ICSV14

Cairns • Australia

9-12 July, 2007

14th International Congress on Sound & Vibration

LOCALISATION OF COHERENT SOUND SOURCES USING A MICROPHONE ARRAY

Chaoying Bao[1], Hayden Albrey[2] and Jie Pan[2]

[1]Building A51, DSTO WA, PO Box 2188, Rockingham, WA 6958, Australia
[2]School of Mechanical Engineering, University of Western Australia, Crawley, WA 6009, Australia
Chaoying.bao@dsto.defence.gov.au

Abstract

Beamforming of a sound field using microphone arrays forms a foundation of acoustic camera technology, which overlays a sound pressure map (acoustic image) with an image of sound sources (visual image) to provide quick source localisation. Recent experiments conducted at the University of Western Australia on the measurement of sound radiation using an acoustic camera have observed significant shifts of the acoustic image from the visual image when the different parts of the sound source are coherent. In this paper, we examine the reason behind this image shifting phenomenon and present an approach that uses adaptive beamforming together with a spatial smoothing technique to eliminate those shifts. Sound radiations from loudspeakers and a music flute measured by the acoustic camera are used to illustrate the effectiveness of the approach.

1. INTRODUCTION

Sound source localisation using microphone arrays is a technology that has recently been used in many practical applications. One of these applications is the acoustic camera technique [1] for quick localisation of dominant sound sources. An acoustic camera typically consists of a microphone array and video camera. It overlays a sound pressure map (acoustic image) measured by the microphone array with an optical image captured by the video camera to provide a rapid identification of sound sources. The acoustic image is often formed by using beamforming technology.

Recent experiments conducted at the University of Western Australia on the measurement of sound radiation using an acoustic camera have observed that the source locations indicated by the acoustic image may shift significantly from those by the video image when the sound sources are coherent. Figure 1 shows an overlaid image obtained from the acoustic camera and illustrates such phenomenon.

The objectives of this paper are to explain the reason behind this phenomenon and to put forward a solution for locating coherent sources. In Section 2, we use a simple two element array to illustrate that the shift observed in Figure 1 is caused by the similarity of wavefronts generated by coherent and incoherent sources when coherent sound sources are involved. In Section 3, we demonstrate by simulation that adaptive beamforming (ABF) together with a de-correlation process called forward-backward averaging (FBA) [2] may solve the problem of incorrect localisation. In section 4, we apply the solution to the data measured by the acoustic camera and demonstrate its effectiveness in identifying the locations of practical coherent sources.

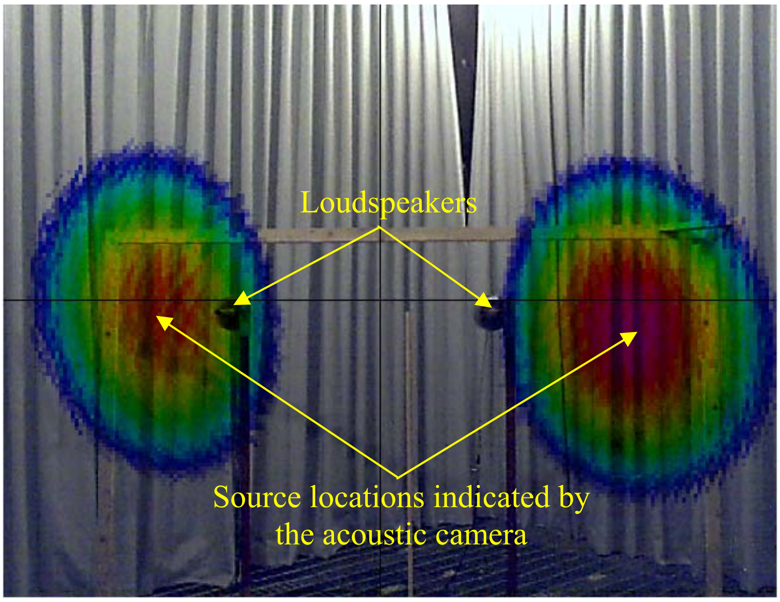


Figure 1. Image from the acoustic camera showing the incorrect localisation when the sources are coherent.

2. SHIFT OF SOUND SOURCE IMAGES IN BEAMFORMING

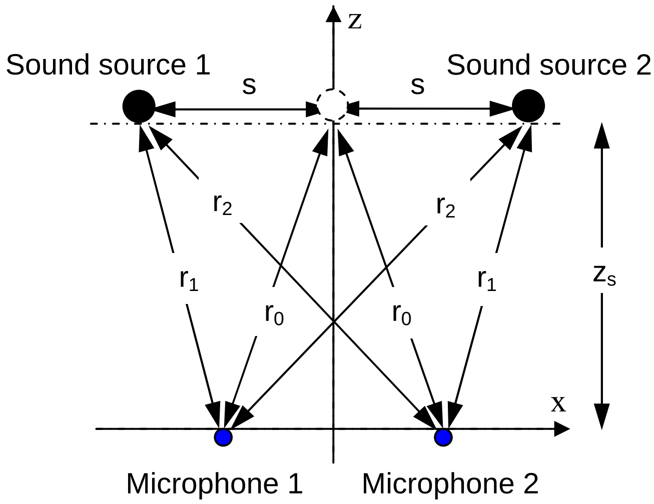


Figure 2. Coordinates of the array and sources.

In order to explain the phenomenon illustrated in Figure 1, let us consider a very simple scenario where a two element array is placed in the free field with two coherent sound sources. The sound produced by the sources is a single frequency pure tone with wavelength λ. Figure 2 shows the coordinates of the array and sound sources. Assuming that the strength and phase of both sources are 1 and 0° respectively, the complex sound pressures at the two microphones are

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} \dfrac{e^{jkr_1}}{r_1} + \dfrac{e^{jkr_2}}{r_2} \\ \dfrac{e^{jkr_2}}{r_2} + \dfrac{e^{jkr_1}}{r_1} \end{bmatrix}, \tag{1}$$

where k=2π/λ, and r_1 and r_2 are the distances between the microphones and sources as illustrated in Figure 2. From (1), we have p_1=p_2. Let us define this pressure has magnitude A and phase B. Thus, (1) can be rewritten as

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} Ae^{jB} \\ Ae^{jB} \end{bmatrix} , \tag{2}$$

where $A = \sqrt{\dfrac{1}{r_1^2} + \dfrac{1}{r_2^2} + \dfrac{2\cos(kr_2 - kr_1)}{r_1 r_2}}$ and $B = \tan^{-1}\left(\dfrac{r_2 \sin(kr_1) + r_1 \sin(kr_2)}{r_2 \cos(kr_1) + r_1 \cos(kr_2)} \right)$.

Now assuming a single sound source at location (0, z_s) with strength Ar_0 and phase (B-kr_0), where r_0 is the distance between the source and the microphones, the sound pressures produced by this single source at the two microphones are

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} \dfrac{Ar_0 e^{j(B-kr_0)} e^{jkr_0}}{r_0} \\ \dfrac{Ar_0 e^{j(B-kr_0)} e^{jkr_0}}{r_0} \end{bmatrix} = \begin{bmatrix} Ae^{jB} \\ Ae^{jB} \end{bmatrix}. \tag{3}$$

Comparing (2) to (3), we see that the complex sound pressures produced by two coherent sources at the two microphones are identical to those of a single source. An identical result is therefore expected for both cases by beamforming. It is known that in the case represented by (3) beamforming will produce a single peak at (0, z_s). Thus, in the case by (2) of two coherent sources, beamforming will also produce a single peak at (0, z_s) instead of two peaks at (-s, z_s) and (s, z_s).

The above example illustrates that the sound pressure pattern or wavefront generated by a set of coherent sources and observed by a microphone array can be identical or very similar to that by a different set of incoherent sources. This is the fundamental reason for incorrect localisation of beamforming when coherent sources are involved.

3. METHOD TO RECTIFY INCORRECT LOCALISATION

As we learnt in Section 2, beamforming does not give correct localisation of coherent sources whenever their sound pressure pattern observed by a microphone array is identical to that of incoherent sources. In this section, we will demonstrate by simulation that when those sound pressure patterns are similar but not identical, the widely used conventional beamforming

(CBF) based on the 'delay-and-sum' method [2] fails in localisation but adaptive beamforming (ABF) together with a spatial smoothing technique may work.

The microphone array considered in the simulation has the same configuration as the array used in the acoustic camera mentioned in Section 1. It is a uniform circular array with a diameter of 0.725 meters and 32 microphones. In the simulation, two coherent tonal point sources are placed in a free field and in the plane parallel to the array. The distance between the source plane and the array varies from 0.7 to 5 meters. The separation between the two sources varies from 0.07 to 4 meters. The frequency of tone varies from 300 to 3000 Hz. The sound speed is 340 m/s. Figure 3 shows the configuration of the array and sources in the simulation.

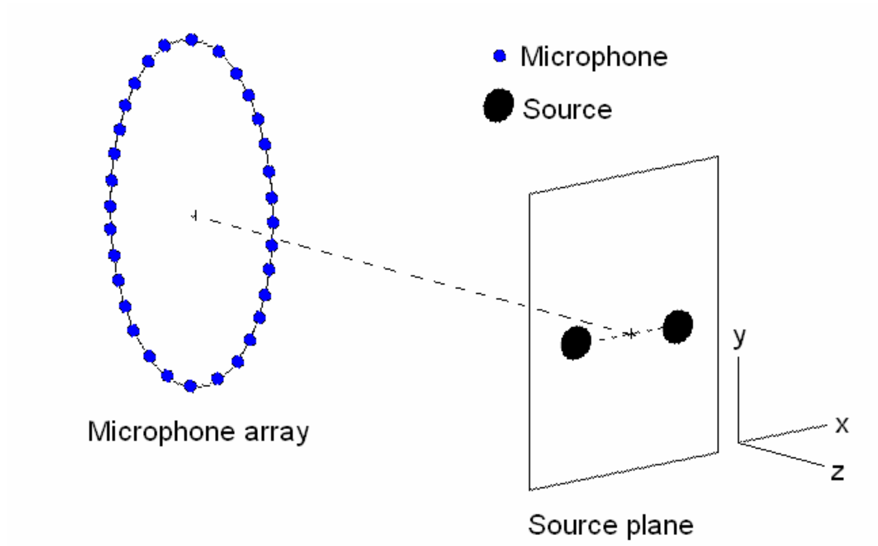


Figure 3. Configuration of the microphone array and sources in simulation.

Two types of beamforming are carried out in frequency domain and compared. One is the CBF used in the acoustic camera. The other is the ABF based on the Minimum Variance Distortionless Response (MVDR) approach [3]. A spatial smoothing technique called forward-backward averaging (FBA) is applied in both beamforming algorithms. FBA is a method of manipulating the cross-spectral matrix of the array input data to reduce correlation in it. It should be noted that FBA is applicable only if the array steering vectors are conjugate symmetric [2]. Whether the condition is satisfied depends on the array geometry. For the array used in the acoustic camera, the condition is satisfied if sources are in the far field. In the near field the condition can only be approximated. The accuracy of the approximation depends on three factors: the distance between the source plane and the array (the larger the better), the elevation angle of the beamforming point (the closer to 90° the better), and the frequency of sound (the lower the better). Fortunately, a large area exists within the working range of the acoustic camera where the accuracy of the approximation is high.

Figure 4 shows a typical result of the simulation where the two sources of equal strength are separated by 0.34 meters. The phase difference between the two sources varies from 0° to 360°. The frequency of sound is 2000 Hz. The array is 3.4 meters from the source plane, and its centre points to the centre of the source pair (see Figure 3).

The lighter lines in Figure 4 are the source locations indicated by CBF as function of the phase difference between the two sources. For an entire range of phase difference, CBF fails to indicate the correct locations of the sources. Particularly in the range of 0° to 125° and 235° to 360°, the output of CBF only shows a single source. The black lines are the source locations reported by MVDR. These locations are correct over most values of phase difference except in the small regions around 0° (360°) and 180°. Examining the eigenvalues of the cross-spectral matrix of the array shows FBA fails to de-correlate the coherent signals in these regions. It is known that FBA does not work if the forward and backward data samples are close to being either identical or symmetrical. In the simulation presented in Figure 4, the data samples are close to identical in the region around 0° and symmetrical in the region around 180°. The solution to the problem is to move the array to break the identical or symmetrical pattern. Figure 5 shows the result in which the array moves aside to a position where the array centre now points to one of the sources. It can be seen that CBF still fails to find the correct source locations in an entire range, but MVDR now reports the source locations correctly in the regions around 0° and 180°.

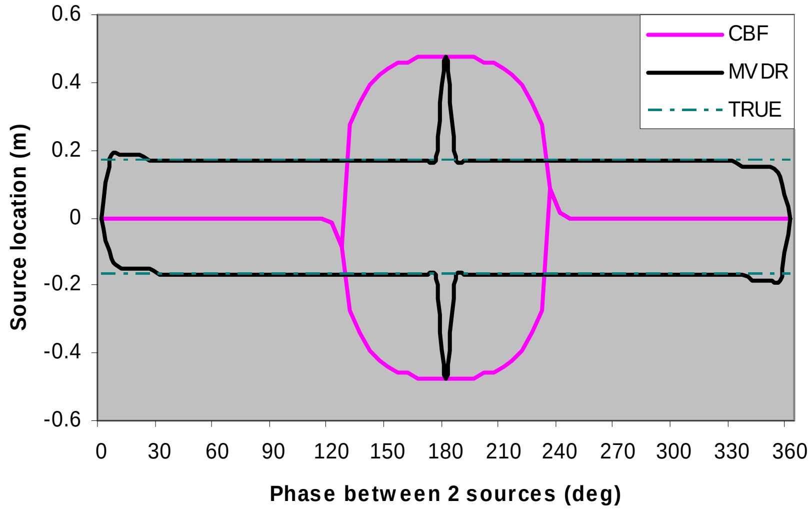


Figure 4. Source locations indicated by CBF and MVDR as function of the phase difference. True locations are (-0.17, 0.17) meters.

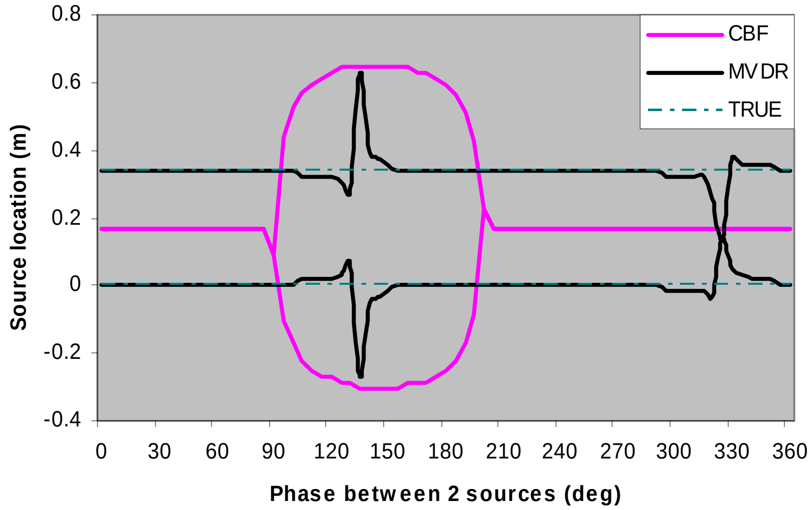


Figure 5. Source locations indicated by CBF and MVDR as function of the phase difference. True locations are (0, 0.34) meters.

In summary, simulation indicates that adaptive beamforming together with the spatial smoothing technique is able to correctly locate the coherent sound sources if the following two conditions are satisfied. Firstly the array steering vectors are close to being conjugate symmetric. Secondly the de-correlation process by the spatial smoothing is successful. Satisfaction to either condition can be affected by adjusting the position of the array relative to the sources. The procedure for satisfying the conditions and therefore successful localisation of coherent sources is summarized as follows:

- Place the array in a position where the sources are about the center of the measurement plane, and check whether the first condition (conjugate symmetric) is satisfied. If not, adjust the distance between the array and the sources until the condition is met.
- Check whether the second condition (de-correlation) is satisfied by computing the eigenvalues of the cross-spectral matrix after FBA. If not, shift the array sideway but parallel to the measurement plane until the condition is met.

4. EXPERIMENTS WITH ACOUSTIC CAMERA

In this section, we apply the approach proposed in the previous section to the data measured by the acoustic camera to demonstrate its effectiveness in locating practical coherent sources.

Experiment with two loudspeakers

This experiment is for locating two well separate point sources where the success of localisation can be easily judged.

In the experiment, two loudspeakers are placed in an anechoic chamber generating a tonal sound. The tonal signals fed into the loudspeakers have a frequency of 1000 Hz and phase difference of 90°. It should be pointed out that due to the differences in the two loudspeaker channels the final phase difference between the two sources may not be 90°. The distance between the two loudspeakers is 0.9 meter and the distance between the acoustic camera and the loudspeaker plane is 2.8 meters. Figure 1 shows the result obtained by the acoustic camera where the delay-and-sum beamformer is used. The localisation provided is obviously off the mark.

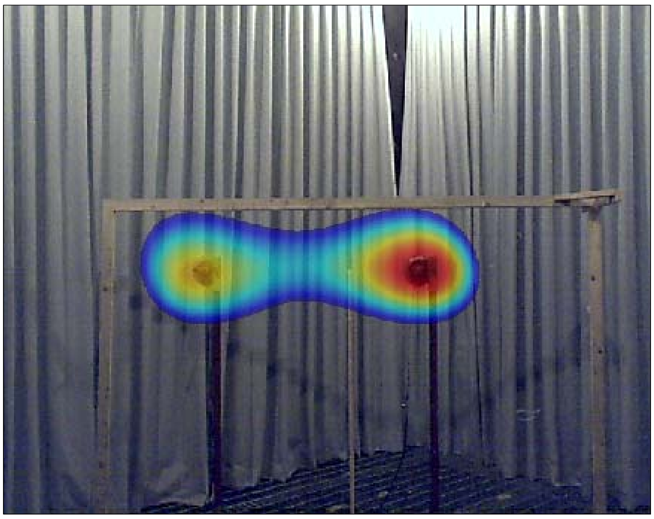


Figure 6. Same data as in Figure 1 but processed using MVDR with FBA showing the correct localisation.

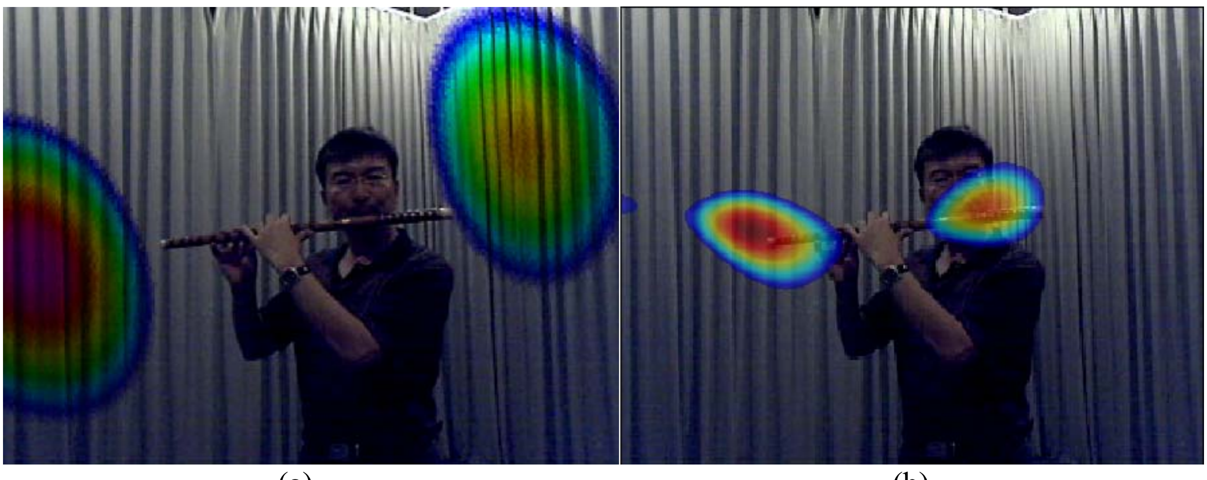
Before applying the proposed approach to the acoustic camera data, let us first examine whether the two conditions mentioned in Section 3 are satisfied. For the first condition, the maximum variation from being conjugate symmetric is 1.6°, within the tolerant range of 5°. For the second condition, the first four largest eigenvalues of the cross-spectral matrix before and after FBA are (1.3, 2.0×10^{-6}, 3.9×10^{-7}, 3.2×10^{-7}) and (1.1, 2.4×10^{-1}, 1.3×10^{-6}, 3.1×10^{-7}), respectively. The cross-spectral matrix after FBA clearly indicates two sources. The de-correlation process is therefore successful. Figure 6 shows the result obtained by using the MVDR beamformer with FBA. The correct localisation is achieved.

Experiment with a music flute

The acoustic camera is used to study the mechanism of sound generation for a music flute. Unlike the previous experiment where there are two separate sources, here a flute should be treated as a group of distributed sources. The aim of this experiment is to demonstrate the performance of the proposed approach in a more complex situation.

In the experiment, a music flute is played with a single note in an anechoic chamber. The fundamental frequency of the note is 772 Hz. The length of the flute is 0.68 meters and the distance between the acoustic camera and the flute is about 1.5 meters. Figure 7 (a) shows the source localisation at the fundamental frequency obtained by the acoustic camera. It is clear that the localisation is incorrect as the sources indicated are not on the flute. However, as incorrect as it is, this acoustic image does contain certain useful information. Comparing this image to that in Figure 1, we may conclude that the flute at this note acts more or less like a dipole-like source.

Let us examine the two required conditions. For the first condition, the maximum variation from being conjugate symmetric is 3°, still within the tolerant range. For the second condition, the first four largest eigenvalues of the cross-spectral matrix before and after FBA are (3.6, 2.9×10^{-3}, 7.0×10^{-5}, 3.2×10^{-5}) and (3.3, 3.1×10^{-1}, 1.9×10^{-3}, 8.7×10^{-4}), respectively. From the eigenvalues of the cross-spectral matrix after FBA, we can see an emerging second source. The de-correlation process is working but might not be at full extent. If an opportunity is present for redoing the measurement, adjusting the position of the acoustic camera is recommended. Figure 7 (b) shows the source localisation achieved by using the MVDR beamformer with FBA. The result seems to be more reasonable, for the sources are now on the flute.



(a) (b)

Figure 7. Source localisation of flute playing obtained by
(a) the acoustic camera and (b) the MVDR beamformer with FBA.

5. CONCLUSIONS

The problem of localisation of coherent sources using a uniform circular microphone array has been studied. It has been found that even with a de-correlation process the CBF based on delay-and-sum fails in locating coherent sound sources. On the other hand, successful localisation is possible by using ABF together with a de-correlation process. This has been demonstrated through both simulation and experimentation with the acoustic camera.

In the paper, a spatial smoothing technique called forward-backward averaging is used for the de-correlation process. The condition for the technique being applicable requires that the array steering vectors are approximately conjugate symmetric. This condition can often be satisfied through manipulating the position of the array relative to the sources. The success of de-correlation also often requires such manipulation.

AKNOWLEDGEMENT

We would like to thank GFaI of Germany for kindly lending the acoustic camera to us. We also thank Mr. Waibel of HW Technologies and Mr. Heilmann of GFaI for supporting our experiments.

REFERENCES

[1] O. Jaeckel and G. Heilmann, “Transient noise source localisation”, *Proceedings of Euronoise 2006*, Tampere, Finland, 30 May–1 June 2006.

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

[3] D.H. Johnson and D.E. Dudgeon, *Array signal processing – concepts and techniques*, Prentice-Hall, Englewood Cliffs, NJ, 1993.