OPTIMIZED SPATIAL RESAMPLING FOR MICROPHONE ARRAY BEAMFORMING

Chao Lin\textsuperscript{1}, Aneline Grand\textsuperscript{2}, Stéphan Tassart\textsuperscript{2}, Jean-Yves Tourneret\textsuperscript{1}, Olivier Besson\textsuperscript{1} and Laurent Saïd\textsuperscript{2}

\textsuperscript{1} TéSA Laboratory, University of Toulouse, 14-16 Port Saint Etienne, 31071 Toulouse cedex 7, France
\textsuperscript{2} ST-Ericsson, 3GP/3GM/PSA Audio Algorithms, 29 Boulevard Romain Rolland, 75669 Paris cedex 14, France

chao.lin@tesa.prd.fr, \{aneline.grand,stephan.tassart,laurent.said\}@stericsson.com, olivier.besson@isae.fr, jean-yves.tourneret@enseeiht.fr

ABSTRACT

Due to the specialty of portable devices, the microphone array configurations usually depend on the device manufacturing constraints, and thus are various and generally irregular. It is then more difficult for the algorithm designers to establish a generic processing. The main contribution of this paper is an optimized spatial resampling procedure for beamforming with irregular 3-D sensor arrays. Beam pattern analysis show that, via spatial resampling, it is possible to apply beamformer coefficients designed for Uniform Linear Array (ULA) to signals captured by arrays having arbitrary configurations and thus carry out beamforming and spatial filtering in both azimuth and zenith directions. Moreover, this algorithm may benefit to other multi-sensor audio processing methods that require specific sensor array configuration.

Index Terms— Microphone array, spatial resampling, beamforming

1. INTRODUCTION

Beamforming, as spatial filtering, is a microphone array processing allowing to capture the signals coming from a desired direction while minimizing the contributions of other directions. For example, the widespread Minimum Variance Distortionless Response (MVDR) beamforming algorithm consists of minimizing the output signal power while preserving the signal arriving from the look direction (design method given in [1]). This method relies on the full knowledge of the array manifold to establish MVDR coefficients. However, the portable device manufacturing industry demands irregular microphone array configurations. The architectural challenge is to separate the embedded array processing in two steps (as in Fig. 1), first including a preprocessing specific to the actual array configuration, then including a general MVDR design for ULA (for instance a third-part component provided for several clients).

The approach of interpolated arrays was initially proposed in [2] and was later extended in [3, 4, 5]. This approach builds a virtual array configuration (e.g., an ULA) through the interpolation of an arbitrary array geometry. The implementations of Root-MUSIC and spatial smoothing with interpolated arrays were for instance found promising in [6]. The combination of interpolated arrays with MVDR-type beamforming methods has been applied to Non-uniform Linear Array (NLA) beamforming problems in [7], where a good reduction of array size and sensor number has been reported. The theoretical background necessary for understanding interpolated arrays as a wavefield resampling problem is considered in [8].

In this paper, we document performance results from an interpolation array design procedure based on the Cook \textit{et al.} transformation (as labelled in [9]). Unlike the azimuth-only system with NLA configuration proposed in [7], the proposed algorithm provides a full azimuth and zenith Field Of View (FOV) system with irregular 3-D array configuration. The proposed algorithm can be interesting for other multi-sensor applications using irregular 3-D sensor configurations and requiring a full discrimination capability in both azimuth and zenith directions.

The paper is organized as follows. In Section 2, notations of array signal models are presented. A brief introduction of spatial resampling is given in Section 3. In Section 4, we present the optimized spatial resampling algorithm for beamforming. Representative simulation results and real sound processing results of 3-D arrays are given in Section 5, followed by concluding remarks in Section 6.

2. SIGNAL MODEL

Any general square integrable sound field $\psi$ is the superposition of incident plane waves. In the Fourier domain, this sound field evaluates at an arbitrary observation point $x$ as

$$\psi(x, \omega) = \int_{|u|=1} \Psi(\omega u) \exp\left(-j\omega u^T x\right) du \quad (1)$$

where $\omega$ is the angular frequency, $k$ is the wave number, $u$ is the incident direction vector and $\Psi(\omega u)$ is the radiation density. We collect the wave number and the incident direction vector into the wave vector $k = ku$. The angular frequency and the wave number are related by the dispersion relation $c^2 k^2 - \omega^2 = 0$, where $c$ is the speed of sound.

The sound field is observed in $x$ by a sensor with a certain
directivity pattern $\gamma(k)$
\[
s(x, \omega) = \int_{|u|=1} \Psi(k) \gamma(k) \exp\left(-j k^T x\right) \, du. \tag{2}
\]
The contribution from the sensor can be separated from the actual sound field $\psi$ with the help of a characterizing steering vector $d(x, k) = \gamma(k) \exp\left(-j k^T x\right)$ as follows
\[
s(x, \omega) = \int_{|u|=1} d(x, k) \Psi(k) \, du. \tag{3}
\]

3. SPATIAL RESAMPLING

In our context, the basic idea of spatial resampling is to evaluate the sound field at location $y$ given an arbitrary sampling array $x = [x_1, x_2, \cdots, x_N]$ (irregular in terms of sensor positioning). This resampling technique can be generalized to a global sensor mapping so as to get a regular array made of $M$ virtual sensors $y$ (e.g., an ULA). Beamforming is then applied to the sensor signals of the virtual array as shown in Fig. 1.

![Sensor mapping as a beamforming preprocessing (through a transformation matrix $T(\omega)$ in frequency domain)](image)

A linear mapping of the physical steering vector $d(x, k) = [d(x_1, k), \cdots, d(x_N, k)]^T$ is related to a desired steering vector as $d(y, k)$:
\[
d(y, k) = d(x, k)^T v(\omega) + e(k) \tag{4}
\]
where $v(\omega)$ is a $N \times 1$ vector. The interpolation error in the Fourier domain denoted as $e(\omega)$ can be expressed as
\[
e(\omega) = \int_{|u|=1} \Psi(k) e(k) \, du \tag{5}
\]
A suboptimal option for minimizing $e(\omega)$ independently from $\Psi(k)$ is to minimize $e(k)$ on every direction of interest (cf. [5]). In the context of spatial filtering, there is no reason to emphasis a specific direction. One can formulate the problem as minimizing the following cost function $J(v, \omega)$ for any $v$
\[
J(v, \omega) = \int_{|u|=1} |e(k)|^2 \, du. \tag{6}
\]
For omnidirectional and equalized sensors, $\gamma(k) = 1$.

The optimization problem is then simplified and integral calculations lead to the spherical bessel function $j_0$ (cf. [8])
\[
\int_{|u|=1} d(x, k) du = 4\pi j_0(\|x\|) = 4\pi \frac{\sin(\|x\|)}{\|x\|} \tag{7}
\]
\[
\int_{|u|=1} d(y, k) du = 4\pi j_0(\|y-x\|) \tag{8}
\]
where $\|\cdot\|$ denotes the vector norm. Using (7) and (8), we obtain the standard least square (LS) representation of (6)
\[
J(v, \omega) = \frac{1}{4\pi} Q(\omega) v - 2 v^H f(\omega) + 1 \tag{9}
\]
where $f(\omega) = \{ j_0(k \|y-x_n\|) \}_{n=1}^N$ is an $N \times 1$ vector and $Q(\omega) = \{ j_0(k \|x_p-x_r\|) \}_{p,r=1}^N$ is an $N \times N$ matrix.

Solving the LS problem (9) for a given $\omega$ yields the solution $v(\omega)$ that minimizes $J(v, \omega)$
\[
v(\omega) = Q^{-1}(\omega) f(\omega). \tag{10}
\]

This principle can be extended to non-omnidirectional or non-equalized sensors. In this case, (4) must be modified in such a way that $d(x, k)$ represents the real physical steering vector including the directivity pattern. Only simple directivity patterns result in a closed form resampling matrix.

4. OPTIMIZED INTERPOLATION ARRAY FOR BEAMFORMING

In order to interpolate an array of sensors located in $y = [y_1, y_2, \cdots, y_M]$, the reasoning from section 3 can be used. In particular, one can look for a linear combination of the physical steering vector synthetizing the desired steering vector $d(y, k)$
\[
d(y, k) = T d(x, k) + e(k). \tag{11}
\]
When physical sensors are equalized and omnidirectional, (10) can be reused (for improved readability, we drop the dependency on $\omega$)
\[
\begin{align*}
T^H &= Q^{-1} F \tag{12}
\end{align*}
\]
with $F = \{ j_0(k \|y_i-x_n\|) \}_{n=1}^N,i=1\cdots M$ an $N \times M$ matrix.

Different authors (cf. [9]) have noted that LS may lead to large norm problems and an ill-conditionned matrix $T$. Robustness can be improved by adding a norm constraint $|t_i| \leq \alpha$ for each line from $T$ (see [9, 7]).

The algorithm solving the constrained LS (CLS) problem is found in [10]. The algorithm relies on a singular value decomposition of $Q$. In essence, CLS substitutes each singular value $\sigma_n$ by $\sigma_n(1 + \lambda\sigma_n^{-2})^{1/2}$, solves the standard LS problem with the modified matrix and chooses the smallest $\lambda \geq 0$ that verifies the norm constraint. Note that large singular values are not significantly modified whereas small singular values are replaced by $\lambda^{1/2}$. 

201
The usage of the $\nu$-reduction of the sampling matrix is reported in [8] in order to bound $|\varepsilon(\omega)|$. The matrix $Q_\nu$ is constructed by removing from $Q$ all singular values smaller than the threshold $\nu \geq 0$. A closer look to this procedure indicates that it may be considered as a sub-optimal alternative to the standard CLS algorithm. Indeed, the $\nu$-reduction procedure keeps unchanged singular values above the threshold $\nu$ and substitutes any other singular value $\sigma_p$ by $\infty$. If the smallest threshold $\nu$ is chosen in order to verify the norm constraint, the reduction procedure becomes an alternative to the standard CLS.

5. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed algorithm by introducing a beam pattern analysis. In the following example, we apply the algorithm to an irregular 3-D array with its 4 sensors positioned at $(0,0,-1), (0.7, 0.5, 1), (0.7, -0.5, 0)$ and $(-1, 0, 1)$ (in cm) within the Cartesian system shown in Fig. 2. Sensor spacing of the virtual ULA is set to $d = 0.75$cm to avoid spatial aliasing up to the frequency $f_{\text{max}} = \frac{c}{2R}$. Concerning the placement of the virtual array, Friedlander et al. proposed in [4] a “rule of thumb”. In essence, the virtual array placement should satisfy $|y_i| \leq R$, where $R$ denotes the radius of the smallest sphere centered at the coordinate origin and enclosing all real arrays. Here, we choose to place the virtual ULA along the DOI and centered at the origin as shown in Fig. 2.

According to our application case, the beamforming weight vectors and the mapping matrix $T$ are computed at frequency points between 0.1kHz and 24kHz with a sampling frequency of 48kHz and FFT length of 256. MVDR weight vectors are computed with a standard MVDR design proposed in [1] for the ULA virtual array (diffuse noise field, diagonal loading of the spatio-spectral correlation matrix) with the direction of interest (DOI) $\phi_d = 0, \theta_d = \frac{\pi}{2}$ in azimuth and zenith directions. As pointed out in [7], the output of an MVDR beamformer is unaffected by a linear transform as long as we use ideal spatio-spectral correlation matrices. Norm-constraint LS algorithm is used and the norm constraint parameter $\alpha$ is set to 2.

Beam pattern measurements are then performed on a 0.1kHz - 24kHz spectrum. The sensors are assumed to be omnidirectional and noise-free. The theoretical far-field beam pattern of the array located in $y$ is given by

$$\|H(k)\|_{2\text{dB}}^2 = -10 \log_{10} \left( \frac{|W(y, k)d(y, k)|^2}{W^{H}(y, k)\Gamma_{\text{VV}}W(y, k)} \right)$$

where $\Gamma_{\text{VV}} = \{e^{-jk(y_r-y_r')}\}_{r,r=1,...,4}$ denotes the wavefront correlation matrix, $W(y, k)$ denotes the beamformer weight vector and $d(y, k)$ denotes the steering vector in the DOI. By representing the vectors in spherical coordinates, the beam pattern is a function of three variables, the frequency $\omega$, the azimuth $\phi$ and the zenith $\theta$.

Fig. 3(a) and Fig. 3(b) show the zenith direction beam patterns obtained with fixed azimuth $\phi = \phi_d$ for the ULA and for the physical array with spatial resampling, while Fig. 3(c) and Fig. 3(d) show the azimuth direction beam pattern obtained with fixed zenith $\theta = \theta_d$ for the two arrays, respectively. The beam pattern of the spatially resampled array with standard MVDR beamforming is close to that of the ULA array in terms of essential characteristics (no attenuation in the DOI, significant attenuation in other directions).

In the following tests, we evaluate the proposed algorithm by using the wideband beam pattern. By separating the contribution of the angular frequency $\omega$ in the wave vector $k$, the theoretical wideband beam pattern of the array $y$ is given by

$$P(u) = \frac{1}{\pi} \int_{-\omega_{\text{max}}}^{\omega_{\text{max}}} |W(y, k)d(y, k)|^2 d\omega.$$  

Fig. 4 shows the wideband beam pattern for ideal ULA and for the irregular 3-D array with spatial resampling in both zenith and azimuth directions. The $-6$dB-mainlobe-width difference between the interpolated array and the ideal ULA is $14^\circ$ in the zenith direction and $28^\circ$ in the azimuth direction.

Note that the selected example just illustrates some aspects and is by no means exhaustive. It should also be noted that the algorithm presented previously is not able to handle all kinds of array configurations with the same performance level. Many physical array configurations may have a direct influence on the results, including the spatial aliasing. Since our mapping system is not capable of recovering lost information in high frequency band, if the physical array placement does not respect the spatial sampling theorem, the virtual ar-
Fig. 3. (a) Beam pattern in zenith direction for ULA (with fixed azimuth $\theta = \theta_d$); (b) Beam pattern in zenith direction for irregular array with spatial resampling (with fixed azimuth $\theta = \theta_d$); (c) Beam pattern in azimuth direction for ULA (with fixed zenith $\phi = \phi_d$); (d) Beam pattern in azimuth direction for irregular array with spatial resampling (with fixed zenith $\phi = \phi_d$).

ray will still suffer from the aliasing problem. Further studies should be conducted on the performance analysis of this optimized interpolation algorithm when applied to various array configurations as well as the criterion which allows array configuration limits to be described.

6. CONCLUSION

In this paper, we proposed an optimized spatial resampling approach for microphone array beamforming dealing with irregular 3-D array configurations. Instead of sampling the sound field, an alternative approach of transforming the physical array observations to that of the ULA virtual array through spatial resampling was introduced. Beam pattern analysis was carried out for the proposed approach. Simulation results show that, with the help of spatial resampling, it is possible to apply the general standard MVDR coefficients of ULA array to the signals captured by arrays with arbitrary configurations and to carry out beamforming in both azimuth and zenith directions.

7. REFERENCES