Jurnal Kejuruteraan 32(4) 2020: 723-728
https://doi.org/10.17576/jkukm-2020-32(4)-23

Affine Projection Algorithm for Speech Enhancement using Controlled Projection Order

Ali O. Abid Noor
Department of Communication Engineering,
University of Technology- Iraq

**Corresponding author: 30212@uotechnology.edu.iq*

Received 27 December 2019, Received in revised form 02 March 2020
Accepted 10 April 2020, Available online 30 November 2020

ABSTRACT

This research presents a development of the affine projection algorithm (APA) in voice communication applications. A method of controlling the parameters of the APA is devised to improve the performance in cancelling various types of ambient noise that could possibly corrupt speech signals in voice communication systems. Indicators are used to identify the type of noise accompanying the target signal. Then the corrupted signal is processed in a noise cancellation setup in such a way that three parameters of algorithm are changed according to the nature of the noise. The spreading of elements in the covariance matrix of the noise is used as an indicator for the type of noise so that the projection order, step-size and filter length are changed at the same time. This way the performance of the canceller is improved rendering lower estimation error with a moderate computational power. The method was tested under various types of noise and showed better convergence performance than the original APA as well as other commonly used algorithms in noise cancellation systems. The MSE of the proposed VPAPA method drops to -65 dB in steady state compared to -20 dB using NLMS and just below -30 dB using standard APA with projection order of 8, while the powerful RLS reaches around -60dB under the same environment. The method can be useful for clearer voice communication in variable environmental noise.

Keywords: Adaptive filtering; noise cancellation; affine projection; speech enhancement

INTRODUCTION

The nature of the surrounding noise can adversely affect voice communication in nowadays communication technologies. Real world environmental noise can vary from high decibels of noise to quietness in places such as airports, streets, stadiums and market places (Kuo et al. 2018). It also can have a negative impact on voice controlled audio systems such as the one developed by Kuan & Bukhori (2019). In such cases, the use of usual filtering techniques to remove the noise is not applicable since noise can change from silence to high level of decibels in one place. For that reason, methods of varying filter characteristics are used to cope with the change in noise characteristics, these filters are named as adaptive filters since they can vary the values of their coefficients i.e. adapt themselves to track any change in noise properties

For many years now, several well-known algorithms such as the least mean square (LMS) and the recursive least squares (RLS) or their variants are used control the coefficients of a digital filter, which can be finite impulse response (FIR) or infinite impulse response (IIR) (Diniz 2008). The use of IIR filters is avoided in speech applications due to phase distortion and instability which causes subsequent distortion in the processed signal. Due to its desirable properties, the LMS algorithm is used to control the coefficients of a FIR filter in adaptive filter systems for speech enhancement (Noor et al. 2018). While it can efficiently cope with white noise, the LMS filter shows high residual noise in the filtered signal when subjected to colored noise. This is due to the large spread in the autocorrelation matrix of the noise signal (Ramli et al. 2017). This can raise problems for noise cancellation in mobile telecommunication systems where the user is changing place, which requires fast adaptive filtering and tracking process. On the other hand, the RLS algorithm can work efficiently in these environments, but the price to be paid is the huge increase in the number of computations (Haykin 2013).

More advanced algorithms such as the affine projection algorithm (APA) are mitigation between the two extreme cases (Jiang & Huang 2019; Zhang & Zhe 2011). However, in high order projections, the APA results in large estimation errors and a large number of computations per iteration. Therefore, the need for an APA that can change its projection order according to changes in noise type became necessary. This way, it can reduce noise with lower steady error at lesser number of computations. The early version of the APA was proposed by Ozeki and Umeda (1984), where the reusing of past input vectors and weights updating of input vectors was proposed. The algorithm showed better convergence behavior than the LMS when subjected to colored noise signals. The APA converges faster when the order of projection and the step-size are both high. However, there has been large residual noise in the output i.e. higher estimation error; meanwhile lower residual noise and slower convergence are obtained when both step-size and projection order are of small value (Samarakan & Beex 2000). These issues have been treated in literature in recent years, an example is the fast APA devised by Zhi., et al. (2016), which increases the convergence rate for low projection orders while a small step-size is used. A second example is the evolutionary APA proposed by Kim, et al. (2009) in which, the order of the projection is calculated according to the amount of error in the output and a preset threshold. This algorithm was further modified by Albu et al. (2010) so as to give less number of calculations by using dichotomous coordinate descent method. Furthermore, the original version of the APA was subjected to variable step-size in other proposals such as those by Vega et al. (2008) and by Mayyas & Momani (2011). In earlier studies, there were techniques sought to select the input vector dynamically, in which the estimation error was relatively lower than the original APA (Kong et al. 2007), however, the error was much higher than the error of the LMS which considered as a bench mark for comparison.

Variable projection order of the APA has not been used widely in literature to cure APA issues, because it needs analytical solutions, which may lead to non-practical solutions that use large number of computations when implemented on digital signal processors. Having said that, there have been attempts in literature to use a varying projection order of the APA, such as that proposed by Arablouei & Dogancay (2012). The strategy is based on using combinations of analytical and empirical procedures. The technique was relied on using extensive number of simulations in working out the order of the algorithm. The method has succeeded to some extent, but the problem associated with it is the need for a lot of background analytical processes, and so many repeated simulations in order to change the algorithm order. Other proposals have used a combination of dynamic selection and variable step-size DSVSS-APA (Motar & Noor 2017), the method was applied to echo cancellation in telecommunications, although it showed success in cancelling echo from voice, it uses a compromise between the two parameters based on trial and errors, also entails large estimation errors in low projection order cases. More recent techniques have utilized analytical solutions based on wave domain analysis, and used for cancelling active noise (Zhang et al. 2018), which is somehow complicated.

Therefore, the objective of this research is to develop a simpler and more practical strategy in order to improve the performance of the APA in voice communication systems. The proposed method takes into account the variations in noise properties, using them as indicators to identify the type of noise, so as to change parameters of the APA occasionally according to the change in noise type. The change in environmental noise results in changing the order, the step-size and the filter length, rendering better convergence performance at moderate number of computations of the algorithm. This makes the algorithm suitable for real time DSP implementation. The developed method is named as the variable parameter affine projection algorithm (VPAPA).

METHODOLOGY

A schematic diagram of the proposed VPAPA noise canceller is displayed in Figure 1. The noise canceller uses two-input model in which the voice and the external noise signals forms the desired input $d(n)$. The error $e(n)$ is formed by subtracting the output of the filter from the desired input and it is used to control the coefficient of a direct form FIR filter using the proposed VPAPA.

In noise cancellation circumstances, it is common to use the LMS or one of its versions such as the NLMS as the controlling mechanism due to its simplicity, however as mentioned in the introduction, the LMS accumulates large amount of estimation errors when the noise is not white. Other algorithms such as the RLS and the original APA lead to a large number of computations if the order increased. Therefore, a modified version of the affine projection algorithm is developed in this research and used as the controlling mechanism in the adaptive noise canceller shown in Figure 1. Three parameters of the APA are varied simultaneously according to the received noise by the filter, these parameters are: projection order, step-size and the filter length. In the following, the development of the method is described. Based on the original APA set of equations, the VPAPA is described as follows:

$$\mathbf{w}(n) = \mathrm{w}(n-1) + \mu_v.\mathbf{x}(n)\mathbf{t}(n) \quad (1)$$

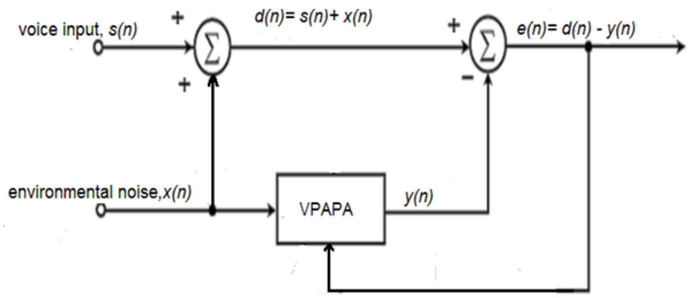

FIGURE 1. The proposed VPAPA noise canceller

where, $\mathbf{w}$(n) is the coefficient weight vector at discrete time n, $\mathbf{x}$(n) is the excitation input matrix ($M_v \times .P_v$), μ_v is a discretely variable step-size, $\mathbf{t}(n)$ and $\boldsymbol{e}(n)$ are given by the following:

$$\mathbf{t}(n) = [\mathbf{x}^T(n)\mathbf{x}(n)+\delta.\mathrm{I}]^{-1}\ \mathbf{e}(n) \tag{2}$$

$$\mathrm{e}(n) = d(n) - \mathbf{x}^T(n)\mathbf{w}(n) \tag{3}$$

where $\boldsymbol{d(n)}$ represents the desired signal, which is the speech signal corrupted with noise and (α,**I**) is a diagonal matrix, with α is a constant along the diagonal which is used to regularize matrix inversion in the algorithm. The parameters M_v and P_v are the filter length the projection order respectively. The subscripts v associated with them is to denote the variable nature of these parameters, in the same way as used for the step-size of the algorithm. T is a transpose natation. For stability μ_v should lie between just above zero and 2, i.e. $0 < \mu_v < 2$. However, values of μ_v larger than 1 leads to instability in the same way as for the LMS. In this work a more conservative limits on μ_v are placed, that is $0 < \mu_v < 1$, this condition is reached after several trials using types of environmental noise that can corrupt voice communication.

In the original version of the APA, the projection order, step-size and filter length are kept fixed throughout the adaptation process, in this research these parameters are made discretely or occasionally variable, in other words they are changed at the same time in a certain pattern for every noise type. The chosen values for these parameters are determined experimentally by prior testing of the noise canceller in various environments. A look-up table is formed for several noise types corresponding to certain noise characteristics. The criterion used to identify a certain type of noise is the spread in eigenvalues in the covariance matrix (**xx**) of a particular noise. Mathematically, this can be expressed as follows:

$$\mathbf{x} = E[\mathbf{x}(n)\mathbf{x}^H(n)] \tag{4}$$

The required criterion is obtained by calculating the eigenvalues from the determinant of the characteristic equation as follows:

$$\det(\mathbf{x} - \text{ë}_i\mathbf{I}) = 0 \tag{5}$$

where $\boldsymbol{I}$ is an identity matrix and λ_i represents the eigenvalues along the diagonal. The target feature i.e. the spread is determined by dividing the largest eigenvalue in equation (5) by the lowest one to render the value of the spread; hence we can identify the type of noise and consequently use a set of parameters in the adaptation process of the proposed noise canceller. For instance if the noise is white, a certain value of the spread is used as a feature, therefore the corresponding values of projection order, filter length and step-size are assigned to the adaptive filter, likewise if the noise is colored the value of the spread is high, a different set of parameters corresponding to this value are assigned. Spreading values were calculated for the several commonly found types of noise.

The aim is to reduce additive external noise from voice communication depending on noise features. The effect of environmental interference is reduced with lower estimation errors at a moderate computational power. In the proposed VPAPA, high projection orders are not used all the time during noise cancellation process. Therefore in cases of white and periodic noise such as engine noise, high orders are not needed, while in colored noise cases, relatively high projection orders are required to achieve desirable convergence. A compromise is achieved by using longer filter lengths and lower step-sizes with lower projection orders to reduce estimation errors. In cases of white and periodic noise, the algorithm is made to act similar to the NLMS in its simplicity and robustness, while for colored types of noise, the algorithm is made to act just like the RLS in its efficiency at lower computations than the RLS. A flow chart of the algorithm is shown in Figure 2. The process starts by acquiring noise data then identifying the type of noise according to the value of the spread, then using a look-up table to set the parameters of the VPAPA, this process continues until the noise samples have been finished.

EXPERIMENTAL RESULTS AND DISCUSSION

The testing of the proposed VPAPA was conducted using human voice subjected to several types of noise assumed to corrupt voice communication in real world. The voice signal is recorded from an utterance by the author counting (One, Two, Three…). This utterance was recorded in a noise-free room using a personal computer microphone, then converted to (.wav) format. The clean segment of the utterance is shown graphically in Figure 3. On the other hand, environmental noise segments are obtained from a trusted data base that is used for conducting experiments to study effects of additive noise on speech recognition systems (Vagra & Steeneken 1993). Segments of various noise types are selected from this data base and used to corrupt the voice utterance. The selected types were; white noise, voice bubbles (also known as cocktail party), engine noise, factory (machinery) noise and colored noise. These segments of different types of noise were concatenated together to form a noise with variable characteristics, then added to the voice signal resulting in a noisy signal as shown in Figure 4. Now, the waveform is ready to be processed by the proposed noise canceller.

Before conducting noise cancellation on the full noisy utterance, the spread values for each type of noise were calculated separately and a table is constituted as in Table 1. Having done that, each noise type was used individually to corrupt the speech segment, then noise cancellation was performed on that particular type of noise. For each type of corrupting noise a set of parameters is obtained for the APA.

These parameters are; projection order, step-size and filter length. The experiments were repeated several times until the best parameters setting were obtained. This has been done for all types of noise used in this research. From these experiments, a second table is constituted as shown in Table 2. In this table, PO is the projection order, SZ is step-size and FL is the filter length. The table was used as a look-up table for the propose VPAPA noise canceller.

Obtaining these tables made the environment for running the proposed noise canceller ready. Noise cancellation experiments using the VPAPA have been carried out on the noisy voice utterance signal of Figure 4. Variable noise is applied to the reference input of the noise canceller shown in Figure 1. The white noise piece is placed at the beginning followed by other types of noise in the following sequence: white, babbles, factory and colored. Data samples from the error output of the noise canceller were stored and processed for performance evaluation of the proposed VPAPA noise canceller. Mean square error MSE plots are used as a measure for the quality of the noise cancellation. To judge the performance of the proposed noise cancellation method, a comparison is made with noise cancellers based on standard APA as well as other commonly used algorithms for noise cancellation, namely the NLMS and the RLS. Figure 5 shows these results. The experiments were repeated several times and MSE samples were averaged over the number of runs and smoothed using a moving average filter.

The main impression from figure 5 is that the proposed VPAPA system has a superior noise cancellation performance compared to other methods. In the standard APA, NLMS and RLS, the filter length is set to 32. All methods showed good start with white noise but when the noise changes to other types the NLMS exhibited a degraded performance, flattening early with large amount of residual noise. The standard APA with projection order of 8 converges better than the NLMS, however it possesses larger amount of estimation error than the proposed VPAPA. The the proposed VPAPA gave the best performance even compared to the RLS which is considered the best for noise cancellation.

The MSE of the proposed PAPA method reduces to nearly -65 dB in steady state compared to -20 dB using NLMS and just below -30 dB using standard APA with projection order of 8, while the powerful RLS reaches only around -60dB on steady state at the expense of higher complexity, proportional to the square of the filter's order. In the proposed method, the VPAPA changes its order, step-size and filter length every time a change in noise characteristics is detected. To confirm the success of the proposed method, a filtered voice utterance using the VPAPA noise canceller is depicted in Figure 6, from which it is clear that the original

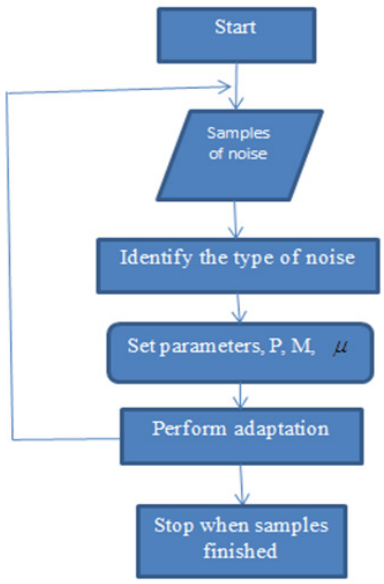

FIGURE 2. Flow chart of the VPAPA

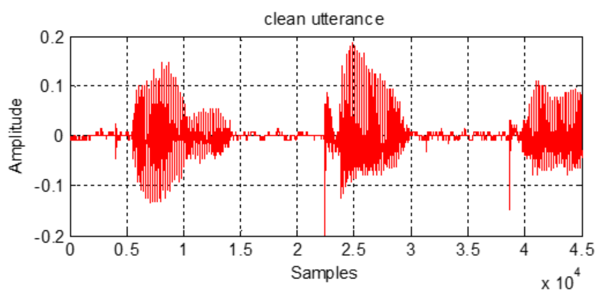

FIGURE 3. Clean voice utterance

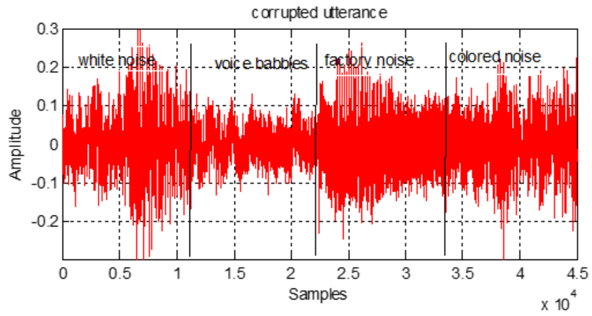

FIGURE 4. Noisy utterance

TABLE 1. Noise types corresponding to spread values

Noise Type	Spread
white	0-5<
Engine, Factory	5-10<
Voice babbles	10-15<
Colored	>15

TABLE 2. Parameters values corresponding to noise types

Noise Type	PO	SZ	FL	Spread	Number of Experiments
White	2	0.1	32	0-5<	8
Engine, Factory	4	0.08	32	5-10<	10
Voice babbles	8	0.02	64	10-15<	12
Colored	16	0.01	64	>15	8

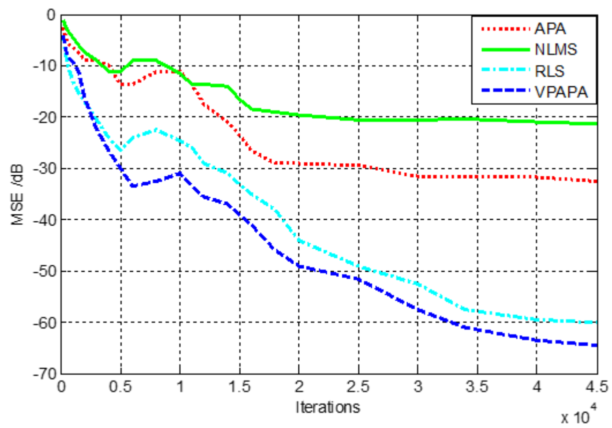

FIGURE 5. Mean square error performance comparison

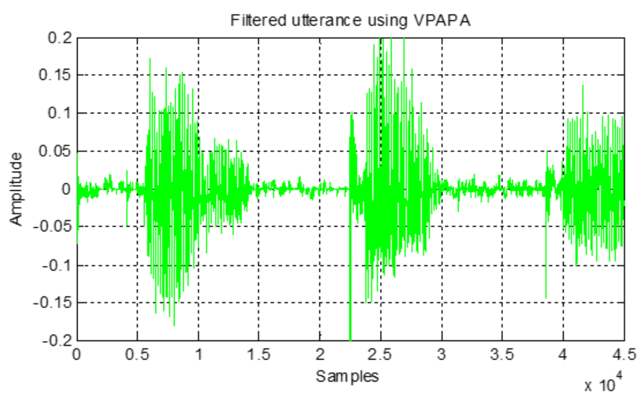

FIGURE 6. Voice utterance filtered using the VPAPA noise canceller

waveform which has been corrupted by noise have now been cleaned nicely. The slight distortion at the wave edges has negligible effect on hearing.

The other main advantage of the proposed VPAPA is the low computation complexity of the proposed method. For types of noise that are hard to remove by the NLMS or by a low order APA, the RLS can perform effectively using a very large number of computations. Given that the RLS noise canceller possesses the maximal number of operations because it involves matrix inversion, the moderate projection order VPAPA performs almost similar if not better than the RLS at a lower number of computations. In the proposed VPAPA, the algorithm processes the four equal segments of noise in equal proportions; therefore it only needs the highest projection order for a limited time segment. Most of the processing time it performs either similar to NLMS when the noise is white i.e. PO equals to 2, or as a low order APA when projection order is 4 or 8.

CONCLUSION

The proposed VPAPA noise canceller possesses the advantages of both good noise cancellation performance and lower complexity than existing techniques. The method has the property that it can adapt to any type of environmental noise. The technique used in this research offers more simplicity than existing methods in recent literature and it can be implemented on digital signal processors for purpose of voice communications in mobile communication or other speech and audio applications.

DECLARATION OF COMPETING INTEREST

None.

REFERENCES

Albu, F., Pale, C. & Benesty, J. 2010. Efficient implementation of a variable projection affine projection algorithm. *Proceedings of the 7th Int. Symposium on wireless communication*, 2010: 369-373.

Arablouei, R. & Dogancay, K. 2012. Affine Projection Algorithm with Variable Projection Order. *Proceedings of the IEEE Int. Conference on communication*, 10-15 June 2012, Ottawa: 3734-3738.

Diniz, S. R. 2008. *Adaptive Filtering: Algorithms and Practical Implementation*. Third edition. Springer Science & Business Media. New York.

Haykin, S. O. 2013. *Adaptive Filter Theory***.** 5th edition. Pearson Press.

Jiang, Z., Li, Y. & Huang, X. 2019. A correntropy-based proportionate affine projection algorithm for estimating sparse channels with impulsive noise. *MPDI Journal of Entropy* (21)555:1-13.

Kim, S. E., Kong, S. J. & Song, W. J. 2009. An affine projection algorithm with evolving order. *Signal Processing* 16(11): 937-940.

Kong, S. J., Hwang, K. Y. & Song, W. J. 2007. An affine projection algorithm with dynamic selection of input vectors. *IEEE Signal Process. Letters* 14(8):529–532.

Kuo, S. M., Chen Y.R., Chang, C.Y. & Lai, C. W. 2018. Development and evaluation of light-weight active noise cancellation earphones. *Journal of Applied Sciences MPDI Applied* Science 8(1178): 1-11.

Mayyas, K. & Momani, F. 2011. An LMS adaptive algorithm with a new step-size control equation. *Journal of the Franklin Institute* 348:589-605.

Motar, S. M. & Noor, A. O. A. 2017. Echo cancellation in telecommunications using variable step-size, dynamic selection, affine projection algorithm. *Iraqi Journal of Computers, Communication and Control & System Engineering* (IJCCCE) 17(1):49-58.

Noor, A. O. A., Al-Hussaini, I. H., Samad, S. A. 2018. Adaptive cancellation of localised environmental noise. *Jurnal Kejuruteraan* 30(2): 179-186.

Ozeki K. & Umeda, T. 1984. An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties. *Journal of Electronics and Communications in Japan* 67-A(2): 126-132.

Ramli, R. M., Noor, A. O. A. & Samad S. A. 2017. Noise cancellation using selectable adaptive algorithm for speech in variable noise environment. International Journal of Speech Technology. *International Journal of Speech Technology* 20(3): 535-542.

Samarakan, M. & Beex, A. 2000. Convergence behavior of affine projection algorithms. *IEEE Transactions on Signal Processing* 48(4): 1086-1096.

Tanaka, M., Makino S. & Kojima, J. 1999. A Block exact fast affine projection algorithm. *IEEE Transitions on Speech and Audio Processing* 7(1):79-86.

Vagra, A. & Steeneken, H. J. M. 1993. Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems. *Speech Communication* 12(3): 247-251.

Vega, L. R., Rey, H., Benesty, J. & Tressens, S. 2008. A new robust variable step size NLMS algorithm. *IEEE Transactions on Signal Processing* 56(5): 1878-1893.

Kuan, W. T. & Bukhori, M. F. 2019. Prototype development of a smart voice-controlled audio system based on the raspberry Pi platform. *Jurnal Kejuruteraan* 31(2): 341-348.

Zhang, J., Abhayapala, T. D., Zhang, W., Samarasinghe, P. N. & Shouda, J. 2018. Active noise control over space: A wave domain approach. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)* 26(4): 774-786.