Models and Analysis of Vocal
Emissions for Biomedical Applications
(MAVEBA 1999)
Firenze, Italy, September 1-3, 1999

AN EXPRESS-RECONSTRUCTION OF DISTORTED SPEECH BY INVERSE FILTERING METHOD*

Helena B. Minervina

Institute of Engineering Cybernetics, National Academy of Sciences of Belarus,
Surganov Str. 6, 220012, Minsk, Belarus, alxkrot@newman.basnet.minsk.by

Abstract. An express-reconstruction of distorted speech signals based on a fast filtering algorithm for inverting linear convolution by sectioning method combined with effective real-valued split radix fast Fourier transform (FFT) algorithms are proposed.

Keywords: inverse filtering, fast algorithms, speech signal reconstruction

Introduction

When speech signals are difficult for understanding, for example, because of the pronunciation deficiencies (lisp, burr) the problem of distorted speech reconstruction is very important. This problem means the determing speech signal without the pronunciation deficiencies when the distored speech signal and impulse response for the filter which has led to mentioned distortion are known [1],[2]. So speech reconstruction problem can be reduced to determing such filter which is inverse of that which led to speech distortion [1],[2]. In turn, the filtering procedure is realized on the basis of inverting a linear convolution (LC). If the values of the impulse response are known, the distored signals can be restored with an absolute accuracy. If there is not this information, the choice of the impulse response is performed on the basis of the known information about the distortion character or by the empirical way. It is proposed to write the distorted speech signals on dictaphone and then to input vocal data in digital form on computer where the express-reconstruction is performed. The received speech signals, filtered from the distortions, are reproduced by speech synthesizer.

For the algorithm for inverting a LC by sectioning proposed in [3], it is assumed that the impulse response of the distorting effect is specified exactly, so that an absolutely accurate solution of the inverse filtering problem can be obtained. Unfortunately, because it does not use FFT algorithms, this algorithm is inefficient.

Our purpose here is to synthesize a fast algorithm for inverting a LC based on the sectioning method combined with the most efficient real-valued FFT algorithms [4]-[8], which will reduce the computational complexity of solving digital speech signal processing problems.

Calculating inverse linear convolutions by sectioning

We will consider the digital filtering of signals by inversion of an LC of the form

$$y_m = \sum_{n=0}^{m} h_{m-n} x_n,\ m = 0,1,\dots,N+M-2, \qquad (1)$$

where x_n is the incoming speech signal, y_m is the distorted signal and h_n is the impulse response which describes a linear filter with a finite impulse response . Before describing the algorithm of calculating the inverse LC let us consider briefly a way of calculating a P-long direct LC of the sequence $\{x_n\}$ with an impulse response $\{h_n\}$ of the length N based on the method of sectionalization of LC - overlap-add method (Fig.1).

First of all we shall divide $\{x_n\}$ into M-long sections (see Fig. 1a) and calculate a number of LC's of these sections with an impulse response:

$$y_m = \sum_{n=0}^{m} h_{m-n} x_n = \sum_{i=1}^{S}\left(\sum_{l=0}^{m} h_{m-l-M(i-1)} x_{M(i-1)+l}\right) = \sum_{i=1}^{S} y_m^{(i)}, \qquad (2)$$

where

$$y_m^{(i)} = \sum_{l=0}^{m} h_{m-l} x_l^{(i)}. \qquad (3)$$

* This work was partiality supported by International Science and Technology Center (ISCT) under the project B-95 and by the University of Firenze

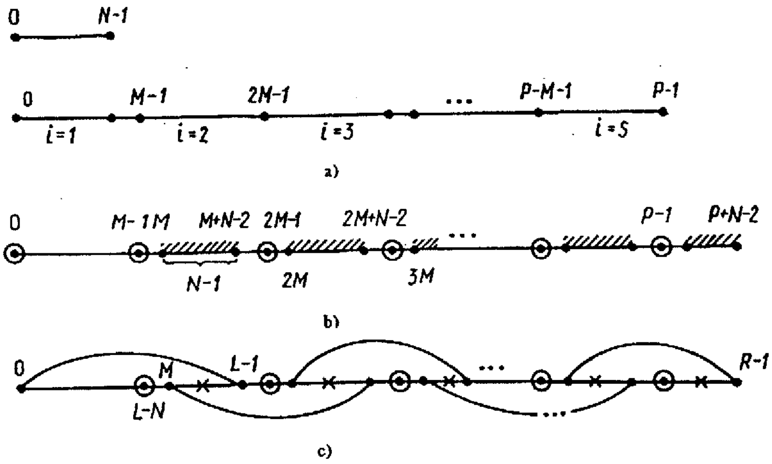

Fig.1. A scheme of calculating a P-long direct LC of sequence $\{x_n\}$ with an impulse response $\{h_n\}$ of the length N based on the overlap-add method

In (3) it is taken into account that the value h_{m-n} in (2) are zero for negative indices, i.e. $h_{m-l-M(i-1)}=0$ for $i\geq 2$, when $M\geq N-1$. According to (2), (3) the calculation of the LC (1) is reduced to calculating S LC's of sequences $\{x_l^{(i)}\} = \{x_{M(i-1)+l}\}$ with $\{h_n\}$ followed by the addition of the results. As is evident from Fig.1b, S sequences $y_m^{(i)}$ resulting from LC $x_l^{(i)}$ and h_n have the length $L=M+N-1$, the addition of values of two adjacent LC's being done in $N-1$ points.

Suppose we are solving the problem of inverting the LC calculated by means of relationships (2), (3) according to the method [3]. This means that it is necessary to determine sequence x_n , i.e. to restore S sections $x_l^{(i)}$ by inverting LC, knowing the result LC y_m (or S sections $y_m^{(i)}$) and sequence h_n. Since transformation (3), in the matrix form, is described by a low-triangular Toeplitz's matrix, the inverse transformation to (3) is described by the low-triangular Toeplitz's matrix as well, therefore it is clear that

$$x_l^{(i)} = \sum_{m=0}^{l} h_{l-m}^{(-1)} y_m^{(i)}, \quad l,m = 0,1,\dots,L-1, \qquad (4)$$

where $h_m^{(-1)}$ - the generating sequence of operator (4). As a rule, $h_m^{(-1)}$ can be calculated beforehand. Let the length L of the LC $y_m^{(i)}$ be known as well. Then, given the length N of the impulse response, one can find the section length $x_l^{(i)}$: $M=L-N+1$. For an effective solution of the inverse problem, deconvolution, one also can use the sectionalization method, for which refer to Fig.1c.

Consider the first L-long section of the LC $y_l^{(1)}$.It is clear from Fig.1c that $x_l^{(1)}$of length $L-N+1=M$ can be calculated by means of (4), i.e.

$$x_l^{(1)} = \sum_{m=0}^{l} h_{l-m}^{(-1)} y_m^{(1)}, l,m = 0,1,\dots,L-1. \qquad (5)$$

But $y_m^{(1)}$ has L points whereas for calculating $x_l^{(1)}$, according to (5), only the first M's are sufficient. However, in subsequent $N-1$ points the results of the LC $y_l^{(1)}$ and $y_l^{(2)}$ overlap, which causes the need for their correction for finding the true values of section $x_l^{(2)}$.

According to the algorithm of [3] the first M values $\{x_0^{(1)},\dots,x_{M-1}^{(1)}\}$ in (5) really form the searched for block $x_l^{(1)}$ of the sequence x_n, the last $N-1$ values $\{x_M^{(1)},\dots,x_{L-1}^{(1)}\} = \{\bar{x}_0^{(2)},\dots,\bar{x}_{N-2}^{(2)}\}$ being kept for the subsequent stage of correction:

$$\bar{y}_m^{(2)} = \sum_{l=0}^{m} h_{m-l}\bar{x}_l^{(2)}, m,l = 0,1,\dots,N-2. \qquad (6)$$

According to (5) and (6), for inverting the LC section $y_m^{(2)}$, a corrected sequence $y_m^{(2)}$ is formed as follows:

$$\{y_m^{(2)}\} = \{\bar{y}_0^{(2)},\dots,\bar{y}_{N-2}^{(2)}, y_{N-1}^{(2)},\dots,y_{L-1}^{(2)}\}, \qquad (7)$$

from which the deconvolution is then calculated:

$$x_l^{(2)} = \sum_{m=0}^{l} h_{l-m}^{(-1)} y_m^{(2)}, \quad m,l = 0,1,\dots,L-1. \qquad (8)$$

The procedure (6)-(8) enables the second section $x_l^{(2)}$ of the initial sequence x_n, i.e. the section corresponding to $i=2$ in Fig.1a, to be restored. Similary, the subsequent sections $x_l^{(i)}$ with $i=3$, ..., S can be restored as well. The number of sections S is expressed in terms of R - the length of the LC $\{y_m\}$, L- the section lengths for the LC and $N-1$ - length of the overlapping range of adjacent sections. Indeed, as is evident from Fig.1a-c, $P=R-(N-1)$ and $M=L-(N-1)$, therefore $S=P/M=(R-N+1)/(L-N+1)$ [3]. It should be noted that from $M \geq N-1$ and $L=M+N-1$ it follows that $L \geq 2(N-1)$.

Proceed now to the evaluation of the computational complexity of the inverse LC calculation algorithm considered. First of all it should be noted that the direct calculation of the inverse LC based on the low-trianqular Toeplitz's matrix of RxR dimension will require a number of real multiplications and additions equal to

$$M(R) =(1/2)R(R+1) = O((1/2)R^2) \quad (9a)$$

$$A(R)=(1/2)R(R-1)=O((1/2)R2) \quad (9b)$$

respectively.

The inverse LC calculation algorithm by means of sectionalizing [3], in restoring the first section $x_l^{(1)}$, will require according to (5) and to the expressions of (9a), (9b), $0.5L(L+1)$ multiplications and $0.5L(L-1)$ additions, in restoring the second section it will require, according to (6)-(8), $0.5N(N-1)+0.5L(L+1)$ multiplications and $0.5(N-1)(N-2)+0.5L(L-1)$ additions, etc. The total number of arithmetic operations for inverting the LC by means of the algorithm presented, with assumption $L=2N$ and $R=SL/2$, will come to:

$$M(R)= O((5/4)(L+2/5)R); \quad (10a)$$

$$A(R)= O((5/4)(L-2+8/(5L))R). \quad (10b)$$

From the comparison of (9a) and (9b) with (10a) and (10b) it follows that the algorithm of [3] will allow to decrease the number of arithmetic operations about $(2R)/(5L)$ times, i.e. the greater the relationship R/L, the more considerable the gain. However, it should be noted that $L \geq 2(N-1)$, where N - the length of an impulse transient response, and, even if $N \sim R/10$, i.e. $L \sim R/5$ then the gain will only be 2 times. It should be added that if, for instance, the given impulse response contains N=32 points, the number of arithmetic operations required for inverting the 320-point LC will be, according to (10a) and (10b), approximately $(5/4)(64+2/5)320=25760$ multiplications and $(5/4)(64-2)320=24800$ additions. There evaluations of the algorithm computational complexity in the case of calculating the 320-point inverse LC are large enough and approximately correspond to those of the 2048 - 4096 - point discrete Fourier transform (DFT), calculated on the basis of one of the most effective algorithms [5]-[8].

The computational complexity of the algorithm of [3] clearly needs to be improved way by using the real-valued split-radix fast Fourier transform algorithm (RVSRFFTA) [5]-[8] to calculate cyclic convolution (CC) and then LC [9]. In connection with the foregoing, let us turn to improving the algorithm of [3] by using an effective LC calculation algorithm by means of a RVSRFFTA. With that purpose a low-triangular Toeplitz's matrix H_L^{-1} of $L \times L$ dimension describing (4) must be presented in the form of the matrix of the corresponding LC. The LC matrix may be obtained from H_L^{-1} by prolonging elements $\{h_0^{(-1)},...,h_{L-1}^{(-1)}\}$ situated on the secondary diagonals to the length on the vertical equal to $2L-1$. As a result, a rectangular LC matrix of $(2L-1) \times L$ dimension tape - type in structure, is obtained. Further the LC matrix can be transformed into a CC square matrix of $2L \times 2L$ dimension by complementing each of sequences $\{h_l^{(-1)}\}$ and $\{y_m^{(i)}\}$ in (5), (8) etc. with zeros to the length $2L$. If $L=2^r$, then such a $2L$-point CC of sequences $\{h_0^{(-1)}, ... , h_{L-1}^{(-1)}, 0, ... ,0\}$ and $\{y_0^{(i)}, ... , y_{L-1}^{(i)}, 0, ... ,0\}$ can be effectively calculated by means of fast algorithms like the RVSRFFTA, though it is described by a matrix two times larger than the initial triangular Toeplitz's matrix of H_L^{-1} type.

As for the low-triangular Toeplitz's matrix H_{N-1} describing relationship (6), it may be transformed into a band LC matrix of N-point sequence $\{h_n\}$ with $(N-1)$-point sequence $\{\bar{x}_l^{(i)}\}$. However, the above mentioned LC matrix only contains the first $(N-1)$ columns and $2N-2$ rows of the H_L matrix, i.e. it is a rectangular one. Thus instead of the low-triangular Toeplitz's matrix of $Nx(N-1)$ dimension we shall consider the $(2N-2) \times (N-1)$ LC matrix. In its turn, we shall immerse the LC matrix into the $2N \times 2N$ CC square matrix, each of the sequences $\{h_n\}$ and $\{x_l^{(i)}\}$ having been complemented with zeros up to the length $2N$. That will result in obtaining the CC matrix two times larger than the initial Toeplitz's one, but which can be calculated by means of fast algorithms like the RVSRFFTA [5]-[8]. Indeed, choosing $N = 2^\alpha$, we find that for calculating the $2N$-point CC sequences

$\{h_0,...,h_{N-1},0,...,0\}$ and $\{\bar{x}_0^{(i)},...,\bar{x}_{N-2}^{(i)},0,...,0\}$ it is necessary to apply the $2^{\alpha+1}$-point real and Hermite SRFFT algorithms (RVSRFFTA and HSRFFTA, respectively) [5]-[8].

The algorithms based on calculating triangular Toeplitz matrices by means of CC of twice the dimensions can be written as follows [8], [9]:

(1) for $i=1$ we compute a $2L$-point CC of the form

$$x_l^{(1)} = \sum_{m=0}^{2L-1} \tilde{h}_{((l-m))}^{(-1)} \tilde{y}_m^{(1)}, l = 0,1,...,2L-1, \quad (11)$$

where $\{\tilde{h}_m^{(-1)}\} = \{h_0^{(-1)},...,h_{L-1}^{(-1)},0,...,0\}$, $\{\tilde{y}_m^{(-1)}\} = \{y_0^{(-1)},...,y_{L-1}^{(-1)},0,...,0\}$, are $2L$-point sequences, $((l-m))=(l-m)mod2L$;

(2) for $i \geq 2$ we form the $(N-1)$-point sequence $\{\bar{x}_l^{(i)}\} = \{x_{M+L}^{(i-1)}\}, l = 0,...,N-2$, and compute the CC of the form

$$\bar{y}_m^{(i)} = \sum_{m=0}^{2N-1} \bar{h}_{((m-l))} \tilde{\bar{x}}_l^{(i)}, l = 0,1,...,2N-1, \quad (12)$$

where $\{\bar{h}_n\} = \{h_0,...,h_{N-1},0,...,0\}$, $\{\tilde{\bar{x}}_l^{(i)}\} = \{\bar{x}_0^{(i)},...,\bar{x}_{N-2}^{(i)},0,...,0\}$, are $2N$-point sequences, $((m-l)) = (m-l) mod\, 2N$;

(3) for $i \geq 2$ we form the L-point sequence $\{y_m^{(i)}\} = \{\bar{y}_0^{(i)},...,\bar{y}_{N-2}^{(i)}, y_{N-1}^{(i)},...,y_{L-1}^{(i)}\}$ and then compute the $2L$-point CC :

$$x_l^{(i)} = \sum_{m=0}^{2L-1} \tilde{h}_{((l-m))}^{(-1)} \tilde{y}_m^{(i)}, l = 0,1,...,2L-1, \quad (13)$$

where $\{\tilde{h}_m^{(-1)}\} = \{h_0^{(-1)},...,h_{N-1}^{(-1)},0,...,0\}$, and $\{y_m^{(i)}\} = \{\bar{y}_0^{(i)},...,\bar{y}_{N-2}^{(i)}, y_{N-1}^{(i)},...,y_{L-1}^{(i)},...,0\}$ are $2L$-point sequences, $((l-m))=(l-m)mod2L$.

The number of arithmetic operations needed to calculate a real LC by means of a $2N$-point CC using RVSRFFTA and HSRFFTA was estimated in [8] and can be expressed by the following relationships:

$$M(N) = N(2 \log_2 N-1)+3; \quad (14a)$$

$$A(N) = N(6 \log 2N-1)+5. \quad (14b)$$

Substituting L for N in (14a) and (14b) we obtain the number of arithmetic operations for calculating the $2L$-point CC sequences $\{h_0^{(-1)}, h_{N-1}^{(-1)},0, ..., 0\}$ and $\{y_0^{(i)}, ..., y_{L-1}^{(i)}, 0, ..., 0\}$. Taking that into account, the total number of arithmetic operations required for inverting LC by the sectionalization method combined with the RVSRFFTA and the HSRFFTA is equal to:

$$M(R) = L(2\log_2 L-1)+3+ +(S-1)\,[N(2\log_2 N-1)+3]+ +(S-1)[L(2\log_2 L-1)+3]; \quad (15a)$$

$$A(R) = L(6\log_2 L-1)+5+ +(S-1)\,[N(6\log_2 N-1)+5]+ +(S-1)[L(6\log_2 L-1)+5]. \quad (15b)$$

Since $L \geq 2(N-1)$, to simplify calculations in (15a), (15b) we shall assume $L=2N=2^{\alpha+1}$, that is, $\gamma=\alpha+1$, also we shall take into consideration, as in deducing (10a) and (10b), that $R = SN = SL/2$. To sum up, relationships (15a) and (15b) will be as follows:

$$M(R) = 2R(2\log_2 L-1)+3S+ +(2\log_2 L-3)+3S-(1/2)L(2\log_2 L-3)-3 = =O((6\log_2 L+(12/L)-5)R), \quad (16a)$$

$$A(R) = 2R(6\log_2 L-1)+5S+ +R(6\log_2 L-7)+5S-(1/2)L(6\log_2 L-7)-5 = = O((18\log_2 L+(20/L)-9)R). \quad (16b)$$

From the comparison of (16a) and (16b) with (10a) and (10b) it is evident that the algorithm improved on the basis of the RVSRFFTA allows to save the number of multiplications more than $0.25(L+0.4)/(1.2\log_2 L-1)$ times, and the number of additions - $2.5(L/2-1)/9(2\log_2 L-1)$ times, as compared to the algorithm of [3]. For the case considered above, when $R \sim 5L$ and $L=2N=64$, the number of arithmetic operations required, according to (16a) and (16b), will be approximately $(36+12/64-5) \cdot 320 \approx 9920$ multiplications and $(108+20/64-9) \cdot 320 \approx 31680$ additions. From the comparison with the algorithm of [3] for this case, it is evident that the proposed algorithm reduces the number of multiplications essentially, *2.6* times, but increases the number of additions *1.28* times. Besides, the proposed algorithm decreases the total number of arithmetic operations by *9200*, as compared to the initial algorithm . As this saving is achieved due to the considerable reduction of multiplications which are more labour consuming than additions, the software or hardware realization of the case considered by means of the proposed algorithm brings about the advantage both in speed and in precision, as compared to the algorithm of [3].

Thus, despite the view held by the author [3], the cost of solving inverse filtering problems using inversion of an LC by sectioning [8]-[10] is reduced by using FFT algorithms(in this case, real-valued split-radix FFT algorithm [5]-[8], one of the best). As a result we comes to the matrix of CC two times larger in size then the initial Toeplitz one, but it can be calculated on the basis of effective fast algorithms.

The proposed fast inverse convolution algorithm for reconstructing distorted signals by sectioning the inverse convolution and using the RVSRFFTA was programmed. A computer experiment was performed in which the distorted speech signals of different sizes were reconstructed, using impulse responses of various lengths. It was shown that speech signals was constructed 1.6 times faster on average by the proposed algorithm than with the approach based on the algorithm of [3]. The advantage is especially pronounced when the sequence is fairly long.

References

1. Dudgeon D.E. and Mersereau R.M. Multidimensional Digital Signal Processing. Prentice-Hall, Englewood Cliffs, NJ, 1984.
2. Pratt W. Digital Image Processing, vol.2, Wiley, NY, 1978.
3. Titkov B.V. "An algorithm for inverting discrete convolution by sectioning". Zh.Vychisl.Mat.i Mat.Fiz. vol.32, pp.199-207, 1992.
4. Krot A.M. "The method of eigentransforms in different fields for computing cyclic convolutions and discrete Fourier transforms", U.S.S.R. Comput. Math and Math. Phys., vol.29, No.3, pp.23-34,1989 (a translation of Zh. Vychisl. Mat. i Mat. Fiz. vol.29, No.5, pp.675-692, 1989).
5. Krot A.M. and Minervina Ye.B. "A fast Fourier transform algorithm for real and Hermitian-symmetrical sequences". Soviet. J. Comm. Tech. Electron., vol.34, No.12, pp.122-129, 1989 (a translaton of Radiotekh.i Elektronika, vol. 34, No.2, pp. 369-376, 1989.
6. Krot A.M. and Minervina Ye.B "Synthesizing symmetric-Hermitian sequences" FFT algorithms on a split base for real and Radioelectronics and Commun.Systems, 1989, vol.32., no.12, pp.10-15 (a translation of Radioelektronika, vol. 32, No.12, pp. 12-17, 1989).
7. Krot A.M. Discrete Models of Dynamic Systems Based on Polynomial Algebra Nauka i Tekh., Minsk, 1990, (in Russian).
8. Krot A.M. and Minervina Ye.B. Fast Algorithms and Programs for Digital Spectral Processing of Signals and Images, Nauka i Tekh., Minsk, 1995 (in Russian).
9. Kas'ko A.T., Krot A.M. and Minervina Ye.B.A fast algorithm for calculating the inverse convolution for signal and image restoration" Comput.Math. and Math. Phys., vol.36, No. 2, pp. 269-277, 1996.
10. Krot A.M. and Minervina H.B. Minimal multiplicative complexity and fast restoration algorithm of digital signals and images. Proc. SPIE, "Signal Processing, Sensor", April 13-15, 1998, Orlando, Florida, USA, vol. 3374, pp. 426-435.