

ACOUSTIC SOURCE SEPARATION USING THE SHORT-TIME QUATERNION FOURIER TRANSFORMS OF PARTICLE VELOCITY SIGNALS

Hüseyin Hacıhabiboğlu

Graduate School of Informatics, Middle East Technical University (METU)
Çankaya, Ankara, TR-06800, Turkey

ABSTRACT

Quaternion Fourier transforms (QFT) provide a powerful tool for the analysis of signals obtained from vector probes. Acoustic particle velocity is one such signal which can be measured with specially designed microphone arrays. This paper presents a time-frequency source separation method based on the short-time quaternion Fourier transform of acoustic particle velocity signals and the k -plane clustering of the vector part of the resulting representation. Two example cases, one with a single and one with two interfering sources are presented.

Index Terms— Acoustic signal processing, microphone arrays, source separation, quaternions, vector clustering

1. INTRODUCTION

Decomposition of a recorded sound field into its constituent parts is an important step in spatial audio object coding [1][2]. Separation of sound sources from a convolutive mixture, which contains not only the sound sources but also their reflections, is an essential part of such acoustic scene analysis.

Different source separation and beamforming strategies exist for different microphone arrays. Linear, planar or rigid spherical microphone arrays can be used to generate directivity patterns that can be used to separate spatially distinct sources by maximising the target signal and minimising the interference signal [3][4]. Signals obtained from microphone arrays which can measure the acoustic intensity by registering the pressure and particle velocity at the recording position can be used for time-frequency source separation [5][6]. Such algorithms typically process the pressure and the axial components of particle velocity separately. After obtaining time-frequency representations of the pressure signal and the axial components of the particle velocity, these algorithms determine the active intensity vector for each time-frequency bin and generate a time-frequency mask based on the source direction and the directional distribution of intensity vectors.

Recent developments in quaternionic signal processing made it possible to process individual channels of multidimensional signals such as signals from vector sensors jointly [7]. Numerical algorithms used for Fourier transforms and adaptive signal processing have been adapted for quaternionic signals [8]. These algorithms can achieve better results than processing each component separately.

The work reported in this paper is supported by the Turkish Scientific and Technological Research Council (TÜBİTAK) Research Grant 113E513 “Spatial Audio Reproduction Using Analysis-based Synthesis Methods”. The data used in this paper was recorded in METU-SPARG Audio Lab at METU-TAF Modelling and Simulation Research Centre (MODSIMMER). This paper is dedicated to the author’s newborn son, Kerem.

This paper presents a time-frequency source separation method based on the short-time quaternion Fourier transforms of particle velocity signals. It is shown that an informed selection of the transform axis allows grouping directional components of the sound field in planar configurations. Vector clustering is then carried out to separate sources with known directions.

While the present work is related to other time-frequency source separation approaches such as those presented in [5][6][9], the proposed algorithm is fundamentally different in the sense that the obtained signals are processed jointly using a new time-frequency representation based on quaternion Fourier transforms. This representation which incorporates directional information about the source increases the separability of the underlying mixture. In addition the present work uses subspace clustering which allows for selecting time-frequency bins with a dominant contribution from the target source more effectively.

The paper is organised as follows. Sec. 2 briefly reviews quaternion algebra and quaternionic Fourier transforms. Acoustic particle velocity is explained in Sec. 3. Sec. 4 introduces the proposed time-frequency source separation approach. Sec. 5 presents examples of the proposed method for different cases. Sec. 6 concludes the paper.

2. QUATERNION FOURIER TRANSFORMS (QFT)

2.1. Quaternion algebra

Quaternions are hypercomplex numbers which are multidimensional extensions of complex numbers [10]. An arbitrary quaternion, $\mathbf{q} \in \mathbb{H}$ can be expressed as $\mathbf{q} = a + bi + cj + dk$ where $a, b, c, d \in \mathbb{R}$ are real parameters and \mathbf{i}, \mathbf{j} and \mathbf{k} are the quaternion basis elements such that:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1, \quad (1)$$

$$\mathbf{ij} = \mathbf{k}, \quad \mathbf{jk} = \mathbf{i}, \quad \mathbf{ki} = \mathbf{j} \quad (2)$$

$$\mathbf{ji} = -\mathbf{k}, \quad \mathbf{kj} = -\mathbf{i}, \quad \mathbf{ik} = -\mathbf{j} \quad (3)$$

As also evident from the above fundamental properties, quaternion algebra is non-commutative e.g. $\mathbf{pq} \neq \mathbf{qp}$ for $\mathbf{p} \neq \mathbf{q}$. Note that the complex imaginary unit is usually denoted as I to differentiate it from quaternionic imaginary units.

Quaternions can be represented in a variety of ways [11]. In this paper, scalar-vector form, $\mathbf{q} = S(\mathbf{q}) + V(\mathbf{q})$, where $S(\mathbf{q}) = a$ and $V(\mathbf{q}) = bi + cj + dk$ is used. The quaternion conjugate is represented as $\bar{\mathbf{q}} = S(\mathbf{q}) - V(\mathbf{q})$ and the quaternion norm and the inverse quaternion are defined as $\|\mathbf{q}\| = \mathbf{q}\bar{\mathbf{q}}$ and $\mathbf{q}^{-1} = \bar{\mathbf{q}}/\|\mathbf{q}\|$, respectively. Product of two full quaternions can be expressed as:

$$\mathbf{pq} = S(\mathbf{p})S(\mathbf{q}) - \langle V(\mathbf{p}), V(\mathbf{q}) \rangle + S(\mathbf{p})V(\mathbf{q}) + S(\mathbf{q})V(\mathbf{p}) + V(\mathbf{p}) \times V(\mathbf{q}) \quad (4)$$

where $\langle \cdot, \cdot \rangle$ and \times denote the inner and cross products, respectively.

Unit quaternions, $\tilde{\mathbf{q}}$, have unit norm. Full quaternions are quaternions with non-zero scalar parts (i.e. $S(\mathbf{q}) \neq 0$) and pure quaternions, $\boldsymbol{\mu} \in \mathbb{C}_\mu$, are quaternions with zero scalar parts, $S(\boldsymbol{\mu}) = 0$. A full quaternion can be expressed as a combination of a real number and a pure unit quaternion:

$$\mathbf{q} = \gamma + \eta\boldsymbol{\mu}, \quad \gamma, \eta \in \mathbb{R} \quad (5)$$

The quaternionic exponential function is defined as:

$$e^{\mathbf{q}} = e^\gamma e^{\eta\boldsymbol{\mu}} = e^\gamma (\cos \eta + \boldsymbol{\mu} \sin \eta). \quad (6)$$

The quaternionic exponential forms the quaternion Fourier basis when \mathbf{q} is a pure quaternion (i.e. $\gamma = 0$).

2.2. Fourier transforms of Quaternionic Signals

Due to the non-commutativity of quaternionic algebra, two types of quaternion Fourier transforms (QFT) can be defined for a quaternionic signal, $\mathbf{f}(t) : \mathbb{R} \rightarrow \mathbb{H}$. These are the left-sided and right-sided transforms given respectively as:

$$\mathbf{F}_\nu^L(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\nu\omega t} \mathbf{f}(t) dt, \quad (7)$$

$$\mathbf{F}_\nu^R(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathbf{f}(t) e^{-\nu\omega t} dt, \quad (8)$$

where $\nu \in \mathbb{C}_\mu$ is the transform axis. Only the left-sided QFT is used in this paper. However, it is straightforward to obtain similar results with the right-sided transform.

QFT pairs and relevant theorems are discussed in detail elsewhere [11]. However, three important properties are repeated here due to their relevance: linearity, symmetry and time-shift.

- *Linearity*: $\mathbf{f}(t) = \sum_i a_i \mathbf{f}_i(t) \Leftrightarrow \mathbf{F}(\omega) = \sum_i a_i \mathbf{F}_i(\omega)$ iff. $\{a_i\} \in \mathbb{R}$ and $\mathbf{f}_i(t) \Leftrightarrow \mathbf{F}_i(\omega)$
- *Symmetry for pure quaternionic signals*: For $\mathbf{f}(t) \in \boldsymbol{\mu}\mathbb{R}$, $\mathbf{F}(\omega) = -\overline{\mathbf{F}(\omega)}$
- *Time shift*: $\mathbf{f}(t - \tau) \Leftrightarrow e^{-\nu\omega\tau} \mathbf{F}(\omega)$

Discrete-time and discrete quaternion Fourier transforms are also defined and fast implementations based on standard FFT algorithms exist [12]. This paper uses the implementation in [13]. The derivations in the following chapters will use continuous-time signals and transforms for the sake of clarity of the exposition.

3. ACOUSTIC PARTICLE VELOCITY

Pressure and particle velocity are two components which specify a local sound field at a given point. Given a pressure field, $p(\mathbf{x}, t)$, the time derivative of particle velocity, $\mathbf{u}(\mathbf{x}, t)$ is related to the pressure, $p(\mathbf{x}, t)$, by Euler's equation:

$$\frac{\partial \mathbf{u}(\mathbf{x}, t)}{\partial t} = -\rho_0^{-1} \nabla p(\mathbf{x}, t) \quad (9)$$

where $\rho_0 \approx 1.225 \text{ kg/m}^3$ is the ambient density. For plane waves, the relation between the pressure and the particle velocity is

$$\mathbf{u}(\mathbf{x}, t) = (\rho_0 c)^{-1} p(\mathbf{x}, t) \mathbf{n} \quad (10)$$

where $c \approx 340 \text{ m/s}$ is the speed of sound and $\mathbf{n} \in \mathbb{R}^3$ is the unit vector in the wave direction. Note that, the actual direction depends not only on this vector but also on the polarity of the pressure signal.

In a typical multipath environment such as a room, the sound field will consist of a superposition of plane waves incident from different directions. The total particle velocity when there are K sources and their $\sum_k L_k$ associated reflections can be given as:

$$\mathbf{u}(t) = \sum_{k=1}^K \sum_{l=0}^{L_k} \mathbf{u}_{kl}(t) \quad (11)$$

where $\mathbf{u}_{kl}(t) = \gamma_{kl} u_{k0}(t - \tau_{k,l}) \mathbf{n}_{k,l}$. Here, γ_{kl} , $\tau_{k,l}$, and $\mathbf{n}_{k,l}$ are the gain, delay and direction associated with the l^{th} reflection of the k^{th} source, respectively, and $u_{k0}(t)$ is the amplitude of the particle velocity from the direct path of the k^{th} source with $\gamma_{k0} = 1$ and $\tau_{k0} = 0$.

Acoustic intensity probes [14][15] typically measure both the pressure and the particle velocity which are then used in the calculation of acoustic intensity. Measurement of particle velocity [16][17] is beyond the scope of this paper and is not discussed further. It is assumed that particle velocity measurements from such arrays are available.

4. TIME-FREQUENCY SOURCE SEPARATION

4.1. Signal Model

The quaternionic signal model used in this paper is constructed by designating the axial components of the particle velocity as the components of a pure quaternionic signal. For each plane wave component:

$$\mathbf{u}_{kl}(t) = u_{kl}(t) \boldsymbol{\mu}_{kl} \quad (12)$$

where $u_{kl}(t) \in \mathbb{R}$ is the (scalar) particle velocity due to the reflection l of source k and $\boldsymbol{\mu}_{kl} = n_{kl,x} \mathbf{i} + n_{kl,y} \mathbf{j} + n_{kl,z} \mathbf{k} \in \mathbb{C}_\mu$ is a pure unit quaternion in the direction of the plane wave.

The linearity property of QFT, when applied to (11) allows expressing the QFT of particle velocity as:

$$\mathbf{U}(\omega) = \sum_{k=1}^K \sum_{l=0}^{L_k} \mathbf{U}_{kl}(\omega) \quad (13)$$

where by using the time-shift property of QFT:

$$\mathbf{U}_{kl}(\omega) = \gamma_{kl} e^{-\nu\omega\tau_{kl}} \mathbf{U}_{k0}(\omega) \overline{\boldsymbol{\mu}_{k0}} \boldsymbol{\mu}_{kl} \quad (14)$$

with $\nu \in \mathbb{C}_\mu$ the transform axis and $\mathbf{u}_{k0}(t) = u_{k0}(t) \boldsymbol{\mu}_{k0} \in \boldsymbol{\mu}\mathbb{R}$ such that:

$$\begin{aligned} \mathbf{U}_{k0}(\omega) \overline{\boldsymbol{\mu}_{k0}} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\nu\omega t} u_{k0}(t) dt \\ &= \alpha_{k0}(\omega) + \nu \beta_{k0}(\omega) \end{aligned} \quad (15)$$

where $\alpha_{k0}(\omega), \beta_{k0}(\omega) \in \mathbb{R}$ are the real and imaginary parts of the real Fourier transform of the source signal $u_{k0}(t) \in \mathbb{R}$ such that:

$$\alpha_{k0}(\omega) = \Re\{U_{k0}(\omega)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \cos(\omega t) u_{k0}(t) dt, \quad (16)$$

$$\beta_{k0}(\omega) = \Im\{U_{k0}(\omega)\} = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sin(\omega t) u_{k0}(t) dt, \quad (17)$$

and $|U_{k0}(\omega)| = [\alpha_{k0}(\omega)^2 + \beta_{k0}(\omega)^2]^{1/2}$. Expressing the phase angle of the real Fourier transform as:

$$\Theta_{k0}(\omega) = -\tan^{-1} \left(\frac{\beta_{k0}(\omega)}{\alpha_{k0}(\omega)} \right), \quad (18)$$

the QFT of the l^{th} reflection due to k^{th} source can be specified using the source signal spectrum:

$$\mathbf{U}_{k0}(\omega) = |U_{k0}(\omega)| \gamma_{kl} e^{-\nu \Phi_{kl}(\omega)} \boldsymbol{\mu}_{kl}. \quad (19)$$

where $\Phi_{kl}(\omega) = \omega \tau_{kl} + \Theta_{k0}(\omega)$. The quaternion factor in (19) is a product of the full quaternion, $e^{-\nu \Phi_{kl}(\omega)}$, with the pure quaternion, $\boldsymbol{\mu}_{kl}$. Denoting $\eta_{kl}(\omega) = \cos \Phi_{kl}(\omega)$ and $\xi_{kl}(\omega) = \sin \Phi_{kl}(\omega)$, this quaternion factor is given as:

$$\psi_{kl}(\omega) = \xi_{kl} \langle \nu, \boldsymbol{\mu}_{kl} \rangle + \eta_{kl} \boldsymbol{\mu}_{kl} + \xi_{kl} (\boldsymbol{\mu}_{kl} \times \nu) \quad (20)$$

It may be observed that the locus of the vector part of the QFT of each plane wave component at each frequency bin is an ellipse with axes defined by $\boldsymbol{\mu}_{kl}$ and $\boldsymbol{\mu}_{kl} \times \nu$. The signal model in the frequency domain can be expressed as:

$$\mathbf{U}(\omega) = \sum_{k=1}^K |U_{k0}(\omega)| \sum_{l=0}^{L_k} \gamma_{kl} \psi_{kl}(\omega). \quad (21)$$

The purpose of the proposed algorithm is to separate $\mathbf{U}_{k0}(\omega)$ given the total particle velocity and the direction of the source, $\hat{\mathbf{n}}_{k0}$.

4.2. Short-time Quaternion Fourier Transform

Short-time quaternion Fourier transform (STQFT) is used here for the time-frequency analysis of the particle velocity signal. Since QFT possesses linearity with real coefficients only, the employed windowed QFT is defined to be very similar to regular overlapped short-time Fourier transform (STFT) using real windows:

$$\mathcal{U}(\omega, \tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\nu \omega t} w(t - \tau) \mathbf{u}(t) dt \quad (22)$$

where $w(t)$ is a real valued window function. The main difference between regular STFT and STQFT as used in this paper is the conjugate symmetry between the positive and negative frequency parts for pure quaternionic signals such that $\mathcal{U}(\omega, \tau) = -\overline{\mathcal{U}(-\omega, \tau)}$.

An important aspect of the proposed method is the selection of the transform axis, ν , for the QFT. If an estimate of the source direction, $\boldsymbol{\mu}_{Kk0}$ is available, ν is selected to be a pure unit quaternion orthogonal to that direction so that:

1. The scalar part of $\psi_{k0}(\omega, \tau)$ is zero since $\langle \nu, \boldsymbol{\mu}_{kl} \rangle = 0$,
2. The vector part of $\psi_{k0}(\omega, \tau)$ is on a circle lying on the plane spanned by $\boldsymbol{\mu}_{kl}$ and $\boldsymbol{\mu}_{kl} \times \nu$. The radius of the circle is determined by the magnitude spectrum of the corresponding source signal.

Fig. 1 shows the linear, elliptical and circular loci of the vector part of STQFT for the same signal $\mathbf{f}(t) = \sin(400\pi t) \boldsymbol{\mu}$, where $\boldsymbol{\mu} = \frac{1}{\sqrt{3}}(\mathbf{i} + \mathbf{j} + \mathbf{k})$, with three different transform axis selections: coincident, non-orthogonal, and orthogonal. Each vector represents the vector part of a single time-frequency bin.

4.3. Vector Clustering and Cluster Selection

The separability of different components in the mixture increases with the STQFT. It was shown above that STQFT of individual plane waves lie on planes spanned by $\boldsymbol{\mu}_{kl}$ and $\boldsymbol{\mu}_{kl} \times \nu$. It is then necessary to separate these components and obtain the source signal. k -plane clustering which is a simple subspace clustering algorithm [18][19] can be used to cluster data into plane (or in the N dimensional case,

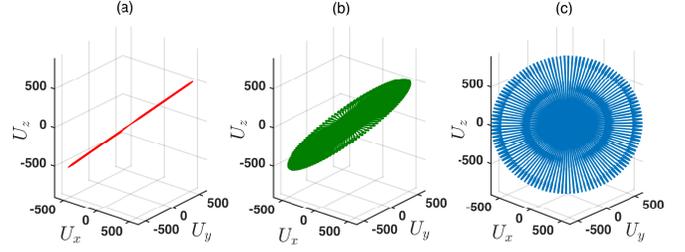


Fig. 1. Loci of the vectorial part of the STQFT of a sinusoidal signal with a frequency of 400 Hz with three different transform axis directions: a) coincident: $\frac{1}{\sqrt{3}}(\mathbf{i} + \mathbf{j} + \mathbf{k})$, b) non-orthogonal: $\frac{1}{\sqrt{2}}(\mathbf{i} + \mathbf{j})$, and c) orthogonal: $\frac{1}{\sqrt{2}}(\mathbf{i} - \mathbf{j})$.

hyperplane) clusters. After convergence, the k -plane clustering algorithm provides k normal vectors corresponding to the plane clusters.

Once clusters are obtained from the STQFT of the particle velocity signal, k -plane clustering is applied and the time-frequency bins that belong to the source are selected. Depending on the choice of the number of clusters in the algorithm, more than one cluster may correspond to the source making it necessary to select the clusters using a thresholding operation. Specifically, clusters for which the magnitude of the dot product between the plane normal vector and the transform axis is below a given threshold close to unity are set to zero. In addition, values of the time-frequency bins for which the ratio of the norms of the vector part and the full quaternion are below another threshold close to unity are also set to zero. The obtained quaternionic time-frequency representation, $\hat{\mathcal{U}}_{K0}(\omega, \tau)$, is then used to reconstruct the separated source particle velocity signal, $\hat{u}_{K0}(t)$ by inverting the transform. The separated source pressure signal is obtained by dividing this signal by the pure quaternion denoting the source direction, i.e. $\hat{u}_{K0}(t) = S(\mathbf{u}_{K0}(t) \bar{\boldsymbol{\mu}}_{K0})$.

5. EXAMPLES

Room impulse responses (RIR) due to source directions positioned in the horizontal plane between 0° and 150° at 30° intervals at a radius of $r = 1.29$ m were measured in the METU-SPARG Audio Lab ($T_{30} \approx 80$ ms) using the open spherical microphone array described in [20] incorporating a total of 13 omnidirectional microphones positioned at the centre and the vertices of a regular icosahedron. The array radius is 3 cm. These RIRs were convolved with 3 s long, energy normalised anechoic samples (female and male speech, music played by viola) from B&O Music for Archimedes CD [21]. The quaternionic particle velocity signals were calculated for individual sources. The signals were then summed to obtain the particle velocity mixture¹. The transform axis is the pure quaternion whose vector part is on the plane normal to the source direction and which has the maximum inner product with the interference source direction.

In order to quantify the separation quality, we use one-at-a-time broadband signal-to-interference ratio (SIR) [22] defined as:

$$SIR = 10 \log_{10} \frac{E\{(\hat{u}_K | u_K)^2\}}{E\{\sum_{j \neq K} (\hat{u}_K | u_j)^2\}} \quad (\text{dB}) \quad (23)$$

¹The mixtures and the separated signals are made available at <http://www.hacihabiboglu.org/samples/icassp16.html>.

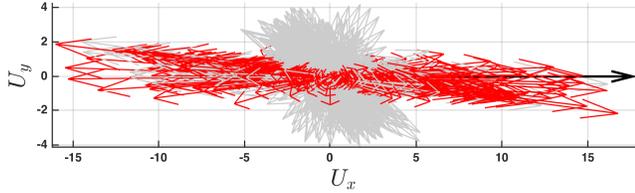


Fig. 2. Top view of the vector part of the STQFT of the two source mixture (grey) and of the separated target sound (red). The black arrow denotes the nominal direction of the target source.

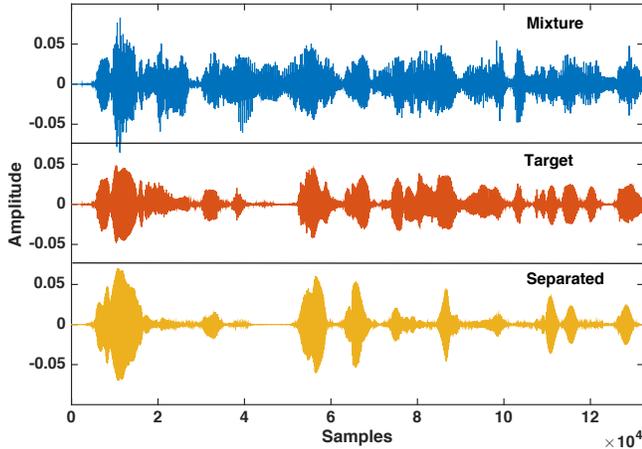


Fig. 3. Mixture, target and separated sound signals for two-source case with female speech as the target and male speech as the interference. The target is at 0° and the interference is at 90° azimuth.

where $E\{(\hat{u}_K|u_K)^2\}$ is the expected value of the output energy of the algorithm for source K when only that source is active, and $E\{\sum_{j \neq K} (\hat{u}_K|u_j)^2\}$ is the expected value of the output energy for source K when all sources except source K is active. In all cases, the target source is the female speech positioned at 0° azimuth. We assume that the source and interference directions are known.

5.1. Two sources: Speech only

There are two concurrent sources in the first case. The interference source is the male speech and is positioned at 30° , 60° , 90° , 120° and 150° . Fig. 2 shows the vector part of the STQFT for frequencies up to 2.2 kHz and the vector part of the time-frequency bins in the cluster set which correspond to the target sound. The actual direction of the target sound is also shown. Each vector in the plot represents the vectorial part of a single time-frequency bin. It may be noted that the selected vectors are aligned well with the source direction. Fig. 3 shows the mixture, target and separated waveforms. It may be observed that the separated source signal is very similar to the target. Informal listening tests revealed that the interference sources are very effectively suppressed and the leakage from the interference source is limited. Fig. 4 shows the SIR improvement for different interference directions. It may be observed that the SIR improvement is above 15 dB which is the preferred lower limit for sound source separation algorithms [23]. The lowest SIR improve-

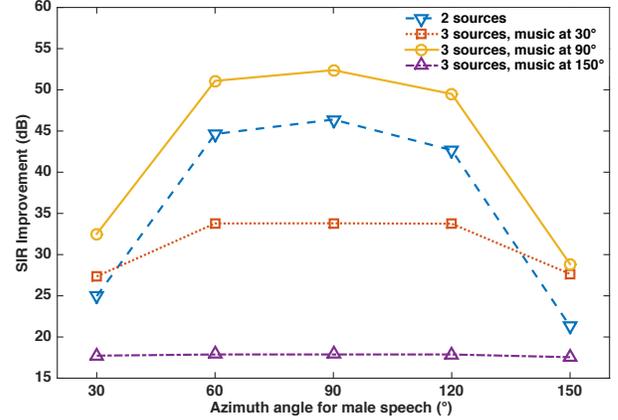


Fig. 4. SIR improvement for the two-source case (blue, dashed line) and three-source case where fixed interference source is at 30° (red, dotted line), 90° (yellow, solid line) and 150° (purple, dash-dot line).

ment for the tested conditions is 21.3 dB when the source and the interference are separated by 150° and can be as high as 46.4 dB as the separation is increased to 90° .

5.2. Three sources: Speech and music

In the second case there are three concurrent sources. The interference sources are the male speech and the music signal. The music source is positioned at 30° , 90° , and 150° and the male speech is positioned at 30° , 90° , 60° , 120° , 150° . Fig. 4 also shows the SIR improvements for these cases. The highest SIR improvements is observed when the interferers are both at 90° . When the angular separation between the target and one of the interferers is 30° or 150° , the SIR improvement is at its lowest. However, this value is still higher than the 15 dB lower limit given in [23]. This performance degradation may be due to the fact that the interference particle velocity causes the source particle velocity to be misaligned from the planar configuration that the proposed method assumes. It is also interesting to note that when there is a strong interference source in the direction orthogonal to the target direction this may increase the SIR improvement even beyond the improvement observed when there is only one interference source. As with the two-source case, informal listening tests showed that the interference sources are effectively suppressed and the leakage from the interference sources is small.

6. CONCLUSIONS

A time-frequency source separation method based on a short-time quaternion Fourier transform of particle velocity signals was proposed in this paper. The method benefits from the joint processing of the three axial components of particle velocity. It was shown that, when the source direction is known, the careful selection of the transform axis will allow for a better separability of the particle velocity components of the source and the interference sources in the short-time quaternion Fourier transform domain. Two examples are presented with two and three source situations. It was shown that the one-at-a-time SIR improvement was above 15 dB for all cases but depends on the angular separation between the target and the interference sources.

7. REFERENCES

- [1] J. Herre, H. Purnhagen, J. Koppens, O. Hellmuth, Jonas Engdegård, Johannes Hilper, Lars Villemoes, L. Terentiv, C. Falch, A. Hölzer, et al., “MPEG spatial audio object coding—the ISO/MPEG standard for efficient coding of interactive audio scenes,” *J. Audio Eng. Soc.*, vol. 60, no. 9, pp. 655–673, 2012.
- [2] J. Vilkamo, T. Lokki, and V. Pulkki, “Directional audio coding: Virtual microphone-based synthesis and subjective evaluation,” *J. Audio Eng. Soc.*, vol. 57, no. 9, pp. 709–724, 2009.
- [3] J. Benesty, J. Chen, and Y. Huang, *Microphone array signal processing*, Springer Verlag, Heidelberg, 2008.
- [4] B. Rafaely, *Fundamentals of Spherical Array Processing*, vol. 8, Springer, 2015.
- [5] B. Günel, H. Hacıhabiboğlu, and A. M. Kondoç, “Acoustic source separation of convolutive mixtures based on intensity vector statistics,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 4, pp. 748–756, 2008.
- [6] X. Chen, W. Wang, Y. Wang, X. Zhong, and A. Alinaghi, “Reverberant speech separation with probabilistic time–frequency masking for b-format recordings,” *Speech Communication*, vol. 68, pp. 41–54, 2015.
- [7] S. Miron, N. Le Bihan, and J. Mars, “Quaternion-MUSIC for vector-sensor array processing,” *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1218–1229, 2006.
- [8] C. C. Took and D. P. Mandic, “The quaternion LMS algorithm for adaptive filtering of hypercomplex processes,” *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1316–1327, 2009.
- [9] O. Yilmaz and S. Rickard, “Blind separation of speech mixtures via time-frequency masking,” *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1830–1847, 2004.
- [10] K. Gürlebeck and W. Sprössig, *Quaternionic and Clifford calculus for physicists and engineers*, John Wiley & Sons, 1997.
- [11] T. A. Ell, N. Le Bihan, and S. J. Sangwine, *Quaternion Fourier Transforms for Signal and Image Processing*, ISTE-Wiley, 2014.
- [12] S.-C. Pei, J.-J. Ding, and J.-H. Chang, “Efficient implementation of quaternion fourier transform, convolution, and correlation by 2-d complex FFT transform,” *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2783–2797, 2001.
- [13] S. J. Sangwine and N. Le Bihan, “Quaternion and octonion toolbox for Matlab,” <http://qtfm.sourceforge.net>, 2013.
- [14] H.-E. de Bree, W.F. Druyvesteyn, E. Berenschot, and M. Elwenspoek, “Three-dimensional sound intensity measurements using Microflown particle velocity sensors,” in *12th IEEE Int. Conf. on Micro Electro Mech. Syst. (MEMS ’99)*, Jan 1999, pp. 124–129.
- [15] S. Oguro, M. Anzai, H. Suzuki, and T. Ono, “A three-dimensional sound intensity probe,” *J. Acoust. Soc. Am.*, vol. 91, no. 4, pp. 2370–2370, 1992.
- [16] F. Jacobsen, “Sound intensity,” in *Springer Handbook of Acoustics*, pp. 1053–1075. Springer Verlag, 1st edition, 2007.
- [17] F. Fahy, *Sound intensity*, E & FN Spon Press, 2nd edition, 1995.
- [18] P. S. Bradley and O. L. Mangasarian, “*k*-plane clustering,” *J. Global Optim.*, vol. 16, no. 1, pp. 23–32, 2000.
- [19] R. Vidal, “A tutorial on subspace clustering,” *IEEE Signal Process. Mag.*, vol. 28, no. 2, pp. 52–68, 2010.
- [20] H. Hacıhabiboğlu, “Theoretical analysis of open spherical microphone arrays for acoustic intensity measurements,” *IEEE/ACM Trans. Audio, Speech, and Lang. Process.*, vol. 22, no. 2, pp. 465–476, 2014.
- [21] Bang & Olufsen, “Music for Archimedes,” Audio CD, 1992.
- [22] D. Schobben, K. Torkkola, and P. Smaragdis, “Evaluation of blind signal separation methods,” in *Proc. 4th Int. Conf. on Indep. Comp. Anal. and Blind Signal Sep. (ICA-1999)*, Aussois, France, Jan. 11-15 1999.
- [23] K. E. Hild, D. Erdogmus, and J. C. Principe, “Experimental upper bound for the performance of convolutive source separation methods,” *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 627–635, 2006.