A QUATERNION LEAST MEAN PHASE ADAPTIVE ESTIMATOR

Sayed Pouria Talebi, Dongpo Xu*, Anthony Kuh†, and D. Mandic**

*Imperial College London, Department of Electrical and Electronic Engineering
† University of Hawaii, Department of Electrical and Electronic Engineering
Emails: {s.talebi12,dongpo.xu, d.mandic}@imperial.ac.uk, kuh@spectra.eng.hawaii.edu,

ABSTRACT

A Quaternion Least Mean Phase (QLMP) algorithm is introduced for phase-only adaptive filtering of quaternion-valued signals. This is achieved by first defining the notion of phase in the quaternion domain based on the exponential representation of quaternion random variables, and then using the $\mathbb{HR}$-calculus to derive a steepest-decent weight update. The advantages of this adaptive algorithm are illustrated in a bearings only tracking scenario, where the QLMP is shown to outperform the amplitude-phase based quaternion Least Mean Square (QLMS).

***Index Terms*—** Quaternions, adaptive signal processing, phase estimation, bearings only tracking.

1. INTRODUCTION

Recent advances in technology have brought to light new three- and four-dimensional sensors, such as inertial body sensors and three-dimensional anemometers. To fully exploit the multidimensional statistics of such data, signals from such sensors should be processed in the multidimensional domain where they reside [1, 2]. Quaternions provide a natural representation for three- and four-dimensional data, and they have become a platform for developing signal processing algorithms for such data sources [3, 4]. Applications include computer graphics, satellite navigation, and aeronautics (see [1–5]).

When dealing with three- and four-dimensional signals, phase becomes particularly interesting, especially in applications where the amplitude information is either not important or is corrupted. Quaternions provide a straightforward representation of rotation and orientation in three dimensions, as in comparison toother methods, such as Euler angles and matrix transformations, quaternion representation is more compact, accurate and helps to avoid gimbal lock [6, 7]. Since for describing rotation and orientation only unit quaternions are needed, the amplitude does not carry any information, and therefore the performance of any estimation method in such a scenario can be assessed based on the error in the phase only.

The Mean Square Error (MSE) is the most frequently used error measure for signal processing algorithms. It forms the basis of mainstream adaptive filtering algorithms, such as the Least Mean Square (LMS). Originally, the LMS was introduced for real-valued signals and was extended to the complex domain in [8]. A quaternion LMS (QLMS) has been introduced only recently in [9] for unified filtering of three- and four-dimensional data, and has been shown to converge faster and to have enhanced stability compared to the multichannel LMS.

In the complex domain, the Least Mean Phase (LMP) adaptive filtering algorithm introduced in [10] was developed based on the phase error of complex variables. Within the cost function designed from the phase error, both the gradient of the squared error and the gradient of the phase error were used to update the filter weights. The algorithm was tested for channel estimation in communications. The Least Mean Magnitude Phase (LMMP) algorithm introduced in [11] decomposes the squared error cost function into two parts: error of the amplitude and a function related to the phase error. The gradient of each part is calculated individually, and separate step sizes are used for the respective weight updates. The LMMP was validated for channel equalization in the presence of Doppler shift due to physical motion and in array processing, outperforming the complex LMS in both scenarios.

Unlike the complex domain, there is no straightforward definition of phase in the quaternion domain, yet phase-based filtering is very natural for quaternion data. To address this, we first give a brief background on quaternion algebra and the $\mathbb{HR}$-calculus. Then a notion of phase in the quaternion domain is introduced, in order to define a cost function based on the phase of quaternion random variables. Using the gradient of this cost function we develop the Quaternion Least Mean Phase (QLMP) adaptive filter. The performance of this adaptive filter is verified using both synthetic data and in a bearings only tracking scenario, where the proposed QLMP outperforms the QLMS.

2. BACKGROUND

A quaternion variable $q \in \mathbb{H}$ consist of a real part $\Re(q) = q_0$, also referred to as the scalar part or $\mathbb{S}_q$, and a vector part $\mathbb{V}_q$, called a pure quaternion $\Im(q) = iq_1 + jq_2 + kq_3$ [12]. The vector part comprises of three orthogonal imaginary units i, j

and k. An arbitrary quaternion q can be expressed as

$$q = \Re(q) + \Im(q) = q_0 + iq_1 + jq_2 + kq_3 \quad (1)$$

where $q_0, q_1, q_2, q_3 \in \mathbb{R}$. The three form an orthogonal basis and follow the multiplication rules

$$ij = k, jk = i, ki = j, ijk = i^2 = k^2 = j^2 = -1. \quad (2)$$

Observe that, as a result, quaternion multiplication is not commutative ($ab \neq ba$).

The quaternion conjugate, q^* is given by $q^* = q_0 - iq_1 - jq_2 - kq_3$, while the norm of q is given by

$$\parallel q \parallel = \sqrt{qq^*} = \sqrt{q_0^2 + q_1^2 + q_3^2 + q_4^2}$$

and corresponds to the Euclidean distance from the origin.

For $q, \mu \in \mathbb{H}$, the involution of q around μ is defined as

$$q^\mu = \mu q \mu^{-1} \quad (3)$$

The geometric interpretation of involution is a rotation of the imaginary part of q around μ by 2θ, where $\mu = |\mu| \left(\cos(\theta) + \frac{\Im(\mu)}{|\Im(\mu)|}\sin(\theta)\right)$ [5]. Of special interest are involutions for $\mu = i, j, k$ which are given by

$$\begin{aligned} q^i &= -iqi = q_0 + iq_1 - jq_2 - kq_3 \\ q^j &= -jqj = q_0 - iq_1 + jq_2 - kq_3 \\ q^k &= -kqk = q_0 - iq_1 - jq_2 + kq_3 \end{aligned} \quad (4)$$

where q_0, q_1, q_2, q_3 are the real and imaginary parts of q defined in (1). These rotations can be used to extract each component of a quaternion as follows

$$\begin{aligned} q_0 &= \frac{1}{4}(q + q^i + q^j + q^k),\ q_1 = \frac{1}{4}(q + q^i - q^j - q^k) \\ q_2 &= \frac{1}{4}(q - q^i + q^j - q^k),\ q_3 = \frac{1}{4}(q - q^i - q^j + q^k) \end{aligned} \quad (5)$$

To perform optimization or adaptive filtering in the quaternion domain, the gradient or the stochastic gradient of a cost function must be obtained; however, the Cauchy-Riemann Feuter (CRF) conditions for quaternion analyticity [13]

$$\begin{aligned} \frac{\partial_{right} f}{\partial q^*} &= \frac{1}{4}\left(\frac{\partial f}{\partial q_0} + \frac{\partial f}{\partial q_1}i + \frac{\partial f}{\partial q_2}j + \frac{\partial f}{\partial q_3}k\right) = 0 \\ \frac{\partial_{left} f}{\partial q^*} &= \frac{1}{4}\left(\frac{\partial f}{\partial q_0} + i\frac{\partial f}{\partial q_1} + j\frac{\partial f}{\partial q_2} + k\frac{\partial f}{\partial q_3}\right) = 0 \end{aligned} \quad (6)$$

impose severe limitations on the range of analytic functions possible. For example, cost functions are real functions of quaternion variables and are not analytic and thus not differentiable in $\mathbb{H}$. Using the equations in (5) and establishing a duality between $\mathbb{R}^4$ and $\mathbb{H}$, a framework for calculating derivatives of quaternion functions, whether analytic or not, has been presented in [4, 14], and was termed the $\mathbb{HR}$-calculus. Within the $\mathbb{HR}$-calculus it was established that the direction of steepest decent is given by $-\frac{\partial f}{\partial q^*}$.

One of the most important functions in signal processing is the exponential function. Generally, the exponential of any number can be represented by the sum

$$e^q = \sum_{n=0}^{\infty} \frac{q^n}{n!}. \quad (7)$$

This also applies to quaternions and this expansion is convergent in the norm for all quaternions [15].

For any quaternion with a non-vanishing imaginary part, the following holds: $\Im^2(q) = -|\Im(q)|^2$. Thus, the ratio $\Im(q)/|\Im(q)|$ is a square root of -1, and for any quaternion q we can write

$$q = |q| (\cos(\theta) + \zeta \sin(\theta)) = |q| e^{\zeta\theta} \quad (8)$$

where $\zeta = \frac{\Im(q)}{|\Im(q)|}$ and $\theta = \arctan(\frac{|\Im(q)|}{\Re(q)})$. This formula allows for a simple description of rotation in $\mathbb{H}$.

Consider the polar representation of q given in (8), where ζ is a normalized projection of q on the vector part of $\mathbb{H}$. Given that $\zeta^2 = -1$, ζ and the real axis define a two-dimensional plane in $\mathbb{H}$, which is isomorphic to the complex domain and contains q. In this subspace, θ denotes the angle between the real axis and q, so that we can consider $\zeta\theta$ to be the phase of q, as it uniquely describes the orientation of q.

3. QUATERNION LEAST MEAN PHASE ESTIMATION (QLMP)

Consider the linear estimator $y = \mathbf{h}^H \mathbf{q}$, where $\mathbf{h}$ is the weight vector, $\mathbf{q}$ is the observation vector, and y is the estimation of the desired signal d in terms of $\mathbf{q}$. The phase error of the estimation at time instant k is given by

$$e_k = \zeta_{d_k}\theta_{d_k} - \zeta_{y_k}\theta_{y_k} \quad (9)$$

Based on the phase error, we can now introduce the following cost function

$$J_\phi(h_k) = e_k e_k^* = |\zeta_{d_k}\theta_{d_k} - \zeta_{y_k}\theta_{y_k}|^2 \quad (10)$$

Our objective is to develop an adaptive algorithm that minimizes J_ϕ by implementing a steepest-decent weight vector update, to give the update in the form

$$h_{k+1} = h_k - \mu \nabla_{h^*} J_\phi(h_k) \quad (11)$$

where μ is the step size and $\nabla_{h^*} J_\phi(h_k)$ is the quaternion gradient of the phase only cost function $J_\phi(h_k)$.

3.1. Stochastic Gradient Derivation

Observe from (9) that the phase error e_k is a pure quaternion, and thus $e_k^* = -e_k$, so that the cost function can be written as $J_\phi(h_k) = -e_k^2$. For simplicity of presentation, we shall ignore time indices in this section.

The cost function is a real-valued function of quaternion variables, thus by using the $\mathbb{HR}$-calculus the gradient of $J_\phi = -e^2$ is $\nabla_{h^*} J_\phi = -(\nabla_{h^*} e)e - e(\nabla_{h^*} e)$ where the term

$$\nabla_{h^*} e = (\nabla_{h^*} \zeta_y)\theta_y + \zeta_y(\nabla_{h^*} \theta_y) \tag{12}$$

Note that the gradient $\nabla_{h^*} e$ is a quaternion, and because quaternions are not commutative $(\nabla_{h^*} e)e \neq e(\nabla_{h^*} e)$, and the only gradients that need to be calculated are $\nabla_{h^*} \zeta_y$ and $\nabla_{h^*} \theta_y$.

We start by calculating $\nabla_{h^*} \Im(y)$ and $\nabla_{h^*} \Re(y)$ using the quaternion conjugate, to give

$$\nabla_{h^*} \Im(y) = \frac{1}{2}\nabla_{h^*}(y - y^*) = -\frac{1}{4}q - \frac{1}{2}q^* \tag{13}$$

$$\nabla_{h^*} \Re(y) = \frac{1}{2}\nabla_{h^*}(y + y^*) = -\frac{1}{4}q + \frac{1}{2}q^* \tag{14}$$

Since $\nabla_{h^*} \Im(y) = -\nabla_{h^*} \Im^*(y)$, by using $|\Im(y)| = \sqrt{|\Im(y)|^2}$, the term $\nabla_{h^*} |\Im(y)|$ now becomes

$$\nabla_{h^*} |\Im(y)| = \nabla_{h^*} \sqrt{|\Im(y)|^2} = \frac{\nabla_{h^*}\big(\Im(y)\Im^*(y)\big)}{2|\Im(y)|} \tag{15}$$

where

$$\nabla_{h^*}\big(\Im(y)\Im^*(y)\big) = \big(\nabla_{h^*}\Im(y)\big)\Im^*(y) + \Im(y)\big(\nabla_{h^*}(\Im^*(y)\big)$$

and the terms $\nabla_{h^*} \Im(y)$ and $\nabla_{h^*} \Im^*(y)$ are given in (13). In order to find $\nabla_{h^*} \theta_y$ keeping in mind that $|\Im(y)|/\Re(y) \in \mathbb{R}$, and using the $\mathbb{HR}$-calculus, we have

$$\begin{aligned}\nabla_{h^*} \theta_y &= \frac{1}{1 + \big(\frac{|\Im(y)|}{\Re(y)}\big)^2} \nabla_h^* \frac{|\Im(y)|}{\Re(y)} \\ &= \frac{1}{1 + \big(\frac{|\Im(y)|}{\Re(y)}\big)^2} \left(\frac{\nabla_{h^*}|\Im(y)|}{\Re(y)} + |\Im(y)|\nabla_{h^*}\frac{1}{\Re(y)}\right)\end{aligned} \tag{16}$$

where $\nabla_{h^*}\frac{1}{\Re(y)} = \frac{-1}{\Re(y)^2}\nabla_{h^*}\Re(y)$, and $\nabla_{h^*}|\Im(y)|$ is given in (15). Since $\zeta_y = \Im(y)/|\Im(y)|$, this yields

$$\nabla_{h^*} \zeta_y = \frac{1}{|\Im(y)|}\nabla_{h^*}\Im(y) + \Im(y)\nabla_{h^*}\big(\frac{1}{|\Im(y)|}\big) \tag{17}$$

where $\nabla_{h^*}\frac{1}{|\Im(y)|} = \frac{-1}{|\Im(y)|^2}\nabla_{h^*}|\Im(y)|$, and $\nabla_{h^*}|\Im(y)|$ is given in (15). Finally, inserting (16) and (17) into (12) gives the gradient of the cost function of the proposed Least Mean Phase (QLMP) adaptive filtering algorithm in the form

$$\begin{aligned}\nabla_{h^*} J_\phi = \\ &-\left[\left(\frac{\nabla_{h^*}\Im(y)}{|\Im(y)|} - \Im(y)\frac{\nabla_{h^*}|\Im(y)|}{\Im|(y)|^2}\right)\theta_y + \zeta_y\nabla_{h^*}\theta_y\right]e \\ &-e\left[\left(\frac{\nabla_{h^*}\Im(y)}{|\Im(y)|} - \Im(y)\frac{\nabla_{h^*}|\Im(y)|}{\Im|(y)|^2}\right)\theta_y + \zeta_y\nabla_{h^*}\theta_y\right]\end{aligned}$$

while the weight update equation is given in (11).

3.2. Stability Analysis

Consider the *priori* and the *posteriori* estimation errors respectively given by

$$\hat{e}_k = \rho(d_k) - \rho(h_k q_k) \tag{18}$$

$$\bar{e}_k = \rho(d_k) - \rho(h_{k+1} q_k) \tag{19}$$

where $\rho(\cdot)$ denotes the phase. To calculate the range of μ for which $|\bar{e}_k|^2 < |\hat{e}_k|^2$ thus ensuring convergence in the mean square sense, we consider the first order Taylor series expansion (TSE) of $|\bar{e}_k|^2$ around $|\hat{e}_k|^2$ is given by

$$|\bar{e}_k|^2 = |\hat{e}_k|^2 + \langle \nabla_{h^*}|\hat{e}_k|^2, \Delta h_k\rangle \tag{20}$$

where the symbol $\langle\cdot,\cdot\rangle$ denotes the inner product and $\Delta h_k = h_{k+1} - h_k$. From (11) we have $\Delta h_k = -\mu\nabla_{h^*}J_\phi(h_k) = -\mu\nabla_{h^*}|\hat{e}_k|^2$, and by replacing into (20) we have $|\bar{e}_k|^2 = |\hat{e}_k|^2 - \mu\langle\nabla_{h^*}|\hat{e}_k|^2, \nabla_{h^*}|\hat{e}_k|^2\rangle$. Furthermore, considering that $\langle\nabla_{h^*}|\hat{e}_k|^2, \nabla_{h^*}|\hat{e}_k|^2\rangle = (\nabla_{h^*}|\hat{e}_k|^2)(\nabla_{h^*}|\hat{e}_k|^2)^*$ and $\nabla_{h^*}J_\phi = -(\nabla_{h^*}\hat{e}_k)\hat{e}_k - \hat{e}_k(\nabla_{h^*}\hat{e}_k)$ the term $|\bar{e}_k|^2$ in (20) can be written as

$$\begin{aligned}|\bar{e}_k|^2 =& |\hat{e}_k|^2 - \mu(\nabla_{h^*}|\hat{e}_k|^2)(\nabla_{h^*}|\hat{e}_k|^2)^* \\ =& |\hat{e}_k|^2 - \mu|(\nabla_{h^*}\hat{e}_k)\hat{e}_k + \hat{e}_k(\nabla_{h^*}\hat{e}_k)|^2\end{aligned} \tag{21}$$

By expanding $|(\nabla_{h^*}\hat{e}_k)\hat{e}_k + \hat{e}_k(\nabla_{h^*}\hat{e}_k)|^2$ and using the triangle inequality $|(\nabla_{h^*}\hat{e}_k)\hat{e}_k|^2 \leqslant |(\nabla_{h^*}\hat{e}_k)|^2|\hat{e}_k|^2$ we have

$$|\bar{e}_k|^2 \leqslant |\hat{e}_k|^2(1 - 4\mu|\nabla_{h^*}\hat{e}_k|^2)$$

Thus, to ensure $|\bar{e}_k|^2 < |\hat{e}_k|^2$ it is sufficient to guarantee

$$\left|1 - 4\mu|\nabla_{h^*}\hat{e}_k|^2\right| < 1$$

giving the stability bound

$$0 < \mu < \frac{1}{2|\nabla_{h^*}\hat{e}_k|^2}$$

where $\nabla_{h^*}\hat{e}_k$ was calculated in the previous section and is given by (12)

4. SIMULATIONS

The performance of the proposed algorithm was first validated for a synthetic signal with unit amplitude. The signal was set to oscillate in the i axis in a sinusoidal fashion in the range of $f = 45 - 55$ Hz , as shown in Fig. 1. The signal was corrupted by Additive White Gaussian Noise (AWGN) with Signal to Noise Ratio (SNR) of 35 dB. The QLMP and QLMS were implemented in a one step prediction setting, where the phase of the weight represents the frequency. The learning rates for both algorithms were set to $\mu = 10^{-3}$. In this experiment the gradients were normalized to provide a fair comparison between the direction of decent in QLMS and QLMP. Fig. 1 shows the results of the frequency estimation, with the squared error averaged over 20 realizations shown in Fig. 2. Observe that the QLMP both outperformed the QLMS and also converged faster.

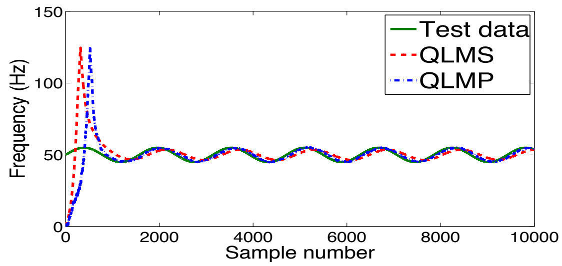

Fig. 1: Frequency estimation of FM-modulated test signal using QLMS and QLMP.

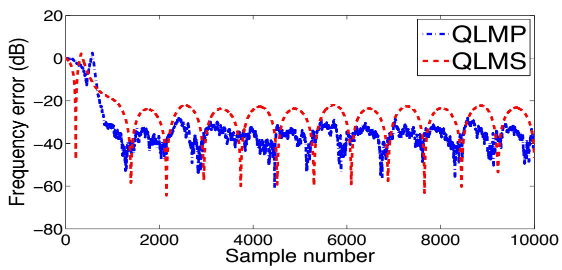

Fig. 2: Squared error of the frequency estimation.

4.1. Bearings Only Tracking (BOT)

Bearings Only Tracking (BOT) tracks the bearing of an object in a three-dimensional space, and is used when reliable information about distances is not available or has been corrupted by noise. Expressing rotations in the Cartesian and polar domains requires nine real-valued coefficients, while only four real-valued parameters (one quaternion) are needed in the quaternion domain.

In a three-dimensional space any point can be uniquely described by a set of Cartesian coordinates (x, y, z). The Cartesian coordinates can be transformed to spherical coordinates (r, δ, λ) where, r, δ and λ are the radius, polar, and azimuthal angle respectively. The coordinates can be transformed into a quaternion as (see [3] and references therein)

$$q = r\big(\sin(\delta)\cos(\lambda) + i\cos(\delta)\sin(\lambda) + j\sin(\delta)\sin(\lambda) + k\cos(\delta)\cos(\lambda)\big) \tag{22}$$

As we are interested in bearings estimation, we normalized the data allowing us to express the trajectory of an object traveling in three-dimensions as a quaternion.

We assumed the bearing measurements to be taken by a sensor located at the origin, which measures the bearings of an object in the three-dimensional space every second. The target was moving at an altitude of 5 km traveling from the bearing of $(-50\text{ km}, 3\text{ km})$ to $(50\text{ km}, 3\text{ km})$ at a constant speed of 600 km/h. The measurements were corrupted by AWGN to give the SNR of 10 dB. Both filters were initiated at zero altitude and bearing of $(1, 0)$. The bearings of the target and QLMP and QLMS estimates are plotted on a unit sphere in Fig. 3. While both QLMP and QLMS performed well, the QLMP exhibited lower bearings error, as shown in Fig. 4.

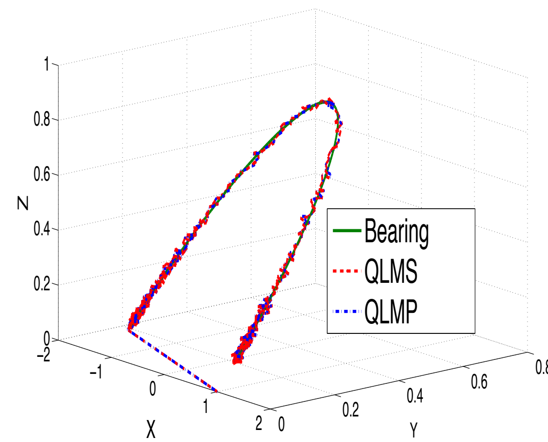

Fig. 3: Bearings only tracking .

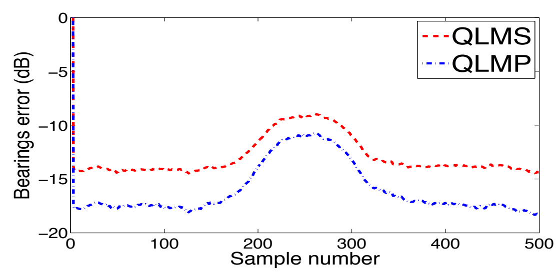

Fig. 4: Squared bearings error (dB).

The differences in bearings errors between QLMP and QLMS were largest for samples in the regions $(1 - 200)$ and $(400 - 500)$ which corresponds to the target approaching and leaving the sensor location. The samples in the region of $(200 - 400)$ correspond to the target passing directly in front of the sensor. In this region the phase changed rapidly causing an increase in bearings error for both the QLMP and QLMS; however, the QLMP still performed better than the QLMS.

5. CONCLUSIONS

We have explored the phase in the quaternion domain in order to define a cost function based on the phase error that is suitable for developing adaptive filtering algorithms. The Quaternion Least Mean Phase (QLMP) algorithm has been derived based on the gradient of such a phase-only cost function, and analysis was carried out to find a range for the step size where the adaptive filter remains stable. The performance of the QLMP has been verified using both synthetic data and in a bearings only tracking scenario and was benchmarked against the magnitude-phase based QLMS.

6. REFERENCES

[1] J. Via, D.P. Palomar, L. Vielva, I. Santamaria, "Quaternion ICA from second-order statistics, "*IEEE Transactions on Signal Processing*, vol.59, no.4, pp.1586-1600, Apr. 2011.

[2] N. Le Bihan and S.J. Sangwine, "Quaternion principal component analysis of color images," *In Proceedings of International Conference on Image Processing (ICIP)*, pp. 14-17 Sept. 2003.

[3] J B. Kuipers, "Quaternions and rotation sequences: A primer with applications to orbits, aerospace and virtual reality," *Princeton University Press*, 2002.

[4] D.P. Mandic, C. Jahanchahi, and C.C. Took, "A quaternion gradient operator and its applications," *IEEE Signal Processing Letters*, vol.18, no.1, pp. 47-50, Jan. 2011.

[5] T.A. Ell and S.J. Sangwine, "Quaternion involutions and anti-involutions," *Comput. & Math Appl.*, vol.53, no.1, pp. 137-143 January 2007.

[6] P. G. Jung, G. Lim, and K. Kong,"Human posture measurement in a three-dimensional space based on inertial sensors," *In Proceeeding of International Conference on Control, Automation and Systems (ICCAS)*, pp. 1013-1016, October 2012.

[7] T. Kadowaki, K. Kobayashi, K, and K. Watanabe, "Rotation angle measurement of high-speed flying object," *In Proceedings of In Proceedings of SICE-ICASE International Joint Conference*, pp. 5256-5259, Oct. 2006.

[8] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 719-720, April 1975.

[9] C.C. Took and D.P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1316-1327, April 2009.

[10] A. Tarighat and A.H. Sayed, "Least mean-phase adaptive filters with application to communications systems," *IEEE Signal Processing Letters*, vol.11, no.2, pp. 220-223, February 2004.

[11] S.C. Douglas and D.P. Mandic, "The least-mean-magnitude-phase algorithm with applications to communications systems," *In Proceedimgs of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4152-4155, May 2011.

[12] C.C. MacDuffee, "Algebras debt to Hamilton," *Scripta Mathematica*, pp. 25-36, 1944.

[13] K. Abdel-Khaalek, "Quaternion analyticity," *Dipartimento de Fisica-Universita de Lecce, Technical Report*,1996.

[14] C. Jahanchahi, C.C. Took, and D.P. Mandic, "On HR calculus, quaternion valued stochastic gradient, and adaptive three dimensional wind forecasting," *In Proceedings of The International Joint Conference on Neural Networks (IJCNN)*, pp.1,5, 18-23 July 2010.

[15] S. Said, N. L. Bihan, and S. J. Sangwine, "Fast complexified quaternion Fourier transform," *IEEE Transactions on Signal Processing*, vol.56, no.4, pp. 1522-1531, April 2008.