

# Proposal of Internet Radio Walking Around Street While Listening to Voice of Virtual Idols

Shiori Furuta<sup>(✉)</sup> and Katsuhiko Ogawa

Keio University, 5322 Endo, Fujisawa-shi, Kanagawa 252-0882, Japan  
{s13791sf, ogw}@sfc.keio.ac.jp

**Abstract.** When walking around the city with another person, there are many interesting things that are much more noticeable than if one is alone. However, nobody can always walk with another person. This research proposes a location based internet radio walking program together with virtual idols. There are two versions of the radio program: voice only and voice with vibration. This research describes design and experiments for both the voice version and voice and vibration version. It also clarifies future research directions.

**Keywords:** Walking · Internet radio · Vibration

## 1 Introduction

When walking around the city with someone, there are many interesting things that are much more noticeable than if one is walking alone. As a substitute for walking with another person, it may be possible to use a smartphone that is not a person. Like RPG games, virtual idols appear in smartphones, and you can walk while watching them on the smartphone. For example, Pokémon GO [1], released in 2016, corresponds to this, although no RPG messages were displayed.

However, “walking smartphone,” or using smartphones while walking, has become a social issue; it involves hitting people walking without looking around. In a survey announced in 2015 [2], 91% of participants in the survey were bothered by others walking with smartphones. Serious accidents have occurred even in Pokémon GO.

In this research, we propose a service—“Walking with Virtual Idols (WaVIs)”—for people walking alone by offering voice together with virtual idols of speech using internet radio. The purpose of this service is to rediscover the charm of the city even when people are walking alone with the help of virtual idols.

As conventional studies of street walking services using radio demonstrate, there are “Podwalk” services [3]. This service is based on the premise that podcast voices are downloaded to playback equipment in advance, and one can walk while watching the printed map. Since the voice is determined by time flow, the walking speed and voice do not match. In addition, they refer to the podwalk from past experience rather than for present time.

This research proposes a service that allows for walking while watching the sceneries in the town by listening to voice guidance by virtual idols, walking at the pace of users.

## 2 Concept of “Walk with Virtual Idols (WaVIs)”

Women are generally sociable and often walk in groups, but men often walk alone. For this reason, women as virtual idols often appear in games. In addition, because there are many female college students in our laboratory, and considering the convenience in the making of the audio content, WaVIs used female college students as virtual idols.

This service creates the effect that when you walk around the city listening to the voice of virtual idols, you can imagine someone virtual from the voice and have a sense of walking together in the city. By doing this, we believe that awareness of a new viewpoint can be gained through the virtual idol pointing out places.

In WaVIs system, voice comments of virtual idols are mapped onto Google Maps before use [4]. Looking at Google Maps on the browser of the smartphone, we automatically acquire location information, and this becomes a mechanism that automatically sounds when it comes to the place of the voice comments. As shown in Fig. 1, it is possible to listen to the voice of the virtual idol independent of the walking speed of the user.



Fig. 1. Map to voices

To use WaVIs, open the page mapped on the smartphone and listen to the sound with the earphone. The user can walk along the route only from the starting point; so, the user walks along the street along with the voice.

## 3 Walk with Virtual Idols Ver.1 - WaVIs1

### 3.1 Design of WaVIs1

WaVIs1 (Walk with Virtual Idols ver. 1) is a town-walking service where only the voice of virtual idols flows. The recording is monophonic. We do not set the character of virtual idols, whose voice is designed to mimic a general female college student using the voice of one author. One voice speaks for around 5 to 15 s, and these voices are mapped in front of the intersection, the turning point, and the store, in order to introduce them. Information on the road is also made by a virtual idol voice. The route is fixed and requires walking in one direction.

### 3.2 Experiment

WaVIs1 experiments were conducted around Roppongi in Tokyo. To create voice comments, we first conducted a “survey.” We walked in the town and determined what content would be used. We then “recorded” the content decided in the “survey.” Finally, we “pasted” the sound “recorded” on the map. Figure 2 is a map-to-voice in Roppongi.



Fig. 2. Map to voices in Roppongi

**Method.** The purpose of this experiment was extraction of improvement points. After four subjects walked using WaVIs1, comments and impression were sought via interview.

**Result.** There were no problems in length and placement place. It seems that individuals felt as if they walked together, as the sound matched the pace of walking. Interviewees claimed it was not an image of a general Japanese female student that they imagined, but rather a calm voice like a radio personality. Interviewees also noted that because virtual idols spoke in an explanatory style, they felt that the virtual idols and distance cannot be measured. They seemed to want to engage with the voice like they would a friend. Also, it seems that emotion was hard to understand because there was no intonation in the voice. Interviewees felt virtual idols and a sense of distance, and

seemed to not understand with much familiarity. Thus, they were less likely to feel emotions toward a virtual entity. However, it is convincing that this virtual idol is calm if a female college student's radio personality has been used.

**Discussion.** WaVIs1 could not create virtual idols. There are two main reasons for this. One is that the interviewees did not feel familiar with the virtual idols. Like a news-caster in a news program, the lack of tone makes explanation difficult; therefore, it is hard to understand emotions because there is no intonation in the voice, as there would be for a machine voice. This made it difficult to feel emotions towards virtual idols.

The second reason was due to a gap with regard to the female college student that each person thought. For example, one person might think of a bright and energetic female college student as a general female college student, while another might imagine a calm female college student thereby resulting in the voice and the female college student imagined not matching.

To solve these problems, WaVIs2 decided to make three improvements. The first was to use binaural recording to create a spatial sound. The second was to set up characters to eliminate the imaginary virtual idol gap. The third was to use a vibration device that uses a sense other than hearing.

## 4 Walk with Virtual Idols Ver.2 - WaVIs2

### 4.1 Design of WaVIs2

Based on the experiments in the previous chapter, we improved the WaVIs1 so that it included binaural recording, a virtual idol character setting, and a vibrating device known as the "Whispering Street Vibration Device (WSVD)".

**Binaural Recording.** Since there were many opinions that WaVIs1's virtual idols were unfamiliar with, leaving a gap, we adopted binaural recording as an improvement measure. When binaural recordings are heard with the human ear, they become a sound with a three-dimensional feeling, so it is as if the sound is moving. Experiments were conducted to determine whether users had the effect of feeling familiarity with virtual idols through this binaural recording.

In the experiment, eight subjects (male and female university students) who were subjects heard a sound for about 30 s. The content included a voice, wherein virtual idols walked together while completing a tour of the city, and also included a voice that whispered when approaching the ears of the subject on the way. As a result of listening to this sound, subjects felt their heart beating when approaching their ears and the result was that they felt a sense of intimacy.

**Pattern of Virtual Idols.** In order to eliminate the virtual idol gap imagined by the user, we provided the virtual idols with characters. About 50 persons participated in the experiments. These participants listened to the sound of WaVIs1 and wrote comments about the virtual idols on research. The experimental results are as shown in Table 1.

**Table 1.** Result of workshop

Appearance	<ul style="list-style-type: none"> <li>• Normal • Some fleshy • Not too cute • Small</li> <li>• Face that looks everywhere • Black hair, with bangs</li> <li>• I do not make a face like RINNA [5], but behind the scenes it is attractive</li> </ul>
Catch copy	•“Mimidoru”
Personality	<ul style="list-style-type: none"> <li>• People who are popular with grandpa and grandma</li> <li>•Gentle • Positive • Smiling</li> </ul>
Hobby	<ul style="list-style-type: none"> <li>• Virtual idol is traveling around a fashionable café</li> <li>• There is a stylish commitment store (get going)</li> </ul>
Image episode	<ul style="list-style-type: none"> <li>• Carrying goods • Honors student • Observes traffic rules</li> <li>• Moves in a red bicycle • She is playing with little girls in the park</li> <li>• She brings a lunch box for lunch • Sento (Japanese bathroom) bytes</li> <li>• Meet at any time • Go to city hall • From the area</li> <li>• Common to shopping districts • Go to local cafes</li> <li>• She is familiar with the area • Grandpa is good for grandpa</li> <li>• The virtual idol’s parent’s store is a bakery/rice store • Signboard girl</li> </ul>
Ideal plan	<ul style="list-style-type: none"> <li>• Tsundere. I hope there is a zone to be drawn when walking.</li> <li>• I want you to follow one step behind • Breath rising on a slope</li> <li>• I am a clerk • Small enough to place in the palm of your hand</li> <li>• Show your face as far as possible from the telephone pole</li> </ul>

We picked up some of the comments made, and the images were developed as per the five below.

The first image uses the catch phrase “Mimidoru.” The second is of a cute female college student. The third is of a girl who thinks that simple is best. The fourth is of a girl familiar with the urban environment. The fifth is of a girl who likes fashionable items like cafes.

The first catchphrase has a role related to making friends through greetings. This girl feels familiar with the way the talent speaks on television gourmet programs concerning the second and third items. The fourth and fifth items are focused on selecting a route with shops within the city.

**Vibration Device.** In WaVIs2, we thought that virtual idols would be easier to imagine if we were to stimulate the user’s five senses in addition to the voice of virtual idols. For example, two of Japan’s most famous virtual idols, Hatsune Miku [6], are used to provide visual (visual) and voice/song (auditory) aids. Based on this, we decided to combine the other five senses and create a vibrating device (WSVD) to make use of the tactile sense. We also thought of using the olfactory sense, but decided to instead use the tactile effect with the least impact, so as not to disturb the user’s own experiences, such as their smelling bakeries on street corners. As a reference, there is a device called “Buru-Navi 3 [7]” by Nippon Telegraph and Telephone Corporation. This is a device that creates a feeling as if one is being pulled on when it is held in the hand, and it is used as a navigation device that shows the direction of travel. This time we made a prototype of a wristwatch-type vibration device WSVD with Arduino. Figure 3 shows an image used when actually going to the town.





Fig. 3. Vibration device on left hand

## 4.2 Experiment

WaVis2 conducted a physical experiment of WSVD. We experimented with the author mapping the pattern of three vibrations. This time, the device was set to Kanagawa, a neighboring city of Tokyo. The reason for choosing this town is that the author knows the area well because there are universities to which the author belongs.

**Method.** The purpose of this experiment was to select the mapping pattern of vibration. We tested which mapping patterns were easiest to imagine the largest number of virtual idols. There were three patterns. Pattern A is a pattern that adds vibration to every-thing before the mapped sound is reproduced (Fig. 4). Pattern B vibrates between the mapped voice and voice (Fig. 5). Pattern C adds vibrations before voices that are desired to be focused on among the mapped voices are reproduced (Fig. 6).



Fig. 4. Pattern A



Fig. 5. Pattern B



Fig. 6. Pattern C

**Result.** Pattern A was too frequent at the point where contents were dense, but in places where there was time between contents, it created the attitude that sound would come. Pattern B vibrated every time, so it got in the way of being disturbed when enjoying one’s time. The vibration was just right for Pattern C. This vibration provided the sensation that the virtual idols had something to share and that they imagined someone virtual.

**Discussion.** We felt that Pattern C was good, out of the three patterns we experienced. The reason for this is that the sound of the vibrated place could be clearly understood as being something that the virtual idol wanted to convey. Pattern C was divided into hours to enjoy virtual idols and town walking and time to enjoy walking alone by one’s self, and we were, in that pattern, able to experience and enjoy both varieties. Patterns A and B were associated with a small amount of discomfort. The reason for this is that there is a lot of vibration and it is disturbed by one person’s time. However, both patterns A and B were considered to be usable patterns when meaning was applied to the vibration.

## 5 Conclusion

In this research, we proposed and developed “Walk with Virtual Idols (WaVIs)” which allows a person to walk with the virtual idol of a female college student when walking alone in the city. In WaVIs1, we used only voice to extract problems. There were two

problems associated with this approach. The first is that one could not develop a feeling of familiarity. Another was that the general Japanese female student imagined by each user was different. To solve these problems, WaVIs2 produced a binaural recording, a workshop for making virtual idol settings, and a vibrating device “Whispering Street Vibration Device (WSVD).” Several small experiments were conducted using these.

Future research ought to improve WSVD. For example, the appearance of the device could be improved. For instance, it could be worn like a necklace, like a wrist-watch, and could be transformed into various shapes. Repeat prototype production, experimentation and improvement will enable completion of WaVIs.

## References

1. Pokémon GO. <http://www.pokemongo.com/>. Accessed 14 Mar 2017
2. General Association of Telecommunications Carriers Association, “Survey on Walking Smartphones” (2015). [http://www.tca.or.jp/press\\_release/pdf/150123sumahochosa.pdf](http://www.tca.or.jp/press_release/pdf/150123sumahochosa.pdf). Accessed 14 Mar 2017
3. Kato, F.: Capture, share, and experience: “Podwalk” as a medium for flaneurs. In: PICS (Pervasive Image Capture and Sharing) Workshop, UbiComp2006, Orange County, CA, 18 September 2006
4. Utsumi, S., Ogawa, K.: The Experiment of Vicarious Experience by Using Place’s Second Sound Channel Media, “Whispath”, The institute of electronics information and communication engineers, LOIS2015-94, pp. 181–185 (2016)
5. Microsoft Japan – RINNA. <http://rinna.jp/>. Accessed 14 Mar 2017
6. Hatsune Miku official. <http://miku.sega.jp/>. Accessed 14 Mar 2017
7. Nippon Telegraph and Telephone Corporation (2014) - Buru-Navi 3. <http://www.kecl.ntt.co.jp/human/burunavi3/>. Accessed 14 Mar 2017