Reconstruction methods for sound visualization based on acousto-optic tomography

Torras Rosell, Antoni; Lyllof, Oliver; Barrera Figueroa, Salvador; Jacobsen, Finn

Published in: Proceedings of INTER-NOISE 2013

Publication date: 2013

Reconstruction methods for sound visualization based on acousto-optic tomography

Antoni Torras-Rosell, Oliver Lylloff, Salvador Barrera-Figueroa, and Finn Jacobsen

1, 3 DFM, Danish National Metrology Institute
Matematiktorvet 307, 2800 Kongens Lyngby, Denmark
2, 4 Acoustic Technology, Department of Electrical Engineering, Technical University of Denmark
Ørsteds Plads 352, 2800 Kongens Lyngby, Denmark

ABSTRACT
The visualization of acoustic fields using acousto-optic tomography has recently proved to yield satisfactory results in the audible frequency range. The current implementation of this visualization technique uses a laser Doppler vibrometer (LDV) to measure the acousto-optic effect, that is, the interaction between sound and light, over an aperture where the acoustic field is to be investigated. By identifying the relationship between the apparent velocity of the LDV and the Radon transform of the acoustic field, it is possible to reconstruct the sound pressure distribution of the scanned area using tomographic techniques. The filtered back projection (FBP) method is the most popular reconstruction algorithm used for tomography in many fields of science. The present study takes the performance of the FBP method in sound visualization as a reference and investigates the use of alternative methods commonly used in inverse problems, e.g., the singular value decomposition and the conjugate gradient methods. A generic formulation for describing the acousto-optic measurement as an inverse problem is thus derived, and the performance of the numerical methods is assessed by means of simulations and experimental results.

Keywords: Acousto-optic, Tomography, Visualization

1. INTRODUCTION
The ability to measure airborne sound by means of the acousto-optic effect, that is, the interaction between sound and light, has recently rendered surprisingly good results for various acoustic applications, namely visualization, localization and identification of acoustic sources. The use of light as a sensing element poses a fundamental advantage over conventional measurement techniques: the light does not change the properties of sound when traveling through the acoustic field, which means that the acousto-optic effect is...
a non-invasive measurement principle. This is particularly beneficial and crucial at high frequencies, where scattering effects normally corrupt measurements carried out with transducer-based techniques. The present work investigates the possibility of enhancing the visualization of sound fields based on acousto-optic tomography by means of numerical methods.  

Acousto-optic tomography is usually implemented with the so-called filtered back projection (FBP) method, which is in fact the most renowned tomographic reconstruction algorithm in many fields of science. The numerical methods are seen in this study as a promising alternative to the FBP method because these methods stem from a completely different mathematical principle, and they have been applied successfully to many other areas of research such as medical imaging, seismology and astronomy.

In the following, a brief review about acousto-optic tomography and its formulation as an inverse problem are given. The numerical methods under investigation are then presented. The performance of these methods is examined by means of simulations. The comparison between the numerical methods is possible by analyzing two different performance indicators, namely the mean square error (MSE) and the structural similarity (SSIM) index. Some preliminary experimental results also illustrate the potential of the examined numerical methods. Most of the source code used in this study is based on the open MATLAB toolboxes Regularization Tools\textsuperscript{4} and AIRtools\textsuperscript{5}.

\section{2. ACOUSTO-OPTIC TOMOGRAPHY}

\subsection{2.1. Measurement principle}

The speed of light depends on the density of the medium where the light is traveling. For instance, light travels fastest in vacuum, it slows slightly down in air and it propagates even slower in water. This is just to illustrate that sound waves, as pressure fluctuations, change indeed the density of the medium, and thus, influence the propagation of light as well. The acousto-optic effect is a tiny phenomenon that within the audible frequency range and for sound pressure levels below the threshold of pain can simply be modeled as a modulation effect on the phase of light: light travels slightly faster/slower when the acoustic pressure reduces/increases. Such a small effect can be measured using interferometry, and in particular, with an LDV. This optical instrument is originally intended for measuring mechanical vibrations of a surface where the laser beam is reflected off. However, it can be shown that by keeping the vibrations of the reflecting point negligible and when the laser beam travels across a sufficiently intense sound field, the output of an LDV corresponds to an apparent velocity that fulfills the following expression in the frequency domain,\textsuperscript{2}

\begin{equation}
V_{\text{LDV}}(\omega) = j\omega \frac{n_0 - 1}{\gamma p_0 n_0} \int_{L} P(\vec{r}, \omega) dl,
\end{equation}

where \(P(\vec{r}, \omega)\) is the temporal Fourier transform of the acoustic pressure, \(L\) represents the path traveled by the laser beam, \(n_0\) and \(p_0\) are the refractive index and pressure of the medium (in this case, air) under static conditions, and \(\gamma\) is the ratio of specific heats. This equation shows that the measurement of the acousto-optic effect does not provide a direct measure of the sound pressure, but rather its line integral along the laser beam. A single line integral is not enough to reconstruct an arbitrary sound field, at least without taking for granted basic assumptions about the nature of the sound field. Instead, a complete reconstruction without incorporating any previous knowledge about the sound field can be achieved using tomography and the Radon transform. The latter is an integral transform named after the Austrian mathematician Johann Radon, who laid down the principles of tomography with his posterior derivation of the inverse Radon transform. The Radon transform of an acoustic field can be expressed as a function of the apparent velocity of the LDV stated in equation (1) as follows,\textsuperscript{2}

\begin{equation}
R_0(\theta, x', \omega) = \int_{-L/2}^{L/2} P(x', y', \omega) dy' = \frac{\gamma p_0 n_0}{n_0 - 1} \frac{V_{\text{LDV}}(\omega)}{j\omega},
\end{equation}

where \(\theta\) is the angle of projection of the acoustic field, \(x'\) and \(y'\) are the Cartesian coordinates according to the projection angle \(\theta\), and \(L\) is the total length of the line integral. A sketch of the measurement principle can be seen in Figure 1. The measurement of the acousto-optic effect over a sufficient number of angles of projection and parallel lines constitutes a well-defined problem that relates the unknown data (sound pressure) with the observed data (Radon transform). The trivial solution to this forward problem is to apply the inverse Radon transform. However, such a naive solution will often lead to terrible results due to the presence of noise. In practice, tomographic algorithms incorporate regularization techniques to stabilize results. The FBP method is probably the most well-known tomographic reconstruction algorithm. For parallel line scanning.
configuration, the sound pressure is reconstructed as follows,
\[
P(\vec{r}, \omega) = \int_0^{\pi} Q(\theta, x', \omega) d\theta,
\]
where \( Q(\theta, x', \omega) \) is the so-called “filtered projection”, which corresponds to the convolution of the measured Radon transform with a filter \( h(x') \) that accounts for both the implementation of the reconstruction algorithm using the two-dimensional Fourier transform (Ram-Lak filter) and the regularization (filtering) imposed in the wavenumber domain (e.g. Shepp-Logan filter) to reduce the influence of noise.

2.2. Tomography as an inverse problem

The reconstruction of sound pressure from the Radon transform of an acoustic field can also be described as an inverse problem using a matrix equation of the following form,

\[
Ax = b,
\]
where \( x \) is a column vector with \( n \) elements representing the unknown, complex sound pressures \( (x \in \mathbb{C}^n) \), \( b \) is a column vector with \( m \) elements corresponding to the measured, complex Radon transform \( (b \in \mathbb{C}^m) \), and \( A \) is a real matrix with dimensions \( m \times n \) that holds the underlying tomographic model \( (A \in \mathbb{R}^{m \times n}) \). Unlike the FBP method, one of the great advantages of using numerical methods is the possibility of modeling any tomographic scheme beyond the classical parallel and fan-shaped scanning configurations. The matrix \( A \) is currently implemented with the MATLAB function `radon`, where the columns are the Radon transform of a black image with a single white pixel that is looped over all the positions of the reconstruction aperture.\(^6\) The inversion of equation (4) constitutes a discrete ill-posed problem that arises from the discretization of the forward problem (the Radon transform) and the noise in the measurement vector \( b \). The rest of this section presents five regularization techniques to circumvent the ill-posedness of the problem.

2.2.1. Singular value decomposition

The singular value decomposition (SVD) is an essential theoretical and computational tool for understanding both the ill-posedness and the effects of regularization on linear inverse problems such the one stated in equation (4). The SVD of \( A \) can be defined as,\(^7\)

\[
A = U \Sigma V^T = \sum_{i=1}^{n} u_i \sigma_i v_i^T,
\]
where \( \sigma_i \) is the \( i \)’th singular value, \( \Sigma \) is the corresponding nonnegative diagonal matrix consisting of the singular values appearing in descending order, \( u_i \) and \( v_i \) are the \( i \)’th left and right singular vectors of \( A \), and \( U \) and \( V \) are the corresponding matrices where the singular vectors are arranged in columns, yielding matrices with orthonormal columns. When \( A \) is inverted and applied to equation (4), the following naive solution is obtained:

\[
x = A^{-1}b = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i.
\]
The fact that the singular values tend to zero and that each term of the summation is inversely proportional to its singular value result in increasing amplification of the potentially noise-dominated vector product $u_i^T b$ as $i \to m$. To overcome this problem, the following two regularization techniques are considered:

**Truncated singular value decomposition (TSVD)** is simply, as the name indicates, a truncated version of SVD that stems from the idea that the last terms of the summation stated in equation (6) correspond to very small singular values, and these singular values can potentially amplify the measurement noise yielding meaningless reconstructions. The approach for regularizing the solution is thus to truncate the number of singular values contributing to the final solution,

$$x_k = \sum_{i=1}^{k} \frac{u_i^T b}{\sigma_i} v_i, \quad \text{where} \quad k \leq n. \quad (7)$$

**Standard Tikhonov regularization** computes the regularized solution by solving the following minimization problem in a least squares sense:

$$\min \left\{ \|Ax - b\|_2^2 + \lambda \|x\|_2^2 \right\}, \quad (8)$$

where $\lambda$ is the relaxation parameter that controls the weight of the regularization term in the overall minimization problem. Note that, in standard form, the regularization term is simply the energy of the solution vector $x$. If we include the SVD of $A$ into the minimization problem, it can be shown that the Tikhonov solution $x_\lambda$ can be written as follows,

$$x_\lambda = \sum_{i=1}^{m} \varphi_i^{[\lambda]} u_i^T b / \sigma_i v_i, \quad \text{where} \quad \varphi_i^{[\lambda]} = \frac{\sigma_i}{\sigma_i^2 + \lambda^2}. \quad (9)$$

In this case, each term of the SVD is weighted with a filter coefficient $\varphi_i^{[\lambda]}$, that basically reduces the undesired amplification of noise caused by the smallest singular values, but with a smoother transition compared to that of the TSVD method ($\varphi_i^{[\lambda]} \approx 1$ when $\sigma_i > \lambda$, and $\varphi_i^{[\lambda]} \approx 0$ when $\sigma_i < \lambda$).

### 2.2.2. Iterative methods

Iterative regularization relies on matrix multiplications or row action methods, which makes these methods often suitable for large scale problems where an explicit factorization of $A$ is too cumbersome. Each iteration of an iterative method yields a regularized solution, initially approaching the sought exact solution, but as the number of iterations increases, the algorithm converges to undesired solutions. This behavior is often referred to as semi-convergence. Hence, the number of iterations plays in these methods the role of regularization parameter. In the following, three different iterative regularization methods are considered:

**Conjugate gradient least squares (CGLS)** is referred to as the most stable implementation of the conjugate gradient method solved in a least squares sense. The core of the algorithm can be summarized in the following five statements:

$$\alpha_k = \|A^T r^{(k-1)}\|_2^2 / \|Ad^{(k-1)}\|_2^2,$$

$$x^{(k)} = x^{(k-1)} + \alpha_k d^{(k-1)},$$

$$r^{(k)} = r^{(k-1)} - \alpha_k A d^{(k-1)},$$

$$\beta_k = \|A^T r^{(k)}\|_2^2 / \|A^T r^{(k-1)}\|_2^2,$$

$$d^{(k)} = A^T r^{(k)} + \beta_k d^{(k-1)}. \quad (10)$$

The algorithm is initialized with $x^{(0)}$, $r^{(0)} = b - Ax^{(0)}$ and $d^{(0)} = A^T r^{(0)}$. With this iteration method, the large singular values seem to converge faster than the rest, meaning that CGLS can effectively regularize the solution when stopping the algorithm long before it converges to the least squares solution.

**Landweber iteration** is one of the classical iterative methods, originally described in 1951. In its simplest form, the algorithm leads to the following regularized solution,

$$x^{(k)} = x^{(k-1)} + \lambda_k A^T r^{(k-1)}, \quad (11)$$

where $\lambda_k$ is a relaxation parameter that can be updated at each iteration $k$, and $r^{(k-1)}$ is the residual vector of the last iteration. In this study, $\lambda_k$ is updated in each iteration following a line search strategy.
Kaczmarz method is one of the most popular row action methods. These are methods that use only one row of $A$ at a time to update the iterative solution, in particular for the Kaczmarz method,\textsuperscript{7}

$$x \leftarrow x + \lambda \frac{b_i - a_i^T x}{\|a_i\|_2^2},$$  \hspace{1cm} \text{(12)}

where $\lambda$ is the relaxation parameter, $a_i^T$ is the $i$’th row of $A$, and for each iteration of the method, all the rows of $A$ are swept from top to bottom, that is, $i = 1, \ldots, m$. Note that in this case, the relaxation parameter is kept constant. Although $\lambda$ was set to 1 in the original paper of Kaczmarz\textsuperscript{9}, other values between 0 and 2 can also be used ($\lambda \in (0, 2)$). In the present investigation, $\lambda$ was set to 0.2 due to convergence rate concerns.

3. PERFORMANCE INDICATORS

The presented numerical methods are assessed by means of the following two performance indicators:

Mean square error (MSE) is probably the most common way of quantifying the difference between an estimate and its corresponding true function. In vector notation, the normalized MSE can be calculated as follows,

$$\hat{\delta}_x = \frac{\|x_{\text{exact}} - x\|_2^2}{\|x_{\text{exact}}\|_2^2}.$$  \hspace{1cm} \text{(13)}

Structural similarity (SSIM) index is a method developed for assessing the similarity between two images, and thus, it provides a measure of quality in image reconstruction problems.\textsuperscript{10} This index can take values between 0 and 1. The closer the value is to 1, the higher the similarity. In the current study, the images correspond to sound pressure reconstructions obtained with the numerical methods under investigation. The nature of the SSIM index is far from the physical meaning underlying the MSE. It alternatively tries to take advantage of known features of the human visual system, which makes it an interesting approach when assessing a sound visualization technique such as acousto-optic tomography.

4. SIMULATION STUDY

Two different and well-defined sound fields are investigated in the following simulation study, namely the sound fields generated when two monopoles are radiating in phase and in antiphase (dipole case), respectively. Hence, one sound field is dominated by the constructive interference of the two monopoles and the other one features a clear destructive interference between the two point sources. This is illustrated on the

![Fig. 2](image-url)
right hand side of figure 2, where the top panels show a instantaneous representation of the acoustic field synthesized by the two monopoles driven in phase and the bottom panels correspond to the dipole case. The different nature of these two acoustic fields makes it possible to draw conclusions about the performance of the numerical methods with a certain independence of the actual sound fields used in the calculations.

The simulation study starts computing the Radon transform of the two acoustic fields under investigation, see left hand side in figure 2. These data are then properly arranged in a column vector that constitutes the right-hand side vector $\mathbf{b}$ defined in the matrix equation (4). Once the matrix $\mathbf{A}$ is also set up, the inverse problem is solved with the regularization methods presented in section 2. The evolution of the regularized solutions as a function of the regularization parameters is shown in figure 3. As can be seen, the overall results obtained for the two sound fields under investigation are approximately the same. This suggests that the results are fairly independent of the actual sound fields used to examine the performance of the numerical methods. The optimal value of the regularization parameter of each method can be identified by looking at the minimum/maximum of the MSE/SSIM curves, respectively. In some cases, this optimal value is not the same for both the MSE and the SSIM index. The two top panels show the results corresponding to the two sound fields under investigation are approximately the same. This suggests that the numerical methods with a certain independence of the actual sound fields used in the calculations.

Fig. 3 – Performance of the numerical methods as a function of their regularization parameters.

(a) Two monopoles in phase.

(b) Dipole.
FBP method in terms of both MSE and SSIM index. The second row of plots in figure 3 shows the TSVD results. The best regularized solution provided by this method yields approximately the same MSE and SSIM index as the CGLS method, and thus, the overall performance is worse than that of the FBP method. The two bottom plots in figure 3 present the results obtained with Tikhonov regularization. In this case, the smoother regularization of the singular values in comparison to the sudden truncation of the TSVD method yields the best results out of the five numerical methods under investigation.

The performance of each method has also been investigated as a function of the signal-to-noise ratio (SNR), see figure 4. As expected, the MSE increases while the SSIM index decreases as the SNR is worsened.

**Fig. 4 – Quality of the reconstructions as a function of the SNR. For each SNR, the optimal regularization parameter is first identified in terms of MSE.**

In general, the results are similar for both the two monopoles radiating in phase and the dipole. Interestingly, the overall values of MSE are slightly higher for the dipole case, while the performance of the reconstruction methods in terms of SSIM index is slightly worse for the two-monopoles case. This indicates the convenience of considering more than one performance indicator when assessing the quality of the reconstructions. The ratio SSIM/MSE is thus presented in the two bottom plots of figure 4 in order to have an idea of the overall performance of the methods. In short, the larger the ratio, the better the performance. Ideally, when SSIM → 1 and MSE → 0, then the ratio SSIM/MSE → ∞. As can be seen, Tikhonov regularization yields the best results for SNRs higher than 12 dB in the two-monopoles case, and 6 dB in the dipole case. In this range of SNRs, the good performance of the Tikhonov regularization is followed by the Landweber and Kaczmarz...
methods. The classical FBP method lies in between this group and the other group characterized by poorer performance consisting of the CGLS method and the TSVD. In fact, these two methods yield almost the same results. At lower SNRs, the Landweber and Kaczmarz algorithms yield better results than the Tikhonov regularization. The last two plots in figure 4 also show the ratio SSIM/MSE for a microphone array that would sample the acoustic field at the same positions where the numerical methods reconstruct the pressure. The microphone results are simply biased by the random noise introduced in the simulations. It is interesting to see that the Tikhonov results are better than the ones obtained with the microphone array for SNRs below 23 dB in the two-monopoles case, and 21 dB in the dipole case. The other reconstruction methods progressively outperform the microphone results as the SNR worsens. It is in fact remarkable that all the reconstruction methods investigated for acousto-optic tomography yield better results than the microphone array for SNRs below 14 dB.

Figure 5 shows the sound pressure maps obtained with the microphone array, the FBP method, and the Tikhonov regularization. The simulation is adjusted to a SNR of 18 dB. The first and second rows of plots correspond to the two-monopoles and dipole cases, respectively. As can be seen, the main features of the two sound fields under investigation are clearly identified in all the reconstructions. The microphone results are though visibly more affected by noise, whereas the other pressure maps are smoother. However, a one-dimensional plot of the diagonal of the pressure maps (see figure 6) reveals that the FBP does not reconstruct the peak values accurately. Instead, the Tikhonov results show the best agreement with the theoretical values of the pressure, which is also in agreement with the results presented in figure 4.

Fig. 5 – Examples of sound field reconstructions for a SNR of 18 dB. The top row corresponds to the two-monopoles case, and the second row shows the results of the dipole.

Fig. 6 – Diagonal of the pressure maps shown in Fig. 5.
5. EXPERIMENTAL RESULTS

A set of preliminary measurements were carried out in an anechoic room of 1000 m$^3$. A picture of the experimental setup can be seen in figure 7. In order to make the measurements somehow comparable to the simulation results, two loudspeakers, whose centers were separated by 20 cm, were driven in phase and in antiphase with a 2.4 kHz pure tone. The measured Radon transforms are shown on the left hand side of figure 8, and they are in good agreement with the instantaneous representation examined in the simulation study (cf. figure 2). The corresponding reconstructions based on the FBP method and Tikhonov regularization are presented on the right hand side of figure 8. The first and second rows correspond to the two-monopoles and dipole cases, respectively. As observed in the simulations, both FBP and Tikhonov yield smooth pressure maps, although it must be noted in this case that the SNR during the measurements was around 30 dB approximately. A closer look, in particular among the dipole reconstructions, shows that the FBP method leads to lower pressures at the peaks of the reconstruction maps. This is in agreement with the simulations results shown in figure 6. The rest of the numerical methods were also examined and led to similar pressure maps as the ones presented in figure 8.

![Image of experimental setup](image_url)

Fig. 7 – Experimental setup used for sound field reconstruction based on acousto-optic tomography.

![Experimental results images](image_url)

Fig. 8 – Experimental results obtained in the anechoic room with a SNR of about 30 dB. The top row shows the results of the two-monopoles case, and the second row shows the results of the dipole case.
6. SUMMARY AND CONCLUSIONS

The reconstruction of acoustic fields based on acousto-optic tomography has been implemented with numerical methods. This is possible by describing the tomographic reconstruction as a discrete inverse problem. Five different regularization techniques have been examined to tackle the ill-posedness of the tomographic problem. The investigated techniques have been assessed by means of two performance indicators, namely the MSE and the SSIM index. The latter is a quality parameter originally designed for image processing, and thus, it assesses the quality of the tomographic reconstructions from a visual point of view. This is a very suitable alternative to the classical MSE parameter, especially if we bear in mind that acousto-optic tomography is basically a sound visualization technique. Both the MSE and the SSIM index identify the optimal regularization parameter of each method in the same range of values, though they do not provide exactly the same value. Hence, this points out that to rely the quality of the reconstruction on a visual inspection procedure of the resulting acoustic image is not a guarantee that the optimum solution is achieved, though it will typically be fairly close to it. The overall results show that the numerical methods can lead to successful sound field reconstructions. In particular, three out of the five methods (Tikhonov regularization being the best, Landweber iteration and Kaczmarz method) have consistently proved superior to the classical FBP method. The numerical methods also seem fairly robust to low SNRs. In fact, when comparing the results to that of a microphone array, Tikhonov regularization yields better reconstructions for SNRs lower than 21 dB approximately, and all the tomographic methods present better SSIM/MSE ratios for SNRs lower than 14 dB.

The main challenge of using numerical methods for reconstructing acoustic fields is the choice of the optimal regularization parameter. Typically, the optimal value differs from problem to problem. Therefore, it is necessary to investigate in the future the use of parameter choice methods that could make this choice automatically without compromising the final reconstruction significantly. On the other hand, it should also be noted that numerical methods can in principle be implemented for any tomographic scheme, while the classical FBP method is normally restricted to parallel and fan-shaped scanning configurations. This makes acousto-optic tomography more versatile, and it can ease the complexity of the measurement setup currently required.

REFERENCES


