

Unmixed Spectrum Clustering for Template Composition in Lung Sound Classification

Tomonari Masada, Senya Kiyasu, and Sueharu Miyahara

Nagasaki University, 1-14 Bunkyo-machi, Nagasaki 852-8521, Japan
{masada,kiyasu,miyahara}@cis.nagasaki-u.ac.jp

Abstract. In this paper, we propose a method for composing templates of lung sound classification. First, we obtain a sequence of power spectra by FFT for each given lung sound and compute a small number of component spectra by ICA for each of the overlapping sets of tens of consecutive power spectra. Second, we put component spectra obtained from various lung sounds into a single set and conduct clustering a large number of times. When component spectra belong to the same cluster in all clustering results, these spectra show robust similarity. Therefore, we can use such spectra to compose a template of lung sound classification.

1 Introduction

Real-world problems provide various spectral data. We often regard a given set of such spectra as a mixture of elementary spectra, which can be computed by some *spectral unmixing* method. In this paper, we apply independent component analysis (ICA) [5] to lung sound power spectra and approximate a set of tens of power spectra with a mixture of a small number of spectra which we call *component spectra*. Further, we filter out less effective component spectra by conducting clustering a large number of times on a set of component spectra obtained from various lung sounds. Component spectra belonging to the same cluster over all clustering results can be regarded as robustly similar with each other and may show an outstanding feature of a specific type of lung sounds. Therefore, we regard only such component spectra as efficient ones and use them to make templates for lung sound classification.

The rest of the paper is organized as follows. Section 2 provide the results of previous researches. Section 3 describes details of our method. In Section 4, we show the settings and the results of our evaluation experiment. Section 5 concludes this paper along with our future works.

2 Previous Work

We are experiencing a new wave of lung sound analysis due to the import of various machine learning techniques in recent years [1]. However, there seem no researches applying ICA not to lung sounds but to their power spectra obtained by FFT. Güler et al. [3] reduce the feature space dimension to two and provide

a visual classification of lung sounds drawn in two-dimensional space. Pelletier [9] regards each power spectrum as a histogram and provides a reduction of the number of bins with principal component analysis. In contrast, we do not reduce the dimension of power spectra. We assume that the number of component spectra, which are mixed to produce a given set of power spectra, is small and obtain a set of such component spectra by applying a spectral unmixing. Other research fields provide this type of approaches. In the field of remote sensing, we obtain a compact representation of a hyperspectral image by extracting component spectra, called *endmember spectra* [8]. Our research follows this line.

3 Details of Our Method

First of all, we obtain power spectra from a given lung sound by using FFT with Hanning window. In other sets of experiments, whose results are not included in this paper, we confirm that Hamming window also gives similar results, but that flattop window is not effective. We make any two consecutive windows share the half of their lengths. Let each window include $2T$ data points. When the sampling frequency is f Hz, each window corresponds to an interval of length $2T/f$ sec. In this paper, we set $T = 2048$. Two key modules are described below.

Spectral Unmixing We call a set of N consecutive windows *frame*. Since two consecutive windows overlap by one-half the window length, each frame is of length $(N + 1)T/f$ sec. In this paper, we set $N = 32$. We make two consecutive frames share $3N/4$ windows. N power spectra obtained from each frame are approximated by a linear mixture of M ($M \ll N$) spectra, called *component spectra*. We use *fastICA* algorithm [5] to obtain component spectra. The update formula proposed in [4] is employed to achieve stable convergence. Since this paper is intended to reveal basic characteristics of our method, we choose the most simple set up for M . That is, we set $M = 2$. We call a pair of component spectra extracted from each frame *component pair*.

Component Spectra Clustering In our setting, we can obtain many component pairs from a lung sound. We put component spectra taken from various lung sounds into a single set and conduct clustering over this set a large number of times. In this paper, we conduct k -means 100 times. We use no information about which two of component spectra come from the same component pair. If two component spectra from the same component pair belong to the same cluster in many of the 100 clustering results, we can conclude that the unmixing method we adopt is not effective, because a large number of executions of clustering provide a robust estimation of similarities between component spectra. We write $\mathbf{q} \sim \mathbf{q}'$ when two component spectra \mathbf{q} and \mathbf{q}' belong to the same cluster over all clustering results. We call this relation *coupling*. Let $(\mathbf{q}_1, \mathbf{q}_2)$ and $(\mathbf{q}'_1, \mathbf{q}'_2)$ be two component pairs obtained from different frames. $\mathbf{q}_1 \sim \mathbf{q}_2$ and $\mathbf{q}'_1 \sim \mathbf{q}'_2$ are undesirable, because component spectra from the same component pair are obtained by ICA and thus are expected to be dissimilar. We call this type of couplings *malicious*. When we use NMF algorithm in [6] in place of fastICA, a large number of couplings turn out to be malicious. This result shows the low quality

of this version of NMF unmixing. By excluding malicious couplings, we have the following cases of similarity between component pairs $(\mathbf{q}_1, \mathbf{q}_2)$ and $(\mathbf{q}'_1, \mathbf{q}'_2)$:

1. Both $\mathbf{q}_1 \sim \mathbf{q}'_1$ and $\mathbf{q}_2 \sim \mathbf{q}'_2$ hold, or both $\mathbf{q}_1 \sim \mathbf{q}'_2$ and $\mathbf{q}_2 \sim \mathbf{q}'_1$ hold.
2. Either $\mathbf{q}_1 \sim \mathbf{q}'_1$ or $\mathbf{q}_2 \sim \mathbf{q}'_2$ holds, or either $\mathbf{q}_1 \sim \mathbf{q}'_2$ or $\mathbf{q}_2 \sim \mathbf{q}'_1$ holds.

For Case 1, we say that both couplings are *perfect*. The couplings in Case 2 are called *imperfect*. Perfect couplings tell a strong similarity between the two frames corresponding to the two component pairs $(\mathbf{q}_1, \mathbf{q}_2)$ and $(\mathbf{q}'_1, \mathbf{q}'_2)$.

4 Experiment

Settings We use lung sounds recorded in the CD accompanying a textbook for nurses [2]. After identifying different sound data corresponding to lung sounds of the same type, we have lung sounds splitted into 39 categories shown in Table 1. When we use each power spectrum in their full range, many malicious couplings are obtained. High-frequency part seems to have an undesirable effect to our method. Therefore, we eliminate high-frequency part and test the following settings for the range of spectra: 1~256, 1~512, 257~512, and 257~1024. Recall that $T = 2048$. As for the last two cases, we also eliminate low-frequency part. However, the third setting results in a few malicious couplings. This shows that the range 513~1024 has some importance when we discard the range 1~256.

At the initialization of k -means, we set the number of clusters to 100. In the course of the execution of a clustering, some clusters eventually get empty. For example, we obtain from 59 to 79 clusters and 69.7 clusters in average for 100 executions of k -means when we use 1st to 512th entries of power spectra.

Results Table 1 includes the results when we use 1st to 512th entries of power spectra. Column A presents the number of frames obtained from the lung sounds of each category. The total number of frames is 3,357. Since ICA provides two component spectra for each frame, we obtain a set of 6,714 component spectra. The results of 100 executions of k -means induce perfect and imperfect couplings. Column B shows the number of frames giving imperfect or perfect couplings. Column C shows the number of frames giving perfect couplings. When the value in Column B is small, we have some trouble in processing the lung sounds of the corresponding category, because we can expect that component spectra from the lung sounds of the same category form at least imperfect couplings. For example, the category of ID 35 has a small value in Column B. This may be due to the fact that the sound is in small volume. It is, however, beyond our scope to propose preprocessing methods appropriate for our method.

Column D (resp. Column E) gives the number of imperfect (resp. perfect) couplings which include at least one component spectra from the corresponding category. All numbers in Column E are even, because every perfect coupling comes with another perfect coupling. When a coupling includes two component spectra from the same category, we call such a coupling *correct*. Otherwise, a coupling is called *incorrect*. For the category of ID 10, all of 10812 perfect couplings turn out to be correct. That is, each of these perfect couplings consists

Table 1. Experimental results obtained when we apply Hanning window in FFT and use 1st to 512th entries of power spectra.

| ID | lung sound type | A | B | C | D | E |
|----|--|-----|-----|-----|-------|-------|
| 01 | normal vesicular breath sounds | 525 | 391 | 210 | 17233 | 8748 |
| 02 | normal bronchovesicular breath sounds | 84 | 81 | 67 | 3522 | 716 |
| 03 | normal tracheal breath sounds | 44 | 42 | 18 | 896 | 206 |
| 04 | decreased breath sounds (atelectasis) | 129 | 114 | 63 | 4882 | 358 |
| 05 | absent breath sounds (hemothorax) | 134 | 59 | 34 | 973 | 220 |
| 06 | increased breath sounds | 154 | 148 | 108 | 8678 | 4280 |
| 07 | bronchial sounds in abnormal locations | 73 | 36 | 2 | 333 | 2 |
| 08 | low pitched rhonchi | 146 | 103 | 40 | 951 | 114 |
| 09 | high pitched rhonchi | 60 | 44 | 10 | 851 | 10 |
| 10 | fine crackle | 179 | 174 | 157 | 2423 | 10812 |
| 11 | coarse crackle | 201 | 192 | 118 | 7958 | 1988 |
| 12 | spontaneous pneumothorax | 29 | 15 | 0 | 151 | 0 |
| 13 | normal vocal sounds (to be compared with 14) | 27 | 22 | 10 | 169 | 58 |
| 14 | spontaneous pneumothorax (vocal sounds) | 30 | 25 | 17 | 123 | 80 |
| 15 | pleural effusion (effusion side) | 43 | 28 | 2 | 1021 | 2 |
| 16 | pleural effusion (non-effusion side) | 38 | 23 | 0 | 1239 | 0 |
| 17 | asthma | 136 | 133 | 88 | 3029 | 444 |
| 18 | asthma (partial recovery) | 56 | 36 | 5 | 412 | 8 |
| 19 | asthma (nearly complete recovery) | 45 | 28 | 6 | 704 | 10 |
| 20 | pulmonary edema | 100 | 82 | 52 | 967 | 938 |
| 21 | pulmonary edema (nearly complete recovery) | 33 | 27 | 8 | 290 | 24 |
| 22 | chest drain sound | 202 | 142 | 47 | 496 | 138 |
| 23 | tracheal stenosis | 56 | 41 | 9 | 183 | 12 |
| 24 | tracheal stenosis (partial recovery) | 75 | 39 | 15 | 123 | 56 |
| 25 | pneumonia | 115 | 93 | 22 | 1815 | 92 |
| 26 | pneumonia (partial recovery) | 59 | 46 | 19 | 802 | 106 |
| 27 | pneumonia (nearly complete recovery) | 43 | 26 | 5 | 200 | 8 |
| 28 | pleuritis | 118 | 115 | 101 | 2285 | 4796 |
| 29 | congestive heart failure (early stage) | 45 | 33 | 2 | 210 | 2 |
| 30 | congestive heart failure (late stage) | 31 | 15 | 0 | 156 | 0 |
| 31 | congestive heart failure (late stage, mouth breathing) | 24 | 22 | 12 | 87 | 30 |
| 32 | congestive heart failure (recovered) | 62 | 62 | 50 | 2511 | 992 |
| 33 | congestive heart failure (recovered, mouth breathing) | 32 | 21 | 7 | 164 | 20 |
| 34 | sputum collection tube (before sputum collection) | 31 | 15 | 5 | 246 | 20 |
| 35 | sputum collection tube (after sputum collection) | 34 | 6 | 0 | 5 | 0 |
| 36 | air leak around tracheal tube (before adding air) | 41 | 31 | 12 | 240 | 26 |
| 37 | air leak around tracheal tube (after adding air) | 23 | 21 | 19 | 71 | 142 |
| 38 | subcutaneous emphysema | 40 | 21 | 7 | 98 | 10 |
| 39 | pneumothorax | 60 | 28 | 2 | 182 | 2 |

of two component spectra taken from the lung sounds of this category. We can conclude that this category exhibits a strong self-similarity. In contrast, categories of ID 12, 16, 30 and 35 provide no perfect couplings. Among these four categories, the sounds of categories 12 and 35 are in small volume. Further, the lung sounds of category 16 have an irregularly long respiration cycle. These categories require appropriate preprocessing methods, which is our important future work. The lung sounds of category 30 seem to show rapid changes in the same respiration cycle. We may need to try other settings for T or for the overlapping widths of windows. The category of ID 17 provides interesting results. While all 444 perfect couplings are correct, the number of imperfect couplings is 3,029, far larger than 444. Moreover, 2,330 among these 3,029 imperfect couplings are correct. We may explain these results by the fact that the lung sounds of this category show poor periodicity. Stable characteristics is reflected by one half of each component pair, and changeable characteristics is reflected by another.

Evaluation of Results We can make equivalence classes by regarding perfect coupling as an equivalence relation. These equivalence classes are, in a sense, “meta-clusters” constructed based on multiple clustering results. Each equivalence class can provide a powerful clue to form templates useful in classification of unknown lung sounds. It is an important future work to devise a procedure of template composition for the categories giving no perfect couplings, e.g. categories of ID 12, 16, 30 and 35, because we can have no meta-clusters for these categories. Table 2 presents an evaluation of this meta-clustering. The number of categories which give at least two perfect couplings is shown in the column marked by “*” for each setting. The number in the column marked by “†” refers to the number of component spectra which are taken from the corresponding category and, at the same time, are included in at least one perfect couplings. Table 2 also presents an evaluation of meta-clustering. We use 39 categories presented in Table 1 as the ground truth. The evaluation measures are microaveraged/macroaveraged precision, microaveraged/macroaveraged recall and their harmonic mean. These measures are often used in the evaluation of clustering and are defined as follows. The *dominating category* of a cluster is the category providing the largest number of component spectra to that cluster. The *precision* of a cluster is equal to the number of component spectra from the dominating category divided by the cluster size, i.e., the number of component spectra included in the cluster. The *recall* of a cluster is the number of component spectra from the dominating category divided by the number of all component spectra from that category. We can summarize the precisions and the recalls of all obtained clusters by microaveraging or by macroaveraging. *Microaveraged precision* is equal to the sum of the numerators of all precisions divided by the sum of the denominators of all precisions. We can obtain *microaveraged recall* in the same manner. *Macroaveraged precision* (resp. *macroaveraged recall*) is computed simply as an arithmetic mean of all precisions (resp. all recalls).

In all settings, we have high precisions and low recalls, because the numbers of clusters, shown in the last column, are far larger than the number of categories. However, our method is designed for template composition in classification task.

Therefore, each category can have several templates. The number of dominating categories is shown in the column “**”. When this number is smaller than that presented in the column “*”, there exist categories which can dominate no clusters. We have no meta-clusters also for such categories. Only when we use 1st to 512th entries of power spectra, every category dominates at least one cluster. We also have the best precision, i.e., 0.9711, for this setting.

Table 2. Evaluation of “meta-clusters” induced by perfect couplings

| range | * | † | precision | | recall | | harmonic mean | | ** | # of clusters |
|----------|----|------|---------------|--------|--------|--------|---------------|--------|----|---------------|
| | | | micro | macro | micro | macro | micro | macro | | |
| 1-256 | 37 | 1276 | 0.9561 | 0.9473 | 0.0778 | 0.1500 | 0.1438 | 0.2590 | 36 | 228 |
| 1-512 | 35 | 1349 | 0.9711 | 0.9688 | 0.0680 | 0.1343 | 0.1272 | 0.2358 | 35 | 252 |
| 257-512 | 31 | 1017 | 0.5320 | 0.9365 | 0.0816 | 0.1519 | 0.1416 | 0.2614 | 28 | 130 |
| 257-1024 | 31 | 657 | 0.8706 | 0.9637 | 0.0546 | 0.1236 | 0.1028 | 0.2191 | 30 | 208 |

5 Conclusion and Future Work

In this paper, we propose a method for detecting robust similarities between short intervals taken from various lung sounds. The results of evaluation show that we can obtain equivalence classes of component spectra with high precision. However, some types of lung sounds do not provide meaningful similarities. We need to propose preprocessing methods for such lung sounds. Further, in the near future, we will provide the procedures for composing templates and will devise a method for comparing unknown lung sounds with the templates.

References

1. <http://www.rale.ca/pub/RRC.PDF>
2. Yonemaru, M., Sakurai, T.: Lung Sound Auscultation Training via CD for Nurses. Nanko-do, Tokyo (2001)
3. Güler, E. C., Sankur, B., Kahya, Y. P., Raudys, S.: Visual Classification of Medical Data Using MLP Mapping. *Computers in Biology and Medicine* **28** (1998) 275–287
4. Hyvärinen, A.: Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks* **10**, Issue 3 (1999) 626–634
5. Hyvärinen, A., Karhunen, J., Oja, E.: *Independent Component Analysis*. Wiley-Interscience (2001)
6. Lee, D. D., Seung, H. S.: Algorithms for Non-negative Matrix Factorization. *Advances in Neural Information Processing Systems* **13** (2001) 556–562
7. Marshall, A., Boussakta, S.: Signal Analysis of Medical Acoustic Sounds with Applications to Chest Medicine. *J. of the Franklin Inst.* **344**, Issues 3-4 (2007) 230–242
8. Miao, L., Qi, H., Szu, H.: Unsupervised Decomposition of Mixed Pixels Using the Maximum Entropy Principle. in *Proceedings of ICPR’06* (2006)
9. Pelletier, C.: *Classification des Sons Respiratoires en vue d’une Détection Automatique des Sibilants*. Master Thesis, Université du Québec à Rimouski (2006)