

Chapter 2

Transmitting Human Presence Through Portable Teleoperated Androids: A Minimal Design Approach

Hiroshi Ishiguro

Abstract In this article, results from the JST/CREST project “Studies on cellphone-type teleoperated androids transmitting human presence” are described. The goal of this project is to develop new robotic communication devices that can provide a strong feeling of human presence, called “sonzaikan”, by distance and to clarify the principles of human nature behind this. We took a synthetic approach, exploring three bidirectional processes: system development, field studies, and evaluation. We first define *sonzaikan*, which forms the core concept of this study, and then in the following sections describe various implementations of *sonzaikan* media prototypes and supporting technologies, evaluations of *sonzaikan* media, and field studies using prototypes we developed. We discuss the results in detail and then conclude this article with a brief summary of our essential findings and a mention of future work.

Keywords Teleoperated android · Minimal design · Human presence · Robotic media · Telenoid · Elfoid · Hugvie

2.1 Introduction

In the past two decades, people have acquired countless new means for communication: cellphones, e-mail, online chat, social networking services, and so on. With such a plethora of communication media, along with progress in information technologies and devices such as the World Wide Web and smartphones, the modern human lifestyle has rapidly changed. We can now talk with others anywhere, anytime and can send and receive not just text or voice data but also images and movies to express our ideas and feelings in finer detail. Such changes have not only increased the bandwidth and relaxed the distance limitations of communication they have also changed how people communicate with each other. Although such technologies seemingly increase opportunities for social relationships, studies have shown that more and

H. Ishiguro (✉)

Advanced Telecommunications Research Institute International,
2-2 Hikaridai, Keihanna Science City, Kyoto 6190288, Japan
e-mail: ishiguro@sys.es.osaka-u.ac.jp

© Springer Japan 2016

T. Nishida (ed.), *Human-Harmonized Information Technology, Volume 1*,
DOI 10.1007/978-4-431-55867-5_2

29

more people feel socially isolated [1]. Our current technologies are failing to create a strong subjective feeling of being with others.

There are now various types of robots being developed and appearing on the market that are designed to work in our daily environment. Some robots can make simple conversation with people autonomously, while other cannot speak but are anthropomorphized by people who talk to them. Still others function as a mobile video chat system. Robots differ from other information devices in that they can physically interact with real world objects. Specifically, they can move around in the world in which we live, can carry things, and can touch people or be touched by them. One can feel a strong presence of the robot itself, and as such, when robots are used as communication devices, they may have advantages over current communication tools.

Recent studies in cognitive science have argued that an object is represented as the integration of multimodal information in the human brain [2, 3]. For example, one study showed that people more accurately and rapidly identify objects when they are defined by a combination of auditory and visual cues than when the same objects are defined by unimodal cues alone [4]. Ernst and Banks proposed maximum likelihood estimation as a general principle of human sensory integration both within and across modalities for robust perception [5, 6]. Interestingly, their model suggests that the variance in an integrated estimate is reduced and the robustness of perception is increased by combining and integrating sensory information not only from within but also across modalities. These studies imply that the integration of sensory cues across different modalities is an alternative approach to facilitate and improve object recognition other than the integration of rich information within a single modality. Besides, we are fine tuned to recognize and interact with other humans. Many functionalities in our brain are specialized toward perceiving and reacting to other people. From this viewpoint, humanlike robots may be a key component for realizing a revolutionary new style of communication that is in harmony with human nature, easy to use, and provides new meaning in communication.

The question then arises: how much in the way of humanlike characteristics must we reproduce on the media to create the feeling of a human presence? Eliminating elements that are not required to induce a feeling of human presence might enhance the feeling because only relevant verbal and non-verbal information for human perception would be presented. In addition, developing a humanlike robot with the essential elements to feel human presence would help us understand how we perceive each other and enable us to design a robot's internal mechanism suited to natural interactions with people.

In this article, results from the JST/CREST project "Studies on cellphone-type teleoperated androids transmitting human presence" are described. The goal of this project is to develop new robotic communication devices that can provide a strong feeling of human presence, called *sonzaikan*, by distance and to clarify the principles of human nature behind this. We took a synthetic approach, exploring three bidirectional processes system development, field studies, and evaluation (Fig. 2.1).

As a first step, we must construct an appropriate robotic medium to understand how people recognize humanlike robotic media created with a minimal design approach.

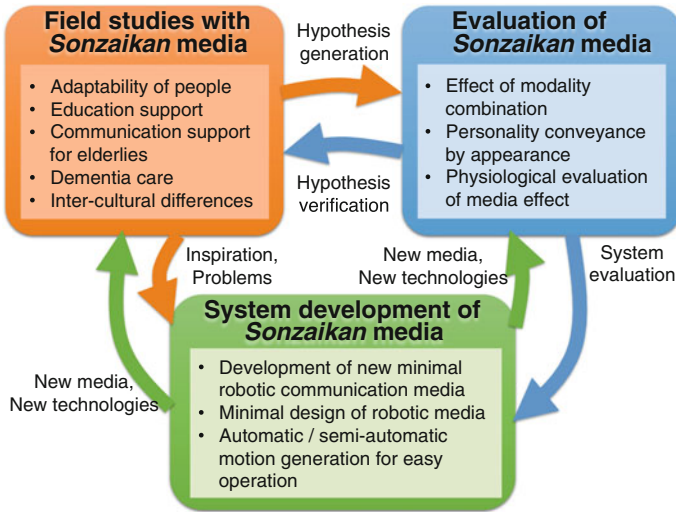


Fig. 2.1 Overview of JST/CREST project “Studies on cellphone-type teleoperated androids transmitting human presence.” A synthetic approach using bidirectional processes including system development, field studies and evaluation is taken

In this process, existing findings in cognitive science and psychology stimulate discussion of possible elements to enhance *sonzaikan* since they provide detail on the impact of each cue in the sensory information relating to humans on human perception. We also need artistic intuition to design an entire system based on such elements.

Besides the design of a robotic medium, we need to develop new technologies to allow people to easily use them as an effective communication media. Normally, teleoperated robots require complex operation and their human interface aspects have rarely been examined, as they are currently used only in specialized areas such as large-scale construction or the exploration of hazard scenes. Besides, when creating communication media for daily use, the robotic body needs to be portable and compact, and only a limited space is available to embed the equipment (such as actuators) that make up a robotic medium. In light of these requirements, we need new means to implement the necessary functionalities for *sonzaikan* media.

After developing a new medium, it remains unclear how people will treat it because it provides new experiences for them. In this case, we must observe a variety of interactions between ordinary people and the medium in actual situations to generate and test several hypotheses. Field studies are essential to investigate which features of the medium contribute to *sonzaikan*. They also provide new ideas about media design and improvement. At the same time, experiments under controlled conditions must be conducted to verify the discovered hypotheses and to evaluate the media’s effect. The verification and evaluation results are useful in terms of identifying possible applications and improving the media to enhance *sonzaikan*.

Once we identify the important features to enhance the feeling and the problems to be solved in the existing media through observation using case studies and evaluation of the media, the process focuses on system development. To this end, we develop new media with fewer features and new technologies to confirm and improve the effect of the features or to explore elements to enhance the feeling. We again observe the effect of the developed media using case studies and evaluate them through laboratory experiments.

In the following sections, we first describe the definition of *sonzaikan*, which forms the core concept of this study. In Sects. 2.3 and 2.4 we describe our implementation of *sonzaikan* media prototypes and their supporting technologies. Section 2.5 introduces studies evaluating *sonzaikan* media, and various field studies using the developed prototypes are described in Sect. 2.6. We then discuss the results and conclude the article with a brief summary of our essential findings and a mention of future work.

2.2 Presence, *sonzai*, and *sonzaikan*

In the past, we have run studies on android robots to create robots with humanlike presence and have developed several androids for this purpose. The most popular android we have developed is the Geminoid HI modeled on Ishiguro (the author), as shown in Fig. 2.2. The operator of the Geminoid HI talks with a visitor by watching the monitors. The operator's voice is sent to the android via the Internet and the operator's computer analyzes the voice and generates corresponding lip movements on the android. The computer also tracks the operator's face and head movements on the basis of images from a USB camera and synchronizes the android's face and head movements to the operator's. That is, the operator can see on the monitor the android moving the lips and head as the operator talks and moves and can make sure that he/she talks with a visitor by using the android body.

In our earlier experiments, we found that both visitor and operator can feel as if the android were a human. When Ishiguro operates the android modeled on himself, the visitor interacts with the android as if it were he. At the same time, Ishiguro,

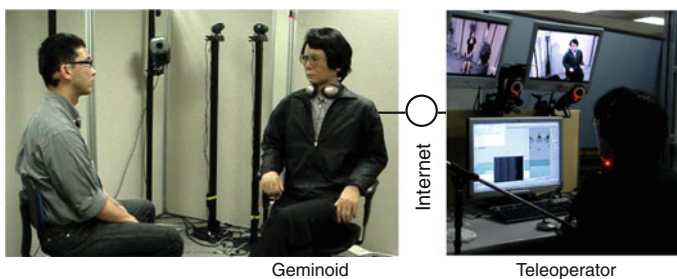


Fig. 2.2 Geminoid™ HI: teleoperated android of an existing person

who is operating the android, feels the android body as his own body. We have studied a teleoperated android called the Geminoid for the purpose of developing a new information media that can transmit human presence to distant places [7, 8]. “Geminoid” means a twin or doppelganger whose presence is not separate from the operator.

The Geminoid is a robot system that transmits the operator’s presence to distant places. In order to investigate and discuss the many interesting phenomena concerning this system, we need to start with a careful study of “presence.” In Japanese, we have two terms, *sonzai* and *sonzaikan*, corresponding to “presence.” *Sonzai* means “presence” or “existence” in English. There is no English word that perfectly matches the notion of *sonzaikan*, but as *sonzaikan* comprises *sonzai* and *-kan*, where *-kan* literally means “feeling”, it might be translated as “a feeling of presence”. However, this translation does not capture the precise meaning of the term.

The differences between *sonzai* and *sonzaikan* are subtle but substantial. *Sonzai* or presence of an object, say a human or a robot, implies the obvious or objective presence of the human or the robot, and many people can equally share its presence. Needless to say, it is difficult to theorize about presence per se, but we submit that whatever is present, e.g., Ishiguro the man or the Geminoid modeled on Ishiguro, many people share an awareness of the presence of them. For example, Ishiguro is considered to be present because anyone can see, hear, or touch him.

In contrast, *sonzaikan* is to feel *sonzai* or presence, and it differs from the obvious presence. An object may have *sonzaikan* even if we cannot see, hear, or touch it. That is, *sonzaikan* is the feeling of presence, and the relevant kind of feeling may not be sensory.

2.2.1 Hypotheses on Recognition and *sonzaikan*

In the Geminoid system, there are two types of *sonzaikan*: the *sonzaikan* visitors feel about the android and the *sonzaikan* operators feel through operating the android. These have some common features. Here, we focus on the first type, i.e., the feeling of a human presence in an android. Before proceeding, we need to consider what recognition is. For something to have *sonzaikan* (to a person) is for its presence to be recognized (by the person) in some sense. Therefore, we have to define “recognition” before considering *sonzaikan*.

Recognition is defined differently in different fields. For example, in psychology it is defined as consciousness of information that comes from the external world and is loaded with meaning, while in informatics it is usually defined as pattern recognition. This definition implies that recognition is to collect data about a particular target from various sensors. While bearing in mind the definitions in psychology and informatics, we propose an original and simpler definition: namely, recognition is to “represent information obtained through more than one modality”. For example, we cannot recognize an orange as such merely on the basis of its shape. The recognition of an orange requires (at least) one more modality, typically smell. If we can perceive the

smell of an object when we see its shape, then we can recognize it as an orange. Recognition of an object involves representing it in more than one modality and linking representations in different modalities with each other. We can apply this general idea of recognition to the specific recognition of human presence. We need to have at least two modalities to feel the reality of human presence or *sonzaikan*. Paradoxical as it may seem, robots or media can have *sonzaikan* if we can feel them in just two modalities.

2.2.2 *The Minimum Design Requirements for Providing sonzaikan*

When we interact with a person, we obtain via various sensors a variety of information related to his or her appearance, movement, voice, smell, and so on and recognize who it is by integrating this information. Then, how many sensors do we need? How many modalities do we need in order to recognize the person? To help with this inquiry, we developed the Telenoid, Hugvie and Elfoid, described in the following.

The Telenoid is different from the Geminoid in that its appearance and movement lack human likeness. The Telenoid was developed in the pursuit of the minimal design of a human being. It looks vaguely like a human, but we cannot tell its age or gender. The Hugvie is even more deprived of humanlike appearance, but even so, we can still feel *sonzaikan* when we are with a Hugvie. The modalities that the Telenoid can provide are much fewer than the Geminoid provides, but it still maintains a basic humanlike appearance, movement, and voice. In contrast, the Hugvie only provides human voice and humanlike touch.

The Hugvie is a human-shaped cushion designed to be used as a smartphone folder. There is a pocket for a smartphone on the head. Users talk while hugging it. That is, users can hear a human voice from the Hugvie and feel its humanlike surface by touch. The Hugvie is totally different from regular smartphones. When we use it, we have a feeling of holding the person who is speaking, even though he or she is nowhere near us. Our hypothesis is that we feel the *sonzaikan* of a human if we feel it through (at least) two modalities. Our earlier experiments with Hugvies contribute to verifying our hypothesis, at least to some extent.

Another interesting finding in our experiments is that we humans always fill in the lack of information with one or another positive interpretation when we feel *sonzaikan* in a limited number of modalities. When we use Telenoids or Hugvies, we identify the person who is speaking from the voice. However, identification by voice may fail. In such cases, we have some mental image of the speaker. If we hear an unknown but appealing voice from a Hugvie, we always imagine that the person speaking is attractive. We never have negative images. Similar effects have been confirmed in Telenoids. This may be why the elderly particularly like Telenoids (as discussed later).

We consider the Telenoid, Hugvie and Elfoid to be “*sonzaikan* media,” meaning media that produce feelings of humanlike presence. *Sonzaikan* media are often more widely accepted than Geminoids that have a humanlike presence. Geminoids, perhaps because of their fixed and detailed appearance, limit one’s imagination, whereas *sonzaikan* media, because they do not have such appearances, leave room for imagining its identity and coming up with positive interpretations thereof.

2.3 System Development of *Sonzaikan* Media

In this section, we introduce the development of three *sonzaikan* media, Telenoid, Hugvie, and Elfoid, based on the minimal design concept of *sonzaikan*.

People tend to attribute humanlike qualities to nonhuman objects to some extent, regardless of whether they resemble humans and whether they have intelligence [9, 10]. Based on this anthropomorphism tendency, researchers have demonstrated that people will treat even simple robots like a human in typical human-robot interactions. Matsumoto et al. developed Muu, which only has one eye, in a minimal design approach that eliminated the nonessential components and kept only the most fundamental functions for human-agent communication [11]. Osawa et al. demonstrated that attaching body parts to an object generates a virtual body image of it [12]. They also developed a ring-shaped robot called Pygmy with eyes and a mouth to anthropomorphize the user’s hands [13]. They designed robot agents by focusing on the facial elements as the minimal parts of a human. Although they enhanced the user’s feeling of being with socially interactive agents, their designs seem inappropriate to convey a human presence since their robots do not resemble humans. As a consequence, the interaction between users and robots is not the same as human-human interaction.

One option to enhance *sonzaikan* is to interact with a robot teleoperated by a person at a distant location. Many teleoperated robots have been developed to facilitate telecommunication. Some robots, which are utilized in elderly care [14], medical care [15], and group work [16], are equipped with a display monitor to project the teleoperator’s image [17]. Others have a movable artificial head rather than a monitor for telecommunication [18, 19]. These robots have achieved some level of success in terms of enhancing *sonzaikan* because they convey teleoperator movements, voices, and images. However, these studies ignore the perception of interaction partners because they aim to provide telepresence for the teleoperators. We address the design issues in terms of the perception of the person who is facing the robots.

2.3.1 Telenoid

Telenoid was made as a test-bed in the pursuit of the minimal design of a human being [20]. The Telenoid robot is 70 cm long and weighs about 3 kg (Fig. 2.3). It has nine degrees of freedom (DoFs), most of which are assigned to control its eyes, mouth, and head with the rest devoted to its left and right hands.



Fig. 2.3 Telenoid™. Photo copyright by Rosario Sorbello

Telenoid’s design emphasizes its human likeness in visual and tactile information to facilitate both human-robot and mediated human-human interactions. To provide a contrast to existing androids such as Geminoid HI [8], we removed as many features less important for communication with a human as possible from Telenoid.

We designed Telenoid in three steps. First we identified the important features for communication with a human and eliminated non-neutral and less important ones. Psychology has shown that non-verbal information such as gaze behavior and bodily gestures plays a crucial role in communication [21, 22], so we considered not only voice but also a humanlike head and body to be important while eliminating non-neutral and less important features such as beards, hair, and eyebrows.

Second, we reevaluated whether the chosen features fit the design requirements by eliminating less important ones. We eliminated legs since we assume that a robotic medium and its partner do not move around. Although fingers are often used for hand gestures, especially pointing, in human communication [21], we eliminated hands and retained arms because pointing is possible if a robot has arms. Even though facial expressions provide fundamental information about emotions, we do not consider them in our current development because vocal information conveys a person’s emotion to some extent [23].

Finally, we integrated the crucial features we chose. We kept the symmetry in Telenoid’s face because a symmetrical face increases attractiveness [24]. Personality biases, including gender and age, cause problems because such personal information about virtual and robotic avatars affects both the teleoperator’s attitude and the partner’s to the avatar [25, 26]. To avoid this, we designed Telenoid’s face and body in our current design as both ageless and sexually neutral.

Telenoid is covered with soft vinyl chloride. Its tactile quality resembles that of humans and it is much more robust and suitable for physical contact than silicon

rubber (as was used for Geminoids). It enables people to feel that they are touching a person during such physical interactions as hugging.

We chose a teleoperation system to control Telenoid because this allows Telenoid to not only be used as a robotic avatar but also to pretend to be an autonomous agent by using a Wizard of Oz approach. We do not implement an autonomous system on Telenoid since our focus is how people react to its appearance and/or behavior, regardless of its behavioral robot architecture and the cognitive processes that might occur inside the robot. In the following experiments, we used Telenoid as a robotic avatar and participants were informed of the existence of a teleoperator behind it. The Telenoid teleoperation system conveys its teleoperator's movements and vocal information. We used the teleoperator's head movements and lip motions to produce these movements for Telenoid. The actual motor commands are computed by face-tracking and lip motion systems and sent to a server by TCP/IP. The teleoperator can also display such predefined behaviors as goodbyes or hugs using GUI buttons on the laptop screen. Some involuntary movements including breathing and blinking are generated automatically so that the interaction partner feels as if Telenoid is alive. The system is easy to use and carry: it requires only a single laptop with a Web camera.

According to our hypothesis that *sonzaikan* is enhanced when sensory information representing a human from at least two different modalities is presented, Telenoid provides rich information even though much information is eliminated since it presents visual, tactile, and audio information. This richness allows us to investigate the effect of possible combinations of modalities on *sonzaikan* in various interactions.

2.3.2 Hugvie

Observations from field studies with Telenoid have shown that physical contact with it, especially hugging, is a primary form of interaction and has a strong psychological impact on a wide age range of users. This implies that a combination of auditory and tactile sensations enhances *sonzaikan*, supporting our hypothesis that the feeling is enhanced when information is presented from at least two different modalities. On the basis of this finding, we developed a human-shaped cushion phone called Hugvie [27] (75 cm, 600 g) as a communication device that focuses on the hugging experience (Fig. 2.4). While Telenoid maintains a minimal humanlike appearance, movement, touch, and human voice, Hugvie focuses on just the human voice and humanlike touch. It is a soft cushion filled with polystyrene microbeads and is covered with spandex fiber, which is often used for microbead pillows. It resembles a person with open arms for a hug and enables us to bring the hug experience into telecommunication by putting a hands-free mobile phone inside a pocket in its head. Since the phone is in the pocket, people can call and talk while hugging Hugvie. As they converse with a distant partner, they become immersed in the vocal and tactile information since they almost do not see Hugvie while they are hugging it. This increases the feeling that they are actually hugging their distant conversation partner.

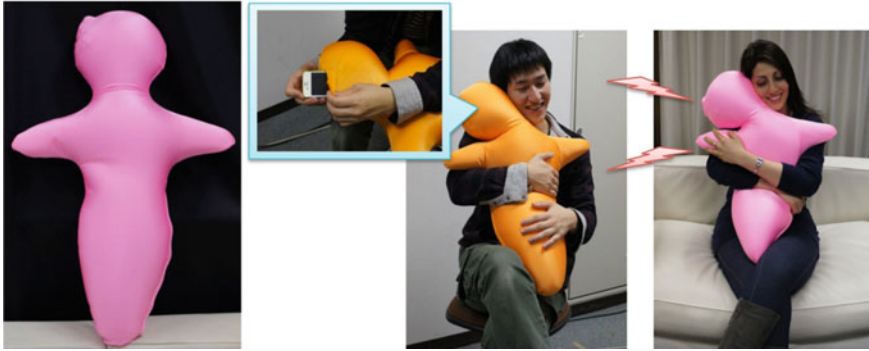


Fig. 2.4 Hugvie™. A user inserts his/her smartphone in a pocket at Hugvie’s head and talks by hugging it

Preliminary studies with a Hugvie revealed that conversations with a female while hugging it stimulated the affections that heterosexual male users felt to her [27, 28].

2.3.3 *Elfoid*

Elfoid is a hand-held version of Telenoid [29]. Figure 2.5 shows a cellular phone version of Elfoid that can connect to a public cellular phone network and enables communication with other telephones. In addition to the cellular phone version, we have developed a Bluetooth version that has functionality similar to a Bluetooth headset for cellular phones.

Similar to Telenoid, Elfoid has a simplified human shape and is designed to transfer the speaker’s voice and gestures using the cellphone network. Unlike Telenoid, Elfoid

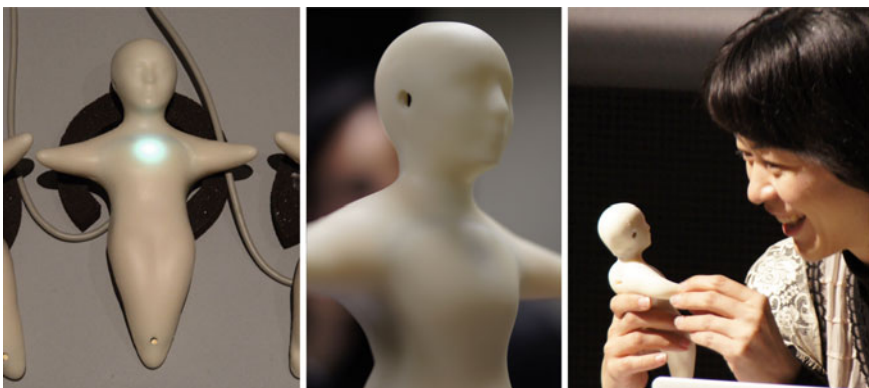


Fig. 2.5 Elfoid™. In contrast to cellphones, people use Elfoid by holding it in front of them and talking to others through it

is covered with urethane gel, which provides a soft, pleasant feeling on touching and holding it. Its functionality is further reduced from that of Telenoid, having a smaller size, fewer appearance features, and reduced embedded equipment. We assume Elfoid to be used as shown in Fig. 2.5, where the user is holding it in hand in front and talking to it. In this way, we expect people to feel stronger *sonzaikan* of others compared to usual cellular phone usage while maintaining the ubiquity of cellular phone, thus enabling people to use it anytime, anywhere.

2.3.4 *Exploration of Human Form from Ancient Human Design*

Although Telenoid, Hugvie, and Elfoid were designed with a minimal design approach, it remains unclear whether their components actually satisfy the minimal requirements to enhance *sonzaikan*. To explore this, we studied the minimal requirements of the human form by investigating the chronological development of Dogū, one of the most ancient examples of attempts to create an artificial human form [30]. The purpose of these small human/animal figures remains unknown, but they were probably meant to represent a human or to communicate with invisible spirits that take human form [31, 32]. We surveyed the development of Dogū and found that the torso, not the face, was considered the primary element for representing a human. Less attention was paid to the arms, legs, hair, and ears, all of which were represented very crudely. On the basis of these survey findings, we examined what kind of body representation is necessary to feel *sonzaikan* by using a conversation task consisting of one speaker and five hand-held avatars whose body forms are different. In the experiment, participants spoke to an experimenter through one of the avatars or the speaker in his/her hand about a topic provided by the experimenter. After the conversation, they rated the degree to which they felt the experimenter's presence on a five-point Likert scale from 0 (not at all) to 4 (very strongly) and then repeated conversations with different avatars on different topics. The experimental results showed that the forms for the torso and head enhance this feeling most significantly, while the arms and legs have less impact. This implies that Telenoid's appearance satisfies the requirements to feel *sonzaikan* and that we can eliminate more elements from it, such as arms and legs.

2.4 Technologies Behind *sonzaikan* Media

The development of a new medium requires new background technologies for the human interface as well as for the internal structures. In this section, we describe some of the technical studies that support our implementation of *sonzaikan* media.

2.4.1 Motion Generation Through Speech Information

In our field studies, we observed that people sometimes had difficulty communicating with Telenoid because the teleoperator's voice is not synchronized with Telenoid's movements when we generated its movements by vision-based head and lip tracking techniques, such as active appearance models. A problem with such image processing techniques is that their performance depends on good lighting conditions and image resolution. This is often crucial for applications in the real world. For example, our system failed to capture a teleoperator's movements in experiments at a shopping mall and at an elderly care facility because of poor lighting conditions. Therefore, complementary approaches are required that use information other than visual. A motion capture system achieves synchronization, but it is too expensive and complex for daily use. If the motions are reproduced from vocal information, synchronization can be easily maintained.

Ishi et al. generated lip motions based on the teleoperator's speech information. They transformed the formant spaces where vowel information is associated with lip shape in strong correlations between the two and demonstrated that the transformed space allows us to generate natural lip motions from the teleoperator's speech without any sensor system for motions [33]. They also analyzed the head motions associated with speech in human conversations and found a strong relationship between head motion and such dialogue acts as affirmative or negative reactions, the expression of emotions like surprise or unexpectedness, and turn-taking functions. On the basis of this finding, they constructed a model to generate head movements from teleoperator's dialogue acts and were able to improve the naturalness of robot head movements [34].

2.4.2 Motion Generation and Emotional Expression Through Visual Stimuli

In use cases with Telenoid, both young and elderly users pointed out that it was too heavy. Its weight (3 kg) is mostly due to the electric motors used as its actuators, although we did try to minimize them. To further reduce the weight, we need another implementation to produce its movements. This approach is especially useful to develop small, cheap, and portable robots because electric motors increase costs and reduce portability. If a robot can help users perceive the illusory motions of its limbs by light, sounds, or vibrations, it can support natural interaction without embedded actuators that move the limbs. Sakai et al. induced the illusion of motion with Elfoid by embedding blinking LEDs in its face [35]. After evaluating several possible patterns, we designed a blinking LED pattern to induce an illusory nodding motion, which is an important nonverbal expression in face-to-face communication. We demonstrated that Elfoid with illusory nodding motions eased participant frustrations more than with a random blinking pattern when participants grumbled at it.

This approach is a new way to achieve a portable robot avatar designed with minimal elements to enhance feelings of *sonzaikan*.

The idea of using blinking LED patterns for motion generation is also applicable to emotional expressions. Fujie et al. explored this possibility [36] by investigating which emotions are conveyed by blinking color patterns. Multi-color LEDs were embedded into Elfoid's face, its torso, and a spherical object. Eight color patterns (red, blue, green, yellow, purple, orange, blue-green, and yellow-green) were displayed with three different blinking patterns (continuous emission, emission at 0.1 s intervals, emission at 1.0 s intervals). Thirty-two participants saw all conditions randomly and evaluated each one with Plutchik's set of eight basic emotions (joy, acceptance, fear, surprise, sadness, disgust, anger, and anticipation) [37] on a six-point scale from 1 (not felt at all) to 6 (felt extremely strongly). While no conspicuous difference was observed between Elfoid and the spherical object with respect to red, blue, green, yellow, and yellow-green light, different impressions were created by purple, orange, and blue-green lights between Elfoid and the spherical object. With Elfoid, a purple light increased negative emotions, an orange light at the chest decreased anger, and a blue-green light increased fear. For both Elfoid and the spherical object, high-speed blinking increased surprise. Elfoid's humanlike appearance induces different emotions conveyed by color patterns compared to simply shaped objects.

2.5 Evaluation of *sonzaikan* Media

2.5.1 Impact of Physical Embodiment on *sonzaikan*

Although Telenoid's physical embodiment apparently conveys *sonzaikan*, it remains unclear which factors play an important role in enhancing that feeling. Tanaka et al. evaluated the impact of several key factors on *sonzaikan*, including physical presence, voice, body motion, and appearance [38, 39]. In their experiments, they controlled the factors with different communication media. Participants explained electronic devices for more than 1 min in the same room to a remote conversation partner through Telenoid, a static Telenoid, an audio speaker, a virtual avatar, a video chat system, or their partner. The partner gave vocal backchannel feedback with head nodding. In this conversation, each medium presented the information shown in Fig. 2.6. To evaluate the impact of the physical embodiment on *sonzaikan*, participants rated how realistically they felt they were with the partner on a nine-point Likert scale.

Results showed that the movements presented by the media enhanced *sonzaikan*. Interestingly, this effect was significantly stronger when a physical entity (an actual person or Telenoid) was presented than when a teleoperator's image or a virtual avatar was projected on a display monitor or only the partner's voice was presented. Telenoid's score was significantly lower than the score of the actual person. On

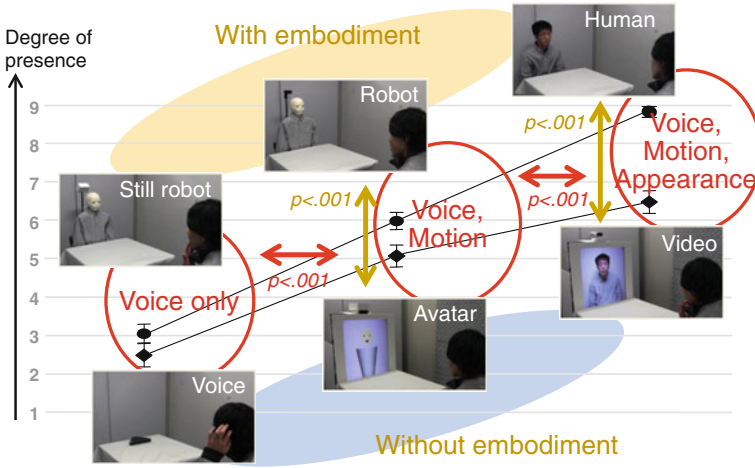


Fig. 2.6 Effect of physical embodiment. Combination of elements were compared to check their contribution to presence transmission

the other hand, its scores nearly equaled those of the video chat system, although it provides less visual information than video chat systems. These results indicate that Telenoid’s humanlike physical presence enhances *sonzaikan* in comparison with vocal information and displayed visual images. Also, *sonzaikan* is more enhanced when Telenoid moves.

2.5.2 Personality Conveyance

Since Telenoid has a humanlike appearance that consists of the important components to represent a human, people can easily recognize it as a teleoperator, unlike other robots whose appearance greatly differs from that of the teleoperator. Kuwamura et al. verified that Telenoid can convey the teleoperator’s personality, hypothesizing that teleoperated robots whose appearances differ from humans distort the teleoperator’s personality more than those that resemble humans because such appearances make different impressions on the teleoperator [40]. Participants conversed with a teleoperator who talked through one of three physical entities with different appearances (Telenoid, a stuffed-bear robot, and a video chat system) in three different face-to-face conversations: free talk, teleoperator’s self-introduction, and interviews by the teleoperator.

The teleoperator’s personality was measured by the Japanese Big Five personality test, a 60-item questionnaire that represents five personality parameters (extraversion, neuroticism, openness to experience, agreeableness, and conscientiousness) that was translated from the Adjective Check List [41, 42]. We evaluated the distortion effect

with the consistency of the answers to questions from each parameter in the measurement. We found that if a physical medium's appearance differs from that of the teleoperator, the answers to questions related to each parameter cause inconsistencies because some questions in each parameter are answered on the basis of impressions of the medium's appearance while others are answered on the basis of those of the teleoperator. The stuffed-bear medium had poor consistency on extraversion in the interview situation and on agreeableness in the self-introduction situation. Such poor consistency was not observed for Telenoid or in the video chat cases. This indicates that the personality transmitted through the stuffed-bear robot was distorted under certain situations, while Telenoid conveyed the teleoperator's personality to a similar extent to what a video chat system can convey. Although we did not directly compare Telenoid with its teleoperator, the results imply that Telenoid can maintain the teleoperator's *sonzaikan*.

2.5.3 Changes in Impression Toward Others

In another study, we examined whether the act of using *sonzaikan* media through can enhance positive feelings toward others [27]. Hugvie was designed so that its users could naturally hug the device in order to listen to their conversation partner's voice. In most cultures, social morals strongly discourage hugging strangers; if you are hugging someone, you are in close relationship with that person, sometimes an intimate relationship. However, using Hugvie, you are naturally led to perform the act of hugging even when you are talking with a stranger. Does this act of hugging change your impression toward the stranger?

In 1974, Dutton and Aron showed that one's affections are sometimes mistakenly evoked [43]. In their experiment, male participants were interviewed either on a fear-arousing suspension bridge or a non-fear-arousing bridge. The participants showed stronger response in the fear-arousing situation, but only when they were interviewed by a female. From these results, Dutton et al. concluded that misattribution of arousal happened to those participants; that is, their raise in tension and heartbeat change due to the fearful situation were mistakenly perceived as due to the attractiveness of the interviewer, which led to a stronger positive impression toward the female. Nishimura et al. showed that similar changes can be induced by controlling the frequency of the heartbeat-like vibration provided to participants [44].

In the experiment, we focused on hugging behavior through media among young people and found that using Hugvie enhances the feeling of being together and being loved compared to a Bluetooth headset. All participants were male university students and were told to interact with the other participant, which was actually a recorded female voice played by the experimenter, and watch a movie together while connected by the media. After the interaction, participants answered a questionnaire and were briefly interviewed. We used a "Loved-Liked scale," a passive version of a questionnaire, to measure positive impression toward others, specifically, participants' impression on how the people on the other side thought about themselves.

We found significant differences in both loved and liked scales: when using Hugvie, people tended to feel that they are more liked/loved by others than when using a normal headset.

We also found that participants with Hugvie were more impressed by a movie scene where a boy says, “I love you.” This is a scene in which one typically feels rather embarrassed even if watching alone, and even more so when watching it with someone else nearby. This seems to indicate that using Hugvie provided participants with a much stronger and realistic feeling of their conversation partner’s presence.

2.5.4 *Physiological Measures*

Even though the results from the above studies seem to suggest that *sonzaikan* media has a strong effect on people, it remains unclear whether the usage of such media can produce physiological responses as observed in real human interactions, or sometimes even stronger responses than toward other people. To address this issue, we investigated whether endocrine changes are observed following a brief conversation through a huggable communication device [45]. This approach enables us to quantitatively evaluate the physiological effects of the mediated touch without relying on subjective reports of affective states.

We hypothesized that communication with a remote person by giving a hug to a physical device would be sufficient to influence the human neuroendocrine system. To test this idea, we examined the changes in cortisol hormone, which is a reliable bio marker of psychological illnesses [46], before and after participants engaged in a human-human conversation mediated by a huggable communication device. We focused on the cortisol hormone because stress relief is one of the most critical issues in providing social support to facilitate recovery from many types of mental and physical illness [47]. Considering the potential applications of communication media for social support, the impact of the media on stress relief is highly relevant.

In the present study, participants had a conversation with a stranger while hugging a Hugvie (Hug group). In a control group, participants went through the same procedure but used a mobile phone instead of Hugvie (Phone group). To assess the neuroendocrine responses to the social interaction with the communication media, we measured cortisol levels before and after the conversation session. We collected the cortisol levels both from the blood and salivary samples since they can be dissociated due to differences in their regulatory mechanisms [46]. We predicted that physical contact with the huggable device would reduce the cortisol levels at a greater rate than the control group in which participants had conversations on a mobile phone without physical contact. We also evaluated the effect of physical contact on subjective psychological states with a post-session questionnaire that assessed positive effect, negative effect, and calmness.

Results showed that hugging the communication medium reduced the cortisol levels in both saliva and blood. These results, which support our hypothesis that physical contact with communication media can produce an effect even at the endocrine level,

suggest that physical contact with such a medium might be effectively used for mental stress relief. To the best of our knowledge, this is the first study that demonstrates an endocrine effect from physical contact with a communication medium.

We also found a reduction of the cortisol levels in both the blood and saliva samples and a positive correlation between the changes in salivary and blood cortisol. This indicates that we can use salivary cortisol, which can be more easily handled than blood cortisol, to evaluate the effect of physical touch with communication media. We expect salivary cortisol to be a promising new measure to assess the effects of physical touch with communication media that have previously been evaluated only with behavioral or psychological measures.

Our results provide us with two important implications. First, they suggest that communication media do not need to actively stimulate a person's skin to reduce cortisol levels. In previous research on interpersonal touching, active touching by others, such as stroking the arm and massaging, was primarily used as tactile stimuli [48–50]. Other studies used a combination of several types of inactive touch, such as holding hands and hugging, along with other factors (e.g., watching romantic videos) [51, 52]. There has been little investigation on the endocrine effect of single inactive touch, aside from one study that reported changes in the heart rate and blood pressure during gentle touches of the wrist [53]. Our results demonstrate that 15 min inactive touching with an inanimate object reduces the levels of stress hormone.

Second, our results indicate that communication media can be used as research tools to investigate the positive effects of physical touch independently of the touching situation and the person doing the touching. The effects of interpersonal touch on physiological responses are affected by how people are touched and by whom [54]. For example, while positive physiological changes are induced by a hug with a friend or family member, such changes do not occur for a hug with a stranger of the opposite sex because it could be taken as sexually offensive. By contrast, our present study shows that hugging an inanimate object reduces cortisol even during conversations with a stranger of the opposite sex. This suggests that communication media allow us to separate the actual effects of physical contact from the effect of intimate relationships in interpersonal touching, which could induce multiple effects.

The ability to reduce cortisol levels seems suitable for improving the quality of intimate social interaction in which trust and bonding are crucial. For example, remote counseling services are widely used to improve patients' psychological states and mental health. The quality of communication with therapists, typically conducted with telephones, the Internet, or videophones, may be enhanced by huggable devices [55].

2.6 Field Studies with *sonzaikan* Media

We brought Telenoid into public places, elderly care facilities, and an elementary school to observe how people react to its appearance and/or behavior and what information substantially enhances *sonzaikan*. In this section, we first report ordinary

people's responses to Telenoid in such public places as a shopping mall and then describe the responses of seniors in one-to-one conversations with Telenoid in elderly care facilities to investigate their impressions. We also compare interaction responses with Telenoid between Japan and Denmark to examine cultural differences. Finally, we report on Telenoid's effectiveness at representing a remote person's presence in a group setting.

2.6.1 Acceptability from People

Since Telenoid is a new robotic medium, it is unclear to what extent it makes people feel *sonzaikan*. Ogawa et al. observed how people responded to it through a demonstration at a shopping mall [20]. Seventy-five people, many of whom were in their 20s, had 5 min conversations with an experimenter through Telenoid. After the conversations, we asked whether Telenoid was better than a telephone for talking to a remote person. More than 70 % felt that Telenoid outperformed the telephone. They were also asked for their impressions of Telenoid. About 36 % admitted that it was strange at first glance. However, about 73 % said that their attitudes became more positive after hugging it. These tendencies were also observed for elderly people. At an open house at Advanced Telecommunications Research Institute International in Kyoto, 47 elderly people had 5 min conversations with Telenoid and were also asked whether Telenoid was better than a telephone, with 66.6 % preferring Telenoid and 88.8 % giving positive comments. Interestingly, all of the elderly users hugged it immediately after they got it. These results show that ordinary people generally accept Telenoid, although some had a cautious first impression. Physical contact with Telenoid might be a primary component to enhance *sonzaikan* because hugging it made the users' attitudes more positive.

2.6.2 Elderly Care with *sonzaikan* Media

The fact that elderly people are attracted to Telenoid suggests possible applications to elderly care. Yamazaki et al. used Telenoid at a residential care facility to observe the reactions of elderly people to it [56]. The participants were ten elderly women with dementia (mean age, 86.6 years), including mildly demented patients who could live independently with supervision and patients with moderate to severe dementia who had difficulty communicating with others.

Each participant had a relaxed, 20 min conversation about her health, hobbies, or family with Telenoid, which was teleoperated by the experimenters or the chief caretaker at the facility (Fig. 2.7, left). The caretaker reported that the overall reactions of participants were quite positive. We also observed that they frequently interacted with Telenoid by talking to and touching it. The verbal and non-verbal responses of nine of the ten participants to Telenoid were positive from the very start. The



Fig. 2.7 Elderlies interacting with Telenoid in Japan (*left*) and in Denmark (*right*)

patients with mild dementia were especially responsive in verbal interactions. At first glance, almost every senior reacted positively to interaction with it, often making such comments as “You are really cute.” Generally, attachment to Telenoid increased when they held it. Although the participants with severe dementia had difficulty maintaining verbal communication, they intermittently caressed its back and arms and slowly interacted with it. Interestingly, some participants asked it questions like, “May I hold you?” Although they might speak to it, such asking behavior is not typical for interaction with a doll, implying that the participants treated Telenoid more like a human. In fact, some seemed to confuse Telenoid with a child; one woman said, “You look about five years old” when the teleoperator asked the woman to guess its age. Perhaps its appearance, which was designed with fewer human elements, provides enough information to enhance *sonzaikan*.

2.6.3 Cultural Differences in Responses Toward Telenoid

Although it has been shown that elderly people accept Telenoid, is this attitude specific to Japan? In a field experiment in Denmark, Telenoid was introduced into care centers and the homes of elderly persons to investigate cultural attitudes to it [57]. In one case, we observed 2-hour free conversations between Telenoid and two participants who were living alone in houses attached to care facilities: a healthy 92-year-old and a 75-year-old with mild Alzheimer’s disease. In both cases, Telenoid was set up in the relaxing environment of a living room (Fig. 2.7, right). Telenoid was teleoperated by experimenters, nursing students, or the patient’s friends. The conversation topics included health, hobbies, family, and a cooperative map game.

As observed in Japan, Telenoid elicited positive responses and behaviors from both participants, who actively talked with it and engaged in conversations from the beginning. For example, the healthy participant entertained Telenoid by playing the piano and singing. The participant with mild Alzheimer’s basically remained quieter and calmer, but he did talk about his interests. Both participants engaged in physical contact, such as touching, holding, hugging, imitating, and kissing. Our observations indicate that both physical contact with Telenoid and its appearance might play an important role in enhancing *sonzaikan* across cultures, although investigation with more participants from various age groups and cultural backgrounds is required.

2.6.4 Education Support with *sonzaikan* Media

While the above studies focused on adults, Yamazaki et al. introduced Telenoid into an elementary classroom’s group activities to observe how children dealt with and adapted to it [58]. A class consisting of 28 children participated in the experiment in a typical Japanese classroom. They were divided into six groups of four or five students to discuss scenarios for a four-frame cartoon. One member of each group teleoperated Telenoid from a small room next to the classroom during the group discussions. We qualitatively assessed the effect of Telenoid by comparing the group work with or without its intervention on the basis of recorded dialogues, the recorded children behaviors, and post-interviews (Fig. 2.8).

We observed each group’s changing structure in the interaction among the children and the Telenoid controlled by a group member. Before Telenoid was introduced, the children’s participation in discussions was limited because they were dominated by a member who had been assigned as the leader; other members performed tasks irrelevant to the group task (partial participation). After Telenoid was introduced, however, all of the children started to negotiate with the operator, who became a newcomer to the group, since they were attracted to the novel tool. However, once group work began, they realized Telenoid’s limited function to cooperate with them



Fig. 2.8 Group-work in elementary school with Telenoid. One of the students is joining the class through Telenoid

and so began to work together to help the operator (cohesion and negotiation). Once Telenoid was accepted as a member of the group with the help of the others, the operator also made valuable contributions to the group. She began to take on the role of coordinator because she was able to keep an objective eye on the behaviors and roles of the others. Interestingly, as discussion continued, some of the children said that Telenoid seemed to be the operator herself. In other words, they felt the operator's presence (full participation). These observations indicate that Telenoid can maintain a classmate's *sonzaikan* due to its physical embodiment. They also demonstrate Telenoid's usefulness for facilitating group work.

In another study, Hugvie was used in school to help increase the concentration of students [59]. This was based on the idea that the stress reduction and human presence enhancement effect of Hugvie would allow younger students, especially first graders, to better concentrate on listening to their teacher. Children this age have not yet learned that listening to others is a very active process and is critical for learning and memory [60]. Despite the importance of listening, many young students have a problem with it; they walk around the classroom, chat with friends, and exhibit other restless/disobedient behaviors during class. The inability to concentrate in class causes many problems for children, especially those in the lower grades of elementary school, because it lowers academic performance in later life. Hugvie can relieve this problem by encouraging children to pay more attention in class.

In this study, we introduced Hugvie into a storytelling context to 33 preschool children who will soon start elementary school. They were given Hugvies and instructed how to use them. Two volunteers told them stories that were illustrated with picture cards.

Somewhat surprisingly, no child walked around the room or chatted with friends during the talking stories of the volunteers. Figure 2.9a shows a typical scene during storytelling. Although we had worried that some children might play with their Hugvies, all of the children calmly listened to the volunteers in both storytellings. About two thirds listened to the volunteer voices from their Hugvies, while the rest used them like cushions. Many children preferred to listen to the storytellers through the Hugvies. Children at the back of the room seemed to listen to the volunteers' voices from Hugvies without any complaints, even though they had difficulty seeing the picture cards. Hugvie seemed to help children continue to pay attention even from the back of the room.

In contrast, while not using Hugvie, the attention of the children in the back of the room shifted to other things, even as the children in the front continued to focus on the volunteers. Some children in the back walked around the room and others started playing with their friends (Fig. 2.9b). This might be because the volunteers sometimes concentrated on themselves without addressing the children. However, since more of the children who showed restless behavior were observed in the back of the room than in the front, perhaps the children in the back had less feeling that the volunteers were talking to them due to the distance.

These results demonstrate that Hugvie has the potential to help children maintain their attention to listening to others by reducing their stress and strengthening the feeling that a storyteller is close. Our observation of storytelling to children supports



Fig. 2.9 a Storytelling with Hugvie and b paper-cutout activity without Hugvie

such potential, suggesting that Hugvie is a useful tool to relieve the educational problem caused by children who show restless and disobedient behavior during class. We believe that concentration on listening can improve learning and memory performances, as suggested in a listening model [61]. We also believe that Hugvie’s effect is useful for children with such developmental disorders as attention deficit and hyperactivity disorder (ADHD), who typically have difficulty maintaining attention in class. We have begun applying our system for storytelling to such special-needs children.

2.7 Discussion

Our field studies showed that ordinary people easily accepted Telenoid and enjoyed talking with it. Their positive impressions seemed enhanced when they hugged it. Studies at a shopping mall and elderly care facilities showed that seniors also quickly accepted Telenoid and actively talked to it. These results highlight the possible contributions of its appearance, movements, voice, and touch for enhancing *sonzaikan*. The combination of humanlike touch and a human voice seems especially important. This finding supports our hypothesis and inspired us to develop Hugvie. Although Hugvie shows sufficient effect to change human attitude, further studies are required in terms of its shape. For example, if it had a normal cushion shape instead of a humanlike shape, would the same results hold? Would a user still have the feeling of

hugging a distant conversation partner in such a case? We will address these questions in future work.

Our observations in an elementary school revealed that introducing Telenoid into group activities changed the group's structure. Its physical embodiment and its movements captivated the children and helped them accept Telenoid as their classmate. Consequently, their distant classmate became more integrated into the group's activities. These results indicate that a medium designed with our approach has enough human information to convey a teleoperator's personality and to encourage involvement in group activities, even though it provides much less information than Geminoids.

Some of our field work findings were evaluated in controlled experiments. Telenoid's physical presence and motions created greater *sonzaikan* feelings than virtual avatar systems and an audio speaker. Its humanlike appearance conveyed a teleoperator's personality with less distortion than a non-humanlike appearance. Therefore, its physical embodiment and motion features provide significant effects of this feeling. Our results confirmed that Telenoid has advantages over existing media in terms of its enhancement of *sonzaikan* and personality conveyance. However, its appearance can be simplified because experimental results based on the findings from ancient human Dogū forms imply that a simpler human appearance has the same effect as Telenoid. It is also worth investigating whether *sonzaikan* provided by Telenoid and a simpler robot is comparable with that by Geminoid and whether they can induce natural responses from people as Geminoid can. Even though in our field studies we found that tactile interaction was observed as a primary factor, no investigation has evaluated the effect of physical contact with the media on this feeling. We have to verify the effect in controlled experiments.

2.8 Conclusion

In this paper, we proposed a minimal human design approach to explore the minimum requirements to enhance the feeling of a human presence, or *sonzaikan*, and gave an overview of our work with Telenoid, which was developed using minimal human design. Since developed communication media always provide new experiences for people, careful observations in exploratory studies are necessary to identify the primary factors to achieve *sonzaikan* as well as evaluation in control conditions and system development. These three processes, system development, field study, and system evaluation, must be repeated to explore the minimal requirements to enhance such feelings of a human presence. From field studies with Telenoid, we found that its physical embodiment and physical contact with it are crucial factors to experience the *sonzaikan* feeling. In controlled experiments, we verified some of our findings, which helped us develop a new medium and improved technologies to enhance the feeling. We also discussed some potential advantages of our approach. The lack of information might not be a disadvantage because it promotes positive social interaction.

We developed a simple but effective communication media for Hugvie based only on the minimal design approach (not on existing approaches) to telepresence media that reproduced the original modalities. Beyond the existing communication media, this approach will also usher in a new fashion of *sonzaikan* media that might, for example, make us feel another person's presence only by auditory and tactile information. We expect that this approach will improve social support systems in our future highly networked societies and help us understand how we recognize a human and how we design an autonomous system that can naturally interact with ordinary people.

Several technical challenges related to problems found in the field work and the evaluation results were addressed. Motion generation from vocal information [34] was easily synchronized between voice and motion, which is often a problem during teleoperation, without a large sensor system such as a motion capture system. The design of motion and emotional expressions with blinking light patterns [35, 36] is a new idea to achieve high portability. Although these methods show effective performance in certain telecommunication aspects, evaluating their effects on *sonzaikan* must be conducted in the future since they have not yet been assessed in this context.

Even though several issues of behavior generation were addressed, less attention has been paid to the development of recognition systems. To enhance *sonzaikan*, the teleoperator and conversation partner states must be extracted as precisely as possible. However, we have to avoid devoting too much computation to recognizing them since this prevents real-time conversation. Image processing forces a system to spend excessive computational loads. Cloud computing is a possible solution to reduce the computational load. We plan to design an effective facial recognition system by exploiting cloud computational resources.

In *sonzaikan* evaluations, many different subjective measures were used. We have to develop a unified measure for future evaluations. For example, we will integrate several measures of telepresence, social presence, and copresence that were used to evaluate the virtual agents [62]. In addition, physiological evaluations, including changes in brain activity and hormonal activities, must be conducted not only to evaluate *sonzaikan* but also to show how feeling *sonzaikan* affects our mental and physical health. We have already started addressing this issue in an investigation of the stress release of Hugvie with hormonal tests [45].

At the time we started the series of studies described here, we were not completely sure what exactly is entailed by transferring human presence to a remote operator. Through the development of *sonzaikan* devices and several basic studies, our ideas on *sonzaikan* have become much more clear and we were able to build a hypothesis on essential elements for *sonzaikan* transfer. First, as described in Sect. 2.2, we clarified our idea in terms of presence, *sonzai* and *sonzaikan*. *Sonzai* denotes explicit status as being human or robot as commonly accepted by many people. We can assume something exists when many people share a common belief on its status. In contrast, *sonzaikan* denotes the feeling that something is present. While being present or not is clear and explicit, there are degrees of *sonzaikan*, from strong to weak. In an extreme case, we can feel the *sonzaikan* of an entity even if it is not visible.

On the basis of these definitions, we composed a hypothesis on the principles embedded in us for establishing the *sonzaikan* of others. That is, the *sonzaikan* of people requires at least two unique modalities of human likeliness. Mere voice transfer is not sufficient; the transfer of an additional modality, such as vision (humanlike appearance or motion) or humanlike tactile sensation, is required. In the studies described here, we have been investigating the minimum necessary modality to represent human *sonzaikan*. Field tests using the developed *sonzaikan* devices, namely Telenoid and Hugvie, seem to support our hypothesis. In most cases, the teleoperated robotic media exhibited a strong and significant effect, sometimes even stronger than interacting with people face to face. Such results are supported not only in Japan but also in European countries, especially in Denmark, and the development of real-world application such as in care facilities has already started. Through investigation, both in basic studies and in practical usage of our robotic media, we aim to establish a firm design guideline for developing further advanced *sonzaikan* media and new forms of human communication.

Acknowledgments This project has been a collaborative work between several researchers from Advanced Telecommunications Research Institute International (ATR), Osaka University, and Tottori University. As principle investigator, I, Hiroshi Ishiguro, deeply appreciate the active collaborative work of everyone in the project. Dr. Shuichi Nishio at ATR took the initiative with the Telenoid field tests in Denmark and Japan. In addition, he helped me to summarize this report. Dr. Takashi Minato at ATR developed various versions of Telenoid and Elfoid. Dr. Carlos T. Ishi developed the software system that generates lip and head movements from the operator's voice. Dr. Hidenobu Sumioka took the initiative with the Hugvie studies in elementary schools. Dr. Ryuji Yamazaki at ATR contributed to the field tests with Telenoid in Denmark. Prof. Hideyuki Nakanishi at Osaka University took the initiative with the psychological studies with Telenoid. Prof. Kazumasa Tanaka at Osaka University worked with Prof. Hideyuki Nakanishi and contributed to the psychological studies. Prof. Aya Nakae at Osaka University worked with Dr. Hidenobu Sumioka on the hormone tests with Hugvie. Prof. Yoshio Iwai at Tottori University developed the vision system of Telenoid. Prof. Hiroki Yoshimura and Prof. Maiya Hori worked with Prof. Yoshio Iwai on the vision system. In addition to these researchers, our students at ATR, Osaka University, and Tottori University made contributions that are much appreciated. This project has been very successful and has obtained a good reputation, especially in Japan and Denmark. We could not have made this achievement without the efforts of all these collaborators. This work has been supported by JST/CREST.

References

1. M. McPherson, L. Smith-Lovin, M.E. Brashears, Social isolation in america: changes in core discussion networks over two decades. *Am. Sociol. Rev.* **71**(3), 353–375 (2006)
2. M.O. Ernst, H.H. Bulthoff, Merging the senses into a robust percept. *Trends Cogn. Sci.* **8**(4), 162–169 (2004)
3. L. Shams, R. Kim, Crossmodal influences on visual perception. *Phys. Life Rev.* **7**(3), 269–284 (2010)
4. M.H. Giard, F. Peronnet, Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cognitive Neurosci.* **11**(5), 473–490 (1999)
5. M.O. Ernst, M.S. Banks, Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**(6870), 429–433 (2002)

6. J.M. Hillis, M.O. Ernst, M.S. Banks, M.S. Landy, Combining sensory information: mandatory fusion within, but not between, senses. *Science* **298**(5598), 1627–1630 (2002)
7. H. Ishiguro, S. Nishio, Building artificial humans to understand humans. *J. Artif. Organs* **10**(3), 133–142 (2007). doi:[10.1007/s10047-007-0381-4](https://doi.org/10.1007/s10047-007-0381-4)
8. S. Nishio, H. Ishiguro, N. Hagita, Geminoid: Teleoperated android of an existing person. in *Humanoid Robots: New Developments* ed. by A.C. de Pina Filho (I-Tech Education and Publishing, Vienna, Austria, 2007) pp. 343–352
9. B.R. Duffy, Anthropomorphism and the social robot. *Robot. Auton. Syst.* **42**(3), 177–190 (2003)
10. B. Reeves, C. Nass, *The Media Equation* (Cambridge university press, Cambridge, 1996)
11. N. Matsumoto, H. Fujii, M. Okada, Minimal design for human-agent communication. *J. AROB* **10**(1), 49–54 (2006)
12. H. Osawa, J. Mukai, M. Imai, Anthropomorphization framework for human-object communication. *J. Adv. Comp. Intell. Intell. Inform.* **11**(8), 1007–1014 (2007)
13. M. Ogata, Y. Sugiura, H. Osawa, M. Imai, in *Proceedings of the Asia Pacific Conference on Computer Human Interaction*. Pygmy: a ring-shaped robotic device that promotes the presence of an agent on human hand (2012), pp. 85–92
14. J.M. Beer, L. Takayama, in *Proceedings of the International Conference on Human-robot Interaction*, Mobile remote presence systems for older adults: acceptance, benefits, and concerns, pp. 19–26 (2011)
15. P. Thacker, Physician-robot makes the rounds. *JAMA: J. Am. Med. Assoc.* **293**(2), 150–150 (2005)
16. M. Lee, L. Takayama, in *Proceedings of the Conference on Human Factors in Computing Systems*. Now, i have a body: uses and social norms for mobile remote presence in the workplace, pp. 33–42 (2011)
17. E. Paulos, J. Canny, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Prop: personal roving presence (1998), pp. 296–303
18. T. Kashiwabara, H. Osawa, K. Shinozawa, M. Imai, in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computer Systems*, Teroos: a wearable avatar to enhance joint activities, pp. 2001–2004 (2012)
19. Y. Tsumaki, Y. Fujita, A. Kasai, C. Sato, D. Nenchev, M. Uchiyama, in *Proceedings of the International Workshop on Robot and Human Interactive Communication*. Telecommunicator: a novel robot system for human communications, pp. 35 – 40, (2002). doi:[10.1109/ROMAN.2002.1045594](https://doi.org/10.1109/ROMAN.2002.1045594)
20. K. Ogawa, S. Nishio, K. Koda, G. Balistreri, T. Watanabe, H. Ishiguro, Exploring the natural reaction of young and aged person with telenoid in a real world. *J. Adv. Comput. Intell. Intell. Inform.* **15**(5), 592–597 (2011)
21. A. Kendon, Do gestures communicate? a review. *Res. Lang. Soc. Interact.* **27**(3), 175–200 (1994)
22. A. Mehrabian, *Silent messages* (Wadsworth Pub, Belmont, California, 1971)
23. K.R. Scherer, Vocal communication of emotion: a review of research paradigms. *Speech Commun.* **40**(1–2), 227–256 (2003). doi:[10.1016/S0167-6393\(02\)00084-5](https://doi.org/10.1016/S0167-6393(02)00084-5)
24. D.I. Perrett, D. Burt, I.S. Penton-Voak, K.J. Lee, D.A. Rowland, R. Edwards, Symmetry and human facial attractiveness. *Evol. Human Behav.* **20**(5), 295–307 (1999). doi:[10.1016/S1090-5138\(99\)00014-8](https://doi.org/10.1016/S1090-5138(99)00014-8)
25. I. Straub, S. Nishio, H. Ishiguro, in *IEEE International Symposium on Robot and Human Interactive Communication*. Incorporated identity in interaction with a teleoperated android robot: a case study, Viareggio, Italy, pp. 139–144, (2010). doi:[10.1109/ROMAN.2010.5598695](https://doi.org/10.1109/ROMAN.2010.5598695)
26. N. Yee, J. Bailenson, The proteus effect: the effect of transformed self-representation on behavior. *Hum. Commun. Res.* **33**, 271–290 (2007)
27. K. Kuwamura, K. Sakai, T. Minato, S. Nishio, H. Ishiguro, in *IEEE International Symposium on Robot and Human Interactive Communication*. Hugvie: a medium that fosters love (Gyeongju, Korea, 2013) pp. 70–75, doi:[10.1109/ROMAN.2013.6628533](https://doi.org/10.1109/ROMAN.2013.6628533)

28. J. Nakanishi, K. Kuwamura, T. Minato, S. Nishio, H. Ishiguro, in *the First International Conference on Human-Agent Interaction*. Evoking affection for a communication partner by a robotic communication medium, (Hokkaido University, Sapporo, Japan, 2013), pp. III-1-4
29. T. Minato, H. Sumioka, S. Nishio, H. Ishiguro, in *the RO-MAN 2012 workshop on social robotic telepresence*. Studying the influence of handheld robotic media on social communications, (France, Paris, 2012), pp. 15-16
30. H. Sumioka, K. Koda, S. Nishio, T. Minato, H. Ishiguro, in *IEEE International Symposium on Robot and Human Interactive Communication*. Revisiting ancient design of human form for communication avatar: design considerations from chronological development of dogu, Gyeongju, Korea, pp. 726-731, (2013a). doi:[10.1109/ROMAN.2013.6628399](https://doi.org/10.1109/ROMAN.2013.6628399)
31. S. Kaner, D. Bailey (2009) The power of Dogū: Ceramic figures from ancient Japan
32. MIHO Museum (ed.) Dogu, a Cosmos (Hatori Press, 2012)
33. C.T. Ishi, C. Liu, H. Ishiguro, N. Hagita, Speech-driven lip motion generation for tele-operated humanoid robots, *Processing of the the International Conference on Audio-Visual Speech* (Volterra, Italy, 2011), pp. 131-135
34. C. Liu, C.T. Ishi, H. Ishiguro, N. Hagita (2013) Generation of nodding, head tilting and gazing for human-robot speech interaction. *Int. J. Humanoid Rob.* **10**(1), 1350,009(1-19) (2013), doi:[10.1142/S0219843613500096](https://doi.org/10.1142/S0219843613500096)
35. K. Sakai, H. Sumioka, T. Minato, S. Nishio, H. Ishiguro, Motion design of interactive small humanoid robot with visual illusion. *Int. J. Innovative Comput. Inf. Control* **9**(12), 4725-4736 (2013)
36. Y. Fujie, M. Hori, H. Yoshimura, Y. Iwai, in *Proceedings of the HRI2013 Workshop on Design of Humanlikeness in HRI from uncanny valley to minimal design*, Emotion transmission by color effects for a teleoperated mobile communication robot, pp. 19-25 (2013)
37. R. Plutchik, *Emotion: A Psychoevolutionary Synthesis* (Harper and Row, New York 1980)
38. K. Tanaka, H. Nakanishi, H. Ishiguro, *Appearance, Motion, and Embodiment: Unpacking Avatars by Fine-Grained Communication Analysis* (Practice and Experience, Concurrency and Computation, 2015a). doi:[10.1002/cpe.3442](https://doi.org/10.1002/cpe.3442)
39. K. Tanaka, H. Nakanishi, H. Ishiguro, Physical embodiment can produce robot operator's pseudo presence. *Frontiers in ICT* **2**(8), (2015). doi:[10.3389/fict.2015.00008](https://doi.org/10.3389/fict.2015.00008)
40. K. Kuwamura, T. Minato, S. Nishio, H. Ishiguro, Personality distortion in communication through teleoperated robots, *IEEE Int. Symp. Robot Hum. Interact. Commun.* (France, Paris, 2012), pp. 49-54
41. H. Gough, A. Heilbrun, *The adjective check list manual* (Consulting Psychologists Press, Palo Alto, 1983)
42. S. Wada, Construction of the Big Five Scales of personality trait terms and concurrent validity with NPI. *Jpn. J. Psychol.* **67**(1), 61-67 (1996). in Japanese
43. D.G. Dutton, A.P. Aron, Some evidence for heightened sexual attraction under conditions of high anxiety. *J. Pers. Soc Psychol.* **30**(4), 510-517 (1974). doi:[10.1037/h0037031](https://doi.org/10.1037/h0037031)
44. N. Nishimura, A. Ishi, M. Sato, S. Fukushima, H. Kajimoto, in *CHI '12 Extended Abstracts on Human Factors in Computing Systems*. Facilitation of affection by tactile feedback of false heratbeat, pp. 2321-2326 (2012). doi:[10.1145/2212776.2223796](https://doi.org/10.1145/2212776.2223796)
45. H. Sumioka, A. Nakae, R. Kanai, H. Ishiguro, Huggable communication medium decreases cortisol levels. *Sci. Rep.* **3**(3034), (2013b). doi:[10.1038/srep03034](https://doi.org/10.1038/srep03034)
46. D.H. Hellhammer, S. Wust, B.M. Kudielka, Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology* **34**(2), 163-171 (2009)
47. S. Cohen, Social relationships and health. *Am. Psychol.* **59**, 676684 (2004)
48. B. Ditzen, I.D. Neumann, G. Bodenmann, B. von Dawans, R.A. Turner, U. Ehler, M. Heinrichs, Effects of different kinds of couple interaction on cortisol and heart rate responses to stress in women. *Psychoneuroendocrinology* **32**(5), 565-574 (2007)
49. T. Field, M. Hernandez-Reif, M. Diego, S. Schanberg, C. Kuhn, Cortisol decreases and serotonin and dopamine increase following massage therapy. *Int. J. Neurosci.* **115**(10), 1397-1413 (2005)

50. J. Holt-Lunstad, W.A. Birmingham, K.C. Light, Influence of a “warm touch” support enhancement intervention among married couples on ambulatory blood pressure, oxytocin, alpha amy-lase, and cortisol. *Psychosom. Med.* **70**(9), 976–985 (2008)
51. K.M. Grewen, S.S. Girdler, J. Amico, K.C. Light, Effects of partner support on resting oxytocin, cortisol, norepinephrine, and blood pressure before and after warm partner contact. *Psychosom. Med.* **67**(4), 531–538 (2005)
52. K.C. Light, K.M. Grewen, J.A. Amico, More frequent partner hugs and higher oxytocin levels are linked to lower blood pressure and heart rate in premenopausal women. *Biol. Psychol.* **69**(1), 5–21 (2005). doi:[10.1016/j.biopsycho.2004.11.002](https://doi.org/10.1016/j.biopsycho.2004.11.002). Current Trends in Women’s Health Research
53. J.L. Edens, K.T. Larkin, J.L. Abel, The effect of social support and physical touch on cardiovascular reactions to mental stress. *J. Psychosom. Res.* **36**(4), 371–381 (1992)
54. W.J. Nilsen, S.R. Vrana, Some touching situations: the relationship between gender and contextual variables in cardiovascular responses to human touch. *Ann. Behav. Med.* **20**(4), 270–276 (1998)
55. P.E. Bee, P. Bower, K. Lovell, S. Gilbody, D. Richards, L. Gask, P. Roach, Psychotherapy mediated by remote communication technologies: a meta-analytic review. *BMC Psychiatry* **8**, 60 (2008)
56. R. Yamazaki, S. Nishio, K. Ogawa, H. Ishiguro, Teleoperated android as an embodied communication medium: a case study with demented elderlies in a care facility, *IEEE Int. Symp. Robot Human Interact. Commun.* (France, Paris, 2012), pp. 1066–1071
57. R. Yamazaki, S. Nishio, H. Ishiguro, M. Nørskov, N. Ishiguro, G. Balistreri, Acceptability of a teleoperated android by senior citizens in Danish society: a case study on the application of an embodied communication medium to home care. *Int. J. Soc. Robot.* **6**(3), 429–442 (2014). doi:[10.1007/s12369-014-0247-x](https://doi.org/10.1007/s12369-014-0247-x)
58. R. Yamazaki, S. Nishio, K. Ogawa, K. Matsumura, T. Minato, H. Ishiguro, T. Fujinami, M. Nishikawa, Promoting socialization of schoolchildren using a teleoperated android: an interaction study. *Int. J. Humanoid Rob.* **10**(1), 1350,007(1–25) (2013), doi:[10.1142/S0219843613500072](https://doi.org/10.1142/S0219843613500072)
59. J. Nakanishi, H. Sumioka, M. Shiomi, D. Nakamichi, K. Sakai, H. Ishiguro, in *Proceedings of the Second International Conference on Human-agent Interaction*. Huggable communication medium encourages listening to others, pp. 249–252 (2014). doi:[10.1145/2658861.2658934](https://doi.org/10.1145/2658861.2658934)
60. A.D. Wolvin, *Listening and Human Communication in the 21st Century* (Wiley Online Library, 2010)
61. A. Wolvin, C. Coakley, *Listening*. (McGraw Hill, New York, 1996)
62. K. Nowak, in *Presence 2001 4th Annual International Workshop*. Defining and differentiating copresence, social presence and presence as transportation (2001)