applied sciences

MDPI

Article

Regularization Factor Selection Method for l_1-Regularized RLS and Its Modification against Uncertainty in the Regularization Factor

Junseok Lim [1,*] and Seokjin Lee [2]

[1] Department of Electrical Engineering, College of Electronics and Information Engineering, Sejong University, Gwangjin-gu, Seoul 05006, Korea

[2] School of Electronics Engineering, College of IT Engineering, Kyungpook National University, Daegu 41566, Korea; sjlee6@knu.ac.kr

* Correspondence: jslim@sejong.ac.kr; Tel.: +82-2-3408-3299

Received: 5 December 2018; Accepted: 4 January 2019; Published: 1 January 2019

Featured Application: This algorithm can be applied to various kinds of sparse channel estimations, e.g., room impulse response, early reflection, and underwater channel response.

Abstract: This paper presents a new l_1-RLS method to estimate a sparse impulse response estimation. A new regularization factor calculation method is proposed for l_1-RLS that requires no information of the true channel response in advance. In addition, we also derive a new model to compensate for uncertainty in the regularization factor. The results of the estimation for many different kinds of sparse impulse responses show that the proposed method without a priori channel information is comparable to the conventional method with a priori channel information.

Keywords: l_1-regularized RLS; sparsity; room impulse response; total least squares; regularization factor

1. Introduction

Room impulse response (RIR) estimation is a problem in many applications that use acoustic signal processing. The RIR identification [1] is fundamental for various applications such as room geometry related spatial audio applications [2–5], acoustic echo cancellation (AEC) [6], speech enhancement [7], and dereverberation [8]. In [9], the RIR has relatively large magnitude values during the early part of the reverberation and fades to smaller values during the later part. This indicates that most RIR entries have values close to zero. Therefore, the RIR has a sparse structure. The sparse RIR model is useful for estimating RIRs in real acoustic environments when the source is given a priori [10]. There has been recent interest in adaptive algorithms for sparsity in various signals and systems [11–22]. Many adaptive algorithms based on least mean square (LMS) [11,12] and recursive least squares (RLS) [14–17] have been reported with different penalty functions. Sparse estimation research, such as that done by Eksioglu and Tanc [17], has proposed a sparse RLS algorithm, l_1-RLS, which is fully recursive like the plain RLS algorithm. The algorithm of l_1-RLS in [17] proposed a proper calculation method for the regularization factor. These recursive algorithms have the potential for sparse RIR estimation; however, the regularization factor should be established prior to applying these algorithms. The regularization factor calculation method requires information about a true sparse channel response for a good performance. The authors in [18,19] have also proposed recursive regularization factor selection methods; however, these methods still need the true impulse response in advance.

In this paper, we propose a new regularization factor calculation method for l_1-RLS algorithm in [17]. The new regularization factor calculation needs no information for the true channel response in advance. This makes it possible to apply l_1-RLS algorithm in various room environments. In addition, we derive a new model equation for l_1-RLS in [17] with uncertainty in the regularization factor and show that the new model is similar to the total least squares (TLS) model that compensates for uncertainty in the calculated regularization factor without the true channel response. For the performance evaluation, we simulate four different sparse channels and compare channel estimation performances. We show that, without any information of the true channel impulse response, the performance of the proposed algorithm is comparable to that of l_1-RLS with the information of the true channel impulse response.

This paper is organized as follows. In Section 2, we summarize l_1-RLS in [17]. In Section 3, we summarize the measure of sparsity. In Section 4, we propose a new method for the regularization calculation. In Section 5, we show that l_1-RLS with uncertainty in the regularization factor can be modeled as the TLS model. In Section 6, we summarize l_1-RTLS (recursive total least squares) algorithm as a solution for l_1-RLS with uncertainty in the regularization factor. In Section 7, we present simulation results to show the performance of the proposed algorithm. Finally, we give the conclusion in Section 8.

2. Summarize l_1-RLS

In the sparse channel estimation problem of interest, the system observes a signal represented by an $M \times 1$ vector $\mathbf{x}(k) = [x_k, \cdots, x_{k-M+1}]^T$ at time instant n, performs filtering, and obtains the output $y(i) = \mathbf{x}^T(k)\mathbf{w}_o(k)$, where $\mathbf{w}_o(k) = [w_k, \cdots, w_{k-M+1}]^T$ is the M dimensional actual system with finite impulse response (FIR) type. For system estimation, an adaptive filter system applies with M dimensional vector $\mathbf{w}(k)$ to the same signal vector $\mathbf{x}(k)$ and produces an estimated output $\hat{y}(k) = \mathbf{x}^T(k)\mathbf{w}(k)$, and calculates the error signal $e(k) = y(k) + n(k) - \hat{y}(k) = \tilde{y}(k) - \hat{y}(k)$, where $n(k)$ is the measurement noise, $y(k)$ is the output of the actual system, and $\hat{y}(k)$ is the estimated output. In order to estimate the channel impulse response, an adaptive algorithm minimizes the cost function defined by

$$\mathbf{w} = \arg\min_{\mathbf{w}} \frac{1}{2}\sum_{m=0}^{k} \lambda^{k-m}\left(e(m)\right)^2. \tag{1}$$

From the gradient based minimization, Equation (1) becomes

$$\mathbf{R}(k)\mathbf{w}(k) = \mathbf{r}(k), \tag{2}$$

where $\mathbf{R}(k) = \sum_{m=0}^{k} \lambda^{k-m}\mathbf{x}(m)\mathbf{x}^T(m)$ and $\mathbf{r}(k) = \sum_{m=0}^{k} \lambda^{k-m}\tilde{y}(m)\mathbf{x}(m)$. This equation is the normal equation for the least squares solution. Especially, $\mathbf{w}_o(k)$ is considered as a sparse system when the number of nonzero coefficients K is less than the system order of M. In order to estimate the sparse system, most estimation algorithms exploit non-zero coefficients in the system [11–17]. In [17], Eksioglu proposed a full recursive l_1-regularized algorithm by the minimization of the object function as shown in Equation (3).

$$J_k = \frac{1}{2}\varepsilon_k + \gamma_k\|\mathbf{w}\|_1, \tag{3}$$

where $\varepsilon_k = \sum_{m=0}^{k} \lambda^{k-m}(e(m))^2$. From the minimization of Equation (3), a modified normal equation was derived as shown in Equation (4).

$$\mathbf{R}(k)\mathbf{w}(k) = \mathbf{r}(k) - \gamma_k \nabla^s\|\mathbf{w}(k)\|_1 = \hat{\mathbf{p}}(k). \tag{4}$$

When we solve Equation (4), we should select the regularization factor as shown in Equation (5).

$$\gamma_k = \frac{2^{\frac{tr\left(\mathbf{R}^{-1}(k)\right)}{M}}}{\left\|\mathbf{R}^{-1}(k)\nabla^s f\left(\mathbf{w}(k)\right)\right\|_2^2} \times [\left(f\left(\mathbf{w}(k)\right)-\rho\right)+\nabla^s f\left(\mathbf{w}(k)\right)\mathbf{R}^{-1}(k)\varepsilon(k)], \tag{5}$$

where $f(\mathbf{w}(k)) = \|\mathbf{w}(k)\|_1$ and the subgradient of $f(\mathbf{w})$ is $\nabla^s\|\mathbf{w}\|_1 = \mathrm{sgn}(\mathbf{w})$. In Equation (5), the regularization factor has the parameter, ρ, which should be set beforehand. In [17], the parameter was set as $\rho = f(\mathbf{w}_{true}) = \|\mathbf{w}_{true}\|_1$ with $\mathbf{w}_{true}$ indicating the impulse response of the true channel. There was no further discussion about how to set ρ. However, it is not practical to know the true channel in advance.

3. Measure of Sparseness

In [20], the sparseness of a channel impulse response is measured by Equation (6).

$$\chi = \frac{L}{L-\sqrt{L}}\left(1-\frac{\|\hat{\mathbf{w}}\|_1}{\sqrt{L}\|\hat{\mathbf{w}}\|_2}\right), \tag{6}$$

where $\|\hat{\mathbf{w}}\|_p$ is the p-norm of $\hat{\mathbf{w}}$ and L is the dimension of $\hat{\mathbf{w}}$. The range of χ is $0 \le \chi \le 1$. That is dependent on the sparseness of $\hat{\mathbf{w}}$. As $\hat{\mathbf{w}}$ becomes sparser, the sparsity, χ, comes close to 1, and as $\hat{\mathbf{w}}$ becomes denser, χ comes close to 0. We often have small and none-zero value of χ, even in a dense channel. For example, Figure 1 shows the relation of the value of χ and the percentage of none-zero components in $\hat{\mathbf{w}}$ with $L = 215$. In Figure 1, we consider all possible cases of none-zero components in $\hat{\mathbf{w}}$.

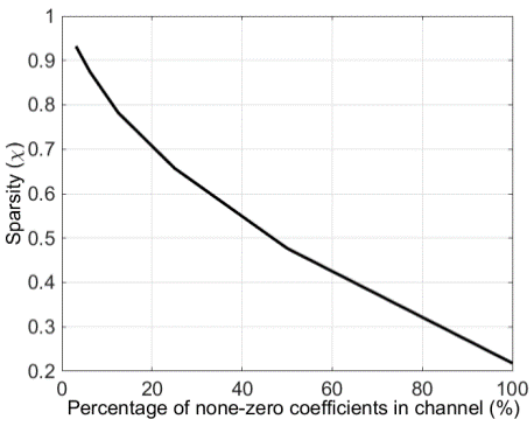

Figure 1. Sparsity (χ) vs. the percentage of none zero coefficients in the channel impulse response.

4. New ρ Selection Method in the Sparsity Regularization Constant γ_k

Section 2 shows that the regularization constant γ_k in Equation (5) needs ρ to be set as $\rho = \|\text{true system impulse response}\|_1 = \|\mathbf{w}_{true}\|_1$. However, we need a new method in the constant selection because Equation (5) is not practical. Therefore, Section 4 proposes a new method to set this constant.

For a practical method for the constant selection, we can consider using the estimated vector $\hat{\mathbf{w}}$ instead of using the true vector $\mathbf{w}_{true}$ because $\hat{\mathbf{w}}$, the solution with l_1-norm, will be closer to the sparse true vector than the solution of the conventional RLS. The more iteration is repeated, the more $\hat{\mathbf{w}}$ converges to the true value. Conventional RLS also converges to the true value; however, the solution with l_1-norm, is closer to the sparse true value. Therefore, we can use sparse estimate $\hat{\mathbf{w}}$ instead of $\mathbf{w}_{true}$ when we set ρ, and the uncertainty arising from this is compensated through a TLS

solution in the next section. When we determine ρ using the estimated $\hat{\mathbf{w}}$, we choose between the average ρ and the current estimate $\|\hat{\mathbf{w}}\|_1$. Table 1 summarizes the ρ selection steps.

Table 1. ρ selection method in the sparsity regularization constant γ_k

Step 1	Sparsity: $\chi = \frac{L}{L-\sqrt{L}}\left(1-\frac{\|\hat{\mathbf{w}}\|_1}{\sqrt{L}\|\hat{\mathbf{w}}\|_2}\right)$ [20] where L is the length of the impulse response.
Step 2	$\rho(k) = 0.99\rho(k-1) + 0.01\left(e^{\chi-1} \times \|\hat{\mathbf{w}}\|_1\right)$
Step 3	$\rho(k) = \begin{cases} \min\left(\rho(k), 0.98\|\hat{\mathbf{w}}\|_1\right), \text{if } x > 0.75 \\ \min\left(\rho(k), 0.999\|\hat{\mathbf{w}}\|_1\right), \text{otherwise} \end{cases}$

The determination method for ρ value shown in Table 1 is as follows. In Step 1, the sparsity of the estimated $\hat{\mathbf{w}}$ is calculated. The sparsity represents the sparseness of $\hat{\mathbf{w}}$ as a number [23]. In Step 2, l_1-norm of the estimated $\hat{\mathbf{w}}$ is scaled and the value is averaged with the previous ρ value. The scaling value approaches 1 as the sparsity, χ, gets close to 1. However, the scaling value gets close to $e^{-1} \simeq 0.37$ as the sparsity, χ, gets close to 0. Therefore, the scaling does not change l_1-norm of $\hat{\mathbf{w}}$ for the sparse $\hat{\mathbf{w}}$. Instead the scaling changes the l_1-norm smaller for the dense $\hat{\mathbf{w}}$. In Step 3, the smaller one between the averaged ρ and the l_1-norm of the estimated $\hat{\mathbf{w}}$ is selected as the new ρ value. In this case, the ρ value becomes completely new if the l_1-norm of the estimated $\hat{\mathbf{w}}$ is selected, otherwise the previous trend is maintained. In Figure 1, the reference value 0.75 used in Step 3 means that less than 16% of all the impulse response taps are not zero.

5. New Modeling for l_1-RLS with Uncertainty in the Regularization Factor

If we set $\rho = \text{constant}$, the regularization factor becomes

$$\begin{aligned}\tilde{\gamma}_k &= \frac{2\frac{tr\left(\mathbf{R}^{-1}(k)\right)}{M}}{\left\|\mathbf{R}^{-1}(k)\nabla^s f\left(\mathbf{w}(k)\right)\right\|_2^2} \times [\left(f\left(\mathbf{w}(k)\right) - \text{constant}\right) + \nabla^s f\left(\mathbf{w}(k)\right) \times \mathbf{R}^{-1}(k)\varepsilon(k)] \\ &= \frac{2\frac{tr\left(\mathbf{R}^{-1}(k)\right)}{M}}{\left\|\mathbf{R}^{-1}(k)\nabla^s f\left(\mathbf{w}(k)\right)\right\|_2^2} \times \left(f\left(\mathbf{w}(k)\right) - \|\mathbf{h}\|_1 + \|\mathbf{h}\|_1 - \text{constant}\right) + \frac{2\frac{tr\left(\mathbf{R}^{-1}(k)\right)}{M}\nabla^s f\left(\mathbf{w}(k)\right)\mathbf{R}^{-1}(k)\varepsilon(k)}{\left\|\mathbf{R}^{-1}(k)\nabla^s f\left(\mathbf{w}(k)\right)\right\|_2^2}.\end{aligned} \tag{7}$$

Then,

$$\tilde{\gamma}_k = \gamma_k + \frac{2\frac{tr\left(\mathbf{R}^{-1}(k)\right)}{M}\left(\|\mathbf{h}\|_1 - \text{constant}\right)}{\left\|\mathbf{R}^{-1}(k)\nabla^s f\left(\mathbf{w}(k)\right)\right\|_2^2} = \gamma_k + \Delta\gamma. \tag{8}$$

Using Equation (8), Equation (4) becomes

$$\mathbf{R}(k)\mathbf{w}(k) = \mathbf{r}(k) - (\gamma_k + \Delta\gamma)\nabla^s\|\mathbf{w}(k)\|_1 \quad . \tag{9}$$

$\nabla^s\|\mathbf{w}\|_1 = \text{sgn}(\mathbf{w})$ is represented as

$$\nabla^s\|\mathbf{w}(k)\|_1 = \begin{bmatrix} \ddots & & \\ & \frac{1}{|w_i|} & \\ & & \ddots \end{bmatrix}\mathbf{w}(k). \tag{10}$$

By applying Equation (10) to Equation (9),

$$\left(\mathbf{R}(k)+\Delta\gamma\begin{bmatrix}\ddots & & \\ & \frac{1}{|w_i|} & \\ & & \ddots\end{bmatrix}\right)\mathbf{w}(k)=\mathbf{r}(k)-\gamma_k\|\mathbf{w}(k)\|_1, \tag{11}$$

where w_i is i-th element of $\mathbf{w}(k)$. Then it is simplified as

$$\left(\mathbf{R}(k)+\Delta\gamma\begin{bmatrix}\ddots & & \\ & \frac{1}{|w_i|} & \\ & & \ddots\end{bmatrix}\right)\mathbf{w}(k)=\hat{\mathbf{p}}(k). \tag{12}$$

Equation (12) is very similar to the system model in Figure 2 that is contaminated by noise both in input and in output. Suppose that an example of the system in Figure 2 is represented as

$$\begin{bmatrix} x_k+n_{i,k} & \cdots & x_{k-N+1}+n_{i,k-N+1} \\ x_{k-1}+n_{i,k-1} & \cdots & x_{k-N}+n_{i,k-N} \\ \vdots & \ddots & \vdots \\ x_{k-N+1}+n_{i,k-N+1} & \cdots & x_{k-2N+2}+n_{i,k-2N+2} \end{bmatrix}\times\mathbf{w}(k)=\begin{bmatrix} y_k+n_{o,k} \\ y_{k-1}+n_{o,k-1} \\ \vdots \\ y_{k-N+1}+n_{o,k-N+1} \end{bmatrix}, \tag{13}$$

where x_k is $x(k)$, $n_{i,k}$ is $n_i(k)$, and $n_{o,k}$ is $n_o(k)$. Equation (13) is simplified as

$$\mathbf{A}\mathbf{w}(k)=\mathbf{b}. \tag{14}$$

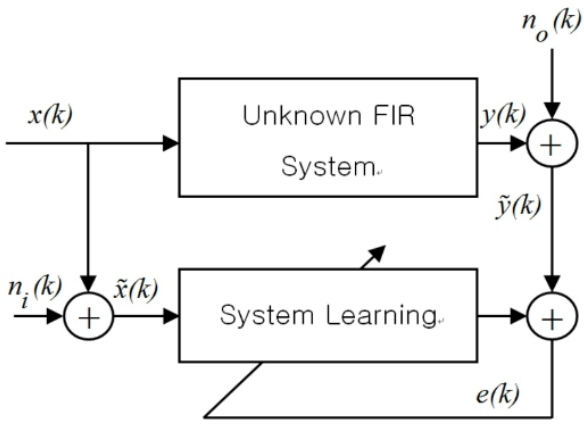

Figure 2. The model of a noisy input and noisy output system.

If we multiply Equation (14) by $\mathbf{A}^H$ and average it, we get

$$E\left(\mathbf{A}^H\mathbf{A}\right)\mathbf{w}(k)=E\left(\mathbf{A}^H\mathbf{b}\right). \tag{15}$$

We can rewrite Equation (15) as follows

$$\begin{bmatrix} r_{xx}(0)+\sigma_n^2 & r_{xx}(1) & \cdots & r_{xx}(N-1) \\ r_{xx}(1) & r_{xx}(0)+\sigma_n^2 & \cdots & r_{xx}(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}(N-1) & r_{xx}(N-2) & \cdots & r_{xx}(0)+\sigma_n^2 \end{bmatrix}\mathbf{w}(k)=\begin{bmatrix} r_{xy}(0) \\ r_{xy}(1) \\ \vdots \\ r_{xy}(N-1) \end{bmatrix} \tag{16}$$

Then, it can be represented as

$$\left(\mathbf{R}+\sigma_n^2\mathbf{I}\right)\mathbf{w}(k)=\tilde{\mathbf{p}}(k). \tag{17}$$

When we compare Equation (12) with Equation (17), the two system models have almost the same form. Therefore, it is feasible that the TLS method can be applied to Equation (12) [24–30]. Therefore, we expect to obtain almost the same performance as l_1-RLS with the true channel response if we apply the TLS method by the regularization factor with the new ρ in Table 1. In the next section, we summarize l_1-RTLS (recursive total least squares) algorithm in [29].

6. Summarize l_1-RTLS for the Solution of l_1-RLS with Uncertainty in the Regularization Factor

Lim, one of the authors of this paper, has proposed the TLS solution for l_1-RLS known as l_1-RTLS [30]. In this section, we summarize l_1-RTLS in [30] for the solution of Equation (11).

The TLS system model assumes that both input and output are contaminated by additive noise as Figure 2. The output is given by

$$\tilde{y}(k) = \tilde{\mathbf{x}}^T(k)\mathbf{w}_o + n_o(k), \tag{18}$$

where the output noise $n_o(k)$ is the Gaussian white noise with variance σ_o^2. The noisy input vector in the system is modeled by

$$\tilde{\mathbf{x}}(k) = \mathbf{x}(k) + \mathbf{n}_i(k) \in C^{M\times 1}, \tag{19}$$

where $\mathbf{n}_i(k) = \left[n_i(k), n_i(k-1), \cdots n_i(k-M+1)\right]^T$ and the input noise $n_i(k)$ is the Gaussian white noise with variance σ_i^2. For the TLS solution, we set the augmented data vector as

$$\bar{\mathbf{x}}(k) = \left[\tilde{\mathbf{x}}^T(k), \tilde{y}(k)\right]^T \in \mathrm{R}^{(M+1)\times 1}. \tag{20}$$

The correlation matrix is represented as

$$\bar{\mathbf{R}} = \begin{bmatrix} \tilde{\mathbf{R}} & \mathbf{p} \\ \mathbf{p}^T & c \end{bmatrix}, \tag{21}$$

where $\mathbf{p} = E\left\{\tilde{\mathbf{x}}(k)y(k)\right\}$, $c = E\left\{y(k)y(k)\right\}$, $\mathbf{R} = E\left\{\mathbf{x}(k)\mathbf{x}^T(k)\right\}$ and $\tilde{\mathbf{R}} = E\left\{\tilde{\mathbf{x}}(k)\tilde{\mathbf{x}}^T(k)\right\} = \mathbf{R} + \sigma_i^2\mathbf{I}$. In [27,28], the TLS problem becomes to find the eigenvector associated with the smallest eigenvalue of $\bar{\mathbf{R}}$. Equation (22) is the typical cost function to find the eigenvector associated with the smallest eigenvalue of $\bar{\mathbf{R}}$.

$$J(k) = \frac{1}{2}\tilde{\mathbf{w}}^T(k)\bar{\mathbf{R}}(k)\tilde{\mathbf{w}}(k), \tag{22}$$

where $\bar{\mathbf{R}}(k)$ is a sample correlation matrix at k-th instant, and $\tilde{\mathbf{w}}(k) = \left[\hat{\mathbf{w}}^T(k), -1\right]^T$ in which $\hat{\mathbf{w}}(\mathrm{k})$ is the estimation result for the unknown system at k-th instant. We modify the cost function by adding a penalty function in order to reflect prior knowledge about the true sparsity system.

$$J(k) = \frac{1}{2}\tilde{\mathbf{w}}^T(k)\bar{\mathbf{R}}(k)\tilde{\mathbf{w}}(k) + \lambda\left(\tilde{\mathbf{w}}^T(k)\tilde{\mathbf{w}}(k-1) - 1\right) + \gamma_k f\left(\tilde{\mathbf{w}}(k)\right), \tag{23}$$

where λ is the Lagrange multiplier and γ_k is the regularized parameter in [13]. We solve the equations by $\nabla_{\hat{\mathbf{w}}} J(k) = 0$ and $\nabla_\lambda J(k) = 0$ simultaneously. $\nabla_{\hat{\mathbf{w}}} J(k) = 0$:

$$2\bar{\mathbf{R}}(k)\tilde{\mathbf{w}}(k) + \lambda\tilde{\mathbf{w}}(k-1) + \gamma_k \nabla^s f\left(\tilde{\mathbf{w}}(k)\right) = 0, \tag{24}$$

$$\nabla_\lambda J(k) = 0: \quad \tilde{\mathbf{w}}^T(k)\tilde{\mathbf{w}}(k-1) = 1, \tag{25}$$

where the subgradient of $f(\tilde{\mathbf{w}}) = \|\tilde{\mathbf{w}}\|_1$ is $\nabla_{\tilde{\mathbf{w}}}^s \|\tilde{\mathbf{w}}\|_1 = \mathrm{sgn}(\tilde{\mathbf{w}})$. From (24), we obtain

$$\tilde{\mathbf{w}}(k) = -\frac{\lambda}{2}\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1) - \gamma_k \bar{\mathbf{R}}^{-1}(k)\nabla^s f\left(\tilde{\mathbf{w}}(k)\right). \tag{26}$$

Substituting Equation (26) in Equation (25), we get

$$\left(-\frac{\lambda}{2}\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1)-\gamma_k\bar{\mathbf{R}}^{-1}(k)\nabla^s f\left(\tilde{\mathbf{w}}(k)\right)\right)^T\times\tilde{\mathbf{w}}(k-1)=1, \tag{27}$$

or

$$\lambda=-2\frac{1+\gamma_k\nabla^s f\left(\tilde{\mathbf{w}}(k)\right)^T\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1)}{\tilde{\mathbf{w}}^T(k-1)\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1)}. \tag{28}$$

Substituting λ in Equation (26) by Equation (28) leads to

$$\tilde{\mathbf{w}}(k)=\frac{1+\gamma_k\nabla^s f\left(\tilde{\mathbf{w}}(k)\right)^T\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1)}{\tilde{\mathbf{w}}^T(k-1)\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1)}\times\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1)-\gamma_k\bar{\mathbf{R}}^{-1}(k)\nabla_{\tilde{\mathbf{w}}} f\left(\tilde{\mathbf{w}}(k)\right). \tag{29}$$

Equation (29) can be expressed in a simple form as

$$\tilde{\mathbf{w}}(k)=\alpha\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1)-\gamma_k\bar{\mathbf{R}}^{-1}(k)\nabla^s f\left(\tilde{\mathbf{w}}(k)\right), \tag{30}$$

where $\alpha=\frac{1+\gamma_k\nabla^s f\left(\tilde{\mathbf{w}}(k)\right)^T\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1)}{\tilde{\mathbf{w}}^T(k-1)\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1)}$. Because asymptotically $\|\tilde{\mathbf{w}}(k)\|=1$ as $k\to\infty$, Equation (29) can be approximated as the following two equations.

$$\tilde{\mathbf{w}}(k)\simeq\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1)-\gamma_k\left(\tilde{\mathbf{w}}^T(k-1)\bar{\mathbf{R}}^{-1}(k-1)\tilde{\mathbf{w}}(k-1)\right)\bar{\mathbf{R}}^{-1}(k)\nabla^s f\left(\tilde{\mathbf{w}}(k-1)\right). \tag{31}$$

$$\tilde{\mathbf{w}}(k)=\tilde{\mathbf{w}}(k)/\|\tilde{\mathbf{w}}(k)\|. \tag{32}$$

Finally, we obtain the estimated parameter of the unknown system as

$$\hat{\mathbf{w}}(k)=-\tilde{\mathbf{w}}_{1:M}(k)/\tilde{w}_{M+1}(k)\ . \tag{33}$$

For Equation (23), we can use the modified regularization factor γ_k in [30]

$$\gamma_k=\frac{2\frac{tr\left(\bar{\mathbf{R}}^{-1}(k)\right)}{M}}{\left\|\bar{\mathbf{R}}^{-1}(k)\nabla^s f\left(\hat{\mathbf{w}}_{aug}(k)\right)\right\|_2^2}\times[\left(f\left(\hat{\mathbf{w}}_{aug}(k)\right)-\rho\right)+\nabla^s f\left(\hat{\mathbf{w}}_{aug}(k)\right)\bar{\mathbf{R}}^{-1}(k)\varepsilon(k)], \tag{34}$$

where $\hat{\mathbf{w}}_{aug}(k)=\left[\hat{\mathbf{w}}^T(k),-1\right]^T$, $\hat{\mathbf{w}}_{aug,RLS}(k)=\left[\hat{\mathbf{w}}_{RLS}^T(k),-1\right]^T$, $\varepsilon(k)=\hat{\mathbf{w}}_{aug}(k)-\hat{\mathbf{w}}_{aug,RLS}(k)$, and $\hat{\mathbf{w}}_{RLS}(k)$ is the estimated parameter by recursive least squares (RLS). As $f(\hat{\mathbf{w}})=\|\hat{\mathbf{w}}\|_1$, the subgradient of $f(\hat{\mathbf{w}}_{aug}(k))$ is

$$\nabla^s\|\hat{\mathbf{w}}_{aug}(k)\|_1=\mathrm{sgn}(\hat{\mathbf{w}}_{aug}(k))\ . \tag{35}$$

As mentioned in Section 4, we apply new constant ρ in Table 1, to the regularization factor γ_k in Equation (34) instead of $\|\mathbf{w}_{true}\|_1$, where $\mathbf{w}_{true}$ is the true system impulse response.

7. Simulation Results

This section confirms the performance of the proposed algorithm in sparse channel estimation. In the first experiment, the channel estimation performance is compared with other algorithms using randomly generated sparse channels. In this simulation, we follow the same scenario in the experiments as [17]. The true system vector $\mathbf{w}_{true}$ is 64 dimensions. In order to generate the sparse channel, we set the number of the nonzero coefficients, S, in the 64 coefficients and randomly position the nonzero coefficients. The values of the coefficients are taken from an $N(0,1/\mathrm{S})$ distribution, where $N()$ is the normal distribution. In the simulation, we estimate the channel impulse response by the proposed algorithms that are l_1-RLS using the ρ in Table 1 and l_1-RTLS using the ρ in Table 1. For the comparison, we estimate the channel impulse response by l_1-RLS using the true channel

response; in addition, we also execute the regular RLS algorithm in an oracle setting (oracle-RLS) where the positions of the true nonzero system parameters are assumed to be known. For the estimated channel results, we calculate the mean standard deviation (MSD), where $\mathrm{MSD} = E\left(\left|\hat{\mathbf{w}} - \mathbf{w}_{true}\right|^2\right)$, $\hat{\mathbf{w}}$ is the estimated channel response and $\mathbf{w}_{true}$ is the true channel response. For the performance evaluation, we simulate the algorithms in the sparse channels for S = 4, 8, 16, and 32.

Figure 3 illustrates the MSD curves. For S = 4, Figure 3a shows that the estimation performance of l_1-RTLS using the regularization factor with the ρ in Table 1 is almost the same as the l_1-RLS using regularization with a true channel impulse response. However, the performance of l_1-RLS using the regularization factor with the ρ in Table 1 is gradually degraded and shows a kind of uncertainty accumulation effect. In the other cases of S, we can observe the same trend in the MSD curves. Therefore, we can confirm that the new regularization factor selection method and the new modeling for l_1-RLS can estimate the sparse channel as good as l_1-RLS using the regularization with the true channel impulse response. In all the simulation scenarios, oracle RLS algorithm produces the lowest MSD as expected.

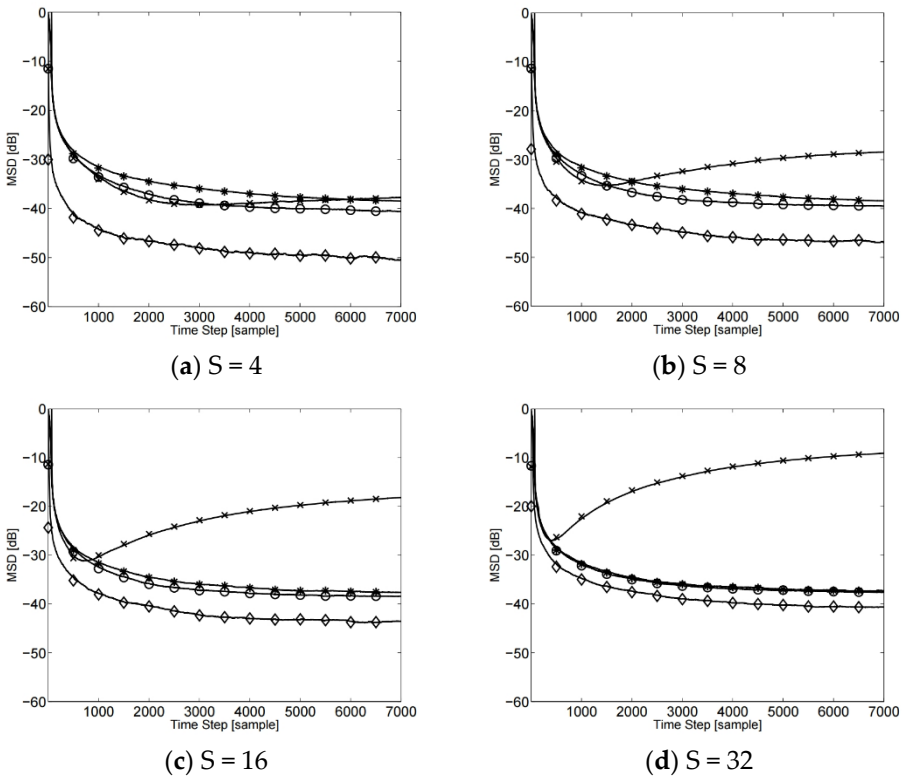

Figure 3. Steady-state MSD for S = 4, 8, 16, and 32 when applying the new ρ method to the regularization factor (-o-: l_1-RLS with the true channel response, -x-: l_1-RLS with the new ρ method, -*-: proposed l_1-RTLS with the new ρ method, - ◊ -: oracle-RLS).

Table 2 summarizes the steady-state MSD values as varying S from 4 to 32. The results show that the proposed l_1-RTLS with the new ρ is comparable to l_1-RLS with the true channel.

In the second experiment, we compare channel estimation performance using room impulse response. The size of the room is (7.49, 6.24, 3.88 m). The position of the sound source is (1.53, 0.96, 1.12 m) and the position of the receiver is (1.81, 5.17, 0.71 m), respectively. T60 is set to 100 ms and 400 ms. The impulse response of the room is generated using the program in [31]. We focus on the direct reflection part and the early reflection part in the RIR because the direct reflection and early reflection part of the RIR has a sparse property. This is the part that is estimated in the AEC applications [32]. This part is also related to localization and clarity in room acoustics [33–35]. Comparing the impulse response (IR) generated by setting T60 = 100 ms to the channel with 65 coefficients used in the first

experiment, it is equivalent to S = 4 in the channel with 65 coefficients. In the same manner, the IR generated by setting T60 = 400 ms is equivalent to S = 10.

Table 2. MSD (mean square deviation) comparison.

Sparsity (S)	Algorithm	MSD
4	l_1-RLS with the true channel	−40.6 dB
	l_1-RLS with the new ρ method	−37.8 dB
	proposed l_1-RTLS with the new ρ method	−38.5 dB
	Oracle-RLS	−50.4 dB
8	l_1-RLS with the true channel	−39.5 dB
	l_1-RLS with the new ρ method	−28.4 dB
	proposed l_1-RTLS with the new ρ method	−38.5 dB
	Oracle-RLS	−46.9 dB
16	l_1-RLS with the true channel	−38.4 dB
	l_1-RLS with the new ρ method	−18.2 dB
	proposed l_1-RTLS with the new ρ method	−37.6 dB
	Oracle-RLS	−43.6 dB
32	l_1-RLS with the true channel	−37.6 dB
	l_1-RLS with the new ρ method	−9.1 dB
	proposed l_1-RTLS with the new ρ method	−37.3 dB
	Oracle-RLS	−40.6 dB

Table 3 summarizes the steady-state MSD values. The results also show the same trend as Table 2. In RIR estimation, the proposed l_1-RTLS with the new ρ is also comparable to l_1-RLS with the true channel.

Table 3. MSD (mean square deviation) comparison in sparse RIR estimations.

Reverberation Time (T60)	Algorithm	MSD
100 ms	l_1-RLS with the true channel	−38.5 dB
	l_1-RLS with the new ρ method	−34.7 dB
	proposed l_1-RTLS with the new ρ method	−35.4 dB
	Oracle-RLS	−45.3 dB
400 ms	l_1-RLS with the true channel	−32.1 dB
	l_1-RLS with the new ρ method	−20.9 dB
	proposed l_1-RTLS with the new ρ method	−30.1 dB
	Oracle-RLS	−36.0 dB

8. Conclusions

In this paper, we have proposed the regularization factor for recursive adaptive estimation. The regularization factor needs no prior knowledge of the true channel impulse response. We have also reformulated the recursive estimation algorithm as l_1-RTLS type. This formulation is robust to the uncertainty in the regularization factor without a priori knowledge of the true channel impulse response. Simulations show that the proposed regularization factor and l_1-RTLS algorithm provide good performance comparable to l_1- RLS with the knowledge of the true channel impulse response.

Author Contributions: Conceptualization, J.L.; Methodology, J.L.; Validation, J.L. and S.L.; Formal analysis, J.L.; Investigation, J.L. and S.L.; Writing—original draft preparation, J.L.; Writing—review and editing, S.L.; Visualization, J.L.; Project administration, J.L.; Funding acquisition, J.L. and S.L.

Funding: This research received no external funding.

Acknowledgments: This research was supported by Agency for Defense Development (ADD) in Korea (UD160015DD).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Benichoux, A.; Simon, L.; Vincent, E.; Gribonval, R. Convex regularizations for the simultaneous recording of room impulse responses. *IEEE Trans. Signal Process.* **2014**, *62*, 1976–1986.
2. Merimaa, J.; Pulkki, V. Spatial impulse response I: Analysis and synthesis. *J. Audio Eng. Soc.* **2005**, *53*, 1115–1127.
3. Dokmanic, I.; Parhizkar, R.; Walther, A.; Lu, Y.M.; Vetterli, M. Acoustic echoes reveal room shape. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 12186–12191.
4. Remaggi, L.; Jackson, P.; Coleman, P.; Wang, W. Acoustic reflector localization: Novel image source reversion and direct localization methods. *IEEE Trans. Audio Speech Lang. Process.* **2017**, *25*, 296–309.
5. Baba, Y.; Walther, A.; Habets, E. 3D room geometry interference based on room impulse response stacks. *IEEE Trans. Audio, Speech Lang. Process.* **2018**, *26*, 857–872.
6. Goetze, S.; Xiong, F.; Jungmann, J.O.; Kallinger, M.; Kammeyer, K.; Mertins, A. System Identification of Equalized Room Impulse Responses by an Acoustic Echo Canceller using Proportionate LMS Algorithms. In Proceeding of the 130th AES Convention, London, UK, 13 May 2011; pp. 1–13.
7. Yu, M.; Ma, W.; Xin, J.; Osher, S. Multi-channel l_1 regularized convex speech enhancement model and fast computation by the split Bregman method. *IEEE Trans. Audio Speech Lang. Process.* **2012**, *20*, 661–675.
8. Lin, Y.; Chen, J.; Kim, Y.; Lee, D.D. Blind channel identification for speech dereverberation using l1-norm sparse learning. In Proceeding of the Twenty-First Annual Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 3–6 December 2007; pp. 921–928.
9. Naylor, P.A.; Gaubitch, N.D. *Speech Dereverberation*; Springer: London, UK, 2010; pp. 219–270.
10. Duttweiler, D.L. Proportionate normalized least-mean-squares adaptation in echo cancelers. *IEEE Trans. Speech Audio Process.* **2000**, *8*, 508–518.
11. Gu, Y.; Jin, J.; Mei, S. Norm Constraint LMS Algorithm for Sparse System Identification. *IEEE Signal Process. Lett.* **2009**, *16*, 774–777.
12. Chen, Y.; Gu, Y.; Hero, A.O. Sparse LMS for system identification. In Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009; pp. 3125–3128.
13. He, X.; Song, R.; Zhu, W.P. Optimal pilot pattern design for compressed sensing-based sparse channel estimation in OFDM systems. *Circuits Syst. Signal Process.* **2012**, *31*, 1379–1395.
14. Babadi, B.; Kalouptsidis, N.; Tarokh, V. SPARLS: The sparse RLS algorithm. *IEEE Trans. Signal Process.* **2010**, *58*, 4013–4025.
15. Angelosante, D.; Bazerque, J.A.; Giannakis, G.B. Online adaptive estimation of sparse signals: Where RLS meets the ℓ_1-norm. *IEEE Trans. Signal Process.* **2010**, *58*, 3436–3447.
16. Eksioglu, E.M. Sparsity regularised recursive least squares adaptive filtering. *IET Signal Process*. **2011**, *5*, 480–487.
17. Eksioglu, E.M.; Tanc, A.L. RLS algorithm with convex regularization. *IEEE Signal Process. Lett.* **2011**, *18*, 470–473.
18. Sun, D.; Liu, L.; Zhang, Y. Recursive regularisation parameter selection for sparse RLS algorithm. *Electron. Lett.* **2018**, *54*, 286–287.
19. Chen, Y.; Gui, G. Recursive least square-based fast sparse multipath channel estimation. *Int. J. Commun. Syst.* **2017**, *30*, e3278.
20. Kalouptsidis, N.; Mileounis, G.; Babadi, B.; Tarokh, V. Adaptive algorithms for sparse system identification. *Signal Process.* **2011**, *91*, 1910–1919.
21. Candes, E.J.; Wakin, M.; Boyd, S. Enhancing sparsity by reweighted minimization. *J. Fourier Anal. Appl.* **2008**, *14*, 877–905.
22. Lamare, R.C.; Sampaio-Neto, R. Adaptive reduced-rank processing based on joint and iterative interpolation, decimation, and filtering. *IEEE Trans. Signal Process.* **2009**, *57*, 2503–2514.
23. Petraglia, M.R.; Haddad, D.B. New adaptive algorithms for identification of sparse impulse responses—Analysis and comparisons. In Proceeding of the Wireless Communication Systems, York, UK, 19–22 September 2010; pp. 384–388.
24. Golub, G.H.; Van Loan, C.F. An analysis of the total least squares problem. *SIAM J. Numer. Anal.* **1980**, *17*, 883–893.

25. Dunne, B.E.; Williamson, G.A. Stable simplified gradient algorithms for total least squares filtering. In Proceeding of 34th Annual Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 29 October–1 November 2000; pp. 1762–1766.
26. Feng, D.Z.; Bao, Z.; Jiao, L.C. Total least mean squares algorithm. *IEEE Trans. Signal Process.* **1998**, *46*, 2122–2130.
27. Davila, C.E. An efficient recursive total least squares algorithm for FIR adaptive filtering. *IEEE Trans. Signal Process.* **1994**, *42*, 268–280.
28. Soijer, M.W. Sequential computation of total least-squares parameter estimates. *J. Guid. Control Dyn.* **2004**, *27*, 501–503.
29. Choi, N.; Lim, J.S.; Sung, K.M. An efficient recursive total least squares algorithm for raining multilayer feedforward neural networks. *Lect. Notes Comput. Sci.* **2005**, *3496*, 558–565.
30. Lim, J.S.; Pang, H.S. l_1-regularized recursive total least squares based sparse system identification for the error-in-variables. *SpringerPlus* **2016**, *5*, 1460–1469.
31. Lehmann, E. Image-Source Method: MATLAB Code Implementation. Available online: http://www.eric-lehmann.com/ (accessed on 17 December 2018).
32. Gay, S.L.; Benesty, J. *Acoustic Signal Processing for Telecommunication*; Kluwer Academic Publisher: Norwell, MA, USA, 2000; pp. 6–7.
33. Swanson, D.C. *Signal Processing for Intelligent Sensor Systems with MATLAB*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2012; p. 70.
34. Kuttruff, H. *Room Acoustics*, 6th ed.; CRC Press: Boca Raton, FL, USA, 2017; p. 168.
35. Bai, H.; Richard, G.; Daudet, L. Modeling early reflections of room impulse responses using a radiance transfer method. In Proceeding of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, USA, 20–23 October 2013; pp. 1–4.