

An Efficient Isolated Speech Recognition Based on the Adaptive Rate Processing and Analysis

Saeed Mian Qaisar

Electrical and Computer Engineering Department, Effat University, Jeddah, KSA
sqaisar@effatuniversity.edu.sa

Abstract—This paper proposes a novel approach, based on the adaptive rate processing and analysis, for the isolated speech recognition. The idea is to smartly combine the event-driven signal acquisition and windowing along with adaptive rate processing, analysis and classification for realizing an effective isolated speech recognition. The incoming speech signal is digitized with an event-driven A/D converter (EDADC). The output of EDADC is windowed with an activity selection process. These windows are later on resampled uniformly with an adaptive rate interpolator. The resampled windows are de-noised with an adaptive rate filter and their spectrum are computed with an adaptive resolution short time Fourier transform (ARSTFT). Later on, the magnitude, Delta and Delta-Delta spectral coefficients are extracted. The Dynamic Time Warping (DTW) technique is employed to compare these extracted features with the reference templates. The comparison outcomes are used to make the classification decision. The system functionality is tested for a case study and results are presented. An 8.2 times reduction in acquired number of samples is achieved by the devised approach as compared to the classical one. It aptitudes a significant computational gain and power consumption reduction of the proposed system over the counter classical ones. An average subject dependent isolated speech recognition accuracy of 96.8% is achieved. It shows that the proposed approach is a potential candidate for the automatic speech recognition applications like rehabilitation centers, smart call centers, smart homes, etc.

Index Terms –Event-Driven Processing, Speech recognition, Adaptive Resolution Analysis, Features extraction, Dynamic Time Warping, Classification.

I. INTRODUCTION

The Human beings naturally communicate via speech. Speech not only carries the message to convey but it also possesses an important information about the speaker.

The automatic speech recognition process can be traced back in 1960s when computer scientists were discovering algorithms and methods to make computers able to record, interpret and comprehend the human speech. After decades of progress, the first systems were introduced in 1980s. However, these early systems were very limited in scope and performance.

The spoken language is a primary way of communication among humans. Therefore, people expect machines and computers, which can recognize the incoming speech and can act accordingly [1]. The automatic recognition of speech involves generating a sequence of words best matches the incoming speech signal.

The main challenges, faced by any automatic speech recognition system are the confusable sounds, speaker variability and homophones words [1].

The performance of speech recognition systems is typically measured in terms of accuracy. It can be presented in terms of the Word Error Rate and the Command Success Rate [1, 2].

The recent technological advancements have revolutionized the speech based human machine interaction [3, 4]. It has diverse applications like smart cities, secure access, rehabilitation centers, criminals investigation analysis, etc. [5-7]. These systems are founded on the machine based speech comprehension and recognition. The principle is to achieve an appropriate extraction of phonetic features form the incoming speech. Later on, these extracted parameters are employed to classify the incoming speech. It allows realizing effective and robust machine based speech recognition systems [8-10].

The classical speech recognition systems are time invariant in nature [8-10]. It can lead towards a useless usage of the system resources and power consumption [11-15]. An efficient system can be realized for such signals by adjusting the acquisition, processing and transmission rates as a function of the intended signal temporal variations [13-15]. In this framework, an EDADC is employed for the intended speech signals acquisition. It works on the opportunistic sampling principle and therefore allows the devised system to overcome the downsides of the counter classical ones up to a certain extent [16]. Therefore, it promises a simplified and power efficient front-end electronics along with a real time data compression [13-16].

This work focuses on intelligently employing the event-driven and adaptive arte signal processing and analysis tools. It is in the objective to achieve an efficient automatic isolated speech recognition.

The proposed system principle is described in Section II. Section II, also discusses the employed materials and methods, used to realize the devised system. A summary of the system performance verification results is discussed in Section III. Section IV, finally concludes the article.

II. THE PROPOSED SOLUTION PRINCIPLE

Figure 1, shows the proposed solution block level diagram. It shows that the intended speech signals are acquired via an EDADC. These pulses are firstly passed through an antialiasing

band-pass filter with $[F_{Cmin}=50; F_{Cmax}=3.5k]$ Hz. The band limited signal $x(t)$ is acquired with a uniform quantization based 5-Bit resolution EDADC [15, 17, 18]. The EDADC output is time windowed with the Activity Selection Algorithm (ASA) [13, 14]. The windowed signal is resampled uniformly and then de-noised with an appropriate adaptive rate filtering algorithm [14]. The de-noised signal spectrum is computed with an ARSTFT [13]. The spectral magnitude and delta coefficients are extracted. These features are later on employed by the DTW module for the intended incoming speech recognition. A detail of different system modules is provided in the following subsections.

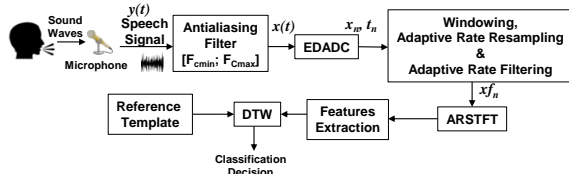


Figure 1: The proposed system block diagram

A. The Event-Driven A/D Conversion (EDADC)

Figure 1, shows that in the studied case the band limited signal $x(t)$ is acquired with a 5-Bit resolution, uniform quantization based, EDADC [15, 17]. The EDADC is developed on the principle of Level Crossing Sampling (LCS) [13-18]. The ADC is an essential component of any signal processing chain [11, 12, 19]. It dictates the overall system performance [19]. The Nyquist sampling and processing theory governs the functionality of traditional ADCs. The signal acquisition is performed at a constant frequency irrespective of its sporadic nature, i.e. without exploiting the local signal variations [13-18]. Consequently, design parameters of the traditional ADCs are chosen for the worst case [11, 19]. Therefore, in the case of low activity random signals like Speech, Seismic signals, etc. such ADCs are not effective [13-18]. The employment of EDADCs can treat this inadequacy up to a certain degree. The EDADCs are founded on the principle of opportunistic sampling and can adjust their sampling frequency according to the local variations of the intended signal. As a result, only the intended information is acquired. Therefore, a drastic reduction in the, post processing and analysis chain, activity and power consumption is achieved. In this case, the sampling frequency is not fixed, and it is piloted by the input signal. However, the signal reconstruction is confirmed by respecting the Bernstein inequality [14]. It is realized by employing some appropriate design parameters of the employed EDADC [15].

B. The Windowing, Resampling and Filtering

The EDADC output can be employed for further non-uniform digital processing. However, the practical system realization necessitates the finite time partitioning of the acquired data. Since a real system has limited resources like memory, processing speed, etc. [11, 12]. In this context, the

ASA is employed for an effective windowing of the EDADC output. It exploits the sampling process non-uniformity to window only the relevant parts of signal. Furthermore, the characteristics of each selected signal part are analyzed in order to extract its local parameters. Later on, these extracted parameters are used to adjust the devised system parameters and processing activity accordingly [13, 14].

The output of ASA is resampled uniformly. The extracted window parameters are employed to decide the resampling frequency [11-12]. The resampler acts as a bridge between the non-uniform and the uniform signal processing domains. It allows to realize smart solutions by taking advantage of both verges [11-15]. The interpolation is required during the resampling process. The artifacts are introduced, because of interpolation, in the resampled signal as compared to the original one. In the suggested system, the resampling error is a function of the interpolation technique employed to resample the data and the EDADC resolution, M , and amplitude dynamics, ΔV [13-15]. The employed EDADC thresholds are distributed uniformly within the range, ΔV . Therefore, its quantum, q , can be calculated as $q = \Delta V / (2^M - 1)$. For this case, the worst interpolation error is bounded by q [20-25]. In the devised system the Simplified Linear Interpolator (SLI) is employed for the resampling purpose.

For the SLI the value of an interpolated sample xr_n corresponding to a resampling instant tr_n is set equal to the average of its prior and next non-uniform samples. The process can be mathematically expressed as: $xr_n = (x_{n-1} + x_n) / 2$. For SLI the worst error per resampled observation is bounded by $q/2$ [20-25].

The resampling frequency of each selected window is adapted by employing the selected window parameters extracted with the ASA [13, 14]. Let Frs^i be the resampling frequency for the i^{th} selected window W^i , then its choice depends on F_{ref} and Fs^i . Here, F_{ref} is the chosen reference sampling frequency in the system, such as it remains greater than and closest to $F_{Nyq} = 2 \cdot f_{max}$. Here, f_{max} is the $x(t)$ bandwidth. $Fs^i = N^i / L^i$ is the sampling frequency for W^i . Here, L^i and N^i respectively present the length in time and the number of non-uniform samples exist in W^i . After resampling, there exist Nr^i uniformly placed samples for W^i .

For the studied case, $f_{max} = 3.5$ kHz. It is because of the employed $F_{Cmax} = 3.5$ kHz. Therefore, $F_{ref} = 16$ kHz is chosen.

The uniformly resampled data is de-noised by applying the adaptive rate filtering approach [14]. In this case, a bank of Finite Impulse Response (FIR) filters is offline designed for a range of sampling frequencies. Later on, during online processing an appropriate filter is chosen, from this bank, during online processing. For the studied case, FIR filters cut-off frequencies are chosen as $[F_{CDmin} = 300; F_{CDmax} = 3k]$ Hz. This choice of F_{CDmin} , removes any DC offset from the incoming speech segment. It improves the classification accuracy [6-10, 26]. Furthermore, it filters out the first harmonics of the pitch information from the incoming speech segment and therefore makes the classification process independent from the speaker's gender.

The filters bank is offline designed by employing the Parks-

McClellan algorithm for a range of sampling frequencies, F_R , between 6kHz to 16kHz and with a uniform frequency step of 1kHz [11]. A summary of the employed filters bank is shown on Table 1. Each filter is designed for the cut-off frequencies of [$F_{CDmin}=300$; $F_{CDmax}=3k$]Hz.

TABLE 1: SUMMARY OF THE REFERENCE FILTERS BANK PARAMETERS

hc_k	$h1_k$	$h2_k$	$h3_k$	$h4_k$	$h5_k$	$h6_k$
FR_c (kHz)	6	7	8	9	10	11
P_c	119	138	158	178	198	217
hc_k	$h7_k$	$h8_k$	$h9_k$	$h10_k$	$h11_k$	
FR_c (kHz)	12	13	14	15	16	
P_c	237	257	277	296	316	

The online filter selection and order adaption feature, for each selected window, allows to achieve the intended signal de-noising with a lesser computational complexity compared to the counter time invariant traditional solutions [14]. It adds to the proposed system performance.

The F_s^i can be specific, therefore an appropriate reference filter is chosen, during online computation, for W^i . The choice of reference filter is made as a function of F_{ref} and F_s^i . If $F_s^i \geq F_{ref}$, then the reference filter, offline designed for F_{ref} , is employed for W^i . Otherwise, if $F_s^i < F_{ref}$ then the reference filter with F_{Rc} closest or equal to F_s^i , is chosen for W^i . The index notation, c , makes a distinction between the chosen reference frequency and F_R . Afterwards, F_{rs}^i is chosen equal to F_{Rc} . For proper filtering operation F_{rs}^i should match to F_{Rc} . The Algorithmic State Machine (ASM) chart of selecting F_{rs}^i and keeping it coherent with F_{Rc} is shown on Figure 2.

This adaptation of F_{rs}^i makes to resample W^i closer to the Nyquist rates or at sub-Nyquist rates [13, 14]. Therefore, it avoids unnecessary interpolations and filtering operations during the data resampling and de-noising processes. As a result, it improves the proposed approach computational complexity and power efficiency.

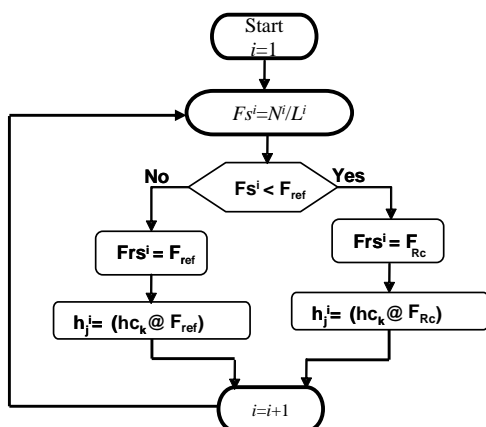


Figure 2: The ASM chart for choosing F_{rs}^i and a filter from the reference filters bank for W^i .

C. The Adaptive Resolution Short Time Fourier Transform (ARSTFT)

The limitation of short time Fourier transform (STFT) is the fixed time-frequency resolution. In this case, the sampling frequency, F_s , and effective window length, L_{ref} , remains fixed regardless of the $x(t)$ local variations [13]. Therefore, the number of samples, N_{ref} , for each windowed segment also remain same. It results in a fixed time $\Delta t = L_{ref}$ and frequency resolution $\Delta f = F_s / N_{ref}$.

It shows that, for a fixed F_s , Δf can be increased by increasing N_{ref} . However, increasing N_{ref} requires to increase L_{ref} which will cause to reduce Δt . Therefore, a larger L_{ref} provides better Δf but poor Δt and vice versa. This conflict between Δf and Δt is the reason for the creation of the multi resolution analysis techniques, which can provide good Δt for high frequency events and good Δf for low frequency events, which is the best suited analysis for most of the real life signals.

The ARSTFT is an appealing substitute of the multi resolution analysis approaches. It makes an adaptive time-frequency resolution analysis, which is not realizable with the classical STFT. It is attained by adjusting the F_s^i , F_{rs}^i , L^i , N^i and N_{ref}^i as a function of the $x(t)$ local variations. As a result, Δt^i and Δf^i can be specific for W^i . Furthermore, the variation of F_{rs}^i and L^i also adds to the computational gain of the ARSTFT compared to the classical one. It is achieved firstly by avoiding the unnecessary samples to process and secondly by avoiding the use of the cosine window functions [13].

D. The Parameters Extraction

The classifiable parameters are extracted from the incoming speech spectra. The spectral coefficients only present the static characteristics of the intended speech word. However, the human auditory system sensitivity is higher for the dynamic characteristics of sound [1-10]. Following this approach the intended word dynamic features are derived from its spectral coefficients. In this context, the Delta and Delta-Delta coefficients are computed. The Delta coefficients are obtained by computing the first order derivative of the spectral coefficients. The Delta-Delta coefficients are computed as the second order derivative of the spectral coefficients. It increases the dimension of the intended speech segment feature vector and thereby increases the classification accuracy.

E. The Classification Method

The considered spectral, Delta and Delta-Delta coefficients are employed for the features matching. The DTW algorithm is employed for the intended speech recognition. It compares computed coefficients of the speech segment under test with the reference template.

A matrix of order M by E is created whose (i, j) element is the distance $d(y_i, r_j)$ between points y_i and r_j of two temporal sequences. Here y is representing the word under test and r is representing one of the reference templates. Euclidean method is employed to calculate the distance between extracted features of the word under test and the saved templates. In this way, the word under test is classified by finding the least distance

between features of word under test and the reference templates [26].

III. RESULTS AND DISCUSSION

In order to demonstrate the proposed system functionality, a case study is conducted. In this context an isolated speech words database is used. The employed database is composed of 5 auditory commands: Forward, Reverse, Left, Right and Center. These are pronounced by 10 different speakers including 5 males and 5 females. The recording is carried out in a noise controlled environment. Each pronounced word is recorded for a duration of 3.0 Seconds. Each speaker 30 times pronounced every command.

Each digitized word is processed and analyzed in order to extract its features. Later on, these features are employed for the classification purpose. The process is clear from Figure 1.

An example of the auditory command ‘Forward’ is considered for the illustration purpose. The digitized word ‘Forward’, acquired in the classical case is shown on Figure 3-a. In classical case, a 16-Bit resolution A/D converter is used. It acquires the signal at a rate of 16k samples per second (SPS). The signal acquired by using the EDADC and selected with the ASA is shown on Figure 3-b. The uniformly resampled and de-noised window is shown on Figure 3-c.

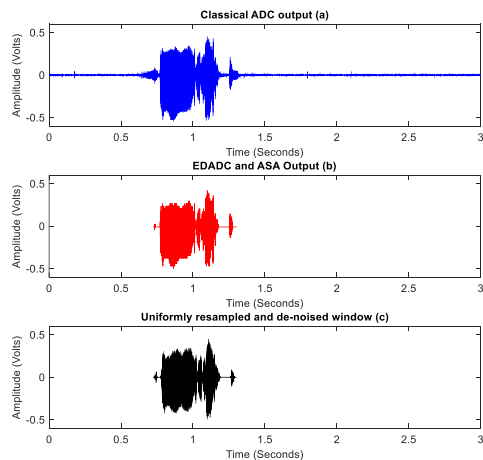


Figure 3: The signal obtained at the output of classical ADC (3-a), the signal obtained at the output of EDADC and ASA (3-b), the uniformly resampled and de-noised window (3-c).

Figure 3, shows that the EDADC and ASA allows to acquire and window only the relevant signal part. Moreover, during acquisition process the EDADC removes the low amplitude noise across the signal base line. This phenomenon is called the noise thresholding and it improves the post speech classification accuracy [7-9]. Furthermore, an additional signal enhancement is achieved with the de-noising process, realized with an adaptive rate filtering method [14]. The process is clear from Figure 4.

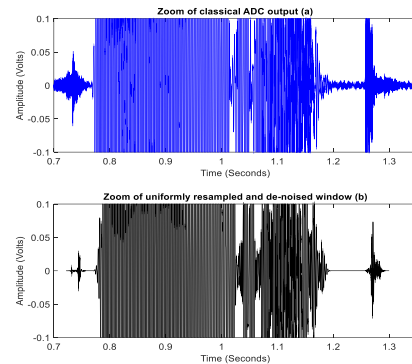


Figure 4: The zoom of signal obtained at the output of classical ADC (4-a), the zoom of uniformly resampled and de-noised window (4-b).

In this case, the superior bound on the selected windows lengths is chosen as $L_{ref} = 1$ seconds [13, 14]. Parameters of selected windows, obtained respectively for an utterance of each intended auditory command, are summarized in Table 2.

TABLE 2: SUMMARY OF THE SELECTED WINDOWS PARAMETERS

Auditory Command	L^i (Sec.)	N^i (Samp.)	Fs^i (Hz)	Frs^i (Hz)	Nr^i (Samp.)
Forward	0.5772	7742	13413	13000	7504
Backward	0.5808	4454	7668	8000	4646
Left	0.4783	5398	11285	11000	5261
Right	0.4605	6176	13411	13000	5987
Center	0.5003	6749	13490	13000	6504
Auditory Command	P^i	Δt^i (Sec.)	Δf^i (Hz)		
Forward	257	0.5772	1.73		
Backward	158	0.5808	1.72		
Left	217	0.4783	2.09		
Right	257	0.4605	2.17		
Center	257	0.5003	1.99		

Table 2 displays the appealing features of the devised solution. The sampling frequency is adjusted as a function of the $x(t)$ local variations. It is achieved due to the smart features of the EDADC and the ASA. The process is clear from values of Fs^i . The adjustment of the resampling frequency for each selected window improves to the computational gain of the proposed technique. It avoids the unnecessary interpolations during the resampling process as clear from values of Frs^i . P^i demonstrates that how the unnecessary operations are reduced during the online filtering process for W^i . Nr^i exhibits that how the adjustment of Frs^i avoids the processing of unnecessary samples during the online filtering and spectral computation processes. L^i shows the dynamic feature of the ASA which is to correlate the window function length with the $x(t)$ local variations. Adjustment of L^i renders an adaptive time-frequency resolution of the ARSTFT. The phenomenon is clear from the values of Δt^i and Δf^i in Table 2.

In the classical case, Fs is chosen equal to F_{ref} . The $x(t)$ is constantly sampled at 16kHz regardless of its local variations. It results in an acquisition of unnecessary samples. Besides, the windowing process is not able to select only the active parts of the sampled signal. Therefore, it causes the system to process

unnecessary samples and as a consequence results in an increased computational activity than the devised solution. In the classical case, $L_{ref}=1$ second remains static and is not able to adapt with the signal local variations. In this studied example the fixed $N_{ref}=16000$ is obtained for each window. It results in a fixed $\Delta t=1$ second and $\Delta f=1$ Hz for all windows.

An overall reduction in the acquired number of samples, achieved by the proposed solution compared to the counter classical one, is 8.2 times. It promises a noticeable decrease in the computational complexity and the power consumption of the proposed solution compared to the classical one.

In the studied case, 1500 sound waves are recorded for 5 considered commands. These 1500 utterances are tactfully employed in order to prepare the training and testing data sets. Later on, the 10-fold cross-validation process is employed while measuring the classification precision.

The obtained subject dependent percentage recognition accuracies are summarized in Table 3.

Table 3: Subject dependent accuracy for the 5 class isolated speech recognition

Auditory Command	Classification Accuracy (%age)	Average Classification Accuracy (%age)
Forward	96.6	96.8
Reverse	97.4	
Left	96.1	
Right	97.1	
Center	96.8	

Table 3, shows that the obtained subject dependent auditory commands classification accuracies are respectively 96.6% for 'Forward', 97.4% for 'Reverse', 96.1% for 'Left', 97.1% for Right and 96.8% for 'Center'. It results into an average subject dependent isolated word recognition accuracy of 96.8%.

The summary of obtained subject independent percentage recognition accuracies is presented in Table 4. It shows that the percentage recognition accuracy is respectively 93.3% for 'Forward', 92.8% for 'Reverse', 91.7% for 'Left', 94.1% for 'Right' and 92.9% for 'Centre'. It results into an average subject independent isolated word recognition accuracy of 95%.

Table 4: Subject independent accuracy for the 5 class isolated speech recognition

Auditory Command	Classification Accuracy (%age)	Average Classification Accuracy (%age)
Forward	93.3	93
Reverse	92.8	
Left	91.7	
Right	94.1	
Center	92.9	

It concludes that the obtained subject dependent recognition accuracy is 3.8% higher than the average subject independent recognition accuracy. It concludes that the proposed system is more suitable for the mono user applications as compared to the poly user ones. Moreover, the proposed speech recognition module performance is comparable in terms of the subject dependent and subject independent recognition accuracy, as compared to the existing state of the art solutions [13, 8, 9, 20-

25].

IV. CONCLUSION

A novel, event-driven signal acquisition and adaptive rate signal processing and analysis based, method is devised for the Arabic isolated speech recognition. It is shown that how the employment of EDADC and ASA have avoided the unnecessary samples to process. It results into an 8.2 times reduction in the number of samples, obtained by the proposed system as compared to the traditional one. It assures that the devised system will lead towards a drastic computational complexity reduction as compare to the counter classic ones. It will lead towards the design and development of low power and efficient automatic speech recognition systems. The output of ASA is resampled uniformly and de-noised respectively by using the adaptive rate SLI and adaptive rate FIR filtering. The EDADC and ASA allows these modules to focus and process only on the important signal parts at an adaptable processing rate. It improves the system computational complexity and processing efficiency. The de-noised signal spectrum is computed by using the ARSTFT. It provides an efficient and adaptive resolution time-frequency representation of the de-noised windows. Later on, the spectral magnitude, Delta and Delta-Delta coefficients are extracted for each window and are employed for the classification purpose.

It has been shown that the proposed speech recognition module achieves an average subject dependent accuracy of 96.8%. The average subject independent isolated word recognition accuracy is 93%. It expresses that the average subject dependent recognition accuracy is 3.8% higher than the average subject independent recognition accuracy. It concludes that the proposed system is a better candidate for mono user applications as compared to the poly user ones.

The devised system performance depends on the type of employed interpolator, de-noising stage, parameters extraction module and the classification algorithm. A study on the devised system performance while employing the high order interpolators like polynomial, spline, etc. is a future work. Moreover, study the system performance while employing other robust classification techniques like Rotation forest, State Vector Machine, Artificial Neural Networks, Random forest, etc. is another future task. Employment of ensemble classifiers can further improve the system classification accuracy but at a cost of an increased processing load. Exploring this approach is another research axis.

V. FUNDING

This project is funded by the Effat University with the decision number UC#7/28Feb 2018/10.2-44f.

VI. ACKNOWLEDGEMENT

Author is thankful to anonymous reviewers for their valuable feedback.

REFERENCES

- [1] Gaikwad, S. K., Gawali, B. W., & Yannawar, P. (2010). A review on speech recognition technique. *International Journal of Computer Applications*, 10(3), 16-24.
- [2] Sarma, M., & Sarma, K. K. (2015). Speech Recognition in Indian Languages—A Survey. In *Recent Trends in Intelligent and Emerging Systems* (pp. 173-187). Springer, New Delhi.
- [3] Bennett, Ian M., et al. "Distributed real time speech recognition system." U.S. Patent No. 9,076,448. 7 Jul. 2015.
- [4] M. E. Deisher, R. P. Banginwar, and R. C. Knauerhase, "Speech recognition command via an intermediate mobile device," Feb. 2007.
- [5] E. C. Anderson, "Automatically storing and presenting digital images using a speech-based command language," Dec. 2002.
- [6] B. Burger, I. Ferrané, F. Lerasle, and G. Infantes, "Two-handed gesture recognition and fusion with speech to command a robot," *Autonomous Robots*, vol. 32, no. 2, pp. 129-147, 2012.
- [7] Wang, K., An, N., Li, B. N., Zhang, Y., & Li, L. (2015). Speech emotion recognition using Fourier parameters. *IEEE Transactions on Affective Computing*, 6(1), 69-75.
- [8] Price, M., Glass, J., & Chandrakasan, A. P. (2015). A 6 mW, 5,000-word real-time speech recognizer using WFST models. *IEEE Journal of Solid-State Circuits*, 50(1), 102-112.
- [9] Li, L., Xu, W., Wu, J., He, S., & Li, X. (2014, December). The Hokkien isolated word recognition system based on FPGA. In *Anti-counterfeiting, Security, and Identification (ASID), 2014 International Conference on* (pp. 1-5). IEEE.
- [10] Liu, L. (2015). Acoustic models for speech recognition using Deep Neural Networks based on approximate math (Doctoral dissertation, Massachusetts Institute of Technology).
- [11] Smith, S. (2013). *Digital signal processing: a practical guide for engineers and scientists*. Newnes.
- [12] Leupers, R. (2013). *Retargetable code generation for digital signal processors*. Springer Science & Business Media.
- [13] Mian Qaisar, S. (2009). *Échantillonnage et traitement conditionnés par le signal: une approche prometteuse pour des traitements efficaces à pas adaptatifs* (Doctoral dissertation, Grenoble INPG).
- [14] Qaisar, S. M., Fesquet, L., & Renaudin, M. (2014). Adaptive rate filtering a computationally efficient signal processing approach. *Signal Processing*, 94, 620-630.
- [15] S.M. Qaisar et al., "An Efficient Signal Acquisition with an Adaptive Rate A/D Conversion", IEEE, ICCAS'13, Kuala Lumpur, Malaysia, 2013.
- [16] Guan, K. M., & Singer, A. C. (2007, April). Opportunistic sampling by level-crossing. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on* (Vol. 3, pp. III-1513). IEEE.
- [17] Akopyan, F., Apsel, A., & Manohar, R. (2008). U.S. Patent No. 7,466,258. Washington, DC: U.S. Patent and Trademark Office.
- [18] Tang, W., Osman, A., Kim, D., Goldstein, B., Huang, C., Martini, B., ... & Culurciello, E. (2013). Continuous time level crossing sampling ADC for bio-potential recording systems. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(6), 1407-1418.
- [19] Van de Plassche, R. J. (2013). *CMOS integrated analog-to-digital and digital-to-analog converters* (Vol. 742). Springer Science & Business Media.
- [20] Qaisar, S. M., Akbar, M., Beyrouthy, T., Al-Habib, W., & Asmatullah, M. (2016, June). An error measurement for resampled level crossing signal. In *Event-based Control, Communication, and Signal Processing (EBCCSP), 2016 Second International Conference on* (pp. 1-4). IEEE.
- [21] Qaisar, S.M., 2011. Electronic management system for rechargeable battery has measuring circuit measuring parameter determining variation of parameter transmitting data to electronic processing unit if variation is higher than predetermined threshold.
- [22] Qaisar, S. M. (2015, June). An event driven technique for filtering computational complexity reduction. In *Event-based Control, Communication, and Signal Processing (EBCCSP), 2015 International Conference on* (pp. 1-4). IEEE.
- [23] Qaisar, S. M., Dallet, D., Benjamin, S., Desprez, P., & Yahiaoui, R. (2013, May). Power efficient analog to digital conversion for the Li-ion battery voltage monitoring and measurement. In *Instrumentation and Measurement Technology Conference (I2MTC), 2013 IEEE International* (pp. 1522-1525). IEEE.
- [24] Qaisar, S. M., Simatic, J., & Fesquet, L. (2017, May). High-level synthesis of an event-driven windowing process. In *Event-Based Control, Communication and Signal Processing (EBCCSP), 2017 3rd International Conference on* (pp. 1-8). IEEE.
- [25] Qaisar, S. M., Ben-Romdhane, M., Anwar, O., Tlili, M., Maalej, A., Rivet, F., ... & Dallet, D. (2017, May). Time-domain characterization of a wireless ECG system event driven A/D converter. In *Instrumentation and Measurement Technology Conference (I2MTC), 2017 IEEE International* (pp. 1-6). IEEE.
- [26] Eljhani, M. M. (2015). *Front-end of Wake-Up-Word Speech Recognition System Design on FPGA*. Florida Institute of Technology.