The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 2021

Volume 16, Pages 153-161

IConTES 2021: International Conference on Technology, Engineering and Science

Optimization Parameter of the 1P Keys Interpolation Kernel Implemented in the Correlation Algorithm for Estimating the Fundamental Frequency of the Speech Signal

Zoran MILIVOJEVIC
Academy of Applied Technical and Preschool Studies

Bojan PRLINCEVIC
Kosovo and Metohija Academy of Applied Studies

Natasa SAVIC
Academy of Applied Technical and Preschool Studies

Abstract: The first part of this paper describes an algorithm for estimating the fundamental frequency F_0 of a speech signal, using an autocorrelation algorithm. After that, it was shown that, due to the discrete structure of the autocorrelation function, the accuracy of the fundamental frequency estimate largely depends on the sampling period T_S. Then, in order to increase the accuracy of the estimation, an interpolation of the correlation function is performed. Interpolation is performed using a one parameter (1P) Keys interpolation kernel. The second part of the paper presents an experiment in which the optimization of the 1P Keys kernel parameter was performed. The experiment was performed on test Sine and Speech signals, in the conditions of ambient disturbances (N8 Babble noise, SNR = 5 to -10 dB). MSE was used as a measure of the accuracy of the fundamental frequency estimate. Kernel parameter optimization was performed on the basis of the MSE minimum. The results are presented graphically and tabularly. Finally, a comparative analysis of the results was performed. Based on the comparative analysis, the window function, in which the smallest estimation error was achieved for all ambient noise conditions, was chosen.

Keywords: Fundamental frequency, Parametric convolution, Convolution kernel, Speech signal.

Introduction

For several decades, digital speech signal processing has been actual in multimedia systems. Algorithms have been developed for: a) speaker recognition, b) semantic speech recognition, c) speaker health analysis, d) language recognition, e) speech extraction from the background noise, f) dereverberations, d) echo suppression, h) speech signal quality corrections, etc. (Qiu et al., 2000). Increasing intelligibility requires increasing the quality of the speech signal (Rao et al., 2000). One way to increase speech intelligibility is by applying an algorithm to reduce dissonant frequencies (Milivojevic et al., 2009). Many speech signal processing algorithms require an estimate of the fundamental frequency of the speech, as well as the dominant harmonics. Algorithms for estimating the fundamental frequency are based on the processing of the speech signal in: a) the time domain, b) the spectral domain and c) the cepstrum domain (Kacha et al., 2004).

Estimation of the fundamental frequency in the spectral domain is based on the application of the DFT and peak peaks in the spectrum (Pang et al., 2000). The spectrum, calculated using DFT, is discrete. The accuracy of the fundamental frequency estimate is directly dependent on the length of the DFT. However, in reality the

- This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 Unported License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

- Selection and peer-review under responsibility of the Organizing Committee of the Conference

© 2021 Published by ISRES Publishing: www.isres.org

fundamental frequency can be located between the spectral components (Milivojevic et al., 2013). In this case, the estimation is performed using interpolations. Interpolation with convolutional kernel is often applied (Keys, 1981). Third-order interpolation kernel, due to their numerically low complexity, allow high estimation speeds (Meijering, 2003).

The time domain estimation algorithms are based on the analysis of time waveforms. If the waveform of the signal is periodic, then the period can be observed and F_0 can be estimated on its basis. The TD algorithms intensively uses autocorrelation functions (Rabiner et al., 1978) to detect the pitch period. In (De Cheveigné et al., 2002), an algorithm, called the YIN algorithm, where the estimation is performed using the autocorrelation function (ACF), was proposed.

The ACF of a discrete periodic signal is a discrete and a periodic function (Milivojevic et al., 2017). The ACF components have a time interval witch equal to the sampling time periods, T_S, of the signal. Determining the period of the discrete signal implies locating the first, dominant peak, at the ACF. Then the fundamental frequency is equal to the reciprocal of the time shift of the peaks in relation to the beginning of the ACF. Here, the problem of estimating of the fundamental frequency arises when the actual dominant peak of ACF is not located on an integer product of T_S, but somewhere between two adjacent components with the highest energy. In this case, estimation of the position is performed by selecting the position of the peak and, thus, a significant estimation error F_0 occurs. The estimation error reduction can be done by applying an interpolation algorithm (Milivojevic et al., 2021).

This paper presents an algorithm for estimating the fundamental frequency in the time domain. The algorithm is based on the application of autocorrelation function and interpolation using the interpolation 1P Keys kernel (Milivojevic et al., 2017). After that, an experiment, in which the effect of window functions on the accuracy of the fundamental frequency estimate was analyzed, is described. First, the Sine test signal and the Speech test signal are processed using window functions. Then, effect of changing of the kernel parameter, α, on the estimation error, was analyzed. Mean square error, MSE, is used as the estimation error. By minimizing the estimate error MSE, the optimal value of the kernel parameter, α_{opt}, is calculated. Analyzes were performed for Test signals with superimposed acoustic interference (N8 Babble noise). The Test signals with SNR = -5 to 5 dB was formed. The results are presented in tables and graphs. Finally, a comparative analysis was performed. Based on the results of the comparative analysis, a window function suitable for implementation for real-time operation was selected.

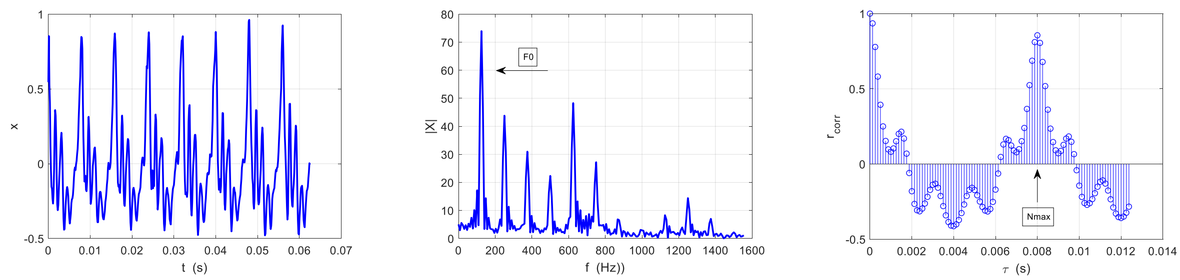
Estimation F_0 Using Autocorrelation Function

Correlation is a measure of the similarity of two signals. It is defined as the similarity of one signal at time k and another at time $k + m$. In this case, the correlation function is cross correlation. The autocorrelation function is a measure of the similarity of the same signal at time k and at time $k + m$. For a discrete signal $x(n)$, whose length is N, an autocorrelation function is defined by (Milivojevic et al., 2021):

$$r_{corr}(m) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)\cdot x(n+m),\ \mathrm{m} = 0,1,2,... \tag{1}$$

On Figure 1.a the Speech test signal $x(n)$ is shown. Its spectrum is shown in Figure 1.b and autocorrelation function r_{corr} is shown on Figure 1.c. The waveform of x can be complex and unsuitable for determining periods. The autocorrelation function r_{corr} is more suitable for calculating the signals period. On Figure 1.b the position of fundamental frequency is denoted by F_0. On Figure 1.c the position of the maximum of the autocorrelation function is denoted by N_{max}. The signal period is $T_0 = N_{max} * T_S$, where T_S is the sampling frequency of the time continuous signal $x(t)$. The fundamental frequency of the signal $x(n)$ is $F_0 = 1 / T_0 = 1 / (N_{max} * T_S)$. Determining the position of the maximum component of the autocorrelation function is realized by the Peak-Picking algorithm.

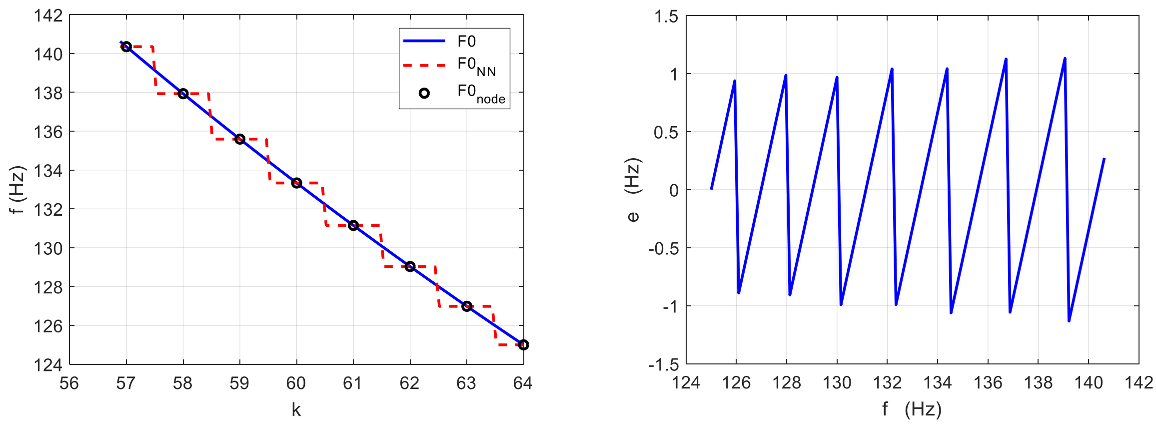
After determining the autocorrelation function and locating the peak, it is possible to accurately estimate the fundamental frequency only for signals whose fundamental frequency is $F_0 = 1 / (k * T_S)$ for $k = 1, 2, 3, ...$ For signals, whose fundamental frequency F_0 is in the interval $(k + 1) * T_S) < F_0 < 1 / (k * T_S)$, the estimation is performed by rounding and, thus, causes an estimation error. The calculation of F_0 is realized using the Nearest Neighbor method.



a) b) c)

Figure 1. Speech test signal: a) time form, b) amplitude characteristic, c) autocorrelation function.

On Figure 2.a shows the actual F_0 of the signal x sampled with F_S = 8 kHz in the range (125 - 140.625) Hz, which corresponds to the components k = 57 - 64 of the autocorrelation function (symbol '-'). Using the Peak-Picking algorithm, the values of F_0 for node k = 57 - 64 (symbol 'o') were calculated. The estimated values of the fundamental frequency, F_{0NN}, determined by applying the Nearest Neighbor method in the interval k = 57 - 64 are shown by the symbol '--'. The estimation error, $e(f)$, is shown on Figure 2.b.



a) b)

Figure 2. a) Fundamental frequency F_0 trajectory between (57-64) autocorrelation components, value $F_{0\mathrm{NN}}$ estimated by rounding and value $F_{0\mathrm{node}}$ in nodes k, and b) estimation error e caused by Nearest Neighbor method.

Reducing the fundamental frequency estimation error, $e(f)$, can be done by applying interpolation. By interpolation, based on the position of the maximum value of the autocorrelation function, N_{max}, a series of $m = \{N_{\max} -1, N_{\max}, N_{\max} + 1, N_{\max} + 2\}$ is formed and the position of the maximum is interpolated and, based on it, the fundamental frequency is calculated.

Algorithm for Fundamental Frequency Estimation

The algorithm for estimation of the fundamental frequency (*Estimation_algorithm*) is applied over the i-th block $\mathbf{x}_\mathrm{I}$ of the signal $\mathbf{x}$, and consists of the following steps (Milivojevic et al., 2021):

Input: $\mathbf{x}_\mathrm{I}$ - frame of discrete signal $\mathbf{x}$. N - frame length. F_0 -real fundamental frequency. T_S - sampling period.
Output: F_e - estimated fundamental frequency. MSE - estimation error.

Step 1: The $\mathbf{x}_\mathrm{I}$ signal is modified by the window function w:

$$\mathbf{x}_{IW} = \mathbf{x}_I * \mathbf{w} , \quad (2)$$

Step 2: Determine the autocorrelation function r_X
Step 3: Using the Peak-Picking algorithm, the position of the maximum of the autocorrelation function, N_{max}, is calculated.

Step 4: By applying parametric interpolation with the interpolation kernel r_{PCC}, the continuous function R_X is determined.
Step 5: By differentiating the function R_X and equalizing with zero, the position of the maximum between the two n_{max} samples is determined. The real position of the maximum is $N_M = N_{max} + n_{max}$.
Step 6: The estimated fundamental frequency is:

$$F_e = 1/\left(\left(N_{\max} + n_{\max}\right) \cdot T_S\right), \tag{3}$$

Step 7: The mean square error of the fundamental frequency estimate is:

$$MSE = \overline{(F_0 - F_e)^2}. \tag{4}$$

In the continuation of this paper, an experiment is described in the framework of which the efficiency of the fundamental frequency estimation algorithm at Sine and Speech signal was tested.

Experimental Results and Comparative Analysis

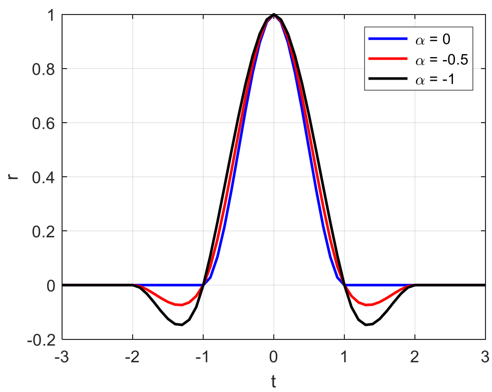
Experiment

An experiment with the aim of determining the efficiency of the fundamental frequency estimation by a correlation algorithm, was performed. Increasing the accuracy of the assessment was achieved by applying window functions for processing the Test signal in the time domain. After that, a further increase in accuracy was achieved with the application of the interpolation algorithms. MSE was used as a measure to analyze the accuracy of the assessment. Using the Test Algorithm, which is explained later in this section, the trajectories of the MSE estimation error were calculated. By minimizing the estimate error, the optimal parameters of the parametric cubic interpolation kernel are determined. Kernel parameters are specified for some standard, time-symmetric, window functions. All analyzes were performed in the presence of background acoustic interference.

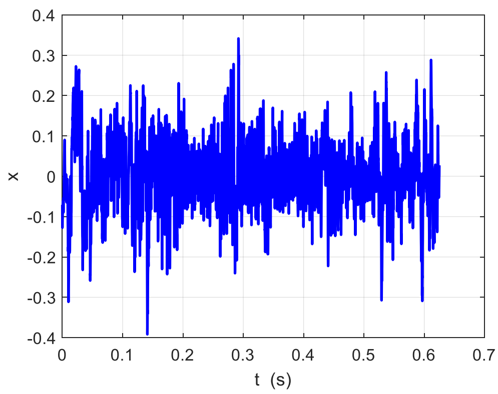
Test signals were created with F_S = 8 kHz and with windows with length N = 256, which assures the analysis of frame that last 32 ms. The results presented further in this paper relate to F_0 = 125-140.625 Hz (frequencies between the eighth and ninth DFT components). Number of frequencies in the specified range for which the estimation is done is M = 100. The Sine test signal is with K=10 harmonics. All further analyzes will relate to a) Hamming, b) Hanning, c) Blackman, d) Rectangular, e) Kaiser and f) Triangular window. N8 Babble noise (SNR = {5, 2.5, 0, -2.5, -5, -7.5, -10} dB) was used as acoustic interference (Figure 3.b). The 1P Keys interpolation kernel (Keys, 1981) was used, which was defined as

$$r(f) = \begin{cases} (\alpha+2)|f|^3 - (\alpha+3)|f|^2 + 1, & |f| \le 1, \\ \alpha|f|^3 - 5\alpha|f|^2 + 8\alpha|f| - 4\alpha, & 1 < |f| \le 2, \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where α is the kernel parameter. The time form of the 1P Keys kernel, for some values of the parameter α, is shown in Figure 3.a.



a)



b)

Figure 3. a) 1P Keys window, b) Babble noise N8.

Test Algorithm

The Test algorithm was implemented in the following steps:

Input: x_I - frame of discrete signal x. N - frame length. F_{0_real} - real fundamental frequency. T_S - sampling period. ($\alpha_{\min}$, $\alpha_{\max}$) parameter limits, $\Delta\alpha$ - step, w - window function,
Output: $\mathrm{MSE}_{\min}$. α_{opt}.

FOR $\alpha = \alpha_{\min}: \Delta\alpha: \alpha_{\max}$
FOR $\mathrm{SNR} = \mathrm{SNR}_{\min}$: Step: $\mathrm{SNR}_{\max}$
Step 1: Create a Test Signal with SNR:

$$x_{Ib} = x_I + k \cdot x_b \,, \tag{6}$$

Step 3: Estimation of the fundamental frequency (algorithm described earlier in the previous section)

$[F_e, \mathrm{MSE_s}] = \textit{Estimation\_algorithm}\ (x_{Ib}, \alpha, N, w, T_s)$

END

$$MSE(\alpha) = \overline{MSE_s} \,, \tag{7}$$

END
Step 4: Determining minimum of the estimate error:

$$MSE_{\min} = \min(MSE). \tag{8}$$

Step 5: Determining the optimal kernel parameter:

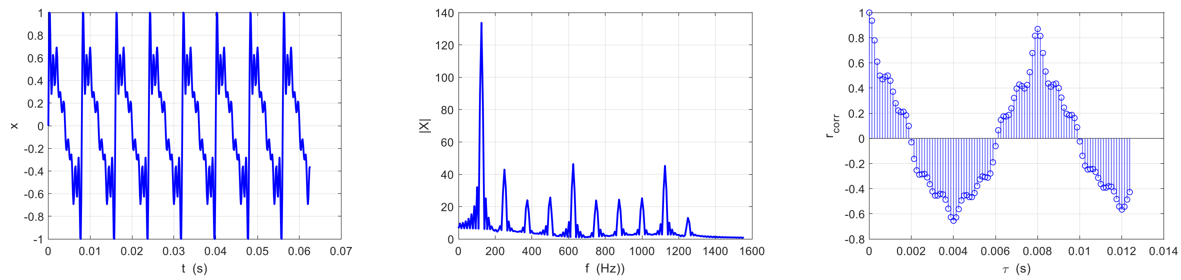
$$\alpha_{opt} = \arg\min_{\alpha}(MSE), \tag{9}$$

Test Signal

PCC algorithm for the fundamental frequency estimation will be applied to: a) Sine test signal, and b) Speech test signal. Simulation Sine signal for testing of PCC algorithm is defined in (Pang et al., 2000):

$$s(t) = \sum_{i=1}^{K}\sum_{g=0}^{M} a_i \sin\left(2\pi i\left(F_o + g\frac{F_s}{NM}\right)t + \theta_i\right), \tag{10}$$

where F_0 is fundamental frequency, □$_i$ and a_i are phase and amplitude of the i-th harmonic, respectively, K is the number of harmonics, and M is the number of points between the two samples in spectrum. The Sine test signal is shown on Figure 4.: a) time form, b) amplitude characteristic and c) autocorrelation function. The real Speech test signal is obtained by recording of a speaker in the real acoustic ambient. The Speech test signal is shown on Figure 1.: a) time form, b) amplitude characteristic and c) autocorrelation function.



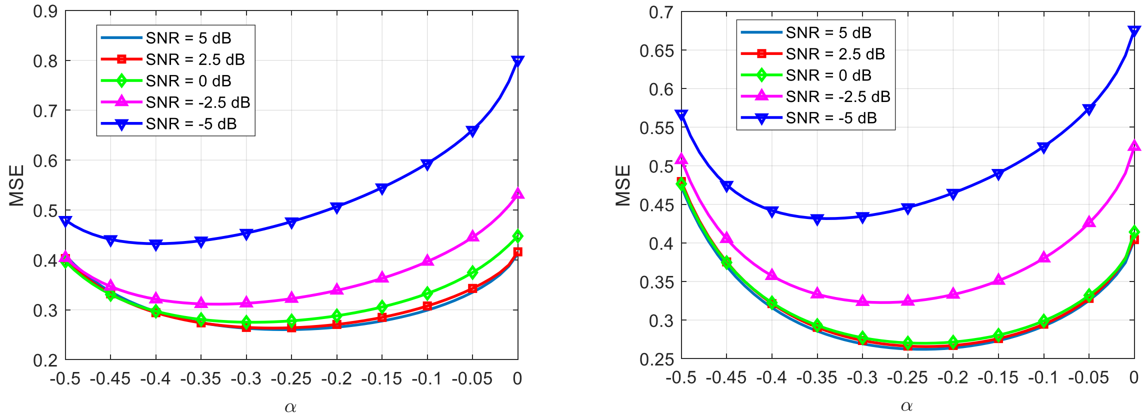
a) b) c)

Figure 4. Sine test signal: a) time form, b) amplitude characteristic, c) autocorrelation function.

Results

Using the Test algorithm, described in the previous section, MSE trajectories for all tested window functions were calculated. The MSE trajectories for the rectangular window are shown on: a) Figure 5.a (Sine test signal) and b) Figure 5.b (Speech test signal). The minimum values of MSE and the optimal kernel parameters, for the Sine test signal and the Speech test signal, are shown in Table 1. The trajectories of the minimum MSE for SNR

= -5 to 5 dB are shown on: a) Figure 6.a (Hamming), a) Figure 6.b (Hann), a) Figure 6.c (Blackman), a) Figure 6.d (Rectangular), a) Figure 6.e (Kaiser) and a) Figure 6.f (Triangular).



a) b)

Figure 5. MSE for Rectangular window: a) Sine test signal, b) Speech test signal.

Table 1. Optimal parameters and estimates error for testing windows

Window	SNR (dB)	Sine test signal α_{opt}	Sine test signal MSE	Sine test signal MSE_{NN}	Speech test signal α_{opt}	Speech test signal MSE	Speech test signal MSE_{NN}
Hamming	5	-0.1250	0.2864	4.4057	-0.1250	0.3133	4.9651
	2.5	-0.1250	0.2908	4.2141	-0.1250	0.3056	4.6656
	0	-0.1250	0.3222	3.8738	-0.1250	0.3355	4.2339
	-2.5	-0.1250	0.4444	3.4930	-0.1250	0.4677	3.6422
	-5	-1.3400	0.5433	2.6137	-1.3400	0.8381	2.8477
	-7.5	-0.5300	11.1040	13.7504	-10.250	7.6133	7.0663
	-10	-10.250	77.2945	69.7670	-0.5300	44.6659	56.9970
			$\overline{MSE}_{sine}$ = 12.8979	$\overline{MSE}_{sin\_NN}$ = 14.5882		$\overline{MSE}_{speech}$ = 7.7913	$\overline{MSE}_{sp\_NN}$ = 12.0597
Hann	5	-0.1250	0.2907	4.4057	-0.1250	0.3226	4.9627
	2.5	-0.1250	0.2961	4.2141	-0.1250	0.3196	4.6747
	0	-0.1250	0.3322	3.8738	-0.1250	0.3662	4.3238
	-2.5	-0.1250	0.4557	3.4182	-0.1250	0.5101	3.6668
	-5	-1.3400	0.6204	2.5644	-1.3400	1.7657	3.5394
	-7.5	-0.5300	12.6568	15.4588	-10.250	7.8694	7.2064
	-10	-0.5300	51.3111	70.2882	-0.5300	44.8402	59.4322
			$\overline{MSE}_{sine}$ = 9.4233	$\overline{MSE}_{sin\_NN}$ = 14.8890		$\overline{MSE}_{speech}$ = 7.9991	$\overline{MSE}_{sp\_NN}$ = 12.5437
Blackman	5	-0.1250	0.2883	4.3133	-0.1250	0.3206	4.8683
	2.5	-0.1250	0.2932	4.0458	-0.1250	0.3380	4.5908
	0	-0.1250	0.3688	3.9673	-0.1250	0.3935	4.1463
	-2.5	-0.5300	0.5326	3.3815	-0.1250	0.5787	3.6113
	-5	-1.3400	0.6519	3.8943	-0.9350	3.1323	3.9570
	-7.5	-0.5300	16.1659	19.8547	-10.250	16.3176	14.9718
	-10	-0.5300	48.0547	71.8709	-0.5300	37.3017	58.8938
			$\overline{MSE}_{sine}$ = 9.4793	$\overline{MSE}_{sin\_NN}$ = 15.9040		$\overline{MSE}_{speech}$ = 8.3403	$\overline{MSE}_{sp\_NN}$ = 13.5770
Rectangular	5	-0.1250	0.2871	4.5004	-0.1250	0.2815	4.5866
	2.5	-0.1250	0.2950	4.4080	-0.1250	0.2840	4.5866
	0	-0.1250	0.3177	4.2426	-0.1250	0.2883	4.5090
	-2.5	-0.1250	0.3784	3.9317	-0.1250	0.3640	4.3959
	-5	-0.5300	0.5379	3.3755	-0.1250	0.5064	3.9829
	-7.5	-1.3400	0.8136	2.8734	-1.3400	3.8830	5.4915
	-10	-1.3400	1.9469	2.8648	-0.5300	6.4847	7.3310

			$\overline{MSE}_{sine}$ = 0.6538	$\overline{MSE}_{sin\_NN}$ = 3.7423		$\overline{MSE}_{speech}$ = 1.7274	$\overline{MSE}_{sp\_NN}$ = 4.9834
	5	-0.1250	0.2912	4.4997	-0.1250	0.3203	5.0716
	2.5	-0.1250	0.2936	4.3133	-0.1250	0.3210	4.8728
	0	-0.1250	0.3294	4.1271	-0.1250	0.3167	4.3127
Kaiser	-2.5	-0.1250	0.4128	3.5589	-0.1250	0.4489	3.8147
	-5	-1.3400	0.5505	2.8017	-1.3400	0.7412	2.9532
	-7.5	-0.5300	7.9480	10.5746	-0.5300	3.2700	6.2004
	-10	-0.5300	61.2133	68.1588	-0.5300	45.0961	52.9036
			$\overline{MSE}_{sine}$ = 10.1484	$\overline{MSE}_{sin\_NN}$ = 14.0049		$\overline{MSE}_{speech}$ = 7.2163	$\overline{MSE}_{sp\_NN}$ = 11.4470
	5	-0.1250	0.2926	4.4997	-0.1250	0.3069	4.9651
	2.5	-0.1250	0.2945	4.3133	-0.1250	0.3397	4.9627
	0	-0.1250	0.3206	4.0458	-0.1250	0.3369	4.4027
Triangular	-2.5	-0.1250	0.4157	3.4930	-0.1250	0.4516	3.8106
	-5	-1.3400	0.5510	2.7381	-1.3400	0.7809	3.0417
	-7.5	-0.5300	8.8860	11.3704	-10.250	6.7040	6.4624
	-10	-0.5300	64.5430	68.4134	-0.5300	41.9750	53.4811
			$\overline{MSE}_{sine}$ = 10.7576	$\overline{MSE}_{sin\_NN}$ = 14.1248		$\overline{MSE}_{speech}$ = 7.2707	$\overline{MSE}_{sp\_NN}$ = 11.5895

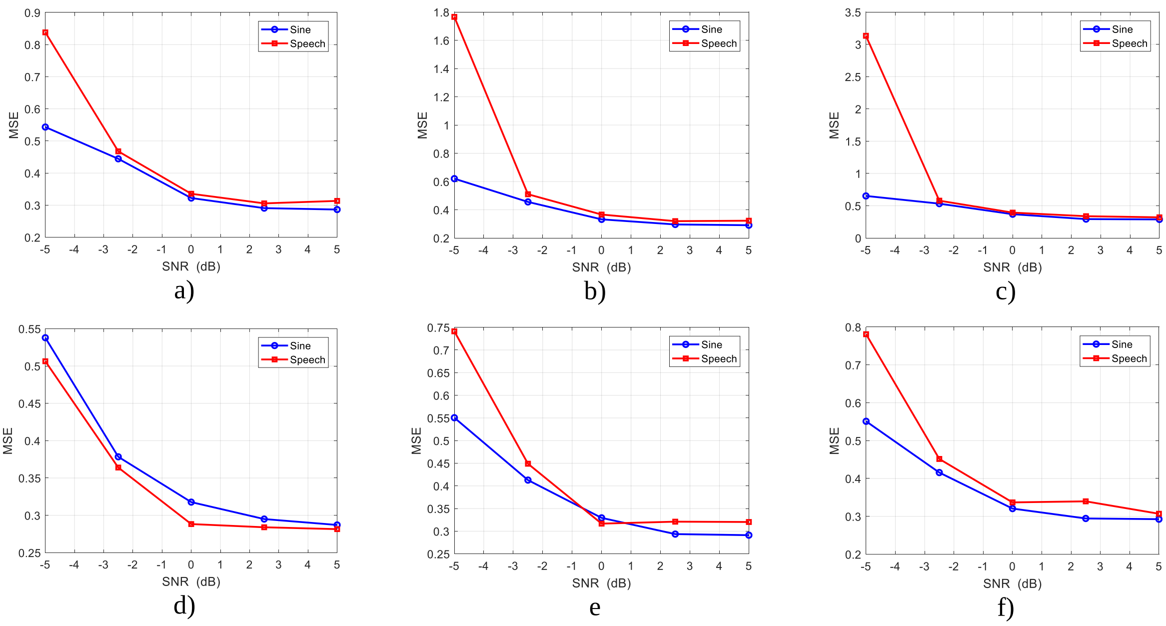


Figure 6. MSE estimation errors, for superimposed noise N8 Babble, for: a) Hamming, b) Hann, c) Blackman, e) Rectangular, e) Kaiser and f) Triangular window.

Analysis of Results

Based on the results shown in Table 1 and graphically on Figure 5-6, it is concluded that:

a) Sine test signal: $\overline{MSE}_{sine}$ is the lowest when applying the rectangular window ($\overline{MSE}_{sine}$ = 0.6538). Compared to the application of other window functions the accuracy is higher: $\overline{MSE}_{sine}$ / $\overline{MSE}_{rect}$ = 12.8979 / 0.6538 = 19.7276 (Hamming), 9.4233 / 0.6538 = 14.4131 (Hann), 9.4793 / 0.6538 = 14.4988 (Blackman), 10.1484 / 0.6538 = 15.5222 (Kaiser) and 10.7576 / 0.6538 = 16.4540 (Triangular) times. Compared to the estimation of the fundamental frequency without interpolation, when applying interpolation, the accuracy is higher $\overline{MSE}_{sin\_NN}$ / $\overline{MSE}_{sine}$ = 3.7423 / 0.6538 = 5.7239 times.

b) Speech test signal: $\overline{MSE_{\text{speech}}}$ is the lowest when applying the rectangular window ($\overline{MSE_{\text{speech}}}$ = 1.7274). Compared to the application of other window functions the accuracy is higher: $\overline{MSE_{\text{speech}}}$ / $\overline{MSE_{rect}}$ = 7.7913 / 1.7274 = 4.5104 (Hamming), 7.9991 / 1.7274 = 4.6307 (Hann), 8.3403 / 1.7274 = 4.8282 (Blackman), 7.2163 / 1.7274 = 4.1776 (Kaiser) and 7.2707 / 1.7274 = 4.2090 (Triangular) times. Compared to the estimation of the fundamental frequency without interpolation, when applying interpolation, the accuracy is higher $\overline{MSE_{\text{sp\_NN}}}$ / $\overline{MSE_{\text{speech}}}$ = 4.9834 / 1.7274 = 2.8849 times.

c) The accuracy of estimate of the fundamental frequency with the Speech test signal compared to the estimates fwith the Sine test signal, using the rectangular window, is $\overline{MSE_{speech}}$ / $\overline{MSE_{\text{sine}}}$ = 1.7274 / 0.6538 = 2.6421 times lower.

d) The optimal value of the kernel parameter with 1P Keys kernel is α_{opt} = $\overline{\alpha_{opt\_\text{Sine}} + \alpha_{opt\_Speech}}$ = -0.4432.

With the fact that the numerical complexity of the 1P Keys kernel is small, and as such, it is suitable for real-time operation. The optimal choice is the implementation of the rectangular window and the kernel parameter $\alpha = \alpha_{\text{opt}} = -0.4432$.

Conclusion

This paper describes an algorithm for estimating the fundamental frequency of a Speech signal in the time domain. The estimation algorithm is based on autocorrelation. The increase in precision was achieved by applying parametric interpolation with a 1P Keys interpolation core. Through experimentation, it has been shown that an additional increase in estimation accuracy can be achieved by processing Speech signals with window functions. MSE was used as a measure of precision. A detailed analysis of the experimental results showed that the rectangular window function was optimal. The error of estimating the fundamental frequency by autocorrelation is higher in relation to the error of estimating by interpolation: a) $\overline{MSE_{\text{sin\_NN}}}$ / $\overline{MSE_{\text{sine}}}$ = 3.7423 / 0.6538 = 5.7239 (Sine test signal) and b) / $\overline{MSE_{\text{sin\_NN}}}$ / $\overline{MSE_{\text{sine}}}$ = 4.9834 / 1.7274 = 2.8849 times. For the Rectangular window, the optimal value of the kernel parameter is $\alpha_{\text{opt}} = -0.4432$. These results recommend the application of the correlation algorithm and interpolation with the 1P Keys core in real-time systems.

Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in EPSTEM journal belongs to the authors.

References

De Cheveigné, A., & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, *111*(4), 1917-1930.

Kacha, A., Grenez, F., & Benmahammed, K. (2005). Time–frequency analysis and instantaneous frequency estimation using two-sided linear prediction. *Signal Processing*, *85*(3), 491-503.

Keys, R. (1981). Cubic convolution interpolation for digital image processing. *IEEE transactions on acoustics, speech, and signal processing*, *29*(6), 1153-1160.

Meijering, E., & Unser, M. (2003). A note on cubic convolution interpolation. *IEEE Transactions on Image processing*, *12*(4), 477-479.

Milivojević, Z. N., & Balanesković, D. Z. (2009). Enhancement of the perceptive quality of the noisy speech signal by using of DFF-FBC algorithm. *Facta universitatis-series: Electronics and Energetics*, *22*(3), 391-404.

Milivojević, Z. N., & Brodić, D. (2013). Estimation of the fundamental frequency of the speech signal compressed by mp3 algorithm. *Archives of Acoustics*, 363-373.

Milivojevic, Z., & Prlincevic, B. (2021). *Estimation of the fundamental frequency of the speech signal using autocorrelation algorithm*. Unitech.

Milivojevic, Z. N., Brodic, D. T., & Stojanovic, V. (2017). Estimation of fundamental frequency with autocorrelation algorithm. http://vtsnis.edu.rs/wp-

Pang, H. S., Baek, S., & Sung, K. M. (2000). Improved fundamental frequency estimation using parametric cubic convolution. *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences*, *83*(12), 2747-2750.

Qiu, L., Yang, H., & Koh, S.-N. (1995). Fundamental frequency determination based on instantaneous frequency estimation. *Signal Processing*, *44*(2), 233-241.

Rabiner, L.R. and Schafer, R.W. (1978) *Digital processing of speech signals.* Prentice Hall.

Rao, P., & Barman, A. D. (2000). Speech formant frequency estimation: evaluating a nonstationary analysis method. *Signal Processing*, *80*(8), 1655-1667.

Author Information

Zoran MILIVOJEVIC
Academy of Applied Technical and Preschool Studies
A. Medvedeva 20, 18000 Niš, Serbia
e-mail: *zoran.milivojevic@akademijanis.edu.rs*

Bojan PRLINCEVIC
Kosovo and Metohija Academy of Applied Studies
Dositejeva bb, 38218 Leposavic, Serbia

Natasa SAVIC
Academy of Applied Technical and Preschool Studies
A. Medvedeva 20, 18000 Niš, Serbia

To cite this article:

Milivojevic, Z., Prlincevic, B., & Savic, N. (2021). Optimization parameter of the 1P keys interpolation kernel implemented in the correlation algorithm for estimating the fundamental frequency of the speech signal. *The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 16,* 153-161.