A DECOMPOSITION TECHNIQUE FOR COMPOSITE SIGNALS

T. V. ANANTHAPADMANABHA AND B. YEGNANARAYANA

(Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore-560012 India)

Received on February 6, 1975

ABSTRACT

A new method for decomposition of composite signals is presented. It is shown that high frequency portion of composite signal spectrum possesses information on echo structure. The proposed technique does not assume the shape of basic wavelet and does not place any restrictions on the amplitudes and arrival times of echoes in the composite signal. In the absence of noise any desired resolution can be obtained The effect of sampling rate and frequency window function on echo resolution are discussed. Voiced speech segment is considered as an example of composite signal to demonstrate the application of the decomposition technique.

Key words: Composite Signal, Echo Resolution, Voiced Speech.

1. INTRODUCTION

Composite signals consisting of multiple wavelets of identical shape overlapping in time can be regarded as superposition of echoes of a basic waveform. If the echo amplitudes and arrival times are expressed relative to the first basic waveform, then the echo characteristics can be represented by an impulse of unit value at the origin followed by impulses of values corresponding to the echo amplitudes occurring at the echo arrival times. The composite signal will then be a convolution of the basic waveform and the echo characteristics. The problem of decomposition of composite signals consists of determining the shape of the basic wavelet and the number of echoes along with the amplitudes and arrival times. Such situations often arise in areas like sonar, radar, speech, electrophysiology, seismology etc. [1].

Digital signal processing techniques have been effectively used for composite signal decomposition although each technique has proved useful under individual circumstances. In the digital inverse filtering technique [1, 2]. for example, the composite signal is passed through a linear time-invariant system so that the output gives pulses at the instants of echo arrivals. Such a system must necessarily have a frequency response which is the reciprocal of the Fourier transform of the signal components to be removed. In general it turns out that the transfer function of the system is the ratio of the Fourier transform of the output pulse to the Fourier transform of the basic waveform. By this method it is possible to obtain any desired echo resolution under noise-free conditions provided sufficient computational time is allowed. The drawback of this technique is that the shape of the basic waveform must be known for designing the inverse filter.

Cepstrum techniques [3], [4], [5] have been suggested for the decomposition of composite signals when the shape of the basic signal is not known. Since the Fourier transform of the composite signal is a product of the Fourier transforms of the basic waveform and the echo characteristics, in in the log-spectrum of the composite signal the periodic components due to the echo characteristics will be superimposed on the log spectrum of the basic waveform. The inverse Fourier transform of the log spectrum called cepstrum will yield peaks at the echo arrival times. The basic waveform is extracted through cepstral smoothing [4], [5].

Cepstrum techniques are effective mainly in the detection and wavelet extraction for single echo case. For multiple echoes peaks appear in the cepstrum not only at the echo arrival times but also at the sums and differences of their multiples. A careful interpretation is required to determine the peaks corresponding to the echo arrival times. The heights of these peaks do not have a simple relation with the amplitudes of the echoes they represent. In fact, the heights depend upon the relative positions of the echoes also. For an unambiguous interpretation of echo data from cepstrum there are restrictions [5] on the number of echoes and their amplitudes and arrival times. Thus the technique is not applicable for a general composite signal.

We present here a decomposition technique that does not require *a priori* knowledge of the basic waveform and also imposes no restrictions on the echo characteristsics. For noiseless signals the technique appears to be superior to earlier ones. In the presence of noise this technique requires a higher signal to noise ratio compared to other techniques for the same echo resolution. A voiced speech segment is chosen to illustrate the application of our technique on real data.

2. The Proposed Decomposition Technique

(a) Problem

Given a signal $x(t)$ that is a finite summation of a basic wavelet $s(t)$ and its echoes:

$$x(t) = \sum_{i=0}^{n} a_i s(t - t_i) \tag{1}$$

where $a_0 = 1$ and $t_0 = 0$, the problem is to determine the echo arrival times t_i and the echo amplitudes a_i for i = 1, 2,.., n and the waveform $s(t)$.

(b) Basis

Equation (1) can be written as

$$x(t) = s(t) * h(t) \tag{2}$$

where

$$h(t) = \sum_{i=0}^{n} a_i \delta(t - t_i) \tag{3}$$

and the * denotes convolution. Taking the Fourier transform of (2) we get

$$X(w) = S(w) H(w) \tag{4}$$

where

$$H(w) = \sum_{i=0}^{n} a_i e^{-jwt_i} \tag{5}$$

H (w) is a finite sum of periodic functions with frequencies t_i. For finite energy signals $s(t)$ having rational spectra, $|S(w)|$ falls off monotonically at high frequencies [6]. It is this difference in the characteristics of $S(w)$ and $H(w)$ that we would like to exploit to extract the information about the echoes even though $S(w)$ is unknown.

In a general deconvolution problem $H(w)$ may have any shape and hence must be known for all w in order to get the impulse response $h(t)$ of $H(w)$. This requires a complete knowledge of $S(w)$ so that an inverse filter of the form $1/S(w)$ can be designed to extract $H(w)$ from $X(w)$. However, in the present case, since $H(w)$ is a finite sum of periodic functions, it is sufficient if we know $H(w)$ over a band of frequencies to determine the periodic components. Once all the periodic components and their amplitudes are identified then $H(w)$ is known completely from (5).

(c) Method

The log spectrum of a typical composite signal is shown in Fig. 1 where the periodicities due to $H(w)$ appear as fluctuations over the log spectrum of $s(t)$ shown by dotted line. The portion of the spectrum in the range 0 to w_0 includes the significant frequency components of the signal $s(t)$. Therefore $s_1(t)$, the inverse Fourier transform of $S(w)$ taken over the band 0 to w_0 corresponds very nearly to $s(t)$ itself. The inverse transform $x_1(t)$ of $S(w)\,H(w)$ over this frequency band will be $s_1(t) * h(t)$ which corresponds almost to the original composite signal $x(t)$ itself. Thus we have

$$S_1(w) = \begin{cases} S(w) & \text{for } 0 \leqslant |w| \leqslant w_0 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$s_1(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_1(w)\, e^{jwt}\, dw \approx s(t) \tag{7}$$

$$\begin{aligned} x_1(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_1(w) H(w)\, e^{jwt}\, dw = s_1(t) * h(t) \\ &\approx s(t) * h(t). \end{aligned} \tag{8}$$

On the other hand, consider the log spectrum of Fig. 1 in the range $w_1 \leqslant |w| \leqslant w_2$ where the spectrum of $s(t)$ is very nearly constant. Then

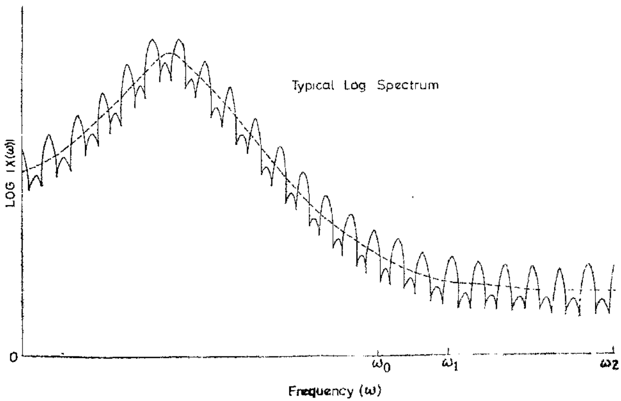


FIG. 1. Log spectrum of a typical composite signal.

$s_2(t)$, the inverse transform of $S(w)$ over $w_1 \leqslant |w| \leqslant w_2$, will have a significant component around $t = 0$ and negligible components at other times. Thus we have

$$S_2(w) = \begin{cases} S(w) & \text{for } w_1 \leqslant |w| \leqslant w_2 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

$$s_2(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_2(w)\, e^{jwt}\, dw \tag{10}$$

The inverse transform $x_2(t)$ of $S(w)\, H(w)$ over the band $w_1 \leqslant |w| \leqslant w_2$ is given by

$$x_2(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_2(w)\, H(w)\, e^{jwt}\, dw$$

$$= s_2(t) * h(t) = \sum_{i=0}^{n} a_i\, s_2(t - t_i) \tag{11}$$

If $s_2(t)$ is sharp enough then all the echoes will be resolved and the a_i and t_i can be obtained from $x_2(t)$. We shall show that this is true for signals possessing rational spectra.

For signals with rational spectrum the asymptotic behaviour of the spectrum at high frequencies is given by $1/w^n$ [6].
Then, for large w, we have

$$|S(w)| \approx K/w^n \tag{12}$$

where K is a constant. For $w = w_c - B$,

$$|S(w_c - B)| = |S(w_c)| \left(1 - \frac{B}{w_c}\right)^{-n} \tag{13}$$

If $nB/w_c \ll 1$, we can regard $|S(w)|$ to be a constant equal to $|S(w_c)|$ in the range $w_c - B \leqslant |w| \leqslant w_c$. Since the phase of $S(w)$ is also very nearly constant in this range, we can write

$$\begin{aligned} S(w) &\approx S(w_c) \text{ for } w_c - B \leqslant w \leqslant w_c \\ S(w) &\approx S^*(w_c) \text{ for } -w_c \leqslant w \leqslant -w_c + B \end{aligned} \tag{14}$$

where $S^*(w_c)$ is the complex conjugate of $S(w_c)$. Define a window function

$$G(w) = \begin{cases} 1 & w_c - B \leqslant |w| \leqslant w_c \\ 0 & \text{otherwise,} \end{cases} \tag{15}$$

Then

$$S_2(t) = \frac{1}{2\pi}\left[\int_{w_c-B}^{w_c} S(w)\, e^{jwt}\, dw + \int_{-w_c}^{-w_c+B} S(w)\, e^{jwt}\, dw\right]$$

$$\approx \frac{1}{2\pi}\left[\int_{0}^{\infty} S(w_c)\, G(w)\, e^{jwt}\, dw + \int_{-\infty}^{0} S^*(w_c)\, G(w)\, e^{jwt}\, dw\right] \quad (16)$$

$S(w_c)$ is a Complex constant as per equation (14).

Let

$$S(w_c) = P - jQ$$

and

$$S^*(w_c) = P + jQ \quad (17)$$

where P is the even real part and Q is the odd imaginary part of $S(w)$ in the range $w_c - B \leqslant |w| \leqslant w_c$. Substituting (17) into (16) we get

$$s_2(t) = \frac{2}{2\pi}\left[\int_{0}^{\infty} PG(w)\cos wt\, dw + \int_{0}^{\infty} QG(w)\sin wt\, dw)\right] \quad (18)$$

Normally either P or Q will be significant. When $P \gg Q$ the envelope of $s_2(t)$ will be of the form shown in Fig. 2 *a*. This occurs when the phase of $S(w)$ at large w is an even multiple of $\pi/2$. When $Q \gg P$ the envelope of $s_2(t)$ will be of the form shown in Fig. 2 *b*. This occurs when the phase of $S(w)$ at large w is an odd multiple of $\pi/2$. In the latter case the function crosses zero at the echo location. But by making Q the even real part and P the odd imaginary part it is possible to get a peak at the echo location as in Fig. 2 *a*. This is achieved by multiplying $X(w)$ with $jSgn(w)$.

(d) Choice of the Window Function

Proper choice of the window function helps in reducing the ambiguity of the echo detection and its amplitude estimation. The transform of a rectangular window function has considerable sidelobe leakage [7]. The first sidelobe peak is about 20% of the main lobe and therefore it is likely to be misinterpreted as another echo. Moreover, the side lobes affect the amplitudes of closely spaced echoes. A suitable window function should be chosen to overcome these problems [7]. By setting the values of the spectrum outside the window width to zero the picket fence effect [8] in $x_2(t)$ is reduced.

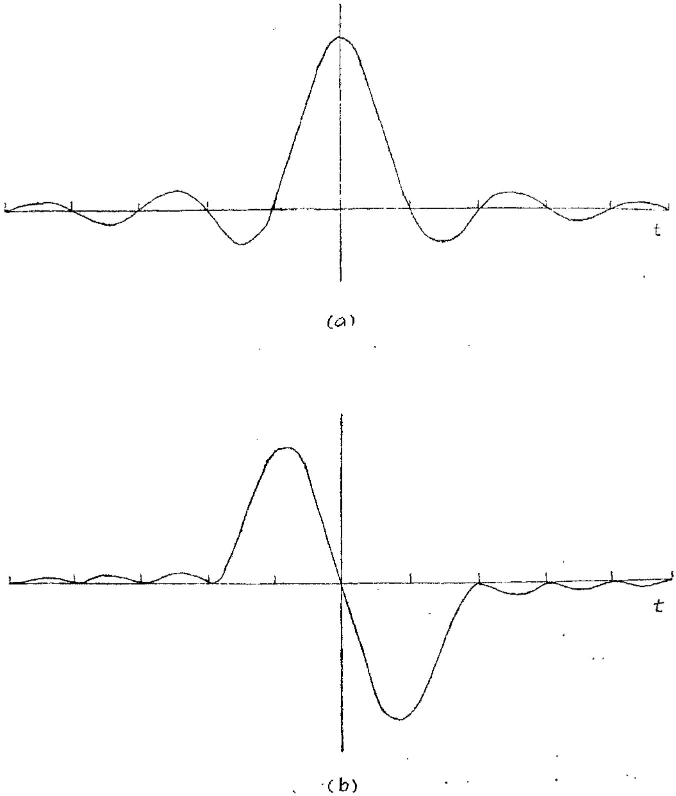


FIG. 2. The shape of the envelope of $s_2(t)$. (*a*) $P \gg Q$. (*b*) $Q \gg P$.

(*e*) *Choice of Window Width and its Implications*

The choice of the window width B for the best representation of the output $y(t)$ depends upon the minimum echo arrival time or the minimum epoch difference. The width B must be greater than $2\pi/t_{min}$ so that several cycles of frequency corresponding to t_{min} are included within B. Therefore for resolving echoes the lowest value of B is given by

$$B = 2\pi/t_{min} \tag{19}$$

where t_{min} is the lower of the minimum echo arrival time and the minimum epoch difference. For equation (14) to be satisfied we require

$$nB/w_c \ll 1. \tag{20}$$

Therefore, we get

$$\frac{n\,(2\pi/t_{min})}{w_c} \ll 1. \tag{21}$$

If w_c is chosen as the folding frequency of the signal $x(t)$, then the inequality to be satisfied by the sampling interval $T(=\pi/w_c)$ is given by

$$T \ll t_{min}/2n. \tag{22}$$

Thus we can stipulate a sampling rate to obtain a desired resolution for a given minimum echo arrival time. Usually the sampling rate required for composite signals is much higher than for the basic waveform.

3. Computational Procedure

The following steps outline the procedure for implementation of the proposed technique.

(*a*) The composite signal is sampled at some convenient rate ($1/T$ Hz).

(*b*) The discrete Fourier transform (DFT) is computed and its log spectrum is plotted.

(*c*) From the log spectrum a suitable frequency band is chosen over the nearly flat portion of $|S(w)|$. The DFT of the composite signal is multiplied by a suitable window function $G(w)$ over the chosen width. In this paper, a Hanning window centered around the folding frequency is chosen for $G(w)$.

(*d*) The inverse DFT of $S(w)\ H(w)\ G(w)$ is computed. Since the discrete points of this inverse DFT alternate in sign they are multiplied by $(-1)^n$ to obtain the output samples $y(t)$. The a_is and t_is are obtained from the amplitudes and locations of the peaks in the output.

(*e*) If the resolution in the output is poor then the original composite signal is sampled at a higher rate and the steps (*a*) through (*d*) are repeated.

(*f*) If the output pulses are of the form shown in Fig. 2*b* then a 90° phase shift [*jSgn* (*w*)] is given to *S* (*w*) and the steps (*c*) and (*d*) are repeated.

4. EXAMPLES

The above procedure is illustrated thorugh examples in this section A basic wavelet of the form $s(t) = t \exp(-at)$ is considered as it is supposed to be representative of a degraded radar or sonar pulse [5]. Consider the composite signal

$$x(t) = s(t) + 0{\cdot}4\, s(t-27) + 0{\cdot}4\, s(t-54) \tag{23}$$

where

$$s(t) = t \exp(-0{\cdot}06\, t) \tag{24}$$

$x(t)$ is sampled at 1/1·5 Hz and the log spectrum is obtained through an FFT subroutine. In the log spectrum (Fig. 3*a*) of $x(t)$ the gross spectral features corresponding to the spectrum of $s(t)$ are shown by dotted line. The periodic fluctuations superimposed on the log spectrum of $s(t)$ represent

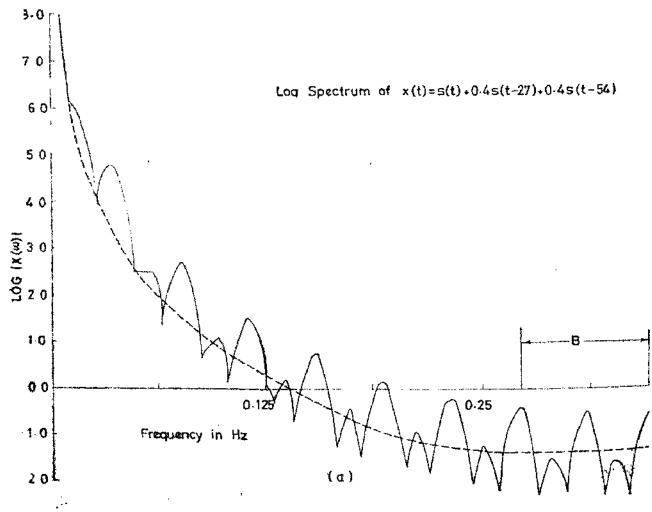


FIG. 3*a*. Log spectrum of a composite signal with two echoes. Sampling frequency = 0·66 Hz,

the echo characteristics. It can easily be seen that the region over which the log spectrum is flat is small for the sampling rate considered. The width B contains only two cycles for the echo arrivl time of 27 seconds. For lower values of echo arrival times fewer number of cycles will be contained ir the width resulting in a poor resolution of echoes. The resolution can be improved by taking a largeı sampling rate. The effect of increasing th· sampling rate fiom 1/1·5 Hz to 1/0·375 Hz is shown in Fig. 3 *b*. Here the width B over which the log spectrum is flat is larger compared to that shown in Fig. 3 *a*. The output (Fig. 4) is obtained by applyıng Hanning window over this width and taking the inverse DFT.

Figuıes 5 *a*, 5 *b* and 5 *c* show the effect of the width B and the sampling rate on echo resolution. Smaller values of B produce poorer ıesolution. On the other hand higher sampling rate allows larger value of B to be chosen thus improving ıesolution. In general high sampling rate is recommended for good resolution especially for low values of interecho intervals.

The application of the technique for the composite signal

$$x(t) = s(t) + 0{\cdot}4s(t-9) + 0{\cdot}4s(t-24) + 0{\cdot}8s(t-3\ 6) + 0{\cdot}2s(t-72) \qquad (25)$$

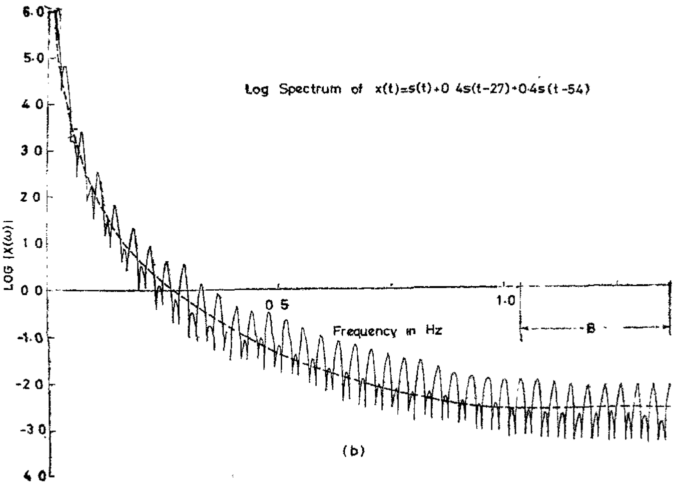


FIG. 3*b* Log spectuam of composite signal with two ecohes. Sampling frequency = 2·66 Hz.

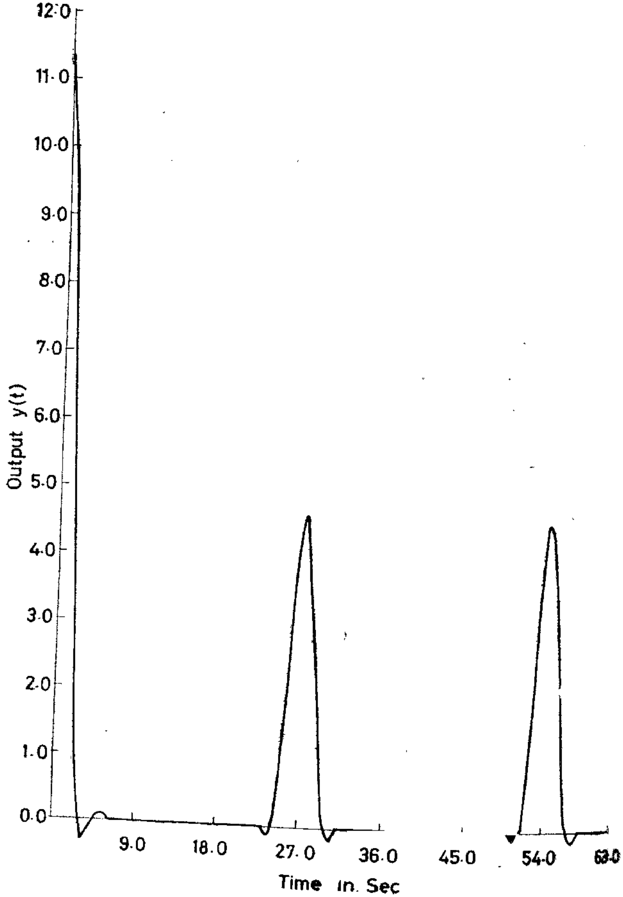


FIG. 4. The output for the composite signal $s(t) + 0 \cdot 4s(t - 27) + 0 \cdot 4s(t - 54)$ and window width shown in Fig. 3b.

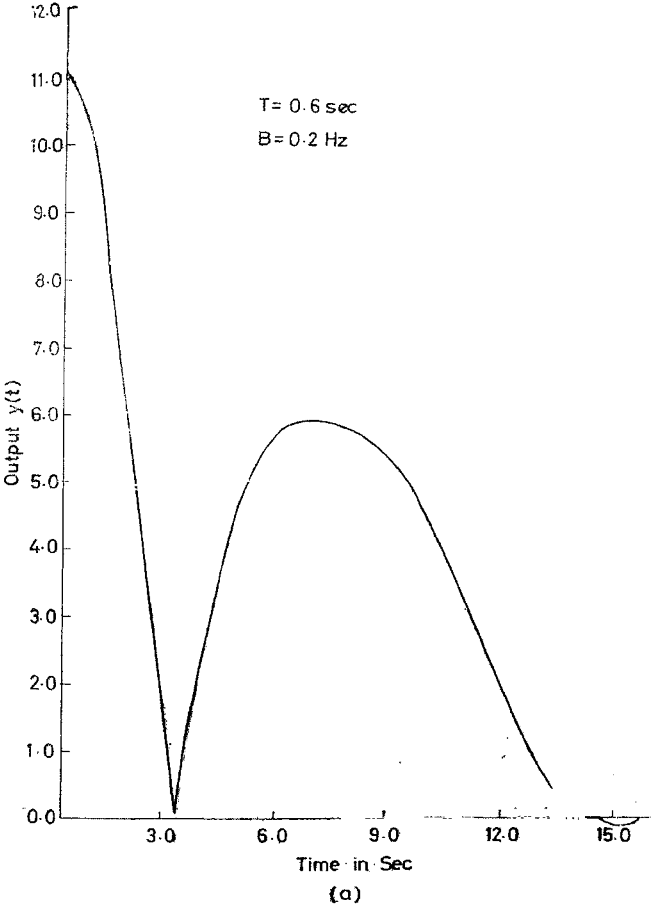


FIG. 5a. The output for the composite signal $x(t) = s(t) + 0{\cdot}4s(t-9) + 0{\cdot}4s\,t(-15)$ for different sampling frequencies and window widths.

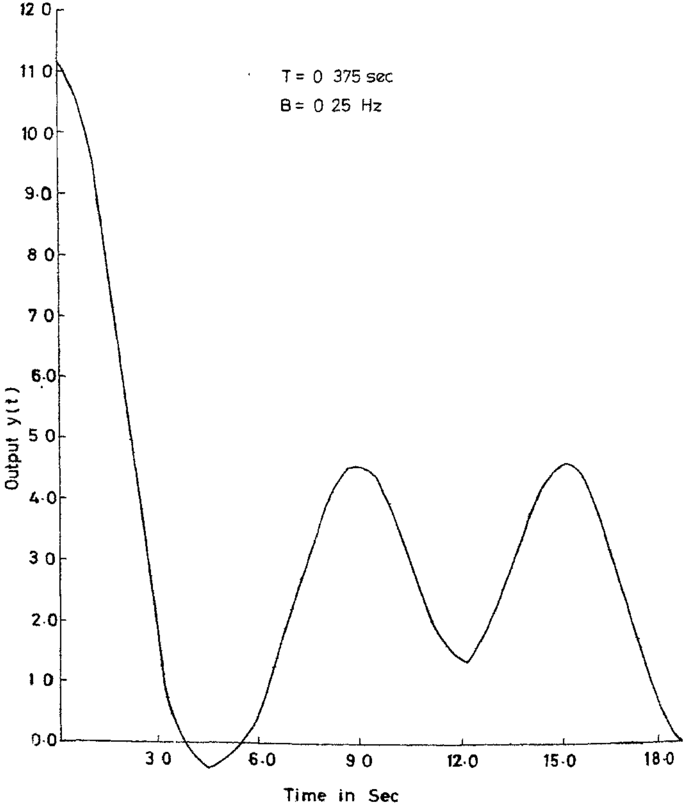
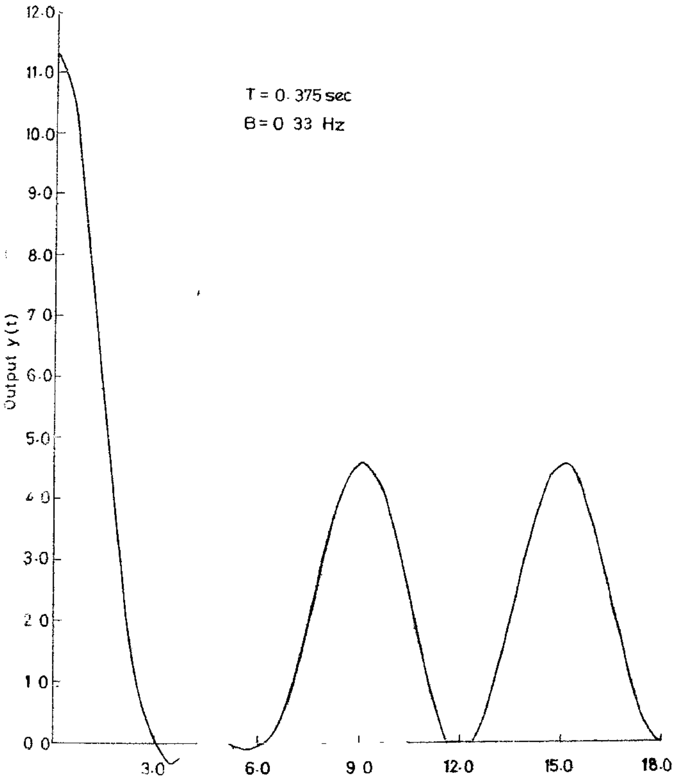

T = 0 375 sec
B = 0 25 Hz
Output y(t)
12 0
11 0
10 0
9·0
8 0
7 0
6·0
5 0
4·0
3 0
2·0
1 0
0·0
3 0
6·0
9 0
12·0
15·0
18·0
Time in Sec

FIG. 5b



12.0
11.0
10.0
9.0
8.0
7 0
6.0
5.0
4 0
3.0
2 0
1 0
0 0
Output y(t)
T = 0. 375 sec
B = 0 33 Hz
3.0
6.0
9 0
12.0
15.0
18.0
Time in Sec

(c)

FIG. 5c

with four echoes is illustrated in Figs. 6*a* and 6*b*. The echo arrival times are obtained exactly from the peaks in the output shown in Fig. 6*b*.

The above examples are typical of the cases where the constant phase at large frequencies is an even multiple of $\pi/2$. The transform of $s(t)$ is

$$S(w) = \frac{1}{(a + jw)^2} = \frac{s^2 - w^2}{(a^2 + w^2)^2} - \frac{2jaw}{(a^2 + w^2)^2} \quad (26)$$

For $w \ll a$,

$$S(w) \approx - 1/w^2 \quad (27)$$

That is, the real part is much greater than the imaginary part. Consider the basic signal

$$s(t) = t^2 e^{-at} \quad (28)$$

The transform of this signal is

$$S(w) = \frac{2}{(a + jw)^3} = 2 \frac{a^3 - 3aw^2}{(a^2 + w^2)^3} - j \frac{(3a^2w - w^3)}{(a^2 + w^2)^3} \quad (29)$$

For $w \gg a$,

$$S(w) = \approx jw^3 \quad (30)$$

That is, the imaginary part is much greater than the real part.

composite signal $x(t)$ is formed with $s(t)$ of Eq. (28) as the basic signal

$$x(t) = s(t) + 0 \cdot 8s(t - 45) + 1 \cdot 2s(t - 54) \quad (31)$$

The output obtained by the application of the present technique for the decomposition of the above composite signal is shown in Fig. 7*a*. It is seen that the output pulses have an odd symmetry and they cross the time axis at the echo arrival times. The output obatined when a 90° phase shift is given to the transform of $x(t)$ over the chosen window width is shown in Fig. 7 *b*. Now the peaks of the pulses occur at the echo arrival times. The relative amplitudes are still maintained. The excessive side lobes ir Fig. 7*b* are due to the fact that over the chosen frequency window the ratio of the maximum to the minimum values of the spectrum is large.

5. DECOMPOSITION IN THE PRESENCE OF NOISE

Composite signal decomposition is in general difficult in the presence of additive noise. The difficulty arises because a better resolution requires

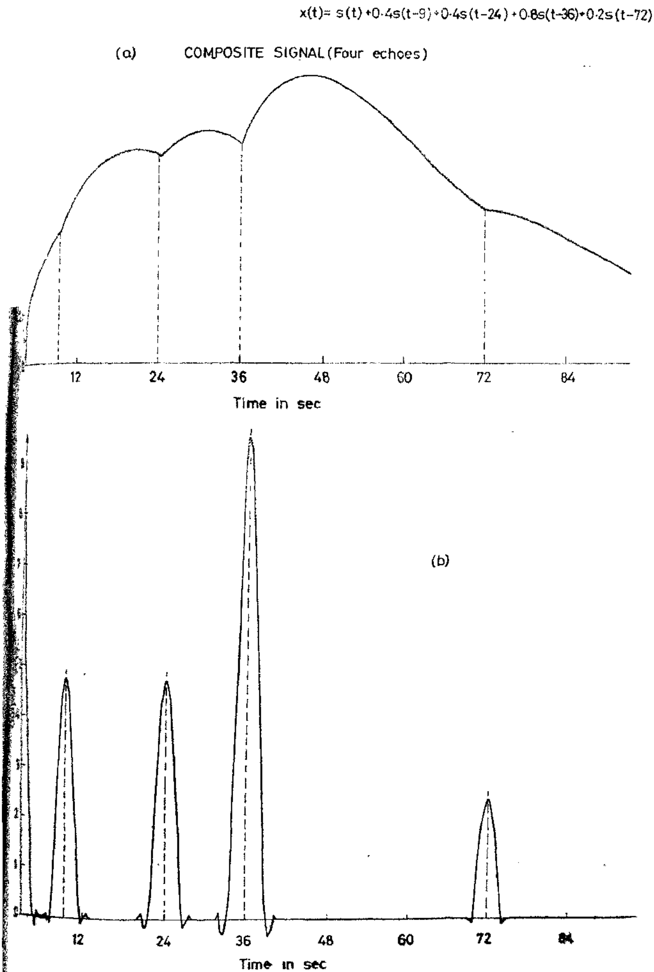


FIG. 6. Decomposition of a composite signal with four echoes. (*a*) The composite signal (*b*) The output.

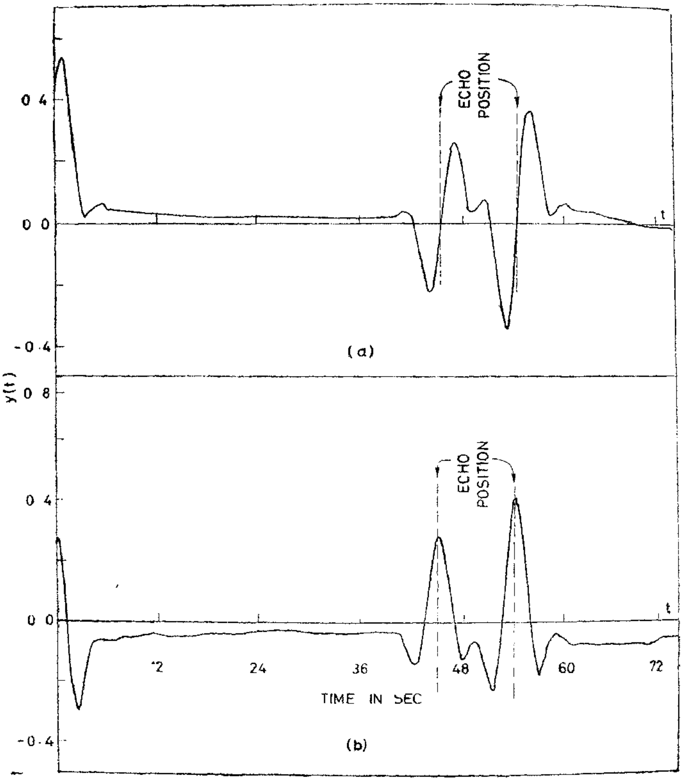


FIG. 7. The output for a composite signal with two echoes [Basic wavelet: $t^2 \exp(-at)$].

a larger bandwidth which in turn allows more noise. Our technique gives a poorer performance in the presence of noise compared to inverse filter [9]. But the disadvantage of the latter, namely, the requirement of the knowledge of the basic wavelet, is not present here.

A noisy composite signal $x(t)$ is generated by adding a band-limited white Gaussian noise [10] to a composite signal:

$$x(t) = s(t) + 0{\cdot}8s(t - 45) + 1{\cdot}2s(t - 54) + n(t) \quad (32)$$

The output obtained for a rms basic signal to rms noise ratio of 40 dB is shown in Fig 8. In the figure additional peaks due to noise are also

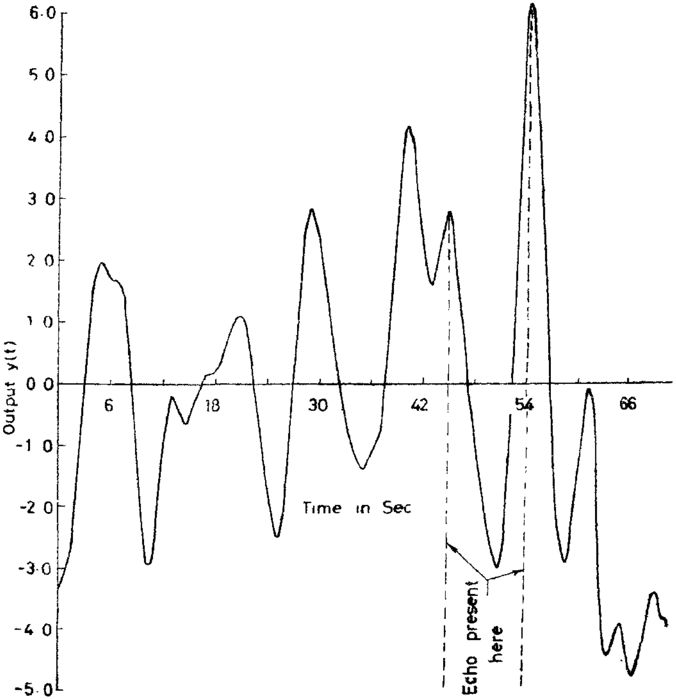


FIG. 8. The output for a noisy composite signal.

seen. When several sample functions of the same process are available the detection of echoes may be improved by taking ensemble average of the outputs of these sample functions.

6. VOICED SPEECH AS A COMPOSITE SIGNAL

As a practical example of a composite signal, voiced speech is considered in this section. The decomposition of the signal using the frequency window ing technique yields an estimate of the pitch initiation points and hence the pitch period.

Voiced speech can be modelled as the output of a quasistationary linear system excited by a quasi-periodic impulse train [11]. The responses of the individual impulses overlap in time making the resultant speech output a composite signal. Figure. 9 shows the log spectrum of a voiced segment of speech (Fig. 10 *a*) of 32 msec duration. Here the gross spectral features (shown by the dotted line) correspond to the impulse response of the system and the periodit fluctuations are due to the impulse train excitation [11]. The outputs obtained using the present technique with and without 90° phase rotation are shown in Figs. 10 *c* and 10 *d* respectively. The pitch initiation points (shown dotted) can be obtained from these figures. The output of a digital inverse filter [12] with 12 coefficients is also shown in Fig. 10 *b* for comparison. It is to be noted that the digital inverse filtering technique can be applied only if the basic signal can be derived as the output of an all pole model.

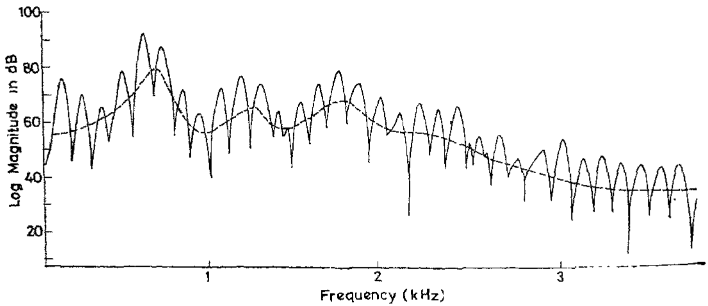


FIG. 9. Log spectrum of a voiced speech segment.

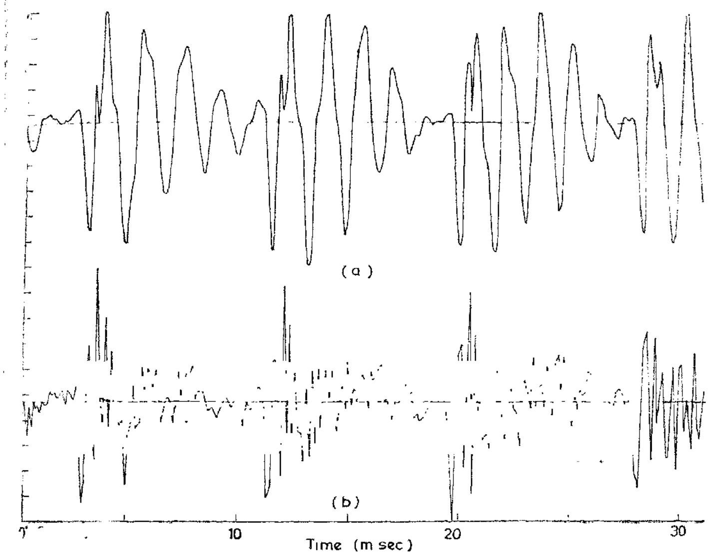


FIG. 10. Decomposition of voiced speech. (*a*) Voiced speech segment. (*b*) Output of the inverse filter.

A detailed analysis of voiced speech for epoch extraction is presented in [13].

7. CONCLUSIONS

An improved technique for composite signal decomposition has been suggested. This technique overcomes the limitation of the digital inverse filtering technique which requires a knowledge of the basic wavelet for its implementation. For basic signals possessing rational spectra the high frequency portion can be effectively used for the decomposition. Any desired resolution can be obtained in the absence of noise. The effect of the phase of the basic signal at high frequencies on the echo detection, and the choice of the sampling frequency and the window width for echo resolution have been illustrated through examples. There are no restrictions on the number

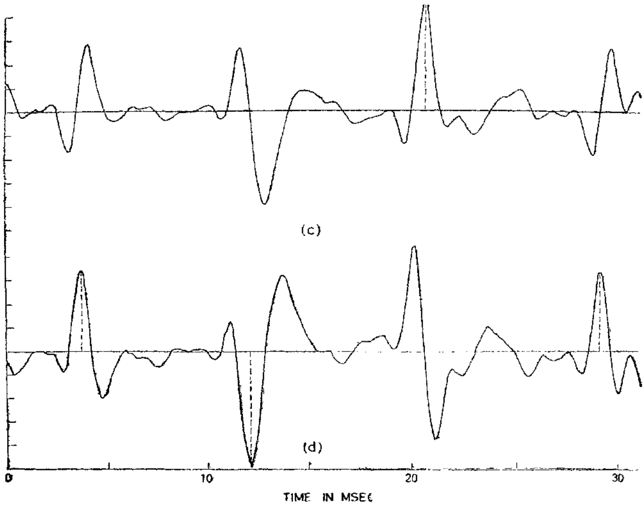


FIG. 10(*c*) Output obtained using the present technique. (*d*) Output after 90° phase rotation.

of echoes and their amplitudes as in the cepstrum techniques. The effect of additive noise on echo detection has also been discussed.

8. ACKNOWLEDGEMENT

The authors wish to thank Professor B. S. Ramakrishna for his constant encouragement and Dr. V. V. S. Sarma for useful discussions. The first author wishes to thank CSIR, India, for the award of Senior Research Fellowship.

REFERENCES

[1] Childers, D. G., Varga, R. S. and Perry, Jr. N. S. — Composite signal decomposition. *IEEE Trans. Audio Electroacoust*, 1970, Vol. AU-18, 471–477.

[2] Treitel, S. and Robinson, E. A. — The design of high resolution digital filters. *IEEE Trans Geo. Sc, Electronics*, 1966, Vol. GE-4, 25–38.

[3] Bogert, B. P., Healy, M. J. and Tukey, J. W. The frequency analysis of time series for echoes: Cepstrum, psudo-autocovariance, cross cepstrum and saphe cracking. In *Proc. Symp. Time Series Analysis*, M. Rosenblatt, Ed., New York: Wiley, 1963, 209–243.

[4] Oppenheim, A. V., Schafer, R. W. and Stockham, Jr., T. G. Nonlinear filtering of multiplied and convolved signals. *Proc. IEEE*, 1968, 56, 1264–1291.

[5] Kemrait, R. C. and Childers, D. G. Signal detection and extraction by cepstrum techniques. *IEEE Trans. Inform. Theory*, 1972, Vol IT–18, 745–759.

[6] Papoulis, A. .. Systems and Transforms with Applications in Optics, New York: McGraw-Hill, 1968, Ch. 7.

[7] Blackman, R. B. and Tukey, J. W. *The Measurement of Power Spectra*, New York, Dover, 1958

[8] Bergland, G. D. .. A guided tour of the fast Fourier transform. *IEEES pectrum*' 1969, 6, 41–52.

[9] Senmoto, S. and Childers, D. G. Signal resolution via digital inverse filtering. *IEEE Trans Aerosp. Electron. Syst.* 1972, Vol. AES-8, 633–640.

[10] Balas, L. F. .. Digital simulation of low pass Gaussian noise. *Simulation*, 1969, 110½–112.

[11] Schafer, R. W. .. A survey of digital speech processing techniques. *IEEE Trans. Audio Electroacoust*, 1972, Vol. AU–20, 28–35.

[12] Markel, J. D. .. Digital inverse filtering, a new tool for format trajectory estimation. *IEEE Trans. Audio Electroacoust.*, 1972, Vol AU–20, 129–137.

[13] Ananthapadmanabha, T. V. and Yegnanarayana B. Epoch extraction of voiced speech. *IEEE Trans. Acoust., Speech and Signal Processing*, 1975, Vol. ASSP–23, 562–570