

- (a) voiced segment classified as unvoiced (VU)
- (b) unvoiced segment classified as voiced (UV)
- (c) gross errors (G) over 10% [2]
- (d) fine errors (F) from 2 to 10%
- (e) voiced segment where T_v is not evaluated (NT_v). The percentage of the segments where errors are detected with respect to the total number of voiced segments is shown in Table 1.

Conclusion: The proposed detector evaluates T_v very accurately from clean speech (without UV errors) and reliably from noisy and telephone speech due to the use of the $R_v(\tau)$, SAF and the preprocessing by means of adaptive thresholds. However, NT_v errors are present (0.2% for clean, 1.7% for noisy and 5.1% for phone speech) because the LSP and CEP detectors fail when the harmonic structure is distorted.

Acknowledgment: Work supported by the National Fund 'Scientific Research' under contract L-18-91 and contract ERB3510PL922096 of the Commission of the European Communities.

© IEE 1993
Electronics Letters Online No: 19931286

7 September 1993
B. Boyanov, T. Ivanov and S. Hadjitodorov (Institute of Biomedical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev Street, Block 105, 1113 Sofia, Bulgaria)

G. Chollet (URA 820, ENST, 46 rue Barrault, 75634 Paris, France)

References

- 1 HESS, W. (Ed.): 'Pitch determination of speech signals' (Springer Verlag, Berlin, 1983)
- 2 INDEFREY, H., HESS, W., and SEESER, G.: 'Design and evaluation of double-transform pitch determination algorithms with nonlinear distortion in the frequency domain-preliminary results'. Proc. Int. Conf. on ASSP, 1985, (Atlanta), pp. 137-140
- 3 MURTHY, H., and YEGANARAYANA, B.: 'Speech processing using group delay functions', *Signal Processing*, 1991, 22, pp. 259-267
- 4 CHILTON, E., and EVANS, B.: 'The spectral autocorrelation applied to the linear prediction residual of speech for robust pitch detection'. Proc. Int. Conf. on ASSP, 1988, (New York), pp. 358-361

Suppression of impulsive disturbances from audio signals

T. Kasparis and J. Lane

Indexing terms: Audio, Filters, Digital signal processing

A simple and effective method is proposed for suppressing impulsive noise from audio signals with minimal high frequency loss. Suppression is accomplished by median filtering the contaminated signal sections. A gating signal that enables local median filtering is generated by an algorithm that involves impulse detection and recursive median filtering.

Introduction: Impulse suppression in image processing is a mature field of study which has produced numerous techniques and algorithms for removing impulse type interference from digital image data [1]. Some of these methods can easily be adapted to digital audio processing and, in particular, for the restoration of damaged phonograph records. Many high quality hi-fi amplifiers incorporate a 'scratch' (lowpass) filter that rolls-off excess high frequencies where most of the energy of ticks and pops caused by damage on the surface of the record is contained. Unavoidably, music frequencies in the same band are also lost. More advanced filtering techniques that minimise this loss have been implemented either using elaborate hardware [2] or complex digital signal processing

algorithms [3]. We propose a simple, yet effective method for the suppression of such impulsive disturbances from audio signals with minimal high frequency loss. The algorithm can be implemented on any home computer equipped with inexpensive sound boards (known as 'sound-blaster'), and thus damaged favourite recordings can be processed at home. The proposed method can successfully discriminate impulsive data from non-impulsive, resulting in a restored signal with vastly improved audio quality.

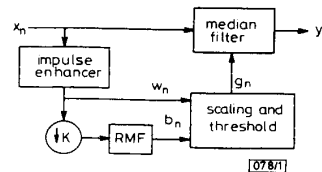


Fig. 1 Block diagram of impulse suppression system

Proposed approach: A block diagram of the proposed approach is shown in Fig. 1. The heart of the system is the median filter (MF) which has been popular in image processing applications for the removal of narrow impulses from otherwise smooth signals [1]. However, direct median filtering will lead to excessive filtering and distortion of the signal. To minimise distortion of scratch-free passages the filter is gated on and off by a signal provided by an impulse detector algorithm as in the following equation:

$$y_n = \begin{cases} m_n & \text{if } g_n = 1 \\ x_n & \text{if } g_n = 0 \end{cases} \quad (1)$$

where

$$m_n = \text{med}\{x_{n-k}, \dots, x_n, \dots, x_{n+k}\} \quad (2)$$

is the median value over a window of length $N = 2k + 1$ samples, x_n and y_n are the input and output sequences and g_n is a gating signal. The discrimination of impulses and generation of a good gating signal is the primary focus of this work. Solutions for g_n are generally easier to express in algorithmic rather than analytical form. For a more robust impulse detection in our approach we first use an impulse enhancement step by applying a second derivative operator on the input signal, i.e.

$$z_n = D^2\{x_n\} = x_{n-1} - 2x_n + x_{n+1} \quad (3)$$

The second derivative behaves like an edge detector and produces an output signal with abrupt transitions (such as a scratches) magnified whereas smooth regions are attenuated. Fig. 2 displays the resulting signal z_n . A typical scratch waveform consists of an initial

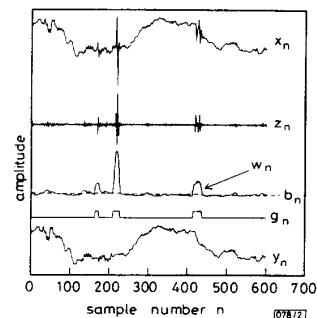


Fig. 2 Restoration example

Input x_n , second derivative z_n , impulse profiles w_n , reference floor b_n , gate g_n , and restored output y_n ; curves are plotted at different scales

pulse followed by decaying oscillations (see Fig. 2) due to the mechanical vibration of the stylus after encountering a scratch. To keep the MF active for the entire duration of the scratch, the gating pulse must be bracketing the scratch waveform. An impulse profile is obtained by applying a local RMS operator over a sliding window of length M on z_n , calculated by

$$w_n = \left[\frac{1}{M+1} \sum_{j=-M/2}^{M/2} z_{n+j}^2 \right]^{1/2} \quad (4)$$

The integration window length M is not critical but it should be large enough to include at least one cycle of the impulsive oscillations. Fig. 2 depicts w_n obtained with $M = 11$, where it is noticed that the impulse duration is clearly defined. The next step in developing g_n is to extract an impulse free background b_n from w_n which can be used as a reference floor for impulse detection. Again, we may call upon the MF for assistance in this task, but a recursive median filter (RMF) was found to be more appropriate because with a single application it can provide an impulse free reference floor, even when multiple impulses close to each other are present. The RMF is identical to the MF with the exception that previous medians are placed in the input buffer and used in the computation of subsequent medians [1]. Compared to eqn. 2 the RMF output rm_n is obtained from

$$rm_n = \text{med}\{rm_{n-k}, \dots, rm_{n-1}, x_n, x_{n+1}, \dots, x_{n+k}\} \quad (5)$$

Because impulse peak widths in w_n may exceed 30 samples, the RMF window length required to extract the background should be over 60 samples, but such large windows are computationally expensive. A simple method to reduce the window size and the computation without significant performance loss, is to decimate w_n and use a smaller RMF window. The final background extraction operator can be written as

$$b_n = \mathcal{RMF}\{w_n; L, K\} \quad (6)$$

where L is the RMF window length, K is the decimation ratio, and KL is the effective window length. The gating signal can be defined from the normalised absolute difference between w_n and b_n as

$$g_n = \begin{cases} 1 & \text{if } d_n > C \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where

$$d_n = \frac{|w_n - b_n|}{b_n} \quad (8)$$

Normalisation by b_n makes d_n insensitive to input conditions (amplitude or spectral content) so that a constant value for the threshold C can be used. Values of C from 2 to 3 seem to work well over a wide range of input conditions. The parameters controlling g_n are the averaging window length M , the decimation ratio K , the RMF window length L as well as the value of C . Fig. 2 shows the resulting b_n and g_n for $M = 11$, $K = 8$, $L = 15$ and $C = 2$.

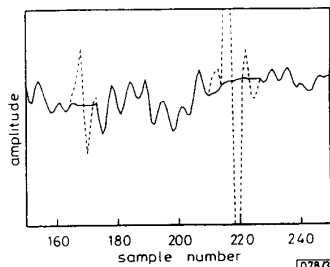


Fig. 3 Close-up of input (dashed) and output (solid) signals

Experimental results: Fig. 2 presents the restored signal using an MF window length of $N = 31$. Three impulsive disturbances have been detected and replaced by median values without degradation to the signal outside of those areas. Fig. 3 is an expanded view of the filtered regions. For short disturbances the listening result of the median replacement seemed to be quite adequate. For extremely wide impulses (50 samples or greater) it is necessary to increase the MF window size N considerably. The use of large MF windows has two drawbacks: the amount of computation increases drastically, and for short impulses the medians will be computed over a greater length of the signal and therefore may

not resemble signal values adjacent to the impulse. In addition, for wide disturbances the median replacement may not produce the best listening result. An adaptive MF window size with some additional post-processing would be the next step in improvement and the subject of our future work.

© IEE 1993

Electronics Letters Online No: 19931250

10 August 1993

T. Kasparis and J. Lane (Department of Electrical and Computer Engineering, University of Central Florida, Orlando, USA)

References

- 1 PITAS, I., and VENETSANOPOULOS, A.N.: 'Non-linear digital filters' (Kluwer Academic Publishers, 1990)
- 2 BURWEN, R.: 'Suppression of low-level impulse noise', *J. Audio Eng. Soc.*, 1979, **27**, (7/8), pp. 568-572
- 3 VASEGHI, S., and FRAYLING-CORK, R.: 'Restoration of old gramophone recordings', *J. Audio Eng. Soc.*, 1992, **40**, (10), pp. 791-801

Performance of novel synchronisation algorithm over AWGN channels

Y.O. Al-Jalili and J.M. Henriquez-Navarro

Indexing terms: Synchronisation, Random noise, Digital communication systems

Recently a novel symbol timing recovery technique that is implemented in the frequency domain has been presented. This technique is briefly described, where the performance of the time-error estimator under noisy conditions is assessed. Four different methods for time-error estimation are discussed, and the simulation results show that highly accurate estimates can be achieved.

Introduction: The role of synchronisation (symbol timing recovery) is to align the time position of each data pulse of the baseband signal with respect to the sampling instants at the receiver. Current literature describes only conventional time domain synchronisation techniques [1], which require one or two samples per symbol. Recently, a new synchronisation technique which is implemented in the frequency domain has been presented [2]. By taking this approach for symbol timing recovery, only one pair of spectral samples from a K -symbol message provides all the timing information required for synchronisation. Hence, this new method offers a simpler and more efficient timing synchronisation algorithm.

The complex envelope of the transmitted signal can be expressed as [2]

$$s(t, \epsilon) = \sum_{k=1}^K a_k g(t - kT - \epsilon T) \exp(j\theta) \quad (1)$$

where a_k are the symbols of a K -symbol message, θ is the carrier phase, ϵ is the timing offset between the transmitter and the reference clock at the receiver and T is the symbol period. After some manipulation, it can be shown [2] that the timing offset is given by

$$\epsilon = \frac{1}{2\pi} [\arg(S(f)) - \arg(S(f + 1/T))] \quad (2)$$

The above shows that the timing offset ϵ can be obtained from the phase difference of any two spectral samples separated by the symbol rate $1/T$. Also, once ϵ is known at the receiver, the signal needs to be time shifted by ϵT to ensure synchronisation. This timing correction can also be achieved in the frequency domain by modifying the phases of the frequency samples such that a zero time offset is obtained when the signal is converted to the time domain.

Fig. 1 shows a simple block diagram of a digital receiver incorporating a frequency domain approach to synchronisation. The received signal is transformed into the frequency domain by means of an FFT. The complex frequency samples are then converted