BLIND ADAPTIVE ESTIMATION OF KLT BASIS VECTORS

Carlos E. Davila

Electrical Engineering Department, Southern Methodist University
Dallas, Texas, 75275-0338, USA, cd@seas.smu.edu

ABSTRACT

An algorithm for estimating the basis vectors used in the Karhunen-Loeve Transform (KLT) is described. The algorithm is "blind" in the sense that it utilizes minimal information about the data vector being encoded. It is capable of estimating the KLT basis vectors using only the KLT coefficients and one additional scalar quantity. This eliminates the need to repeatedly encode and transmit the KLT basis vectors.

1. INTRODUCTION

The Karhunen-Loeve Transform (KLT) is known to be the optimum transform for signal compression [1, 2]. Unfortunately, the KLT basis functions, which are the eigenvectors of the data correlation matrix, are data dependent. Hence the basis functions must also be encoded and transmitted which reduces compression and leads to increased data rates. For this reason, the KLT has found limited use in data compression applications. Let

$$x_n = \left[\begin{array}{cccc} x(nN-1) & x(nN-2) & \cdots & x(N(n-1)) \end{array} \right]^T \tag{1}$$

be the N-dimensional signal frame to be encoded. We assume that x_n has correlation matrix $R = E\left[x_n x_n^T\right]$ having rank $r \leq N$. This means that x_n can be represented as a linear combination of the eigenvectors of R given by $q^1, q^2, \ldots, q^r$, corresponding to eigenvalues $\lambda^1 \geq \lambda^2 \geq \ldots \geq \lambda^r > 0$, respectively. Let $Q = \left[\begin{array}{cccc} q^1 & q^2 & \cdots & q^r \end{array} \right]$ be an $N \times r$ matrix whose columns are the KLT basis vectors (eigenvectors of R). The transform coefficients, given by $\alpha_n = Q^T x_n$, can then be quantized as $\hat{\alpha}$, encoded, and transmitted. If the receiver has knowledge of the basis vectors Q, x_n can be recovered as $\hat{x}_n = Q\hat{\alpha}_n$. If the signal $x(n)$ is statistically stationary then the eigenvectors need only be estimated and transmitted once, which would not lead to much loss of compression, however in practice, the eigenstructure of most signals tends to vary considerably over time. Hence the eigenvectors of R need to be constantly retransmitted which is why the KLT is not often used. In this paper, we give a method of determining the basis vectors for the KLT directly from the KLT coefficients given only very limited knowledge of $x(n)$. This eliminates the need to retransmit the KLT basis vectors.

2. TRACKING KLT BASIS VECTORS

Blind estimation of the KLT basis vectors can be accomplished using ideas from the subspace tracking literature. Let $\hat{R}_n$ be an estimate of $E\left[x_n x_n^T\right]$ at time n that is updated using,

$$\hat{R}_n = \gamma \hat{R}_{n-1} + x_n x_n^T \tag{2}$$

where $0 < \gamma < 1$. Let

$$\hat{Q}_n = \left[\begin{array}{cccc} \hat{q}_n^1 & \hat{q}_n^2 & \cdots & \hat{q}_n^r \end{array} \right] \tag{3}$$

and $\hat{\Lambda}_n = diag(\hat{\lambda}_n^1, \hat{\lambda}_n^2, \ldots, \hat{\lambda}_n^r)$ be *estimates* of the eigenvectors and eigenvalues, respectively, of $\hat{R}_n$. Then $\hat{R}_n \approx \gamma \hat{Q}_{n-1} \hat{\Lambda}_{n-1} \hat{Q}_{n-1}^T + x_n x_n^T$. The eigenvector estimates can be updated as follows:

1. Solve the generalized eigenvalue problem

$$F W_n = G W_n \Pi_n \tag{4}$$

   where $F = \overline{Q}_n^T \left(\gamma \hat{Q}_{n-1} \hat{\Lambda}_{n-1} \hat{Q}_{n-1}^T + x_n x_n^T \right) \overline{Q}_n$, $G = \overline{Q}_n^T \overline{Q}_n$, and W_n and the diagonal matrix Π_n are the respective generalized eigenvectors and eigenvalues. The matrix $\overline{Q}_n = \left[\begin{array}{cc} \hat{Q}_{n-1} & v_n \end{array} \right]$ has dimension $N \times (r+1)$ and v_n is a *search direction vector.*

2. Update the eigenvector estimates as

$$\hat{Q}_n = \overline{Q}_n W_n(1:r) \tag{5}$$

   where $W_n(1:r)$ are the eigenvectors corresponding to the maximum r eigenvalues in Π_n.

3. The eigenvalue estimates are updated as $\hat{\Lambda}_n = \Pi_n(1:r, 1:r)$.

If, in the above algorithm, the search direction is set to $v_n = x_n$, then the algorithm is a standard *subspace averaging* algorithm used for subspace tracking [3, 4]. Note that if we treat the columns of $\hat{Q}_{n-1}$ as the KLT basis vectors, then the KLT coefficients are contained in the first r elements of $\overline{Q}_n^T x_n$ (see (4)), hence the algorithm never explicitly uses x_n. We will show that if v_n is a white noise vector, independent of x_n, then the eigenvectors of $\hat{R}_n$ can still be tracked. This implies that the above algorithm can be run by both the sender and the receiver concurrently using the same initial conditions. If the search direction vectors are known to both the sender and the receiver, then the receiver can also track the KLT basis vectors having only knowledge of the KLT coefficients and the additional scalar coefficient, $v_n^T x_n$. In this context, the algorithm is "blind" since the receiver requires no explicit knowledge of the signal $x(n)$ to track the KLT basis vectors.

3. ALGORITHM CONVERGENCE

It can be shown that the algorithm attempts to find the subspace in the column space of $\overline{Q}_n$ which is closest in terms of Euclidean distance to the subspace spanned by the r eigenvectors of $\gamma \hat{Q}_{n-1}\hat{\Lambda}_{n-1}\hat{Q}_{n-1}^T + x_n x_n^T$; this subspace is often referred to as the *signal subspace*. The vectors in W_n give the linear combinations of the columns of $\overline{Q}_n$ which are closest to the signal subspace. A suitable measure of the algorithm's convergence is

$$\epsilon(n) \equiv trE\left[P_{\overline{Q}_n}^{\perp}\left(\gamma \hat{Q}_{n-1}\hat{\Lambda}_{n-1}\hat{Q}_{n-1}^T + x_n x_n^T\right)P_{\overline{Q}_n}^{\perp}\right] \tag{6}$$

where $P_{\overline{Q}_n}^{\perp} = I_N - P_{\overline{Q}_n}$, $P_{\overline{Q}_n}$ is a projection matrix onto the column space of $\overline{Q}_n$, and tr is the matrix trace operation. Some straight-forward calculations give

$$\epsilon(n) = \sum_{i=1}^{r} \lambda^i E\|P_{\overline{Q}_n}^{\perp} q^i\|^2, i = 1, \ldots, r \tag{7}$$

Where $q^i, i = 1, \ldots, r$ is the eigenvector of R corresponding to eigenvalue λ^i. We note that $\epsilon(n) = 0$ only when q^i is contained in the column space of $\overline{Q}_n$ for $i = 1, \ldots, r$. Next we note that

$$P_{\overline{Q}_n} q^i = P_{\hat{Q}_{n-1}} q^i + \frac{P_{\hat{Q}_{n-1}}^{\perp} v_n v_n^T P_{\hat{Q}_{n-1}}^{\perp}}{v_n^T P_{\hat{Q}_{n-1}}^{\perp} v_n} q^i \tag{8}$$

Some additional manipulations lead to

$$\epsilon(n) = E\left(1 - \frac{v_n^T A v_n}{v_n^T B v_n}\right)\sum_{i=1}^{r} \lambda^i q^{iT} P_{\hat{Q}_{n-1}}^{\perp} q^i \tag{9}$$

where $A = \frac{P_{\hat{Q}_{n-1}}^{\perp} q^i q^{iT} P_{\hat{Q}_{n-1}}^{\perp}}{q^{iT} P_{\hat{Q}_{n-1}}^{\perp} q^i}$ and $B = P_{\hat{Q}_{n-1}}^{\perp}$. Hence, the reduction in each mode of $\epsilon(n)$ at each time step is determined by,

$$E\left[v_n^T A v_n / v_n^T B v_n\right]$$

which has an upper bound of 1. This expectation can be evaluated using the method developed by Bershad for analyzing the normalized LMS algorithm [5]. If we assume that v_n is a zero mean, unit variance Gaussian white noise vector then

$$E\left[\frac{v_n^T A v_n}{v_n^T B v_n}\right] = \frac{1}{(2\pi)^{N/2}}\int \frac{v_n^T A v_n}{v_n^T B v_n} exp\left\{\frac{-v_n^T v_n}{2}\right\} dv \tag{10}$$

Let

$$g(\beta) \equiv \frac{1}{(2\pi)^{N/2}}\int \frac{v_n^T A v_n}{v_n^T B v_n} exp\left\{-\beta v_n^T B v_n\right\} exp\left\{\frac{-v_n^T v_n}{2}\right\} dv \tag{11}$$

Then

$$\begin{aligned} \frac{dg}{d\beta} &= \frac{-1}{(2\pi)^{N/2}}\int v_n^T A v_n exp\left\{-\beta v_n^T B v_n\right\} exp\left\{\frac{-v_n^T v_n}{2}\right\} dv \qquad (12)\\ &= -|C|^{1/2}\left[\frac{1}{(2\pi)^{N/2}|C|^{1/2}}\int v_n^T A v_n exp\left\{\frac{-v_n^T C^{-1} v_n}{2}\right\} dv\right] \end{aligned}$$

where $C^{-1} = (I + 2\beta B)$. The quantity in the square brackets is the expectation of $v_n^T A v_n$ where v_n is a zero-mean Gaussian vector with covariance matrix C. Since $A = aa^T$ is a constant rank-1 matrix, $E\left[v_n^T A v_n\right] = trAE\left[v_n v_n^T\right] = trAC = a^T C a$ hence

$$\frac{dg}{d\beta} = -|C|^{1/2} a^T C a \tag{13}$$

where $a = P_{\hat{Q}_{n-1}}^{\perp} q^i / \sqrt{q^{iT} P_{\hat{Q}_{n-1}}^{\perp} q^i}, i = 1, \ldots, r$. Therefore

$$\rho(N, r) \equiv E\left[\frac{v_n^T A v_n}{v_n^T B v_n}\right] = \int \frac{dg}{d\beta} d\beta\bigg|_{\beta=0} + D \tag{14}$$

where the integration constant $D = 0$ [5]. The integral in (14) can be evaluated numerically. It can be shown that $\rho(N, r)$ is dependent only on N and r, and is independent of $P_{\hat{Q}_{n-1}}^{\perp}$ and $q^i, i = 1, \ldots, r$. Figure 1 shows $\rho(N, r)$ for a range of values of N, and r; it is seen to increase with decreasing N and increasing r.

4. INCREASING CONVERGENCE SPEED

As seen in Figure 1, the reduction in the eigenvector estimation error $\epsilon(n)$ can be slow for a white noise search direction, particularly for larger values of N. One way of improving convergence speed is with a codebook search. If both the sender and the receiver have

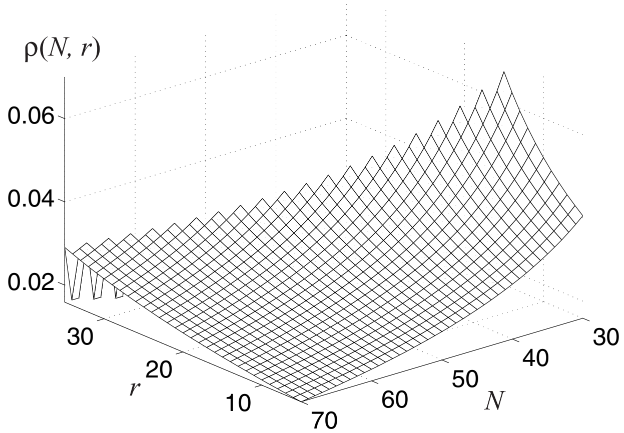

Figure 1: $\rho(N, r)$ is proportional to the reduction in the error measure $\epsilon^i(n)$ per time step.

the same codebook, then since each are running the same algorithm with the same initial conditions, each can search the codebook for the search direction which maximizes $tr\Pi_n(1 : r, 1 : r)$. This eliminates the need to transmit the codebook index for the best search direction.

5. EXPERIMENTS

The following signal was generated

$$x(n) = \begin{cases} cos(0.3\pi n) + cos(0.7\pi n + 0.35\pi), & n = 1, \ldots, 999 \\ cos(0.6\pi n) + cos(0.8\pi n + 0.35\pi), & n = 1000, \ldots, 2,000 \end{cases} \quad (15)$$

and the adaptive KLT (AKLT) algorithm described above was applied to this signal with $\gamma = 0.7$ and $r = 4$. The search direction v_n was set to a zero-mean Gaussian white noise vector. To measure the algorithm's performance the mean squared error between the original and reconstructed data frame was estimated as

$$\epsilon(n) = |x_n - \hat{Q}_n \hat{Q}_n^T x_n|^2 \quad (16)$$

Figure 2 shows $\epsilon(n)$ for different values of N. As predicted, the convergence speed increases with decreasing N. Then experiment was then repeated, but this time, using a 1000-word Gaussian white noise codebook to find the best search direction as described above. The resulting mean squared error plots are shown in Figure 3. The codebook search gives a dramatic improvement in convergence speed. Next we applied the algorithm to speech data sampled at 8 kHz, 16-bits per sample using $N = 20$, and $\gamma = 0.5$. Figure 4(a) shows $\epsilon(n)$ for two runs, one with $r = 5$, and the other with $r = 15$. In both cases a 1000 word Gaussian white noise codebook was used to select the best search direction. The error for $r = 15$ is much lower than for $r = 5$, this reflects the fact that the algorithm converges faster for larger values of r. So if for each frame, nearly all of the KLT coefficients must be transmitted for accurately tracking the basis vectors of speech, the usefulness of this approach as a data compression method may seem questionable. However, for the KLT, most of the energy is contained in the first few transform coefficients. Hence only the (few) larger KLT coefficients need be allocated a significant number of bits, the majority of smaller coefficients can be allocated fewer bits [2]. This raises the possibility of very low-rate coding.

6. SUMMARY

An algorithm for blindly tracking the KLT basis vectors using only the KLT coefficients and one additional scalar coefficient was described. Convergence of the KLT basis vector tracker was proven for a Gaussian white noise search direction. A codebook search method was described which greatly improves the convergence rate of the algorithm. Several experiments designed to demonstrate the feasibility of blind adaptive KLT coding were described. Future work involves the use of codebooks more ideally suited for the signal being encoded, and quantizing the KLT coefficients.

REFERENCES

[1] M. Loéve, *Probability Theory.* Van Norstrand, 1955.

[2] N. S. Jayant and P. Noll, *Digital Coding of Waveforms.* Prentice Hall, 1984.

[3] I. Karasolo, "Estimating the covariance matrix by signal subspace averaging," *IEEE Trans. Acoust. , Speech, Signal Processing*, vol. 34, pp. 8–12, Feb. 1986.

[4] P. Comon and G. Golub, "Tracking a few extreme singular values and vectors in signal processing," *Proc. IEEE*, vol. 78, pp. 1327–1343, Aug. 1990.

[5] N. J. Bershad, "Analysis of the normalized LMS algorithm with Gaussian inputs," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 793–806, 1986.

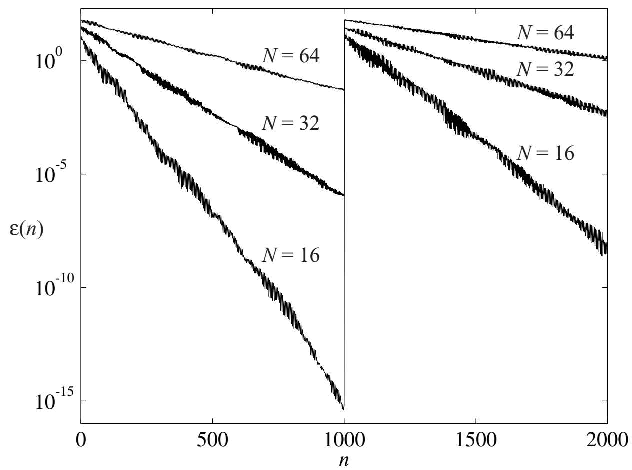

Figure 2: Mean squared error for sinusoid reconstruction for different frame lengths N using a white noise search direction.

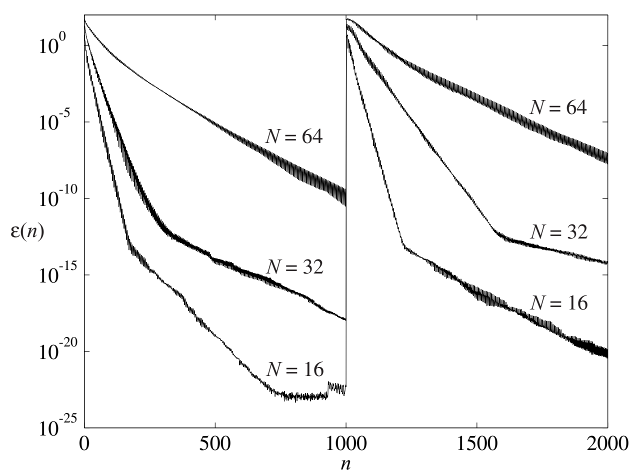

Figure 3: Mean squared error for sinusoid reconstruction using a codebook search to find the optimum search direction. Codebook searching leads to a significant increase in convergence rate.

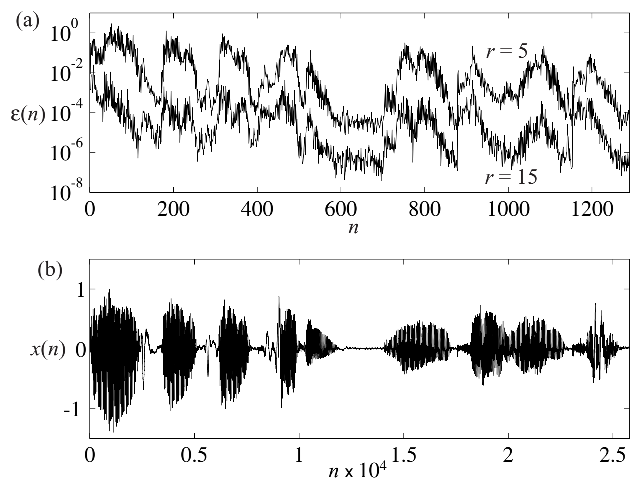

Figure 4: (a) Mean squared error for speech reconstruction experiment for $r = 5$ and $r = 15$ dimensional signal subspace; (b) Speech segment used.