Title	Influence of Gait-Induced Upper Body Motion by Moving Wheeled Android on Human Perception and Behavior
Author(s)	八木, 聡明
Citation	大阪大学, 2022, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/88102
rights	
Note	

Influence of Gait-Induced Upper Body Motion by Moving Wheeled Android on Human Perception and Behavior

SATOSHI YAGI

MARCH 2022

Influence of Gait-Induced Upper Body Motion by Moving Wheeled Android on Human Perception and Behavior

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

SATOSHI YAGI

MARCH 2022

Abstract

Due to COVID-19, new lifestyles are being accepted to our society which encourage non-contact and non-face-to-face, such as a remote work and an online meeting. This social situation change has led to grow a demand for mobile robot application to replace labors where worked not only in factories but also in human living environments.

As robots become more widely used in our living environment, we expect that the future society in which humans and robots' symbioses, where teleoperated mobile robots controlled by a remote human operator or autonomous mobile robots move on the same sidewalks in a city. However, it is still a big challenge for robots to move freely in a human living environment.

The existence of humans makes harder robots to move around. It is more difficult for robots to recognize irregularly moving pedestrians and predict their future trajectories than to simply avoid static obstacles. Even if a robot can successfully recognize human actions and complete path planning, this is just one direction of the solutions for the robot navigation. As the other direction, it is important that human sides perceive the robot's behavior as well and move cooperatively without getting in the way of the robot.

The interaction between mobile robots and humans should be to keep a minimum since a moving robot is essentially an obstacle for humans. Most of the humans in the same environment with the robot have different purposes, and there is almost no benefit to interacting with the robot. Even if the robot tries to send some cues to the humans nearby, humans might not make any effort to catch the signals.

Gait, which is a pattern of limb movements made during locomotion, is a main factor to enable human pedestrian collaborative mobility in a society. Humans convey various information to the surrounding others from gaits while walking, and observers can perceive and understand meanings from their gait. This interaction is performed in unconscious. With this, pedestrians spontaneously form a walking flow and achieve high-density and high-efficiency mobility while avoiding collisions with each other.

This thesis aims to propose that a wheeled android expresses an upper body motion induced by a human gait during movement and to validate whether humans who see the body expressions can perceive the meanings or change their behavior.

We have developed a wheeled mobile android, *ibuki*, who is equipped with a mechanism that replicates a human-like upper body oscillating motion induced by gait in the mobility unit. We call *Vertical-Oscillation Mechanism* for this mechanism. With the combination of motions through the vertical-oscillation mechanism and joints of the upper body, the android

generates a human gait-like motion while moving despite being wheel driven. We call this gait-like upper body motion a *gait-induced upper body motion* in this thesis. As I named the title of this thesis “Perception and Behavior”, we evaluated the effectiveness of the gait-induced upper body motion of the android from two perspectives: emotion perception and gait synchronization.

For the emotion perception, firstly we show an importance of the body even for a human-like facial expression, which is the most characteristic part of an android. We validated that an ambiguous facial expression of an android could be perceived more clearly by viewers when body postures and movements were added. The experiment result shows that the facial valence distribution of the Intense emotion had the highest entropy with high intensity, as has been reported in previous human studies. This facial valence of Intense could be clearly distinguished by adding body postures and movements. Viewers were able to distinguish the ambiguous facial valence with specific combinations of neck, arm, and vertical motion.

Secondly, we implemented in the android three types of emotional gait-induced upper body motions (anger, happiness, and sadness), and validated how viewers perceive the emotion and their confidence level in their answer. The experiment result shows that the emotional expression of happiness using the vertical oscillation was better perceived by viewers (with higher recognition rates and higher confidence levels) than that of without vertical oscillation. The incongruent application of anger/happiness vertical oscillation decreased the recognition rate of happiness/anger emotional expression. Furthermore, the emotional expression of happiness using the vertical oscillation for anger decreased the confidence levels; however, the emotional expression of anger using the vertical oscillation for happiness increased the confidence levels.

For the gait synchronization, we validated the synchronization of gait phases between a human and a wheeled android. We measured the gait phase difference between the android and a human walking behind. The moving android operated its upper body in three distinct types of motion patterns: no-motion, arm-swinging, and arms-swinging with a periodic vertical oscillation. The experiment result shows that a significant distributional bias due to the phase locking by the gait synchronization was confirmed when participants were walking only with the android controlled in the third motion pattern with the vertical oscillation.

Lastly, we discussed the implicitly control for both of robots and humans as a new social navigation strategy. This strategy regards the entire environment surrounding the robot as a system having weak controllability. In order to successfully move not only robots but also the surrounding humans, it is important to increase the controlability of human side. For this, robots must have capability to convey the intention to humans through such as the emotional empathy and gait synchronization which we investigated in this thesis.

Table of Contents

Chapter 1

Introduction

1.1 Mobile robots in our daily lives

In response to the rapid growth of global economy and the workforce shortage against it due to the declining birthrate, aging society and rising worker wages, various countries all over the world are investing in the field of robotics to develop robots which become the alternatives of humans. One market research company reported that the global robotics market is valued at 27.7 billion USD in 2020 and is expected to grow up to 74.1 billion USD by 2026 [1]. And they also analyzed that there are significant increasing investments in industrial robots across the globe.

The reason why the investments are concentrated in industrial robotics applications in a factory is due to the ease of designing the working task and environment. Let's take an arm robot as an example of a typical industrial robot, although the grasping ability of robot hands has not yet reached the level of humans, we can predetermine what to be grasped and adopt the suitable grippers as a robot hand to effectively grasp certain objects. In the case of mobile robots in a factory, to improve the transportation efficiency, we can adopt special travel lanes and specially designed cargo racks to make it easier and faster for robots to carry products. Designing an efficient work environment is also easier in a factory. To increase the work speed, even if each robot works at a limited slow speed, an operator just needs to install multiple robots as many as meeting the required work speed. Or a robot can move as much fast as required, if it works in which separated from human workers where unexpected collisions are not occurred. Furthermore, as many sensors as needed for the robot operation such as cameras for vision can be installed anywhere whether on the robot body or outside.

Due to COVID-19, new lifestyles are being accepted to our society which encourage non-contact and non-face-to-face, such as a remote work and an online meeting. This social situation change has led to grow a demand for the mobile robot application to replace labors where previously humans worked not only in factories but also in human living environments. In particular, the demand for mobile service robots is rapidly increasing to replace for professional jobs such as cleaning, building maintenance, field research, and logistic, for household jobs such as automatic vacuum cleaners, security monitoring devices and nursing care, and for entertainment applications such as hobbies and education usages.

As robots become more widely used in our living environments, we expect that the future society in which humans and robots' symbioses, where teleoperated mobile robots controlled by a remote human operator or autonomous mobile robots move on the same sidewalks in a city. Automobiles and trains are representative examples of the symbiosis between humans and machines in the current society. Cars go on the road and trains run on the rail; their working environment of such mobile machines is designed clearly separated from where humans walk. However, it is necessary that mobile robots, which are required to work as alternatives of humans, can move in the same space where humans are as shown in Figure 1.1. As other examples, a robot which displays items in a supermarket needs to move in a narrow aisle while avoiding shoppers. A service robot which carries goods to a guest room in a hotel is required to move around the floor where humans come and go, and to ride an elevator to reach the destination. Therefore, the boundaries which have separated the working environments of humans and robots will become even more blurred, and technology will be required to realize symbiosis between humans and robots in the next era.

It is still a big challenge to achieve robots moving freely in a large pedestrian crowd such as Figure 1.1. The existence of humans makes harder robots to move around. It is much more difficult for robots to recognize irregularly moving pedestrians and predict their future trajectories than it is for them to simply avoid static obstacles. When the moving robot is surrounded by many humans, it is necessary to consider how far away pedestrians should be recognized and included in the robot's path planning. In addition, there are often unexpected situations where the robot cannot accurately recognize the surrounding humans nearby, such as when the robot's visibility is poor due to bad weather or when there is a wall at a corner. Furthermore, changing environments caused by humans are also critical issues for robots. Robots must deal with unexperienced situations such as an obstacle placed by a human on the way of robot's path and a construction work which requires a detour. What much more difficulty for mobile robots is that robots are required to move in the same way without disturbing the pedestrian flow as where pedestrians move in a high dense and fast flow such

Fig. 1.1 Where robots should be operated in the next era. This figure shows a large crowd of humans in an underground space of a train station rushes to walk towards their destinations. Both of humans and robots need to move together within the same crowd.

as a train station or an airport. These are the major reasons why it is difficult to realize a robot which can move freely in a human living environment and why it is still being actively researched.

Robot navigation strategies in human living environments are mainly categorized into two approaches. The first strategy is for a long term, related to path planning and self-localization necessary for a robot to move to its destination. This strategy has been actively researched at the topic of Simultaneous Localization and Mapping (SLAM) and has applied to industrial mobile robots and self-driving cars. The second strategy is for a short term required a more real time control, regarding as a collision avoidance. Although this strategy is not directly related to the main robot goal of moving to the destination, it is necessary for achieving safe mobility in our society. Therefore, collision avoidance methods for human safe have been researched, by estimating a human position and future trajectory which is difficult to predict.

For example, a simple navigation method accompanying to a specific person who seems to go to the same destination [2, 3], a path planning method to calculate an artificial potential field based on human positions around the robot and generate a safe trajectory without contacting a person [4], and a pedestrian's path prediction method using a neural network [5, 6] have been proposed. Although enormous research has been made to model pedestrian behaviors and to use it for a robot navigation, there are still issues difficult to solve such as the freezing robot problem, in which a robot falls in a deadlock state and move nowhere as a result of recognizing pedestrian around it.

Considering practical usage, those approaches described above which the robots precisely perceive human behaviors and behave appropriately are just one direction of the solutions for a robot navigation in a human living environment. As the other direction, it is important that human sides perceive the robot's behavior and move cooperatively without getting in the way of the robots. While the robot is perceiving the human's behavior and taking action to avoid collision, the human is also perceiving the robot's behavior and taking the next step as the same way of the robot to avoid collision. Although this human cognitive property has often been ignored in traditional mobile robot research, some studies have successfully taken advantage of utilizing this cognitive property for mobile robots. As an example, that is easy to implement in a robot, a method has been proposed that uses technical expressions such as light cues to make humans aware of the robot's approach and destination [7, 8].

From the human side, a moving robot is essentially an obstacle. Therefore, the interaction between a mobile robot and a human should be to keep a minimum. Except to operators and collaborators who are directly involved with the robot's task, most of the humans in the same environment with the robot have different purposes, and there is almost no benefit to interacting with the robot. Even if the robot tries to send some cues to the humans nearby, humans might not make any effort to catch the signals. Taking an example of the freezing robot problem described earlier, when there are many pedestrians moving around and the robot stops on the spot to avoid contact with pedestrians, it could be solved if some pedestrians around the robot open a space for the robot to move. However, in the most cases, humans do not care about the robot's behavior and keep walking, so the robot cannot escape from that situation.

There is another factor which makes it difficult for mobile robots to cooperate with humans. In the case of home appliances and cars, humans spend time to learn how to use those, however, the human facing the robot is not given the opportunity to learn the principles of its operation and must successfully interact with the robot in the situation of first meeting it.

In general, it takes several hours for a human to learn how to operate a home appliance such as a washing machine, and several dozen hours to learn how to operate a car. Normally, a training period is necessary for successful interaction between humans and machines, but we do not have the opportunity to train smoothly behaving with a robot which we meet on the street in advance. Since it is not possible to learn in advance what cues the robot shows, in research on communication robots, it has been reported that it is easier to convey intentions to a human facing to the robot expresses itself in a human-like manner such as gestures, rather than using robot-specific functions such as glowing eyes [9].

Why can pedestrians successfully collaborate to walk in the crowd? Gait is a main factor to enable the pedestrian collaborative mobility in a society. A gait is a pattern of limb movements made during locomotion. Humans convey various information to the surrounding others from gaits while walking, and observers can perceive and understand meanings from their gait. With this, pedestrians spontaneously form a walking flow and achieve high-density and high-efficiency movement while avoiding collisions with each other. The information conveyed by a human gait can be classified into physical property such as the orientation of the destination, a walking speed, and a walking cycle, as well as psychological property such as the personal space and the purpose at there. From the viewpoint of each part of the body during walking, the orientations of the head and torso implicitly indicate the destination. On the other hand, the orientation of the lower limbs indicates the short-term destination in which the legs will step next. It is known that the personal space, which is the space where we feel uncomfortable when we are approached by others, increases as the walking speed increases. The pitch, stride and the joint amplitude of arm swing depend on gender and age [10, 11, 12, 13, 14, 15, 16, 17, 18, 19]. In addition, a gait changes greatly with emotions. Roughly, positive emotions accelerate the gait cycle, incline the upper body backward, and increase the body swing when stepping. Vice versa for negative emotions [20, 21, 22, 23, 24]. Interestingly, those characteristics are expressed almost unconsciously. As an interesting example, from images taken by a security camera installed in a store, it is possible to detect shoplifters with analyzing their suspicious gaits. Since a gait is expressed with the whole body, it can be perceived from a long distance. Therefore, it is said to be able to identify individuals with their gaits far from 50 to 100 meters, depending on the resolutions of the camera used for the gait recognition, compared to a fingerprint recognition and face recognition which are effective only at a short distance.

1.2 Research topics

If mobile robots can influence the perception and behavior of humans who firstly meet the robot by performing a body expression comparable to a human gait (gait-like motion), it will be useful to convey information from the robot to humans. This gait-like motion is intuitive even for humans who are interacting with the robot for the first time, without having any specialized knowledge about robots. And it is possible to convey various information continuously and extensively to the surrounding humans at once, as described above. The gait-like motion can be expected to implement for whichever robots if it has a human-like body by simply controlling the robot's body periodically as it moves. By using the robot gait-like motion while moving, humans can recognize the meaning of the robot's action. With using this, the robot can elicit human actions which the robot wants. Furthermore, it can be a natural starting point for the next interaction while performing the task of moving. For example, moving while scurrying and looking around can tell the others that the robot is not familiar with where it is and expects human's help.

We believe that a gait-like motion by a mobile robot during moving can be applied as a new navigation strategy in mid-term of Figure 1.2. This mid-term strategy enables the robot to interact with humans farther away in advance before performing the short-term strategy. In Figure 1.2, all the mobile robot and humans have each long-term navigation task and move to different destinations each other under the same environment (green circle). During this, it is sometimes unavoidable to approach others (red circle). Therefore, the robot must plan a short-term navigation strategy to avoid collisions with others. If the robot can elicit human actions which the robot want through the gait-like motion in advance (blue circle), the robot can minimize short-term collision avoidance events which may occur due to the unexpected action of the robot and humans and concentrate on performing a long-term navigation task.

Can wheeled mobile robots, which are more common than bipedal legged robots, express the gait induced by the leg movements? We can easily imagine the gait-like of a bipedal robot, however most of mobile humanoid robots designed to be applied in human living environments have a wheeled mechanism as the lower body and the anthropomorphic face and arms as the upper body. It is worthwhile to study this question and the application of the gait-like motion for mobile humanoid robots. The wheel mechanism is safe and easy to control as a robot mobile mechanism. Wheels have stability with a low risk of falling and keeps standing even when the power is off. In the case of a human size robot, the joint structure of

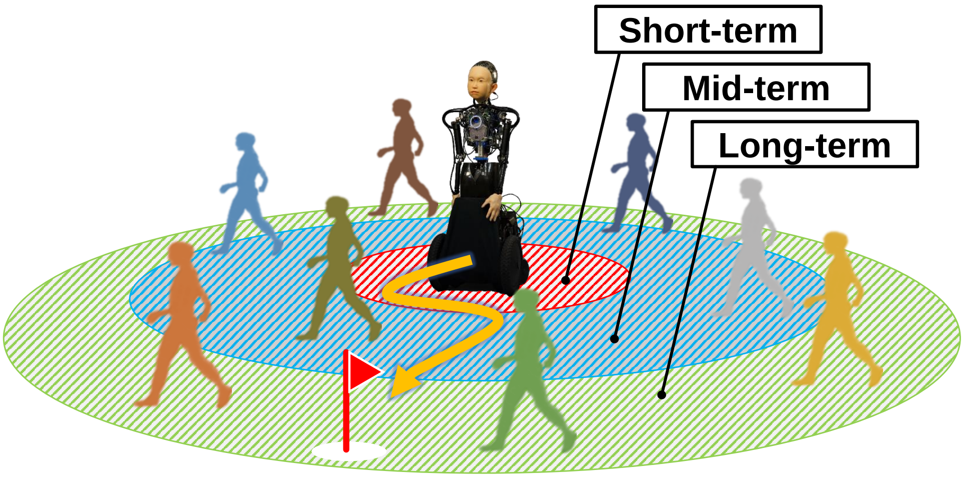


Fig. 1.2 Navigation strategies with three stages for a mobile robot in human living environments. Our proposed body expression during moving can be categorized into a strategy in mid-term.

the wheeled mechanism is simpler than that of the legged and the energy efficiency during movement is usually higher. Thus, most mobile robots have no legged but wheeled. Due to the kinematic wheel-axis stability, wheeled robots do not swing the upper body which human legs induce during walking. Furthermore, the anthropomorphic upper body is useful for working in a human living environment because many things around us are designed to be easy for humans to handle. Taking the door-opening task as an example, a human needs to manipulate a doorknob by grasping it with the hand and pulling it open with the arm. In addition, the face can be useful for interpersonal communication and teleoperation. For these reasons, a wheeled humanoid with the anthropomorphic upper body is a practical body configuration for applications in a human living environment. Therefore, investigating whether a wheeled humanoid can express the upper body motion induced by a gait is useful for real world robot applications.

This thesis aims to propose that a wheeled android expresses an upper body motion induced by a human gait during moving and to validate whether humans who see the upper body motions can perceive the meanings or change their behavior. Firstly, we have developed a wheeled mobile android, *ibuki* [25], who is equipped with a vertical oscillation mechanism which creates upper body motion like that of the human upper body while moving in its mobility unit. With the combination of motions through the vertical oscillation mechanism and the joints of the upper body, the android generates a human gait-like motion while moving despite being wheel driven. We call this gait-like upper body motion *gait-induced upper body motion* in this thesis. From the viewpoint of practicality, we evaluated the effectiveness of the gait-induced upper body motion of the android from two perspectives: emotion perception and gait synchronization, with three research.

For emotion perception (Research 1 & 2), we focused on that human gaits are emotionally expressive. Humans walk in various ways depending on their emotions and viewers can perceive the emotions from others through their gaits. We believe that these various expressions through the gait are crucial to achieve a smooth crowd movement as for forming a pedestrian flow or adjusting the distance between each other. Before implementing emotion expressions into the upper body motion, we verified whether the body is essential for a human-like facial expression, which is the most characteristic part of androids (Research 1). And then, we implemented in the android emotion expressions of gait-induced upper body motions (Research 2).

At Research 1 of Chapter 4 shows that an ambiguous facial expression of an android can be perceived more clearly by viewers when body postures and movements are added. It is widely recognized that facial expressions are important for judging human emotions. However, recent research has found some cases in which emotions may not be clearly distinguishable from facial expressions [26]. Aviezer found that when participants were shown peak expressive reactions to winning and losing points in professional high-stakes tennis matches, it was difficult for participants to correctly judge whether the emotion was positive or negative through intense facial expressions alone, however, they could perceive emotions when these were seen with the body [27]. It was found that humans recognize facial emotions by unconsciously considering factors other than facial expression. In previous research on robots that have emotional expression functions, facial expressions have been used to convey emotions on the assumption that facial expressions have a one-to-one correspondence with certain emotions. Thus, it still has not been investigated whether robots can express the ambiguous facial expressions with no clear valence which humans sometimes show, and what these look like. Furthermore, it has also not been investigated whether the robot's emotional valence can be clearly determined from such expressions by adding body expressions. This research claims that the importance of the body for the android and enforces the importance of investigating a gait-induced upper body motion to express emotions.

At Research 2 of Chapter 5, we implemented in the android three types of gait-induced upper body motions (anger, happiness, and sadness), and validated how viewers perceive the emotion and their confidence level in their answer. Many studies have been conducted to investigate the characteristics of emotional gaits [20, 21, 24, 28]. For human-robot interaction, considering the expression of various gaits, including *emotional* gaits, is valuable to build deep human-robot relationships. In the field of bipedal robots, some research has been conducted to realize human-like gaits [29] and further emotional gaits [30, 31]. However, to our best knowledge, the human perception of emotion expressions of wheeled mobile humanoid robots while its moving has not been studied yet. If the robot could express its intentions and emotions while moving, humans could act appropriately according to that information. As a result, robots and humans could achieve orderly interactions by being able to predict and understand each other's behavior in various situations.

For the gait synchronization, at Research 3 of Chapter 6, we investigated the synchronization of gait phases between a human and a wheeled humanoid robot. We measured the gait phase difference between the robot and a human walking behind the robot. It is known that humans walking in large crowds can walk collectively, as they unconsciously synchronize their gait with the surrounding pedestrians [32, 33, 34]. Previous studies on pedestrian gait synchronization have reported that gait phase synchronization is more likely to occur in high-density groups and that pedestrian synchronization also maximizes the walking flow efficiency [35, 36]. If humans can synchronize a gait phase with robots, we can imagine applications with smooth robot-human mobility in our daily lives. In particular, if wheeled robots can synchronize gait phases by utilizing upper body motions, robots can move by taking advantage of the human ability to walk skillfully in the crowd.

1.3 Thesis outline

The outline of this thesis is followed; Chapter 1 described the challenges for social implementation of mobile robots especially when moving in a large pedestrian crowd, the importance of the human-like gait expression, and the three research topics of this thesis. Chapter 2 introduces related research focusing on body expressions, including gaits, in mobile robots.

Chapter 3 describes a development of the wheeled child android (*ibuki*). Chapter 4 shows that an ambiguous facial expression of an android can be perceived more clearly with adding body postures and movements (Research 1). Chapter 5 shows that the implementation three types of emotional gait-induced upper body motions and investigated whether viewers can perceive the original emotion with high confidence (Research 2). Chapter 6 verifies whether the gait phases of humans who follow the android which performed the periodic upper body motions while moving can be synchronized for that of the android (Research 3). Chapter 7 discusses the importance of the proposed gait-induced upper body motion for wheeled mobile robots and shows the future direction of this thesis. Chapter 8 summarizes this thesis at the end.

Chapter 2

Related research

This chapter introduces related research focusing on body expressions of mobile robots. For this purpose, firstly we will briefly explain the history of robotics and challenges which needed to be solved in our current society, focusing on the field of mobile robots and humanoid robots.

2.1 Current situation of robotics

The term of "Robot" was first used in R.U.R., a play by Czech writer Karel Čapek in 1920 [37]. In the play, robots were used as machines which imitated the human appearance to work as alternatives of humans. The definition of robot has been described in various fields since then, but generally speaking, it is safe to say that a robot is a machine that has the three elements of calculator, sensor, and actuator. For example, an industrial arm robot working in a factory is a robot because it works by repeating a process in which a control computer (calculator) estimates the working environment and the current state of the robot, based on information from joint angle sensors and camera image sensors (sensor), and sends the motion commands necessary for performing the task to the motors that drive each joint (actuator). A flying drone in the air is also a robot because it uses IMU, gyro, and GPS sensors to sense its own posture and position in the air, and based on that information a control computer sends the motor commands to the motors that drive its propellers in order for balancing itself or flying to the destination. Nowadays, robotics, as a discipline which deals with technologies related to robots, sometimes covers a wider range of subjects, such as robots which do not have a calculator, sensor, or actuator [38, 39, 40]. Although many robots have been studied, the realization of robots which can replace humans is still one of the ultimate goals of robotics research, even after 100 years have passed since the term was first coined. Next, we introduce the development of the robotics field which aimed to this goal, especially mobile robots in a human living environment. And also, we will clarify what has been achieved so far in robotics technology and what issues need to be addressed for the future social implementation.

In the 1960s, SRI International developed the world's first mobile robot, which senses an environment and navigates itself, known as Shakey. Through the development of this robot, essential technologies for autonomous mobile robots, such as environment recognition and path planning, were studied [41]. On the other hand, a science fiction anime Astro Boy was broadcast in Japan and the U.S., which is one of the most successful manga and anime franchises in the world and have sold over 100 million copies worldwide. Its popularity led to the cultural familiarization of robots. While research on robots which could move in human living environments began in earnest, robots got culturally accepted.

In the 1970s, Waseda University started research on a human-size humanoid robot, WABOT-1 [42]. It was developed under the design concept that a robot working in a place of humans needs to have the same body and perception as humans, with the ultimate goal of creating "My Robot" that is as close to a human as possible. In the 1980s, Honda started to develop the E series of bipedal robots, which later became Asimo, and through the research of the E series robots, the basis of the currently widely used walking control algorithms for legged robots, such as the idea of static walking, dynamic walking, and the target zero moment point control, was developed. In the 1990s, Sony announced the concept of aibo which was a household pet robot [43].

In the 2000s, Honda developed Asimo [44], which gave a significant impact on later humanoid research, and the way Asimo poured water into a cup and moved upstairs in situations which replicated our living environments minded us of the availability and challenges of robots in our society. In the same decade, aibo, the household pet robot, and Rumba, iRobot's automatic vacuum cleaner, went on sale, and robots which could operate in our house became popular.

In 2011, the Tohoku Earthquake and the Fukushima Daiichi nuclear disaster occurred, where Packbot and Quince were used to investigate the inside of the nuclear power plant building at which radiation leaked [45, 46]. In response to this accident, the DARPA Robotics Challenge was held from 2013 to 2015, where several humanoids performed tasks required of humans, such as opening doors, opening valves, driving cars, and walking rough terrain [47].

In recent years, humanoid robots with human-like appearance and expressive capability have been presented mainly in the entertainment field, such as Audio-Animatronics of Disney [48, 49], Sophia by Hanson Robotics [50], and Pepper by Softbank Robotics [51]. In the field of mobile robots, numerous autonomous mobile robots have been developed for applications such as product delivery in hotels and disinfection cleaning in hospitals. Furthermore, response to needs for robot applications in the service industry, several mobile humanoid robots, which are designed to interact with humans and work in a place of humans, have been developed. Telexistence Model-T and Model-H are humanoids designed to work in environments where humans work. OriHime-D is a tele-operated mobile robot which can express emotions, although it does not have much freedom of physical expressions [52, 53, 54].

2.2 Importance of robot's body expressions

Many studies have pointed out the importance of nonverbal communication for robots to communicate smoothly with humans in society. The modalities used for nonverbal communication are diverse, including facial expression, breathing, gaze, gesture, posture, and distance. In addition, nonverbal communication is essential not only for humans, but also between humans and animals (e.g. pets and their owners). It is the same for different species of animals (e.g. predators and preys in nature). Nonverbal communication is usually expressed and conveyed unconsciously to the other human and the person also infer unconsciously the intentions and feelings. The expressions of nonverbal communication are difficult to control by oneself, and humans tend to regard that the nonverbal expressions express the true intentions and feelings more than the verbal. To utilize this embodiment-based communication between humans and robots (or computers), enormous efforts have been made in both directions: one direction is that robots attempt to recognize nonverbal expressions by humans and the other opposite direction is that robots attempt to express in nonverbal to humans. In recent years, with the significant improvement of computing power, image and voice recognition technologies based on deep reinforcement learning technology have been applied into practical use in our society. This has made computers possible to recognize such as the meaning of human nonverbal gestures and the emotions from human faces or voices. With this, many robotics studies have been conducted by recognizing human nonverbal communication for a robot's behavior. On the other hand, there are relatively less studies on robots expressing nonverbal expressions to humans, especially focused on body expressions.

The most well-studied situation in which a robot tries to convey something to a human through a nonverbal expression is where a human and a robot face up each other with having their interaction already started. It is considered that robots which interact with humans should physically express internal states such as reactions and emotions to regulate human behavior who see the expressions for improving the task performance and impression. As the physical expression, facial expression, gaze, posture, and gesture are mainly focused to research in robotics. Compared to the facial expression and gaze which require a precise control of expressions, body expressions of posture and gesture can be easily implemented even with a limited degree of freedom if a robot has an anthropomorphic appearance.

Emotional expressions in nonverbal communication are crucial for deep relationship between robots and humans. Previous research has been studied methods to convey emotions through body expressions by humanoids. For example, it is reported that the implementation method of emotional expressions for a tabletop humanoid robot Nao by combining human-like body motions with robot-specific sound and light functions [55]. As research on implementing emotion expressions for a real human size humanoid, a bipedal humanoid Kobian is known [30]. The research team investigated that the recognition rates of Kobian's emotion expressions by the face-only, body-only, and combining both as the whole body [56]. Later, they developed an additional humanoid, Habian, in which the upper body was the same with Kobian but the legs were replaced by a wheeled unit, to investigate the recognition difference of emotional expressions by legged and wheeled humanoids [57].

From the viewpoint of effectiveness against human actions, body expressions by humanoids have also been studied, such as combining gestures and utterances when explaining directions [58]. Even handing and receiving an object by a humanoid arm, the applications of human-like body expressions have been reported effective. It is reported that a human reacted faster for an object when the person was receiving it from a humanoid hand which the hand trajectory was controlled in a minimum jerk as the way a human does [59]. In addition, it is also reported that a human felt more comfortable when the person was handing an object to a humanoid robot which reacts it with a time delay in perceptual response like a human [60]. For dance teaching, a wheeled mobile humanoid has been developed which utilizes the human-like upper body and waist joints with torque-controlling those joints while dancing with a human practician [61, 62].

Several studies validated effectiveness of body expressions by humanoid robots while moving. Emotion expressions through gaits by bipedal humanoid robots have been reported [30, 31, 63]. It is reported that a humanoid robot enables to express the intention to require a person standing in its path makes room for the robot, by raising the arms forward during moving. In this method, if the person still does not notice the robot intention, the robot directly touches and moves the human to make space for the robot [64]. A robot's backward motion for giving way was also studied. This motion shows the intention, by a human-like manner, to avoid collisions when the moving trajectory interferes with that of a human [65]. (Note that, the experiment was conducted with using a simple wheeled mobile robot attaching a monitor on top.) In addition, it has been reported that when a robot and a person pass each other on a narrow path, such as a store aisle, the impression of the robot was improved if the robot rotates the body to open a space [66].

There has been no research on body expressions of a walking manner by a wheeled humanoid during moving. Previous study was limited to study the emotion perception by changing parameters of the moving speed and direction [67]. As mentioned in the previous chapter, if a wheeled humanoid expresses the upper body motion which induced by human gait and humans can perceive the meanings and make behave as the robot wants, this expression is useful as the way interacting humans around the robot while moving. Most of commercialized robots do not equip upper limbs because what robots can do is limited compared to the cost of implementation. Or even those with upper limbs have difficulties in performing quick and smooth human-like movements due to the heavy weight and limited motor power supplied by batteries. In the first place, a mobile robot equipped with the upper body in human-size was expensive and as big as making difficult to handle in research. Therefore, it was necessary to develop a new wheeled mobile humanoid robot which can perform dynamic upper body motion as like a human for this thesis. In the next chapter, we introduce the wheeled humanoid robot which we developed to validate the usefulness of the aforementioned body expression, and in particular, we explain the mechanism to realize the upper body motion induced by a human gait during moving.

Chapter 3

Development of *ibuki*

This chapter shows an electrically driven childlike android named *ibuki* equipped with a wheeled mobility unit and a human-like upper body. *ibuki* can perform various human-like body expressions and move in human living environments with using 46 degrees of freedom. Since the mobility unit includes a vertical oscillation mechanism, *ibuki* can replicate the movements of the human center of mass, and can express human-like upper-body movements even when moving by wheels. In the latter half of this chapter, we will describe this mechanism and the *gait-induced upper body motion* which performs the upper body behave as if the motion were induced by leg walking.

3.1 Background

An android, which has a human-like appearance, can interact with humans naturally owing to various features. For example, the expressive face, which has multiple degrees of freedom covered with human-like skin, can convey its internal states and emotions through changing facial expressions. Humans use such expressions in the daily communication, and they can therefore intuitively understand these expressions even if which are expressed by robots. This rich-expressiveness ability should contribute to establishing a deeper relationship between humans and robots. Various androids have been introduced so far [68, 69], most of them representing adult persons who has a specific role, such as a fashion model [70], a conversation partner [71], and a salesperson [72], to embed androids in a daily life. Due to implement androids into a pre-designed situation as a particular role, most of androids were fixed a place in the working space and do not have a capable of mobility.

Since most androids do not require mobility for the applications, hydraulic and pneumatic

Fig. 3.1 Introducing *ibuki*: a childlike android with mobility.

actuators with an external large compressor are commonly employed to realize robots' joint drive. Relatively less electric-drive androids with a moving mechanism have been introduced. A cybernetic human HRP-4C is a humanoid robot that has a human-like face and hands and can walk using legs [73]. EveR-4 is an android with a wheeled mobile mechanism [74]; however, the studies dedicated to this robot have mainly focused on analyzing its facial expressions, and the details concerning its ability to perform wheel movement have not been reported.

Figure 3.1 shows our developed electrically driven childlike android with a mobility unit, named *ibuki* (*ibuki* means *breathing* in Japanese). To investigate human–android interaction in our daily living, we adopted a childlike appearance for *ibuki*. We expect the child android to be naturally accepted in society without the need for a specific role as children do. To ensure the mobility and stability in a human living environment, we adopted a wheeled mobile mechanism as its lower body. A wheel mechanism of a large diameter enables to move not only on a robot-friendly flat floor but also on a slightly rough outdoor road. It is essential for androids to provide not only a human-like appearance but also human-like behavior. We developed a mobility unit, which enables *ibuki* to mimic movements corresponding to human gait-induced upper body motion and to imitate walking even though it is wheel-driven.

In sum up, we believe that *ibuki* would have advantages in terms of the following two aspects: (1) Childlike appearance to easy acceptance in human society. As expectations corresponding to children's abilities are below than those of adults, a childlike android may be acceptable even if it is inferior to humans. This covers technological gap between what the current robots can do and what people expect for robots. (2) High mobility of a wheeled mobility unit. The wheeled mobility is deemed more stable than bipedal locomotion. Even if power is not supplied, a robot can keep standing still without falling. When unexpectedly contacting with a human, the childlike body reduces this risk owing to its small size and weight.

3.2 Child android *ibuki*

The child-like mobile android (Figure 3.1) that we developed called *ibuki*, is 120 cm tall and is comprised of two parts, a mobility unit (lower part) and the upper body, which is designed based on a Japanese boy. The face and hands are covered with silicone skin to have a human-like appearance. An electric motor drives each joint, and mobile batteries are used as the power supply. Table 3.1 and Table 3.2 outline the basic specifications and the degrees of freedom of *ibuki*.

The dimensions of *ibuki*'s body are specified according to the average parameters of a ten years old Japanese boy [75]. We decided to cover only the minimum areas with skin, the head and both hands, so as not to lose the human-likeness of *ibuki*. Facial appearance is crucial to express emotions. Hands with flexible skin can be utilized to touch humans, as when shaking hands or during other such interactions. In the design of other body areas, the mechanical parts are intentionally exposed to reduce the uncanny. To achieve lightweight and high strength, we adopted glass fiber reinforced nylon and carbon fiber reinforced plastic parts for the body. Also we reduced the number of parts by using a 3D printer to print several parts in one piece.

Figure 3.2 represents the mechanical structure of *ibuki*. It has a total of 46 DoF, including two driven wheels (See Table 3.2). Table 3.3 provides the principal joint specifications. The joint range shows a structurally movable angle range at the time of design. All actuators are DC geared motors and motor driver modules equipped with current sensors enable *ibuki* to

Table 3.1 Basic Specifications of the Mobile Android *ibuki*

Height		1200 [mm]
Weight with batteries		38.6 [kg]
Processor		Jetson AGX Xavier (NVIDIA Corp.)
Microcontrollers		mbed LPC1768 (ARM Ltd.)
Sensors	Joint	- Potentiometer (SVK3A103AEA01 and SVM4A103A0L17R00, Murata Manufacturing Co., Ltd.) - Current sensor (INA226, Texas Instruments Inc.)
	Head	- Built-in RGB camera × 2
	Upper-body	- RGB-D camera (Intel RealSense Depth Camera D435, Intel Corp.) - IMU sensor module × 2 (RT-USB-9axisIMU2, RT Corp.)
	Mobility unit	- LRF (RPLIDAR A3, Shanghai Slamtec Co., Ltd.) - Microphone array (ReSpeaker Mic Array v2.0, Seeed Technology Co., Ltd.) - IMU sensor module - Wheel: Absolute magnetic rotary encoder (AEAT-6012, Broadcom) - Linear motion: Linear potentiometer (LP-150FJ-1K, MIDORI PRECISIONS)
Audio	Upper-body	- Speaker (JBL Clip2, Harman International Industries, Inc.)
Batteries	Actuation + USB power source (5V)	Li-ion DC 25.2 V battery × 2 (7LPL0678G8C1-1P01, Maxell, Ltd.)
	Xavier	Li-ion AC 100V battery (RP-PB055, Sunvalleytek International, Inc.)

Table 3.2 Degrees of Freedom of the Mobile Android *ibuki*

Description	DOF
Total	46
Head	15
Neck	3
Arm	6×2
Hand	5×2
Waist	3
Mobility unit	Wheel 1×2 + Linear motion 1

control its joint torque. The central computer sends reference values to each microcomputer via Ethernet. The angle of each joint is controlled at 100 Hz.

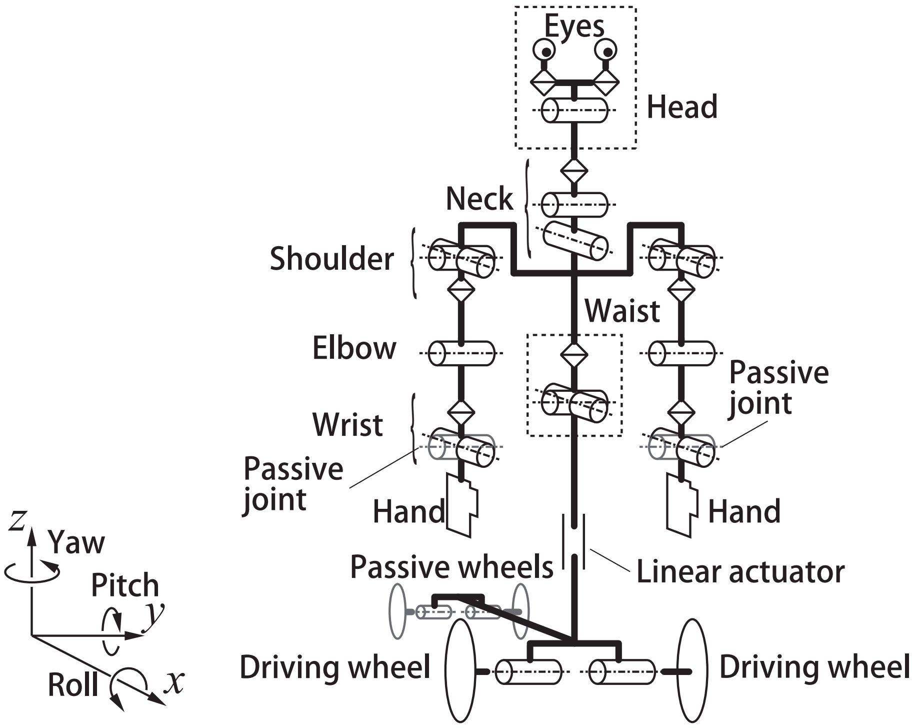


Fig. 3.2 The mechanical structure of *ibuki*. The wrist roll axes of both arms are specified as passive joints. Mechanical springs are attached to each joint.

Figure 3.3 (a) represents a 3D CAD image of the waist; (b) represents the schematic image of the structure; and (c) shows the photograph of the joint. We designed *ibuki*'s waist with three DOF joints with three orthogonal intersecting axes. For the roll and pitch joints,

Table 3.3 Principal Joint Specifications

Description		Joint range [degree]			Max torque [Nm]
Eyes	Pitch	-20	to	10	0.3
	Yaw	-40	to	40	0.1
Neck	Roll	-12	to	5	0.3
	Pitch	-30	to	40	1.5
	Yaw	-90	to	90	4.4
Shoulder	Roll	-162	to	11	4.9
	Pitch	-180	to	180	6.7
	Yaw	-90	to	90	1.8
Elbow	Pitch	-132	to	90	1.8
Wrist	Roll	-106	to	115	1.0
	Yaw	-80	to	80	1.0
Waist	Roll	-38	to	48	7.3
	Pitch	-20	to	23	7.3
	Yaw	-180	to	180	3.6

mechanical springs and dampers are attached to reduce passive vibration of the upper body while moving. For the yaw joint, a motor driving is transmitted by a belt from a motor whose rotational axis is placed along the x-axis. For further details of the upper body joint structure, please refer to our paper [25].

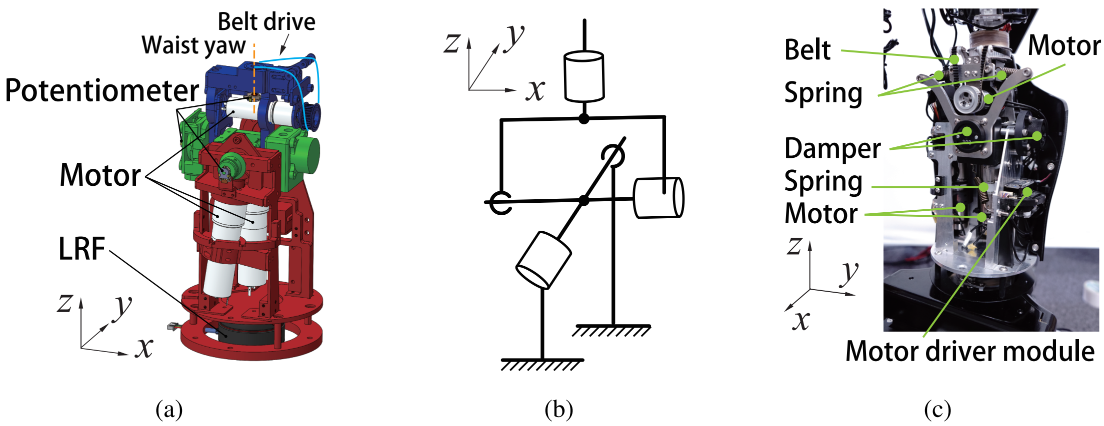


Fig. 3.3 Waist joints: (a) three-dimensional computer-aided design of the waist joint mechanism; (b) a schematic image of the waist joint; (c) a photograph of the waist joint.

Figure 3.4 represents the electrical system embedded into *ibuki*. The utilized on-board main computer, Jetson AGX Xavier, is mounted on the mobility unit. Since *ibuki* was developed to move around in a human living environment, it is necessary to ensure *ibuki* working even when the network connection becomes unstable. Thus, this on-board computer is used for applications where relying on the computational power of the cloud or server would delay the response time. The other additional calculations are performed on an external computer. Each motor driver module is connected via Ethernet, and the sensors are connected via USB.

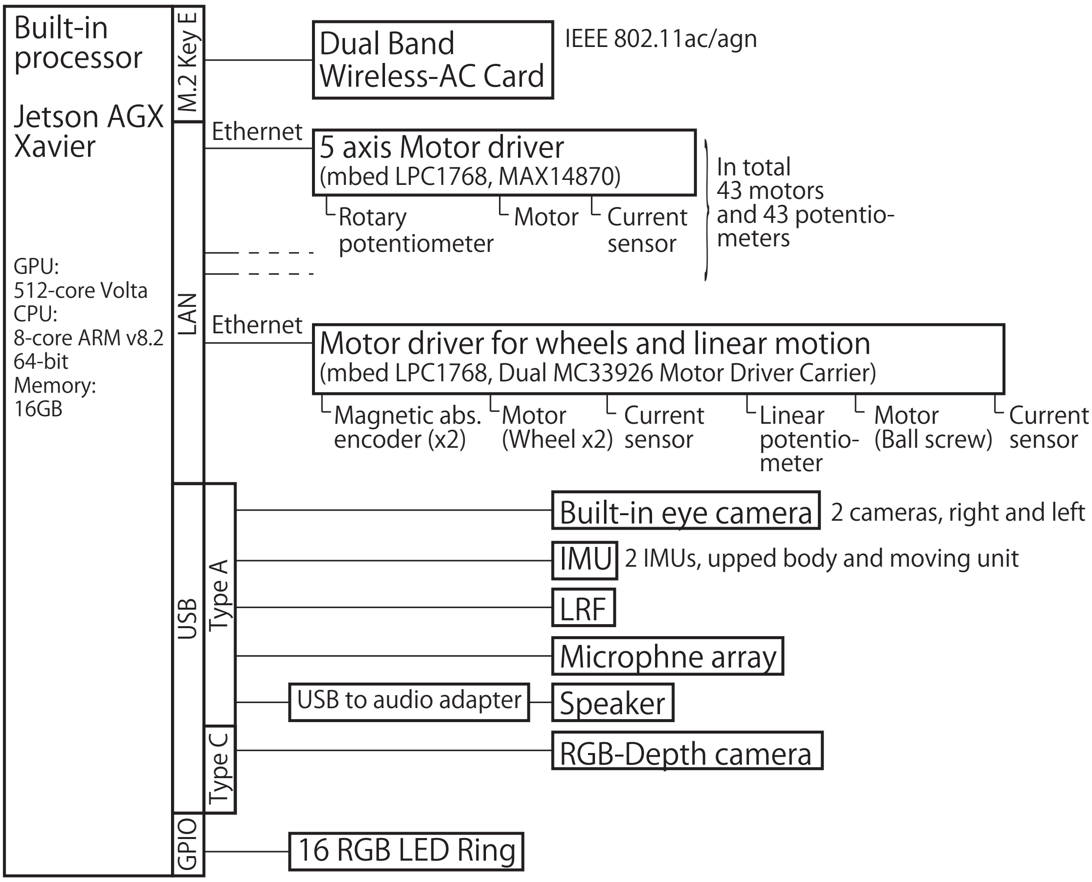


Fig. 3.4 Electrical system of *ibuki*.

Figure 3.5 outlines the arrangement of sensors and other devices within *ibuki*. The laser range finder (LRF) senses the surrounding obstacles and people. The RGB-Depth camera on the torso detects a person in front of the android. A small USB camera mounted inside each eye detects a face or an object and execute the gaze control.

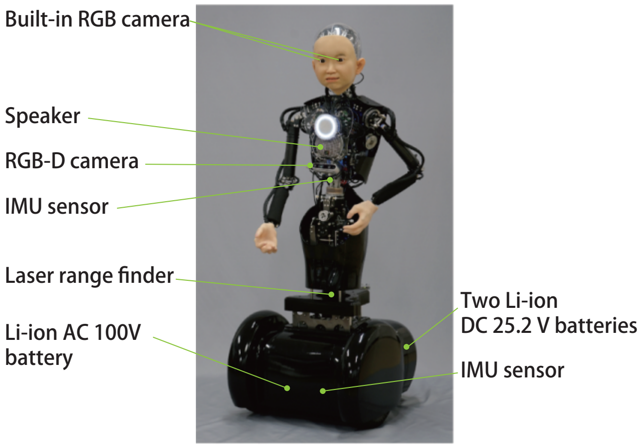


Fig. 3.5 Sensors and other devices embedded into *ibuki*.

3.3 Mobility unit

Figure 3.6 shows a three-dimensional computer-aided design (3D CAD) image of the mobility unit composed of a wheel unit and a vertical oscillation mechanism (VOM). The wheel unit comprises two driving wheels at the front and two omnidirectional wheels as auxiliary wheels at the rear. The rotation angle of each driving wheel is measured by a magnetic rotary encoder (AEAT-6012, Broadcom). Figure 3.7 shows a 3D CAD image of VOM. The vertical oscillation is generated by a linear actuator using a motor-driven slide screw. The displacement is measured by means of a linear potentiometer (LP-150FJ-1K, Midori Precisions). Table 3.4 shows the specifications of the mobility unit.

Table 3.4 Principal Specifications of the Mobility Unit

Description		Value
Outline dimensions	Height (z-axis) [mm]	370
	Width (y-axis) [mm]	441
	Depth (x-axis) [mm]	503
Driving wheel	Radius [mm]	153
	Max torque [Nm]	6.0
VOM	Stroke [mm]	124 (150*)
	Max force [N]	733

* Stroke [mm] without stopper.

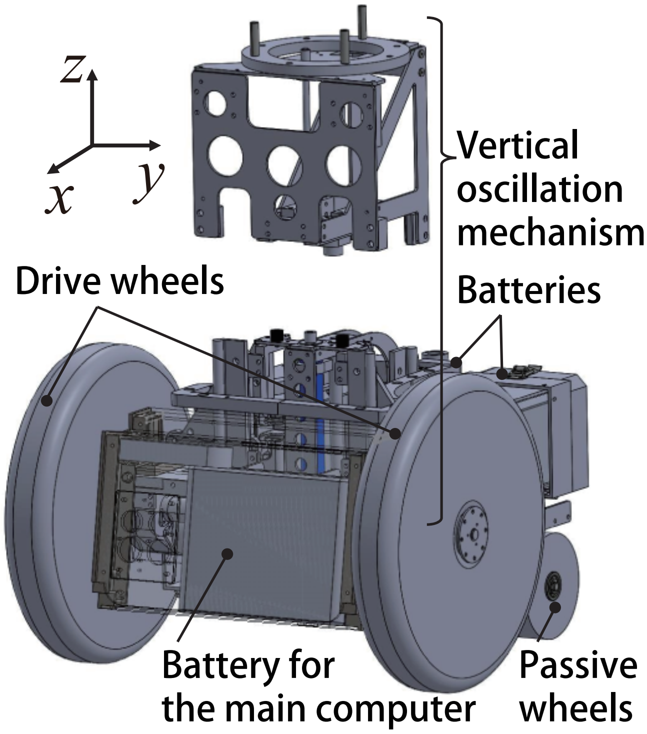


Fig. 3.6 Three-dimensional computer-aided design image of the mobility unit.

3.4 Gait-induced upper body motion

Before we introduce the implementation of gait-induced upper body motion in our mobile humanoid robot using the VOM, we describe the principle of human walking, especially focus on the movement of center of mass.

When humans walk, all body parts move simultaneously except in particular situations. As the CoM is located near the center of the pelvis, its position is affected by the lower limbs. We modeled the trajectory of the human CoM during walking and controlled the position of the mobile humanoid robot's apparent-CoM (aCoM) to achieve human gait-induced upper body motion (Figure 3.8). The aCoM is essentially the same as the root (null) joint used widely in robot and animation motion generation, but adjusted according to the analogy with characteristics of human walking, which is why we use the term aCoM.

One gait cycle starts at the heel strike of a foot and continues until the heel strike of the same foot for the next step. Thus, a human moves two steps forward in one gait cycle. In

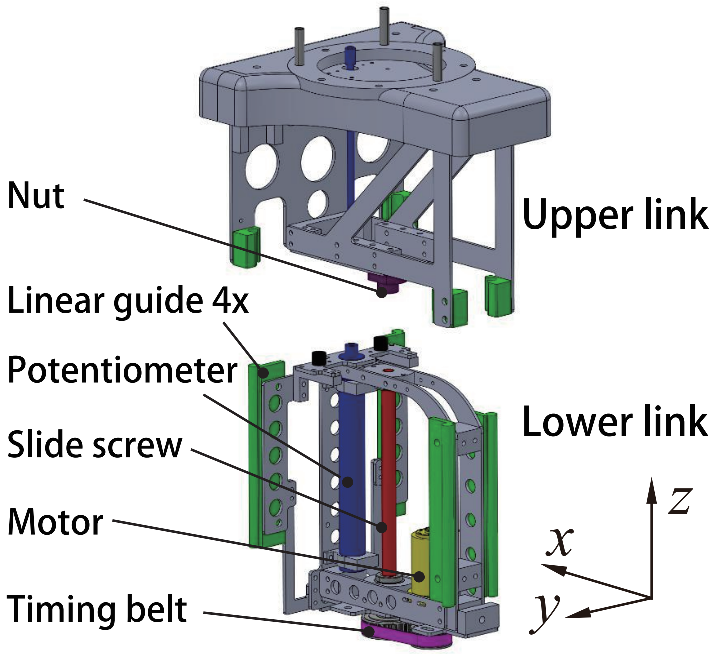


Fig. 3.7 Three-dimensional computer-aided design image of vertical-oscillation mechanism.

each single support phase, the ankle and knee joints of each support leg flex and then extend, and the pelvis rotates slightly. As a result, the human CoM oscillates twice on the sagittal plane in one gait cycle. The trajectory of the CoM oscillation on the sagittal plane can be modeled as a cosine or cycloid curve (Figure 3.8 (A)) [76, 77]. The amplitude is about 40 mm in adult males [78]. Due to the characteristics of the CoM movement, the horizontal moving displacement of the aCoM and the vertical oscillation caused by the gait described above are calculated according to the target gait stride and speed (or frequency). Thus, we can generate the gait-induced upper body motion by controlling wheel angles and VOM displacement based on aCoM (Figure 3.8 (B)). Note that the displacement of the robot's head driven by the VOM is larger than that of an actual human during walking. In the case of human walking, the spine and neck joints of the body compensate for the upper body motion to keep the eye sight at a constant height and angle.

Here, we show our implementation of gait-induced upper body motion. When we approximated the trajectory of the human CoM on the sagittal plane during walking by a cosine wave, the position of the VOM $d(t)$ can be expressed as follows:

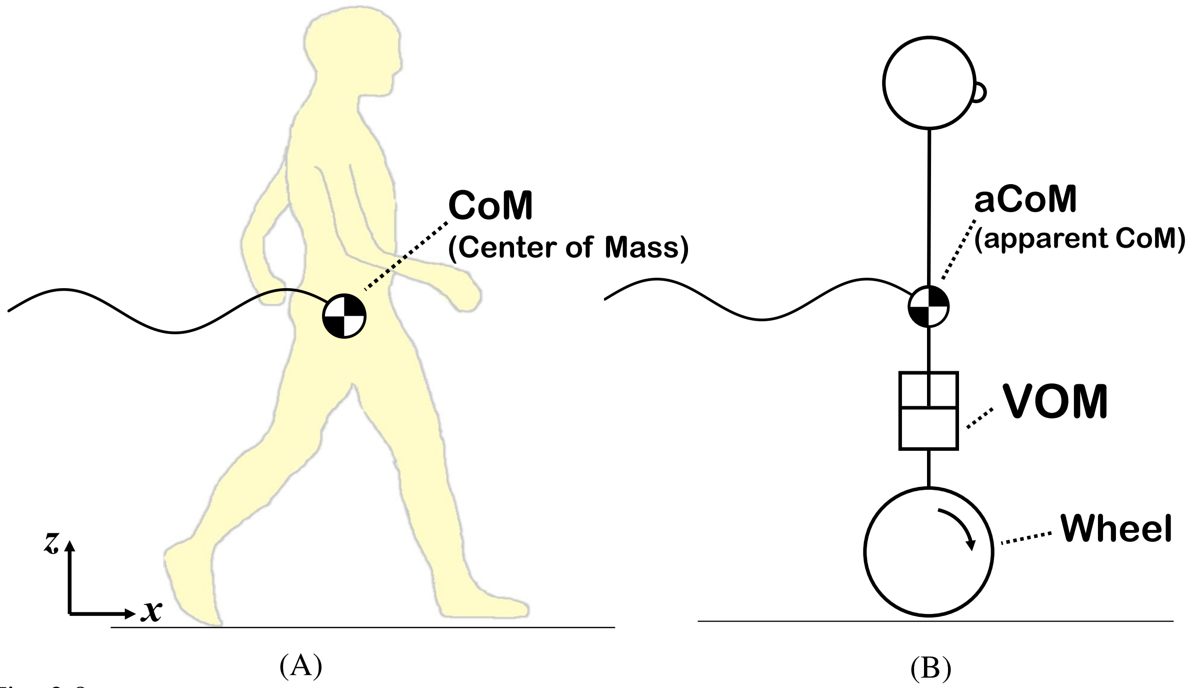


Fig. 3.8 Kinematic relationship between (A) human gait and (B) gait-induced upper body motion of mobile humanoid robot equipped with vertical oscillation mechanism.

$$d(t) = -d_A \cos(4\pi pt) + d_0 \quad (3.4.1)$$

where d_A and d_0 are the amplitude and baseline of the VOM position, respectively. Here, p is the walking pitch of walking; t is time. In a single gait cycle, aCOM oscillates once per step. Therefore, the frequency of VOM is twice as large as that of the gait. The arm and torso motion (joint angle $\theta(t)$) is defined as follows:

$$\theta(t) = \theta_A \sin(2\pi pt + \phi) + \theta_0 \quad (3.4.2)$$

where θ_A and θ_0 are the amplitude and baseline; respectively, ϕ is phase difference.

Figure 3.9 represents the movements of *ibuki*. We used a prototype model of ibuki's mobility unit in this experiment. This mobility unit had two eccentric wheels for realizing motion in a coronal plane, i.e., lateral swinging, by the wheels rotating. If normal wheels were employed, the waist roll joint could be utilized to realize the same lateral swinging. For the joint control, since one wheel rotation corresponded one gait cycle, we calculated pt in Equation 3.4.1 and Equation 3.4.2 base on the measured wheel angle.

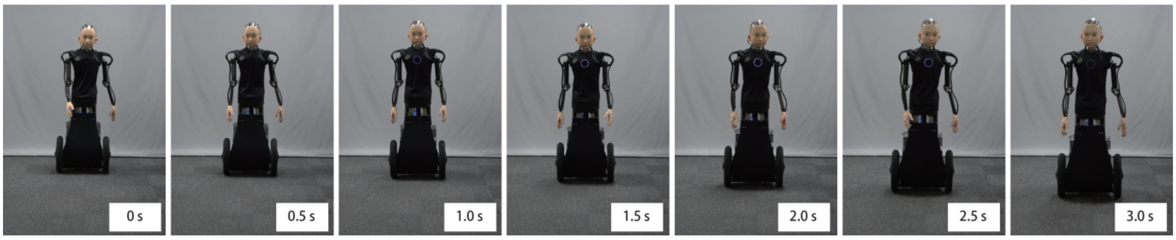
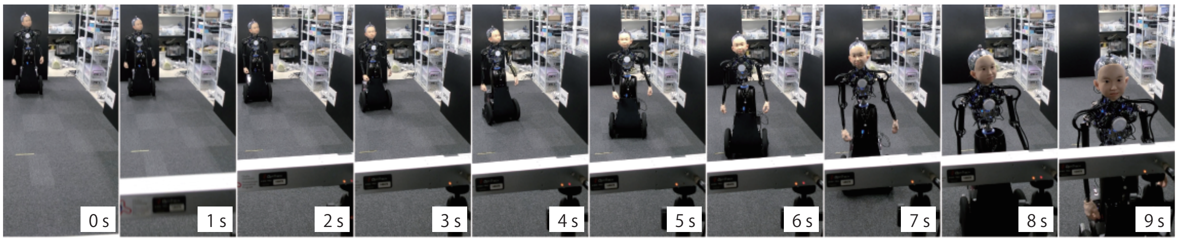
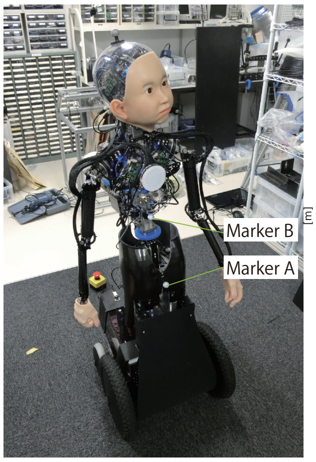


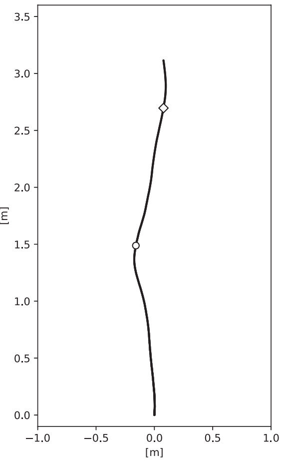
Fig. 3.9 Time series photographs of the gait-induced upper-body motion.

In addition, we measured the trajectories of the upper body and the mobility unit on the coronal plane during movement. Figure 3.10 (a) shows the time series photographs during the movement. A motion capture system (OptiTrack V120: Trio, NaturalPoint, Inc.) was used for this measurement. Figure 3.10 (b) shows the arrangement of two reflection markers (A and B). The marker A was for the mobility unit and the marker B was for the upper body. *ibuki* moved towards the motion capture camera. The trajectory of the movement on the horizontal plane is shown in Figure 3.10 (c). The trajectory of each marker was examined in the interval beginning with a circle and ending with a diamond. Figure 3.10 (d) and Figure 3.10 (e) show the trajectories of the markers A and B projected on the coronal plane, respectively. These trajectories are as seen from the diamond side. The trajectory of marker B is longer horizontally than that of marker A because of the passive swing of the roll axis of the waist during movement.

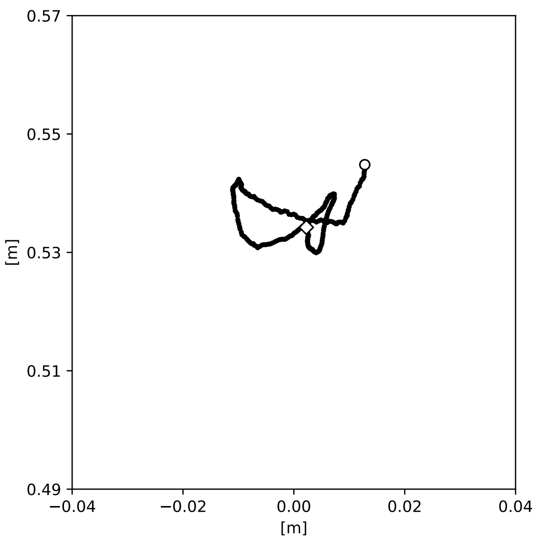


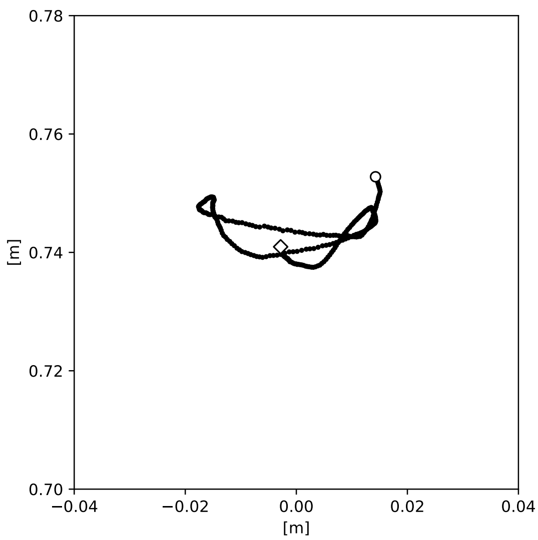
(a)





(b) (c)





(d) (e)

Fig. 3.10 The measurement of the gait-induced upper-body motion: (a) time series photographs of the motion; (b) arrangement of the reflection markers; (c) trajectory of the *ibuki*'s movement on the horizontal plane; (d) trajectory of the marker A; (e) trajectory of the marker B.

Chapter 4

Research1

This chapter shows that an ambiguous facial expression of an android can be perceived more clearly by viewers when body postures and movements are added. Recent research on human behavior reveals that some emotional expressions, such as the emotion “intense”, are difficult to judge as positive or negative by just looking at the facial expression alone. We conducted three experiments and online surveys. In Experiment 1, we validated facial expressions which cannot be clearly discriminated to be either positive or negative, and in Experiment 2 and 3, we validated the possibility of clearly perceiving ambiguous facial expressions of the android by adding postures or movements. In order to investigate only the perception of the peak moment of the android’s emotional expressions, we showed posture photos to participants in Experiment 2. Then, in order to investigate the perception of practical emotional expressions, we showed movement movies to participants in Experiment 3.

4.1 Background

Numerous studies have pointed out that nonverbal communication from robots to humans is as important for robots to communicate smoothly with humans in society as it is in human-to-human communication [79, 80, 81, 82]. Facial expressions and gestures are representative of the nonverbal communication that people often use to convey their emotions [83, 84]. Therefore, it has been actively studied how robots’ facial expressions and gestures can convey emotions to achieve meaningful communication with humans [85, 86, 87]. For example, humanoids’ human-like upper body gestures (nonverbal communication) leads people to perceive a higher animacy of the robot, and the gestures also affect the emotional state and self-disclosure, compared to robot-specific nonverbal behavior (such as LED-eyes color changes) [9]. Also, humanoids' facial expressions lead to an increase in people's desire to interact again with the robot [88]. While there are many findings showing that nonverbal expressions by robots play an important role in human-robot communication, it is also reported that slight differences in expression can lead to the conveyance of significant misinformation [89]. Thus, there is still room for further investigation into the implementation of robotic facial expression, and the human perception of such expressions.

It is widely recognized that facial expressions are important for judging human emotions objectively and clearly. In the 19th century, Darwin already described how facial expressions are associated with certain emotions [90]. Subsequently, Ekman suggested that facial expressions had universality, and classified six types of basic emotions by different facial expressions [91]. Russell explained the correspondence of facial expressions and emotions by using two dimensional affective scales: valence (positive to negative) and arousal (high to low) [92]. All of this research focused on mapping facial expressions to basic emotions, assuming that there are specific facial muscle activation patterns for each emotional expression, and that other people can perceive these emotions by reading these patterns. This method of mapping facial expressions to certain emotions has also been used in research that enables robots to communicate their emotions to humans. Breazeal pioneered research on machines imitating emotional expressions, arguing for the importance of facial expressions and eye gaze for a social robot [93]. Later humanoid robots, such as Kobian, Flobi, Bert2, and iCub, would also use this mapping method, with all four having human-like facial features that allowed for basic emotional expressions [56, 94, 95, 96]. The android robots developed in our research group, such as Geminoid, took this one step further, with a realistic appearance created by using silicon-made skin and an original actuation mechanism for the face [97]. This allowed the android to express emotions through facial expressions to an almost human degree [98].

Recent research has found some cases in which emotions may not be clearly distinguishable from facial expressions [26]. Meeren reported that emotional perception is hindered when facial expression differs from body expression (for example, an angry face, but a frightened body posture or vice versa) [99]. Van den Stock also reported that the perception of emotions from facial expressions is strongly influenced by body posture [100]. Aviezer found that when participants were shown peak expressive reactions to winning and losing points in professional high-stakes tennis matches, it was difficult for participants to correctly judge whether the emotion was positive or negative (i.e. a win or a loss) through intense facial expressions alone, however, they could perceive emotions when these were seen with the body [27]. Aviezer stated that although the faces were inherently ambiguous, viewers erroneously reported perceiving valence in the face and this process seemed to be automatic as participants had little awareness of the actual facial ambiguity and the original diagnostic source of the valence. In addition, later research also showed that intense emotions cannot be judged solely by facial expressions [101].

It was found that humans recognize facial emotions by unconsciously considering factors other than facial expression. In previous research on robots that have emotional expression functions, facial expressions have been used to convey emotions on the assumption that facial expressions have a one-to-one correspondence with certain emotions. Thus, it still has not been investigated whether robots can express the ambiguous facial expressions with no clear valence which humans sometimes show, and what these look like. Furthermore, it has also not been investigated whether the robot's emotional valence can be clearly determined from such expressions by adding body expressions.

What we validate in this chapter are the following: 1) what kind of facial expression of an android is indistinguishable (ambiguous) in the scale of positive to negative emotions?, and 2) with an indistinguishable facial expression, do postures and movements solve the facial ambiguity? We believe that the investigation of ambiguity expressed by the facial expressions of an android and the clear emotional expressions (positive and negative) achieved by adding body modality can help to realize smooth communication between social robots and humans.

4.2 Method

The protocol was approved by the ethics committee for research involving human subjects at the Graduate School of Engineering Science, Osaka University (#R1-6). We recruited separate samples of participants using Amazon Mechanical Turk. The number of participants in each assessment is as follows: 94 people in Experiment 1 (41 females and 53 males, Mean age (M) = 35.5 years old, Standard Deviation (SD) = 10.72), 114 people in Experiment 2 (44 females and 70 males, M = 34.9, SD = 9.93), and 114 people in Experiment 3 (53 females and 61 males, M = 35.4, SD = 9.92).

In Experiment 1, participants were asked to answer the following two questions, corresponding to the emotional state consisting of two dimensions: valence and intensity, after viewing *ibuki*'s facial expressions:

1. To rate the robot's facial emotion (negative to positive valence) from a scale of -4 to

4.
2. To rate the intensity of the robot's facial expression from a scale of 1 to 9.

In Experiment 2 and 3, participants were asked to watch a photo or movie that showed the moment *ibuki* reacted to the result of a game and instructed to guess whether *ibuki* had won or lost the game based on its facial expression. In addition to the questions from Experiment 1, two more questions were added (interpretation of the game result and human-likeness):

1. To guess whether they think the robot has won or lost the game.
2. To rate the robot's facial emotion (negative to positive valence) from a scale of -4 to 4.
3. To rate the intensity of the robot's facial expression from a scale of 1 to 9.
4. To rate the human-likeness of the robot's facial expression from a scale of 1 to 9.

The question about human-likeness was added to confirm that the human likeness of *ibuki* did not affect the assessment of the valence, since the postures and movements were manually created by the authors.

4.2.1 Procedure

In Experiment 1, we investigated whether there are ambiguous facial expressions which cannot be determined by the facial expression alone. For facial expressions, we created nine facial expressions, namely: anger, disgust, fear, happiness, sadness, surprise, contempt, neutral, which are representative emotions in facial expression studies, and intense. For facial expressions except intense, we operated actuators with reference to the Emotional Facial Action Coding System, which explains the characteristics of actual human facial movements. For intense, we created the intense facial expression as an expression of high muscle activity that expresses the excitement immediately after a game result was decided. It was created to look like an emotion in which the eyes are closed strongly and the mouth expresses shouting (See Figure 4.1) based on the paper of Aviezer [27].

From those eight expressions except neutral, we selected one facial expression with the highest entropy of the valence in order to use it in Experiment 2 and 3. Since entropy is a physical property that represents a state of disorder of the system, for this research, we defined the indistinguishable (ambiguous) facial expression as having the highest entropy of the facial valence distributions. The entropy S of each facial expression is calculated by Equation 4.2.1:

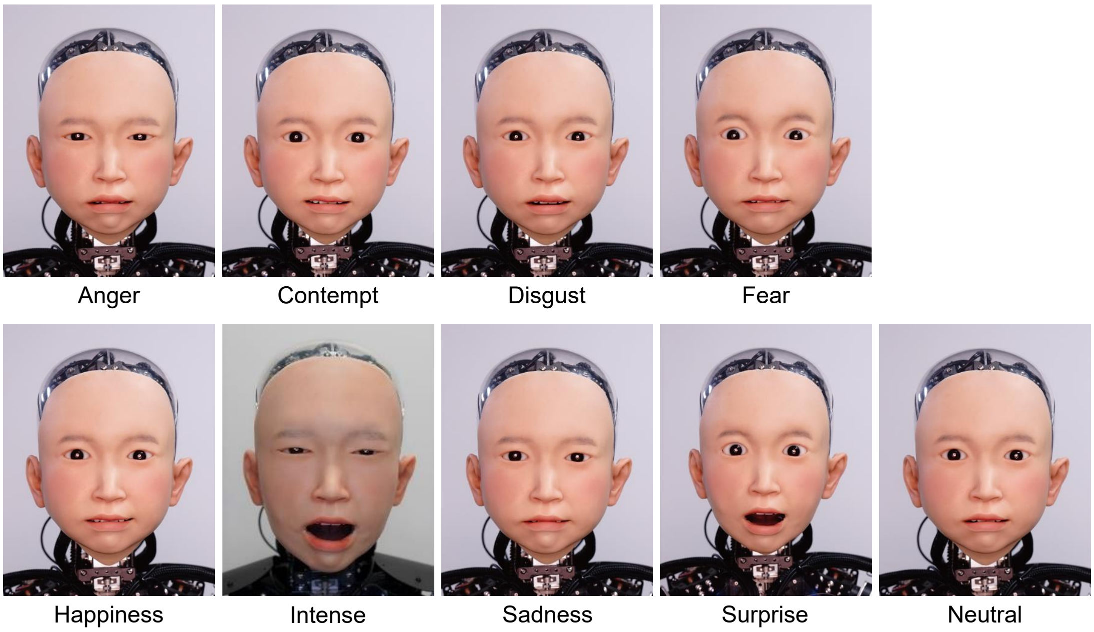


Fig. 4.1 **Nine facial expressions of *ibuki*.**

$$S = -\sum_{x_i=-4}^{4} P(x_i)\log_2 P(x_i) \cdots\cdots (4.2.1)$$

where $x_i = \{-4, -3, \cdots, 3, 4\}$ is the possible facial valence, $P(x_i)$ is the probability which the facial valence is rated as x_i.

Taking the example of intense facial expressions, the probability $P(x_i)$ to be assessed as valence x_i was shown in Table 4.1. In this case, the entropy S of Equation 4.2.1 is calculated as follows:

$$\begin{aligned} S &= -\sum_{x_i=-4}^{4} P(x_i)\log_2 P(x_i) \qquad (4.2.2)\\ &= -0.096\log_2 0.096 - 0.160\log_2 0.160 - 0.149\log_2 0.149\\ &\quad - 0.085\log_2 0.085 - 0.117\log_2 0.117 - 0.170\log_2 0.170\\ &\quad - 0.128\log_2 0.128 - 0.074\log_2 0.074 - 0.021\log_2 0.021\\ &= 3.032 \end{aligned}$$

In Experiment 1, one photo of the eight emotions (350 x 450 pixels size) was displayed beside the neutral face on each page in a random order. After each viewing, participants were

Table 4.1 **The probability of participants' answers for each valence of intense**

x_i	-4	-3	-2	-1	0	1	2	3	4
$P(x_i)$	0.096	0.160	0.149	0.085	0.117	0.170	0.128	0.074	0.021

asked to answer the facial valence and intensity by answering the two questions mentioned in the Materials and Methods section, taking into account that the valence and intensity of the neutral face was 0 and 1, respectively.

Together with the highest entropy facial expression in Experiment 1, we took photos of 15 types of postures by *ibuki* for Experiment 2. In addition, we took movies of 30 types of movements performed by *ibuki*. Firstly, we constructed three arm poses (AP; labeled as A, B, and C) and five head angles (HA; labeled as -43, -17, 0, 17, and 43, which are head inclination degrees at the sagittal plane). HA consist of the neck joint angle θ_1 and waist joint angle θ_2 as shown in Figure 4.2. Table 4.2 shows the angles of the head, neck, and waist.

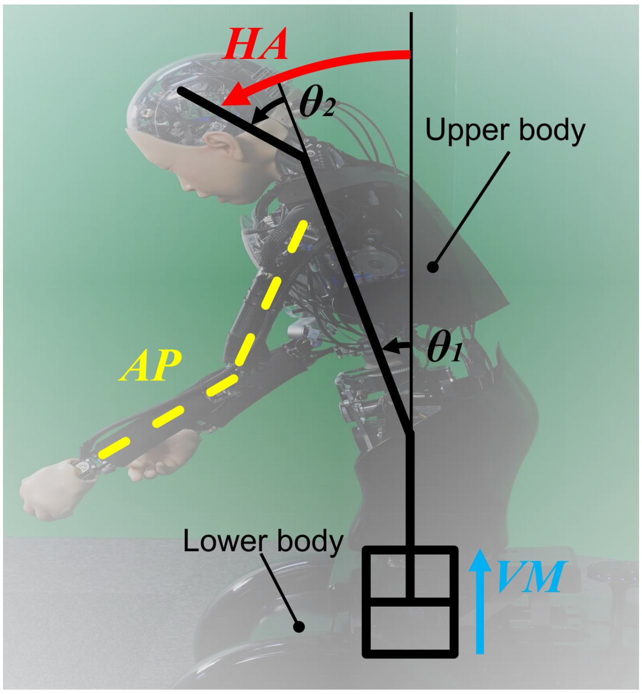


Fig. 4.2 **The body expression with the ambiguous facial expression and the configuration of HA, AP, and VM.**

Secondly, we created 30 movements by combining the previous 15 postures and two types of vertical motions (VM) of the upper body: one was a +40 mm upward motion, the other was a -40 mm downward motion (VM; labeled as 40 and -40). Then, we placed two cam-

Table 4.2 **The relationship among head angle, the neck, and waist joint. All angles are shown in degrees.**

Head angle HA	-43	-17	0	17	43
Neck angle θ_1	-30	-30	0	30	30
Waist angle θ_2	-13	13	0	-13	13

eras – in front of and 40 degrees diagonally to *ibuki* – to take photos and movies of *ibuki*'s emotional expressions. Both cameras were installed at a height of 100 cm – assuming the position of an adult on the knees looking at a child 140 cm away (Figure 4.3). Behind *ibuki* was a neutral green screen. For Experiment 2, we took two photos of each of *ibuki*'s postures from the front and diagonal angles, resulting in 30 photos. Figure 4.4 shows 15 postures from the front view. For Experiment 3, we took two movies of each movement from the front and diagonal angles, resulting in 60 movies. Each movie was edited to a length of three seconds. The first 0.5 seconds of the movie showed *ibuki* in the neutral posture before moving towards the target posture. Again, the same fixed facial expression was used for both Experiment 2 and 3.

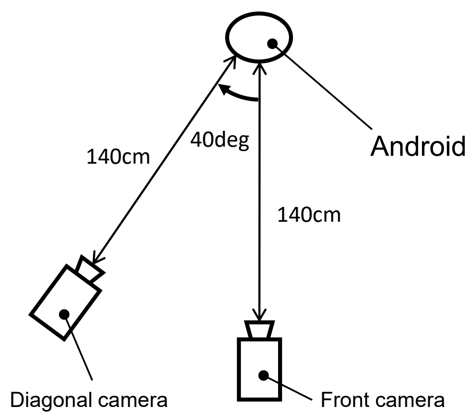


Fig. 4.3 **An overview of shooting environment.**

In Experiment 2 and 3, we investigated whether the indistinguishable facial expression can be distinguished by adding body postures and movements. Participants were asked to look at the photos or movies. To eliminate the influence on the assessment of valence due to viewpoint changes, one participant group (57 people) only saw the front angle photos, and the other participant group (57 people) looked at photos taken from the diagonal angle. In order for participants to understand the contents of the questions, we used a context that evaluate the reaction of the child android at the moment when the win or loss of a game

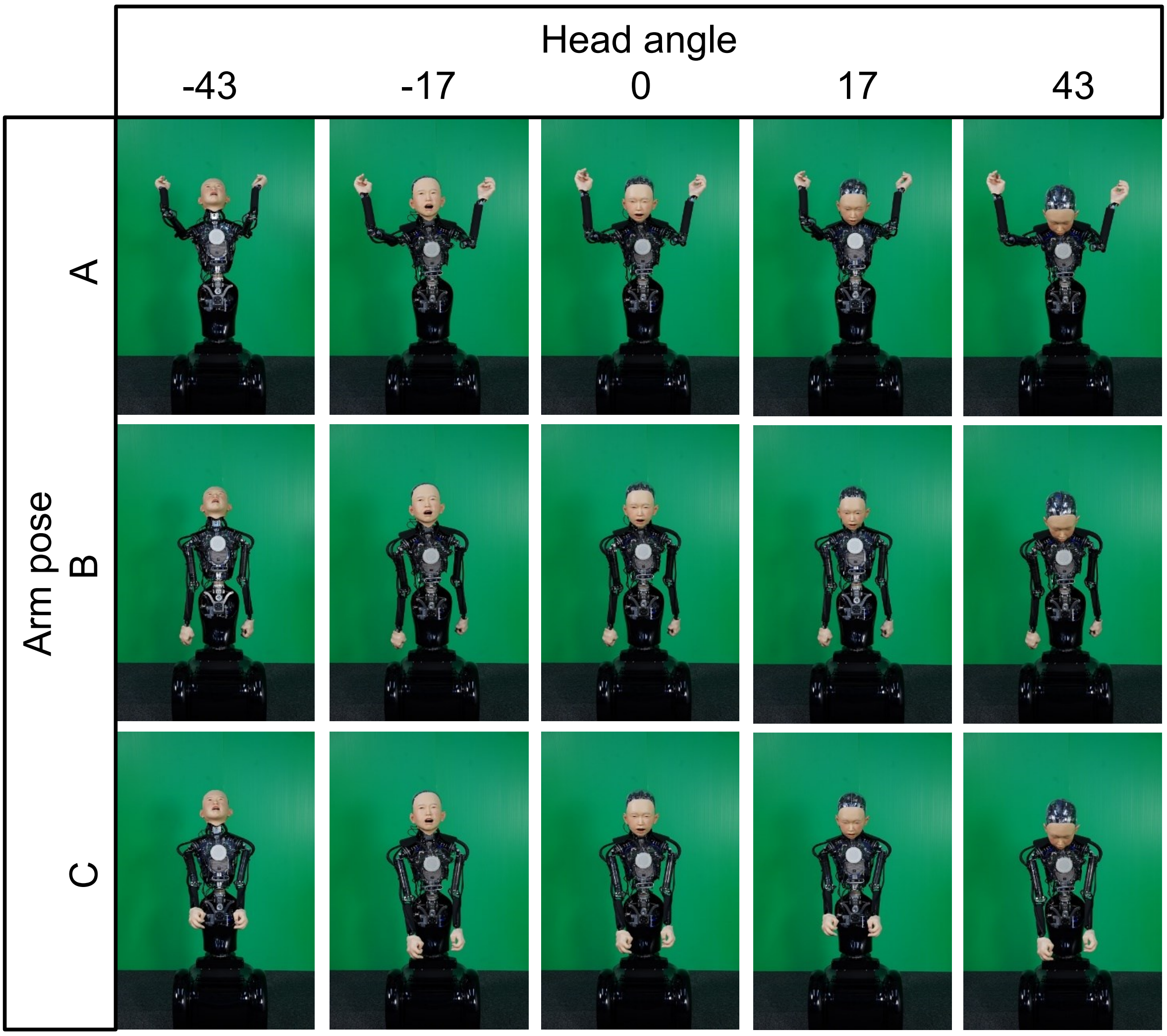


Fig. 4.4 **15 postures from the front view.**

was decided. At the beginning of each assessment, the neutral upright posture of *ibuki* was shown as a reference. For both experiments, a photo or movie was displayed on each page in random order one by one. After each viewing, participants were asked to answer the aforementioned four questions (interpretation of the game result, facial valence, intensity and human-likeness).

For the processing of the data of each experiment, only participants who answered the manipulation check questions correctly were counted as valid answers (this question asked participants to answer with a certain score as was instructed in the movie to check if the participants properly watched the movies). In addition, in Experiment 2, the neutral posture with the neutral face that was shown at first as a reference was shown twice during the task, and participants who rated the valence of the neutral posture as -4 or 4 had their answers excluded (resulting in 10 exclusions). In Experiment 3, the movie of condition HA/AP/VM: -17/A/40 was shown three times during the task, and participants who provided answers with a more than five points difference were excluded (resulting in 23 exclusions).

4.2.2 Analysis

In Experiment 1, distributions of both valence and intensity were analyzed to check the characteristics of *ibuki*'s facial expressions. Then we calculated the entropy for the facial valences of each expression to verify the indistinguishable facial expressions, which have a high entropy.

In Experiment 2, the mean facial valences and entropy were analyzed to check the influence of android body expressions. Two-way repeated measures analyses of variance (ANOVA) were conducted with HA and AP as two within-subjects factors. The significance level was set at 0.05. Partial eta-squared (${\eta_p}^2$) was reported to demonstrate the effect size in ANOVA. Then Tukey's HSD test (HSD test) was performed for multiple comparisons to verify our hypothesis that adding postures or movements contributes to the clear perception of the ambiguous facial expression. In addition, the percentage that assessed as *won* at each posture was also calculated to verify the distinguishable face due to body expressions.

As an android can control each component of the body individually – e.g. facial expression, head angles and arm poses – instead of in conjunction with other joints and other parts of the body like humans, the postures and movements of *ibuki* cannot be guaranteed to be intense and human-like. Therefore, we calculated correlations between facial valence with intensity and human-likeness to confirm whether the effect of intensity and human-likeness on the assessment of facial valence was small or not.

In Experiment 3, we analyzed the data following the same analysis procedure used in Experiment 2.

4.3 Result

4.3.1 Experiment 1

Figure 4.5 shows the distribution of assessed facial valence for eight emotions in Experiment 1. The horizontal axis shows the assessed facial valence (-4 to 4) and the vertical axis shows the normalized number of responses. The dotted line of the histogram shows the mean facial valence. In order to select the most ambiguous face, we calculated the entropy of each distribution, which is shown in the upper left of the histogram. The highest was intense (3.032). With this result, we decided to use intense as the ambiguous facial expression for the next experiments.

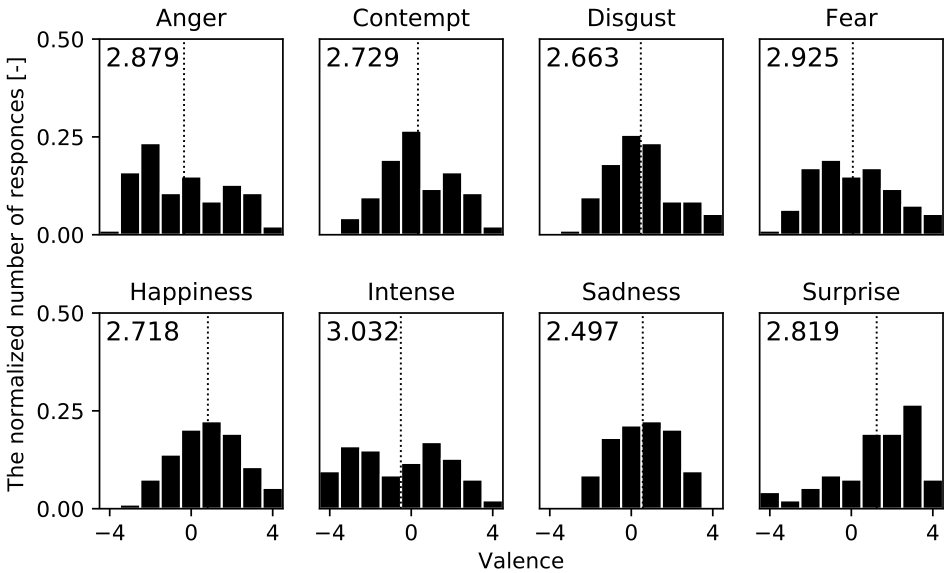


Fig. 4.5 **Distributions of the assessed facial valence in Experiment 1.**
The horizontal axis shows the assessed facial valence (-4 to 4) and the vertical axis shows the normalized number of responses. The dotted line shows the mean facial valence. The entropy is shown in the upper left of each histogram.

Figure 4.6 shows the distribution of assessed facial intensity for the eight emotions in Experiment 1. As in the previous valence result (Figure 4.5), the horizontal axis shows the assessed facial intensity (1 to 9) and the vertical axis shows the normalized number of responses. The dotted line in the histogram shows the mean facial intensity.

4.3.2 Experiment 2

In Experiment 2, *ibuki*'s posture photos were shown to participants in order to investigate only the perception of the peak moment of emotional expressions. We validated that postures contribute to the perception of the indistinguishable facial expression. Figure 4.7 shows the distribution of assessed facial valence in Experiment 2. The horizontal axis shows the

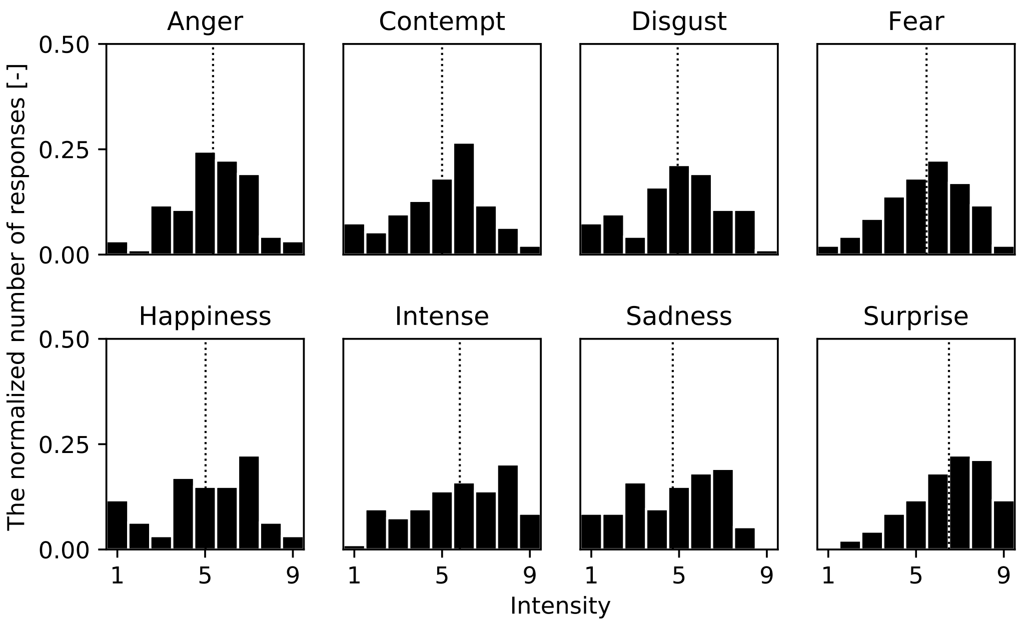


Fig. 4.6 **Distributions of the assessed facial intensity in Experiment 1.**
The horizontal axis shows the assessed facial intensity (1 to 9) and the vertical axis shows the normalized number of responses. The dotted line shows the mean facial intensity.

facial valence as assessed by participants and the vertical axis shows the normalized number of answers for each posture. The red line histogram shows the result of the intense facial expressions in Experiment 1 as a reference. We calculated the entropy of each distribution, which is shown in the upper left of the histogram. The highest was -43/B and the lowest was 0/A.

In order to investigate whether the mean values difference of facial valences were affected by HA and AP, we ran a two-way repeated measures ANOVA with HA and AP. For the assessment of facial valence, there were significant main effects of HA (F (4, 113) = 23.86, p < 0.001, $\eta_p{}^2$ = 0.034) and AP (F (2, 113) = 86.16, p < 0.001, $\eta_p{}^2$ = 0.202) and there was also a significant interaction effect of the HA/AP (F (8, 113) = 9.89, p < 0.001, $\eta_p{}^2$ = 0.055). Then, we ran HSD test to verify the difference of the mean values for the 15 postures. Table 4.3 shows the mean differences for each posture. The highest valence (most positively perceived) posture was HA: -17 degrees and AP: A (Figure 4.8a, Mean value = 1.89, Standard Error = 0.19), while the lowest valence (most negatively perceived) posture

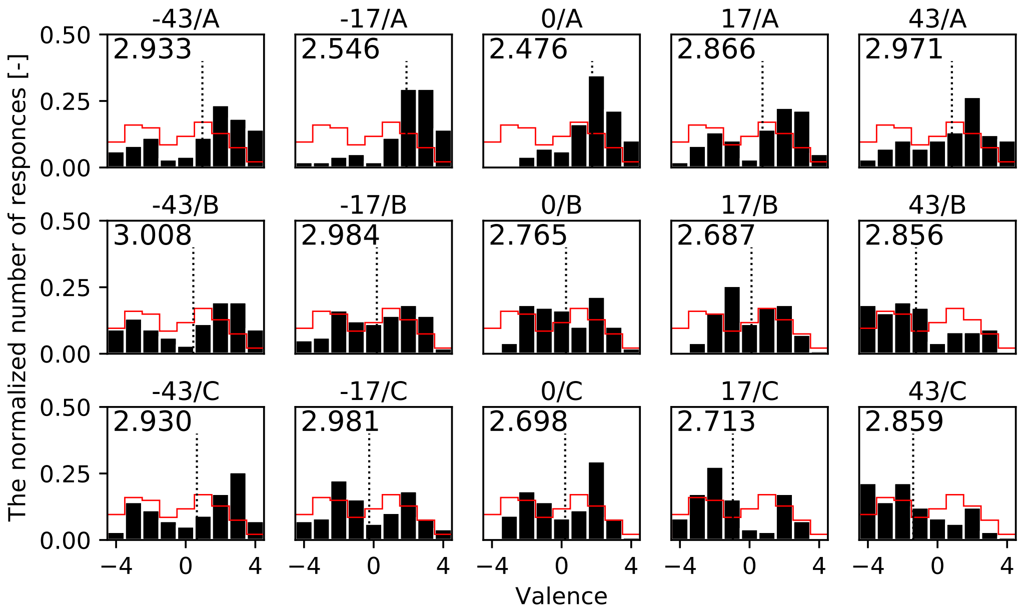


Fig. 4.7 **Distributions of the assessed facial valence in Experiment 2.** The horizontal axis shows the assessed facial valence (-4 to 4) and the vertical axis shows the normalized number of responses. The dotted line shows the mean facial valence. The red line on the histogram shows the result of the intense facial expressions in Experiment 1 as a reference. The entropy is shown in the upper left of each histogram.

was HA: 43 degrees and AP: C (Figure 4.8b, M = -1.40, SE = 0.22), with a significant difference of 3.29 between these two postures ($p < 0.001$).

Table 4.4 shows the percentages of people that assessed each posture as *won* in Experiment 2. AP: A was assessed as *won* between 54.1 % and 83.7 %, in contrast AP: B and AP: C were assessed as lost (lower than 50 %). The result is also consistent with the assessment tendency of the facial valence.

Looking at HA: -17, the maximum value (83.7%) was under AP: A, on the contrary, the same HA: -17 was also the minimum value (27.6%) under AP: C. Interestingly, there was a 56.1 points gap between AP: A and AP: C even with the same HA.

Figure 4.9 and Figure 4.10 show the distribution of facial intensity and human-likeness in Experiment 2. The total average of intensity was 6.14 with SE = 0.04 and human-likeness was 6.43 with SE = 0.04. Furthermore, the facial valence had a low correlation with intensity ($r = 0.163$, $p < 0.001$) and with human-likeness ($r = 0.138$, $p < 0.001$).

Table 4.3 **The mean difference of facial valence among 15 postures in Experiment 2.**

	-43/A	-43/B	-43/C	-17/A	-17/B	-17/C	0/A	0/B	0/C	17/A	17/B	17/C	43/A	43/B	43/C
-43/A		0.52	0.32	-0.93	0.78	1.21**	-0.80	0.70	0.76	0.21	0.85	1.92***	0.13	2.19***	2.36***
-43/B			-0.20	-1.45***	0.26	0.69	-1.32**	0.18	0.23	-0.31	0.33	1.40***	-0.39	1.67***	1.84***
-43/C				-1.24**	0.46	0.90	-1.11*	0.39	0.44	-0.10	0.53	1.60***	-0.18	1.88***	2.04***
-17/A					1.70***	2.14***	0.13	1.63***	1.68***	1.14*	1.78***	2.85***	1.06*	3.12***	3.29***
-17/B						0.44	-1.57***	-0.07	-0.02	-0.56	0.07	1.14*	-0.64	1.42***	1.58***
-17/C							-2.01***	-0.51	-0.46	-1.00	-0.37	0.70	-1.08*	0.98	1.14*
0/A								1.50***	1.55***	1.01	1.64***	2.71***	0.93	2.99***	3.15***
0/B									0.05	-0.49	0.14	1.21**	-0.57	1.49***	1.65***
0/C										-0.54	0.09	1.16*	-0.62	1.44***	1.60***
17/A											0.63	1.70***	-0.08	1.98***	2.14***
17/B												1.07*	-0.71	1.35**	1.51***
17/C													-1.79***	0.28	0.44
43/A														2.06***	2.22***
43/B															0.16
43/C															

Each label shows HA/AP (e.g. -43/A means HA: -43 and AP: A). For example, the top left mean difference 0.52 is calculated as -43/A - -43/B. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

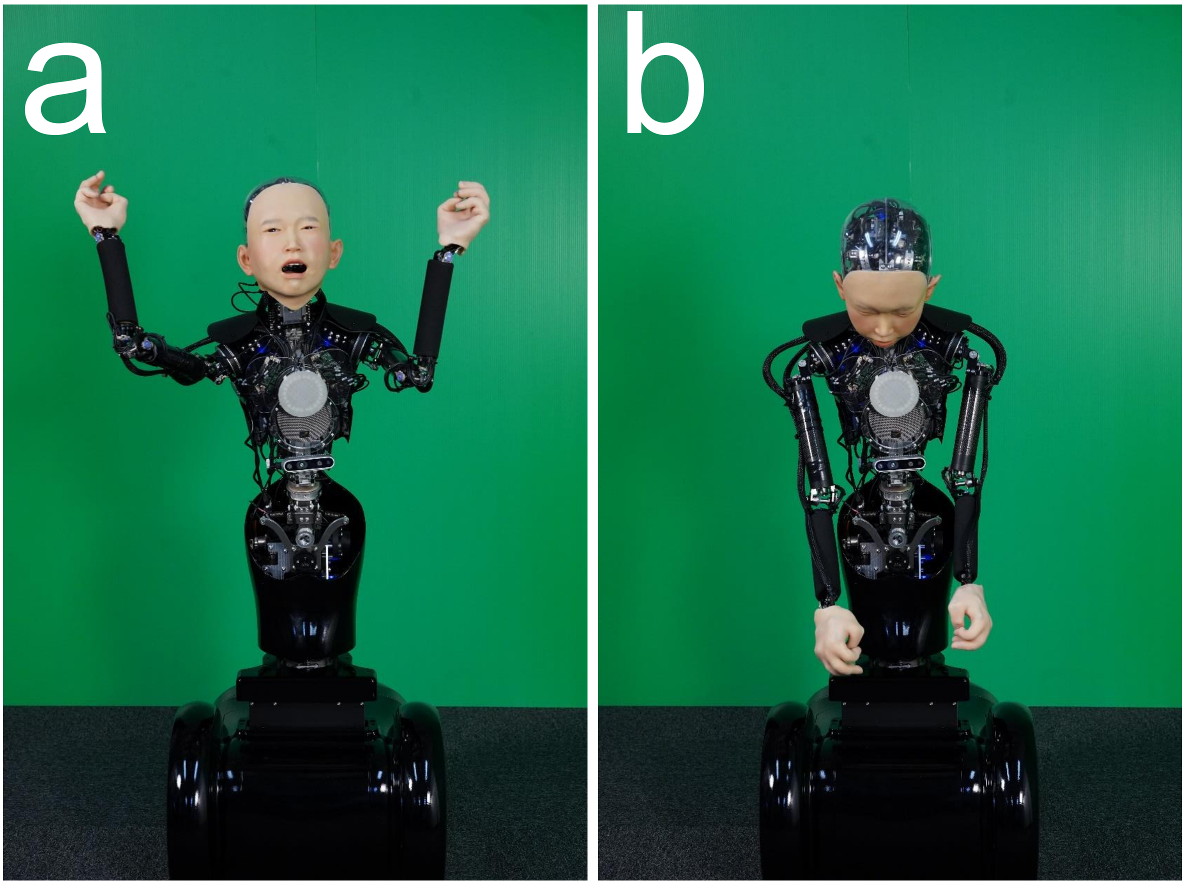


Fig. 4.8 **Two representative postures which have distinguishable facial expressions by adding body expressions.**
(a) The posture -17/A at the highest facial valence / (b) The posture 43/C at the lowest facial valence.

Table 4.4 **Percentage of participants that judged each posture as won a game in Experiment 2.**

		HA				
		-43	-17	0	17	43
AP	A	61.2	83.7	82.7	54.1	55.1
	B	50.0	36.7	35.7	25.5	5.1
	C	49.0	27.6	42.9	6.1	7.1

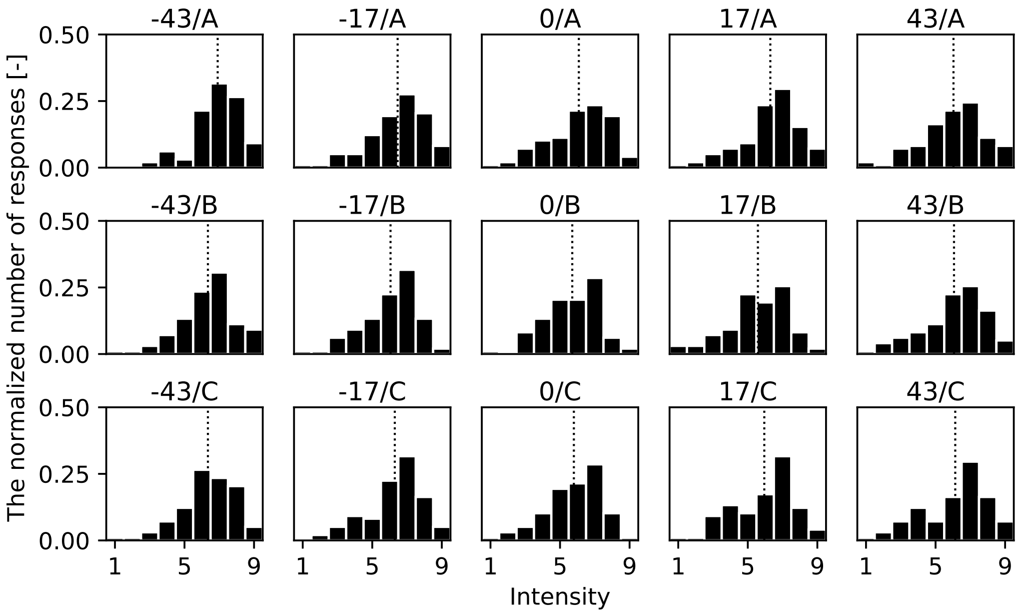


Fig. 4.9 **Distributions of the assessed facial intensity in Experiment 2.**
The horizontal axis shows the assessed facial intensity (1 to 9) and the vertical axis shows the normalized number of responses. The dotted line shows the mean facial intensity.

4.3.3 Experiment 3

In Experiment 3, *ibuki*'s movement movies were shown to participants in order to investigate the perception of emotional expressions. We validated that movements contribute to the perception of the indistinguishable facial expression. Figure 4.11 shows the distribution of assessed facial valence for the condition VM: 40 and -40 mm in Experiment 3. The horizontal axis shows the facial valence as assessed by participants and the vertical axis shows the normalized number of answers for each movement. The red line histogram shows the result of intense facial expression in Experiment 1 as a reference. We calculated the entropy of each distribution, which is shown in the upper left of the histogram. The highest was -43/C/-40 and the lowest was 43/A/40.

In order to investigate whether the mean values difference of facial valence were affected by HA, AP, and VM, we ran a three-way repeated measures ANOVA with HA, AP, and VM. For the assessment of facial valence, there were significant main effects for all three: HA (*F*

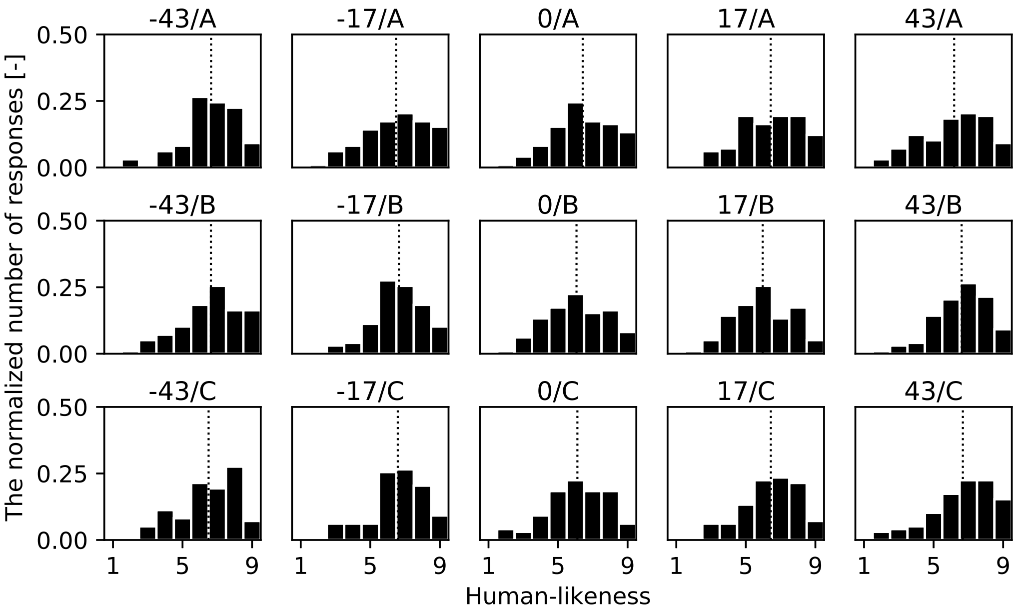
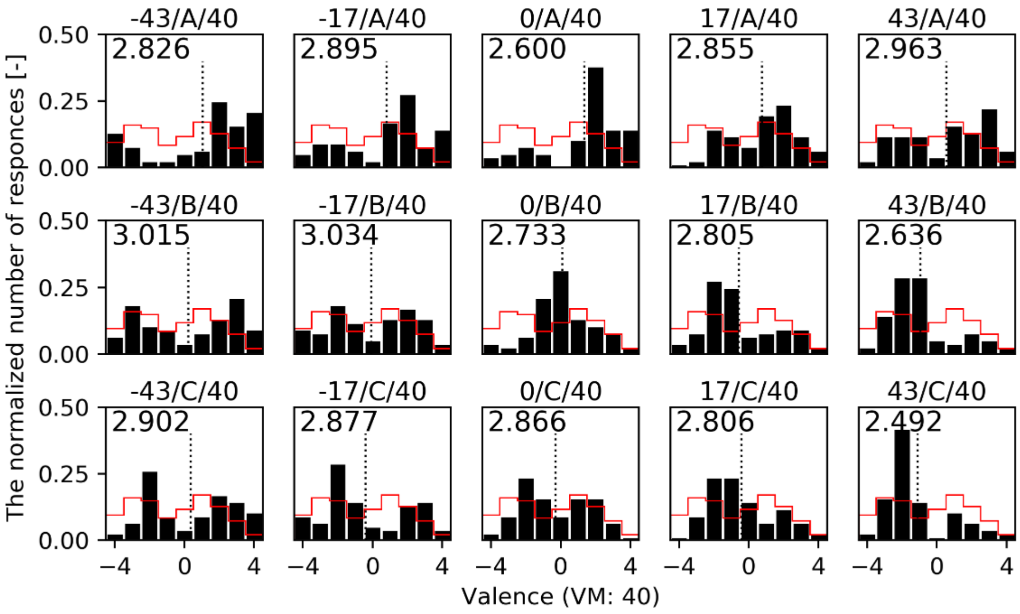


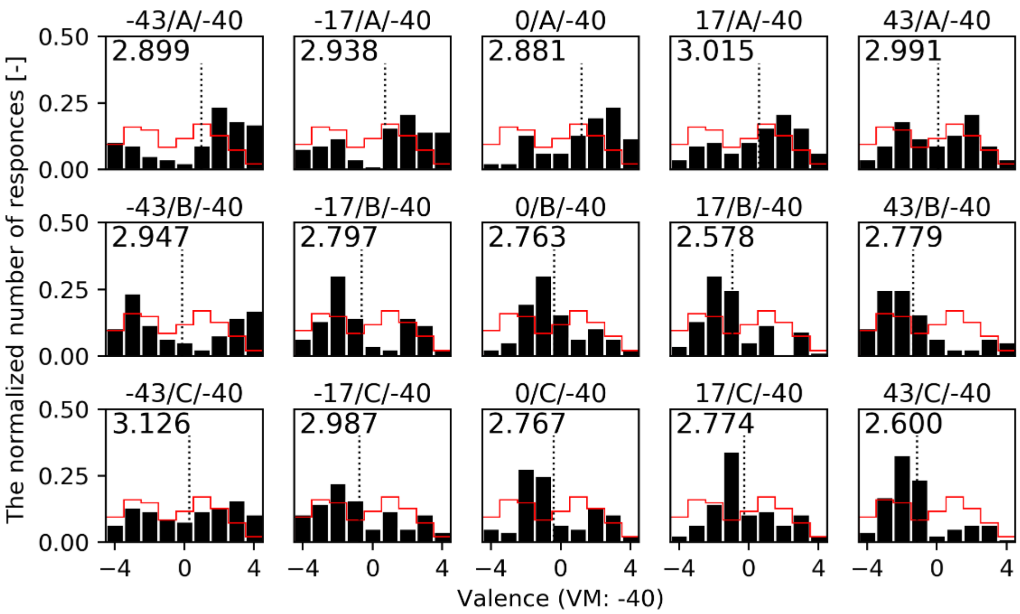
Fig. 4.10 **Distributions of the assessed facial human-likeness in Experiment 2.** The horizontal axis shows the assessed facial human-likeness (1 to 9) and the vertical axis shows the normalized number of responses. The dotted line shows the mean facial human-likeness.

(4, 113) = 11.68, $p < 0.001$, $\eta_p{}^2 = 0.020$), AP (F (2, 113) = 52.82, $p < 0.001$, $\eta_p{}^2 = 0.044$), and VM (F (1, 113) = 17.35, $p < 0.001$, $\eta_p{}^2 = 0.007$). Also there were significant interaction effects of the HA/AP (F (8, 113) = 3.29, $p = 0.001$, $\eta_p{}^2 = 0.011$) and the AP/VM (F (2, 113) = 3.99, $p = 0.021$, $\eta_p{}^2 = 0.003$). On the other hand, no significance was found in the HA/VM (F (4, 113) = 0.48, $p = 0.748$, $\eta_p{}^2 = 0.001$) or the HA/AP/VM (F (8, 113) = 0.57, $p = 0.806$, $\eta_p{}^2 = 0.002$).

Thus, we ran Tukey's HSD tests to verify the difference of the mean values at each interaction effect of HA/AP and AP/VM. The result of HA/AP is shown in Table 4.5 and AP/VM is shown in Table 4.6. Under the condition of HA and AP, the highest valence movement was HA: 0 degrees and AP: A (M = 1.28, SE = 0.18), while the lowest valence movement was HA: 43 degrees and AP: B (M = -1.14, SE = 0.17), with a significant difference of 2.42 between these two movements ($p < 0.001$). Also under the condition of AP and VM, the highest valence movement was AP: A and VM: 40 mm (M = 0.90, SE = 0.12), while the lowest valence movement was AP: B and VM: -40 mm (M = -0.69, SE = 0.12), with a



(a) condition VM: 40 mm



(b) condition VM: -40 mm

Fig. 4.11 **Distributions of the assessed facial valence in Experiment 3.**
The horizontal axis shows the facial valence as assessed by participants and the vertical axis shows the normalized number of answers at each movement. The dotted line shows the mean facial valence. The red line on the histogram shows the result of intense facial expression in Experiment 1 as a reference. The entropy is shown in the upper left of each histogram.

significant difference of 1.59 between these two movements ($p < 0.001$).

Table 4.7 shows the percentages of people that assessed each movement as *won* in Experiment 3.

As in Experiment 2, different AP leads to the exact opposite judgement for the game result. For example, under HA: 17, there is a 39.5 points gap between A/40 (63.2 %) and B/40 (23.7 %) and a 51.3 points gap between A/-40 (65.8 %) and B/-40 (14.5 %).

Figure 4.12 and Figure 4.13 show the distribution of facial intensity and human-likeness in Experiment 3. The total average of intensity was 6.00 with SE = 0.04 and human-likeness was 6.02 with SE = 0.04. Furthermore, the facial valence had a low correlation with intensity ($r = 0.164$, $p < 0.001$) and with human-likeness ($r = 0.196$, $p < 0.001$).

4.4 Discussion

In the past, research on robots expressing emotions has been based on the assumption that a certain expression is equivalent to a certain emotion (e.g. a smiling indicates happiness). However, there has been no investigation into whether robots can also express ambiguous facial expressions with no clear valence. Additionally, for ambiguous facial expressions, there has been no investigation into whether the addition of expressions such as body expression can make the robot's emotional valence clearer to humans.

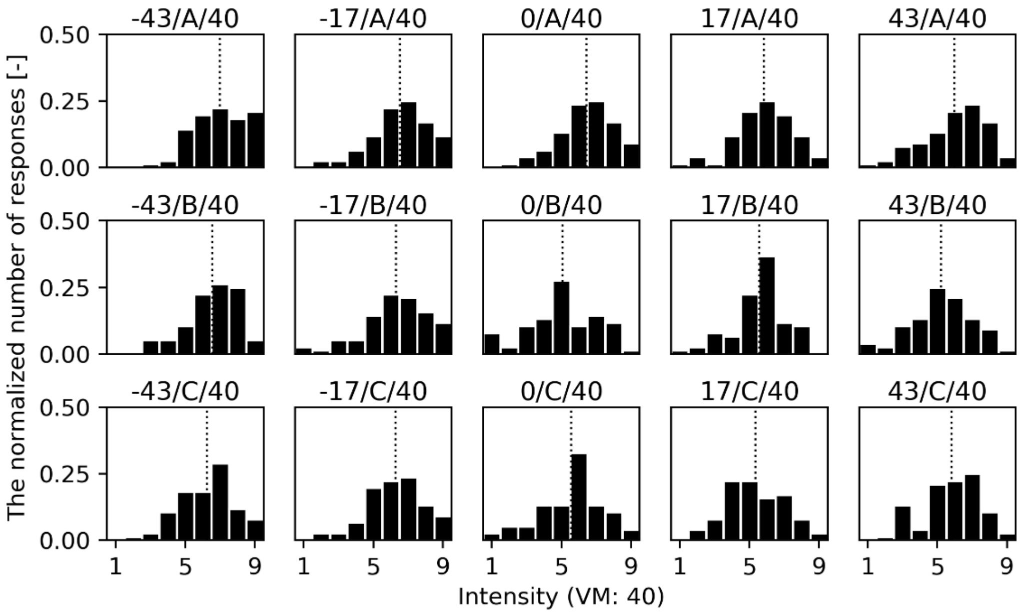
In Experiment 1, we validated that the distributions of valence and intensity were different depending on the facial emotion expressions. In particular, according to the results of the entropy of facial valences, participants could not distinguish the facial valence of the intense emotion as positive or negative. The previous study by Aviezer found that it is difficult to judge the valence of intense facial expressions when only the facial expression can be seen [27]. Our experiment corroborates this phenomenon in the case of androids. The histogram of the facial valence of intense emotion in Figure 4.5 seems to have two peaks in both sides positive (1) and negative (-3). We believe that this is because the participants were mainly divided into two groups who perceived the intense facial expression positively or negatively. Thus, the entropy of intense got the highest value of the eight emotions.

In Experiment 2 and 3, it was found that both body postures and body movements changed the facial valence of the intense emotion and contributed to the lower entropy in the valence assessments. In other words, postures and movements can improve the perception of the ambiguous facial expression. As mentioned in the result, the maximum difference in the mean facial valences was 3.29 for the posture and 2.42 for the movement. Although the number

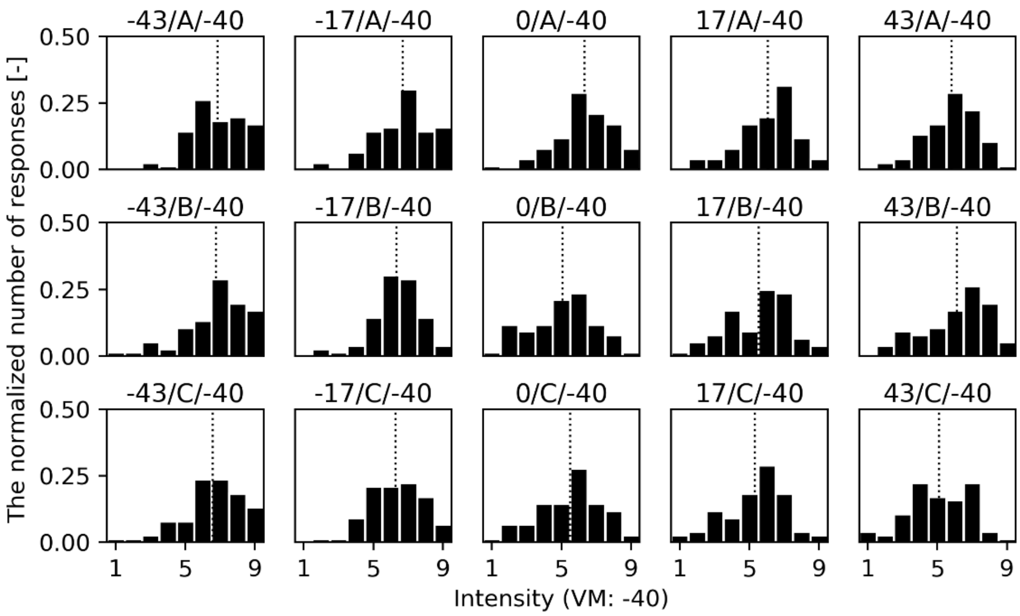
Table 4.5 **The mean difference of facial valence among head angles and arm poses in Experiment 3.**

	-43/A	-43/B	-43/C	-17/A	-17/B	-17/C	0/A	0/B	0/C	17/A	17/B	17/C	43/A	43/B	43/C
-43/A		0.95*	0.67	0.24	1.35***	1.58***	-0.28	1.14**	1.34***	0.32	1.76***	1.35***	0.70	2.14***	2.12***
-43/B			-0.28	-0.72	0.39	0.62	-1.23***	0.18	0.39	-0.63	0.80	0.39	-0.26	1.19***	1.16***
-43/C				-0.43	0.68	0.91*	-0.95*	0.47	0.67	-0.35	1.09**	0.68	0.03	1.47***	1.45***
-17/A					1.11**	1.34***	-0.51	0.90*	1.11**	0.09	1.52***	1.11**	0.46	1.91***	1.88***
-17/B						0.23	-1.62***	-0.21	-0.01	-1.03**	0.41	-0.00	-0.65	0.80	0.77
-17/C							-1.86***	-0.44	-0.24	-1.26***	0.18	-0.23	-0.88*	0.57	0.54
0/A								1.41***	1.62***	0.60	2.03***	1.62***	0.97*	2.42***	2.39***
0/B									0.20	-0.82	0.62	0.21	-0.44	1.01**	0.98*
0/C										-1.02**	0.41	0.01	-0.64	0.80	0.78
17/A											1.43***	1.03**	0.38	1.82***	1.80***
17/B												-0.41	-1.06**	0.39	0.36
17/C													-0.65	0.80	0.77
43/A														1.45***	1.42***
43/B															-0.03
43/C															

Each label shows HA/AP. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

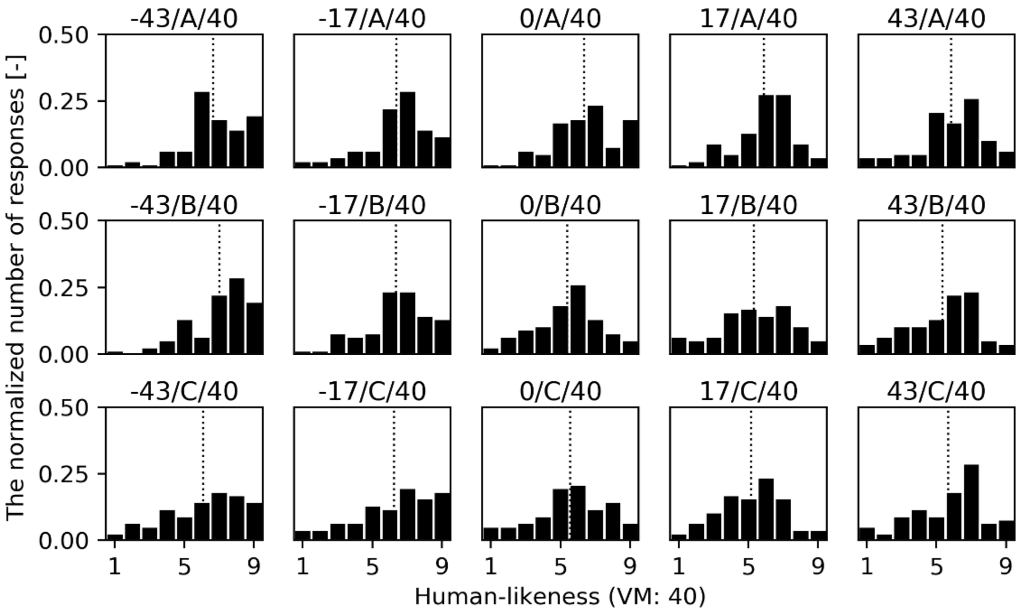


(a) condition VM: 40 mm

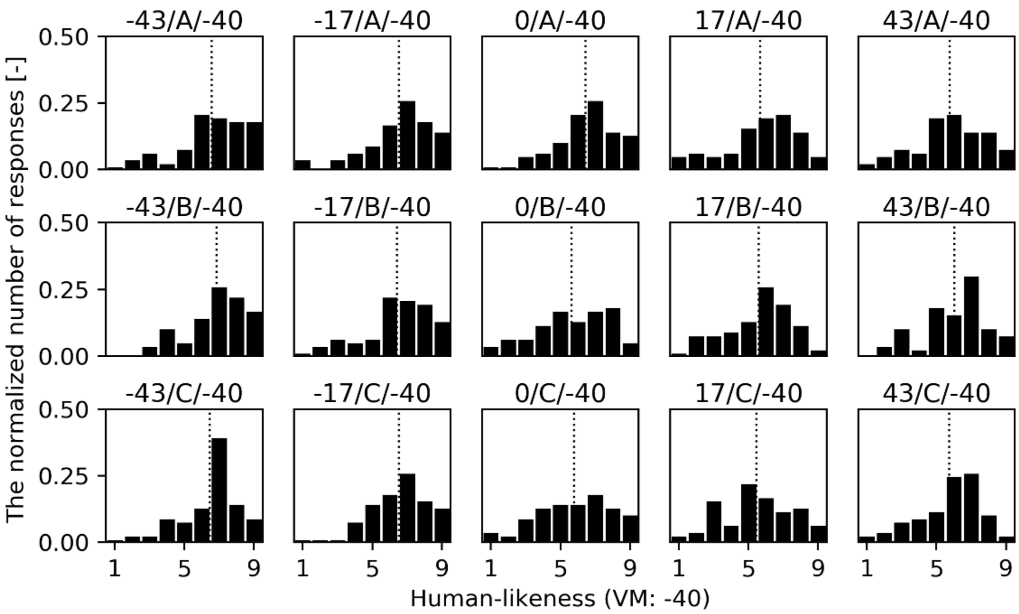


(b) condition VM: -40 mm

Fig. 4.12 **Distributions of the assessed facial intensity in Experiment 3.**
The horizontal axis shows the facial intensity as assessed by participants and the vertical axis shows the normalized number of answers at each movement. The dotted line shows the mean facial intensity.



(a) condition VM: 40 mm



(b) condition VM: -40 mm

Fig. 4.13 **Distributions of the assessed facial human-likeness in Experiment 3.**
The horizontal axis shows the facial human-likeness as assessed by participants and the vertical axis shows the normalized number of answers at each movement. The dotted line shows the mean facial human-likeness.

Table 4.6
The mean difference of facial valence among arm poses and vertical motions in Experiment 3.

	A/40	A/-40	B/40	B/-40	C/40	C/-40
A/40		0.19	1.14^{***}	1.59^{***}	1.27^{***}	1.35^{***}
A/-40			0.96^{***}	1.40^{***}	1.08^{***}	1.16^{***}
B/40				0.44	0.12	0.21
B/-40					-0.32	-0.24
C/40						0.08
C/-40						

Each label shows AP/VM. ***: $p < 0.001$.

Table 4.7
Percentage of participants that judged each movement as won a game in Experiment 3.

		HA				
		-43	-17	0	17	43
AP/VM	A/40	69.7	64.5	76.3	63.2	50.0
	B/40	47.4	43.4	46.1	23.7	11.9
	C/40	52.6	28.9	40.8	23.7	18.4
	A/-40	68.4	63.2	75.0	65.8	43.4
	B/-40	39.5	30.3	25.0	14.5	17.1
	C/-40	51.3	31.6	30.3	30.3	13.2

of facial expressions and postures used in the experiment is different, as a reference, those mean differences are as close as the mean difference of 2.3 points in the previous human study by Aviezer in which they exchanged win and lose body postures against intense face in a tennis match. Looking at the histograms of facial emotional values in Figure 4.7 of Experiment 2, under postures which had lower entropy than that of the intense facial expression in Experiment 1, such as -17/A (0.486 lower) and 0/A (0.556 lower), those distributions only had a one side peak. On the other hand, under postures which had little lower entropy, such as -43/B and -17/B, those distributions still had two peaks in both positive and negative sides as in Experiment 1. However, -17/B/-40 (the movement of -17/B with VM: -40) tended to be perceived more negatively due to the effect of the downward vertical motion in Experiment 3. Thus, -17/B/-40 had a further lower entropy than that of Experiment 2.

From the ${\eta_p}^2$ values in ANOVA, we can deduce that participants' assessments were most affected by AP, then HA and, finally, by VM. Our findings are consistent with several previous studies on human emotion expressions. As for AP, when the arms are located vertically upwards or move upwards from the body, the posture tends to be perceived as joy [102, 103] and joy corresponds to positive valence [92]. As for HA, raised head angles are associated with positive emotions such as joy [104]. As for VM, rising upper body motion corresponds to happiness, and sinking body motion corresponds to sadness [105]. However, in this research, we only controlled the displacement and not the speed and acceleration of the movement. It has been found that the human perception of robots is negatively impacted if human-like robots perform robot-like movements which contradict its appearance [106]. This might be the reason why VM comes last. Motion activity, including how fast and smooth a person or android is moving, is said to be one of the important factors for emotion recognition [107].

Since interaction effects were confirmed between AP and HA, and AP and VM from the results of ANOVA, we argue that robot developers should not separately design facial expressions from body expressions when designing robots for the application of human-robot interaction, as there are expressions that cannot be distinguished by the facial expressions alone, such as the intense emotion discussed in this chapter. Strong emotions influence human social decision making [108], enhancement on memory [109], and time perception [110]. In the same way, strong emotion expressions by robots are expected to enhance the relationship between humans and robots [111].

One limitation was that *ibuki* was based on a young Asian male child. In past studies on recognition of human emotions, it has been reported that age, gender, and ethnicity can influence emotional perception [112, 113]. Therefore, it is necessary to further investigate how our finding applies to other common humanoids and robots universally. Previous research on humans shows that humans can perceive emotions even from point-light displayed facial expressions or body movements [83, 114]. We expect that this capability is also applicable for the emotion perception of robot faces.

Chapter 5

Research2

This chapter shows *ibuki*'s three types of gait-induced upper body motions to express anger, happiness, and sadness. Our hypothesis is that a gait-induced upper body motion with vertical oscillation enhances human perception of a mobile robot's emotional expressions. Thus, we confirmed whether the presence of *ibuki*' s motion with vertical oscillation enhances the participants' perception in Experiment 1 and whether the congruence of *ibuki*' s motion with vertical oscillation enhances the participants' perception in Experiment 2.

5.1 Background

Humans express their intentions and emotions with facial expressions and gestures [27], even while walking with various gaits (a person's manner of walking). These emotional expressions are thought to contribute significantly to building interpersonal relationships and our interactions with others [115]. Some researchers have previously investigated human intentions expressed through human gait. Frohnwieser showed that people gave different allowances of personal space according to differences in walking behavior [11]. Park found that human body posture indicated their destination and direction [13].

Human gaits are emotionally expressive, and people may walk in various ways depending on their emotions. People can perceive the emotions from others through their gait. For example, when you see your friend walking with a perceived happy gait, you might ask, "Did something good happen?". When you perceive a friend's step as mournful, you might put a comforting hand on her or his shoulder. Many studies have been conducted to investigate the characteristics of emotional gaits [20, 21, 24, 28]. Karg reported that various parts of the upper body, such as the arms and head, express internal states during walking [116].

Michalak measured the motions of humans in sad and happy moods, and found that there is a significant difference in vertical upper body movement when expressing these emotions [22].

For human-robot interaction (HRI) in real-world environments, considering the expression of various gaits, including *emotional* gaits, is valuable to build deep human-robot relationships. In the field of bipedal robots, some research has been conducted to realize human-like gaits [29] and further emotional gaits [30, 31]. Destephe created emotional gait, which expressed happiness, sadness, and neutral emotion, by using the bipedal humanoid Wabian-2R, and investigated the human perceptions and psychological influences by using a questionnaire survey [30]. Izui reported the characteristics of emotional expression based on the human walking motion. They implemented those characteristics on the small humanoid Nao as emotional gaits and investigated the recognition rates of the person who saw gait motions expressing emotions [31].

Since wheel-based motion has moving efficiency and stability on flat ground, it has been actively used for robots operating in real-world environments. However, while some of these mobile robots have a human-like upper body, in most cases, their human-like movements, such as gestures, are realized without wheel-based motion [51, 58, 117].

In research of a robot's expression with wheel-based motion, Nakata investigated a mobile robot's joint angle and velocity parameters based on the Laban theory. They theorized that those parameters are useful in allowing a wheeled robot to express emotions [118]. Granados has developed a wheeled dance-teaching robot that has a waist joint that uses a parallel-link mechanism to allow dancing with a human partner. Furthermore, other studies have examined the emotional expression by wheeled humanoid robots [61]. Zecca created expressions of emotions based on human data by using a wheeled humanoid Habian. They reported recognition rates for emotional body expressions performed by the standing robot [57]. Tsiourti created emotion expressions through five modalities (face, head, body, voice, locomotion) by using wheeled humanoids and confirmed that even a simple locomotion modality can convey a robot's emotion [67].

However, to our best knowledge, the human perception of emotion expressions of wheeled mobile humanoid robots while its moving has not been studied yet. If the robot could express its intentions and emotions while moving, humans could act appropriately according to that information. As a result, robots and humans could achieve orderly interactions by being able to predict and understand each other's behavior in various situations.

5.2 Method

In this section, we describe the implementation of gait-induced upper body motion in our mobile humanoid robot using a vertical-oscillation-mechanism (VOM). The movement of this mechanism was based on human center-of-mass (CoM) motion. We also describe how we applied human-motion data to enable *ibuki* to move with emotional expressions.

5.2.1 Motion generation

We chose anger, happiness, and sadness as they are some of the basic human emotional expressions. The purpose of this research is to confirm whether a wheeled android can express emotions by the gait-induced upper body motions while moving. Therefore, we decided to implement emotions that people can express while walking. A previous study has investigated four emotions (anger, happiness, sadness, and fear) and they concluded the perception of fear depends on a specified context [20]. Thus, we decided to exclude fear and used three expressions of emotions: anger, happiness, and sadness in this research.

The author filmed himself walking a 5 m distance on a flat floor in our laboratory while expressing these emotions and neutral emotion as a reference. For these expressions, we used [20] as a reference to determine inclination angles of head and torso and walking speeds of the human body. Walking motions were filmed on the sagittal plane of the body for all three emotions, with a camera fixed at 0.9 m in height and 2.6 m away from the middle of the route (filming was repeated three times per emotion).

Joint position time series data was obtained by human pose recognition software Openpose [119]. The sampling rate was 30 Hz and all the data were filtered with a 3 Hz low-pass filter. Reference joint angle values for two gait cycles of the neck, right shoulder, right elbow, and waist in the pitch axis of *ibuki* were calculated from this data. The reference joint angle of the left shoulder $\theta_L(t)$ was calculated by the following equation $\theta_L(t) = 2\bar{\theta} - \theta_R(t)$, where $\theta_R(t)$ is the joint angle of the right shoulder and $\bar{\theta}$ is the average angle at all times. The reference joint angle of the right elbow was calculated in the same way. In order to achieve the aCoM trajectory, we obtained the human CoM height $z(t)$ on the waist position. We determined that *ibuki*'s aCoM located at the center of the waist (545 mm height) like in humans. In this case, we could control the height by using the VOM joint displacement $d(t)$. The displacement of VOM $d(t)$ was reduced to 0.71 times by taking the ratio of the height of human and *ibuki*. Due to the limitation of the joint range (maximum was 40 and minimum was -40 mm), we adjusted the average height of a CoM to be $d(t) = 0$. Thus, we calculated the VOM displacement by the following equation: $d(t) = 0.71(z(t) - \bar{z})$.

Figure 5.1 shows comparisons of reference joint values and measured values of three gait-induced upper body motions for each emotion. The upper body is tilted forward for anger and sadness, while in happiness, it is tilted backward. The upper body and neck are tilted forward with a comparatively large angle for sadness. As reported by [22], amplitudes for anger and happiness are larger than that of sadness. The error between the reference and the measured value in Figure 5.1 were caused by factors such as the joint friction, dampers installed to reduce unnatural vibrations on the upper body, and the delay due to P-D control for joint angles. Although there was a delay in the entire orbit, we concluded that the difference in waveform between the reference and the measured value was small. Therefore, we conducted the experiments using those motions.

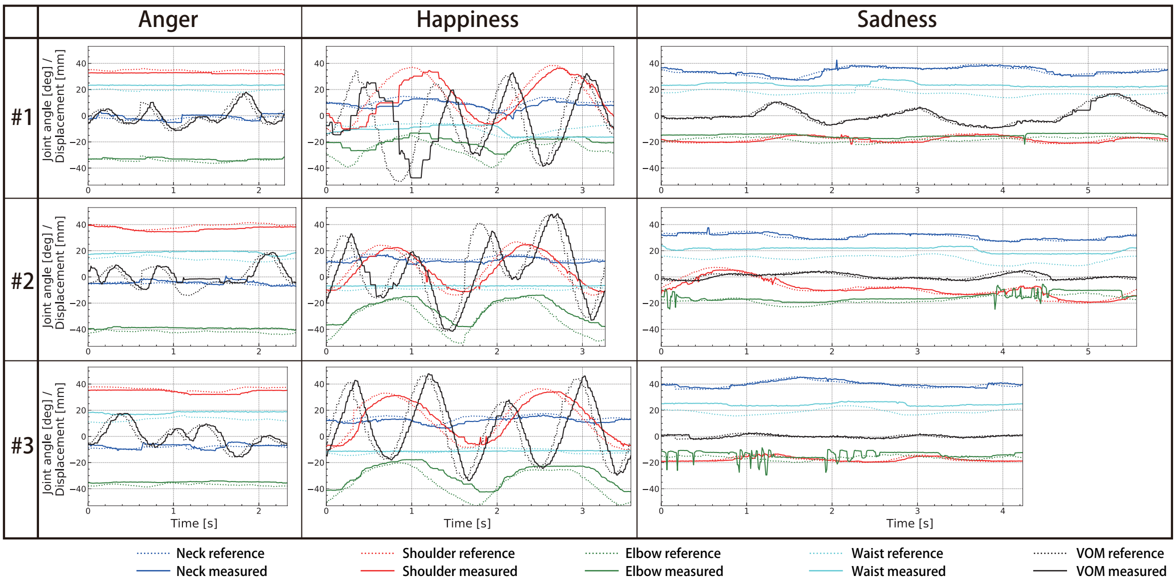


Fig. 5.1 Gait-induced upper body motions of anger, happiness, and sadness emotion. The vertical axis represents joint angles [deg] about two cycle gaits of the neck, right shoulder, right elbow, waist, and displacement [mm] of the vertical oscillation. The horizontal axis represents time [s]. The dotted lines show the reference joint angles, and the solid lines show the measured angles. (The fluctuation of the measured sensor value was due to the electrical noise.)

The human walking speed in the traveling direction is not usually constant due to the reaction force on the sole of the foot. However, for a practical and easy integration with other robot systems, the angular velocities of wheels were approximated as a constant speed and we used an average horizontal speed of aCoM except for the neutral emotion.

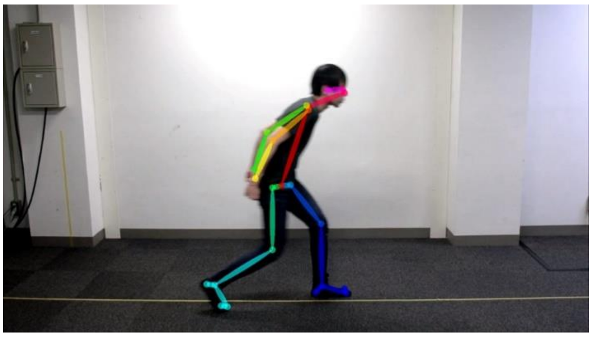
We filmed *ibuki*'s emotional gait-induced upper body motions by using the motion data above (Figure 5.2). The same shooting environment as for the original human motion data was used. Two cameras were installed at the front and side. The front camera was at a height of 0.9 m, placed in front of the walking route. The side camera was at a height of 0.9 m, placed 2.6 m away from the middle of the walking route. The output of the motor that drives the VOM at happiness motion was insufficient, so in actual shooting, the robot was moved at 1/3 speed, and 3x speed videos were shown to participants, thus mimicking the same speed as humans when walking. *ibuki*'s face was covered with paper to ensure participants did not judge the emotion based on the facial expression.

5.2.2 Procedure

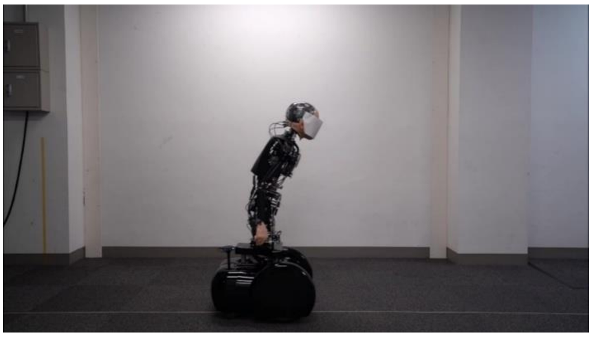
The purpose of this research is to investigate the human perception of the gait-induced upper body motions with the VOM to achieve emotional expressions when a wheeled android is moving. In Experiment 1, we investigated whether the presence or absence of the vertical oscillation improves the human perception of *ibuki*'s emotional expressions through the gait-induced upper body motions.

In both Experiment 1 and 2, we recruited all the participants by using Amazon Mechanical Turk. To control the regional influence on the survey, we limited the survey target to North America because it had the largest available pool of participants. In our preliminary survey, we chose participants from other areas and confirmed that the perception result did not have any large changes. Before starting the experiment, we explained the experiments' purpose to all the participants and obtained consent forms. The protocol was approved by the ethics committee for research involving human subjects at the Graduate School of Engineering Science, Osaka University (#R1-6).

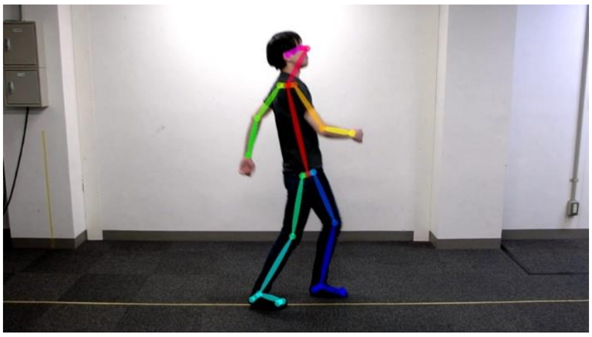
Firstly, we checked the recognition rates and confidence levels of side views of the original human gait motions used for *ibuki*. 79 people participated in this survey (29 females and 50 males, average age = 36.2, Standard Deviation (SD) = 10.1). At the beginning of the survey, the neutral emotion gait motion was shown as a reference. After that, a video of the nine original human gait (3 emotions x 3 motions) was displayed on each page in random order one at a time. Each video was looped three times. After each viewing, participants were asked to: 1) Assess which emotion was expressed from four choices: anger, happiness, sadness, and other; 2) Indicate the confidence level of their choice from a range of totally not confident (1) to highly confident (5).



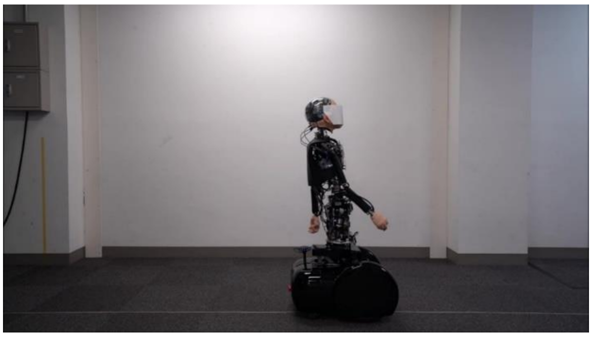
(A) Anger



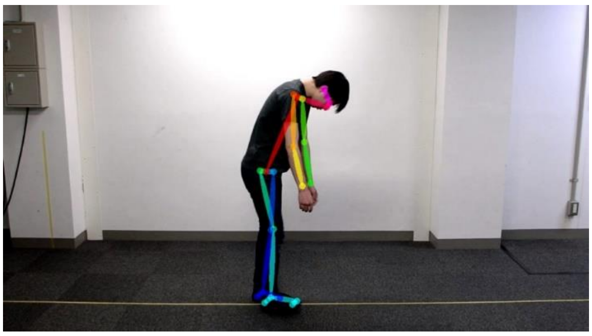
(B) Anger of *ibuki*



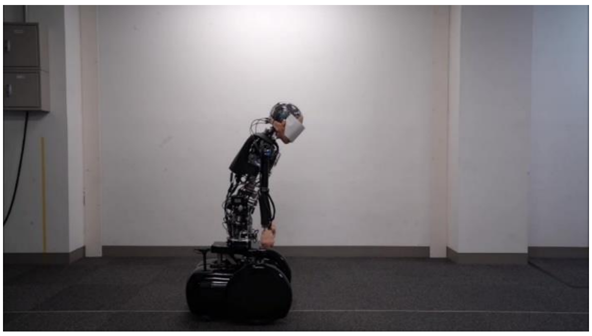
(C) Happiness



(D) Happiness of *ibuki*



(E) Sadness



(F) Sadness of *ibuki*

Fig. 5.2 Representative frames of the original human emotional expressions and *ibuki*'s gait-induced upper body motions expressing emotions. Anger (A) (B), happiness (C) (D), and sadness (E) (F). (A), (C), and (E) also show the results of pose recognition.

Secondly, we investigated the recognition rates and confidence levels of *ibuki*'s emotion expressions by the gait-induced upper body motions. As we assumed the hypothesis that the presence of a vertical oscillation enhances the participants' perception, we conducted the survey under two condition of emotion expressions with and without the vertical oscillation. The significance level was 5.000 % (the marginally significance level was 10.000 %), the detection power was 90 %, and we conducted the survey with a sample size of 150 participants. We divided the new participants in two different groups: one group of 178 people (76 females and 102 males, average age = 36.3, SD = 9.68) saw the videos with vertical oscillation, and the other group of 177 people (68 females and 109 males, average age = 36.6, SD = 10.6) saw the videos without vertical oscillation.

For both with and without vertical oscillation conditions, we prepared nine videos of *ibuki*'s gait-induced upper body motions, combining three motions per emotion from a side view. To prevent participants from just watching side view motions, we also prepared nine videos from the front view as dummy trials which are not used in the result analysis. In total, each participant watched 18 videos.

Like in the previous survey with the human data, at the beginning of each condition, the gait-induced upper body motion of neutral emotion without vertical oscillation was shown as a reference. For both with and without vertical oscillation conditions, a video was displayed on each page in random order one at a time. Each video was looped three times. After each viewing, participants were asked to: 1) Assess which emotion was expressed from four choices: anger, happiness, sadness, and other, 2) Indicate the confidence level of their choice from a range of totally not confident (1) to highly confident (5).

5.3 Result

5.3.1 Experiment 1

For the original human gait motions, the average recognition rates of anger, happiness, and sadness were 97.0 %, 93.2 %, and 96.2 %, respectively. The average confidence levels of anger, happiness, and sadness expressing motions were 4.51 (SD = 0.70), 4.44 (SD = 0.66), and 4.49 (SD = 0.73), respectively. Table 5.1 shows the confusion matrices of the recognition rates for the three emotions.

For the *ibuki*'s gait-induced upper body motions, Table 5.2 and Table 5.3 show the confusion matrices of the average recognition rates for the three emotions (anger, happiness and sadness) for both with and without vertical oscillation conditions. In the with vertical oscillation condition (Table 5.2), recognition rates are: 56.0 % for anger, 77.7 % for happiness, and

Table 5.1 Confusion matrix of recognition rates of the original human emotional expressions in Experiment 1. Expressed emotions are shown in columns and selected emotions are shown in rows. Each value indicates the average recognition rate [%].

	Anger	Happiness	Sadness	Other
Anger	**97.0**	1.3	0.4	1.3
Happiness	0.4	**93.2**	0.8	5.5
Sadness	0.0	0.4	**96.2**	3.4

97.0 % for sadness. On the other hand, in the without vertical oscillation condition (Table 5.3), recognition rates are: 57.6 % for anger, 63.7 % for happiness, and 95.9 % for sadness. Chi-square tests were performed to check whether participants could distinguish the three different emotions during their assessment. As a result, both conditions were statistically significant: $\chi^2(6) = 343.4$, $p < 0.001$ in the with vertical oscillation condition and $\chi^2(6) = 314.8$, $p < 0.001$ in the without vertical oscillation condition. Mann-Whitney U-tests were performed to check the difference in recognition rates between with and without vertical oscillation for each emotion. To reduce the possibility of false positives due to multiple tests (for three emotions), we adjusted the significance level from 5.000 % to 1.667 % according to the Bonferroni adjustment. As a result, there was significant difference for happiness ($U = 1.62 \times 10^5$, $p < 0.001$, effect size Cohen's $d = 0.313$). There was no significant difference for anger ($U = 1.39 \times 10^5$, $p = 0.591$, $d = 0.033$) and sadness ($U = 1.43 \times 10^5$, $p = 0.313$, $d = 0.062$).

Table 5.2 Confusion matrix of recognition rates under the with vertical oscillation condition in Experiment 1. Expressed emotions are shown in columns and selected emotions are shown in rows. Each value indicates the average recognition rate [%].

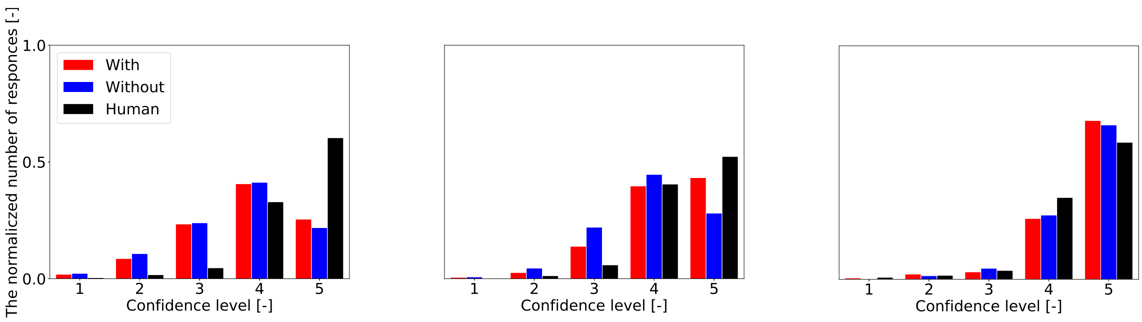
	Anger	Happiness	Sadness	Other
Anger	**56.0**	9.0	18.7	16.3
Happiness	6.9	**77.7**	4.9	10.5
Sadness	0.7	0.7	**97.0**	1.5

Figure 5.3 shows the distributions of the confidence level in the with and without conditions (red and blue) and the original human gait (black). The vertical oscillation condition is indicated by the red color. The averages of participants' confidence level of assessed emotions are: 3.79 (SD = 0.98) for anger, 4.22 (SD = 0.82) for happiness, and 4.59 (SD = 0.71)

Table 5.3 Confusion matrix of recognition rates under the without vertical oscillation condition in Experiment 1. Expressed emotions are shown in columns and selected emotions are shown in rows. Each value indicates the average recognition rate [%].

	Anger	Happiness	Sadness	Other
Anger	**57.6**	10.0	16.8	15.6
Happiness	4.5	**63.7**	9.6	22.2
Sadness	0.9	2.8	**95.9**	0.4

for sadness. Without vertical oscillation condition is indicated by the blue color. The averages of participants' confidence level of assessed emotions are: 3.70 (SD = 1.00) for anger, 3.95 (SD = 0.86) for happiness, and 4.58 (SD = 0.67) for sadness. Mann-Whitney U-tests were performed to check the difference in confidence levels between with and without vertical oscillation for each emotion. As a result, there was a significant difference for happiness (U = 1.68 × 10^5, p <0.001, d = 0.329). There was no significant difference for anger (U = 1.49 × 10^5, p = 0.125, d = 0.096) and sadness (U = 1.44 × 10^5, p = 0.535, d = 0.012).



(A) Anger (B) Happiness (C) Sadness

Fig. 5.3 Histograms of confidence level of (A) Anger, (B) Happiness, and (C) Sadness emotions in Experiment 1. The vertical axis represents the normalized number of responses and the horizontal axis represents each confidence level. Red, blue, and black color indicate confidence levels of the with vertical oscillation condition, the without vertical oscillation condition, and the original human gaits, respectively.

5.3.2 Experiment 2

In Experiment 1, we compared two conditions - the presence or absence of vertical oscillation. The result of happiness emotional expression supported the hypothesis that gait-induced upper body motion with vertical oscillation enhances human perception for robot emotional expressions. In Experiment 2, we investigated whether the congruence of the vertical oscillation enhances the human perception of *ibuki*'s emotional expressions through the gait-induced upper body motions. Therefore, an additional group of participants evaluated the gait-induced upper body motions with exchanged vertical oscillations (incongruent motions). After this, we analyzed the perception difference between the incongruent motions (newly obtained with this assessment) and the previous congruent motions (the with vertical oscillation condition in Experiment 1).

First, we describe how we created an incongruent motion with an exchanged vertical oscillation. We selected similar gait phase pairs from three types of anger and happiness emotions in Figure 5.1. One pair was Anger 1 and Happiness 1, which motions began at the right leg pre-swing phase (50 % in a gait cycle) and also ended at the right leg pre-swing phase (50 % in a gait cycle). Thus, the phase transitions of those two vertical oscillation were almost matched. Another pair was Anger 2 and Happiness 2. Anger 2 began at the right leg pre-swing phase (50 % in a gait cycle) and also ended at the right leg pre-swing phase (50 % in a gait cycle). On the other hand, Happiness 2 began at the right leg stance phase (0 % in a gait cycle) and also ended at the right leg stance phase (10 % in a gait cycle). As mentioned before, the human CoM oscillates twice on the sagittal plane in one gait cycle, the phase transitions of those two vertical oscillations were also matched. However, the remaining pair of Anger 3 and Happiness 3 motions could not be exchanged due to an inconsistency in gait phases. Anger 3 began at 50 % and ended at 30 % in a gait cycle. In contrast, Happiness 3 began at 10 % and ended at 20 % in a gait cycle. As a result, we created two incongruent anger motions with happiness vertical oscillations and two incongruent happiness motions with anger vertical oscillations as shown in Figure 5.4. Also, sadness motions were not exchanged as the recognition rate of the robot's emotional expression of sadness was as high as that of the rate in the human data - with or without vertical oscillation.

Therefore, we prepared two pairs of incongruent anger and happiness motions with exchanged vertical oscillations, one pair of original anger and happiness motions, and three original sadness motions. Like in Experiment 1, we also prepared nine videos from the front view as dummy trials. In total, we prepared 18 videos for Experiment 2. Similar to the way we conducted Experiment 1, at the beginning of each condition, the gait-induced upper body motion of neutral emotion without vertical oscillation was shown as a reference. After that, a video was displayed on each page in random order one at a time. Each video was looped three times. After each viewing, participants were asked to: 1) Assess which emotion was

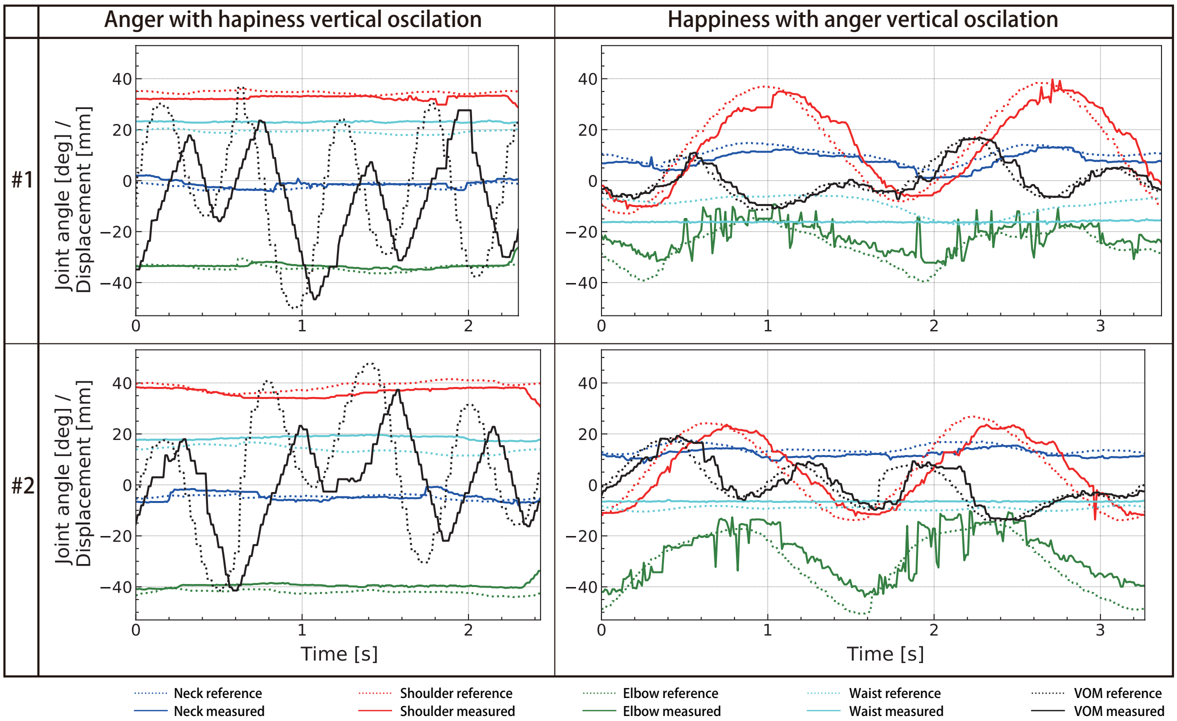


Fig. 5.4 Gait-induced upper body motion under the exchanged vertical oscillation condition. The graph shows two anger motions with happiness vertical oscillation and two happiness motions with anger vertical oscillation. The vertical axis represents joint angles [deg] about two cycle gaits of the neck, right shoulder, right elbow, waist, and displacement [mm] of vertical oscillation. The horizontal axis represents time [s]. The dotted lines show the reference joint angles, and the solid lines show the measured angles. (The fluctuation of the measured sensor value was due to the electrical noise.)

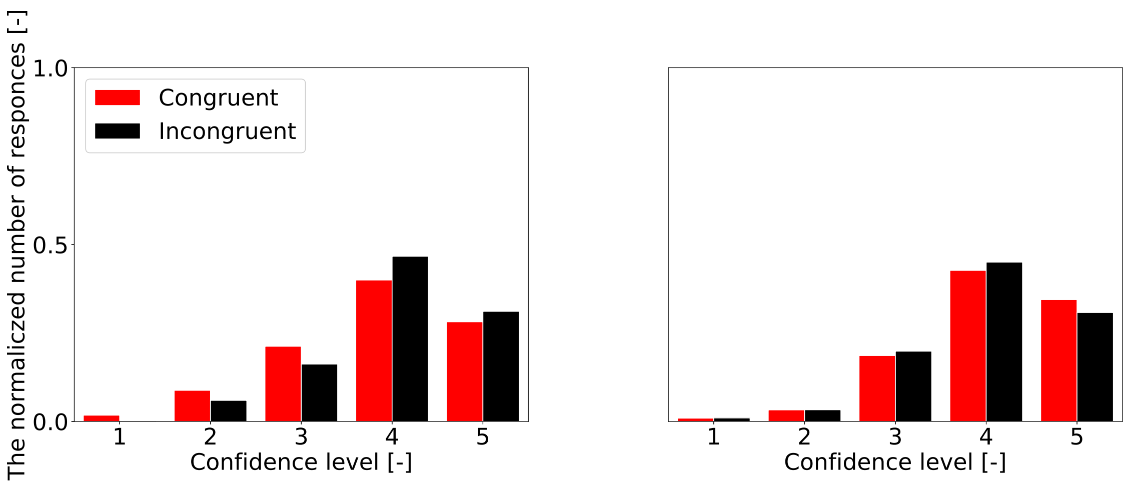
expressed from four choices: anger, happiness, sadness, and other, 2) Indicate the confidence level of their choice from a range of totally not confident (1) to highly confident (5).

A total of 151 people (70 females and 81 males, average age = 35.1, SD = 9.27) participated in Experiment 2. As in Experiment 1, the significance level was 5.000 % (the marginally significance level was 10.000 %), the detection power was 90 %, and we conducted the survey with a sample size of 150 participants.

Table 5.4 show the confusion matrices of the recognition rates for the incongruent motions. Figure 5.5 show the histograms of confidence level of (A) anger motions and (B) happiness motions. Red and black color indicate confidence levels of congruent and incongruent motions. For the anger gait-induced upper body motion with the happiness vertical oscillation, the recognition rate was 48.3 % and the average confidence level was 4.03 (SD = 0.84). For the happiness gait-induced upper body motion with anger vertical oscillation, the recognition rate was 58.6 % and the average confidence level was 4.01 (SD = 0.85). On the other hand, for congruent motions in Experiment 1, the recognition rate and the average confidence level were 48.3 % and 3.79 for anger, and 58.6 % and 4.22 for happiness.

Table 5.4 Confusion matrix of recognition rates for the incongruent motions in Experiment 2. Expressed emotions are shown in columns and selected emotions are shown in rows. Each value indicates the average recognition rate [%].

	Anger	Happiness	Sadness	Other
Anger	**48.3**	27.5	13.6	10.6
Happiness	6.0	**58.6**	12.9	22.5



(A) Anger (B) Happiness

Fig. 5.5 Histograms of confidence level of (A) anger motions and (B) happiness motions. Red color indicates confidence levels of congruent motions which were assessed in the with vertical oscillation condition of Experiment 1 (Red histogram in Figure 5.3). Black color indicates confidence levels of incongruent motions with exchanged vertical oscillation which were assessed in Experiment 2. The vertical axis represents the normalized number of responses and the horizontal axis represents each confidence level.

Mann-Whitney U-tests were performed to check the difference in recognition rates and confidence levels between congruent and incongruent motions. To reduce the possibility of false positives due to multiple tests, we adjusted the significance level from 5.000 % to 2.500 % according to the Bonferroni adjustment. Between the incongruent and congruent motions of anger and happiness, there was a marginally significant difference in recognition rates for anger (U = 7.45 × 10^4, p = 0.033, d = 0.154) and significant difference for happiness (U = 6.52 × 10^4, p <0.001, d = 0.429). There were also significant differences in confidence levels for both anger (U = 9.10 × 10^4, p = 0.001, d = 0.255) and happiness (U = 6.91 × 10^4, p <0.001, d = 0.254).

From these results, we concluded that the recognition rates decreased when the vertical oscillation did not match the emotional human movement. Interestingly, participants' confidence level in anger with happiness vertical oscillation gait-induced upper body motion was significantly higher than that of anger gait-induced upper body motion with anger vertical oscillation in Experiment 1. Thus, the results show that easily recognizable motion patterns and high confidence motions patterns may not always match - as seen in the higher confidence level of anger with happiness vertical oscillation condition compared to anger with the original vertical oscillation condition.

5.4 Discussion

We implemented gait-induced upper body motions that express three emotions using the VOM of the mobile android *ibuki*. We also confirmed that the recognition rate and confidence level of the emotional expression of happiness were higher with vertical oscillation. From the incongruent motion of anger and happiness, we can see that the incongruence of the vertical oscillations of anger and happiness decreased recognition rates. However, the confidence level of anger was increased by using happiness vertical oscillation.

For reference, we summarize our results and previous research of emotional expression of robot walking motion here. Table 5.5 shows the robot height, the survey scale, the number of question answer options, and recognition rates of emotional expression of robot walking in previous research: *ibuki*'20 [120], Wabian-2R [30], CG of Wabian-2R [121], Nao [31], and CG of human animation [20].

In previous research [120], we investigated *ibuki* gait-induced upper body motions based on Destephe's data [63], adding simple cosine wave VOM movements. Recognition rates at *ibuki*'20 in Table 5.5 are under the condition of 0.8x speed and simple cosine wave vertical oscillation (there were four conditions by combining [0.6x, 0.8x speed] x [with, without vertical oscillation]). Note that, in the previous study, only recognition rates with a confidence level of more than 2 degrees were counted in the results. Because the gait-induced upper body motions were consisted of the same joints, it is suggested that vertical oscillation of the

Table 5.5 Summary of emotional expression recognition by walking motions of robots in previous researches: this research (Human and *ibuki*), the previous *ibuki*'20 study by [120], Wabian-2R [30]) and the original CG by [121], and Nao by [31] and the original CG by [20]. Height indicates a robot or a CG avatar height as a representative size scale. Participants indicate people who participated the survey. Choices indicate emotion options in the survey. Each value indicates the average recognition rate [%]. * indicates a value which is not shown the exact value in a paper. (The author read the value from a graph.)

	Height [cm]	Participants	Choices	Anger	Happiness	Sadness	Fear	Neutral
Human	170	79	4	97.0	93.2	96.2	N/A	N/A
ibuki	120	178	4	56.0	77.7	97.0	N/A	N/A
ibuki '20	120	17	3	43	62	88	N/A	N/A
Wabian-2R	148	26	6	N/A	55.8	92.3	N/A	80*
Nao	57	59	5	51.8	34.7	66.1	15*	63.9
CG (Destephe)	N/A	16	6	61.9	65.5	76.2	85.7	N/A
CG (Venture)	N/A	26	6	90*	65*	90*	90*	90*

human emotional gait, and not a simple cosine wave, enhances the human perception for the gait-induced upper body motions. In addition, the average confidence level for all emotions was 3.43 in the previous study. In contrast, in this research, it was 4.07 for the with vertical oscillation condition. Again, it is considered that the application of the human emotional gait for the vertical oscillation contributed to the increase in confidence level.

For the perception of anger, the gap in recognition rates between humans and robots is pronounced. In the case of humans, emotional expression of anger was recognized with the same accuracy as other emotions. However, in robots they were about 50 %. We consider that velocity and acceleration of walking motions are important for the correct recognition of angry emotion. Montepare reported that the anger gait is recognized by a heavy-footed gait characteristic [23]. In our experiments, as we used time series joint angle data for emotional gait-induced upper body motions, the lack of consideration of velocity and acceleration might have caused the lack of recognition. We can see this also happening if we review [20] and [31]. In [20], where she considered joint angles and velocities of human walking while creating walking animations, the recognition rate of anger was 78 %. Even though [31] used that as a basis to move their robot, the recognition rate decreased to 51.8 %. Furthermore, Ikeda and Watanabe reported that humans are better able to perceive anger than happiness [122]. In the natural world, angry creatures, either animals or humans, are objects of fear, and observers must be sensitive in taking action (e.g. fight or flight). Therefore, it is natural to have an advantage in perceiving anger. Nevertheless, it is difficult to perceive anger expressed by robots. To achieve a high level of perception for a robot's expression of anger, we must sufficiently reproduce a wider range of components, from the joint movement features to the appearance humans use to convey their anger. On the other hand, the results show that confidence level for the anger emotion was not correlated to the original motion. For that reason, the vertical oscillation of happiness could increase the confidence level of anger expression in Experiment 2.

For the perception for happiness, our results suggested that vertical oscillation is important for both recognition and confidence aspects. This is consistent with [22] which identified faster gait speed and large vertical movement as characteristics of happiness. We calculated the approximate ratio of vertical movement compared to the heights of humans, *ibuki*, *ibuki*'20, Wabian-2R, and Nao in Table 5.5. The ratio of the original human walking motion and *ibuki* were both 0.06. In other words, the 120 cm high *ibuki* was swinging the upper body about 7 cm while moving. In the previous research, *ibuki*'20 ration was about 0.04 (5 cm vertical movement / 120 cm height). Wabian-2R was about 0.007 (1 cm / 148 cm, referred from [123]) and Nao was 0 (0 cm / 57 cm, as inferred from [124]). From these results, it is suggested that there is a positive correlation between the ratio of vertical movement per height and the recognition rate of happiness.

For the perception of sadness, our results show that the effects of the vertical oscillation were small for both recognition and confidence. This is consistent with [30] that the locomotor unit are important for the expression of happiness in the walk but does not have mush influence on the expression of Sadness. As can be seen in the study by [20], sadness has different gait characteristics compared to anger and happiness, such as the neck and torso bending forward. As a result, recognition rates for humanoid robots (*ibuki*, Wabian-2R) were as high as for human gaits. Since those robots have a human-like appearance, participants were easily able to perceive sadness with a high level of confidence.

The limitations are mainly that this research used online video surveys with a questionnaire that had a pre-designed response range. Therefore, this method may miss some details, like participants' further perceptions when they watched the emotional gait-induced upper body motions. Further real-time robot evaluation by participants is needed for further understanding. From the author's experience, we expect that emotional perception is made stronger by watching real robot motions. In addition, humans also express gait characteristics on the roll or yaw axis joints. However, the gait-induced upper body motions in this research were limited only to the sagittal plane.

Chapter 6

Research3

This chapter shows the synchronization of gait phases between a human and a wheeled humanoid robot. We implemented a periodic upper body motion in a wheeled child-like android robot, which oscillates its upper body vertically while moving, and measured the gait phase difference between the robot and a human walking behind the robot. In the experiment, participants walked in a circle behind the robot in a line, and their gaits were captured by a camera to obtain the phase difference between the two. For comparison, we conducted paired walking under four conditions: participants walking with a human, behind the robot with No-Motion, behind the robot that its Arms-Swinging, behind the robot that its Arms-Swinging adding Vertical-Oscillation of the upper body. After the experiment, we analyzed the bias in the phase difference distributions measured at each gait cycle. We then verified whether the gait phase synchronization occurred under each of the four conditions.

6.1 Background

Although robots designed to be integrated into factories are widely used in an industrial setting, it has been challenging to place robots in human living environments. This issue has been actively researched [125, 126]. Previous robotics research has contributed to navigation methods to predict surrounding people's trajectories and avoid collisions [2, 3, 5, 127, 128, 129, 130, 131, 132].

These technologies are one-directional solutions from the robot's perspective on how to perceive human action and act accordingly. The other direction is how humans recognize the robot's actions and intentions accurately. There has been research on methods for communicating the robot's movement intentions [14, 64, 65, 66, 133, 134]; however, the people around the working robot may disregard them and instead feel that the robot is an obstacle [135]. Therefore, the effort of people required to recognize the robot' s action and intention should be minimized, and ideally, the level of effort should be unconscious [136].

It is known that humans walking in large crowds can walk collectively, as they unconsciously synchronize their gait with the surrounding pedestrians [32, 33, 34]. This pedestrian synchronization is induced by a various factors such as hand-holding [137], vibrations transmitted via a bridge [138], and ambient music [139]. Even a small cognitive load, such as listening to a story when walking, enhances a gait synchronization [140]. Previous studies on pedestrian gait synchronization have reported that gait phase synchronization is more likely to occur in high-density groups and that pedestrian synchronization also maximizes the walking flow efficiency [35, 36]. Additionally, it has been reported that there is a psychological aspect to synchronized side-by-side walking, as it improves the people' s impression of each other [141].

If humans can synchronize a gait phase with robots, we can imagine applications with smooth robot-human mobility in our daily lives. In particular, if both the bipedal robots and the more common wheeled robots can synchronize gait phases by utilizing upper body motions, robots can move by taking advantage of the human ability to walk skillfully in the crowd.

6.2 Methods

6.2.1 Participant

26 participants (Mean age 24.1 ± SD 7.9, 11 females and 15 males) were recruited for the experiment. All participants gave written informed consent and agreed to publish the data and images in an online open-access publication. The protocol was approved by the ethics committee for research involving human subjects at the Graduate School of Engineering Science, Osaka University (#R1-6).

6.2.2 Motion generation

We extracted characteristic joint amplitudes and frequencies from a human gait to implement the robot's gait-induced upper body motion. The author (176 cm) filmed himself walking on a treadmill at 0.83 m/s (3.0 km/h) treadmill speed for three minutes in 30 frames per second. At this speed, we confirmed beforehand that the robot was able to stably and safely move indoors with a human.

From the recorded video, joint position time series data was obtained by human pose recognition software MediaPipe [142]. After that, we calculated a time-series joint angle of the neck, the right shoulder and the right elbow in the pitch axis, the waist in the yaw axis, and a vertical displacement of the waist's center position. It is known that human gait motions can be well described by the sum of sinusoidal functions using fast Fourier transformation [143]. Therefore, after all the data were filtered with a 3.3 Hz low-pass filter, we used fast Fourier transformation on these time-series data to acquire Fourier coefficients up the second terms for the amplitudes and frequencies. As the result, the robot's joints $\boldsymbol{\theta}(t)$ at time t were controlled by Equation 6.2.1, where α was set as 0 or π depending on the right or left body and the averages of the original joint angles $\bar{\boldsymbol{\theta}}$ were added.

$$\boldsymbol{\theta} = \boldsymbol{A}\cos(2\pi t\boldsymbol{f} + \boldsymbol{\alpha}) + \bar{\boldsymbol{\theta}} \quad (6.2.1)$$

Table 6.1 shows the calculated parameters for the gait-induced upper body motion. The left column shows the body joints which we controlled and each row shows parameters including amplitudes, frequencies, a phase difference, and a center angle of oscillation. The rotation directions of the pitch and yaw axes coincide with the rotations around the y and z-axes as shown in Figure 6.1.

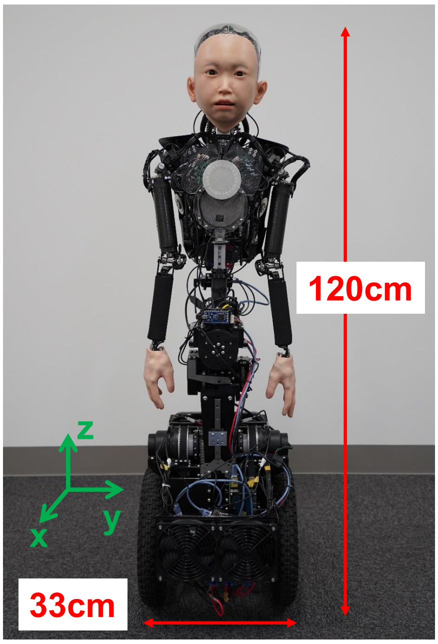


Fig. 6.1 The overview of the coordinate system

For comparison, we conducted paired walking under four conditions in the experiment: participants walking with a human (“HU condition”), behind the robot with No-Motion (NM condition”), behind the robot that its Arms-Swinging (“AS condition”), behind the robot that its Arms-Swinging adding Vertical-Oscillation of the upper body (“AS+VO condition”). For each condition, we controlled the following joints for the robot motions. At NM condition, we did not operate any joints. At AS condition, we operated both sides of the shoulders and elbow joints. At AS+VO condition, we operated the neck, waist, and VOM in addition to the shoulders and elbows. In all conditions, the wheels were driven at 0.83 m/s by a velocity control.

θ	[unit]	A	f	α	$\bar{\theta}$
Neck (pitch)	[deg]	0	-	-	-28.3
Shoulder (right, pitch)	[deg]	13.4	0.575	π	-14.1
Shoulder (left, pitch)	[deg]	13.4	0.575	0	-14.1
Elbow (right, pitch)	[deg]	5.47	0.575	π	-28.6
Elbow (left, pitch)	[deg]	5.47	0.575	0	-28.6
Waist (yaw)	[deg]	6.76	0.575	π	0
VOM	[mm]	18.7	1.15	$\frac{\pi}{2}$	0

Table 6.1 The calculated parameters of each joint θ, the amplitude A, frequency f, phase difference α, and $\bar{\theta}$ for the gait-induced upper body motion.

6.2.3 Procedure

The experiment consisted of five walking sessions. In order, a participant walked behind a human (HU condition), participant walked alone (this was not an experimental condition but for measuring the gait speed when walking freely), walked behind the robot with no motion (NM condition), walked behind the robot with the arms swinging (AS condition), walk behind the robot with the Arm-Swing adding Vertical-Oscillation of the upper body (AS+VO condition).

Each walking session (condition) lasted two minutes. Participants walked into the indoor experiment room shown in Figure 6.2. There was a fixed base in the experiment room with a rotatable lightweight beam at the center. The beam ensured that the distance between the center of rotation and the robot was kept 2.3 m for all conditions. Furthermore, it held the rotating camera in order to record the human gait. The camera was connected and controlled by a single board computer to capture photos of a participant's gait (480 x 680 pixels, 30 fps) with time information during the experiment. We also recorded the robot's joint reference angles and measured angles on 10 Hz in conditions with the robot.

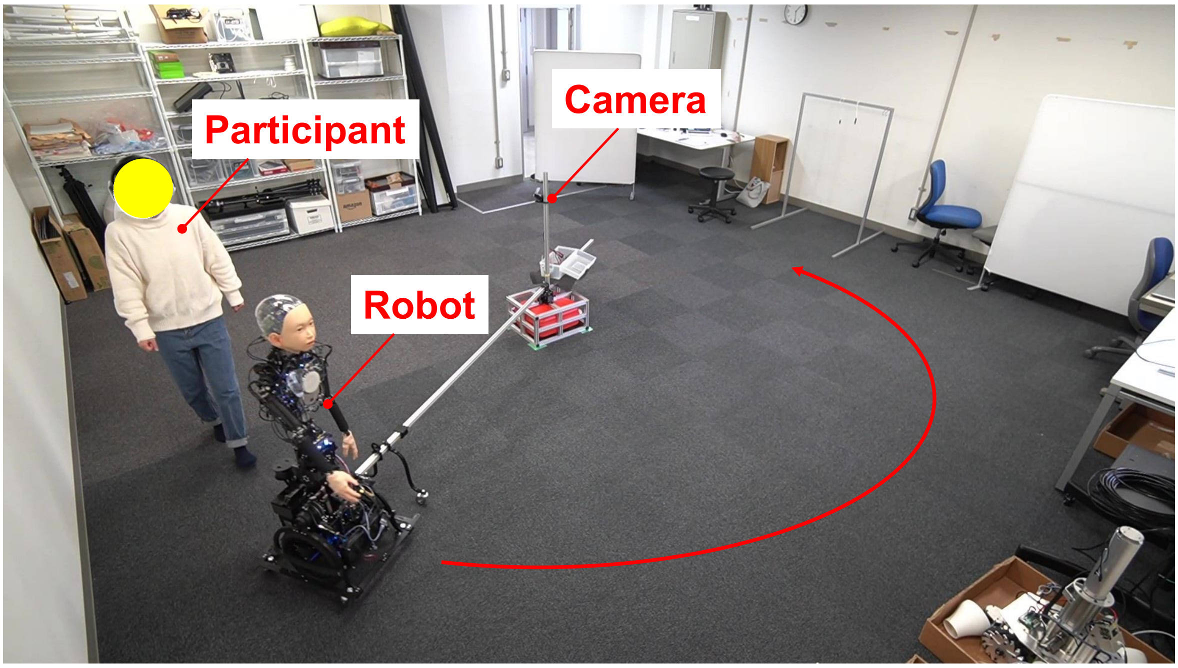


Fig. 6.2 The overview of the experiment. A participant walked behind the robot with three types of upper body motion patterns.

In HU condition, the participant walked behind a human (the author), leading and pushing the beam with the attached camera. In the free-walking session, a participant walked alone and freely without any constraints, except for the author filming with the camera attached to the manually controlled beam. In the conditions with the robot, the robot led and the participants walked behind it. In these sessions, the beam was connected to the robot's mobility unit. The order of these three sessions was changed for every participant to ensure counterbalancing.

The human gait synchronization can be influenced by the surrounding sounds [139, 144]. In order to eliminate the influence of the motor drive noise while the robot was moving, the participants wore noise-cancelling headphones and listened to white noise while walking. In addition, a participant's gait was considered to be symmetrical for the analysis [145].

6.2.4 Analysis

One cycle of the gait phase (0 to 360 deg) is defined as follows: heel-strike (0 deg) in which the right heel touches the ground, middle-stance (around 90 deg) in which the right foot supports the upper body and the left foot swings forward, pre-swing (180 deg) in which the left heel touches the ground and the right foot leaves the ground, middle-swing (around 270 deg) in which the left foot supports the upper body and the right foot swings forward, and then back to the first heel-strike (360 deg) where the right heel touches the ground [146].

For each condition, the gait phase is calculated using the left leg angle from the captured human gait images. Firstly, two positions of a participant's left waist and left knee were estimated by using image pose recognition (Mediapipe) in every frame to obtain the time-series with the left leg angle against the vertical direction. Next, a 3.3 Hz low-pass filter was applied to remove high-frequency noise from the time-series data. After that, we obtained a time tn that took the maximum of the left leg angle of the n-th gait cycle, and set $\phi_n = 180$ deg as the gait phase at this time.

As the robot does not have legs, we obtained the gait phase $\phi_n = 180$ deg as the gait phase at a time t_n of the n-th gait cycle from a maximum of a virtual gate phase increasing over time, which takes $\omega_l t$ when $0 < \omega_l t < 180$ deg and 0 when $180 < \omega_l t < 360$ deg for NM condition and a minimum of the robot's measured left upper arm angle for both of AS and AS+VO condition. Setting the robot's gait cycle which was closest to the n-th gait cycle of the participant as the n'-th gait cycle, the phase difference $\Delta\phi$ is given by Equation 6.2.2. We calculated an angular frequency ω_n of n-th gait cycle as $\frac{360}{t_n - t_{n-1}}$, and the natural angular frequency ω_0 of each participant was the average value of ω_n obtained when the person walked alone.

$$\Delta\phi = \begin{cases} 360\frac{t_n - t_{n'}}{t_{n'+1} - t_{n'}} & (t_{n'} \geq t_n) \\ \\ 360\frac{t_n - t_{n'-1}}{t_{n'} - t_{n'-1}} & (t_{n'} < t_n) \end{cases} \quad \cdots\cdots (6.2.2)$$

6.3 Result

Figure 6.3 shows one of the representative participants who spontaneously synchronized their gait phase with the human / robot under both the HU and AS+VO conditions. The horizontal axis shows the measurement time (s), and the vertical axis shows the phase difference (deg). No synchronization, i.e. constant phase locking, was observed for NM (Figure 6.3, green) and AS (Figure 6.3, blue) conditions. The phase difference decreased monotonically in every gait cycle. Interestingly, for AS+VO (Figure 6.3, red) condition, we can observe synchronization between the human and the robot, similar to the human control condition (Figure 6.3, yellow). Under HU condition, the gait phase difference was M = 127.6 ± SD 16.4 deg (Mean ± Standard Division) during the entire measurement. Under the robot AS+VO condition, the difference was even kept M = 70.2 ± SD 15.8 deg. As can be seen in Figure 6.3, when the gait phase synchronization occurs under a certain condition, the distribution of the gait phase difference is biased. In contrast, when the synchronization does not occur, the gait phase difference is uniformly distributed. Next, we analyzed whether and if so which under condition the distribution of the gait phase difference is biased.

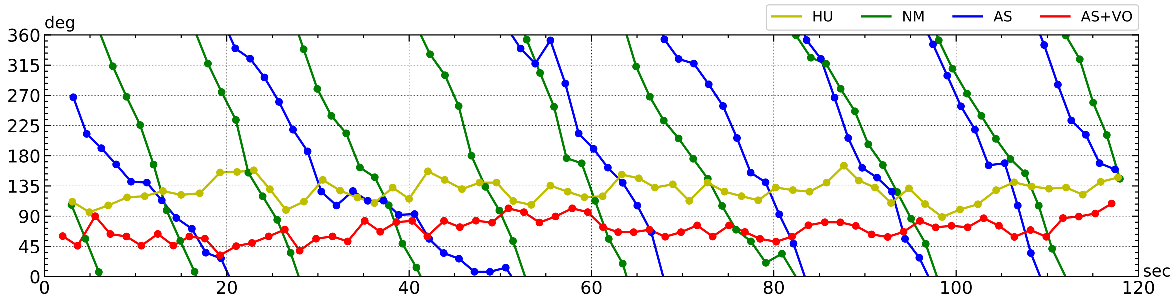


Fig. 6.3 One of the representative obtained time-series phase difference graphs under four condition (Participant 3 in Figure 6.5). The horizontal axis shows the measurement time (s) and the vertical axis shows the phase difference (deg). Yellow: HU, green: NM, blue: AS, and red: AS+VO condition.

Figure 6.4 shows the gait phase difference distributions of 1960, 1993, 2000, and 1989 gait cycles in total from 26 participants in HU, NM, AS, and AS+VO condition, respectively. The circle histogram shows the proportions of obtained phase differences at every five deg. The height in radius direction shows the normalized range of 0 to 12 percent proportionally. For HU condition (Figure 6.4 (a)), the distribution seems to be biased to the directions of 0 and 180 deg, indicating that the gait phases were synchronized in-phase or anti-phase. In the same way, in the AS+VO condition, the distribution seems to be biased to the directions of 45 and 225 deg. To test the distributional bias of the phase differences, we performed Rayleigh test for each condition. We set the p-value to less than 0.05 to be statistical significance and confirmed that the distributional bias was significant for both HU ($p < 0.0001$) and AS+VO ($p = 0.0485$) conditions. At the same time, there was no significant distribution bias for NM (p = 0.9566) and AS (p = 0.0919) conditions.

To determine the degree of synchronization for each participant and condition, we calculated the Phase Locking Index (PLI) [147, 148]. PLI is defined by Equation 6.3.1 using N phase difference data. The value of PLI falls between 0 and 1, with 0 being completely unsynchronized and 1 being perfectly synchronized.

$$PLI = \frac{1}{N}\left|\sum_{i=1}^{N} e^{i\Delta\phi_i}\right| \cdots\cdots \quad (6.3.1)$$

Figure 6.5 shows a heatmap of Phase Locking Index (PLI). The four conditions are indicated in rows from top to bottom, and the participants in columns from left to right (in order of higher PLI under HU condition). It can be seen that some participants in AS+VO condition gave a high PLI as well as HU condition. Especially, there were eleven / five participants whose PLI was higher than 0.5 in HU / AS+VO condition. Figure 6.6 shows the distributions of the gait phase differences of those participants, whose PLI was higher than 0.5 under each condition. Note that, in Figure 6.6 (b), the robot gait phase was calculated based on the VOM displacement, instead of the left upper arm as the way of Figure 6.4 (d). We calculated a time t_n when the VOM displacement took a minimum value as $\phi_n = 225$ deg and we adjusted the gait phase difference by -45 deg. Both distributions have biases to the directions of 0 and 180 deg, indicating that the gait phases were synchronized in-phase or anti-phase.

At the end, Figure 6.7 shows the time-series phase difference graphs for all the participants. In addition, Figure 6.8 shows how participant 1 in Figure 6.5 started to synchronize the gait at the beginning of AS+VO condition, and Figure 6.9 (a)(b) show the gait which was not synchronized under condition AS and synchronized under AS+VO condition.

6.4 Discussion

We investigated the synchronization of gait phases between humans and a wheeled humanoid robot. We implemented a periodic upper body motion in a wheeled child-like android robot and measured the gait phase difference between the robot and a human walking behind the robot under four conditions. We then analyzed the bias in the gait phase difference distribution under each condition by Rayleigh test. As a result, we confirmed a significant distributional bias under the human condition and the robot condition with vertical-oscillation of the upper body.

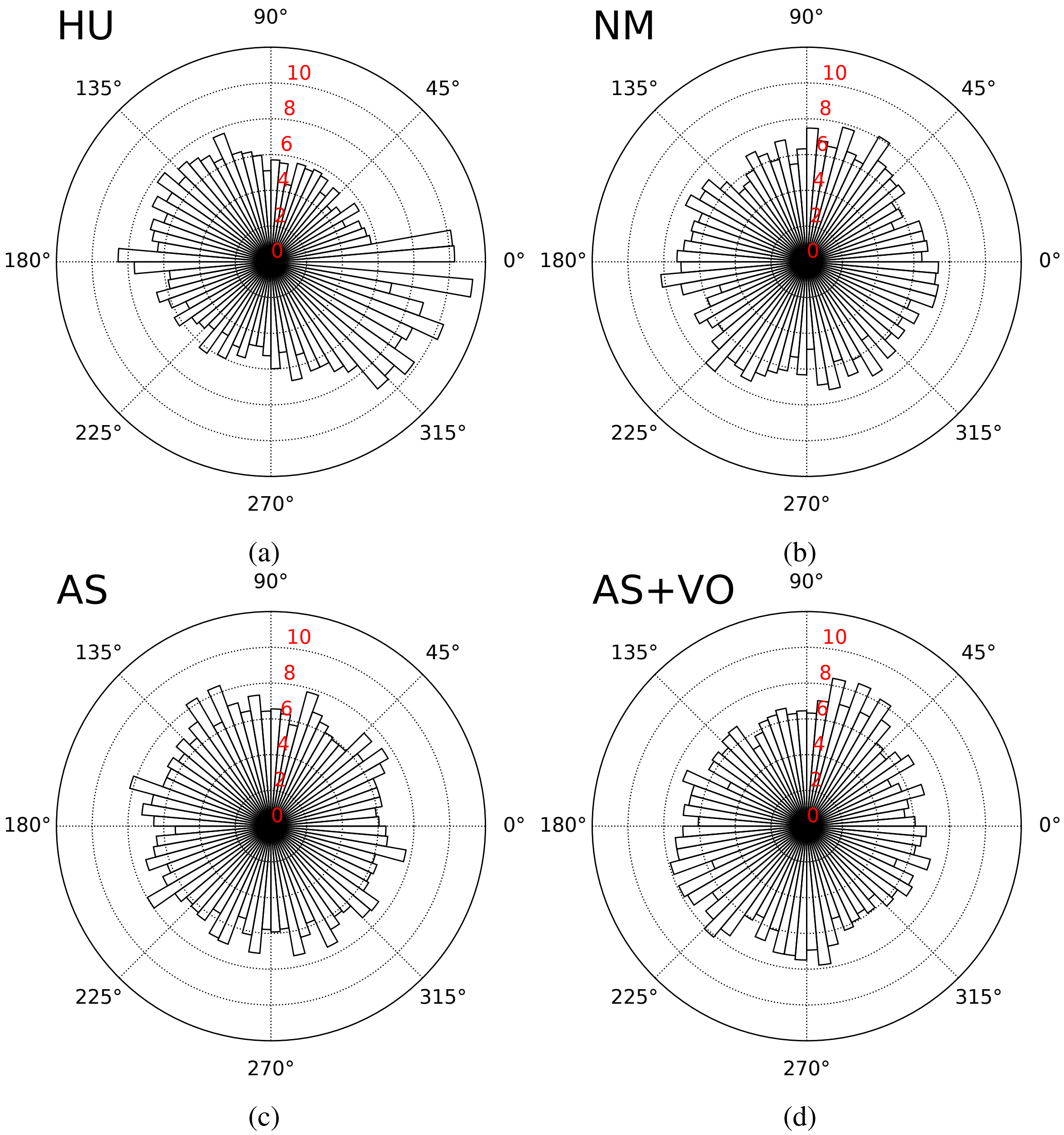


Fig. 6.4 The circle histograms of four conditions which show the proportion of obtained phase differences at every five deg. The height in radius direction shows the range of 0 to 12 percents in the normalized proportion. The distributional bias was significant for both HU ($p < 0.0001$) and AS+VO ($p = 0.0485$) conditions. At the same time, there was no significant distribution bias for NM ($p = 0.9566$) and AS ($p = 0.0919$) conditions.

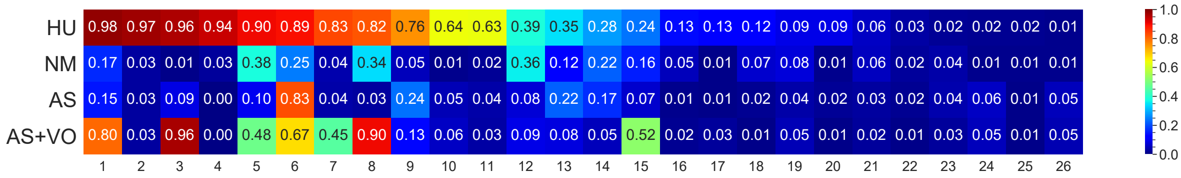


Fig. 6.5 Heatmap of Phase Locking Index for each participant (row) and condition (column).

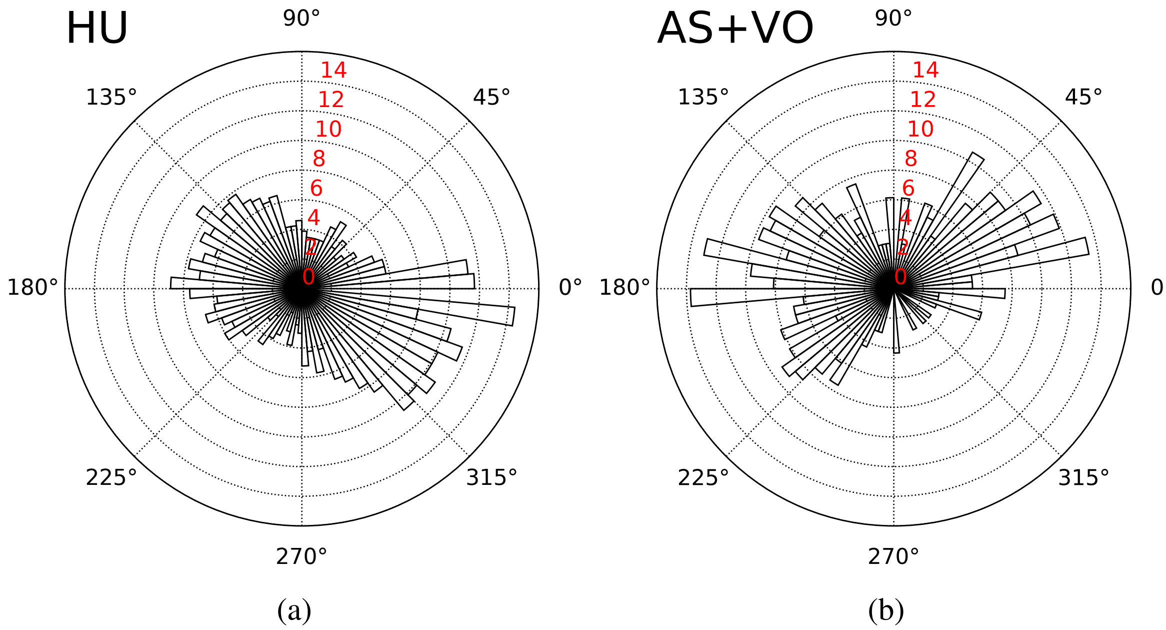
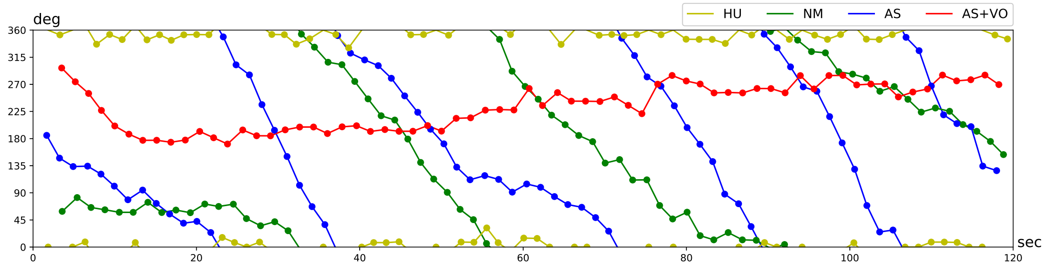


Fig. 6.6 The circle histograms of the synchronized participant only (PLI > 0.5) in HU and AS+VO conditions which show the proportion of obtained phase differences at every five deg. The height in radius direction shows the normalized range of 0 to 16 percents proportionally.

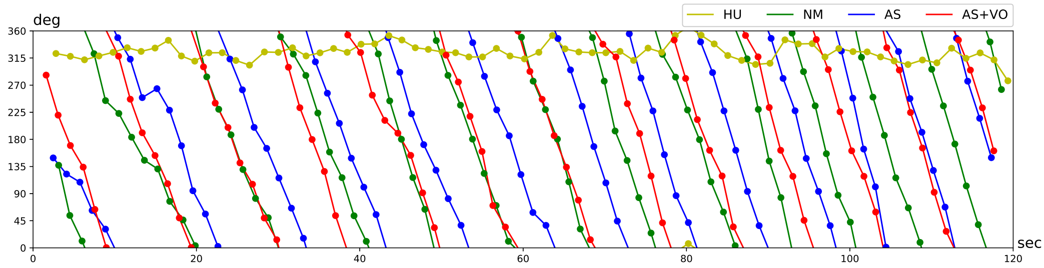
The histogram of the gait phase based on the displacement of VOM in the synchronized participants (Figure 6.6 (b)) showed two types of gait phase synchronization, 0 deg in-phase and 180 deg anti-phase, as in humans. Furthermore, there was a phase delay in the synchronization of the walking phase based on the angle of the left shoulder joint compared to Figure 6.4 (d) and Figure 6.6 (b). This might be because the participants were synchronizing their gaits based more on the VOM than on the arm swing. This is consistent with the previous research which claimed the importance of the upper body swing at a human interpersonal coordination [149, 150, 151]. Since we did not control to align the phases of the robot joints in the experiment, we confirmed the control delay which was about 0.2 seconds in the displacement of the VOM relative to the left shoulder joint. Considering that we controlled the robot by the natural angular frequency of 360 x 0.575 = 207 deg/s, there was a gait phase difference of 225 x 0.2 = 41.4 deg between the shoulder joint and the VOM. This is why the peak of the distribution shown by the synchronization in Figure 6.4 (d) is tilted to the 45 deg direction.

We also investigated the factors that caused the difference in the number of participants who had PLI ¿ 0.5 in the HU and AS+VO conditions. The mean natural angular frequency of the gait phase when a participant walked alone was M = 261.9 ± SD 35.4 deg/s and the mean / controlled angular frequency of the gait phase of the human / robot that moved in front of the participants in the HU and AS+VO conditions were M = 226.2 ± SD 13.3 deg/s and 207.0 deg/s, respectively. The participants who had a closer natural angular frequency, which means a smaller detuning [152], with the human or the robot, synchronized their gait more frequently. Considering that the difference between the angular frequency of the robot and the participant was 54.9 on average, we believed that the robot's gait phase was driven faster by that amount and a higher frequency of synchronization could be observed. For practical use, we need a method to estimate the natural angular frequency from the gait phases of people walking around robots, in order to adjust the robot's gait and to minimize the error of those two angular frequencies of the gait cycles.

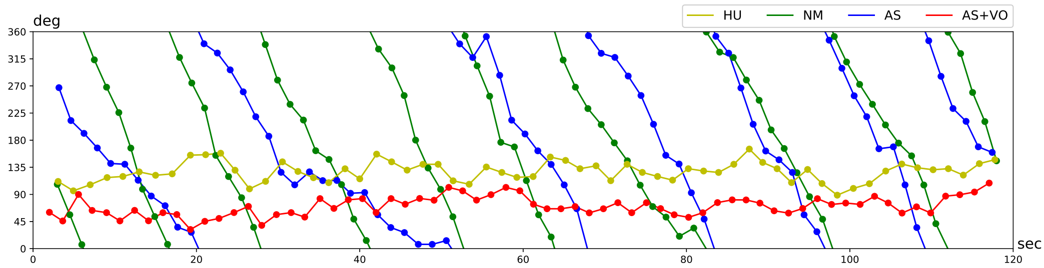
Gait phase synchronization between humans and robots is expected to improve the efficiency of pedestrian flows in settings where both humans and robots are mobile. The mechanism implemented in this research is easy to implement, it only requires oscillating the upper body of a wheeled robot when moving. The number of people that synchronized their gait with the robot in the experiment we conducted was limited. However, we expect that the gait synchronization can spread further even if a robot can synchronize its gait with only a few people around it as it will be conveyed from human to human. Further research is needed to investigate how gait synchronization between humans and robots can be improved to a higher degree and how robot motion can be adjusted based on the inputs from surrounding human gait phase feedback.



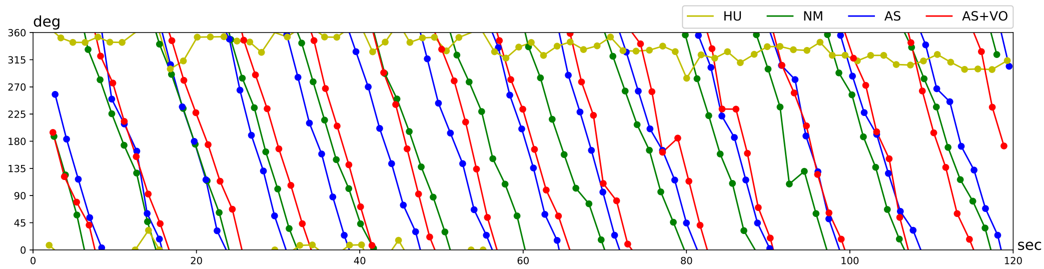
(a) Participant 1



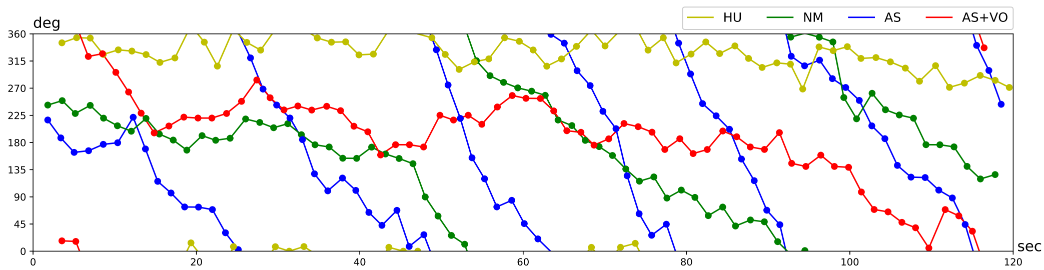
(b) Participant 2



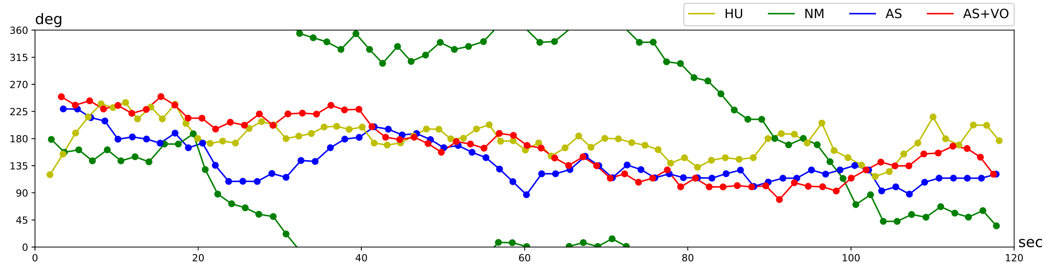
(c) Participant 3



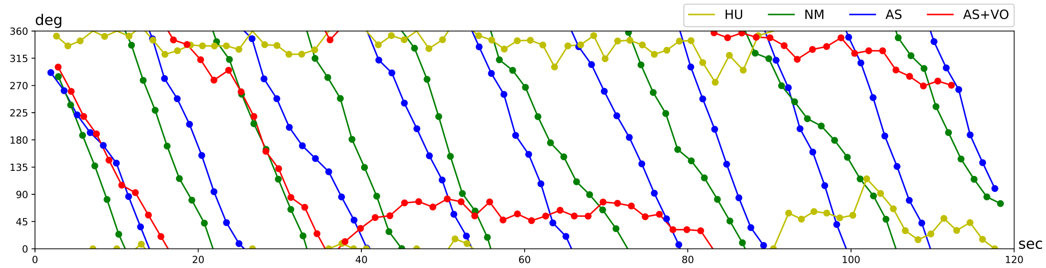
(d) Participant 4



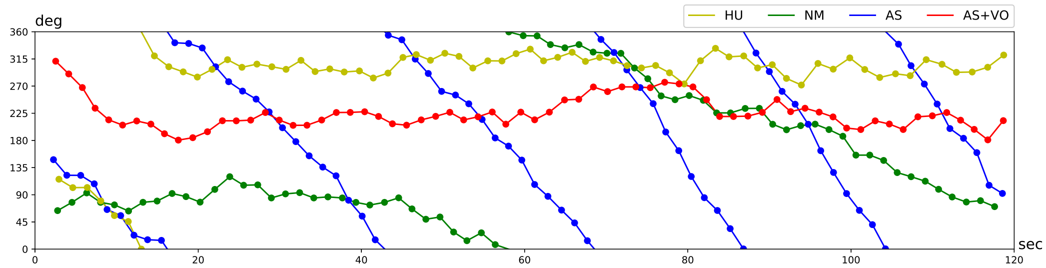
(e) Participant 5



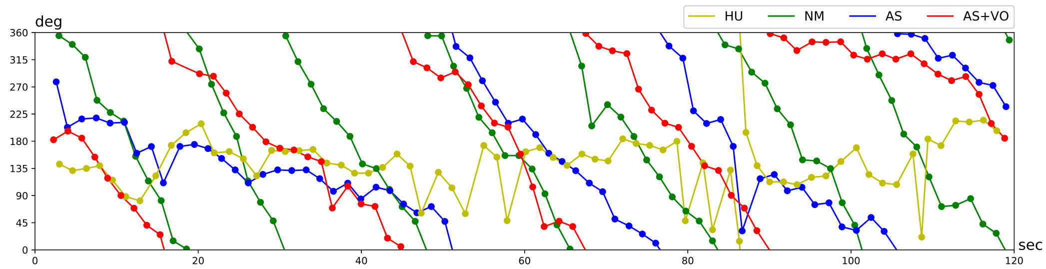
(f) Participant 6



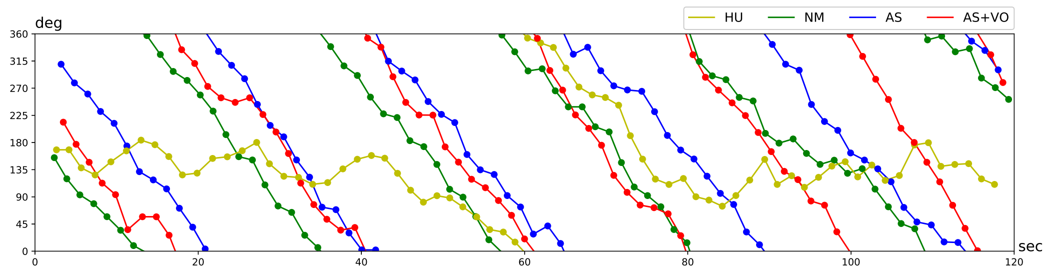
(g) Participant 7



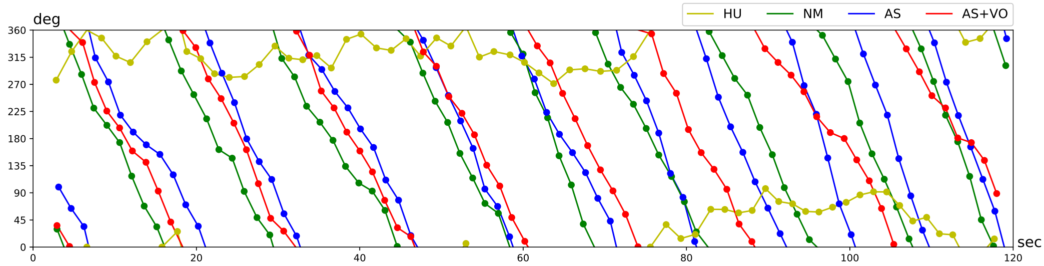
(h) Participant 8



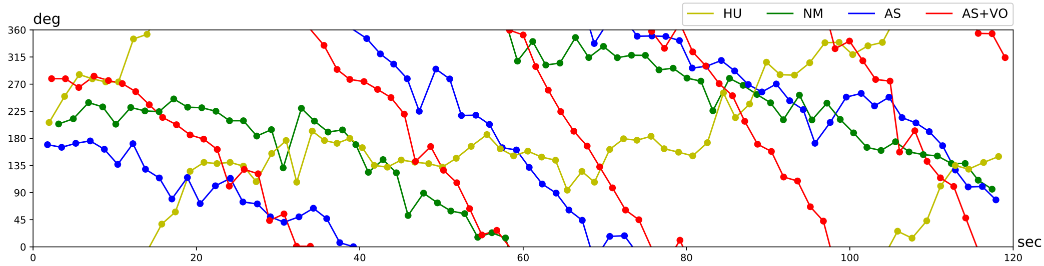
(i) Participant 9



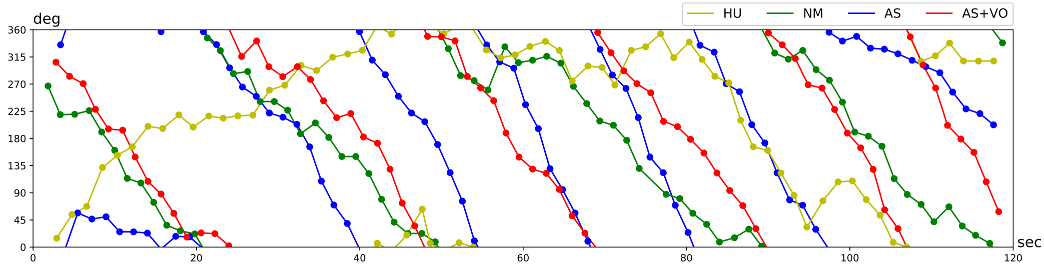
(j) Participant 10



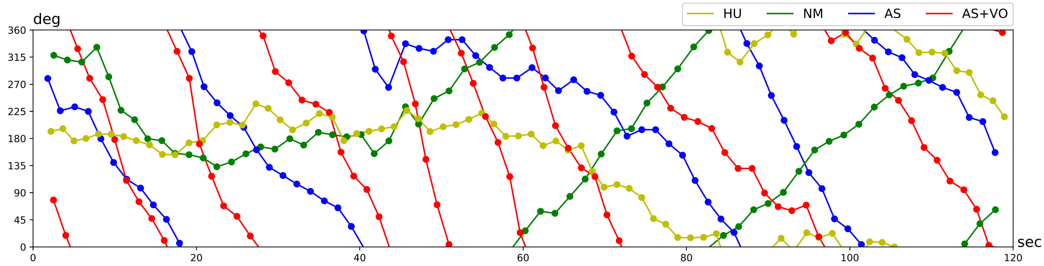
(k) Participant 11



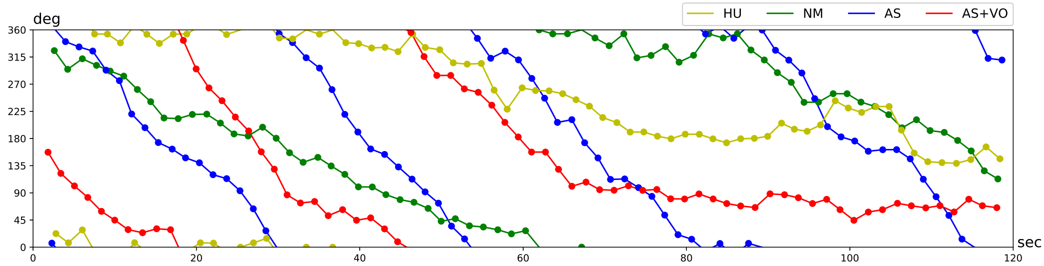
(l) Participant 12



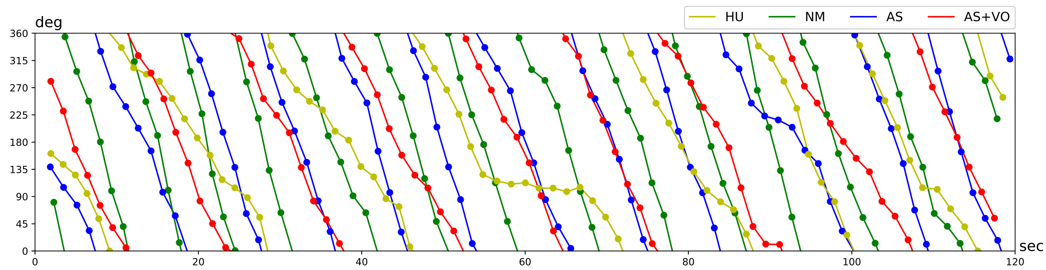
(m) Participant 13



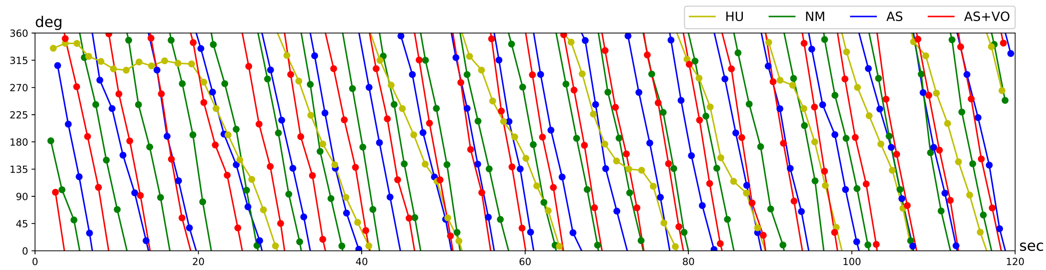
(n) Participant 14



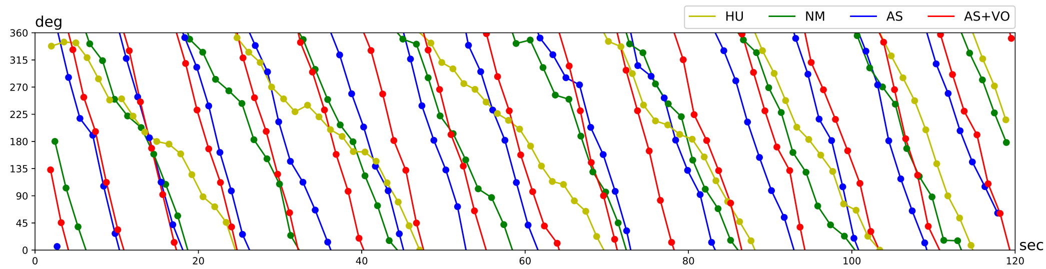
(o) Participant 15



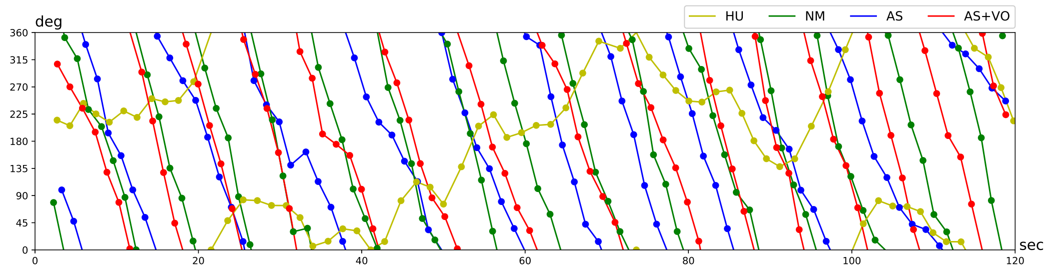
(p) Participant 16



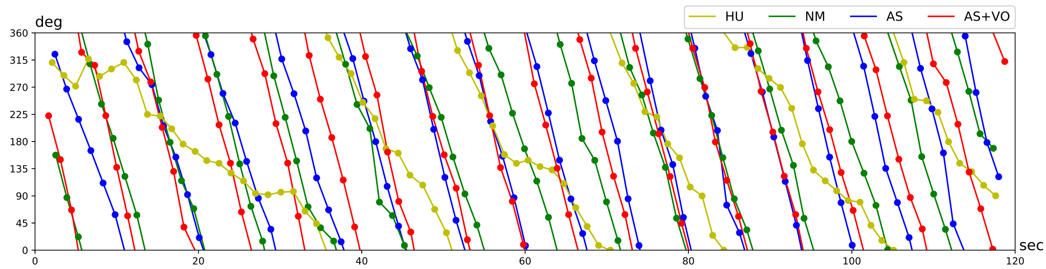
(q) Participant 17



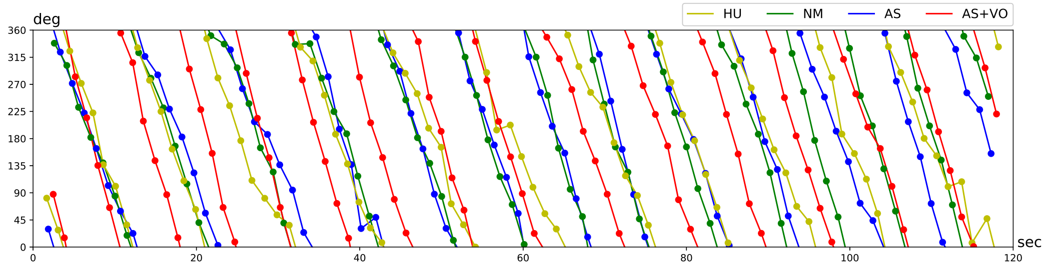
(r) Participant 18



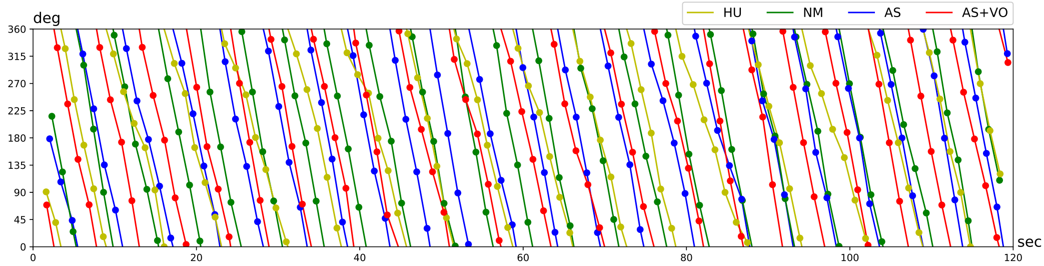
(s) Participant 19



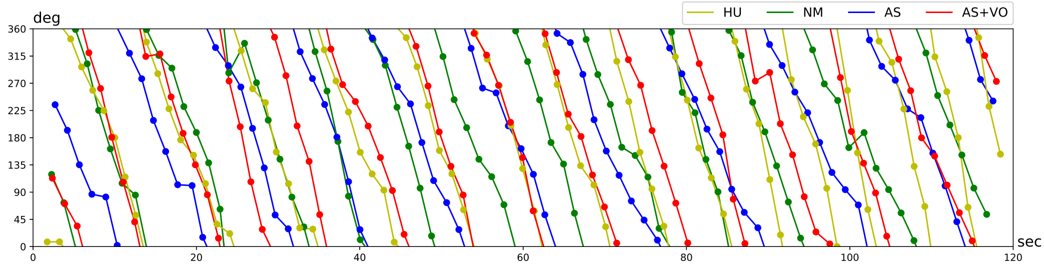
(t) Participant 20



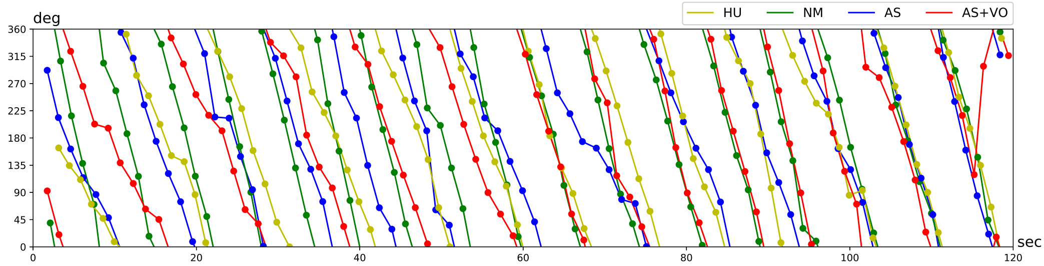
(u) Participant 21



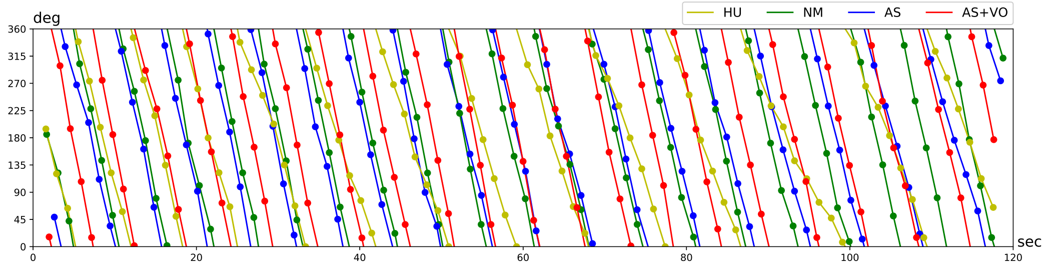
(v) Participant 22



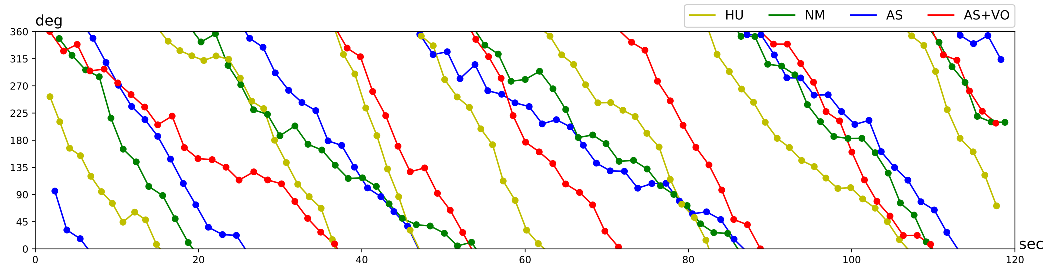
(w) Participant 23



(x) Participant 24



(y) Participant 25



(z) Participant 26

Fig. 6.7 Time-series phase difference graphs under four condition for all 26 participants. Graphs are shown in order of higher PLI at HU condition. The participant number corresponds to the one in Figure 10. The horizontal axis shows the measurement time (s) and the vertical axis shows the phase difference (deg). Yellow: HU, green: NM, blue: AS, and red: AS+VO condition.

Fig. 6.8 A participant's gait spontaneously synchronizes with the gait phase expressed by *ibuki*'s gait-induced upper body motion. Each small picture shows the posture of Participant 1 taken every ten frames (30 frames per second). Time proceeds from left to right, and the next picture for the rightmost one corresponds to the leftmost one in the row below it.

(a) Condition AS

(b) Condition AS+VO

Fig. 6.9 Two types of walking manners captured by a camera in the center of the room. (a) A participant was walking without synchronized under the AS condition and (b) The same participant was walking with synchronized under the AS+VO condition. Each small picture shows the posture of Participant 1 taken every ten frames (30 frames per second). Time proceeds from left to right, and the next picture for the rightmost one corresponds to the leftmost one in the row below it.

Chapter 7

Discussion

This chapter summarizes the importance of body expressions for mobile robots thorough the three research, discusses new insights from being compared to the previous studies and limitations of our methodology, and finally shows the future direction of this thesis.

7.1 Summary of research findings

We implemented the gait-induced upper body motion of the wheeled mobile android during moving and verified the influence of human perception and behavior on the motions, aiming to apply robot's body expressions during moving to interact with humans. The previous research achieved emotional expressions in the walking control of a biped robot, and experimentally validated that viewers could understand the expressed emotions. In addition, research on mobile robots has been reported on the expression of the intention to give way by wheeled robots with retreating the body or rotating the torso, and contrary to ask to give way using the arms touch. However, the influence of gait, which is usually expressed continuously by humans while walking, have not been investigated whether available to wheeled mobile humanoid robots.

Thanks to the spread of advanced mechanical components such as high-performance motors and lightweight and high-strength materials, recently humanoid robots have actively been developed, which aim to safely interact with humans in our society. The android *ibuki* was developed based on the body dimensions of a 10-year-old Japanese boy. We adopted glass fiber reinforced nylon and carbon fiber reinforced plastic to achieve lightweight and high strength. In addition, the number of parts was reduced by using a 3D printer to print several parts in one piece, which also helped to reduce the machine size. These designs enabled us to use a low power but low noise motor with a plastic gear reducer. Previous advanced humanoids, Wabian-2R and HRP-4C consisted of mainly metal skeletons for bipedal walking. This precise metal mechanical structure enables to increase the joint output for its dynamic motion. However, the metal frame is too heavy and dangerous to be used in a human living environment where unexpected contacts between humans and robots might occur. Therefore, in practical use, their motions were often controlled conservatively at low speed. If without taking into account the constraint of a robot being a real human size, It is reported that a small size mobile humanoid AlterEgo has both of mobility and expressiveness in a human living environment [153]. As human-size wheeled humanoid robots, Pepper and Modle-H are designed with less metal and more plastic materials to operate in human living environments, but seeing the practical use, they do not have high power motion performance and high mobility.

Gait-induced upper body motion, which we proposed as a body expression for wheeled mobile robots during moving, is a walking-like motion that focuses on the characteristics of the vertical oscillation of the upper body caused by the extension and flexion of the legs during human walking. Specifically, a mechanism that drives the upper body vertically, named Vertical Oscillation Mechanism (VOM), is implemented in the mobility unit, and the upper body is oscillated twice by the VOM while the robot's wheel rotates for one walking cycle distance. The proposed VOM is easy to implement and control because its hardware design requirement is just to install one joint which drive the upper body linearly on the mobility unit. In particular, the motion was achieved as the following mechanical design of the VOM in this thesis. A slide screw is rotated by a motor to move the nut up and down, which is connected to the bottom of the upper body. A timing belt is wound by a pulley connected to a motor to move the upper link up and down, thus making the upper body of the android oscillate.

This thesis validated the influence of gait-induced upper body motion from two perspectives: perception and behavior. From the viewpoint of perception, firstly, we dealt with the perception of emotional expressions, and created an ambiguous facial expression of the android with a high entropy of valence, as cases where humans cannot perceive from facial expressions alone. As the result, the highest entropy was found for the Intense emotion with high intensity, as has been reported in previous human studies. In contrast, the facial valence tended to be clearly distinguished for expressions with low Intensity. Next, we investigated whether the facial valence of Intense expression could be clearly distinguished by adding body postures and movements. The results showed that viewers were able to distinguish the facial valence with specific combinations of neck, arm, and vertical motion. It was found that both body postures and body movements changed the facial valence of the intense emotion and contributed to the lower entropy in the valence assessments. In other words, postures and movements can improve the perception of the ambiguous facial expression.

Secondly, we measured three human gait motions (attributed to the emotions anger, happiness, and sadness) and implemented these findings as upper body motions in the mobile android. These motions under one condition include a vertical oscillation base on the movement of CoM induced when humans walk. The experiment result showed that the emotional expression of happiness using the VOM was well perceived by viewers (with higher recognition rates and higher confidence levels) and better than that of without vertical oscillation. In addition, the incongruent (switched vertical oscillations between anger and happiness emotions) application of anger/happiness vertical oscillation decreased the recognition rate of happiness/anger emotional expression (significance/marginally significant, respectively). Furthermore, the emotional expression of happiness using the vertical oscillation for anger decreased the confidence levels; however, the emotional expression of anger using the vertical oscillation for happiness increased the confidence levels (both were significance).

From the viewpoint of behavior, we confirmed that the gait phase of a person walking behind the android performing the gait-induced upper body motion is spontaneously synchronized with the gait phase expressed by the gait-induced upper body motion. In the experiment, we compared the gait phase synchronization between humans and the android, which moved its upper body in three distinct types of motion patterns: (1) no-motion, (2) arm-swinging (as is common for typical mobile humanoids), and (3) arms-swinging in addition to periodic vertical-oscillation similar to the human upper body movement while walking. We found that a significant distributional bias due to the phase locking by the gait synchronization was confirmed when participants were walking only with the robot controlled in the third motion pattern.

According to our experiment result, the phenomena that have been reported in human behavior studies were experimentally confirmed in wheeled androids. Research 1 confirmed that ambiguous facial expressions can be clearly distinguished by adding the body. Research 2 confirmed that emotions can be perceived from motions during moving. Research 3 confirmed that the gait phase spontaneously synchronizes with that of others. These findings have all been actively studied in human behavior study, but there are not much more examples of research showing the applications to robotics due to such as the difficulty in handling mobile humanoids as aforementioned. There are two factors why this thesis was able to reproduce the same phenomenon in the robot. The first factor is the robot's expressiveness. The android robot for this thesis has a human appearance, although its lower body mechanism is different from that of a human, participants seemed to naturally regard the android as if it is a human due to its upper body appearance. In fact, in Research 1, participants who saw the android postures and movements evaluated the human-likeness with high values. In addition to the appearance, the gait-induced upper body motion also contributed to help recognizing of the android as like a human. We believe that the bi-directional pursuit of human-like appearance and behavior has led to the reproduction of phenomena known from the human behavior study. The second factor is the high level of human recognition capability. Humans can communicate with other species using the body by nature. Furthermore, it is known that humans do not need as much information to recognize human motions. In the research of biological motion, it has been shown that in periodic body motions such as walking, humans are capable of estimating the gender, age, and mood of a walker by simply observing the movement of markers where major joint positions [143]. This is why that the human communication capability worked for not an animal but a humanoid robot which even if participants encountered for the first time.

7.2 Research limitations

The ratio of humans influenced by the gait-induced upper body motion is not 100 % for both perception and behavior, and it is not guaranteed this method works absolutely. Therefore, it is necessary to consider the failure recovery for the practical use when humans do not behave as intended by a robot. In the case of emotional expressions during moving (Research 2), anger had the lowest recognition rate of 56.0 %. In the synchronization of gait phase, the percentage of participants who synchronized with the robot's gait phase was 19.2 %. Therefore, not only improving the expressiveness of the robot, but we also need to design the next action which the robot should take when a human fails to recognize the robot intention. For emotional expressions during moving, the robot can additionally use a facial expression and robot-specific expressions (LED and utterance) in a situation where humans are approaching closer to the robot. We showed only around ten seconds of video in the experiment, but in reality, humans and robots interact continuously, so it would be better if the robot could express a single target emotion by giving various cues step by step. As for the gait synchronization, even if the robot can only synchronize with few numbers of humans around the robot (19.2 % in the experiment), we expect that the synchronization will spread to others from those who successfully synchronized with the robot.

The appearance of the robot is also essential for effective body expressions. Since the android *ibuki* had the human-like head, arms, and hands, viewers could easily recognize what the android's motion was expressing as like a human does, even though the lower body of the android was completely different from what a human has. To develop a robot with simple and cost-effective, it is necessary to verify the minimum factors of the robot's appearance which enable humans to recognize the robot's motion as the same way with that of a human. For example, a wheeled mobile robot with only anthropomorphic arms and no head might be one kind of practical designs for performing tasks in a human living environment, however it is not clear whether humans can recognize their gait-induced upper body motions during moving as the same way with our result. It would also be interesting to investigate an influence of adding a human-like upper body exterior and a VOM to actuate the upper body for a simple wheeled mobile robot originally without an upper body in original.

Although the sound was not the focus of this thesis, it is necessary to propose motors, gear reducers, and joint structures which cause less noise in order to interact with humans in a safe manner. Motor noise caused by dynamic motions might enforce discomfort and fear to the humans around. For example, one solution would be to actively adopt low noise mechanical elements (e.g. plastic material), however we have to pay an attention that this would result in lower output at the same time.

Although a robot's body expressions are expected to regulate unexpected human actions toward a robot and consequently reduce the burden of robot's perception and behavior, dynamic movements for body expressions consume a greater amount of energy. Therefore, the energy consumption devoted to whether the physical body expression as like this thesis and the perception and behavior of the robot self as like traditional robotics must be carefully designed, considering who can efficiently solve the problem by a human side or the robot own. Of course, one solution to this trade-off is to improve the efficiency of the robot's motor output.

For reference, here we compare the energy expenditure for the mobility between the android and human in Research 3. We found that participants walking behind *ibuki* synchronized the gait phase with that of expressing gait-induced upper body motion in AS+VO condition. When the gait phase synchronization occurred, the human gait cycle and angular frequency were locked to a constant value. On the other hand, participants did not synchronize their gait phase against *ibuki* without the gait-induced upper body motion in NM condition. This was because the participants adjusted the acceleration and deceleration of their gaits to maintain the safe distance from the slowly moving *ibuki* in front. We consider that the human side consumes energy with this gait adjustment instead of *ibuki*. For a clear comparison, we estimate the energy consumption for the five participants who well-synchronized with *ibuki* in higher PLI than 0.5 from all the 26 participants of Research 3.

The ratio of the walking energy in condition NM with respect to condition AS+VO can be estimated as Equation 7.2.1. We put $\omega_{NM,\,i}$ as the angular frequency of the i-th measured gait in condition NM and I as the representative inertia of the body joints, the kinetic energy which the walking person had was $\frac{1}{2}\ I\ \omega_{NM,\,i}{}^2$. Here we assume all the body joints move at this same angular velocity and neglect the difference of the translational kinetic energy between two conditions. What we are interested in now is the kinetic energy consumption for every gait cycle. When we assume that the participants consumed energy in the same way both when they increase or decrease the walking speed, the energy consumption from $i-1$-th gait cycle to i-th one was calculated as $\frac{1}{2}\ I\ |\omega_{NM,\,i}{}^2 - \omega_{NM,\,i-1}{}^2|$. The sum of the energy consumption obtained during the 120-second measurement in the experiment was taken as $\sum_i$. The same is true for AS+VO condition.

As a result of the calculation, the average η for the five participants was 1.51 (For the remaining 21 participants, the average η was 1.17.). We can say that a participant changed 1.51 times for the joint's kinetic energy when following the android under the NM condition than that of the AS+VO condition.

$$\eta = \frac{\sum_i \frac{1}{2}\ I\ |\omega_{NM,\,i}{}^2 - \omega_{NM,\,i-1}{}^2|}{\sum_i \frac{1}{2}\ I\ |\omega_{AS+VO,\,i}{}^2 - \omega_{AS+VO,\,i-1}{}^2|} \cdots\cdots (7.2.1)$$

According to Zarrugh's report [154], the energy consumption E ($\frac{cal}{min\ kg}$) of walking at velocity v ($\frac{m}{min}$) can be estimated by Equation 7.2.2. The participant walked at the same speed as *ibuki*, 0.83 m/s, and the total energy consumption by a person weighing 60 kg in two-minute experiment was about 11000 J. On the other hand, the total amount of energy consumed by ibuki's gait-induced upper body motion for two minutes was measured about 14000 J.

$$E = 32 + 0.0050v^2 \cdots\cdots (7.2.2)$$

It was found that the energy expenditure was the same order, which required by a human following *ibuki* whose upper body does not move during moving, and which required by *ibuki* performing the proposed gait-induced upper body motion. If there are a lot of humans walking behind a robot and they are expected to synchronize the robot's gait phase, the gait-induced upper body motion is beneficial because it requires less energy expenditure for humans. However, in order to successfully induce synchronization, the robot will need to expend additional energy on sensing, such as estimating a human gait by image recognition. On the other hand, if the number of humans is less than that of robots, robots do not need to move the body in the viewpoint of the energy consumption of the entire system, and have the human side pay that burden. Of course, it must cause an unavoidable safety risk when we force humans to do so.

7.3 Future directions

As the end of this thesis, we describe the future direction of this thesis. In navigation research on robotics so far, only a robot has controllability in an environment, and we tried to model the working environment as well as possible for higher observability. By improving the expressiveness of the robot's body during moving, it is possible to implicitly control not only the robot own but also the humans around. This viewpoint proposes a new control strategy which regards the entire environment surrounding the robot as a system having weak controllability. In order to successfully move not only the robot but also the surrounding humans, it is important to increase the controlability of human side. For this, the robot must convey the intention to humans as much as possible through the emotional empathy, gait synchronization, and so on.

We expect that humans and robots can adjust their distance from each other by expressing and perceiving the emotions for each other. In this way, a pedestrian flow can be formed even though biped humans and various kinds of wheeled robots are crowded in the same place. Furthermore, by synchronizing their gait phases, the pedestrian flow can be improved smoother and faster.

Lastly, We did not consider the cultural difference on the human perception and behavior since that was not the focus of this thesis. We consider that every cultural area has a difference transportation system and humans also perceive, behave, and formulate a pedestrian flow depended on their systems. It would be important to take cultural differences into account when implementing mobile robots into their societies.

Chapter 8

Conclusion

Chapter 1 introduced the overview of this thesis. It is still a big challenge to achieve robots moving freely in a human living environment. The existence of humans makes harder robots to move around. The key points were those:

- Robot navigation strategies in human living environment are mainly categorized into two approaches. One is related to path planning and self-localization necessary for a robot to move to its destination. The other one is regarding as a collision avoidance. This strategy is not directly related to the main robot goal of moving to the destination, however it is necessary for achieving safe mobility in our society.
- Considering practical usage, those approaches described above which the robots precisely perceive human behaviors and behave appropriately is just one direction of the solutions for the robot navigation at a human living environment. As the other direction, it is important that human sides recognize the robot's behavior and move cooperatively without getting in the way of the robot.
- Gait is a main factor to enable the pedestrian collaborative mobility in a society. Humans convey various information to the surrounding others from the body posture and movements while walking, and observers can perceive and understand meanings from their gait.

Therefore, it is worthwhile to study if wheeled robots, which are more common than legged robots, can express the upper body motion induced by the legged gait, and the application of the gait-like motion for mobile robots. This thesis aimed to propose that a wheeled android expresses an upper body motion induced by human gait during movement and to validate whether humans who see the body expressions can perceive the meanings or change their behavior.

Chapter 2 traced the development of mobile robots and humanoids which operate in human living environments. Those robots are expected to become alternatives of humans, which is still the ultimate goal since the term of robot was first used. The body expressions of robots have been widely investigated at the field of HRI, and one of the significant examples was the emotional gait by the bipedal humanoid. However, there was no research that implemented the human gait-induced body motion to the wheeled humanoid robot upper body motion during moving and validated its influence on human perception and behavior.

Chapter 3 describes the development of the wheeled child android *ibuki*. *ibuki* is 120 cm tall and is comprised of two parts, a mobility unit and the upper body, which is designed based on a Japanese boy. The face and hands are covered with silicone skin to have a human-like appearance. An electric motor drives each joint, and mobile batteries are used as the power supply. Next, we explained the gait-induced upper body motion, which we proposed as a body expression for wheeled mobile robots during moving, is a walking-like motion that focuses on the characteristics of the vertical oscillation of the upper body caused by the extension and flexion of the legs during human walking. Specifically, a mechanism that drives the upper body vertically, named Vertical Oscillation Mechanism (VOM), is implemented in the mobility unit, and the upper body is oscillated twice by the VOM while the robot's wheel rotates for one walking cycle distance.

Chapter 4 validated that an ambiguous facial expression of an android can be perceived more clearly with adding body postures and movements. The main contributions are following:

- We found that the facial valence distribution of the Intense emotion had the highest entropy with high intensity, as has been reported in previous human studies.
- We found that the facial valence of Intense could be clearly distinguished by adding body postures and movements. Viewers were able to distinguish the ambiguous facial valence with specific combinations of neck, arm, and vertical motion.

Chapter 5 showed that the implementation three types of emotional gait-induced upper body motions and investigated whether viewers could perceive the original emotion with high confidence. The main contributions are following:

- We measured three human gait motions (anger, happiness, and sadness) and implemented these findings as upper body motions in *ibuki*.
- The emotional expression of happiness using the VOM was better perceived by view-

ers (with higher recognition rates and higher confidence levels) than that of without vertical oscillation.

- The incongruent application of anger/happiness vertical oscillation decreased the recognition rate of happiness/anger emotional expression. Furthermore, the emotional expression of happiness using the vertical oscillation for anger decreased the confidence levels; however, the emotional expression of anger using the vertical oscillation for happiness increased the confidence levels.

Chapter 6 verified whether the gait phases of humans who follow the android which performed the periodic upper body motions while moving can be synchronized for that of the android. In the experiment, we compared the gait phase synchronization between humans and the android, which moved its upper body in three distinct types of motion patterns: (1) no-motion, (2) arm-swinging (as is common for typical mobile humanoids), and (3) arms-swinging in addition to periodic vertical-oscillation similar to the human upper body movement while walking. We found that a significant distributional bias due to the phase locking by the gait synchronization was confirmed when participants were walking only with the robot controlled in the third motion pattern.

Chapter 7 summarized the contributions of this thesis and the importance of body expressions for mobile robots. Research limitations were also described as below:

- The ratio of humans influenced by the gait-induced upper body motion is not 100 % for both perception and behavior, and it is not guaranteed this method works absolutely. Therefore, it is necessary to consider the failure recovery for the practical use when humans do not behave as intended by the robot.
- Dynamic movements of human-like body expressions consume a greater amount of energy. Therefore, the energy consumption devoted to whether the physical body expression or the robot perception and behavior must be carefully designed, considering who can efficiently solve the problem by a human side or the robot own.

We also compared the energy expenditure for the mobility between the android and human for reference. In the end, we described about the implicitly control for both of robots and humans. This new control strategy which regards the entire environment surrounding the robot as a system having weak controllability. In order to successfully move not only the robot but also the surrounding humans, it is important to increase the controlability of human side. For this, the robot must enable to convey the intention to humans through such as the emotional empathy and gait synchronization which we investigated in this thesis.

Acknowledgement

I am grateful for Professor Hiroshi Ishiguro to allow me to join in Ishiguro labolatory and to absorbed in this wonderful research opportunity. I would like to thank the advisors who supervised my research during the Ph.D. course, Associate Professor Yoshihiro Nakata and Associate Professor Yutaka Nakamura. I would also like to thank the members of Ishiguro labolatory. Especillay, Huthaifa Ahmad, Naoki Ise, Genki Jogan, and Ren Yamaki for technical assistance with my research. I am grateful for Humanware Innovation Program for Leading Graduate Schools of the Ministry of Education, Culture, Sports, Science and Technology, Japan. This research was supported by JSPS KAKENHI Grant Numbers JP19J20127, and JST ERATO JPMJER1401. Finally, I am sincely thankful to my wife Nami and my cutey pets Legop, Atari, Urachi, Estery, Nimy, and Charmy for supporting my daily research life.

Bibliography

1) Mordor Intelligence, "Global robotics market - growth, trends, covid-19 impact, and forecasts (2021 - 2026)." `https://www.mordorintelligence.com/industry-reports/robotics-market`. Last accessed 25 November 2021.
2) Y. Morales, T. Kanda, and N. Hagita, "Walking together: Side-by-side walking model for an interacting robot," *Journal of Human-Robot Interaction*, vol. 3, no. 2, pp. 50–73, 2014.
3) E. Repiso, A. Garrell, and A. Sanfeliu, "Adaptive side-by-side social robot navigation to approach and interact with people," *International Journal of Social Robotics*, vol. 12, no. 4, pp. 909–930, 2020.
4) R. Mead and M. J. Matarić, "Autonomous human–robot proxemics: socially aware navigation based on interaction potential," *Autonomous Robots*, vol. 41, no. 5, pp. 1189–1201, 2017.
5) Y. F. Chen, M. Everett, M. Liu, and J. P. How, "Socially aware motion planning with deep reinforcement learning," 2018.
6) J. Wu, H. Woo, Y. Tamura, A. Moro, S. Massaroli, A. Yamashita, and H. Asama, "Pedestrian trajectory prediction using birnn encoder–decoder framework," *Advanced Robotics*, vol. 33, no. 18, pp. 956–969, 2019.
7) A. D. May, C. Dondrup, and M. Hanheide, "Show me your moves! conveying navigation intention of a mobile robot to humans," in *2015 European Conference on Mobile Robots (ECMR)*, pp. 1–6, IEEE, 2015.
8) N. J. Hetherington, E. A. Croft, and H. M. Van der Loos, "Hey robot, which way are you going? nonverbal motion legibility cues for human-robot spatial interaction," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5010–5015, 2021.
9) A. M. Rosenthal-von der Pütten, N. C. Krämer, and J. Herrmann, "The effects of humanlike and robot-specific affective nonverbal behavior on perception, emotion, and behavior," *International Journal of Social Robotics*, vol. 10, no. 5, pp. 569–582, 2018.

10) M. Moussaïd, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz, "The walking behaviour of pedestrian social groups and its impact on crowd dynamics," *PloS one*, vol. 5, no. 4, p. e10047, 2010.
11) A. Frohnwieser, R. Hopf, E. Oberzaucher, *et al.*, "Human walking behavior: the effect of pedestrian flow and personal space invasions on walking speed and direction," *Human Ethology Bulletin*, vol. 28, no. 3, pp. 20–28, 2013.
12) D. Helbing and P. Molnar, "Social force model for pedestrian dynamics," *Physical review E*, vol. 51, no. 5, p. 4282, 1995.
13) J. H. Park, F. A. Rojas, and H. S. Yang, "A collision avoidance behavior model for crowd simulation based on psychological findings," *Computer Animation and Virtual Worlds*, vol. 24, pp. 173–183, May 2013.
14) H. Khambhaita, J. Rios-Martinez, and R. Alami, "Head-body motion coordination for human aware robot navigation," in *9th International workshop on Human-Friendlly Robotics (HFR 2016)*, p. 8p, 2016.
15) F. Farina, D. Fontanelli, A. Garulli, A. Giannitrapani, and D. Prattichizzo, "Walking ahead: The headed social force model," *PLOS ONE*, vol. 12, p. e0169734, 01 2017.
16) K. Kitazawa and T. Fujiyama, "Pedestrian vision and collision avoidance behavior: Investigation of the information process space of pedestrians using an eye tracker," in *Pedestrian and evacuation dynamics 2008*, pp. 95–108, Springer, 2010.
17) Y. Makihara, H. Mannami, and Y. Yagi, "Gait analysis of gender and age using a large-scale multi-view gait database," pp. 440–451, 11 2010.
18) B. Nigg, V. Fisher, and J. Ronsky, "Gait characteristics as a function of age and gender," *Gait & posture*, vol. 2, no. 4, pp. 213–220, 1994.
19) T. Öberg, A. Karsznia, and K. Öberg, "Basic gait parameters: reference data for normal subjects, 10-79 years of age," *Journal of rehabilitation research and development*, vol. 30, pp. 210–223, 1993.
20) G. Venture, H. Kadone, T. Zhang, J. Grézes, A. Berthoz, and H. Hicheur, "Recognizing emotions conveyed by human gait," *International Journal of Social Robotics*, vol. 6, pp. 621–632, Nov. 2014.
21) C. L. Roether, L. Omlor, A. Christensen, and M. A. Giese, "Critical features for the perception of emotion from gait," *Journal of vision*, vol. 9, pp. 1–32, June 2009.
22) J. Michalak, N. Troje, J. Fischer, P. Vollmar, T. Heidenreich, and D. Schulte, "Embodiment of sadness and depression–gait patterns associated with dysphoric mood," *Psychosomatic Medicine*, vol. 71, pp. 580–587, June 2009.

23) J. M. Montepare, S. B. Goldstein, and A. Clausen, "The identification of emotions from gait information," *Journal of Nonverbal Behavior*, vol. 11, pp. 33–42, Mar. 1987.
24) M. Destephe, T. Maruyama, M. Zecca, K. Hashimoto, and A. Takanishi, "The influences of emotional intensity for happiness and sadness on walking," in *Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (Osaka, Japan), pp. 7452–7455, IEEE, July 2013.
25) Y. Nakata, S. Yagi, S. Yu, Y. Wang, N. Ise, Y. Nakamura, and H. Ishiguro, "Development of 'ibuki' an electrically actuated childlike android with mobility and its potential in the future society," *Robotica*, p. 1–18, 2021.
26) C. Bartneck and J. Reichenbach, "Subtle emotional expressions of synthetic characters," *International journal of human-computer studies*, vol. 62, no. 2, pp. 179–192, 2005.
27) H. Aviezer, Y. Trope, and A. Todorov, "Body cues, not facial expressions, discriminate between intense positive and negative emotions," *Science*, vol. 338, pp. 1225–1229, Nov. 2012.
28) M. M. Gross, E. A. Crane, and B. L. Fredrickson, "Effort-shape and kinematic assessment of bodily expression of emotion during gait," *Human movement science*, vol. 31, no. 1, pp. 202–221, 2012.
29) K. Miura, M. Morisawa, F. Kanehiro, S. Kajita, K. Kaneko, and K. Yokoi, "Human-like walking with toe supporting for humanoids," in *Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, (San Francisco, CA, USA), pp. 4428–4435, IEEE, Sept. 2011.
30) M. Destephe, M. Brandao, T. Kishi, M. Zecca, K. Hashimoto, and A. Takanishi, "Emotional gait: effects on humans' perception of humanoid robots," in *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication*, (Edinburgh, UK), pp. 261–266, IEEE, Aug. 2014.
31) T. Izui, I. Milleville, S. Sakka, and G. Venture, "Expressing emotions using gait of humanoid robot," in *Proceedings of the 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, (Kobe, Japan), pp. 241–245, IEEE, Aug.-Sept. 2015.
32) N. R. van Ulzen, C. J. Lamoth, A. Daffertshofer, G. R. Semin, and P. J. Beek, "Characteristics of instructed and uninstructed interpersonal coordination while walking side-by-side," *Neuroscience letters*, vol. 432, no. 2, pp. 88–93, 2008.
33) A. Z. Zivotofsky, L. Gruendlinger, and J. M. Hausdorff, "Modality-specific commu-

nication enabling gait synchronization during over-ground side-by-side walking," *Human movement science*, vol. 31, no. 5, pp. 1268–1285, 2012.

34) C. Chambers, G. Kong, K. Wei, and K. Kording, "Pose estimates from online videos show that side-by-side walkers synchronize movement under naturalistic conditions," *PloS one*, vol. 14, no. 6, p. e0217861, 2019.
35) A. Jelić, C. Appert-Rolland, S. Lemercier, and J. Pettré, "Properties of pedestrians walking in line: Fundamental diagrams," *Physical review E*, vol. 85, no. 3, p. 036111, 2012.
36) Y. Ma, E. W. M. Lee, M. Shi, and R. K. K. Yuen, "Spontaneous synchronization of motion in pedestrian crowds of different densities," *Nature human behaviour*, vol. 5, no. 4, pp. 447–457, 2021.
37) K. Capek, *RUR (Rossum's universal robots).* Penguin Classics, 2004.
38) A. Rosendo, S. Nakatsu, K. Narioka, and K. Hosoda, "Pneupard: A biomimetic musculoskeletal approach for a feline-inspired quadruped robot," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1452–1457, IEEE, 2013.
39) Y. Masuda, K. Naniwa, M. Ishikawa, and K. Osuka, "On brainless-control approach to soft bodies: a novel method to generate motion patterns by pneumatic reflex devices," in *21th IFAC World Congress*, 2020.
40) S. H. Collins, M. Wisse, and A. Ruina, "A three-dimensional passive-dynamic walking robot with two legs and knees," *The International Journal of Robotics Research*, vol. 20, no. 7, pp. 607–615, 2001.
41) N. J. Nilsson, "A mobile automaton: An application of artificial intelligence techniques," tech. rep., Sri International Menlo Park Ca Artificial Intelligence Center, 1969.
42) Waseda University, "Development of waseda robot." `https://www.humanoid.waseda.ac.jp/booklet/kato_2.html`. Last accessed 1 December 2021.
43) Sony, "Aibos history." `https://www.sony-aibo.com/aibos-history/`. Last accessed 1 December 2021.
44) S. Shigemi, A. Goswami, and P. Vadakkepat, "Asimo and humanoid robot research at honda," *Humanoid robotics: A reference*, pp. 55–90, 2019.
45) B. M. Yamauchi, "Packbot: a versatile platform for military robotics," in *Unmanned ground vehicle technology VI*, vol. 5422, pp. 228–237, International Society for Optics and Photonics, 2004.
46) T. Yoshida, K. Nagatani, S. Tadokoro, T. Nishimura, and E. Koyanagi, "Improvements to the rescue robot quince toward future indoor surveillance missions in the fukushima

daiichi nuclear power plant," in *Field and service robotics*, pp. 19–32, Springer, 2014.

47) E. Krotkov, D. Hackett, L. Jackel, M. Perschbacher, J. Pippine, J. Strauss, G. Pratt, and C. Orlowski, "The darpa robotics challenge finals: Results and perspectives," *Journal of Field Robotics*, vol. 34, no. 2, pp. 229–240, 2017.
48) The Disney wiki, "Audio-animatronics." `https://disney.fandom.com/wiki/Audio-Animatronics`. Last accessed 1 December 2021.
49) M. K. Pan, S. Choi, J. Kennedy, K. McIntosh, D. C. Zamora, G. Niemeyer, J. Kim, A. Wieland, and D. Christensen, "Realistic and interactive robot gaze," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 11072–11078, IEEE, 2020.
50) Hanson Robotics, "Sophia." `https://www.hansonrobotics.com/sophia/`. Last accessed 1 December 2021.
51) A. K. Pandey and R. Gelin, "A mass-produced sociable humanoid robot: Pepper: The first machine of its kind," *IEEE Robotics & Automation Magazine*, vol. 25, no. 3, pp. 40–48, 2018.
52) Telexistence, "Telexistence Model-T & Model-H." `https://tx-inc.com/en/technology/`. Last accessed 1 December 2021.
53) Ory laboratory, "Orihime." `https://orylab.com/en/`. Last accessed 1 December 2021.
54) K. Takeuchi, Y. Yamazaki, and K. Yoshifuji, "Avatar work: Telework for disabled people unable to go outside by using avatar robots," in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 53–60, 2020.
55) M. Häring, N. Bee, and E. André, "Creation and evaluation of emotion expression with body movement, sound and eye color for humanoid robots," in *2011 RO-MAN*, pp. 204–209, 2011.
56) M. Zecca, Y. Mizoguchi, K. Endo, F. Iida, Y. Kawabata, N. Endo, K. Itoh, and A. Takanishi, "Whole body emotion expressions for kobian humanoid robot—preliminary experiments with different emotional patterns—," in *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 381–386, IEEE, 2009.
57) M. Zecca, G. Macrì, Y. Mizoguchi, V. Monaco, N. Endo, K. Itoh, P. Dario, and A. Takanishi, "Evaluation of the kobian and habian emotion expression humanoid robots with european elderly people," in *ROMANSY 18 Robot Design, Dynamics and Control*, pp. 449–456, Springer, 2010.

58) Y. Okuno, T. Kanda, M. Imai, H. Ishiguro, and N. Hagita, "Providing route directions: design of robot's utterance, gesture, and timing," in *Proceedings of the 2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, (La Jolla, CA, USA), pp. 53–60, IEEE, Mar. 2009.
59) M. Huber, M. Rickert, A. Knoll, T. Brandt, and S. Glasauer, "Human-robot interaction in handing-over tasks," in *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 107–112, IEEE, 2008.
60) M. K. Pan, E. Knoop, M. Bächer, and G. Niemeyer, "Fast handovers with a robot character: Small sensorimotor delays improve perceived qualities," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 6735–6741, IEEE, 2019.
61) D. F. P. Granados and K. Kosuge, "Design of a male-type dance partner robot for leading a physical human-robot interaction," in *2015 IEEE International Conference on Mechatronics and Automation (ICMA)*, (Beijing, China), pp. 1234–1240, IEEE, Sept. 2015.
62) D. F. P. Granados, B. A. Yamamoto, H. Kamide, J. Kinugawa, and K. Kosuge, "Dance teaching by a robot: Combining cognitive and physical human–robot interaction for supporting the skill learning process," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1452–1459, 2017.
63) M. Destephe, *Analysis, Modeling and Application of Emotional Gait to Humanoid Robots*. Doctoral dissertation, Waseda University Repository, Japan, 2015. (Accession No. 6930).
64) M. Kamezaki, A. Kobayashi, Y. Yokoyama, H. Yanagawa, M. Shrestha, and S. Sugano, "A preliminary study of interactive navigation framework with situation-adaptive multimodal inducement: pass-by scenario," *International Journal of Social Robotics*, vol. 12, no. 2, pp. 567–588, 2020.
65) J. Reinhardt, L. Prasch, and K. Bengler, "Back-off: Evaluation of robot motion strategies to facilitate human-robot spatial interaction," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 10, no. 3, pp. 1–25, 2021.
66) E. Senft, S. Satake, and T. Kanda, "Would you mind me if I pass by you? socially-appropriate behaviour for an omni-based social robot in narrow environment," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 539–547, 2020.
67) C. Tsiourti, A. Weiss, K. Wac, and M. Vincze, "Designing emotionally expressive

robots: a comparative study on the perception of communication modalities," in *Proceedings of the 5th international conference on human agent interaction*, pp. 213–222, 2017.

68) H. Ishiguro, "Scientific issues concerning androids," *The International Journal of Robotics Research*, vol. 26, no. 1, pp. 105–117, 2007.
69) S. Park, H. Lee, D. Hanson, and P. Y. Oh, "Sophia-hubo's arm motion generation for a handshake and gestures," in *2018 15th international conference on ubiquitous robots (UR)*, pp. 511–515, IEEE, 2018.
70) K. Miura, S. Nakaoka, S. Kajita, K. Kaneko, F. Kanehiro, M. Morisawa, and K. Yokoi, "Trials of cybernetic human hrp-4c toward humanoid business," in *2010 IEEE Workshop on Advanced Robotics and its Social Impacts*, pp. 165–169, IEEE, 2010.
71) P. Milhorat, D. Lala, K. Inoue, T. Zhao, M. Ishida, K. Takanashi, S. Nakamura, and T. Kawahara, "A conversational dialogue manager for the humanoid robot erica," in *Advanced Social Interaction with Agents*, pp. 119–131, Springer, 2019.
72) M. Watanabe, K. Ogawa, and H. Ishiguro, "Can androids be salespeople in the real world?," in *Proceedings of the 33rd annual ACM conference extended abstracts on human factors in computing systems*, pp. 781–788, 2015.
73) K. Kaneko, F. Kanehiro, M. Morisawa, K. Miura, S. Nakaoka, and S. Kajita, "Cybernetic human hrp-4c," in *2009 9th IEEE-RAS International Conference on Humanoid Robots*, pp. 7–14, IEEE, 2009.
74) H. S. Ahn, D. Lee, D. Choi, D.-Y. Lee, M. Hur, and H. Lee, "Appropriate emotions for facial expressions of 33-dofs android head ever-4 h33," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, pp. 1115–1120, IEEE, 2012.
75) Research Institute of Human Engineering for Quality Life, "Construction of a database of physical characteristics of children to improve the safety of machine products and research on visualization simulation technology of human body damage." `https://www.hql.jp/hql/wp/wp-content/uploads/2018/01/children_data2008.pdf`. Last accessed 19 December 2021.
76) W. Zijlstra and A. L. Hof, "Displacement of the pelvis during human walking: experimental data and model predictions," *Gait & posture*, vol. 6, no. 3, pp. 249–262, 1997.
77) J. Carpentier, M. Benallegue, and J.-P. Laumond, "On the centre of mass motion in human walking," *International Journal of Automation and Computing*, vol. 14, no. 5,

pp. 542–551, 2017.

78) M. S. Orendurff, A. D. Segal, G. K. Klute, J. S. Berge, E. S. Rohr, and N. J. Kadel, "The effect of walking speed on center of mass displacement.," *Journal of Rehabilitation Research & Development*, vol. 41, no. 6, 2004.
79) R. B. Adams Jr and R. E. Kleck, "Effects of direct and averted gaze on the perception of facially communicated emotion.," *Emotion*, vol. 5, no. 1, p. 3, 2005.
80) B. Mutlu, T. Kanda, J. Forlizzi, J. Hodgins, and H. Ishiguro, "Conversational gaze mechanisms for humanlike robots," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 1, no. 2, pp. 1–33, 2012.
81) S. Sheikholeslami, A. Moon, and E. A. Croft, "Cooperative gestures for industry: Exploring the efficacy of robot hand configurations in expression of instructional gestures for human–robot interaction," *The International Journal of Robotics Research*, vol. 36, no. 5-7, pp. 699–720, 2017.
82) C. Thepsoonthorn, K. Ogawa, and Y. Miyake, "The relationship between robot's nonverbal behaviour and human's likability based on human's personality," *Scientific reports*, vol. 8, no. 1, pp. 1–11, 2018.
83) C. Bidet-Ildei, A. Decatoire, and S. Gil, "Recognition of emotions from facial point-light displays," *Frontiers in Psychology*, vol. 11, p. 1062, 2020.
84) A. P. Atkinson, W. H. Dittrich, A. J. Gemmell, and A. W. Young, "Emotion perception from dynamic and static body expressions in point-light and full-light displays," *Perception*, vol. 33, no. 6, pp. 717–746, 2004.
85) G. Trovato, T. Kishi, N. Endo, M. Zecca, K. Hashimoto, and A. Takanishi, "Cross-cultural perspectives on emotion expressive humanoid robotic head: recognition of facial expressions and symbols," *International Journal of Social Robotics*, vol. 5, no. 4, pp. 515–527, 2013.
86) M. S. Erden, "Emotional postures for the humanoid-robot nao," *International Journal of Social Robotics*, vol. 5, no. 4, pp. 441–456, 2013.
87) A. Deshmukh, B. Craenen, A. Vinciarelli, and M. E. Foster, "Shaping robot gestures to shape users' perception: The effect of amplitude and speed on godspeed ratings," in *Proceedings of the 6th International Conference on Human-Agent Interaction*, pp. 293–300, 2018.
88) J. Kedzierski, R. Muszyński, C. Zoll, A. Oleksy, and M. Frontkiewicz, "Emys—emotive head of a social robot," *International Journal of Social Robotics*, vol. 5, no. 2, pp. 237–249, 2013.

89) C. Tsiourti, A. Weiss, K. Wac, and M. Vincze, "Multimodal integration of emotional signals from voice, body, and context: effects of (in) congruence on emotion recognition and attitudes towards robots," *International Journal of Social Robotics*, vol. 11, no. 4, pp. 555–573, 2019.
90) C. Darwin and P. Prodger, *The expression of the emotions in man and animals*. Oxford University Press, USA, 1998.
91) P. Ekman, W. V. Friesen, M. O'sullivan, A. Chan, I. Diacoyanni-Tarlatzis, K. Heider, R. Krause, W. A. LeCompte, T. Pitcairn, P. E. Ricci-Bitti, *et al.*, "Universals and cultural differences in the judgments of facial expressions of emotion.," *Journal of personality and social psychology*, vol. 53, no. 4, p. 712, 1987.
92) J. A. Russell, "A circumplex model of affect.," *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.
93) C. Breazeal, "Emotion and sociable humanoid robots," *International journal of human-computer studies*, vol. 59, no. 1-2, pp. 119–155, 2003.
94) I. Lütkebohle, F. Hegel, S. Schulz, M. Hackel, B. Wrede, S. Wachsmuth, and G. Sagerer, "The bielefeld anthropomorphic robot head "flobi"," in *2010 IEEE International Conference on Robotics and Automation*, pp. 3384–3391, IEEE, 2010.
95) D. Bazo, R. Vaidyanathan, A. Lentz, and C. Melhuish, "Design and testing of a hybrid expressive face for a humanoid robot," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5317–5322, IEEE, 2010.
96) N. Churamani, F. Cruz, S. Griffiths, and P. Barros, "icub: learning emotion expressions using human reward," *arXiv preprint arXiv:2003.13483*, 2020.
97) H. Ishiguro, "Android science," in *Robotics Research*, pp. 118–127, Springer, 2007.
98) C. Becker-Asano and H. Ishiguro, "Evaluating facial displays of emotion for the android robot geminoid f," in *2011 IEEE workshop on affective computational intelligence (WACI)*, pp. 1–8, IEEE, 2011.
99) H. K. Meeren, C. C. van Heijnsbergen, and B. de Gelder, "Rapid perceptual integration of facial expression and emotional body language," *Proceedings of the National Academy of Sciences*, vol. 102, no. 45, pp. 16518–16523, 2005.
100) J. Van den Stock, R. Righart, and B. De Gelder, "Body expressions influence recognition of emotions in the face and voice.," *Emotion*, vol. 7, no. 3, p. 487, 2007.
101) Z. Witkower and J. L. Tracy, "Bodily communication of emotion: evidence for extrafacial behavioral expressions and available coding systems," *Emotion Review*, vol. 11, no. 2, pp. 184–193, 2019.

102) M. De Meijer, "The contribution of general features of body movement to the attribution of emotions," *Journal of Nonverbal behavior*, vol. 13, no. 4, pp. 247–268, 1989.
103) M. M. Gross, E. A. Crane, and B. L. Fredrickson, "Methodology for assessing bodily expression of emotion," *Journal of Nonverbal Behavior*, vol. 34, no. 4, pp. 223–248, 2010.
104) A. Mignault and A. Chaudhuri, "The many faces of a neutral face: Head tilt and perception of dominance and emotion," *Journal of nonverbal behavior*, vol. 27, no. 2, pp. 111–132, 2003.
105) T. Shafir, R. P. Tsachor, and K. B. Welch, "Emotion regulation through movement: unique sets of movement characteristics are associated with and enhance basic emotions," *Frontiers in psychology*, vol. 6, p. 2030, 2016.
106) A. P. Saygin, T. Chaminade, H. Ishiguro, J. Driver, and C. Frith, "The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions," *Social cognitive and affective neuroscience*, vol. 7, no. 4, pp. 413–422, 2012.
107) D. Glowinski, N. Dael, A. Camurri, G. Volpe, M. Mortillaro, and K. Scherer, "Toward a minimal representation of affective gestures," *IEEE Transactions on Affective Computing*, vol. 2, no. 2, pp. 106–118, 2011.
108) J. S. Winston, B. A. Strange, J. O'Doherty, and R. J. Dolan, "Automatic and intentional brain responses during evaluation of trustworthiness of faces," *Nature neuroscience*, vol. 5, no. 3, pp. 277–283, 2002.
109) S. Hamann, "Cognitive and neural mechanisms of emotional memory," *Trends in cognitive sciences*, vol. 5, no. 9, pp. 394–400, 2001.
110) S. Droit-Volet and W. H. Meck, "How emotions colour our perception of time," *Trends in cognitive sciences*, vol. 11, no. 12, pp. 504–513, 2007.
111) M. Tielman, M. Neerincx, J. Meyer, and R. Looije, "Adaptive emotional expression in robot-child interaction," in *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 407–414, IEEE, 2014.
112) O. Houstis and S. Kiliaridis, "Gender and age differences in facial expressions," *The European Journal of Orthodontics*, vol. 31, no. 5, pp. 459–466, 2009.
113) R. E. Jack, O. G. Garrod, H. Yu, R. Caldara, and P. G. Schyns, "Facial expressions of emotion are not culturally universal," *Proceedings of the National Academy of Sciences*, vol. 109, no. 19, pp. 7241–7244, 2012.
114) A. A. Sokolov, P. Zeidman, M. Erb, F. E. Pollick, A. J. Fallgatter, P. Ryvlin, K. J. Fris-

ton, and M. A. Pavlova, “Brain circuits signaling the absence of emotion in body language,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 34, pp. 20868–20873, 2020.

115) M. Karg, A. Samadani, R. Gorbet, K. Kühnlenz, J. Hoey, and D. Kulić, “Body movements for affective expression: A survey of automatic recognition and generation,” *IEEE Transactions on Affective Computing*, vol. 4, pp. 341–359, Nov. 2013.

116) M. Karg, K. Kuhnlenz, and M. Buss, “Recognition of affect based on gait patterns,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, pp. 1050–1061, Aug. 2010.

117) A. Parmiggiani, L. Fiorio, A. Scalzo, A. V. Sureshbabu, M. Randazzo, M. Maggiali, U. Pattacini, H. Lehmann, V. Tikhanoff, D. Domenichelli, A. Cardellino, P. Congiu, A. Pagnin, R. Cingolani, L. Natale, and G. Metta, “The design and validation of the r1 personal humanoid,” in *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (Vancouver, BC, Canada), pp. 674–680, IEEE, Sept. 2017.

118) T. Nakata, T. Sato, T. Mori, and H. Mizoguchi, “Expression of emotion and intention by robot body movement,” in *Proceedings of the 5th international conference on autonomous systems*, pp. 352–359, IOS Press, June 1998.

119) Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh, “Openpose: Realtime multi-person 2d pose estimation using part affinity fields,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

120) S. Yagi, N. Ise, S. Yu, Y. Nakata, Y. Nakamura, and H. Ishiguro, “Perception of emotional gait-like motion of mobile humanoid robot using vertical oscillation,” HRI ’20, (New York, NY, USA), p. 529–531, Association for Computing Machinery, 2020.

121) M. Destephe, A. Henning, M. Zecca, K. Hashimoto, and A. Takanishi, “Perception of emotion and emotional intensity in humanoid robots gait,” in *2013 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1276–1281, IEEE, 2013.

122) H. Ikeda and K. Watanabe, “Anger and happiness are linked differently to the explicit detection of biological motion,” *Perception*, vol. 38, pp. 1002 – 1011, 2009.

123) T. Otani, A. Iizuka, D. Takamoto, H. Motohashi, T. Kishi, P. Kryczka, N. Endo, L. Jamone, K. Hashimoto, T. Takashima, *et al.*, “New shank mechanism for humanoid robot mimicking human-like walking in horizontal and frontal plane,” in *2013 IEEE International Conference on Robotics and Automation*, pp. 667–672, IEEE, 2013.

124) S. Dalibard, A. El Khoury, F. Lamiraux, A. Nakhaei, M. Taïx, and J.-P. Laumond,

"Dynamic walking and whole-body motion planning for humanoid robots: an integrated approach," *The International Journal of Robotics Research*, vol. 32, no. 9-10, pp. 1089–1103, 2013.

125) A. Pandey, S. Pandey, and D. Parhi, "Mobile robot navigation and obstacle avoidance techniques: A review," *Int Rob Auto J*, vol. 2, no. 3, p. 00022, 2017.

126) S. G. Tzafestas, "Mobile robot control and navigation: A global overview," *Journal of Intelligent & Robotic Systems*, vol. 91, no. 1, pp. 35–58, 2018.

127) V. V. Unhelkar, C. Pérez-D'Arpino, L. Stirling, and J. A. Shah, "Human-robot co-navigation using anticipatory indicators of human walking motion," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6183–6190, IEEE, 2015.

128) H. S. Shanee, K. Dror, O. Tal, and E. Yael, "The influence of following angle on performance metrics of a human-following robot," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 593–598, IEEE, 2016.

129) A. Ravankar, A. A. Ravankar, Y. Hoshino, M. Watanabe, and Y. Kobayashi, "Safe mobile robot navigation in human-centered environments using a heat map-based path planner," *Artificial Life and Robotics*, vol. 25, no. 2, pp. 264–272, 2020.

130) E. Repiso, A. Garrell, and A. Sanfeliu, "People's adaptive side-by-side model evolved to accompany groups of people by social robots," *IEEE Robotics and Automation Letters*, vol. 5, pp. 2387–2394, 2020.

131) F. W. Siebert, J. Klein, M. Rötting, and E. Roesler, "The influence of distance and lateral offset of follow me robots on user perception," *Frontiers in Robotics and AI*, vol. 7, p. 74, 2020.

132) S. Sekiguchi, A. Yorozu, K. Kuno, M. Okada, Y. Watanabe, and M. Takahashi, "Human-friendly control system design for two-wheeled service robot with optimal control approach," *Robotics and Autonomous Systems*, vol. 131, p. 103562, 2020.

133) T. Matsumaru, "Experimental examination in simulated interactive situation between people and mobile robot with preliminary-announcement and indication function of upcoming operation," in *2008 IEEE International Conference on Robotics and Automation*, pp. 3487–3494, IEEE, 2008.

134) S. Yamashita, T. Kurihara, T. Ikeda, K. Shinozawa, and S. Iwaki, "Evaluation of robots that signals a pedestrian using face orientation based on analysis of velocity vector fluctuation in moving trajectories," *Advanced Robotics*, vol. 34, no. 20, pp. 1309–1323,

2020.

135) B. Zhang, J. Amirian, H. Eberle, J. Pettré, C. Holloway, and T. Carlson, "From hri to cri: Crowd robot interaction—understanding the effect of robots on crowd motion," *International Journal of Social Robotics*, pp. 1–13, 2021.

136) C. Vassallo, A. Olivier, P. Souères, A. Crétual, O. Stasse, and J. Pettré, "How do walkers behave when crossing the way of a mobile robot that replicates human interaction rules?," *Gait & posture*, vol. 60, pp. 188–193, 2018.

137) F. Sylos-Labini, A. d'Avella, F. Lacquaniti, and Y. Ivanenko, "Human-human interaction forces and interlimb coordination during side-by-side walking with hand contact," *Frontiers in physiology*, vol. 9, p. 179, 2018.

138) S. H. Strogatz, D. M. Abrams, A. McRobie, B. Eckhardt, and E. Ott, "Crowd synchrony on the millennium bridge," *Nature*, vol. 438, no. 7064, pp. 43–44, 2005.

139) M. Leman, D. Moelants, M. Varewyck, F. Styns, L. van Noorden, and J.-P. Martens, "Activating and relaxing music entrains the speed of beat synchronized walking," *PloS one*, vol. 8, no. 7, p. e67932, 2013.

140) A. Z. Zivotofsky, H. Bernad-Elazari, P. Grossman, and J. M. Hausdorff, "The effects of dual tasking on gait synchronization during over-ground side-by-side walking," *Human Movement Science*, vol. 59, pp. 20–29, 2018.

141) M. Cheng, M. Kato, J. A. Saunders, and C. Tseng, "Paired walkers with better first impression synchronize better," *PLoS one*, vol. 15, no. 2, p. e0227880, 2020.

142) C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C. Chang, M. G. Yong, J. Lee, *et al.*, "Mediapipe: A framework for building perception pipelines," *arXiv preprint arXiv:1906.08172*, 2019.

143) N. F. Troje, "Decomposing biological motion: A framework for analysis and synthesis of human gait patterns," *Journal of vision*, vol. 2, no. 5, pp. 2–2, 2002.

144) D. Yanagisawa, A. Tomoeda, and K. Nishinari, "Improvement of pedestrian flow by slow rhythm," *Physical Review E*, vol. 85, no. 1, p. 016111, 2012.

145) Y. Akiyama, S. Okamoto, H. Toda, T. Ogura, and Y. Yamada, "Gait motion for naturally curving variously shaped corners," *Advanced Robotics*, vol. 32, no. 2, pp. 77–88, 2018.

146) J. Rose and J. G. Gamble, *Human walking*, vol. 3. Williams & Wilkins Baltimore, 1994.

147) J. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, "Measuring phase synchrony in brain signals," *Human brain mapping*, vol. 8, no. 4, pp. 194–208, 1999.

148) B. Schack and S. Weiss, “Quantification of phase synchronization phenomena and their importance for verbal memory processes,” *Biological cybernetics*, vol. 92, no. 4, pp. 275–287, 2005.

149) A. Chang, S. R. Livingstone, D. J. Bosnyak, and L. J. Trainor, “Body sway reflects leadership in joint music performance,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 21, pp. E4134–E4141, 2017.

150) A. Chang, H. E. Kragness, W. Tsou, D. J. Bosnyak, A. Thiede, and L. J. Trainor, “Body sway predicts romantic interest in speed dating,” *Social Cognitive and Affective Neuroscience*, vol. 16, no. 1-2, pp. 185–192, 2021.

151) C. Scheer, L. Horn, and P. Jansen, “Moving in synchrony with an avatar–presenting a novel and unbiased body sway synchronization paradigm,” *Current Psychology*, pp. 1–7, 2021.

152) A. Pikovsky, M. Rosenblum, J. Kurths, and A. Synchronization, “A universal concept in nonlinear sciences,” *Self*, vol. 2, p. 3, 2001.

153) G. Lentini, A. Settimi, D. Caporale, M. Garabini, G. Grioli, L. Pallottino, M. G. Catalano, and A. Bicchi, “Alter-ego: a mobile robot with a functionally anthropomorphic upper body designed for physical interaction,” *IEEE Robotics & Automation Magazine*, vol. 26, no. 4, pp. 94–107, 2019.

154) M. Zarrugh, F. Todd, and H. Ralston, “Optimization of energy expenditure during level walking,” *European journal of applied physiology and occupational physiology*, vol. 33, no. 4, pp. 293–306, 1974.

Publications

Journal Papers

1) Satoshi Yagi, Yoshihiro Nakata, Yutaka Nakamura, and Hiroshi Ishiguro, “Spontaneous Gait Phase Synchronization of Human to a Wheeled Mobile Robot with Replicating Gait-induced Upper Body Oscillating Motion,” [Article under review].
2) Satoshi Yagi, Yoshihiro Nakata, Yutaka Nakamura, and Hiroshi Ishiguro, “Perception of Emotional Expression of Mobile Humanoid Robot Using Gait-Induced Upper Body Motion,” IEEE Access, vol. 9, pp. 124793–124804, 2021.
3) Satoshi Yagi, Yoshihiro Nakata, Yutaka Nakamura, and Hiroshi Ishiguro, “Can an android’s posture and movement discriminate against the ambiguous emotion perceived from its facial expressions?” PLoS ONE 16(8): e0254905, 2021.
4) Yoshihiro Nakata, Satoshi Yagi, Shiqi Yu, Yifei Wang, Naoki Ise, Yutaka Nakamura, and Hiroshi Ishiguro, “Development of ’ibuki’ an Electrically Actuated Childlike Android with Mobility and Its Potential in the Future Society,” Robotica, pp. 1–18, 2021.
5) Haruhiko Adachi, Makoto Ozawa, Satoshi Yagi, Makoto Seita, and Shigeru Kondo, “Pivot burrowing of scarab beetle (Trypoxylus dichotomus) larva,” Scientific Reports 11, 14594, 2021.

Reviewed Conference Papers

1) Satoshi Yagi, Yoshihiro Nakata, Hiroshi Ishiguro, “Android Printing: Towards On-Demand Android Development Employing Multi-Material 3-D Printer,” IEEE International Conference on Humanoids, 2021.
2) Satoshi Yagi, Naoki Ise, Shiqi Yu, Yoshihiro Nakata, Yutaka Nakamura, Hiroshi Ishiguro, “Perception of Emotional Gait-like Motion of Mobile Humanoid Robot Using Vertical Oscillation,” ACM/IEEE International Conference on Human-Robot

Interaction (HRI), 2020.

3) Shiqi Yu, Yusuke Nishimura, Satoshi Yagi, Naoki Ise, Yifei Wang, Yoshihiro Nakata, Yutaka Nakamura, Hiroshi Ishiguro, “A Software Framework to Create Behaviors for Androids and Its Implementation on the Mobile Android ‘ibuki’,” ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2020.
4) Satoshi Yagi, Sunhwi Kang, Shiqi Yu, Jingcheng Shen, Hamed Mahzoon, “Evaluation of Shape-Changing Tensegrity Structure Robot for Physical Human-Robot Interaction,” IEEE International Conference on Advanced Robotics and its Social Impacts (ARSO), 2019.
5) Masashi Hamaya, Takamitsu Matsubara, Jun-ichiro Furukawa, Yuting Sun, Satoshi Yagi, Tatsuya Teramae, Tomoyuki Noda, Jun Morimoto, “Exploiting Human and Robot Muscle Synergies for Human-in-the-loop Optimization of EMG-based Assistive Strategies,” IEEE International Conference on Robotics and Automation (ICRA), 2019.
6) 八木聡明, 仲田佳弘, 中村泰, 石黒浩, “車輪移動アンドロイドの上体の揺動が人間の歩容に与える影響の検討,” 第 23 回ロボティクスシンポジア, 2018.

Non-Reviewed Conference Papers

1) 濱屋政志, 濱屋政志, 松原崇充, 松原崇充, 古川淳一朗, Sun Yuting, 八木聡明, 寺前達也, 野田智之, 森本淳, “ユーザ嗜好に基づく EMG を用いた運動支援制御器のベイズ最適化,” 計測自動制御学会システムインテグレーション部門講演会, 2018.
2) 八木聡明, 姜淳熙, 沈靖程, 于士キ, 中野賢, “昆虫モデルのロボットの歩容が人間に与える印象の評価,” ロボティクス・メカトロニクス講演会, 2018.
3) 八木聡明, 仲田佳弘, 中村泰, 石黒浩, “車輪移動アンドロイドの上体の揺動が人間の歩容に与える影響の検討,” 第 23 回ロボティクスシンポジア, 2018.
4) 植村充典, 八木聡明, 平井宏明, 宮崎文夫, 川村貞夫, “先端化学繊維ロープを用いた軽量脚・腕ロボット,” 第 35 回日本ロボット学会学術講演会, 2017.
5) 八木聡明, 植村充典, 平井宏明, 宮崎文夫, “足首周りの角運動量に基づく脚ロボットバランス制御の実験的検討,” ロボティクス・メカトロニクス講演会, 2017.