A machine sound monitoring for predictive maintenance focusing on very low frequency band

Kazuki Tsujii, Shota Imaia, Ryota Takaoa, Tomonori Kimurab, Hitoshi Kondob and Yukihiro Kamiya

aDepartment of Information Science and Technology, Aichi Prefectural University, Aichi, Japan; bCosmotec Co. Ltd., Nagoya, Aichi, Japan

ABSTRACT
The monitoring of machines is one of key issues in the Industry 4.0 era. Particularly, the monitoring realized by non-contact sensors is drawing attention since it is easy to install and to avoid any problems caused by sensors accidentally dropping down into the machines. For example, sound monitoring satisfies this requirement. In this paper, we propose to apply the Accumulation for Real-time Serial-to-parallel Converter (ARS) for the monitoring of machine sounds to analyse low frequency bands which have not been sufficiently investigated so far. The machine sounds captured in a real factory are analysed so that the change of the machine sounds which varies in accordance with machine status is detected. It is verified that ARS successfully detects the difference as precise as wavelet transform (WT) with Morlet wavelet even though its computational load is significantly lower than that of WT.

ARTICLE HISTORY
Received 6 January 2020
Accepted 13 September 2020

KEYWORDS
Sound monitoring; machine sound; ARS; FFT; low frequency

1. Introduction

Needless to say, the Internet of Things (IoT) is a key technology for the realization of Industry 4.0 [1–3]. It is expected that the efficiency of industries is drastically improved by taking advantage of big data collected by IoT. One of the applications of the big data is predictive maintenance [4,5]. It is to detect machine problems before they cause a serious damage to the productivity.

The predictive maintenance is often realized by the monitoring of machine vibrations as well as the quantity of machine oil, the temperature etc. [6,7]. Particularly, it is empirically known that the vibration of machines contains information regarding machine faults. For example, experienced machine operators can detect the machine faults by listening to machine sounds [8,9]. Since the sound is caused by the machine vibrations, it is natural that the sound varies in accordance with the machine status. Therefore, it is expected that the analysis of the machine sounds can contribute to the predictive maintenance. In reality, the sound monitoring is also advantageous because it is not necessary to put any sensors or electrodes on the machine surface so that we can avoid any problems caused by the sensors accidentally dropping down into the machine. However, the predictive maintenance by the machine sounds should be realized without relying on experiences or intuitions of the machine operators.

Although the application of machine learning techniques for the analysis has been proposed [10–12], it is difficult to collect a vast amount of training data. Definitely, it must be perfect if we can detect the machine faults immediately from the results of the analysis, without feeding the captured sound data into neural networks, etc.

The analysis of the sounds has been typically performed through the fast Fourier transform (FFT). To cope with time-variant signals, the short-time Fourier transform (STFT) based on FFT is also applied. However, it is well-known that FFT is not suitable for the analysis of low-frequency signals [13]. In order to improve the resolution of the low frequency band, FFT requires a huge number of signal samples. It results in not only a huge amount of memory and computational load but also long time to capture the signal samples.

Aiming to reduce the computational load of FFT and to improve the resolution in the low frequency band, the Accumulation for Real-time Serial-to-parallel converter (ARS) [14] was proposed. Actually, the low computational load and the high resolution in the low frequency band are often important in IoT scenarios. For example, in case of the analysis of vital signs, the very low frequency is interested since the frequency of respirations or heartbeats is typically lower than 1 Hz where FFT is not advantageous.

In the machine monitoring, we can expect that the low frequency bands which have not been sufficiently investigated might consist of remarkable information relevant to the machine faults. To the best of the authors’ knowledge, there is no extant researches dealing with the analysis of the low frequency band of the machine sounds captured in a real factory.
sound. In addition, if we can apply ARS in this field, the computational complexity needed for the machine monitoring is reduced.

In this paper, the performance of the machine monitoring through the sounds using ARS is examined not only by computer simulations but also experiments conducted in a real factory. In this factory, the sounds of a press machine are captured before and after its overhaul. Therefore, the objective of the experiment is to verify if the sounds before the overhaul differ from that after the overhaul.

The sounds are analysed through ARS, STFT and wavelet transform (WT). It is verified that ARS and WT with the Morlet wavelet achieves the best performance in terms of the histogram correlation even though ARS requires lower computational load than WT.

The rest of the paper is organized as follows: The motivation of this research is explained in Section 2. Then, signals are formulated in Section 3 followed by the overview of ARS explained in Section 4. Before going to the experiments in Section 5, Section 4.3 provides several numerical examples analysed by ARS. Through discussions in Section 6, this paper is concluded in Section 7.

2. Research Motivation

Press machines are widely used in factories to shape materials by pressing with a very high pressure. The movement of the press machine is relatively slow, typically one stroke of the pressing takes around 1 s.

Through the overhaul of the press machine used in this experiment, the following two problems are found.

(1) The chipping of gears
(2) The slippage of the rotation axis of crank shafts caused by abrasions

Since the movement of the gears and the crank shafts are very slow, it is expected that the two problems affect very low frequency band of the sounds.

For the sound monitoring of press machines, to the best of our knowledge, no extant researches which suggest requirements on the sound monitoring for press machines have been found among publications.

Extant researches are found if we focused on the monitoring for the vibrations of bearings [15–17] and the sound of motor [18]. Although an extant research in [19] deals with a press machine, the target of the monitoring is products produced by the press machines.

It means that there is no a priori knowledge as requirements of the measurement, such like a frequency of the sound corresponding to a certain phenomenon of the press machine. So, as the first challenge, this paper tries to verify if the changing of the sound is detected by the signal processing before and after the overhaul where the above-mentioned machine problems are fixed. Although it is very basic, this paper suggests the following points in the sound monitoring.

(1) The difference of machine sounds between in a normal status and in a status under machine abnormalities are quantitively distinguished without being depending on human experiences or intuitions.
(2) The sound analysis should not employ machine learning techniques. This is because such techniques need a number of training data for each of many cases. It costs a lot in terms of time and effort.

In the following sections, it will be shown that the two points above will be fulfilled by the sound analysis with ARS.

The advantages and disadvantages of the machine sound analysis aiming at the fault detection are briefly discussed. The sound monitoring is advantageous due to the following reasons:

- It is easy to capture the machine sounds using microphones as relatively cheap sensors.
- There is no need to install sensors directly onto a machine surface.

In contrast, it is disadvantageous that the sound is easily contaminated by unwanted sounds [6]. Therefore, the fault detection by the sound monitoring has not been well considered compared with other methods [7].

In this paper, the sound analysis by ARS has been proposed and compared with DFT and WT as a first step for press machines which are one of the most common machinery in a variety of factories. Thinking about the real field where the sound monitoring is conducted, the contamination by unwanted sounds will be the first to be considered. The unwanted sounds might be mitigated by using properly the directivity of the microphone, or a microphone array. However, the verification of their performance improvement will be topics of our future considerations.

3. Formulations of signals

Let \( x(t) \) denote the sound recorded by a microphone. The output of the microphone is sampled with the sampling period \( T_S \) as follows:

\[
x[k] = x(kT_S)
\]  

where \( k (k = 0, \ldots, K - 1) \) is an integer as a discrete time index.

The sampled sound is formulated as follows:

\[
x[k] = s[k] + \eta[k]
\]
where \( s[k] \) and \( \eta[k] \) denote the sound and noise, respectively. It is assumed that the sound \( s[k] \) contains multiple periodical waveforms formulated as follows:

\[
s[k] = \sum_{u=0}^{U-1} s_u[k] \tag{3}
\]

\[
s_u[k] = \sum_{v=0}^{V_u-1} \sqrt{P_S^{(u)} \phi_u[k - vT_u - \tau_u]} \tag{4}
\]

where \( s_u[k] \) denotes the \( u \)th (\( u = 0, \ldots, U-1 \)) waveform formed as a chain of a corresponding fundamental waveform \( \phi_u[k] \) whose period is \( T_u \). In addition, \( V_u \) is the number of the fundamental waveform included in \( s_u[k] \) while \( P_S^{(u)} \) is the power of the \( u \)th waveform. An unknown delay is denoted by \( \tau_u \).

### 4. ARS overview and computer simulations

In this section, ARS is mathematically defined in order to clearly identify the parameters in computer simulations and experiment explained later.

Figure 1 shows the configuration of ARS. An intuitive explanation of ARS is provided in [14]. The input sample sequence \( x[k] \) is fed into the multiple SPCs. At each of the SPCs, the SPC outputs are accumulated so that the maximum value is detected from the accumulated waveform. Finally, the maximum values are compared. It should be noted that the period of the input sequence corresponds to the number of output ports of the SPC which yields the maximum value among the SPCs.

#### 4.1. Formulations of ARS

Suppose that there are \( N \) SPCs. The SPCs are identified by SPC ID \( \# n \) \((n = 1, \ldots, N)\). The number of output ports of SPC \( \# n \) denoted by \( M_n \) is defined as follows:

\[
M_n = L_{\text{min}} + n - 1 \tag{5}
\]

where \( L_{\text{min}} \) is the number of output ports at SPC \( \# 1 \). The \( g \)th \((g = 0, \ldots, G-1)\) output of the SPC \( \# n \) is defined as a vector \( x_n[g] \) of size \((M_n \times 1)\) as

\[
x_n[g] = \begin{pmatrix} x[gM_n] \\ x[gM_n + 1] \\ \vdots \\ x[(g + 1)M_n - 1] \end{pmatrix} \tag{6}
\]

The accumulation at the SPC \( \# n \) is formulated as follows:

\[
y_n = \sum_{g=0}^{G_n-1} x_n[g] \tag{7}
\]

where \( G_n \) is the number of the output waveforms at SPC \( \# n \) defined as follows:

\[
G_n = \left\lfloor \frac{K}{M_n} \right\rfloor \tag{8}
\]

where \( \lfloor a \rfloor \) is the maximum integer which is less than or equal to \( a \).

Finally the maximum amplitude of the entities of \( y_n \) is identified as follows:

\[
\hat{y}_n = \max_{m=1, \ldots, M_n} |y_{n,m}| \tag{9}
\]

where \( y_{n,m} \) denotes the \( m \)th entity of \( y_n \).

Actually, the number of the accumulated waveforms at SPC \( \# n \) denoted by \( G_n \) is not equal to that of the other SPCs. In order to compensate the difference, the following averaging is performed.

\[
z_n = \frac{\hat{y}_n}{G_n} \tag{10}
\]

Then, the maximum value of \( z_n \) \((n = 1, \ldots, N)\) is identified as follows:

\[
\hat{z} = \max_{n=1, \ldots, N} z_n. \tag{11}
\]

Suppose that \( \hat{n} \) is the index corresponding to the \( z_n \) i.e.

\[
\hat{z} = z_{\# \hat{n}}. \tag{12}
\]

Now we estimate the period of the \( u \)th periodic signal \( T_u \) as \( \hat{n} \).

#### 4.2. Application of ARS to the machine monitoring

In the experiment, aiming to observe time-variant signals, we apply a short-term ARS formulated as
explained in this section. As depicted in Figure 2, the \( K \) samples are divided into \( R \) slots defined as follows:

\[
R = \left\lfloor \frac{K}{Q} \right\rfloor. \tag{13}
\]

where \( Q \) denotes the number of samples contained in a slot. So the first sample of the \( r \)th slot \((r = 0, \ldots R - 1)\) is identified by the discrete time index \( k \) as

\[
k = rQ \tag{14}
\]

while the last sample of the slot is as follows:

\[
K_r = (r + 1)Q. \tag{15}
\]

For each of the \( R \) slots, the procedure introduced in Section 4.1 is executed.

Owing to the principle of ARS, such like FFT, it is not suitable to analyse non-stationary signals. However, it is considered that the machine sound is stationary since its movement is generated by the rotation of crank shafts and gears. Therefore, the sound is periodic and stationary.

### 4.3. Computer simulations

Prior to experiments, we conduct a computer simulation for performance verification of ARS. In this simulation, we generate two signals \( s_1[k] \) and \( s_2[k] \) formed as a chain of the corresponding fundamental waveforms \( \phi_u[k], u = 1, 2 \), respectively, as depicted in Figure 3. The periods are defined as \( T_1 = 99 \) while \( T_2 = 101 \). [AE-1] The two fundamental waveforms are generated so as to be identical except the period through as follows:

\[
\phi_u[k_u] = b[k_u]c[k_u] \tag{16}
\]

\[
b[k_u] = 0.95^{-k_u} \tag{17}
\]

\[
c[k_u] = \cos \pi k_u \tag{18}
\]

where \( k_u = 0, 1, \ldots, (T_u - 1), u = 1, 2. \)

By using the generated fundamental waveforms \( \phi_1[k] \) and \( \phi_2[k] \), we generate \( s_1[k] \) and \( s_2[k] \), respectively, and allocate them on the discrete time axis as shown in Figure 4. As shown in the figure, there are \( s_1[k] \) and \( s_2[k] \), partially overlapping on the time axis. Obviously, this is a very difficult situation for FFT to detect the two signals. The conditions are listed in Table 1 and the waveform of \( x[k] \) is shown in Figure 5.

Figure 6 shows \( z_n \) at the \( r \)th slot corresponding to the color bar. The high values of \( z_n \) tend to be shown in red. It is clearly shown that \( s_1[k] \) and \( s_2[k] \) which are intermittently allocated on the time axis depicted in Figure 4 are detected.
Figure 6. $z_n$ obtained by ARS.

### Table 1. Simulation conditions.

| Parameter                          | Value               |
|------------------------------------|---------------------|
| SNR                               | $10\log_{10}(P_1/N)$ |
|                                    | $10\log_{10}(P_2/N)$ |
| Signal power ratio                 | $0 \text{ dB}$      |
| The number of fundamental waveforms | $V_1 = 50, V_2 = 49$ |
| The number of SPC                  | $N = 50$            |
| The number of SPC output ports     | $L_{\text{min}} = 75$ |
| Slot length                        | $Q = 1250 \text{ samples}$ |

5. Experiments

This section provides the experimental results applying ARS to the analysis of machine sounds aiming at non-contact fault detection.

5.1. Settings of the experiment

Figure 7 illustrates the configuration of the experimental system. The sound of the machine is recorded using a microphone connected with a personal computer (PC) via an audio interface. The microphone is installed at 2 m away from the target machine. The sound is sampled at the audio interface and stored in the PC for the analysis using MATLAB. Figure 8 shows the overview while Table 2 lists the equipment in the experimental system.

![Figure 7. A configuration of the experimental system.](image)

![Figure 8. A configuration of the experimental system.](image)

### Table 2. Equipment in the experimental system.

| Equipment            | Model          |
|----------------------|----------------|
| Audio interface      | Steinberg UR12 |
| Sampling frequency   | 8000 [Hz]      |
| Resolution           | 16 [bit]       |
| Microphone           | SHURE PGA48-XLR |
| PC                   | HP ProBook 430 G2 Notebook PC |
| Press machine        | AIDA C2-16     |

### Table 3. Machine conditions at the time of the recordings.

| Day | Conditions          |
|-----|---------------------|
| 1   | Before overhaul     |
| 2   | After overhaul 1    |
| 3   | After overhaul 2    |

As shown in Table 3, the recording of the machine sounds was conducted three times. The first recording was conducted before the overhaul, i.e. the machine was not in a perfect condition at this time. Then, after the overhaul, the second recording was conducted. In this recording, the machine condition was adjusted better. For the verification purpose, one more recording was conducted in the same machine status in another day. The samples of the recorded sounds were analysed by ARS as introduced in Section 4.2 as well as the Short-Time Fourier Transform (STFT) and the wavelet transform (WT). WT employs the three types of wavelets consisting of Morse, Morlet and **Bump [AE-1]** functions. Table 4 lists the analysis conditions of ARS and STFT. As shown in the table, the total number of samples to be analysed is set at 960,000. Since the sampling frequency is set at 8 kHz, it corresponds to the sound during 2 min.
Table 4. Analysis conditions of ARS and STFT.

| Parameter                  | Value          |
|----------------------------|----------------|
| Number of SPC              | \( N = 7000 \) |
| Number of SPC output ports at SPC #1 | \( L_{\text{min}} = 5000 \) |
| Slot length                | \( Q = 80,000 \) |
| Number of slots            | \( R = 12 \)  |
| Total number of samples    | \( K = QR = 960,000 \) |

5.2. Results of the experiment

In total, 15 experimental results are obtained by the analysis using ARS, STFT, and WT. Table 5 clarifies the date of the recording and the analysis method for Figures 9–23. It is observed that the result of the machine before overhaul shown in Figure 9 differs from Figures 10 or 11 which are obtained by ARS. However, we cannot find remarkable differences among the results obtained by STFT shown in Figures 12–14 due to very low frequency resolution. The results of analysis by WT is shown by Figures 15–17 for Morse, by Figures 18–20 for Morlet, by Figures 21–23 for Bump. It is visually shown that WT achieves higher resolution compared with STFT. However, on the frequency axis, it is observed that the resolution is still lower than ARS.

5.3. Evaluations of the experimental results

In order to quantify the similarity of the figures listed in Table 5, we calculated the histogram correlation (HC) [20] as follows: First of all, histograms of the pixel values are derived consisting of 100 bins from the pixel values, as illustrated in Figure 24, of the figures listed in Table 5. The similarity between two figures are measured by the HC \( R \) formulated as follows:

\[
R = \frac{\sum_{h=1}^{H} (\xi_h - \bar{\xi}) (\zeta_h - \bar{\zeta})}{\sqrt{\sum_{h=1}^{H} (\xi_h - \bar{\xi})^2 \sum_{h=1}^{H} (\zeta_h - \bar{\zeta})^2}} \tag{19}
\]

where \( \xi_h \) and \( \zeta_h \) are the value of the \( h \)th \( (h = 1, \ldots, H) \) bin of the histogram corresponding to the compared
two figures while $\bar{\xi}$ and $\bar{\zeta}$ are given as follows:

$$\bar{\xi} = \frac{1}{H} \sum_{h=1}^{H} \xi_h$$  \hspace{1cm} (20)

$$\bar{\zeta} = \frac{1}{H} \sum_{h=1}^{H} \zeta_h$$  \hspace{1cm} (21)

Table 6 lists the HC $R$ obtained by (19). It is obvious that the Figures 10 and 11 obtained by ARS after the overhauling are similar while these two figures are less similar comparing with Figure 9 obtained before the overhaul.
We can visually confirm these relationships looking at Figures 9–11.

However, the results obtained by STFT given as Figures 12–14 are less correlated compared with the results obtained by ARS. It means that ARS detects the difference produced by the overhaul better, compared with STFT.

In Table 6, it is observed that the performance in terms of the HC between WT with Morlet wavelet and ARS are similar. However, it should be noted that ARS achieves higher value in Day 2–3 compared with WT with Morlet wavelet. For realization of the non-contact fault detection, it is required that only the correlation coefficient of the machines in the same condition is high. In addition, as precisely evaluated in Section 6.1, the computational load of ARS is lower than that of WT even though ARS achieves almost the same performance compared with WT.
Table 5. Relationship between figures, dates of recording and signal analysis.

| Figure | Date of recording | Signal analysis |
|--------|------------------|-----------------|
| Figure 9 | Day 1            | ARS             |
| Figure 10 | Day 2           | ARS             |
| Figure 11 | Day 3           | ARS             |
| Figure 12 | Day 1           | STFT            |
| Figure 13 | Day 2           | STFT            |
| Figure 14 | Day 3           | STFT            |
| Figure 15 | Day 1           | Morse WT        |
| Figure 16 | Day 2           | Morse WT        |
| Figure 17 | Day 3           | Morse WT        |
| Figure 18 | Day 1           | Morlet WT       |
| Figure 19 | Day 2           | Morlet WT       |
| Figure 20 | Day 3           | Morlet WT       |
| Figure 21 | Day 1           | Bump WT         |
| Figure 22 | Day 2           | Bump WT         |
| Figure 23 | Day 3           | Bump WT         |

Table 6. A correlation coefficients between the results of each method.

|            | ARS     | STFT    | Morse WT | Morlet WT | Bump WT |
|------------|---------|---------|----------|-----------|---------|
| Day 1–2    | 0.8951  | 0.6223  | 0.9458   | 0.8885    | 0.9124  |
| Day 1–3    | 0.8837  | 0.7672  | 0.7975   | 0.8959    | 0.8629  |
| Day 2–3    | 0.9967  | 0.5930  | 0.8951   | 0.9378    | 0.8620  |

6. Discussions

6.1. Computational loads

Through the evaluation of the experimental results, it was clarified that WT using Morlet wavelet achieves almost equivalent performance in terms of HC, compared with ARS. In this section, the computational complexity is compared between the two methods.

6.1.1. ARS

The number of additions of ARS is calculated by (5) and (8) as follows:

\[
\theta^{(ARS)}_{add} = \sum_{n=1}^{N} G_n M_n = \sum_{n=1}^{N} \left( \frac{K}{L_{\text{min}} + n - 1} \right) (L_{\text{min}} + n - 1) \quad (22)
\]

Note that \( K \) is set at \( Q \) in this case since ARS is applied to one slot.

Equation (10) contains a division. However, this is interpreted as a multiplication of a constant \( 1/G_n \). Therefore, the number of multiplications contained in ARS \( \theta^{(ARS)}_{mul} \) is equal to the number of SPC, \( N \).

In addition, it should be noted that these additions and multiplications are repeated for \( R = 12 \) slots as specified in Table 4. Recall that one slot consists of \( Q = 80,000 \) samples.

Eventually, the total number of the additions \( \theta^{(ARS)}_{add} \) and the multiplications \( \theta^{(ARS)}_{mul} \) needed for an analysis to obtain a result such like Figure 9 is calculated as follows:

\[
\theta^{(ARS)}_{add} = R \theta^{(ARS)}_{add} = 12 \sum_{n=1}^{7000} \left( \frac{80,000}{5000 + n - 1} \right) (5000 + n - 1) = 6,353,674,896 \quad (23)
\]

\[
\theta^{(ARS)}_{mul} = R \theta^{(ARS)}_{mul} = 12 \times 7000 = 84,000 \quad (24)
\]

6.1.2. WT

In WT, the bandwidth ranging from 0.6667 to 1.6000 Hz is divided into 13 subbands. The bandwidth is corresponding to ARS SPCs as follows: The SPC equipped with the minimum number of output ports is 7000. Since the sampling frequency is 8 kHz, the frequency is given as \( 8000/7000 = 0.6667 \) Hz. Likewise, The SPC equipped with the maximum number of output ports is 12,000. The corresponding frequency is given as \( 8000/12,000 = 1.6000 \) Hz.

The Morlet wavelet is given as follows [21]:

\[
\psi (t) = \frac{1}{\sqrt{\pi}} e^{2\pi f_C t} e^{-\frac{t^2}{2}} \quad (25)
\]

where \( f_C \) is the centre frequency.

This is digitized by the sampling period \( T_S \) and the discrete time index \( k \) as follows:

\[
\psi [k] = \psi (kT_S) = \frac{1}{\sqrt{\pi}} e^{2\pi f_C T_S} e^{-\frac{(kT_S)^2}{2}} \quad (26)
\]

By using \( f_C \), the bandwidth from 0.6667 to 1.6000 Hz is divided into 2 octaves with 10 voices, i.e. 1 octave is divided into 2 subbands as shown in Table 7. It is shown that the target bandwidth is divided into 13 subbands.

By increasing the number of the subbands, the frequency resolution is improved even though the increase causes huge computational load. In each subbands, the input sample sequence \( x[k] \) is convolved with an Morlet wavelet. It is well-known that the computational load...
of the convolution is drastically reduced by converting it to multiplications in the frequency domain through FFT [22]. The procedure is given as follows:

1. To apply FFT, the length of $x[k]$ and the wavelet $\psi[k]$ are adjusted by appending zeros to be $2^\alpha$ where $\alpha$ is given as follows:

   $$\alpha = \lceil \log_2 K \rceil$$  

   Note that $[a]$ denotes the minimum integer which is larger than or equal to $a$.

2. Perform FFT for $x[k]$ and for the wavelet $\psi[k]$ to obtain $X[f]$ and $\Psi[f]$, respectively, where $f = 0, \ldots, \alpha - 1$ is a frequency index.

3. Calculate the product for all $f$ as follows:

   $$Y[f] = X[f] \Psi[f]$$  

4. Perform the inverse FFT (IFFT) to recover the time-domain signal $y[k]$ from $Y[f]$.

Therefore, WT requires 2 FFTs in addition to $K$ multiplications and an IFFT for each of the subbands. It is also known that an FFT contains $(2^\alpha/2) \log_2 2^\alpha = \alpha 2^{\alpha - 1}$ complex multiplications and $2^\alpha \log_2 2^\alpha = \alpha 2^\alpha$ complex additions [23], while IFFT contains $(2^\alpha/2) \log_2 2^\alpha + (2^\alpha/2) = 2^{\alpha - 1}(\alpha + 1)$ complex multiplications and the same number of the complex additions in FFT [24]. Therefore, the total number of the complex multiplications is obtained as follows:

$$\alpha = \lceil \log_2 960,000 \rceil = 20$$

$$\Theta_{\text{mul}}^{\text{FFT}} = 20 \times 2^{19} = 10,485,760$$

$$\Theta_{\text{mul}}^{\text{IFFT}} = 21 \times 2^{19} = 11,010,048$$

$$\Theta_{\text{mul}}^{\text{WT}} = 13 \left( \Theta_{\text{mul}}^{\text{FFT}} + \Theta_{\text{mul}}^{\text{IFFT}} + 2^\alpha \right)$$

$$= 429,391,872$$  

where $\Theta_{\text{mul}}^{\text{FFT}}$ and $\Theta_{\text{mul}}^{\text{IFFT}}$ denote the number of complex multiplications in FFT and IFFT, respectively. Note that a complex multiplication contains four multiplications and the same number of the complex additions in FFT [24]. Therefore, the total number of the complex multiplications is obtained as follows:

$$\Theta_{\text{add}}^{\text{add}} = 20 \times 2^{20} = 20,971,520$$

$$\Theta_{\text{add}}^{\text{add}} = \Theta_{\text{add}}^{\text{add}}$$

$$\Theta_{\text{add}}^{\text{WT}} = 13 \times \left( 2\Theta_{\text{add}}^{\text{add}} + \Theta_{\text{add}}^{\text{add}} \right) = 817,889,280$$  

where $\Theta_{\text{add}}^{\text{add}}$ and $\Theta_{\text{add}}^{\text{add}}$ denote the number of complex additions in FFT and IFFT, respectively. Note that a complex addition contains two additions. In addition, one complex multiplication yields two additions. Eventually, the total number of the additions is calculated as follows:

$$2\Theta_{\text{add}}^{\text{add}} + 2\Theta_{\text{add}}^{\text{add}} = 2 \left( 817,889,280 + 429,391,872 \right)$$

$$= 2,494,562,304$$  

### 6.1.3. Comparison

Table 8 compares the number of multiplications and additions between ARS and WT.

According to the table, it is verified that ARS reduces the number of multiplications 1/20,000 even though it needs 2.5 times more additions compared with WT. Now, the numbers of the multiplications and the additions listed in Table 8 are evaluated. Each of the multiplications is implemented by a barrel shifter and an adder, and realized by a repetition of the additions [25]. In case of an $N_B$-bit multiplication, $N_B$ additions are required as the maximum. As the minimum, only one addition is enough. Now, for the evaluation, it is assumed that $N_B/2$ additions are required for a multiplication. So the total numbers of the $N_B$-bit addition for ARS and WT are formulated as follows:

$$\Delta_{\text{ARS}} = 84,000 \times \frac{N_B}{2} + 6,353,674,896$$  

$$\Delta_{\text{WT}} = 717,567,488 \times \frac{N_B}{2} + 2,494,562,304$$  

Figure 25 shows $\Delta_{\text{WT}} / \Delta_{\text{ARS}}$ versus $N_B$ (bit). According to this figure, $\Delta_{\text{WT}} / \Delta_{\text{ARS}} = 2.014$ at $N_B = 12$ bits. It means that ARS reduces the half of the computational load of WT with Morlet wavelet at 12-bit resolution.

| Multiplications | Additions  |
|-----------------|------------|
| ARS             | 84,000     | 6,353,674,896 |
| WT              | 1,717,567,488 | 2,494,562,304 |

---

K. TSUJI ET AL.

---

**Figure 25.** Histogram generation from two figures.
By using the double precision with 64-bit calculations, $\Delta_{WT}/\Delta_{ARS} = 9.039$. It is shown that ARS reduces the computational load compared with WT with Morlet as a function of the bit length.

7. Conclusion

This paper focused in very low frequency band of the press machine sound captured in a real factory toward the predictive maintenance by IoT. The sounds were recorded several times, before and after the overhaul. The analysis showed that ARS and WT with Morlet wavelet are superior to other methods including STFT and WT with Morse and Bump wavelets, in terms of the histogram correlation. In addition, it was clearly shown that ARS costs significantly low computational load compared with WT with Morlet wavelet to achieve approximately the same performance.

Further researches include implementations of sensor nodes containing an ARS processor aiming to realize edge computing with long battery life taking advantage of its simplicity.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Kazuki Tsuji is a student at Graduate School of Computer Science and Technology, Aichi Prefectural University, pursuing his master’s degree. His research interests include digital signal processing and its application for vital sensing techniques.

Shota Imai is a student at Graduate School of Computer Science and Technology, Aichi Prefectural University, pursuing his master’s degree. His research interests include digital signal processing for IoT systems.

Ryota Takao is a student at Graduate School of Computer Science and Technology, Aichi Prefectural University, pursuing his master’s degree. His research interests include digital signal processing and its application for various sensing techniques.

Tomonori Kimura joined Symantec corp. in 2007 as an Account Manager. Through EMC Japan Co., Ltd., he is currently a head of sales department and information technologies at Cosmotec Co. Ltd.

Hitoshi Kondo joined Aida Engineering Ltd. in 1971. Currently he is a senior corporate executive officer at Cosmotec Co. Ltd.

Yukihiro Kamiya received MS and PhD degrees in information engineering and information electronics from Nagoya University in 1994 and 2000, respectively. He joined Kokusai Denshin Denwa (KDD) Co., Ltd. in 1994. From October 1998 to February 2001, he was a researcher at ATR Adaptive Communications Research Laboratories. He was a post-doctoral research fellow at Ecole Nationale Superieure d’Ingenieurs de Constructions Aeronautiques (ENSICA), Toulouse, France in 2001, then he was with PCOM:13 ApS, Aalborg, Denmark as a project manager till 2003. He is currently an associate professor at Aichi Prefectural University. His interests include digital signal processing applicable to IoT systems. He is a member of IEEE.

References

[1] Fernandez-Carames TM, Fraga-Lamas P. A review on human-centered IoT-connected smart labels for the industry 4.0. IEEE Access. 2018;6:25939–25957.
[2] Chen B, Wan J, Shu L, et al. Smart factory of industry 4.0: key technologies, application case, and challenges. IEEE Access. 2018;6:6505–6519.
[3] Canizo M, Conde A, Charramendieta S, et al. Implementation of a large-scale platform for cyber-physical system real-time monitoring. IEEE Access. 2019;7:52455–52466.
[4] Yan J, Meng Y, Lu L, et al. Industrial big data in an industry 4.0 environment: challenges, schemes, and applications for predictive maintenance. IEEE Access. 2017;5:23484–23491.
[5] Wu Z, Luo H, Yang Y, et al. K-PdM: KPI-oriented machinery deterioration estimation framework for predictive maintenance using cluster-based hidden Markov model. IEEE Access. 2018;6:41676–41687.
[6] Zhou W, Habetler TG, Harley RG. Bearing condition monitoring methods for electric machines: a general review. 2007 IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives; Cracow; 2007. p. 3–6.
[7] Henriquez P, Alonso JB, Ferrer MA, Travieso CM. Review of automatic fault diagnosis systems using audio and vibration signals. IEEE Trans Syst Man Cybern Syst. 2014 May;44(5):642–652.
[8] Nishino T, Takeuchi S, Saito T, Watanabe I. Fault sound simulations from normal sounds for data-driven prognosis based on human expert and vibration knowledge. Proceedings of 2017 IEEE International Conference on Prognostics and Health Management; 2017. p. 266–272.
[9] Chong U, Lee S, Sohn C. Fault diagnosis of the machines in power plants using LPC. Proceedings of The 8th Russian-Korean International Symposium on Science and Technology, Vol. 1; 2004. p. 170–174.
[10] Xia M, Li T, Xu L, et al. Fault diagnosis for rotating machinery using multiple sensors and convolutional
neural networks. IEEE ASME Trans Mechatron. 2018 Feb;23(1):101–110.

[11] Wei Y, Li Y, Xu M, et al. Intelligent fault diagnosis of rotating machinery using ICD and generalized composite multi-scale fuzzy entropy. IEEE Access. 2019;7:38983–38995.

[12] Wang W. An intelligent system for machinery condition monitoring. IEEE Trans Fuzzy Syst. 2008 Feb;16(1):110–122.

[13] Smith WW. Handbook of real-time fast Fourier transforms: algorithms to product testing. New York: Wiley; 1995.

[14] Kamiya Y. A simple parameter estimation method for periodic signals applicable to viral sensing using Doppler sensors. SICE J Control Meas Syst Integr. 2017 Sept;10(5):378–384.

[15] Song L, Wang H, Chen P. Vibration-based intelligent fault diagnosis for roller bearings in low-speed rotating machinery. IEEE Trans Instrum Meas. 2018 Aug;67(8):1887–1899.

[16] Kahaei MH, Torbatian M, Poshtan J. Detection of bearing faults using Haar wavelets. IEICE Trans Fund. 2006 Mar;E89-A(3):757–763.

[17] Pang B, Tang G, Tian T. Complex singular spectrum decomposition and its application to rotating machinery fault diagnosis. IEEE Access. 2019;7:143921–143934.

[18] Huda F, Anggriawan A, Rusli M. Meifal: the using of sound signal and simple microphone to detect damages in induction motor. IOP Conf Ser Mater Sci Eng. 2019;539:012034. DOI:10.1088/1757-899X/539/1/012034.

[19] Song J, Kim S, Liu Z, et al. A real time nondestructive crack detection system for the automotive stamping process. IEEE Trans Instrum Meas. 2016 Nov;65(11):2434–2441.

[20] Imamura K, Kuroda H, Fujimura M. Origin image preparation for search method for digital watermarking image using correlation coefficient between pixel value histograms. IPSJ SIG Tech Rep. 2012 Sept;2012-AVM-78(9).

[21] Addison PS. Introduction to redundancy rules: the continuous wavelet transform comes of age. Philos Trans R Soc A. 2018. DOI:10.1098/rsta.2017.0258.

[22] Zhang Z, Nakahori T, Kawabata H. Fast wavelet transform and its application to electroencephalogram. J Acoust Soc Jpn. 1999 May;65(633):181–187.

[23] Hayes MH. Statistical digital signal processing and modeling. New York: Wiley; 1996.

[24] Salih SM. Fourier transform – signal processing. New York: inTech; 2012.

[25] Patterson DA, Hennessy JL. Computer organization and design. New York: Morgan Kaufmann; 2016.