Sound source localisation and signal extraction with multiple microphone arrays

Chaoying Bao (1,2), Ling Jia (2), Brendan Coral (2), David Mathews (1,2), Hongmei Sun (2) and Jie Pan (2)

(1) Defence Science and Technology, Australia
(2) Department of Mechanical Engineering, University of Western Australia, WA, Australia

ABSTRACT

The idea of applying the microphone array technology for machine condition monitoring of a mine site has been investigated at the University of Western Australia. As a key part of the investigation, this paper focuses on the sound source localisation and signal extraction using the beamforming technique. Instead of a single large array, multiple small arrays are used, as practically, they can be easily deployed. In terms of array processing, they are normally processed as independent arrays. However, with synchronised sampling they can also be processed as a single combined array. These two processing approaches are termed the independent array approach (IAA) and the synchronised array approach (SAA), respectively. The main advantage of SAA is its greater array gain and higher beam resolution. However, the arrays need to be placed close to one another to avoid producing aliasing lobes. The IAA on the other hand has no such restrictions and the arrays can be placed independently in any position. This may provide gains in signal to interference and noise ratio (SINR) for individual sound sources in certain circumstances. In order to compare the performance of the two approaches under various conditions, numerical simulations are carried out. The performances are evaluated in terms of the accuracy of source localisation and the fidelity of signal extraction. The general guideline for using the two approaches is provided.

1 INTRODUCTION

The ability to anticipate failure of equipment is a key area of concern for mining operators. Equipment maintenance costs range from 20% to over 35% of total operating costs (Dhillon, 2008). The goal of machine condition monitoring is to repair or replace the devices before any significant interruptions to production. These devices may include pumps, generators, compressors and more. All mechanical systems generate noise and the sound radiated can indicate the condition and health of the device (Ravetta, Muract, & Burdisso, 2007). Therefore, the ability to incorporate microphone arrays to detect when a piece of equipment is approaching failure will provide significant benefits to industry.

A microphone array comprises a set of microphones that are positioned in such a way that spatial information can be captured for processing (Benesty, Chen, & Huang, 2008). Noise generated from a number of locations can make isolating specific sources problematic. Therefore, the objective of the signal processing aspect is to estimate the parameters required for extracting the signal of interest (Benesty, Chen, & Huang, 2008). There are different techniques available to achieve this. One of these is beamforming and is the technique used in this paper.

In general, the number and geometry of microphone arrays play a role, depending on the nature of the application. For the application of machine condition monitoring of a mine site, we adopt the multiple small array strategy, as it is easier for array deployment and provides better flexibility and reliability. There are two approaches in terms of array processing. In one approach, beamforming takes place independently for each small array. It is termed the independent array approach (IAA). In the other, with synchronised sampling the small arrays are combined to form a single larger array for beamforming. It is termed the synchronised array approach (SAA). The main advantage of SAA is its greater array gain and higher beam resolution. However, the arrays need to be placed close to one another to avoid producing aliasing lobes. The IAA on the other hand has no such restrictions and the arrays can be placed independently in any position. This may provide gains in signal to interference and noise ratio (SINR) for individual sound sources in certain circumstances. In order to compare the performance of the two approaches under various conditions, numerical simulations are carried out. The aim of this paper is to obtain the general guidelines for using the two approaches through a simulation study.

This paper is organised as follows. In Section 2, we divide the beamforming task into two phases and explain the beamforming algorithms employed for those two phases, respectively. In Section 3, we first define the environment and parameters of the numerical simulation. We then present the simulation results and thereby obtain the general guidelines for using the SAA and IAA as well as a rule of thumb for the placement of arrays for both approaches. We conclude the paper in Section 4.

2 BEAMFORMING ALGORITHMS

We assume two dimensional beamforming to be adequate for the current stage of the application. In order to achieve higher beam resolution with a small number of sensors, we choose a uniform linear array (ULA) as the basic form of the multiple arrays. Although a single ULA suffers from right/left ambiguities, the problem can be mitigated with the use of multiple ULAs.

In our investigation, we divide the beamforming task for condition monitoring into two phases. In Phase 1 we scan the whole area of a selected site to find exact locations of sound sources of interest. This phase is termed the source localisation in this paper. In Phase 2 we extract a signal of interest in time series form by beamforming only to the source location which is obtained in Phase 1. This phase is termed the signal extraction.

2.1 Beamforming algorithm for source localisation

The source localisation is achieved by examining the intensity map of the scanning area obtained through beamforming. In this phase, the beamforming output required is in the form of intensity. To this end, frequency domain adaptive beamforming (ABF) algorithms are a better choice because of their superior performance. In this study, the well-known algorithm of MVDR (Minimum Variance Distortionless Response) with diagonal loading is used (Van Trees, 2002).

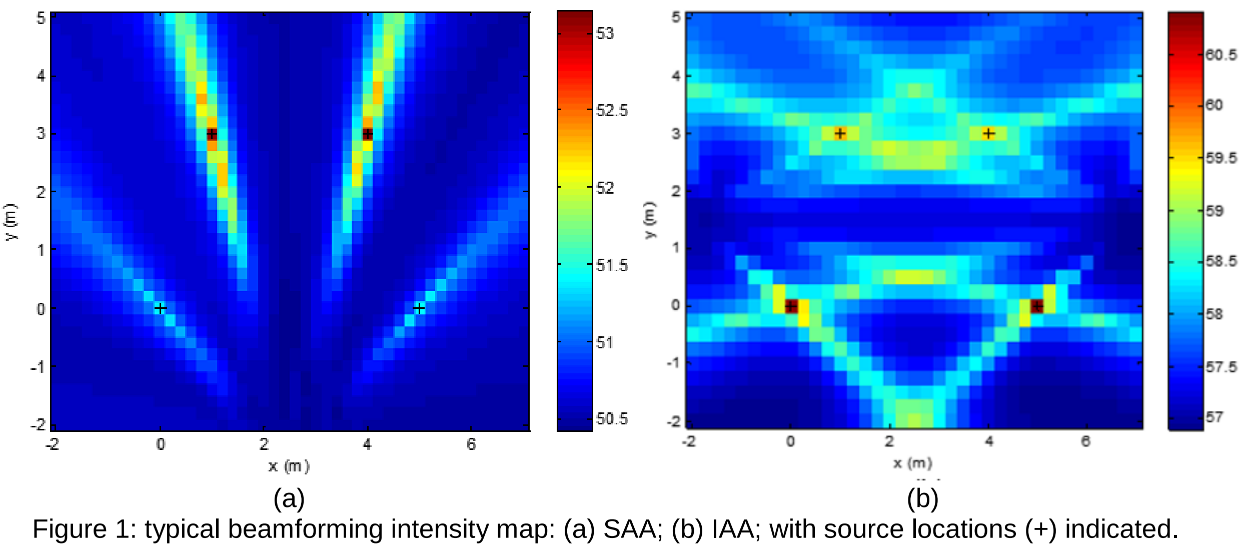

Figure 1: typical beamforming intensity map: (a) SAA; (b) IAA; with source locations (+) indicated.

For IAA, each of the independent arrays produces its own intensity map and the final intensity map is obtained by simply adding the corresponding intensities from those individual maps. Figure 1(a) and 1(b) show a typical beamforming intensity map for SAA and IAA, respectively, where there are four sources in the area and two arrays are used. It can be clearly seen from the figures that simply sorting out the four global maxima in the intensity map does not give the correct answer for the source locations. This is because the beam paths that lead to different sources often have different strengths, as shown in Figure 1(a) where the strengths of the two middle beam paths are stronger than those of the two side ones. Thus the intensity levels of many points on the stronger beam path will be greater than that of the correct point on the weaker beam path. Fortunately for an application of machine condition monitoring, the location of a sound sources is roughly known, as the exact location of a piece of machinery is known. It should be pointed out that, however, there are situations where the centre of the noise may differ to that of the machine. Therefore, sorting out the local maximum in the neighbourhood of the location of a piece of machinery is the mechanism used in the source localisation.

2.2 Beamforming algorithm for signal extraction

The beamforming requirements in the signal extraction phase are different to those in the source localisation phase. Firstly, there is no need to beamform the whole area but only the estimated location of a source. Secondly, the required beamforming output is a waveform or a signal in the time domain. There are many adaptive beamforming algorithms that can fulfil the second requirement. Depending on what domain that beamforming is carried out, they can be categorised in two classes. One class is time domain beamforming where the whole process including the beamforming part takes place in the time domain. The other is time-frequency domain beamforming (TFDBF) where beamforming is carried out in the frequency domain. In the previous studies (Bao, 2005, Bao, 2014), the performances of several ABF algorithms from both classes were examined and evaluated. According to the studies, TFDBF is a better option for this application where a certain amount of latency can be tolerated.

TFDBF begins with time series data. This time domain data is Fourier transformed to the frequency domain and processed by a frequency domain beamformer. The output of that process is then inverse-Fourier transformed back to the time domain for a waveform output. Figure 2 shows the schematic representation of TFDBF, where FFT stands for Fast Fourier Transform and IFFT for Inverse-FFT.

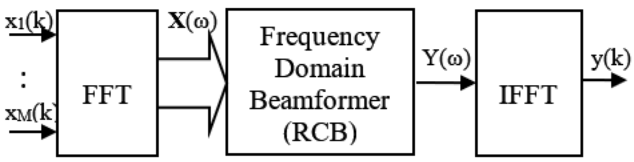

Figure 2: Schematic representation of TFDBF.

The algorithm used in TFDBF is the Robust Capon Beamforming (RCB) algorithm developed by Li etc. (Li, Stoica & Wang, 2003), a type of diagonally loaded MVDR algorithm. It should be noted that the original purpose of the RCB algorithm was to make it robust to errors in the signal model, and the appropriate value for the loading is determined by the anticipated signal mismatch. For a waveform application, however, loading is introduced for another purpose (Bao, 2005). Some degree of loading is always required even in the absence of any signal mismatch.

It should be noted that, for IAA, beamforming produces as many signals as the number of independent arrays for each source. Methods for optimally combining these signals require further investigation. It is assumed for now that the array closest to the source produces the cleanest signal, and the others are discarded.

3 NUMERICAL SIMULATION

The aim of the simulation is to compare the performance of the SAA and IAA, and thereby obtain the general guidelines for using them. To this end, we would like to keep the complexity of the environment and parameters of the simulation minimal so long as it does not distort the comparison and subsequently the conclusion obtained.

In the simulation, two ULAs consisting of eight omnidirectional microphones with an inter-element spacing of 0.2 meters are used. The sound speed is 340 m/s. Thus, the arrays cover a frequency range from 100 to 850 Hz. The sampling frequency is 8000 Hz.

Four incoherent monopole sound sources are assumed in a free field and located in an area of 40 meters x 3 meters. It should be noted that ignoring the ground reflection in the simulation has a negligible effect on the performance comparison. Each source emits a different narrow band signal consisting of two frequencies, one around 100 Hz and the other around 800 Hz, covering both ends of the frequency range. Independent and identically distributed (IID) random noise is also added at each sensor to simulate all other noises in the field. A reference microphone is placed at the geometric centre of the four sources to adjust the power of IID noise for both approaches in order to make the comparison valid. The signal to noise ratio (SNR) listed in the simulation is defined as the sum of all signal powers to the power of IID noise at the reference microphone, unless mentioned otherwise. The signal power ratio (SPR) which specifies the intensity of each source is referenced to unity.

Three typical scenarios in terms of source location configurations are considered in the simulation. In Scenario 1, the four sources are located close to each other in an area of 4 meters x 3 meters. In Scenario 2, the four sources are divided into two groups. Each group consists of two sources located close to each other, and the two groups are separated by a large distance of about 40 meters. In Scenario 3, the four sources are divided into three groups. One of the groups consists of two sources located close to each other. The spacing of these three groups is about 20 meters. Table 1 lists the coordinates of the sources and other simulation parameters of the three scenarios.

Table 1: Parameters of 3 scenarios

	Scenario 1	Scenario 2	Scenario 3
Source coordinate (m) (x, y)	(0,0) (1,3) (3,3) (4,0)	(0,0) (1,3) (39,3) (40,0)	(0,0) (19,3) (21,3) (40,0)
Signal frequency (Hz) (f_1, f_2)	(110,800) (120,810) (130,820) (140,830)	(110,800) (120,810) (130,820) (140,830)	(110,800) (120,810) (130,820) (140,830)
SPR (dB)	0,0,0,0	0,0,0,0	0,0,0,0
SNR (dB)	0	0	0

In a real life application, there may be restrictions on array placement set up by mining operators. These restrictions may affect the performances of both source localisation and signal extraction. Without any prior knowledge, the only restriction applied in the simulation is that the arrays should be placed at least 2 meters away from any sound sources to be monitored.

All results presented in this section are obtained from the average of 200 independent runs. The phase relationships among the signals of the four sources and also among the frequencies within a signal are randomly selected in each independent run. The IID noise added is also randomly generated in each run.

3.1 Results of source localisation

The performance of source localisation is evaluated in terms of a metric of root mean square error (RMSE) defined as

$$\mathrm{RMSE} \triangleq \sqrt{\frac{1}{N}\sum_{n=1}^{N}(\mathrm{x}_s - \hat{x}_s(n))^2 + (\mathrm{y}_s - \hat{y}_s(n))^2} \tag{1}$$

where x_s is the x coordinate of the source location, $\hat{x}_s(n)$ the estimate of x_s from the *n*th independent run, y_s the y coordinate of the source location, $\hat{y}_s(n)$ the estimate of y_s from the *n*th run, and *N* is the number of independent runs.

One of the factors affecting the accuracy of source localisation is the array placement. In the simulation, the optimal array placement for each approach and each scenario is obtained manually by trial-and-error. Table 2 shows the results from the source localisation.

Table 2: Results of source localisation

	RMSE for each source (m)	
	SAA	IAA
Scenario 1	0, 0, 0, 0	0, 0, 0, 0
Scenario 2	0, 0, 0, 0	0, 0, 0, 0
Scenario 3	0, 0, 0, 0	0.9, 0, 0, 0.8

It can be seen from Table 2 that for Scenarios 1 and 2 both SAA and IAA are capable of estimating all four source locations accurately without any error. As for Scenario 3 where there are three separated groups of sources, the SAA still works with 100% accuracy, whereas the IAA sometimes makes errors in estimating the locations of Sources 1 and 4. The reason for the reduced performance of IAA in Scenario 3 can be explained as follows. As mentioned in Section 1, the SAA has a better array gain and beam resolution. The IAA can compensate this shortfall by moving the arrays close to the sources. However, there are three groups of sources and only two arrays in Scenario 3. Thus, a trade-off is needed. As it is more important to accurately estimate the locations of the two sources located close to each other for the sake of signal extraction, the arrays need to be placed close to these two sources. As a result, the input signal to interference and noise ratios (SINR) at the arrays from the two side sources are not high enough for obtaining the correct estimation.

3.2 Results of signal extraction

The performance of signal extraction is evaluated in terms of a metric of coherence. The coherence γ is based on the correlation coefficient, defined in a way that accounts for a time delay between the signal *s(t)* and the signal extraction output *y(t)*,

$$\gamma = \max_{\tau} \frac{|\mathrm{COV}\{s(t)y(t-\tau)\}|}{\sqrt{\mathrm{COV}\{s^2(t)\}\mathrm{COV}\{y^2(t-\tau)\}}} . \tag{2}$$

Here COV{} denotes the covariance. Note that γ has a value between 0 and 1, with value 0 indicating no correlation between the signal and the output of signal extraction, and value 1 indicating that the two waveforms are proportionally identical.

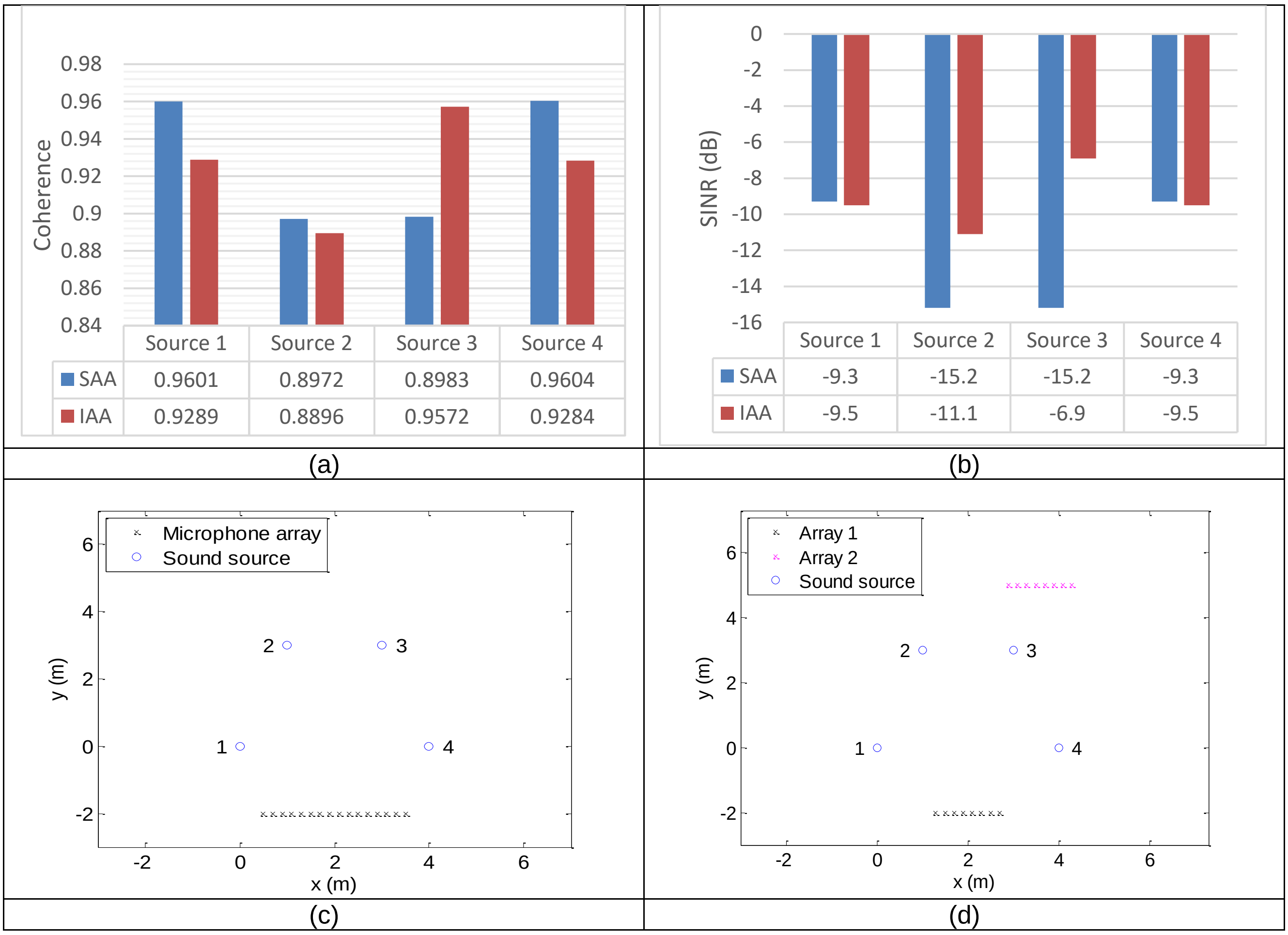

Figure 3: Results of signal extraction for Scenario 1: (a) Coherence; (b) SINR; (c) Array placement for SAA; (d) Array placement for IAA.

The estimated source locations are used in TFDBF for signal extraction. In the case where there are estimation errors, the best estimate and the worst estimate from the 200 independent runs are used to produce two sets of coherence.

Figure 3 shows the results of signal extraction for Scenario 1, where Figure 3(a) compares the coherences of the two approaches, Figure 3(b) shows input SINR from each source at the array where the coherences are obtained for both approaches, and Figure 3(c) and 3(d) illustrate the array placement for SAA and IAA relative to the source locations, respectively. Several observations can be made. Firstly, from Figure 3(a), the coherences achieved by SAA for all sources except for Source 3 are higher than those of IAA. This can be explained by the advantage of SAA in the array gain and beam resolution. Secondly, for SAA, the coherences achieved for Sources 1 and 4 are noticeably higher than those of Sources 2 and 3. This can be explained as follows. It can be seen from Figure 3(c) that the distances from the array to Sources 1 and 4 are shorter than those to Sources 2 and 3. As we know, for a monopole source in the free field the power of the source will be attenuated 6 dB per double the distance. Thus, the SINRs from Sources 2 and 3 at the array are lower than those from Sources 1 and 4 as shown in Figure 3(b), resulting in the lower coherences for the former. Thirdly, the coherence of IAA for Source 3 is markedly higher than that of SAA. This can also be explained by the power distance relationship mentioned above. It can be seen from Figure 3(c) and 3(d) that the distance from Source 3 to Array 2 of IAA is much shorter than that to the array in SAA, resulting in the SINR of the former is about 8 dB higher than that of the latter, as shown in Figure 3(b) for Source 3. Although the array gain of SAA is higher than that of IAA, it is not high enough to compensate the bigger loss in SINR in this case. Despite this individual poor performance, on balance, the overall performance of SAA is better than that of IAA in this scenario.

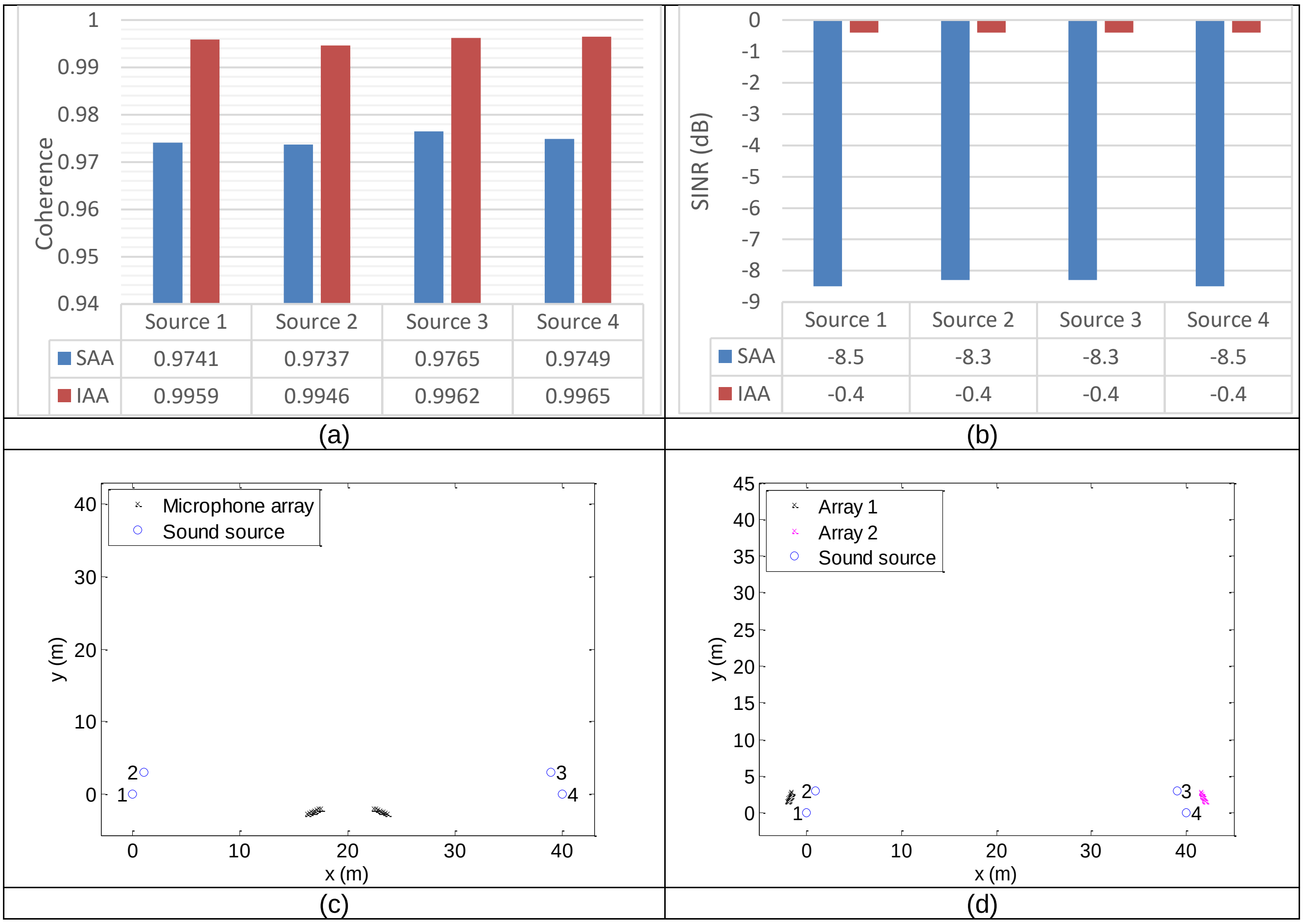

Figure 4: Results of signal extraction for Scenario 2: (a) Coherence; (b) SINR; (c) Array placement for SAA; (d) Array placement for IAA.

Figure 4 shows the results of signal extraction for Scenario 2 where the four sources are divided into two groups about 40 meters apart. It can be seen from Figure 4(a) that the coherences achieved by IAA for the four sources are all around 0.995 and markedly higher than those of SAA. This significant better performance of IAA is due to the fact that for IAA the two arrays can be placed independently in any position (apart from the restriction set up by mining operators). Thus, one array is placed close to one group of sources and the other one close to the other group, as shown in Figure 4(d). This results in higher SINRs at the array for the respective sources, and consequently higher coherences. Whereas for SAA the two arrays need to be placed close to each other to avoid producing aliasing lobes. Not favouring any particular source, the array is placed in the middle location from all sources, as shown in Figure 4(c). As a result, the SINRs of SAA are about 8 dB lower than those of IAA, as shown in Figure 4(b). Although the array gain of SAA is higher than that of IAA, it is not high enough to compensate for the loss in SINR in this case. The performance of IAA in this scenario is clearly better than that of SAA.

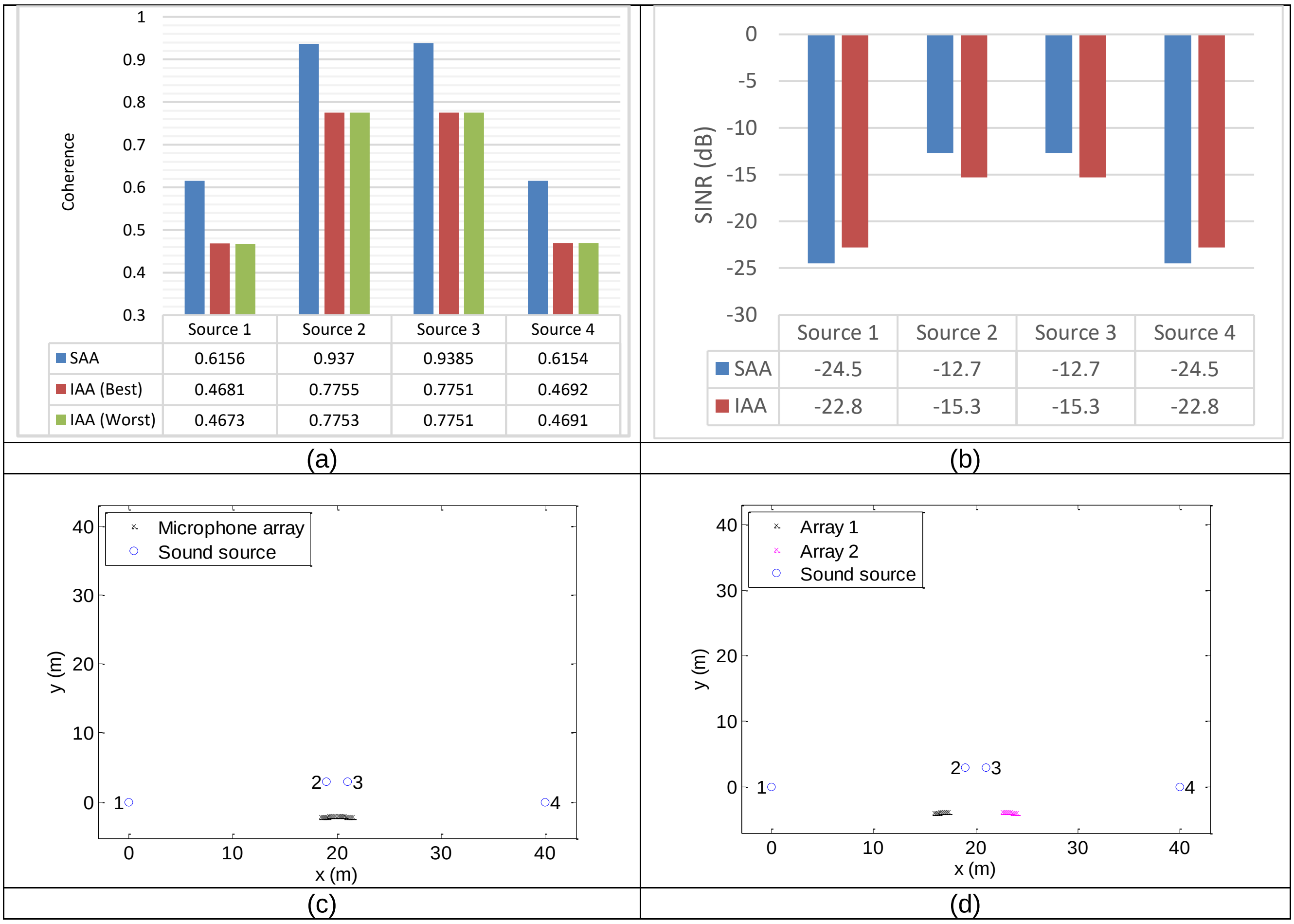

Figure 5: Results of signal extraction for Scenario 3: (a) Coherence; (b) SINR; (c) Array placement for SAA; (d) Array placement for IAA.

Figure 5 shows the results of signal extraction for Scenario 3 where the four sources are divided into three groups. Because there are errors for source localisation for IAA in this scenario, two sets of coherence for IAA are shown in Figure 5(a). One is for the best case scenario where the best estimate of source location (which is the actual source location in this case) is used for signal extraction. The other is for the worst case scenario where the worst estimate is used. Figure 6 shows the beamforming map of IAA with the indicators of actual source locations and their worst estimates. Several observations can be made. Firstly, as shown in Figure 5(a), the coherences of IAA for best case and worst case are very similar. Only the coherences for Sources 1 and 4 are in consideration, as the errors arise only for those two sources. As can be seen from Figure 6, although there is an error of approximately 1.5 meters in the estimation, the estimates are still on the beam paths of the corresponding sources. As beamforming is insensitive along the beam path and there are no other interference sources nearby, the beam-forming with the error estimates still achieves a very similar result to that with the accurate estimate. Secondly, the coherences achieved by SAA for all four sources are noticeably higher than those by IAA, thanks to the advantage of SAA in the array gain and beam resolution. Thirdly, the coherences of both SAA and IAA for Sources 1 and 4 are quite low, as shown in Figure 5(a). This can be explained as follows. As it is more important for signal extraction to accurately estimate the locations of Sources 2 and 3 which are located close to each other (as illustrated evidently in Figure 6), the arrays need to be placed close to these two sources, as shown in Figures 5(c) and 5(d). As a result, the SINRs at the arrays from Sources 1 and 4 become very low, less than -22 dB, as shown in Figure 5(b). The array gain of beamforming is not high enough to overcome such low SINRS to achieve higher coherences. Furthermore, these low SINRs are almost entirely due to the strength of IID noise in the system which beamforming is less effective at attenuating. Table 3 compares SNRs, which are purely due to the IID noise, with SINRs for Sources 1 and 4. These are the factors causing very low coherences. It should be noted that with such low coherences the signals obtained by signal extraction may be inadequate for the purpose of condition monitoring. In order to achieve a better coherence, SINR needs to be increased. This might only be achievable by adding another array. With only two arrays, the performance of SAA is better than that of IAA.

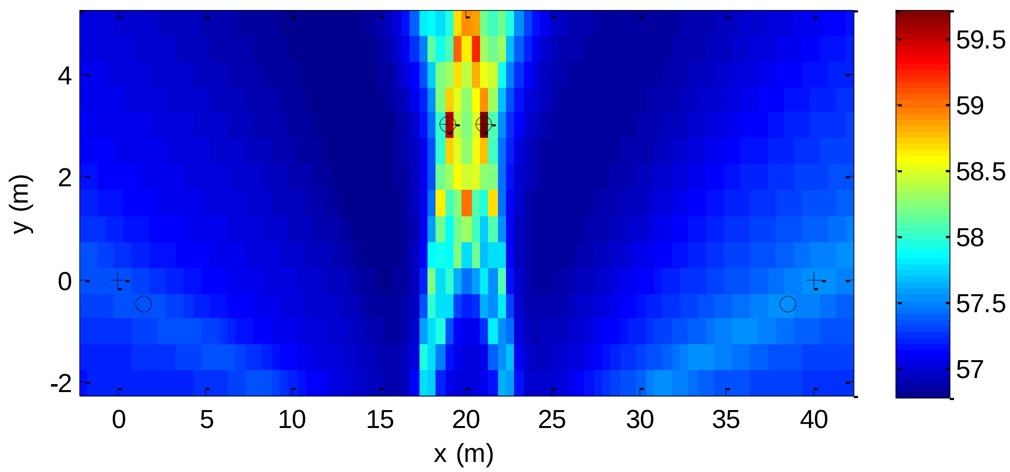

Figure 6: Beamforming map of IAA with indicators of actual source locations (+) and the worst estimates (O).

Table 3: Comparison of SINR and SNR at the array for Sources 1 and 4 in Scenario 3

	Source 1		Source 4	
	SAA	IAA	SAA	IAA
SINR (dB)	-24.5	-22.8	-24.5	-22.8
SNR (dB)	-24.0	-22.6	-24.0	-22.6

3.3 Guidelines

From the simulation study, the following general guidelines for using the two approaches may be provided.

- For the case of a single group of closely placed sources, SAA should be used.
- For the case of multiple groups of sources, if the number of groups is equal to the number of arrays, IAA should be used, and if the number of groups is greater than the number of arrays, SAA should be used.

There are also some rules of thumb guiding the placement of arrays for both approaches.

- Place the array as close as possible to the sources to be monitored. This increases the signal power at the array and thus SINR.

- Place the array in such a way that the beam paths in the beamforming map are well separated. This reduces the leakage from the other sources into the beamforming output and thus produces a cleaner signal.
- Place the array in such a way that reduces the number of beam paths which have a small angle parallel to the array. Those beams have poorer resolution and therefore admit more noise and potentially interferences into the beamforming output. By having less of those beams, a cleaner signal will be achieved.

4 CONCLUSION

To explore the idea of applying the microphone array technology for machine condition monitoring of a mine site, this paper focuses on the sound source localisation and signal extraction using the beamforming technique. The multiple array strategy has been adopted, as it is easier for array deployment and provides better flexibility and reliability. There are two approaches in terms of array processing, i.e., the independent array approach and the synchronised array approach. In order to compare the performance of these two approaches under various conditions, numerical simulation has been carried out. The performances have been evaluated in terms of the accuracy of source localisation (using a metric of RMSE) and the fidelity of signal extraction (using a metric of coherence). The general guidelines for using the two approaches as well as the rule of thumb for the placement of arrays are provided.

REFERENCES

Bao, C. 2005. ‘Adaptive beamforming for sonar audio’. In *Proceedings of Acoustics 2005*, 483-485. Busselton, Australia.

Bao, C. 2014. ‘Performance of time domain and time-frequency domain adaptive beamformers with moving sound sources’. In *Proceedings of Inter-noise 2014*, Melbourne, Australia.

Benesty, J., Chen, J., and Huang, Y. 2008. *Microphone Array Signal Processing.* Berlin: Springer.

Dhillon, B. 2008. *Mining Equipment Reliability, Maintainability and Safety.* London: Springer.

Li, J., Stoica, P., and Wang, Z. 2003. ‘On robust capon beamforming and diagonal loading’. *IEEE Trans. Signal Processing*, 51 (7): 1702-1715.

Ravetta, P., Muract, J., and Burdisso, R. 2007. ‘Feasibility study of microphone phased array based machinery health monitoring’. *Mecanica Computacional*, Vol XXVI: 23-37.

Van Trees, H. L. 2002. *Optimum array processing*. Wiley-Interscience.