RECONSTRUCTION FROM FOURIER TRANSFORM PHASE WITH APPLICATIONS TO SPEECH ANALYSIS

B. Yegnanarayana, S. Tanveer Fathima and Hema A. Murthy

Department of Computer Science and Engineering
Indian Institute of Technology, Madras 600036, India.

ABSTRACT

This paper addresses the problem of signal reconstruction from Fourier transform phase. In particular, we examine two aspects of this problem. First, we discuss signal reconstruction from the phase spectrum of the short-time Fourier transform(STFT). Next, we examine the problem of signal recovery from partial phase information. We present the results of our studies on reconstruction from partial phase and discuss the application of these results in speech analysis and coding.

INTRODUCTION

In recent years, considerable interest is being shown in signal reconstruction from the Fourier transform magnitude or phase. This interest is due to the usefulness of such a procedure in several practical problems. In speech enhancement, for example, if speech could be synthesised from its short-time Fourier transform(STFT) magnitude alone, then the STFT phase spectrum of the noisy speech is not required[1]. Also, in applications such as signal coding, the phase spectrum alone need be coded for most finite duration signals[2]. Therefore, efforts have been made to reconstruct such signals from the phase spectra of their Fourier transforms. Several algorithms for reconstruction of a signal from its magnitude or phase spectrum have been proposed[1]-[3]. A different approach to this problem using group delay functions was given in [4],[5].

In processing signals such as speech, short-time processing is appropriate. Nawab et al.,[1] present conditions under which a signal can be uniquely reconstructed from its STFT magnitude. In this paper, we address the problem of signal reconstruction from the phase spectrum. Specifically, we consider two aspects of this problem. We first discuss the reconstruction of a signal from the phase spectrum of its STFT. We see that only mild conditions need be imposed on the signal to reconstruct it uniquely from its STFT phase. We then address the problem of signal recovery from partial phase information. Here, we show that the missing phase information can be compensated by the knowledge of a certain number of signal samples.

SIGNAL RECONSTRUCTION FROM STFT PHASE

The conditions for unique recovery of a signal from its phase spectrum are given in [2],[3]. These conditions state that a signal is uniquely specified, to within a scale factor, by its phase spectrum, if (i) the signal is of finite duration, (ii) the z-transform of the signal does not have zeros in reciprocal pairs, and (iii) the z-transform has no zeros on the unit circle. These conditions also hold good for reconstruction from the STFT phase. The basic method used to recover the signal from the STFT phase is a sequential extrapolation. An iterative algorithm which reconstructs the first section using the STFT phase information alone is given in [2]. Knowledge of one signal sample in this section helps in the exact recovery of this segment of the signal. The other sections are reconstructed by shifting the rectangular analysis window successively across the signal, and making use of the samples in the region of overlap.

We notice that although the signal can be reconstructed using the phase spectrum alone, extra information about the signal samples can be used to reduce the reconstruction error. Specifically, as the number of known samples increases, error(error/unknown sample) decreases for a given number of iterations. Such a behaviour is understandable, because, as the number

CH2396-0/87/0000-0301 $1.00

of known samples increases, fewer samples remain to be reconstructed in that section. Fig.1 illustrates this behaviour for four segments(64 samples per segment) of a speech signal. In each case there is an optimum number of known samples beyond which increasing the number of samples does not reduce the reconstruction error significantly.

SIGNAL RECONSTRUCTION FROM PARTIAL PHASE

In several practical situations, the phase information may only be partial. In order to investigate the conditions under which reconstruction with partial phase is possible, we examine the following cases :

A. Partial phase and partial signal is available.
B. Effect of distribution of partial phase data.
C. Partial phase and a few magnitude spectral samples are available.

A. Reconstruction from Partial Phase and Partial Signal Samples

We know that complete phase information is required if a finite duration signal is to be reconstructed uniquely. Therefore, if some of the phase samples are missing, it is not possible to recover the signal. However, we would like to see if the missing phase samples can be compensated by the knowledge of signal samples. In our studies, we have found that there exists a certain relationship between the number of phase samples and the number of signal samples for a unique reconstruction of the signal.

Suppose only m of the phase samples, i.e., $\theta_x(k)$, for $k = 1, 2, \ldots, m$ and N-m signal values, i.e. $x(n)$, for $n = 0, 1, \ldots, N-m-1$ are known. We now present an algorithm to reconstruct the signal from its partial phase and partial signal samples. Let us assume that the signal $x(n)$ is zero outside the interval $0 \leq n \leq N-1$. Let its M-point DFT be denoted by $X(k)$, where $M \geq 2N$. Then

$$X(k) = |X(k)| \exp[j\theta_x(k)], \quad \text{for } k = 0, 1, \ldots, M-1 \quad (1)$$

The iterative algorithm is described as follows :

Step 1: We begin by forming a signal

$$y_0(n) = \begin{cases} x(n), & n \in I_T \text{ and } 0 \leq n \leq N-1 \\ 0, & n \notin I_T \text{ and } 0 \leq n \leq M-1 \end{cases} \quad (2)$$

where I_T is the set of time instants at which the signal samples are known.

Step 2: We compute the M-point DFT $Y_0(k)$ of $y_0(n)$. From this we form the function $X_1(k)$ using the known phase samples in place of the phase samples of $Y_0(k)$ at the appropriate frequency points. That is

$$X_1(k) = |Y_0(k)| \exp[j\theta_1(k)] \quad (3)$$

where

$$\theta_1(k) = \theta_x(k), \; k \in I_p$$
$$= \theta_{Y_0}(k), \; k \notin I_p \text{ and } 0 \leq k \leq M-1 \quad (4)$$

where I_p is the set of frequency points at which the phase samples are known. Computing the inverse DFT of $X_1(k)$ provides the first estimate $x_1(n)$ for $x(n)$. Since an M-point DFT is used, $x_1(n)$ is an M-point sequence which is, in general, nonzero for $N \leq n \leq M-1$.

Step 3: From $x_1(n)$ form another sequence, $y_1(n)$, defined by

$$y_1(n) = x(n), \quad n \in I_T \text{ and } 0 \leq n \leq N-1$$
$$= x_1(n), \quad n \notin I_T \text{ and } 0 \leq n \leq N-1 \quad (5)$$
$$= 0, \quad N \leq n \leq M-1$$

Using $y_1(n)$, Step 2 is performed to get a new estimate $x_2(n)$. Steps 3 and 2 are then repeated for a specified number of iterations, or, until a specified level of reconstruction is reached.

In this iterative procedure we can show that [6] the total squared error between $x(n)$ and its estimate is nonincreasing with each iteration.

Therefore, we conclude that it is possible to exactly compensate for the missing phase information with the signal samples. In practice, however, we may not have the required number of signal samples known. In such cases, it is interesting to see the effect of under compensation of the phase samples by the signal samples. In the study conducted, we found that the performance was poor when only very few signal samples were available. It improves steadily as the number of known signal samples increases until the required number of samples are reached (N-m for m known phase samples). As the number of known signal samples increases beyond N-m, then we do not see much change in the quality of reconstruction. However, the reconstruction is now much faster, i.e., a given error limit is reached earlier for an over compensated case than for the exact compensated case. Thus we find that there is no abrupt transition from exact reconstruction to poor reconstruction as the number of signal samples falls below the required number.

We now illustrate the effect of compensation of phase samples by signal samples. Fig.2a shows the original signal of length 256 divided into four segments each of length 64. For each segment 32 distinct phase(uniformly distributed) samples out of the required 64 samples and 16 signal samples are assumed to be known. This corresponds to the case of under compensation. Fig.2b shows the reconstructed signal after 100 iterations for each segment. Fig.2c illustrates the reconstruction of the signal for the exact compensated case(32 phase and 32 signal samples).

Fig.2d illustrates the reconstruction of the signal for the over compensated case(32 phase and 48 signal samples). From these figures we note that the reconstruction is poor for the undercompensated case, but improves gradually as the number of signal samples are increased. For the over compensated case the error in reconstruction was found to be smaller than that of the exact compensated case. Fig.3 shows the error in reconstruction per unknown sample as a function of the number of the known samples.

B. Effect of Distribution of Partial Phase Data

If exact reconstruction is not required, then certain distributions of the partial phase samples for some special cases of signals such as the narrow band signals, could give a reconstruction which is good enough for practical purposes. As an illustration consider the segments of the speech signal shown in Fig.4a. If the partial phase samples are so distributed that they cover the frequency regions where the magnitude spectrum peaks, then we obtain interesting results. About 50% of the phase samples and 20% of signal samples are assumed to be known. The reconstructed segments after 50 iterations using partial phase for two different distributions of the phase samples are shown in Figs.4b and 4c. Fig.4b shows the reconstructed signal for the case when the phase samples are uniformly distributed. Fig.4c shows the reconstructed signal for the case when the phase samples cover the frequency regions corresponding to the peaks in the magnitude spectrum. From these figures we conclude that reasonably good reconstruction can be obtained if the phase samples are known in the bands where the magnitude spectrum peaks occur.

C. Reconstruction from Partial Phase and Partial Magnitude spectra

In section A we have seen that the missing phase samples can be exactly compensated by the knowledge of signal samples. It appears that the missing information cannot be compensated by the magnitude spectral samples. We have examined this case in detail. From our studies, we conclude that, in general, it is not possible to reconstruct the signal exactly when partial phase spectrum and partial magnitude spectrum are given. However, a close approximation to the original signal can be obtained for certain special cases of magnitude spectrum sample distributions. Also, the nature of reconstruction is affected by the distribution of magnitude spectral samples. This is unlike the earlier case of compensation with signal samples, wherein the distribution of signal samples did not affect the nature of reconstruction.

CONCLUSIONS

In this paper, we have addressed the problem of signal reconstruction from phase information. We first discussed the reconstruction of a signal from its short-time Fourier transform phase. Next, we addressed the problem of signal recovery from partial phase information. We have shown that, while it is possible to compensate for the missing phase samples by the signal samples, it is not possible to do so by the magnitude spectral samples. From this study, it appears that, as far as the reconstruction of a finite duration signal is concerned, the signal and the phase spectral samples are more important than the magnitude spectral samples. We are currently exploring the possibility of using these results for processing noisy speech and also for efficient coding of speech.

REFERENCES

[1] S.H.Nawab, T.F.Quatieri, and J.S.Lim, "Signal Reconstruction from Short-time Fourier Transform Magnitude", IEEE Trans. Acoust., Speech, Signal Processing, Vol.ASSP-31, pp.986-998, August 1983.

[2] M.H.Hayes, J.S.Lim, and A.V.Oppenheim, "Signal reconstruction from phase or magnitude", IEEE Trans. Acoust., Speech, Signal Processing, Vol.ASSP-28, pp.672-680, December 1980.

[3] G.A.Merchant and T.W.Parks, "Reconstruction of signals from phase: efficient algorithms segmentation, and generalizations", IEEE Trans. Acoust., Speech, Signal Processing, Vol.ASSP-31, pp.1135-1147, October 1983.

[4] B.Yegnanarayana and A.Dhayalan, "Noniterative technique for minimum phase signal reconstruction from phase or magnitude", in Proc., IEEE Int.conf. Acoust., Speech, Signal Processing, Boston, MA, pp.639-642, 1983.

[5] B.Yegnanarayana, D.K.Saikia and T.R.Krishnan, "Significance of group delay functions in signal reconstruction from spectral magnitude and phase", IEEE Trans. Acoust., Speech, Signal Processing, Vol.ASSP-31, pp.1286-1293, June 1984.

[6] S.T.Fathima, "Studies on a class of information recovery problems", M.S. Thesis, Department of Computer Science and Engineering, Indian Institute of Technology, Madras, India, 1986.

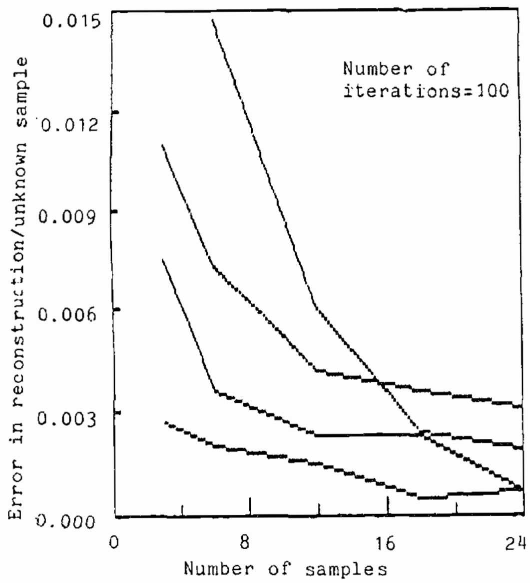

Fig.1. Signal reconstruction from full phase. Error in reconstruction per unknown sample as a function of number of known signal samples. Plots for four different speech segments are shown.

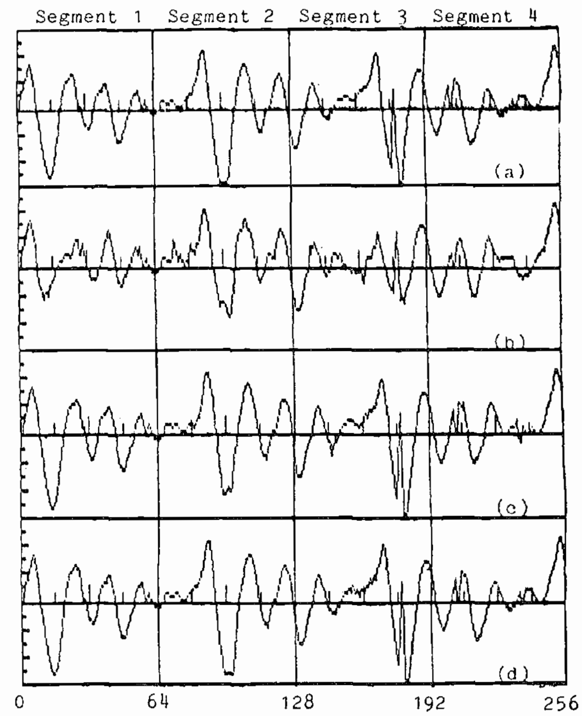

Fig.2. Illustration of reconstruction from partial phase for different cases of known signal samples. (a) Original signal (b) Under compensated (32 phase, 16 signal samples/segment). (c) Exact (32 phase, 32 signal samples/segment). (d) Over compensated (32 phase, 48 signal samples/segment).

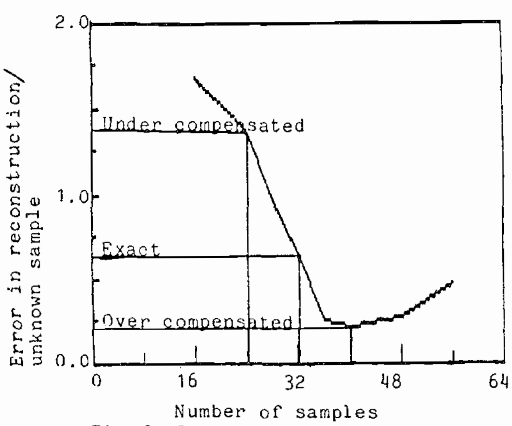

Fig. 3. Error in reconstruction/unknown sample (for partial phase) as a function of number of known samples.

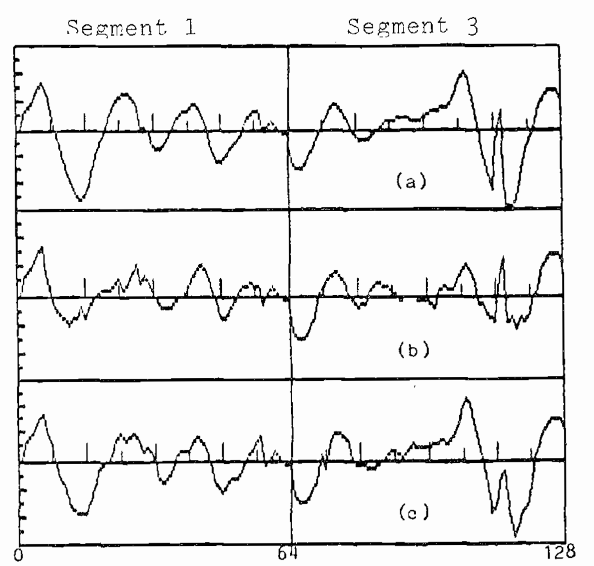

Fig.4. Signal reconstruction from partial phase. 32 phase and 12 Signal samples are known/section. (a) Original signal (b) Signal reconstructed with known phase samples uniformly spaced (c) Signal reconstructed with known phase samples placed at the peaks in the magnitude spectrum.