Journal of Information Processing Vol.22 No.3 547–555 (July 2014)

[DOI: 10.2197/ipsjjip.22.547]

Regular Paper

SensorTank: An Interactive Water Vessel

K.S. Lasith Gunawardena[1,a) Koki Kimura[1] Masahito Hirakawa[1]

Received: August 20, 2013, Accepted: March 7, 2014

Abstract: The study presents an interaction system named SensorTank, which can detect the position and volume of objects such as human feet inserted into water. SensorTank is designed to be used as an interactive water vessel interface in which three-dimensional interactions occur. Feedback is provided by visual, auditory, and thermal sensations. For detection of inserted objects, combinations of phototransistors and red lasers that form sensing units are arranged on four sides of the vessel. Signals given from the phototransistors are interfaced via an Arduino Uno microcontroller and a multiplexer circuit. An elementary application has been implemented to illustrate the use of the tank for foot gestures, and early experimental results suggest that the proposed mechanism is both feasible and practical even under murky water conditions.

Keywords: water interaction, arduino, three-dimensional interaction.

1. Introduction

In computing, interaction with the real world can be achieved using various methods. The most common approach, as is often the case with personal computers, is the use of keyboards and mice, while other devices such as cameras and microphones have become increasingly ubiquitous in many computing platforms aimed at consumers.

However, science has yet to incorporate the complete spectrum of human senses in interfacing with computer systems. Our investigations suggest that there is an increased focus in the incorporation of more direct forms of human interaction such as haptics. A closely associated yet different form of interaction lies in recognizing gestures performed by human body. Gestures performed using the upper torso, including the head, hands, and arms, are an important nonverbal communication channel in human to human communication. Capturing natural interaction in such behaviors is an active research field. This trend is partly due to introduction of sophisticated hardware devices, such as the Microsoft Kinect, Softkinetic's DepthSense, Leap Motion's Leap Motion Controller, and MYO of Thalmic Labs, which are all aimed at detecting gestures.

In comparison, less research has been conducted on foot-based interactions. One possible reason for this is, when compared with hand movements for the same task, foot movements are less accurate, require more execution time, and are probably less satisfying [1], [2]. However, it should be noted that the opportunity for using foot interaction in non-accurate spatial tasks remains [1]. Moreover, considering the essential fact that human beings are bipedal it is worth investigating foot-based interaction.

Interaction among people in public with technology is an interesting domain. Products such as touch tables, touch walls, and interactive exhibits are becoming increasingly popular. Our innate affinity with water is frequently seen in the way people come together around fountains, which are commonly located in city centers. In Japan, public places where people can bathe their feet, better known as Ashiyu, are quite common, and this is part of regular cultural activity. There is a strong potential for highly user-friendly (or invisible) interfaces using water as an interface medium for computing devices that help people interact. The unique sensation provided by water can have a soothing effect on the body, while fatigue attributed to gesture movements in air and on surfaces has been identified as an issue that requires investigation [3], [4], [5], [6].

In our paper, we introduce SensorTank, a user interface that enables interaction with a volume of water, based on observations obtained through previous studies [7], [8], [9], [10]. The system is organized by a water vessel at which combinations of a laser and a phototransistor are arranged as sensing units. An LCD monitor is placed at the bottom of the vessel to display computer-generated images to the user. SensorTank detects the 3D positions of objects (for example feet and hands) that are inserted into the water and provides visual, auditory, and thermal feedback to the user through the LCD monitor, embedded speakers, and a heating element, respectively.

A key difference between existing studies and our study is that, while existing studies have concentrated on interactions in air (using various body parts or walking movements) or interactions at the surface of a volume of water, our study treats the volume of water as a 3D interaction space. The proposed system provides thermal sensations and natural water friction in addition to the visual and auditory opportunities that are provided in existing interactive systems.

The paper is organized as follows. In Section 2, we provide a brief overview of existing research, which provides the background to this study. We also discuss limitations of the aforementioned studies when applied to our problem domain. In Section 3,

---

[1] Interdisciplinary Graduate School of Science and Engineering, Shimane University, Matsue, Shimane 690–8504, Japan

a) research@lasith.com

we explain our previous work, which has led to the present research, followed by an explanation of the hardware design, software design, and prototype applications of the proposed system. Section 4 discusses experiments for evaluating the system performance, and in Section 5, we conclude by outlining the potential for further research in this area.

2. Background and Related Work

2.1 Bodily Gestures and Related Detection Technologies

In examining human gesture research, investigation of gesture movements performed by the upper torso, including hands, arms, and head occupies a prominent space. There are many gestures in use today; some are universally understood while others are associated with culturally different meanings [11]. Common gestures performed on 2D surfaces, such as tablets and smart phones, include tapping, holding, dragging, pinching (performed by a single hand), and dragging two opposite corners (using both hands). Although all of these gestures may be applied extensively to a 3D space, they do not fall under intuitive 3D gestures.

Various techniques have been employed by researchers to capture gestures accurately. As cited by Dhawale et al. [12], the methods include data gloves, bats/wands, and video cameras.

While cameras have been commonly used for gesture and motion detection in existing trials [3], [7], [8], [9], [12], [13], [14], [15], [16], [17], the performance of object detection depends on the lighting in a working environment. Some researchers have proposed illumination of objects to improve object detection performance using visible light, such as blue light emitting diodes (LEDs) [8], [9], [15], infrared red (IR) LEDs [14], and IR lasers [13], [16]. In the case of IR illumination, web cameras must be modified to enable the detection of IR light. Overlay touch frames are a category of add-on hardware mounted on an LCD monitor to enable multi-touch user interaction. This also uses IR LED technology.

The use of LEDs, either in the IR or visible spectrum, is practical and economical for the detection of target objects, such as a finger in a normal environment (i.e., air medium), but poses a problem when used to detect objects in water due to attenuation issues.

Meanwhile, gesture movements by feet could work for performing certain tasks, although the foot does not offer the same degree of precision and dexterity for selection as the wrist and hand, as observed by Scott et al. [18]. In their study, Scott et al. considered four experimental conditions for foot-based interaction space, as illustrated in **Fig. 1**:

- *Dorsiflexion*: rotation of the ankle such that the angle between the shin and foot decreases
- *Heel rotation*: internal and external rotation of the foot and leg with respect to the midline of the body while pivoting the rotation on the heel
- *Plantar flexion*: rotation of the ankle such that the angle between the shin and foot increases
- *Toe rotation*: internal and external rotation of the foot and leg while pivoting the rotation on the toe.

In this paper, we propose adopting foot-centric gestures in human-computer interaction as a trial to extend computer applicability.

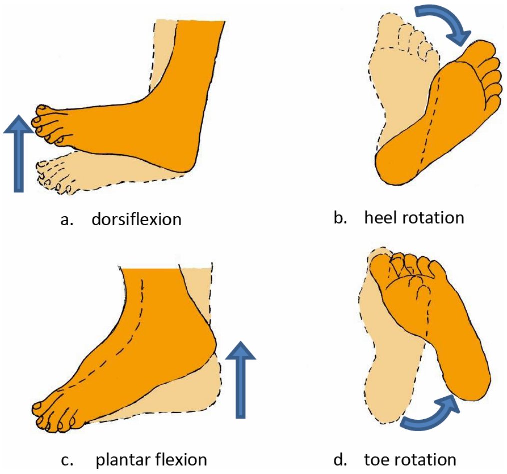

Fig. 1 Four basic foot movements; perspective for (b) and (d) are from above the right foot [18].

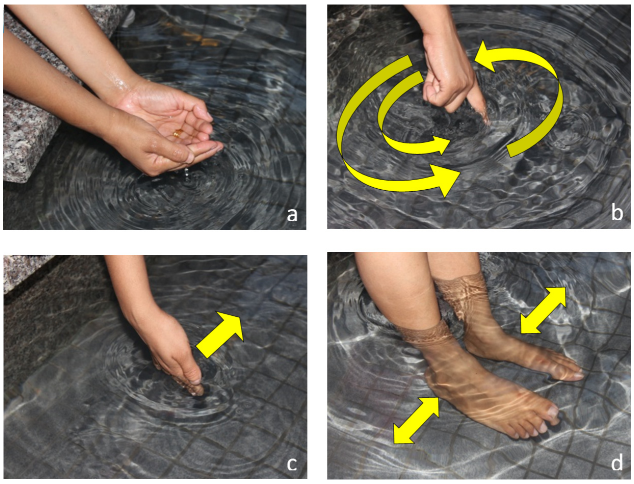

Fig. 2 Natural interaction with water: a. two handed scoop b. twirling c. paddling with hand d. paddling with feet.

2.2 Tangible/Tactile Interfaces and Water Interaction

Most efforts in existing trials have focused on hand and/or body gestures in air as input to interactive systems. However, this interaction scheme lacks proper feedback from the physical world. Tangible and tactile user interfaces are a promising approach to overcome this drawback and have come a long way since the pioneering article by Ishii and Ullmer [19]. Our study of human-computer interaction through water extends this category of research.

When we consider our day to day interactions with water, a few unique gestures can be identified (**Fig. 2**). Scooping, paddling, and twirling are basic hand gestures. Relatively complex gestures include washing hands. When considering foot movements, paddling is a natural behavior.

When considering research in water interaction, some studies have focused on the dynamic aspects of water, including water movement, pressure, and flux [15]. A number of applications dealing with interaction with fluids such as water have also been presented [7], [8], [9], [10], [14], [17], [20]. Using ferromagnetic fluids Koh et al., presented a tangible and malleable interface that allowed the user to produce a 3D response [21]. However, most of the trials concern interaction with water (or another liquid) at the surface level. Touché [22] facilitated the detection of gestures performed in water and in air using a swept frequency capacitive sensing technology; however, this was not designed to provide any information on the 3D space (positional data) in which the gesture was performed. Gurgle [23] is an augmented public space that augments an existing water fountain with watery reflections and sound to motivate human behavior change.

The Nintendo Wii Remote [24], which is a controller for the Nintendo Wii gaming console, was used as a sensor for water level measurement [25]. The Microsoft Kinect motion detector device [26] was originally designed for use with full body gaming, but its application in scanning a dynamic water surface [27] and depth (up to 0.203 m) [28] has also been presented. However, these experiments were not intended to be applied as techniques for human-computer interaction. AquaTop [29] uses the Kinect to detect gestures performed at the surface of cloudy water.

In a more general sense, sonar is a well-known methodology that has been used for object detection in water. Sonar used in the marine industry utilizes instruments such as hydrophones, but they are costly and not designed for use with short distances of less than one meter. Low-cost acoustic sensors that adopt ultrasonic sound waves are available. However, most these devices have a narrow detection angle such as 20–30 degrees and a minimum detection distance in the order of 10–20 cm. In addition, they only detect the distance to an object: therefore to generate information on the shape of the object, multiple sensors operating at different frequencies may be required. In comparison, low power laser modules are quite reasonably priced, and when paired with phototransistors they provide a simple detection mechanism which can be scaled according to the detection space.

In a summary, little research has focused on issues related to object tracking or gesture detection within a limited aquatic space such as a water vessel.

3. Prototype System Description

3.1 Background

One of the authors has presented interactive systems using water as a medium [7], [8], [9]. O-Key [7], which using a web camera, video projector, tub, and a personal computer detects the movement of hands at the surface level (2D) to identify a scooping gesture. The subsequent experiments [8], [9] made use of Frustrated Total Internal Reflection [30] as a technique for sensing hands submersed in a tub. The system consisted of an acrylic tub, two web cameras, a video projector, and a personal computer. The depth positions of the hands in a 3D volume (i.e., water) and their spatial positions can be obtained using a stereo camera.

However, camera-based approaches require significant space to set up as cameras must be positioned away from targets so that they can capture an entire search space. For a water tub with a base size of 50 cm × 37 cm [8], the distance from the cameras to the tub is 64 cm. In the case where a larger water tub (interaction area) is required, the camera must be positioned further away. This may become a burden if the system is applied in practical environments, especially when considering foot interactions. As noted in experiments by Sangsuriyachot and Sugimoto [31], capturing foot interaction using a camera-based system demands that the interaction space be elevated considerably, or the floor must be suitably modified to embed devices. Furthermore, to avoid deterioration of position measurement performance, cameras must be positioned with care.

We experimented on whether the Kinect motion detector device could work for the detection of gestures in water. We found that, when the unit is positioned above the water surface and gestures are performed within water, the ripples that are generated act as a barrier to successful detection. Moreover, when we mount the Kinect on the side of the water body using a clear acrylic tank, detection is only successful within 5 cm of the tank wall. We further experimented using another similar device, i.e., the Softkinetic DepthSense 311 Camera. This device gave similar results. One reason why these devices do not perform as expected in a fluid medium can be attributed to the use of low power IR illumination. IR is attenuated in water; therefore, it can be harder to detect objects as depth and distance increase.

3.2 Hardware Design

We propose using a sensor array to overcome the issues mentioned in the previous subsection. A sensing unit that organizes the sensor array is comprised of a red laser and a phototransistor. The use of lasers is advantageous because they provide a coherent light source compared with LEDs and attenuation is relatively negligible, even in water.

The interaction tank was built using transparent acrylic panels of 1.5 cm thickness with tank dimensions of 20 cm × 88.4 cm × 50 cm (H × L × W). In previous experiments [7], [8], [9], we employed a video projector over the water tank, which was used to present computer generated images to the user. However, the use of the video projector presented a weakness in the system setup space, i.e., the system became larger. Therefore, we decided to use an LCD monitor as a display device. An LCD monitor with built-in stereo speakers was placed at the bottom of the tank using a support structure. The use of an overhead video projector may be considered in situations where a bottom view of the LCD monitor is obstructed due to objects presented in the water. The size of the whole system (excluding the PC) was 37 cm × 99 cm × 80 cm (H × L × W). The dimensions of the experimental pilot system (**Fig. 3**) were determined so that the feet or hands of several people can be inserted simultaneously. This size can be made smaller depending on the type of interaction required.

We used lasers and phototransistors for detection of an object inserted into water. The modules were arranged such that the pairs of lasers and phototransistors were placed facing each other at the sides of the tank, as is illustrated in **Fig. 4**. When the laser beam is blocked by an object, it can be detected by the associated phototransistor. To eliminate false positives, we selected a load resistor value that was not triggered by external light sources.

A total of 78 sensor units were mounted at a separation of 5 cm (horizontal) and 3 cm (vertical) between modules on all four walls of the tank using three mounting layers, i.e., 26 units per layer (**Fig. 5**).

The resolution is rather coarse compared to other existing

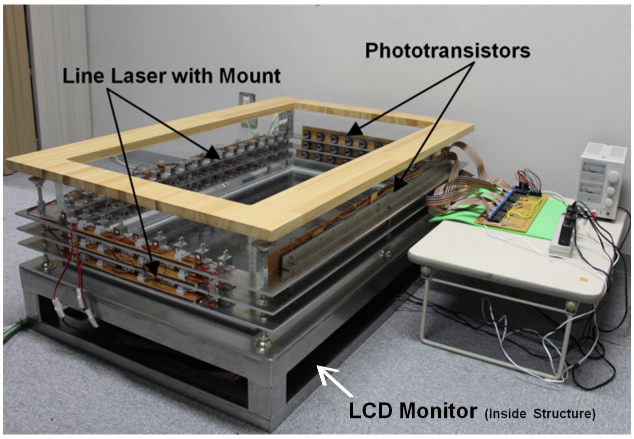

Fig. 3 Side view of SensorTank.

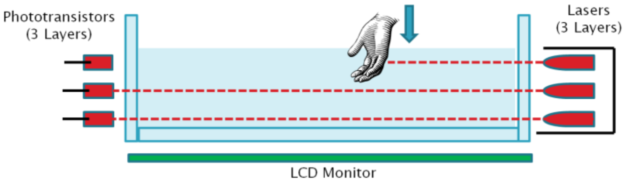

Fig. 4 Cross-section diagram of SensorTank.

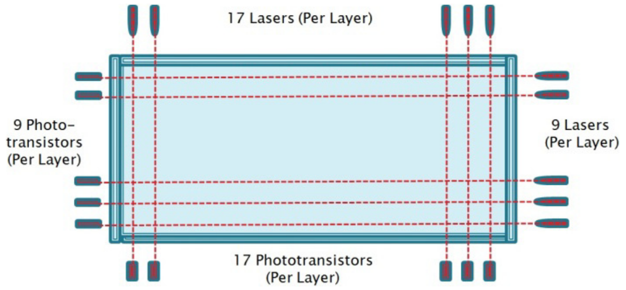

Fig. 5 Overhead diagram of SensorTank.

touch sensing devices as the position of a fingertip on a touch display is not our intention. We aim to apply the system to domains in which a body part, such as a foot or hand with 3D volume, is used for interaction with a computer. We considered the following hand and foot anthropometry data [32] when we had to determine the laser spacing (points). We considered the lower fifth percentile of the population to validate our resolution as practiced in ergonomics.

1. The fifth percentile foot length of males and females at nine years of age is 18.5 cm for each; this distance occupies 3 to 4 laser points.
2. The fifth percentile foot breadth of males and females at nine years of age is 7 and 6.9 cm, respectively; these distances occupy 1 to 2 laser points.
3. The fifth percentile hand length of males and females at nine years of age is 12.81 and 12.97 cm, respectively; these distances occupy 2 to 3 laser points.
4. The fifth percentile hand breadth of males and females at nine years of age is 5.36 and 5.39 cm, respectively; these distances occupy 1 to 2 laser points.

Since the horizontal separation between lasers is 5 cm, theoretically it is possible that a hand or toes are inserted between two adjacent lasers. They would not be detectable at that instant. However, we expect the interaction to be dynamic in nature. A target such as a hand or toes would be detected without trouble once they are moved in either direction. It is also possible to place a middle layer in such a way that its lasers are placed horizontally right between the lasers at the upper and lower layers.

Ideally, by adding more layers and/or modules, higher resolution can be obtained. We estimated that it is physically possible to have maximum horizontal and vertical resolutions of 2 cm using the same components as we used.

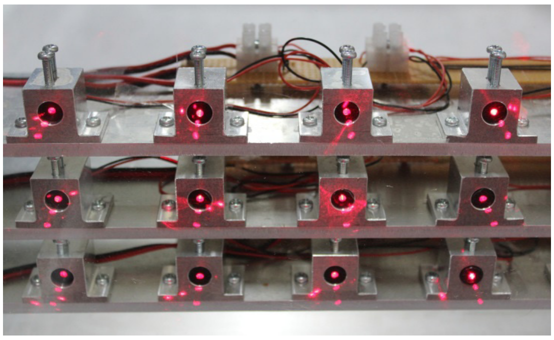

Fig. 6 Lasers with their mounts.

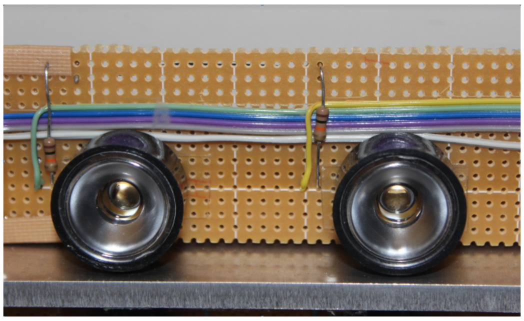

Fig. 7 Lens to capture laser beam.

Figure 6 shows a close-up of the laser configuration. For each laser module, the beam is focused to a very tiny spot and does not transfer to phototransistors that are not associated with it. Even a small error in mounting the laser module can cause a problem. Therefore, the laser modules were positioned in a custom-built mount, as shown in Fig. 6, with two separate screws to fasten each module. Furthermore, a lens is attached in front of each phototransistor so that the laser beam can be captured properly even if the beam misses the center of the phototransistor due to inaccurate laser module mounting (**Fig. 7**).

Although previously reported research [13], [16] used IR lasers to illuminate gesturing objects, IR lasers are class 3 in terms of laser safety classification, which carries a risk of injury to the exposed eye as the beam is invisible. Thus, for safety reasons, we employed class 2 lasers for our prototype system. The blink reflex limits accidental exposure to the eyes when using class 2 lasers, which are visible. Our selected class 2 lasers provide a red dot pattern, with a wavelength of 650 nm and power less than 1 mW.

The laser beams are not recognizable from above under normal conditions; however, the dot matrix will illuminate any object immersed in the tank. **Figure 8** illustrates the laser matrix, which was specially illuminated for demonstration purposes by filling the tank with water mixed with a clouding agent.

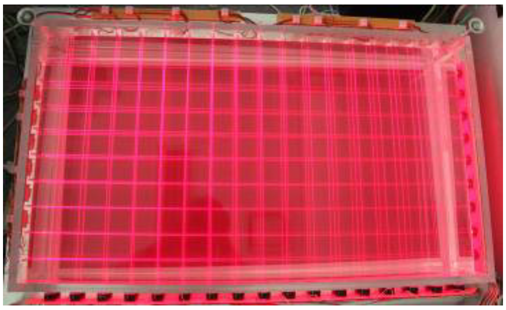

Fig. 8 Overhead view of SensorTank with visible laser matrix.

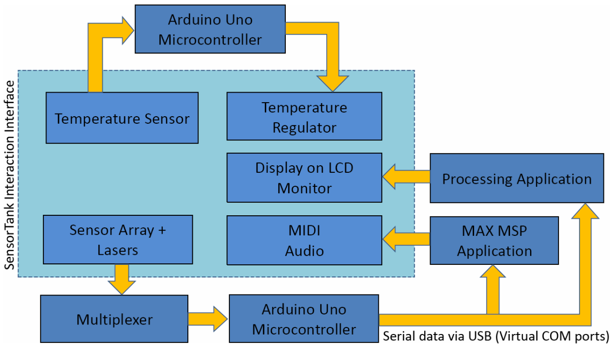

Fig. 9 Schematic flow diagram.

Since there were 78 sensors to be read, the output was combined through a multiplexer circuit and connected to an Arduino Uno Microcontroller Board [33]. Additional explanation about processing the signals is provided in the next subsection. A schematic flow diagram is given in **Fig. 9**.

A separate temperature sensor was also connected to an Arduino Microcontroller to enable temperature regulation in the vessel by controlling a heating element. Ideally, it would be beneficial to have a cooling element to lower the temperature of water in the vessel; this will be a consideration in future work.

3.3 Software Design

The Arduino was programmed using its proprietary programming language to provide a stream of serial data corresponding to the coordinates of the 3D space. The positional information is coded by length (x-axis) and width (y-axis) coordinates of a tracking object, along with a depth value that corresponds to the sum of weighted values given for the layers at which the object is present. Specifically, consider that layers 1, 2, and 3 took values of 1, 2, and 4, respectively. This ensured that the summation of any weight combination can be decoded easily to determine which layers contained an object. When an object covered layers 1 and 3, for example, the depth information was expressed as 5 (= 1 + 4). This ensured that the data stream is optimized for performance. This 2D matrix is transformed into a 2D matrix with binary data, with 0 indicating no object at the given length and width in any laser layer (1–3). A connected component analysis (blob detection) runs on the 2D matrix to determine different objects in the tank. Noise filtering is performed to extract those objects suitable for further analysis. The process is repeated to identify the gestures involved. Once a shape has been identified, it is possible to output data, such as the centroid of the object. The algorithmic architecture for gesture detection is displayed in **Fig. 10**. Approaches undertaken to detect gestures such as those mentioned in Figures 1 and 2 will be reported in the future.

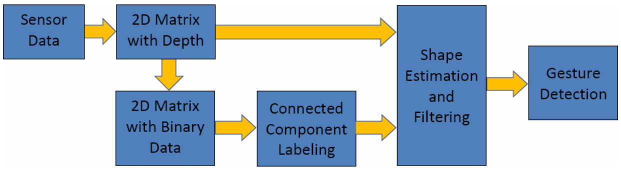

Fig. 10 Algorithmic architecture for gesture detection.

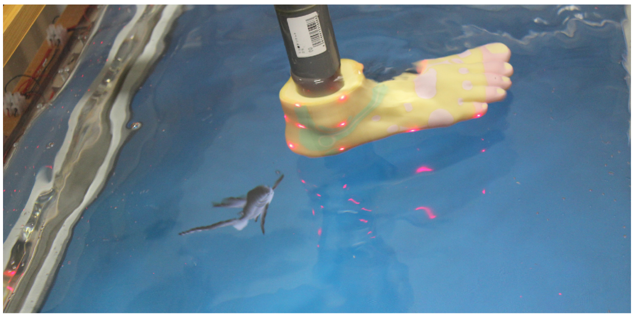

Fig. 11 Fish following the foot.

3.4 Prototype Applications

Using the positional data of the points located inside the target object's volume, we developed several applications that are suitable for different application scenarios. The data was simultaneously used by independent program modules written in Processing and MAX/MSP using virtual COM ports to present both visual and audio feedback (Fig. 9).

We have implemented a prototype system to be applied in a foot bath activity (Ashiyu) to exhibit the performance of our framework. Ashiyu has been a social activity for a long time in Japan; people come to meet and talk with others. Incorporating interactive capabilities of the system into the foot bath makes people happy as well as relaxed, and generated music tones create a calming atmosphere.

In one scenario, the system calculates the centroid of an immersed foot and displays a fish so that it follows the foot movements, as shown in **Fig. 11**. The fish changes its direction and depth position depending on the inserted foot position. A MIDI tone is generated on the basis of the position and depth of the centroid of the foot. In this environment, dorsiflexion and plantar flexion gestures may be used to increase/decrease the temperature of water in the tank, respectively, resulting in the change of fish type (e.g., a tropical fish) to be displayed. A temperature controlling module will serve the user by producing thermal sensation as well.

Another scenario in which this system could be of use is in replacing traditional faucets in a bathtub environment. The opening and closing of a faucet can be mapped to the heel rotation. The temperature of the flowing water can be mapped to dorsiflexion (increase) and plantar flexion (decrease). Draining water from the bathtub can be mapped to the toe rotation or paddling with both feet.

Application of the system to physical rehabilitation also shows promise. The system could be applied to assist a patient with rehabilitation, wherein a target symbol is presented to the patient so

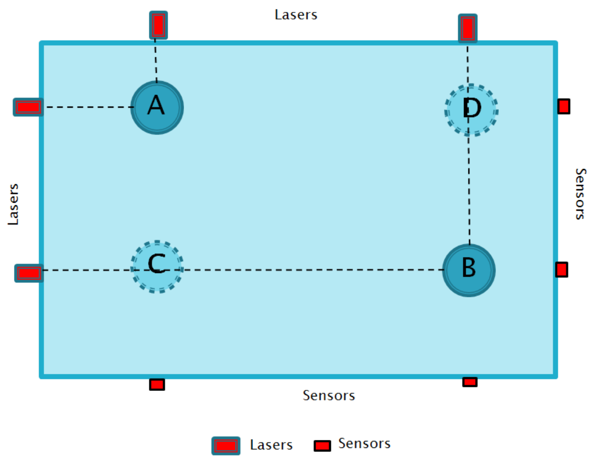

Fig. 12 Objects A and B are real objects. During detection, the detection process detects ghost objects C & D (Overhead view).

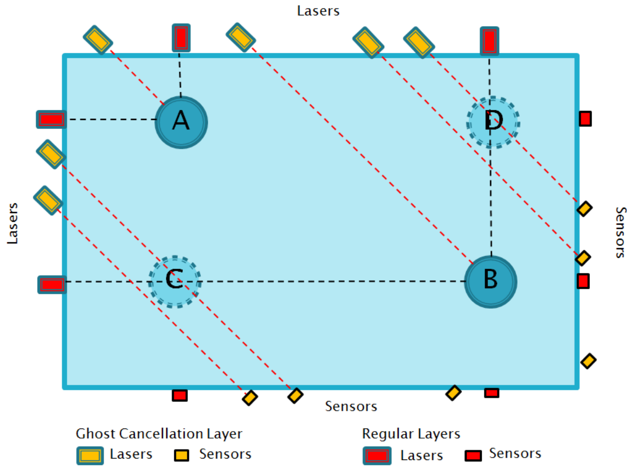

Fig. 13 Ghost cancellation with ghost cancellation layer.

that they may attempt to follow it. This movement (direction and speed) may change depending on the status of rehabilitation, and a performance score could be given to provide motivation. When a patient moves their hand or arm in the water, force feedback is given due to the resistance of water. Using liquid with higher viscosity could yield even higher load. Moreover, the capability to adjust the temperature of the water (liquid) could ease the physical strain on the patient.

3.5 Ghost Point Cancellation

We have implemented a program to detect multiple objects placed in the tank. While the system can easily monitor the position of multiple objects in the same column or row (such as a single person inserting both hands or feet), random simultaneous movement in both directions along the x and y axis creates a situation in which ghost objects are detected due to occlusion (**Fig. 12**).

To solve this problem, we propose using an additional detection layer (referred to as a ghost cancellation layer), which emits laser beams at a 45 degree angle to the beams at the regular laser-phototransistor layers, as illustrated in **Fig. 13**. The ghost cancellation layer having 25 laser-phototransistor pairs is mounted over the top of the regular layers and above the acrylic tank. **Figure 14** shows the setting of the ghost cancellation layer. The beams of this layer do not travel through acrylic material. As such there is no effect of the 45 degree angle, as no refraction occurs.

Fig. 14 Ghost cancellation layer.

Each intersection point in laser-phototransistor sensing units of the regular layers is connected with a certain diagonal laser beam of the ghost cancellation layer. If the tank detects multiple objects in its original estimation, then an additional check is conducted for each of the possible object points detected. For example, in Fig. 13, four possible points exist: A and B are the ones at which objects are truly positioned, and C and D are the ghost points. For A and B, laser beams of the associated sensing modules at the ghost cancellation layer are blocked and the system can thus identify them as true object positions. On the other hand, the sensing modules that are associated with C and D do not detect objects.

The use of the ghost cancellation layer does not provide a perfect solution. As the number of target objects increases, the possibility of recognition failure increases. Furthermore, since the ghost cancellation layer is slightly above the water surface, object positions detected may not be the same as those detected in the regular sensing layers, because of the insertion of a slanting object. To cope with this problem we allow for a margin of error in matching

4. Early Testing

An experiment was conducted with test users to assess the feasibility of the system setup. We asked the users to generate paddling and scooping motions as well as the four basic foot motions suggested by Scott et al. [18]. It was observed that moving hands and feet forcefully created ripples, which rose to 40 mm above the regular water level of 145 mm; however, this had no effect on detection as long as the detection plane was below the water surface. During the trials, we varied the temperature of the water in the vessel between 10°C and 40°C without any negative effect on the sensing performance. The experiment was carried out under ambient lighting conditions, and the system was not exposed to direct sunlight. We plan to implement a filter in front of the phototransistors to reduce the influence of environmental lighting conditions.

Considering an application in a foot bath environment, one concern is system performance when the water is murky (i.e., cloudy). We used a turbidity meter and evaluated the extent that the system could function with no error in detection. We used a commercial bath salt as the clouding agent.

We measured the performance of sensors (i.e., phototransis-

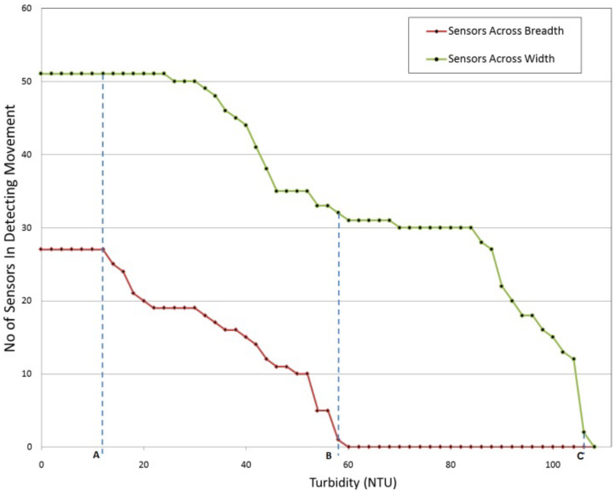

Fig. 15 Detection threshold for phototransistors in murky water.

(a) Turbidity at point A (12 NTU)

(b) Turbidity at point B (56 NTU)

(c) Turbidity at point C (106 NTU)

Fig. 16 Turbidity at three reference points.

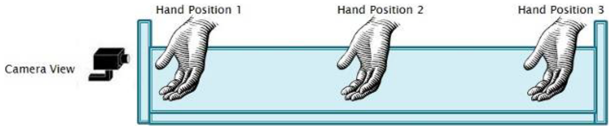

Fig. 17 Reference points for photographic comparison.

(a) Display at Turbidity 0 NTU

(b) Display at Turbidity 12 NTU

(c) Display at Turbidity 56 NTU

Fig. 18 Clarity of the displays at different turbidities.

tors) at different turbidity levels without the insertion of any object into the sensor tank. While increasing the turbidity, we plotted the number of sensors providing false positive detections. Since the distance travelled by the laser light was different depending on the direction, 88.4 cm for 27 phototransistors across the length and 50 cm for 51 phototransistors across the width, we recorded the values as two datasets. **Figure 15** shows the experimental results.

As indicated in the graph, until turbidity reached 12 nephelometric turbidity units (NTU) at point A, all sensors performed as expected. However, at 13 NTU, the first false positive detection was observed, i.e., the sensor did not receive the laser beam at sufficient intensity due to the murky water. Thus, it indicated incorrectly that an object was placed between the laser and the phototransistor. At 58 NTU (point B), all sensors apart from the one across the length generated false positives. At 60 NTU, all sensors gave false positives. At 106 NTU (point C), only one sensor across the width was functioning as expected, while 77 sensors provided false positives. Although we would expect all of the sensors on each side to fail at a given turbidity, the observed variation could be attributed to potential uneven laser power and minute alignment differences between laser and phototransistor. Therefore, we presume that for the given dimensions, our system could potentially work successfully up to a turbidity level of 58 NTU.

Figure 16 shows a visual comparison of states A, B, and C to help assess the cloudiness (i.e., turbidity). A camera was placed at the outer side of the sensor tank on its longer dimension, and a hand was positioned at 1 cm, 43 cm, and 85 cm distances from the camera, as illustrated in **Fig. 17**.

Meanwhile, as the murkiness increases, the visibility of a displayed image decreases. At a turbidity level of 12 NTU where all sensors work without hindrance, there is no serious influence on visibility, as shown in **Fig. 18**. Even if the level increases up to 56 NTU, the displayed image is still visible. However, for better visibility, the display monitor may be replaced with an overhead projector as used in Ref. [29].

5. Conclusion

This paper presented the compact structure of the interactive SensorTank system, in which water is used as a medium in the interaction between the user and the system. Combinations of lasers and phototransistors are arranged at four sides of a water tank. The system detects the position of objects, such as feet and hands inserted into the water. Visual feedback is given through an LCD monitor placed at the bottom of the tank. Auditory feedback is obtained through speakers embedded in the LCD monitor. A thermal sensor and regulator are enclosed to detect and regulate temperature in the vessel.

An experiment demonstrated that the system can provide tracking of an object immersed in the tank as well as visual and auditory responses corresponding to the object's position. The technique presented overcomes hurdles (such as the effects of waves and volume of water) encountered when using more established movement recognition devices in water. We also confirmed that the technique can be used in practical situations in which the water may be cloudy.

Therefore, it is worth pursuing this research further to address detection in multiuser, multi-object environments. In order to avoid ghost point problems, we have discussed the use of a separate laser plane above the water surface, at which additional sensing units (lasers and phototransistors) are mounted at a diagonal angle to the layout of the existing sensing units.

In this study, we use visible red lasers in the prototype to make the various experiments easier. In the future, we may replace the red lasers with similar powered IR lasers so as not to disturb the user during their use of the system. The resolution of position sensing in the prototype is rather coarse; therefore, an increase in the number of sensing units may make the resolution higher, making gesture detection more feasible. While the tank size is a major factor in determining the overall system size for use in a public space, our system could be made more compact by using small factor computing platforms such as Raspberry Pi [34], which has built-in audio, video, and USB support. In the future, we will carry out studies to include more accurate gestures, such as paddling, grabbing, and scooping, to assess a broader range of practical application scenarios.

Acknowledgments This study was partly funded by Japan Science and Technology Agency (A-STEP, Exploratory Research).

References

[1] Pakkanen, T. and Raisamo, R.: Appropriateness of foot interaction for non-accurate spatial tasks, *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems, Extended Abstracts on Human Factors in Computing* (*CHI EA'04*), pp.1123–1126 (2004).

[2] Pearson, G. and Weiser, M.: Of moles and men: The design of foot controls for workstations, *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems* (*CHI'86*), pp.333–339 (1986).

[3] Lenman, S., Bretzer, L. and Thuresson, B.: Computer vision based hand gesture interfaces for human-computer interaction, Technical Report CID-172, Center for User Oriented IT Design, Stockholm (2002).

[4] Ito, T.: Hand gesture-based manipulation of a personalized avatar robot in remote communication, Human Interface and the Management of Information: Interacting with Information, *Lecture Notes in Computer Science*, Vol.6771, pp.425–434, Springer-Verlag, Berlin, Heidelberg (2011).

[5] Van Beurden, M.H.P. H., Ijsselsteijn, W.A. and de Kort, Y.A.W.: User experience of gesture based interfaces: A comparison with traditional interaction methods on pragmatic and hedonic qualities, Gesture and Sign Language in Human-Computer Interaction and Embodied Communication, *Lecture Notes in Computer Science*, Vol.7206, pp.36–47, Springer-Verlag, Berlin, Heidelberg (2011).

[6] Yee, W.: Potential limitations of multi-touch gesture vocabulary: Differentiation, adoption, fatigue, Human-Computer Interaction: Novel Interaction Methods and Techniques, *Lecture Notes in Computer Science*, Vol.5611, pp.291–300, Springer-Verlag, Berlin, Heidelberg (2009).

[7] Tanabe, C. and Hirakawa, M.: O-Key: A scoopable interface, *Proc. Eight IEEE Symposium on Multimedia* (*ISM 2006*), pp.185–192 (2006).

[8] Ikeda, M., Nagira, N. and Hirakawa, M.: A multi-dip interface with water, *Proc. International Conference on Complex, Intelligent and Software Intensive Systems* (*CISIS'09*), pp.169–176 (2009).

[9] Ikeda, M. and Hirakawa, M.: A multi-dip interface with water, *Journal of Information Processing*, Vol.51, No.11, pp.2103–2111 (2010). in Japanese.

[10] Gunawardena, K.S.L., Kimura, K. and Hirakawa, M.: SensorTank: A water vessel for interaction, *Proc. 10th Asia Pacific Conference on Computer Human Interaction* (*apCHI'12*), pp.759–760 (2012).

[11] DeVito, J.A.: Human communication: The basic course, 9th Ed., p.496, Pearson Education, Boston (2002).

[12] Dhawale, P., Masoodian, M. and Rogers, B.: Bare-hand 3D gesture input to interactive systems, *Proc. 7th ACM SIGCHI New Zealand Chapter's International Conference on Computer-Human Interaction: Design Centered HCI* (*CHINZ'06*), pp.25–32 (2006).

[13] Mistry, P. and Maes, P.: Mouseless, *Proc. 23rd Annual ACM Symposium on User Interface Software and Technology* (*UIST 2010*), pp.441–442 (2010).

[14] Okude, E. and Kakehi, Y.: Rainterior: An interactive water display with illuminating raindrops, *Proc. ACM International Conference on Interactive Tabletops and Surfaces* (*ITS'11*), pp.270–271 (2011).

[15] Pier, M.D. and Goldberg, I.R.: Using water as interface media in VR applications, *Proc. 2005 Latin American Conference on Human-Computer Interaction* (*CLIHC'05*), pp.162–169 (2005).

[16] Takeoka, Y., Miyaki, T. and Rekimoto, J.: Z-touch: An infrastructure for 3D gesture interaction in the proximity of tabletop surfaces, *Proc. ACM International Conference on Interactive Tabletops and Surfaces* (*ITS'10*), pp.90–94 (2010).

[17] Mann, S., Janzen, R. and Huang, J.: "WaterTouch": An aquatic interactive multimedia sensory table based on total internal reflection in water, *Proc. 19th ACM International Conference on Multimedia* (*MM'11*), pp.925–928 (2011).

[18] Scott, J., Dearman, D., Yatani, K. and Truong, K.N.: Sensing foot gestures from the pocket, *Proc. 23rd Annual ACM Symposium on User Interface Software and Technology* (*UIST'10*), pp.199–208 (2010).

[19] Ishii, H. and Ullmer, B.: Tangible bits: Towards seamless interfaces between people, bits and atoms, *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems* (*CHI'97*), pp.234–241 (1997).

[20] Watanabe, J.: VortexBath: Study of tangible interaction with water in bathroom for accessing and playing media files, *Proc. 12th International Conference on Human-Computer Interaction: Interaction Platforms and Techniques* (*HCI'07*), pp.1240–1248 (2007).

[21] Koh, J.T.K.V., Karunanayaka, K., Sepulveda, J., Tharakan, M.J., Krishnan, M. and Cheok, A.D.: Liquid interface: A malleable, transient, direct-touch interface, *ACM Computers in Entertainment*, Vol.9, No.2, Article 7 (2011).

[22] Sato, M., Poupyrev, I. and Harrison, C.: Touché: Enhancing touch interaction on humans, screens, liquids, and everyday objects, *Proc. SIGCHI Conference on Human Factors in Computing Systems* (*CHI'12*), pp.483–492 (2012).

[23] Arroyo, E., Bonanni, L. and Valkanova, N.: Embedded interaction in a water fountain for motivating behavior change in public space, *Proc. SIGCHI Conference on Human Factors in Computing Systems* (*CHI'12*), pp.685–688 (2012).

[24] Nintendo Wii (online), available from ⟨http://www.nintendo.com/wii⟩ (accessed 2013-08-14).

[25] Hut, R.W., Weijs, S.V. and Luxemburg, W.M.J.: Using the Wiimote as a sensor in water research, *Water Resources Research*, Vol.46, W12601, AGU 1-5 (2010).

[26] Microsoft Kinect (online), available from ⟨http://www.xbox.com/en-US/kinect⟩ (accessed 2013-08-14).

[27] Hut, R.W.: Scanning dynamic water surfaces using a Kinect game console controller as a sensor, *Geophysical Research Abstracts*, Vol.13, EGU 2011-10096 (2011), available from ⟨http://meetingorganizer.copernicus.org/EGU2011/poster_programme/6551⟩ (accessed 2013-08-14).

[28] Mankoff, K.D., Russo, T.A., Norris, B.K., Hossainzadeh, S., Beem, L., Walter, J.I. and Tulaczyk., S.M.: Kinects as sensors in earth science: Glaciological, geomorphological, and hydrological applications, *American Geophysical Union Fall Meeting 2011*, Abstract #C41D-0442 (2011).

[29] Matoba, Y., Takahashi, Y., Tokui, T., Phuong, S., Yamano, S. and Koike, H.: AquaTop display: A true "immersive" water display system, *Proc. ACM SIGGRAPH 2013 Emerging Technologies* (*SIGGRAPH'13*), Article 4 (2013).

[30] Han, J.Y.: Low-cost multi-touch sensing through frustrated total internal reflection, *Proc. 18th Annual ACM Symposium on User Interface Software and Technology* (*UIST'05*), pp.115–118 (2005).

[31] Sangsuriyachot, N. and Sugimoto, M.: Novel interaction techniques based on a combination of hand and foot gestures in tabletop environments, *Proc. 10th Asia Pacific Conference on Computer Human Interaction* (*APCHI'12*), pp.21–28 (2012).

[32] Japanese body size data 1992-1994, Research Institute of Human Engineering for Quality Life (1997).

[33] Arduino Uno Microcontroller (online), available from ⟨http://arduino.cc/en/Main/ArduinoBoardUno⟩ (accessed 2013-08-14).

[34] Rasberry Pi Single Board Computer (online), available from ⟨http://www.raspberrypi.org⟩ (accessed 2013-08-14).

K.S. Lasith Gunawardena received his B.Sc. from University of Sri Jayewardenpura, Sri Lanka in 1999 and M.Sc. from University of Colombo, Sri Lanka in 2009. He is currently a Doctoral student at the Interdisciplinary Graduate School of Science and Engineering, Shimane University, Japan. His research interests include human-computer interaction and e-learning. He is a member of the IEEE Computer Society, ISOC and ACM.

Koki Kimura received his B.E. degree from Shimane University in 2012. He is a Master's student in Interdisciplinary Graduate School of Science and Engineering, Shimane University. His research interests include human-computer interaction.

Masahito Hirakawa graduated from Hiroshima Institute of Technology in 1979, and received his M.E. and Ph.D. degrees from Hiroshima University in 1981 and 1984, respectively. He has been a professor at Shimane University since 2002. His research interests include human-computer interaction and e-learning. He is a member of IPSJ, IEICE and ACM, and a senior member of IEEE.