High-resolution clean-sc

Sijtsma, Pieter; Snellen, Mirjam

Publication date
2016

Document Version
Accepted author manuscript

Published in
6th Berlin Beamforming Conference

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.
HIGH-RESOLUTION CLEAN-SC

Pieter Sijtsma\textsuperscript{1,2} and Mirjam Snellen\textsuperscript{2}
\textsuperscript{1}PSA3, Prinses Margrietlaan 13, 8091 AV, Wezep, The Netherlands
\textsuperscript{2}Faculty of Aerospace Engineering, TU Delft, The Netherlands

ABSTRACT

In this paper a high-resolution extension of CLEAN-SC is proposed: HR-CLEAN-SC. Where CLEAN-SC uses peak sources in “dirty maps” to define so-called source components, HR-CLEAN-SC takes advantage of the fact that source components can likewise be derived from points at some distance from the peak, as long as these “source markers” are on the main lobe of the Point Spread Function (PSF). This is very useful when sources are closely spaced together, such that their PSFs interfere. Then, alternative markers can be sought in which the relative influence by PSFs of other source positions is minimised. For those markers the source components better agree with the actual sources, which allows for better estimation of their locations and strengths. This paper outlines the theory needed to understand this approach and discusses applications to 2D and 3D microphone array simulations with closely spaced sources.

NOMENCLATURE

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CB</td>
<td>Conventional Beamforming</td>
</tr>
<tr>
<td>CSM</td>
<td>Cross-Spectral Matrix</td>
</tr>
<tr>
<td>PSF</td>
<td>Point Spread Function</td>
</tr>
<tr>
<td>$A_k$</td>
<td>source power</td>
</tr>
<tr>
<td>$\hat{A}_k$</td>
<td>source power estimate</td>
</tr>
<tr>
<td>$a_k$</td>
<td>source amplitude</td>
</tr>
<tr>
<td>$C$</td>
<td>CSM</td>
</tr>
<tr>
<td>$K$</td>
<td>number of incoherent sources</td>
</tr>
<tr>
<td>$g_j$</td>
<td>steering vector</td>
</tr>
<tr>
<td>$h_k$</td>
<td>source component</td>
</tr>
<tr>
<td>$p_k$</td>
<td>source vector</td>
</tr>
<tr>
<td>$u_k$</td>
<td>source marker weight vector for CB</td>
</tr>
<tr>
<td>$w_j$</td>
<td>general weight vector for CB</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

The deconvolution technique CLEAN-SC [1] starts with an acoustic image obtained with Conventional Beamforming (CB) and features the iterative removal of those parts of the acoustic image that are coherent with the peak source. For each iteration step the removed part of the image is related to a “source component”, which represents measured microphone data due to a single coherent source. Each source component is represented by an artificial “clean beam” at the peak location in a new acoustic image: the “clean map”. The levels of the clean beams are calculated from the source components.

A convenient feature of CLEAN-SC is that the determination of source components is not very sensitive to the location that is marked as peak. In other words, if the scan grid is too coarse or if it’s out of focus, a small error may be made in the peak location, but the corresponding source levels remain correct. Thus, CLEAN-SC provides levels at a higher reliability than CB.

CLEAN-SC is not a typical high-resolution technique, since it is not able to separate sources that are spaced within the Rayleigh limit. If two sources are too close to each other, the CB peak location is somewhere in between both sources, and the corresponding source component is a linear combination of the two individual sources.

In those cases, it can be advantageous to move the “source marker” away from the actual peak location, to a location where the CB result is dominated by either one of the sources (see Figure 1). This allows for significantly better estimates of the source levels. Improved estimations of the locations of the sources are obtained by applying CB to the source component.

To determine the best marker locations, knowledge about the actual source positions is required. But these positions are not always known a priori. However, we will demonstrate that an iterative procedure, starting with the standard CLEAN-SC solution, also leads to an increase in resolution, based on the idea of optimising marker locations.

In the following section the theory is outlined. In Section 3 applications to 2D and 3D microphone array simulations with closely spaced sources are discussed. The conclusions are summarised in Section 4.

2 THEORY

2.1 The Cross-Spectral Matrix

The starting point for frequency-domain beamforming methods with microphone arrays is the Cross-Spectral Matrix (CSM). It is assumed here that the CSM can be written as a summation of contributions from $K$ incoherent sources:

$$C = \sum_{k=1}^{K} p_k p_k^*.$$  (1)

Herein, $p_k$ are $N$-dimensional “source vectors” ($N$ being the number of microphones) representing the Fourier components of the signals from the $k$-th source. The asterisk stands for the complex conjugate transpose. The assumption of Eq. (1) is valid under the following conditions:

- The CSM is calculated from a large number of time blocks, so that the ensemble averages of the cross-products $p_l p_l^*$, $k \neq l$, can be neglected.
- There is no decorrelation of signals from the same source between different microphones (e.g., due to sound propagation through turbulence).
- There is no additional incoherent noise.

2.2 Steering vectors

Beamforming methods make use of “steering vectors” \( \mathbf{g} \). These vectors contain the microphone responses of potential sources. It’s quite common to assume point sources represented by free-field Green’s functions of the Helmholtz equation. Then the \( n \)-th component of \( \mathbf{g} \) reads

\[
g_n = \frac{-\exp\left(-2\pi i f \| \hat{x}_n - \hat{\xi} \| / c \right)}{4\pi \| \hat{x}_n - \hat{\xi} \|},
\]

where \( f \) is the frequency, \( c \) the sound speed, \( \hat{x}_n \) the position of the microphone and \( \hat{\xi} \) the source location. However, \( \mathbf{g} \) can represent any sound source mechanism, like plane waves or point sources with non-uniform directivity, or extended source regions. Also it can include acoustic propagation through non-uniform media [2].

For beamforming, a so-called “scan area” is defined, which is basically a set of steering vectors coupled to potential sources. The scan area should comprise all sources that produce the CSM. It is favourable, in general, to put as much as possible of the physics into the steering vectors, in other words, to maximize the likelihood of steering vectors \( \mathbf{g} \) being proportional to the source vectors \( \mathbf{p}_k \). Ideally, the scan area contains steering vectors \( \mathbf{g}_k \) that are exactly proportional to the source vectors:

\[
\mathbf{p}_k = a_k \mathbf{g}_k,
\]

where \( a_k \) is a (complex-valued) source amplitude.

For many array measurements, however, there is no exact proportionality. Deviations between actual source vectors and theoretical steering vectors can be due to the source not being a true point source, a non-uniform directivity, errors in the scan plane, errors in the microphone locations, errors in the sound propagation model, or errors in the microphone sensitivity.

The aim of beamforming is to detect sources and to determine associated source powers, i.e., to decompose the CSM like

\[
\mathbf{C} \approx \sum_{k=1}^{K} A_k \mathbf{g}_k \mathbf{g}_k^*.
\]

For ideal source vectors \( \mathbf{p}_k \), for which Eq. (3) holds, the source powers \( A_k \) must be

\[
A_k = |a_k|^2.
\]

2.3 Conventional Beamforming

The expression for calculating source power estimates with CB [3] is

\[
\tilde{A}_k = \mathbf{w}_k^* \mathbf{C} \mathbf{w}_k,
\]

featuring the “weight vector” \( \mathbf{w}_k \):

\[
\mathbf{w}_k = \frac{\mathbf{g}_k}{\| \mathbf{g}_k \|}.
\]

Application of CB to the CSM assumed in Eq. (1) yields
\[ \tilde{A}_k = w_i^* \left( \sum_{j=1}^{K} p_j p_j^* \right) w_k = \sum_{j=1}^{K} \left| p_j^* w_k \right|^2. \]  \hspace{1cm} (8)

For ideal source vectors, Eq. (3), we have
\[ \tilde{A}_k = A_k + \sum_{j \neq k}^{K} A_j |g_j^* w_k|^2. \]  \hspace{1cm} (9)

For a reasonable estimate of the source power, the terms with \( j \neq k \) in the right hand side of Eq. (9) need to be small. If \( A_j \gg A_k \), then it’s necessary that \(|g_j^* w_k|^2 \ll 1\). \hspace{1cm} (10)

The expression in the left hand side of Eq. (10) is known as the “Point Spread Function” (PSF) of source \( j \). Eq. (10) states that the sources \( j \) and \( k \) must be sufficiently far away from each other, and that side lobes of the PSF should be small [4,5]. The PSF characteristics thus limit the application of CB. The aim of deconvolution methods, like the “CLEAN” methods discussed in the next section, is essentially to correct for the PSFs.

2.4 **CLEAN-PSF and CLEAN-SC**

The classical CLEAN [6] deconvolution algorithm (referred to as “CLEAN-PSF” in [11]) works as follows. Let \( p_i \) be the highest-level source vector and \( g_i \) the best matching steering vector (e.g., found as the one for which CB yields the maximum source power estimate). Then the CSM is rewritten as follows:
\[ C = C_{\text{clean}} + C_{\text{dirty}}, \]  \hspace{1cm} (11)

with
\[ C_{\text{clean}} = \lambda g_i^* g_i, \]  \hspace{1cm} (12)
\[ C_{\text{dirty}} = \left( p_i p_i^* - \lambda g_i g_i^* \right) + \sum_{k=2}^{K} p_k p_k^*, \]  \hspace{1cm} (13)

where \( \lambda \) is the source power estimate \( \tilde{A}_i \) multiplied with a “loop gain” factor. The \( C_{\text{clean}} \) results, featuring a source location associated with \( g_i \) and a corresponding amplitude, are coupled to a “clean map”, while \( C_{\text{dirty}} \) is used as input for the next iteration step.

The CLEAN-SC [1] counterparts of Eqs. (12) and (13) are
\[ C_{\text{clean}} = \lambda h_i h_i^*, \]  \hspace{1cm} (14)
\[ C_{\text{dirty}} = \left( p_i p_i^* - \lambda h_i h_i^* \right) + \sum_{k=2}^{K} p_k p_k^*, \]  \hspace{1cm} (15)

with the “source component” \( h_i \) defined by
\[ h_i = \frac{1}{\tilde{A}_i} C w_i = \frac{p_i w_i}{\tilde{A}_i} \left( p_i + \frac{1}{p_i w_i} \sum_{k=2}^{K} (p_i^* w_i) p_k \right). \]  \hspace{1cm} (16)

where \( w_i \) is the weight vector associated with \( g_i \) via Eq. (7).

If the best-matching steering vector \( g_i \) is proportional to the source vector, then CLEAN-PSF is a perfect deconvolution method, since the contribution of the principal source \( p_i \) can be removed completely from the “dirty” CSM, Eq. (13). With the right choice for \( \lambda \), the first term in the right hand side of (13) can be completely annihilated. This does not hold for CLEAN-SC. For proportional \( g_i \) and \( p_i \) (ideal source vectors) we obtain for the source component:
\[ h_i = \frac{A_i}{A} \left\{ g_i + \frac{1}{A} \sum_{k=2}^{K} A_k \left( g_i w_k \right) g_k \right\}. \] (17)

Thus, the dirty CSM gets polluted by contributions of the principal source. These contributions, however, can be relatively small, dependent on the values of \( g_i w_k \), i.e., on PSF values.

In the non-ideal case, when \( g_i \) is not proportional to \( p_i \), the dirty matrix obtained with CLEAN-PSF, Eq. (13) can contain significant contributions of \( p_i \). On the other hand, with CLEAN-SC no additional errors are introduced when Eq. (17) is replaced by Eq. (16). This is one of the main advantages of CLEAN-SC compared to other deconvolution methods, namely that it does not lose performance if steering vectors don’t exactly match with source vectors.

### 2.5 High-Resolution CLEAN-SC

The fact that the best-matching steering vector \( g_i \) and the source vector \( p_i \) don’t need to be proportional can be exploited for optimizing the CLEAN-SC results. The source component \( h_i \), Eq. (16), provides the best (proportional) estimate for \( p_i \) if the following cost function is minimized:

\[ F(u_j) = \frac{\left\| \sum_{k=2}^{K} (g_i u_j) p_k \right\|^2}{\|p_k u_j\|^2 \|p_k\|^2}, \] (18)

where \( u_j \) is the weight vector associated with an alternative “source marker”, as discussed in the Introduction.

To evaluate Eq. (18) we need an initial set of source vector estimates \( p_j = a_j g_j \), which can be obtained with the standard CLEAN-SC method. Then, updated marker locations are obtained through successive optimization of

\[ F_j(u_j) = \frac{\left\| \sum_{k \neq j} A_k^2 \left( g_k u_j \right) g_k \right\|^2}{A_j^2 \|g_k u_j\|^2 \|g_k\|^2}. \] (19)

Source components are then defined by

\[ h_j = \frac{C u_j}{u_j^* C u_j} \] (20)

and updated source locations are found by maximizing

\[ G(w_j) = \|w_j h_j\|^2. \] (21)

Finally, source power estimates are calculated by

\[ \tilde{A}_j = \left( u_j^* C u_j \right) \|w_j h_j\|^2 \] (22)

With this information we can proceed with the next update by minimizing Eq. (19).

However, what may happen is the following. If sources of equal strength are spaced closely together (closer than the Rayleigh limit), then CLEAN-SC distributes the acoustic energy unequally over the source components. Thus, the weakest source contributes the least to the cost function, Eq. (19). This may lead to an optimum in which the weak sources remain weak. Therefore, we will omit the amplitudes in the cost function and use
2.6 CSM diagonal removal

Another feature of CLEAN-SC is that it works well with CSM diagonal removal. In that case, the expression for the source component reads [1]:

$$h_i = \frac{1}{1 + u_i^T H u_i} \left( \overline{C} u_i + H u_i \right),$$  \quad (24)

in which $\overline{C}$ is the “tripped” CSM with the diagonal elements replaced by zeros. Further, $H$ is a diagonal matrix, consisting of the diagonal elements of $h_i h_i^T$. The weight vector $u_i$ is now calculated from the steering vector by [3]

$$u_i = \frac{g_i}{\left( \|g_i\|^2 - \sum_{n=1}^{N} |g_{i,n}|^2 \right)^{1/2}}. \quad (25)$$

It’s not straightforward to derive cost functions similar to Eq. (19) or Eq. (23). However, a reasonable estimate of Eq. (23) is

$$F_j(u_j) = \left| \sum_{k=1, k \neq j}^{K} (g_k^T u_j g_k) \right|^2 \cdot \frac{1}{\|g_k\|^2}. \quad (26)$$

The removed diagonal equivalent of Eq. (26) is

$$F_j(u_j) = \left| \sum_{k=1, k \neq j}^{K} (u_j^T g_k g_k^T u_j) \right|^2 \cdot \frac{1}{\|g_k\|^2}. \quad (27)$$

The optimisation process is then equivalent to Section 2.5. The source components $h_j$ are calculated as in Eq. (24). Updated source locations are found by maximizing

$$G(w_j) = w_j^T h_j h_j^T w_j \quad (28)$$

and source power estimates are calculated by

$$\tilde{A}_j = (u_j^T \overline{C} u_j) (w_j^T \overline{h}_j \overline{h}_j^T w_j). \quad (29)$$

3 EXAMPLES WITH SIMULATED ARRAY DATA

In this chapter we consider examples of 2D and 3D array measurements. To avoid division by zero in Eq. (23), we have put the following constraint on the marker location $u_j$:

$$\left| g_j^T u_j \right|^2 \geq 0.25. \quad (30)$$

This means that the PSF-value at the marker location is not more than 6 dB below the peak. With this constraint the source marker will normally stay on the main lobe of the PSF.
3.1 2D simulation

Array measurements were synthesized with a linear array of 2 m length, consisting of 101 microphones, uniformly spaced at 2 cm. The sound field consisted of plane waves, arriving from directions characterised by polar angles \( \varphi_k \):

\[
p_k(x, y) = a_k \exp \left\{ 2\pi if \left[ \sin(\varphi_k) y - \cos(\varphi_k) x \right]/c \right\}.
\]

The array is on the \( y = 0 \) line, so for the “measured” pressures we have

\[
p_{k,n} = p_k(x_n, 0) = a_k \exp \left\{ -2\pi if \cos(\varphi_k) x_n/c \right\}.
\]

The simulated sound field consisted of two plane waves at 500 Hz, both at unit strength (94 dB), with polar angles 80° and 89°. Figure 2 shows the angles and the amplitudes of the sources, the CB array response and the contributions from the sources individually (the PSFs). The angular spacing between the two sources is less than half the Rayleigh limit, which reads in two dimensions:

\[
\Delta \varphi = c/Lf,
\]

where \( L \) is the length of the array. Therefore, CB is not able to resolve both sources, as is obvious from Figure 2. Consequently, CLEAN-SC can’t resolve the two sources either.

The HR-CLEAN-SC method, outlined in Section 2.5, would give a perfect reconstruction of the PSFs if the source markers \( u_j \) would be at the locations where the PSF value of the other source is minimal. Those locations, which are indicated in Figure 2, can only be determined if the PSFs and, hence, the source locations are known in advance, which is not always the case. However, we can iterate to the correct solution, as demonstrated in the following.

The first step in the iteration process is made with standard CLEAN-SC. Herewith, 2 sources and corresponding source components are found as indicated in Figure 3. The first source is found halfway the two physical sources, and at a higher level. The second source has a much lower level and is found at an unphysical location. However, this first estimate of the source positions is used to find first estimates of the source markers, by searching for minima of the PSFs associated with the source location estimates. The result is shown in Figure 4.

The next step is to calculate source components, Eq. (20), starting from these markers. Updated source locations are then found by searching the maximum value of the source components. This is illustrated in Figure 5. By considering the PSFs of these updated source locations we can find new marker locations, as illustrated in Figure 6. With these new markers, we can determine new source components and update the source estimates, as shown in Figure 7, and so on. By comparing Figure 3, Figure 5, and Figure 7, we see that the source estimates clearly move into the directions of the true sources. After 8 iterations the process is fully converged and the source estimates coincide with the true sources. This is shown in Figure 8.

With this simulation we demonstrated that the spatial resolution can be increased by a factor 2 compared to the Rayleigh limit. The gain in resolution that can be attained depends on the constraint defined in Eq. (30). In fact, with the constraint value of 0.25 (6 dB), a gain by a factor 2 is about the maximum attainable. More gain is possible by reducing the right hand side of Eq. (30).

3.2 3D simulations

Simulations in three dimensions were made with an acoustic array in the \( z = 0 \) plane. The array consisted of 133 microphones distributed uniformly over a disk with radius 0.6, as depicted in Figure 9. Acoustic fields of multiple point sources in the \( z = 1 \) plane were simulated, each
source with unit strength (94 dB). A series of 6 fixed frequencies were considered, from 250 to 1500 Hz.

3.2.1 Two sources

The first simulation was made with 2 sources on the \(x = 0\) line, spaced 20 cm apart. The CB-results (without CSM diagonal removal) are shown in Figure 10, the CLEAN-SC results in Figure 11, and the HR-CLEAN-SC results in Figure 12. The dashed lines indicate the positions of the sources. The HR-CLEAN-SC results were obtained with 5 iterations, which was sufficient for convergence.

A comparison of Figure 12 against Figure 11 clearly shows the improvement of HR-CLEAN-SC, both in the location of the sources and in their estimated levels. The quality of the 500 Hz image in Figure 12 is comparable to the 1000 Hz image in Figure 11, and the same holds for the respective images at 750 Hz and 1500 Hz. Thus, it seems justified to conclude that the spatial resolution has increased by a factor 2, as in the 2D simulation of Section 3.1.

The beamforming images in Figure 10 to Figure 12 were obtained with the full CSM. However, in many beamforming applications it’s necessary to remove the diagonal. As outlined in Section 2.6, the HR-CLEAN-SC approach without the CSM diagonal is less exact. This is confirmed by Figure 13, which is the removed diagonal equivalent of Figure 12. The results without diagonal are a little worse (especially at the lower part of the frequency range) than with full CSM. Nevertheless, there is still significant improvement compared to standard CLEAN-SC.

In the remaining part of this section we consider beamforming only with the full CSM.

3.2.2 More than two sources

When the HR-CLEAN-SC method is applied to more than two sources, then the summation in the denominator of Eq. (23) is done for more than one steering vector \(g_k\). Thus, there is less freedom in minimizing Eq. (23). Consequently, the improvements obtained with the HR-CLEAN-SC are expected to be less than with two sources.

First, a simulation was made with 3 sources, again on the \(x = 0\) line and with 20 cm spacing. The CB-results are shown in Figure 14, the CLEAN-SC results in Figure 15, and the HR-CLEAN-SC results in Figure 16. The HR-CLEAN-SC results are still significantly better than the standard CLEAN-SC results, but the resolution improvement is no longer by a factor 2.

The trend of reduced added value is continued when 4 sources are closely spaced. This can be concluded from beamforming simulations shown in Figure 17 (CB), Figure 18 (CLEAN-SC) and Figure 19 (HR-CLEAN-SC).

4 CONCLUSIONS

The HR-CLEAN-SC method, proposed in this paper, is a high-resolution extension of CLEAN-SC. It's in particular suitable for pairs of closely-spaced sources. Then the spatial resolution can be increased by a factor 2. The features of the standard CLEAN-SC method [1] are fully preserved.

Obviously, HR-CLEAN-SC needs more computation time than CLEAN-SC. However, the most time-consuming part, which is CB at the start of the iteration process, doesn’t need to be done more often. Consequently, when only a few CLEAN-SC iterations are needed, i.e., when the number of sources \(K\) is small, the additional computation time is limited.
Thus, HR-CLEAN-SC is especially suitable when the number of sources is small. For beamforming applications with many sources (e.g., airframe noise measurements in wind tunnels) HR-CLEAN-SC is not expected to give much added value.

HR-CLEAN-SC can be applied with and without removal of the CSM. In both cases significant increase in resolution is found compared to the standard CLEAN-SC method. Without CSM removal, i.e., with the full CSM, the best results are obtained. When it’s necessary to remove the diagonal, HR-CLEAN-SC may benefit from reconstruction methods [7].

REFERENCES


FIGURES

![Figure 1 Sketch of the main idea of HR-CLEAN-SC](source_marker)
Figure 2 Sources at 80° and 89°, 500 Hz; Array response and ideal source marker locations

Figure 3 Sources at 80° and 89°, 500 Hz; CLEAN-SC solution
Figure 4  Sources at 80° and 89°, 500 Hz; 1st source marker estimates

Figure 5  Sources at 80° and 89°, 500 Hz; 1st update of source estimates
Figure 6  Sources at 80° and 89°, 500 Hz; 2nd source marker estimates

Figure 7  Sources at 80° and 89°, 500 Hz; 2nd update of source estimates
Figure 8 Sources at 80° and 89°, 500 Hz; 8th update of source estimates

Figure 9 Array for 3D simulations
Figure 10 CB results with 2 sources

Figure 11 CLEAN-SC results with 2 sources

Figure 12 HR-CLEAN-SC results with 2 sources; obtained with 5 iterations
Figure 13  HR-CLEAN-SC results with 2 sources; obtained with 5 iterations; CSM diagonal removed

Figure 14  CB results with 3 sources

Figure 15  CLEAN-SC results with 3 sources
Figure 16 HR-CLEAN-SC results with 3 sources; obtained with 5 iterations

Figure 17 CB results with 4 sources

Figure 18 CLEAN-SC results with 4 sources
Figure 19  HR-CLEAN-SC results with 4 sources; obtained with 5 iterations