IMPROVED SPECTRAL SUBTRACTION FOR SIGNAL RESTORATION*

H. Nawab, A. V. Oppenheim, J. S. Lim

Massachusetts Institute of Technology
Lincoln Laboratory and Research Laboratory of Electronics
Room 36-615, Cambridge, Massachusetts 02139

ABSTRACT

Major artifacts of the spectral subtraction signal restoration technique are the presence of harmonic patterns and blurring effects for images, and the presence of short tone bursts of varying frequency for speech signals. In this paper, we discuss several techniques developed to reduce these artifacts. Examples are shown to illustrate the performance of these techniques.

I. INTRODUCTION

A variety of techniques which have been proposed and investigated for restoring signals degraded by additive noise can be described within a general framework referred to as spectral subtraction. These techniques have generally been reported to exhibit favorable properties for high S/N ratios such as above 10db when applied to speech [1] and images [2]. However, for signals with relatively low S/N ratios, the processing introduces a number of artifacts and distortions that affect the quality of the restored signals. An example of typical artifacts associated with spectral subtraction is illustrated in Figure 1. Specifically, Figure 1a represents an original, undegraded image. Figure 1b corresponds to the result after applying spectral subtraction after having added 6 db of white noise. Evident in Figure 1b are two types of distortion. One is the presence of an apparently harmonic pattern which is particularly evident in the large high brightness region of the picture. Also noticeable in Figure 1b are "ripple" blurring effects near high contrast sharp edges such as between the clock and the background. Although generally not as severe as the distortion represented by the harmonic pattern, this is also a quality limiting artifact. Similar distortions are also apparent in applying spectral subtraction to speech. As has been observed in many such systems, spectral subtraction applied to speech degraded by additive noise typically results in the presence of objectionable short tone bursts of varying frequency. It was the desire to suppress such artifacts that motivated our investigation, which in turn resulted in the techniques reported in this paper.

Before describing the apparent sources of the artifacts and suppression techniques, we briefly summarize in the next section the standard short-time spectral subtraction procedure. Further details and derivations may be found in references [1] and [2].

II. SPECTRAL SUBTRACTION

In the basic spectral subtraction algorithm, a spectrum proportional to the average noise spectrum, which is assumed known, is subtracted from the spectrum of the noisy signal. More specifically, if $\hat{S}(\omega)$ is the Fourier transform of the estimated signal, $Y(\omega)$ the transform of the noisy signal, and $P_d(\omega)$ the noise spectrum, $\hat{S}(\omega)$ is specified by

$$|\hat{S}(\omega)|^2 = \begin{cases} |Y(\omega)|^2 - \alpha \cdot P_d(\omega) & \text{if } |Y(\omega)|^2 > \alpha \cdot P_d(\omega) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

and

$$\measuredangle\hat{S}(\omega) = \measuredangle Y(\omega) \quad (2)$$

In the case of both speech and images, $Y(\omega)$ and $\hat{S}(\omega)$ are taken as short-time spectra and the processing of

*This work was supported in part by the Advanced Research Projects Agency monitored by ONR under Contract N00014-75-C-0951-NR 049-328 at R.L.E. and in part by the Department of Air Force under Contract F19628-81-C-0002 at Lincoln Laboratory.

CH1610-5/81/0000-1105 $00.75 © 1981 IEEE

y(t) and re-synthesis of s(t) carried on a short-time basis because of the time-varying character of the signal. A procedure that has previously been applied both to speech [1] and images [2,3] divides the signal into equal length sections with a sequence of windows, $w_i(n)$. These windows may overlap in time, and are chosen subject to the condition that

$$\sum_i w_i(n) = 1 \qquad (3)$$

Spectral subtraction is performed on each windowed section. After the windowed sections of the signals are processed they are added together with the appropriate overlap to obtain the processed signal. Typical windows which, with correct overlap, can satisfy equation (3) include rectangular, triangular, Hanning and Hamming windows. In the case of two-dimensional signals, separable windows of these types have been used.

III. ARTIFACTS IN SPECTRAL SUBTRACTION

As indicated above, there are specific artifacts which are characteristic of processing based on spectral subtraction, specifically, the presence of a harmonic pattern and both ripple and blurring effects near sharp discontinuities. Based on empirical observations and study, we have concluded that the harmonic patterns or tones arise primarily because of a few, large amplitude narrowband peaks of noise energy remaining after spectral subtraction. Specifically, the actual short-time spectra of the noise process deviate randomly about the assumed a priori noise spectrum. Such deviations result in residues of noise energy remaining after spectral subtraction. For wideband random noise with values of α in the range generally used [1,2] in applying spectral subtraction, the residues tend to be dominated by a few narrowband peaks of relatively large amplitude. Of these peaks, the most undesirable are the ones at frequencies where there is little or no signal energy. These give rise to very discernable harmonic variations in the short-time signal sections. Since the noise component of the spectrum has independent deviations from section to section, the dominating frequency of the sinusoidal patterns also varies from section to section. This leads to short-time tone-bursts in speech and a harmonic pattern in images. In the next sections, specific techniques are proposed and illustrated for suppressing this artifact. One of the techniques, based on multi-window spectral smoothing, also reduces the rippling effect near large discontinuities. As would be expected, this particular artifact is due to the inherent blurring associated with signal characteristics which change rapidly in relation to the duration of the short-time window.

IV. ARTIFACT SUPPRESSION

In this section, we present three techniques for suppressing spectral subtraction artifacts. All of these techniques are successful in suppressing the harmonic pattern or tone noise, and the multi-window technique also helps to reduce the blurring artifacts.

A. Multi-Pass Spectral Subtraction

The implementation of this technique for suppressing artifacts consists of repeated application of the spectral subtraction procedure. Specifically, on each pass, the signal is sectioned and resynthesized after applying spectral subtraction. With a total of K passes, (α/K) is used in place of α in applying eq. (1) on each pass.

The key to this technique seems to lie in the post-subtraction synthesis at the end of each pass. Based on our experiments we have conjectured that this causes a spectral smoothing between overlapping sections. Thus, as noise energy is being subtracted, a smoothing process is taking place simultaneously between overlapping sections. Furthermore, as K increases, each short-time spectrum begins to affect the smoothing of spectra an increasing distance away. The idea of smoothing between the spectra of different sections is more directly explored in the Neighborhood Spectral Smoothing technique described next.

B. Neighborhood Smoothing

This approach is based on the assumption that the spectral magnitude of the degraded signal have large deviations between neighboring sections as compared to those of the undegraded signal. Thus, if the spectral magnitude of neighboring sections is averaged, then, in principle, the effect of the noise will be reduced. This, in effect, corresponds to time smoothing of the short-time spectral magnitude. This neighborhood smoothing can be carried out using either linear smoothing or median smoothing [4] techniques.

C. Multi-Window Spectral Smoothing

This technique capitalizes on the

flexibility in the design of the windowing structure used for short-time processing. In its most general form, the idea is to obtain signal estimates with different windowing structures (but the same amount of subtraction) and then to perform some type of spectral smoothing between them. In particular, we have found that using the same window shape but shifted locations in each pass is particularly successful. In our experiments we used short-time median smoothing between the different estimates to obtain the final result. The technique reduces not only the harmonic pattern artifacts but also the blurring, apparently because the different windows result in different blurring artifacts that tend to cancel each other in the smoothing process.

V. EXAMPLES

In this section, we illustrate the performance of each of the three different techniques discussed in section IV. All the examples shown in this section are based on additive white noise degradation and have been processed using overlapping separable triangular windows of size 32x32.

Figure 2 shows the result of multi-pass spectral subtraction. Specifically, the image in Figure 1b is the result of spectral subtraction with $\alpha=3$ when the S/N ratio of the degraded image is 6 db. The image in Figure 2 is the result of repeatedly applying spectral subtraction three times with $\alpha=1$. We have empirically observed that application of spectral subtraction more than three times introduces an undesirable amount of image data smoothing. Comparison of Figure 2 with Figure 1b shows the apparent reduction of the harmonic type of degradation.

The neighborhood smoothing technique has also been applied to the degraded image with S/N ratio of 6 db. The result obtained is very similar to the image shown in Figure 2. In general, we have empirically observed that the results of neighborhood smoothing are very similar to those obtained by the multi-pass spectral subtraction technique.

Figure 3 shows the result of the multi-window spectral smoothing technique applied to the degraded image with S/N ratio of 6 db. In generating the image in Figure 3, two different signal estimates have been obtained by applying spectral subtraction with $\alpha=3$ for each of the two windows which are shifted in location by 16x16 points relative to each other, and then the two signal estimates have been combined by choosing the smaller spectral magnitude estimate at each frequency. Comparison of Figure 3 with Figure 1b shows reduction of both the harmonic type of degradation and the blurring, "ripple" effects near large discontinuities.

The techniques discussed in Section IV may be combined to further reduce the artifacts of the spectral subtraction technique. For example, the multi-pass spectral subtraction technique and the multi-window spectral smoothing technique may be combined by using the multi-window spectral smoothing technique in each pass of the multi-pass spectral subtraction technique. Such an approach appears to reduce the artifacts of spectral subtraction more than using one technique alone. An example of such processing is illustrated in Figure 4 where the result has been obtained by applying both the multi-pass spectral subtraction and the multi-window spectral smoothing techniques to the degraded image with S/N ratio of 6 db.

Each of the three techniques has also been applied to speech signals with similar results. Specifically, each of the three techniques reduces the objectionable short tone bursts of varying frequency which are typical artifacts of spectral subtraction applied to speech signals.

In this paper, we have discussed three specific techniques to reduce the artifacts of the spectral subtraction signal restoration technique. It is clear from the above discussions that there are a variety of other heuristic approaches that may also be used to reduce the artifacts. Some of these other approaches are currently under investigation.

VI. REFERENCES

[1] J.S. Lim and A. V. Oppenheim, "Enhancement and Bandwidth Compression of Noisy Speech," Proceedings of the IEEE, vol. 67, no. 12, pp. 1586-1604, December 1979 (invited).

[2] J.S. Lim, "Image Restoration by Short Space Spectral Subtraction," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-28, no. 2, pp. 191-197, April 1980.

[3] H.J. Trussell and B.R. Hunt, "Sectioned Methods for Image Restoration," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-26, p. 157, April 1978.

[4] L.R. Rabiner and B. Gold, Theory and Application of Digital Signal Processing, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1975.

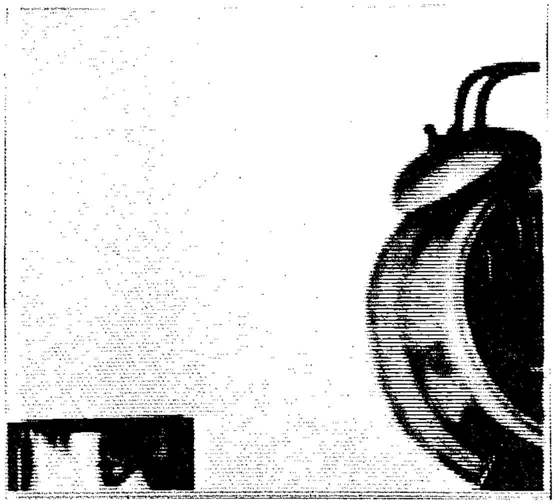

Figure 1a: Original Image

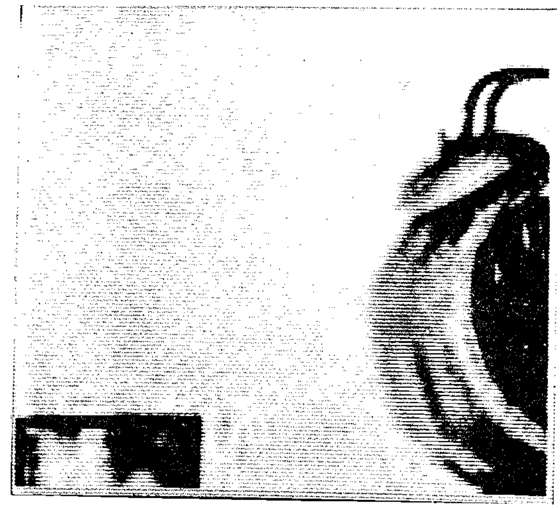

Figure 1b: Result of Spectral Subtraction, $\alpha=3$, at S/N=6 db.

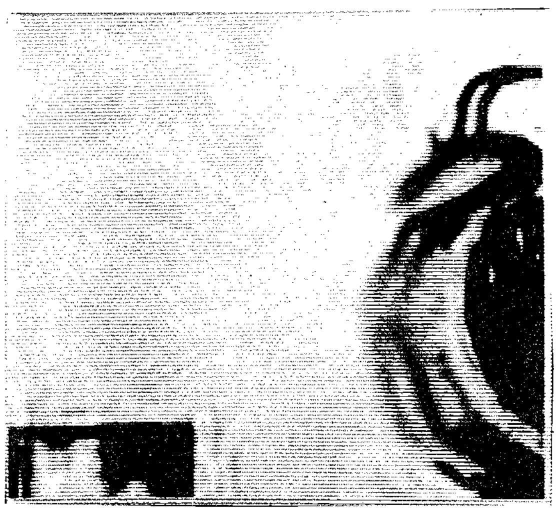

Figure 2: Result of Multi-pass Technique, $\alpha_1=\alpha_2=\alpha_3=1$, at S/N=6 db.

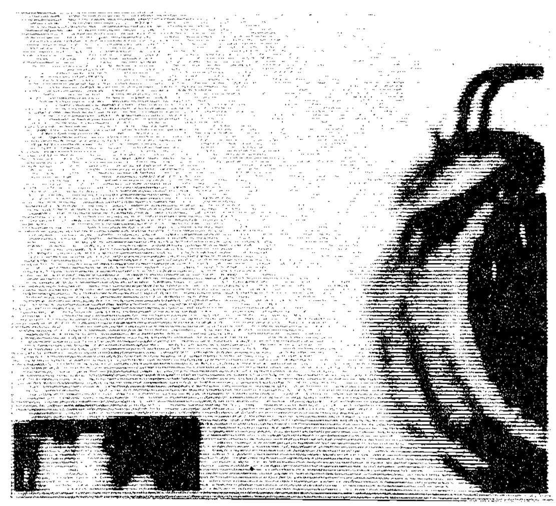

Figure 3: Result of Multi-window technique, $\alpha=3$, at S/N=6 db.

Figure 4: Result of combined Multi-pass and Multi-window techniques, $\alpha_1=\alpha_2=\alpha_3=1$, at S/N=6 db.