feel, to seriously think through hypothetical interactive situations which they too might one day face. (b) People can exert control over books, songs, movies, and pets (for example, putting a book down or skipping to a different chapter), without thinking that they should do the same toward other people; not being able to exert control might actually be more damaging to our relationships, for example, if we could not pause a movie to talk with someone. (c) People's feelings can also be hurt by, for example, offensive words in a book, song, or movie, or a pet which bites or runs away from them when they seek to interact; such experiences, in cautioning that we cannot control everything in life, could even be desirable from the perspective of the second concern above. Thus, we believe that designers should be aware of such important potential ethical problems, but we think that people in general will gain from interactions with robots in the same way that we believe existing media such as books, songs, and movies, and relationships with animals have positively affected our lives.

(5) An interrelationship between well-being, enjoyment, and affection was furthermore indicated by some results based on objective data: hormones associated with well-being have been found to be released in conjunction with enjoyable and affectionate stimuli. For example, vasopressin has been associated with enjoyment and affection [52]. The “love hormone” oxytocin [53], which supports well-being and improved health [54] has been released during affectionate touching [55]. Endorphins which provide well-being by reducing pain [56] have been suggested to be released due to enjoyable music and exercise [57] and to serve a function in affectionate relationships [58]. Serotonin, which facilitates happiness [59], has been measured in non-humans, after enjoyment from finding food [60] and affection [61]. Dopamine, tied by the World Health Organization and researchers to well-being when heightened via addictive drugs [62-63], has been measured after enjoyable experiences involving food and affectionate sexual experiences [64]. Thus, a connection has been indicated, which implies that improved well-being may result if enjoyable and affectionate experiences can be realized.

In summary, we believe that enjoyment and affection are related to social well-being based on the logical, subjective, and objective evidence described. Despite this evidence, however, we cannot conclusively state with absolute certainty that enjoyment and affection will help cause social well-being. Therefore at this stage of our research, we avoid making causal claims, and only describe social well-being as our desired end goal, and enjoyment and affection as a focus of investigation which we ourselves expect will offer progress toward our goal.

Regardless of how these factors relate to well-being, enjoyment and affection have been also acknowledged to be important also as goals in their own right. Enjoyment is a fundamental design criterion which should be maximized [65] and not disregarded when designing functionality [66]. It promotes acceptance of new technologies, including robots, as in ISO 9241-11 [67-68]. Affection is a crucial constituent of human communicative behavior which plays a central role in the establishment, sustenance, and enriching of bonds [69-70, 35]. People have a need for affection [34] which in many individuals is not sufficiently met [36]. Such a perceived lack of affection has been linked to highly detrimental health outcomes [71-72]. Furthermore, affection is not only shown to other people or pets; people have been observed to also hug and stroke robots [73, 37]. As well, some studies have reported promising results that even interactions within a simple context can be therapeutic [74-76].

The importance of enjoyment and affection for humans can also be noted in positive effects due to associated neurotransmitters. For example, oxytocin may assist bonding, improve sleep, facilitate the healing of wounds, reduce blood pressure and lower perceived anxiety and stress [55]; serotonin supports well-being, plays a role in the regulation of sleep, appetite and mood, and mitigates depression [59]. Thus, we expect benefits if enjoyment and affection can be provided.

But are enjoyment and affection truly distinct? And how do they relate to emotion in general? In response to the first question, we understand enjoyment and affection to be closely related. For example, affection may not play a role in enjoyable experiences such as watching a movie, but enjoyment can be perceived from affectionate experiences such as being patted on the back. The main difference in how we use these terms arises from their focus: enjoyment results from an activity or stimulus, while affection is communicated by a person. As an example, perceiving satisfaction from a massage does not entail feeling devotion toward a masseuse; also, love may be felt toward a family member even when a pleasant interaction is not occurring.

Thus, we distinguished between these two ideas in this dissertation. By splitting our problem into two parts, our investigation became easier, which was desired due to the highly challenging nature of our problem. Also, this separation was natural. It occurred in our main sources [38-39], and we could furthermore imagine the difference in our case of robotic interactions. A person could seek enjoyment in causing a robot to dance, or show affection by stroking it; to elicit a person’s enjoyment or affection, a robot could perform a hand-stand (which might not be affectionate) or cry like a child (which might not be enjoyable to hear but could evoke a care-taking response in the human).

We also clarify how we believe enjoyment and affection relate to emotion in general. Emotion can be modeled via a small number of dimensions [77] or in terms of discrete elements. For the former case, in the general sense in which enjoyment and affection can be described as positive feeling toward something or someone respectively, they could be compared with "valence"; valence is an important, fundamental quality describing the positivity of a signal [78], which has also been used to assess the effects of individual channels in behavior involving multiple signals [79]. A difference is that valence can be used to describe any signal, whereas enjoyment and affection relate to something or someone respectively. For example, smiling to one's self is a signal of positive valence which can arise intrinsically without being caused by some external stimulus or person. In the sense that emotions may also be understood as a set of discrete constructs, enjoyment and affection can be related to a few emotions such as joy or love, but cannot alone account for many other emotions such as determination, surprise, confusion, pride, acceptance, doubt, embarrassment, envy, excitement, hunger, interest, or sadness. Thus, enjoyment and affection could be described as one part of emotion.

With this understanding of our target for investigation, we next needed to consider how enjoyment and affection can be evoked in an interaction with a robot, which is not readily apparent. To identify the important components involved, we adapted a standard communication model from Schramm [80], which modifies the Shannon-Weaver stipulation [81] to describe continuous dyadic communication, as shown in Figure 1. In our model, a human interacts based on some behavior strategy (an internal state including reasons for interacting relating to enjoyment or affection), which is recognized by a robot. The robot in turn acts based on its own strategy, which is then recognized by the human. A similar model for interactions was used in robotics by Yohanan and MacLean [82]. This model suggests that a robot's recognition and strategy play a central role in determining how an interaction affects a person, and that improving these factors could influence a person's state of well-being.

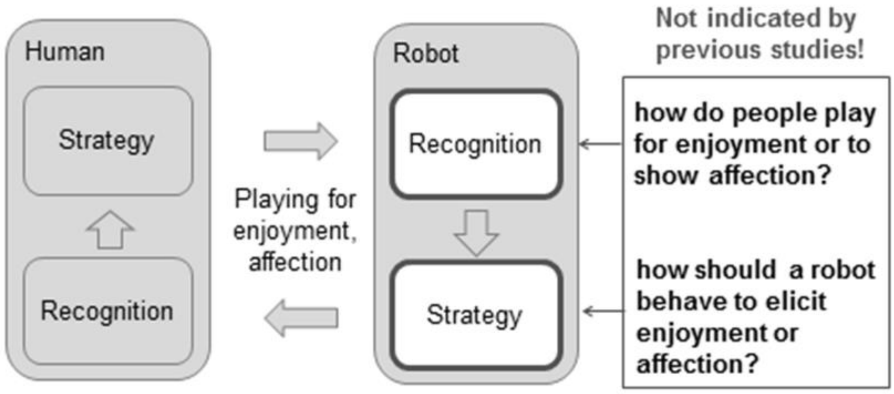


Figure 1. Process flow during an interaction between a human and a robot.
Based on a standard dyadic communication model [69]; by focusing on "Recognition" and "Strategy" on the right, we seek to affect the human's state of well-being (the left side).

We clarify our terminology relating to this model. First, a robot must recognize social "signals" directed toward it. A signal in general is a "perceivable stimulus … produced by an emitter" [83] which can be "anything that serves to indicate, warn, direct, command, or the like, as a light, a gesture, an act" [84]; by this definition a signal comprises both intended communications and non-socially directed "behavior". Behavior is "observable activity" [85] comprising one or more signals. We also use the terms "action", "channel", and "modality" to describe behavior at different levels of detail as follows. An action is "something done or performed; (an) act; (or) deed" [86]. A channel is a means for performing some related actions [87]; some examples of channels are the face, body and tone of voice [83]. A modality is "one of the primary forms of sensation, as vision or touch" [88]. Thus, a smile could be described as an action associated with the channel of the face and principally transmitted through the modality of vision. A person's behavior may also carry meaning. We refer to behavior which possesses many meanings as "polysemous", or having a "diversity of meanings" [89].

In addition to recognizing, a robot requires some strategy which is used to decide a robot's behavior based on the recognition results and other interaction history. Behavior can involve an internal change of state (the robot's intentions), reactive responses, and proactive suggestions. We use the term "response" to describe robot behavior possibly comprising motions or sounds which follows soon after, is causally associated with, and provides feedback related to a recognized human behavior; a "suggestion" is a robot behavior which occurs spontaneously, especially when a person is not interacting.

But are a robot's recognition and behavior strategy truly important? We present some additional evidence suggesting the benefits of improving this functionality. First, recognizing and responding in a socially intelligence way to a person's behavior has been linked with enjoyment and increased technological acceptance [68]. Furthermore, perceiving not only the contents but also the significance of behavior has also been described as a desirable factor for realizing improved communication [90]. On the other hand, shortcomings may be observed when such functionality is limited. For example, the use of scripts by a robot to guide an interaction reduces a person's control; but perceived control is an important correlate of enjoyment [32]. Also, simply imitating a person's behavior can be impossible or undesirable in some cases, such as when a robot is picked up or stopped from performing an undesirable action. In addition, recognizing a set of arbitrary signals risks missing important cues if the set is too small and never showing some responses if the set is too large. Likewise, a need for human operators for all companion robots would be highly costly. Figure 2 portrays our expectation that interactions will be greatly affected if a robot cannot recognize and respond appropriately to a person's behavior.

Within this designated framework, we investigated the act of play involving touch with a humanoid companion robot, which we believed to be the most important focus. "Play" is a fundamental "activity for amusement or recreation" involving "fun or jest" [91], which is free and engrossing [92], and offers intrinsic enjoyment to both the young and old [93]. For a robot, engaging in play can be acceptable when circumstances are not serious [94]. At such a time, a robot can act either like a toy, providing enjoyment directly itself, or as an affectionate partner playing with a human. Robots also do not need to be only capable of playing; we expect people to play with all kinds of robots in the same way that people play with other people and objects including those intended to serve other purposes (e.g., twirling a pen, kicking a chestnut, reflecting light using a watch, or tossing a can).

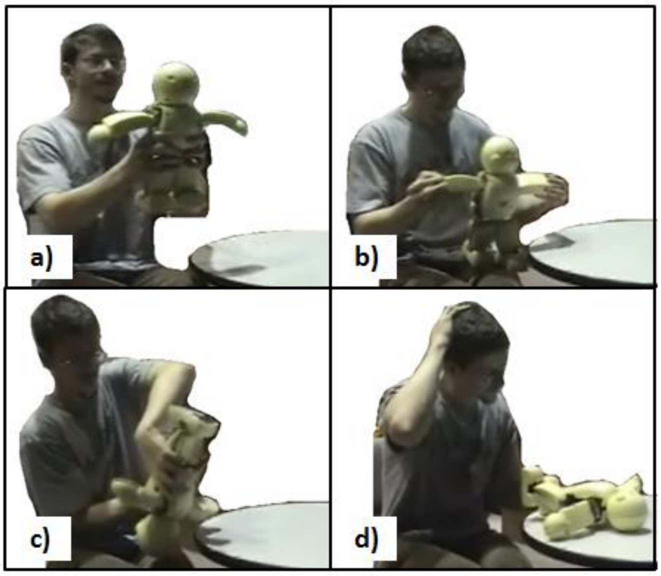


Figure 2. Importance of recognition and behavior strategy during play.
Our concept: a) people will sometimes seek to play with co-existing robots, b) people will watch and enjoy a robot's behavior such as walking, c) people will also expect a humanoid robot to recognize their own behavior and respond appropriately d) problems with either capability will limit the degree of enjoyment and affection felt.

Furthermore, although many modalities may be important for enjoyment and affection, we decided to focus on touch first for two reasons. First, touch is fundamentally associated with both enjoyment and affection. The role served by touch for enjoyment may be seen in the great number of toys which involve touch and movement, such as dolls, bricks, balls, disks and ropes; the importance of touch in the communication of affection is also well known [95-98]. Like the toys listed above, a robot can be touched and moved, and itself move to provide feedback. Some examples of robots with touch sensing capability include Huggable, Robovie, Macra, and Haptic Creature [99-102]; some examples of robots which move and are moved include Keepon, Roillo, and Roball [103-106]. A second reason for first investigating touch is that touch has not been well studied compared to other modalities (e.g., [95, 107]); by investigating it we could fill an important gap in the literature. (Because other modalities may also play an important role, however, we have also started to investigate them in some work which has not yet been published).

Moreover, the case of an interaction with a humanoid robot is important to explore. Although the most well-known companion robots in the past, AIBO and Paro [8, 76], have been designed to be animal-like to leverage people's lower expectations of such robots' capabilities, we think that humanoid robots offer the potential for rich interactions. This is based on the fact that the humanoid form is a familiar interface which people find easy-to-use, and which will therefore be perceived as enjoyable [108].Thus, our basic focus is shown in Figure 3: we seek to develop touch-based recognition and a behavior strategy for a humanoid companion robot to interact enjoyably and affectionately, toward contributing to people's social well-being.

As a specific focus, we selected an interaction in a controlled indoors environment, with both human and robot willing and capable of interacting, and no objects provided. For our target demographic we decided to focus on young adult Japanese. Young adults were selected as our target because children could act unpredictably and elderly could be easily injured in an accident. Japanese participants were recruited because our group is situated in Japan; however, we expect that this is also not a bad choice of target demographic because the robotics industry has been described as more important in Japan than in any other country in the world [81].

In summary, we defined some important terms used in this dissertation, how we believe the designated topics interrelate, and why we feel they are important; next we consider in more detail the specific problem scenario, why it is challenging, and how we approached solving it.

1.2 Our Approach

In order to further clarify the situation we would like to address, we describe one example of a usage scenario involving a hypothetical user, as follows:

> Masha wasn't feeling too well for a while. It wasn't just because he got sick, but also because he had a lot of time to himself in the hospital alone, which just wasn't any fun. It would have been nice if he could have brought his cat at least, but the nurses told him immediately it wouldn't be possible for hygienic reasons. That's when his parents surprised him with a really great gift: his own companion robot, Gozilla! The name was a joke because Gozilla wasn't monstrous, but rather small and cute. Masha found that he enjoyed playing with

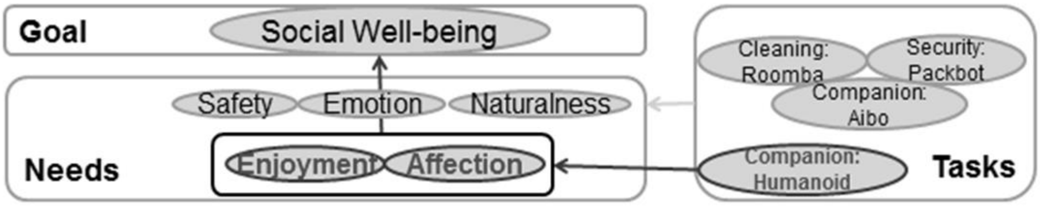


Figure 3. Our basic focus shown within a general context of HRI.
We seek to provide enjoyment and affection in a touch-based interaction with a humanoid companion robot by developing the robot's recognition and strategic capabilities, toward contributing to the social well-being of interacting persons.

Gozilla and quickly grew fond of it, almost as if it were a little friend. Sometimes he liked to tease it sort of like he played with his cat by holding it down or play-fighting a little with it. But he always took care to help it to stand again and to not be too rough because otherwise Gozilla would start to cry, which made him feel bad. Then he had to quickly give the robot a hug, which made it feel better, and felt nice for him too, because the robot was soft. For that reason Masha sometimes liked holding it at night, when the lights were out. He also liked that he didn't have to go get Gozilla each time; it could actually whiz over to him when it got lonely, and even suggest ways they could play together! Another cool thing was that Gozilla didn't always do the same thing; it acted a little bit different each time, so it didn't get boring. Gozilla even surprised Masha a few times with how smart it was: it could even understand a few simple words and gestures he made. Like once, he called out to Gozilla when it seemed like it was going crash into the wall and it actually turned around and came over to him! Or when he showed Gozilla a new watch, the robot actually looked at it. Masha is looking forward to soon getting out of the hospital, so he can show off what Gozilla can do to his friends.

The persona-based narrative above suggested to us some topics which we believe should be addressed. Gozilla needed to appear fun and likable to Masha. Gozilla also needed to recognize playing for enjoyment when Masha laid it down or play-fought and affectionate touches such as hugging. Moreover, Gozilla needed to know how to behave in an enjoyable or affectionate manner by crying, showing happiness, or coming close to Masha. Gozilla also should be able to vary its behavior over the long-term, and recognize and respond to some cues in other modalities such as simple words and gestures. Thus, we believe the most important topics included how a robot's embodiment is perceived, how a robot can recognize and respond to a person's behavior, and how to deal with other modalities than touch. Sub-topics for perception of embodiment include how people perceive robots of different size and form (large vs. small; humanoid vs. non-humanoid), sub-topics for recognition and strategy include enjoyment (which could involve objects) and affection, and for strategy a long-term policy is also important. Some of the identified topics and sub-topics have been partially addressed by previous work, such as perception of embodiment and long-term strategy (explained in Chapter 1.3).

In particular we expect that a long-term adaptive strategy will be important for sustaining long interactions once the novelty effect has passed, and not boring people with repeated behavior. This expectation is supported by the psychology literature, which warns of "hedonic adaptation" (that people quickly revert to their usual level of well-being after one-time events such as winning a lottery or being in an accident) [110] and suggests that frequent social interactions are strongly associated with higher well-being [14]. Therefore, our end goal is to realize not only single interactions, but to be able to build a longer term relationship spanning the course of a number of interactions. We did not focus on this topic in our first endeavor for two reasons. First, because our problem is highly challenging, before addressing long-term interactions we needed to investigate the easier problem of short-term interactions. This yielded insight into some possible fundamental principles for providing enjoyment and affection, which could be used for realizing longer interactions. Second, some work has already been conducted on this topic; e.g., in regard to personalization, identifying people, and encoding motions as primitives (e.g., [111-114]). We wished to focus on what we felt were the most fundamental unexplored topics.

Therefore we selected some sub-goals which have mostly not been investigated, to address in four steps, as follows:

Study 1) *Recognizing enjoyment-motivated play*: first, we focused on the case in which a person plays with a small held robot by moving its body. Sub-topics involved identifying typical kinds of such playful behavior and how to recognize them automatically from sensor data.

Study 2) *Enjoyable strategy*: second, we needed to know how to use recognition results to structure the robot's behavior for enjoyment. We desired to formulate the strategy in terms of a set of guidelines, because guidelines are useful for complex cases in which people's behavior cannot be fully predicted [115]. In doing so, we wished to avoid being too specific (reducing the applicability of the concepts) or too general (thereby not addressing topics); toward this, we followed an example in past work by seeking to specify guidelines as general concepts accompanied by specific details [116].

Study 3) *Recognizing affection-motivated play*: we focused on the case in which a robot recognizes typical touches. Our goal was to identify behavior, its basic affectionate meaning in terms of a single value representing the degree of like or dislike conveyed, and how to recognize it, which involved exploring suitable modalities and features.

Study 4) *Affectionate strategy*: previous studies suggested that affection can be felt from simple contact or touches like stroking [69, 107], even from objects [35]. Therefore, we believed that the important problem for showing affection was that current robots, unlike children and pets, cannot themselves typically get close to people to touch and start interactions. In particular, a humanoid robot with legs or wheels can fall or get stuck when there are obstacles on the ground. To overcome this problem, we proposed a new approach involving flight. For this, we needed to build a proof-of-concept prototype and derive a strategy for how a robot could get close.

We describe Study 1 in Chapter 3, Study 2 in Chapter 4, Study 3 in Chapter 5, and Study 4 in Chapter 6. Aside from these four sub-goals, it should be noted that we consider providing well-being to be a very broad and complex area. Many problems, not just the ones noted above, will have to be solved. The sub-goals we addressed are shown along with other topics we consider important in Figure 4.

To accomplish the sub-goals we proposed, the foremost obstacle was the vast complexity of human signaling conventions. Dautenhahn *et al*. characterized human behavior as "very complex and subtle" and the task of meeting people's

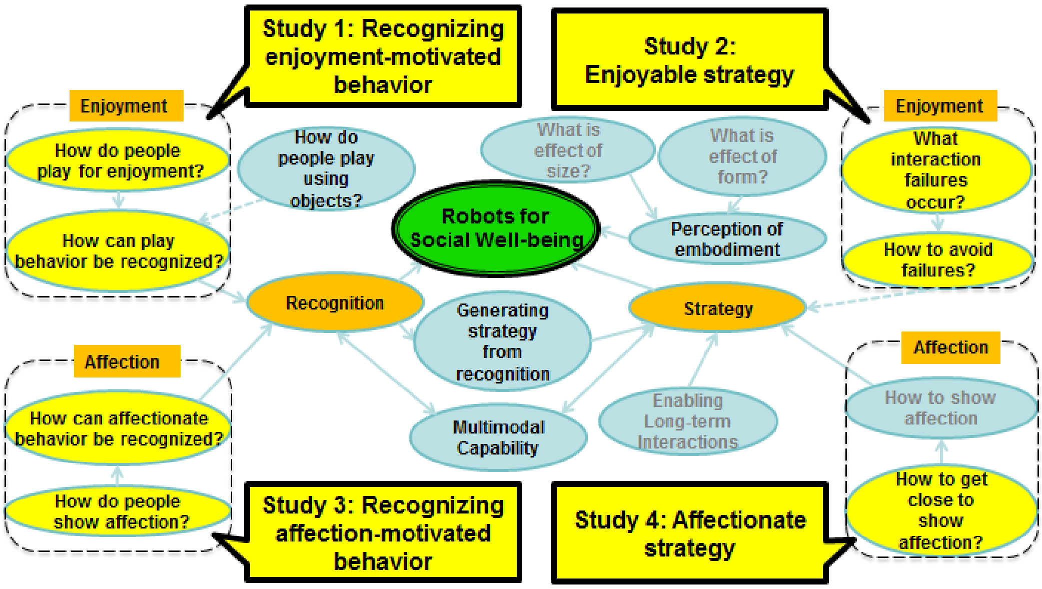


Figure 4. Some important topics for developing robots for social well-being. Our approach comprising four steps is depicted in this context: (1) recognizing enjoyment-motivated behavior, (2) finding an enjoyable behavior strategy, (3) recognizing affection-motivated behavior, and (4) finding an affectionate behavior strategy. The green ellipse shows our main goal. Orange and yellow shapes show topics and sub-topics which we focus on. Blue ellipses are not focused on in the current work. Gray font indicates topics which have been partially addressed by previous work, and thus which are not part of our focus.

expectations in designing a robot as "one of the most challenging open issues" [117]. Pantic and her colleagues also described the exploration of how human beings interact as "one of the greatest scientific challenges" [83]. Vinciarelli, Salamin, and Pantic referred to perceiving behavior as "extremely difficult" and estimating reasons for behavior, which we wish to begin to do by predicting degrees of affection communicated, as a "virtually unexplored area of research" [118]. Some challenges include that an indication is missing of how people engage in enjoyable and affectionate play with a robot, how such behavior can be recognized, and how recognition capability can be utilized to elicit enjoyment and affection. Such problems are not trivial because human behavior exhibits high variance, occurs in complicated multi-channeled patterns, and can signify different meanings. It is not clear how to identify useful signals, how they can be recognized, or how to attribute meaning to conflicting signals. Also, analyzing behavior is costly and prohibits haphazard investigation. Likewise, the challenge for using flight to allow a robot to get close to a person to show affection is that a vast range of possible flight behavior exists; we have no knowledge of how such a robot should fly in proximity to humans or what meaning will be attributed to such flight.

Our approach, stated concisely, involved preparing robot prototypes, observing people playing with the prototypes, and analyzing people's subjective assessments to obtain further insight with regard to how interactions were perceived. First, to design prototypes, intuitive knowledge was drawn from previous designs and products described in the literature with regard to embodiments, behavior, and interaction strategies. Then we constructed mid-level fidelity prototypes, to avoid two problems of high-level and low-level fidelity prototypes: namely, employing a platform which may not meet our requirements, and not being able to observe actual interactions [119]. Second, to observe interactions, we followed a general approach described as the "observational approach" [120] or the "grounded theory method" [121]. To deal with the high variance in people's behavior, we identified typical behavior performed frequently by multiple persons. We also asked participants to interact for typical reasons to acquire some insight into the polysemy, or meanings, of typical signals. Recognition systems were developed iteratively using Support Vector Machines (SVMs) and heuristic thresholds on inertial data and hybrid touch/vision sensor data. Then behavior strategies were designed by seeking typical patterns using a design pattern approach [122-123], which included finding typical failures and their causes, and typical perceptions. To acquire feedback, we obtained subjective measurements including questionnaire scores and interview comments. Although objective measurements may be generally preferred to subjective ones, in some cases such as ours only the latter means allowed determining how people perceived various aspects of the interactions [124]. This data allowed us to acquire some understanding of the "usage and coding" [125] of behavior, and other perceptions directed toward a humanoid robot.

As a result of surmounting the described challenge, we were able to accomplish the objective of this research—gaining new insight into key issues associated with providing enjoyment and affection through touch-based play—toward contributing to people's social well-being via interactions with a companion robot.

1.3 Related Work

In this sub-chapter we seek to describe the novelty of our work in relation to some similar previous studies. Our work touches on a number of aspects: (1) establishing a theoretical background for seeking to provide social well-being via enjoyment and affection during human-robot interaction (our work is compared with other related work in Chapter 1.3.1); (2) identifying what kind of embodiment could be appropriate in such a scenario (Chapter 1.3.2); (3) building a robot's behavior including motions and sounds (Chapter 1.3.3); (4) investigating what behavior should be recognized and how to recognize it (Chapter 1.3.4); (5) exploring how a robot can behave based on recognizing people's behavior (Chapter 1.3.5); and (6) also investigating how a robot could use flight to approach people for showing affection (Chapter 1.3.6).

1.3.1 Providing well-being, enjoyment, and affection with a robot

As described in Chapter 1.1, well-being was proposed as a crucial topic and debated thousands of years ago—likewise, the basic idea of building something to fulfill social needs is not new. Dolls created even in Ancient Egypt and Ancient Greece [126] provided comfort and affection [127]. Furthermore, stories of friendly inanimate artifacts existed, such as an account of a statue which came to life and married its creator [128]; moreover, proposals were made for how to build robot-like servants and entertainers designed to provide enjoyment and comfort [129]. In modern times, many ideas for fictional robots which can recognize social signals and assist humans have been described in the popular media in works such as Astro Boy, Doraemon, Star Wars, and Short Circuit; some relevant scenarios involve an adolescent child playing with a humanoid robot in Terminator II and a detective's affectionate relationship with an android in Bladerunner. However, advances in robotics have only recently allowed researchers to start paying attention to realistically developing such robots.

First studies relating to well-being in robotics have focused mostly on conceptualization and healthcare. Picard first formally expressed the concept in robotics that artifacts could recognize and express emotional cues [11], which was later discussed in relation to touch-based interactions [130]. Such intelligence, it was proposed, could be required in robots to help reduce people's labor [131].

Reddy stated that robotic technologies could "profoundly impact the well being (sic) of our society" [132]. Much important work has also been done in healthcare. For example, Graf *et al*. built a robot which could offer drinks to facilitate sufficient liquid intake [133]. Furthermore, Kawamoto and colleagues reported on an exoskeletal assistive platform for paralyzed individuals, HAL [134]. Also, Mukai *et al.* described a robot built to hold and transport people, RIBA [135]. Thus, such studies have described advantages if a robot could facilitate aspects of a person's well-being, and some healthcare applications, but fewer studies have investigated principles for providing enjoyment and affection.

To provide enjoyment, Takeda, Kosuge and Hirata constructed a humanoid robot which can dance with a human [136]. Another humanoid robot, TOPIO, was built to play Ping-Pong [137]. A playful scenario using flying robots was described by Müller, Lupashin and D'Andrea, in which the robots can play catch with a person [138]. An enjoyable exercise scenario was reported on by Fasola and Matarić reported on another enjoyable scenario in which a humanoid robot exercises with a human [139]. Hansen, Bak and Risager observed enjoyable interactions in which an elderly person chased or was chased by a humanoid robot [140]. Thus, these previous studies focused on specific scenarios which had been decided ahead of time.

Some other studies have focused on affection. In 1965, Grandin built a hugging machine which could provide calm therapeutic touch [74]. DiSalvo and colleagues proposed a design for a minimal humanoid form capable of transmitting hugs to a remote person [141]. Samani and colleagues created a teleoperation artifact intended to transmit affectionate hugs and kisses [142]. Thus, these previous studies developed human-operated tools for communicating affection.

In summary, studies relating to enjoyment and affection in robotics did not show how an autonomous robot could be enjoyable and affectionate in a scenario in which a person was free to choose how they wished to play.

1.3.2 Robot embodiments

In developing prototypes for enjoyment and affection, we drew some indirect inspiration from various previous embodiments, also including some toy robots.

Some were humanoid (e.g., My Real Baby, Robosapien, Nao, E.M.A., FT Robot, and QRIO [143-144]); others resembled toys (Huggable [99]) and Roball [105]), common pets (AIBO, NeCoRo, and Dream Cat), or creatures with which people don't normally interact (Paro [76] and Pleo).

Some robots have also been designed to elicit affection. A "neotenous" or juvenile appearance featuring relatively a large cranium and eyes may attract care-giving [145]. Such an appearance may be seen in robots such as Huggable, Kismet [146], Leonardo [147], Keepon [103], and Paro; as well as toys such as Furby and My Real Baby, virtual agents in games such as Love Plus and Tamagotchi, and popular characters such as Hello Kitty. Additionally, care-taking behavior could be affected by a robot's height, which appears to influence perceived dominance or submissiveness [148]. Some evidence in human science has also suggested a link between softness of an embodiment and perceived affection. Bowlby proposed that humans have a fundamental and innate need to touch and cling [69]; Harlow furthermore found evidence that our close relatives, monkeys, preferred to cling affectionately to soft mother surrogates as opposed to hard ones, even when the hard ones offered alternative rewards such as nutrition [35]. Thus, these clues from previous work were considered when developing our platforms and are not new knowledge.

1.3.3 Robot behavior

Previous products were also considered in constructing behavior for our prototypes. Robots' behavior appears to be tied to affordances from a robot's appearance. For example, humanoid robots E.M.A. and Robosapien dance and walk like humans, cat-like robots NeCoRo and Dream Cat purr and meow like cats, and the ball-like Roball rolls and spins like a ball. Some robots seem to have many motions, such as AIBO or Robovie. Because many robot motions have been described in the past we did not focus on this in our current work.

1.3.4 Recognition capability

Some novel work we conducted involved identifying people's behavior toward a robot, what it meant, and how to recognize it. The first problem of how people behave directly depends on the previous topics we have discussed, robotic embodiment and behavior. We believe this is an important question because both what meanings can be conveyed, and properties of a recognition system such as modality and features depend on what people do. Also, we believe that no advantage is obtained from the capability to sense behavior which a person will never perform; likewise, not being able to sense behavior which people perform reduces a robot's capability to react and display social intelligence, which is undesirable. Therefore, we believe a first question which should be asked when developing recognition capability for an interaction is *what* should be sensed.

1.3.4.1 Identifying behavior A number of studies have addressed this problem for the case of touches directed toward a humanoid robot. Noda and his colleagues seem to have been first in robotics to systematically approach the problem of what to recognize in an interaction by observing interactions with a humanoid robot and manually identifying a number of categories to represent typical playful touch behavior [149]. Knight *et al*. soon after published a useful list of playful touches directed toward a teddy bear robot at three levels of "granularity" as categories, gestures, and sub-gestures [150]. Yohanan and MacLean provided a different list of touches which were expected to be directed toward a cat-like robot, based on the human and animal science literature [82]. What remained to be studied was how people move a humanoid robot's body to play and touch to show affection.

1.3.4.2 Estimating the meaning of behavior Not only the "contents" of communication, but also the underlying meaning is important to recognize for good interaction [90]. That is to say, it is desirable not only to have some idea of which signals should be recognized, but also how they can be interrelated, ranked, or otherwise understood; in other words, if a signal can be sensed, then how should a robot interpret its meaning? Such inference can act as a first step for informing the design of a behavior strategy.

In regard to playing for enjoyment, we did not find any previous work which suggested the degree of enjoyment underlying touches could be recognized, which is intuitive because people enjoy playing in different ways. For example, it would be difficult to assert that raising a robot high into the air indicates more or less enjoyment than causing a robot to dance. However, several previous studies have report results which can be related to estimating the affection shown by different kinds of touches. A pioneering work by Stiehl and Breazeal first reported on a force-based classification system built to distinguish “pleasant” and “painful” touches [151]. François, Polani, and Dautenhahn also constructed a system to discriminate between “gentle” and “strong” touches [152]. Hornfeck, Zhang, and Lee likewise chose to distinguish between “gentle” and “harsh” touches for each touch sensor on their robot based on heuristic thresholds [153]. Thus, some studies described how to distinguish between soft and hard touches, which could be related to affection; however, they did not consider the effect of other factors which could be important, such as the location of touch.

Some alternative models were described in the psychology literature. One slightly more complex conceptualization than those discussed above was proposed over a century ago by Hall and Allin, who reported that not all soft touches are affectionate: highly soft touches can be irritating [154]. Therefore, an improved model might group hard touches and highly soft touches together as unpleasant or comprise three categories. Alternatively, a much more challenging categorization scheme could be constructed from the list of typical reasons for touching reported by Jones and Yarbrough [155] and a set of basic emotions which can be communicated via touches identified by Hertenstein *et al*. [107]; such a scheme would comprise many discrete categories representing specific reasons for touching such as thanking or showing fear. However, the complexity of this scheme might be prohibitive given that this topic is only beginning to be explored in robotics. Thus, some studies in human science suggested interesting possibilities but did not clarify what would be observed in human-robot interaction.

In summary, what was missing was a study which addressed not only the force of touch, but also the role of location of touch on the humanoid form; and described the meanings of typical touches in such a way that they can be interrelated and that robot responses could be easily associated with detected human behavior.

1.3.4.3 Recognizing behavior A third problem was how to actually recognize a human’s touch behavior from sensor data. In regard to recognizing motion-based play behavior toward a robot, some work was previously conducted using an inertial sensor. According to its specification, an accelerometer inside Omron’s cat robot, NeCoRo, "allows it to know its position when cradled or spun around" [156]. Salter and colleagues furthermore described an algorithm which could detect if a ball-like robot was left alone, carried, or spun, based on some statistical features calculated from accelerometer and tilt sensor data [106]. Moreover, Lee *et al*. reported on the development of recognition capability for a small teddy bear robot, Huggable, in which frequency-based features were used to recognize three proprioceptive signals: pick up, bounce, and rock [99]. Also, the specification sheet for the toy robot Pleo states that it has a tilt sensor which can be used to recognize "six possible states, shaking, dropping, etc..(sic)" [157]. Thus, previous work offered some useful insights, such as potentially useful features, but did not indicate how people's typical play behavior when moving a small held humanoid robot could be recognized.

Some work has also focused on recognizing touches which did not involve motion. For example, Naya, Yamato and Shinozawa reported on recognizing some elementary kinds of touches such as rub or pat [158]. Yoshikai and his colleagues described recognizing some different three-dimensional touches, pinch and twist [101]. Some other studies which recognize using touch sensors have been discussed in surveys [159]. Moreover, commercial products such as AIBO, Paro, Furby, and Pleo possess some touch sensing capability although it has not been made publically known what they can sense. Thus, previous work has focused on recognizing some touches using "touch" sensors which measure force, velocity, or close proximity, but little attention has been given to other possible sensor modalities.

Although touch sensors have some advantages such as conceptual simplicity, substantial disadvantages also exist. For example, a touch sensor system is typically not portable and cannot be exchanged easily between robots with different sizes or shapes. Also, such systems tend to involve many sensor units with much wiring which are physically pushed in interactions; therefore such systems are prone to breaking [149]. Moreover, touch sensors must be typically integrated into the surface of a robot, which adds difficulty when seeking to design an attractive and soft exterior. Furthermore, because touch sensors capture data from only the instant of contact (and in some sensors slightly before and after), much data which could be useful for recognition is lost; for example, touch sensors cannot capture the trajectory through space of a touching part, or movement of the touched part.

A study by Hertenstein *et al*. in recent years suggested to us an interesting alternative [107]. In this study, people were able to recognize some fundamental emotions being transmitted via touch merely from a video, without actually feeling the touches themselves. This suggested to us that vision sensors could be used to capture visual data of touch gestures for recognition purposes. In this way, problems associated with touch sensors could potentially be avoided. However, vision sensors possess their own disadvantages. It may be difficult to determine when a person is touching a robot, data may be highly altered by changes in illumination, and occlusions and shadows may be encountered, especially because in the case of touching, a robot and human behave in close proximity to one another. Thus it was unclear if either approach, or a hybrid strategy, would be most useful, suggesting that a comparison should be conducted.

1.3.5 Behavior strategies

Part of our work also involved exploring how a robot should adapt its behavior based on its recognition. Although some few studies have touched upon ways to realize long-term interactions via adaptation and personalization [160, 111]—which could require recognizing individuals automatically through vision (e.g., [112-113])—most work has focused on proposing some general principles for developing a playful interaction. Fujita stated that the following principles would be desired to show that a robot is lifelike: a complex configuration with many degrees of freedom, motions triggered at various timings to sensed stimuli or based on internal intentions or emotions, and non-repetition of behavior patterns [8]. Michaud and Caron felt that their robot should be able to operate in various environments and be robust, autonomous, and interactive [161]. Robins *et al.* showed that a robot's timing and delays during play affected people's behavior [162]. Kozima *et al*. suggested that a minimal design, as well as rhythm and timing were important [103]. Some additional insight could possibly be obtained from the results of a study in which people were asked which emotions a cat-like robot should express in reaction to being touched [82], although it is not clear if these results would apply also to a humanoid robot and how a robot should specifically respond.

Some conclusions could also be drawn from observing previous toy designs. Robot toys such as Tickle Me Elmo, Robosapien, and E.M.A. have a simplified turn-based interaction style in which buttons can be pressed to trigger motion responses. Behavior may be divided into reactions and proactive suggestions. Toys directed toward children such as Tickle Me Elmo may feature large exciting responses; other toys for older people such as Paro may have a range of smaller and larger responses and some capability to adapt the amplitude of their responses based on a person's behavior. Proactive behavior may include idling motions which can be used to make robots appear lifelike (e.g., Pleo, Dream Cat, or Paro). We were not able to find any work showing how to design a robot's suggestions and other proactive behavior for play.

Thus, general principles and rules of thumb had been proposed as described above, but these were not sufficient by themselves for realizing an enjoyable interaction. However we were able to use these as a starting point in our investigation, by refining our initial design based on people's comments and other subjective evaluations. As such, our approach was similar to the spiral design employed by Michaud and Caron [161].

1.3.6 Flight capability for a companion robot

Another new point of our work addressed how flight could be used in a companion humanoid robot. The basic idea of a flying humanoid is not new; stories of flying humanoids have been told since ancient times in various cultures. For example, Ancient Greek mythology described the flight of Icarus and Daedalus; Japanese myth told of Tennyo, flying beings which resembled beautiful women; and Norse and Eastern Europe legends contained accounts of valkyries and vampires. In modern times, some stories of flying humanoids have involved robots; e.g., Astroboy and Gundam.

Recently many flying robots have been built, although their designers have mostly focused on realizing excellent mobility for purposes other than human-robot interaction. One company, Festa, created a flying penguin, jellyfish, and bird. Furthermore, a research team used flying robots in a theater play; although communication was problematic for such robots, which lack a face or gestural ability, excellent expressive capability was realized in this study by pairing the robots with human actors [163]. Another study used a Wizard-of-Oz approach to assess how people felt about commanding a flying robot with gestures, finding some trepidation existed [164]. We believe that the closest resemblance to the flying humanoid robot we proposed may be found in the "Air Swimmers" toy, as well as "Flytech Tinker Bell" and "Flutterbye Flying Fairy", two flight-capable dolls which lack communicative degrees of freedom usually found in humanoid robots. Although we believe that merely possessing a face creates some increased capability for expression as compared to most current flying robots, we expected that better communication could be realized via a humanoid robot solution, which had not been built before.

We also investigated how such a flying robot could approach people. To predict how a robot can move in three dimensions around stationary and moving humans in a safe manner, some intuition was extracted from previous work addressing the two dimensional case in which people may walk but not move up or down. Hall, in introducing proxemics, described a circular model of interactive distance; entering inside the circle results in a qualitatively different interaction than when an agent remains outside [165]. Furthermore, Kendon observed that people do not tend to position and orient themselves randomly; rather, some positional equilibria can be considered to understand how people tend to position themselves relative to one another: "vis-à-vis", "L", or "side-by-side" [166]. Robotics researchers have also uncovered anisotropies in people's perceptions of distances; for example, people may wish a robot to be somewhere other than behind them [167], and approaching from their front sides may be less preferred than approaching from the side [168]. In regard to how a robot can move in conjunction with a moving person, Helbing and Johansson modeled the way persons moved during an evacuation scenario, describing a simplest form of Social Force Model (SFM), called the "Circular Specification" (CS) [169]. A slightly more complex specification, Collision Prediction (CP), was proposed to better model pedestrian motion in a non-emergency scenario [170]. However, none of these studies indicated how a robot can provide a safe and natural impression in flying near people.

Likewise, any other impressions conveyed by typical types of flying were unknown. Ekman and Friesen distinguished some categories of human motions by function which included, but were not restricted to, emotional expressions [125]. Although it could be possible to implement some flying motions corresponding to each of these categories, this would not resolve the fundamental question of how typical flight would be perceived. Several other studies suggested how some useful representative flying motions could be selected for investigation. For example, complicated three dimensional motions could be understood in terms of a sequence of primitive motion constituents [171]. Another work investigated people’s perceptions of motions which were varied in regard to three key descriptors: magnitude, velocity, and posture [172]. This latter work bore some slight similarity to the portion of our own work which explores kinesics. However, we did not investigate how a given motion could be varied to seem joyful or sorrowful; instead we sought to expose basic patterns in the meanings people attribute to flight motions. Thus, previous work did not indicate how to create a flying humanoid robot or how people will perceive such a robot’s flight motions.

1.4 Novel Contributions

The main novel achievement of the current work is describing how a humanoid companion robot can be designed to interact with people by recognizing behavior and responding appropriately within the context of enjoyment and affection, toward facilitating people’s social well-being. Each study described in the current dissertation has been presented before in a number of articles and conference papers; therefore the original contribution of the dissertation is providing an in-depth account of all studies in one place with some added details. Contributions for the individual studies are listed below organized based on sub-topic as follows:

*Study 1) *Recognizing enjoyment-motivated play:* a first recognition system which can recognize how people move a humanoid robot to play with it: based on a list of typical motion-based play touches performed toward a humanoid robot, the system uses heuristics, automatically selected features calculated on a short window of inertial data, and Support Vector Machines (SVMs) to realize good online performance.

**Study 2) Enjoyable strategy*: first guidelines for using recognition capability to realize an enjoyable motion-based play interaction by avoiding typical failures, based on analyzed feedback from participants and verified via experiments; as well as a new robot capable of providing enjoyment in such interactions, Sponge Robot.

* *Study 3) Recognizing affection-motivated play*: a first recognition system which can recognize people's affection in touches; we present a list of typical touches performed toward a humanoid robot for various typical reasons in an interaction; proposals for a mechanism for conveying affection through touch; analysis of useful modalities for recognizing touch, and a hybrid touch/vision system which can recognize all typical touches.

* *Study 4) Affectionate strategy*: a new approach for how a companion robot could use flight to approach people to show affection: we report implementation details for a first flying humanoid robot, Angel, as well as guidelines for safe flight and emotional cues conveyable through three dimensional motion (we call these new problems "z-proxemics" and "z-kinesics" respectively).

1.5 Organization of this Dissertation

Organization of remaining material is as follows. Chapter 2 describes four robotic prototypes built to investigate specific aspects of how a robot could support enjoyment and affection toward facilitating social well-being. Chapter 3 reports on recognizing touch-based play involving moving the full body of a small humanoid robot, and Chapter 4 describes how this recognition capability may be used to provide enjoyment. Chapter 5 describes recognizing affection through touches. Chapter 6 reports on an approach for proactively showing affection by approaching via flight. Chapter 7 discusses and summarizes results.

Chapter 2

Robots

"The animated figures stand adorning every public street

And seem to breathe in stone, or move their marble feet."

Pindar, in the seventh Olympic Ode, c. 522–443 BC describing the automata of ancient Rhodes

"Here I sit, forming humans in my image; a people to be like me,

To suffer, to weep, to enjoy and to delight themselves."

Johann Wolfgang von Goethe, in Prometheus, 1772–1774

To conduct the four studies described in the previous chapter—investigating recognition capability and a behavior strategy for enjoyment and affection—we required appropriate platforms. This was necessary to avoid "tunnel vision" which can arise when using a platform which may not be appropriate [119]. Therefore we developed and prepared four prototypes: Sponge Robot, Elfoid and Kirin, and Angel. Sponge Robot was used for Studies 1-2, recognizing play and providing enjoyment, because it could move like many toys. For Study 3, recognizing affection, high robustness was the important concern because we expected unaffectionate touching would involve striking or slapping a robot; for this we used Elfoid and Kirin. For Study 4, we required a platform which could approach a person to show affection; we built Angel because our other platforms could not do this. These prototypes are shown in Figure 5 and, the interaction scenario we envisioned for each is shown in Figure 6.

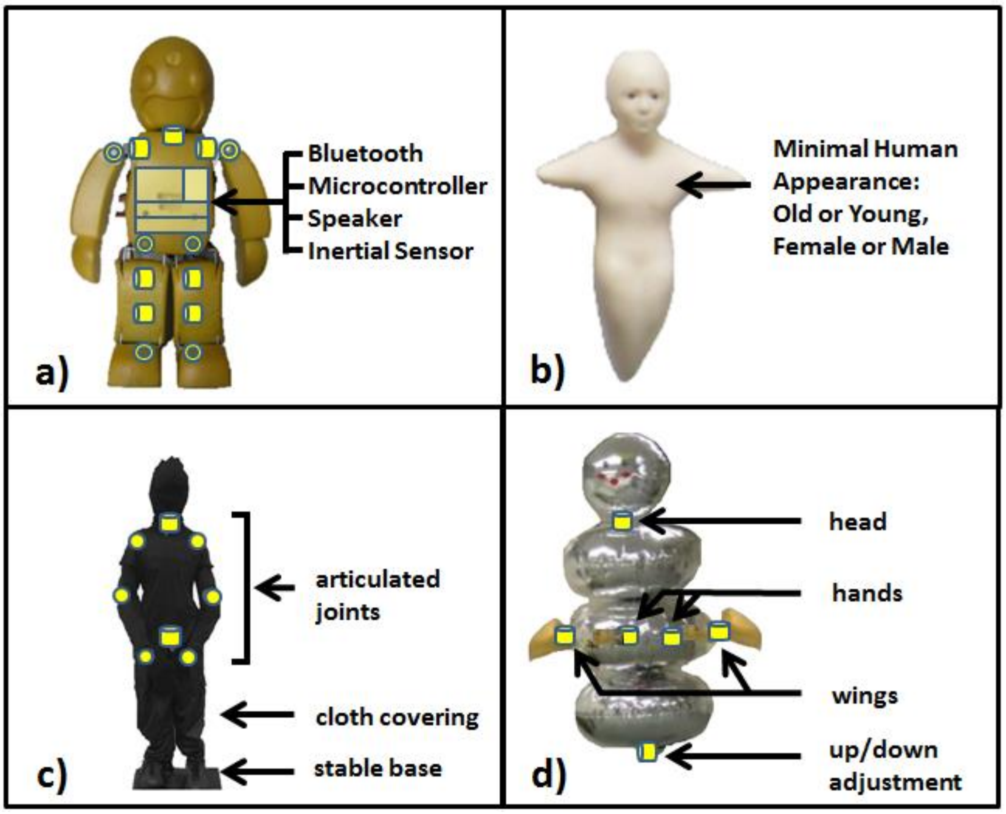


Figure 5. Robot prototypes used to investigate providing well-being.
a) Sponge Robot, b) Elfoid, c) Kirin, d) Angel. Degrees of Freedom and articulated joints are indicated as cylinders or spheres: cylinders indicate that rotation occurs about their long axis; spheres indicate rotation is possible in any direction.

Figure 6. Predicted interaction scenarios with the robot prototypes.
a) Sponge Robot's full body will be moved during play b) Elfoid will be touched while held and Kirin while standing c) Angel will be able to fly to approach a person.

2.1 Sponge Robot

Sponge Robot is a small light humanoid robot (37cm high × 20cm wide × 10cm depth; 1.4kg; comprising a head, body, arms, and legs), which was created by modifying a commercially available kit (“Robovie-X” produced by VStone Co., Ltd.). Figure 6 shows the interaction scenario we envisioned, with Sponge Robot responding to being raised upward in play by a human. In this scenario, the robot’s soft molded urethane covering first entices interaction, such as picking up the robot and moving it around. During this time, the robot’s acceleration and angular acceleration are measured at 12Hz by an inertial sensor (three-axis accelerometer, two-axis gyro) and sent wirelessly via a Bluetooth module to an external computer for processing. From this data, our algorithm detects motion-based behavior performed by a human, such as raising the robot. Recognition results are considered in planning the robot’s behavior in accordance to its interaction strategy. Possible actions include responding to a person’s behavior, performing a proactive behavior, or changing some internal state in the robot. For example, a happy response could be appropriate after being raised up. Commands are thus sent back from the external computer to the robot, which executes motions with thirteen degrees of freedom (one in its head, four in its arms, and eight in its legs) and plays back Adaptive Delta Pulse Code Modulation (ADPCM) sounds via a speaker located near its belly. To enact motions and sounds, power is drawn from batteries inside Sponge Robot, allowing it to be completely wireless.

2.2 Humanoid Robot Forms: Kirin and Elfoid

In observing how people show affection through touch, we did not want participants to be concerned about possibly breaking a robot. Therefore we prepared two mock-up humanoid robot “forms”, Elfoid and Kirin, which appeared to be robots but did not have any dangerous or breakable parts. We use the word “form” to state that they were mock-ups with the appearance of a humanoid robot but without actuators and electronics.

The first mock-up, Elfoid, was a soft creature-like form which did not have any breakable parts: it consisted of hollow Polyvinyl chloride [173]. Its appearance was intended to not seem old, young, female, or male.

A second humanoid robot form was constructed to be able to distinguish typical behavior toward a humanoid robot form from behavior which might be limited only to Elfoid. Kirin was built with a typical humanoid form (comprising a head, legs and hands), a height between average male and female human heights (168cm), and an indistinct appearance with specific features covered by dark fabric. We also ensured that Kirin's base was heavy and stable. Eight joints, two of which could only be rotated and six of which could allow motion in any direction, were used in Kirin's head, shoulders, elbows, hands, and waist to allow people to move the form as they wished. Thus, Kirin was fashioned to be soft, safe, and easy-to-touch.

2.3 Angel

To show affection, we proposed that a humanoid robot could get close to people, but our previous prototypes could not do this; therefore we built a new prototype called Angel which could approach people by flying over obstacles. In doing so, we built both flying and interactive capabilities. Flying was realized by using a lighter-than-air solution (a body comprising four helium balloons) and three degrees of freedom for locomotion: two flapping wings and a sliding center of mass (COM). Using an aerostatic solution allowed our robot to be soft and slow enough to be safe in flying around humans. Beating one wing caused the robot to rotate, and two wings could be moved to allow the robot to advance or retreat. The sliding center of mass (COM) was used, in conjunction with propulsion from the wings, for the robot to ascend or descend, as in the Air Swimmers toy. Stability was achieved by situating heavy components as much as possible at the bottom of the robot. Because this robot was developed as a proof-of-concept, our initial platform was quite large, but we intend to reduce its size in the future.

To enable interactions, Angel was created with a humanoid form comprising a head, arms, and base with three degrees of freedom. Angel could rotate its head to indicate a focus of attention, and rotate either or both of its arms for simple gesturing. Figure 6 shows our expected interaction scenario, in which Angel could fly over an obstacle to get close to a person and show affection. Through this configuration, Angel would be able to communicate in a more familiar manner than typical flying robots and also be able to fly over obstacles which could pose a problem to typical wheeled or legged humanoid robots.

Chapter 3

Recognizing Motion-based Play (Sponge Robot)

"In every real man a child is hidden that wants to play."
Friedrich Nietzsche, in Also sprach Zarathustra, 1883

After a prototype had been built, the next step was to use it in interactions to acquire new knowledge. This chapter describes work using our first prototype, Sponge Robot, a small light humanoid robot described in detail in Chapter 2, toward completing our first objective toward facilitating people's social well-being: of building a system which can recognize people's enjoyment-motivated play behavior. This material has been previously presented [1, 5-6]. A logical premise for this first objective rested on a number of related ideas:

- Social well-being may be facilitated by experiencing enjoyment.
- Enjoyment may be derived from playing.
- Playing often involves moving an artifact.
- The way in which an artifact is moved can be captured using an internal inertial sensor.

The first two points are supported by the literature described in Chapter 1. The third point follows from the many toys which people play with in a motion-based fashion: e.g., dolls (stuffed animals and action figures), ropes (swings, yo-yos), balls, and attachable pieces (building blocks and puzzle pieces). The fourth point can be illustrated by considering the capability of an accelerometer to capture an object's orientation due to the acceleration of gravity, and acceleration caused by an outside force (applied by a person) to move the object. The importance of realizing such a scenario is suggested by the user scenario in Chapter 1.2; the developed functionality will allow a robot like Gozilla to be able to play with a person like Masha.

We note that, as described in Chapter 1.3, we do not seek to directly recognize people's degree of enjoyment through their touches. This is because play is a free experience and people have various preferences. Rather we focus on the first step of gaining insight into the fundamental question of what behavior people perform.

Our approach using Sponge Robot involved two stages: (1) identifying how people attempt to play with a small humanoid robot and (2) recognizing people's behavior online and automatically. (How to use recognition capability to provide enjoyment is described in Chapter 4.)

3.1 Identifying Motion-based Play Behavior

Recognizing a person's behavior is useful because people should not be expected to always bother to study a manual before attempting to play with an artifact, or to have to conform to unnatural rules. By enabling playful interactions, recognition capability can also then be used to proceed to the next step of investigating how a robot can provide enjoyment. Thus, toward allowing for an intuitive enjoyable interaction, our first step was to identify how people attempt to play.

The challenge was that people's behavior is highly complex. We followed the basic spirit of a standard technique for identifying new theory, called the grounded theory method [121]. This method involves analyzing observed phenomena, and has been used in previous work to reveal what behavior should be recognized [120].

Thus, we gave our robot to 17 young adult Japanese, and asked them to freely play for approximately five minutes each. During this time, Sponge Robot's joints were locked, the robot did not move, and we did not make any objects available (such as balls). The reason the robot's limbs could not be moved was because our intent was to investigate how people would move a humanoid robot form, without expecting it to have some predefined number and arrangement of degrees of freedom. The robot was not made to move for a similar reason, because we wished to determine in general which affordances would be presented by a humanoid robot form; the use of motions would also have been complicated by the lack of any indication for which motions to use, the fact that humans usually restrict their movements when they are being touched (as in turn-taking interactions, because simultaneous activity may interfere), and the lack of a precedent in the related literature (other studies described in Chapter 1.3 also used a motionless robot to identify people's behavior). Likewise, objects such as balls were not presented because it would not be possible to provide every object, and because we were interested in the fundamental simplest scenario which does not rely on objects.

During these initial sessions, the experimenter recorded written notes in simple language, and video footage was obtained. The latter was verified after sessions to ensure the completeness of the notes. Codification entailed replacing natural language with short descriptive labels such as "Walk" or "Dance". In this way, much variation in physical behavior could be accounted for by each individual label. For example, "Inspect" was used to describe any one of various motions in which the robot was tilted to be better viewed.

The behavior we observed was highly complicated. One user for example in a short space of time rotated the robot in various directions, embraced it, moved it over to his lap, raised and lowered it, and rested it in a prone orientation. In the midst of such streams of complex movements however, we were able to determine some representative motion patterns. 13 *typical* motions were enacted by more than one participant; these are shown in Figure 7. Inspection of the robot from a number of perspectives was extremely common, in addition to raising or lowering the robot, resting it as if sleeping, and causing it to stand upright. Atypical behavior performed only by one participant was also noted: playing ball games, massaging, causing the robot to climb a wall or a person or hop, giving the robot first-aid and pressing an ear to its chest, using the robot to rub one's head, showing it to someone else, rotating it on its head, and playing hide-and-go-seek. Also some behavior which could not be detected with the inertial sensor was noted: some participants rubbed the top of Sponge Robot's head, squeezed its hand, hung their glasses or bag on the robot, or conveyed a greeting. This noting of many different patterns of play provided some support for our expectation that people would be able to play with our robot in the designated scenario.

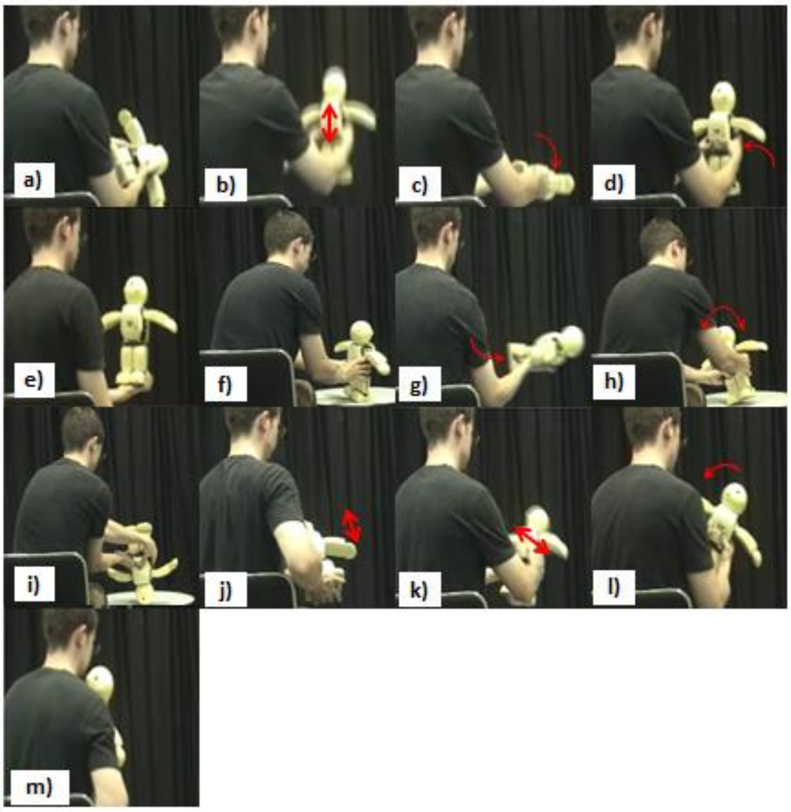


Figure 7. Typical full-body play gestures performed.
The proportion of participants who performed each motion is indicated in parentheses: a) Inspect: viewing the robot from various angles (100%), b) Up-Down: raising or lowering (94%), c) Lay Down: placing in a reclining pose (94%), d) Stand: causing the robot to be upright (88%), e) Balance: holding in an unsteady grasp, e.g., on one hand (64%), f) Walk: making the robot shuffle somewhere (59%), g) Airplane: moving the robot like a toy airplane (47%), h) Dance: making the robot perform a jig (47%), i) Upside-down: inverting (47%), j) Rock Baby: rocking in one's arms (41%), k) Back and Forth: shaking back and forward (35%), l) Fight: play-fighting, e.g., with the robot punching (35%), m) Hug: embracing (29%).

It was also noticed that some behavior was familiar in the context of playing with a small infant, such as "Hug" or "Rock Baby"; other behavior would appear more likely to be directed toward a toy or action figure, such as "Upside-down" or "Fight". Participants described in their feedback impressions of the robot as appearing like an infant or as merely a machine. These impressions are in keeping with previous work which illustrated a difference in the degree to which people are willing to suspend their disbelief in perceiving a robot as a lifelike being [37].

Thus, the predominance of the typical patterns indicated a high possibility that people will attempt these motions during a motion-based play interaction with a small humanoid robot such as we wish to enable. Therefore, we selected these 13 patterns as our classification target.

3.2 Recognizing Typical Motion-based Play Behavior

The next stage in our work involved gathering inertial data, devising a basic recognition approach, determining features useful for our case, and verifying that the developed system yielded sufficient performance.

3.2.1 Acquiring inertial data

First inertial data were obtained for the identified patterns. Because differing quantities of data for each pattern could impact recognition capability, we desired the same number of samples per motion; also, we wished to have data arranged in a convenient format. Thus, we asked 21 young adult participants to enact the 13 typical motions.

Sessions lasted ~15 minutes. During this time, participants sat on the floor with our robot in front of them; the floor was covered with straw mats called “tatami”, which are common in Japanese houses. After receiving simple written instructions with regard to the target patterns, participants were asked to perform them in a random order. Participants were not instructed with regard to how motions were to be enacted. Some scenes of participants touching the robot and inertial data are shown in Figure 8.

In addition to acquiring data, we expected that the robot’s own motions would interfere with correct recognition and we wished to gain some insight into the degree to which this would be a problem. Therefore we manipulated the factor of the robot’s own motions, generating four motion conditions.

Condition 1) Still: the robot did not move at all

Condition 2) Slight, constant motion: the robot continuously moved; random Gaussian noise was used to permute the values of the robot’s joint angles

Condition 3) Quick, unbalanced motion: the robot kicked out a leg and bent one arm to create a torque about its vertical axis, which could cause it to turn or fall

Condition 4) Large motion: the robot performed large abductive and adductive motions with its arms and legs which could interfere with a person’s grasp

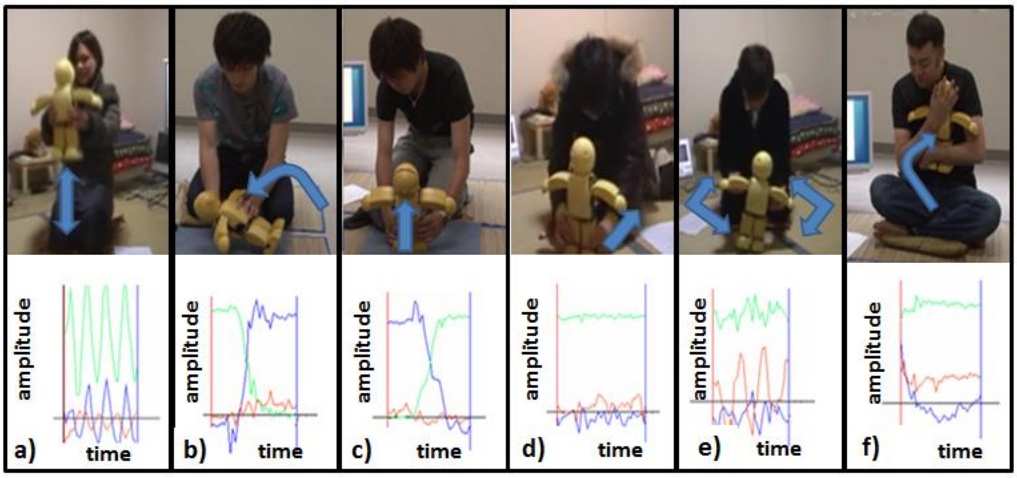


Figure 8. Gestures being performed and accelerometer data.
Gestures (top) with data (below): a) Up Down, b) Lay Down c) Stand, d) Walk, e) Dance, f) Hug; the accelerometer x axis is blue, the y axis is red, and the z axis is green.

In all conditions, the robot's limbs could not be bent, as in the first sessions. In the first two conditions, motion or lack thereof persisted the entire time; in the latter two conditions, the robot was commanded to move once during a person's rendition of each motion. The conditions are shown visually in Figure 9.

During the sessions, inertial data and video footage were acquired. To be able to determine when motions commenced or finished, the current time was displayed on a nearby computer monitor. Afterwards, this was used to label 1328 motions manually.

3.2.2 Basic framework for recognition

Online recognition required detecting play events and discriminating which manner of play was involved. Our approach involved making a judgement based on a stream of classification outputs updated each time-step, as well as some empirically determined heuristic thresholds. As a basis for recognition, we adopted a three-second-long scrolling data window, which was updated—a new inertial data point added and an old data point removed—every 80 milliseconds. The use of such a window does not mean that the system required three seconds to respond to behavior, as a decision can be made each time a new point is added. The recognition delay involves some tradeoff between accuracy and speed, which may be partially decided by the designer. For example, we could define the motion "Hug" as commencing just before Sponge Robot is first raised up until after the robot is released, thereby capturing much information which could raise

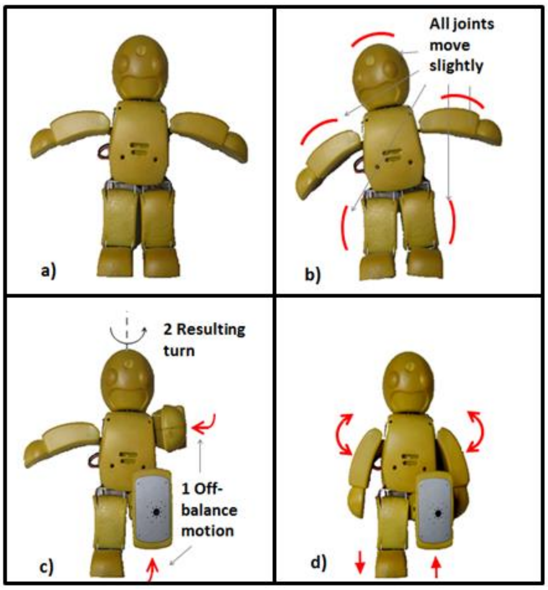


Figure 9. The four motion conditions investigated.
a) Sponge Robot does not move, b) all of Sponge Robot's joints move slightly and continuously, c) Sponge Robot's arm and leg quickly move so the robot loses balance and turns to its left, d) Sponge Robot waves its arms and legs up and down in a large motion.

classification accuracy; or, we could define it as just part of this, which could increase speed of recognition. We sought to preserve a reasonable balance.

As noted, our system performed classification of the data window at every time-step. For this, we employed Support Vector Machines (SVMs), a high-performance machine learning approach [174]. This method consists of finding a "margin" or hyper-plane boundary, between two classes in data space. The margin is optimized by maximizing the distance between data of heterogeneous classes while permitting some relaxation to account for an occasional error. A quick calculation can then be used to determine on which side of the margin a new data point falls, where the margin itself is represented using some important points called "support vectors". Some judgements were required to employ SVMs with our problem in regard to multiclass classification, "kernel" function and hyper-parameters. As SVMs intrinsically handle a binary classification, typically multiple SVMs are used in a "one-vs.-all" or "one-vs.-one" configuration; the former approach involves utilizing one SVM per class, and the latter one SVM per every two classes. A kernel function renders some problems which are linearly inseparable in lower dimensions solvable by projecting the data into higher or infinite dimensions. Hyper-parameters C and γ define the extent to which the margin can handle misclassifications (a penalizing cost parameter) and the degree of over-fitting or under-fitting respectively.

We used a free downloaded library, LIBSVM ("A Library for Support Vector Machines") in conjunction with a one-vs.-one approach for high accuracy, and a Radial Basis Function (RBF) for advantages in dealing with linearly inseparable problems, avoidance of some numerical difficulties, and the few hyper-parameters which require determining [175]. Our program called a free tool contained with the library to conduct a grid search to find C and γ per fold during cross-validation; values of 8 and 0.5 were obtained for all data.

Detecting behavior embedded in the continuous time stream required some additional calculations. Specifically, the system had to recognize only when confidence in the stability of the signal was sufficient, and garbage data involving no playing by a person had to be rejected. We found that, although complex solutions have been proposed in the literature, in our case a simple solution sufficed. For the first problem, some thresholds were used to verify that a class label had been recognized a certain number of times, thereby "stabilizing the stream". For the latter problem, we found that the motion pattern classes recognized naturally acted to distinguish active from passive playing; for example, hugging or dancing are more active than merely inspecting or balancing the robot. Because sometimes the robot's own motions could affect the inertial data in such a way that it appeared as if a person were interacting, we added a second check. This involved recording data while the robot performed its motions in various orientations to find codebook vectors and thresholds. We found these measures allowed Sponge Robot to reliably determine when a person was interacting.

3.2.3 Features

It was desirable to find some useful features to reduce the size of the data and prevent over-fitting, generate accurate results, and gain insight into the underlying nature of the data, but it was not clear from the related literature which features should be used. We first compared some different types of features to find a good set, before selecting individual features from the latter.

For types of features, we examined time-domain, frequency-domain, and time/frequency-domain features. Salter *et al*. for Roball had used time-domain features such as mean values [106]. Lee *et al*. for Huggable had described frequency-domain features, Discrete Fourier Transform coefficients [99]; the applicability of such features was also indicated by repeated patterns in some motions like walking, dancing, or shaking. Speed, ease of implementation, and the capability of capturing qualities of both domains suggested the advantage of using Haar Transform coefficients as time/frequency-domain features, which have been used elsewhere in recognition studies [176].

For time domain coefficients we calculated means, standard deviations, changes from start to finish, medians, minimums, and maximums for each inertial data axis. Standard deviations were calculated as follows:

$$\mathrm{sd} = \sqrt{\left(\frac{1}{n-1}\sum_{k=0}^{n}(x_k - \bar{x})^2\right)} \tag{1}$$

Where x_k is a data value, $\bar{x}$ is the mean, and n is the number of data points in the window for an inertial data axis. For frequency-domain features, we calculated the Discrete Fourier Transform coefficients as follows:

$$c_j = \sum_{k=0}^{n-1} x_k(cos\theta - isin\theta),\ \theta = \frac{2\pi jk}{n}, j = 0, \dots, n\text{-}1 \tag{2}$$

Where c_j is the jth coefficient and n is the number of points for one inertial data axis. For the time/frequency domain, we calculated Haar Wavelet features using the following two equations:

$$c_i = \frac{s_i - s_{i+1}}{2} \tag{3}$$

$$ave_i = \frac{s_i + s_{i+1}}{2} \tag{4}$$

Where c_i, a wavelet coefficient, is a difference between two neighboring values at the designated scale, s_i and s_{i+1}, and ave_i is the average of these values.

A “wrapper-based” algorithm using SVMs and cross-validation was employed to rate each kind of feature. Time-domain features yielded a highest accuracy score, as shown in Table 2. We believe that frequency domain features suffered from the fact that cyclic motions were performed at various inconstant speeds involving accelerations. This problem may have also reduced the accuracy

Table 2. Three different kinds of features compared

Domain	Type	Previous work using such features	Cross-validation accuracy (%)
Time	Statistics (averages, etc.)	Roball [92]	74.3
Frequency	Discrete Fourier Transform coefficients	Huggable [93]	62.1
Time/Frequency	Haar Wavelet Transform coefficients	Object detection [164]	51.4

attained using the Haar Wavelet features, which moreover could not capture differences between even and odd groupings.

From the time-domain group our algorithm automatically selected features as follows. The full set consisted of five values for each feature (one for every data axis, which meant three for the accelerometer and two for the gyro sensor). Removing groups of features (in each case, three accelerometer-based values or two gyro-based values) produced a final set of 19 features with a small improvement in recognition accuracy, as determined by cross-validation. This result is shown in Table 3.

We believe the mean features are useful for classifying motions which involve much movement in one axis like laying down or standing. Standard deviation values further helped to identify cyclic motions for which mean values could be near zero such as cyclic x axis motion for walking, cyclic y axis motion for dancing, and cyclic z axis motion for moving the robot up and down. Changes from start to finish could help to distinguish between motions such as standing and laying down, in which mean values might be equal, and detecting when the robot was turned upside-down. Gyro maximums could capture large and possibly sudden motions found in fighting and vertical or horizontal shaking. Medians were useful in finding the true centers of distributions when means were affected by noise in the data.

The completed recognition system is shown in Figure 10. To summarize, at each time step, the current window of inertial data was checked using heuristics to confirm that interaction was taking place; if so, the 19 features were calculated, and the human’s play behavior was classified using the SVMs. Finally, confidence in the output was checked, after which results could be passed to the robot’s behavior strategy.

Table 3. Features selected automatically by our algorithm

Type of Statistics Feature	Type of Inertial Data	No. of Features
Mean	accelerometer x, y, z	three
Standard Deviation	accelerometer x, y, z; gyro x, y	five
Change from start to finish	accelerometer x, y, z	three
Median	accelerometer x, y, z	three
Minimum	accelerometer x, y, z	three
Maximum	gyro x, y	two

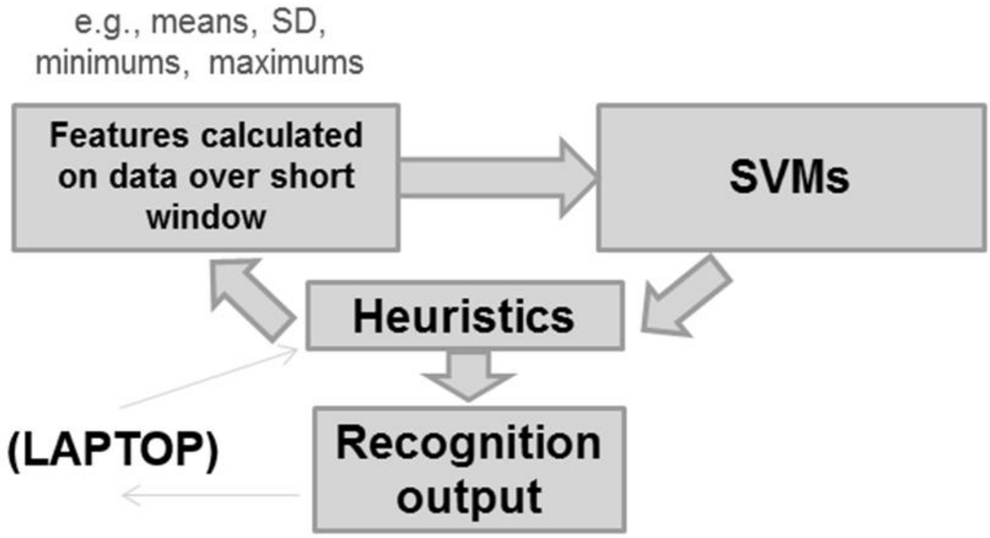


Figure 10. Developed system for recognizing enjoyment-motivated play. Inertial data from the robot is sent via Bluetooth to an external laptop, then the data is checked using heuristic thresholds to detect interaction, statistical features are calculated over a short window of data, SVMs are used to classify what the person is doing to the robot, a second set of heuristics is used to ensure confidence in the result, and finally output is remitted to the robot's behavior strategy to decide what the robot should do.

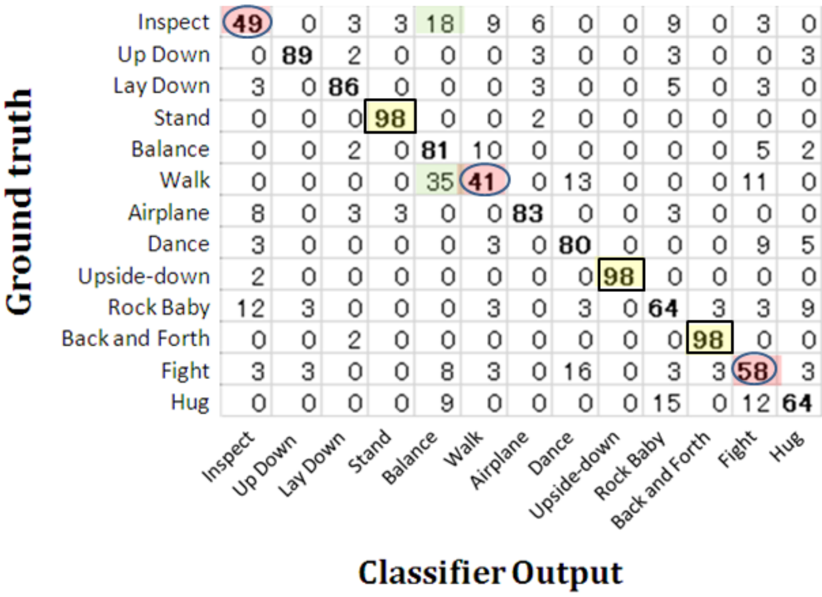
3.2.4 Recognition performance

It was desirable to ensure that recognition performance was sufficient to allow for motion-based play. For this purpose, performance analysis was conducted using the metric of accuracy, defined by dividing the number of true positives by the number of samples. Accuracy, although intuitive to understand, may yield skewed results in some cases when the number of samples per class differs greatly; in the current case, we ensured that the number of samples was the same, which allowed accuracy to be used as an evaluation measure. Leave-one-out cross-validation, a typical form of cross-validation, conducted on the non-motion dataset showed that the 13 target patterns could be recognized together with an accuracy of 77%.

Results are presented in the form of a confusion matrix in Figure 11. Easy-to-recognize, hard-to-recognize, and confusing motions are indicated in outlined, darkly highlighted, and lightly highlighted cells respectively. Some motions such as Back and Forth, Upside-down, and Stand could be classified with high accuracy. Others such as Fight, Inspect, and Walk presented more difficulty. One source of confusion was that Inspect and Walk were sometimes confused with Balance.

Confusion resulted from the large variance in participants' interpretations of how to enact motions, as shown in Figure 12. Some people simply pushed the robot or made it glide, thus leaving only a slight inertial footprint; for Inspect, users performed many different motions, including simply placing the robot on the ground. This is why in some cases these motions were mistaken for Balance. Also shown are some ways participants chose to enact Hug and Fight.

We also conducted analysis on the data in which the robot itself moved. Comparison was necessary because robots should move during an interaction, and if such movements interfere too strongly with the recognition process a different approach would be required. Direct comparison was difficult because the motion dataset was several times larger than the non-motion dataset, which could act as a confounding factor. Therefore, random sampling without replacement was performed to make the two sets the same size. This was conducted ten times and an average accuracy calculated. As a result, we found that accuracy was reduced to 56% (a 21% reduction from 77%). Lowered accuracy was due to quick and large motions interrupting people's motions and adding noise to the detected data. Consequently, motions which leave a fine trace in the inertial data such as Walk, Hug, and Balance appeared similar to high-energy motions such as Fight; drops in accuracy for these three gestures were -33%, -40%, and -74% respectively. An example of how the inertial data changed is depicted in Figure 13. A motion by the robot in this example almost caused a participant to drop it while performing the motion, Stand, creating multiple jagged peaks in the inertial data.



Ground truth \ Classifier Output	Inspect	Up Down	Lay Down	Stand	Balance	Walk	Airplane	Dance	Upside-down	Rock Baby	Back and Forth	Fight	Hug
Inspect	49	0	3	3	18	9	6	0	0	9	0	3	0
Up Down	0	**89**	2	0	0	0	3	0	0	3	0	0	3
Lay Down	3	0	**86**	0	0	0	3	0	0	5	0	3	0
Stand	0	0	0	98	0	0	2	0	0	0	0	0	0
Balance	0	0	2	0	**81**	10	0	0	0	0	0	5	2
Walk	0	0	0	0	35	41	0	13	0	0	0	11	0
Airplane	8	0	3	3	0	0	**83**	0	0	3	0	0	0
Dance	3	0	0	0	0	3	0	**80**	0	0	0	9	5
Upside-down	2	0	0	0	0	0	0	0	98	0	0	0	0
Rock Baby	12	3	0	0	0	3	0	3	0	**64**	3	3	9
Back and Forth	0	0	2	0	0	0	0	0	0	0	98	0	0
Fight	3	3	0	0	8	3	0	16	0	3	3	58	3
Hug	0	0	0	0	9	0	0	0	0	15	0	12	**64**

Figure 11. Confusion matrix for recognizing the 13 typical play touches.
The three highest true positive scores are shown in yellow inside rectangles, the three lowest accuracy scores are shown in pink inside circles, and the greatest sources of confusion are shown lightly highlighted in green.

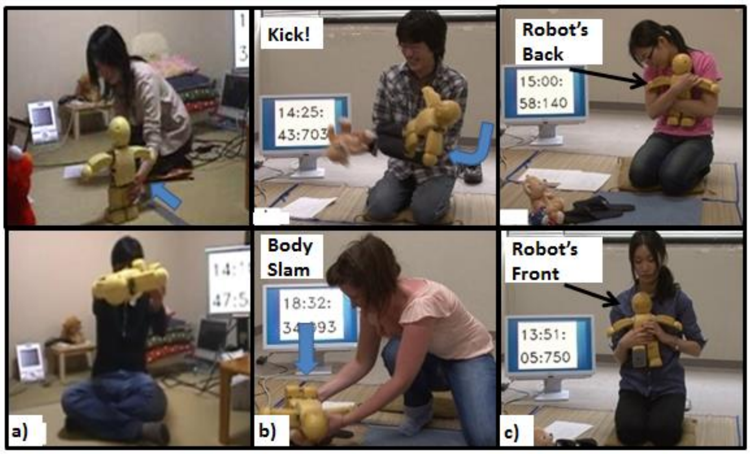


Figure 12. Challenging variety in the ways people chose to perform motions.
E.g., a) some participants pushed or glided the robot over the floor instead of moving its feet forward one at a time for Walk (top) and rotated the robot in any direction for Inspect (bottom), b) moved the robot as if it were kicking (top), or wrestling for Fight (bottom), and c) and hugged the robot's front (top) or back (bottom).

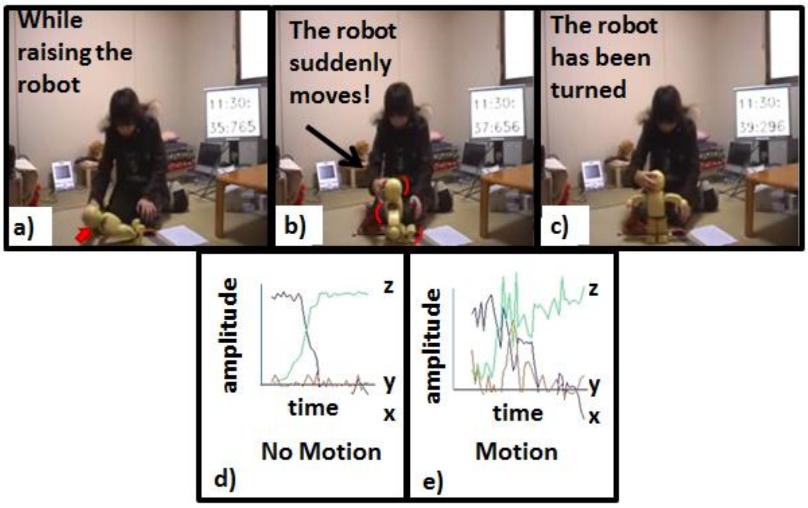


Figure 13. Effect of a robot's motion on inertial data.
In addition to a person moving, a robot itself is expected to move; however this can affect recognition accuracy. For example, for Stand: a-c) a robot's motion caused a person to almost drop the robot, which d-e) had a clearly visible effect on the inertial data.

Although accuracy was affected, the influence of the robot's motions did not forbid us from continuing with our approach because we did not intend to include motions aimed to interfere with a user's grasp, and complete accuracy was not necessary for a robot to play with a person; it was sufficient if the robot seemed consistent and reactive.

3.3 Summary

In summary, people typically attempted to move our small humanoid robot in 13 ways, in order to play. We were able to design a system capable of distinguishing people's play motions with 77% accuracy by calculating statistical features over a short window of data and classifying the set of features via Support Vector Machines (SVMs); empirically determined heuristics aided the reliability of online recognition results.

Chapter 4

Providing Enjoyment via Touch-based Play (Sponge Robot)

"It is requisite for the relaxation of the mind that we make use, from time to time, of playful deeds and jokes."
Thomas Aquinas, in Summa Theologica,
quaestio 168, art. 2, 1265–1274

The goal of the second study of this work was to provide enjoyment; here we report material which has been presented previously [2, 5-6]. Recognition capability alone, described in Chapter 3, was insufficient to provide enjoyment; a strategy was also required to make use of the recognition results. The challenge was that people's perceptions of what is enjoyable are highly complex. Our approach involved (1) observing people playing with a robot with a simple strategy to identify failures and how they could be avoided, (2) conducting an experiment to find some important guidelines for avoiding failures and providing enjoyment, and (3) verifying that using the guidelines resulted in providing more enjoyment than our initial strategy. To do so, we once again used Sponge Robot, described in detail in Chapter 2. By following this approach, we aimed to prepare a robot such as Gozilla for responding enjoyably in a touch-based interaction with a person such as Masha, as in the user scenario described in Chapter 1.2.

4.1 Learning from Failures with a First Strategy

To start we developed some behavior for our robot comprising motions and sounds, and used a simple first strategy to map recognition results to behavior;

this strategy was built based on incorporating intuitive knowledge from the literature into an initial design. Thus, a form of case-based reasoning informed our initial effort to design an enjoyable system [177]: we sought to incorporate some heuristics from previous cases, observe system performance, and then revise the system according to requirements specific to the case of motion-based play.

4.1.1 First strategy

Heuristics observed in previous work were followed, as outlined in Chapter 1.3. The interaction was formulated in a minimal turn-based fashion, and behavior was designed which fit with Sponge Robot's child-like semblance. For the former requirement, Sponge Robot was set to perform proactive behavior to suggest a manner of play, then respond with crying or laughter, depending on a person's behavior. Suggestions were selected based on a person's most recently performed behavior to entice enjoyable repeated patterns. "Response" means robot behavior following and caused by a person's recent behavior; "suggestion" means a spontaneous, proactive robot behavior based on the interaction history. Figure 14 shows this difference graphically. Crying and laughter were restricted to a single sound for each. When a person did not interact, Sponge Robot was caused to perform an idling motion, suggesting its lifelikeness. Before playing, participants were only told that they were free to interact as they wished, and that Sponge Robot could detect motions involving its full body.

To find appropriate behavior corresponding with Sponge Robot's appearance, we examined other humanoid robots such as E.M.A., Jingle Bell Rock Santa, Robosapien, QRIO, and Tickle Me Elmo, which led us to include common motions such as walking and dancing. Motions were constructed to appear infant-like, moderate in amplitude, and smoothly flowing.

4.1.2 Revising the first strategy

The first strategy was tested by observing some interactions between young adult Japanese and our robot. We found that providing enjoyment was not trivial, and that sometimes participants did not experience enjoyment. Examples are shown in Figure 15. In these example cases participants did not understand why a robot was moving the way it was, interpreted negative meanings, couldn't get a response to hand-waving, and tilted the robot without a goal.

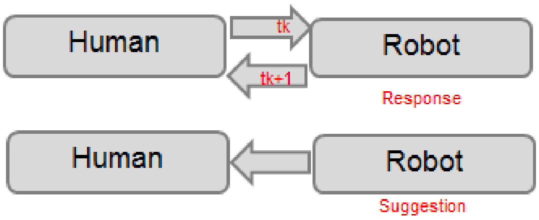


Figure 14. Responses versus suggestions.
Sponge Robot performed responses when a person's behavior was recognized (top), and suggestions elsewise (bottom).

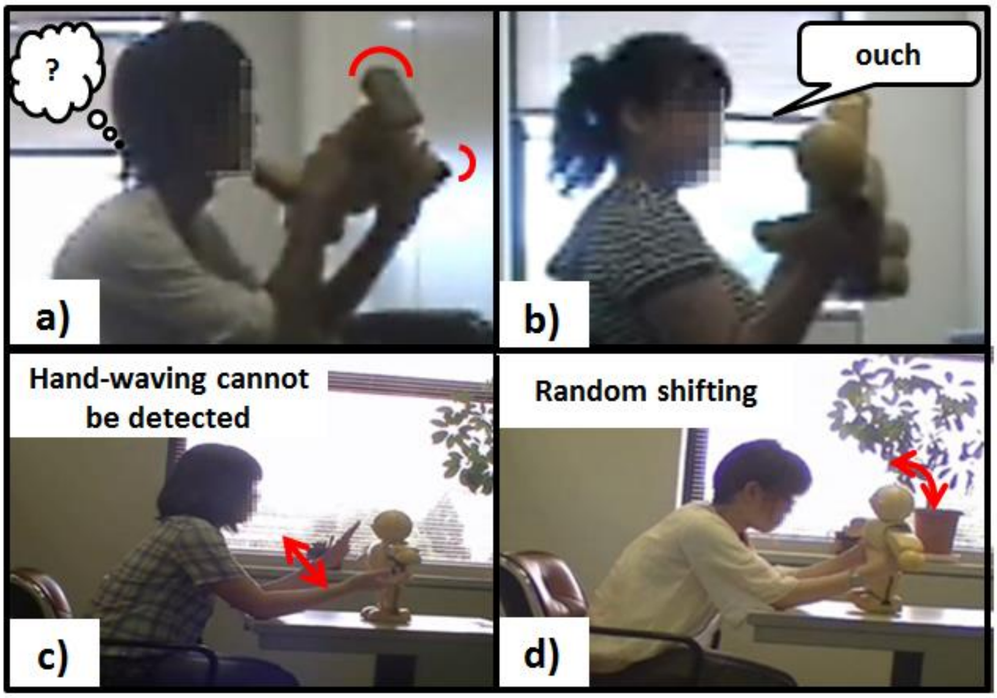


Figure 15. Some examples of failed interactions.
a) the robot's movements appear to be without meaning, b) the robot's motion seems to have a negative meaning, c) the robot cannot recognize a person's hand gesture, d) a person cannot think of anything to do with the robot.

One problem is that participants' comments were not consistent in regard to their views on how to fix the design. Participants indicated different modalities and functionality which could be used, and were not decided on, for example, whether the robot should be more energetic or more peaceful. What could be identified more clearly was when interactions failed. Therefore we adopted a pattern finding approach to determine typical failing patterns, queried participants with regard to causes, and came up with solutions. Such an approach had been discussed and applied previously in HRI [122-123].

A bottom-up approach was used, first identifying specific failure incidents, then grouping these into seven initial patterns, and finally constructing three higher-level categories based on similarities: “Meaningless motions” involved problems in the robot’s behavior; “Robot ignores me” was due to sensing problems; and “Just moving the robot” could result from either. The typical failing patterns and their causes are described in detail next.

4.1.2.1 Meaningless motions

The intended meaning of some motions was not understood. These motions caused participants to think that Sponge Robot had not noticed them; one observation was that the robot seemed like a moving insect paying no attention to them. Sounds reminded of cat calls even though they had been obtained from real children on YouTube, did not demonstrate enough variation, and were not noticed. Also, the outcome for some motions was unpleasant; for example, a pushing motion which occurred when a participant tried to hug the robot. The cause was difficulty in interpreting messages communicated only by motion, and lack of experience with interacting with robots. For the former problem, we asked one person to merely watch motions and state what the robot was doing; his guess was only in-line with our intentions one time out of four. Walking appeared to show a desire for assistance. Dancing appeared to be vehement speaking accompanied by gestures. Hugging, in which the robot inclined itself forward and lifted its hands, appeared to express a desire for reclining or hand-shaking. For the latter problem, lack of experience with robots was a cause; participants were not aware that the robot could respond, suggest, laugh, and cry (for the latter two actions, facial degrees of freedom not present in Sponge Robot could have aided clarity of communication). Motions which were interpreted negatively included responses which distanced (like a push or a kick) and suggestions which did not succeed at showing a positive intention in the robot to play.

4.1.2.2 Robot ignores me

Some participants hardly moved Sponge Robot, e.g., merely waving their hand in front of the robot’s face, or lifting the robot only very slightly or very slowly. The result was that the robot did not react to their behavior, which was not enjoyable. The reasons for this failure related to reactions, suggestions, and instructions. Reactions when noted did not appear more rewarding than the robot's idling movements: participants did not perceive a reason for them to actively interact. Suggestions from the robot did not help participants to conceive of another enjoyable way of playing. Instructions had to mention not only what the robot could sense, but also that the robot could not recognize haptic, visual, or aural stimuli. Some participants who did move the robot, but only slightly, were afraid of breaking the robot or themselves being injured by some quick movement. The reasons for this feeling stemmed from anxiety which some people experience at the thought of playing with a robot [178], apprehension which some people feel toward handling a moving creature such as a dog or cat, and simple lack of being accustomed to playing with a robot. The robot's responses and suggestions failed to convince people that there was no reason for worry and illustrate that they could move the robot with quickly or in a large fashion; in particular, because responses were rarely observed, the cycle of merely watching the robot was not easily halted. Instructions likewise failed to successfully communicate that the robot was safe and would not break.

4.1.2.3 Just moving the robot

Sometimes interactions even failed when Sponge Robot was moved adequately; this involved moving the robot in a fragmented and ambiguous way. For example, the system had trouble when people halted a motion midway then continued. Other times participants constantly moved Sponge Robot without looking if the robot was performing some motion. Again reasons for the failure involved responses, suggestions, and instructions. Perceiving no meaning in motions meant that participants felt no impetus to change the way in which they were seeking to play. Suggestions could not be comprehended or were not followed. Instructions provided no assistance. Importantly, participants did not understand the simplified turn-based style of interacting: that they were supposed to respond to the robot's suggestions and observe its responses.

4.1.2.4 Comparing results

In addition, comparison was conducted with the results of another HRI study [123], in which a large humanoid robot approached people. We found that *Robot ignores me* could be represented by “Unreachable” and “Unsure” in the latter study, two patterns in which a robot does not interact. *Just moving the robot* corresponded to “Unaware”, a pattern in which a robot did not receive attention. We did not observe a correspondence with one pattern, “Rejected”, which occurred when a person did not wish to interact, due to the nature of our study (all of our participants were required to interact). Thus, despite the large difference in the topics addressed by the current dissertation and this other paper, considerable overlap was encountered, indicating that similar failures could occur in various settings, and that guidelines for dealing with them were required.

4.1.2.5 Guidelines

Thus, four categories of reasons contributed to each failure: motions, responses, suggestions, and instructions. For each category, we came up with guidelines to prevent failing points based on user comments, advice in the literature, and our own opinions. To avoid problems with being too general or too specific (missing important information or detailing specifications in a fashion which was too complex to be helpful) we formulated the guidelines as general ideas accompanied by specific application details.

Guidelines apply to development at the system and at the component level, and deal with a robot’s behavior and interactive strategy (1-3), as well as instructions conveyed before interactions (4). Novelty varies as follows. Some proposals require justification and were verified via an experiment (ways for the robot to offer a perception of reward or exert influence over the manner of play). Other statements are for the most part evident but might not be noticed during the design stage (for example, linking easily triggered recognition results to every major body part of a robot). Last, some items are clear but important to consider (for example, responding with good timing). Guidelines are summarized generally below, with italics indicating topics which required testing to verify:

- **Meaningful motions**: Semantics testing, sound stream pairing, set-level assessment, timeliness, and non-repetition
- **Rewarding Responses:** Large, positive; *maximum or progressive*

- **Inspiring Suggestions:** Indicating robot's capabilities and possible play objectives, timely, incomplete, cancellable**,** *shifting or persisting*
- **Fulfilling Instructions** Disclosing non-evident capabilities (sensory-behavioral and robustness)

Specific details explaining guidelines are provided next.

4.1.2.6 Meaningful Motions

To provide enjoyment, it appears as if it is not enough to only string together enjoyable motions: the motions presented over the course of an interaction should also be understood. If motions are not understood, people may cease to bother observing what a robot is doing. This problem is complicated by the lack of a simple scientific way to differentiate intelligible and unintelligible motions. The approach we advise is heuristically guided. We propose that motions should be (1) based on a robot's affordances, (2) matched with sounds, (3) tested for communicative clarity with naïve users, (4) evaluated for set-wise appropriateness by the designer, and (5) performed at appropriate timing in a non-repeating fashion.

In detail, the process of grounding motions in meaning can be first facilitated by designing them based on a robot's affordances; for example, we expect a swimming motion will be easier to recognize as such in a robot with fins and scales than one with wings (in which case it might appear like flying). Motions in living creatures such as humans or animals are often accompanied by sound (e.g., laughing, snoring, yawning, or crying); people will expect such motions in a robot to also be audible. Perhaps most importantly, motions should be tested with naïve users, who should be asked to simply state what a robot is doing. Motions which are not recognized should be redesigned. As well, a designer should evaluate motions expected to occur together in an interaction as a set (e.g., different forms of walking or dancing); in doing so, attention should be paid to three factors in particular, as follows. (1) Is each motion associated with a unique meaning readily discerned from other physically similar motions? (For example, can walking motions be clearly distinguished from dancing motions?) (2) Are motions with similar meanings physically close in all aspects except what makes them unique? (A robot's movements may be faster when running than walking, but causing the robot to bend forward or smile in one case and not the other could convey some unintended meanings.) (3) And, is the set of similar motions together complete in the sense that the designer can convey all the messages which should be conveyed? (For instance, is there a happy or sad walk; a slow and fast step; or a left and right step and transitions from a neutral pose, if these are required?) A closely related concern to the last item is: can a person readily cause the robot to move all of its major limbs? If a robot has arms or legs which appear movable, people will expect to see them move. In terms of rendering motions meaningful at the systems-level, people will expect a robot to move with a human-like delay in response to some stimulus or new intention. Motions which are repeated often also may lose meaning over time, due to the phenomenon of habituation. We recommend that motion patterns be varied or staggered to avoid such an effect. One way in which this problem could be approached is to design motions using motion primitives [114].

In our robot, motions were designed which correspond to touches whose meanings could be easily understood. Persons asked to observe the motions correctly stated what the robot was doing 90% of the time. We also videotaped and watched semantically related motion sequences to determine how individual motions would be perceived in context. Laughter and crying sounds were paired with enthused wiggling and recalcitrant writhing respectively for clear communication; sound volumes were selected to be easily heard but not frightening. We also verified that people found the laughter motions provided enjoyment.

4.1.2.7 Rewarding Responses

A robot’s motions not only have to be meaning; responses should be such that an attempt by a person to interact should also be perceived as rewarding. Otherwise, with no incentive, people will not interact. In the case where a robot is primarily designed for active interaction, enjoyment then becomes difficult to provide. Participants’ comments suggested two possible vehicles for providing reward. Exaggerated, big responses which occur quickly present the impression that a person exerts much control over the robot and the interaction. Appropriately positive-seeming episodes may also be perceived as rewarding, such as if a robot demonstrates enjoyment when a user plays in a gentle way with it. Such a conceptualization by itself however is not sufficient for structuring responses throughout an interaction. At the most basic level, it was unclear if all responses should be big and happy. As described previously in the discussion of related work in Chapter 1.3, toys designed for children such as Tickle Me Elmo present mostly big stimulating responses. The appropriateness of such an exaggerated response schema for robots with restricted communication capability is supported also by previous work [179]. However, therapy robots such as Paro exhibit a range of small and big responses which would seem to be somehow structured to fit with a person's behavior.

Therefore, we implemented in our robot two different strategies for providing reward through responses: "maximum reward" and "progressive reward". The former strategy always entails presenting big happy responses: the robot wags its legs, arms, and head substantially and laughs loudly and happily. The latter strategy involves presenting small noncommittal responses at first which grow in magnitude and positiveness: first our robot wags only its legs or arms, then all appendages, and at last all appendages and its head, while its laughter becomes progressively louder and happier. The last response of the latter strategy is identical to the maximum response employed in the former strategy.

4.1.2.8 Inspiring Suggestions

Successful use of an interface is facilitated by knowing how to use it. In the belief that people should not be made to read a manual to interact with robots, we believe that a capability to provide advice on how to play like a recommender system, or instruct people like a tutor in regard to what can be done to a robot, could be highly useful.

A fundamental difficulty arises in ensuring people can recognize a robot's suggestions for what they are. To distinguish suggestions from responses, we suggest that the former should be "discordant" in some way which elicits a person's behavior; they can be incomplete or small. In this way, the undesired case in which a user simply observes a system designed for active interaction and does not experience it in the way the designer has intended may be discouraged. Likewise, suggestions should not interrupt when a person is seeking to interact.

Toward this, on the systems-level, a robot can suggest when active interaction is not occurring. Suggestions can then be speedily terminated if a person starts to interact, as people might not always wait for the robot to finish what it is doing. However some challenges exist. People might not be interested in what the robot is suggesting; therefore, a robot should be able to suggest not just one, but a variety of possible ways to play. Also, if suggestions are made to be smaller than responses it may be difficult to ensure that they are noticed. As a possible remedy, we expect that the most easily comprehended suggestions may be presented first, especially those which are associated with easily perceived gestures, and repeated more than once.

Given this framework, we were not sure which aspect of the suggesting process should be emphasized: showing a variety of possibilities, or ensuring each possibility was clearly conveyed. This led us to implement in our robot two different strategies for structuring proactive suggestions to facilitate enjoyment: "shifting suggestions" and "persisting suggestions". In the former strategy, our robot quickly attempts to convey a large range of possible play behaviors in random order. In the latter strategy, our robot presents each suggestion several times in an easy-first order (Up-Down, Lay-Down, Dance, Walk, Hug). The robot is deemed to have repeated a suggestion sufficiently if two minutes go by, if the robot is turned upside-down for a substantial time (indicating a person is not satisfied), or if all of the robot's responses related to that suggestion have been observed. When the robot deems it has repeated a suggestion sufficiently it signals a change of intention by attempting to shake hands. Given the demographic targeted by our work, genuflection could have been another option for a signal. We chose hand-shaking because it involves touch and moving the robot; thereby, a user's acknowledgement could be detected by the inertial sensor. For both strategies, attempts by a person to actively interact were identified via some empirically determined thresholds and codebook vectors representing common orientations of the robot. Figure 16 shows some suggestions from our robot.

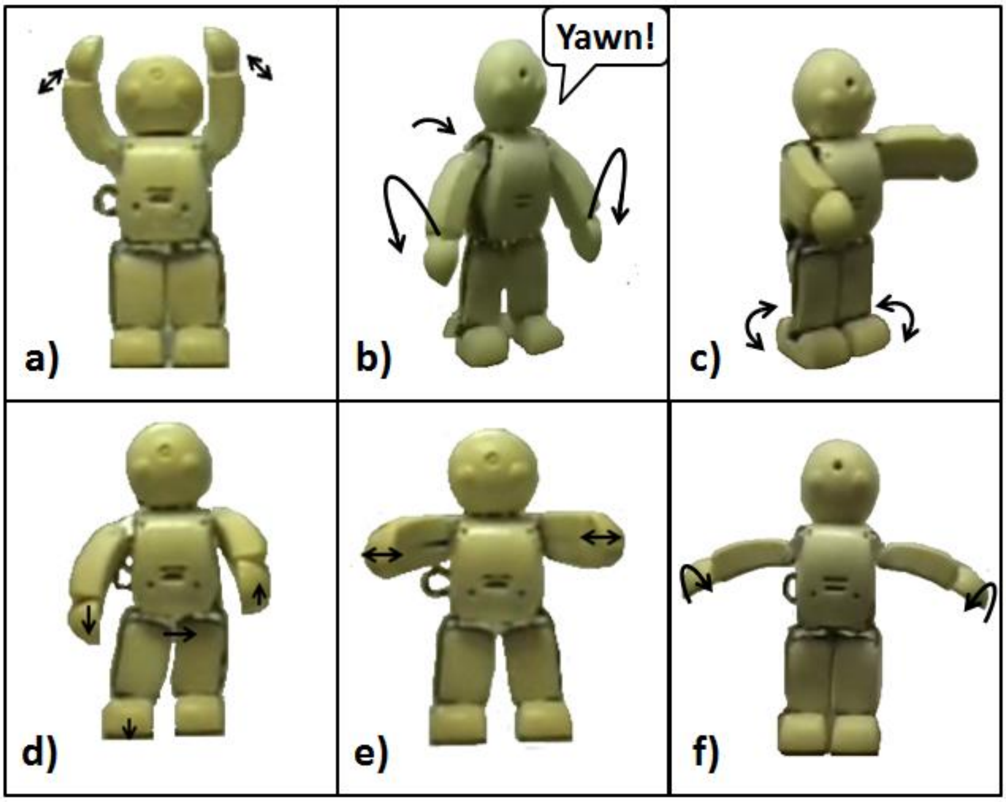


Figure 16. Some suggestions by our robot.
a) Up Down: the robot raises its arms upward and wiggles its hands b) Lay Down: the robot bends forward sleepily, then straightens itself and yawns moving its arms largely, c) Dance: the robot does a dance step in a ballroom dancing pose, d) Walk: the robot shifts its weight and moves a foot forward a bit while the arms move in an opposite direction, e) Hug: the robot opens its arms wide held out to a person, closes them, then reopens them, f) Stand: the robot turns its hands toward the floor and tries to push itself upward.

4.1.2.9 Fulfilling Instructions

People in the current day are not used to interacting with robots and may possess many doubts, apprehensions, and assumptions which cannot all be addressed by a robot during an interaction at zero acquaintance. Some important information should be communicated via a few minimal instructions such that interactions progress smoothly, although, as noted, we do not expect that people should have to read a manual. The information which should be conveyed is readily found from asking participants themselves what they wished they had known before interacting and what they would say to the next person to interact with the robot. For our robot, we learned that a need existed to explicitly state that the robot was strong and would not be damaged by regular interaction, that the robot’s behavior comprised both proactive and reactive components, and that the robot could not see, hear or feel touches which did not involve motion.

4.2 Completing the Guidelines for Reward and Suggestions

In order to finish the guidelines, two important problems had to be answered with regard to how a robot can offer reward and suggest a manner of play. We performed an experiment, also processing participants' feedback to understand what facets of play with a robot afforded perception of enjoyment.

4.2.1 Participants

The experiment involved the participation of 20 young adult Japanese participants (9 females and 11 males; average age = 20.3 years, SD = 2.1 years). They were paid for their time, it was their first time interacting with our robot, and their responses were not used to construct the guidelines of the preceding section.

4.2.2 Conditions

We formulated the experiment as a two by two within-subjects factorial design, with "reward" and "suggestions" as factors, each with two conditions as follows.

C1 Maximum reward: enacting a motion in-line with the robot's intention caused the robot to wave its head, arms, and legs while laughing loudly

C2 Progressive reward: enacting a motion in-line with the robot's intention caused the robot to first wave its arms or legs, then its arms and legs, then its head, arms, and legs while laughing increasingly loudly

C3 Shifting suggestions: the robot's suggestions shifted quickly

C4 Persisting suggestions: the robot's suggestions were each repeated several times

Combining conditions, detailed in the preceding section, resulted in four robot designs, which participants encountered in a counterbalanced order. The robot's embodiment and recognition system were the same for all designs. Suggestions were performed when the robot was standing upright on the table.

4.2.3 Procedure

Participants were admitted to an enclosed space surrounded by partitions and seated at a desk. Written instructions were provided, with content similar to the back of a toy box; e.g., they were told that the robot would react if moved, that the robot would suggest ways to play with it if placed on the desk, and that they should not hurl the robot or set it aflame. A brief movie clip was shown to participants in which Sponge Robot's legs juddered; this happened at times as a result of the large weight placed on the upper leg motors. Participants then had a chance to practice holding and touching Sponge Robot, to reduce order and novelty effects. Participants were not informed of what the robot was capable of reacting to or what play behavior they should perform. Then participants played one-by-one with the four robot designs. They were free to play as they wished, using the desk, the floor, their laps, or holding the robot; and if they desired, they could stand or move around the space. Sponge Robot reacted when moved and suggested when standing on the desk. Each interaction with a design was determined to be over when a participant felt they had played long enough. Immediately thereafter, participants received a questionnaire to evaluate the design they had just finished playing with. At the end, a short interview was carried out.

4.2.4 Questionnaire

Participants completed a questionnaire after playing with each version of the robot regarding their perception of enjoyment and some other related constructs. One previous study had used a questionnaire to evaluate enjoyment [68], but this questionnaire was specific to verbal interactions and gaze and could not be used in our case. The items below were answered via a seven-point scale.

· *How to Play* – Did you understand or not understand how to play with the robot?

· *Perceived Variety* – How rich or not rich in variety were the robot's reactions to your behavior?

· *Control* – Did you feel or not feel a sense of control (as if you controlled the flow and contents of the play in the way you wished)?

· *Intentions* – Did you understand or not understand what the robot was trying to do?

· *Enjoyment* – Was playing with the robot enjoyable or not enjoyable?

4.2.5 Predictions

Given the limited capability of robots to move and the tendency of people to lose interest in a repeated stimulus, we predicted that variety would be a crucial factor in reactions. We also predicted that, due to the difficulty of playing with a robot for the first time, novice participants would prefer easy-to-understand suggestions. Formally worded, our predictions were as follows:

Prediction 1: Participants would perceive greatest variety and enjoyment in the progressive reward condition.

Prediction 2: Participants would perceive most clearly how to play, and greatest enjoyment in the persisting suggestions condition.

4.2.6 Results

To interpret the results of the questionnaire, a two-way repeated measures analysis of variance (ANOVA) was carried out with reward and suggestions as the two within-subject factors. Results are presented in Figure 17.

First we considered the effects of reward. A significant effect of progressive reward was noted toward variety ($F(1, 19) = 6.0$, $p = .024$, $\eta^2 = .240$), and a nearly significant effect was observed toward grasping how to play ($F(1, 19) = 4.0$, $p = .059$, $\eta^2 = .175$). Significant effects were not noted for the other measured items (control: $F(1, 19) = 1.1$, $p = .30$, $\eta^2 = .055$; intentions: $F(1, 19) = 1.9$, $p = .18$, $\eta^2 = .092$; enjoyment: $F(1, 19) = 1.7$, $p = .21$, $\eta^2 = .082$). Thus, Prediction 1, that progressive reward would be perceived as most abundant in variety and enjoyable was partially supported. Next, for suggestions, persisting suggestions were perceived to significantly contribute to all items compared with shifting suggestions (how to play: $F(1, 19) = 37.3$, $p < .001$, $\eta^2 = .663$; variety: $F(1, 19) = 45.2$, $p < .001$, $\eta^2 = .704$; control: $F(1, 19) = 10.2$, $p = .005$, $\eta^2 = .35$; intentions: $F(1, 19) = 19.0$, $p < .001$, $\eta^2 = .50$; enjoyment: $F(1, 19) = 71.1$, $p < .001$, $\eta^2 = .789$). Thus, Prediction 2, that persisting suggestions would most allow participants to grasp how to play and feel the greatest amount of enjoyment, was supported. Interaction effects were not significant (how to play: $F(1, 19) = .609$, $p = 0.445$, $\eta^2 = .031$; variety: $F(1, 19) = .717$, $p = .408$, $\eta^2 = .036$; control: $F(1, 19)$

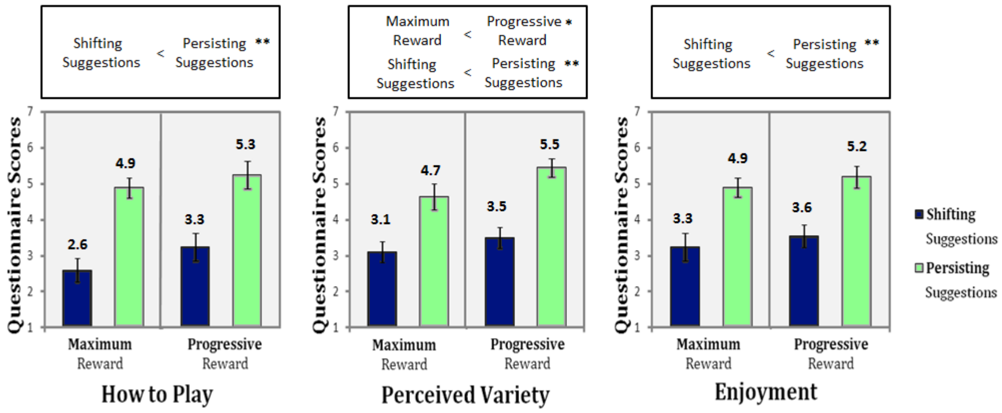


Figure 17. ANOVA results for reward and suggestions.
**** p < .001, * p < .05.**

= .39, p = .54, η^2 = .020; intentions: F(1, 19) = .012, p = .91, η^2 = .001; enjoyment: F(1, 19) = .000, p = 1.000, η^2 = .000).

Feedback was analyzed to understand why the designs were perceived as they were. With regard to the nearly significant effect of reward on understanding how to play, one participant stated that progressively larger motions and the changing volumes of sounds indicated how to play; another participant stated that observing identical responses each time caused her to doubt that she was playing in the intended manner. Regarding variety, three participants who often observed the robot's responses reported varied motions in the progressive condition; two remarked that the magnitude of a motion was a function of the number of times the motion had been enacted, and one participant accurately described which parts of the robot moved for each stage of the progression. In the maximum condition, three participants stated that the robot had repeated an identical motion. Two participants however mistakenly thought that only the robot's arms moved in the progressive condition; therefore, they stated that the robot's reactions in the maximum condition were richer, as the robot's whole body moved. We believe that this was due to the novel nature of the interaction; after playing more than once or for a longer duration, participants would note the reactions more often and correctly understand the difference between conditions.

Comments also shed light on the substantial difference in people's perceptions of the shifting and persisting suggestion conditions, indicating that trouble was experienced in understanding the robot and provoking reactions in the former case (shifting suggestions). In detail, eight persons described problems comprehending what the robot was doing and its wishes. Five persons stated that our robot did not always react to their gestures. Three persons stated that the robot's reactions did not change, two indicated that our robot's sounds were heard less, one professed a lack of certainty that the robot was satisfied, and another stated that the robot's reactions were small. These statements suggested that these persons did not frequently see big happy reactions from our robot. A potential reason is that the interaction may have been perceived to proceed at an excessively quick pace. Five persons stated that they initially did not know how to interact and that experience was required, three persons stressed that the robot had moved and suggested often, one indicated observing greater variety, and another stated that our robot varied its motions too quickly. In the progressive condition, however, participants were able to observe our robot's responses to a given gesture multiple times. One person stated that he was able to be confident about how the robot would react to his gesture only after several attempts. The ability to make such associations inspired understanding, facilitated observation of various responses, and thereby contributed to enjoyment.

4.2.7 Sources of enjoyment

To identify reasons for why people perceived enjoyment, participant's comments were additionally processed. Two reasons were found. We introduce these reasons by starting first with an example of a successful interaction. Figure 18 depicts one participant's experience playing with Sponge Robot. We will call this participant "Tony" although this is not his actual name. Tony shook hands with the robot to start the interaction. Then he watched the robot raise both its arms upward and shake them. He did not at first comprehend the meaning of this suggestion which he observed or how he could interact. However, Sponge Robot repeated its suggestions several times and gave him time to think. Tony decided to try raising the robot up high. He was rewarded by some laughter. Again he raised the robot, which laughed louder and louder and moved its arms, legs, and head happily. Tony indicated in his interview that he found this episode to be enjoyable.

Figure 18. Some moments participants reported to have enjoyed.
a) raising the robot high in the air like an infant, b) watching the robot do push-ups to try to stand, c) helping the robot to walk across a desk.

To investigate further how such a scenario was tied in with perception of enjoyment, we report some feedback from participants. 14 participants described various ways of playing which provided enjoyment: four as with Tony enjoyed raising the robot high up into the air like a child, four enjoyed dance motions, four enjoyed suggestions which the robot performed when it wished to be placed upright in which the robot seemed to be exercising or struggling to rise up by its own power, one liked how the robot struggled when it was placed upside-down, and one found it entertaining to lay down the robot. In terms of responses and suggestions, six participants reported liking when the robot responded to their gestures and two individuals liked knowing how to play as a result of the robot's suggestions. Thus, participants enjoyed being afforded the opportunity to conduct a variety of different play behavior with the robot.

As well, 14 participants volunteered that the robot had seemed happy (without being asked by the experimenter). Four mentioned that playing with Sponge Robot caused the robot to laugh. Another four participants indicated that their actions caused the robot to be happy. We asked the participants who described the robot as happy their reason for this belief. The reported cause was the robot's laughter and voice for 13 participants and the robot's happy-seeming motions for seven participants. We also asked how participants felt about the robot seeming happy. Nine participants felt good, happy, or that it was enjoyable. Two participants noted how the robot seemed to be feeling without investing any emotions themselves. One participant recounted feeling accomplishment. Thus, a feeling of control over the robot's emotional state, especially comprising an altruistic joy over causing it to appear happy, also seems to have contributed to enjoyment.

4.3 Evaluating the Guidelines for Providing Enjoyment

After completing the proposed design, an experiment was required to test that following the guidelines truly results in more enjoyable interactions. To acquire some evidence to support our findings, we compared our final proposed system with our initial naïve design, which had been based only on some intuitive knowledge drawn from existing products.

4.3.1 Participants

21 young adult Japanese (8 females and 13 males; average age = 21.8 years, SD = 3.0 years) participated in our experiment and were remunerated for their time.

4.3.2 Conditions

The following conditions were experienced in counterbalanced order.

Naïve design: the simplified first design for our robot described in Chapter 4.1.1 based on intuition from previous products. This design featured a turn-based style of interaction with idling, infant-like behavior corresponding with the robot's appearance, and simple instructions. The robot suggested always the last gesture performed by a person then responded by laughing if the person followed its suggestion or crying otherwise.

Proposed design: our final design incorporating meaningful motions, progressive responses, persistent suggestions, and fulfilling instructions, as described in Chapter 4.2.6. "Progressive responses" means that when a person did what the robot wanted, the robot wiggled first its arms or its legs, then both its arms and its legs, then everything, with progressively louder and more cheerful laughter. "Persisting suggestions" means that the robot, when not being moved, repeated each suggestion more than once.

4.3.3 Procedure

As in the experiment described in Chapter 4.2.3, participants sat at a desk in a partitioned-off space, received a short handout with simple instructions, and played with the different versions of our robot. They could choose to sit, stand, or otherwise move. After playing with each design, participants completed a questionnaire. At the end, a short interview was conducted. Two differences from the previous experiment was that there were only two designs and that participants were asked by the experimenter to stop playing if ten minutes went by.

4.3.4 Questionnaire

The three most important of the measures used in the previous experiment were utilized: "How to Play", "Perceived Variety", and "Enjoyment".

4.3.5 Predictions

For the reasons provided in Chapter 4.2.6, the proposed design was expected to yield higher scores for all three measured items that our naïve design.

4.3.6 Results

To understand how participants perceived playing with each of the two design conditions, we performed a one-way repeated measures analysis of variance (ANOVA). Figure 19 depicts questionnaire scores and ANOVA results. As a consequence of analysis, we were able to verify that our prediction was supported. Participants better understood how to play, perceived greater variety and felt more enjoyment with the proposed design than with our first design: How to Play: $F(1,20) = 26.5$, $p < .001$, $\eta^2 = .570$; Perceived Variety: $F(1,20) = 23.3$, $p < .001$, $\eta^2 = .538$; Enjoyment: $F(1,20) = 18.0$, $p < .001$, $\eta^2 = .473$). Thus, the guidelines helped to provide enjoyment.

4.4 Summary

In summary, we were able to fulfill our first objective of realizing a way to provide enjoyment via motion-based play with a small humanoid robot. Key findings are presented again:

- A simplified approach toward incorporating recognition capability to structure a robot's behavior, consisting of turn-taking, lifelike idling, motions corresponding to a robot's appearance, and minimal instructions, does not always work. Typical failures include the robot's motions seeming

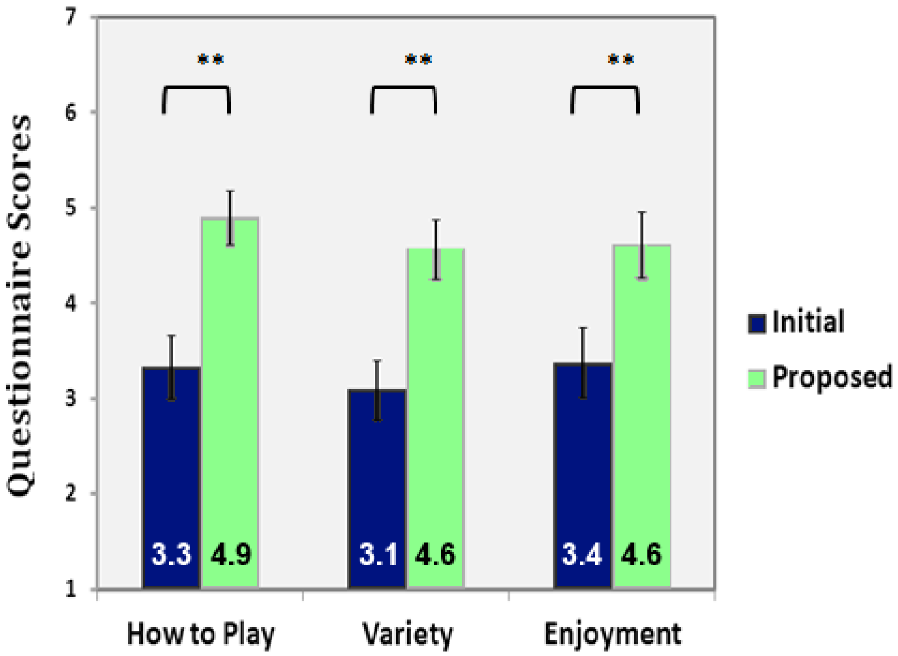


Figure 19. ANOVA results for the initial and proposed systems. ** means p<.001.

meaningless, and the user not interacting much or in the expected manner. Four guidelines were proposed, advocating the use of meaningful motions, rewarding responses, inspiring suggestions, and fulfilling instructions.

- We compared two strategies each for providing enjoyment through a robot's responses and suggestions. A progressive response strategy was found to increase perceived variety compared to a strategy which always used large responses. Suggestions which persisted for a time instead of shifting immediately were observed to facilitate comprehension, enable variety to be perceived, and provide enjoyment. Participants described feeling enjoyment for two reasons: successfully causing the robot to appear happy, and being able to play in various ways.
- A second experiment provided evidence that following the proposed guidelines facilitated comprehension and the perception of variety and enjoyment.

By shedding light on how a robot can provide enjoyment in interactions, we believe that this undertaking moves one step closer to providing a feeling of social well-being in interacting persons, due to the evidence for a connection between enjoyment and well-being discussed in Chapter 1.1.

Chapter 5

Recognizing Affection via Touch (Kirin and Elfoid)

“Then lips to lips he join'd; now freed from fear,
He found the savour of the kiss sincere:
At this the waken'd image op'd her eyes,
And view'd at once the light, and lover with surprize.”
Ovid, in Metamorphoses X, 8 AD
of a statue, recognizing the kiss
of her creator, Pygmalion, and coming to life

The goal of the third study conducted was to build a system to recognize affection, toward allowing a humanoid robot to engage in affectionate interactions; this work has been previously presented [3, 6]. Our logic for this work stemmed from the following suppositions.

- Although people’s behavior is complex, certain touches will be often observed
- The type of touch a person performs will allow a robot to judge the degree of affection a person feels toward the robot.
- The robot can recognize touches via touch or vision sensors.

The first two points represent our expectations based on our knowledge of human-human touching. We know that people’s touches are not completely arbitrary but that common pattern such as hugging and pushing exist. We also know that touches are not meaningless; e.g., a different attitude is expressed by warmly hugging someone or roughly shoving them away. The results of a study we describe in the current section support our expectations, as will be discussed later. We also expected these suppositions to be valid based on some similar previous studies which were described in Chapter 1.3; we note that this differs from the case of enjoyment-motivated behavior, which is characterized by great variety in people's preferences for how to play.

The third point is obvious because touches may be either felt or seen or both; we do not need to be hugged ourselves to visually recognize someone else being hugged. This scenario had to be investigated because, as described in previous sections, affection is an important determinant of well-being and touch is a fundamental modality for expressing affection. By investigating this phenomenon, we aim to prepare a robot such as Gozilla for eliciting affection in a touch-based interaction with a person such as Masha, as in the user scenario described in Chapter 1.2.

Our work to address this scenario involved two stages: identifying typical touches and their meanings, and acquiring data to build a recognition system. Additionally, Chapter 6 describes how a robot can seek to show affection by approaching a person; also, some unpublished work of ours has further investigated strategies for eliciting affection, but this is not discussed here.

5.1 Identifying Typical Touches and their Meanings

Our first task was to determine what a robot should be able to recognize; this included identifying typical touches and confirming that they were meaningful in the context of communicating affection. We expected this to be challenging because people's behavior is extremely complex. As before, we adopted an observational approach in which participants were free to interact.

Using only one humanoid robot form would not inform of the generality of the identified behavior (i.e., whether we can expect people to behave the same toward a differently appearing robot); therefore we used two humanoid robot forms and defined "typical touches" as those performed by two or more participants to each robot. As well, we did not know how people's behavior would change between a big or small robot; therefore we checked two scenarios: one in which people sit and touch an infant-sized robot, and one in which people stand and touch an adult-sized robot.

Before acquiring data we also checked to see how we could elicit a wide variety of touches. We asked people in pre-trials to touch a robot form with affection, neutral affection, and no affection. We learned that people had trouble knowing what to do; they wanted to know why they were touching a robot form. To help them, we compiled a list of typical reasons for which people touch other people from human science references, and modified some of these to fit our scenario in which a person touches a robot. Also, we decided to ask participants themselves to evaluate how much affection a robot should perceive from their touches; a seven-point scale was used with one representing hate (no affection) and seven representing liking (high affection).

5.1.1 Conditions

Participants touched two different robot forms, Kirin and Elfoid, each for 14 typical reasons. We refresh the reader's memory with regard to the two robot forms, which were described previously in Chapter 2.2:

Elfoid: a roughly hand-sized, light, hollow plastic form with a head and arms but no legs; it was designed to seem capable of being interpreted as old or young, male or female.

Kirin: an adult-sized, mannequin-like form clothed in dark fabric, with many degrees of freedom, and a head, arms, and legs.

"Form" means they were mock-ups with the appearance of a humanoid robot but without actuators and electronics. Both forms were very robust and had highly different appearances. Participants touched the forms in counterbalanced order.

The humanoid robot forms we used, Kirin and Elfoid, were not designed to move while participants touched them. This follows an established pattern in previous studies [150, 82]. Our logic for following this pattern stemmed from several expectations. First, a robot's motions may alter a person's behavior in complex ways via the entrainment phenomenon or by interrupting. Second, robots, like humans, will not always move; this non-motion case is also appealing and fundamental because it can apply to robots with various motion capabilities. Third, we had no indication that a robot's motion was needed to receive affection. Fourth, there was no means to test all possible motions which can be performed by a humanoid robot, or any indication if there was any systematic way of selecting a representative subset of motions which could be tested.

Two human science studies mentioned in Chapter 1.3 were used as a basis for identifying reasons for touching [155, 107]. This approach was based on our expectation that humanoid robots will be treated much like people; such an assumption has been followed previously for clarifying how people are likely to touch a cat-like robot [82]. Most reasons were obtained from the first source listed above. The second source indicated several fundamental emotions which can be communicated via touch; we felt that people might seek to display such basic emotions toward robots incapable of comprehending more complex messages. Therefore we added expressing such emotions to the other list of reasons. Support and sympathy seemed to be less different than other reasons and similar; for simplicity these were merged. Because our sources did not consider the case of human-robot interaction, we attempted when possible to adapt reasons to our scenario. Thus, “control” was conceived to mean “controlling a robot to move” and a task-related reason of inspecting a robot was employed. Also, because our focus was on affectionate interactions, we wanted to know not only how people will seek to express themselves but also how they might seek affectionate attention from a robot. Therefore, we broke two reasons which we felt would be most related to affection, love and sympathy, each into two versions: showing the emotion, and seeking it from a robot. The employed compilation of reasons for touching such as greeting or thanking is presented in Table 4.

5.1.2 Participants

21 young adult participants were recruited to touch the humanoid robot forms for typical reasons (9 females and 12 males; average age = 24.1 years, SD = 4.4 years). Our participants comprised 19 Japanese and 2 non-Japanese; due to the exploratory nature of the data acquisition we did not restrict participants to be of Japanese nationality, as our goal was to note a large spectrum of touches.

5.1.3 Procedure

The experimenter admitted participants into a wide room with a desk and two humanoid robot forms, Elfoid and Kirin. Sitting at the desk, participants read instructions on a handout stating that they would be required to touch the two

Table 4. Reasons for touching a humanoid robot adapted from human-human interaction studies.

Reason	Concept	Reason	Concept
Reassurance[1,2]	Poor robot, it's okay	Play[1]	Let's play
Appreciation[1]	Thanks	Control[1*]	Move your body/arm/head
Inclusion[1]	We're friends	Greeting/Parting[1]	Hello/Goodbye
Attraction[1]	How handsome/cute	Task-related[1*]	Let me inspect you
Love[1,2]	I love you	Show Anger[2]	I hate you
Seek Sympathy[1*]	I'm tired; my head/belly/throat hurts	Show Disgust[2]	Ew, gross
Seek Love[1*]	I'm lonely/sad	Show Fear[2]	I'm scared

* Adapted or added for the context of affectionately interacting with a robot.

Sources: 1 [140], 2 [101].

forms for various reasons; the reasons, such as playing or thanking a robot, were listed and the robot forms described. It was also written that they would be asked to assess each of their touches by assigning a number from one to seven to describe how much affection (like or dislike) should be felt by the robot. The experimenter confirmed that participants understood the instructions and answered any questions the participants had. Then the participant stood before Kirin or sat holding Elfoid and the experimenter verbally indicated reasons for interacting one by one. For each, participants received some time to consider how they might behave, touched the robot form when ready then assigned a score representing how affectionate their touch had been. After participants finished touching one robot form, the process was repeated with the remaining form. Participants could touch the forms any way they wished. During sessions, the experimenter took written notes on what participants did. Video footage was also obtained. After participants had interacted with both forms for all reasons, they relayed their impressions to the experimenter in a brief interview. Sessions were roughly 30 minutes in duration.

5.1.4 Making sense of the observed interactions

As in our previous study on proprioceptive playing, participants behaved in an extremely complicated fashion, but some patterns presented themselves. We noted soft touches which were not affectionate, confirming our expectation that softness by itself is not enough to infer affectionate significance. Some touches were observed to occur either in isolation or in conjunction with other touches, such as hugging and back-patting which occurred alone or simultaneously. Sometimes, before touching in a meaningful way, participants also performed preparatory behavior involving touches such as steadying Kirin or holding Elfoid, moving to Kirin's side, or tilting Elfoid. Some touches seemed to involve certain directions and not others: for example stroking the robot's cheek never involved a front to back touch. Because the robot forms and participants both possessed a bilaterally symmetric form (the humanoid form), touches exhibited many variations; for example, shaking hands was performed with right, left, or both hands of either the robot form and the human. Several touches were rare and only performed by one participant such as seizing the chin of a robot form or punching with an upper-cut. Other touches were performed to only one of the robot forms and not the other. Some participants rubbed Elfoid's belly, hugged it while rubbing its head, inverted it, and even tossed it. The last two touches were not enacted because Kirin was too large and heavy. Some participants also pushed Kirin's shoulders and brought its hands up to be able to high-five it. This was not possible with Elfoid because it lacked shoulders, articulated joints, and hands. In general, people touched both forms in a similar way to how people touch other people, performing touches such as hugging the forms and shaking hands. One reason for this was that the humanoid form in both caused people to attribute a human-like quality to them, as was shown by comments during the interviews: two remarked that Kirin resembled a friend and human, and five mentioned that Elfoid seemed like a child or baby. However this perception was not universal. Two participants stated that Elfoid was like a small animal or doll.

In order to better understand the underlying structure of the phenomenon in terms of typical touch patterns, the experimenter's written notes were completed by reviewing the video recordings and codified via brief words to yield 920 touch concepts of which 239 were distinct. Database and SQL queries were used for processing. Touches performed to both forms by more than one participant were automatically extracted. Spatiotemporal similarity and a small distance in affection values were taken into account to build a number of higher-level categories. Thus, processing revealed a conceptual structure of 20 typical touches, and eight categories, sorted as affectionate, neutral or unaffectionate. Affectionate touches included stroking, hugging, and pressing; neutral touches included checking, patting, and controlling; and unaffectionate touches included hitting and distancing (reminding of a fight-or-flight response). The 20 typical touches are shown in Figure 20 sorted by affection and percentage of participants who performed them.

5.1.5 Comparing results

Some previous work described in Chapter 1.3 produced lists of touches either observed or expected to be enacted toward robots. Noda *et al*. produced a list of play gestures toward a large humanoid robot, Robovie, which included both specific touches and abstract categories such as "whee" [149]. We were not able to relate our findings to the latter kind of abstract category, but all specific touches mentioned were also present in our results: hugging, hand-shaking, and head-patting.

Knight *et al*. also made available a list of playful touches toward a small robot with the appearance of a teddy bear [150]. We noticed some of these touches such as head-patting, hand-shaking, and shoulder-tapping; we did not observe feeding, tickling, and foot-rubbing. Feeding was not observed because objects were not provided. We believe we did not observe tickling because (1) we did not focus on observing interactions only for play but for various typical reasons, (2) Kirin was adult-sized which may have dissuaded tickling, and (3) Elfoid was made of plastic and did not feel biological (as if it could feel ticklish). Touches involving feet were not observed because Kirin's feet would be difficult to touch from a standing orientation and because Elfoid lacks feet.

Yohanan and MacLean published a list of touches which was gathered from sources on human-human and human-animal touching [82]. Thus, this list was generated in a slightly different manner from the other studies including our own, in the sense that it did not stem from manually labeling participants' freely chosen touches during interactions (although participants were also asked to demonstrate touches they thought they would be most likely to perform). Some items in the presented touch dictionary were noted in our data acquisition such as Hug, Kiss, Stroke and Rub, Toss, Hit, Push, and Poke (which we called Minimal Touch). Some other items included proprioceptive motions, which we had not considered in the current data acquisition, such as Rock, Shake, and Lift; we had noted these

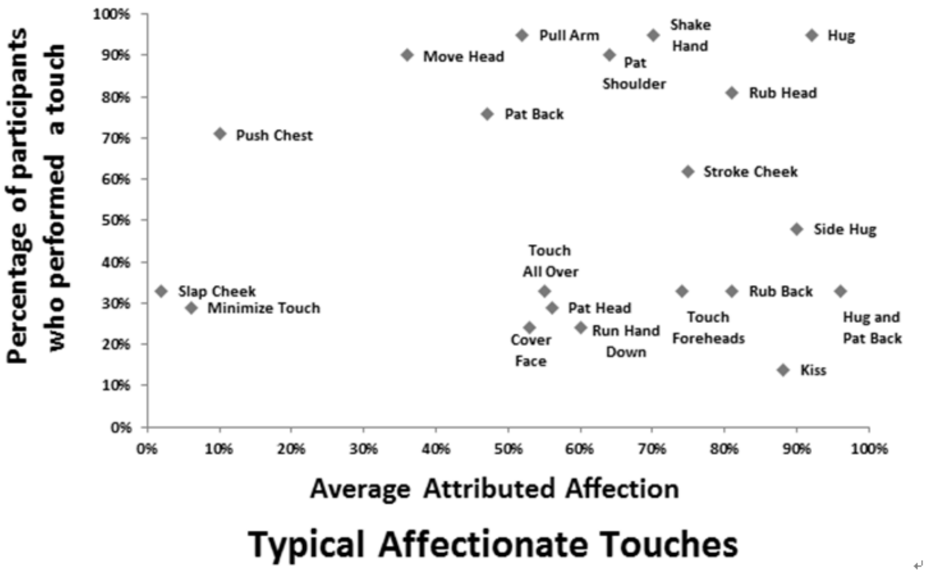


Figure 20. Typical touches performed toward two humanoid robot forms. Performed by more than one participant, organized by affection and frequency.

in our results in Chapter 3.1 in regard to people's motion-based behavior toward a robot. We did not observe some items such as Tremble, Finger Idly, Scratch, Pick (fur-pinching), and Massage. The first touch, tremble, came from a study in which people were asked to show fear by touching another person's hand; the difference in settings may have been the cause for not observing this toward our two robot forms. Finger Idly, Scratch, and Pick were related to touches toward an animal with fur. We believe massage may not have been observed because our participants did not imagine the hollow and mechanical bodies of our robot forms could ache or required kneading. In addition to the touch dictionary, it was reported that participants most frequently touched the back of the cat-like robot used in the study; we can add to this the observation that people most often touched our humanoid forms' chests, arms, heads and backs; very few touches occurred to the lower body of the forms. In summary, the results shared by our study and other similar studies suggested that people will perform such touches to various robots and that they should be recognized.

In terms of identifying the meaning of touches, we found some differences in our study and related ones. As noted in Chapter 1.3, other studies seeking to distinguish meaning based on softness of touch used a canine-shaped robot and arm segment and did not test the effect of using a complete humanoid form. However, our observations suggested that the meaning of a touch on a humanoid form also depends on location-related factors. This offers an explanation for why soft touches such as Minimize Touch, Cover Face, and some softly enacted Push Chests were construed as being unaffectionate. Minimize Touch was characterized by restricting greatly the location of touch, creating distance, and pushing a robot form away. Cover Face likewise was a minimum location touch to a vulnerable-seeming area of a humanoid form's eyes, discouraging communication. Push Chest also enlarged the space between the robot form and human by driving the form away. Moreover, some hard pats communicated affection when they were directed toward a robot's seemingly robust shoulders or back, thereby sometimes also bringing a robot form closer to a person. However, despite this difference in our study and the two related studies, we note that our conceptualization using a single value of affection between zero and one can also be used with a simple threshold value to label touches as belonging to one of two categories, like in the related studies. In summary, we expect that our results can inform in regard to what touches we can expect people to perform directed toward other robots, as well as the meaning of these touches; thus, this information would be useful to recognize.

5.2 Recognizing Touches

Determining what a robot should recognize was not enough to realize our goal of affectionate interactions; we also bore a responsibility to show how these touches could be recognized. Toward this, as described in Chapter 1.3, two sensor modalities, touch and vision, seemed promising, because humans can recognize touches by either feeling them, seeing them, or both. Therefore we investigating all three options—using touch, vision, or both touch and vision—in conjunction with Kirin, a humanoid robot form (a mock-up robot) described in Chapter 2.2. Kirin was selected instead of Elfoid because participants described the former as easier to touch.

5.2.1 Touch and vision sensors

To recognize the complete set of typical affectionate touches with touch sensors, we selected an approach based on guidelines proposed in the literature indicating that touch sensors should be tough, soft, light, adequate in terms of sensitivity and surface covered, readily built, and capable of being attached to round substrates [180-181]. To fulfill these requirements we implemented an idea from previous work, based on a store-bought component, a SHARP GP2S60 photointerrupter [181]. This mechanism works as follows: a photo-emitter beams infrared light onto the reflective inner surface of a sensor, which is reflected down to a photo-receiver. When a person touches the soft deformable sensor covering, pushing it downward, the distance changes, resulting in more light reflected back to the receiver. This in turn creates a measurable change in voltage. The GP2S60 data sheet showed that the sensor is highly sensitive at close range, reacting to very small changes in distance (less than 0.5mm) [182]. These properties and others for our implemented sensors are summarized in Table 5.

A problem we faced was how to design the covering. We started with the observation that physically two types of touches seemed to exist: perpendicular patting and lateral rubbing, with the former most common and simplest to detect. Therefore we constructed two different types of touch sensor: one simple one which detects only patting and one slightly more complex one which detects either patting or rubbing. The coverings and internal structure of the touch sensors are shown in Figure 21. We created forty sensors in total and placed the sensors on Kirin where we expected people to perform each basic kind of touch. Eight sensors were attached at each of five locations: head, chest, back, and arms. This number of sensors is greater than the number on AIBO and Paro [76], and slightly less than the number on Haptic Creature [102] and Sensate Bear [150]. In detail, the constructed sensors were composed of GP2S60 photo-interrupter units, foam, and connecting parts such as hook-and-loop fasteners, bolts and nuts. Their size was large (~15 x 6cm and ~6 x 6cm) to cover the wide area of Kirin we expected people to touch, but they were light (~21g and ~11g), soft with no sharp corners, and robust enough to be treated forcefully without being damaged. Values for all forty sensors could be obtained at ~20 Hz, which we deemed adequate for our purpose.

Table 5. Touch sensor properties.

Property	Description
Measured Quantity	Displacement of the soft covering on contact
Internal Element	SHARP GP2S60 Photo-interrupter
	Dimensions: 3.2×1.7×1.1mm
	Detecting Distance : 0.5mm
Size	~15×6cm (pat/rub),~6×6cm (pat)
Weight	~21g (pat/rub), ~11g (pat)
Sensing rate	~20 Hz (for 40 sensors)
Placement	Kirin's chest, head, left arm, right arm, and back (eight sensors each)
Area coverage	~1700cm^2 from 35 pat sensors and 5 pat/rub sensors
Sensitivity	0.5N (500Pa)- 5N (1350Pa)

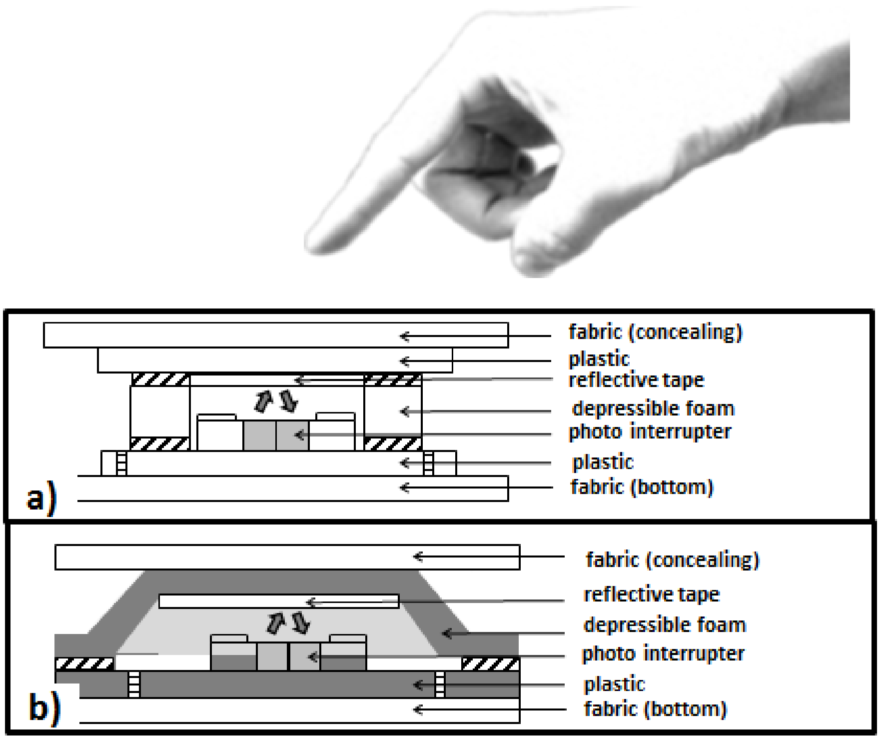


Figure 21. Touch sensors built.
a) a sensor which recognizes only perpendicular touches (pats) b) a sensor which recognizes either perpendicular or lateral touches (pats or rubs).

One potential problem was that affectionate touches could be fairly soft (although we expected them not to be too soft [154]) and thus difficult to detect. We especially expected kisses to be very soft. On the other hand, other touches like pushing could be hard. Therefore, to confirm that the sensitivity of the sensors was sufficient to recognize the designated typical touches and identify a range of forces and pressures which could be measured, we conducted a test. For this, we verified that a kiss and a heavier touch could be perceived by simulating the touches using two weights placed onto a sensor. This required us to have an idea of how much force was required and the area of contact, which would also allow the pressure to be calculated. First we found based on a source in the literature, that the lightest touch we expected, a kiss, typically involves 500Pa of pressure [183]. To find how much area of contact would be involved, the area of a confederate's lips was extracted. To do so, the confederate used a beverage to wet his lips and kissed a sheet of cardboard; then the wettened area was cut out from the cardboard sheet and measured. Because the area was irregular, it was approximated using primitive shapes such as rectangles, triangles, and trapezoids, and found to be $10.05cm^2$. Thus, the force of a kiss was calculated to be approximately 0.5N. To simulate the force of a kiss, a weight placed on top of the sensor thus needed to be approximately 50g. Therefore we prepared a 50g weight and attached it to the top of the cardboard lip-shape. The weight was then suspended on top of the sensor, with a string attached from the weight to a frame above to prevent the weight from falling. To simulate a heavier touch such as a push or slap, we also prepared a second weight of 500g, ten times heavier than the light weight; with an area of $36cm^2$ (approximately palm or fist-sized), this resulted in a pressure of 1350Pa.

Our test involved two phases. First we tested detecting "pats" by placing the light "kiss" weight on the sensor twice, then the heavy weight. Then we tested detecting "rubs" by pushing the light and heavy weight over the sensor. The test apparatus is shown in Figure 22, and the sensor output in Figure 23. The results showed that sensor data was more affected by the pats than the rubs, because some lift was applied to the weights during the rubbing action to stop them from sticking due to friction; although in one case of rubbing the kiss weight created only a weak peak, this was acceptable because we expect kisses to be "pats" rather than "rubs". Thus, we confirmed that the sensors were sensitive enough to detect both light and heavier touches.

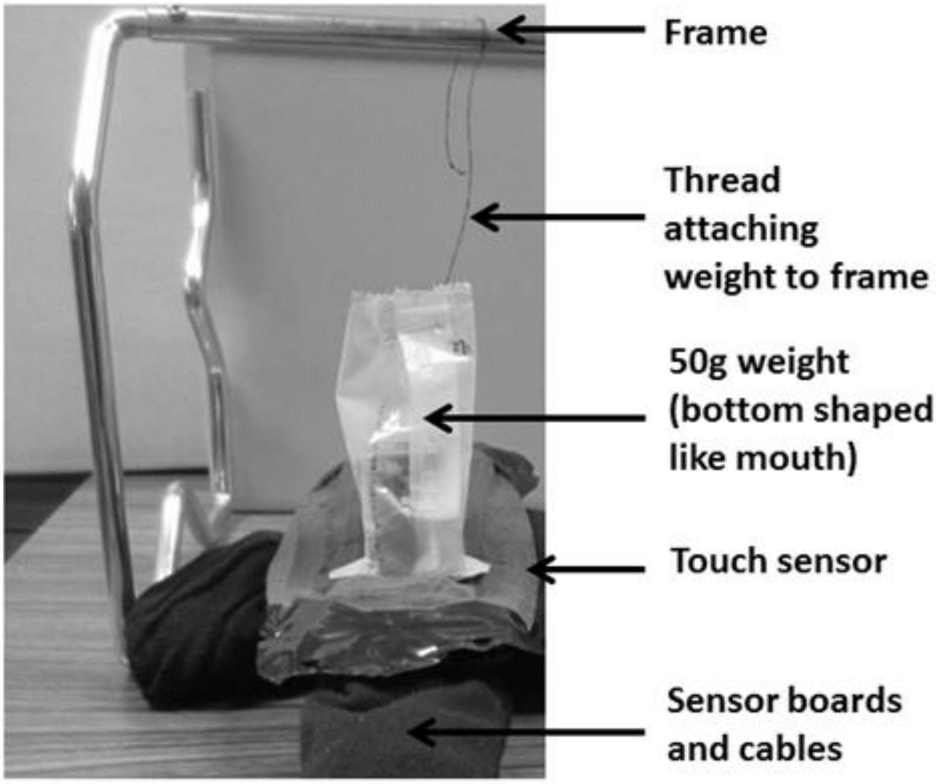


Figure 22. Apparatus for testing touch sensors.
Two weights, 50g and 500g, were alternately placed on top of and pushed across a sensor to check that very light and heavier patting or rubbing touches could be detected; the 50g weight was used to simulate a kiss, and the heavier weight a push or slap.

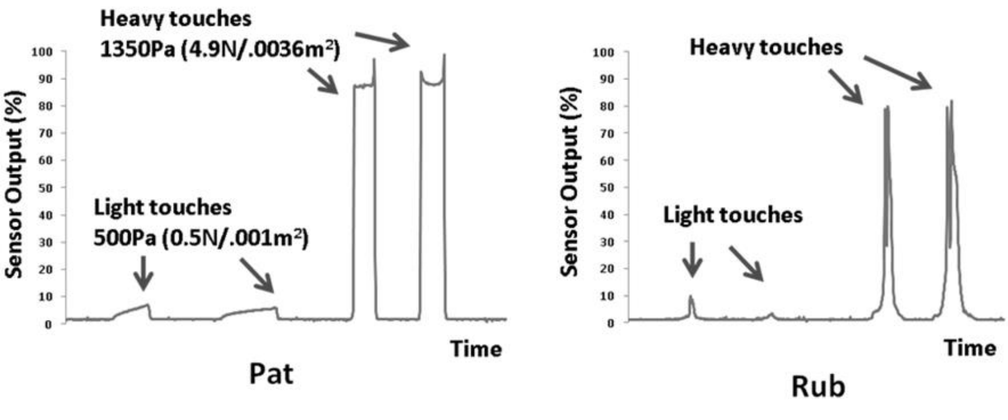


Figure 23. Effects of touches on the touch sensor data.
Two weights, 50g and 500g, were alternately placed on top of and pushed across a sensor to check that very light and heavier touches could be detected; the 50g weight was used to simulate a kiss, and the heavier weight a push or slap.

The touch sensors also needed to be attached to Kirin. We chose to not mount them directly to Kirin but rather to attach them to a sensor suit. The reason for doing so was that we wished to be able to easily attach and remove the suit so the developed touch sensor system could possibly be used with different humanoid robots. Placements for the touch sensors are shown in Figure 24.

In order to sense touch gestures visually, we started by noting that one common way of recognizing gestures is to compute people's postures or skeletons [184]. Motion capture is one way to accomplish this, but requirements are considerable: systems can be expensive, a wide, high space is required, and we expected markers would prevent natural touching. Therefore, we set up a system using Microsoft Kinect and Open NI with Prime Sense. We adopted a scheme which required people to perform a calibration pose; although introducing some unnaturalness, this helped to gain some extra accuracy in pose detection. To avoid requiring a complex approach to deal with potential sensor interference from multiple sensors, we decided to use only a single sensor, which we situated 2.3m behind and to the right of where we expect people to stand when interacting with our robot form (this distance was indicated as an optimal distance for detecting skeletons). The sensor was placed on the right based on our expectation that most people will be right-handed (and thus use their right hands to touch a robot form). This placement is shown in Figure 25.

5.2.2 Recognition approach

Our goal was to recognize typical affectionate touches, from which a robot can make conclusions about the attitude of a person toward the robot. As before, we implemented a solution based on Support Vector Machines (SVMs), described in detail in Chapter 3.2.2. To check that our results were reasonable, we also implemented a k-Nearest Neighbor (k-NN) approach [185]. K-NN was used because it usually yields fair results which may at times exceed results from more complex approaches [186]. The standard version of this technique involves no training phase (all samples provided for training are used as is). To recognize which class a test sample belongs to, the algorithm identifies the closest k training

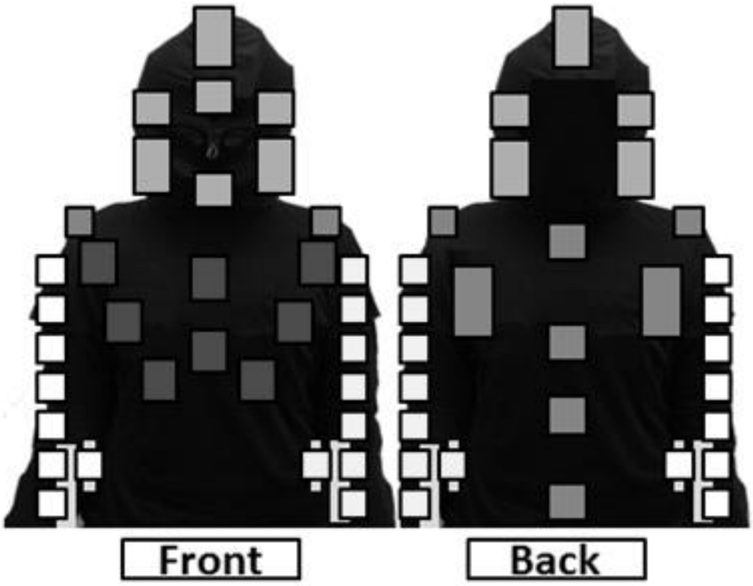


Figure 24. Placements of touch sensors on Kirin.
Square sensors detect perpendicular touches (pats); rectangular sensors detect perpendicular or lateral touches (pats and rubs).

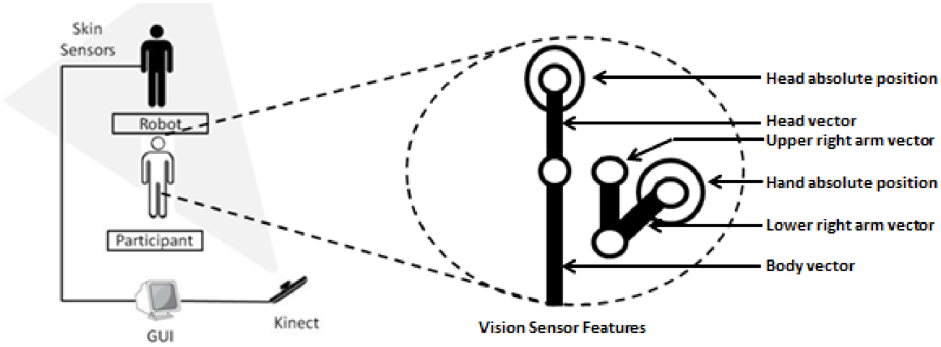


Figure 25. Both Kinect data and touch sensor were obtained.
(Left) data acquisition environment, (right) Kinect data features computed.

samples in feature space. Then the class label which is most common in the associated data points is attributed to the new test sample. Aside from our focus on recognizing typical touches, we also sought to investigate if we could recognize the affectionate meaning directly from the data; for this we used Support Vector Regression (SVR), which is simply the application of SVMs for regression; i.e., instead of providing a class label, the SVMs provide a continuous value output representing the affection which should be attributed to a test sample.

SVMs and SVR were realized as in our previous study by using LIBSVM ("A Library for Support Vector Machines"). Parameters were selected as follows. C =

32.0, γ = .03125 for touch, C = 8.0, γ = .125 for vision and C = 8.0, γ = .03125 for touch/vision. We selected a typical number for k, using a window of three closest neighbors to investigate (k = 3).

In regard to features, we calculated features from a short window of several seconds of data. Vision data was first pre-processed. We computed translation invariant vectors representing a person's pose, following typical procedure for visual recognizing [184, 187]. However, we retained the values for the position of a person's right hand and head in global coordinates in the expectation that this would have some meaning for our problem. (Unlike typical gesture recognition scenarios which seek to answer what a lone person is doing, we wished to identify social gestures which are performed relative toward a robot; given that the robot is stationary, the absolute position could help to determine which part of the robot is being touched.) Next, as in our previous study, simple statistical features (e.g., mean, standard deviation, and median) were calculated from the touch and preprocessed vision sensor data. Per each pre-processed feature, eight statistics were calculated.

The feature set, consisting of touch, vision, or both types of features, was then passed on to the SVMs and k-NN algorithm. In other words, we used "early fusion", the simplest way of combining different types of data. Thus, 320, 144, and 464 features were input to each classification system (touch, vision, and both, with SVMs or k-NN). The process line for feature calculations is shown in Figure 26. Next, to determine how the systems performed, data were required.

5.2.3 Data collection

Touch and vision sensor data for the twenty typical affectionate touches were obtained from 17 adult participants (10 females and 7 males; 8 Japanese and 9 non-Japanese; average age = 31.8 years, SD = 6.3 years). Each touch was performed five times, resulting in 1700 samples.

Participants were admitted to a wide room by the experimenter, seated at a desk, and asked to read a handout of instructions. The objective of data collection was described, as well as the touch and vision sensors, and the list of typical

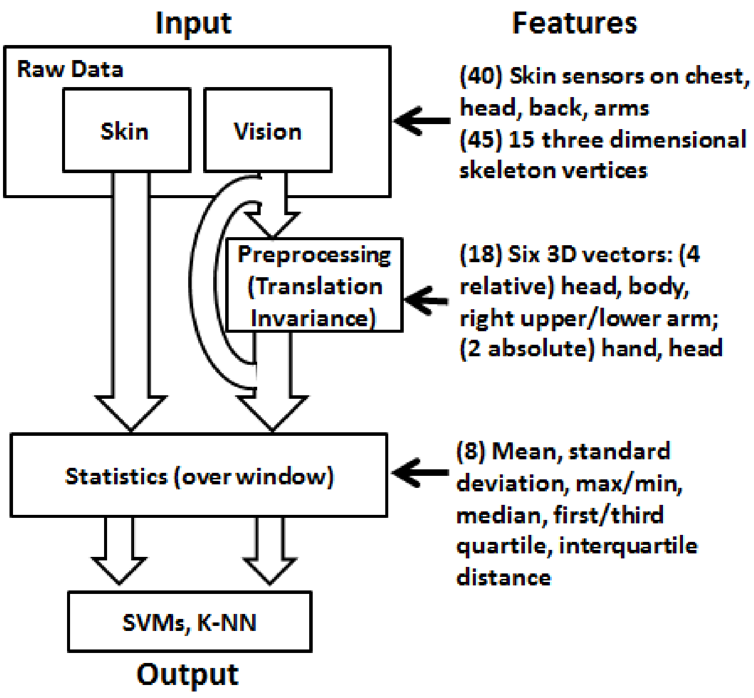


Figure 26. Process flow for recognizing people's touches. Features computed at each stage are indicated.

affectionate touches was presented with explanations. Participants had a chance to ask questions and the experimenter confirmed that the participants understood what they would be asked to do. Before beginning, participants obtained a mask to wear; this was intended to ensure hygiene and their privacy when participants kissed Kirin. To ensure participants behaved naturally, they were given the chance to touch Kirin before data was recorded. They were not told where touch sensors were situated although this could be felt via either touch or examining Kirin. To make our task simpler, we requested that participants touch with either their right hands or both hands; we consider it clear that we could accommodate left-handed persons merely by placing the sensor on the other side. Then participants were asked to perform touches in random order; random orders were pre-generated and presented by our graphic user interface (GUI) program. Before recording a touch, participants were asked to do a calibration pose to aid accurate estimation of skeleton poses. Then, a WAV file was played to indicate to the participants to touch. Finally participants again performed the calibration pose; then they described how much affectionate a robot should perceive as a result of their touch using a seven-point scale. During data acquisition, it was possible using the GUI, to view sensor data in real time and play it back. The program also helped by inserting into each data file labels indicating the name of each touch and times, and counting how many times each touch had been recorded. This facilitated the task of the experimenter in preparing the data to be used with our recognition system.

5.2.4 Evaluation

To determine to what degree the typical affectionate touches could be recognized, we determined recognition accuracies for all versions of our system using leave-one-out cross-validation. Accuracy was defined as before as the number of touches recognized correctly divided by the total. Accuracy could be used because the number of samples was the same for all touches. For SVR, the linear error in predicting affection values was computed for all folds, and the results were used to find an average.

A number of unexpected results were observed. (1) The touch/vision system with SVMs yielded 91% accuracy, which we felt was accurate given the challenge faced in terms of the many touches which needed to be recognized, the high degree of variation in people's interpretations, and some touches which closely resembled. (2) This accuracy was higher than that for the touch or vision systems by 18.6% and 13.0%, suggesting that combining modalities was useful. (3) Also, the vision system yielded good results (5.6% higher accuracy than the touch system) despite the ease of its implementation. The results from the k-NN systems, although exhibiting slightly lower accuracies, also followed this pattern. We attempted to gain some further insight into these results.

(1) The hybrid system with SVMs performed accurately for Side Hug (98%), Shake Hand (96%), and Push Chest (95%). Most trouble ensued in recognizing Move Head (83%), Hug (82%), Hug and Pat Back (81%), and Touch All Over (79%). Large confusion resulted from Hug and Pat Back, and Hug (17%); the physical similarity between these touches, along with nearly identical perceived affection averages, suggests that they might best be merged. Also some confusion resulted from Stroke Cheek and Move Head (7%), and Pat Head and Rub Head (5%). Some further information is displayed in the confusion matrix in Figure 27.

	Kiss	Hug and Pat Back	Hug	Touch Foreheads	Stroke Cheek	Side Hug	Rub Back	Rub Head	Shake Hand	Pat Back	Run Hand Down	Pat Shoulder	Pat Head	Pull Arm	Move Head	Touch All Over	Cover Face	Minimize Touch	Push Chest	Slap Cheek
Kiss	92	0	0	3	0	0	0	1	1	0	1	0	0	0	0	1	0	0	1	0
Hug and Pat Back	2	81	16	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Hug	0	17	82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Touch Foreheads	1	1	3	93	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
Stroke Cheek	0	0	0	0	93	0	0	0	0	0	0	0	0	0	3	1	1	1	0	1
Side Hug	0	0	0	0	0	98	0	0	0	1	0	1	0	0	0	0	0	0	0	0
Rub Back	1	0	1	2	0	0	92	0	0	4	0	0	0	0	0	0	0	0	0	0
Rub Head	0	0	0	0	0	0	0	90	0	0	0	0	5	0	1	1	3	0	0	0
Shake Hand	0	0	0	0	0	0	0	0	96	0	0	0	0	4	0	0	0	0	0	0
Pat Back	0	1	0	0	0	0	3	0	0	94	0	0	0	0	0	1	0	1	0	0
Run Hand Down	0	0	0	1	0	0	0	0	0	0	87	1	1	2	4	4	0	0	0	0
Pat Shoulder	0	0	0	0	0	0	0	0	0	0	3	94	0	0	0	1	0	1	0	1
Pat Head	0	0	0	0	0	0	0	4	0	0	1	1	94	0	0	0	0	0	0	0
Pull Arm	1	0	0	1	0	0	0	0	4	0	0	0	0	91	0	2	0	1	0	0
Move Head	0	1	0	0	7	1	0	3	0	0	0	1	0	0	83	0	3	1	0	0
Touch All Over	0	1	1	1	0	1	3	0	1	0	3	1	0	2	5	79	1	0	1	0
Cover Face	0	0	0	1	4	0	0	0	0	0	0	0	0	0	0	1	92	1	1	0
Minimize Touch	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	93	4	0
Push Chest	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	95	0
Slap Cheek	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0	1	0	1	1	94

Figure 27. Confusion matrix for touch and vision system.
Our best-performing system which used Support Vector Machines: the three highest accuracy touches are shown circled in red; the three lowest accuracy touches are circled in green; high confusion is indicated in outlined rectangular cells.

(2) In regard to the complementary nature of the information yielded by both touch and vision-based systems, it appears each had advantages and disadvantages. The confusion matrices for the case of each sensor modality alone are shown in Figure 28 and 29 respectively. In the case of touch data only with SVMs, Push Chest (88%) and Pat Head (85%) were accurately recognized in general. Push Chest, Hug, and Rub Back were more accurately recognized than with the vision-system (+12, +9%, and +8% respectively). However, difficulty was encountered for Hug and Pat Back (64%), Touch All Over (61%) and Kiss (45%). Confusion for the touch system arose due to soft touches not confined to a single location: Kiss involved Kirin's mouth, cheeks, or forehead; Minimize Touch was confused with many other touches such as Pat Shoulder (11%), Kiss (11%), Touch Foreheads (12%) and Shake Hand (17%); and soft slaps, stemming from a reluctance to strike the robot hard in some participants, were confused with Stroke Cheek (13%).

	Kiss	Hug and Pat Back	Hug	Touch Foreheads	Stroke Cheek	Side Hug	Rub Back	Rub Head	Shake Hand	Pat Back	Run Hand Down	Pat Shoulder	Pat Head	Pull Arm	Move Head	Touch All Over	Cover Face	Minimize Touch	Push Chest	Slap Cheek
Kiss	45	0	3	5	14	4	0	0	1	2	2	1	1	2	0	5	1	11	0	4
Hug and Pat Back	3	64	11	0	0	5	4	0	0	5	0	2	0	1	0	2	0	2	1	0
Hug	0	12	77	0	2	2	0	0	0	1	1	1	0	1	0	1	0	1	1	0
Touch Foreheads	0	0	1	66	0	4	0	0	0	1	3	5	0	0	0	0	8	12	0	0
Stroke Cheek	8	1	1	0	67	0	0	0	3	0	0	0	0	0	0	1	1	5	0	13
Side Hug	0	3	0	0	0	66	3	0	1	5	1	6	0	2	0	5	1	7	0	0
Rub Back	0	1	0	0	0	3	78	0	0	7	1	0	0	0	0	5	0	5	0	0
Rub Head	1	0	0	1	0	0	0	76	0	0	0	0	14	0	1	3	0	4	0	0
Shake Hand	2	0	1	0	0	0	0	0	73	0	1	0	0	4	0	1	1	17	0	0
Pat Back	1	4	0	3	0	3	8	0	1	74	0	0	0	1	0	0	0	5	0	0
Run Hand Down	4	0	0	0	0	5	0	0	0	0	66	4	2	2	5	6	1	5	0	0
Pat Shoulder	4	0	0	1	0	1	0	1	1	0	5	71	0	1	0	3	1	11	0	0
Pat Head	1	0	0	0	0	0	0	9	0	0	5	0	85	0	0	0	0	0	0	0
Pull Arm	1	0	0	1	0	0	2	0	4	0	0	4	0	78	0	3	0	6	0	1
Move Head	3	0	0	1	6	0	1	3	1	0	5	0	1	0	69	1	1	4	0	4
Touch All Over	3	1	4	1	0	1	4	2	1	1	5	2	0	2	2	61	0	4	6	0
Cover Face	0	0	0	10	5	1	0	0	1	0	0	0	0	0	1	1	76	5	0	0
Minimize Touch	3	1	1	1	1	0	2	0	5	0	4	0	0	1	0	1	0	79	1	0
Push Chest	0	2	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	7	88	0
Slap Cheek	4	0	0	0	9	2	0	0	1	0	1	0	0	0	1	4	0	0	0	78

Figure 28. Confusion matrix for touch system.
Using Support Vector Machines: the highest accuracy touches are shown circled in red; the three lowest accuracy touches are circled in green; high confusion is indicated in outlined rectangular cells.

	Kiss	Hug and Pat Back	Hug	Touch Foreheads	Stroke Cheek	Side Hug	Rub Back	Rub Head	Shake Hand	Pat Back	Run Hand Down	Pat Shoulder	Pat Head	Pull Arm	Move Head	Touch All Over	Cover Face	Minimize Touch	Push Chest	Slap Cheek
Kiss	86	0	1	2	0	0	0	0	1	0	4	0	0	0	0	5	1	0	0	0
Hug and Pat Back	1	63	30	0	0	1	3	0	0	1	0	0	0	0	0	1	0	0	0	0
Hug	3	26	68	0	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0
Touch Foreheads	2	1	1	82	0	0	0	0	1	1	0	3	1	1	0	5	0	1	1	0
Stroke Cheek	0	0	0	0	71	0	0	2	0	0	0	0	0	0	6	4	15	0	1	1
Side Hug	0	0	0	0	0	92	2	0	0	5	0	0	0	0	0	1	0	0	0	0
Rub Back	3	3	0	2	1	1	70	0	0	10	0	1	0	0	2	6	1	0	0	0
Rub Head	0	0	0	0	3	1	0	81	0	1	0	2	8	0	4	0	0	0	0	0
Shake Hand	0	0	0	0	0	0	0	0	88	0	0	0	0	9	0	3	0	0	0	0
Pat Back	0	0	0	1	0	5	9	0	0	81	0	2	0	0	1	1	0	0	0	0
Run Hand Down	2	0	0	1	0	1	0	0	0	0	75	0	0	3	6	11	1	0	0	0
Pat Shoulder	1	0	0	1	0	0	1	1	0	0	4	82	1	0	3	2	1	3	0	0
Pat Head	1	0	0	0	0	0	0	13	0	0	1	1	79	0	4	0	0	0	0	1
Pull Arm	4	0	0	0	0	0	0	0	3	0	2	0	0	86	1	2	0	1	1	0
Move Head	0	0	0	0	9	1	4	6	0	1	1	0	3	0	65	3	7	0	0	0
Touch All Over	5	2	0	4	1	0	4	0	1	2	6	2	0	4	2	65	2	0	0	0
Cover Face	0	0	0	1	14	0	0	3	0	1	1	0	0	0	5	0	68	2	4	1
Minimize Touch	0	0	0	0	0	0	0	0	0	1	0	3	0	1	1	0	1	81	10	2
Push Chest	1	0	1	1	3	0	0	0	1	0	1	3	0	0	0	1	5	5	76	2
Slap Cheek	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	3	2	92

Figure 29. Confusion matrix for vision system.
Using Support Vector Machines: the highest accuracy touches are shown circled in red; the three lowest accuracy touches are circled in green; high confusion is indicated in outlined rectangular cells.

In contrast, the vision-based system accurately recognized Side Hug (92%), Slap Cheek (92%), and Shake Hand (88%), and performed better than the touch-based system at recognizing touches involving a unique preparatory phase (e.g., Slap Cheek or Side Hug: +14%, +26%) and a unique part of the human such as facial touching (e.g., Touch Foreheads or Kiss: +16%, +41%). Yet difficulty was experienced with Touch All Over (65%), Move Head (65%), and Hug and Pat Back (63%). The problems stemmed from what the vision sensor could not detect. For example, when a participant touched the robot for a long time, the sensor had difficulty distinguishing which part of the contact area belonged to the participant and which part to the robot; this resulted in confusion between long renditions of Stroke Cheek and Cover Face (14%) or Move Head (9%). Another example is when the sensor could not see the person’s hand behind the robot’s back; this led to confusion between Hug, and Hug and Pat Back (30%).

Another thing we noted was that affectionate touches could be recognized by the touch-only system at rates much higher than random chance (5%) even though we had expected some affectionate touches to be quite soft. We think this was because affectionate touches were soft but not too soft; as Hall and Allin proposed in the literature, highly soft touches can be unpleasant and ticklish [154].

(3) The better-than-expected performance of the vision system suggested that, depending on the desired interaction, vision can be a fast, cheap, and simple alternative to traditional touch sensor-based systems. For example, a vision system might be useful if occlusions are not expected and a robot’s location is readily ascertained. However, we note that using both sensor modalities offers best accuracy, suggesting the wisdom of drawing insight from the way humans recognize touches, if costs can be met. Table 6 summarizes our results using the hybrid touch/vision system with SVMs: typical touches, associated average affection scores, and recognition accuracy.

Along the side, we also explored an alternative option to inferring affection from recognizing touches. Average linear error using SVRs and all data was computed to be 0.13, indicating that a system could also recognize affection directly, without using the subjective categories we identified for touches.

Table 6. Affectionate touches summarized
Average attributed affectionate meanings between zero and one (standard deviations in parentheses) and recognition accuracies for typical touches toward a humanoid robot form

Touch	Affection	Recognition	Touch	Affection	Recognition
Kiss	**0.95 (0.076)**	92%	Run Hand Down	0.60 (0.22)	87%
Hug and Pat Back	**0.93 (0.087)**	81%	Pat Shoulder	0.57 (0.16)	94%
Hug	**0.92 (0.091)**	82%	Pat Head	0.56 (0.14)	94%
Touch Foreheads	0.82 (0.16)	93%	Pull Arm	0.50 (0.19)	91%
Stroke Cheek	0.81 (0.11)	93%	Move Head	0.38 (0.21)	83%
Side Hug	0.76 (0.14)	**98%**	Touch All Over	0.29 (0.16)	79%
Rub Back	0.75 (0.14)	92%	Cover Face	0.24 (0.21)	92%
Rub Head	0.75 (0.18)	90%	Minimize Touch	0.085 (0.11)	93%
Shake Hand	0.66 (0.10)	**96%**	Push Chest	0.078 (0.10)	**95%**
Pat Back	0.60 (0.085)	94%	Slap Cheek	0.039 (0.16)	94%

Bold font indicates highest three values. Highlight indicates lowest three values. Attributed affection values were obtained from subjective measurements by 17 participants during acquisition of touch and vision sensor data. Recognition accuracies are for the best-performing system using Support Vector Machines and both touch and vision sensor data.

However, we feel that both sources of information will be useful in informing a robot's responses during affectionate interaction. By recognizing both how a person is touching and what the touches mean, robots will be able to engage in affectionate interactions with people. We expect this to contribute to people's social well-being, based on the evidence presented in Chapter 1.1 which suggests affection and well-being are linked.

Chapter 6

Affectionate Play via Flight (Angel)

"... all eyes were gratified with seeing it rise majestically from above the trees, and ascend gradually above the buildings, a most beautiful spectacle."

Benjamin Franklin, 1783

of the first manned ascent of a hot-air balloon

"There shall be wings! …man …shall have wings"

Leonardo da Vinci, 1452-1519

The goal of the fourth study was to find an affectionate strategy; our work on this topic has been presented previously [4, 6]. Although touching is effective for communicating affection, there is a considerable gap between children and pets, which can get close to people to initiate interactions, and current robots, whose proactive capabilities are typically highly limited. This is a problem especially for our work, because people will expect a humanoid robot to be capable of behaving in a biological, socially intelligent manner.

We attempted to gain some better insight into this problem. In humans affectionate communication occurs mutually and may be initiated by either party. For example, when seeking affection, infants may demonstrate separation anxiety and produce "social releaser" behavior such as crying or smiling; they may also seek to approach their attachment figures themselves to seek proximity [188]. It has also been proposed that infants may bond less with the persons who spend most time with them, raise them, or change their diapers, and more with the people who play and communicate with them appropriately [189]; this suggests that people will expect infant-like robots as well to seek proximity with the people who play with them.

Such behavior is observed with animals as well. Harlow and Zimmermann described how frightened monkeys ran to and clung to “mothers” constructed of cloth, to whom they felt an affectionate bond [35]. Herrick reported on how cats can travel great distances to be with loved ones [190]. Informed by these descriptions of how people and animals do not merely wait for affection but also proactively seek it, we wished for a robot to be capable of doing the same. By not forcing people to come to a robot, but by itself being capable of approaching and accompanying, we think a robot can facilitate affectionate exchange.

A great challenge is that current legged and wheeled humanoid robots are typically not designed to move about in cluttered human environments featuring stairs, barriers, or objects lying on the ground. In fact we may not even want such robots to touch or even come close to our possessions; we can imagine robots thumping under tables, bumping bottles, and rolling over our groceries. Such a case is depicted visually in Figure 30.

Thus, in this study, we based our logic on the following suppositions:

- One way to realize excellent mobility for a humanoid robot could be to provide it with flight capability, which would allow it to fly over problematic ground obstacles.
- Because this idea has never been tested before, a number of challenges arise: we must verify (1) as a proof-of-concept that a flying humanoid robot can even be made; (2) how such a humanoid robot can safely move and position itself relative to a human while interacting; (3) how a humanoid robot will be perceived if it flies (we call these latter two new problems z-proxemics and z-kinesics respectively).

The first point, in regard to the usefulness of flight in a cluttered environment, follows from elementary physics: objects under the influence of gravity tend to occupy positional equilibria which minimize their potential energies, leading to more objects on floors than on walls or ceilings. The second point is justified by a number of supporting arguments. (1) The humanoid form is not aerodynamic; but if we cannot make a flying humanoid robot, then the usefulness of our work is lost. (2) Safety is possibly the greatest concern in introducing robots into human environments; if safety is not perceived people will not accept flying humanoid

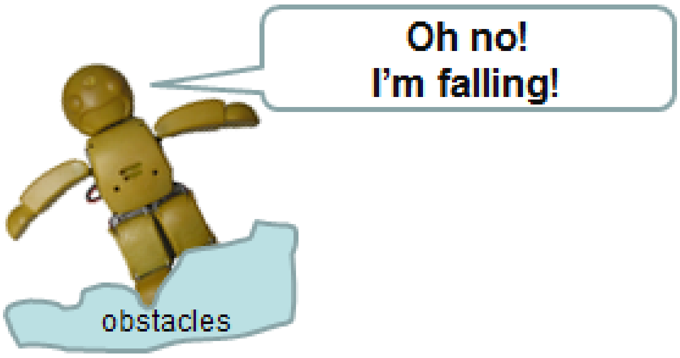


Figure 30. Approaching is difficult in environments with obstacles.

robots as a new technology. (3) If flying introduces unintended communicational cues interactions could fail, as we noted in reporting failed interactions involving meaningless motions in Chapter 4.1. By investigating these points, in regard to the user case described in Chapter 1.2, we believe we can allow a robot like Gozilla to come close to a person like Masha to engage in affectionate interactions.

Because the idea of using flight with a companion robot has received little to no investigation in the past, some ambiguity still exists with regard to how such a robot is expected to fly in proximity with a human. To clarify the usefulness of the scenario we envision, we present another user scenario based on a fictional persona, as we did previously in Chapter 1.2, as follows:

> Gordo's life seemed to be nothing but work; at night, he could sometimes hear the neighbors talking and laughing, which made him feel lonely. As a child he'd always owned birds, which he thought were great; but, as an adult who sometimes had to travel, he couldn't guarantee that he would always be there to take care of a bird, feed it, clean it, and give it company. That was when the robot companion boom started. Gordo bought a small humanoid flying robot and it was great. He decided to call it Swoop because of the way it flew over to him the first time he saw it. Like a bird it looked beautiful, could recognize him, and could fly over to sit close to him. Because of its human-like appearance, it could also do some fun things that a bird couldn't like smiling, nodding, giving him a hug, or even bringing him something light like a pen or a ball. He could really understand what it was feeling because it was so expressive. As a robot, Swoop could also do some other useful things for him, like waking him up in the morning, telling him the weather forecast and his

schedule for the day, suggesting what he could wear, and telling him if his family or friends wanted to talk. In addition to playing with Swoop when it came over to him, Gordo also sometimes liked to go out for a bit of fresh air with Swoop on a tether. He appreciated that Swoop didn't have any problems with rough surfaces, slopes and curbs and such, which other robots could have trouble with; it was also nice that if he got lost, Swoop could point out the way for him. If he stayed out a bit late, Swoop could act as a flashlight and fly out so he could see road signs; or even go and take photos and then show him. Also Swoop could deal with rain and water so he didn't have to worry about it. Gordo felt better because Swoop was with him.

This user case scenario involves many different problems, some of which were also mentioned in the user scenario in Chapter 1.2. Some problems are in the process of being tackled by other research groups, such as talking in a seemingly intelligent way [191] and recognizing people [112-113]. We think the two fundamental problems associated with flight are ones of proxemics and kinesics, for how a flying humanoid robot such as Swoop can approach and accompany a person like Gordo. That is to say, safety is a form of minimal requirement [10]. A robot must be able to not only circumvent obstacles but also distance itself in a safe manner when moving around a human. Also, a robot should be able to move in a way which appears otherwise appropriate in terms of communicational cues. Thus two topics were selected to be investigated as follows:

- *Z-proxemics* (Chapter 6.1): A flying robot should be capable of positioning itself near a stationary human and moving with a human in a clearly safe fashion.

- *Z-kinesics* (Chapter 6.2): A flying robot should be capable of flying in a way which is appropriate in terms of how it is perceived by people, avoiding unintended cues and providing a consistent and expressive impression.

As noted, the described scenario involves facing many different challenges which must be surmounted. The two problems we focused on represent only one central part of this, which required investigation to advance possibilities within the designated topic.

6.1 Z-Proxemics

Our concept for this study would be invalidated if a flying humanoid robot could not move around people in such a way as to convey a sense of safety. This includes flying in such a way as to keep an appropriate distance from both stationary and moving persons. As such, it is a question of proxemics. However, whereas proxemics research commonly concerns itself with positions in two dimensions (e.g., x, y positions on a floor map), we must now consider the third dimension of height; we call this new problem, "z-proxemics". Extending existing theory is not trivial, even at prima facie. It is evident that people can adjust a distance in two dimensions even if a robot's chosen position is not entirely comfortable. In z-proxemics, a person's ability to adjust the distance becomes more restricted; causing a person to constantly jump or maintain a ducked position would not be comfortable and should not be required. In this sub-chapter, we propose extensions to existing theory, which are verified in Chapter 6.3. Two scenarios are dealt with: in one, a robot seeks to position itself in the vicinity of a human partner who is not moving; in the second, the robot seeks to accompany, without colliding with, a human partner who is moving.

6.1.1 Approaching a stationary human

Positioning tasks may be conceptualized as balancing attractive "sociopetal" and repulsive "sociofugal" forces, in order to facilitate the sensing and transmission of communicative behavior while avoiding restricting a partner's capabilities or posing danger. At the same time, it is usually desirable to minimize the cost of moving, unless there is some reason to engage in so-called costly-signaling (e.g., "showing off"); here we will not consider this concern. Thus, the simplest solution to model such a situation in two dimensions might involve imagining a circle around a person, as described by Hall [165]. In this case, a robot could stand anywhere along the circle. Moving within the circle might be too intimate or unsafe. Moving outside the circle might hinder communication. A slightly more complex solution might employ an annular model or probability gradient (indicating an acceptable region and possibly not merely a binary condition), but for simplicity we also do not consider this here. In human science, Kendon observed some typical positions (equilibria) where people tend to position themselves relative to another: “vis-à-vis”, “L”, or “side-by-side” [166]. This could result in a simple circle model with markers to represent common positions. Then, the simplest extension of this model to three dimensions could result in a sphere about a person; with no additional information, stripes could connote common positions.

We find some problems with such a model. First, proxemics researchers are uncovering anisotropies in people’s perceptions of distances; for example, people may wish a robot to be somewhere other than behind them [167], and approaching from their front sides may be preferred less than approaching from the side [168]. This suggests that there might be similar directionality preferences in three dimensions. Also, in terms of finding positional equilibria, stripes might not constitute an appropriate modeling. Based on this reasoning, we sought to identify possible errors in our simplest model. To address these concerns we proposed a revised model as described below and shown in Figure 31.

Prediction 1: Shape of the proxemics contour: A sphere-shaped model is not appropriate for representing humans (and larger humanoid robots) in three dimensions, because the humanoid form is long in the superior-inferior or “z” axis and short otherwise. Thus, a cylindrical model is possibly more appropriate. This can be seen by imagining a robot with its center of mass at the top of the sphere. Its legs, draping downward, could kick a person’s head. Such a scenario could pose significant threat, because the head contains our most important sensing and processing areas (the eyes, ears, nose, mouth, and brain). In fact, it could be the case that people do not want a robot flying above their heads, even at a distance, because technical trouble could result in injury if the robot falls or by accident drops something on them. Even beyond the dangers posed, people may not wish a robot’s feet to touch their heads because in some cultures the feet are regarded as unclean.

Although the head is perhaps the most significant reason for employing a cylindrical rather than a spherical model, it is not the only reason. A small robot too close to a human’s feet poses a danger of colliding or tripping a person if the person moves, and could be hard to see from the perspective of the person.

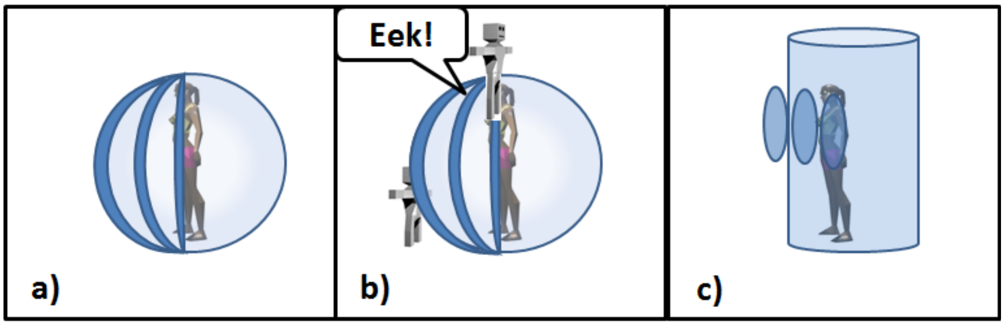


Figure 31. Positioning for a flying robot when approaching a human.
a) a simple sphere model (stripes represent F-formation equilibria where a robot could position itself), b) problems with the sphere model (a robot could get too close to a person's head or feet and its face might not be visible) c) our proposed cylindrical model with good positions for a robot shown as ellipses.

In addition with our prediction that a cylindrical shape is more appropriate, we also modify the shape based on the sources noted to deter a robot from positioning itself behind a person, and allow closer distances at the side than at the front.

Prediction 2 Equilibria: We predict that non-central or extreme positions along "stripes" could be inconvenient for communication because a robot's face and motions might not be visible. If not visible, these communication cues serve no purpose and communication is obstructed. Moreover, forcing a person to bend their heads can cause pain [192]. Thus, we predict that a robot should usually position its face at approximately the same height as a person's face. We also expect differences in height to also convey some attitudinal meanings [148] and therefore advise that a robot's height should be calculated with care to avoid presenting unintended cues.

6.1.2 Accompanying a moving human

The predictions made for the case of a stationary human are employed to extend an established model for collision avoidance. If these predictions are found to hold true, then the logic for this extension will follow. In general our scenario for a robot accompanying a human is as follows. A robot should keep a distance of approximately one meter from a person's side; the angle from navigation goal to human to robot should be approximately 90 degrees. Its velocity should be comparable to the human's; this would typically entail a speed of approximately one meter per second. To present a natural impression of focused attention, a robot's gaze and orientation may be directed at any given time toward the navigation goal, the human, or areas of saliency (for example, areas in view featuring sudden motion or significant colors such as skin tones or rare colors, or in the direction of sudden sounds). Some information regarding the two dimensional case of a robot accompanying a person has been presented in a previous study [193].

Real human environments where we would like to introduce robots feature additional challenge. It may not be possible for a robot to maintain its velocity and position relative to a person due to obstacles including other people, robots, and objects (either stationary or moving without their own intentions). In our work, we seek to employ familiar human cues to realize rich communication; in locomotion as well, a human-like capability to avoid collisions is desired. In other words, people should not be required to work to avoid robots; and people accompanying a robot should be able to predict how it will move to avoid collisions. Toward this, we selected a collision avoidance strategy called the Social Force Model (SFM) which has been employed to model people's walking behavior [169]. This model describes pedestrian locomotion in terms of a basic attractive force toward a navigation goal, resulting in motion at a specified speed and direction. Perturbations occur due to repulsive forces generated by other pedestrians and obstacles. The simplest form of the SFM, known as the "Circular Specification" (CS), which was originally proposed to describe how persons moved during an evacuation scenario, mathematically described such forces as circles. In order to allow a robot to use its ability to move in three dimensions, we may attempt to simply extend this model to a Spherical Specification (SpS):

$$\boldsymbol{f}_{rh}^{SpS} = Ae^{-d_{rh}/B}\frac{\boldsymbol{d}_{rh}}{d_{rh}} \tag{5}$$

where a robot experiences a repulsive force away from a pedestrian's center of mass with A as the "interaction force"; B as the distance sought for protection; and d_{rh} as the magnitude of a three dimensional vector $\mathbf{d_{rh}}$ from the pedestrian's center of mass to the robot's center of mass representing how far apart they are.

This model is simple to understand, but because it was designed specifically to model natural human behavior in an evacuation scenario it is not optimal for the ordinary case we envision, of a robot accompanying a person. For example, if two robots approach each other head-on, according to this model they would both experience a repulsive force which would cause them to slow down and halt, which is undesirable. We would prefer the robots to be capable of moving to the side to avoid one another while retaining their speed and continuing to move toward their destinations. A slightly more complex specification, Collision Prediction (CP), could be used to model such behavior. This specification instead of using an actual current sensed displacement vector uses the predicted displacement vector **d′(**t′**)** at a future time t′ in which the robot and pedestrian find themselves at their closest distance d′(t′). This can be extended to the three dimensional case, which we call 3CP, with the following equation:

$$\boldsymbol{f}^{3CP}_{rh} = \frac{v}{t'} A e^{-d'_{rh}/B} \frac{\boldsymbol{d}'_{rh}}{d'_{rh}} \tag{6}$$

The initial term, v/t′, where v is the robot’s speed, allows a robot to decelerate quickly when a collision seems imminent; the two dimensional version of this specification is described in detail in previous work [170]. Our extension simply takes into account a three dimensional predicted displacement vector between centers of mass instead of a two dimensional one.

We propose one additional modification to take into account our predictions from the preceding sub-chapter. Although a circle can be used to represent humans walking along a flat surface in two dimensions (the x and y axes), we believe that this approximation breaks down when we consider the third dimension (the z axis). This is because a standing human is much taller than they are wide. If the sphere around a person’s center of mass representing repulsive force is too small, a robot could collide with a person’s head or feet, causing injury. On the other hand, if the sphere around a person is large enough to protect the head, this could cause a robot to make unreasonably large detours to the side, which may not be possible in corridors or narrow spaces. Thus, as before, we advise a cylindrical model about the z axis traversing a person’s center of mass, in place of a sphere model, to ensure a robot moves in a safe, natural fashion.

In the event of any trouble, the robot should furthermore at no account collide with a person’s head, which could cause blindness or severe injury. To protect this area especially, we suggest the use of a second repulsive force from the center of a person's head to the robot's center of mass. The Head-Body Collision Prediction (HBCP) specification follows from this logic:

$$\boldsymbol{f}^{HBCP}_{rh} = \frac{v}{t'}\left(A^{H} e^{-d'_{rH}/B^{H}} \frac{\boldsymbol{d'}_{rH}}{d'_{rH}} + A^{Z} e^{-d'_{rZ}/B^{Z}} \frac{\boldsymbol{d'}_{rZ}}{d'_{rZ}} \right) \quad (7)$$

Where $\mathbf{d'_{rH}}$ and $\mathbf{d'_{rZ}}$ are predicted displacement vectors from the center of a person's head to the center of a robot in three dimensions and from the center of a person along the z axis to the center of a robot in two dimensions; A^H, A^Z, B^H, and B^Z are values specific to forces generated from the head or body (z axis). Figure 32 shows the different specifications.

To implement this model in an actual robot, the robot should minimally be able to recognize its location relative to a human's and the human's orientation, in order to be able to approach a stationary human. For accompanying, the robot should also be able to recognize velocities. The proposed model can be altered to allow a robot to drift forward when it wishes to be easily seen by a person (e.g., when it is gesturing, displaying emotions, or expressing some turn-taking behavior). Objects such as bags, balloons, and umbrellas could be modeled by a spherical repulsion force weaker than that attributed to humans (because human safety should be most important). We also suggest incorporating and extending a wall avoidance calculation, described in the CP specification, to allow a robot to avoid walls, floors and ceilings.

Thus we have proposed a new specification which will allow robots to know how to move in a natural human-like way to avoid collisions while accompanying humans; this specification is based on previous work and a number of predictions we made, which we verify in Chapter 6.3.

6.2 Z-Kinesics

Movements convey meanings [194]. In order to avoid signaling undesired, potentially confusing cues, how people perceive a robot's motions should be investigated. The patterns found will also be useful to realize rich communication. Thus we wished to address a problem of kinesics. In particular, incorporating

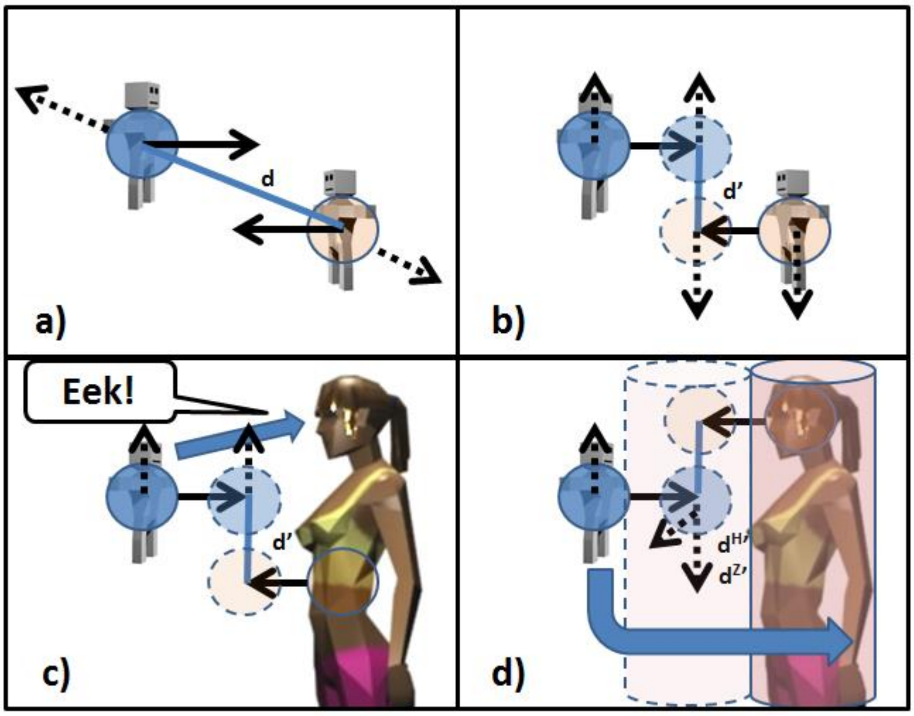


Figure 32. Avoiding collisions in three dimensional flight.
Modeled using a social force model: a) a simplest model (SpS) would result in inefficient flight in some cases, b) such a problem could be solved by modifying the model (3CP) c) another problem affects 3CP in which a robot could collide with a person's head, d) our proposed model using a cylinder and sphere to ensure a robot's flight near people is safe.

flight expands the set of body motions whose perceived meaning must be determined: a flying robot can perform both body language cues possible for non-flying robots and some others which would not be possible. We refer to this new kinesics study as "z-kinesics", because motion along the z axis becomes less restricted. Our investigation entailed first identifying fundamental representative motion primitives and predicting how they will be perceived; evaluation is conducted in the next chapter.

The approach we followed, inspired by the literature described in Chapter 1.3, involved identifying both primitive motions [171] and important describing properties of motions [172] then predicting meanings which included, but were not limited to, emotional expression [125]. Gestures such as waving which have already been investigated for the standard non-flight case and do not change as a result of flight were not considered. We selected as the set of primitive motions all rigid body movements (translations and rotations) about a flying robot's center of mass, as these put together may describe any motion of an entire body in three dimensions. We call the three primitive translations along a humanoid robot's dorsoventral, mediolateral, and superior-inferior axes "surge", "sway", and "heave", and the corresponding three rotations about these axes "roll", "pitch", and "yaw". These six primitive motions can express qualitative information about the manner of motion including how a robot's orientation changes, but do not describe the basic quality of the posture of a robot or the timing of its motion. This information may be conveyed by some other important descriptors. For the former quality, objects moving in three dimensions may or may not display orientational equilibria. For example, a thrown object may be constantly spinning. An airplane moving toward a destination has an orientation which can be described in terms of its "heading", "trim", and "wing level". Thus, intentional motion seems to be non-random in orientation; therefore we sought to investigate this latter case, under the assumption that a robot will also have intentions. We hypothesized that a humanoid form which moves would be likely to adopt a standard human posture such as standing, sitting, or lying. In addition to orientation, how a flying motion occurs can be described by displacement, velocity, and acceleration. These descriptors can be used to distinguish quantitatively large, fast, and jerky flight from small, slow, and smooth flight in only a few dimensions. Thus, jittery insect-like flight and swooping avian-like flight can be described. Primitive motions, orientational equilibria, and quantitative descriptors are shown in Figure 33.

Based on the target motions to investigate, we formed three predictions.

Prediction 1 (Primitive motions): Upward heave will express emotionally positive valence and dominance; downward heave will express depression and submissiveness. Swaying flight will seem irritating or playful in the sense that the robot may be seen as trying to obstruct a person. Pitch will express concurrence by appearing as a bow or nod. Roll will express ennui or indecision. Yaw will express contention by appearing as if the robot were shaking its head.

Prediction 2 (Orientations): Postures other than standing will be seen as odd or comical; a reclining posture will also appear lethargic and unexcited; standing will not express any meaning.

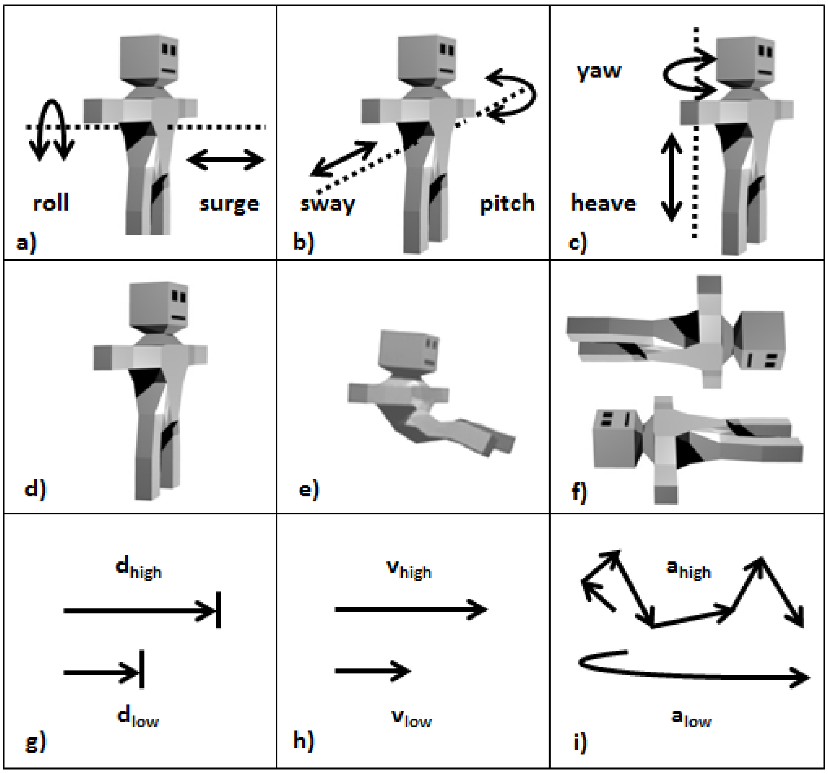


Figure 33. Representative types of flight for a humanoid robot.
a-c) primitive translations and rotations, d-f) typical humanoid postures (standing, sitting, prone, supine) g-i) manner of flying: high or low displacement, velocity, acceleration.

Prediction 3 (Quantitative descriptors): High displacement, velocity, and acceleration will express excitement.

In summary, we predicted that flight motions will allow a humanoid robot to express playfulness, arousal, and happiness in a new manner; these predictions needed to be checked.

6.3 Evaluation

To check predictions regarding z-proxemics and z-kinesics made in Chapter 6.1 and 6.2, we obtained feedback from 10 young adult participants (6 females and 4 males; average age = 29.4 years, SD = 7.8 years).

6.3.1 Approach

Three problems were faced. (1) We did not know how participants would feel about a flying humanoid robot's flight; therefore, we could not use standard techniques such as forced-word choices and specific questionnaires. (2) Our prototype, Angel, had a highly unique appearance which could cause participants to perceive motions in a different way from robots of more typical appearance. (3) Angel could not perform some motions such as heave (to ascend or descend Angel must fly forward while altering its center of mass). To deal with the first problem we adopted the Think Aloud method, which allowed us to gain insight into how people felt without knowing what categories existed beforehand [195]. To deal with the second and third problems, we used a set of animations. This allowed us to keep the simulated robot's appearance highly generic (for example, we do not specify how the robot is flying); and the simulated robot could perform exactly the motions we wished to test and in the same way for each participant.

6.3.2 Materials

Six animated clips comprising three clips each to test proxemics and kinesics predictions were watched by participants in random order. The content of the clips is partially shown in Figure 34 and described below:

Animation A1: a flying humanoid robot and person appeared to be communicating, where the height of the robot caused its face to be (a) hard to see or (b) readily seen.

Animation A2: flying humanoid robots (a) flew near a person's head and overhead or (b) kept their distance

Animation A3: flying humanoid robots (a) flew near a walking person's feet or (b) kept their distance

Animation A4: a flying humanoid robot positioned in front of a person carried out six primitive motions comprising (a-c) translations and (d-f) rotations.

Animation A5: a flying humanoid robot flew (a) standing, (b) sitting, or (c) lying down.

Animation A6: a flying humanoid robot flew with high or low (a) displacement, (b) velocity, and (c) acceleration.

6.3.3 Predictions

The animated clips were used to test a number of our expectations from Chapters 6.1 and 6.2. To refresh the reader's memory, we restate these and indicate to which animated clip they apply:

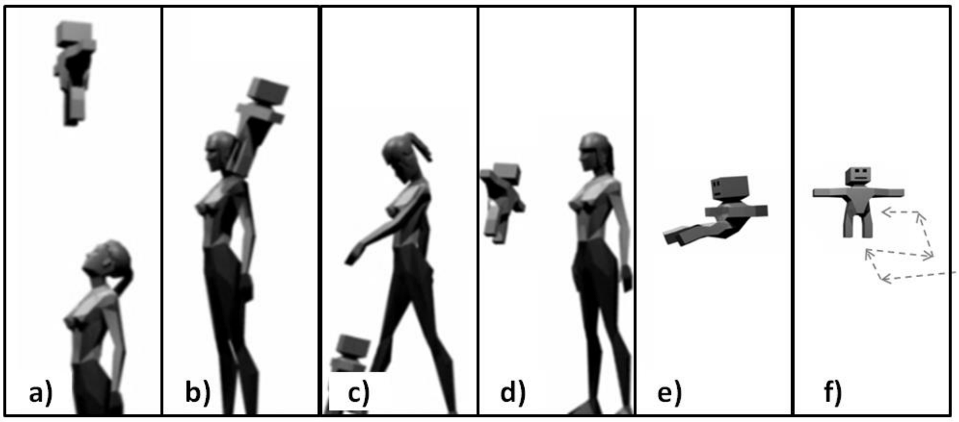


Figure 34. Some screen captures of the animated clips.
a) a robot's face is difficult to see (animation A1) b) a robot flies near a person's head (A2) c) a robot flies near a person's feet (A3) d) robot turns about the roll axis (A4) e) robot sits while flying (A5) f)high acceleration flight (A6).

Z-proxemics

Prediction P1 (Animation A1): A humanoid robot should fly so that its head and a person's head are roughly the same height when interacting so that its interactive behaviors can be clearly and easily observed.

Prediction P2 (Animation A2): A humanoid robot should keep its distance from a person's head and also not fly overhead.

Prediction P3 (Animation A3): A humanoid robot should keep its distance from a walking person's feet.

Z-kinesics

Prediction P4 (Animation A4): Translations and rotations will express emotional qualities such as joyfulness and tristesse, supremacy and submissiveness, concurrence and contention, and facetiousness.

Prediction P5 (Animation A5): Standing will seem ordinary.

Prediction P6 (Animation A6): High displacement, velocity, and acceleration will express excitement.

6.3.4 Procedure

Participants were admitted by the experimenter into a large empty room with a desk and laptop computer. Then participants were instructed that they would observe six animated clips in which a human and a flying robot were interacting. They were asked to continuously say anything that came to their minds as they watched the robot move; for example, how the robot was moving and what it meant. They then observed the animated clips which had been ordered randomly for each participant beforehand, and spoke freely about what they saw. The experimenter allowed them to speak as long as they wished while watching clips and afterwards; they could also pause and replay clips. The experimenter recorded participants' words as they spoke and reminded them to comment on the meaning of what they saw if they forgot. After sessions, an audio record was checked to ensure the accuracy and completeness of the experimenter's transcripts before coding.

6.3.5 Coding

Transcripts, or "protocols", were coded to transform the data into a useful form for understanding how participants perceived the robot's flying motions. This involved two steps: (1) extracting subjective impressions, and (2) gathering similar impressions under a single code label. We present an example of one portion of a participant's voiced thoughts below:

> "Oh my.
> The robots came close, then one passed to the left and
> one flew over the top of a person's head.
> The person didn't get *startled*.
> *Normally you'd crouch down* a bit.
> Thinking they were *going to collide*.
> Oh…!
> The robots came closer (than before).
> That *surprised* me.
> They came very close and avoided at the last moment."

Non-italic text above merely relates what this participant saw. Italicized text represents subjective perceptions, which we were interested in. Of this latter text, we note that this participant used two different words "surprised" and "startled" which have roughly the same meaning. This occurred frequently in various forms: e.g., "robot blocked her", "hard to walk", and "bother"; or "talking" and "communicating". Coding these different words with the same label allowed us to identify typical perceptions which resulted from watching the stimuli (defined here as impressions voiced by two or more participants).

6.3.6 Results

Typical impressions are indicated per animation clip below (numbers in parentheses indicate how many participants described an impression):

A1 Participants felt there was no communication (6) and no eye contact (2) when the robot's face was difficult to see (a), but that communication (10) and eye contact did occur (2) in a friendly way when the robot's head height was the same as a person's (b).

A2 When robots flew near a person's head (a) it seemed scary (6), dangerous (6), disrespectful (2), and mischievous (2); with more distance (b) it seemed safe (3) or that the person and robots were indifferent to one another (2).

A3 When robots flew near a walking person's feet (a), it appeared dangerous (8), abnormal (3), and bothersome (2), although some felt the robots seemed like children (3) because it was a small humanoid moving at a height near the floor at which children could move; with more distance (b) the scene appeared to be safe (6).

A4 (a) When the robot flew upward, it seemed as if it had been commanded to move (3) or was moving to perform a task (2); downward flying as well seemed like the robot had been commanded to do so (4), or that the robot was evading the person or that the interaction had ended. (b) The robot when approaching the person appeared to have been commanded to do so (5), wished to communicate (4), or couldn't hear what the person was saying (2); increasing the distance appeared to result from a person's command (4), lack of desire to interact (3), an interaction having ended (2), fear felt by the robot toward the person (2), or the robot having to perform a different task (2). (c) Sidewise "sway" indicated that the robot wished to evade something (6), had been ordered to move (4), or desired to observe something (3). (d) Nod-like pitch rotations appeared to result from the robot concurring (4), playing (3), saying hello (3), or expressing happiness (2). (e) Yaw rotations in the same direction as head-shaking indicated disagreement (6), playing (2), and looking around (2). (f) Roll rotations like leaning left and right appeared to show happiness (3), playing (2) or a desire for attention (2).

A5 The robot flying while standing (a) was perceived as normal (4) or as if it were walking (3). Flying while sitting (b) seemed like the robot was sliding (5), comfortable (3), funny (2), or hyperactive (2). (c) Flying in a supine position indicted relaxing (5), being pulled (5), or that the robot was deceased (2); a prone posture seemed as if the robot was observing something below itself (3), being pulled (3), sleepy (2), down-spirited (2), or deceased (2).

A6 (a) Low displacement suggested that something had happened while the robot was moving (3); high displacement seemed like the robot had an unfinished objective (3). (b) Low velocity appeared irritating in the sense that it was too slow (2); high velocity seemed frightening (6). (c) Low acceleration flight appeared to show happiness (3) or a gentle disposition (3); high acceleration flight seemed as if the robot were hopping because its direction changed frequently (9), or that the robot seemed joyous (6), strange (2), erratic (2), or possessing some wish to do something (2).

We compared these typical impressions with our predictions as follows. (P1) When the robot was at a height which made its face difficult to observe communication did not appear to occur (A1a), but the person and robot did seem to be conversing when the robot's head and human's head were approximately at the same height (A1b). (P2-3) Robots which neared a person's head and feet (A2a, A3a) were considered bothersome, scary, and unsafe; robots which flew at a distance were perceived as safe. (P4) Translations (4a-c) were not perceived as conveying emotions but instead seen as serving useful functions such as initiating and terminating interactions, which we had not expected; rotations (4d-f) were perceived by some participants as expressing facetiousness or as a nod or head-shake even though the motions performed by the robot involved its full body and not only its head. (P5) Around half of the participants described standing as ordinary (A5a); sitting and lying were perceived as resulting from extraordinary situations such as sliding or being pulled (A5b-c). (P6) Some impressions of large and quick flight suggested excitement: fast flight seemed scary or wild (A6b-c), while smooth flight was seen as gentle (A6c).

In summary, some of the participants' coded comments supported our predictions relating to z-proxemics (P1-3): that robot should fly near head height away from a person's face and feet. Prediction 4, that primitive motions could be used to show valence, dominance, agreement and playfulness, was only partially supported (valence and dominance were not perceived) because translations were not seen to be emotional but rather as serving some useful function. Predictions 5-6 were supported by some comments. Thus, our predictions from Chapter 6.1 were supported, and our predictions from 6.2 were partially supported. This new knowledge of how people perceive a humanoid robot's flight will help designers to construct companion robots with excellent communication and locomotive capability which can behave in such a fashion as to create an impression of safety and intentional consistency; as such, it enables us to move closer to making affectionate play between a person and robot more feasible, toward contributing to people's social well-being.

Chapter 7

Discussion and Conclusions

"You were created by the magicians; return to your dust."
Rav Zeira, in Sanhedrin 65b, c. 6th century AD
to a *gavra* (golem)

"To light a candle is to cast a shadow."
Ursula K. Le Guin, in A Wizard of Earthsea, 1968
(with new answers, come new questions)

This section is structured in three parts: first we summarize again our findings, then we discuss limitations, and finally we propose promising areas for future work.

7.1 Summary of Findings

The current dissertation chronicled some new work toward contributing to a feeling of social well-being in persons interacting with a robot.

In the first study performed, we described a first approach for a robot to recognize enjoyment-motivated touches with fair accuracy in real time. Second, we described a first strategy for a robot to provide enjoyment, in the form of a set of heuristic guidelines indicating how a robot's recognition capability can be incorporated to structure its behavior while playing. Third, we described a first way to enable a humanoid robot to engage in affectionate play with a human; the contribution which allows this is a list of typical touches people perform, along with their affectionate meanings, as well as a description of a system which can recognize touches online. In the fourth study, we investigated a new way for a companion robot to approach a person to show affection via flight. These results are summarized in Figure 35. Details are described below.

- Contributions to developing recognition capability (social intelligence) in humanoid companion robots involved identifying 13 typical proprioceptive motions and 20 typical touches; in the latter case we were also able to obtain subjective measurements of the perceived affectionate quality of each touch. Proposed systems used a three-axis accelerometer and two-axis gyro sensor to capture inertial data; forty touch sensors built from photo-interrupters and soft housings of two different types for recognizing pats and rubs; as well as a Microsoft Kinect device to visually acquire postural data. We reported on features which could be calculated from short windows of raw data. In the first case of enjoyment, we compared three different types of features, finding time-domain statistical features were most useful for our task; we reported 19 features automatically selected by our algorithm. In the second case for affection, we described features used for recognizing touches. In both cases, we used SVMs to classify behavior based on the calculated features. In the first case, our system recognized typical behavior with 77% accuracy which dropped by 21% when the robot moved, altering the data from its sensor. In the second case, we obtained an accuracy of 91% using both touch sensors and the vision sensor.
- Regarding interactive strategies, we determined that interactions with a simplified turn-based design featuring idling and plausible behavior sometimes failed to provide enjoyment. Three common failure patterns were found which involved the robot's motions appearing to lack meaning, the robot not recognizing highly passive interactions, and goal-less moving of the robot. Feedback was used to propose four guidelines: *meaningful motions, rewarding responses, inspiring suggestions,* and *fulfilling instructions*. A maximum and progressive strategy for responses, and a shifting and persisting strategy for suggestions, were evaluated via an experiment, yielding the result that progressive responses were perceived as featuring increased variety, and persisting suggestions aided comprehension, facilitated the observation of variety, and provided enjoyment. Enjoyment was afforded by allowing

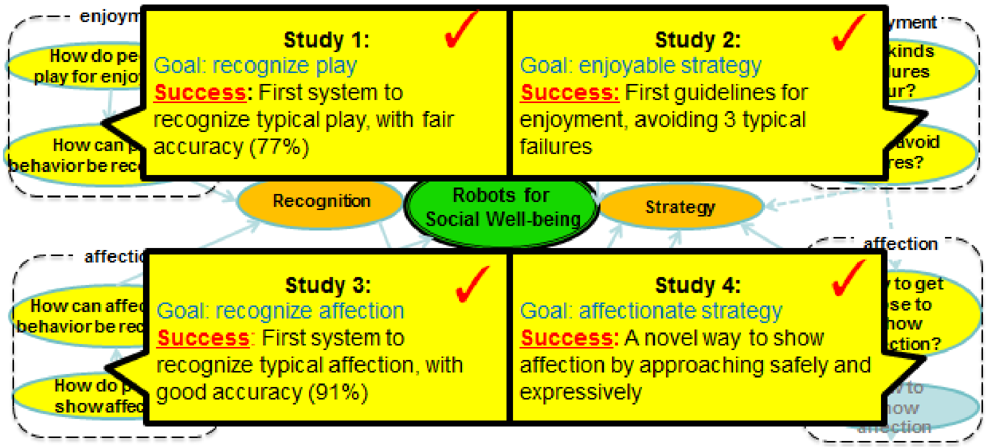


Figure 35. Goals for all four studies were successfully met. (1) recognizing enjoyment-motivated touches (2) enjoyable strategy, (3) recognizing affection-motivated touches, (4) affectionate strategy.

participants the freedom to play with a robot in various ways and the capability to make the robot appear happy as a result of their actions. A second experiment confirmed that the proposed system incorporating the guidelines was perceived as more enjoyable than the initial system. From another perspective, we demonstrated that understanding is vital in motion-based playing with a robot. Participants' feedback in Chapter 4.1.2 indicated that understanding is important in all aspects of an interaction: clear motions, responses unequivocally indicating joy in interacting, suggestions which unambiguously teach, and instructions. The experiment in Chapter 4.2 highlighted particularly the advantages of communicating the robot's desires clearly. Chapter 4.3 indicated that a naïve system may not be understood.

In terms of moving toward a scenario in which a robot is not just a toy which must passively wait for a person, but can proactively initiate affectionate interactions, we proposed one way for a robot to approach a person to show affection via flight. We reported on three important considerations: embodiment, z-proxemics, and z-kinesics. We built a first flying humanoid prototype, Angel, which used a rotating head and arms for communication, and two wings and a center-of-mass changing system to achieve flight. Although merely a simplified prototype, Angel could pass by flight over ground obstacles which would be difficult for previous humanoid robots; and Angel could communicate in a familiar way not possible for previous flying robots by using human cues such as pointing its arms or rotating its head. We proposed how such a flying humanoid robot could move when approaching and accompanying a person; for the latter we extended existing theory via a Head-Body Collision Prediction (HBCP) specification. We also proposed how a humanoid robot's flight motions would be perceived. Simulations and the Think Aloud Method were used to obtain support for our predictions. Participants' thoughts spoken out loud indicated that such a robot should position its face near a person's head-height when communicating, but avoid getting too close to a person's head or feet or flying above a person's head. Some comments indicated that rotations appeared to indicate playfulness and agreement; that a standing posture seemed noncommittal; and that high acceleration and velocity could convey excitement.

Designers can use the above knowledge toward constructing companion robots capable of engaging in and initiating interactions, which we expect will contribute not only to facilitating the introduction of robotic technologies into everyday human environments, but also to contribute to people's social well-being, due to the evidence in Chapter 1.1 which suggests a connection between well-being, enjoyment, and affection.

7.2 Additional Implications

As a result, we also have a deeper knowledge of some important concepts in HRI, which allow us to adapt our definitions. Social well-being is a subjectively perceived long-term state of happiness, which we feel is not only an important concept in human science but also an important goal in HRI.

Enjoyment is an agreeable response involving pleasure, satisfaction, or gratification to a meaningful robotic stimulus which may involve large, positive, progressive responses and clear persisting suggestions, and arises from playing with a robot in various ways and making it happy.

In particular, we feel that the importance of clarity we observed is informative for the design of enjoyable interactions. This is because the previous literature had described both cases in which confusion aided and detracted from playful interactions. For example, Salter and colleagues reported how a mistaken comment by a robot on the ground asking to be put down elicited greater engagement and increased interaction [106]. Turkle and colleagues described both kind and irritated reactions from children in response to a perception of having been ignored by a robot [37]. Our second study added support for an intuition that communicative clarity is important in interactions with robots involving restricted communication capabilities.

Such a trend may also hold true outside of robotics. The fact that people attribute meaning even where none has been intended is well-known [196-197], which could cause confusion from misunderstandings. And, Csíkszentmihályi described a connection between enjoyment and understanding the objectives of an activity [31]. However, as in robotics, the problem is not clear-cut. In some cases, mental discord can also yield positive effects; e.g., people who receive less reward for accomplishing an uninteresting activity may end up with a better impression of it [198]. We suggest that in general, understanding is helpful toward experiencing enjoyment; yet, this trend may be most true in robotics due to the restricted communication abilities of robots, which may invite misunderstandings more than in human-human interactions.

Much other knowledge was also obtained. We found that affection is an important factor involving fond attachment, devotion, or love toward someone, which is conveyed through touch to a humanoid robot in a manner similar to how it is conveyed toward humans, by hugging, stroking, and pressing with appropriate force and location. Furthermore, our results showed that touches could be more accurately recognized by using a vision sensor in conjunction with the traditional approach of using only tactile sensors. In terms of the meanings of a robot’s motions, we learned that people tend to attribute functional meanings to translations and emotional meanings to rotations. Some new knowledge obtained in our work is shown visually in Figure 36.

7.3 Limitations

Our results are restricted by (1) the robot prototypes employed, (2) our specific scenarios, (3) our user demographic, and (4) methodology. (1-2) are shown in Figure 37. (1) Using a robot with a different size or shape and appearance (e.g., not a humanoid robot) may yield different results. (2) Our focus also involved a specific scenario of a touch-based dyadic interaction between a robot and a human

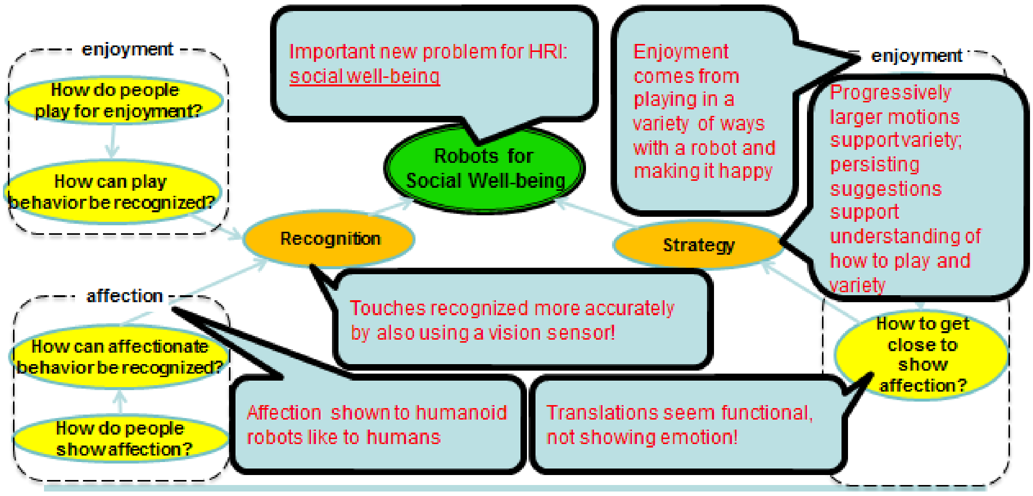


Figure 36. Deeper knowledge of some important concepts in HRI.

conducted without objects (we did not study playing with a ball or other device). (3) We used mainly adult Japanese in our studies. We focused on adults because children's behavior could be unpredictable because they might not have learned rules which adults know, and an accident with an elderly person such as falling could cause serious injury because bone strength and healing could be reduced. Japanese participants were used for convenience because our lab is located in Japan; nonetheless, Japan is one country in which the robotics industry plays an important role, including companion robots for enjoyment such as AIBO. (4) Subjective measurements including questionnaires were often employed in our work, which involve subjective judgements and therefore may not be as convincing to some as objective measurements; however, in our case, only subjective measurements were practical for determining how participants felt.

Despite these restrictions, we expect our findings may offer some understanding for how robots may provide social well-being in other contexts. First, our proposals are scalable; more recognized touches, motions, and sounds can be accommodated. Second, our proposals are independent of degrees of freedom and other hardware. Third, our focus is not restricted to simple prototypes. The interactive modality which we focused on, touch, is fundamental for communication, playing for enjoyment and communication of affection in human-human interactions [95-98]; thus we can expect that people will

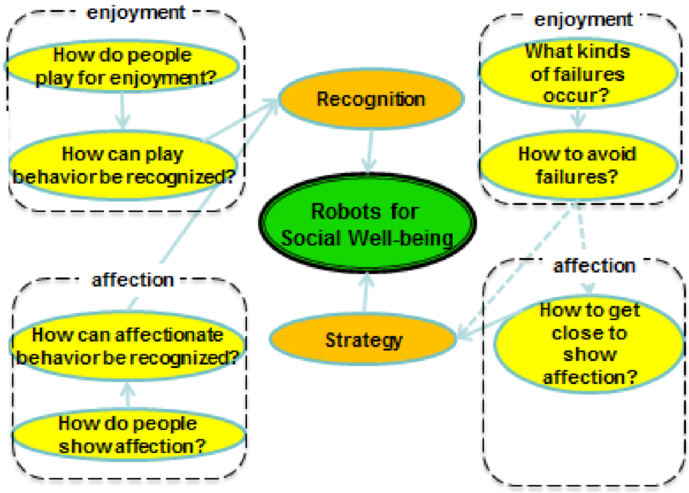


Figure 37. Our results may be applicable to other contexts.
Although we focused on the scenario depicted above, our proposed designs are scalable, hardware-independent, and fundamental.

also touch more complicated human-like robots. Therefore, we expect our findings will apply to other scenarios.

As an example, we describe how existing products described in Chapter 1.3 could gain from the acquired knowledge. Current small robots intended for play could respond to typical behavior such as making the robot walk around or dance; and suggest through motions that this is what the robot wishes a person to do. Affectionate robots could be built which can recognize touches such as placing one's forehead on the robot, stroking its cheek, or pushing it; and know how affectionate their responses should be. Flying versions of previous humanoid robots could approach people in environments where the ground may not be free of objects and indicate emotions in a new way via three dimensional full-body motions. The capability to act in a "social intelligent" and lifelike manner will allow such robots to move beyond effortful interactions managed minutely by a person to one which is natural, rich, and easy for people.

7.4 Future Work

We have conducted some work which is not reported here, in relation to how a robot can elicit affection in a touch-based interaction and how a robot can interpret the meaning of a person's behavior in a multi-modal interaction involving touch, vision, and sound; this is not discussed because it has not yet been published. Our current work involves leveraging all previous results to construct a flying humanoid robot which can recognize multi-modal cues to elicit enjoyment and affection during play.

Due to the highly complex nature of people's perceptions related to well-being, many other important problems remain to be addressed. Future work will involve acquiring objective data such as brain scans, along with many more subjective data. Such data will offer additional support for our hypothesis that enjoyment and affection are correlated with well-being, as well as shed light on how perception of other possible determinants can be facilitated, and how people perceive embodiments, adapted behavior, and other scenarios. To remind the reader, these important topics were not investigated in the current dissertation for the following reasons. Before acquiring additional support for the relationship between enjoyment, affection, and social well-being, we needed to create enjoyable and affectionate interactions. Before investigating other factors, we wished to investigated enjoyment and affection first because we believed them to be fundamentally important. Before investigating topics which have already been partially investigated such as embodiments and behavior adaption, we wished to focus on new problems. And, before investigating more complex scenarios including ones with objects we wished to investigate a fundamental dyadic scenario without objects.

We present some specific examples of important next steps. How to facilitate perception of some other proposed determinants of well-being including health, engagement, meaning, and achievement will be explored. Work on embodiments will involve comparing different shapes, sizes, and appearances of robots, such as comparing people's behavior toward humanoid vs. non-humanoid robots. Adaptation will involve two steps. First we will identify key features which decide whether behavior appears to be enjoyable or affectionate and determine how to combine these in an effective manner. Second, directly recognizing the degree to which a person is enjoying an interaction or feeling affection toward a robot will act as a useful feedback signal for adapting robotic behavior and sustaining long-term interactions, after participants have experienced interactions with the robots and no longer feel effects of novelty. Another enjoyable scenario will involve playing with objects. Our primary future plans are summarized visually in Figure 38.

As noted however, the problem of providing well-being is very complex and, although we have sought to show some of the most important steps, there is no doubt that many other possible topics exist. For example, improvements in base technologies such as batteries and sensors will contribute greatly to making robotic interactions more practical. For all of the projects we conducted, we will have to acquire more results from a greater number of participants of various cultures and ages. Due to the exploratory nature of our work we used our own measurements to gain insight into those factors which we believed most important; subsequent efforts will have to use standard tools for assessing people’s emotional appraisals of interactions, attractiveness, or system usability such as PANAS [199], AttrakDiff [200], and the SUS questionnaires [201] respectively. This will help to compare results with other studies and extend our results by investigating from new perspectives. Field interactions in natural and non-dyadic settings should also be investigated.

Moreover, much work remains to be conducted toward designing companion robots which can approach and accompany humans to facilitate interactions. For flying humanoid robots such as the one we built, controlled flight in the presence of drafts is not trivial. A starting point for dealing with this problem may involve the use of ego-motion detection and anemometers. Helium is also not practical, as it is expensive and escapes over time; some alternatives to achieving such lighter-than-air flight could involve heat-trapping, the use of materials which can be controlled to become liquids or gases, and also incorporating aerodyne features such as propellers. Knowledge of the way insects fly could be useful for generating efficient flight for small robots [202]. Dexterous manipulation will be another challenge. Many useful tasks could be accomplished with such capability, such as opening doors and handling tools such as fire extinguishers in the event of some trouble; highly developed robots could provide CPR or life support. However, developing such capability, especially for a lighter-than-air robot, will be highly challenging. Verbal communication will also require work. A robot should factor in knowledge of wind and altitude effects on the propagation of

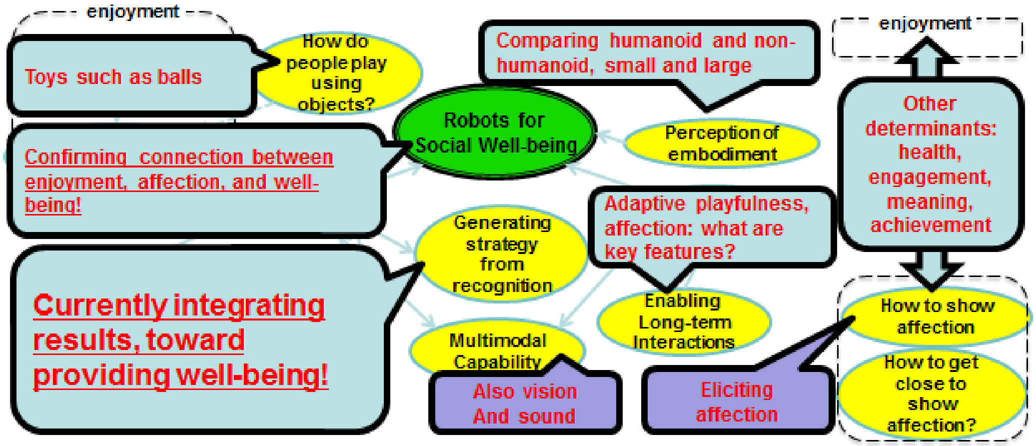


Figure 38. Some next steps for creating robots for social well-being.
We have conducted some new studies shown in purple which have not yet been published and hence are not discussed here, and are currently working to integrate all results; next we want to investigate other determinants of well-being, enjoyable scenarios, perceptions of embodiment, and adaptation.

sound to adapt its speaking; in crowded situations, or when a robot is separated by some distance, a wireless communicator carried by a person might be of value. In terms of the questions we explored, future work will require the use of an actual robot instead of simulations. In regard to z-proxemics, a capability of humanoid robots to fly at various orientations to pass in narrow venues, possibly by employing a mathematical model to predict how to avoid collisions, could be useful. Such a scenario is shown in Figure 39. In regard to z-kinesics, we should investigate the difference in how a single rotation or translation is perceived versus a repeated motion; we predict that the former may be seen as more functional and the latter as emotional displays.

In summary, although we feel an important step has been made in the current work toward acquiring the knowledge to create robots for social well-being, many interesting questions remain which when answered will step-by-step lead to truly rich and meaningful interactions with robots.

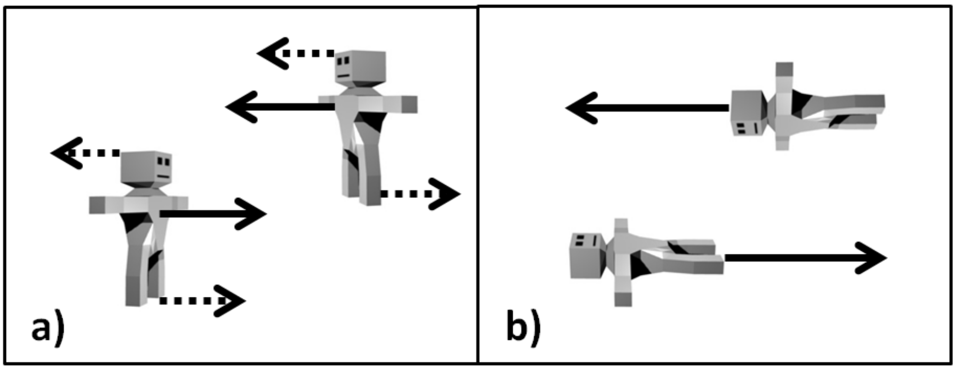


Figure 39. A capability to fly in different orientations could be useful.
A suitable model for collision avoidance which considers "torque" exerted on non-point-source FHRs could allow for efficient passing in tight spaces.

Acknowledgements

I'd like to thank the advisors who supervised my research during the Doctor's Course, Hiroshi Ishiguro (2008-2013), Takayuki Kanda (2008-2011), and Shuichi Nishio (2012-2014), for many years of mentorship. Prof. Ishiguro introduced me to the interesting idea that studying robots can help to better understand the way we humans behave, and treated me kindly and seriously even though I was only a young student. Dr. Kanda spent much time and effort teaching me how to formulate and present research, and gave me the good advice to always seek to tackle large problems. Dr. Nishio suggested not shying away from new learning experiences, gave me some freedom to make decisions about the research, and stressed the importance of logical flow and clearly describing methodology in presenting research. Also, I want to thank the remaining members of my thesis committee: Prof. Tatsuo Arai and Prof. Shogo Nishida, for their time and consideration.

During my time as a PhD student, I had the opportunity to meet many nice and interesting people and work with some of them. I would like to thank collaborators on work represented in the current thesis: Christian Becker-Asano, Aris Alissandrakis, and Francesco Zanlungo. The research environments at Osaka U. and ATR, where I was stationed, were stimulating because they afforded opportunities to converse with many people from various backgrounds researching many different topics.

I am especially grateful to friends, who took the time to eat lunch, buy coffee, play soccer or table tennis, and otherwise accompany me on various adventures. I'm also grateful to my family, my parents and brother, whom I could always count on to listen to many stories and much belly-aching during the preparation of this dissertation.

Finally I would like to thank everyone else who helped with the work described here. My studies in Japan were made possible by a MEXT (Monbusho) Japanese government scholarship, as well as research funding from the Ministry of Internal Affairs and Communications of Japan and JST, CREST.

Bibliography

[1] Cooney MD, Becker-Asano C, Kanda T, Alissandrakis A, & Ishiguro H. (2010) Full-body Gesture Recognition Using Inertial Sensors for Playful Interaction with Small Humanoid Robot. In Proceedings of the IROS 2010, pp. 2276-2282.

[2] Cooney MD, Kanda T, Alissandrakis A, & Ishiguro H (2011) Interaction Design for an Enjoyable Play Interaction with a Small Humanoid Robot. In: 2011 IEEE-RAS International Conference on Humanoid Robots, Humanoids 2011, pp 112-119.

[3] Cooney MD, Nishio S, & Ishiguro H (2012a) Recognizing Affection for a Touch-based Interaction with a Humanoid Robot. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2012, pp. 1420-1427).

[4] Cooney MD, Zanlungo F, Nishio S, & Ishiguro H (2012b) Designing a Flying Humanoid Robot (FHR): Effects of Flight on Interactive Communication, 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2012).

[5] Cooney MD, Kanda T, Alissandrakis A, & Ishiguro H (2013a) Designing Enjoyable Motion-Based Play Interactions with a Small Humanoid Robot, International Journal of Social Robotics. http://dx.doi.org/10.1007/s12369-013-0212-0.

[6] Cooney MD, Nishio S, & Ishiguro H (2013b) Designing Robots for Well-being: Theoretical Background and Visual Scenes of Affectionate Play with a Small Humanoid Robot. Lovotics. (Accepted).

[7] Takayama L, Ju W, & Nass C (2008) Beyond Dirty, Dangerous and Dull: What Everyday People Think Robots Should Do. In: Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction, HRI

2008, pp 25-32.

[8] Fujita M (2004) On Activating Human Communications with Pet-type Robot AIBO. Proc IEEE, 92(11):1804-1813. http://dx.doi.org/10.1109/JPROC.2004.835364.

[9] Carpenter J (2013) The Quiet Professional: An Investigation of U.S. Military Explosive Ordnance Disposal Personnel Interactions With Everyday Field Robots. (Thesis).

[10] Asimov I (1950) I, Robot, New York: Doubleday & Company.

[11] Picard R (1995) Affective Computing MIT Technical Report #321.

[12] Scheutz M, Briggs G, Cantrell R, Krause E, Williams T & Veale R (2013) Novel Mechanisms for Natural Human-Robot Interactions in the DIARC Architecture.

[13] well-being [Def. 1]. (n.d.) In Dictionary.com, Retrieved September 10, 2013, from http://dictionary.reference.com/.

[14] Helliwell JF, Putnam RD (2004) The social context of well-being. Phil. Trans. R. Soc. Lond. B. Biol. Sci. 2004 359, 1435-1446 http://dx.doi.org/10.1098/rstb.2004.1522.

[15] Steel P, Schmidt J, Shultz J (2008) Refining the Relationship Between Personality and Subjective Well-Being. Psychological Bulletin 134(1): 138-161. http://dx.doi.org/10.1037/0033-2909.134.1.138.

[16] Diener EF (1984) Subjective well-being, Psychological Bulletin 95(3) 542–575. doi:10.1037/0033-2909.95.3.542.

[17] Aristotle. (2008) In Stanford Encyclopedia of Philosophy, Retrieved December 13, 2013, from http://plato.stanford.edu/entries/aristotle/.

[18] Thomae de Aquino (2011) In Summa Theologiae: prima pars secundae partis, quaestio VII Retrieved January 16, 2014, from http://www.corpusthomisticum.org/sth2001.html (in Latin).

[19] Aquinas (2009) In Stanford Encyclopedia of Philosophy, Retrieved January 16, 2014, from http://plato.stanford.edu/entries/aquinas/.

[20] John Stuart Mill (2007) In Stanford Encyclopedia of Philosophy, Retrieved

November 25, 2013 from http://plato.stanford.edu/entries/mill/.

[21] Haggbloom SJ et al (2002) The 100 Most Eminent Psychologists of the 20th Century. Review of General Psychology 6 (2): 139 – 152. doi:10.1037/1089-2680.6.2.139.

[22] Seligman MEP, Csíkszentmihályi M (2000) Positive Psychology: An Introduction. American Psychologist 55(1) 5 – 14. doi:10.1037/0003-066X.55.1.5.

[23] McVeigh T (2011) "David Cameron measuring 'wrong type of happiness'" In theguardian.com, Retrieved January 16, 2014, from http://www.theguardian.com/politics/2011/apr/10/david-cameron-wrong-type-happiness/print .

[24] "Make people happier, says Cameron" (2006) In BBC NEWS, Retrieved January 16, 2014, from http://news.bbc.co.uk/2/hi/uk_news/politics/5003314.stm.

[25] Davidson RJ & Begley S (2012) The emotional life of your brain: How its unique patterns affect the way you think, feel, and live -- and how you can change them. New York, NY: Hudson Street Press.

[26] Frey BS (2011) Happy People Live Longer. Science, 331 (6017): 542–543.

[27] Eriksson L, Rice JM, & Goodin RE (2007) Temporal Aspects of Life Satisfaction, Social Indicators Research, 80(3) 511-533.

[28] Vorderer P, Klimmt C, & Ritterfeld U (2004) Enjoyment: At the heart of media entertainment. Commun Theory 14, 388-408. http://dx.doi.org/10.1111/j.1468-2885.2004.tb00321.

[29] enjoyment. [Def. 1, 2, 3]. (n.d.). In Collins English Dictionary - Complete & Unabridged 10th Edition. Retrieved September 10, 2013, from Dictionary.com website: http://dictionary.reference.com/browse/enjoyment.

[30] Seligman MEP (2004) Can Happiness be Taught? Daedalus journal.

[31] Csíkszentmihályi M (1975) Beyond Boredom and Anxiety: experiencing flow in work and play. Jossey-Bass, San Francisco, CA.

[32] Csíkszentmihályi M, Rathunde K (1993) The measurement of flow in everyday life: Towards a theory of emergent motivation. In: Jacobs JE (ed) Developmental perspectives on motivation. Nebraska symposium on motivation, University of Nebraska Press, Lincoln.

[33] affection [Def. 1]. (n.d.) In Dictionary.com, Retrieved September X, 2013, from http://dictionary.reference.com/.

[34] Maslow AH (1943) A theory of human motivation. Psychological Review, 50(4) 370–96. http://dx.doi.org/10.1037/h0054346.

[35] Harlow H (1958) The Nature of Love. American Psychologist, 13(12), 573–685. http://dx.doi.org/10.1037/h0047884.

[36] Prescott JW (1975) Body pleasure and the origins of violence. In The Bulletin of The Atomic Scientists, 10-20.

[37] Turkle S, Breazeal C, Scassellati B, & Dasté O (2004) Encounters with Kismet and Cog: Children's Relationships with Humanoid Robots. In: 2004 IEEE-RAS International Conference on Humanoid Robots, Humanoids 2004.

[38] Seligman MEP (2011) Flourish. N.Y.: Simon & Schuster.

[39] AARP Research & Strategic Analysis, Heart and Mind Strategies, (2012) Beyond Happiness: Thriving. Summary of Findings 06.04.2012, from Surveys and Statistics.

[40] Kahneman D, & Krueger AB (2006) Developments in the Measurement of Subjective Well-Being. Journal of Economic Perspectives 20(1):3–24.

[41] Ancient Ethical Theory. (2009) In Stanford Encyclopedia of Philosophy, Retrieved December 13, 2013, from http://plato.stanford.edu/entries/ethics-ancient/.

[42] Jowett B (trans) (1892) The Dialogues of Plato, Vol. 1, Plato's Protagoras retrieved Dec. 18, 2013 from http://ancienthistory.about.com/library/bl/bl_text_plato_protag_1.htm.

[43] Epicurus. (2009) In Stanford Encyclopedia of Philosophy, Retrieved

December 13, 2013, from http://plato.stanford.edu/entries/epicurus/#2.

[44] Hutcheson (1726) In An Inquiry into the Original of Our Ideas of Beauty and Virtue, Treatise II, Section III, Paragraph VIII, ed. Wolfgang Leidhold (Indianapolis: Liberty Fund, 2004). Retrieved January 16, 2014, from http://oll.libertyfund.org/simple.php?.

[45] Bentham J (1789). The Principles of Morals and Legislation. Oxford: Clarendon Press.

[46] Mill, John Stuart (1998). Crisp, Roger, ed. Utilitarianism. Oxford University Press. p. 56. ISBN 0-19-875163-X.

[47] Erikson EH (1950) Childhood and Society. New York:Norton.

[48] Rogers C (1961). On Becoming a Person: A Therapist's View of Psychotherapy. London: Constable.

[49] Allport GW (1961) Pattern and Growth in Personality. New York: Holt, Rinehart, & Winston.

[50] Ryff CD (1989) Happiness is everything, or is it? Explorations on the meaning of psychological well-being. Journal of Personality and Social Psychology, 57(6):1069-1081.

[51] Young JR (2011) Programmed for Love. In The Chronicle Review, Retrieved January 20, 2014, from http://chronicle.com/article/Programmed-for-Love-The/125922/.

[52] Walum H, Westberg L, Henningsson S, et al. (2008) Genetic variation in the vasopressin receptor 1a gene (AVPR1A) associates with pair-bonding behavior in humans. Proc. Natl. Acad. Sci. U.S.A. 105(37) 14153–6. doi:10.1073/pnas.0803081105.

[53] De Dreu CK, Greer LL, Van Kleef GA, Shalvi S, & Handgraaf MJ (2011) Oxytocin promotes human ethnocentrism. Proc. Natl. Acad. Sci. U.S.A. 108(4) 1262–6. doi:10.1073/pnas.1015316108.

[54] Gouin JP, Carter S, Pournajafi-Nazarloo H, Glaser R, Malarkey WB, et al. (2010) Marital Behavior, Oxytocin, Vasopressin, and Wound Healing. Psychoneuroendocrinology 35 (7): 1082 – 1090.

http://dx.doi.org/10.1016/j.psyneuen.2010.01.009.

[55] Uvnäs-Moberg K (1998) Oxytocin May Mediate The Benefits of Positive Social Interaction and Emotions, Psychoneuroendocrinology, 23(8) 819-835.

[56] Hartwig AC (1991) Peripheral beta-endorphin and pain modulation. Anesthesia progress 38(3): 75-8. PMID 1814247.

[57] Goldstein A (1980) Thrills in Response to Music and Other Stimuli, Physiological Psychology 8, 126-129.

[58] Dunbar R (2003) Evolution of the social brain, Science 302(5648) 1160-1161.

[59] Young SN (2007) How to increase serotonin in the human brain without drugs, Rev. Psychiatr. Neurosci 32(6) 394–99.

[60] Sawin ER, Ranganathan R, & Horvitz HR (2000) C. elegans locomotory rate is modulated by the environment through a dopaminergic pathway and by experience through a serotonergic pathway, Neuron 26(3) 619–31. http://dx.doi.org/10.1016/S0896-6273(00)81199-X.

[61] Meaney MJ & Szyf M (2005) Environmental programming of stress responses through DNA methylation: life at the interface between a dynamic environment and a fixed genome. Dialogues Clin Neurosci 7(2) 103-23.

[62] World Health Organization (2004) Neuroscience of psychoactive substance use and dependence. Retrieved September 11, 2013 from http://www.who.int/substance_abuse/publications/en/Neuroscience_E.pdf.

[63] Arias-Carrión O, Pöppel E (2007). Dopamine, learning and reward-seeking behavior. Act Neurobiol Exp 67 (4): 481–488.

[64] Spanagel R, Weiss F (1999) The dopamine hypothesis of reward: past and current status. Trends Neurosci. 22(11): 521-7. http://dx.doi.org/10.1016/S0166-2236(99)01447-2. PMID 10529820.

[65] Nakatsu R, Rauterberg M, Vorderer P (2005) A New Framework for Entertainment Computing: From Passive to Active Experience. In: Kishino

F, Kitamura Y, Kato H and Nagata N (eds) Entertainment Computing-ICEC, Lecture Notes in Computer Science (LNCS), 3711.

[66] Cramer H, Mentis H, Fernaeus Y (2010) Serious work on playful experiences: a preliminary set of challenges. In: ‘Fun, seriously?’ Workshop at CSCW 2010.

[67] Van der Heijden H (2004) User acceptance of hedonic information systems, MIS Quarterly 28(4) 695–704.

[68] Heerink M, Kröse BJA, Wielinga BJ, & Evers V (2008) Enjoyment, Intention to Use And Actual Use of a Conversational Robot by Elderly People. In: Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction, HRI 2008, pp 113-119.

[69] Bowlby J (1958) The nature of the child’s tie to his mother. International Journal of Psychoanalysis, 39(5) 350–73.

[70] Floyd K (2002) Human affection exchange: V. Attributes of the highly affectionate. Communication Quarterly, 50, 135-154.

[71] Cacioppo J & Patrick W (2008) Loneliness: Human Nature and the Need for Social Connection. NY: W. W. Norton & Co.

[72] Wilson RS, et al (2007) Loneliness and Risk of Alzheimer Disease. Arch Gen Psychiatry 64, 234-240.

[73] Ogawa K, Nishio S, Koda K, Balistreri G, Watanabe T, & Ishiguro H (2011) Exploring the Natural Reaction of Young and Aged Person with Telenoid in a Real World. Journal of Advanced Computational Intelligence and Intelligent Informatics, 15(5) 592-597.

[74] Grandin T (1992) Calming Effects of Deep Touch Pressure in Patients with Autistic Disorder, College Students, and Animals. Journal of Child and Adolescent Psychopharmacology, 2(1) 63–72. http://dx.doi.org/10.1089/cap.1992.2.63.

[75] Kanamori M, et al (2003) Pilot study on improvement of quality of life among elderly using a pet–type robot. In Proceedings of the IEEE

international symposium on computational intelligence in robotics and automation (CIRA, pp. 107‑112).

[76] Wada K, & Shibata T (2007). Living With Seal Robots - Its Sociopsychological and Physiological Influences on the Elderly at a Care House. IEEE Trans. Robot., 23(5) 972-980. http://dx.doi.org/10.1109/TRO.2007.906261.

[77] Russell, J (1980) A circumplex model of affect. Journal of Personality and Social Psychology 39, 1161‑1178. http://dx.doi.org/10.1037/h0077714.

[78] Barrett, LF (2006) Valence as a basic building block of emotional life. Journal of Research in Personality 40, 35‑55.

[79] Mehrabian A, & Ferris SR (1967) Inference of Attitudes from Nonverbal Communication in Two Channels. Journal of Consulting Psychology 31(3). http://dx.doi.org/10.1037/h0024648.

[80] Schramm W (1954) How Communication Works, in The Process and Effects of Communication, ed. Wilbur Schramm (Urbana: University of Illinois Press, 1954).

[81] Shannon, CE & Weaver W (1949). The Mathematical Theory of Communication. Urbana, IL: University of Illinois Press.

[82] Yohanan S, & MacLean KE (2012) The Role of Affective Touch in Human-Robot Interaction: Human Intent and Expectations in Touching the Haptic Creature. Int J Soc Robot 4(2) 163-180. http://dx.doi.org/10.1007/s12369-011-0126-7.

[83] Pantic M, Cowie R, D’Errico F, et al. (2011) Social Signal Processing: The Research Agenda, Laser Physics Review, 511-538.

[84] signal [Def. 1]. (n.d.) In Dictionary.com, Retrieved August 21, 2013, from http://dictionary.reference.com/browse/signal?s=b.

[85] behavior [Def. 2]. (n.d.) In Dictionary.com, Retrieved August 21, 2013, from http://dictionary.reference.com/browse/behavior?s=t.

[86] action [Def. 2]. (n.d.) In Dictionary.com, Retrieved August 21, 2013, from

http://dictionary.reference.com/browse/action?s=t.

[87] channel [Def. 4]. (n.d.) In Collins English Dictionary, Retrieved August 21, 2013, from http://www.collinsdictionary.com/dictionary/english/channel?showCookiePolicy=true.

[88] modality [Def. 2]. (n.d.) In Dictionary.com, Retrieved August 21, 2013, from http://dictionary.reference.com/browse/modality?s=t.

[89] polysemy (n.d.) In Dictionary.com, Retrieved August 21, 2013, from http://dictionary.reference.com/browse/modality?s=t.

[90] Matsumoto N (2007) A Reference of Intentions for Better Medical Communication. In Proceedings of the IEEE/ICME International Conference on Complex Medical Engineering (CME, pp. 453 – 456). http://dx.doi.org/10.1109/ICCME.2007.4381775.

[91] play [Def. 4]. (n.d.) In Dictionary.com, Retrieved August 21, 2013, from http://dictionary.reference.com/browse/play?s=t.

[92] Huizinga J (1950) Homo Ludens: a study of the play element in culture. Roy Publishers, the Beacon Press, Boston.

[93] Ellis MJ (1973) Why people play. Prentice-Hall, Englewood Cliffs, New Jersey.

[94] Goetz J, Kiesler S, & Powers A (2003) Matching robot appearance and behavior to tasks to improve human-robot cooperation. In: 2003 IEEE International Workshop on Robots and Human Interactive Communications, RO-MAN 2003, pp 55-60.

[95] Geldard FA (1960) Some neglected possibilities of communication. Science, 131:1583-1588. http://dx.doi.org/10.1126/science.131.3413.1583.

[96] Barnett K (1972) A theoretical construct of the concepts of touch as they relate to nursing. Nurs Res, 21:102-110. http://dx.doi.org/10.1097/00006199-197203000-00002.

[97] Williams LE, Bargh JA (2008) Experiencing physical warmth promotes interpersonal warmth. Science, 322, 606-607.

http://dx.doi.org/10.1126/science.1162548.

[98] Kraus MW, Huang C, Keltner D (2010) Tactile communication, cooperation, and performance: An ethological study of the NBA. Emotion 10, 745-749. http://dx.doi.org/10.1037/a0019382.

[99] Lee JK, Toscano RL, Stiehl WD, and Breazeal C (2008) The Design of a Semi-Autonomous Robot Avatar for Family Communication and Education. Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN).

[100] Miyashita T, Tajika T, Ishiguro H, Kogure K, and Hagita N (2005) Haptic communication between humans and robots. In Proceedings of the 12th International Symposium of Robotics Research.

[101] Yoshikai T, Hayashi M, Kadowaki A, Goto T, and Inaba M (2009) Design and Development of a Humanoid with Soft 3D-Deformable Sensor Flesh and Automatic Recoverable Mechanical Overload Protection Mechanism. in IROS 2009.

[102] Yohanan S et al. (2009) Affect-Driven Emotional Expression with the Haptic Creature, UIST '09.

[103] Kozima H, Michalowski MP, & Nakagawa C (2009) Keepon: A Playful Robot for Research, Therapy, and Entertainment. Int J Soc Robot 1(1) 3-18. http://dx.doi.org/10.1007/s12369-008-0009-8.

[104] Michalowski MP, Sabanovic S, Michel P (2006) Roillo: Creating a social robot for playrooms. In: 2006 IEEE International Workshop on Robots and Human Interactive Communications, RO-MAN 2006, pp 587 – 592. http://dx.doi.org/10.1109/ROMAN.2006.314453.

[105] Salter T, Michaud F, Dautenhahn K, Letourneau D, Caron S (2005) Recognizing interaction from a robot ’ s perspective. In: 2005 IEEE International Workshop on Robots and Human Interactive Communications, RO-MAN 2005, pp 178-183.

[106] Salter T, Michaud F, Letourneau D, Lee DC, Werry IP (2007) Using Proprioceptive Sensors for Categorizing Human-Robot Interactions. In:

Proceedings of the 2nd ACM/IEEE International Conference on Human Robot Interaction, HRI 2007, pp 105-112.

[107] Hertenstein MJ, Keltner D, App B, Bulleit B, & Jaskolka A (2006) Touch communicates distinct emotions. Emotion 6, 528-533. http://dx.doi.org/10.1037/1528-3542.6.3.528.

[108] Brooks AG, & Arkin RC (2007) Behavioral overlays for non-verbal communication expression on a humanoid robot. Auton Robot 22(1) 55-74. http://dx.doi.org/10.1007/s10514-006-9005-8.

[109] Thomas L (2009) What's Behind Japan's Love Affair with Robots?. Time. Retrieved August 21, 2013 from http://content.time.com/time/world/article/0,8599,1913913,00.html#ixzz0j5ZxitJX.

[110] Brickman P, Coates D, & Janoff-Bulman R (1978) Lottery winners and accident victims: is happiness relative? J. Personality Social Psychol. 36, 917–927.

[111] Lee MK, Forlizzi J, Kiesler S, Rybski P, Antanitis J, Savetsila S (2012) Personalization in HRI: A longitudinal field experiment. HRI 2012.

[112] Turk MA and Pentland AP (1991) Face recognition using eigenfaces.Computer Vision and Pattern Recognition. Proceedings CVPR'91.

[113] Mazloom M, Kasaei S, & Alemi N (2009) Construction and Application of SVM model and Wavelet-PCA for Face Recognition, ICCEE.

[114] Kulic D, Ott C, Lee D, Ishikawa J, Nakamura Y (2012) Incremental Learning of Full Body Motion Primitives and their Sequencing through Human Motion Observation. Int J Rob Res, 31(3) 330-345. http://dx.doi.org/10.1177/0278364911426178.

[115] Lewis C, & Rieman J (1993) Task-Centered User Interface Design: A Practical Introduction. Internet: HCI Bibliography: Human-Computer Interaction Resources http://www.hcibib.org/tcuid/index.html Accessed July 25 2013.

[116] Nielsen J, Molich R (1990) Heuristic evaluation of user interfaces. In: Proc. ACM CHI’ 90 Conf, pp 249-256. http://dx.doi.org/10.1145/97243.97281.

[117] Dautenhahn K, Nehaniv CL, Walters ML, Robins B, Kose-Bagci H, Mirza NA, & Blow M (2009) KASPAR - A Minimally Expressive Humanoid Robot for Human-Robot Interaction Research. Appl Bionics Biomech, Special Issue on Humanoid Robots 6(3), 369-397.

[118] Vinciarelli A, Salamin H, & Pantic M (2009) Social Signal Processing: Understanding Social Actions through nonverbal behaviour analysis, in IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) (Miami, FL, USA).

[119] Egger FN (2000) Lo-Fi vs. Hi-Fi Prototyping: how real does the real thing have to be? Teaching HCI workshop, OzCHI.

[120] McCowan I, Gatica-Perez D, Bengio S, Lathoud G, Barnard M. & Zhang D (2005) Automatic Analysis of Multimodal Group Actions in Meetings. IEEE Trans. on Pattern Analysis and Machine Intelligence, 27(3) 305-317. http://dx.doi.org/10.1109/TPAMI.2005.49.

[121] Glaser BG & Strauss AL (1967) The discovery of grounded theory: strategies for qualitative research. Chicago: Aldine.

[122] Kahn PH, Freier NG, Kanda T, Ishiguro H, Ruckert JH, Severson RL, Kane SK (2008) Design Patterns for Sociality in Human-Robot Interaction. In: Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction, HRI 2008, pp 97-104.

[123] Satake S, Kanda T, Glas DF, Imai M, Ishiguro H, & Hagita N (2009) How to approach humans?: strategies for social robots to initiate interaction. In: Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction, HRI 2009, pp 109-116.

[124] Cushman WH, & Rosenberg DJ (1991) Human factors in product design. New York: Elsevier.

[125] Ekman P, & Friesen, WV (1969) The repertoire of nonverbal behavior: Categories, origins, usage, and coding. Semiotica 1, 49-98.

[126] Fraser A (1973). Dolls. Octopus books. ISBN 0-7064-0056-9.

[127] Winnicott DW (1953) Transitional Objects and Transitional Phenomena−A Study of the First Not-Me Possession. Int. J. Psycho-Anal., 34:89-97.

[128] Ovid (2004). Metamorphoses: Book 13. Poetry X, Edited by Jough Dempsey. Retrieved 22 Aug. 2013 from http://poetry.poetryx.com/poems/4817/.

[129] Rosheim ME (1994) Robot Evolution: The Development of Anthrobotics, Wiley-IEEE, p. 36, ISBN 0-471-02622-0.

[130] McGee K (2004) A touch of the future: contact-expressive devices, Multimedia, IEEE, 11(1) 8-11.

[131] Fong T, Nourbakhsh I (2005) Interaction Challenges in Human Robot Space Exploration. ACM Interactions 12(2): 42-45. http://dx.doi.org/10.1145/1052438.1052462.

[132] Reddy R (2006) Robotics and Intelligent Systems in Support of Society. IEEE Intelligent Systems 21(3): 24-31. http://dx.doi.org/10.1109/MIS.2006.57.

[133] Graf B, Reiser U, Hägele M, Mauz K, & Klein P (2009) Robotic Home Assistant Care-O-bot® 3 - Product Vision and Innovation Platform. In: IEEE Workshop on Advanced Robotics and its Social Impacts - ARSO 2009. Piscataway, NJ: IEEE, 139-144.

[134] Kawamoto H, Taal S, Niniss H, Hayashi T, Kamibayashi K, Eguchi K, & Sankai Y (2010) Voluntary motion support control of Robot Suit HAL triggered by bioelectrical signal for hemiplegia, Conf Proc IEEE Eng Med Biol Soc, 462-6.

[135] Mukai T, Hirano S, Nakashima H, Kato Y, Sakaida Y, Guo S, & Hosoe S (2010) Development of a nursing-care assistant robot RIBA that can lift a human in its arms, IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 5996-6001.

[136] Takeda T, Kosuge K, & Hirata Y (2005) HMM-based Dance Step Estimation for Dance Partner Robot -MS DanceR-, Proceedings of the 2005

IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 1602-1607.

[137] TOPIO: TOSY Ping Pong Playing Robot. Retrieved September 10, 2013, from http://topio.tosy.com/about.shtml.

[138] Müller M, Lupashin S, & D’Andrea R (2011) Quadrocopter Ball Juggling. IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS.

[139] Fasola J, & Matarić MJ (2012) Using Socially Assistive Human-Robot Interaction to Motivate Physical Exercise for Older Adults. Proceedings of the IEEE 100(8) 2512-2526.

[140] Hansen ST, Bak T, & Risager C (2012) An Adaptive Game Algorithm for an Autonomous, Mobile Robot - a Real World Study with Elderly Users, RO-MAN, 892-897.

[141] DiSalvo C, Gemperle F, Forlizzi J, & Montgomery E (2003) The Hug: An Exploration of Robotic Form for Intimate Communication. Ro-Man 2003.

[142] Samani H, Teh J, Saadatian E, & Nakatsu R (2013) XOXO: Haptic Interface for Mediated Intimacy In IEEE International Symposium on Next-Generation Electronics (ISNE), 256-259.

[143] Tanaka F, Movellan JR, Fortenberry B, & Aisaka K (2006) Daily Hri Evaluation at a Classroom Environment: Reports from Dance Interaction Experiments. In: Proceedings of the 1st ACM/IEEE International Conference on Human Robot Interaction, HRI 2006, pp 3-9.

[144] Tanaka F, Cicourel A, & Movellan JR (2007) Socialization between Toddlers and Robots at an Early Childhood Education Center. In: Proceedings of the National Academy of Sciences of the USA (PNAS), pp 17954-17958. http://dx.doi.org/10.1073/pnas.0707769104.

[145] Lorenz K (1971) Studies in Animal and Human Behavior. Cambridge, MA: Harvard Univ. Press.

[146] Breazeal C, & Aryananda L (2002) Recognition of Affective Communicative Intent in Robot-Directed Speech, Autonomous Robots

12(1) 83-104.

[147] Brooks, AG, & Breazeal C (2006) Working with Robots and Objects: Revisiting Deictic Reference for Achieving Spatial Common Ground. In Proceedings of the ACM SIGCHI/SIGART Conference on Human-Robot Interaction (HRI 2006, pp. 297-304).

[148] Rae I, Takayama L Mutlu B The influence of height in robot-mediated communication, HRI 2013.

[149] Noda T, Miyashita T, Ishiguro H, & Hagita N (2007) Map acquisition and classification of haptic interaction using cross correlation between distributed tactile sensors on the whole body surface. In Proceedings of IROS 2007.

[150] Knight H, Chang A, Stiehl WD, Toscano R, Wang Y, & Breazeal C (2009) Robot Design Rubrics for Social Gesture Categorization and User Studies with Children. In Proceedings of the ACM/IEEE International Conference on Human Robot Interaction (HRI 2009).

[151] Stiehl, WD, & Breazeal, C (2004) Affective Touch for Robotic Companions. In Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII LNCS 3784, pp. 747-754). http://dx.doi.org/10.1007/11573548_96.

[152] François D, Polani D, & Dautenhahn K (2007) On-line behaviour classification and adaptation to human-robot interaction styles. In Proceedings of the 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI07, pp. 295-302).

[153] Hornfeck K, Zhang Y, Lee K (2012) Philos: a sociable robot for human robot interactions and wireless health monitoring. SAC 2012: 293-294.

[154] Hall GS, & Allin A (1897) The psychology of tickling, laughing and the comic. The American Journal of Psychology 9, 1 – 42. http://dx.doi.org/10.2307/1411471.

[155] Jones SE, & Yarbrough AE (1985) A naturalistic study of the meanings of touch. Communication Monographs 52, 19-56.

http://dx.doi.org/10.1080/03637758509376094.

[156] ‘Is this a real cat?’ A robot cat you can bond with like a real pet -- NeCoRo is born. Retrieved September 10, 2013, from http://www.machinebrain.com/articles/omroncat101601.html.

[157] Ugobe, Specification sheet for Pleo. Retrieved September 10, 2013, from http://www.pleo-web.de/images/stories/pleo/specificationsheetpleo.pdf.

[158] Naya F, Yamato J and Shinozawa K (1999), Recognizing Human Touching Behaviors using a Haptic Interface for a Pet-robot. Proc. SMC'99, pp. II-1030-1034.

[159] Lee MH, and Nicholls HR (1999) Tactile sensing for mechatronics-a state of the art survey, Mechatronics, Elsevier 9: 1-31.

[160] Lee, MK, Forlizzi J, Rybski PE, Crabbe F, Chung W, Finkle J, Glaser E, and Kiesler S (2009) The Snackbot: Documenting the design of a robot for long-term human-robot interaction. In Proceedings of HRI 2009, 7-14.

[161] F Michaud and S Caron (2002) Roball, the Rolling Robot. Autonomous Robots, 12(2): 211-222.

[162] Robins B, Dautenhahn K, Boekhorst RT, Nehaniv CL (2008) Behaviour Delay and Robot Expressiveness in Child-Robot Interactions: A User Study on Interaction Kinesics. In: Proceedings of HRI 2008.

[163] Murphy R et al (2011) A Midsummer Night’s Dream (with flying robots), Autonomous Robots, 30(2) 143-156.

[164] Ng WS, & Sharlin E (2011) Collocated interaction with flying robots, in IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), (IEEE Press, Atlanta, Georgia), pp. 143-149.

[165] Hall ET (1966) The Hidden Dimension. (Anchor Books, New York).

[166] Kendon A (1990) Conducting Interaction: Patterns of Behavior in Focused Encounters, Cambridge: Cambridge University Press.

[167] Sisbot EA, Marin-Urias LF, Alami R, & Simeon T (2007) A Human Aware Mobile Robot Motion Planner, IEEE Transactions on Robotics, 23(5): 874-883.

[168] Koay KL, Syrdal DS, Walters ML, & Dautenhahn K (2007) Living with Robots: Investigating the Habituation Effect in Participants' Preferences During a Longitudinal Human-Robot Interaction Study, RO-MAN.

[169] Helbing D, & Johansson A (2010) Pedestrian, Crowd and Evacuation Dynamics, Encyclopedia of Complexity and Systems Science 16, 6476-6495.

[170] Zanlungo F, Ikeda T, & Kanda T (2011) Social force model with explicit collision prediction, Europhysics Letters Journal 93-68005.

[171] Lopes M, & Santos-Victor J (2007) A Developmental Roadmap for Learning by Imitation in Robots, SMC - B, 37(2): 308–321.

[172] Nakagawa K, Shinozawa K, Ishiguro H, Akimoto T, & Hagita N (2009) Motion modification method to control affective nuances for robots, IROS, pp. 5003–5008.

[173] Minato T, Nishio S, Ogawa K, & Ishiguro H (2011) Development and Study on a Cellular Phone-Type Teleoperated Android, 'Elfoid'. Robotics Society of Japan, 302-4 (in Japanese).

[174] Vapnik V, Golowich S, and Smola A (1996) Support vector method for function approximation, regression estimation, and signal processing, Adv Neural Inform Process Syst 9: 281–287.

[175] Chang C, & Lin C (2011) LIBSVM: a library for support vector machines, ACM Transactions on Intelligent Systems and Technology (ACM TIST), 2:27:1-27:27. http://dx.doi.org/10.1145/1961189.1961199.

[176] Viola P, Jones M (2001) Robust real time object detection. In: IEEE ICCV Workshop Statistical and Computational Theories of Vision, Vancouver, Canada.

[177] Aamodt A, & Plaza E (1994) Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches, Artificial Intelligence Communications 7(1) 39-52.

[178] Nomura T, Suzuki T, Kanda T, & Kato K (2006) Measurement of Negative

Attitudes toward Robots, Interact Stud, 7(3) 437-454. http://dx.doi.org/10.1075/is.7.3.14nom.

[179] Bates J (1994) The Role of Emotion in Believable Agents. Communications of the ACM 37: 122-125. http://dx.doi.org/10.1145/176789.176803.

[180] Cannata G, Maggiali M, Metta G, and Sandini G (2008) An Embedded Artificial Skin for Humanoid Robots, MFI 2008, pp. 434-438.

[181] Ohmura Y, Kuniyoshi Y, & Nagakubo A (2006) Conformable and Scalable Tactile Sensor Skin for Curved Surfaces. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA, pp. 1348-1353). http://dx.doi.org/10.1109/ROBOT.2006.1641896.

[182] SHARP GP2S60 Retrieved January 8, 2014 from http://sharp-world.com/products/device/lineup/data/pdf/datasheet/gp2s60_e.pdf.

[183] Spears JD & Zollman D (1985) Fascination of Physics. Benjamin-Cummings Pub Co.

[184] Li C, Zheng SQ, and Prabhakaran B (2007) Segmentation and recognition of motion streams by similarity search, TOMCCAP Vol. 3(3).

[185] Cover TM, & Hart PE (1967) Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1) 21 - 27. http://dx.doi.org/10.1109/TIT.1967.1053964.

[186] Zhang H, Berg AC, Maire M & Malik J (2006) SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition, in Proc. CVPR.

[187] Sigalas M, Baltzakis H, and Trahanias PE (2010) Gesture recognition based on arm tracking for human-robot interaction, IROS 2010: pp. 5424-5429.

[188] Bowlby J (1969) Attachment. Attachment and Loss, New York: Basic Books.

[189] Schaffer HR & Emerson PE (1964) The Development of Social Attachments in Infancy. Monographs of the Society for Research in Child Development, 29, 94.

[190] Herrick FH (1922) Homing Powers of the Cat, The Scientific Monthly

14(6): 525-539.

[191] Weizenbaum J (1966) ELIZA − A Computer Program For the Study of Natural Language Communication Between Man And Machine, Communications of the ACM 9(1) 36–45.

[192] Straker L, Pollock C, Burgess-Limerick R, Skoss R, & Coleman J (2008) The impact of computer display height and desk design on muscle activity during information technology work by young adults, Journal of Electromyography and Kinesiology 18(4), 606-617.

[193] Morales Y, Satake S, Huq R, Glas D, Kanda T, & Hagita N (2012) How Do People Walk Side-By-Side?- Using A Computational Model of Human Behavior For A Social Robot, HRI 2012.

[194] Birdwhistell R (1970) Kinesics and Context, University of Pennsylvania Press, Philadelphia.

[195] Van Someren MW, Barnard YF, & Sandberg JAC (1994) The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes, San Diego: Academic Press.

[196] Garfinkel H (1967) Studies in Ethnomethodology. Prentice-Hall, New Jersey.

[197] Heider F, & Simmel M (1944) An experimental study of apparent behavior. Am J Psychol, 57:243–259. http://dx.doi.org/10.2307/1416950.

[198] Festinger L (1957) A theory of cognitive dissonance. Stanford University Press, Stanford, CA.

[199] Watson D, Clark LA, Tellegan A (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. J Pers Soc Psychol 54(6) 1063 – 1070. http://dx.doi.org/10.1037/0022-3514.54.6.1063.

[200] Hassenzahl M, Burmester M, & Koller F (2003) AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Mensch & Computer, Interaktion in Bewegung, pp 187-196 (in German).

[201] Brooke J (1996) SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL (eds), Usability Evaluation in Industry. Taylor and Francis, London.

[202] Lehmann FO, Sane SP, Dickinson M (2005) The aerodynamic effects of wing-wing interaction in flapping insect wings. The Journal of Experimental Biology 208: 3075‑3092.

Publications

Journal Articles

Martin Cooney, Takayuki Kanda, Aris Alissandrakis, & Hiroshi Ishiguro (2013a) Designing Enjoyable Motion-Based Play Interactions with a Small Humanoid Robot, International Journal of Social Robotics. (23 pages.) http://dx.doi.org/10.1007/s12369-013-0212-0

Martin Cooney, Shuichi Nishio, & Hiroshi Ishiguro (2013b) Designing Robots for Well-being: Theoretical Background and Visual Scenes of Affectionate Play with a Small Humanoid Robot. Lovotics. (Accepted, 18 pages.)

Martin Cooney, Shuichi Nishio, & Hiroshi Ishiguro (n.d.) Affectionate Interaction with a Small Humanoid Robot Capable of Recognizing Social Touch Behavior. (Submitted Manuscript (Current Status: Accepted with Major Revision), 33 pages.)

Martin Cooney, Shuichi Nishio, & Hiroshi Ishiguro (n.d.) Multi-modal Affectionate Playing with a Small Humanoid Robot. (Submitted Manuscript, 42 pages.)

Conference Papers

Martin Cooney, Francesco Zanlungo, Shuichi Nishio, & Hiroshi Ishiguro (2012) Designing a Flying Humanoid Robot (FHR): Effects of Flight on Interactive Communication. In Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2012, pp. 364 - 371). http://dx.doi.org/10.1109/ROMAN.2012.6343780

Martin Cooney, Shuichi Nishio, & Hiroshi Ishiguro (2012) Recognizing Affection for a Touch-based Interaction with a Humanoid Robot, In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2012, pp. 1420-1427). http://dx.doi.org/10.1109/IROS.2012.6385956

Martin Cooney, Takayuki Kanda, Aris Alissandrakis, & Hiroshi Ishiguro (2011) Interaction Design for an Enjoyable Play Interaction with a Small Humanoid Robot. In Proceedings of the 2011 11th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2011, pp 112-119). http://dx.doi.org/10.1109/Humanoids.2011.6100847

Martin Cooney, Christian Becker-Asano, Takayuki Kanda, Aris Alissandrakis, & Hiroshi Ishiguro (2010) Full-body Gesture Recognition Using Inertial Sensors for Playful Interaction with Small Humanoid Robot. In Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010, pp. 2276-2282). http://dx.doi.org/10.1109/IROS.2010.5650081 (NTF Award Finalist for Entertainment Robots and Systems)