Straw-like User Interface (II): a new method of presenting auditory sensations for a more natural experience

Yuki HASHIMOTO[1], Masahiko INAMI[2] and Hiroyuki KAJIMOTO[1]

[1,2] The University of Electro-Communications
1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan
[1] {hashimoto, kajimoto}kaji-lab.jp
[2] inami@hi.mce.uec.ac.jp

Abstract. We have proposed a new type of audio-tactile interface called the Straw-like User Interface (SUI) that allowed users to virtually experience the sensations of drinking with straw. The sensations were created based on data of pressure, vibration and sound recorded during drinking with a straw. The device enabled us to develop many unique interfaces, facilitating extension of research fields related to tactile displays, from medical care to entertainment. However, with the previous system, the recorded data were replayed at the same speed without reference to the suction power, and so the sensation became unnatural. In this paper we propose a simple technique to preserve naturalness. We dynamically change the play list in accordance with the user's behavior. We believe that the proposed method provides a simple approach to record-and-replay of audio-tactile information without needing a physical model.

Keywords: Drinking Sensation, Tactile Interface, Food Texture, Virtual Reality Sound Effect, Augmented Reality

1 Introduction

Drinking is a multimodal experience. The popping noise of water bubbles is transmitted to the ear both from in the air and the inner mouth. Vibration, pressure and temperature are sensed by the lips, tongue, internal mouth and throat. Smell and taste are sensed by nose and tongue. Consequently all the human senses, except vision, are involved. Therefore, it makes an interesting challenge to fabricate a device that facilitates a virtual drinking experience.

There have been some works that replayed the "eating" experience [1], but before our Straw-like User Interface (SUI) [2], no other work had succeeded in the reconstruction of drinking. Our SUI concentrated on the reconstruction of the experience of drinking with a straw. The sensations were created based on data of actual pressure, vibration and sound recorded during drinking with a straw.

By following the simple principle of "record-and-replay", realistic auditory and tactile sensations were reproduced. These enabled users to distinguish and identify the substance that they were drinking. Another unique aspect of the device was that we could record-and-replay the drinking experience of something that we ordinary do not take in through a straw (popcorn and caviar, for example). Interestingly, users could quite easily accept and understand these peculiar situations. The device enabled us to develop many unique interfaces, facilitating an extension of research fields involving tactile displays, from medical care to entertainment.

This paper is in three parts. In the next two sections, we introduce the SUI system and show mechanical and psychophysical evaluations. After that we discuss a problem with the previous system. This was that recorded data needed to be replayed at the same speed without reference to the suction power; we then give our solutions.

2 Straw-like User Interface

In this section, we describe the mechanism of the proposed SUI and show an evaluation of its performance.

2.1 Drinking Sensation

The SUI treats three sensations related to drinking with a straw (Fig.1).

- *The pressure change in the mouth generated when food blocks the straw*
- *Vibrations of the collisions and friction of food*
- *Sounds at the time a vibration occurs*

Although there are some other sensations related to drinking, such as smell and taste, the presentation of these sensations needs chemical substances, and may involve sanitary problems. Therefore, we limited our system to treating the above three sensations.

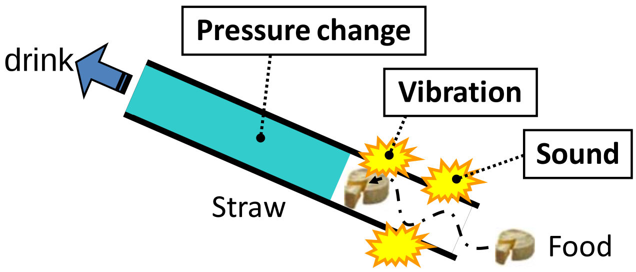

Fig. 1. Overview of the drinking sensation treated by the SUI

2.2 Composition of the SUI

The SUI can reproduce the pressure change in the mouth and transmit vibrations to the lips. For sound, which is one of the three elements of this sensation, it is necessary to have a loud, clear volume; thus we output sounds from an external speaker synchronized with the operation of the SUI. The straw used for the SUI can be any ordinary straw, thus it can be disposed of after use; eliminating any concern in respect to hygiene.

The SUI is composed of a push-pull solenoid, an RC servomotor (Yoshioka model, Atom49), a valve, a cam, a speaker and an air pressure sensor (Fujikura, XFPM-115KPAR).

The pressure change is created by appropriate control of the opening and closing of a valve installed in the SUI. The solenoid and the RC servomotor are used to control the valve. The solenoid is used to completely open and close the valve and generates a high frequency wave for the pressure change. The RC servomotor is used to generate the low frequency waves of the pressure changes. Note that this pressure sensation is displayed "passively". The valve opens and closes, but the pressure is displayed as a result of the user's activity of suction. Therefore, the pressure sensation generated is naturally interactive.

In addition to the pressure sensation, the speaker attached to the straw transmits a high-frequency vibration to the lips. The lips are known to have a similar level of sensitivity as fingers, and hence such a vibration is indispensable for reality. To make the SUI interactive, drinking behavior is detected by the pressure sensor and this then triggers the whole system to start working. Thus, the SUI becomes an intuitive interface without any complex operation requiring buttons.

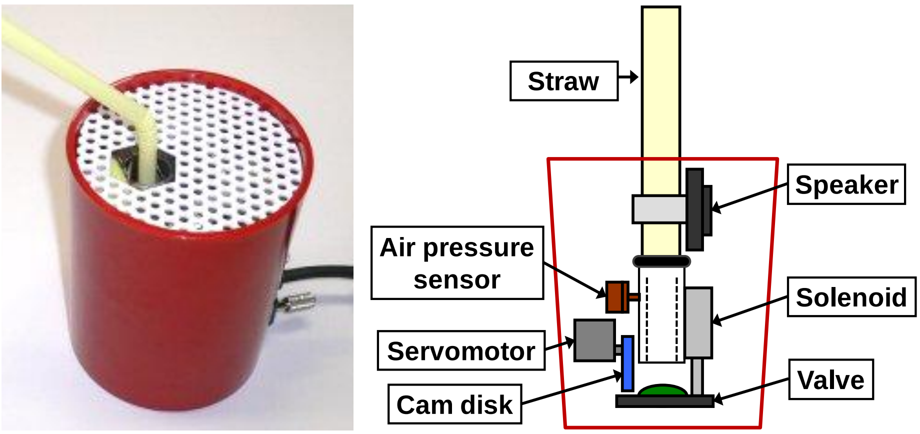

Fig. 2. Straw-like User Interface

2.3 Data Recording and presentation

The SUI follows the simple principle of "record-and-replay". For the data to be used for the operation of the SUI, we recorded pressure changes and sounds in a straw pipe. To record these data, a pressure sensor (Fujikura, XFPM-115KPAR) and a microphone were attached to a straw and we drank actual food. We have so far recorded data relating to about 60 kinds of food. The types of food were not limited to liquids; they included solids, gels, and soluble and gooey things.

We observed that there were similarities in the temporal sequence of pressures. We postulated that it was possible to roughly classify the data into three food groups. Time-series of the pressure of representative foods (raw egg, shake, and caviar) for each of the three classifications are shown in the left column of Fig.3.

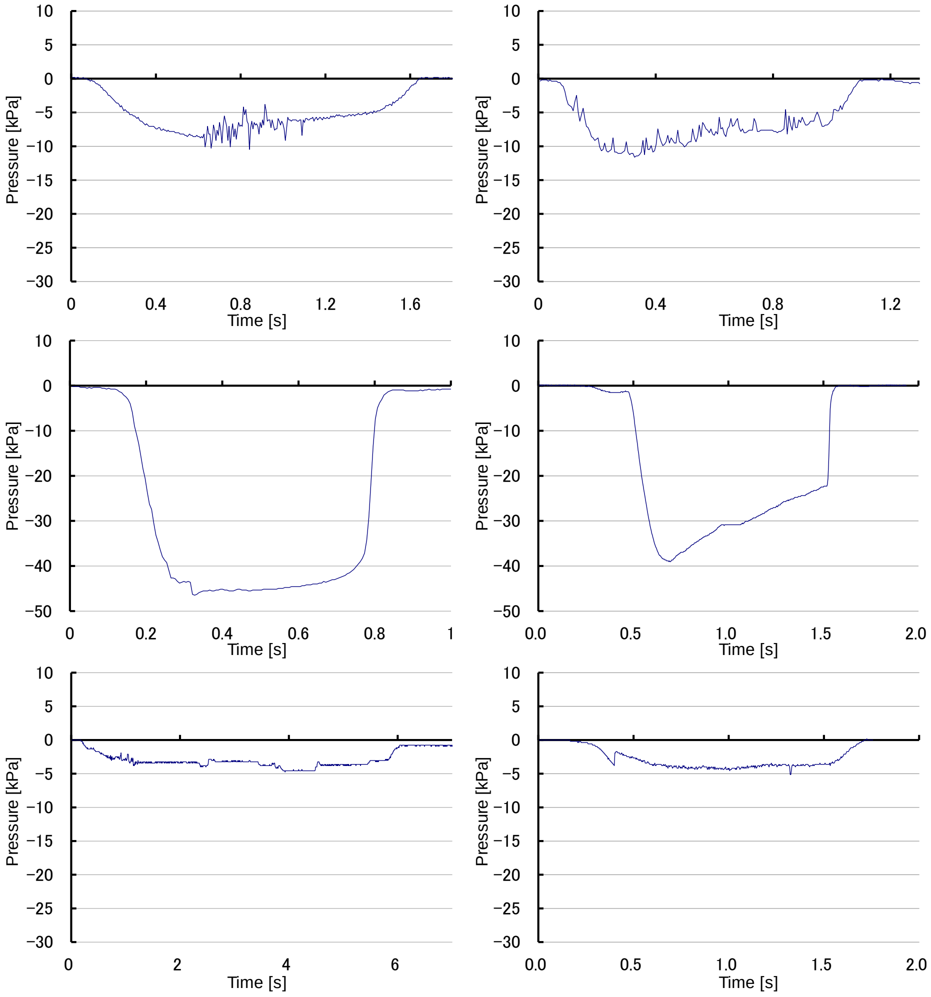

Fig. 3. Recorded air pressure data

Upper left: Original values of raw egg — Upper right: Reproduced value of raw egg

Middle left: Original values of shake — Middle right: Reproduced value of shake

Lower left: Original values of caviar — Lower right: Reproduced value of caviar

A feature of raw egg is that there were high frequency pressure change waves. This is thought to be because of the mixture of solid and gel. Fermented soybeans, curry and rice, and Chinese noodles were foods with similar features. The amount of pressure changes for the shake was dramatically large. The resistance at the time of drinking appeared to be very strong, as did the tendency to have to drink repeatedly. Whipped cream, rice cake, and cheese were foods with similar features. The waveform for caviar showed the least changes, because caviar is composed of very small particles that do not tend to block the straw. Sesame, salmon flakes, and juice were foods with similar features.

As each category showed typical features, we presumed that the drinking sensation could be displayed with a certain level of reality that would remind the user of some particular food. This, we postulated, should be achievable by replaying the recorded pressure changes.

To verify the reproducibility of the pressure changes in the SUI, we conducted an experiment to compare pressure changes between the actual data and that generated by the SUI. In this experiment, we fixed the presentation time regardless of the strength of the user's suction power (right column of Fig.3).

From graphs, we confirmed that the maximum reproduced pressures were similar between the two groups. (Raw egg; Original: 11.6kPa, Reproduced: 10.7kPa, Shake; Original: 49.7kPa, Reproduced: 39.0kPa, and Caviar; Original: 4.8kPa, Reproduced: 5.1kPa). Moreover, the envelopes for the graphs almost matched each other. This demonstrated that the SUI could reproduce the time-series of pressure change by a passive method.

3 Food Discrimination by SUI

In a previous paper [2], we conducted experiments for the discrimination of food using the SUI to show that the device achieved a certain level of reality. An outline of the experiments was as follows:

- *The sensations of a food were presented under three conditions and participants were asked to choose the name of the food from the presented group.*

  Condition 1: Presentation of pressure and vibration
  Condition 2: Presentation of sound
  Condition 3: Presentation of pressure, vibration and sound

Six foods were chosen from the three classifications as described in section 2.3, two from each category. The selected foods were "Chinese noodles and raw egg", "cola and popcorn", and "shake and rice cake". Each condition was presented 20 times. Five participants (five males, ages 23-26) participated in the experiment. The results are shown in Table.1.

Table. 1. Experimental results

	Condition 1	Condition 2	Condition 3
Average of correct answer rate (six foods)	24[%]	57[%]	57[%]
Average of correct answer rate (three categories)	46[%]	63[%]	72[%]

The average rate of correct answers was less than 60 percent. Particularly, the rate of correct answers for condition 1 was about half of that of condition 2. Pressure and vibration were presented in condition 1, and sound was presented in condition 2. Thus we can say that sound plays more important role than pressure and vibration. Indeed, it seems that we are seldom conscious of pressure and vibration under the usual eating conditions, while we always clearly hear the sound. Therefore, it seems that it was difficult to correctly determine food from among the several foods presented by just pressure or vibration.

The rate of correct answers did not increase under condition 3, though it had more information available than condition 2. As we selected two foods from each category, there was a possibility that the test subjects were confused between the two foods that belong to the same category. Therefore, we recalculated the answer rates ascribing that the answer was correct if it belonged to the same category. The correct answer rates as recalculated for each condition were 46%, 63% and 72%, showing that presentation of pressure, vibration and sound marked the highest score.

4 Playback of the recorded data with arbitrary speed

4.1 Problem of current system

We have exhibited the SUI system in many countries around the world (SIGGRAPH2006, Ars Electronica 2006, WIRED Next Fest 2006 and World Haptics Conference 2007 etc.). While the exhibits confirmed that the SUI was a new fun interface for most people, beyond age, nationality and gender, some complained that the sensations were not natural. The complaints were especially about the sound when the user's drinking behavior (speed or power) was different from our assumed pattern. As we noted in section 2.3, we replayed the recorded sound and vibration at the same speed as we recorded. When the suction power of the user was different, the amplitude of the tactile sensation naturally changed with the power, because the tactile sensation was produced "passively" with the opening and closing of the air valve. For this reason the tactile sensation was not the subject of complaints.

However, auditory sensation is keener than the tactile one, and our replay of the sound was not automatically adjusted according to the suction power. Therefore, we sought a more sophisticated method to replay the sound that accorded to the user's suction power, without losing naturalness. As we use the same sound data for both auditory and tactile display, if we resolved the problem, the tactile sensation should also be improved.

4.2 Pilot study

For improving the auditory sensation, we focused on the sound of bubbling water, which we think is one of the funniest yet most familiar situations when drinking.

We tried three methods to replay the sound according to the suction power.

(1) The first method was just changing the volume of the sound in accordance with the negative air pressure.

(2) The second method was just changing the playback speed. In this method, if the suction power becomes smaller, the playback speed becomes lower, and hence the frequency also becomes lower.

(3) Third method was changing the play back speed while preserving the pitch. This is a common feature of current sound processing software [3][4].

We used these methods to generate sound with at arbitrary speed from a recording of bubbling sounds with strong suction power. We compared the generated sounds with true sounds of different suction powers that we recorded.

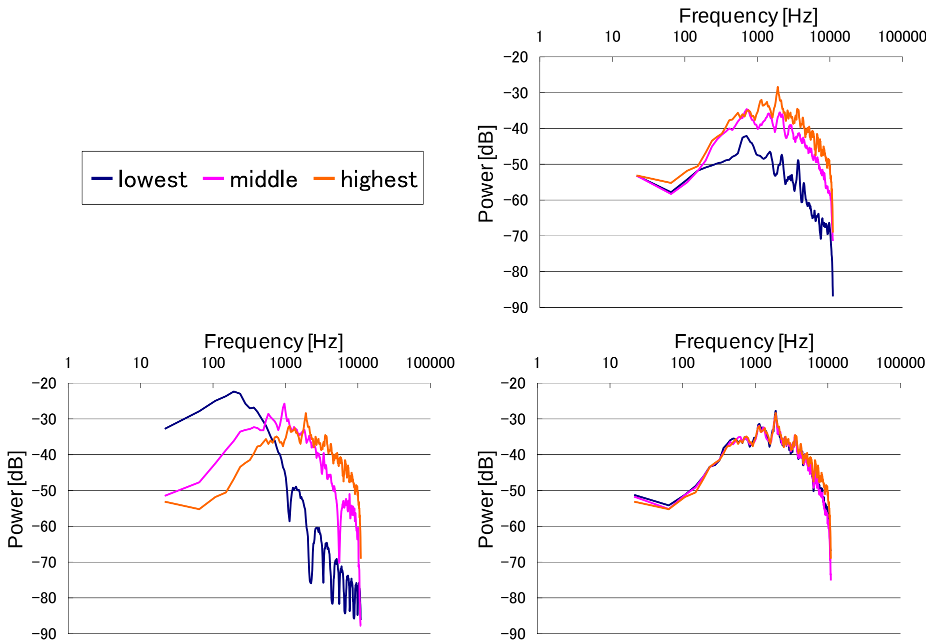

Fig. 4. The power spectrum of the sounds of water bubbling when suction power is low, middle or high. Upper right: real data. Lower left: just changing the playback speed. Lower right: changing the playback speed while preserving pitch.

However, all three methods failed to present the natural sound of bubbling water. In the first method (volume change), it was unnatural especially for low power case because the generated sounds did not have the feature of the sound of lower suction power. In the real sound, the power of low frequency waveform area becomes relatively small in accordance with the suction power (Upper right of Fig.4). Thus, we felt there was a sense of incongruity attached to the generated sound from the lack of balance of the power spectrum for lower suction power sounds.

In the second method (speed change), it became even more unnatural because the power spectrum of the generated sound was far from the real data (Lower left of Fig.4).

In the third method (speed change while preserving pitch), the power spectrum of the generated sound didn't change and so was also unnatural (Lower right of Fig.4). In conclusion, all three methods, which seemed the simplest ways, failed to replay the bubbling sound naturally.
When we display haptic sensation, the general method is to construct a physical model [5][6]. However, this is not practical in our case because the bubbling phenomenon includes dynamics of water and air around the straw, which is rather hard to model.

4.3 Our method for the replay of auditory sensation with arbitrary speed

In response to the above results, we temporary concluded that there was no simple method to generate natural sounds in different situations from one source. Thus we looked back and proposed a more basic method. This involved recording many sounds with different drinking power. This method doesn't have the spectrum problem mentioned above. Moreover, it is quite simple and didn't need a complex technique.
An overview of our method is as follows:

- *For recording, we prepare numerous sound tracks with different suction powers.*
- *For replay, we use the air pressure sensor to make a preliminary measurement of the highest drinking power of the user.*
- *Set the threshold level for each sound track.*
- *Change the track being played in real-time according to the user's drinking power.*

We recorded 10 steps of sound data with different suction powers.

4.4 Experiment

To confirm the availability of this method and determine the optimal number of steps for the sound, we conducted an experiment as follows;

- *Experience real bubbles with a real straw. We asked participants to repeatedly breath at a constant speed that accorded to the rhythm of a metronome (60BPM).*
- *Measure the highest drinking power of the participant while drinking in our system.*
- *Repeatedly carry out the drinking action, guided by the rhythm of the metronome, and rate the naturalness of the sound sequence on a scale from 1 to 5 (bad=1-2-3-4-5=good).*

We prepared five kinds of condition with different step numbers, and presented each condition randomly. Five participants (3 males, 2 females, aged 22-23) participated in the experiment. The evaluation was done three times for each step. Figure 6 shows the results.
From the graph, we observed that four steps of sound were needed to achieve a natural change of sound without a noticeable alteration. One participant commented

that he could not notice the shift of the states, and the sound was quite natural. Another participant commented that this sound experience with suction power was better than the experience using the previous system.
Interestingly, with 10 steps the evaluation became worse. We thought that this would have been caused by the noise that occurred more often when the sound was switched. We considered two solutions for this problem. One was to observe the waveform of the sound playing and change the sound when value of the waveform became zero. The other was to find “fundamental period” for each sound track, one which fully represented the characteristics of the track. If the sound was to change during that period it would lead to an unnatural switch phenomenon; therefore, we keep the track playing until the period was over. We have not yet the various fundamental periods, or indeed how many such periods exist; this will be part of our future work.

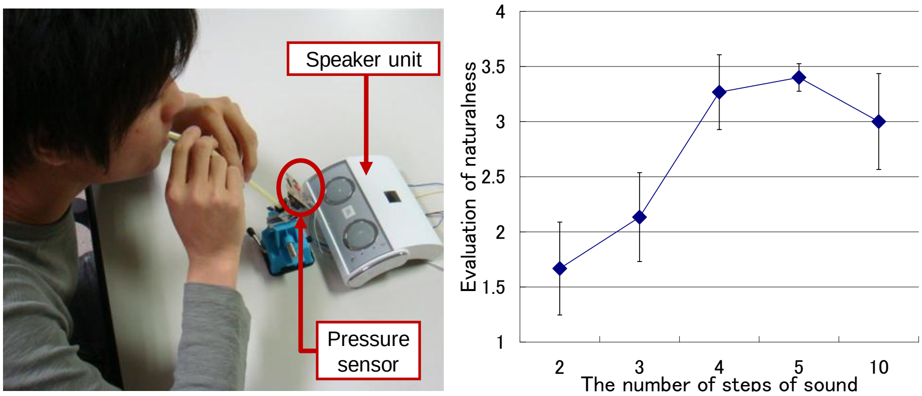

Fig. 5. Left: Experimental scene
Right: Evaluation of naturalness according to the number of sound steps

5 FUTURE WORK

The SUI can be proposed for use in various entertainment applications that can utilize a drinking sensation. The SUI can be used in many kinds of games; players will be able to experience a high level of reality in drinking through the technique of using the mouth with the SUI. In the field of communication, the SUI could present feelings (comfortable, evocative, yucky, etc.). Then, like Edible Bits [7], the drinking sensation can be converted into some information because lips and mouth are very sensitive regions of the human body. Therefore, if this sensation could be converted into words and emotions, we can receive the content of all messages from a unique drinking action and expand the breadth of communication. Moreover, it will be possible to add a new sensation to a real drink, so that ordinary water might feels like soda water, introducing a new type of augmented reality.

6 CONCLUSION

In this paper, we introduced a new version of our SUI that enables users to experience drinking through a straw. We showed the structure of our system, and evaluated it mechanically and psychophysically.
We discussed a problem that occurred with our earlier system; that recorded data needed to be replayed at the same speed without reference to the suction power. Some easy methods to overcome this problem failed. We thus proposed a method that involved preparing many sound tracks and switching them in accordance with the suction power. Our evaluation experiment showed that four or more steps were sufficient.
We will improve the SUI to enable even better presentations of the reality of the drinking sensation. We also want to apply this device to various fields such as for medical treatment of diseases related to the mouth and lips, or to produce a mixed reality situation to modify the usual drinking experience.

References

1. H.Iwata, H.Yano, T.Uemura and T.Moriya: Food Texture Display, 12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (2004) 310-315
2. Y.Hashimoto, N.Nagaya, M.Kojima, S.Miyajima, J.Ohtaki, A.Yamamoto, T.Mitani and M.INAMI: Straw-like User Interface: Virtual experience of the sensation of drinking using a straw, Proceedings of ACM SIGCHI ACE 2006 (2006)
3. J.Laroche and M.Dolson: New phase-vocoder techniques for pitch-shifting, harmonizing and other exotic effects, Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (1999)
4. E. Ravelli, M. B. Sandler and J.P. Bello: Fast implementation for non-linear time-scaling of stereo signals, In Proceedings of the 8th Conference on Digital Audio Effects (2005)
5. H.Yano, H.Igawa, T. kameda, K.Muzutani and H.Iwata: Audiohaptics: audio and haptic rendering based on a physical model, In Proceedings of Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (2004) 250–257
6. D.DiFilippo and D.Pai: The AHI: An audio and Haptic interface for contact interactions, Proceedings of UIST'00 (2000) 149-158
7. Dan Maynes-Aminzade: Edible Bits: Seamless Interfaces between People, Data and Food, CHI2005 (2005)