journal of electrical engineering ISSN 1582-4594

ENHANCED BLIND SPEECH SIGNAL DEREVERBERATION BY UTILIZING A CUCKOO-ICA ALGORITHM

**1 Jasmine J.C Sheeja, AP/ECE, JCT College of Engg and Tech, Coimbatore, mail: jasminejcsheeja@gmail.com*

2 Dr. B. Sankara Gomathi, Prof/ECE, NEC, Kovilpatti, sankara_gomathi@rediffmail.com

**Corresponding Author: jasminejcsheeja@gmail.com*

Abstract

The natural ability of humans to express themselves with their voice has further triggered the development of human-machine interfaces, which allow the control of technical devices using voice commands. Speech dialog systems interact with the user by recognizing and interpreting the meaning of the received commands. The ICA goes to the category of blind source severance (BSS) and the ICA prominently determined by the key assumption of the physical world character. The BSS estimates the original signal using the mixed signal information observed from the input channel. The proposed method avoids the drawback of separated sounds with improper localization, directivity and spatial quality of separate sources. The wavelet filter and multi-step linear prediction coding (mLPC)for extraction of coefficients in the late reverberation. The reverberated signals are eliminated using backward differentiation. Cuckoo is an optimization technique is used to improve the effectiveness of ICA. To reduce the redundant bit and hardware cost the new FPGA was proposed to improve the reliability. Hence the reliability and SNR values are increased. Various methods are compared with proposed Cuckoo search algorithm to demonstrate the efficiency of frequency, time delay and power consumption with reduced area utilization.

Keywords: ICA, FPGA, Speech, Cuckoo

Introduction

Blind source separation using different methods is an important and trending technology due to its good budding tenders in signal processing such as medical image processing, telecommunication and speech signal processing. One of the methods is Independent Component Analysis (ICA) in which the reflected components of all multi-source signals are separated into additive sub elements. Basically, a receiver may receive multiple source components with the linear mixture of reflected signals. When all signals are mixed with each other it is critical to predict the actual speech signal, which can be handled by using ICA algorithm. On the other hand, ICA reduces the higher-order static components and always try to recover the independent components from the mixed signal. Effective outcomes in EEG, fMRI, discourse acknowledgment and confront acknowledgment frameworks demonstrate the control and idealistic trust within the modern worldview.

Along with this method an optimization algorithm called as Cuckoo search algorithm was proposed in this chapter to upgrade the selection strategy of individual components. Concept behind this Cuckoo algorithm is to choose the best coefficients from the adaptive equalization unit.

FPGA implementation of this technique offers additional upshot on this chapter. The ICA and Cuckoo search algorithms are utilized to implement the source separation from the mixture of nonlinear signals. Resource utilization, time delay and power consumption parameters are weighted in the final outcomes of this chapter. FPGA is implemented in XILINX Virttex-7.

Literature review

(Francisco J. Ibarrola, Ruben D. Spies and Leandro Ezequiel 2019) proposed a dereverberation technique with one example. Dereverberation components are achieved by, learning the spectral structure individually, and dedicated to reverberation. (Ina Kodrasi and Simon Doclo 2016) proposed a RPMINT theorem is used to achieve the high-quality speech signal dereverberation in the presence of room impulse response worries but this may add more amplified additive noise in the output of the speech signal. . The result shows that the MWF-DNR technique yields more outcome compared with RPM-DNR technique. (Alejandro Cohen et al 2017)proposed a dereverberation which is based on two paradigms . First paradigm uses a long-term correlation of the signal to reduce it, second paradigm treats the speech signal into a statistically independent data to avoid the noise present in the speech signal. The primary organize comprises of the WPE strategy, and the moment arrange comprises of a Minimum Variance Distortion Less Response (MVDR) beamformer for treating the leftover reverberant component. (Donald S. Williamson and DeLiang Wang 2017) proposed a numerous speech division plans to move forward the greatness reaction of noisy speech. outcomes show that stage is critical for dereverberation, which complex proportion veiling outflanks related strategies. (Tzu-Hao Chen, Chun Huang, and Tai-Shih Chi 2017) proposed a Deep neural network (DNN) based dereverberation algorithm in speech signal will vary the temporal variation of speech content. Recreation result show up that the proposed rate-domain DNN calculation is more able of recovering high-intelligible and high-quality discourse from reverberant discourse than the compared state-of-the-art dereverberation calculation. (Long Zhang et al 2017) proposed a single channels recorded speech severely affected by a noisy signal and the spectrum efficiency of the system is not a good, leading to reduce the quality of the system. The advantages of this method are it can easily enhance the speech signal and the time needed for the RIR is decomposed. (Francisco Ibarrola, Leandro Di Persia and Ruben Spies 2017) proposed a degradation technique. Comparisons of the comes about against state-of-the-art strategies are displayed, appearing significant advancement. (Hideaki kagami, Hirokazu Kameoka and Masahiro Yukawa 2018) proposed that the extension of multichannel non-negative matrix factorization simultaneously solves source separation and dereverberation. (Shahab pasha and Christian Ritz 2015) proposed that for an unsupervised multi-channel approach ad-hoc microphone array for removing microphones with high level of reverberation we use the dereverberation method based on subset of microphones with low level of dereverberation. Based on the kurtosis of linear prediction residual signals the microphones that are closed to the source are detected and utilized for dereverberation process. This method is the clustered enhancement method which can use any dereverberation algorithm. (Yang sun, WenwuWang and Jonathon A.Chambers 2018) proposed the enhanced time-frequency masking by using neural networks for monaural source separation in noisy environment The dereverberation mask has been used for the elimination of reverberations and the ideal ratio mask will be used for the purpose of denoising.to generate speech mixtures for evaluations the IEEE and TMIT corpora with real room responses are used. (Deepak Baby and Hugo Van hamme 2017) proposed a single-channel speech enhancement in noisy reverberant technique which plays an important role in noisy environment. the proposed technique has a great impact for noisy channel and the final simulation outcome shows that this method is better than other previous methods.

Independent Component Analysis (ICA)

Independent component analysis (ICA) may be a strategy for observing essential variables or apparatuses from multivariate (multi-dimensional) arithmetical information. Free component examination is utilized to perform the division by calculating the division lattice values. This works based on considering that the source signals are autonomous so as it were chosen ICA examination. There are a few strategies to execute ICA investigation, the esteem of sources can be extricated from the blend signals. After, changing over flag into frequency domain for preprocess the ICA weighting of the flag is performed. Typically, moreover known as the half portion in ICA, where the relationship between the two signals is removed and the signals are autonomous. This is by calculating the eigenvectors of the covariance lattice of the blended flag and the network with eigenvectors.

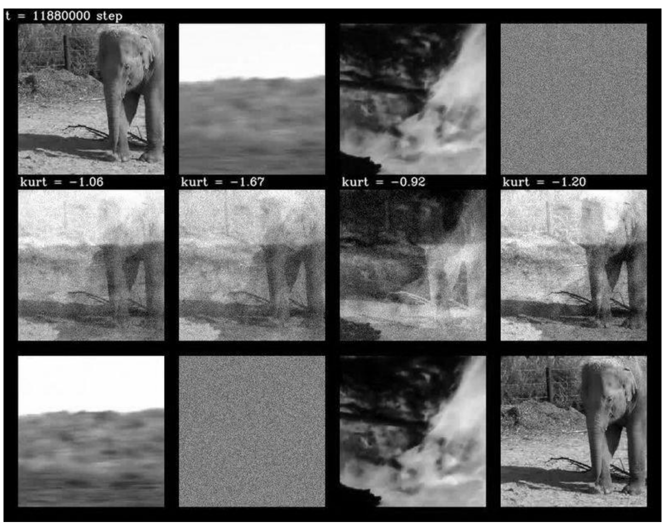

Fig 1: Example for ICA on randomly mixed images

In figure 5.1 four input images are merged together thereby results mixture of different source signals with low resolution.

CS Parameters Optimization ICA

Cuckoo Search Algorithm

Cuckoo Search is another metaheuristic Algorithm for tackling enhancement issue. Cuckoo is a Blood Parasites; they never build its own specific home and lay their eggs in the home of another host winged creature or species.

Each species of cuckoo builds up its very own system to grow the bring forth likelihood of its own eggs
It works on three rules:

(i) Each cuckoo lays one egg at some random minute and dumps its egg in a haphazardly picked home.

(ii)The best egg with the high caliber of eggs will continue to the people to come.

(iii) A Number of assessable hosts homes are settled, an egg laid by a cuckoo are found by the host winged animal with likelihood Pa £ (0, 1)

The point of using cuckoo search calculation is to have better arrangements that are in the settled host. In view of the over three a standard, the CS calculation invigorates the flying creature's best area search way, and they communicated as pursues,

$$X_i^{t+1}=X_i^t+\alpha \oplus Levy \qquad \text{..........} \; (1)$$

Where, X_i^t represent the position of the i^{th} nest at repeatation t, $\oplus$ represent entry-wise multiplication, α is stepwise parameter, L is the Levy flight utilized for a random walk, Xi^{t+1} is generated by utilizing levy fight.

$$x_i^{t+1} = x_i^{(t)} + \alpha \oplus Levy(\lambda) \qquad \text{....} \qquad (2)$$

In the event that $\alpha>0$ is the progression estimate which depends upon the kind of issues, for most case $\alpha=1$ is chosen. Toll flight is essentially an arbitrary walk. Duty can be communicated as,

$$\text{Levy}=0.01\times\frac{\mu}{|v|^{1/\beta}}\times(g_{best}-x_i^t) \quad \ldots \quad (3)$$

Where μ,ν - normal distribution, g_{best} - best nest.

$$\mu\sim M(0,\delta_\mu^2),\nu\sim M(0,\delta_\nu^2)$$

$$\delta_\mu=\left\{\frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma[(1+\beta)/2]\beta 2(\beta-1)/2}\right\}^{\beta/2} \quad \ldots. \quad (4)$$

Where, $\beta=1.5$

CS Algorithm simply has two parameters (M and P_a). M is settled and P_a, controls the balance amid random and local search.

CS based on ICA parameters optimization

In this method, we are using, ICA technique, adaptive equalization and Cuckoo algorithm. These tree strategies are used to achieve the speech signal dereverberation with best performance compared with other techniques. The parameters are evaluated based on power and SNR, resource utilization, finally total proposed work is implemented in FPGA system. Field Programmable Gate Arrays (FPGAs) are semiconductor devices that are created near a matrix of configurable logic blocks (CLBs) joined through programmable interconnects. FPGAs can be reprogrammed to chosen appliance or functionality requirements later manufacturing.

This characteristic differentiates FPGAs from Application Specific Integrated Circuits (ASICs), which are custom factory-made for specific design tasks. Though one-time programmable (OTP) FPGAs are presented, the leading categories are SRAM based which can be reprogrammed.

The regular procedure of CS-ICA is illuminated in the flowchart shown in figure 5.2. The CS algorithm is utilized to optimize the ICA parameters C and γ as take after,

1. Initiate the cuckoo search algorithm and set the number of the nest (M), probability parameters (P_a), the maximum iterations (t_{max}), and the ranges of C and γ.
2. Randomly generate the nest position by utilizing $q_1^0=[x_1^0, x_2^0\ldots x_n^0]^T$. Each nest relates to a set of parameters (C,γ). The fitness evaluation function is defined as follows,

$$I=\sum_{i=1}^{n}(Y\hat{}(i)-Y(i))^2/n$$

Where, Y(i) is the actual value and $Y^{\wedge}(i)$ is the prediction value, and n is the number of preparing samples.

3. Calculate the fitness value of each nest to discover the present best solution,
4. Record the minimum fitness value and its corresponding position.
5. Keep the best solutions from the past generation. Record the position of other nests and calculate the fitness values

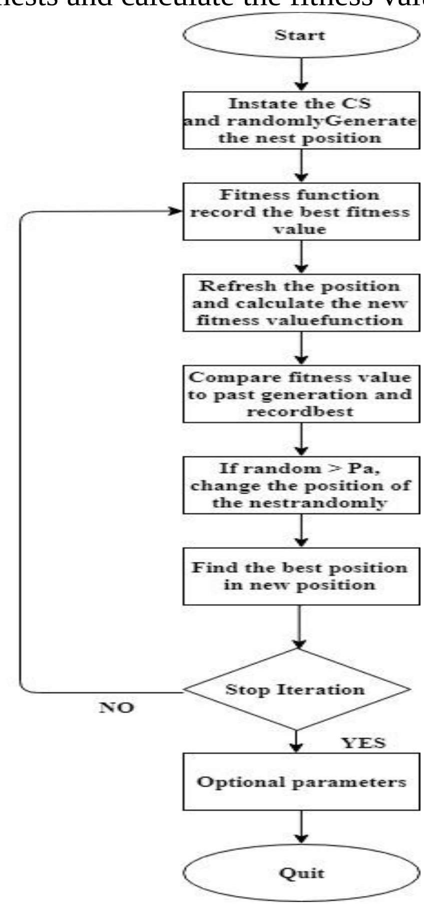

Fig 2: Flow diagram of CS algorithm for ICA Parameter

6. Obtain the better solution for the previous generation which is enhanced than that of the previous generation.
7. Keeps an account of the position of the best nest.
8. Initialize a random number as the probability of egg position.
9. Contrast it with Pa.
10. If random > Pa, alter the location of the nest randomly to obtain a new set of locations.
11. Obtain the best nest location in step (5).
12. Stop searching when the highest iteration limit is reached.
13. output the best position to achieve the optimal parameter value; otherwise, return to step (3).

PROPOSED METHOD

The mixed signals are implied as Y_i recorded through the microphones. The process of signal separation and dereverberation is explained in following steps.

Time to Frequency conversion

Signals in time instants are converted to frequency domain signals. The conversion is offer by using the FFT. The conversion of signals from time to a frequency domain makes the process in frequency-domain BSS approach. After, this signal is whitened for preprocess of ICA.

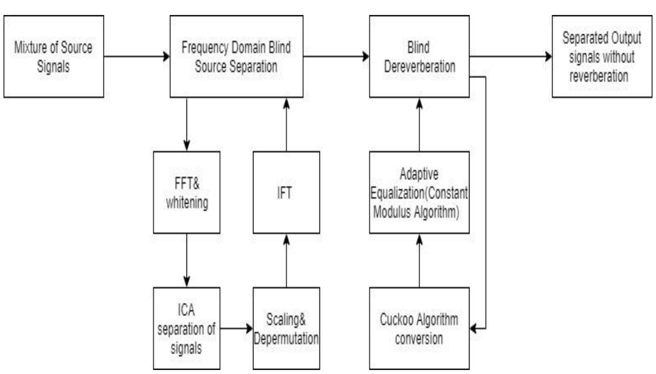

Figure 3: proposed ICA with Cuckoo algorithm

Source separation

Signals are separated by using the ICA analysis. Source signals from the different source mix together and have to perform separation of signals without knowledge of source signal. This can be known as the Blind separation. After separation of the signals again it is converted to the time signals by using IFT.

Alignment of signals

The separated outputs should be aligned by scaling and permutation. Otherwise, signals will be in the form of the shuffled manner.

Dereverberation

To remove the reverberant effect of the separated signals by adaptive equalization technique the CMA Algorithm is employed. After source separation of signals this forms to remove the effect of echoes and reverberation. This is called as the blind dereverberation. The general adaptive filtering method in which the digital filter carries filtering on the original signal, produce an filtered signal.

Cuckoo optimization

Cuckoo is an optimization technique is used to improve the effectiveness of ICA. The dereverberated signals are initially fed into the Cuckoo algorithm convertor in order to lay down the coefficients. The random n number of components are arrived at the convertor with arbitrary signal estimation values. The error estimated signal is now adapted with respect to the channel characteristics by providing a training sequence. A known value of training sequence is transmitted to the receiver through the channel to estimate the channel parameters. By using known transmitter value and the predicted receiver value the channel characteristics are evaluated.

journal of electrical engineering ISSN 1582-4594

With the help of predicted channel parameter, the unknown source corresponding to the original speech signal is estimated thoroughly.

Adaptive Equalization

Thusly, adaptive filter automatically goes on a proposition based the quality of the original signal and the required signal. This strategy of adaptive filter can be altered to the global place by these signals. The CMA is predominantly utilized calculation in adaptive filtering.

It is an inclination plunge calculation; it controls the adaptive filter taps directing them by a sum similar to the quick surmised to the slope from the mistake surface. Its iterative procedure incorporates registering the yield of a Finite Impulse Response (FIR) channel framed by a lot of channel coefficients.

Source Signal Separation

The shuffled speech signals are estimated via channel estimator or adaptive equalizer. At the output node multi-reflected components or reverberated speech signals are received and it might be eradicating the unwanted linear mixture of signals. The dereverberated signal is finally obtained at the output with less amount of noise.

FPGA Implementation

The Cuckoo search optimization implemented in FPGA using Virtex-7 2000T FPGA. This is to improve the low power consumption about 50 percentage than the previous Virtex-6 algorithms. In addition, Virtex-7 2000T FPGA has twice its memory transfer speed compared to past era Virtex FPGAs with 1866 Mbit/s memory meddle execution and over two million rationale cells. It is well-known that the FPGA usage of the Quick ICA calculation is carried out utilizing XILINX virtex5-XC5VLX50t FPGA chip. The LX50t chip has predominant speed and bigger zone over the other virtex5 family. The plan is executed utilizing VHDL dialect.

Since the framework is outlined to account for higher arrange information (four sensors), chain of command is embraced all through the plan to supply distant better; a much better; a higher; a stronger; an improved">a much better control over the generally equipment structure and to screen the flood and sub-current of each square. Besides, execution of DSP frameworks utilizing floating-point math requires a tremendous equipment range and may lead to wasteful plan particularly for FPGA usage on the other hand, fixed-point representation comes about in productive equipment plan.

In this proposal, two's complement fixed-point math, is utilized. The word length was chosen based on a few recreation endeavors. Most of the comes about were flawed when a little word length was utilized since the little word length was not adequate to speak to the values. After a few simulations' endeavors, the choice of the word length was chosen not to be the same for different usage pieces.

For illustration, the QR deterioration piece, the I/O and the middle signals word lengths were set to (26:13) which shows 26 bits with 13 bits speaking to the numbers portion and 13 bits speaking to the fragmentary bits.

This way, the numbers portion can speak to numbers within the run of 2^13 = 8192

.For the Centering and Covariance squares, the word length was set to 16 bits since the calculation of the Centering and the Covariance were not complex and 16 bits were sufficient to speak to for the middle factors like signals and capacity components inside the execution pieces.

RESULT AND DISCUSSION

The error estimation is calculated using output signal $y(n)$ as following equation,

Error signal is, $e(n) = (|y(n)|^2 - 1)^2$ ….(5)

Where, $y(n)$ is the output of the filter and $e(n)$ is the error output.

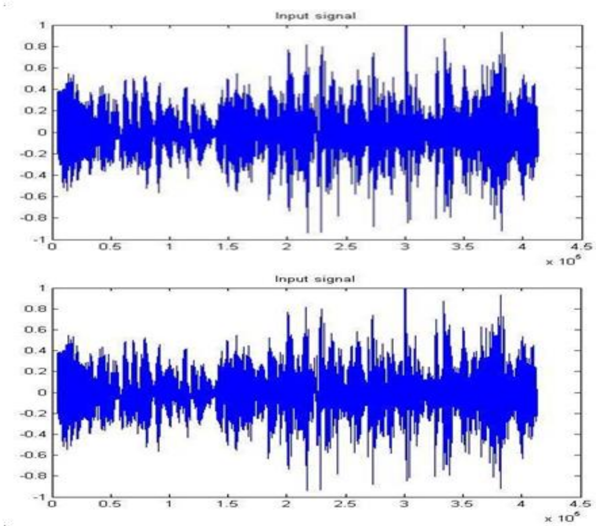

Fig 4: input speech signals

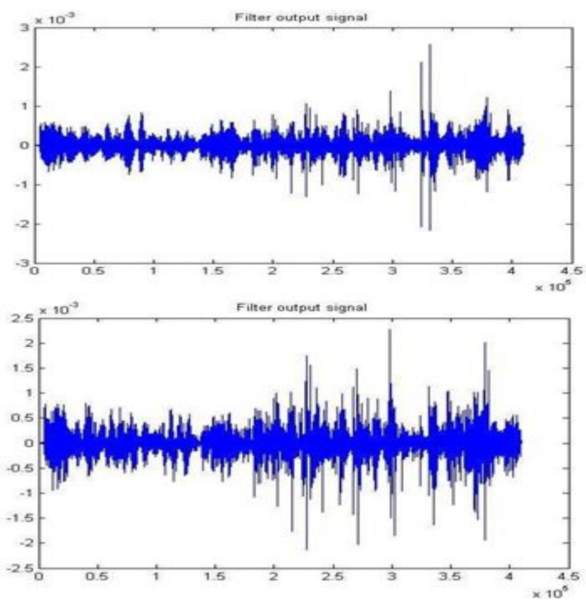

Fig 6: Filtered output signals

Signal to Noise Ratio (SNR)

SNR stands for signal to noise ratio, it is defined as the ratio between signal power to the noise power, which is used to estimate the performance evaluation using the following expression, with input SNR and output SNR, as:

$$SNR_o = 10\, log_{10} \frac{\sum_j [B_j^2]}{\sum_j [B_j - O_j]^2} \quad \dots \quad (6)$$

Where, O_j is the output signal and B_j is the resynthesized signal. The original SNR is given as,

$$\Delta SNR = SNR_o - SNR_i \quad \dots. \quad (7)$$

Table 1: SNR comparison

Window size	SNR_i	SNR_o	LSNR	LSNR
2^8	1.21	7.81	6.60	6.39
2^9	0.98	8.12	7.14	6.70
2^{10}	1.2	7.9	6.7	6.39
2^{11}	1.2	6.91	5.71	5.53
FFT	SNR_i	SNR_o	LSNR	LSNR
2^9	1.04	7.92	6.88	6.34
2^{10}	1.02	7.94	6.92	6.69
2^{11}	1.1	8.31	7.21	6.70
Reverberation Time	SNR_i	SNR_o	LSNR	LSNR
40	1.21	14.21	13.00	12.16
60	1.22	12.41	11.19	9.95
80	1.18	10.51	9.33	8.49
100	1.14	8.41	7.27	6.70
120	1.07	7.81	6.74	5.64
140	0.98	6.01	5.03	4.88
Noise(dB)	SNR_i	SNR_o	LSNR	LSNR
-10	1.04	7.2	6.16	5.9
-20	1.12	7.72	6.6	6.5
-30	1.13	7.81	6.68	6.55
-40	1.12	7.91	6.79	6.59

Area, power and time delay:

During the implementation of FPGA, it is important to consider some parameters such as resource utilization, power consumption of the device and frequency utilization.

The various performance parametric assigned to proposed Cuckoo ICA (Independent Component Analysis) of FPGA implementation is tabulated as shown below,

Technology of ICA	FPGA	FPGA	FPGA	Cuckoo + ICA
Corresponding algorithms	Infomax ICA	ICA	Fast ICA	CICA
Speed in MHz	0.111	10.2	12.29	56
Power consumption	98.56mW	NA	24.992mW	162µW
Resource utilization	35	42	28	19

Table 2: Comparison Results of Different ICA Implementations

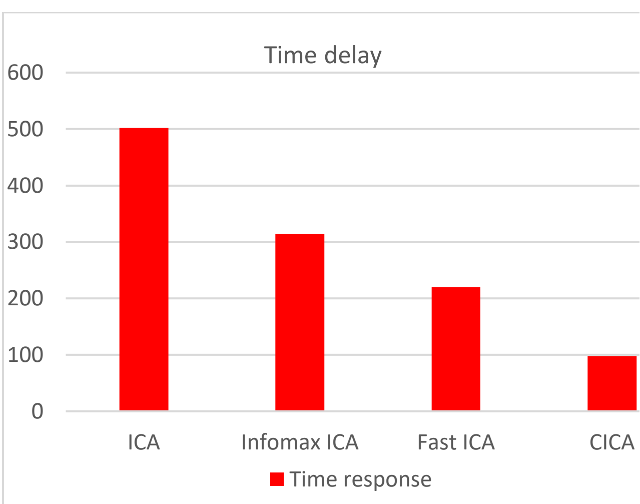

Fig 7: Comparison of time response

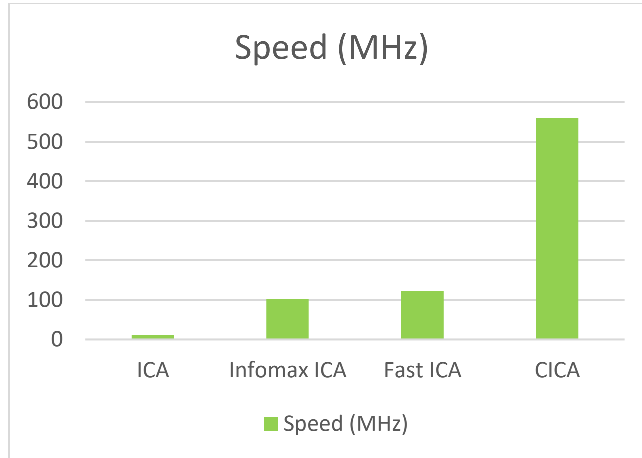

Fig 8: comparisson of speed in MHz

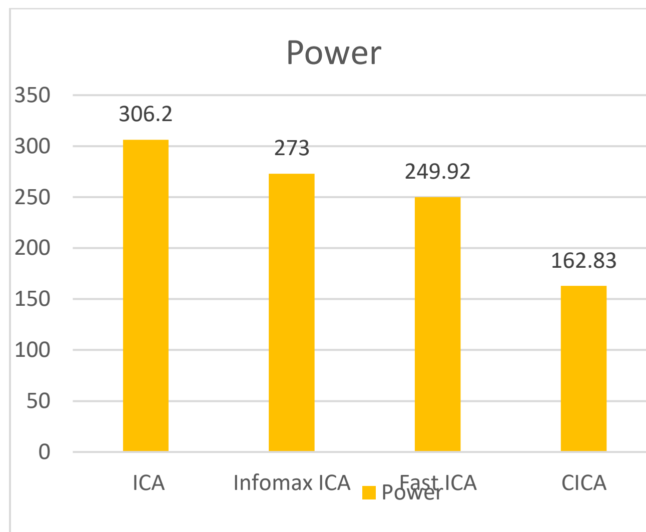

Fig 9: power consumption

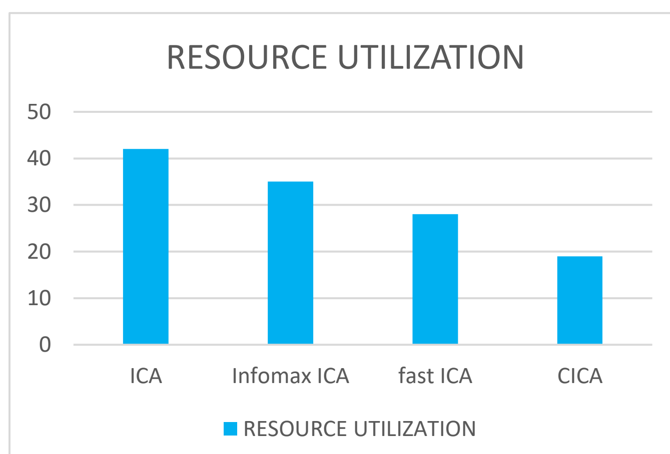

Fig 10: Resource utilization in MHz

When comparing with existing algorithms the proposed Cuckoo algorithm delivers an efficient performance by means of reducing the time delay, power consumption and area utilization as shown in above diagram.

CONCLUSION

In this work an efficient implementation of FPGA architecture has been presented with Cuckoo Algorithm and the result was evaluated. In addition, it gains good area and power consumption for further reduction of reverberation on nosy channel. In previous methods, training sequences was transmitted to know the characteristics of channel. To reduce the redundant bit and hardware cost the new FPGA were proposed and discussed in chapter 5 to improve the reliability. Hence the reliability and SNR values are increased. Various methods are compared with proposed Cuckoo search algorithm to demonstrate the efficiency of frequency, time delay and power consumption with reduced area utilization.

REFERENCES

1. Francisco, J, Ibarrola, Ruben D. Spies, Leandro Ezequiel and E, Di Persia 2019, ‘Switching Divergences for Spectral Learning in Blind Speech Dereverberation’, IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), Vol. 27 (5), Page 881-891.
2. Ina Kodrasi and Simon Doclo 2015,’ Joint Dereverberation and Noise Reduction based on Acoustic Multi-Channel Equalization’, IEEE/ACM Transactions on Audio, Speech, and Language Processing, DOI 10.1109/TASLP.2016.2518804
3. Alejandro Cohen, Georg Stemmer and Seppo Ingalsuo, 2017, ‘Combined Weighted Prediction Error and Minimum Variance Distortionless Response for Dereverberation’, IEEE- ICASSP 2017, pp. 446-450.
4. Donald, S, Williamson, DeLiang Wang 2017, ‘Speech Dereverberation And Denoising Using Complex Ratio Masks’, IEEE- ICASSP 2017, pp.2260-5594
5. Tzu-Hao Chen, Chun Huang and Tai-Shih Chi 2017, Dereverberation Based on Bin-Wise Temporal Variations of Complex Spectrogram’, IEEE, pp.5635-5639.
6. Long Zhang, Jiaxu ehen and You Luo et al 2017, ‘Supervised Single-Channel Speech Dereverberation and Denoising Using a Two-Stage Processing’, 2nd International Conference on Image, Vision and Computing, pp. 818-822.
7. Francisco Ibarrola, Leandro Di Persia and Ruben Spies 2017, On the use of convolutive nonnegative matrix factorization with mixed penalization for blind speech dereverberation’, IEEE.
8. Hideaki Kagami, Hirokazu Kameoka and Masahiro Yukawa, 2018,’ Joint Separation and Dereverberation Of Reverberant Mixtures with Determined Multichannel Non-Negative Matrix Factorization’, ICASSP 2018, pp. 31-35
9. Shahab Pasha and Christian Ritz 2015, ‘Clustered Multi-channel Dereverberation for Ad-hoc Microphone Arrays’, Proceedings of APSIPA Annual Summit and Conference 2015, pp.274-278.
10. Yang Sun, WenwuWang and JonathonA.Chambers et al 2018, ‘Enhanced Time-Frequency Masking By Using Neural Networks For Monaural Source Separation In Reverberant Room Environments’, 26th European Signal Processing Conference (EUSIPCO), pp.1647-1651
11. Deepak Baby, Hugo Van hamme, S 2017, ‘Joint Denoising and Dereverberation using Exemplar-Based Sparse Representations and Decaying Norm Constraint’, IEEE/ACM Transactions on audio, Speech and language Processing, Vol. 25(10), pp.2024-2035.