Extensive Evaluation Experiments for the Accumulated Cross-Power Spectrum Methods for Time Delay Estimation

Radu-Sebastian Marinescu[*,#], Andi Buzo[*], Horia Cucu[*], Corneliu Burileanu[*]
[#]Reasearch & Development, Rohde & Schwarz Topex
[*]SpeeD Laboratory, University Politehnica of Bucharest
Bucharest, Romania
radu-sebastian.marinescu@rohde-schwarz.com, {andi.buzo, horia.cucu, corneliu.burileanu}@upb.ro

Abstract — **In numerous real time signal processing applications, time delay estimation (TDE) is still a hot topic. A recently proposed solution, based on Generalized Cross Correlation (GCC) method, is evaluated further in this paper. Experimental results show that among traditional variants of GCC methods the *accumulated ρ-cross power spectrum phase* proposed by us is the most accurate. Another important aspect is that the computing time of this method is comparable with those of other popular GCC methods.**

Keywords—Time delay Estimation; Generalized Cross Correlation; Accumulated ρ-Cross Power Spectrum Phase

I. INTRODUCTION

Time Delay Estimation (TDE) is an active part in various systems. It is used as a solution for echo cancelling, for acoustic, radar and sonar localization, in pattern recognition, seismic, medical and speech processing. The variety of applications with use TDE has different requirements regarding optimization and integration. Because of this it is inhibited the development of a unique solution for TDE.

Based on applications specific demanding, the various approaches to time delay estimation can be grouped into three main categories: a) generalized cross-correlation (GCC), b) least mean square (LMS) adaptive filtering and c) adaptive eigenvalue decomposition (AEVD). Each category has its own advantages, which make it optimal for particular applications. The latter of these three approaches (AEVD) was proved to be very efficient in reverberant environments [1]. The second class of TDE methods, the adaptive filtering [2-8], is a useful solution for applications with a high accuracy demanding and easy constraints on the adaptation time. But, in applications with a fast response the solution is provided by the first category based on generalized cross-correlation. It does not need any adaptation time, and offers a reasonable accuracy.

The generalized cross-correlation method was proposed initially by Knapp and Carter [9], in 1976. They also presented a particular derivation of the GCC, with weighting, CSP-m. Gradually, over the time, based on this work, multiple variations of the GCC weighting function were proposed or integrated in this technique: ROTH [10], SCOT [11], Eckart [12], Phase Transform (PHAT) or Cross-power Spectrum Phase (CSP) [9], HT (ML) [13], HB [14], Wiener [15], ρ-CSP [16], accumulated CSP (acc-CSP) [17], ρ-CSPC [18]. Over the time were presented several papers which studied these methods. In one of them [19], the methods were compared based on the mean value and root mean square deviation of the estimated delay and.

As a further development, in [20] we presented two new methods: *accumulated ρ-cross power spectrum with coherence (acc-ρCSPC)* and *accumulated ρ-cross power spectrum (acc-ρCSP)*. They benefit from the lower computational load and robustness of the *acc-CSP* method [17] and the higher accuracy, which characterized the *ρ-CSP* [16] and *ρ-CSPC* [18]. In [20] it was shown that *acc-ρCSPC* and *acc-ρCSP* outperform previous methods in computation time, because of the accumulation of cross-power spectrum phase in frequency domain, introduced in [17]. This leads to only one Inverse Discrete Fourier Transform (IDFT) for any number of accumulated frames used. Also, in [20], it was shown that the first method *(acc-ρCSPC)* generally has a higher accuracy than *(acc-ρCSP),* but, after a proper calibration, the second method achieves practically the same accuracy as the *acc-ρCSPC* at a lower computational load.

In [21], we continued to evaluate the proposed *acc-ρCSPC* method over their previous primitives' methods. The obtained results from [21] showed that the new combination of *acc-ρCSPC*, based on previous *acc-CSP* and *ρCSPC*, offers higher accuracy rate and faster computational speed. Another remark was the *acc-ρCSPC*, which is performed in the frequency domain, leads to an accuracy that is more than twice than the accuracy obtained with the previous methods, which average the final results in the time domain. Also, the accuracy of the *acc-ρCSPC* increases with the number of frames, while for previous methods that use the average in time, it decreases as the number of frames grows [21].

For a complete study of *acc-ρCSP* and *acc-ρCSPC* methods, we continued to evaluate them, among other well known generalized cross-correlation methods. In this way, this work appends to the previous evaluations, from [20] and [21], offering detailed results across the variety of GCC functions.

The paper is organized as follows. In first part of Section II we present the time delay estimation problem and various

GCC approaches. The next part is dedicated to our recently proposed TDE methods. Experiments and results discussions stand on Section III. Conclusions and further work are presented in Section IV.

II. Generalized Cross-Correlation for Time Delay Estimation

A. Background

In several applications, a transmitted signal $x(t)$ could be received by two (or more) noisy and delayed signals $y_1(t)$ and $y_2(t)$. The problem of finding the relative delay between these signals is called time delay estimation. To solve this, over the time were proposed different methods, but a time-efficient and wildly are those based on the cross-correlation. They are called generalized cross-correlation, named by the work of Knapp and Carter from [9], where it is incorporated a filtering function:

$$R_{y_1y_2}^{g}(t) = \int_{-\infty}^{\infty} \Psi(f) \cdot G_{y_1y_2}(f) \cdot e^{j2\pi ft} df \qquad (1)$$

where $\psi(f)$ is a general frequency weighting and $G_{y_1y_2}(f)$ represents the cross-power spectrum. The insertion of the weighting function is expected to emphasize spectral information that can depend on source and noise characteristics in order to improve the delay estimation. Thus, the estimated delay is found at the argument t which maximizes the cross-correlation function.

Table I represents various weighting functions used in this work, where $G_{y_1y_1}(f)$ and $G_{y_2y_2}(f)$ are auto power spectrum of the noisy signals and $\gamma_{y_1y_2}^2(f)$ is the signal's coherence function:

$$\gamma_{y_1y_2}^2(f) = \frac{\left|G_{y_1y_2}(f)\right|^2}{G_{y_1y_1}(f) \cdot G_{y_2y_2}(f)} \qquad (2)$$

For the Cross-Correlation the weighting function $\psi(f)$ is 1. This is the basic and the fastest computing GCC, because practically it has not any weighting operation.

The processor proposed by Roth in 1971, has desirable effect of suppressing those frequency regions where $G_{y_1y_1}(f)$ is large and the estimate of $G_{y_1y_2}(f)$ is more likely to be in error [9].

The SCOT (Smoothed Coherence Transform) was introduced by Carter, Nuttall and Cable in 1973, to reduce the influence of a strong tonal. However, for smoothed signal and noise spectra, Hassab and Boucher [14],[22] have noted that the additional SCOT weighting function has weakened the performance of the basic cross correlator while other functions have improved it [23].

PHAT was developed purely as an ad-hoc technique to avoid spreading of the above two presented operators [9]. Ideally, PHAT does not suffer the spreading that other weighting functions do. Also, because it weights $G_{y_1y_2}(f)$ as the inverse of $\left|G_{y_1y_2}(f)\right|$, the errors are accentuated where signal power is smallest.

TABLE I. GCC Weighting functions

GCC name	Weighting function
Cross-Correlation	1
ROTH	$1/G_{y_1y_1}(f)$
SCOT	$1/\sqrt{G_{y_1y_1}(f) \cdot G_{y_2y_2}(f)}$
CSP (PHAT)	$1/\left\|G_{y_1y_2}(f)\right\|$
CSP-m	$1/\sqrt[m]{G_{y_1y_1}(f) \cdot G_{y_2y_2}(f)}$
Eckart	$\frac{\left\|G_{y_1y_2}(f)\right\|}{\left[G_{y_1y_1}(f) - \left\|G_{y_1y_2}(f)\right\|\right] \cdot \left[G_{y_2y_2}(f) - \left\|G_{y_1y_2}(f)\right\|\right]}$
HT (ML)	$\frac{\left\|\gamma_{y_1y_2}(f)\right\|^2}{\left\|G_{y_1y_2}(f)\right\| \cdot \left[1 - \left\|\gamma_{y_1y_2}(f)\right\|^2\right]}$
HB	$\left\|G_{y_1y_2}(f)\right\| / G_{y_1y_1}(f) \cdot G_{y_2y_2}(f)$
Wiener	$\left\|\gamma_{y_1y_2}(f)\right\|^2$
ρ-CSP	$1/\left\|G_{y_1y_2}(f)\right\|^{\rho}$
ρ-CSPC	$\frac{1}{\left\|G_{y_1y_2}(f)\right\|^{\rho} + \min\left[\gamma_{y_1y_2}(f)^2\right]}$

The Eckart filter derives its name from the work done in this area in [12] published in 1951 and it aims to maximize the deflection criterion [9].

In 1971 it was proposed HT processor. It assigns greater weight in regions of frequency domain where the coherence is large. In [9], it was shown that HT processor is a maximum likelihood (ML) estimator for time delay under usual conditions. Under a los signal-to-noise ratio restriction, the HT processor is equivalent to Eckart prefiltering and cross-correlation.

Eight years later, it was presented the HB processor. For highly dynamic spectra, in attempt to reject strong tonals from the observations high and low SNR regions are suppressed from the cross-spectral estimate [14].

The Wiener processor [15] was proposed in 1985. Based on channel's linearity, it tries to estimate the original signal as best it can from the observation $y_1(t)$ and channel output signal from $y_2(t)$, by minimizing the mean-square errors. In this way, given the channel characteristics, the solution results in Wiener filters, which yield the Wiener weighting function.

In 1996 it was presented a new weighting function, the ρ-CSP. It adds to the normal CSP (PHAT) the tuning parameter ρ (with values between 0 and 1) as a whitening parameter, which discards the non-speech portion (below 200Hz) of the CSP.

Relatively recently, in addition to the above work, in 2009 was proposed ρ-CSP. The presence of the minimum of the coherence function in the weighting function helps to reduces errors for relatively small energy signals.

To compute a TDE, the GCC methods use an analysis window. Because of the way the signals are received, the analysis window is usually split in several frames. Then the cross-power spectrum is calculated for each frame. An optimized implementation involves 3 Fast Fourier Transform (FFT) computations. First two are used to calculate the spectra of the input signals, followed by the weighting operation in frequency domain. The third FFT, which is in fact an Inverse Fast Fourier Transform (IFFT), is used to return to the time domain in order to estimate the delay of the current frame. Finally, the overall estimate is obtained by averaging all frame estimates.

Due to the fact that the response time is important in almost all TDE applications, the focus should be on two factors: processing time and window length. The use of a larger frame leads to a higher accuracy rate for correct TDE, but the downside is the computing time, which depends on the frame size. In 2006, Matassoni and Svaizer proposed acc-CSP [17], as a variant to compute the CSP over multiple frames. The novelty of the method was that the time delay is estimated by averaging CSP over all frames, in the frequency domain, while for the previous methods the average is computed in the time domain. In this way the computing time decreases because the cross-power spectrum phase is accumulated over multiple frames, remaining only one IFFT to be computed after the last accumulation. In the frequency domain this is expressed as follows:

$$G_{acc-CSP}(f) = \sum_{k=1}^{K} \frac{G_{y_1 y_2,k}(f)}{\left|G_{y_1 y_2,k}(f)\right|} \quad (3)$$

where K represents the number of accumulated frames.

The previous methods propose to compute the TDE as the average of all partial estimated delays of each frame from the analysis window. While in previous approaches the total FFT operations is $3 \times K$, for the accumulation approach from (3) it is reduced to only $2 \times K + 1$. In this way the computation speed is increased and also, for fixed delay during the analysis window, the delay estimation is enhanced because of the intrinsic integration, how it is shown in [17].

B. Accumulated ρ-Cross-Power Spectrum Phase Methods

Recently, in [20], we proposed to combine *acc-CSP* with *ρCSP* and *ρCSPC*, resulting the new methods *accumulated ρ-Cross Power Spectrum Phase* (*acc-ρCSP*) and *accumulated ρ-Cross Power Spectrum Phase with Coherence* (*acc-ρCSPC*). In frequency domain we can express these, in the same order, by (4) and (5):

$$G_{acc-\rho CSP}(f) = \sum_{k=1}^{K} \frac{G_{y_1 y_2,k}(f)}{\left|G_{y_1 y_2,k}(f)\right|^{\rho}} \quad (4)$$

$$G_{acc-\rho CSPC}(f) = \sum_{k=1}^{K} \frac{G_{y_1 y_2,k}(f)}{\left|G_{y_1 y_2,k}(f)\right|^{\rho} + \min\left[\gamma^2_{y_1 y_2,k}(f)\right]} \quad (5)$$

where $\gamma^2_{y_1 y_2,k}(f)$ is the signal's coherence function for the k-th frame, computed easily with (2).

These new methods incorporate the advantages of their primitives: accuracy improvements from ρCSP and ρCSPC, and the accumulation scheme from acc-CSP. These were analyzed in [20] and [21], concluding in a higher accuracy rate in low SNR conditions.

The whitening parameter ρ emphases the speech frequencies from the spectrum, reducing the noise impact over the non-speech frequencies of the spectrum. The addition of minimum coherence term in *acc-ρCSPC* prevents the errors introduced by a small denominator in the case of small energy signal parts [18].

If *acc-ρCSP* and *acc-ρCSPC* are adequately calibrated, as it is proposed in [20], their accuracy rate are practically equal, but the first method is faster because it does not need to compute the minimum coherence term.

III. Experimental Results and Discussions

A. Experimental Setup

In previous papers, [20] and [21], we evaluated the new methods, comparing them with theirs primitive methods. In this work, we continued a further evaluation and discussions for these methods, but we extended the comparison over all well-known and frequently used GCC methods.

All the methods used in next experiments were presented in sections II.A and II.B. They were implemented in Matlab and C, for accuracy, and respectively for computing time evaluation. We used the international Noizeus [24] speech corpus as the primary database. It has 30 sentences, from 6 persons (3 males and 3 females), on which were added 8 different real-world noises (suburban train, babble, car, exhibition hall, restaurant, street, airport and train station noise), from the AURORA database [25] at 4 different SNRs (0, 5, 10 and 15 dB), with the sampling frequency of 8 kHz.

Summing all combinations of noise types yields a total $C_2^8 =$ 28 possibilities. For relevant results we present the comparison for 0 and 15 dB SNR. In the next experiments we used $\rho = 0.73$ for the whitening parameter, which was configured as mentioned in [20]. Additionally to this configuration, no overlapping was used this time. Despite the fact that for the *CPS-m* weighting function m is chosen in generally between 1 and 2, we obtained better results with $m = 4$.

In this work we used more metrics: accuracy, relative error, standard deviation of the relative error and computing time. The accuracy rate is defined as the ratio between the number of perfectly estimated delays and the total number of estimations performed. This metric was the most relevant one because we intend to apply this TDE to multi-channel speech enhancement applications which need a perfect delay estimation. The relative error represents the relative difference between the estimated and correct delay. For the processing time we implement the methods in C language, compiled with Gnu Compiler Collection (gcc) version 4.7.3, and run tests on a machine with Intel Core i5 processor and Ubuntu 13.04 operating system.

B. *Evaluation of the Generalized Cross-Correlation Methods*

We evaluated the above methods for 10 different artificially introduced delays (from 5 to 50 ms) on frame size of 1024 samples, and $k = 4$. Taking into account the combinations described in the previous subsection, for every different signal-to-noise level (0, 5, 10 and 15 dB) the total number of tested pairs was 28x30x10=8400. Obtained results for 0 and 15 dB SNR are presented in Table II and III. The accuracy, relative error and standard deviation represent the average on all 8400 tested pairs.

The last three rows from Table II and III contains results for methods proposed by us which estimate time delay by averaging CSP over all frames, in the frequency domain (acc-CSP, acc-ρCSP and acc-ρCSPC). For these three methods we obtained high accuracy rates, and for the last two 100%. The impressive differences between these methods and the above, which estimate delay by averaging in time domain, impose them as a real solution, which offer high accuracy.

Better results are obtained in Table III, because the higher signal-to-noise ratio is, more accurate time delay estimation is achieved. It is notable that how the accuracy for CSP and HB is practically the same at different signal-to-noise ratios, but mean relative error and standard deviation are better for CSP. This is explained by the similarities between the weighting functions.

From the group of GCC methods which does not use the accumulating scheme, proposed by Matassoni and Svaizer in [17], ρCSP and ρCSPC have the highest accuracy ration, followed by the CSP-m. This confirms again our decision to apply the accumulation scheme to ρCSP and ρCSPC to obtain improved accuracy for acc-ρCSP and acc-ρCSPC.

Over all, the results show that the proposed acc-ρCSPC outperforms the other previous methods by obtaining higher accuracy and a lower relative error in delay estimation. In addition, the lower values for relative error and standard deviation of the relative error show that even if the signals are not perfectly aligned the gap is smaller than in the other methods. This means that acc-ρCSP and acc-ρCSPC are more efficient even for applications where the perfect alignment is not compulsory, for example signal spotting.

TABLE II. EVALUATION RESULTS - 0DB

GCC name	Accuracy (%)	Relative Error (%)	Standard deviation of relative error
Cross-Correlation	8.8	34.5	42.6
ROTH	0.4	86.6	169.0
SCOT	0.3	51.6	118.0
CSP	5.5	48.2	148.0
CSP-m	12.5	32.2	40.7
Eckart	0.4	84.0	153.0
HT (ML)	0.1	74.7	172.0
HB	5.5	102.7	148.0
Wiener	0.9	48.6	64.4
ρCSP	17.2	31.6	41.0
ρCSPC	15.1	32.3	39.0
acc-CSP	88.4	11.0	60.0
acc-ρCSP	**~100**	**~0**	**~0**
acc-ρCSPC	**~100**	**~0**	**~0**

TABLE III. EVALUATION RESULTS – 15DB

GCC name	Accuracy (%)	Relative Error (%)	Standard deviation of relative error
Cross-Correlation	25.3	32.0	46.0
ROTH	17.4	50.7	119.5
SCOT	3.0	30.9	65.9
CSP	47.6	16.9	71.1
CSP-m	53.6	26.6	44.5
Eckart	5.1	54.5	115.5
HT (ML)	1.9	62.9	143.2
HB	47.6	34.2	72.5
Wiener	4.5	46.4	53.0
ρCSP	63.8	20.4	43.9
ρCSPC	60.9	19.4	45.8
acc-CSP	99.2	0.8	7.0
acc-ρCSP	**~100**	**~0**	**~0**
acc-ρCSPC	**~100**	**~0**	**~0**

Table IV and V present important aspects regarding the accuracy and the computing time, in different implementation variants. In some applications it is possible to use a large analysis window, without dividing it into smaller frames. The advantage here it is a higher accuracy rate. On the other hand, the processing time increases. Also, the longer the analysis window, the more difficult to track correctly unstable delay. In this condition, the final choice must consider all these aspects, and we need a good knowledge of all the possible solution. For this reason, we present in the next tables the accuracy (for 0 and 15 dB signal-to-noise ratio) and computing time for two implementations. First divides the analysis window in 4 smaller frames, of 1024 samples, while the second performs the time delay estimation once, on the entire analysis window of 4096 samples.

In Table IV it is shown that the increase of the signal-to-noise ratio and analysis frame lead to a higher accuracy rate for almost all methods. For the single frame analysis, the accumulation methods reduce to their primitives. I.e. acc-CSP, acc-ρCSP and acc-ρCSPC are reduced to CSP, ρCSP and ρCSPC.

TABLE IV. ACCURACY COMPARASION

GCC name	Accuracy – 0 dB (%)		Accuracy – 15 dB (%)	
	Short frames	Long frame	Short frames	Long frame
Cross-Correlation	8.8	100	25.3	100
ROTH	0.4	77.4	17.4	87.8
SCOT	0.3	48.0	3.0	50.4
CSP	5.5	99.8	47.6	100
CSP-m	12.5	100	53.6	100
Eckart	0.4	74.5	5.1	91.4
HT (ML)	0.1	55.6	1.9	75.0
HB	5.5	99.8	47.6	100
Wiener	0.9	78.7	4.5	96.5
ρCSP	17.2	100	63.8	100
ρCSPC	15.1	100	60.9	100
acc-CSP	88.4	100	99.2	100
acc-ρCSP	100	100	100	100
acc-ρCSPC	100	100	100	100

While for a short frame size of 128 ms the maximum introduced delay of 50 ms represents almost 40%, for a 4 times larger frame it represents around 10%. From table IV we observe that the large analysis frame leads to higher accuracy rate for the almost all methods, while the smaller multi-frame approach could not offer any reliable results for any methods, with an exception for accumulated methods. Moreover, *acc-ρCSP* and *acc-ρCSPC* offer highly accurate estimation even for larger delays of about 40% of the frame size length. Thus, this remark makes them very useful for real time solution of time delay estimation.

TABLE V. COMPUTING TIME EVALUATION

GCC name	Computing time (µs)		Difference (%)
	Short frames	Long frame	
Cross-Correlation	184	224	21.74
ROTH	228	268	17.54
SCOT	295	335	13.56
CSP	282	314	11.35
CSP-m	593	688	16.02
Eckart	538	558	3.72
HT (ML)	528	566	7.20
HB	353	383	8.50
Wiener	422	432	2.37
ρCSP	599	630	5.18
ρCSPC	826	850	2.91
acc-CSP	247	314	27.13
acc-ρCSP	560	630	12.50
acc-ρCSPC	788	850	7.87

The computing time for all the methods is presented in Table V. We compare 2 integration schemes. For the first one we use 4 frames of 1024 samples each, and for the second we use one large analysis frame of 4096 samples. Last column from Table V is reserved for the relative computing time difference between the two integration schemes.

It is easy to see that the lowest computing time is achieved for the basic cross-correlation method. It is the fastest one because this method does not have to apply any weighting. Depending on the complexity of each weighting function, computing time increases accordingly, and the slowest method is ρCSPC. It is about four times slower than the basic cross-correlation, but in the same time, it offers a higher accuracy rate.

Results from Table V confirm that the implementations which use a large analysis frame, is slower than those which divided the large analysis frame in more smaller frames. This is explained by the complexity of the FFT For a large frame size of N samples the FFT complexity is $O(N \cdot \log N)$. If this large frame of N samples is divided in K frames of n samples, then the complexity is $O(K \cdot n \cdot \log(n)) = O\left(N \cdot \log\left(\frac{N}{K}\right)\right)$. This reduction applies for all methods. But, the computing time for the accumulating methods is reduced more, due to the benefit offers by this scheme, which reduces the FFT numbers, form $3 \times K$ to $2 \times K +$ 1, as we presented in Section II.A.

We have to point out that these values can change if different systems are used. Digital signal processor (DSP), Field-Programmable Gate Array (FPGA) or different processor family can improve processing time for the above methods, but they can also change the proportionality between the methods presented in Table V.

The *acc-ρCSP* method has a computing time comparable with all other methods, offering also a high accuracy rate. Comparing with *acc-ρCSPC,* it yields practically the same accuracy, but with a faster computing time. Despite the various constrains that each application imposes for processing time and accuracy, *acc-ρCSP* performances recommend it as an interesting and efficient solution for time delay estimation.

IV. CONCLUSION

In this work, we further evaluated our proposed TDE methods, *accumulated ρ-Cross Power Spectrum Phase* (*acc-*ρCSP) and *accumulated ρ-Cross Power Spectrum Phase with Coherence* (*acc-ρCSPC*), among others well known generalized cross-correlation methods. We used the international *Noizeus* speech corpus database. The results from this work show that our recently proposed combinations between previous *acc-CSP, ρCSP* and *ρCSPC* improve the accuracy rate and the processing time.

We proved that *acc-ρCSP* and *acc-ρCSPC* outperforms popular GCC methods, with a higher accuracy rate. They are also able to estimate accurately even relatively longer delays, around 40% of the frame size length. The methods are also robust and offer accurate results for a large range of signal-to-noise ratio.

When *acc-ρCSP* is calibrated adequately, it practically reaches the same accuracy as *acc-ρCSPC*, but with almost 30% reduction of the processing time. Thus, we propose the *acc-ρCSP* in TDE applications where processing speed and accurate estimations are very important. In this way it provides efficient results in various digital signal applications like echo canceling, speech enhancement, realign noisy signals, pattern detection, radar and sonar localization, medical and seismic processing.

Future work will involve development of *acc-ρCSPC* and *acc-ρCSP* applications for the VoIP environment. Specific analysis will also involve methods characterization for different system implementations.

REFERENCES

[1] J. Benesty, "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization", J. Acoust. Soc. Am. Volume 107, Issue 1, Jan. 2000, pp. 384-391.

[2] B. Widrow and S.D. Strearns, "Adaptive signal processing", Penitence-Hall, ISBN 0130040290, USA, 1985.

[3] S.N. Lin and S.J. Chern, "A new adaptive constrained LMS time delay estimation algorithm", Signal Processing, Volume 71, Issue 1, Nov. 1998, pp. 29-44.

[4] A.W.H. Khong and P.A. Naylor, "Efficient Use Of Sparse Adaptive Filters", In Proceedings of ACSSC '06, Fortieth Asilomar Conference on Signals, Systems and Computers, article ID 10.1109/ACSSC.2006.354982, Nov. 2006, pp. 1375-1379.

[5] R.A. Dyba, "Parallel Structures for Fast Estimation of Echo Path Pure Delay and Their Applications to Sparse Echo Cancellers", In Proceedings of CISS 2008, 42nd Annual Conference on Information Sciences and Systems, article ID 10.1109/CISS.2008.4558529, Mar. 2008, pp. 241-245.

[6] D. Hongyang and R.A. Dyba, "Efficient Partial Update Algorithm Based on Coefficient Block for Sparse Impulse Response Identification", In Proceedings of CISS 2008, 42nd Annual Conference on Information Sciences and Systems, article ID 10.1109/CISS.2008.4558527, Mar. 2008, pp. 233-236.

[7] D. Hongyang and R.A. Dyba, "Partial Update PNLMS Algorithm for Network Echo Cancellation", In Proceedings of ICASSP 2009, IEEE International Conference on Acoustics, Speech and Signal Processing, article ID 10.1109/ICASSP.2009.4959837, Apr. 2009, pp. 1329-1332.

[8] K. Sakhnov, E. Verteletskaya, and B. Simak, "Partial Update Algorithms and Echo Delay Estimation," Communications – Scientific Journal of the University of Zilina, Zilina – Slovakia, vol. 13, no. 2, Apr. 2011, pp. 14-19.

[9] C. Knapp and G.C. Carter, "The Generalized Correlation Method for Estimation of Time Delay", IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 24, issue 4, Aug. 1976, pp. 320-327.

[10] P.R. Roth, "Effective measurements using digital signal analysis", IEEE Spectrum, vol 8, Apr 1971, pp. 62-70.

[11] G.C. Carter, A.H. Nutall, and P.G. Cable, "The Smoothed Coherence Transform", Proc. IEEE (Lett.), vol. 61, Oct 1973, pp 1497-1498.

[12] C. Eckart, "Optimal Rectifier System for Detection of Steady Signals", Univ. California, Scripps Inst. Oceanography, Marine Physical Lab. Rep SIO 12692, SIO Ref 52-11, 1952.

[13] E.J. Hannan and P.J. Thomson, "The Estimation of Coherence and Group Delay", Biometrika, vol. 58, 1971, pp. 469-481.

[14] J.C. Hassab, R.E. Boucher, "Optimum Estimation of Time Delay by Generalized Correlator", IEEE Transaction on Acoustics, Speech, and Signal Processing, vol. assp-27, no. 3, Aug. 1979, pp. 373-380.

[15] A.O. Hero and S.C. Schwartz, "A New Generalized Cross Correlator", IEEE Transactions, Acoustics, Speech and Signal Processing, vol. 33, Feb 1985, pp. 38-45.

[16] D.V. Rabinkin, R.J. Renomeron, A. Dahl, J.C. French, J.L. Flanagan, and M.H. Bianchi, "A DSP Implementation of Source Location Using Microphone Arrays", The Journal of the Acoustical Society of America, Volume 99, Issue 4, Apr 1996, pp. 2510-2527.

[17] M. Matassoni and P. Svaizer, "Efficient Time Delay Estimation Based on Cross-Power Spectrum Phase", European Signal Processing Conference (EUSIPCO), Florence - Italy, Sep 2006.

[18] M. Shean and H. Liu, "A Modified Cross Power-Spectrum Phase Method Based on Microphone Array for Acoustic Source Localization," IEEE International Conference on System, Man and Cybernetics, San Antonio, TX, USA, Oct. 2009, pp. 1286 – 1291.

[19] K. Sakhnov, E. Verteletskaya, and B. Simak, "Echo Delay Estimation Using Algorithms Based on Cross-correlation," Journal of Convergence Information Technology, Volume 6, Number 4, Apr. 2011, pp. 1 – 11.

[20] R.S. Marinescu, A. Buzo, H. Cucu, and C. Burileanu, "Fast Accurate Time Delay Estimation Based on Enhanced Accumulated Cross-Power Spectrum Phase", EUSIPCO 2013, "in press"

[21] R.S. Marinescu, A. Buzo, H. Cucu, and C. Burileanu, "New Considerations for Accumulated ρ-Cross Power Spectrum Phase with Coherence Time Delay Estimation", In Procedings of ICDT 2013, The Eight International Conference of Digital Telecommunications, Apr 2013, pp. 55-59.

[22] J.C. Hassab, R.E. Boucher, "A Cuantitative Study of Optimum and Suboptimum filters in Generalized Correlator", IEEE 1979 Int. Conf. Acoust., Speech, Signal Processing Conf. Rec. 79CH 1379-7 ASSP, 1980, pp. 124-127.

[23] J.C. Hassab, R.E. Boucher, "Performance of the Generalized Cross Correlator in the Presence of a Strong Spectral Peak in the Signal", IEEE Transaction on Acoustics, Speech, and Signal Processing, vol. assp-29, no. 3, Jun. 1981, pp. 549-555.

[24] NOIZEUS: A noisy speech corpus http://www.utdallas.edu/~loizou/speech/noizeus/ [retrieved: Jun, 2013]

[25] AURORA database, http://www.elda.org/article52.html [retrieved: Jun, 2013]