QUATERNION COMMON FACTOR DECOMPOSITION OF HEAD-RELATED IMPULSE RESPONSE

Zhixin Wang and Cheung Fat Chan

City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong
e-mail: ustc2000@gmail.com, itcfchan@cityu.edu.hk

ABSTRACT

The hypercomplex number quaternion is an extension to the complex number and is widely used in computer graphic, image processing and multiple dimensional linear time-invariant systems. In this paper, the quaternion algebra is applied to head-related impulse response (HRIR) modeling. Four HRIRs measured at different elevations or for different ears are used to construct a quaternion impulse response. A two dimensional quaternion common factor decomposition (QCFD) algorithm is developed to represent each quaternion impulse response as the convolution of two factor impulse responses. By using the proposed QCFD algorithm, quaternion impulse responses with the same elevation will share the same elevation factor while quaternion impulses with the same azimuth will share the same azimuth factor. Experimental results show that the QCFD algorithm has better performance as compared to the traditional two-dimension common factor decomposition (CFD) algorithm.

Index Terms— HRIR, Quaternion, CFD

1. INTRODUCTION

Quaternion algebra introduced by Hamilton in 1843 is a high dimension extension to the traditional complex number [1]. Due to its compact representation and efficient implementation of rotations [2][3], quaternion has been widely used in computer graphic. Recently, quaternion processing is extended to multiple dimensional linear time-invariant (LTI) systems modeling because each component of a quaternion can be used to represent each aspect of a system [4]. For example, in image processing the red, green and blue components of each pixel are used to construct a quaternion number and in seismology the seismic wave along the three orthogonal directions are used as the three imaginary components of a quaternion signal [5]. By exploiting the correlation between different components of the signal, quaternion analysis usually has better performance.

Head-related impulse response (HRIR) or its frequency domain counterpart head-related transfer function (HRTF) describes the filtering effect of human torso, head and pinna to a sound propagating from a specific spatial position to the eardrum of a listener and is the core part in virtual 3D sound synthesis [6]. A lot of research works have been focused on HRIR modeling to reduce the storage requirement of the virtual 3D sound system. Nevertheless, some similarities between HRIRs measured at different positions or for different ears are not exploited by the traditional methods. For example, a pair of HRIRs measured at two positions which are symmetric along the front plane (the vertical plane that goes through both ears) is usually very similar. The similarity of such HRIR pair is the main cause for extensive front-back errors reported in subjective listening tests [7]. Besides, the pair of HRIRs measured at two symmetric positions for different ears are also similar due to the symmetry of human body. To exploit the similarity between different HRIRs, the quaternion algebra is applied in the proposed HRIR model. HRIRs which are measured at different positions and are similar to each other are used to construct a quaternion impulse response. Then a quaternion common factor decomposition (QCFD) algorithm is developed to represent each quaternion impulse response as the convolution of two factor impulse responses. Compared with the real number two-dimension common factor decomposition (CFD) modeling [8], QCFD modeling allows the HRIRs which are similar to each other to share both elevation factor response and azimuth factor response and helps to remove more redundance.

The remaining parts of the paper are organized as follows. Section 2 gives a brief introduction of the quaternion algebra. Section 3 develops the quaternion common factor decomposition algorithm. The developed algorithm is evaluated in Section 4.

2. QUATERNION ALGEBRA

There are extensive literature on the algebra of quaternion numbers and only a brief introduction is given in this section. A quaternion number q is defined as

$$q = q_s + q_x \boldsymbol{i} + q_y \boldsymbol{j} + q_z \boldsymbol{k} \tag{1}$$

where q_s, q_x, q_y and q_z are four real numbers and $\boldsymbol{i}, \boldsymbol{j}$ and $\boldsymbol{k}$ are three imaginary units which satisfy

$$\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -1 \tag{2}$$

The addition and multiplication of two quaternion numbers $p = p_s + p_x\boldsymbol{i} + p_y\boldsymbol{j} + p_z\boldsymbol{k}$ and $q = q_s + q_x\boldsymbol{i} + q_y\boldsymbol{j} + q_z\boldsymbol{k}$ are defined in Eq. (3) and Eq. (4), respectively.

$$p + q = p_s + q_s + (p_x + q_x)\boldsymbol{i} + (p_y + q_y)\boldsymbol{j} + (p_z + q_z)\boldsymbol{k} \quad (3)$$

$$\begin{aligned} pq = \; & p_s q_s - p_x q_x - p_y q_y - p_z q_z \\ & +\boldsymbol{i}(p_s q_x + p_x q_s + p_y q_z - p_z q_y) \\ & +\boldsymbol{j}(p_s q_y - p_x q_z + p_y q_s + p_z q_x) \\ & +\boldsymbol{k}(p_s q_z + p_x q_y - p_y q_x + p_z q_s) \end{aligned} \quad (4)$$

It is worth noticing that the quaternion multiplication is not commutative, which means that pq doesn't equal to qp.

Given the definition of quaternion addition and multiplication, the convolution $h[n]$ of two quaternion impulse responses $p[n] = p_s[n] + p_x[n]\boldsymbol{i} + p_y[n]\boldsymbol{j} + p_z\boldsymbol{k}$ and $q[n] = q_s[n] + q_x[n]\boldsymbol{i} + q_y[n]\boldsymbol{j} + q_z\boldsymbol{k}$ is defined in a similar way as real convolution

$$h[n] = \sum_{l=-\infty}^{\infty} p[l]q[n-l] \quad (5)$$

Due to the non-commutativity of quaternion multiplication, the quaternion convolution is also non-commutative.

3. QUATERNION TWO-DIMENSION COMMON FACTOR DECOMPOSITION

In this section a quaternion two-dimension common factor decomposition (QCFD) algorithm is derived based on minimizing the mean square error (MSE) between the four components of the original quaternion impulse response and that of the reconstructed quaternion impulse response.

3.1. Quaternion Factor Decomposition

Given a Length-L quaternion impulse response $q[n] = q_s[n] + q_x[n]\boldsymbol{i} + q_y[n]\boldsymbol{j} + q_z\boldsymbol{k}$, two quaternion factor impulse responses length-L_1 $c[n] = c_s[n] + c_x[n]\boldsymbol{i} + c_y[n]\boldsymbol{j} + c_z\boldsymbol{k}$ and length-L_2 $d[n] = d_s[n] + d_x[n]\boldsymbol{i} + d_y[n]\boldsymbol{j} + d_z\boldsymbol{k}$ can be found to model it so that $c[n] \otimes d[n]$ approximates $q[n]$. The quaternion convolution defined in Eq. (5) can be expanded and implemented via 16 real convolutions as illustrated by Eq. (6)

Eq. (6) can be reformulated in matrix form as

$$\begin{aligned} \boldsymbol{q} =& C_s\boldsymbol{d}_s - C_x\boldsymbol{d}_x - C_y\boldsymbol{d}_y - C_z\boldsymbol{d}_z \\ & +\boldsymbol{i}(C_s\boldsymbol{d}_x + C_x\boldsymbol{d}_s + C_y\boldsymbol{d}_z - C_z\boldsymbol{d}_y) \\ & +\boldsymbol{j}(C_s\boldsymbol{d}_y - C_x\boldsymbol{d}_z + C_y\boldsymbol{d}_s + C_z\boldsymbol{d}_x) \\ & +\boldsymbol{k}(C_s\boldsymbol{d}_z + C_x\boldsymbol{d}_y - C_y\boldsymbol{d}_x + C_z\boldsymbol{d}_s) \end{aligned} \quad (7)$$

where $\boldsymbol{d}_s = [d_s[1] \; \cdots \; d_s[L_2]]'$ and $\boldsymbol{d}_x, \boldsymbol{d}_y, \boldsymbol{d}_z$ and $\boldsymbol{q} = \boldsymbol{q}_s + \boldsymbol{i}\boldsymbol{q}_x + \boldsymbol{j}\boldsymbol{q}_y + \boldsymbol{k}\boldsymbol{q}_k$ are constructed in a similar way and

$$C_s = \begin{bmatrix} c_s[1] & & & \mathbf{0} \\ \ddots & c_s[1] & & \\ c_s[L_1] & \ddots & \ddots & \\ & c_s[L_1] & \ddots & c_s[1] \\ & & \ddots & \ddots \\ \mathbf{0} & & & c_s[L_1] \end{bmatrix}_{L \times L_2}$$

and C_x, C_y and C_z are constructed in a similar way.

Eq. (7) can be expressed by component as

$$\begin{cases} \boldsymbol{q}_s = C_s\boldsymbol{d}_s - C_x\boldsymbol{d}_x - C_y\boldsymbol{d}_y - C_z\boldsymbol{d}_z \\ \boldsymbol{q}_x = C_s\boldsymbol{d}_x + C_x\boldsymbol{d}_s + C_y\boldsymbol{d}_z - C_z\boldsymbol{d}_y \\ \boldsymbol{q}_y = C_s\boldsymbol{d}_y - C_x\boldsymbol{d}_z + C_y\boldsymbol{d}_s + C_z\boldsymbol{d}_x \\ \boldsymbol{q}_z = C_s\boldsymbol{d}_z + C_x\boldsymbol{d}_y - C_y\boldsymbol{d}_x + C_z\boldsymbol{d}_s \end{cases} \quad (8)$$

Eq. (8) can be re-organized as

$$\tilde{\boldsymbol{q}} \overset{def}{=} \begin{bmatrix} \boldsymbol{q}_s \\ \boldsymbol{q}_x \\ \boldsymbol{q}_y \\ \boldsymbol{q}_z \end{bmatrix} = \begin{bmatrix} C_s & -C_x & -C_y & -C_z \\ C_x & C_s & -C_z & C_y \\ C_y & C_z & C_s & -C_x \\ C_z & -C_y & C_x & C_s \end{bmatrix} \begin{bmatrix} \boldsymbol{d}_s \\ \boldsymbol{d}_x \\ \boldsymbol{d}_y \\ \boldsymbol{d}_z \end{bmatrix} \overset{def}{=} C\tilde{\boldsymbol{d}} \quad (9)$$

Given $q[n]$ and one of the factor response $c[n]$, the other factor response $d[n]$ that minimizes the total MSE between all components of $q[n]$ and $c[n] \otimes d[n]$ is calculated by

$$\tilde{\boldsymbol{d}} = C^{\dagger}\tilde{\boldsymbol{q}} \quad (10)$$

where $(\bullet)^{\dagger}$ denotes pseudo-inverse.

Due to the non-commutativity of quaternion multiplication and convolution, when given $q[n]$ and $d[n]$, the best $c[n]$ that minimizes the total MSE between all components of $q[n]$ and $c[n] \otimes d[n]$ can not be calculated in the same way, which is different from the real number common factor decomposition [8]. Nevertheless, the formula to optimize $c[n]$ from given $q[n]$ and $d[n]$ can be derived in a similar way.

3.2. Quaternion Common Factor Decomposition

Given a set of quaternion impulse responses $q_k[n]$, $k = 1, \cdots, K$ and a quaternion factor impulse $d_k[n]$ for each of them, another quaternion factor impulse $c[n]$ that minimizes the total MSE between $q_k[n]$ and $d_k[n] \otimes c[n]$ for all k should satisfy

$$\tilde{\boldsymbol{q}}_k = D_k\tilde{\boldsymbol{c}} \quad k = 1, \cdots, K \quad (11)$$

where D_k is constructed in the same way as C, $\tilde{\boldsymbol{c}}$ and $\tilde{\boldsymbol{q}}_k$ are constructed in the same way as $\tilde{\boldsymbol{q}}$ and the best $c[n]$ is calculated by

$$\tilde{\boldsymbol{c}} = D^{\dagger}\tilde{\boldsymbol{q}} \quad (12)$$

$$\begin{aligned} c[n] \otimes d[n] = & c_s[n] \otimes d_s[n] - c_x[n] \otimes d_x[n] - c_y[n] \otimes d_y[n] - c_z[n] \otimes d_z[n] \\ & + \boldsymbol{i}(c_s[n] \otimes d_x[n] + c_x[n] \otimes d_s[n] + c_y[n] \otimes d_z[n] - c_z[n] \otimes d_y[n]) \\ & + \boldsymbol{j}(c_s[n] \otimes d_y[n] - c_x[n] \otimes d_z[n] + c_y[n] \otimes d_s[n] + c_z[n] \otimes d_x[n]) \\ & + \boldsymbol{k}(c_s[n] \otimes d_z[n] + c_x[n] \otimes d_y[n] - c_y[n] \otimes d_x[n] + c_z[n] \otimes d_s[n]) \end{aligned} \tag{6}$$

where $D = [D_1' \ \cdots \ D_K']'$ and $\tilde{\boldsymbol{q}} = [\tilde{\boldsymbol{q}}_1' \ \cdots \ \tilde{\boldsymbol{q}}_K']'$.

Once again, the best $c[n]$ that minimizes the total MSE between $q_k[n]$ and $c[n] \otimes d_k[n]$ for all k can not be calculated using the same formula due to the non-commutativity of quaternion multiplication and convolution. However, it can be calculated in a similar way.

3.3. Quaternion Two-Dimension Factor Decomposition

Given a 3-dimensional quaternion dataset $\mathbf{Q}(n, \theta, \phi) \in \mathbf{H}^{N \times I \times K}$, $k = 1, \cdots, K$, where $\mathbf{H}$ denotes the set of quaternion numbers, a quaternion 2D-CFD (QCFD) algorithm is proposed to represent the dataset with two sets of common factors. The QCFD algorithm is implemented in an iterative way similar to the real number CFD [8] as shown in Table 1.

Table 1. Optimization Procedure of QCFD

For HRIR set of each azimuth $\mathbf{Q}(n, \phi, \theta_k)$ $k = 1, \cdots, K$
Apply Quaternion CFD to this set of HRIRs and get the CF
End for
All K CFs make up a matrix $A \in \mathbf{R}^{N_1 \times K}$
Repeat
For each elevation ϕ_i $i = 1, \cdots, I$
Get the set of impulse responses in this elevation $\mathbf{H}(n, \phi_i, \theta)$
Use CFs in A as the IFs of this HRIR set
Extract the CF for this elevation
End for
All I CFs make up a matrix $E \in \mathbf{R}^{N_2 \times I}$
For HRIR set of each azimuth $\mathbf{Q}(n, \phi, \theta_k)$ $k = 1, \cdots, K$
Use CFs in E as the IFs of this HRIR set
Extract the CF for this azimuth
End for
All K CFs make up a matrix $A \in \mathbf{R}^{N_1 \times K}$
Until Convergence

After the QCFD modeling, the original quaternion impulse response can be reconstructed by

$$\mathbf{Q}(n, \phi, \theta) = E(n, \theta) \bigotimes A(n, \phi) \tag{13}$$

and the reconstructed HRIR can be obtained from each component of the reconstructed quaternion impulse response.

It is worth mentioning that the QCFD modeling can be implemented in another way so that the original quaternion impulse response can be reconstructed by

$$\mathbf{Q}(n, \phi, \theta) = A(n, \phi) \bigotimes E(n, \theta) \tag{14}$$

. However, there is no essential difference between these two representations and the modeling errors for these two representations are the same.

4. EXPERIMENT AND RESULT

4.1. Experiment Setup

In this section, the proposed QCFD algorithm is applied to model the Subject 3 HRIR in CIPIC HRIR database [9]. Four HRIRs are used to construct a quaternion impulse response. The modeling distortion is evaluated by Spectral Distortion (SD) score and Waveform Fit score defined as

$$SD = \sqrt{\frac{1}{L}\sum_{k=1}^{L}(20\log_{10}\frac{|H_k|}{|\hat{H}_k|})^2} \quad Fit = (1 - \frac{\sum_{n=1}^{L} e[n]^2}{\sum_{n=1}^{L} h[n]^2}) \times 100\%$$

where $h[n]$ and $\hat{h}[n]$ are the measured and the modeled HRIRs, H_k and $\hat{H}_k$ are their spectra, respectively, $e[n] = h[n] - \hat{h}[n]$.

4.2. Quaternion Impulse Response Construction

It is desirable that the four HRIRs are similar to each other so that the QCFD will have a better performance. In this section, three schemes are proposed to construct a quaternion impulse response by using four HRIRs measured at different positions or for different ears. To facilitate specifying the position clearly, the inter-aural pole coordinate system as described in [9] is adopted, in which $(0°, 0°)$, $(0°, -90°)$ and $(90°, 0°)$ stand for the front, left and up, respectively, and two azimuthes θ and $-\theta$ are symmetric along the median plane while two elevations ϕ and $180° - \phi$ are symmetric along the frontal plane.

- **Adjacent** scheme: four HRIRs measured for a single ear at the same azimuth and at adjacent elevations are used.
- **FB-SP** scheme: two left ear HRIRs measured at (ϕ, θ) and $(180° - \phi, \theta)$ and two right ear HRIRs measured at (ϕ, θ) and $(180° - \phi, \theta)$ are used.
- **FB-DP** scheme: two left ear HRIRs measured at (ϕ, θ) and $(180° - \phi, \theta)$ and two right ear HRIRs measured at $(\phi, -\theta)$ and $(180° - \phi, -\theta)$ are used.

4.3. Experiment Result

In experiment I, QCFD is applied to a set of 340 HRIRs with $\phi \in [30° \ 145°]$ and $\theta \in [-45° \ 45°]$. The quaternion impulse response is constructed by using the Adjacent scheme. The modeling error in MSE achieved in each iteration step of the QCFD algorithm is displayed in Fig. 1. It is seen that the modeling error decreases steadily with iterations.

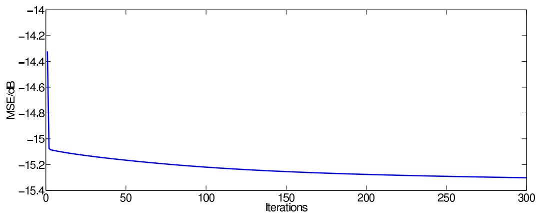

Fig. 1. Convergence curve for the QCFD algorithm

In experiment II, QCFD is applied to a set of 608 pairs of HRIRs with elevation $\phi \in [-45° \ 45°]$ in the front and $\phi \in [135° \ -225°]$ in the back and with azimuth $\theta \in [-45° \ 45°]$. A set of 152 quaternion impulse responses are constructed by using each of the three schemes proposed and QCFD algorithm with varying elevation lengths is applied to model the quaternion impulse responses. The distortion of the proposed modeling is displayed in Fig. 2 and in Fig. 3. For comparision, the real number CFD modeling which is better than PCA-based methods and CAPZ-based methods [8] is also applied to the same HRIR set and its distortion is also displayed in these two figures. It is observed from Fig. 2 and Fig. 3 that QCFD algorithm outperforms real number CFD algorithm no matter which scheme is adopted in quaternion impulse response construction. This shows that the QCFD can exploit the inter-HRIR similarity to reduce the storage for HRIR. Meanwhile, the spectrum difference between front-back HRIR is mainly retained, as show in Fig. 4.

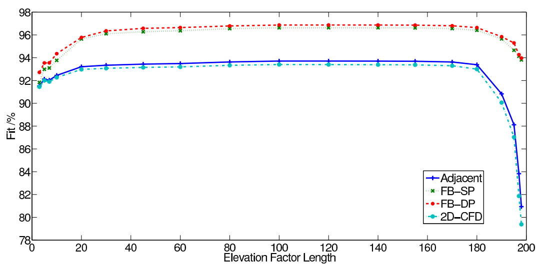

Fig. 2. Model Accuracy in Fit of Quaternion Common Factor Decomposition for each Quaternion Construction Scheme and varying Elevation Factor Length

In Experiment III, QCFD is applied to the set of quaternion impulse responses constructed by using FB-DP scheme from a set of 500 pairs of HRIRs with $\phi \in [34° \ -146°]$ and $\theta \in [-80° \ 80°]$. The distortion of the modeling for different elevation factor lengths are listed in Table 2. The storage requirement in term of number of parameters needed to represent the whole evaluated HRIR dataset is also calculated and shown in Table 2. For comparison, the distortion and storage

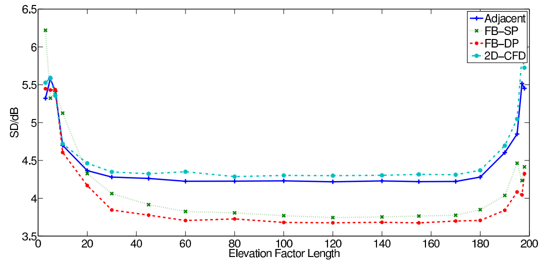

Fig. 3. Model Accuracy in SD of Quaternion Common Factor Decomposition for each Quaternion Construction Scheme and varying Elevation Factor Length

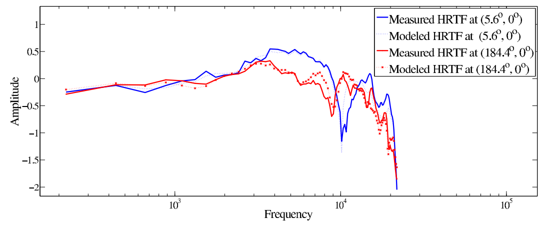

Fig. 4. Spectra of measured and reconstructed HRTF pairs

requirement of the real number CFD modeling is also listed in Table 2. It is seen that QCFD modeling can achieve higher model accuracy with comparable storage requirement.

Table 2. Distortion and storage requirement of QCFD algorithm and real number CFD algorithm with different elevation factor length L_e

L_e	Fit/%		SD/dB		Storage	
	QCFD	CFD	QCFD	CFD	QCFD	CFD
5	94.8	94.1	5.16	5.41	19800	10000
10	96.1	95.3	4.56	4.71	19500	9950
70	97.4	96.4	3.87	4.17	15900	9350
130	97.5	96.5	3.80	4.10	12300	8750
190	97.6	96.2	3.79	4.13	8700	8150
197	96.6	94.9	3.94	4.25	8280	8080

5. CONCLUSION

A HRIR model based on quaternion algebra is proposed to reduce the storage of HRIR dataset. Four HRIRs which are similar to each other are used to constructed a quaternion impulse response. A quaternion common factor decomposition (QCFD) algorithm is derived to represent the quaternion impulse response set, which are 3-dimensional dataset, with two sets of factor responses to reduce the storage. Compared with real number CFD algorithm, QCFD algorithm can exploit the similarity between HRIRs of the same quaternion impulse response and achieves higher modeling accuracy with comparable storage.

6. REFERENCES

[1] William R. Hamilton, “On Quaternions, or on a New System of Imaginaries in Algebra,” *Philosophical Magazine*, vol. 25, no. 3, pp. 489–495, 1844.

[2] Ken Shoemake, “Animating rotation with quaternion curves,” *SIGGRAPH Comput. Graph.*, vol. 19, no. 3, pp. 245–254, July 1985.

[3] K. Spring, “Euler parameters and the use of quaternion algebra in the manipulation of finite rotations: A review,” *Mechanism and Machine Theory*, vol. 21, no. 5, pp. 365–373, 1986.

[4] Thomas Bulow and Gerald Sommer, “Multi-dimensional signal processing using an algebraically extended signal representation,” in *In G. Sommer, J.J. Koenderink (Eds.), Proc. Int. Workshop on Algebraic Frames for the Perception-Action Cycle, Kiel, LNCS*. 1997, pp. 148–163, Springer.

[5] Nicolas Le Bihan and Jérôme Mars, “Singular value decomposition of quaternion matrices: A new tool for vector-sensor signal processing,” *Signal Process.*, vol. 84, no. 7, pp. 1177–1199, July 2004.

[6] J. Blauert, *Spatial Hearing: The Phychophysics of Human Sound Localization*, 55 Hayward Street Cambridge, MA02142-1315 USA, MIT Press, 1983.

[7] Wenzel E.M., M.; Kistler D.J. Arruda, and Wightman F.L., “Localization using nonindividualized head-related transfer functions,” *The journal of Acoustical Society of America*, vol. 94, no. 1, pp. 111–123, July 1993.

[8] Zhixin Wang and Cheung-Fat Chan, “Two-dimension common factor decomposition of head-related impulse response,” *Signal Processing Letters, IEEE*, vol. 20, no. 8, pp. 795–798, 2013.

[9] V.R. Algazi, R.O. Duda, D.M. Thompson, and C. Avendano, “The cipic hrtf database,” in *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*, 2001, pp. 99 –102.