Proceedings of the International MultiConference of Engineers and Computer Scientists 2010 Vol I,
IMECS 2010, March 17 - 19, 2010, Hong Kong

Rule-based Classification for Audio Data Based on Closed Itemset Mining

Takahiro Tada, Tomomasa Nagashima and Yoshifumi Okada*

Abstract— **Automatic audio classification is a major topic in the fields of pattern recognition and data-mining. This paper describes a new rule-based classification method (cREAD: classification Rule Extraction for Audio Data) for multi-class audio data. Typically, rule-based classification requires considerable computation cost to find rules from large datasets because of the combinatorial search problem. To achieve efficient and fast extraction of classification rules, we take advantage of a closed itemset mining algorithm that can exhaustively extract non-redundant and condensed patterns from a transaction database within a reasonable time. The notable feature of this method is that the search space of classification rules can be dramatically reduced by searching for only closed itemsets constrained by "*class label item.*" In this paper, we show that our method is superior to the other salient methods in its classification accuracy of a real audio dataset with the "underdetermined problem" that the number of samples is overwhelmingly fewer than that of attributes.**

Index Terms—**audio data, rule-based classification, closed itemset, LCM.**

I. Introduction

Automatic audio classification is a major topic in the fields of pattern recognition and data-mining, and it has been used in various applications to date, such as musical instrument sound identification [1], music retrieval [2] or personal recognition [3]. The most popular approach in audio classification is a statistical machine learning technique such as the support vector machine (SVM) [4]. This approach constructs a discriminant mathematical model to discriminate among different classes using feature quantities (*e.g.*, power spectrum data) derived from audio sample data [5]. Although the statistical machine learning approach exhibits high classification accuracy in many cases, it is not easy to interpret the resulting models since the learning process is a "black box."

On the other hand, a rule-based approach identifies distinctive patterns (classification rules) between different classes. The advantage of a rule-based classifier is the explicitness and comprehensibility of the resulting model in addition to relatively high classification accuracies [6]. Thus, we can not only infer important features to characterize each class, but also use them as useful knowledge. However, the rule-based approach has some drawbacks. Typically, extraction of classification rules requires heavy computation because of the combinatorial search in large-scale data. Thus the number of rules generated will be very large, so further processing is necessary to select only discriminative rules between classes. Therefore, existing rule-based classification methods have focused on finding rules as possibly local solutions by stochastic or heuristic approaches [7, 8].

In recent years, much attention has been paid to discovering co-occurrence patterns called *closed itemsets* from transaction databases (see Section II) in the field of data-mining. Frequent pattern mining such as the Apriori algorithm [9] searches for all frequent patterns (even if these are included in other patterns), while closed itemset mining can extract maximal and condensed patterns by excluding redundant patterns that have inclusive relations.

Here we propose a new rule-based classification method for audio data based on an efficient closed itemset mining algorithm called Linear time Closed itemset Miner (LCM) [10, 11]. In this method (cREAD: classification Rule Extraction for Audio Data), audio power spectra of multi-classes are transformed into a transaction database that includes a "c*lass label item*" in each transaction. Classification rules are extracted by the following two processes: 1) an exhaustive search of closed itemsets having the class label item by the LCM algorithm, and 2) the greedy rule-selection approach. The notable feature of this method is to drastically reduce the search space by omitting a search for unnecessary closed itemsets that have no class label item. In this paper, we show the results of a performance comparison to two salient methods and the effect of the pruning operation using a real audio sample dataset and some well-known benchmark datasets from the UCI machine learning repository [12].

This paper is organized as follows. Section 2 defines a closed itemset and briefly describes the LCM algorithm. Section 3 describes the computational procedure of cREAD. Section 4 explains the datasets used in this study and the experimental method to evaluate the performance of cREAD.

Manuscript received December 28, 2009. This work was supported in part by Grant-in-Aid for Young Scientists (B) No.21700233 from MEXT Japan and A Research for Promoting Technological Seeds from JSTA.

T. Tada is with Department of Information and Electronic Engineering, Muroran Institute of Technology, 27-1, Mizumoto-cho, Muroran 050-8585, Japan (e-mail: tada@cbrl.csse.muroran-it.ac.jp)

T. Nagashima is with College of Information and Systems, Muroran Institute of Technology, 27-1, Mizumoto-cho, Muroran 050-8585, Japan (e-mail: nagasima@csse.muroran.-it.ac.jp).

Y. Okada is with College of Information and Systems, Muroran Institute of Technology, 27-1, Mizumoto-cho, Muroran 050-8585, Japan (corresponding author to provide phone: +81-143-5498; fax: +81-143-5498; e-mail: okada@csse.muroran.-it.ac.jp).

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

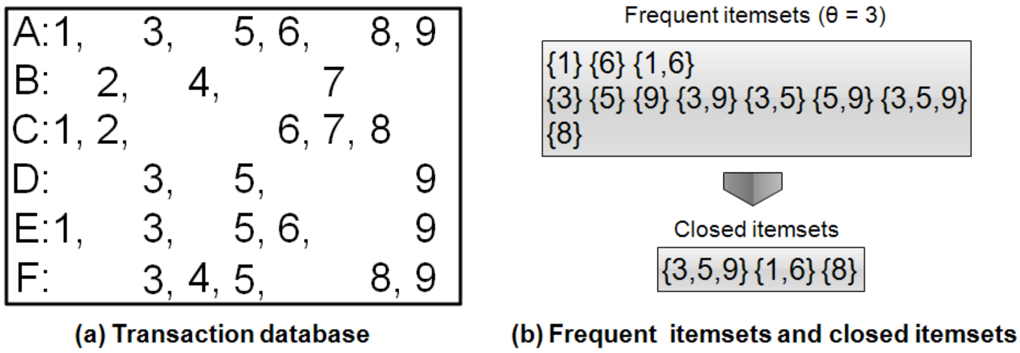

Figure 1: Enumeration of closed itemsets in transaction database

Section 5 shows the results and discussion of the performance. Finally, Section 6 summarizes our conclusions and suggests future work.

II. Closed itemset mining and the LCM algorithm

A. *Transaction Database and Closed Itemset*

A transaction database is a set of transactions with a set of items, as shown in Fig. 1a. Each transaction with an item *I* is called an *occurrence* of *I* and the set of occurrence of the item *I* is termed an *occurrence set*. For a given constant $\theta \geq 0$, a frequent itemset is a set of items that is included in at least θ transactions. A closed itemset is defined as the maximal itemset (with respect to inclusion) among the set of itemsets that have the same frequency and appear in the same transactions. As shown in Fig. 1b, we see that the closed itemsets are non-redundant and condensed patterns in contrast to the frequent patterns.

B. *LCM Algorithm and Its Extension to Rule-Based Classifier*

Typically, closed itemset mining requires considerable computation time because of its combinatorial search. The LCM proposed by Uno *et al.* is the fastest algorithm that can enumerate frequent closed itemsets in linear time of the size of database [10, 11]. This algorithm uses the *prefix-preserving closure extension*, which is an extension of a closed itemset to another closed itemset. Since this extension generates a new frequent closed itemset from the previously obtained itemset without duplications by the depth-first search technique, it enables us to completely prune the unnecessary non-closed frequent itemsets.

Our basic idea is to perform exhaustive classification rule enumeration based on LCM by adding items representing class labels (*class label items*) into each transaction. In this way, we can omit the search for unnecessary closed itemsets without the class label items; hence, we can drastically reduce the computational time as well as the search space in enumerating candidates of classification rules.

III. Method

For simplicity, we describe the method of classification rule extraction in a two-class audio sample. Fig. 2 illustrates the procedure of cREAD.

A. *Preprocessing*

Suppose that we have Fourier transformed power spectra with *p* frequency points for each audio sample in two classes (n_1 samples in class 1 and n_2 samples in class 2). First, all powers in each frequency f_{ij} are normalized to have a mean of 0 and a variance of 1 over the two classes, where $i = 1,2,\ldots, n_1 + n_2$ and $j = 1,2,\ldots, p$. Next, these normalized powers (nf_{ij}) are quantized to *q* levels by uniformly dividing the range of the maximum and minimum values. In this way, each power spectrum is transformed into a *p*-dimensional vector having one of the quantized values in each element.

B. *Generation of Transaction Database*

The transaction database is created by reference to the itemization table, as depicted in Fig. 2b. In the itemization table, each item corresponds to a uniquely given number for a quantized value in each dimension (*i.e.* each frequency point) of the quantized vector. Note that the items “1” and “2” in Fig. 2b are used as the class label items. According to the itemization table, we transform each quantized vector into an itemized vector having a class label item. Finally, these itemized vectors are summarized into the transaction database in which each itemized vector corresponds to each transaction.

C. *Rule Enumeration*

Subsequently, closed itemsets are exhaustively enumerated using LCM from the transaction database generated in the previous section. Fig. 2c shows the closed itemsets (and their rule representations) extracted from the transaction database in Fig. 2b. The aim here is to extract rules such as “*A and B → class1*” whose antecedent and consequent include itemsets and a class label item, respectively. This process means to find only closed itemsets with class label items by eliminating unnecessary closed itemsets without any class label items. Note that we can easily transform each item in the antecedent of the rule into a quantized value (or a normalized power) in a corresponding frequency point by reference to the itemization table.

D. *Rule Selection*

We select final discriminative classification rules from the enumerated rules by the following greedy approach:
1) Select the most frequent rule (*i.e.* the most frequent closed

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

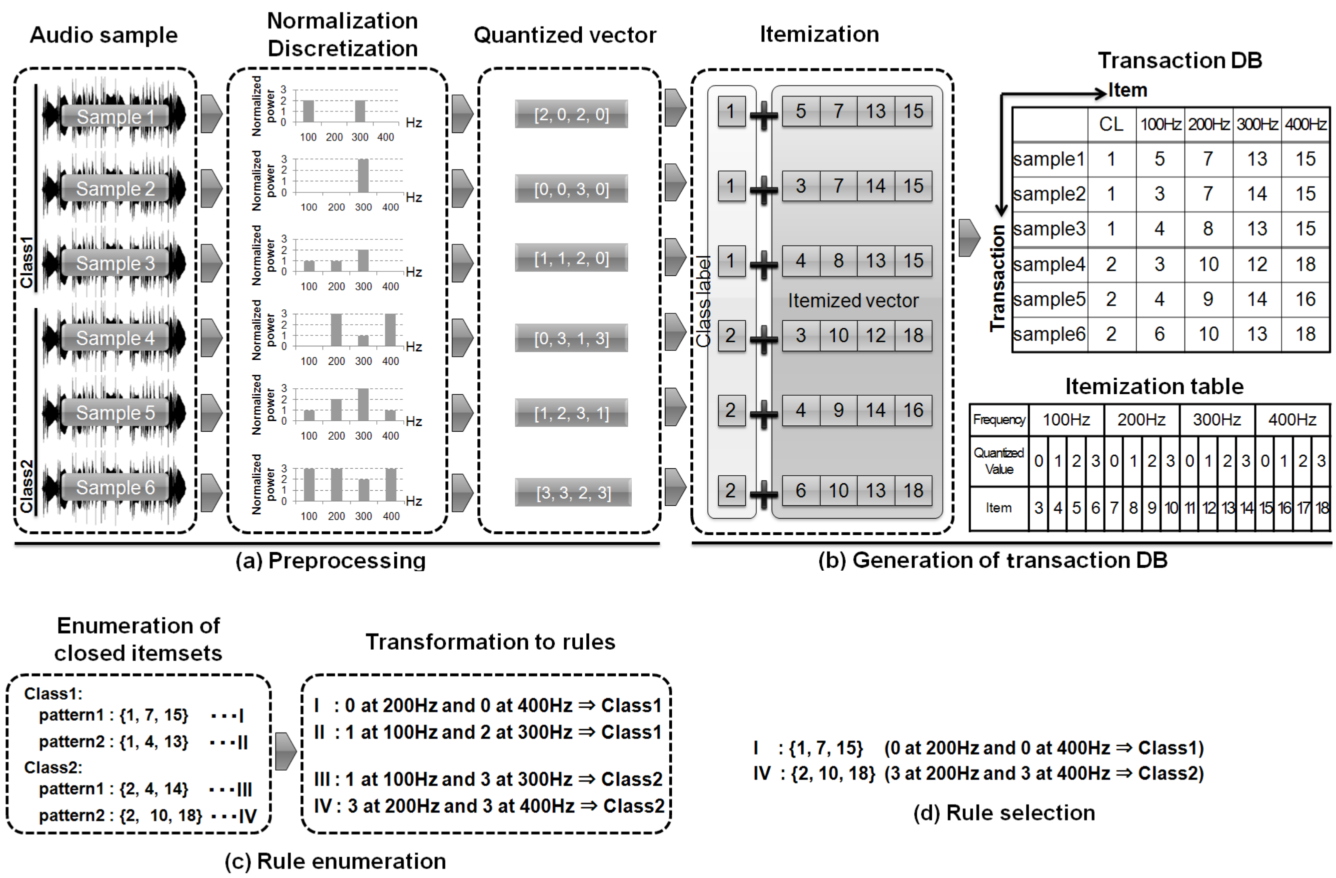

Figure 2: Procedure of cREAD

itemset) in the transaction database. If there are multiple rules with the same occurrence frequency, then we perform the following steps with respect to those rules:

1-1) Pick up a rule and calculate an inner product between the quantized values of the rule's antecedent and those (in the same frequency points as the rule) of each vector of other classes.

1-2) Calculate the mean value of the inner products obtained in 1-1).

1-3) Repeat 1-1) and 1-2) for all rules in step 1).

1-4) Select a rule with the minimal absolute value in the mean values of the inner products.

2) Exit if the union of the occurrence sets of the rules selected in the step 1) covers all transactions, repeat 1) otherwise.

E. *Classification*

Here, we explain the classification method for unknown (or test) samples using the finally-obtained rules from the above procedure. First, each unknown sample is transformed into a quantized vector with p dimensions in the same manner as the above *Preprocessing*. Second, the quantized values of an unknown sample are compared with those of the classification rules obtained from each class by the manner of k-nearest neighbor (k-NN) approach. In this process, we calculate the Euclidian distances between quantized values of the classification rules and those of each unknown sample. Lastly, we extract the top-k classification rules in ascending order of their Euclidian distances and assign each unknown sample into a class by the majority vote of its neighbors.

IV. EXPERIMENTS

A. *Datasets*

As a real audio sample dataset, we use the baby cry dataset produced by Wang *et al.* [13] to evaluate the usefulness of our method. This dataset consists of two classes: a class of cries belonging to the genetic disease called Ankyloglossia with Deviation of the Epiglottis and Larynx (ADEL) [14] and the other class of normal cries. The ADEL class and normal class include 17 and 22 audio samples (waveform data), respectively. These samples are transformed into the frequency domain by Fourier analysis in the frequency range 0–22,050 Hz and a resolution of 43 Hz (513 frequency points).

In addition, we use three well-known benchmark datasets (iris, wine, and segmentation) from the UCI machine learning repository. (see [12] for detailed information about these datasets). Since these datasets are not audio data, we perform only the processes from B to E in Section III after normalization and discretization (quantization) of each sample of the datasets.

B. *Evaluation of Classification Accuracy*

The evaluation is conducted by Leave-One-Out Cross Validation (LOOCV) [15]. In LOOCV, first, we extract one sample (*i.e.*, a quantized vector with p dimensions) as a test sample from the dataset and generate classification rules using the remaining samples. Second, the test sample is assigned to

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table1: Classification results of ADEL dataset

Method	Classification accuracy(%)	Computation time(sec.)
C4.5	38.5	17
cREAD	69.2	178
L-SVM	66.7	30
R-SVM	61.5	17

Table2: Classification results of UCI benchmark datasets

Data set	Attributes	Samples	Classes	C4.5(%)	cREAD(%)
iris	4	150	4	95.3	91.3
wine	13	178	13	93.2	87.6
segmentation	19	210	7	88.6	73.8

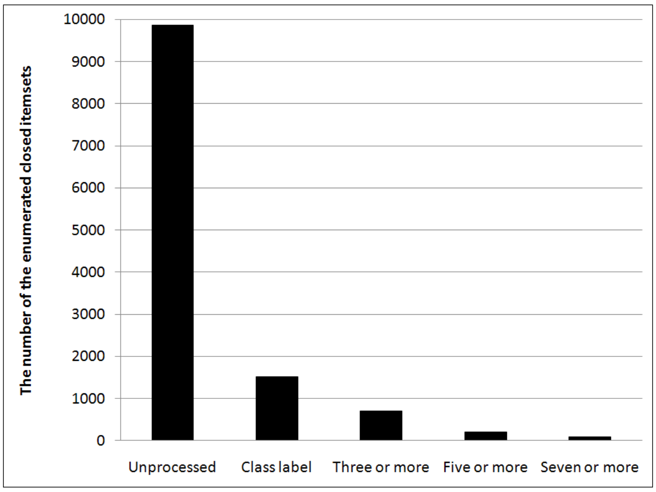

Figure3: Solution space of ADEL dataset

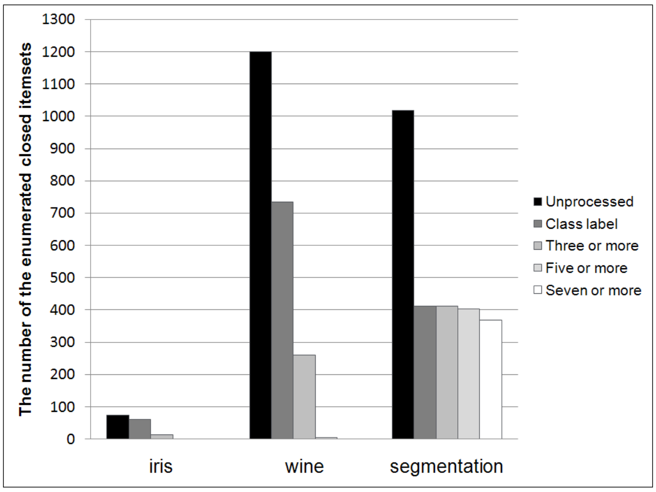

Figure 4: Solution space of UCI benchmark datasets

a class by our classification method. We repeat these processes for all samples and calculate the rate of correctly classified samples. In the evaluations, the number of quantization bins and the minimal frequency of closed itemset mining are set to 21 and 4, respectively.

The classification accuracy is compared to those of the two salient classifiers, C4.5 [16] and support vector machine (SVM) [2]. C4.5 is a statistical classifier based on a decision tree, and SVM is one of the latest and most successful kernel-based machine learning methods. We download the free software available in [17] and [18, 19], respectively, and employ them with their default parameters.

V. Results and Discussion

A. *ADEL Dataset*

Table 1 shows the results of LOOCV on cREAD, C4.5, and SVM. For SVM, we use two kernel functions: the linear kernel (L-SVM) and the radial basis function kernel (R-SVM). As a result, cREAD exhibits the best accuracy (69.2 %) in the LOOCV test. Notably, this score is far superior to that of C4.5, which is one of the rule-based classifiers. This result suggests that cREAD is an effective method for datasets with the "underdetermined problem" that the number of samples is overwhelmingly fewer than that of attributes. In this paper, the final classification rules are selected by a simple greedy approach. Toward further accuracy improvement, we will incorporate a rule optimization process for selecting more sophisticated rules.

cREAD requires considerable computation time compared to the other methods. This is majorly because the current version scans all closed itemsets including ones without the class label items, since the pruning process has not been implemented at present. However, it is possible to estimate the size of the search space before and after the pruning process by checking if each closed itemset includes a class label item. Fig. 3 is a graph representing the size of the search space before and after the pruning of unnecessary closed itemsets. In this figure, "unpruned" indicates the number of all closed itemsets enumerated without pruning, and "class label" is the number of closed itemsets having the class label items. "*X* or more" indicates the number of closed itemsets having at least *X* items other than a class label item. As shown in this figure, the search space is drastically reduced by the pruning process. For example, 9,867 closed itemsets before pruning are reduced to 1,515 in the case of "class label." Thus, it is expected that the pruning process can strikingly reduce the computation time as well as the search space of closed itemsets.

B. *UCI Benchmark Datasets*

Table 2 shows the details of the three datasets (iris, wine and segmentation) together with the classification accuracies. The datasets are all composed of three or more classes. Thus, we show only the comparative results of cREAD and C.4.5 because SVM is basically a binary classifier. As a result, C4.5 presents better accuracies than cREAD in all the datasets in contrast to the considerably low accuracy (38%) in the ADEL dataset. In these datasets, we see that the number of samples is significantly larger than that of attributes. It is important to provide a better score in such well-conditioned datasets as well as in the unbalanced datasets such as the ADEL dataset. To enhance the robustness of the method, we need further technical improvements such as parameter tuning and rule optimization.

Fig. 4 shows the sizes of search space for these datasets. We can see the drastic reduction of the search space by the pruning process in the same way as in the ADEL dataset.

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

VI. Conclusion

In this paper, we proposed a new rule-based classification method (cREAD) for audio data based on an ultra fast and exhaustive closed itemset mining algorithm. cREAD can be applied to different types of datasets other than audio data if those data are appropriately normalized and discretized in advance. In this study, we applied this method to real audio sample data and the UCI benchmark datasets and compared the classification accuracies with the salient classifiers, C4.5 and SVM. As a result, cREAD presented the best accuracy in the real audio sample dataset among the three methods. In contrast, the accuracies on the UCI benchmark datasets were lower than those of C4.5. In addition, it was shown that the search space is drastically reduced by pruning unnecessary closed itemsets that have no class label items.

From these result, we conclude that cREAD is a potentially robust and effective method for unbalanced datasets with the “underdetermined problem,” even though further technical improvements such as parameter search and rule optimization still remain. In the future work, we will not only tackle these technical issues but also verify the performance using various audio sample datasets.

Acknowledgment

The authors would like to express their thanks to Dr. Mukai for his advice and cooperation. This work was supported in part by Grant-in-Aid for Young Scientists (B) No.21700233 from MEXT Japan and A Research for Promoting Technological Seeds from JSTA.

References

[1] L. Chen, S. Gündüz and M. T. Özsu, “Mixed Type Audio Classification with Support Vector Machine,” *International Congress on Mathematical Education 2006*, 2006, pp.781-784.

[2] R. Dannenberg, W. Birmingham, B. Pardo, C. Meek, N. Hu and G. Tzanetakis, “A comparative evaluation of search techniques for query by humming using the MUSART testbed,” *j. am. Soc. Inf. Sci. Technol.*, 58, 2007, pp. 687-701.

[3] I. Pollack, J. M. Pickett and W. H. Sumby, “On the Identification of Speakers by Voice,” *J. Acoust. Soc. Am*, 26(3), 1954. pp. 403-412.

[4] B. E. Boser, I. M. Guyon and V. N. Vapnik, “A Training Algorithm for Optimal Margin Classifiers,” *5th COLT*, 1992, pp. 144-152.

[5] B. H. Juang and L. R. Rabiner, “Hidden Markov Models for Speech Recognition,” *Technometrics*, Vol. 33, No. 3, 1991, pp. 251-272.

[6] R. L. Lawrence and A. Wright, “Rule-Based Classification Systems Using Classification and Regression Tree (CART) Analysis,” *American Society for Photogrammetry & Remote Sensing*, Vol. 67, No. 10, 2001, pp. 1137-1142.

[7] H. Ishibuchi, I. Kuwajima and Y. Nojima, “Multiobjective Association Rule Mining.,” *Proceedings of PPSN Workshop on Multiobjective Problem Solving from Nature*, 2006

[8] X. L. Li, Z. L. Du, T. Wang and D. M. Yu, “Audio Feature Selection Based on Rough Set,” *International Journal of Information Technology*, Vol.11, No.6, 2005, pp.117-123

[9] R. Agrawal and R. Srikant, “Fast algorithm for mining association rules in large databases,” *Proceedings of Very Large Data Bases 1994*, 1994, pp. 487-499.

[10] T. Uno, T. Asai, Y. Uchida and H. Arimura, “An Efficient Algorithm for Enumerating Closed Patterns in Transaction Databases,” *Lecture Notes in Artificial Intelligence*, 3245, 2004, pp. 133-141.

[11] T. Uno, LCM (Linear time Closed itemset Minor). Available: http://research.nii.ac.jp/~uno/code/LCM.htm

[12] D. J. Newman, “UCI Machine Learning Repository,” Available: http://archive.ics.uci.edu/ml/

[13] X. Wang, T. Nagashima, K. Fukuta, Y. Okada, M. Sawai, H. Tanaka and T. Uozumi, “Statistical method for classifying cries of baby based on pattern recognition of power spectrum,” to be published.

[14] S. Mukai and K. Asaoka, “Congenital ankyloglossia with deviation of the epiglottis and larynx: symptoms and respiratory function in adults,” *Annals of Otolaryxgol*, Vol. 100, 1991, pp. 1-20.

[15] P. A. Lachenbruch and M. Goldstein, “Discriminant Analysis,” *BIOMETRICS 35*, 1979, pp. 69-85.

[16] J. R. Quinlan, “C4.5 Programs for Machine Learning,” *Morgan Kaufmann*, 1993.

[17] P. Winston, “C4.5Tutorial,” Available: http://www2.cs.uregina.ca/~hamilton/courses/831/notes/ml/dtrees/c4.5/tutorial.html

[18] J. Thorsten, “SVM-Light Support Vector Machine,” Available: http://svmlight.joachims.org/

[19] C. C. Chang and C. J. Lin, “LIBSVM -- A Library for Support Vector Machines,” Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm/

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)