

Fundamental Frequency Estimation of the Speech Signal Compressed by MP3 Algorithm Using PCC Interpolation

Zoran N. MILIVOJEVIC¹, Milorad D. MIRKOVIC², Sladjana M. MILIVOJEVIC³

¹Technical College, Aleksandra Medvedeva 20, Nis, Serbia.

Faculty of Technical Science, Kosovska Mitrovica, Serbia.

³Airport "Constantin the Great", Nis, Serbia

zoran.milivojevic@jotel.co.rs

Abstract—In this paper the fundamental frequency estimation results of the MP3 modeled speech signal are analyzed. The estimation of the fundamental frequency was performed by the Picking-Peaks algorithm with the implemented Parametric Cubic Convolution (PCC) interpolation. The efficiency of PCC was tested for Catmull-Rom, Greville and Greville two-parametric kernel. Depending on MSE, a window that gives optimal results was chosen.

Index Terms—Fundamental frequency, Speech compression, Speech processing, Signal representation

I. INTRODUCTION

The rising trend of multimedia communications has imposed the need for audiovisual information archiving and transferring. The amount of data which is supposed to be archived or transferred is very large [1, 2]. For instance, audio record rate in stereo technique at the sampling frequency $f_s=44.1$ kHz is 10.584 MB/min. Transferring this number of bits is a very slow process even in very fast communication media. For that reason, a need occurred for developing compressing techniques. A number of algorithms for audio signal compressing appeared. The most used is MP3 algorithm with a compression degree of 1:12. Such a compression ratio allowed the archiving of digitalized audio signal as well as its transferring by multimedia systems. MP3 became particularly popular in internet applications [3].

MP3 is the shortened name for a coding algorithm derived from the standard MPEG-1, Layer III, developed by the German Technology Group. It was standardized by ISO (International Standards Organization). MP3 does the compression tasks eliminating redundancy, similarly to zip algorithm in accordance with a psycho-acoustic model that describes mechanisms of sound perception of man. Technically seen, MPEG-1 Layer III and MPEG-2 Layer III are declared as MP3 standard. MPEG-1 Layer III is used for 32, 44.1 and 48 kHz of sampling frequency, while MPEG-2 Layer III is used for 16, 22.05 and 24 kHz of sampling frequency. The standard broadening with a sign MPEG 2.5 is used for 8 and 11 kHz [3]. MP3 compression algorithm is based on a combination of several techniques whose function is to maximize the relation between the perceived quality and the necessary storage volume (file size). Spectrum of an audio signal is divided into 32 equally spaced frequency sub-bands. After that a modified discrete cosine transformation MDCT is applied. Precision of

MDCT coefficient is reduced by the process of quantization. Further the signal is processed according to the psycho-acoustic model. This model emulates the human perception, i.e. the masking effects (auditory and temporal masking) [3]. After the signal processing according to the psychoacoustic model, i.e. perceptual coding, Huffman's coding is performed. This coding additionally performs reduction of file size for 20%.

A number of old music and speech records are digitalized and compressed by MP3 algorithm, which makes it possible for creating of huge collection of precious music and speech material. In many multimedia applications it is necessary to process audio records additionally in order to improve the quality, intelligibility of speech, verification of the speaker etc. A characteristic example is the quality improvement of the speech signal by reducing dissonant frequencies [4-6].

In processing of music and especially speech signal it is necessary to determine the fundamental frequencies. A number of algorithms for determining the fundamental frequency are developed where the processing is performed in the time and frequency domains [7-12]. The frequently applied method for determination of the fundamental frequency is based on the Picking Peaks of the amplitude characteristic in the specific frequency range. This method is used for analyzing the signal values in the spectrum on frequencies on which the Discrete Fourier Transform (DFT) was calculated. Most often the real value of the fundamental frequency is not there on the frequencies where DFT is calculated, but lies between the two spectrum samples. That causes the frequency estimation error that lies in the interval $[-(F_s/(2N)) \text{ Hz}, (F_s/(2N)) \text{ Hz}]$, where F_s is the sampling frequency and N is the DFT window size. One way of reducing the error is determination of the interpolation function and estimation of the spectrum characteristics in an interval between the two samples. This procedure gives the reconstruction of the spectrum on the base of DFT. The spectrum parameters are then determined by analytic procedures (differentiation, integration, extreme values,...).

Calculation of the interpolation function by using Parametric Cubic Convolution (PCC) was represented in [13,14]. The special case of PCC interpolation applied in computer graphics was called Catmull-Rom interpolation [15]. The detailed analysis of the fundamental frequency estimation, as well as the advantage of the PCC interpolation, which can be seen in the speed of determining the interpolation function parameters, is described in the

paper [16]. The results of the application of PCC interpolation for determining of the fundamental frequency in the conditions of application of some window in the processing the discrete speech signal are presented in [17]. Through some simulation procedures algorithm efficiency analyses have been done where, as a quality measure of an algorithm, the Mean Square Error (MSE) has been used. The best results were shown by the algorithm with the implemented Blackman window. The algorithm efficiency analysis in the conditions of the changeable S/N relation in the presence of a number of important harmonics of the fundamental function, shown in [18], confirmed the efficiency of the algorithm with the Blackman window. In [19] an analysis of the PCC interpolation algorithm efficiency is made for the case where Greville two-parametric cubic convolution kernel (G2P) was implemented is made. The window was determined and the kernel parameters were calculated (α, β) where the minimal MSE was generated (in relation to Catmull-Rom kernel the error was smaller for 58.1%). The new method of speech signal modeling called A Novel Systematic Procedure to Model Speech Signals via Predefined "Envelope and Signature Sequences" (SYMPES) was presented in the paper [20]. The results of the fundamental frequency estimation of the speech signal modeled by SYMPES method are shown in [21].

Further on in this paper some results of the speech signal analysis modeled by MP3 algorithm method in order to determine the fundamental frequency using the PCC algorithm will be presented. These results will be compared to those obtained in determining the fundamental frequency in speech signals: a) where no compression algorithm was applied [19] and b) which are modeled by SYMPES method [21]. As the algorithm quality measure the MSE will be applied. MSE will be determined for Catmull-Rom, Greville and Greville two-parametric kernel (G2P). On the base of minimal values of MSE optimal kernel parameters will be determined for the case of implementing of some characteristic windows.

This paper is organized as follows: In Section II there is a description of the PCC algorithm. In Section II.a there are definitions of interpolation kernels. In Section II.b the algorithm for determination of optimal kernel parameters is presented. In Section III numerical MSE results in the estimation of fundamental frequency of the speech signal modeled by the MP3 method are presented. The comparative analysis of the results and the choice of the optimal kernel and window function are shown in Section IV. Section V represents the conclusion.

II. ALGORITHM OF FUNDAMENTAL FREQUENCY ESTIMATION

Algorithm of estimate of fundamental frequency [16] is shown in Fig. 1. It can be realized throughout a few steps:

Step 1: The spectrum is calculated by using DFT at the discrete signal $x(n)$ that is achieved by time sampling of a continuous signal $s(t)$:

$$X(k) = DFT(x(n)) \quad (1)$$

The spectrum is calculated in discrete points $k=0, \dots, N-1$, where N is the length of DFT. The real spectrum of signals

$x(n)$ is continuous, whereas DFT defines the values of the spectrum at some discrete points.

Step 2: By using Peak Picking algorithm, the position of the maximum of the real spectrum that is between k -th and $(k+1)$ -th samples are determined, where the values $X(k)$ and $X(k+1)$ are the highest in the specified domain.

Step 3: The position of the maximum of the spectrum is calculated by PCC interpolation. The reconstructed function is:

$$X_r(f) = \sum_{i=k-L}^{k+L+1} p_i \cdot r(f-i), \quad k \leq f \leq k+1 \quad (2)$$

where $p_i = X(i)$, $r(f)$ is the kernel of interpolation and L the number of samples that participate in interpolation.



Figure 1. Algorithm of estimate of fundamental frequency.

The quality of the algorithm for the estimate of fundamental frequency can be also expressed by MSE:

$$MSE = \overline{(f_0 - f_e)^2} \quad (3)$$

where f_0 is true fundamental frequency and f_e is fundamental frequency estimate.

A. Interpolation Kernel

Next, we give definitions of the interpolation kernels which are tested in this paper:

a) **Catmull-Rom** interpolation kernel:

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

The maximum of the reconstructed function $X_r(f)$ is found by differentiating in spectrum domain and equalizing the first derivative with zero. The position of the maximum is:

$$f_{\max} = \begin{cases} k - \frac{c}{2b}, & a = 0 \\ k + \frac{-b - \sqrt{b^2 - ac}}{a}, & a \neq 0 \end{cases} \quad (5)$$

where:

$$\begin{aligned} a &= 2(\alpha p_{k-1} + (\alpha + 2)p_k - (\alpha + 2)p_{k+1} - \alpha p_{k+2}) \\ b &= -2\alpha p_{k-1} - (\alpha + 3)p_k + (2\alpha + 3)p_{k+1} + \alpha p_{k+2} \\ c &= -\alpha p_{k-1} - \alpha p_{k+1} \end{aligned} \quad (6)$$

b) **Greville** interpolation kernel:

$$r(f) = \begin{cases} \left(\alpha + \frac{3}{2}\right)|f|^3 - \left(\alpha + \frac{5}{2}\right)|f|^2 + 1; & \text{if } 0 \leq |f| \leq 1, \\ \frac{1}{2}(\alpha - 1)|f|^3 - \left(3\alpha - \frac{5}{2}\right)|f|^2 + \\ \left(\frac{11}{2}\alpha - 4\right)|f| - (3\alpha - 2); & \text{if } 1 \leq |f| \leq 2, \\ -\frac{1}{2}\alpha|f|^3 + 4\alpha|f|^2 - \frac{21}{2}\alpha|f| + 9\alpha; & \text{if } 2 \leq |f| \leq 3, \\ 0; & \text{if } 3 \leq |f|. \end{cases} \quad (7)$$

c) **Greville** two-parametric cubic convolution kernel (G2P) [15]:

$$r(f) = \begin{cases} \left(\alpha - \frac{5}{2}\beta + \frac{3}{2}\right) \cdot |f|^3 - \left(\alpha - \frac{5}{2}\beta + \frac{5}{2}\right) \cdot |f|^2 + 1; & \text{if } 0 \leq |f| \leq 1, \\ \frac{1}{2}(\alpha - \beta - 1) \cdot |f|^3 - \left(3\alpha - \frac{9}{2}\beta - \frac{5}{2}\right) \cdot |f|^2 + \\ \left(\frac{11}{2}\alpha - 10\beta - 4\right) \cdot |f| - (3\alpha - 6\beta - 2); & \text{if } 1 \leq |f| \leq 2, \\ -\frac{1}{2}(\alpha - 3\beta) \cdot |f|^3 + \left(4\alpha - \frac{25}{2}\beta\right) \cdot |f|^2 - \\ \left(\frac{21}{2}\alpha - 34\beta\right) \cdot |f| + (9\alpha - 30\beta); & \text{if } 2 \leq |f| \leq 3, \\ -\frac{1}{2}\beta \cdot |f|^3 + \frac{11}{2}\beta \cdot |f|^2 - 20\beta \cdot |f| + 24\beta; & \text{if } 3 \leq |f|. \end{cases} \quad (8)$$

In the eq. (4-8) there are α and β parameters. The optimal values of these parameters will be determined by the minimal value of MSE, for Catmull-Rom, Greville and G2P kernel. For the first two of them:

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

and for the G2P kernel

$$(\alpha_{opt}, \beta_{opt}) = \arg \min_{\alpha, \beta} (MSE) \quad (10)$$

The detailed analysis in [16,17] showed that the minimal value of MSE depends on the application of window by which signal processing $x(n)$ is carried out in time domain. MSE will be defined for: a) Hamming, b) Hanning, c) Blackman, d) Rectangular, e) Kaiser and f) Triangular window.

B. INTERPOLATION KERNEL PARAMETERS

Algorithm for determination of interpolation kernel parameters α and β is realized in the following steps:

Step 1: continual time signal $s(t)$ is sampled with f_s and modified by the window whose length is N ,

Step 2: Spectrum $X(k)$ is determined by the application of DFT,

Step 3: Reconstruction of the continual function that represents spectrum $X(f)$ is performed by the application of PCC interpolation,

Step 4: MSE is calculated for various values of parameters α and β depending on the implemented window,

Step 5: α_{opt} and β_{opt} are determined for which minimal value of MSE is obtained.

III. NUMERICAL RESULTS

The algorithm for estimation of the fundamental frequency will be applied on the simulation signal which is compressed by MP3 algorithm. After that, the obtained results will be compared to the simulation signal which was not processed at all (original signal, OS) [19] and to the signal which was compressed by SYMPES algorithm [20]. The simulation signal for testing of PCC algorithm is defined in [16]:

$$s(t) = \sum_{i=1}^K \sum_{g=0}^M a_i \sin \left(2\pi \left(f_o + g \frac{f_s}{KM} \right) t + \theta_i \right) \quad (11)$$

where f_o is fundamental frequency, θ_i and a_i are phase and

amplitude of the i -th harmonic, K is the number of harmonics, M is the number of points between the two samples in spectrum where PCC interpolation is being made. In the simulation process f_o and θ_i are random variables with uniform distribution in the range [G2 (97.99 Hz), G5 (783.99 Hz)] and $[0, 2\pi]$. Signal frequency of sampling is $f_s = 16$ kHz and the length of window is $N = 512$, which assures the analysis of subsequences that last 32 ms. Results presented further in this paper relate to $f_o = 187.5$ Hz, $K = 10$ and $M = 100$. All further analyses will relate to a) Hamming, b) Hanning, c) Blackman, d) Rectangular, e) Kaiser and f) Triangular window. Further in the text results of the PCC algorithm application on the simulation signal eq. (11), which was previously coded and decoded by MP3 algorithm, are presented.

A. Catmull-Rom Kernel

Applying the algorithm for determination of parameters of Catmull-Rom kernel diagrams MSE(α) have been drawn (Fig. 2) and values α_{opt} have been determined for a) Hamming ($\alpha_{opt} = -1.01$), b) Hanning ($\alpha_{opt} = -0.8825$), c) Blackman ($\alpha_{opt} = -0.8024$), d) Kaiser ($\alpha_{opt} = -1.125$) and e) Triangular ($\alpha_{opt} = -1.028$) window.

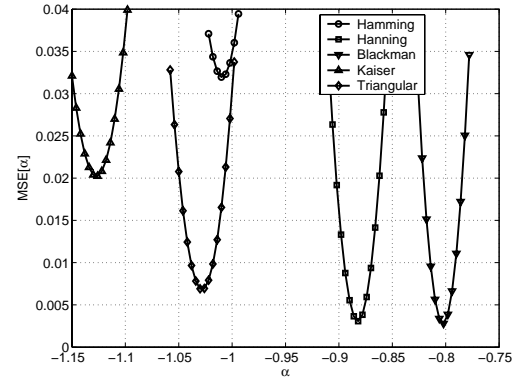


Figure 2. MSE(α) for the application of: a) Hamming, b) Hann, c) Blackman, d) Kaiser and e) Triangular windows in Catmull-Rom PCC interpolation.

B. Greville Kernel

Applying the algorithm for determination of parameters of Greville interpolation kernel diagrams have been drawn (Fig. 3) MSE(α) and values α_{opt} have been determined for a) Hamming ($\alpha_{opt} = -0.575$), b) Hanning ($\alpha_{opt} = -0.45$), c) Blackman ($\alpha_{opt} = -0.42$), d) Kaiser ($\alpha_{opt} = -0.66$) and e) Triangular ($\alpha_{opt} = -0.575$) window.

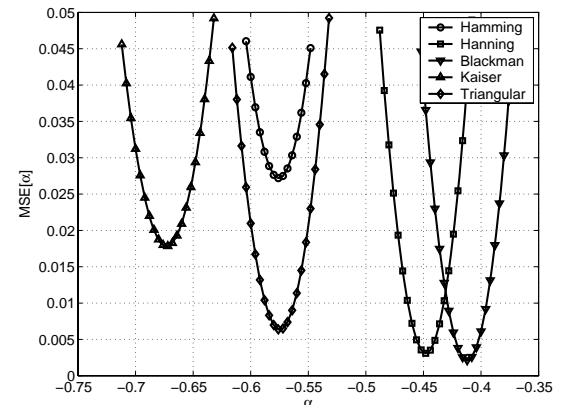


Figure 3. MSE(α) for the application of: a) Hamming, b) Hann, c) Blackman, d) Kaiser and e) Triangular windows in Greville PCC interpolation.

C. G2P Kernel

MSE(α, β) charts for G2P kernel are calculated (eq. 8) and determined α_{opt} and β_{opt} values are calculated for: a) Hamming ($\alpha_{opt}=-0.59$, $\beta_{opt}=0.006$), b) Hanning ($\alpha_{opt}=-0.46$, $\beta_{opt}=0.006$), c) Blackman ($\alpha_{opt}=-0.42$, $\beta_{opt}=-0.002$), d) Rectangular ($\alpha_{opt}=-2.3$, $\beta_{opt}=-0.0265$), e) Kaiser ($\alpha_{opt}=-0.66$, $\beta_{opt}=0.006$) and f) Triangular ($\alpha_{opt}=-0.568$, $\beta_{opt}=-0.003$) window. MSE(α, β) diagram for Blackman window is shown in Fig. 4. In fig. 5 positions of MSE($\alpha_{opt}, \beta_{opt}$) minimum in (α, β) plane for Blackman (point A) and G2P (point B) interpolation kernel, are shown. Vector **AB** shows the position change of $\min(\text{MSE}(\alpha_{opt}, \beta_{opt}))$.

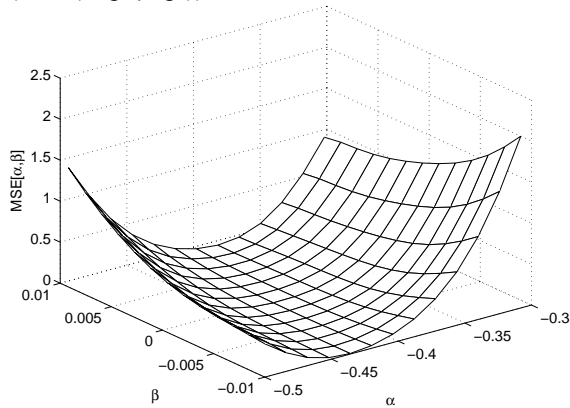


Figure 4. MSE(α, β) for the application of Blackman windows in G2P PCC interpolation.

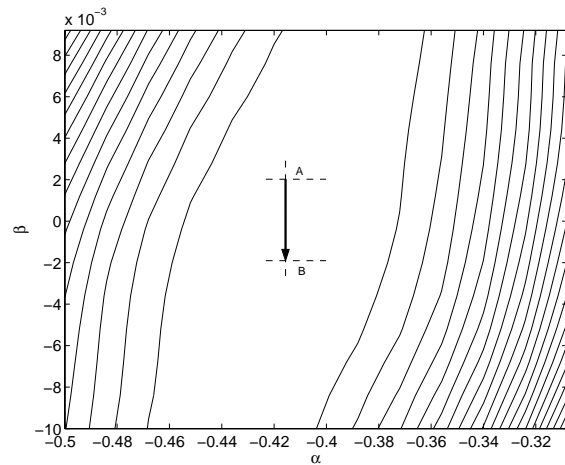


Figure 5. Positions of $\min(\text{MSE}(\alpha_{opt}, \beta_{opt}))$ in plane (α, β) of Blackman windows for Greville (point A) and G2P PCC (point B) interpolation.

IV. NUMERICAL RESULTS

Comparative precision analysis of the estimated fundamental frequency among an OS signal without processing (results are presented in [19]), signal coded by SYMPES method (results are presented in [21]) and signal compressed by MP3 algorithm will be performed on the base of the minimal values of MSE. The minimal value of MSE is determined on the base of a diagram in the Fig. 2. (Catmull-Rom), Fig.3 (Greville) and Fig.4 (G2P) and presented in the Table I (MSECRmin, MSECRSmin, MSECRMP3min), Table II (MSEGMmin, MSEGSMmin, MSEGMP3min,) and Table III (MSE2Dmin) respectively.

TABLE I. MSE MINIMAL VALUES AND α_{opt} WITH THE APPLICATION OF CATMULL-ROM INTERPOLATION.

Window	OS	
	α_{opt}	MSE _{CRmin}
Hamming	-1.005	0.023
Hanning	-0.885	0.004
Blackman	-1.801	0.001
Rectangular	-2.61	0.515
Kaiser	-1.125	0.02
Triangular	-1.028	0.0028
	SYMPES	
	α_{opt}	MSE _{CRSmin}
Hamming	-0.96	4.7521
Hanning	-0.84	5.9784
Blackman	-0.761	8.0173
Rectangular	-2.23	4.6945
Kaiser	-1.055	4.0495
Triangular	-0.97	4.0739
	MP3	
	α_{opt}	MSE _{CRMP3min}
Hamming	-1.0100	0.0320
Hanning	-0.8825	0.0031
Blackman	-0.8024	0.0028
Rectangular	-2.5500	0.4388
Kaiser	-1.1250	0.0203
Triangular	-1.0280	0.0068

TABLE II. MSE MINIMAL VALUES AND α_{opt} WITH THE APPLICATION OF GREVILLE INTERPOLATION.

Window	OS	
	α_{opt}	MSE _{Gmin}
Hamming	-0.57	0.0175
Hanning	-0.449	0.0027
Blackman	-0.415	0.0009
Rectangular	-2.254	0.4054
Kaiser	-0.6676	0.0124
Triangular	-0.575	0.002
	SYMPES	
	α_{opt}	MSE _{Gmin}
Hamming	-0.53	4.9934
Hanning	-0.39	6.1186
Blackman	-0.37	7.7194
Rectangular	-1.944	6.3222
Kaiser	-0.6	4.4923
Triangular	-0.52	4.3984
	MP3	
	α_{opt}	MSE _{GMP3min}
Hamming	-0.5750	0.0272
Hanning	-0.4500	0.0032
Blackman	-0.4200	0.0037
Rectangular	-2.2000	0.3966
Kaiser	-0.6600	0.0207
Triangular	-0.5750	0.0064

TABLE III. MSE MINIMAL VALUES, α_{opt} AND β_{opt} WITH THE APPLICATION OF G2P INTERPOLATION.

Window	Signal	α_{opt}	β_{opt}	MSE _{2Dmin}
Hamming	OS	-0.55	0.03	0.0046
	SYMPES	-0.495	0.003	4.7047
	MP3	-0.5900	-0.0060	0.0270
Hanning	OS	0.5	0.015	0.0018
	SYMPES	-0.29	0.0055	4.6195
	MP3	-0.4600	-0.0060	0.0029
Blackman	OS	-0.42	0.002	0.000377
	SYMPES	-0.39	-0.007	5.8612
	MP3	-0.4200	-0.0020	0.0022
Rectangular	OS	-2.272	0.005	0.2244
	SYMPES	-1.9	-0.0065	4.8314
	MP3	-2.3000	-0.0265	0.3855
Kaiser	OS	-0.681	0.001	0.0096
	SYMPES	-0.58	0.025	4.3869
	MP3	-0.6600	0.0060	0.0178
Triangular	OS	0.6	-0.001	0.001
	SYMPES	-0.7	-0.0105	3.1740
	MP3	-0.5680	0.0030	0.0060

Comparing the values of MSE from the Tables I-III it may be concluded that:

a) The optimal choice for OS is Blackman window for all interpolation kernels. In comparison to Catmull-Rom kernel, Greville kernel generates 10% whereas G2P generates 62.23% less of MSE.

b) The optimal choice for SYMPES signal is the Triangular window. In comparison to Greville kernel Catmull-Rom kernel generates 7,94% whereas G2P generates 27,48% less of MSE.

c) The optimal choice for MP3 signal is the Blackman window. In comparison to Greville kernel Catmull-Rom kernel generates 24.32% whereas G2P generates 40.54% less of MSE.

d) By comparing MSE for G2P kernel for OS (Blackman window) and SYMPES (Triangular window) a relation $(MSE_{2D\min})_{SYMPES}/(MSE_{2D\min})_{OS} = 8419.098$ has been obtained. On the base of this relation it has to be concluded that the estimation of the fundamental frequency for the signal modeled by SYMPES method is very imprecise.

e) By comparing MSE for G2P kernel for OS (Blackman window) and MP3 (Blackman window) a relation $(MSE_{2D\min})_{MP3}/(MSE_{2D\min})_{OS} = 5.83$ has been obtained.

In accordance to the derived conclusion, the application of the algorithm for further processing of SYMPES signal with algorithms based on the estimated fundamental frequency (automatic verification of a speaker, recognition of the speech etc.) would not bring satisfactory results [21], while processing of MP3 signal causes significantly smaller error. The obtained results recommend using PCC algorithm with G2P kernel in preprocessing signals compressed by MP3 method for further processing by algorithms which require a precise determination of the fundamental frequency.

V. CONCLUSION

This paper presents the analysis of the results obtained for estimation of the fundamental frequency of the speech signal modeled by MP3 method. The estimation of the fundamental frequency has been made by Picking Peaks algorithm with implemented PCC interpolation. Experiments have been performed with Catmull-Rom, Greville and Greville two-parametric kernels. In order to minimize MSE some windows have been implemented. The detailed analysis has shown that the optimal choice is Greville two-parametric kernel and the Blackman window implemented in PCC algorithm. In relation to Greville kernel, Catmull-Rom kernel generates 24.32% whereas G2P generates 40.54% less MSE. Comparing the obtained results to the results of the estimation of the fundamental frequency in the speech signal that is not modeled by MP3 method, a relation of minimal MSEs 5.83 has been obtained. The obtained results recommend using of PCC algorithm with G2P kernel in preprocessing of signals compressed by MP3 method for further processing by algorithms which require a precise determination of the fundamental frequency

(automatic verification of the speaker, recognition of the speech etc.).

REFERENCES

- [1] K. Brandenburg, G. Stoll, Y. F. Dehery, J. D. Johnston, L. V. Kerkhof, E. F. Schroeder, The ISO/MPEG Audio Codec: A Generic Standard for Coding of High Quality Digital Audio, 92nd. AES-convention, preprint 3336, Vienna, 1992.
- [2] ISO/IEC 13818-3, Information Technology. Generic Coding of Moving Pictures and Associated Audio: Audio. International Standard, 1994.
- [3] S. Hacker, MP3: The Definitive Guide, O'Reilly & Associates, Sebastopol, CA 95472, 2000.
- [4] B. Joen, S. Kang, S. J. Baek, K. M. Sung, "Filtering of a Dissonant Frequency Based on Improved Fundamental Frequency Estimation for Speech Enhancement", *IEICE Trans. Fundamentals*, vol. E86-A, No. 8, pp. 2063-2064, August 2003.
- [5] S. Kang, "Dissonant frequency filtering technique for improving perceptual quality of noisy speech and husky voice", *IEEE Signal Processing*, vol. 84, pp. 431-433, 2004.
- [6] S. Kang, Y. Kim, A Dissonant Frequency Filtering for Enhanced Clarity of Husky Voice Signals, *Lect Notes Comp Science*, Springer, vol. 4188, pp. 517-522. Berlin, 2006.
- [7] A. Cheveigne, H. Kawahara, "YIN, a fundamental frequency estimator for speech and music", *J. Acoust. Soc. Am.* vol. 111 (4), pp. 1917-1930, 2002.
- [8] L. Qiu, H. Yang, S. N. Koh, "Fundamental frequency determination on instantaneous frequency estimation", *IEEE Signal Processing*, vol. 44, pp. 233-241, 1995.
- [9] S. C. Sekhar, T. V. Sreenivas, "Effect of interpolation on PWVD computation and instantaneous frequency estimation", *IEEE Signal Processing*, vol. 84, pp. 107-116, 2004.
- [10] P. Rao, A. Barman, "Speech formant frequency estimation: evaluating a nonstationary analysis method", *IEEE Signal Processing*, vol. 80, pp. 1655-1667, 2000.
- [11] Z. M. Hussain, B. Boashash, "Adaptive instantaneous frequency estimation of multicomponent signals using quadratic time-frequency distributions", *IEEE Trans. Signal Process.* vol. 50 (8), pp. 1866-1876, 2002.
- [12] A. Kacha, F. Grenet, K. Benmahammed, "Time-frequency analysis and instantaneous frequency estimation using two-sided linear prediction", *IEEE Signal Process.* vol. 85, pp. 491-503, 2005.
- [13] R.G. Keys, "Cubic convolution interpolation for digital image processing", *IEEE Trans. Acoust., Speech & Signal Process.*, Vol. 29, No. 6, pp. 1153-1160, Dec. 1981.
- [14] K. S. Park, R. A. Schowengerdt, "Image reconstruction by parametric cubic convolution", *Comput. Vision, Graphics & Image Process.* vol. 23, pp. 258-272, 1983.
- [15] E. Meijering, M. Unser, "A Note on Cubic Convolution Interpolation", *IEEE Trans. on Image Process.*, vol. 12, No. 4, pp. 447-479, April 2003.
- [16] H. S. Pang, S. J. Baek, K. M. Sung, "Improved Fundamental Frequency Estimation Using Parametric Cubic Convolution", *IEICE Trans. Fund.* Vol. E83-A, No. 12, pp. 2747-2750, Dec. 2000.
- [17] Z. Milivojevic, M. Mirkovic, P. Rajkovic, "Estimating of the fundamental frequency by the using of the parametric cubic convolution interpolation", *Proceedings of International Scientific Conference UNITECH '04*, Session: Electronics and Communication Engineering, pp. 138-141, Gabrovo, Bulgaria, 2004.
- [18] M. Mirkovic, Z. Milivojevic, P. Rajkovic, "Performances of the system with the implemented PCC algorithm for the fundamental frequency estimation", *Proceedings of XII Telecommunications Forum TELFOR '04*, Section. 7, Signal processing, Beograd, 2004.
- [19] Z. Milivojevic, M. Mirkovic, S. Milivojevic, "An Estimate of Fundamental Frequency Using PCC Interpolation - Comparative Analysis", *Inf Technol Control*, vol. 35, No. 2, pp.131-136, 2006.
- [20] B. Yarman, U. Guz, H. Gurkan, "On the comparative results of SYMPES: A new method of speech modeling", *Int J Electron Commun (AEU)* 2006; 60:421-7.
- [21] Z. Milivojevic, D. Mirkovic, "Estimation of the fundamental frequency of the speech signal modeled by the SYMPES method", *Int J Electron Commun (AEU)*, vol 63, pp. 200-2008, (2009).