Missouri University of Science and Technology

Scholars' Mine

Electrical and Computer Engineering Faculty Research & Creative Works

Electrical and Computer Engineering

01 Jan 2004

A Multiple Principal Components Based Adaptive Filter

Steven L. Grant

Missouri University of Science and Technology, sgrant@mst.edu

Follow this and additional works at: https://scholarsmine.mst.edu/ele_comeng_facwork

Part of the Electrical and Computer Engineering Commons

Recommended Citation

A Mutiple Principal Components Based Adaptive Filter[*]

Steven L. Grant and Steven L. Gay
University of Missouri-Rolla
228 EECH, 1870 Miner Circle
Rolla MO 65401
email: sgrant@umr.edu

ABSTRACT

Proportionate normalized least mean squares (PNLMS) is an adaptive filter that has been shown to provide exceptionally fast convergence and tracking when the underlying system parameters are sparse. A good example of such a system is a network echo canceller. Principal components based PNLMS (PCP) extends this fast convergence property to certain non-sparse systems by applying PNLMS while using the principal components of the underlying system as basis vectors. An acoustic echo canceller is a possible example of this type of non-sparse system. Simulations of acoustic echo paths and cancellers indicate that PCP converges and tracks much faster than the classical normalized least mean squares (NLMS) and fast recursive least squares (FRLS) adaptive filters. However, when a basic parameter, like room temperature, changes, the underlying acoustic structure of the room changes as well and principal components of the room responses at one temperature are very different from those at another. This paper addresses this problem by using multiple sets of principle components as basis vectors and performing PNLMS in each basis set. Each set of principle components are derived from the room at a different temperature. The new algorithm, multiple principal components PNLMS (MPCP) is a generalization of PNLMS++. Simulations show the potential effectiveness of the approach.

1. INTRODUCTION

Proportionate normalized least mean squares (PNLMS) algorithm [1, 2, 3] is an adaptive filter that converges and tracks much faster than the classical normalized least mean squares (NLMS) adaptive filter when the solution is sparse in non-zero terms. Developed independently, the exponentiated gradient (EG) adaptive filter [4] is very similar to PNLMS. The connection between the two algorithms is demonstrated by Benesty [5].

PCP (principal components PNLMS), exploits the fact that the sparseness of a vector is a function of the chosen basis set. It is well known that a given set of random vectors may be expressed most succinctly as a linear combination of the principal components [6] of the generating random process, i.e., the eigenvectors of the random vectors' covariance matrix. By choosing the principal components of a statistical sampling of solution vectors as a basis set, the solutions may be represented more sparsely than otherwise. Thus, PNLMS operating under such a basis set converges faster than under the original basis set.

In [7] the efficacy of PCP was demonstrated by applying it to a simulation of the acoustic echo cancellation (AEC) problem. One problem with applying PCP to AEC is that the speed of sound in air changes by as much as 1% with as little as five degrees centigrade change in temperature. This dramatically alters the room impulse response statistics; meaning the principal components at one temperature are quite different from those at another. Therefore, PCP will not perform well over even a modest temperature range. We address this by generalizing the technique used in PNLMS++ [2]. There, both PNLMS and NLMS adaptive techniques were used to adapt the coefficients of a telephone network echo canceller. The concern was that in rare cases where the echo response was not sparse, PNLMS converged slower than NLMS. By efficiently implementing both techniques, we obtained fast convergence for sparse echo paths and standard convergence for dispersive echo paths. With MPCP, we perform multiple adaptations using PCP with basis sets obtained from the room at different temperatures, thus obtaining fast convergence over a range of environments.

2. PNLMS

First, we briefly review the PNLMS adaptive filter under the guise of the echo cancellation problem. The signals,

---

[*] This work was performed under the Wilkens Missouri Endowment.

vectors, and matrices used in this paper are as follows: $x(n)$ is the far-end signal which excites the echo path, $\mathbf{x}(n) = \left[x(n), \cdots, x(n-L+1)\right]^T$ is the excitation vector, the true echo path impulse response vector is $\mathbf{h}_{ep} = \left[h_0, \ldots, h_{L-1}\right]^T$, v_n is the near-end signal, or near-end noise, $y(n) = \mathbf{x}^T(n)\mathbf{h}_{ep} + v(n)$ is the combination of the echo and the near-end signals, $\mathbf{h}(n) = \left[h_0(n), \ldots, h_{L-1}(n)\right]^T$ is the adaptive filter coefficient vector, $e(n) = y(n) - \mathbf{x}^T(n)\mathbf{h}(n-1)$ is the error or residual-echo signal, $\mathbf{G}(n) = diag\left\{g_0(n), \ldots, g_{L-1}(n)\right\}$ is the diagonal individual step-size matrix, μ, is the "stepsize" parameter and is chosen in the range, $0 < \mu < 1$, and δ is a small positive number known as the regularization parameter.

An NLMS adaptive filter iteration involves the following two steps:

$$e(n) = y(n) - \mathbf{x}^T(n)\mathbf{h}(n), \qquad (1)$$

the error calculation, and

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \mu\mathbf{x}(n)\left[\mathbf{x}^T(n)\mathbf{x}(n) + \delta\right]^{-1}e(n), \qquad (2)$$

the coefficient vector update. PNLMS is similar, except that in the coefficient vector update a *proportionate* diagonal matrix, $\mathbf{G}(n)$, whose elements are roughly proportionate to the magnitude of the coefficient vector, $\mathbf{h}(n-1)$, is used to window the excitation vector, $\mathbf{x}(n)$,

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \\ + \mu\mathbf{G}(n)\mathbf{x}(n)\left[\mathbf{x}^T(n)\mathbf{G}(n)\mathbf{x}(n) + \delta\right]^{-1}e(n) \qquad (3)$$

where,

$$\mathbf{G}(n) = f\left(\rho, \delta_p, \mathbf{h}(n-1)\right) \qquad (4)$$

and $f\left(\rho, \delta_p, \mathbf{h}(n-1)\right)$ is a nonlinear function described by the series of steps in Table 1. The computational complexity of the steps of Table 1 is approximately L operations per sample period as step (a) need not be done every sample period and the normalization in step (d) can be absorbed in to the relaxation parameter, μ in equation (3).

Table 1: $f\left(\rho, \delta_p, \mathbf{h}(n-1)\right)$

Step	Calculation
a	$L_{max} = \max\left\{\delta_p, \left|h_0(n-1)\right|, \ldots, \left|h_{L-1}(n-L+1)\right|\right\}$
b	$\gamma_i = \max\left\{\rho L_{max}, \left|h_i(n-1)\right|\right\} \quad 0 \le i \le L-1$
c	$L_1 = \sum_{i=0}^{L-1} \gamma_i$
d	$\left[\mathbf{G}(n)\right]_{i,i} = \gamma_i / L_1 \qquad 0 \le i \le L-1$

When $\mathbf{h}_{ep}$ is sparse and $\mathbf{h}(n-1)$ converges to it, the filter becomes effectively *shorter* because of $\mathbf{G}(n)$'s windowing of $\mathbf{x}(n)$. Since shorter adaptive filters converge faster than longer ones, PNLMS's convergence is accelerated. On the other hand, when $\mathbf{h}_{ep}$ is dispersive, PNLMS has no advantage over NLMS, in fact its convergence is significantly slower. PNLMS++ [2] and IPNLMS [3] were designed to improve the convergence rate for dispersive impulse responses so that they converge at least as fast as NLMS.

3. PNLMS WITH ARBITRARY BASIS SET

In this section, we investigate how PNLMS may operate under any arbitrary basis set. In the next section, we discuss how to find that basis set which makes the solution that PNLMS seeks sparse.

Let us rotate the acoustic impulse response with a linear transformation, say,

$$\mathbf{b}_{ep} = \mathbf{U}\mathbf{h}_{ep} \qquad (5)$$

where $\mathbf{U}$ is an L-by-L unitary matrix. We may then make the following substitution in equations (1) and (2),

$$\mathbf{h}(n-1) = \mathbf{U}^T\mathbf{b}(n-1) \qquad (6)$$

and

$$\mathbf{x}(n) = \mathbf{U}^T\mathbf{s}(n) \qquad (7)$$

Then, (1) can be rewritten as

$$e(n) = y(n) - \mathbf{s}^T(n)\mathbf{b}(n-1) \qquad (8)$$

and (2) as

$$\mathbf{U}^T\mathbf{b}(n) = \mathbf{U}^T\mathbf{b}(n-1) + \\ + \mu\mathbf{U}^T\mathbf{s}(n)\left[\mathbf{s}^T(n)\mathbf{U}\mathbf{U}^T\mathbf{s}(n) + \delta\right]^{-1}e(n) \qquad (9)$$

Multiplying both sides of (13) from the left by $\mathbf{U}$ and recalling that for unitary matrices, $\mathbf{I} = \mathbf{U}\mathbf{U}^T$, where $\mathbf{I}$ is the identity matrix,

$$\mathbf{b}(n) = \mathbf{b}(n-1) + \mu\mathbf{s}(n)\left[\mathbf{s}^T(n)\mathbf{s}(n) + \delta\right]^{-1}e(n) \qquad (10)$$

Equations, (8) and (10) represent NLMS in the transformed domain. We may now apply PNLMS by defining a proportionate diagonal matrix as a function of $\mathbf{b}(n-1)$,

$$\mathbf{M}(n) = f\left(\rho, \delta_p, \mathbf{b}(n-1)\right) \qquad (11)$$

In addition, the PNLMS coefficient update in the transform domain becomes,

$$\mathbf{b}(n) = \mathbf{b}(n-1) + \\ + \mu\mathbf{M}(n)\mathbf{s}(n)\left[\mathbf{s}^T(n)\mathbf{M}(n)\mathbf{s}(n) + \delta\right]^{-1}e(n) \qquad (12)$$

To summarize, the PCP algorithm is shown in Table 2 where we have added the intermediate step,

$$\mathbf{p} = \mathbf{M}(n)\mathbf{s}(n), \qquad (13)$$

to specify a more efficient calculation of (12).

946

Table 2: PCP

Step	Computation	Complexity
a	$\mathbf{s}(n) = \mathbf{U}\mathbf{x}(n)$	L^2
b	$e(n) = y(n) - \mathbf{s}^T(n)\mathbf{b}(n-1)$	L
c	$\mathbf{M}(n) = f(\rho, \delta_p, \mathbf{b}(n-1))$	L
d	$\mathbf{p} = \mathbf{M}(n)\mathbf{s}(n)$	L
e	$\mathbf{b}(n) = \mathbf{b}(n-1) + \mu\mathbf{p}\left[\mathbf{s}^T(n)\mathbf{p} + \delta\right]^{-1} e(n)$	$2L$

4. FINDING PROPER BASIS VECTORS

Up to this point we have viewed the solution vector, $\mathbf{h}_{ep}$, as fixed. In practical applications though, it is almost always time varying. Therefore, we now add a time index to the solution vector $\mathbf{h}_{ep}(n)$ and view it as the output of a random process. As such, we may use the well know tool of PCA (principal components analysis) to find the most efficient, i.e. sparse, representation of $\mathbf{h}_{ep}(n)$. That is, letting the L-length column vector, $\mathbf{w}_k$, be the kth principal component of the random process, we can express $\mathbf{h}_{ep}(n)$ as

$$\mathbf{h}_{ep}(n) = \sum_{k=0}^{L-1} \mathbf{w}_k b_{k,ep}(n) = \mathbf{W}\mathbf{b}_{ep}(n) \tag{14}$$

where $\mathbf{W}$ is an L-by-L unitary matrix whose k^{th} column is $\mathbf{w}_k$. According to PCA, if we define the solution process covariance matrix as

$$\mathbf{R}_{hh} = E\left\{\mathbf{h}_{ep}(n)\mathbf{h}_{ep}^T(n)\right\} \tag{15}$$

and its diagonal decomposition as

$$\mathbf{R}_{hh} = \mathbf{W}\mathbf{D}\mathbf{W}^T \tag{16}$$

where $\mathbf{D}$ is diagonal and $\mathbf{W}$ is unitary, then the columns of $\mathbf{W}$ represent the principal component vectors and the corresponding diagonal elements of $\mathbf{D}$ are proportional to the probable contribution of that component. Most importantly, $\mathbf{b}_{ep}(n)$ tends to be sparser than $\mathbf{h}_{ep}(n)$. Therefore, a good choice for $\mathbf{U}$ in (5) is $\mathbf{W}^T$.

So to find $\mathbf{U}$ for the acoustic echo cancellation problem, we can build a sample covariance matrix from many observations of $\mathbf{h}_{ep}(n)$,

$$\hat{\mathbf{R}}_{hh} = \sum_{k=0}^{P} \mathbf{h}_{ep}(k)\mathbf{h}_{ep}^T(k) \tag{17}$$

and define

$$\mathbf{U} = \hat{\mathbf{W}}^T \tag{18}$$

where $\hat{\mathbf{W}}$ is the unitary matrix in the diagonal decomposition of $\hat{\mathbf{R}}_{hh}$.

5. MPCP

Finally, we turn to updating the coefficient vector under multiple sets of basis vectors. As indicated above, an example application where multiple basis sets are useful is the acoustic echo problem where each basis set corresponds to a different ambient temperature. Let us consider the case where we wish to adapt the coefficients at K distinct temperatures. At each temperature, k, we will develop off-line, and in the manner of section 4, an optimal rotation matrix, $\mathbf{U}_k$. The idea is to update the coefficient vector under each of these rotations. Under one of the rotations, the coefficient vector is sparse and PNLMS will yield fast convergence. Under the other rotations, the coefficient vector is not sparse, and convergence will be slower. Overall, the speed of convergence will be dominated by the sparse rotation.

As with PNLMS++ [2] we may implement the different coefficient updates by 1) alternating between them in different sample periods, or 2) updating the vector using all the rotations in a singe sample period, or a combination of 1) and 2). Here, we describe the method of alternating between updates in different sample periods.

First let us define the coefficient vector under the kth rotation matrix, $\mathbf{U}_k$, as

$$\mathbf{b}_k(n) = \mathbf{U}_k\mathbf{h}(n) . \tag{19}$$

With this definition it is easy to see that we can express $\mathbf{b}_k(n)$ in terms of $\mathbf{b}_{k-1}(n)$ by applying

$$\mathbf{b}_k(n) = \mathbf{T}_{k,k-1}\mathbf{b}_{k-1}(n) \tag{20}$$

where

$$\mathbf{T}_{k,k-1} = \mathbf{U}_k\mathbf{U}_{k-1}^T . \tag{21}$$

The $\mathbf{T}_{k,k-1}$ matrices can be computed off-line when the $\mathbf{U}_k$ s are calculated. MPCP is summarized in Table 3.

Table 3: MPCP

Step	Computation	Complexity
a	$k = (n)_{\bmod K}, \quad m = (n-1)_{\bmod K}$	$\sim$
b	$\mathbf{s}_k(n) = \mathbf{U}_k\mathbf{x}(n)$	L^2
c	$\mathbf{b}_k(n-1) = \mathbf{T}_{k,m}\mathbf{b}_m(n-1)$	L^2
d	$e(n) = y(n) - \mathbf{s}_k^T(n)\mathbf{b}_k(n-1)$	L
e	$\mathbf{M}(n) = f(\rho, \delta_p, \mathbf{b}_k(n-1))$	L
f	$\mathbf{p} = \mathbf{M}(n)\mathbf{s}_k(n)$	L
g	$\mathbf{b}_k(n) = \mathbf{b}_k(n-1) + \mu\mathbf{p}\left[\mathbf{s}_k^T(n)\mathbf{p} + \delta\right]^{-1} e(n)$	$2L$

6. SIMULATIONS

We simulated the process of an object moving about a room with the image-derived model for finding acoustic impulse responses [8]. The microphone and speaker positions were fixed (as is often the case in normal speaker-phone use) and a perfectly absorbing spherical object was located at random positions about the room. The room was perfectly rectangular with different reflection coefficients on each wall, ceiling, and floor. With each new location, the impulse response was measured and added into the sample covariance matrix

Fig. 1 shows a typical impulse response from the speakerphone simulation. The optimal transform of the vector is shown in Fig. 2. Clearly, the transformed version is sparser. The room simulation parameters were as follows:

- the room dimensions were 6.4' by 8' by 6.4',
- the reflection coefficients of the 6 walls were 0.91, 0.87, 0.95, 0.85, 0.8, 0.6,
- the radius of the spherical absorbing object was 1.2',
- the source (loudspeaker) was located at coordinates, (0.64', 3.2', 3.2'), where the origin, (0,0,0) was at the front lower left corner (the positive directions were back, up, and right),
- the microphone was located at coordinates (0.64',4.8',3.2').

The sampling rate in most simulations in this paper was 8 kHz. The only exception was the internal sampling rate of the room simulation, which was 80 kHz, but the final impulse response was sub-sampled to 8 kHz.
Fig. 3 we use the sparseness measure,

$$S_{21}(\mathbf{x}) = \frac{\|\mathbf{x}\|_2}{\|\mathbf{x}\|_1} \qquad (22)$$

to compare the sparseness of 100 room impulse responses before and after having been transformed by the room's principal components. Note: $S_{21}(\mathbf{x})$ ranges from $\frac{1}{\sqrt{L}}$ to one for maximally dispersive and maximally sparse vectors, respectively. The 100 different responses were generated by moving the absorbing object randomly around in the room. The transformed impulse responses were consistently sparser than the untransformed ones.

Fig. 4 shows the convergence of PCP compared to NLMS and FRLS [9]. The algorithms in these simulations use the following parameters: $L = 512$, $\mu = 0.2$, $\rho = .01$, $\delta_p = .01$. The effective length of the forgetting factor, λ, in the FRLS algorithm is $2L$. The initial convergence of PCP is much faster than NLMS and FRLS even though the untransformed impulse response is dispersive. At 2.5 seconds, the echo path is changed. Fig. 4 shows that the

re-convergence of PCP is again faster than NLMS or FRLS.

Figure 5 shows the convergence of MPCP compared to PCP and NLMS. MPCP operates with two rotation matrices, one trained at 20°C and another at 21°C. PCP operates only at 20°C. At 2.5 seconds the temperature in the room is abruptly changed from 20°C to 21°C. The initial convergence of PCP and MPCP is the same, both faster than NLMS, indicating that there is little penalty in alternating the types of updates. When the temperature changes, however, PCP converges slower than NLMS, while MPCP maintains fast convergence.

7. COMPUTATIONAL ISSUES

In the simulations above, we first calculated the solution covariance matrix off-line and then used an eigen-decomposition to find the transforms to make the solutions sparse. The calculation of the covariance matrix can be done in real time by periodically adding outer products of good echo path estimates to the current solution covariance matrix. Finding the eigenvectors of the solution covariance matrix is an $0\{L^3\}$ problem, but the eigenvectors do not need to be calculated very often. If the locations of the microphone and loudspeaker do not change, perhaps the calculation only needs to be done once per day or so.

8. CONCLUSIONS

The simulations presented in this paper provide evidence that MPCP can be used effectively in the problem of acoustic echo cancellation in environments with some temperature variation. Though the computational complexity is high compared to NLMS and FRLS, there is a distinct advantage in the speed of convergence over both of these established algorithms.

REFERENCES

[1] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancellers," *IEEE Trans. Speech Audio Processing*, vol. 8, pp. 508-518, Sept. 2000.

[2] S. L. Gay, "An efficient, fast converging adaptive filter for network echo cancellation," *Proc. of Asilomar*, Nov. 1998.

[3] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Orlando, Florida, May, 2002.

[4] J. Kivinen and M. Warmuth, "Exponentiated gradient versus gradient descent for linear predictors," Inform. Comput., Vol. 132, pp. 1-64, Jan. 1997.

[5] J. Benesty and Y. Huang, "The LMS, PNLMS, and exponentiated gradient algorithms," in *Proc. EUSIPCO*, 2004.

[6] G. Strang, "Linear Algebra and its Applications (second edition)," pp. 225, Academic Press, New York, NY, 1980.

[7] S. L. Grant, "A solution space principal components based adaptive filter," IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Montreal, Canada, May, 2004.

[8] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, no. 4, pp. 943-950, Apr. 1979.

[9] J. Ciofii and T. Kailath, "Fast, recursive-least-squares transversal filters of adaptive filtering," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 304-337, Apr. 1984.
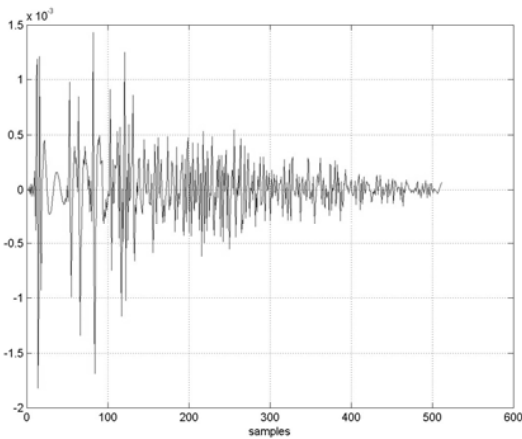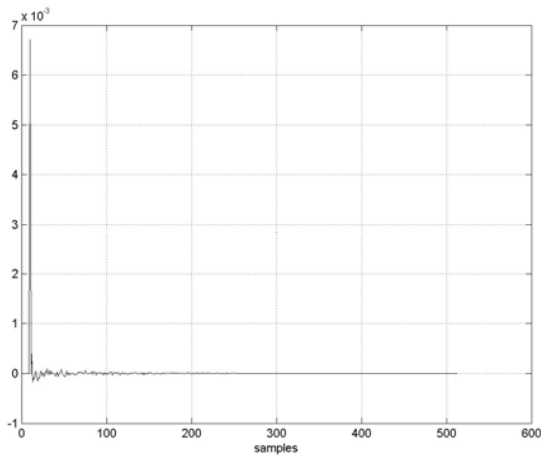
Figure 2: Transform of figure 4
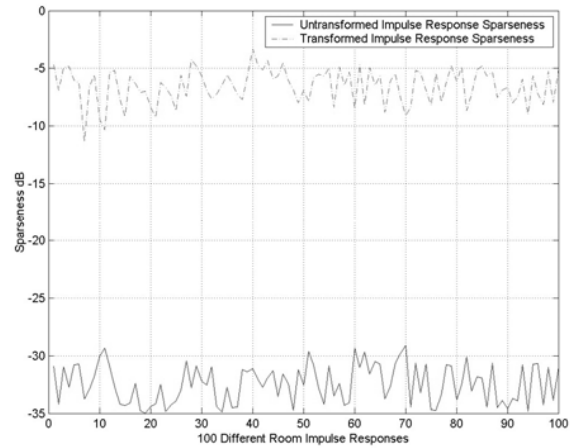
Figure 3: Comparison of sparseness of 100 untransformed and transformed room impulse responses

Figure 1: Impulse response from image derived acoustic model
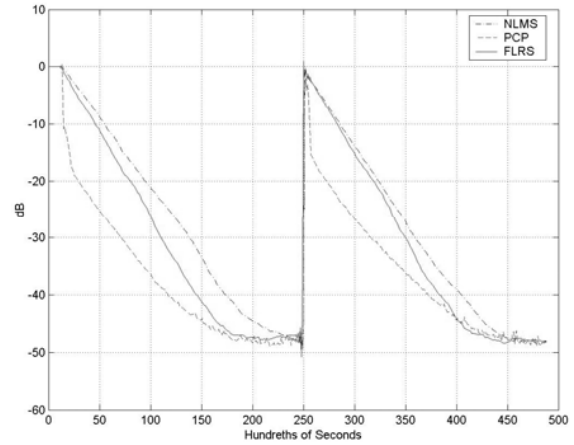
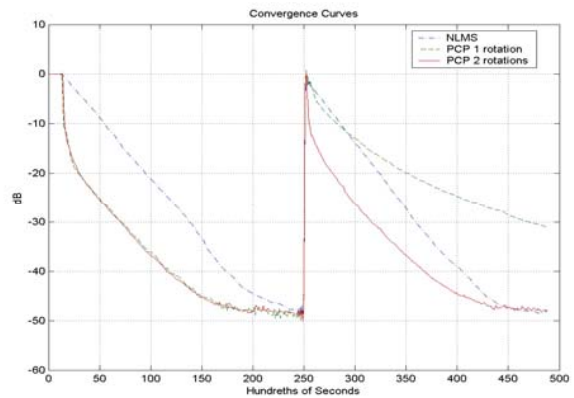

Figure 4: Convergence curves of NLMS, FRLS and PCP

Figure 5: Convergence curves for NLMS, PCP (PCP 1 rotation), and MPCP (PCP 2 rotations)