Smoothing Approaches to Reconstruction of Missing Data in Array Processing

Jennifer A. Fulton

Telstra Mobile, Liverpool Street, Sydney NSW 2000, Australia

Robert R. Bitmead

Cooperative Research Centre for Robust & Adaptive Systems, Department of Systems Engineering, RSISE, Australian National University, Canberra ACT 0200, Australia

Robert C. Williamson

Department of Engineering, Australian National University, Canberra ACT 0200, Australia

Abstract

Kalman smoothing techniques are migrated from downsampling in speech coding applications to provide new approaches to array processing with missing elements. The key principles are the use of signal models to develop interpolation methods for the reconstruction of missing data streams. Performance is studied through received array response patterns and is compared to forwards/backwards Linear Predictive methods. This analysis is extended to include the estimation of the models from the data and performance is compared for a variety of signal to noise ratios.

Key words: Kalman smoothing, array processing, data reconstruction, beamforming, linear predictive coding

1 Introduction

Array sensors arise in radar and sonar signal processing for the determination of number and location of signal sources. The geometry we consider is a linear,

[1] The authors acknowledge the funding of the activities of the Cooperative Research Centre for Robust & Adaptive Systems by the Australian Commonwealth Government under the Cooperative Research Centres Program.

equispaced array of elements with beamforming performed by delay-and-sum techniques to steer the array in particular look directions at particular frequencies [9,5,6]. With the failure of one or more of these sensors, the signal gain of the array over the background noise is affected because of the absence of the elements' signals and because of the inappropriate use of the available sensors' data. Since these delay-and-sum beamformers rely on the constructive interference of sensor signals to achieve the look direction, modified processing is called for with absent sensors. These modifications can be computed by explicit optimisation, which is time-consuming and direction dependent. Our aim here is to provide a systematic approach to the development of this modified computation using the received data based on Kalman smoothing ideas from speech compression.

Kalman smoothing is an outgrowth of Kalman filtering in which data up to and subsequent to the time being estimated is used to improve the estimate. The methodology of smoothing evolves from time series analysis, where prediction and filtering may be extended to a variety of smoothing paradigms [1]. Here we develop fixed-lag spatial smoothers to accommodate the failure of known sensors in antenna arrays. Our approach is through the reconstruction of the absent time domain signals at the failed sensor's location using signals from neighbouring functioning sensors on both side of the fault. This approach permits the direct design of techniques to accommodate multiple failures.

Alternative reconstruction approaches have been proposed based on linear predictive coding (LPC) methods [8] or Fourier interpolation of excitations (FIX) [2]. Performance of these latter methods is comparable and will be compared with *optimal* Kalman smoothing techniques, which also use the available data to reconstruct missing signals. Still other data-independent techniques can be used which compute the suitably modified weights for steering the incomplete array. These latter methods are based on linear programming and can be computationally demanding.

The outcome of our comparison between LPC and the Kalman smoother is that, with the exact knowledge of the signal data generating model, the Kalman filter outperforms that LPC at all SNRs. However, when the received data is also used to generate the signal model from correlations, the LPC approach proves better at very low SNR. This is due to the reliance of the Kalman smoother's performance on an accurate knowledge of the signal's spatial correlation structure. Our approach to the development of smoothing solutions is by comparison with methods of modelling and reconstruction from speech coding.

2 Reconstruction Techniques for Autoregressive Models

2.1 Reconstruction in Speech Coding

Speech coding is a digital signal compression technique in which the total bit-rate is minimised subject to perceptual performance requirements of the encoder-decoder pair. For signals of telephony quality, two main approaches are evident; Adaptive Differential Pulse-Coded Modulation (ADPCM) and Codeword Excited Linear Prediction (CELP) methods. One of the central distinctions between these approaches is that ADPCM operates on data synchronous with the speech signal sampling while CELP performs block analysis.

Recently Ramabudran and Sinha [7] proposed an approach to speech coding via ADPCM in which bit-rate reduction is achieved by omitting every second sample from the transmission but increasing the accuracy (i.e. the number of bits per sample) for each sample transmitted. This creates an algorithm which is halfway between ADPCM and CELP. The central aspect of this method of bit-rate reduction lies in the reconstruction of the missing samples. Ramabudran and Sinha do this via predictive methods using signal models. This has been extended by the authors through the introduction of the Kalman smoother for the reconstruction from the received data [3].

The use of signal models is critical for the application of these smoothing methods and to alleviate the apparent violation of Nyquist's sampling theorem. Interestingly, both ADPCM and CELP coders compute signal models routinely as part of their operation. Furthermore, in most implementations, these models are computed from reconstructed data alone in order not to require the separate side transmission of model parameters. This is known as *backwards adaptation.*

The usual structure of models in speech coding is autoregressive. That is, the speech signal $\{S_k\}$ is modelled as the output of an autoregressive (AR) filter,

$$S_k = -a_1 S_{k-1} - a_2 S_{k-2} - \cdots - a_n S_{k-n} + e_k, \tag{1}$$

where $\{e_k\}$ is a white noise sequence and the parameters $\{a_1, a_2, \cdots, a_n\}$ are the AR coefficients, with the polynomial

$$A(z) = 1 + a_1 z + a_2 z^2 + \cdots + a_n z^n$$

having roots strictly inside the unit circle in the complex z-plane.

Such AR speech models are chosen because they may be easily and efficiently computed from covariance data using the Levinson-Durbin algorithm, see [4].

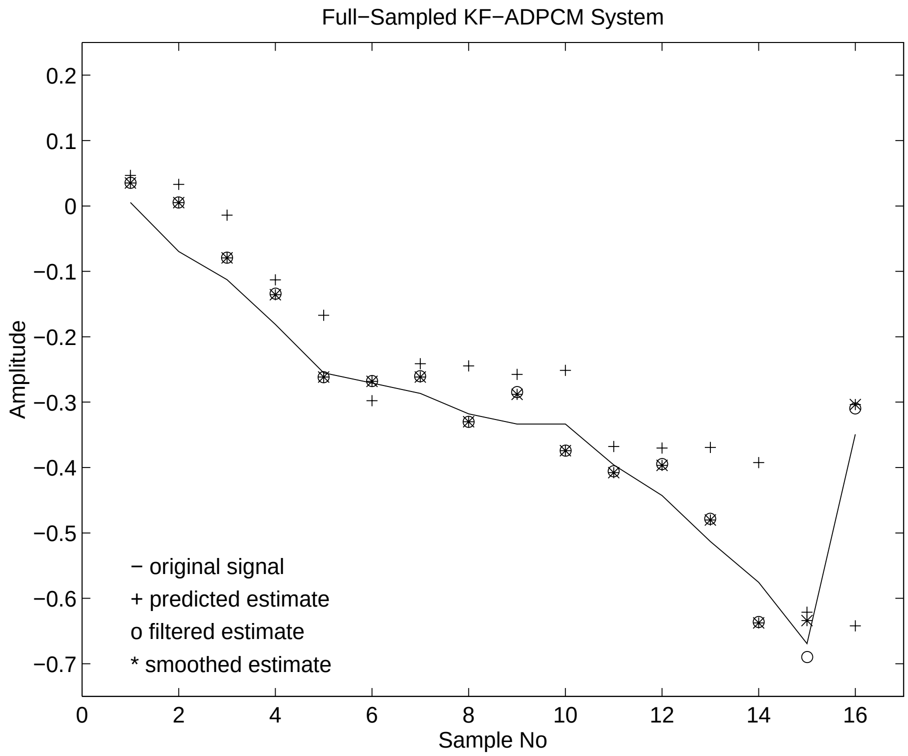

Fig. 1. Decimated speech reconstruction using prediction, filtering and smoothing.

In this algorithm, models of successive orders are computed recursively from the first (defining) row of the covariance matrix. These models are guaranteed to be stable and the order of computation is $O(n^2)$. These factors are central to their applicability in speech coding where models are refined roughly every 20 milliseconds. We shall utilise these properties and the simplicity of smoother computation for these models when used in array beamforming.

Figure 2.1 shows a close-up of the reconstruction of decimated real speech data using Kalman prediction (+), Kalman filtering (o), and 3-step fixed-lag Kalman smoothing (*). Several features are noteworthy here.

- Predicted signal values are not as good a fit to the original as filtered signals.
- At every second time sample, filtering has no advantage over prediction. This is due to the absence of further information because of the decimation.
- Filtered signals are not as good a fit as smoothed signals. In particular, 3-step smoothed signals are able to use the three subsequent data to refine the reconstruction for all times.

The work in [3] goes on to consider and compare the behaviour of smoothers with decimated accurate encoded data versus full-rate encoded data with half as many bits per symbol. Here we shall consider the application of these autoregressive signal reconstruction techniques in the array problem.

From (1), which defines an autoregressive model, the standard signal predictor is immediate. Denoting by $\hat{S}_{k+1|k}$ the one-step-ahead prediction of S_{k+1} from signal data up to and including time k, we have an immediate recursion for a predictor,

$$\hat{S}_{k+1|k} = -a_1\hat{S}_{k|k-1} - a_2\hat{S}_{k-1|k-2} - \cdots - a_n\hat{S}_{k-n|k-n-1} + \varepsilon_k, \tag{2}$$

$$\varepsilon_k = S_k - \hat{S}_{k|k-1}. \tag{3}$$

This is known as the Linear Predictive Coder (LPC) and defines a recursive formula for the one-step-ahead predictions. Its simplicity is its immediate attraction, both in terms of its numerical complexity and of its transparent dependence on the autoregressive model. It is actually the optimal predictor, or Kalman predictor, for this problem in which the speech signal S_k is perfectly measured. Indeed, the optimal filter and smoother are also simple: $\hat{S}_{k|k} = S_k$ and $\hat{S}_{k|k+m} = S_k$.

If, however, the measured signal is the speech corrupted by white noise, a different formulation is appropriate. Rewriting (1) as a state-space equation one has the following variant of the signal model.

$$x_k = \begin{pmatrix} S_k \\ S_{k-1} \\ S_{k-2} \\ \vdots \\ S_{k-n+1} \end{pmatrix} \tag{4}$$

$$x_{k+1} = \begin{pmatrix} -a_1 & -a_2 & -a_3 & \cdots & a_n \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} x_k + \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} e_k \tag{5}$$

$$y_k = S_k + v_k = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \end{pmatrix} x_k + v_k. \tag{6}$$

Here the state x_k embodies the collection of past signal values used in the autoregression and the aim is to use measured data $\{y_k\}$ to reconstruct these signals.

Kalman filtering is associated with the best linear unbiased estimator of the state x_k (and hence of S_k) from the signal measurements y_k. Kalman prediction constructs the entire one-step-ahead state prediction vector $\hat{x}_{k+1|k}$ from data up to time k. This is given by

$$\hat{x}_{k+1|k} = (F - K_k H)\hat{x}_{k|k-1} + K_k y_k, \tag{7}$$

where F is the square system matrix in (5), H is the output matrix of (6) and K_k is the Kalman gain computed from the solution to the Riccati Difference Equation,

$$K_k = F\Sigma_k H^T(H\Sigma_k H^T + R_k)^{-1} \tag{8}$$
$$\Sigma_k = F\Sigma_k F^T - F\Sigma_k H^T(H\Sigma_k H^T + R_k)^{-1}H\Sigma_k F^T + Q, \tag{9}$$

where Q is the variance of w_k and R_k is the time-varying variance of v_k.

In normal, time-invariant applications, the Kalman gain K_k may be taken as fixed and the covariance matrix Σ_k is also fixed. However, for the Ramabudran and Sinha signal decimation approach, the reconstruction is required for all signal samples even though data are provided only at every second instance. This is accommodated in the Kalman filter (or smoother if we use AR models) by taking measurement error covariance R_k as fixed R for times when data is provided (say at even time instants) and as infinity when the data sample is absent (at odd time instants). The result is that the Kalman gain K_k cycles between two distinct values. See [3] for more detail including a theoretical analysis of performance through the consideration of Σ as the state estimation error covariance matrix. It should be noted that infinite measurement noise variance R is simply accommodated computationally in (8-9).

There are three key points to be made here:

- Linear Predictive Coding (2-3) may be written in a form reminiscent of the Kalman predictor with K_k replaced by the vector

  $$K_{LPC} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \end{pmatrix}^T.$$

  These LPC predictions are no longer least-mean-square optimal when measurement noise is present, which is reserved for the Kalman filter with its more appropriate vector K_k.
- The optimal Kalman state prediction $\hat{x}_{k+1|k}$ is a vector whose components are as follows.

  $$\hat{x}_{k+1|k} = \begin{pmatrix} \hat{S}_{k+1|k} \\ \hat{S}_{k|k} \\ \hat{S}_{k-1|k} \\ \vdots \\ \hat{S}_{k-n+1|k} \end{pmatrix}.$$

  That is, the Kalman filter yields the filtered signal estimate and smoothed signal estimates up to lag $n-1$ collaterally with the prediction.

- For the signal decimation approach of Ramabudran and Sinha, the Kalman smoothing techniques may be applied with the measurement covariance R_k alternating between R and infinity.

While the maximum price of computation of K_k is the solution of an $n \times n$ matrix eigenvector problem, there are a number of short cuts applicable with autoregressive models [10]. The Kalman filter would appear to provide a clear performance advantage over LPC for these models.

The expected performance improvement from smoothing over prediction is due to the property that signal prediction effectively uses prior data plus model to extrapolate $\hat{S}_{k+1|k}$, while the smoother can take advantage of subsequent data to interpolate the desired signal $\hat{S}_{k|k+m}$.

3 Reconstruction of Missing Data in Array Processing

Consider a linear, equispaced array of M elements in which certain sensors have failed. These sensors can be identified by comparing the broadband power at each sensor to the average over all sensors. The issue is to use time series data from the spatially neighbouring sensors to reconstruct the failed sensor's time series data. Thereafter, beamforming or other applications may use this reconstructed data. Because the data is arriving at each sensor at a fixed known frequency but with unknown direction of arrival, the AR modelling phase describes the spatial relationship between neighbouring sensors' signals. Accordingly, the reconstruction may take place using either the time domain data at each sensor, $x_{\ell,t}$, or using the time-domain Fourier transform at the signal frequency, X_ℓ.

Linear Prediction Algorithm

The linear prediction method proposed in [8] can be applied to estimate the data at the failed sensors. The algorithm linearly combines the both the forwards (in space) and backwards (in space) linear predictions when they are available. The forwards spatial prediction for the faulty kth sensor is given by

$$\hat{X}_k^f = \sum_{j=1}^{p} X_{k-j} a_j, \qquad p+1 \leq k \leq M, \tag{10}$$

where M is the number of sensors and the a_j are the coefficients of a pth order model relating the sensor signals. The backwards spatial prediction is given by

$$\hat{X}_k^b = \sum_{j=1}^{p} X_{k+j} a_j^\star, \qquad 1 \le k \le M - p. \tag{11}$$

The reconstructed estimate is given by a convex linear combination of the forwards and backwards predictions,

$$\hat{X}_k = \alpha \hat{X}_k^f + (1-\alpha)\hat{X}_k^b, \qquad \alpha \in [0,1]. \tag{12}$$

This method can be used with unidirectional data via $\alpha \in \{0,1\}$ if necessary or it can accommodate unequal weightings and AR models from each direction, as would arise when the failed sensor was near one end of the array.

The derivation of the a_j AR coefficients for the LPC model (and for the subsequent Kalman smoother model) is performed via the computation of the signal correlation matrix and the application of the Levinson-Durbin algorithm. The sequence of operations to arrive at the model are as follows.

- Collect a block of data at operable sensors $\{x_{\ell,t},\ \ell = 1,\ldots,n,\ t = 1, N\}$.
- Compute the time Fourier transform component of each sensor signal at the source frequency to yield complex values $\{X_\ell,\ \ell = 1,\ldots,n\}$.
- Compute the sample correlation function between pairs of sensors $\tau_i = \langle X_\ell X_\ell^\star - i\rangle$.
- Apply the Levinson-Durbin algorithm to the correlation values to produce the coefficients, $\{a_j\}$, of the AR model of required order.

As remarked earlier, this algorithm generates AR model coefficients all orders up to the final value. Moderate order AR models can be constructed without needing to rely on data from failed sensors.

Kalman Filtering Algorithm

In parallel to the speech coding problem, Kalman filtering may be used to extend predictive extrapolation methods to interpolation approaches. The same AR signal model may be used but, instead of computing separate forwards and backwards predictions, the one smoothed estimate is calculated. Following the speech coding lead, this Kalman (spatial) predictor for the AR model yields directly spatially smoothed estimates. Missing data due to failed sensors are accommodated by taking the measurement noise variance infinite.

The Kalman filter recursion (7) splits into two tasks,

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k y - k, \quad \text{(measurement update)} \quad (13)$$
$$\hat{X}_{k+1|k} = F\hat{X}_{k|k} \quad \text{(time/space update).} \quad (14)$$
$$(15)$$

Taking infinite R corresponds to conducting two successive space updates before continuing with a measurement update, since $R_k = \infty$ implies $K_k = 0$. The Riccati equation (9) iterates with the non-positive term zero. Compared to the speech coding model, *subsequent* data corresponds to data further down the array for a forwards model and to data up the array for the backwards model.

4 Computational Example

Here we compare the performance of the above Kalman smoother and LPC algorithms for reconstructing measurements at failed sensors. We also examine performance of the array with the measurements from the failed sensors set to zero. All results are compared to the full sensor case.

The linear equispaced array used in the simulations consists of 20 sensors with an inter-sensor spacing of 0.4λ (where λ is the wavelength of the narrowband source). One sensor, in known location 12, has failed. Autoregressive models are chosen to be order five and are fitted from the correlation function of the data. The array response pattern is computed using Fourier methods. The data is windowed with a Hamming window. We consider the case where a narrowband plane wave is arriving from and angle of -0.51 radians. This corresponds to a phase difference between sucessive sensors of -1.2 radians which is obvious from the figures described below. The results are presented for different signal-to-noise ratios.

Figures 2–4 represent the normalized beam power obtained with a single simulation run at the indicated SNR level. Each run comprised 80 data points at each of the 20 sensors. (We present data for the single run case as averaging over many runs tends to confuse the picture; the runs present are "typical" though.) The dash-dot line represents the response with all sensors functioning. The dashed line arises when the failed sensor's signal is set to zero. The other two curves portray the LPC and Kalman smoother approaches. From these plots we can see that there is a substantial advantage in using either the Kalman smoothing or the LPC algorithm over setting the measurements from the failed sensor to zero. Both methods regain most of the performance loss from the failed sensor for SNR of 10dB and 20dB. At levels of 3dB the recovery diminishes. Qualitatively similar results were achieved for multiple source problems.

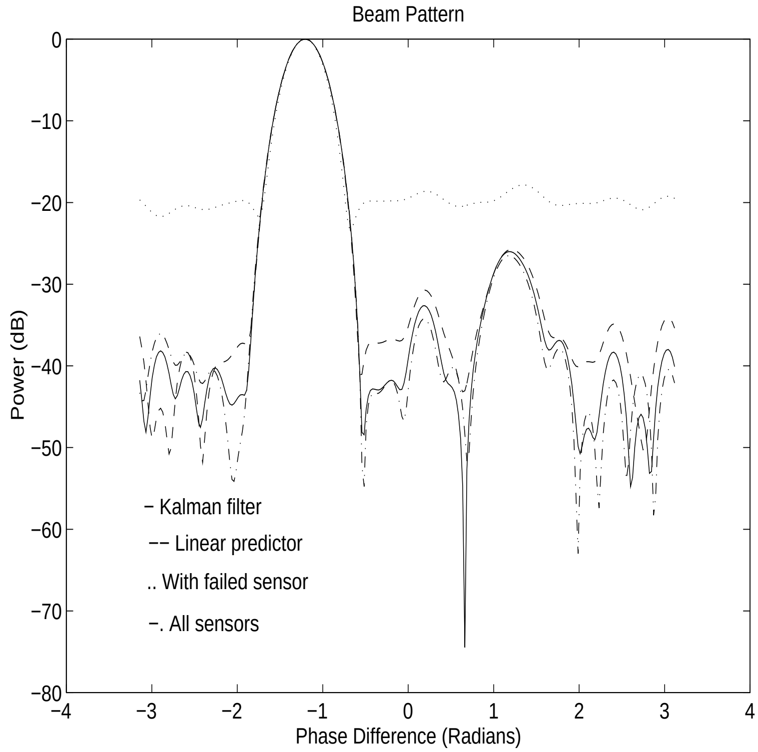

Fig. 2. Normalized Beam Power versus Phase Difference for SNR 20dB

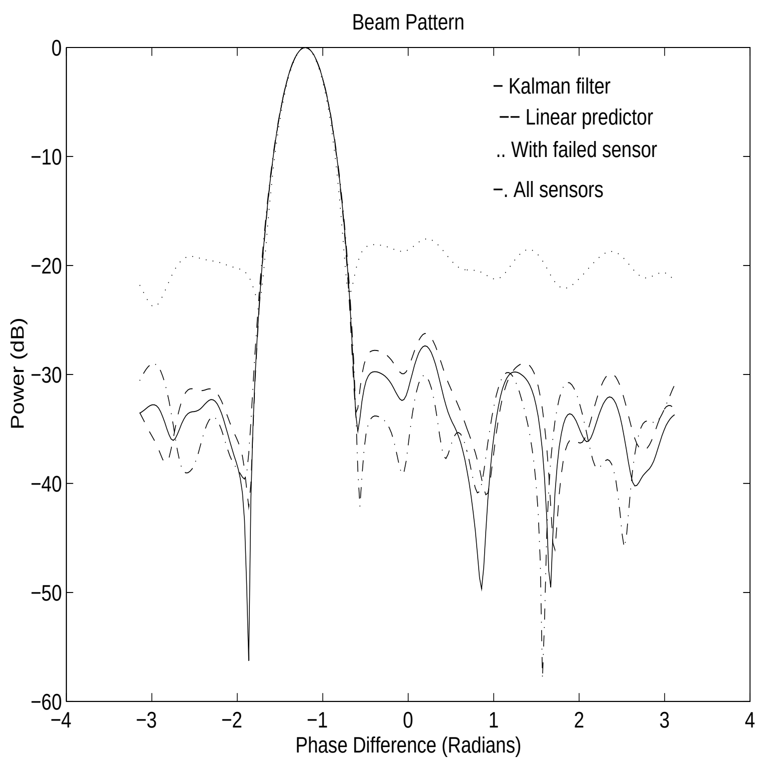

Fig. 3. Normalized Beam Power versus Phase Difference for SNR 10dB

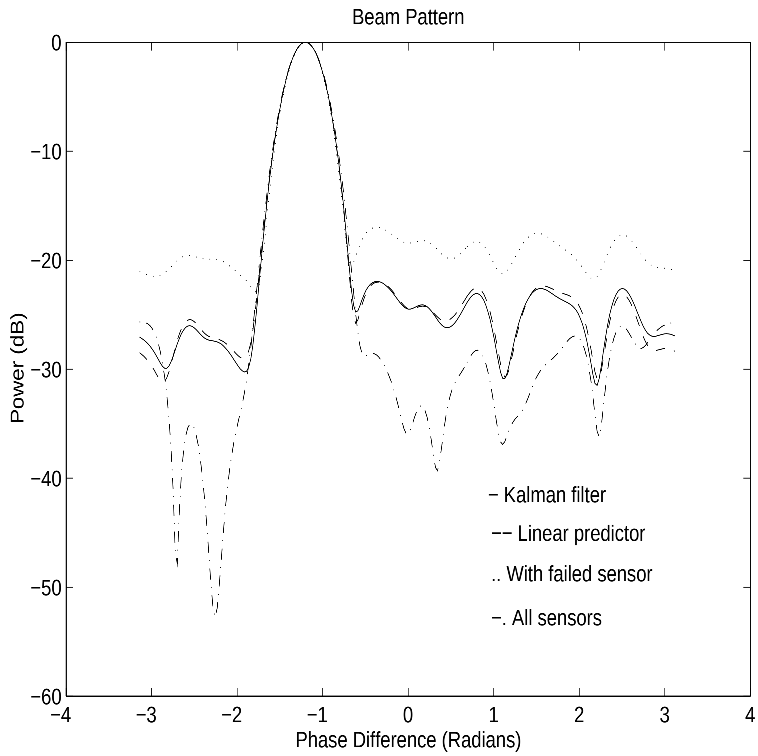

Fig. 4. Normalized Beam Power versus Phase Difference for SNR 3dB

In order to quantify the comparison between the methods, a measure of performance is needed. The exact figure of merit one is interested in will vary from application to application. In order to have *some* simple index of performance we chose to examine the total power of the normalised array response for various SNRs. This figure corresponds to the area under the response pattern and thus gives a (crude) indication of main lobe width. This comparison is shown in Figure 4 with the figure of merit (area under the pattern) versus SNR. This indicates that, for high SNR, the Kalman smoothing approach outperforms LPC. However, as the SNR drops the curves cross indicating that the LPC approach performs better than the *optimal* Kalman smoother. This indicates that, at these signal levels, the AR model derived from the data might be deficient either as a description of the correlation behaviour or as an estimate of the true value. To test this, the data from Figure 4 were recomputed with both the LPC and Kalman smoother being provided precise model information, as opposed to estimated values. Both the reconstruction method perform similarly close to optimal.

5 Conclusion

We have presented the migration of an approach to decimated signal interpolation over to the domain of missing sensor signal reconstruction in array

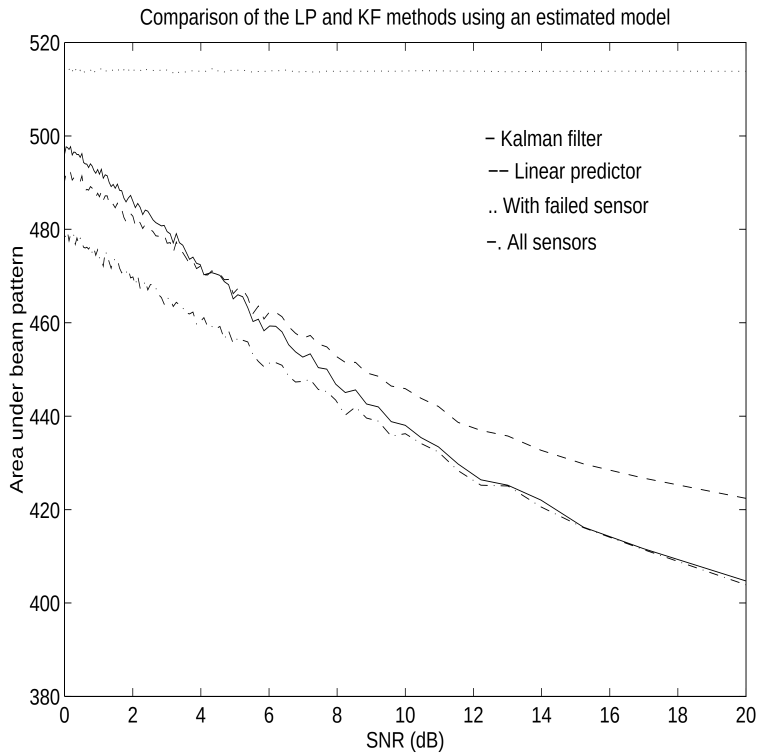

Fig. 5. Comparison of LPC and Kalman smoothing algorithms for various SNRs using an estimated model

processing. While the domains are different, the techniques of Kalman smoothing were applicable and these extended earlier uses of linear predictive coding approaches. Both methods rely on the provision or estimation of signal models, which we have taken in autoregressive form. The analysis of relative performance with signal-to-noise ratio indicates that there is an inherent sensitivity to model quality of the more highly performing smoothing methods.

References

[1] B.D.O. Anderson and J.B. Moore, *Optimal Filtering* (Prentice-Hall, Englewood-Cliffs NJ, 1979).

[2] J.B. Farrell, An array processor implementation of the FIX missing sensor compensation technique, *Defence Research Establishment Atlantic* Canada, Technical Memorandum, 1984.

[3] J.A. Fulton, R.R. Bitmead and R.C. Williamson, Sampling rate versus quantisation in speech coders, *Signal Processing* vol. 56, pp. 209–218, 1997.

[4] S. Haykin, *Adaptive Filter Theory* (Prentice-Hall, Englewood Cliffs, 1991).

[5] J.E. Hudson, *Adaptive Array Principles* (The Institute of Electrical Engineers, London, 1981).

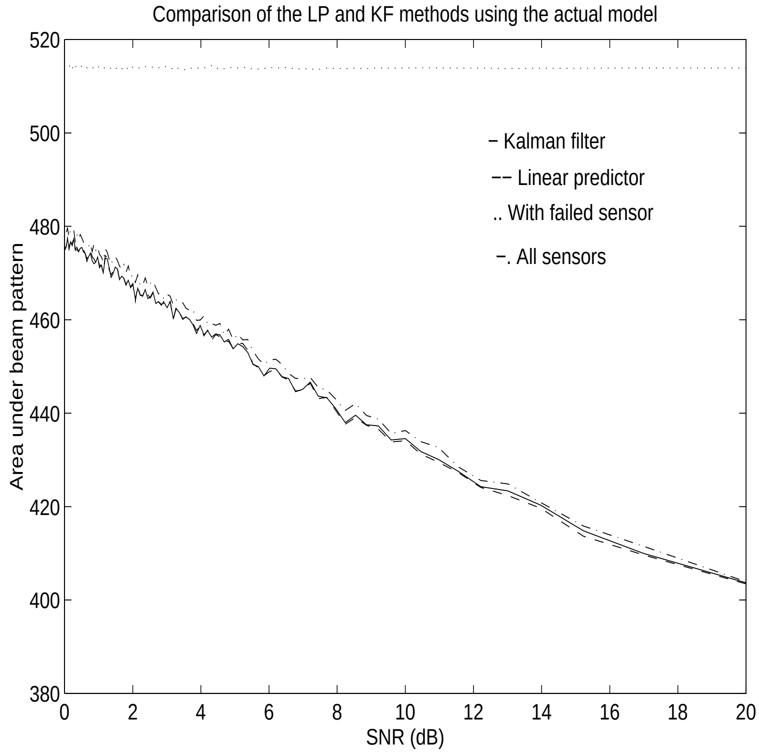

Fig. 6. Comparison of LPC and Kalman smoothing algorithms for various SNRs using the exact model

[6] R.A. Monzingo and T.W. Miller, *Introduction to Adaptive Arrays* (John Wiley and Sons, New York, 1980).

[7] T.V. Ramabudran and D. Sinha, Speech coding using least-squares estimation, *Advances in Speech Coding* (Kluwer Academic Publishers, Boston 1991).

[8]] D.N. Swingler and R.S. Walker, Line-array beamforming using linear prediction for aperture interpolation and extrapolation, *IEEE Trans Acoustics, Speech and Signal Processing* vol. 37, pp. 16–30, 1989.

[9] B.D. Van Veen and K.M. Buckley, Beamforming: a versatile approach to spatial filtering, *IEEE Trans Acoustic, Speech and Signal Processing Magazine* 1988.

[10] C.R. Watkins, R.R. Bitmead and S. Crisafulli, Reduced complexity Kalman filtering for signal coding, *Intelligent Signal Processing and Communications Systems International Workshop* Sendai Japan, 1993.