Instant Learning Sound Sensor: Flexible Real-world Sound Event Recognition System

Yuya Negishi[1] and Nobuo Kawaguchi[1]

Graduate School of Engineering, Nagoya University
1, Furo-Cho, Chikusa-ku, Nagoya, 464-8601, Japan
negishi @ el.itc.nagoya-u.ac.jp, kawaguti @ nagoya-u.jp

Abstract. We propose a smart sound sensor for building context-aware systems that can instantly learn and detect events from various sound information with a small and low-cost device. Using the proposal sensor, a developer of a ubiquitous service can easily utilize a real world sound, like an event trigger to control appliances or for a detailed presence service, without a signal processing programming.

1 Introduction

Context aware systems are beginning to play an important role to support human activities in the real world. As the way to obtain the context information into the system, a lot of input devices such as accelerometer, pressure and temperature sensor are used. Small network devices with these sensors are developed such as Mote[1]. There are also some researches and systems using signal processing with a time series data from sensors to obtain context information. However, in signal processing, a design of the recognition algorithm of a complex pattern is not easy, because an analysis of a feature quantity requires a lot of time. It is hard for anyone to utilizing real world sounds for building a context-aware system using a life sound recognition. In this paper, we propose an "Instant Learning Sound Sensor" which can instantly learn signal patterns of environmental and life sounds on the fly, using a small and low-cost device.

2 Smart Sound Recognition Sensor

Obtaining information from the real world is one of the most important issues in ubiquitous computing. To make the connection between the real world and digitized world, current research trend is mostly focused on specialized device or algorithm for each real world event. However, there are enormous number of real world event types which should be recognized by these specialized device or algorithms. It is almost impossible to cover all of these events by studying one by one. Additionally, it is too difficult to develop a signal processing for an ordinary person.

In this paper, we propose an instant learning sound sensor built upon low-cost device. "Sound" is one of the most rich context media. We believe, by only using the smart sound sensor, most of real-world events such as walking, door open/close, vacuum cleaner, TV watching, tea pouring, bath room activities, etc. can be recognized. Of course, some of these events can easily recognized by other sensing devices such as mechanical switches or motion detectors. However, we think it is costly and bothersome to develop a device or a method for each event. While, our smart sound sensor is a single device and low-cost. Additionally, it is very flexible and it has a wide range of application, such as do-it-yourself smart spaces and rapid prototyping of context-aware systems.

The main contributions of this paper are follows.

(1)Proposing a simple sensor which can only recognize a few sounds. This enables a sound recognition in a low-cost device, which has less processing power and memory. If you want to recognize several events, just use several sensors.

(2)Consideration of on-site configuration of sensing algorithm and parameters. This enables a single device to be flexibly utilized. It enables ordinary person to utilize ubiquitous computing.

3 Design of Instant Learning Sound Sensor

3.1 System Architecture

For easily utilizing real world sounds as events, we designed the proposal system to have following three features.

(1) Instant Learning: A user is only required to record the target event sound a few times, then the system automatically analyzes the sound, and chooses the most appropriate algorithm and parameters to extract a feature quantity. Finally, it generates a lightweight recognition program for sound sensor devices.

(2) Simple Device: The generated process is possible to run on low-cost devices such a microcontroller. The system makes it as simple as possible.

(3) Smart Sensor: Sensor can process the signal by itself. It is possible to cooperate and distribute with other devices easily.

The proposal system consists of following two components.

(a) Sensor Configurator: This software supports event learning phase. This has a function to select an appropriate recognition process to sensor devices for each target events.

(b) Sound Sensor Device: This is a device, which actually runs a sound recognition process in event detection phase.

To make the sensor device simple, we implement the sensor configurator on a separated PC.

3.2 Light-weight Sound Recognition Processing

As an example of a light sound recognition process, we considered a signal processing using DP matching[2]. As feature quantities for a target sound, we use some characteristics such as frequency and amplitude characteristic through time. To calculate a similarity between a captured sound and a target event sound, we use a DP matching with a change of the feature quantities.

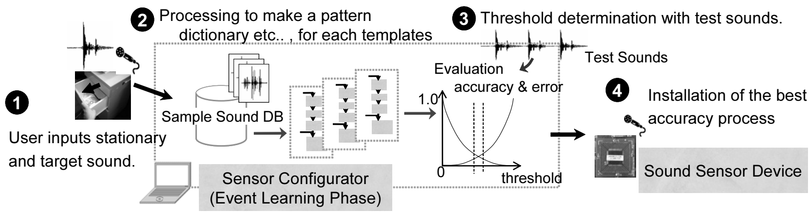


Fig. 1. Process in the Event Learning Phase

In this research, our aim is an automatic generation of unknown life-sound recognition program for small and low-cost devices. So, we have to think configurations of the parameters that is light processing and less memory consumption, such as less vector dimensions and less code book size. This requires a search for specific and appropriate parameter configuration for each target event sounds. So, we designed a flexible configuration method of the feature quantities to select a proper set for the target sound.

3.3 Instant Learning

In consideration of the previous section, we designed the sensor configurator that processes the event learning phase using the sample sounds database. This database has sample sounds and corresponding parameter configurations as templates. The template is an appropriate set of parameters for light recognition process, and it was optimized in advance.

In event learning phase, the sensor configurator searches some templates of similar sample sounds with a target event sound, and it estimates parameters for the target using the templates. We show steps of the process in the following and Figure 1.

(1) User inputs a target event sound, then the sensor configurator calculates similarities with each sample sounds about frequency and amplitude characteristics. Then, it selects some templates which has high similarities.

(2) Based on each parameter configurations of templates, it makes code book, using LBG algorithm[3], and code pattern dictionary of a target sound.

(3) The configurator determines a threshold of cost in DP matching in each parameter configurations, using some target event sounds inputted by user and non target sounds. As non-target sounds, it uses other sample sounds in the database.

(4) Finally, the configurator evaluates rates of accepted non-target sounds in each parameter configurations. After evaluations, it chooses the lowest error configuration, and installs the parameter configuration and recognition program to sound sensor devices.

By the sake of above process, user is only required to prepare some recorded target event sounds, and then the sensor configurator automatically selects an appropriate recognition process.

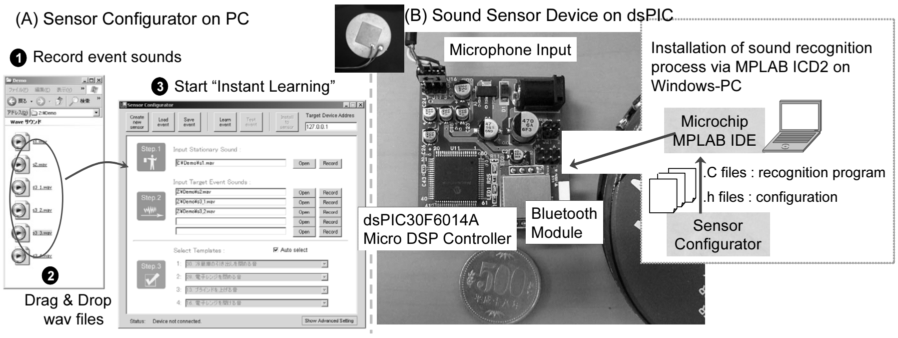


Fig. 2. Prototypes of Sensor Configurator and Sound Sensor

4 Prototype

Based on the above design, we implemented a first prototype of a sensor configurator and a sound sensor on PC and Microchip dsPIC micro-DSP controller (dsPIC30F6014A with 8KB RAM) as a small and low-cost device. We use a cheap piezoelectric device as a microphone, and SENA Parani-ESD200 Bluetooth module for an event notification. Figure 2 shows the prototype.

To make the sample sounds and parameter templates database, we have collected 30 events, and adjusted parameter templates as simple as possible by our hands in this implementation. Using the prototype on PC, we evaluated an accuracy of recognition process automatically generated by "Instant Learning". As a result, the accuracy of recognizing sounds of opening a window shade is 96.9%, opening a drawer is 83.3%, and writing on clipboard is 85.6%.

About memory consumption, the required RAM size can also come to less than 4KB in case of a sound of turning on a faucet, and the code size including const data is about 10KB.

5 Conclusion

We propose a smart sound sensor system. A user is only required to input target event sounds, and the system automatically generates recognition processes for small and low-cost devices. So anyone can build ubiquitous services utilizing real world sounds as event information without a signal processing programming.

References

1. Crossbow Technology, MOTE, http://www.xbow.com/Products/Wireless_Sensor_Networks.htm
2. H. Sakoe, S. Chiba : A Dynamic Programming Algorithm Optimization for Spoken Word Recognition, IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. 26, No. 1, pp.43-49, Feb.1978.
3. Yoseph Linde, Andres Buzo, Robert M.Gray : An Algorithm for Vector Quantizer Design, IEEE Trans.Commun.,COM − 28, 1, pp.84-95, Jan.1980.

Instant Learning Sound Sensor :
Flexible Real-world Sound Event Recognition System

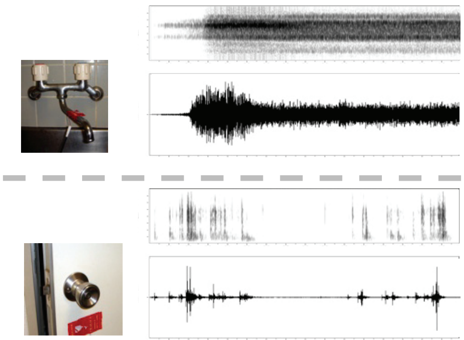
Yuya Negishi , Nobuo Kawaguchi (Graduate School of Engineering, Nagoya University)

We propose a smart sound sensor for easy building context-aware systems utilizing real-world sounds. The proposal system can instantly learn and detect life-sounds as events with a small and low cost device.

DESIGN OF INSTANT LEARNING SOUND SENSOR

<u>(1) Instant Learning :</u>

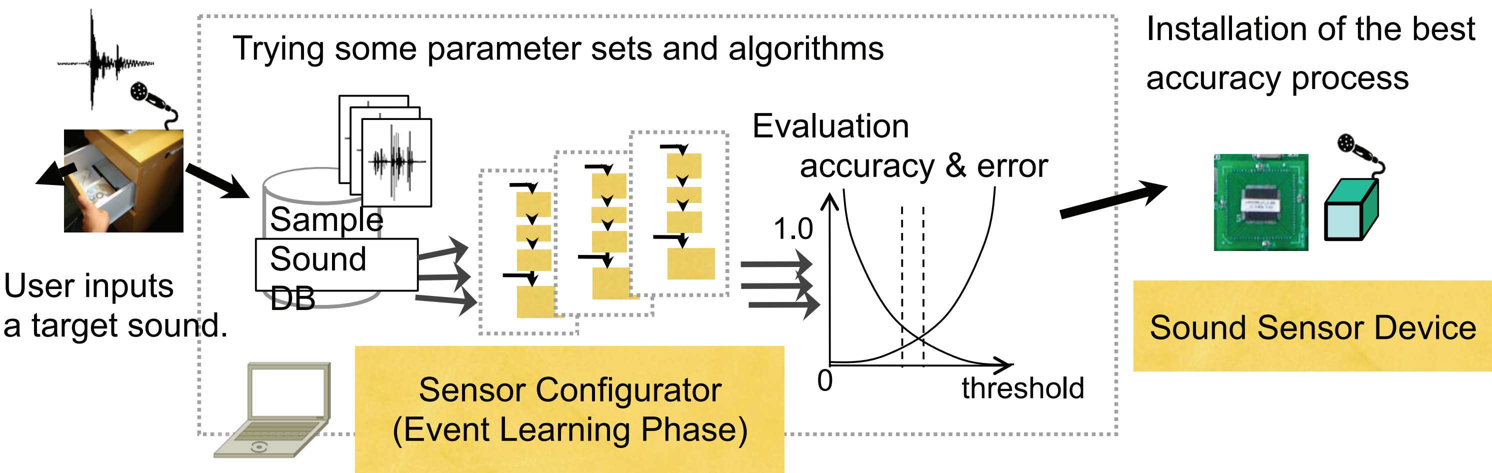
A user is only required to input a target event sound.
The sensor configurator automatically analyzes the sound, and chooses the appropriate algorithm and parameters.
Finally, it generates a recognition program for target.



(Fig.A) Example of Life-Sounds

<u>(2)Simple Device :</u>

The generated program is possible to run on low cost devices such a micro-DSP chip.
The configurator makes the programs as light as possible.

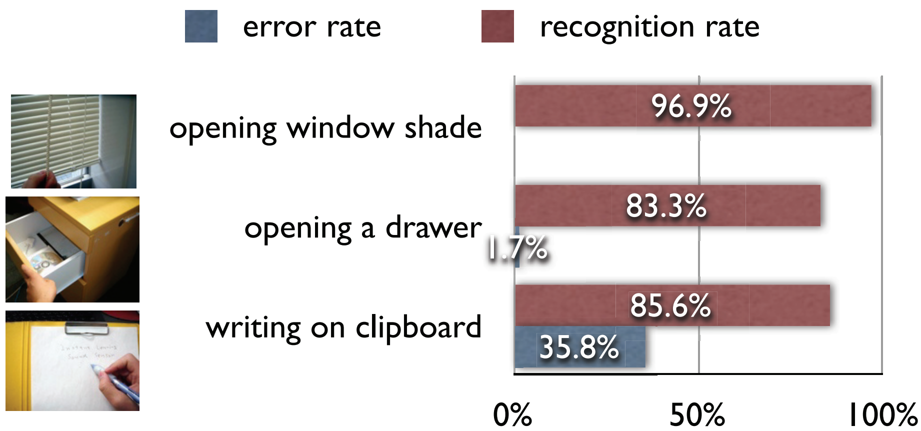


EXAMPLE OF USE

(1)Event Triggers for Appliances Control
(2)Detailed User Presence Service ,
(3)Rapid Prototyping of ubicomp system

EVALUATION

Recognition rate of automatically generated recognition process by the Instant Learning.
Almost over 83% accuracy.



(Fig.B) Evaluation of automatic generated recognition program by Instant Learnig