Research of Information Recognition in PC Operation with Auditory Signal

Hirohiko Honda*, Kazuko Tobita**

** Shonan Institute of Technology, honda@sc.shonan-it.ac.jp*
*** Shonan Institute of Technology, tobitak@center.shonan-it.ac.jp*

Abstract: By effectively introducing auditory signals and substituting visual information with auditory information, we designed a system in which screen status can be understood and transition can be recognized. By introducing the concept of soundscape into the field of engineering, we were able to utilize it in producing materials for auditory signals applied in this study. We make auditory signals selectable according to user characteristics by modifying artificial auditory signals that are similar to environmental sounds and combining them with voice synthesis. To make the system proposed this time to work effectively also for users with a progressing disease, visual-dependent information was gradually substituted with auditory information in accordance with the progress of symptoms. In this manner, we can expect that the performances using PC does not have to be reduced as much as possible if it can be adjusted gradually with synchronization with disability state. In practice, we applied the design of this study to a application with typical form, tested with intractable disease patients, and successfully verified its effectiveness.

Key words: Human Interface, Assistive Technology, Adaptation, Soundscape

1. Introduction

Some patients with intractable diseases such as muscular dystrophy, Spinal Progressive Muscular Atrophy (SPMA), Fibrodysplasia Ossificans Progressive (FOP), and Amyotrophic Lateral Sclerosis (ALS) have range of motion of the body that gradually loses freedom as the disease progresses, and depending on the condition of the disease it becomes difficult for them to work continuously or maintain posture. Use of PC is highly beneficial to them that they can use Braille and voice synthesis to input and output various information by themselves regardless of screen.

On the other hand, current PCs that depend highly on GUI, require users to operate a mouse while viewing the screen, thus no work can be done unless the user understands the screen configuration. There have been attempts to overcome this problem such as GUI access technology [1] called a screen reader with which screen information is read out and various other input assist devices [2].

However, even though vision-independent operations are possible, those methods require the screen to be read out in sequence for the user to select and determine the operations and therefore advantages of the GUI, i.e., list display and intuitive operations, are not fully utilized. There is a study on a technique to express the GUI screen layout with a three-dimensional sound configuration for intuitive operations [3] but the screen status can only be determined by mouse operations and cannot be determined in a static state where nothing is operated. Patients having difficulty in maintaining their posture cannot keep setting in front of the display for more than a certain length of time, and thus a length of time that they take their eyes off the screen becomes longer. This means that they need to reconfirm the state of where the pointer is focusing, thereby reducing efficiency from viewing the screen again to determine the state and decide on the action.

One of the characteristics of intractable disease patients have in common is progression of disability. In particular, for patients who have difficulty in maintaining their posture, it is difficult to maintain the current status of operation devices and applications. Therefore designing of a device that can be adjusted with the user's operation characteristics in mind is desired. There are many prior studies for a method in which the machine adjusts to human [6] such as self-adjusting keyboard in view of the users' operation characteristics [4] or a skill enhancing effect by machine characteristic modification [5]. However, those studies do not present specific measures for intractable disease patients with a progressive disease.

In view of those issues, this study intends to assist users in accessing to a PC by presenting a part of GUI's visual information as acoustic information. In particular, for patients with growing difficulty in keep viewing the screen as the disease condition develops, acoustic information is gradually added so that the users can understand the state even when they are not viewing the screen.

The approach of this study allows the user to select and determine next action without viewing the screen, thereby expecting the efficiency of the PC operation to be maintained as much as possible.

format is to be used for submissions that are published in the conference proceedings. We wish to give this volume a consistent, high-quality appearance. Therefore, we ask that authors follow some simple guidelines. Essentially, you should format your paper exactly like this document. The easiest way to do this is simply to download a template from the conference web site, and replace the content with your own material. The template file contains specially formatted styles naming (e.g., Abstract, Heading, Contact/e-mail, Author, Figure No, Table No, Reference, Text, Table Text) that will reduce your work in formatting your submission

2. Information Presentation using Auditory Signals of a Screen

2.1 Target of This Study

This study targets at patients who have difficulty in maintaining a posture due to progressive diseases and thus cannot keep sitting and maintaining the same posture in front of a PC. Although they are not totally incapable of viewing the screen, they have difficulty in keeping their eyes on the screen as they have to turn their heads or twist their bodies when changing posture. They will restore physical strength after a while and it becomes possible to face the PC again but the length of time they take their eyes off the screen gets inevitably longer.

Even when they are not facing the display, they can move their hands and other parts of the body, and thus can operate the mouse and the keyboard. For this reason, acquiring of screen-independent information by voice information using a screen reader is possible, but to such targets it is not rational to simply recommend a system that is for the visually impaired. It is well known that in terms of being screen-independent, a screen reader with voice synthesis device is useful. However, we have subjective evaluations indicating that it is rather bothersome for the users who can also acquire visual information to receive all the GUI-independent screen information via voice or sound.

Thus the main target of this study is not the visually impaired who are totally unable to view the screen or patients who use alternative devices such as head-mounted display. Basically, this study targets at patients who do not rely on special assist devices for the disabled and wish to continue to use general devices as long as possible as healthy people.

2.2 Introduction of the Concept of Soundscape

When using a PC, patients who have difficulty in maintaining a posture particularly wish to understand the state while they are not viewing the screen. In order to allow these users to understand the screen status, this study intends to build a system to notify the user of such information as to what application is running, what is being focused on, and what stand-by state as acoustic information of not by voice but by auditory signals even in a static state where nothing is operated. OSs such as Windows have a customizing function to output various auditory signals when an event is raised but its application range is limited and there is no effective tool to recognize screen information in a state where nothing is operated.

Auditory signals that users may feel as bothersome, noisy and annoying cannot assist a comfortable operation. This study is introduced with a concept of soundscape, in which sound is handled in the context of relationship between environment and human [7]. The word soundscape is a compound of “sound” and “scape”, a suffix meaning “scene/scenery”, thus meaning “sound scenery”. This concept has been introduced into not only academic fields but also environmental designing or art and cultural activities and the like.

Auditory signals in this study are output but preferably with a background at a level that doesn’t bother users unless they become conscious of it. For example, a room ventilation fan makes negligible sounds but not really bothering. Even if some of such auditory signals overlap, at the time the users becomes conscious, they can immediately distinguish the sounds and can understand the state of the PC screen by the difference in meanings given to the sounds. Consequently, in this study, auditory signals are an important element. This coincides with the concept of soundscape where sounds are regarded as scenery. We expect to use some aspects of soundscape, i.e., emitting a sound that gently appeals to human consciousness and planting them as auditory signals and try to introduce them into a new field, the field of engineering.

2.3 Development of the Test Tool using the Auditory Signals

Purposes of using PC in the daily life for people with disabilities and patients are similar to those of many healthy people: Internet browsing using a browser, acquiring and transmitting information using a mailing tool, preparing documents, playing games, and appreciating and editing music and videos. As verified in the previous section, static state appears for a long time also in a mailer where text display and character entry are the main purposes of use. So in most applications, presentation of screen information in a static state is expected to operate effectively. If the test subject only grasps the structure on the screen and the state of the window in focus or the application, he can select and determine the next action, thereby allowing himself to predict resultant event raising and state transition.

On the other hand, a GUI of each application has wide variation including menu, tools, various buttons, the number of settings, arrangement, hierarchic structure, functions, and so on. However, displays or functions presented in there are determined unambiguously and do not undergo changes like the Internet or texts in e-mails. Therefore the text contents need to be read out in a voice but tool buttons and menus can be replaced with auditory information by sound effects. However, a current GUI has numerous operation environmental conditions, and thus it is very difficult to transmit everything as acoustic information. So, in this study, we created a test application with a standard menu and a few command buttons and verified whether the test subject could grasp and discriminate the screen status only with the auditory information in a limited number of operation environmental conditions.

3. Verification Experiment with Test Tool

The test application has two types of menus, each of which is provided with three hierarchic structures. In the window, four command buttons with different functions are arranged. In practice, a click does not raise any event other than a quit button. The test application was developed and produced in Microsoft Visual C# 2010 Express Program Development Environment in a PC with Windows 7 OS.

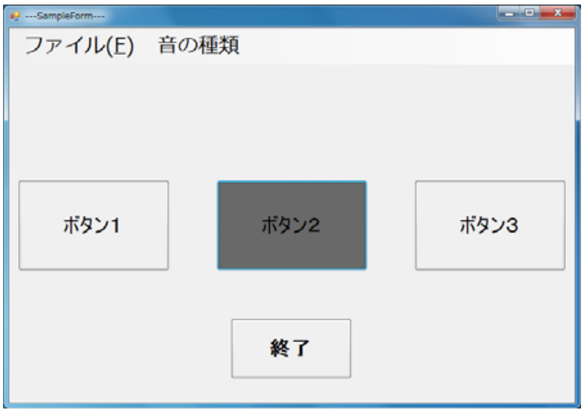

Fig. 1 Test Application

The auditory signals are assigned as follows.

- Menu tag identification = 2 types
- Menu item identification = 3 * 2 = 6 types
- Button identification = 4 types
- During application power up = 1 type
- In event raising by click = 1 type

By assigning a total of 14 types of auditory signals, the situation in the window when the application is activated is estimated to be possible. It needs to be constantly notified that the application is in operation and the selection status also by keeping a loop playback of an auditory signal assigned to the target while the target is focused. For this reason, multiple auditory signals can be output simultaneously.

The auditory signals are not single tones but three patterns of basic tones prepared in WAV format. The sound sources were selected from a sample library of tones produced by a software synthesizer and modified by an oscillator and a filter. The three patterns were introduced with the concept of soundscape with the image of comfortable tones that are thought likely to arouse audio sensitivity and thoughts for scenery and nature. This is because applying the concept of soundscape to auditory signals in this study will coincide with one concept of sound environmental design, i.e., to make people notice the presence of sounds that they were never conscious of before and discover various sounds. More specifically, we prepared the following three patterns of basic tones : science fiction related tones that make people image a vast cosmic space, tones of nature that are close to the sounds of insects or winds, and tones of musical instruments that remind people of church bells. Types of tones were prepared with a minor difference in between the items in the menu such as changing the pitch of a sound. The volume of the auditory signals is appropriate when it is at a barely-noticeable level during playback as background sound and the user himself is to adjust the volume with the knob.

We asked the user to select auditory signals in his preference from the above three basic patterns, then actually operate the application for a few minutes after adjusting the volume. We waited until he got accustomed to the application to some extent and then tested his operations performed without viewing the screen at all. The verification items are as follows.

1: Whether he can recognize a focused target (by replying which menu item, button number and form the focus is on).

2: Whether he can raise an event in a designated menu item (by using the Tab key or the cursor key to go round and focus and then clicking on the key)

3: Whether he can raise an event in a designated button number (by using the Tab key or the cursor key to go round and focus and then clicking on the key)

4: Whether he can quit using the quit button (by using the Tab key or the cursor key to go round and focus and then clicking on the key)

The test subject faces the screen in the opposite direction, and without viewing the screen at all, starts the task from a state set by the questioner. The test subject has to reply within five seconds after he was given a question. However, the first question was to be given at random during operation. In the above condition, we gave him questions 1 to 3 ten times each and a question 4 in the end and calculated the percentage of correct answers. This was set as one rotation and three rotations were conducted per person with time intervals.

Experiments were first targeted at healthy people. We conducted three rotations each to 14 males in their twenties who were accustomed to PC operation. The results are shown in Table 1.

Table 1. Experiment Results

	Percentage of correct answer to question 1	Percentage of correct answer to question 2	Percentage of correct answer to question 3	Percentage of correct answer to question 4
14 healthy people (averages)	86%	82%	94%	98%

The experiment results indicate that percentage of correct answers on grasping the focused target, the selection status of menus with a hierarchic structure, and the selection status of the command button are more than 80% each. The percentage of correct answer to question 2 was slightly low. We infer that it is because the tones of auditory signals to distinguish two types of menus were alike, and thus it took time to determine by going round the focus with the Tab and failed to process within the five minutes' limit. We expect that replacing of auditory signals will enable the users to distinguish tones with higher accuracy. We did not see any particular difference in the percentage of correct answers by differences in the three basic auditory signal patterns.

As a result, we successfully verified in these experiments that the test subjects can recognize where the focus is on in a state when they are not viewing the screen or even in a static state when they are touching nothing, and active key operations allow them to sufficiently distinguish and recognize the number of conditions as they intended and to operate as they desired.

4. Design of a System that Substitutes Visual Information

4.1 Evaluations by Patients with Intractable Disease

System of this study, it may be validity if applied to a typical Windows application, whose arranged some buttons and menus. Therefore, We developed the original application with reference to the generic Windows applications in which this system was embedded. We asked two patients with intractable diseases (distal Myopathy patient A and Fibrodysplasia Ossificans Progressiva patient B) who have difficulty in maintaining posture to try out the mailer, conducted a simple verification, and asked for their comments.

With reference to a Windows application used by the test subjects in their everyday life, we simplified the functions by narrowing the minimum functions such as click of a button and menu. To simplify the operations, we prepared a system with limited options so that the test subjects will not be at a loss as to what to do next. Hence, the mailer is operated only with the cursor key, the Tab key, and the Enter key. The mailer was developed and produced on a PC with Windows 7OS in Microsoft Visual C# 2010 Express program development environment. We also confirmed in Windows XPOS environment.

Fig. 2 Form of Application

Auditory signals are assigned as follows.

- While main screen is actuated = 1 type
- Identification of main buttons = 2 types
- Identification of main menus = 12 types
- While sub screen is actuated = 2 type
- Identification of sub buttons = 4 types
- Identification of sub menus = 14 types
- While sub sub screen is actuated = 4 type

- Identification of sub sub menus = 4 types
- When button is clicked = 4 type
- Quit identification = 4 type

A total of 51 types of auditory signals are assigned. We produced sound materials by using and adjusting the three patterns produced in Section 3.

In the verification, we measured the length of time until they achieved the following instructions.

1: Starting from the main screen, entering into a menu mode, click to select the specified item. Then click on the sub-screen mode by clicking on the button 1 to select the specified item in the menu mode of the sub-screen. Then, the subject exits back to the main screen press the Close button. Repeated three times given a different menu with different buttons this operation.

2: Proceed in the same way as before verification. However, without having to go back to the main screen from the sub-screen, navigate to the screen under the hierarchical structure of the sub-screen further. Click the item that is specified in the menu mode, then return to the main screen and sub-screen through. Repeated three times given a different menu with different buttons this operation.

We asked the users to operate for about five minutes and then we tested in a state where the users do not view the screen at all. After they quit, they were given a certain time of interval, and we asked them again to use the mailer for another five minutes in a state where they view the screen, and we conducted a verification experiment as done previously. In this way we conducted a total of three tests for each person. The verification results are shown in Table 2.

Table 2. Experiment Results

Test subject	Verification items	First measurement time	Second measurement time	Third measurement time
A	1	58 sec	52 sec	44 sec
A	2	97 sec	92 sec	85 sec
B	1	49 sec	45 sec	38 sec
B	2	67 sec	57 sec	51 sec

To acquire subjective evaluations, we asked for feedbacks from the test subjects in the form of questionnaire on operational feelings of the system, discrimination of auditory signals, and enjoyment during operation. As a result, we had one comment that after repeating a few times he became less lost and could process relatively smoothly. According to what he said, he was at a loss at the beginning in assigning information by sounds, but after a few minutes' use he came to understand the meaning of the sounds, and was capable of appropriately proceeding the process. There was another comment that the test subject could work with a sense of enjoyment and never got bored.

Table 2 indicates that as the tests progressed, the amount of time tended to reduce for both test subjects A and B. We could judge from the comments made by the test subjects, a major factor of the time reduction owes significantly to their adaptation to operation. The test subject B is more accustomed to the use of PC and can operate it faster. Although the test subject A also has a long experience in using PC, his disability in movements keeps him from moving positions and hands smoothly like healthy people, thereby taking relatively long time.

In this verification experiment, we conducted measurements basically in a state where the test subjects were not viewing the screen at all. Hence, we have not verified as far as to whether the lack of visual information has been complemented gradually with acoustic information. However, similar to the experiment results in Section 3, the test subjects were able to recognize the screen status without viewing the screen even in a static state and were able to sufficiently distinguish acoustic information even with a relatively large number of operation environmental conditions as this time and we have successfully verified that it is possible for them to conduct desired operations.

4. Conclusions

By effectively introducing auditory signals and substituting visual information with auditory information, we designed a system in which screen status can be understood and transition can be recognized. By introducing the concept of soundscape into the field of engineering, we were able to utilize it in producing materials for auditory signals applied in this study.

We analyzed the utilization situation of the users while the application was actuated and found out that there was a static state for a long length time in reality even with a tool whose main uses are text display and entry. This study focused on this static state and discussed promotion of efficiency of operations by providing screen information in this state.

To make the system proposed this time to work effectively also for users with a progressing disease, visual-dependent information was gradually substituted with auditory information in accordance with the progress of symptoms. In this manner, we can expect that the performances using PC does not have to be reduced as much as possible if it can be adjusted gradually with synchronization with disability state. In practice, we applied the design of this study to a mailer, tested with intractable disease patients, and successfully verified its effectiveness.

Progress of disease conditions often gives patients mental pains and we hope that this system can play even a small role in relieving those pains.

References

[1] Ishikawa, J., “GUI Access Software for the Blind Current Stage and Europe :in North American and Europe”, Information Processing Society of Japan, 36-12, 1133-1139, (1995).

[2] Hatakeyama, T., Koizumi, Y., Umeda, Y., and Uchio, M., “The Design and Implementation of Pointing Device for a Multimedia Users with Severe Physical Disability”, The journal of the Institute of Image Information and Television Engineers, 52-9, 137-143, (1998).

[3] Aritsuka,T., Hataoka,N., “GUI Representation System Using Three-dimensional Sound for Visually Disables”, Journal of IEICE, 84-5, 333-336, (2001).

[4] Fukuda, S. and Hiwaki, K., “Proposal of Automatic Adjusting Pushdown Keyboard-Toward an Age When Machines Adjust to Humans-“, The Japan Society of Mechanical Engineers, 15, 372-375, (2005).

[5] Igarashi, H., “Skill Enhancing by Subliminal Calibration for Machine Operation Skill Quantification using Input Characteristics”, Proc. Conference on Information, Intelligence and Precision Equipment (IIP2010), 10-7, 102-105, (2010).

[6] Kiguchi, K., Hayashi, Y., and Kose, Y., "An upper-limb power-assist system with perception-assist in accordance with a handling tool and movement", Proc. Conference on Information, Intelligence and Precision Equipment(IIP2010), 10-7, 82-85, (2010).

[7] Nishimura Atsushi,"The Concept and Technique of a Soundscape Design", Bulletin of Okinawa National College of Technology No.2, 127-130, (2008).