A NOVEL APPROACH TO CALCULATION OF LINE SPECTRAL FREQUENCIES BASED ON INTER-FRAME ORDERING PROPERTY

V. Semenov

Delta S.P.E.
27 Academician Krymsky St., Kiev, 03142 Ukraine
e-mail: semenov@deltacorp.net

ABSTRACT

A problem of calculation of line spectral frequencies (LSF) is considered. The investigation of mutual LSF location on adjacent quasi-stationary frames is performed. It was found that in majority of cases LSF inter-frame ordering property takes place. On this basis a new approach to LSF calculation is proposed. The LSF localization is mainly reduced to verification of inter-frame ordering property. The computational expenses are reduced in 3.4 times in comparison with widely used Kabal's method. Besides, the maximum number of operations is lower than the minimum expenses of accelerated Kabal's method. The method was implemented on fixed-point DSP and showed stable performance.

1. INTRODUCTION

The majority of modern speech processing methods are based on the autoregressive (AR) model where the vocal tract is modelled by all-pole filter with coefficients $a_k, k = 1,..., p$. However in real speech coding applications AR coefficients are typically transformed to LSF $\omega_k, k = 1,..., p$ to decrease spectral quantization errors. The calculation of LSF is connected with root finding procedures [1, 2] which are undesired for a majority of computational devices, especially for fixed-point DSPs since they usually cause unpredictable delays and error accumulation. In paper [3] the authors proposed a new method of LSF calculation based on developed universal method of transcendental equations' solution. Although the proposed approach had several essential advantages over existing ones, it did not use physical LSF features which, obviously, must contain a good reserve for the improvement of the method's efficiency. Many other algorithms also neglect to include information on LSF distribution when doing their calculations.

That is why the aim of this investigation is the construction of a computationally efficient LSF calculation method which optimally exploits the features of their time distribution. In the second section the investigation of mutual LSF location on adjacent frames is performed and the condition of LSF inter-frame ordering is introduced. In the third section a new algorithm of LSF calculation based on inter-frame ordering property is proposed. In the experimental section the proposed method is compared with method [3], Kabal-Ramachandran's method and its accelerated modification.

2. INVESTIGATION OF MUTUAL LSF LOCATION ON ADJACENT FRAMES

Taking into account the connection of LSF with vocal tract resonances [4], one can assume that LSF on adjacent quasi-stationary frames must not differ too much. To verify this assumption, consider a speech signal pronounced by a male speaker and digitized with sampling frequency of 8000 Hz. The spectrogram of this signal is represented at fig. 1.

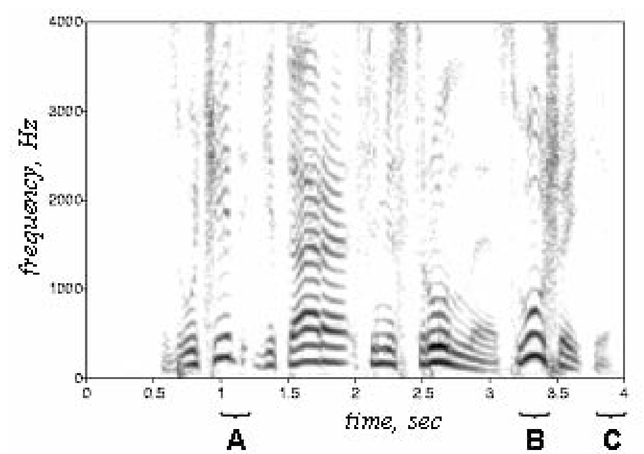

Figure 1. The spectrogram of test speech signal.

To illustrate mutual LSF location on adjacent frames, consider three parts of this signal marked at fig.1 as A, B, C. Time plots of these segments and corresponding LSF plots are presented at figures 2, 3, and 4 respectively. The calculation of LSF was performed on 20 ms frames for the order of AR model $p = 10$.

Fig. 2 shows the situation corresponding to the end of vowel "e" and the beginning of consonant "t". It can be seen that though each LSF vary in quite a wide range, almost in all situations LSF with number i lies between LSF of previous frame with numbers $i-1$ and $i+1$. The only exclusion is the third LSF of seventh frame which is slightly higher (on 4 Hz) than a forth LSF of a previous frame. As is seen from the time plot, this situation corresponds to the beginning of a new sound (consonant “t”).

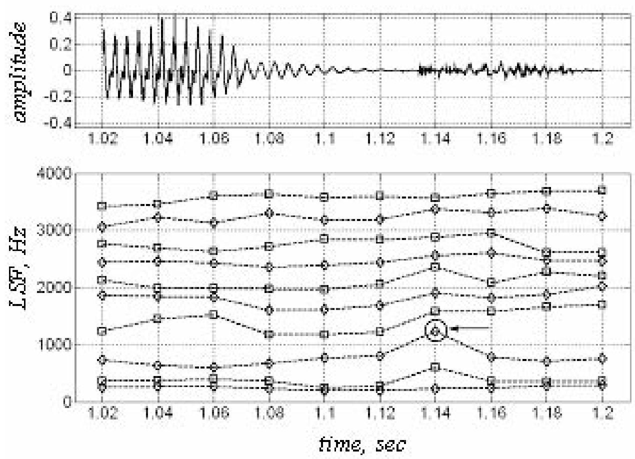

Fig.2. Example of LSF distribution on boundary of two sounds.

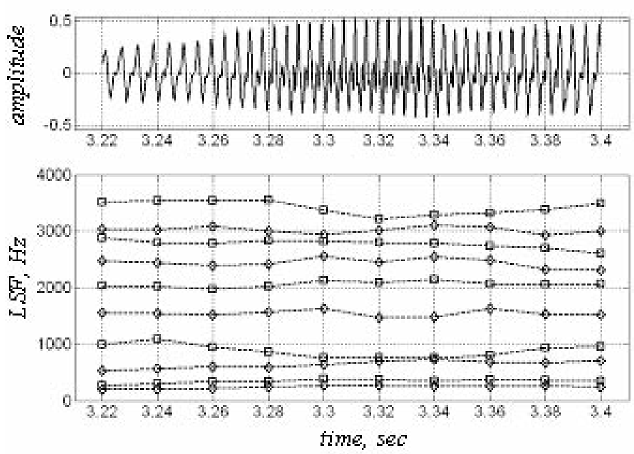

Fig. 3. Example of LSF distribution during the continuous vowel.

Fig. 3 shows LSF distribution during a continuous sound “l”. On all frames of this speech fragment LSF ω_i lies between the previous frame LSF with numbers i and $i-1$, i.e. the following inequality is satisfied:

$$\omega_{i-1}^{(n-1)} < \omega_i^{(n)} < \omega_{i+1}^{(n-1)},\ i = 2,..., p\ , \tag{1}$$

where the upper indices denote the numbers of frames.

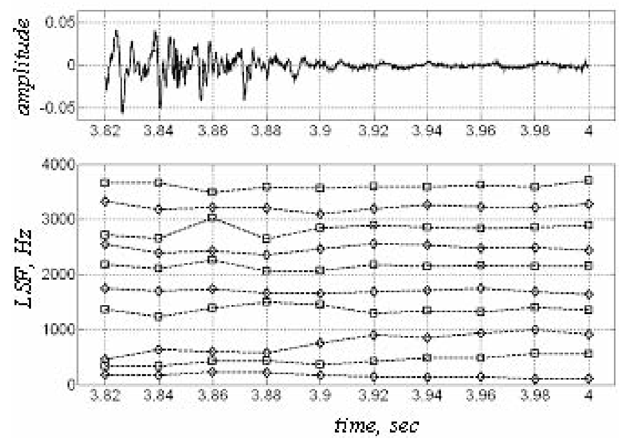

Fig.4. Example of LSF distribution on boundary of vowel and pause.

At fig. 4 one can see the end of the vowel and a following pause filled by background noise. During the decaying of the vowel, the LSF spans a relatively wide range. Nevertheless, LSF with number i constantly lies between LSF of previous frame with numbers $i-1$ and $i+1$. On the time fragment containing background noise LSF have a uniform almost time-invariant distribution and, obviously, satisfy to inequality (1). The fact that the pauses usually take not less than 40-50 % of speech duration, additionally enforces the assumption that condition (1) is met in a majority of practical situations.

To verify this hypothesis we tested the condition (1) on a speech database of 8 speakers with a total duration of 8 min (the sampling frequency was equal to 8000 Hz). At the preliminary stage we deleted pauses between phrases, which allowed to the estimate the lowest probability of condition (1). We considered AR model orders from 8 to 20. The results of investigation are summarized in Table 1. It shows percentages of cases when condition (1) did not fail (n_0), or it failed for one, two or three LSF (n_1, n_2 and n_3 respectively).

Table 1. Percentage of cases when condition (1) is met (n_0) and when it is not met (n_1, n_2 and n_3).

	p=8	*p=10*	*p=12*	*p=14*	*p=16*	*P=18*	*p=20*
n_0	96.7	95.6	94.3	93.1	91.9	90.1	88.8
n_1	2.8	3.6	4.6	5.4	6.3	7.6	8.1
n_2	0.3	0.5	0.7	0.9	1.1	1.4	1.9
n_3	0.1	0.1	0.1	0.2	0.3	0.4	0.5

As follows from Table 1 inequality (1) is true for a majority of cases. Corresponding percentage lies from 88.85 % at $p = 20$ to 96.78 % at $p = 8$. It can therefore be stated that LSF localization problems are primarily reduced to verification of **inter-frame ordering property** (1).

3. ALGORITHM FOR LSF COMPUTATION

According to the above mentioned investigation of LSF mutual placing, the following algorithm for LSF computation is proposed. It is assumed that LSF cosines from previous frame $\{x_1^{(n-1)}, x_2^{(n-1)}, ..., x_{p-1}^{(n-1)}, x_p^{(n-1)}\}$ are already known. At first, equations for the cosines of LSF are formed [1]:

$$\sum_{k=0}^{M} c_{m,k} x^{M-k} = 0 \qquad (m = 1,2) . \qquad (2)$$

Verification of condition (1) can be easily performed by comparison of polynomial function signs at points $\{-1, x_2^{(n-1)}, x_4^{(n-1)}, ..., x_{p-2}^{(n-1)}, x_p^{(n-1)}, 1\}$ when the roots of the first equation (2) are computed or on the grid $\{x_1^{(n)}, x_3^{(n)}, ..., x_{p-3}^{(n)}, x_{p-1}^{(n)}\}$ for the second equation (2) (it is assumed that AR model order is an even number). This is illustrated by fig. 5, where the plot of the first polynomial (2) is depicted and the cosines of $\{x_i^{(n-1)}\}$ are shown. As can be seen, considered function changes sign in points $\{-1, x_2^{(n-1)}, x_4^{(n-1)}, ..., x_{p-2}^{(n-1)}, x_p^{(n-1)}, 1\}$ and, thus, necessary number of polynomial function calls is just $(p+4) = 14$ (which is a strong contrast to 111 calls for the Kabal-Ramachandran's method [1]).

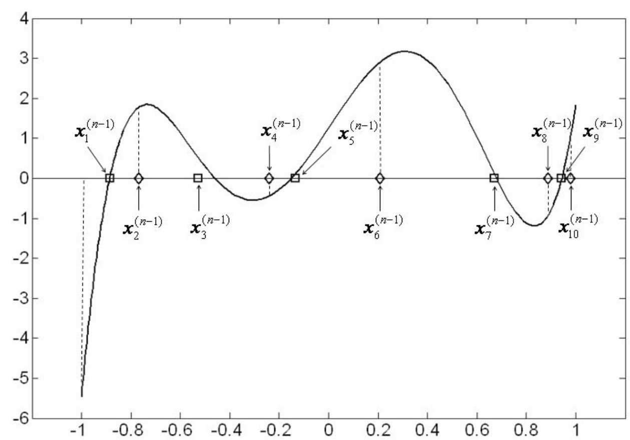

Fig. 5. Example of LSF cosines' localization.

If for some LSF with number i condition (1) is not met (i.e. polynomial function has the same signs in points $[x_{i-1}^{(n-1)}, x_{i+1}^{(n-1)}]$), the verification of root presence is verified by algorithm [3].

After the localization of LSF cosines, their exact values are determined by Newton's method. At this point it was found that for the computation of cosine $x_k^{(n)}$, a good initial approximation is provided by the corresponding value of previous frame $x_k^{(n-1)}$. From fig. 5 one can see that odd LSF cosines of the previous frame are a good approximations for the odd cosines of the current frame. Finally, LSF $\{\omega_k\}$ are obtained from roots $\{x_k\}$ by transformation $\omega = \arccos(x)$.

4. EXPERIMENTAL RESULTS

The effectiveness of the proposed method was verified for different AR model orders p: 8,10,…,20. For this purpose we used a base of five-minute records of four male and two female speakers digitized with $f_s = 8000$ Hz. LSF were computed for every 20 ms of these signals.

Table 2 shows average computational expenses (in Mflops), corresponding to method described in [3] (in brackets) and to the method proposed in this paper. For the objectivity of comparison initialization of the Newton's method was taken the same as in [3]. The condition for algorithm stop was $|f(x)| < 10^{-6}$.

Table 2. Computational expenses (Mflops) of method (3) (in brackets) and proposed method.

AR model order						
8	*10*	*12*	*14*	*16*	*18*	*20*
0.03 (0.05)	0.04 (0.08)	0.06 (0.12)	0.08 (0.18)	0.10 (0.24)	0.13 (0.31)	0.16 (0.38)

From table 2 it can be seen that proposed algorithm provides a reduction of computational expenses from 1.7 times (for $p = 8$) to 2.4 times (for $p = 10$). This is explained by more perfect LSF localization procedure proposed in this paper.

Since the test speech signals were characterized by a variety of speakers and were almost free of pauses, provided computational characteristics of proposed method can be considered as corresponding upper bounds. During a real work of speech coding devices expenses must be lower.

Now let's make a comparison of proposed algorithm with Kabal-Ramachandran's method [1], since it is most widely used in speech processing applications. Since Kabal-Ramachandran's method exploits bisection method of root refinement, we also considered its "accelerated" version corresponding to root refinement by Newton's method.

Table 3 shows average, minimal and maximum numbers of operations necessary for the computation of ten LSF at one frame by Kabal-Ramachandran's method, its accelerated version, method [3] and the method proposed in this paper. The convergence criterion based on the uncertainty of root position was used: $|x_k - x_{k-1}| < 10^{-3}$, where x_{k-1}, x_k - approximate root values obtained at successive iterations.

Table 3. Comparison of LSF computation methods.

	Average number of operations	Minimum number of operations	Maximum number of operations
K.-R. method	2150	2150	2150
Acc. K.-R. method	1498	1491	1557
Method [3]	1309	1066	1936
Proposed method	629	428	1428

From Table 3 follows that proposed algorithm provides reduction of computational expenses in comparison with Kabal-Ramachandran's method in 3.42 times. The gain over accelerated Kabal-Ramachandran's algorithm is equal to 2.38 times.

One of the main characteristics of method [3] was that its peak computational expenses were lower than (time-invariant) expenses of Kabal-Ramachandran's method. However, the method proposed in this paper has a stronger property: its **maximum** number of operations (1428) is lower than that for Kabal-Ramachandran's method (2150), but is also lower than the **minimum** number of operations for the accelerated Kabal-Ramachandran's method. This fact additionally suggests the advantage of application of proposed method in real-time systems. The method was implemented into a fixed-point ADSP2191 vocoder (2.4 kbps). During an hour of continuous testing with different speakers the algorithm failed only in 2 of 144000 cases due to insufficient accuracy of polynomial function evaluation in fixed-point 16-bit arithmetic.

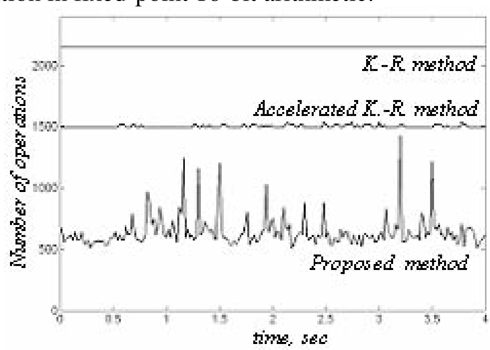

Fig. 6. Example of computational expenses distribution for different LSF calculation methods.

Fig. 6 shows the numbers of operations of different methods for the calculation of LSF of test speech signal, the spectrogram of which was shown at fig.1 This clearly shows the advantage of the proposed method in the context of the number of operations. The minimum computational expenses take place in pauses filled by stationary background noise, while the peaks in the distribution of operations take place at abrupt transitions from one sound to another.

5. CONCLUSIONS

In given paper a simple and effective method of LSF calculation was proposed. The main task was to answer a question whether LSF of previous time frame can be effectively used for the localization of LSF at current time frame. It was found that in majority of situations (from 88.85 % of cases at AR model order $p = 20$ to 96.78 % of cases at $p = 8$) a property of "inter-frame ordering" (1) takes place. This means that LSF localization task can be mainly reduced to the verification whether LSF cosines satisfy to inter-frame ordering property. For the localization of LSF which do not satisfy this condition, universal algorithm of roots' localization [3] is used.

During experimental verification of the proposed method for different speakers and AR model orders the following results were obtained.

1. The resulting computational savings in comparison with method [3], which did not use features of LSF distribution, are from 1.8 to 2.5 times (at different AR model orders).

2. For AR model order $p = 10$ the proposed approach reduces computational expenses in comparison with most widely used Kabal-Ramachandran's method in 3.4 times. Also there is a 2.4 times gain over the accelerated combination of Kabal-Ramachandran's method with Newton's method.

3. It was found that the **maximum** number of operations of the proposed method is lower not only than the time-invariant expenses of Kabal-Ramachandran's method, but is also lower than the **minimum** number of operations of the accelerated combination of Kabal-Ramachandran's method and Newton's method.

These facts tell about the advantage of application of proposed method in real-time systems. The method was realized on fixed-point 16-bit DSP and showed stable work.

6. REFERENCES

[1] P. Kabal and R. P. Ramachandran, "The computation of line spectral frequencies using Chebyshev polynomials", IEEE Trans. ASSP, vol. 34, pp. 1419-1425, 1986.

[2] J. Rothweiler, "A rootfinding algorithm on line spectral frequencies", Proc. ICASSP-1999, Phoenix, vol.2, pp. 661-664, 1999.

[3] V. Semenov, A. Kalyuzhny and A. Kovtonyuk, "Efficient calculation of line spectral frequencies based on new method for solution of transcendental equations", Proc. ICASSP-2003, Hong Kong, Vol. 2, pp. 457-460, 2003.

[4] K.K. Paliwal, B.S. Atal, "Efficient vector quantization of LPC parameters at 24 bits/frame", IEEE Trans. Speech Audio Processing, vol. 1, p. 3–14, 1993.