Smart Audio Sensor on Anomaly Respiration Detection using FLAC features

Jiaxing YE
Computer Science Dept. Univ. of Tsukuba, Japan

Takumi Kobayashi
AIST, Japan

Tetsuya Higuchi
AIST, Japan
Computer Science Dept. University of Tsukuba, Japan

IEEE Sensor Application Symposium, Brescia, Italy, Feb. 8th, 2012

Content overview

- ◇Research background
- ◇The proposed approach
  - The FLAC feature extraction method
  - The anomaly detection framework
- ◇Experimental validations
- ◇The generalized application of proposed scheme
- ◇Conclusions

Research background

The facts of respiration symptoms from World Health Org.(WHO)

1. Many kinds of respiration diseases, worldwide spread, e.g. 235million people suffering from asthma.
2. Chronic disease or cannot be cured, by well treatment works for higher life quality.
3. Early and effect diagnosis is of crucial importance

Research background

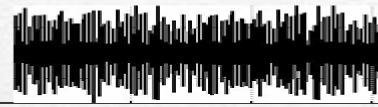
Res
man

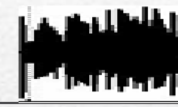
Har (neglect)
Erro ng
Tim + symptom)

Computerized
Symptom
targeting

Research background

Goal: supporting respiration symptom early diagnosis





Respiration sounds: healthy (Common)+ symptom (Seldom)

Anomaly detection by the proposed system

Normal parts

Discard the healthy sections

Abnormal sections recording: suspected symptoms

Further diagnosis by doctor

Research background

Requirements for computerized respiration diagnosis

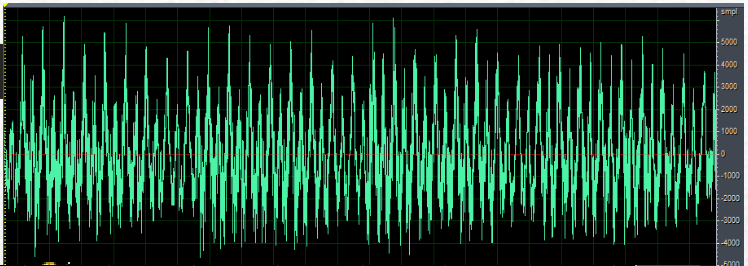
1. Suitable for each individual with variation adaptation
2. High detection performance, especially for early diagnosis
3. High efficiency for real-time operation and hopefully low cost for wide applications.

Algorithm interpretation:

1. Unsupervised manner for different individual
2. Online scheme for self-adaptation
3. Algorithm effectiveness and efficiency

The proposed approach

The FLAC feature extraction m

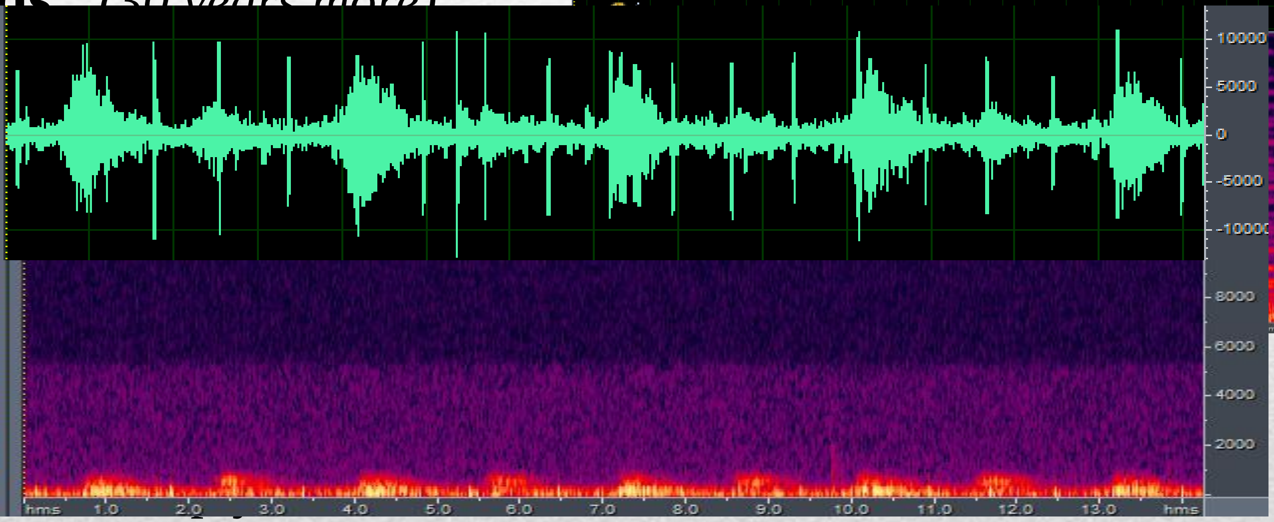


Acoustic feature extraction:

Structured sounds: *(30 years more)*

Sp

M



- **u**

V

pre-defined model: respirations

The proposed approach
The FLAC feature extraction method

Acoustic signals:

- **Structured sounds:** *(30 years more)*

Features:

MFCC (Mel-frequency cepstral coefficients)
LPC (Linear Predictive Code), Subband Energy
ZCR (Zero-Cross Rate), Pitch, Spectrogram…

- **Unstructured sounds:** *(new, 2005 --)*

Features:

MPEG 7.

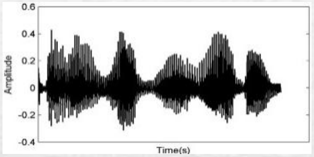
Utilizing the new feature for respiration analysis (FLAC)

J. Ye el al, "Audio-based Sports Highlight Detection by **F**ourier **L**ocal **A**uto-**C**orrelations." INTERSPEECH 2010: 2198-2201

The proposed approach

The FLAC feature

☆ 1. Taking advantage of **complex** spectrogram, not rely on the magnitude only.



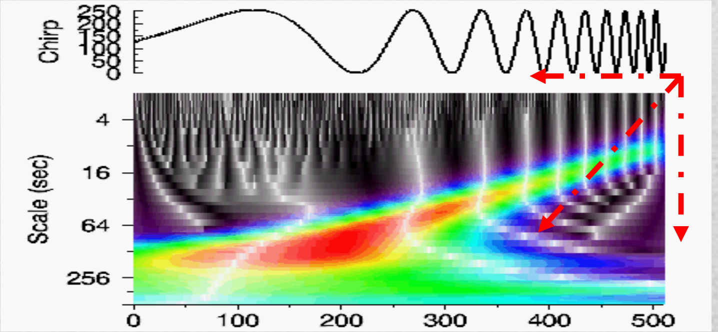
FFT

$$a + bi = \underline{A} \cdot e^{\underline{j\varphi}}$$

Phase

Magnitude

☆ 2. Extracting the time-frequency domain dynamics features for representing unstructured sounds.



The proposed approach

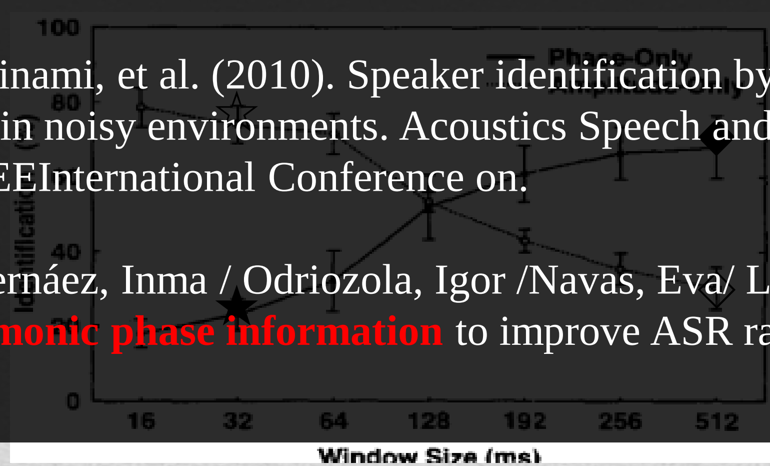
The FLAC feature extraction method

Recent advancements on using phase information

Rajan, P., S. H. K. Parthasarathi, et al. (2009).Robustness of **Phase based Features** for Speaker Recognition. Proc. INTERSPEECH, 2009.

Longbiao, W., K. Minami, et al. (2010). Speaker identification by **combining MFCC and phase information** in noisy environments. Acoustics Speech and Signal Processing (ICASSP), 2010 IEEEInternational Conference on.

Saratxaga, Ibon / Hernáez, Inma / Odriozola, Igor /Navas, Eva/ Luengo, Iker / Erro, Daniel (2010): "**Using harmonic phase information** to improve ASR rate", In INTERSPEECH-2010, 1185-1188.

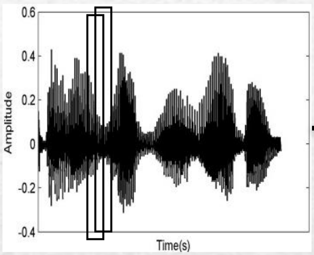


Do magnitude + phase!

Source : L.Liu, J.He and G.Palm, ”Effects of phae on the perception of intervocalic stop consonants”, Speech Communication, Vol.22 No. 4, pp.403-417, Sept. 1997

The proposed approach

The FLAC feature extraction method



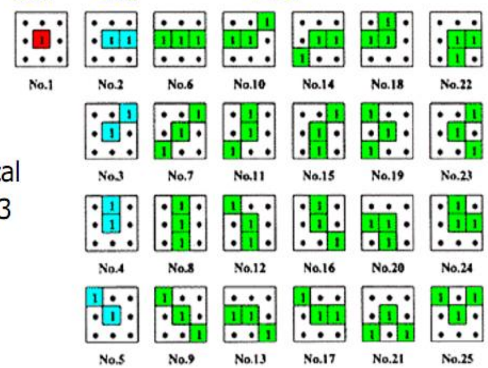
$y[n] = x[n$

Pre-

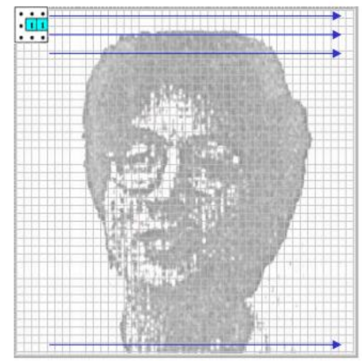
HLAC (Higher-order Local Auto-Correlation)

$$x(a_1, \cdots, a_N) = \int I(r)I(r+a_1)\cdots I(r+a_N)dr$$

N- th Order



x: dimension = 35 (gray) , 25 (binary)



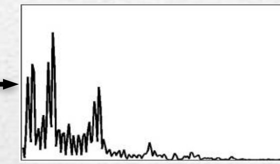
Scanning image and summing up the products

Goudail, F., et al. (1996). "Face recognition system using local autocorrelations and multiscale integration." IEEE-PA MI , Vol. 18, Issue 10, pp. 1024-1028, 1996

$$a+bi|^2 = |A \cdot e^{j\varphi}|^2 = A^2$$

Power Spectrum

FT

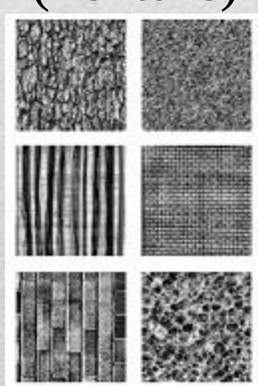


Frequency

1.STFT (S

e spectrogram

(Texture)



Extra
dyna
on t-f
by lo
descriptor

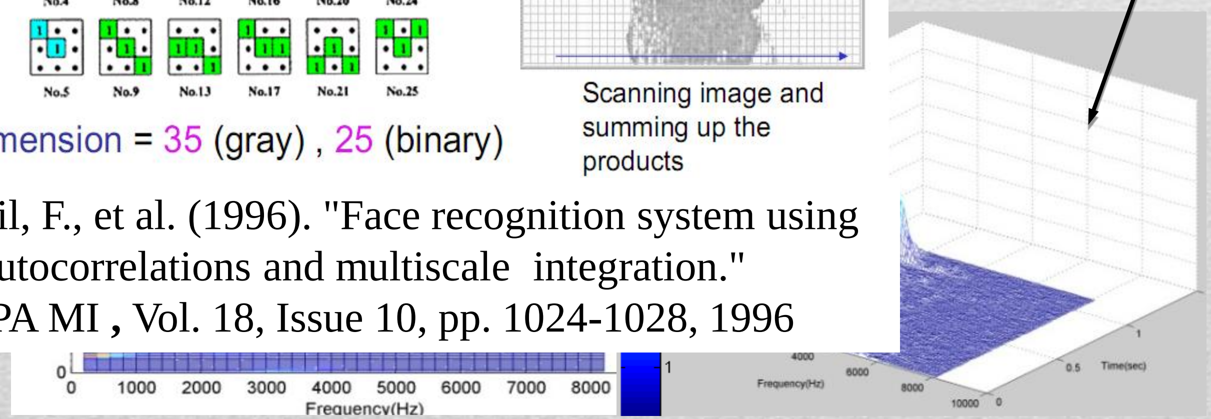


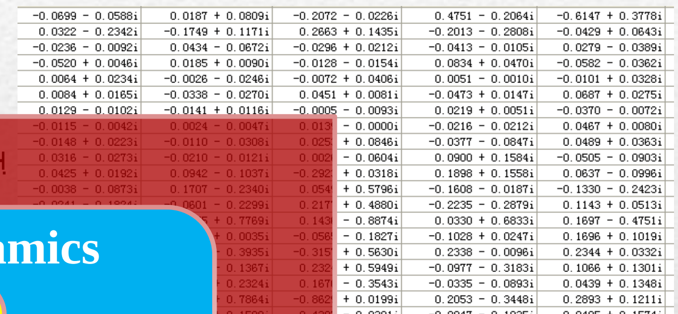
Image processing on (complex) spectrogram

The proposed approach

The FLAC feature extraction method

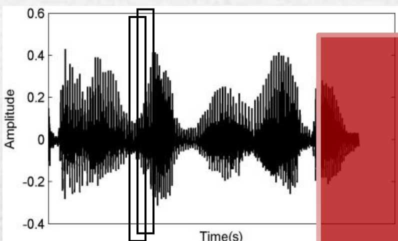
- Details in FLAC feature extraction

$a+bi=A\cdot e^{j\varphi}$ *Phase+magnitude*

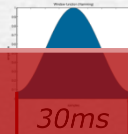


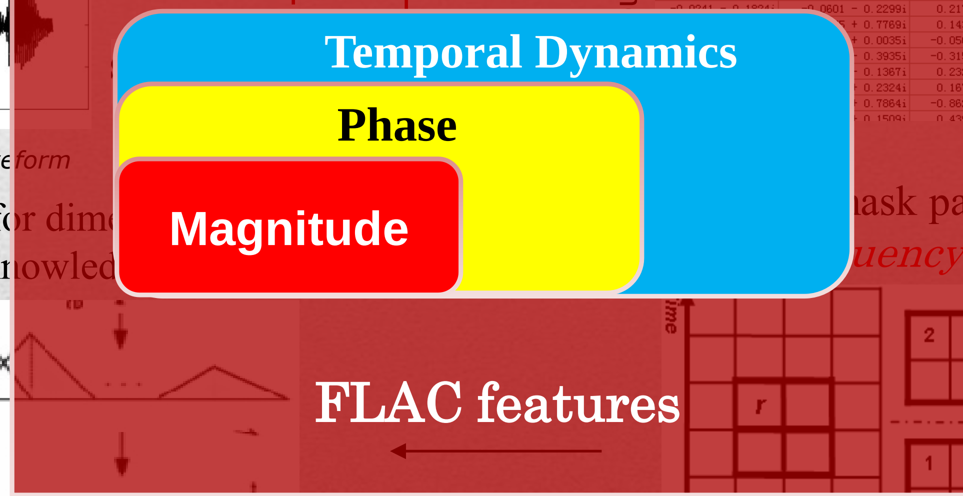
Frequency (d-dim.)

***a*. Short-time Fourier Transform**

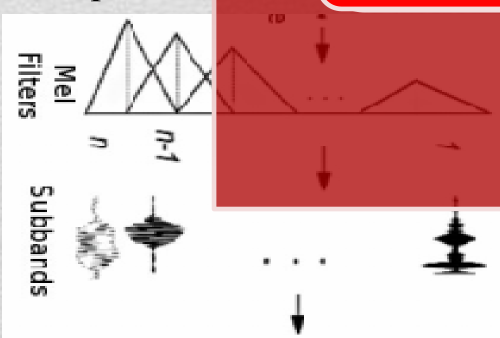


Time-waveform



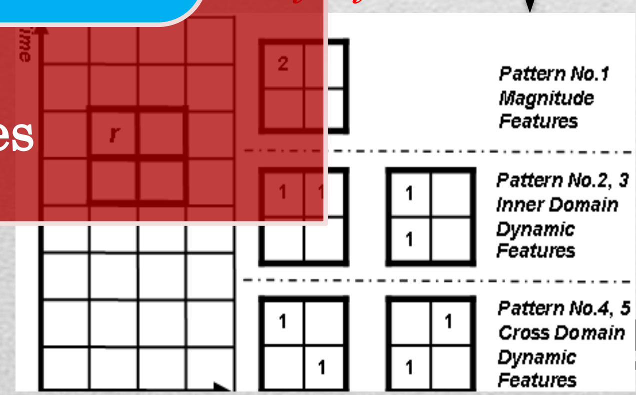


c. Mel-filter bank for dim... based on prior-knowled...



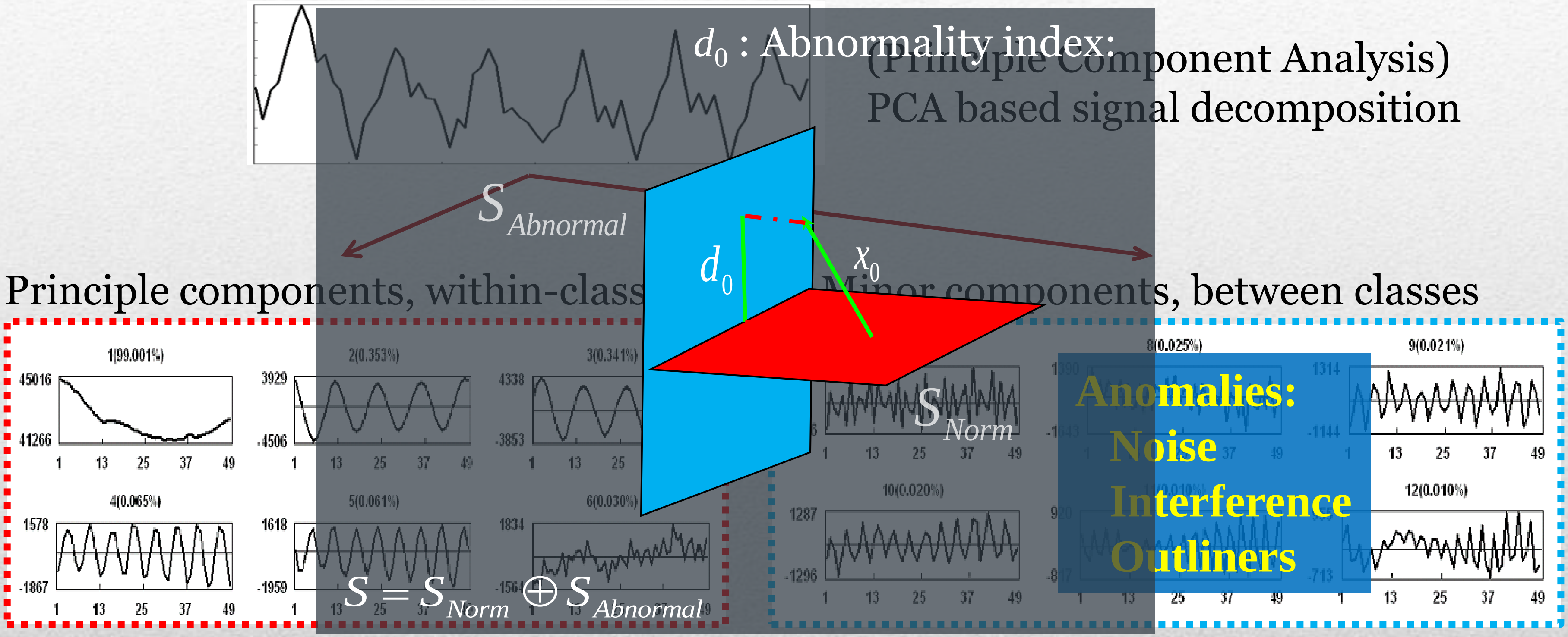
...ask patterns

...uency dynamics



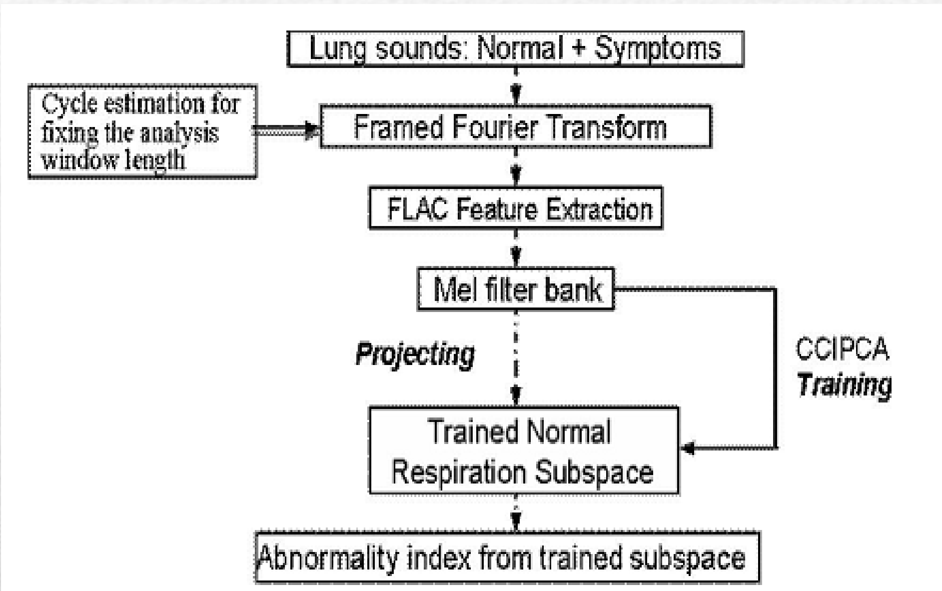
The proposed approach

The anomaly detection framework



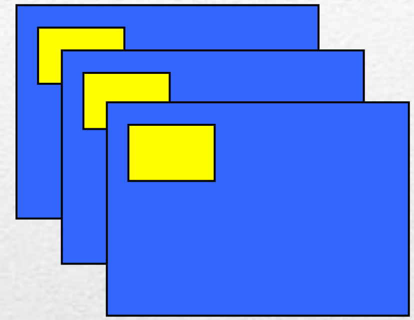
Source: Hassani, Hossein (2007): *Singular Spectrum Analysis: Methodology and Comparison.* Journal of Data Science , Vol. 5, No. 2 (01. April 2007): pp. 239-257.

The proposed approach
The anomaly detection framework

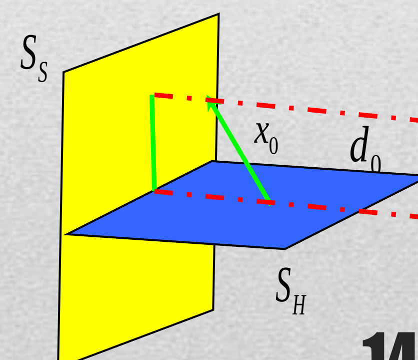


Detection framework

Sequential input



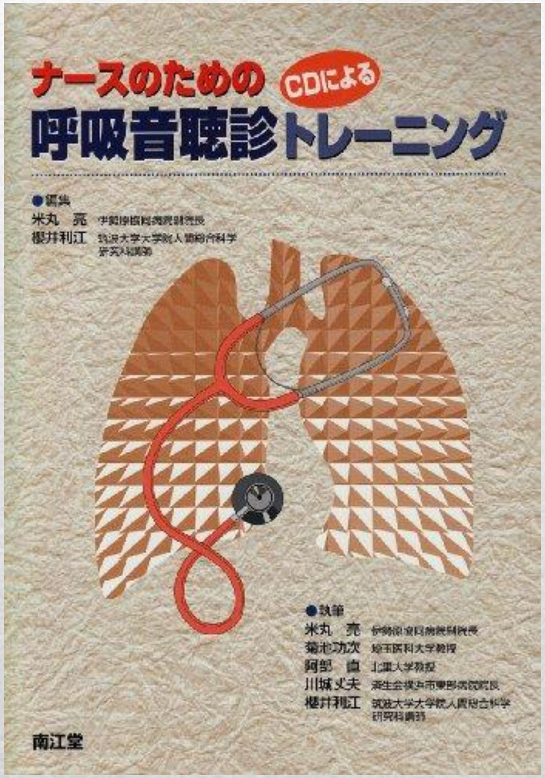
Timely updating



Experimental validation

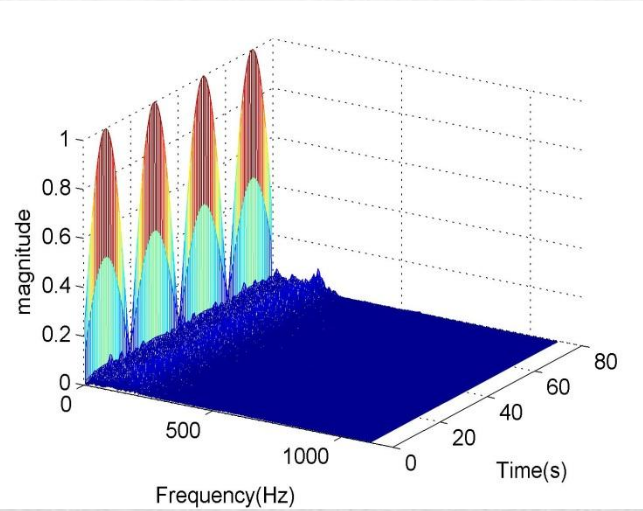
Dataset Introduction: stethoscope recordings at 44.1 kHz with 16bit quantization.

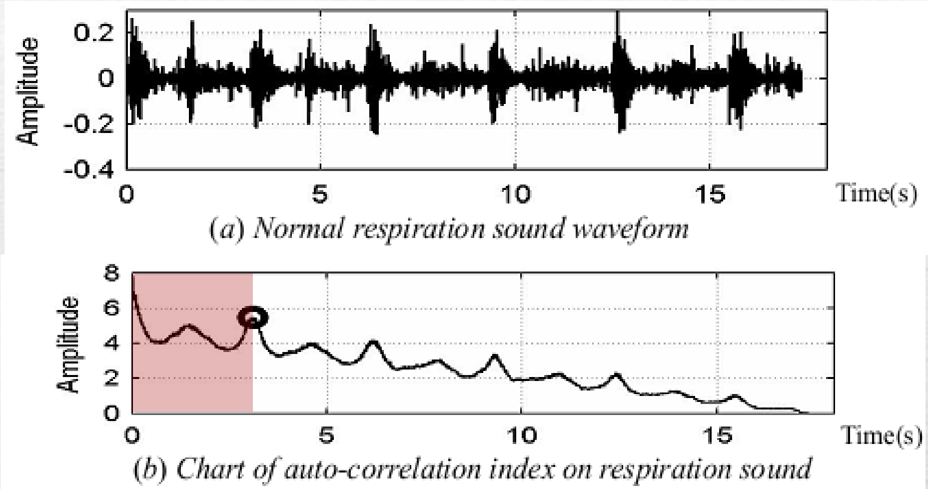
Healthy respiration sounds	Symptom respiration sounds	
Class1 Normal Symptom respiration sound data	Adventitious symptoms	Breath symptoms
	Fine crackle, Bronchiectasis, Bronchial stenosis, Asthma, Tracheobronchial stenosis, Pulmonary fibrosis (Interstitial pneumonia), Coarse crackle, Rhonchus Adult respiration distress syndrome (ARDS), Wheeze, Pneumonia, Expectoration Congestive heart failure, Pulmonary edema stethocatharsis sedimentation	Spontaneous pneumothorax, Retention of pleural effusion Atelectasis, Tracheal stenosis Hemopneumothorax, Enhanced respiration **Class2 Abnormal**



Experimental validation

Parameter setup





(a) Normal respiration sound waveform

(b) Chart of auto-correlation index on respiration sound

Frequency location
$f < f_{max}$

Fourier analysis window length
$L_{Fourier}$ = Cycle of respiration

Experimental validation

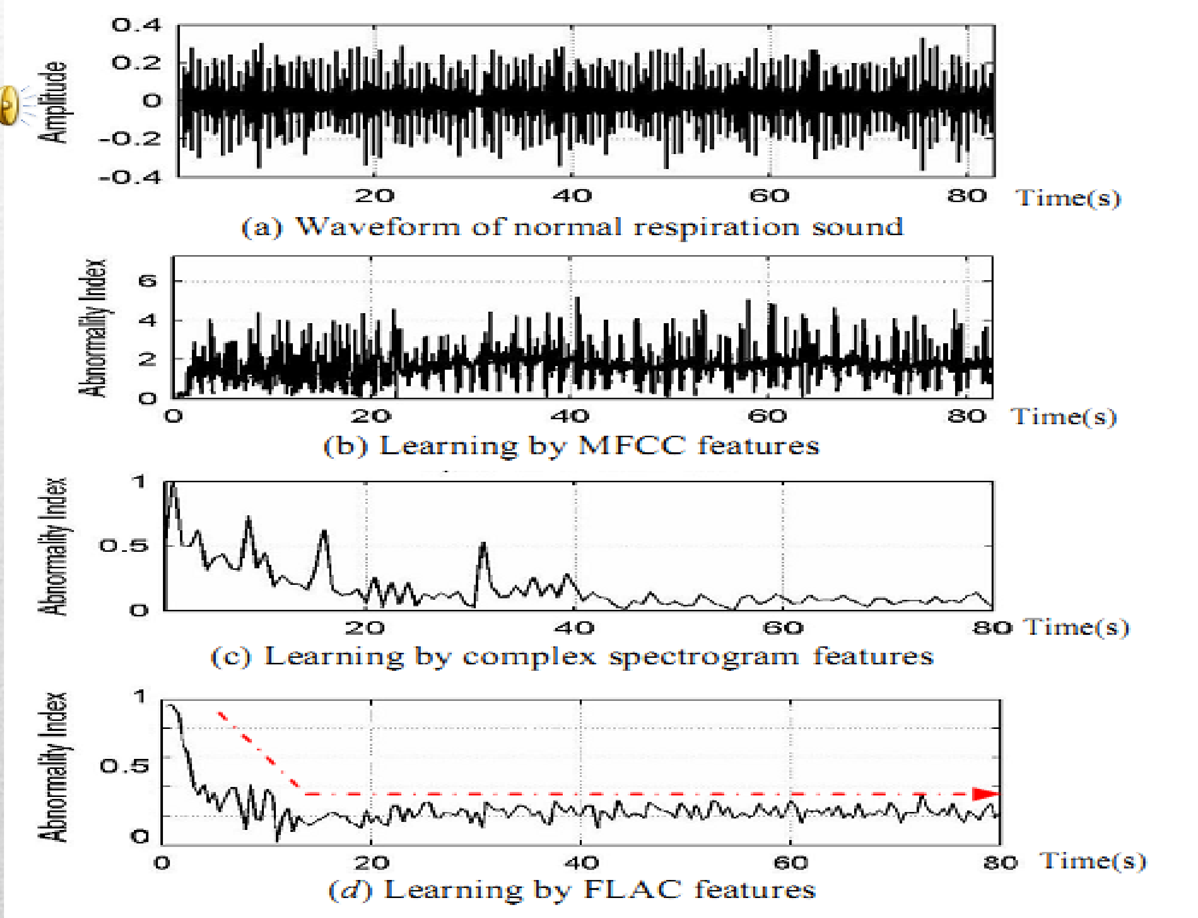
Modeling normal respirations

Data Input

MFCC

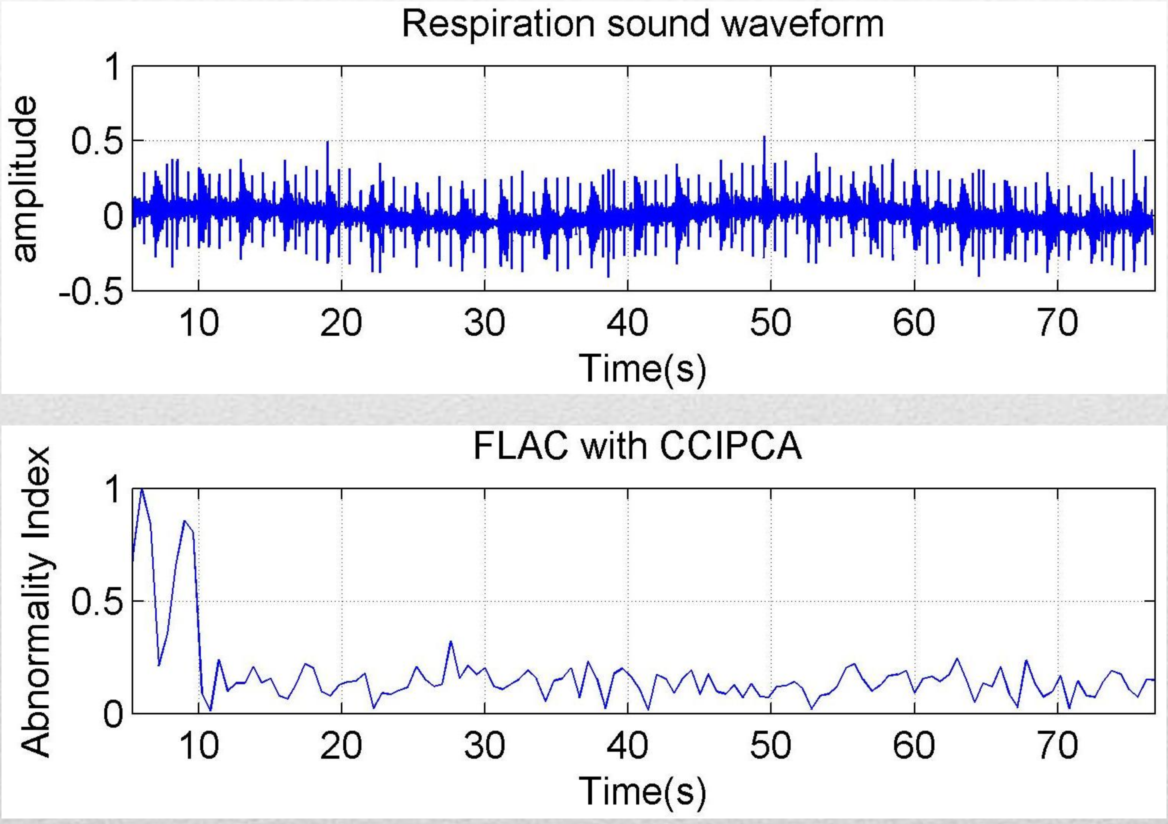
Complex spectrogram

FLAC



Experimental validation

Modeling normal respirations: Base line wandering effect

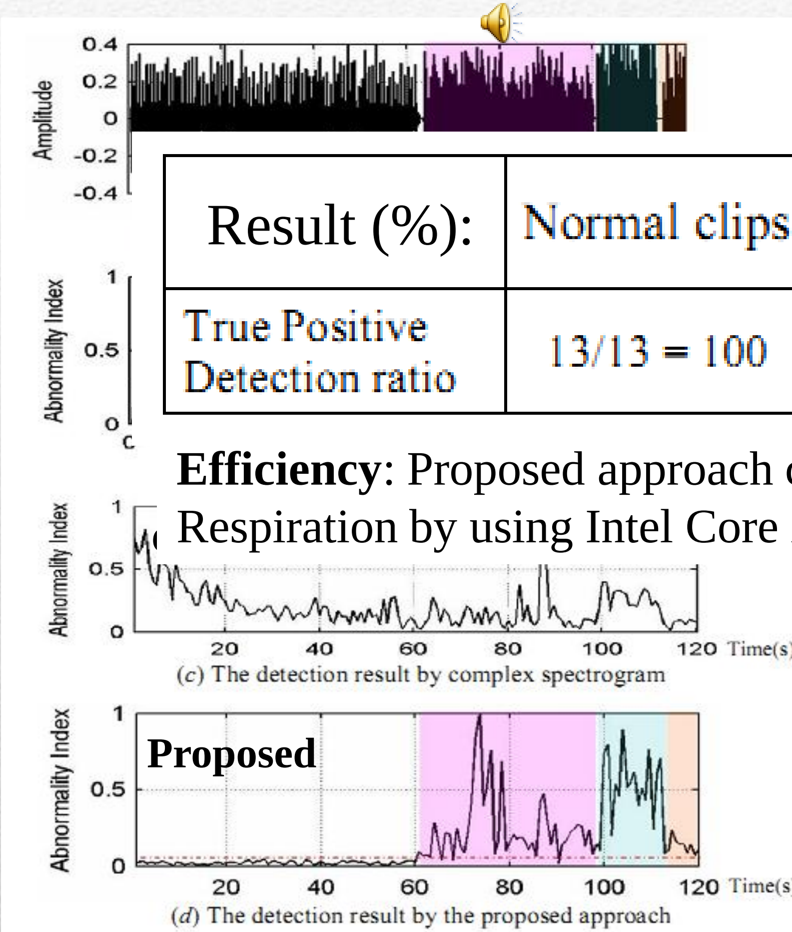


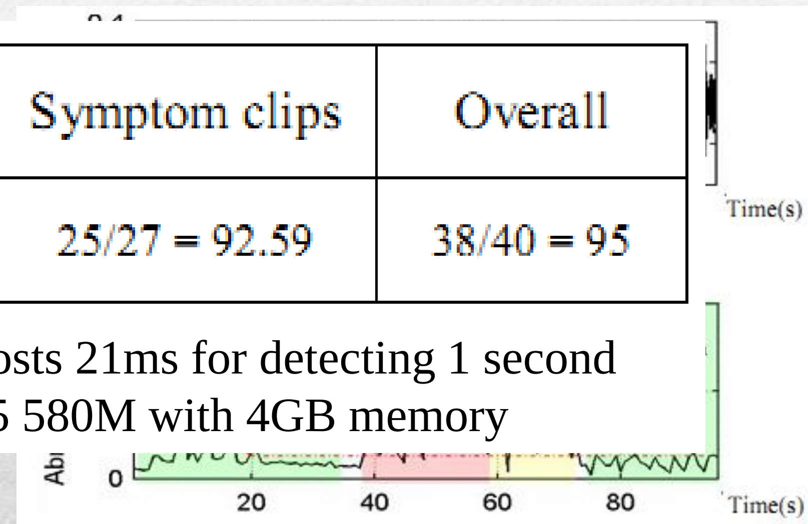
Experimental validation

Anomaly detection in respiration sound

Case 1:

Case2: symptoms among normal respirations



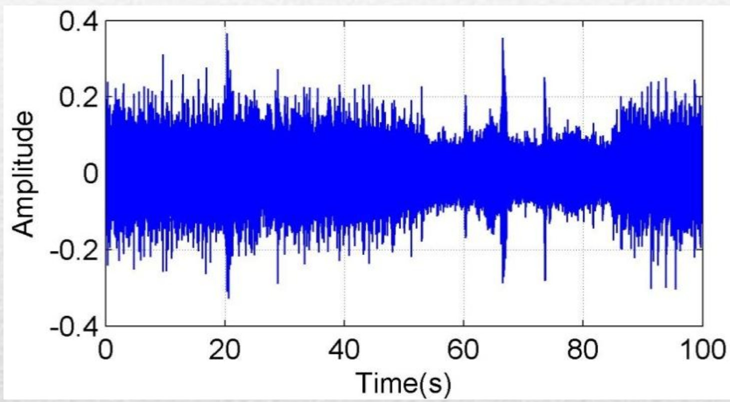


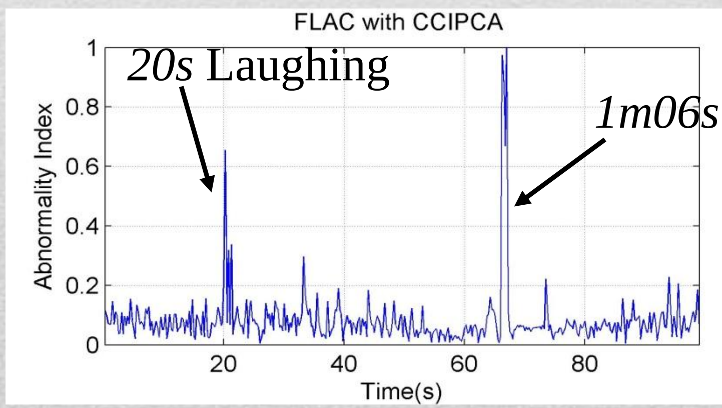
Result (%):	Normal clips	Symptom clips	Overall
True Positive Detection ratio	13/13 = 100	25/27 = 92.59	38/40 = 95

Efficiency: Proposed approach costs 21ms for detecting 1 second Respiration by using Intel Core i5 580M with 4GB memory

Generalized application: foreground acoustic event detection

Acoustic environment modeling for acoustic event detection

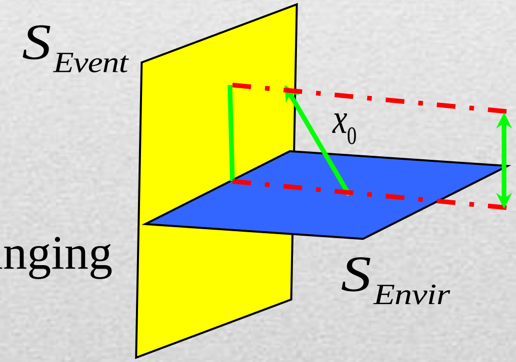




20s Laughing

1m06s Bell ringing

Acoustic event in noisy factory environment



Conclusions

Unsupervised learning for individuals respiration character

Adaptation to the variations in respiration in an online manner

High anomaly respiration detection performance

High efficiency (FFT with Local Auto-correlation)

Thank you very much for your attention !

Questions