Wearable Intelligence for Musical Performance and Handiwork

Naofumi OTANI[1], Goh YAMAMOTO[1], Yutaka SAKANE[2], and Yoichi TAKEBAYASHI[2]

[1] Graduate School of Informatics, Shizuoka University
3-5-1 Johoku, Hamamatsu 432-8011, JAPAN
{cs9016, cs9094}@cs.inf.shizuoka.ac.jp
[2] Faculty of Informatics, Shizuoka University
3-5-1 Johoku, Hamamatsu 432-8011, JAPAN
{sakane, takebay}@cs.inf.shizuoka.ac.jp

Abstract. This paper describes a new ubiquitous learning system that makes knowledge acquisition and sharing possible. The system enables learners who want to play music and handiwork to see their predecessors' know-how through multimodal knowledge contents which consist of movie clips annotated by text descriptions. The system also films the users' talks, music, sound and various behaviors on their job and contributes to knowledge accumulation. Players and workers can carry these contents by using wearable technologies, and learn from them anywhere.

1 Introduction

The development of mobile devices and electric musical instruments enables a musician to give a performance anywhere they want without bothering about moving to a studio or a concert hall. In order to express sounds and melodies with mobile devices, many electric instruments and interfaces have been developed[1-3]. However, it is still hard for people who don't know much about music or maintenance techniques to play musical instruments.

In order to convey knowledge or know-how about work procedures, paper media like instruction manuals or instructional books are often used. However, it is hard to express sounds or motions and to cover everything that happened in working time. Thus it is necessary to provide learning materials that make it possible to learn pragmatical knowledge.

In order to realize new learning materials that allow learners to understand sounds or motions, we use multimodal knowledge contents, which consist of audio-visual information. The multimodal knowledge contents consist of multimedia contents, text descriptions in natural language, and various sensor data. The system involved can reconstruct multimodal knowledge contents into the form which enables learners to easily understand them.

By using mobile or ubiquitous devices, we realize a new environment that can provide learning materials efficiently and can generate new contents while working. In this environment, the system can create learning materials, and at the same time, learners can learn. Thus they can improve their motivation of learning and working.

In this paper, we focused on two domains, learning how to play a musical instrument and do electric handiwork, and discuss the experimental results and the importance of resultant know-how.

2 Use and Accumulation of Multimodal Knowledge Contents

In order for novices to learn music performances and techniques for repairing and musical instruments maintenance we propose a system that uses mobile technologies. The system we propose enables learners to browse knowledge contents without interruption, and get feedback of performance itself or sounds on the moment.

Fig.1 shows the scene of playing an instrument and electric handiwork. Both tasks need both hands to hold an instrument or to manipulate industrial tools. In this case, it is necessary not to disturb their vision or handiwork while browsing learning materials[4].

We propose a concept of the model called wearable intelligence to realize the creation of learning materials learners can learn. It is shown in Fig.2. While browsing contents, we realize the way of using wearable devices, which makes it possible to record learners' working processes while browsing learning contents and to expand learning contents more variedly. In the following sections, we describe the use of contents and environments that generate learning contents.

(a) (b)

Fig. 1. Playing the contrabass (a) and handiwork (b)

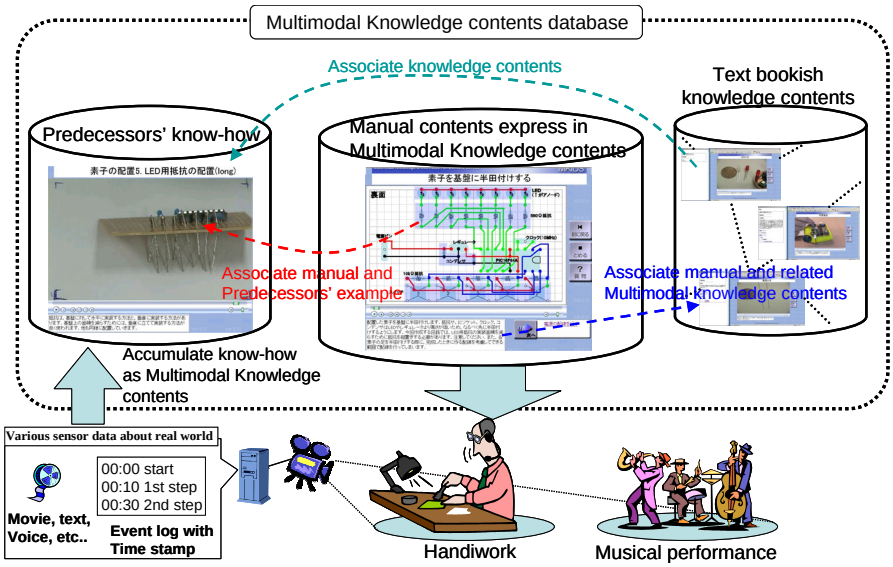


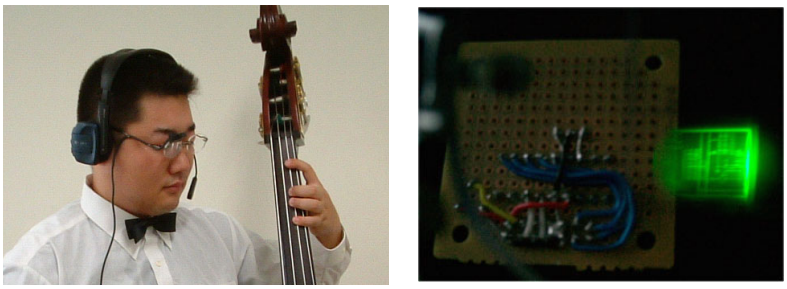
Fig. 2. The structure of a new ubiquitous learning system using multimodal knowledge contents.

2.1 The Use of Multimodal Knowledge Contents

Music performance and manual works need both hands of learners in learning or working. Thus it is important not to disturb their activity while bringing learning contents or contents of necessity.

We employed Multimodal Knowledge and Information on Demand Service (MKIDS), which was developed by one of the authors[5]. Users can access knowledge using a hands-free speech recognizer made with Bluetooth technologies[6], as shown in Fig.3(a). By using a speech recognition interface, the proposed system enables learners to browse or search for related contents, or to record their speech with their hands-free. Furthermore, by using a text-to-speech reading interface, learners can hear information about learning contents without hindering their vision. Learners can see the learning materials and objects on wearable display, as shown in Fig.3(b). The system provides multimodal knowledge contents through practical experiences in musical performances and handiwork.

In searching related contents, we use a MySQL database server that accumulates substances of contents and relational information between contents. By using the title and substance of a content that appeared in knowledge contents, the system provides various kinds of knowledge contents such as predecessors' works or explanations of technical terms.



(a) (b)

Fig.3. The wireless headset and wearable display enables the user to access knowledge in hands-free and eye-free.

2.2 Extraction of Multimodal Knowledge Contents

In order to produce pragmatically effective learning materials, it is important to record what happened in their work and accumulate various kinds of examples. Thus we have installed multiple cameras around each worker in the environment to record his or her working process and make it into contents. Fig.4 shows the environment we constructed to collect working contents.

With regard to the contents of the workers' speech, we accumulate them as audio data recorded with headset and as text data converted by a speech-to-text software. And this makes it possible to record the progress of their work or the timing of their speech.

By using the contents, we can take meaningful information from the flood of data recorded by multiple cameras. The created contents are related with existing learning materials, and used as new learning materials for others.

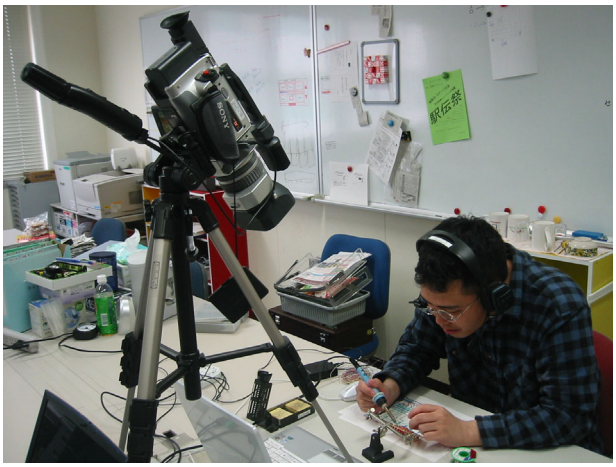


Fig. 4. The system films the worker's talks and motions.

3 Experiments

We conducted two experiments to evaluate multimodal knowledge utilization and accumulation by using multimodal knowledge contents freely. One was aimed at five people playing the contrabass. And the other was aimed at four people working on handiwork. As a result, we got five-hour movie clips of contrabass performance and eleven-hour ones of handiwork. Fig.5 shows examples of accumulated multimodal knowledge contents.

After the experiments, the subjects reviewed the filmed contents, and were asked questionnaire survey. From the questionnaire results, we found there are some useful contents which have none of their activities filmed due to the scope limit of the cameras but which have many useful text descriptions.

The experimental result suggested that it is not important for learners to increase the number of sensing devices such as sensors or cameras, but to increase that of clips annotated by expert workers or players.

From the viewpoint of the use of contents, the system can create more contents in proportion to the number of users and time. We must relate contents with each other in order for users to learn many things from any content that they choose.

(a) (b)

Fig. 5. Examples of multimodal knowledge contents that consist of movie clips and text descriptions.

4 Conclusions

There have been many researches and development on composing music and creating a new style of performances with mobile technologies. It is, however, far important to create new environments that make more people become familiar with music and instruments. Without activating musical communities in which people enjoy playing instruments, the proposed interfaces would be used only by limited people, and valuable technologies would be buried dead.

The proposed system has the potential to provide and produce various knowledge contents of learning, such as sports, arts, cultures, hobbies and so on. It is obvious

that there is no single well-defined recipe for developing successful multimodal knowledge contents. We must continue to focus on building rich and useful knowledge contents to support various human activities including musical performance.

We have to set out a vision of future ubiquitous society and application systems fitting social needs, and coordinate researches for infrastructure technologies, application technologies and evaluations. Furthermore, we must incorporate more and more knowledge to increase the range of our ubiquitous computing systems and knowledge creation.

References

1. Philippe Stanislas Zaborowski: "ThumbTec: A New Handheld Input Device," International Conference on New Interfaces for Musical Expression (NIME04), pp.112-115 (2004).
2. Arianna Bassoli, Cian Cullinan, Julian Moore, and Stefan Agamanolis: "TunA: A Mobile Music Experience to Foster Local Interactions," Proceedings UBICOMP 2003 (2003).
3. Andrea Polli: "Active vision:Controlling sound with Eye Movements," Leonardo, Vol.32, Issue 5, Seventh Annual New York Digital Salon, pp.405-411 (1999).
4. Marc von Waldkirch, Paul Lukowicz, and Gerhard Tröster: "Spectacle-Based Design of Wearable See-Through Display for Accommodation-Free Viewing," Pervasive 2004, pp.106-123 (2004).
5. Y. Takebayashi, and M. Doi: "Multimodal human computer interaction research at Toshiba research and development center," Computer-Human Interaction Extended Abstracts, pp.79-80, Computer-Human Interaction 97 (1997).
6. http://www.toshiba.co.jp/about/press/2003_05/pr_j1201.htm