AI & SOCIETY (2022) 37:1563–1573
https://doi.org/10.1007/s00146-021-01243-8

OPEN FORUM

Robot and *ukiyo-e*: implications to cultural varieties in human–robot relationships

Osamu Sakura[1,2]

Received: 9 April 2021 / Accepted: 21 June 2021 / Published online: 4 October 2021

Abstract
The social and cultural causes behind the widespread use and acceptance of robots in Japan are not yet completely understood. This study compares humans and robots in images gathered through Google searches in Japanese and in English. Numerous pictures obtained by the search in Japanese were found to have a human and a robot looking together at something else ("third item"), whereas many of the images acquired by search in English show a human and a robot facing each other. This is similar to the composition of mother and child in paintings: in *ukiyo-e* that was painted mainly in the Edo period of Japan, the mother and child are often depicted together viewing something other than themselves, whereas this is not the case in Western paintings of mother and child. It has also been pointed out that, in modern Western paintings, the world inside the picture is separated from the outside world, forming an independent microcosmos, whereas the inside and outside are continuous in Japanese paintings. These may indicate that, in Japanese society, robots are to a certain extent regarded as fellow human beings who can share the third item. In Western society, on the other hand, no code is embedded that can fix robots' superiority or inferiority to humans, which would easily trigger an antagonistic view toward artificial intelligence (AI)/robots as threatening entities, as shown in most of Western literature and movies. We suggest that such cultural characteristics of Japanese society can contribute to enhance coexistence with AI/robots.

Keywords Robots · Humans · Relationship · Cultural differences · Direction of gaze · Iconography

1 Introduction

Japan is one of the leading countries in the world in robot use, but why this is so has not yet been convincingly explained. Japan had the largest number of industrial robots in operation in the world until 2015, when China took the top spot (Japan Robot Association 2019). South Korea ranks third, so East Asian countries occupy the top three places in the world for robot use, which suggests that the cultural and social background of the East Asian region may influence the social acceptance of robots. In terms of robot density per 10,000 employees, the top three nations are Singapore (where Chinese people are the main ethnic group), South Korea, and Japan (International Federation of Robotics 2020).

Exploring the cultural influences behind the social acceptance of technology is also necessary to predict future trends in science and technology. The process and factors for the diffusion and establishment of emergent technologies such as artificial intelligence (AI) and robotics in society cannot be properly understood by looking only at the development and supply side of the technology. This is because, as a consequence of such technologies, other functions that developers did not anticipate have emerged, creating unexpected situations (Sakura 2019, 2020). As a result of such actions from the social side, changes are made to the technologies. Such mutual interactions between society and technology are called the "social shaping of technology (SST)" (MacKenzie and Wajcman 1999). Variables that significantly impact on technology in SST include the cultural and historical background of the society. In recent years, the importance of ethical, legal, and social issues (ELSIs) is often emphasized when considering the social diffusion of emergent technologies, including AI and robotics. However, considering the

✉ Osamu Sakura
osamu.sakura@riken.jp; sakura@iii.u-tokyo.ac.jp

[1] Science, Technology and Society Research Team, RIKEN Center for Advanced Intelligence Project (AIP), Tokyo, Japan

[2] Interfaculty Initiative in Information Studies, The University of Tokyo, Tokyo, Japan

SST process in which technology and society interact with each other, it is not sufficient to look at ELSIs alone because the cultural and historical aspects of society and the ecological dynamics of technology are also important (Sakura 2019, 2020).

The public image and public attitude toward certain emergent technologies provides insights into such dynamics. Previous studies have made international comparisons of the public images and public attitudes of robots have (Bartneck 2008; Bartneck et al. 2007; Haring et al. 2014; MacDorman et al. 2008; Nomura 2017; Nomura et al. 2015; Shibata et al. 2004; Syrdal et al. 2013; Trovato et al 2013; Trovato and Eyssel 2017). However, most of these are empirical studies examining cultural differences, rather than the possible causes of such differences. Most of these studies have found no significant cultural differences in dislike or attachment to robots. Moreover, contrary to popular belief, the hypothesis that robots are popular in Japan because Japanese people are attached to robots has not been confirmed—in fact, a few studies report that dislike of robots is slightly stronger in Japan than in Western countries (Bartneck et al. 2007; Haring et al. 2014; Nomura et al. 2015; Nomura 2017). Besides, most of the comparisons are between Japan and the "West," with only a few comparative studies conducted in the East Asian cultural sphere.

Complementing these quantitative studies are qualitative considerations, such as Frumer (2018) examining the social perception of robots in Japan in a cultural-historical context; Kovacic (2018) discussing the relationship with Japan's "manufacturing" culture; Ito (2010) and Nakao (2014) analyzing pop culture, such as science fiction novels and *manga*, from a science-historical perspective as the root of robot acceptance in Japan; whereas Nishigaki (1988, 2018) points out differences in religious backgrounds between East and West. Allison (2006) has proposed the notion of "Techno-Animism" as an explanation of how the animistic mindset in Japan extends to the field of technology, with Helmreich (1998/2000) discussing the possibility of cultural pluralism in artificial life research in light of the Western (more specifically, Caucasian and masculine) bias in the concept of life. Kaplan (2004) compares the Japanese and western cultures to point out that nature and technology are not treated as opposites in Japan, and innovative technology has been "tamed" by this traditional relationship between technology and nature. On the contrary, in the west, nature and artifacts are in opposition to each other, and this is expressed in the difference in the way robots are perceived in the two cultures. All these studies are important and interesting but, until now, there have been few perspectives that consider the linkages between actual human relationships in society and robot acceptance.

Based on these previous studies, this paper assumes that the ready acceptance of robots in Japan is related to the norms of interpersonal relationships in Japanese society, and examines the theoretical framework and methodology to verify this assumption.

It is not unreasonable to assume that the perception of the relationship between people and robots is reflected in images such as photographs and paintings. Iconographical analysis is a well-established method in historical ethnography, which uses paintings of people to infer what kind of social perceptions operated in the past (e.g., Ariès 1960). This is because the perception of the person who painted the picture or took the photograph is unconsciously reflected in the composition. Therefore, an analysis of the composition of photographs and/or paintings of people and robots together allows a cultural comparison of their public images. This paper makes a preliminary analysis of the composition of photos and illustrations on the web showing people and robots together and how they are positioned.

When people are facing each other, the direction they look to, whether together or independently, reflects their interpersonal relationships and their respective cognitive abilities. Newborns in the first few months of life are unable to follow their mothers' gaze when the mother looks at another object (Moore and Dunham 1995; Reddy 2005; Scaife and Bruner 1975). The newborn is unable to recognize that the object of the mother's gaze is relevant to them. This is because looking in the same direction as a communication partner requires similar cognitive capacity. The Japanese psychiatrist Kitayama (2005a) extended this joint-attention theory to a more general framework for comparative cultural study of parent–child and interpersonal relationships, renaming it the "viewing together hypothesis." He pointed out that a communication pattern in which two people share a *ba*, or field, by gazing at a common third term and reaffirming the strong relationship between them is unique to Japanese interpersonal relationships. This paper applies Kitayama's theory to the relationship between humans and robots and discusses the cultural differences in social epistemology toward robots.

The second tool this paper utilizes to analyze human–robot relations is an examination of the relationship between the depicted world inside the painting and the real world outside. In Japanese painting, the boundary between the inner world and the outer world is not as clear as in Western painting, and the frame of a picture does not function as a strict border (Sasaki 2013; Takashina 1996). The use of linear perspective is also different (Kishi 1994; Sasaki 2013). This suggests that Japanese painting may invite the viewer to have a more intimate relationship with the inner world of the painting. The same argument may be applied to paintings and photographs of humans and robots.

In the following, the relationship between humans and robots will be analyzed using these two approaches.

Table 1 The composition pattern of humans and robots appeared in pictures found in the Google search using the two sets of the keywords (see text) in English or Japanese

Pattern of composition		Keyword Set 1			Keyword Set 2		
		English	Japanese	Total	English	Japanese	Total
humans + robots	opposite (a)	39	12	51	24	11	35
	same (b)	12	21	33	7	16	23
	multi (c)	3	3	6	6	4	10
	unclear (c)	24	8	32	8	8	16
robot only		15	50	65	52	59	111
human only		7	6	13	3	2	5
Total		100	100	200	100	100	200

The patterns of pictures with both humans and robots were significantly different for both Sets 1 and 2 (cells enclosed by dotted line). Parenthesized alphabets in the column "Pattern of composition" represent correspondence patterns shown in Fig. 1

2 Empirical data of differences in the compositions in the pictures and illustrations of humans and robots

2.1 Materials and methods

To investigate the composition of photos and illustrations of people and robots on the web, two Google searches were conducted on February 8 and June 12, 2021, using "robot, human, relationships" (Set 1) and "robots, people, interpersonal" (Set 2) as the sets of keywords. The same sets of keywords were searched for in English and Japanese ("ロボット, 人, 関係" [Set 1] and "ロボット, 人々, 対人関係" [Set 2], respectively) and the results were compared.

The retrieved photos or illustrations were collated in the order in which they appeared in the search results, excluding those that showed neither people nor robots, for a total sample of 100 photos (or illustrations) that showed either people or robots.

The photos (or illustrations) were divided into those in which a person and a robot were depicted together and those in which only the robot or the person was depicted. The former were further divided into three categories based on the gaze of the person and the robot: (1) the person and the robot are facing each other; (2) both looking in the same direction, including at the viewer and multiple patterns; and (3) those in which the direction of gaze is ambiguous and difficult to judge. Each pattern was quite clear and there was no ambiguity in classification.

2.2 Results

Table 1 summarizes the results. In the case of images found through Set 1 of the English keywords, there were 78 photos or illustrations of people and robots together, 54 of which had a clear direction of gaze. Of these, 39 were people and robots facing each other (Fig. 1a); 12 were people and robots looking in a common direction (Fig. 1b); and 3 showed multiple people and robots with multiple gaze patterns (Fig. 1c). In 24 images, it was difficult to determine the line of gaze of the person and the robot (Fig. 1c). In 15 cases, only the robot and no person was included, and in 7 cases, only the person and no robot was included.

In the searches where Japanese was used as Set 1 of the keywords, there were 12 cases in which people and robots were shown or depicted together facing each other (Fig. 1a); there were 21 cases with people and robots looking in a common direction (Fig. 1b); and there were 3 cases in which multiple people and robots were shown and multiple gaze patterns were included (Fig. 1c). In eight cases, the direction of gaze was not clear, and it was difficult to determine the line of gaze of the person and the robot (Fig. 1c). There were 50 cases in which only the robot was included in the composition without a person, which was the largest number of photos and illustrations found in the Japanese search. By

Fig. 1 **a** Human and robot looking at each other (copyright free), **b** looking the "third item" together (with permission of Professor Hiroshi Ishiguro), **c** multiple and/or hard to identify (drawn from Creative Commons, by Paul Houle, https://www.flickr.com/photos/paul_houle/4379880536)

contrast, a mere six cases did not include robots. Of the 50 robot-only compositions that did not include people, 2 were of robots and pet animals, which was not found in English searches.

On comparing the results between the English and Japanese searches, the distribution patterns were significantly different ($p < 0.0001$, $\chi 2 = 42.4$, df = 4; Table 1, "multi" and "unclear" were combined together because of the low frequency in the "multi" pattern). Another comparison of the differences in the cases in which both humans and robots were depicted (cells enclosed in dotted lines under the column "Keyword Set 1" in Table 1) showed a significant difference between the results ($p < 0.0001$, $\chi 2 = 16.7$, df = 1; Table 1).

The research using the second set of keywords showed similar results (Table 1). The search in English found 24 cases of opposing configurations of humans and robots and 7 cases of same direction gazes, whereas the search in Japanese resulted in 11 cases of opposing configurations and 16 cases of same direction gazes, which was a significant difference between the results ($p < 0.005$, $\chi 2 = 16.8$, df = 1; cells enclosed by dotted lines under the column "Keyword Set 2" in Table 1). However, the differences between the total pattern in the searches using English and Japanese keywords were not apparent ($p > 0.05$, $\chi 2 = 9.15$, df = 4; Table 1).

To summarize these results, there were more pictures in which a person and a robot were facing each other when searching in English, whereas in the Japanese search, a person and a robot were facing the same direction in most of the images. The results do not seem to be artifacts of Google's algorithms, because there are no reasonable factors affecting the relation between the composition and usage of languages. Even if the results were produced with some biases in the algorithms, those are most likely to have been caused by characters from data used to train the program that reflect tendencies observed in real human society.

3 Interpretations from the theory of "Looking Together"

The above results are sufficient to reveal differences between the two cultures toward human–robot relations. The difference can be interpreted as the relationship between the human and the robot being complete within the two actors or being open to other agents, including the photographer and the viewer. The fact that the composition is open to something other than the two actors means that people and robots in the pictures share the "third term." In other words, the results of the English search suggest that the human–AI/robot relationship is considered as a binary relationship without the third item, whereas the results of the Japanese search suggest that it is viewed as a ternary relationship.

In developmental psychology, joint attention, in which mothers and children not only look at each other but also focus on a common third term, is considered an important milestone in the mother–child relationship (Moore and Dunham 1995; Reddy 2005; Scaife and Bruner 1975). This usually happens around 9 months of age, when the infant's cognitive abilities undergo accelerated development (Endo 2006). This suggests that the Japanese composition shows that AI/robots are treated as "something like fellows that are capable of paying attention to the common third term together." That is probably how the photographer and the human subjects unconsciously feel.

Kitayama (2005b), a psychiatrist who investigated the depiction of mother–child relationships in Japanese *ukiyo-e* prints, reported that about 20% (or about half, if we broaden the definition a little) of 213 *ukiyo-e* mother and

Fig. 2 A pair of mother and infant are gazing together the "third item," drawn in traditional Japanese painting, *ukiyo-e*, by Shôun Yamamoto (1870–1965). *Ima-sugata: Goran nasai (A Beautiful Woman: Look at This)* (1907) (copyright expired, drawn from the website of the National Diet Library Digital Collection, Japan)

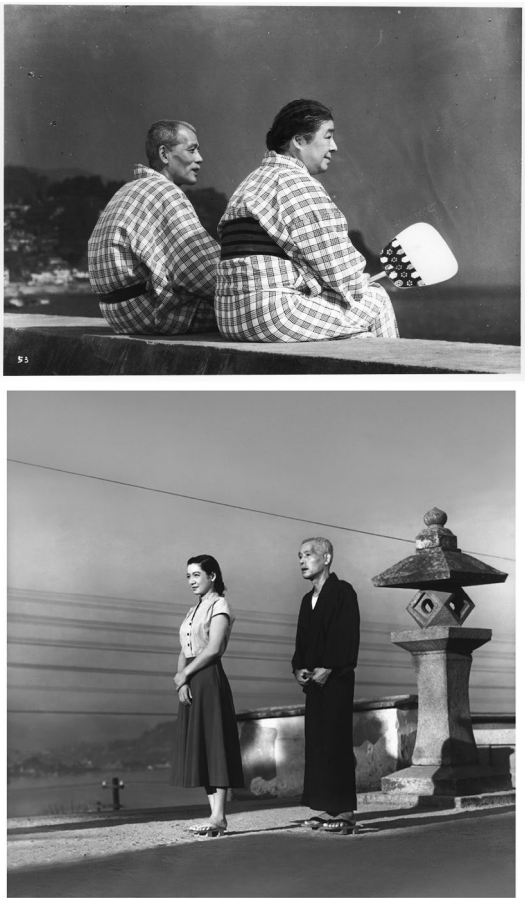

Fig. 3 Scenes from *Tokyo Monogatari (TOKYO STORY,* 1953). Director: Yasujiro OZU. Characters are not looking at each other, rather viewing together the "third item/direction." ©1953 Schochiku Co., Ltd. Reproduciton prohibited

child images demonstrate joint attention of viewing (Fig. 2), which is significantly more than the 5% or so of similar compositions in Western paintings (see also Yamaguchi 2005). Kitayama introduces a new term, "viewing together," to describe the joint attention of mother and child depicted in this way, applying it to the analysis of human relationships, including adults in Japanese society. In many cases, the third items that are the focus of the shared gazes of mother and child are transient and fleeting presences, such as falling flower petals or the hazy moon, and this may be the reason for the features of human relationships in Japanese society: at best, they are sensitive and subtle, and at worst, they require "reading the atmosphere" and others' feelings. The same kind of "viewing together" composition has been used effectively to depict relationships among adults in films by Japanese director Yasujiro Ozu (Miura 2005; Yamada 2005) (Fig. 3). This supports Kitayama's considerations that "viewing together" may be deeply related to the characteristics of human relationships in Japanese society in general.[1]

Abe (1995), a Western historian in Japan, developed a discussion of the characteristics of Japanese society in which interpersonal relationships are based on concern for the actions of third parties, using the keyword "*seken* (世間)." Here, the word "*se* (世)" means "world," and "*ken* (間)" means "in between" or "relation." According to Abe, *seken* differs from the Western "public" space and "society" because it is not a relationship based on individual or private

[1] Shigehiko Hasumi (1983/2003), one of the most powerful advocates of Yasujiro Ozu, criticizes the use of the term "Japanese" to describe Ozu's films. He points out that Ozu's filmmaking techniques are not confined to such a local cultural context. This is true, but Hasumi's point is about Ozu's grammar in film making, not about the human relationships and social characteristics depicted in his films. Ozu genuinely succeeded to depict his topics of human relations in truly realistic manners, especially those within the family of 1950s Japan, even including paternalism and misogyny.

space. *Seken* is rather vague and difficult to verbalize, but it has a strict power to constrain the actions and thoughts of individuals. Western society is constituted by the individual, but the Japanese individual is defined by *seken*. However, the causal relation is not clear: is it because the mother–child relationship is dominated by the sympathetic relationship of "viewing together" that we develop ambiguous relationships such as *seken*? Or is it because the *seken* relationship is embedded in Japanese society in the first place that Japanese mother–infant share the third item to empathize with each other? Or do the two mutually reinforce each other in a loop effect?

At any rate, it can be said that the composition of Japanese pictures, in which people often share their gazes with those of robots, unconsciously assumes that robots are agents capable of sharing attention to the third item—that is, they are fellow members of the same *seken*. On the other hand, the composition in Europe and the United States, where people and robots are often placed facing each other, may reflect a view that robots are in a subordinate relationship with people, rather like pet animals.

4 Interpretations from the relation of in- and out-worlds of paintings

Tanaka (2005), a researcher of Edo culture, argues that the "viewing together" relationship depicted in *ukiyo-e* is intended to draw the viewer into the world of the painting.[2] Tanaka's argument, which is central to this essay, is that perspective drawing, which was introduced into Japan from Europe around the eighteenth century, was not used in Japan to enhance realism but to realize "the intention of connecting the world in the picture with the world of the viewer" (Tanaka 2005, p.49). In other words, Western art appreciation places the subject (the viewer) and the object (the painting) as two separate parties (binary relationship), whereas in Japanese art appreciation the two are not independent and the painting is created in such a way that the viewer is encouraged to "enter the painting through the act of participation in the scene and assimilation to the characters" 2005 (Tanaka 2005, p.69).

Tanaka's argument is based on art historian Kishi's (1994) analysis of *uki-e*, which is confusingly similar to *ukiyo-e* and is actually a type of *ukiyo-e*, but embodies enhanced realism using Western perspective drawings. This technique was named *uki-e*, which means "pop-up," because it emphasized perspective and made the foreground appear to be floating. *Uki-e* was a popular form from the first half of the eighteenth century to the mid-nineteenth century. Kishi (1994), focusing on the works of the painter Masanobu Okumura (1686–1764), notes that *uki-e*, which mainly depicts the *kabuki* stage, mixes the newly introduced Western perspective drawing (linear perspective) and the ancient Japanese non-linear perspective drawing (parallel, or bird's-eye view drawing). Parallel drawing is a method in which parallel lines do not terminate at a vanishing point (Kishi 1994). It is commonly believed that Western linear perspective drawing and parallel drawing were mixed in one painting simply because Masanobu Okumura and other Japanese painters of the time did not acquire proficiency in Western perspective techniques. However, Kishi argues that linear perspective drawing and parallel drawing were deliberately used in different ways. Parallel drawing was used to portray the facial expressions and other details of the actors on the *kabuki* stage, whereas linear perspective was used to illustrate the excitement of a large audience (Fig. 4).

Through this technique, the eighteenth century Japanese *uki-e* tried to incorporate the viewer into the inner world of the painting. In Japanese painting techniques, the world inside the painting and the world outside the painting are treated as a continuum (see also Sasaki 2013).

Takashina (1996), another art historian, calls this "shifting viewpoints." In Western painting, the "painter's ego" is the fixed point of view and the painting is drawn in a cut-out form within a certain frame of the canvas—thus the painting exists as something independent of its surroundings. In Japanese painting, by contrast, the artist's point of view is not fixed but moves around in space according to the artist's interest, sometimes enlarging parts of the object and, in some cases, freely going back and forth along the timeline. All of the materials gathered from every scene are put together into the same plane, as a single picture. The work is not completely independent from the outer world as a painting, and the frame that separates the painting from its surroundings does not function as a complete border between the inner and outer worlds. "In fact, the frame of the work is determined by the actual screen, such as *byobu* (a folding screen) or *fusuma* (sliding door), but the world expressed within the frame is not completed there, but tends to expand outward beyond the frame" (Takashina 1996, p.21).

Figure 5 shows a piece of *ukiyo-e*, painted by Hiroshige Utagawa (1797–1858), one of the great masters of *ukiyo-e*, which is titled as *Haneta no Watashi: Bennten no Yashiro (A Passenger Boat in Haneda: Benten Shrine)*, comprising one of the serials of *Meisho Edo Hyakkei (One Hundred Famous Views of Edo)*. The painting depicts a ferryboat moving toward the shrine on the other side of the river, but the body of the rowing boatman is not represented inside the painting except for a part of his feet and his hands. His body

[2] Tanaka (2005) also pointed out that the relationship between mother and child drawn in *ukiyo-e* is a substitute for the adult male–female relationship, which is related to the pornographic function of the paintings.

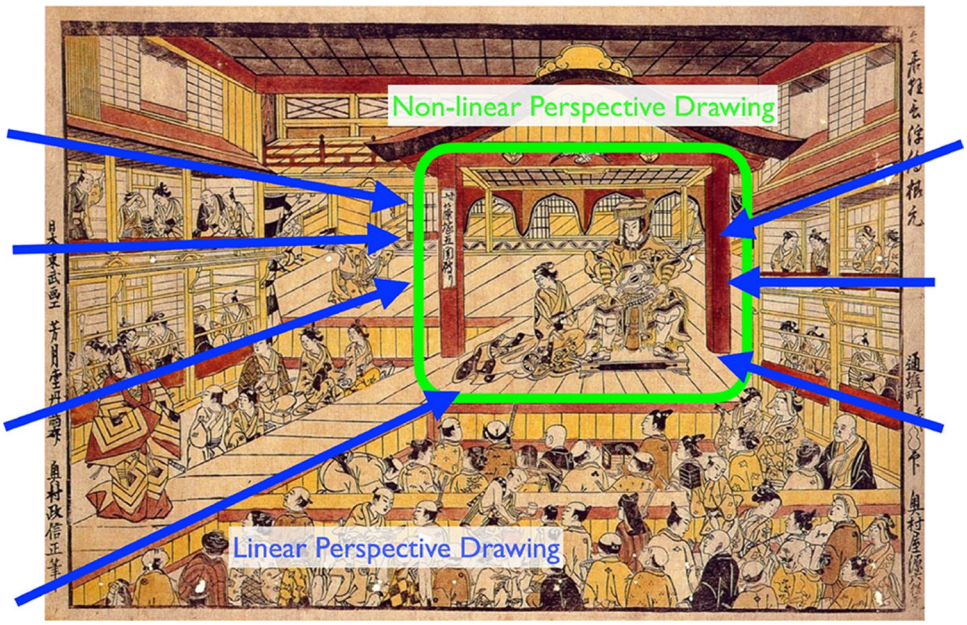

Fig. 4 Masanobu Okumura's *Shibai Kyôgen Uki-e Nemoto (Japanese Comic Short Drama, Kyôgen*, Drawn in the style of pop picture, *uki-e*) (1743?), with analysis about the usage of perspective drawings, by Kishi (1994) (Copyright of original painting has expired. Other symbols were added by the author)

is supposed to exist in the space outside the painting, implying that the inside and outside of the painting are contiguous. This implies that there is a continuity between the inside and the outside of the painting, even evoking the image of the viewer (= we) who are outside the painting, as if the viewer were riding on this boat. Such a motif of "continuity between outside and inside" is often seen in Japanese paintings in, for example, "hanging down" where branches of trees and bunches of plants fall from outside above the picture. This approach was not observed in Western painting until Claude Monét applied the motif in his *Les Nymphéas* in the late nineteenth century (Takashina 1996). Monét, notably, was known to be greatly influenced by Japanese *ukiyo-e*.

Takashina (1996) also points out that the scroll form of painting, which flourished in Japan, represents the fact the irrelevance of the frame in Japanese painting. Although the scroll form was not completely absent in Western painting, it was not a favored medium. However, scrolls have been always important in the history of Japanese painting, especially in biographical depictions of the lives of great figures and travelogues, suggesting that the method of viewing a work as a continuous story was well established.

Applying these concepts to photographs in which both humans and robots are the subjects, it seems that the composition in which humans and robots face each other side-by-side does not only indicate that they are viewing the third item together, but is also an invitation to the viewer to be a component of their world. In other words, the robot is not only on the same level as the human subject, but also on the same level as the viewers or close to such.

5 General discussions and perspectives

5.1 Dominance and subordinate relation between humans and robots

Based on what has been described above, it seems that in Japanese photographs and illustrations, robots are treated as some kind of comrades to human beings. They are entities that can stare at the third item together and share its meaning; moreover, they belong to the world that includes the viewer. It is uncertain, of course, if the photographers and illustrators intentionally designed the images in this way, but such a sense is nonetheless reflected in the composition, even if unconsciously. This explanation is consistent with the results of previous studies showing that compared to Westerners, Japanese people are more favorable to humanoid robots than non-humanoid ones (Haring et al. 2014), whereas European people expect robots to aid them in practical tasks (Ray and Mondada 2008).

This is also supported by the fact that when searching in Japanese, the most common composition included no people and only had robots although the search was conducted using the term "humans." There is hardly any reasonable foundation to assume that Google's algorithms reacted differently enough to cause such a huge dissimilarity between English and Japanese. This suggests that, in Japan, robots are regarded as somewhat independent entities, and it is not always necessary to talk about them in relation to people. It also implies that the distance between robots and companion animals is closer in the Japanese image because, in two cases out of 50 that included no humans, robots and pets were

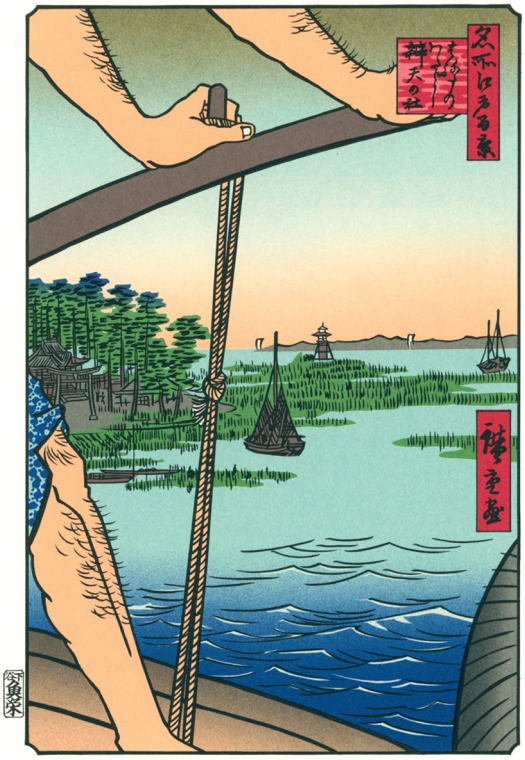

Fig. 5 Hiroshige Utagawa's (1797–1858) *Haneta no Watashi: Bennten no Yashiro (A Passenger Boat in Haneda: Benten Shrine)*, one of the serial of *Meisho Edo Hyakkei* (*One Hundred Famous Views of Edo*) (copyright expired)

depicted together in the Japanese search, whereas no such composition was found in the English search.

The fact that they are almost equal to each other means that there is no great difference in superiority or inferiority between humans and robots. They are "fellows" of similar rank and belong to the same "world." This is in line with Kaplan's (2004) assertion that in Japan, rather than confronting new technologies, people have been taming and coexisting with them ("technology taming"). The technological singularity theory (Kurzweil 2005), which predicts that the capabilities of AI will surpass humanity's in the mid-twenty-first century and that humanity will eventually be dominated by AI, plus the arguments that humanity faces the risk of extinction by AI (e.g., Bostrom 2014), do not seem to have had as much resonance in Japan as in Western society. This may arise partly from the differing view of robots. It is also possible to imagine the opposite direction of causality that the difference revealed here came from hostility to, and fear of, robots in Western society.

It is difficult to identify the reasons for such differences. One possibility is differences in both culture's views of nature and the distance between humans and animals. Photos of humans with apes and monkeys in Japan and the West seem to represent cultural differences similar to those observed in photos with robots. Japanese primatologists are often in "viewing together" relationships with gorillas and chimpanzees, whereas European and American researchers are more likely to be face-to-face with apes (unpublished preliminary data by the author). If this is the case, it suggests that there may be similarities between the "human–robot relation" and the "human–ape relation." Both robot and ape are "human-like but non-human" beings. It is reasonable to speculate that there is a common cultural difference in how we perceive our relationship with such neighbors. Japanese primatology has explored the complexity of society of non-human primates using a rather anthropomorphic approach to monkeys and apes, which could never be revealed in the mechanical paradigm that has dominated Western natural science (Fedigan and Asquith 1991; de Waal 2001; Sakura 2010). In Japan, a stronger anthropomorphism may also be applied for robots. More systematic analysis is required for further discussion, but the method used in this paper is difficult to apply for analysis of the composition with humans and monkeys because the tern "relation" fetched illustrations showing the phylogenetic relationship between humans and non-human primates. An alternative method is, therefore, needed.

Nonetheless, it is safe to say that this may be due to differences in Japanese and Western views of nature. It is generally accepted that the East Asian view of nature, including the Japanese, as represented by Taoism, emphasizes the unity and continuity of nature and human beings. This contrasts with the Judeo-Christian view, which emphasizes the differences between the two, and moreover, places man as dominating nature (Asquith and Kalland 1997). American zoologist Morse (1917), who was instrumental in introducing Charles Darwin's theory of evolution to Japan, gave several public lectures on Darwinism. He was surprised at the almost total lack of anti-evolutionary reactions from the Japanese public (Morse 1917). In the US, Morse's supervisor at the Peabody Museum of Natural History at Harvard University, Louis Agassiz, was a strong anti-Darwinist, so Morse inevitably experienced a lot of friction. Japanese primatologists were less resistant to the anthropomorphic treatment of monkeys and apes, which may be because the Japanese were less influenced by Christian culture (Umesao 1960). Anthropomorphism has been often criticized as "unscientific," but Japanese primatologists have also revealed that non-human primates have many characteristics in common with humans, such as complex society, proto-cultural behaviors, and proto-language abilities (Fedigan and Asquith 1991; de Waal 2001).

5.2 Limitations of this study and future perspectives

Some issues need to be addressed to develop a more accurate discussion from the perspective described in this study.

The first requirement is a more systematic investigation of the use of perspective and the way of viewing paintings in Japan. As the author does not have professional skills in this area, it is necessary to seek input from appropriate experts.

Second, it is necessary to deepen the consideration of animism. Suwa (1998) emphasizes the relationship between animism and Japanese painting and viewing methods based on the shifting of viewpoint, but his argument is hardly convincing. A more systematic examination is necessary. On the other hand, Allison (2006), an anthropologist who analyzed Japanese pop culture, found a link between cutting-edge technology and animism in Japanese society, naming this "techno-animism" as a major feature of contemporary Japanese culture. Although this is an interesting point, the claim needs to be examined more closely.

The third requirement is a more careful examination of Western thought and views of nature. In the Western tradition of knowledge, there is a deep-rooted way of thinking that emphasizes the close relationship between the subject and the environment, rather than a clear distinction between subject and object (e.g., Henri Bergson and Jakob von Uexküll). There are also many discourses that focus on the importance of embodied knowledge (e.g., M. Merleau-Ponty). It is necessary to re-examine how these discourses can be placed in the context of human–AI/robot relations.

Fourth, it is necessary to study the cultural differences in the public image and social perception of robots *within* East Asian countries (Japan, Korea, Taiwan, and China). These regions share common natural ecological environments based on forests and the same social roots in ancient Chinese culture. It is necessary to conduct a systematic comparison based on the characteristics of the cultural sphere, beyond a limited comparative study of "Japan versus Euro-America."

More systematic comparative cultural researches will provide a more constructive perspective on the relationship between humans and robots, based on the ecological and cultural characteristics of East Asia. Wang et al. (2010) reported that robot users have greater trust in robots that behave according to their cultural norms (e.g., suggesting implicitly rather than explicitly) and collaborate with such robots more effectively. In addition, Okada and his colleagues (Okada 2012, 2017; Yamaji et al. 2011) have shown that "weak" robots with some deficiencies facilitate easier communication with humans compared to robots designed to be more powerful and to enhance humans. This idea may have come from the Japanese view of robots as "friends." A better symbiotic relationship between humans and robots can be fostered by developing the suggestions in this paper.

Acknowledgements I am grateful to Jun Shozawa for his suggestion of the "viewing together" theory. I also appreciate Hiroshi Nakagawa, Shin'ichi Fukuzumi, Yueh-Hsuan Weng, Emitiyaz Kahn for their helpful comments and discussions, as well Takuya Mizukami, Yoko Fujishima, Haruka Maeda, Tomohiro Inokuchi. I am equally grateful to Hiroshi Ishiguro for the permission to use his photo and to Machiko Sakai for advice about copyright. I would like to thank Enago (www.enago.jp) for the English language review. This research was supported by KAKENHI from the Japan Society for the Promotion of Sciences (JSPS) Nos. 19H01228 and 19K21604, Research Institute of Science and Technology for Society (RISTEX) of Japan Science and Technology Agency (JST): Human Information Technology Ecosystem (HITE) focus area project "PATH-AI: Mapping an Intercultural Path to Privacy, Agency, and Trust in Human-AI Ecosystems," and cooperative research program with NEC Corporation. The author declares that there is no conflict of interest. All data are given on the text of the paper.

Funding JSPS-KAKENHI 19H01228, 19K21604, JST RISTEX: HITE focus area project "PATH-AI," cooperative research program with NEC Corporation.

Availability of data and materials All data are given on the text of the paper.

Code availability Not applicable.

Declarations

Conflict of interest There is no conflict of interest.

References

Abe K (1995) "Seken" towa nanika? (What is the "seken": or, how Japanese human relations are characterized?). Kodansha, Tokyo (in Japanese)

Allison A (2006) Millennial monsters: Japanese toys and the global imagination. The University of California Press, Birkley

Ariès P (1960) L'enfant et la vie familiale sous l'ancien régime. Plon, Paris

Asquith PJ, Kalland A (eds) (1997) Japanese images of nature: cultural perspectives. Curzon Press, London

Bartneck C (2008) Who like androids more: Japanese or US Americans? Proc 17th IEEE Int Symp Robot and Human Interactive Communication, pp 553–557. https://doi.org/10.1109/ROMAN.2008.4600724

Bartneck C, Suzuki T, Kanda T, Nomura T (2007) The influence of people's culture and prior experiences with Aibo on their attitude towards robots. AI Soc 21:217–230. https://doi.org/10.1007/s00146-006-0052-7

Bostrom N (2014) Superintelligence: paths, dangers, strategies. Oxford University Press, Oxford
de Waal F (2001) The ape and the sushi master: cultural reflections of a primatologist. Allen Lane, London
Endo T (2006) Discussing the origin of mind through the analysis of visual attention: a review. In: Endo T (ed) Yomu me, yomareru me: shisen rikai no shinka to hattatsu no shinrigaku (Eyes reading others, eyes read by others: evolution and developmental psychology of visual attention). University of Tokyo Press, Tokyo, pp 11–66 (in Japanese)
Fedigan LM, Asquith PJ (eds) (1991) The monkeys of Arashiyama: thirty-five years of research in Japan and the West. State University of New York Press, New York
Frumer Y (2018) Cognition and emotions in Japanese humanoid robotics. Hist Technol 34(2):157–183. https://doi.org/10.1080/07341512.2018.1544344
Haring KS, Mougenot C, Ono F, Watanabe K (2014) Cultural differences in perception and attitude towards robots. Int J Affect Eng 13(3):149–157. https://doi.org/10.5057/ijae.13.149
Hasumi S (1983) Kantoku Ozu Yasujiro (Director Yasujiro Ozu), Chikuma Shobo, Tokyo (in Japanese) (Enlarged edn, 2003, Chikuma Shobo. French version is available)
Helmreich S (1998/2000) Silicon second nature: culturing artificial life in a digital world. University of California Press, Berkley
International Federation of Robotics (2020) https://ifr.org/ifr-press-releases/news/robot-race-the-worlds-top-10-automated-countries. Accessed 17 Mar 2021
Ito K (2010) Robots, A-bombs, and war: cultural meanings of science and technology in Japan around World War II. In: Robert J (ed) Filling the hole in the nuclear future: art and popular culture respond to the bom, Kindle edn. Lexington Books, Lanham
Japan Robot Association (2019) https://jara.jp/data/dl/Operational_stock_2018.pdf. Accessed 17 Mar 2021
Kaplan F (2004) Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots. Int J Humanoid Robot 1(3):465–480. https://doi.org/10.1142/S0219843604000289
Kishi F (1994) Edo no enkinho: ukie no shikaku (Perspectives in Edo era: vision realized in uki e) Keiso Shobo, Tokyo (in Japanese)
Kitayama O (ed) (2005a) Kyôshiron: boshi zo no shinrigaku (Viewing together: psychological analysis of pictures of mother-infant pairs). Kondansh, Tokyo (in Japanese)
Kitayama O (2005b) Considering from the pictures of mother-infant pairs. In: Kitayama (ed) (2005a), pp 7–46 (Kodansha) (in Japanese)
Kovacic M (2018) The making of national robot history in Japan: *monozukuri*, enculturation and cultural lineage of robots. Crit Asian Stud 50(4):572–590. https://doi.org/10.1080/14672715.2018.1512003
Kurzweil R (2005) The singularity is near: when humans transcend biology. Viking, New York
MacDorman KF, Vasudevan SK, Ho C-C (2008) Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures. AI Soc 23:485–510. https://doi.org/10.1007/s00146-008-0181-2
MacKenzie D, Wajcman J (eds) (1999) The social shaping of technology, 2nd edn. Open University Press
Miura K (2005) The structure of visual attention. In: Kitayama (2005a), pp 129–158 (in Japanese)
Moore C, Dunham P (1995) Joint attention: its origins and role in development. Lawrence Erlbaum Associates
Morse ES (1917) Japan day by day, 2 vols. Houghton Mifflin, Boston
Nakao M (2014) Robots in Japanese popular culture. In: Funk M, Irrgang B (eds) Robotics in Germany and Japan: philosophical and technical perspectives. Peter Lang, Frankfurt am Main, pp 113–124. https://doi.org/10.3726/978-3-653-03976-4
Nishigaki T (1988) AI: jinko chino no konseputo (AI: cultural analysis of its concept). Kodansha, Tokyo (in Japanese)
Nishigaki T (2018) AI genron: kami no shihai to ningen no jiyu (The principles of AI: god's rule and human freedom). Kodansha, Tokyo (in Japanese)
Nomura T (2017) Cultural differences in social acceptance of robots. In: 26th IEEE Int Symp on Robot and Human Interactive Communication, pp 534–538. https://doi.org/10.1109/ROMAN.2017.8172354
Nomura T, Syrdal DS, Dautenhah K (2015) Differences on social acceptance of humanoid robots between Japan and the UK. In: Procs 4th Int Symp New Frontiers in Human-Robot Interaction. http://hdl.handle.net/2299/16345
Okada M (2012) Yowai robotto (Weak robots). Igaku Shoin, Tokyo
Okada M (2017) "Yowai robotto" no shiko (Thoughts inspired by the "weak robots"). Kodansha, Tokyo
Ray C, Mondada F (2008) What do people expect from robots? IEEE-RSJ Int Conf Intell Robots Syst. https://doi.org/10.1109/IROS.2008.4650714
Reddy V (2005) Before the 'third element': understanding attention to self. In: Eilan N, Hoerl C, McCormack T, Roessler J (eds) Joint attention: communication and other minds. Oxford Scholarship Online. https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780199245635.003.0005
Sakura O (2010) Style, field and structure: the reasons why Takasakiyama no Saru is exciting even today. Notes to paperback renewed edition of Takasaki-yama no Saru (Monkeys in Takasaki-yama), written by Itani J. Kodansha, Tokyo (in Japanese)
Sakura O (2019) Considering autonomous driving from the STS point of view. Jidosha Gijyutsu (Automob Technol) 73(3):10–15 (in Japanese)
Sakura O (2020) Kagaku towa nanika? (How should we treat science and technology?). Kosandha, Tokyo (in Japanese)
Sasaki K (2013) Perspectives East and West. Contemp Aesthet 11 https://contempaesthetics.org/newvolume/pages/article.php?articleID=670#FN4
Scaife M, Bruner JS (1975) The capacity for joint visual attention in infant. Nature 253:265–266. https://doi.org/10.1038/253265a0
Shibata T, Wada K, Tanie K (2004) Subjective evaluation of seal robot in Brunei. In: Proc IEEE Int Workshop on Robot and Human Interactive Communication, pp 135–140. https://doi.org/10.1109/ROMAN.2004.1374744
Suwa H (1998) Nihonjin to enkinhô (Japanese and perspective drawing). Chikuma Shobo, Tokyo (in Japanese)
Syrdal DS, Nomura T, Dautenhahn K (2013) The Frankenstein Syndrome questionnaire: results from a quantitative cross-cultural survey. In: Herrmann G et al (eds) Social robotics. CSR 2013. Lecture notes in computer science, vol 8239. Springer, Cham, pp 270–279. https://doi.org/10.1007/978-3-319-02675-6_27
Takashina S (1996) Nihon bijyutu wo miru me: higashi to nishi no deai (How to analyze Japnese fine art: encounter of the East and West), New edn. Iwanami-Shoten, Tokyo (1st edn 1991) (in Japanese)
Tanaka Y (2005) Edo culture of the *ba*, or field. In: Kitayama (2005a), pp 47–70 (in Japanese)
Trovato G., Eyssel, F (2017) Mind attribution to androids: a comparative study with Italian and Japanese adolescents. In: 26th IEEE Int Symp on Robot and Human Interactive Communication, pp 561–566. https://doi.org/10.1109/ROMAN.2017.8172358
Trovato G, Zecca M, Sessa S, Jamone L, Ham J, Hashimoto K, Takanishi A (2013) Cross-cultural study on human-robot greeting interaction: acceptance and discomfort by Egyptians and Japanese. Paladyn 4(2):83–93. https://doi.org/10.2478/pjbr-2013-0006
Umesao T (1960) Nihon Tanken (Exploration of Japan). Chuokoronsha, Tokyo (in Japanese)
Wang L, Patrick Rau P-L, Evers V, Robinson BJ, Hinds P (2010) When in Rome: the role of culture & context in adherence to

robot recommendations. In: 5th ACM/IEEE Int conference on human-robot interaction (HRI), pp 359–366. https://doi.org/10.1109/HRI.2010.5453165.

Yamada Y (2005) Looking together and talking together: human relation of positioning side by side and sharing the third item. In: Kitayama (2005a), pp 73–87 (in Japanese)

Yamaguchi H (2005) Visual attention in a vertical society. In: Kitayama (2005a), pp 159–175 (in Japanese)

Yamaji Y, Miyake T, Yoshiike Y, De Silva PRS, Okada M (2011) Stb: Child-dependent sociable trash box. Int J Soc Robot 3(4):359–370. https://doi.org/10.1007/s12369-011-0114-y