

Regular Paper

Improved Feature Extraction Method for Sound Recognition Applied to Automatic Sorting of Recycling Wastes

TATSUJI MUNAKA^{1,a)} FARZAD SAMIE² LARS BAUER² JÖRG HENKEL²

Received: October 28, 2019, Accepted: February 4, 2020

Abstract: As an effort to realize a sustainable society, Tokai University Takanawa Campus has used a smart garbage collection service called BigBelly™ since 2016 for recovering wastes such as cans, bottles, and plastic bottles for recycling. The objective of this demonstration experiment is to clarify how much waste has been correctly separated and collected. As a result of this demonstration about three years, it found that about 30% of the recycling wastes had not correctly sorted. To improve this situation, we propose an automatic sorting function using sound recognition. In many types of research for voice recognition, Mel Frequency Cepstral Coefficient (MFCC) has been used as an algorithm for extracting features used for machine learning Support Vector Machines (SVMs). One reason is that MFCC extracts valuable features that focus on the low dimension for the human voice. However, the sounds of recycling wastes have features of frequency components found in higher dimensions. Based on this characteristic, we propose an improved method of MFCC suitable for sounds, rather than voice recognition for identifying recycling wastes and show the results of the automatic sorting of recycling wastes.

Keywords: Internet of Things (IoT), edge computing, smart garbage collection services, sound recognition, feature extraction, Mel Frequency Cepstral Coefficient (MFCC), Support Vector Machine (SVM)

1. Introduction

Internet of Things (IoT) becomes a big keyword for achieving “everything as a service.” In Germany, IoT attracts attention as a keyword to realize a new industrial revolution called Industrie 4.0. IoT has enabled advanced monitoring and control services to improve efficiency and reduce costs in various application domains such as smart cities [1], [2]. Conventional waste management is generally costly due to the inefficient use of wastes. For instance, most discarded wastes could be recycled if there were correctly separated. Based on this concept, Tokai University has contrived a demonstration experiment for the recovery of recycling wastes using a smart garbage collection service called BigBelly™ [3] since 2016. As a result, it found that about 30% of collected recycling wastes have been mistakenly sorted [4]. This result calls for smart garbage collection services with an automatic sorting function for recycling wastes.

Due to a large number of IoT devices and their constant data collection, cloud computing is not reliable for real-time IoT applications. The main reason is the unpredictable latency caused by the multi-hop and congested network. To address this issue, edge computing, an emerging paradigm in IoT, pushes the data processing to the edge of the IoT system where the data is collected (i.e., embedded devices) [5].

In this paper, we propose an improved feature extract algorithm

for automatic sorting, using sound recognition in the context of edge computing. We have evaluated the automatic sorting with the improved feature extract algorithm. The evaluation system is equipped with a microphone sensor and an embedded processor that runs the classification algorithm for recycling wastes. The classifier is trained based on a machine learning model (SVM) with the data samples that were collected at the training time.

The rest of this paper is organized as follows. In Section 2, we provide an overview of related works. Section 3 and 4 provide background on smart garbage collection services and sound analysis. Then, we present the details of our proposed algorithm in Section 5. Evaluations and experimental results are presented in Section 6, while Section 7 concludes the paper.

2. Related Work

For achieving automatic sorting functions, there have been many similar research efforts. Gundupalli et al. [6] review some research in the area of automated sorting techniques and systems for recycling wastes. Their objective is to improve the overall efficiency of the recycling process. In this review, techniques of sorting wastes are divided into two categories: direct sorting techniques utilize material properties like magnetic susceptibility, electrical conductivity, and density for heavy media separation, indirect sorting techniques utilize sensors to detect wastes. Indirect sorting techniques are eddy current based sorting, laser-induced breakdown spectroscopy, X-ray based sorting, optical-based sorting, and spectral imaging-based sorting. In this paper, we only discuss indirect sorting because direct sorting is a

¹ Tokai University, Minato, Tokyo 108-8619, Japan

² Karlsruhe Institute of Technology, Karlsruhe, Germany

^{a)} munaka@tsc.u-tokai.ac.jp

mechanical-based solution and does not utilize information technologies.

Seri [7] proposed a two-module smart box. Each basket was able to know its full status, detect any unexpected objects inside, and notify the user when an intervention is necessary. A weight sensor was used in this box, and it made estimates using historical data and detected unexpected wastes. Naveen et al. [8] proposed a proper segregation method using IR sensors and moisture sensors for identifying wastes. Navghane et al. [9] proposed an IoT based smart garbage and waste collection box. It works with a combination of sensors that are a weight sensor and an IR sensor. Aleena et al. [10] studied an automatic waste segregator and monitoring system, which sorts wastes into three main categories that are metallic, organic, and plastic. In their studies, Near Infrared sensors and X-ray technology are used to segregate wastes. Gupta et al. [11] used some sensors such as SONOR sensors, temperature sensors, and weight sensors along with metal detectors. The main interest in these studies is which sensors were most cost-effective for separating wastes. However, they did not show how accurately they sort garbage.

Chandramohan et al. [12], Donovan et al. [13], García [14], and Bhor et al. [15] prototyped a system for predicting the type of recycling waste by using machine learning-based image classification research. Their research only uses a camera sensor without using multiple sensors. Chandramohan has achieved an identification rate of 98.3% using KNN (K-Nearest Neighbor). Rabano et al. [16] used MobileNet to generate a model that classifies wastes into the following categories: glass, paper, cardboard, plastic, metal, and other waste. A dataset of more than 2500 images was used for the training. The model used transfer learning from a model trained on the ImageNet. They tried image classification using deep convolutional networks running on a server and archived 87.2% accuracy.

Yang et al. [19] aim to automate waste sorting by applying machine learning techniques to recognize the type of waste from their images only. Two popular learning algorithms were used: deep learning with CNN (Convolutional Neural Network) and SVM. The objective is to classify recycling waste into six classes consisting of glass, paper, metal, plastic, cardboard, and trash. SVM achieved better results than CNN, and the classification accuracy is about 63%. Sakr et al. [20] also try to apply CNN and SVM for classifying wastes. In their study, wastes are separated into three main categories: plastic, paper, and metal, using a colored image of the waste. SVM achieved a high classification accuracy of 94.8%, while CNN achieved only 83%. We know that many sophisticated neural networks are provided as cloud services, however many neural networks require much memory and GPUs. Even if we could use machine learning algorithms on the cloud, it would be better to operate on limited resources. SVMs, which are well-known classical machine learning, can run on limited resources in edge computing, so many types of research have used this algorithm [17], [18], [19], [20].

Zhang et al. [21] and Zhu et al. [22] extracted five features, including MFCC, from voice and used SVM and DBN to identify emotional states of voice, each with an average accuracy of about more 95%. Their papers show that MFCC is useful as a feature

extraction algorithm. However, MFCC is mainly a method for extracting voice features. Since non-voice sound has many features in both the low-frequency dimensions and high-frequency dimensions, the extraction method needs to be improved.

There are many studies related to sound recognition. Ouchi et al. [23] recognize the activities of the elderly in homes by using sounds generated by some action by the elderly. These use an acceleration sensor and a sound sensor in a smartphone to know to monitor the health status of elderly persons. After recognizing the movement of a person with the acceleration sensor, the sound sensor starts recording. Based on the recorded sound data, it identifies the activity of the person, by way of a dishwasher, vacuum cleaner, dryer, and so on. In the processing, MFCC is used as the feature extraction algorithm, and SVM is used as the classifier. The average identification rate was over 90% in their verification. G. Laput et al. [24] performed research into recognizing the activities of people, not only in homes but also in schools with built-in devices with multiple sensors installed. Among these sensors, sound and acceleration sensors have been shown to play significant roles. They also use MFCC for extracting features of sounds of daily life, and machine learning SVM for the classifier. However, they propose a method to raise the identification rate by combining a sound and an acceleration sensor, and they do not show the degree of accuracy just by way of a sound sensor. As shown in these studies, when using a sound sensor, it is common to classify using feature values extracted by MFCC as training data and using machine learning SVM.

Inamura et al. [25] introduced a sorting robot that automatically classifies recycling wastes such as aluminum cans, steel cans, bottles, and plastic bottles using a sound sensor and an image sensor. It classifies each waste from image recognition and identifies aluminum cans and steel cans, which cannot be recognized by images, using sounds. For sound recognition, features are extracted by cepstrum features and classified by HMM (Hidden Markov Model). The identification rate was 93.3% for aluminum cans, 96.7% for steel cans, and 93.3% for plastic bottles. Their method uses the sound generated when crushing a can with a robot arm, and the sound generated when a grasped arm is loosened. Since these features have a time series, it is considered that the use of HMM is suitable. However, the method using a robot arm is intended for use in a large-scale garbage collection site and not in a trash box where located in many places. When recycling wastes such as cans, bottles, and plastic bottles are separated using free-falling sounds in a trash box, the sounds are much shorter than the sound crushed by the robot, and because there is no time series, further evaluation is indispensable.

As a previous study classified using sound sensors, Inoue et al. [26] modeled daily life sounds by HMM, and they try to make clear what kind of events are generated daily sounds, for example, in the case of sink sounds, curtain opening and closing. It also recognizes the sound of machines, for example, a vacuum cleaner. However, the sounds differ depending on the sources, and it is challenging to generate a standard model, and a large amount of data is required to create a model for correctly recognizing events.

Korucu et al. [27] classify recycling wastes into four cate-

gories: metal, plastic, glass, and cardboard, using the free-falling sound. They have set up a use case to collect what was sold on a vending machine, so cardboards are included in the classification. Among the four types, cardboards are significantly different from other materials and have different sound characteristics, and they are relatively classified. Their prototype system used a high-quality condenser microphone called MXL CR89. In their research, the normal MFCC is used for feature extraction, and SVM and HMM are used for machine learning. They use the ordinal MFCC because it relies on features from extracting a high-quality and noise-free sound signal recorded through a high-quality condenser microphone. So, they also do not mention in detail the characteristics of why MFCC, which is typically used in voice recognition, works effectively in distinguishing falling sounds. As a result of evaluation using 44 wastes in a total of 4 types, both SCM and HMM show that all wastes are correctly classified. One of the reasons that they have accurate results is to use a high-quality microphone. However, even if they use it, their study does not show the distinction between steel and aluminum cans. In the recovery of renewable resources such as bottles, cans, and plastic bottles, the distinction between steel cans and aluminum cans is critical due to subsequent processing differences. In our preliminary evaluation using the ordinal MFCC with a regular quality microphone, the recognition rate of steel cans and aluminum cans was about 80%. For this reason, we focus on the feature value of the sound and improve the feature value extraction algorithm MFCC to distinguish steel from aluminum cans with a normal-quality microphone. We clarify the signal characteristics of the falling waste sounds and propose a feature extraction algorithm for distinguishing between steel cans and aluminum cans.

In this paper, we try to classify recycling wastes such as aluminum cans, steel cans, bottles, and plastic bottles using sound recognition. We use SVM as the classifier for the features that are extracted by an improved MFCC algorithm. In the features for identifying sounds of recycling wastes, it shows the characteristics of the frequency component even in dimensions other than the low-frequency dimensions. Based on this characteristic, we could increase the identification rate by improving MFCC and clearly distinguish steel cans and aluminum cans. In this paper, we explain how to improve an MFCC algorithm and show the evaluation results.

3. Smart Garbage Collection Services

Figure 1 shows the system configuration of a smart garbage collection service called BigBelly™ [3] at Takanawa campus, Tokai University. This system consists of a smart trash collection box located in the campus, a smart garbage collection management server on the cloud and smartphones or PCs for monitors. If it exceeds the limitation of collecting trash, the box sends information to the smartphones or the PCs through the server.

BigBelly™ has some functions such as monitoring the garbage in a trash box, compressing, and notifying the amount of garbage in it.

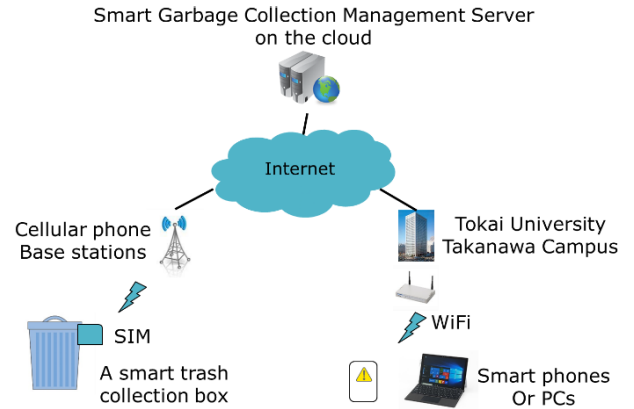


Fig. 1 System configuration of the smart garbage collection service at Takanawa Campus Tokai University.

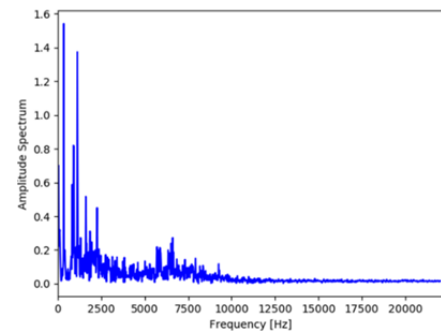


Fig. 2 Amplitude spectrum.

4. Feature Extracting Algorithm: MFCC

Voice recognition usually uses MFCC, which is a feature considering human auditory properties. As the frequency of the sound signal increases, the amplitude decreases, making it challenging to extract frequency components. To make it easier to extract high-frequency components, we perform pre-emphasis processing to emphasize frequency components. The pre-emphasis filter can be applied to a signal s using the first-order filter in the following Eq. (1).

$$y(n) = s(n) - p * s(n - 1) \quad (1)$$

With the number of samples n in $s(n)$ as a variable, the difference between the value one sample before and the value of the current sample is taken to emphasize the entire frequency. The pre-emphasis coefficient p is to be 0.97 for voice recognition. Next, the pre-emphasis-processed tone signal is digitized, and framing is performed to pick up a certain length. The value 0.8 [s] is usually used for the frame length. After framing, we perform FFT (Fast Fourier Transform) on the framed digital signal. By performing FFT, it is possible to convert the horizontal axis of the digital signal to the frequency representing the pitch of the sound over time. It can obtain the amplitude spectrum by FFT. In **Fig. 2**, the vertical axis represents amplitude, and the horizontal axis represents frequency. A Mel Filter Bank is performed on the amplitude spectrum to extract features closer to human auditory information. **Figure 3** is a Mel Filter Bank based on the acoustic properties of human, with a low-frequency dimension being exceptional and a high-frequency dimension being a coarse filter setting Mel scale. The number of filters of the Mel Filter Bank

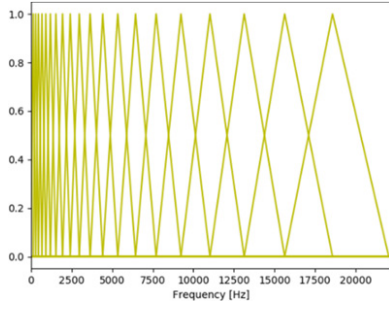


Fig. 3 Mel Filter Bank.

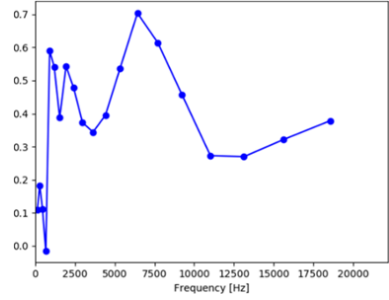


Fig. 4 Logarithmized Mel band spectrum.

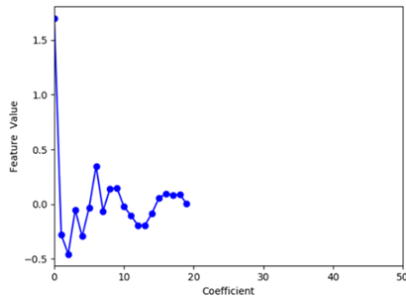


Fig. 5 Cepstrum.

is 20. A Mel Filter Bank gives a Mel band spectrum close to human auditory information in **Fig. 4**. Since the Mel Filter Bank describes the Mel band spectrum, the frequency component is found more than the discrete cosine transform (2).

$$Ci = \sqrt{\frac{2}{L}} \sum_{l=1}^L \log m(l) \cdot \cos \left\{ \left(l - \frac{1}{2} \right) \frac{i\pi}{L} \right\} \quad (2)$$

In Eq. (2), L represents the number of filter bank channels, and $\log m(l)$ represents the amplitude of the logarithmic filter bank. The information obtained from the discrete cosine transform is the cepstrum shown in **Fig. 5**. MFCC is the information extracted from low-dimensional 12 dimensions in which voice features appear from the cepstrum in **Fig. 6**. However, the free-falling sounds of recycling wastes have features of frequency components looking in higher dimensions. To extract more aspects focusing on higher dimensions, we introduce an improved MFCC.

5. Proposal of MFCC to Extract High Dimensional Frequency Components

5.1 Characteristic of the Feature for Voice and Sound

MFCC is a feature extraction algorithm used for voice recognition in general. Frequency components linked with human voice characteristics are easily recognizable in low-frequency dimensions, so MFCC extracts low-frequency dimensions. However,

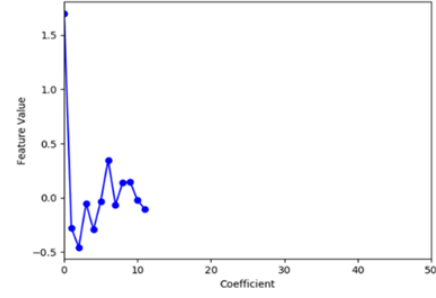


Fig. 6 MFCC.

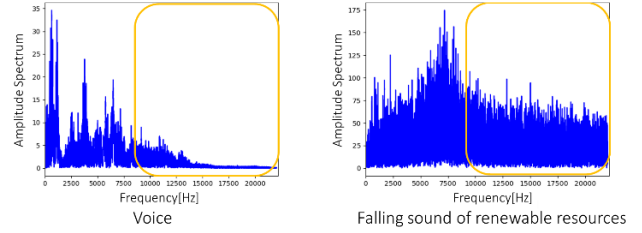


Fig. 7 Amplitude spectrum of a voice and a free-falling sound of a recycling waste.

free-falling sounds of recycling wastes have the characteristic of the frequency component, not only the low-frequency dimensions but also the high-frequency dimensions. **Figure 7** shows the difference between the frequency components of a voice and a free-falling sound of recycling waste. The frequency of the voice has a characteristic in low-frequency dimensions, and not in high-frequency dimensions. On the other hand, the free-falling sound of a recycling waste shows that the characteristic of the frequency component that can easily be seen in both the low-frequency dimensions and the high-frequency dimensions. Therefore, in the case of free-falling sounds of recycling wastes, it can be expected that the identification rate can be increased by increasing the dimension number of the features and extracting the frequency component of the high-frequency dimensions.

5.2 Improved Method of MFCC

Based on the result described in the previous section, we propose a method for improving the features extraction algorithm that aims to extract essential features by MFCC not only from the low-frequency dimensions but also from the high-frequency dimensions.

(1) First stage

Purpose:

- Expand the frame size of a sound
- Emphasis on frequency components of a sound
- Extract high-frequency components of a sound

Methods:

- Adjust a frame length to be framed
- Change the number of samples of FFT
- Change a pre-emphasis coefficient
- Increase the number of filters in a Mel Filter Bank

(2) Second stage

Purpose:

- Reduce dimensions that are not important

Method:

- Optimize the number of dimensions by PCA

The improved method consists of two stages. In the first stage, the emphasis is set so that the features of the frequency components present in the original sound signal appear, and the number of dimensions is increased so that features can be extracted up to the high-frequency dimensions. The parameters used in the first stage of processing are shown in **Table 1**. In the second stage, thirty-three dimensions were extracted from the forty-five dimensions extracted in the first stage using PCA (Principal Component Analysis). The reason is that we learned through the evaluation that SVM tends to decrease the identification rate when the number of feature dimensions is large. The reason for choosing SVM as machine learning is that Korucu, the comparison target of this study, uses SVM, to show the effect of our proposed MFCC directly, we chose the same SVM as Korucu's evaluation.

5.3 Verification of Improved MFCC

Using a free-falling sound of aluminum cans which are one type of recycling wastes, we evaluate the result of the improved MFCC by comparison with the conventional MFCC. At first, it can be confirmed that frequency components are further emphasized by increasing the sample number, and the pre-emphasis coefficient of the FFT handled in the extraction calculation of the amplitude spectrum shown in **Fig. 8**. **Figure 9** shows a comparison of Mel Filter Banks. It can be seen that the filter is precisely set also in the high-frequency dimensions. The Mel band spectrum extracted by multiplying the amplitude spectrum by the Mel Filter Bank is shown in **Fig. 10**. Since a Mel Filter Bank discretizes a Mel band spectrum, it is necessary to convert to the original frequency to extract frequency components. Therefore, in the improved MFCC, frequency components are extracted from discrete sine transformation, not discrete cosine one used in the normal MFCC.

Figure 11 shows the difference between the discrete cosine transformation and the discrete sine transformation. We can confirm that more significant features are extracted by using a discrete cosine transformation. **Figure 12** shows a comparison of the features between the ordinary MFCC and the improved MFC. We can see that the number of dimensions extracted from the sound sign and the size of the features is different. Finally, we apply PCA to reduce the extra features. PCA is a principal component analysis that extracts important dimensions from multidimensional features. In the improved MFCC, features that are extra as a result of many extracted dimension numbers are obtained correspondingly.

Figure 13 and **Fig. 14** are three-dimensional plots in which the multidimensional features are divided into three dimensions. In these three-dimensional plots, a value obtained by using training data as features is used. The red plot represents aluminum cans(a), the yellow plot represents steel cans(b), the blue plot represents bottles(c), and the green plot represents plastic bottles(d). Figure 13 shows plots of features generated by the normal MFCC, but it can be confirmed that the plots of each discrimination target are dispersed. In the improved MFCC shown in Fig. 14, the plots are not distributed as a whole, and the plots of each discrimination target are within a specific range for each type, so the useful features have been successfully extracted.

Table 1 Parameters used in the improved MFCC.

Parameters	Value
Frame length	3.5 [s]
Number of FFT samples	50,000
Pre-emphasis coefficient: p	0.99
Number of filters in a Mel Filter Bank	45
Number of dimensions to extract	45

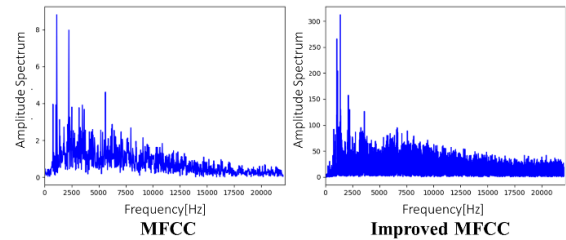


Fig. 8 Comparison of the amplitude spectrum.

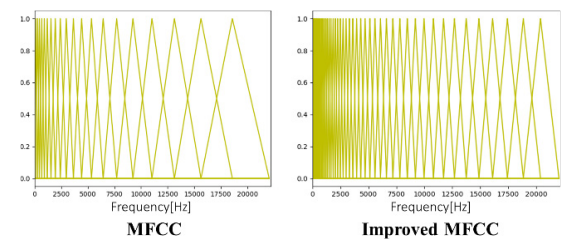


Fig. 9 Comparison of Mel Filter Banks.

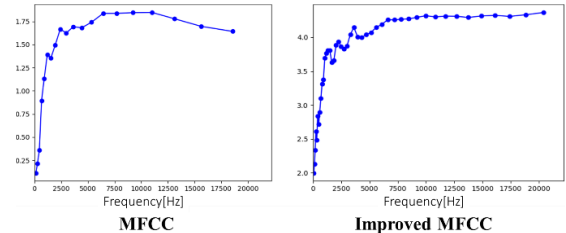


Fig. 10 Comparison of logarized Mel band spectrum.

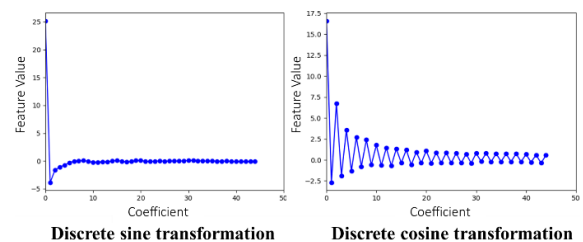


Fig. 11 Difference of discrete transformation.

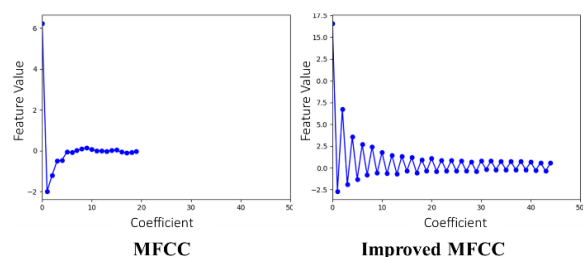


Fig. 12 Cepstrum of MFCC and improved MFCC.

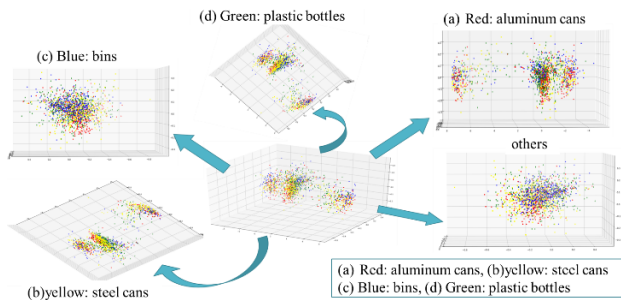


Fig. 13 Map of features by MFCC.

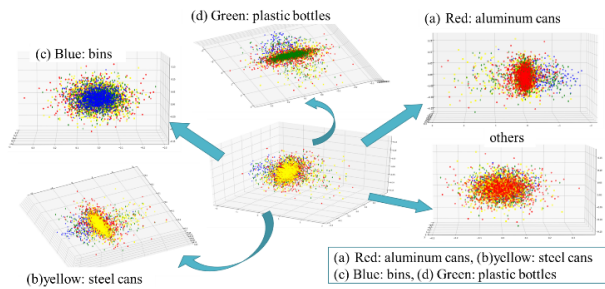


Fig. 14 Map of features by the improved MFCC.

6. Evaluation of Automatic Sorting Using the Improved MFCC Algorithm

6.1 Evaluation System

Figure 15 shows the evaluation system. This system is equipped with a condenser microphone and a Raspberry Pi that runs the classification algorithm for recycling wastes. When a recycling waste falls into the hard-plastic plate in the box shown in Fig. 15 (a), the sound sensor embedded starts the record of the sound shown in Fig. 15 (b). After recording sound data, it converts to digital data, and the features are extracted from the digital data by the MFCC algorithm. We use SVM as the classifier for the features that are extracted by MFCC as the same as Refs. [8], [11], and [26]. In the evaluation with the prototype machine, we confirm that steel cans and aluminum cans were correctly distinguished using the LED installed in the box. Because the size of the plastic box used was not large enough, those have been stored in one place (the second row in Fig. 15 (c)).

SVM classifies an object based on the training data stored in advance; in this case, the features of a sound extracted by MFCC; therefore, the result of identification strongly depends on the features. We identify a recycling waste from a free-falling sound with the training data stored in a Raspberry Pi in the box. After identifying the recycling waste, it was sorted automatically in the box. Figure 16 shows the software configuration.

In this evaluation, we prepared 100 training data for each of the four kinds of recycling wastes that are aluminum cans, steel cans, bottles, and plastic bottles. The training data were collected in the environment of a system that consists of a PC, a microphone, and a hard-plastic plate, and installed into the Raspberry Pi in the trash box shown in Table 2.

6.2 Evaluation Results and Discussion

We evaluated 100 recycling wastes such as actual bottles, steel cans, aluminum cans, and plastic bottles totaling 400 in the eval-

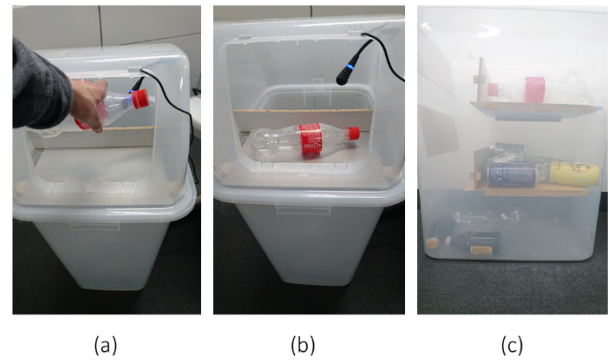


Fig. 15 Evaluation system; (a) Fall a waste into the box, (b) Recording sound, (c) sorted recycling wastes in the box.

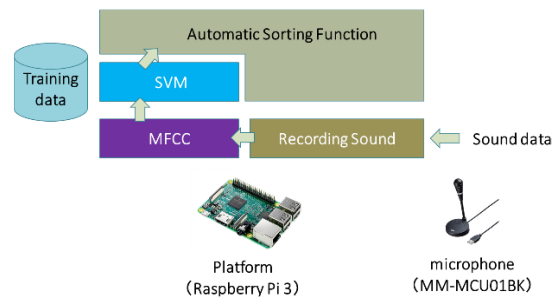


Fig. 16 Software configuration.

Table 2 Training data and its collecting environment.

Recycling wastes	The number of training data
Bottles	100
Steel cans	100
Aluminum cans	100
Plastic bottles	100
Collecting environment	
A microphone connected to a PC, a hard-plastic plate	

uation box. Parameters are adjusted via grid search and the combination of parameters with the best results that means the highest identification rate was used in the evaluation. The identification rate was 84.0% for aluminum cans, 75.0% for bottles, plastic bottles for 83.0%, and 69.0% for steel cans when using the conventional MFCC as the features in Fig. 17. The table in Fig. 17 shows that eighty-four aluminum cans are correctly identified as that; however, the remaining sixteen were not. One was identified as a bottle, two were as plastic bottles, and thirteen were as steel cans. The average identification rate was 77.7%. This result shows that it is not able to distinguish between steel and aluminum can clearly when using an inexpensive condenser microphone and a normal MFCC.

According to the confusion matrix shown in Fig. 18, all three types of aluminum cans, steel cans, and bottles to be identified were correctly identified for 100 pieces of real wastes, but in plastic bottles, erroneous identification was made twice did. Two errors were aluminum can and bottle once. So, the average accuracy of plastic bottles cans can be calculated as: $98/(1 + 1 + 98) = 98.0\%$. The average accuracy of aluminum cans, bottles, and steel cans are 100%. Therefore, the overall accuracy was 99.5%. As we can see in the verification result, we were able to recover the average identification rate to about 20% higher than that of

True label	Alcan-	84	1	2	13
	Bottle-	1	75	14	10
	Pet-	7	8	83	2
	Stcan-	15	8	8	69
		Alcan-	Bottle-	Pet-	Stcan-
		Predicted label			

Fig. 17 Confusion matrix from the SVM using MFCC.

True label	Alcan-	100	0	0	0
	Bottle-	0	100	0	0
	Pet-	1	1	98	0
	Stcan-	0	0	0	100
		Alcan-	Bottle-	Pet-	Stcan-
		Predicted label			

Fig. 18 Confusion matrix from the SVM using improved MFCC.

the conventional MFCC. It was confirmed that steel cans and aluminum could be clearly distinguished, which was not clearly mentioned in the results of Korucu et al. [27]. As our proposal only uses an inexpensive condenser microphone for the classification, the system is simpler and more cost-effective than Korucu's prototype.

Besides, this result shows that this system maintains the reusability of the training data set. In this evaluation, the training data set was able to be applied to another trash box with the same structure, even though it was created in a PC environment. We installed the same data set into another trash box, and we could see that it achieved the same identification percentage as the other trash boxes. The reason is that the materials used in trash boxes are the same, and the environment where the waste is thrown in is minimal space, making it difficult to receive noise from the outside. In addition, our improved MFCC algorithm successfully extracts features suitable for the identification of recycling wastes.

7. Conclusion and Future Work

In this research we aim to present a solution for detecting recycling wastes such as cans, bottles, and plastic bottles. We propose an improved feature extract method. The idea is based that the frequency component of free-falling sounds of recycling wastes appears in the high-frequency dimensions. We evaluated our proposed algorithm using the collected training and verification data from the trash box. As a result, the average accuracy of the model

was 99.5%, which shows a more than 20% improvement compared to the classical MFCC. It was confirmed that not only bottles and plastic bottles, but also steel cans and aluminum cans could be distinguished, this result was not mentioned in the results of Korucu which is an excellent previous study.

Rafferty et al. [28] indicate that the learning mechanisms are usually divided into two types of approaches that are generative and discriminative, depending on the modeling strategy employed. Generative approaches have the drawback of requiring large amounts of data to produce the complete set of probabilistic representations to provide proper functionality. Discriminative approaches can produce results using a less exhaustive dataset compared to generative approaches. In our evaluation system, the classifier SVM runs on the embedded processor of the trash box at runtime for the inference, as suggested by the edge computing paradigm. Since IoT devices at the edge are usually resource-constrained (i.e., limited energy budget, memory, and computation capability), we chose SVM as a machine learning and evaluated the improved MFCC with it. However, in future work, we will evaluate the effect of other classification models such as HMM, or naive Bays belonging to generative approaches. We also need to evaluate other future extract algorithms, such as pitch, formant, short-term zero-crossing rate, short-term energy [22], linear predictive coding coefficients (LPCC) and wavelet transform (WT). From the viewpoint of applications, we apply this algorithm to the activity recognition of people in houses to monitor the safety life of the elderly.

References

- [1] Samie, F., Bauer, L. and Henkel, J.: IoT technologies for embedded computing: A survey, *International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS)* (2016).
- [2] Alturki, B. and Reiff-Marganiec, S.: Towards an off-the-cloud IoT data processing architecture via a smart car parking example, *Proc. 2nd International Conference on Internet of things, Data and Cloud Computing (ICC '17)*, pp.37–42, ACM (2017).
- [3] BigBelly: available from (<http://bigbelly.com>).
- [4] Munaka, T. and Oshima, K.: On the results of demonstration experiments on renewable resources using smart trash boxes (in Japanese), *ISPJ Symposium, DiCoMo2017* (July 2017).
- [5] Yu, W., Liang, F., He, X., Hatcher, W.G., Lu, C., Lin, J. and Yang, X.: A Survey on the Edge Computing for the Internet of Things, *IEEE Access*, Vol.6, pp.6900–6919 (2018).
- [6] Gundupalli, S.P., Hait, S. and Thakur, A.: A review on automated sorting of source-separated municipal solid waste for recycling, *Waste Management*, Vol.60, pp.56–74, ELSEVIER (2017).
- [7] Seri, F.: Implementation of a smart bin for waste monitoring and recognition, Master thesis, Politecnico di Milano (2018).
- [8] Naveen, B., Kavya, G.K., Kruthika, S.N., Ranjihta, K.N. and Sahana, C.N.: An Economic Automatic Waste Segregator using ARDUINO, *International Journal of Advanced Engineering and Research Development*, Vol.5, No.5 (2018).
- [9] Navghane, S.S., Killedar, M.S. and Rohokale, V.M.: *IoT Based Smart Garbage and Waste Collection Bin, IJARRECE*, Vol.5, No.5 (2016).
- [10] Aleena, V.J., Balakrishnan, K., Rosmi, T.B., Krishna, S., Sreejith, S. and Subha, T.D.: Automatic Waste Segregator and Monitoring System, *Journal of Microcontroller Engineering and Applications*, Vol.3, No.2 (2016).
- [11] Gupta, P.K., Shree, V., Hiremath, L. and Rajendran, S.: The Use of Modern Technology in Smart Waste Management and Recycling: Artificial Intelligence and Machine Learning, *Recent Advances in Computational Intelligence*, pp.173–188, Springer (2019).
- [12] Rafeeq, M., Atequrahman, Alam, S. and Mikdad: Automation of plastic metal and glass waste materials segregation using Arduino in scrap industry, *Communication and Electronics Systems (ICCES) International Conference*, pp.1–5 (2016).
- [13] Donovan, J.: Auto-trash sorts garbage automatically at the

TechCrunch Disrupt hackathon (online), available from <http://techcrunch.com/2016/09/13/auto-trash-sorts-garbage-automatically-at-the-techcrunch-disrupt-hackathon/>.

- [14] García, A.T., Aragón, O.R., Gandara, O.L., García, F.S. and González-Jiménez, L.E.: Intelligent Waste Separator, *Computación y Sistemas* (2015).
- [15] Bhor, V., Morajkar, P., Pandya, D. and Deshpande, A.: Smart Garbage Management System, *International Journal of Engineering and Technical Research*, Vol.4, No.3 (2015).
- [16] Rabano, S.L., Cabatuan, M.K., Sybingco, E., Dadios, E.P. and Calilung, E.J.: Common Garbage Classification Using MobileNet, *IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, pp.1–4 (2018).
- [17] Mahdavinchad, M.S., Rezvan, M., Barekatain, M., Adibi, P., Barnaghi, P. and Sheth, A.P.: Machine learning for the internet of things data analysis: A survey, *Digital Communications and Networks*, Vol.4, pp.1161–175 (2018).
- [18] Samie, F., Bauer, L. and Henkel, J.: From Cloud Down to Things: An Overview of Machine Learning in Internet of Things, *IEEE Internet of Things Journal*, Vol.6, No.3 (2019).
- [19] Yang, M. and Thung, G.: Machine Learning, Classification of Trash for Recyclability Status, *CS229* (2016).
- [20] Sakr, G.E., Mokbel, M., Darwich, A., Khneisser, M.N. and Hadi, A.: Comparing deep learning and support vector machines for autonomous waste sorting, *IEEE IMCET2016*, pp.207–212 (2016).
- [21] Zhang, W., Zhao, D., Chai, Z., Yang, L.T., Liu, X., Gong, F. and Yang, S.: Deep learning and SVM-based emotion recognition from Chinese speech for smart affective services, *Software: Practice and Experience*, Vol.47, No.8, pp.1127–1138, Wiley Online Library (wileyonlinelibrary.com), DOI: 10.1002/spe.2487 (2017).
- [22] Zhu, L., Chen, L., Zhao, D., Zhou, J. and Zhang, W.: Emotion Recognition from Chinese Speech for Smart Affective Services Using a Combination of SVM and DBN, *Sensors*, Vol.17, No.7, p.1694 (2017).
- [23] Ouchi, K. and Doi, M.: Living Activity Recognition Technology Using Sensors in Smartphone, *Toshiba Review*, Vol.68, No.6 (2013).
- [24] Laput, G., Zhang, Y. and Harrison, C.: Synthetic Sensors: Towards General-purpose Sensing, *CHI2017* (May 2017).
- [25] Inamura, T., Kojo, N., Hatao, N., Tokutus, S., Fujimoto, J., Sonoda, T., Okada, K. and Inaba, M.: Realization of Trash Separation of Bottles and Cans for Humanoids using Eyes, Hand and Ears, *Journal of the Robotics Society of Japan*, Vol.25, No.6, pp.813–821 (2007).
- [26] Inoue, Y., Wakahara, J., Yagiu, T. and Takahama, N.: Consideration of life sound analysis by HMS, *The Society of Instrument and Control Engineers, System and Information Symposium*, pp.313–316 (2004).
- [27] Korucu, M.K., Kaplan, Ö., Büyük, O. and Güllü, M.K.: An investigation of the usability of sound recognition for source separation of packaging wastes in reverse vending machines, *Waste Management*, Vol.56, pp.46–52, ELSEVIER (2016).
- [28] Rafferty, J., Nugent, C.D., Liu, J. and Chen, L.: From Activity Recognition to Intention Recognition for Assisted Living Within Smart Homes, *IEEE Trans. Human-machine Systems*, Vol.47, No.3 (June 2017).



Tatsuji Munaka is a Professor of the Graduate School of Information and Telecommunication Engineering, Tokai University. He joined Mitsubishi Electric Corporation in 1986 and had worked until 2015. During the period, he had managed some new research projects such as information technology, information security, and wireless communication. He received the D.E. degree in science and engineering from Shizuoka University, Japan, in 2003. One of his research is the realization of health care services using IoT technology. He is a senior member of the IEEE Communication Society, a senior member of the Information Processing Society of Japan, and a member of the Institute of Electronics, Information and Communication Engineers.



Technology, Iran in 2008 and 2010, respectively. His main research work is focused on Internet-of-Things (IoT) with a special interest in Edge Computing and healthcare domain.



Lars Bauer received his M.Sc. and Ph.D. degrees in computer science from the University of Karlsruhe, Germany, in 2004 and 2009. He is currently a research group leader and lecturer at the Chair for Embedded Systems (CES) at the Karlsruhe Institute of Technology (KIT). Dr. Bauer received two dissertation awards (EDAA and FZI), two best paper awards (AHS'11 and DATE'08) and several nominations. His research interests include architectures and management for adaptive multi-/many-core systems.



Jörg Henkel (M'95–SM'01–F#15) is the Chair Professor for Embedded Systems at Karlsruhe Institute of Technology. He received the Diploma and the Ph.D. (summa cum laude) degrees in electrical engineering from the Technical University of Braunschweig, Braunschweig, Germany. His research work is focused on co-design for embedded hardware/software systems with respect to power, thermal and reliability aspects. He served as the Editor-in-Chief for the ACM Transactions on Embedded Computing Systems and is currently the Editor-in-Chief of the IEEE Design&Test Magazine. He has led several conferences as a General Chair incl. IC-CAD, ESWeek and serves as a Steering Committee chair/member for leading conferences and journals for embedded and cyber-physical systems. Prof. Henkel coordinates the DFG program SPP 1500 “Dependable Embedded Systems” and is a site coordinator of the DFG TR89 collaborative research center on “Invasive Computing”. He is a Fellow of the IEEE.