A SOLUTION SPACE PRINCIPLE COMPONENT BASED ADAPTIVE FILTER

Steven L. Grant, aka Steven L. Gay*

ABSTRACT

This paper introduces a new type of adaptive filter called PCP, principle component proportionate normalized least mean squares. It is an extension of PNLMS (proportionate normalized least mean squares), an adaptive filter that has been shown to provide exceptionally fast convergence and tracking when the underlying system parameters are sparse (as in network echo cancellation). PCP extends the application of PNLMS to certain non-sparse systems by applying it while using the principle components of the underlying system as basis vectors. Room acoustic echo systems are possible examples of such non-sparse systems. 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.
.

1. INTRODUCTION

This paper discusses a new class of adaptive filters that exploits known statistics of the solution. The idea is an extension of the proportionate normalized least mean squares (PNLMS) algorithm [1, 2, 3]. PNLMS has been shown to convergence and track much faster than the classical normalized least mean squares (NLMS) adaptive filter when the solution is sparse in non-zero terms. The sparseness of a vector is a function of the chosen basis vectors. It is well known that a given set of random vectors may be expressed most succinctly as a linear combination of the principle components [4] of the generating random process, i.e., the eigenvectors of the random vectors' covariance matrix. By choosing the principle 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. We call this new algorithm principle component PNLMS, or PCP.

We demonstrate the efficacy of PCP by applying it to the acoustic echo cancellation (AEC) problem. The solution of the AEC problem is the acoustic impulse response of a room as measured between the loudspeaker and microphone. This impulse response is generally not sparse. We simulate samples of room impulse responses using the so-called image derived model [5] where the positions of the microphone and the loudspeaker are fixed and a perfectly absorbing sphere is located in random positions about the room. This represents a first order model of a person using a speakerphone.

When this acoustic model is used, the room impulse responses are much more sparse when the eigenvectors of the room impulse response covariance matrix are used as the basis vectors. In addition, PCP converges much faster than NLMS and FRLS [6] (fast recursive least squares) when PCP operates under the new basis set. However, we must be somewhat cautious of any results when the image-derived model is used. It only takes into account specular reflections and completely ignores diffraction. The success of the experiments described in this paper does not necessarily prove that the method under test will work in a real room, but its failure would strongly indicate that it would not.

2. PNLMS

First, we briefly review the PNLMS adaptive filter under the guise of the echo cancellation problem. The signals, vectors, and matrices used in this paper are as follows: x_n is the far-end signal which excites the echo path, $\mathbf{x}_n = [x_n, \cdots, x_{n-1}]^T$ is the excitation vector, the true echo path impulse response vector is $\mathbf{h}_{ep} = [h_0, ..., h_{L-1}]^T$, v_n is the near-end signal, or near-end noise, $y_n = \mathbf{x}_n^T \mathbf{h}_{ep} + v_n$ is the combination of the echo and the near-end signals, $\mathbf{h}_n = [h_{0,n}, ..., h_{L-1,n}]^T$ is the adaptive filter coefficient vector, $e_n = y_n - \mathbf{x}_n^T \mathbf{h}_{n-1}$ is the error or residual-echo signal, $\mathbf{G}_n = diag\{g_{0,n}, ..., g_{L-1,n}\}$ 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

$$e_n = y_n - \mathbf{x}_n^t \mathbf{h}_{n-1}, \tag{1}$$

the error calculation, and

$$\mathbf{h}_n = \mathbf{h}_{n-1} + \mu \mathbf{x}_n \left(\mathbf{x}_n^t \mathbf{x}_n + \delta \right)^{-1} e_n, \tag{2}$$

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

* MIT, Lincoln Laboratory, 244 Wood St., Lexington, MA, email: slg@ll.mit.edu The author was previously with the Acoustics Research department of Bell Laboratories, Lucent, where this work was performed.

proportionate to the magnitude of the coefficient vector, $\mathbf{h}_{n-1}$, is used to window the excitation vector, $\mathbf{x}_n$,

$$\mathbf{h}_n = \mathbf{h}_{n-1} + \mu \mathbf{G}_n \mathbf{x}_n \left(\mathbf{x}_n^t \mathbf{G}_n \mathbf{x}_n + \delta\right)^{-1} e_n \tag{3}$$

where,

$$\mathbf{G}_n = f(\rho, \delta_p, \mathbf{h}_{n-1}). \tag{4}$$

and $f(\rho, \delta_p, \mathbf{h}_{n-1})$, the series of steps that generate $\mathbf{G}_n$. See Table 1.

Table 1: PNLMS $\mathbf{G}_n$ calculation

$L_{\max} = \max\left\{\delta_p, \lvert h_{0,n-1}\rvert, \ldots, \lvert h_{L-1,n-1}\rvert\right\}$
$L_1 = \sum_{i=0}^{L-1} \gamma_i$
$(\mathbf{G}_n)_{i,i} = \gamma_i / L_1 \qquad 0 \le i \le L-1$

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.

This property of PNLMS's convergence speed being dependent on the sparseness of the solution, $\mathbf{h}_{ep}$, determines the proper application of PNLMS. Impulse responses of echo paths in telephone networks tend to be sparse while those of acoustic echo paths in offices and conference rooms tend to be dispersive. As a result, PNLMS has been successfully used in network echo cancellers, but not in acoustic echo cancellers.

3. PNLMS WITH ARBITRARY BASIS SET

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

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

$$\mathbf{b}_{ep} = \mathbf{U}\mathbf{h}_{ep}. \tag{5}$$

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

$$\mathbf{h}_{n-1} = \mathbf{U}^T \mathbf{b}_{n-1} \tag{6}$$

and

$$\mathbf{s}_n = \mathbf{U}^T \mathbf{x}_n. \tag{7}$$

Then, (1) can be rewritten as

$$e_n = y_n - \mathbf{s}_n^T \mathbf{b}_{n-1} \tag{8}$$

and (2) as

$$\mathbf{U}^T \mathbf{b}_n = \mathbf{U}^T \mathbf{b}_{n-1} + \mu \mathbf{U}^T \mathbf{s}_n \left(\mathbf{s}_n^t \mathbf{U}\mathbf{U}^T \mathbf{s}_n + \delta\right)^{-1} e_n. \tag{9}$$

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

$$\mathbf{b}_n = \mathbf{b}_{n-1} + \mu \mathbf{s}_n \left(\mathbf{s}_n^t \mathbf{s}_n + \delta\right)^{-1} e_n. \tag{10}$$

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

$$\mathbf{M}_n = f(\rho, \delta_p, \mathbf{b}_{n-1}). \tag{11}$$

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

$$\mathbf{b}_n = \mathbf{b}_{n-1} + \mu \mathbf{M}_n \mathbf{s}_n \left(\mathbf{s}_n^t \mathbf{M}_n \mathbf{s}_n + \delta\right)^{-1} e_n. \tag{12}$$

To summarize, the Transformed PNLMS algorithm is shown in Table 2.

Table 2: Transformed PNLMS

$\mathbf{s}_n = \mathbf{U}^T \mathbf{x}_n$.	(a)
$e_n = y_n - \mathbf{s}_n^T \mathbf{b}_{n-1}$	(b)
$\mathbf{M}_n = f(\rho, \delta_p, \mathbf{b}_{n-1})$	(c)
$\mathbf{b}_n = \mathbf{b}_{n-1} + \mu \mathbf{M}_n \mathbf{s}_n \left(\mathbf{s}_n^t \mathbf{M}_n \mathbf{s}_n + \delta\right)^{-1} e_n$	(d)

4. FINDING PROPER BASIS VECTORS

Now that we know how to perform PNLMS under an alternate set of basis vectors, we turn to the problem of defining what those basis vectors should be. To this point we have viewed the solution vector, $\mathbf{h}_{ep}$, as fixed. In practical applications though, it is almost always time varying. Otherwise, an adaptive filter would be unnecessary. The manner in which $\mathbf{h}_{ep}$ varies with time is, of course, application dependent. In network echo cancellation $\mathbf{h}_{ep}$ changes very little during the period of a given telephone connection, but changes abruptly when a new connection is made. In acoustic echo cancellation $\mathbf{h}_{ep}$ changes throughout the communication session as objects in the acoustic echo path (such as the conversation participants) move about the room.

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 (principle 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 k^{th} principle component of the random process, we can express $\mathbf{h}_{ep}(n)$ as

$$\mathbf{h}_{ep}(n) = \sum_{k=0}^{L-1} \mathbf{w}_k b_{i,ep}(n) = \mathbf{W}\mathbf{b}_{ep}(n), \tag{13}$$

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

$$\mathbf{R}_{hh} = E\left\{\mathbf{h}_{ep}(n)\mathbf{h}_{ep}^T(n)\right\} \tag{14}$$

and its diagonal decomposition as

$$\mathbf{R}_{hh} = \mathbf{W}\mathbf{D}\mathbf{W}^T, \tag{15}$$

where $\mathbf{D}$ is diagonal and $\mathbf{W}$ is unitary, then the columns of $\mathbf{W}$ represent the principle 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 $\mathbf{h}_{ep}(n)$,

$$\hat{\mathbf{R}}_{hh} = \sum_{k=0}^{P}\mathbf{h}_{ep}(k)\mathbf{h}_{ep}^T(k) \tag{16}$$

and define

$$\mathbf{U} = \hat{\mathbf{W}}^T \tag{17}$$

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

5. SIMULATIONS

Collecting hundreds to thousands of room impulse responses is a somewhat daunting task. Instead, as a first order approximation to the real thing, we simulated the process of an object moving about a room with the image-derived model for finding acoustic impulse responses. 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 transform of the vector is shown in Fig. 2. Clearly, the transformed version is sparser. The salient parameters in the room simulation are as follows:

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

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

Fig. 3 we use the sparsness measure,

$$S_{21}(\mathbf{x}) = \frac{\|\mathbf{x}\|_2}{\|\mathbf{x}\|_1} \tag{18}$$

to compare the sparseness of 100 room impulse responses before and after having been transformed by the room’s principle components. Note: $S_{21}(\mathbf{x})$ ranges from $1/\sqrt{L}$ 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 are consistently sparser than the untransformed ones.

Fig. 4 shows the convergence of PCP compared to NLMS and FRLS. 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 $2L$. That is, $\lambda = 2L/2L + 1$.

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.

5. COMPUTATIONAL ISSUES

In the simulations above, we first calculated the solution covariance matrix off-line and then used an eigen-decomposition to find the transform 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. Using a forgetting factor on the old solution covariance matrix will help the adaptive filter to progressively forget old echo paths that may no longer represent the room statistics should they change with time, e.g., if the loudspeaker and microphone are moved. Of course, there is still the problem of finding the eigenvectors of the solution covariance matrix. This 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.

The vector/matrix multiply in step (a) of table 2 is $0\{L^2\}$. At the price of a block delay of L, this complexity can be reduced to $0\{L\log(L)\}$ by using FFT techniques. Additional computational savings can be realized by not using every eigenvector in $\mathbf{U}^T$.

6. CONCLUSIONS

The simulations presented in this paper provide evidence that PNLMS can be used effectively in the problem of acoustic echo cancellation under a change of basis vectors. Though the computation al complexity is high compared to NLMS and even FRLS, there is a distinct advantage in the speed of convergence over both of these established algorithms.

We must be somewhat cautious with these results however, since the image derived model is only a first order approximation of the behavior of sound in a real room. However, if the simulations had shown that the new method failed, further work would clearly not be warranted. However, since the simulations showed PNLMS to be very effective using PCA derived basis vectors, we propose further experiments using data collected from real rooms.

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

[5] 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.

[6] 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.

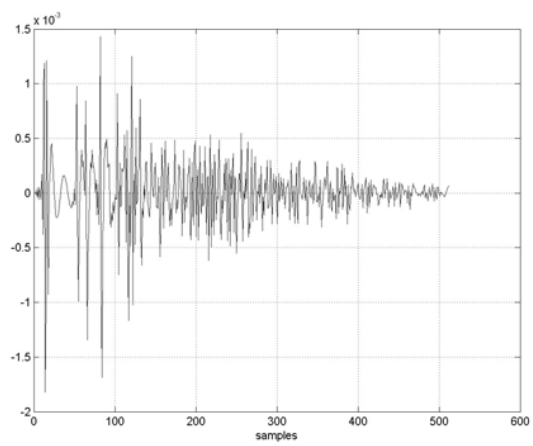

Figure 1: Impulse response from image derived acoustic model

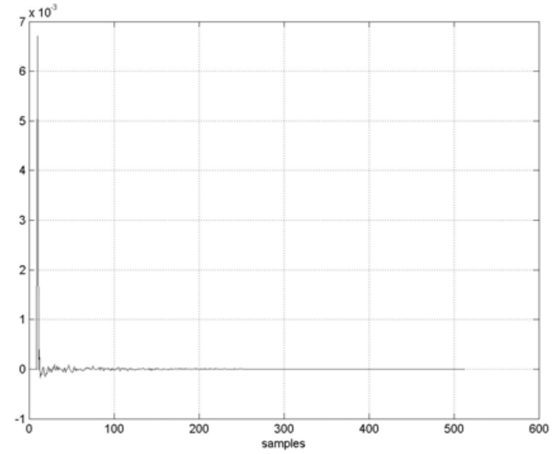

Figure 2: Transform of figure 4

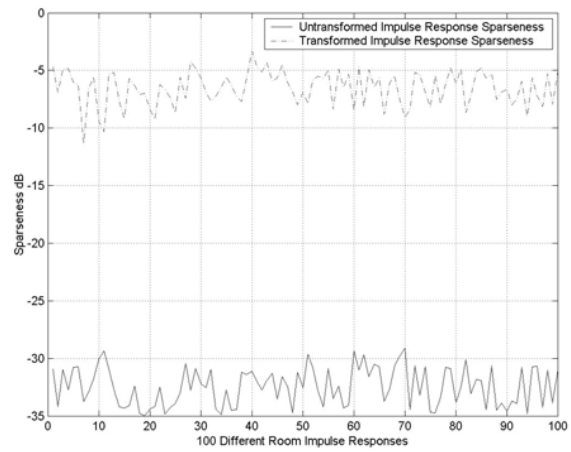

Figure 3: Comparison of sparseness of 100 untransformed and transformed room impulse responses

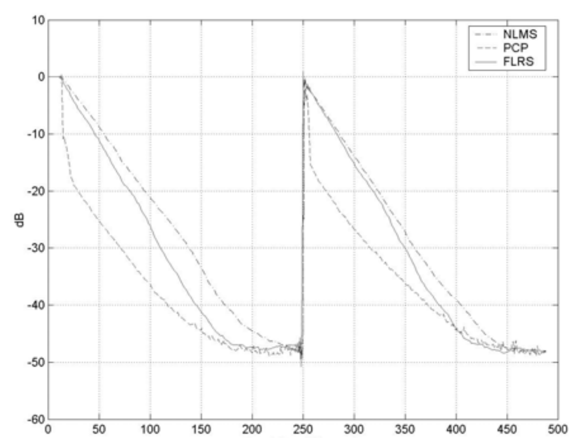

Figure 4: Convergence curves of NLMS, FRLS and PCP