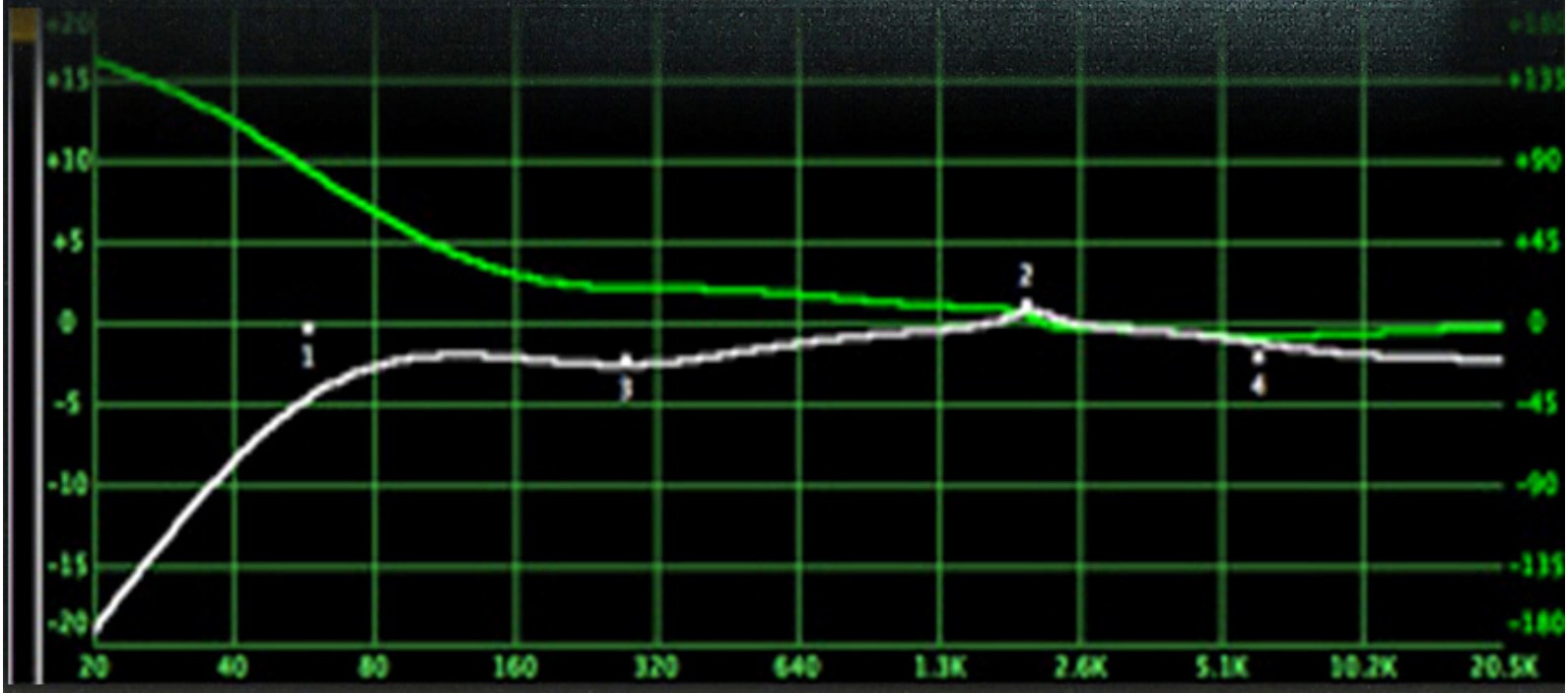
Signal Processing:
An International Journal (SPIJ)

ISSN : 1985-2339

VOLUME 4, ISSUE 6

PUBLICATION FREQUENCY: 6 ISSUES PER YEAR



Signal Processing: An International Journal (SPIJ)

Volume 4, Issue 6, 2011

Edited By
Computer Science Journals
www.cscjournals.org

Editor in Chief Professor Hu, Yu-Chen

Signal Processing: An International Journal (SPIJ)

Book: 2011 Volume 4 Issue 6
Publishing Date: 08-02-2011
Proceedings
ISSN (Online): 1985-2339

SPIJ Journal is a part of CSC Publishers
http://www.cscjournals.org

Published in Malaysia

Typesetting: Camera-ready by author, data conversation by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers

Editorial Preface

This is fifth issue of volume four of the Signal Processing: An International Journal (SPIJ). SPIJ is an International refereed journal for publication of current research in signal processing technologies. SPIJ publishes research papers dealing primarily with the technological aspects of signal processing (analogue and digital) in new and emerging technologies. Publications of SPIJ are beneficial for researchers, academics, scholars, advanced students, practitioners, and those seeking an update on current experience, state of the art research theories and future prospects in relation to computer science in general but specific to computer security studies. Some important topics covers by SPIJ are Signal Filtering, Signal Processing Systems, Signal Processing Technology and Signal Theory etc.

This journal publishes new dissertations and state of the art research to target its readership that not only includes researchers, industrialists and scientist but also advanced students and practitioners. The aim of SPIJ is to publish research which is not only technically proficient, but contains innovation or information for our international readers. In order to position SPIJ as one of the top International journal in signal processing, a group of highly valuable and senior International scholars are serving its Editorial Board who ensures that each issue must publish qualitative research articles from International research communities relevant to signal processing fields.

SPIJ editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for SPIJ. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. SPIJ provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members
Signal Processing: An International Journal (SPIJ)

Editorial Board

Design and Implementation of Low Ripple Low Power Digital Phase-Locked Loop

M. Saber mohsaber@kairo.csce.kyushu-u.ac.jp
Department of Informatics
Kyushu University 744 Motooka,
Nishi-ku, Fukuoka-shi, 89-0395, Japan

Y. Jitsumatsu jitsumatsu@inf.kyushu-u.ac.jp
Department of Informatics
Kyushu University 744 Motooka,
Nishi-ku, Fukuoka-shi, 89-0395, Japan

M. T. A. Khan tahir@apu.ac.jp
Ritsumeikan Asia Pacific University, College of Asia pacific Studies
1-1 Jumonjibaru, Beppu, Oita, 874-8577, Japan

Abstract

We propose a phase-locked loop (PLL) architecture which reduces double frequency ripple without increasing the order of loop filter. Proposed architecture uses quadrature numerically–controlled oscillator (NCO) to provide two output signals with phase difference of $\pi/2$. One of them is subtracted from the input signal before multiplying with the other output of NCO. The system also provides stability in case the input signal has noise in amplitude or phase. The proposed structure is implemented using field programmable gate array (FPGA) which dissipates 15.44 mW and works at clock frequency of 155.8 MHz.

Keywords*:* Digital Phase-Locked Loop (DPLL), Field Programmable Gate Array (FPGA), Numerically-controlled Oscillator (NCO), Read Only Memory (ROM), Look-Up Table (LUT).

1. INTRODUCTION

A PLL is a closed-loop feedback system that sets fixed phase relationship between its output phase and the phase of a reference input. It tracks the phase changes that are within its bandwidth. Tasks performed by PLL include carrier recovery, clock recovery, tracking filters, frequency and phase demodulation, phase modulation, frequency synthesis and clock synchronization. PLLs are also used in radio, television, every type of communications (wireless, telecom, datacom), all types of storage devices and noise cancellers [1].

A PLL block diagram is shown in Fig 1. It has three main components which are, phase detector (PD), a low pass filter (LPF) called as the loop filter and a voltage controlled oscillator (VCO). PLL operates as a negative feedback loop. VCO generates a signal at center frequency, multiplied by the input signal in PD and the resultant is passed through LPF to eliminate double frequency ripple. The filter output is fed back to VCO to adjust its generated frequency and phase. This process continues until no phase or frequency difference exists and the PLL is said to be “frequency locked”.

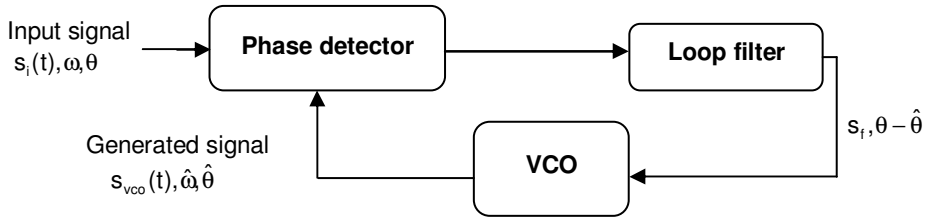


FIGURE 1: Phase Locked Loop

PLL behaves as a narrow band tracking filter and its LPF output has a characteristics of frequency discriminator. The linearity of VCO affects the overall linearity of PLL in case of analog implementation [2]. Hence, digital PLLs can solve some of the limitations of analog ones. In addition, the digital tangent method can compute frequency from the ratio of in-phase and quadrature (I-Q) signals [3,4,5]. Some other problems associated with the analog loops like sensitivity to d.c. drifts and the need for initial calibration and periodic adjustments can also be alleviated using DPLL. Nonuniform sampling DPLLs are the most ones because they are simple to implement and easy to model [6]. Digital tanlock loop (DTL), proposed in [7], has introduced several advantages over other nonuniform sampling digital phase locked loops. It allows a wider locking range of the first-order loop and a reduced sensitivity of the locking conditions to the variation of input signal power [7].

Noise is an extremely important issue in the field of PLL applications and is of two main types. Most common is the additive white Gaussian noise (AWGN) which is added to the signal at every component of communication system. Fig. 2 shows an input sinusoidal signal with AWGN at 35 dB signal to noise ratio (SNR). Second type of noise in PLL is called as the double frequency ripple, which is generated by phase detector. Since phase detector multiplies two sinusoidal signals, its output will be two terms, one is at low frequency and the other is at higher frequency (2ω). The situation will be critical when the two noise sources are involved with the feedback loop. This will affect the operation of PLL especially when it is used as a synthesizer, as the generated frequencies will have much noise or jitter.

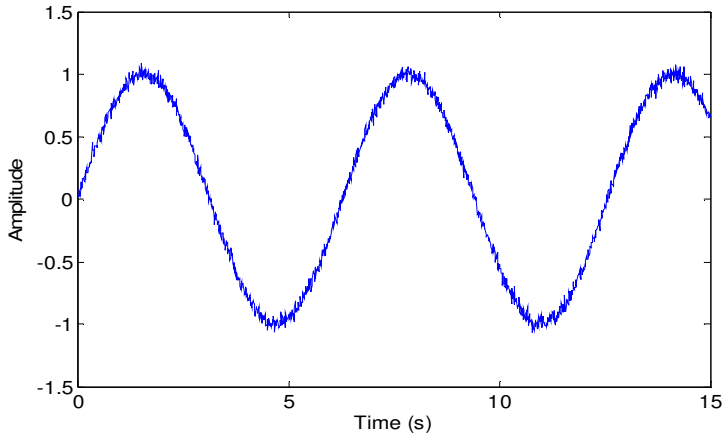


FIGURE 2: Input signal with AWGN at 35dB SNR

LPF in the loop is used to eliminate double frequency terms resulting from multiplying the input signal with generated signal from VCO. In practical circuits, there is no ideal LPF and a residual noise will be present at the output of phase detector. Using higher order loop filters offer better noise cancellation. However, in this case generally the PLL becomes unstable since its order is always higher by 1 than that of loop filter. Noise cancellation performance can also be improved

by lowering cut-off frequency of first order LPF. This will cause lower bandwidth and lower damping of the PLL and will also have a negative impact on its locking ability.
This paper presents an improved PLL which helps in suppressing noise without increasing the order of LPF. Computer simulations performed using Matlab show significant performance improvements over conventional PLL. Furthermore, the system is also modeled in VHDL and implemented using FPGA. This hardware implementation further confirms the results obtained through computer simulations.

2. Mathematical Analysis of PLL

In this section we'll perform mathematical analysis of a conventional PLL shown in Fig. 1. We assume an input sinusoidal signal

$$s_i(t) = A_i \sin(\omega t + \theta(t)) = A_i \sin\psi(t), \quad (1)$$

where ω is the angular frequency and $\theta(t)$ is the unknown phase of input signal. The signal generated by VCO is

$$s_{vco}(t) = A_o \cos(\hat{\omega} t + \hat{\theta}(t)) = A_i \cos\hat{\psi}(t), \quad (2)$$

where $\hat{\omega}$ is the estimation of angular frequency of VCO and $\hat{\theta}(t)$ is the estimated phase of VCO.

There are several designs and construction methods for phase detector. For the present discussion, we assume that PD is a multiplier. Input signal is multiplied by the VCO output, then

$$\begin{aligned} s_d(t) &= k_m s_i(t) \times s_{vco}(t) \\ &= \frac{A_i A_o k_m}{2}[\sin((\omega-\hat{\omega})t + \theta(t) - \hat{\theta}(t)) + \sin((\omega+\hat{\omega})t + \theta(t) + \hat{\theta}(t))] \\ &= k_d\,[\sin(\psi(t) - \hat{\psi}(t)) + \sin(\psi(t) + \hat{\psi}(t))], \end{aligned} \quad (3)$$

where k_m is the gain of phase detector with dimension [1/V], $k_d = \frac{A_i A_o k_m}{2}$.

In the simplest case we assume that low-pass filter removes the upper sideband with frequency $\omega+\hat{\omega}$ but passes the lower sideband $\omega-\hat{\omega}$. VCO's tuning voltage will be

$$s_f(t) = k_d \sin(\psi(t) - \hat{\psi}(t)) = k_d \sin\psi_e(t), \quad (4)$$

where $\psi_e(t)$ is the phase difference between input and output VCO signals

$$\psi_e(t) = \psi(t) - \hat{\psi}(t). \quad (5)$$

This difference will be used to control the frequency and phase generated by VCO. If the error signal is zero, VCO produces just its free running frequency (ω_c, center frequency). If the error signal is other than zero, then VCO responds by changing its operating frequency.

$$\hat{\omega}(t) = \omega_c + k_o s_f(t), \quad (6)$$

where the constant k_o is the gain of VCO in units (2πHz / V). After integration of the above equation and substituting into (5), the phase difference is

$$\psi_e(t) = \psi(t) - \omega_c - \int_{-\infty}^{t} k_o s_f(\tau)\, d\tau. \quad (7)$$

This can be rearranged as follows:

$$\psi_e(t) = \omega\, t - \omega_c t - \int_{-\infty}^{t} k_o k_d \sin\psi_e(\tau)\, d\tau. \tag{8}$$

Differentiating (8) w.r.t. 't' gives

$$\frac{d}{dt}\psi_e(t) = \Delta\omega - K\sin\psi_e(t) \tag{9}$$

where $\Delta\omega = \omega - \omega_c$ and $K = k_o k_d$ is the gain of PLL having units $[2\pi Hz]$.

The PLL continues to vary the phase of VCO $\hat{\omega}$ until locked, that is, frequency and phase of the input signal are the same as those generated by VCO. After getting locked, PLL follows the changes in frequency and phase of input signal [9,10].
It can be concluded from the above analysis that the phase lock arrangement is described by the nonlinear equation (9). Solution of this equation is not known for arbitrary values $\Delta\omega$, and k. Without an aperiodic solution, the feedback system (PLL) cannot achieve phase stability, i.e., output frequency of VCO $\hat{\omega}$ will never be equal to input frequency (ω). Simplifications are needed to solve the equation. One solution is the linear solution, in which we assume that, for small values of $\psi_e(t)$

$$\sin\psi_e(t) \approx \psi_e(t) \tag{10}$$

Substituting in (9)

$$\frac{d}{dt}\psi_e(t) = \Delta\omega - K\psi_e(t) \tag{11}$$

If $\frac{d}{dt}\theta(t) = 0$, solution of this differential equation is

$$\psi_e(t) = e^{-Kt}(\psi_{e0} - \frac{\Delta\omega}{K}) + \frac{\Delta\omega}{K} \tag{12}$$

where $\psi_{e0} = \psi_e(0) = \theta(0) - \hat{\theta}\,(0)$.
For steady state, that is, for $t \to \infty$, the left hand side of (11) is equal to zero, with the result that

$$\psi_{e\infty} = \frac{\Delta\omega}{K}. \tag{13}$$

In the above analysis we assumed that LPF completely suppressed the high frequency term in (3). This operation is not easy for a first order low pass filter and requires a higher order. However, the LPF is part of a feedback control system and instability of the system will increase in this case. Our proposed improvement of PLL aims to reduce noise inside the loop without using higher order LPF.

3. Proposed PLL

We propose modification in conventional PLL to suppress noise without increasing the order of LPF (first order LPF). Fig. 3 shows the block diagram of our suggested architecture. It uses a quadrature numerically controlled oscillator (NCO) to generate two signals $\sin(\hat{\omega}\,t+\hat{\theta}\,(t))$ and $\cos(\hat{\omega}\,t+\hat{\theta}\,(t))$. The phase detector subtracts $\sin(\hat{\omega}\,t+\hat{\theta}\,(t))$ from the input signal and multiplies the resultant by $\cos(\hat{\omega}\,t+\hat{\theta}(t))$ before passing on its output to LPF.

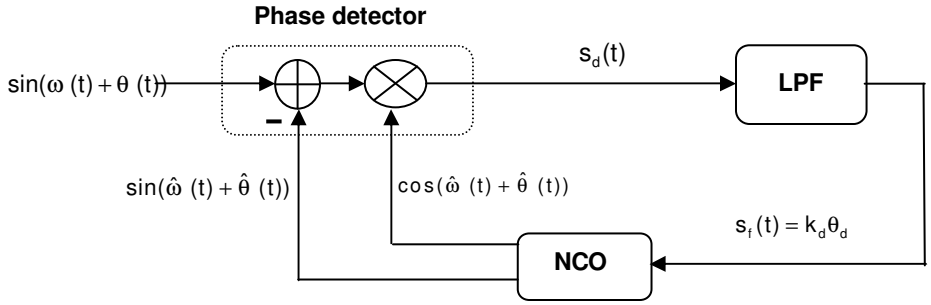


FIGURE 3: Block diagram of proposed PLL

This operation eliminates almost the entire high frequency term and if there is some residual, the first order LPF can remove it easily. Following equations describe the operation of proposed PLL in time domain.

Output signal from the phase detector is

$$\begin{aligned} s_d(t) &= k_d[\sin\psi\,(t)-\sin\hat{\psi}\,(t)]\cos\hat{\psi}\,(t) \\ &= k_d\,[\sin(\psi\,(t)-\hat{\psi}\,(t))+\sin(\psi\,(t)+\hat{\psi}\,(t))-\sin(2\,\hat{\psi}\,(t)] \end{aligned} \tag{14}$$

A comparison of (3) and (14) clearly establishes that in the proposed structure, the high frequency term $\sin(\psi\,(t)+\hat{\psi}\,(t))$ is subtracted by the term $\sin(2\,\hat{\psi}\,(t))$ before passing on the resultant to LPF. Thus even a first order LPF can suppress the noise easily. In conventional PLL, LPF is responsible for removing the term $\sin(\psi\,(t)+\hat{\psi}\,(t))$. First order filter cannot suppress all this term and its output will have large residual high frequency component.

Proposed modification will not have any negative impact on the PLL's locking performance. The term $k_d \sin\psi_e(t)$ is passed to VCO and the process continues like a conventional PLL.

4. Simulation Results

In this section we compare the performances of proposed and conventional PLL's using computer simulations. Two types of simulations are done. In the first group of simulations, we use a first order LPF in both architectures. Both architectures receive a signal with frequency difference (results shown in Fig. 4), phase difference, or both frequency and phase are changed (results shown in Fig. 6). In the second group of simulations, proposed PLL uses a first order LPF, while conventional PLL uses a second order LPF and performances of the two are compared.

4.1 Proposed versus Conventional PLL (Using first order LPF)

The parameters for both PLLs are:

$k_m = 1\ v^{-1}$, $k_o = 1500$ rad / v.s, $f_s = 1.0\times10^5$ Hz, $f_{vco}=1.0\times10^4$ Hz.

LPF: first order with cut-off frequency =1000 Hz,

VCO generates $s_{vco}(t) = \cos(2\pi f_{vco} t)$

First of all we perform comparison of the two PLL structures in case of input signal having frequency difference. For $\omega = 2\pi\times10100$ rad/s and $\theta(t) = 0$ in (1), Fig. 4 shows the response of both PLLs for the input signal. It can be seen that the proposed PLL suppresses much more noise than the conventional one. In Fig. 4 the response of proposed PLL converges to the mean value 0.0389 with variance 1.3926×10^{-4}. While the conventional PLL has the mean value 0.0389 with a high variance of 4.4758×10^{-4}.

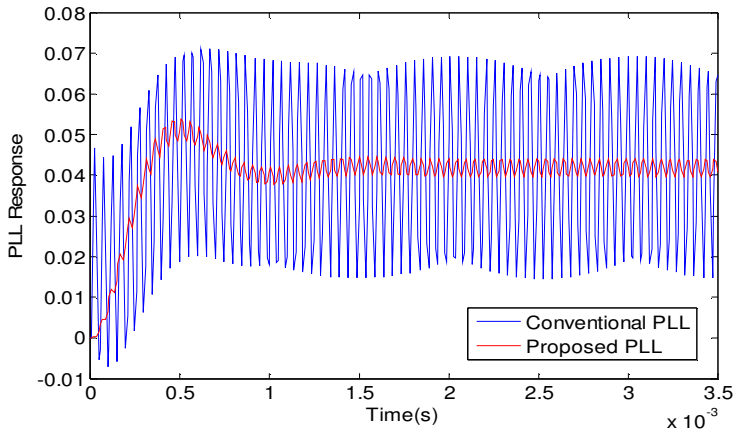


FIGURE 4: Response in case of frequency difference

Fig. 5 shows amplitude spectrum of output signals from each phase detector and each LPF for the two architectures. As can be seen in Fig. 5(a) the proposed architecture attenuates the high frequency term at $f = 2\times10^4$ Hz to an amplitude of 0.05 V before passing signal to the LPF which eliminates this term completely as shown in Fig. 5(b). While in conventional PLL the amplitude of high frequency term is 10 V which is almost equal to the amplitude of low frequency term. After passing signal through LPF high frequency component residual noise still has amplitude of 0.025 V as shown in Fig. 5(d).

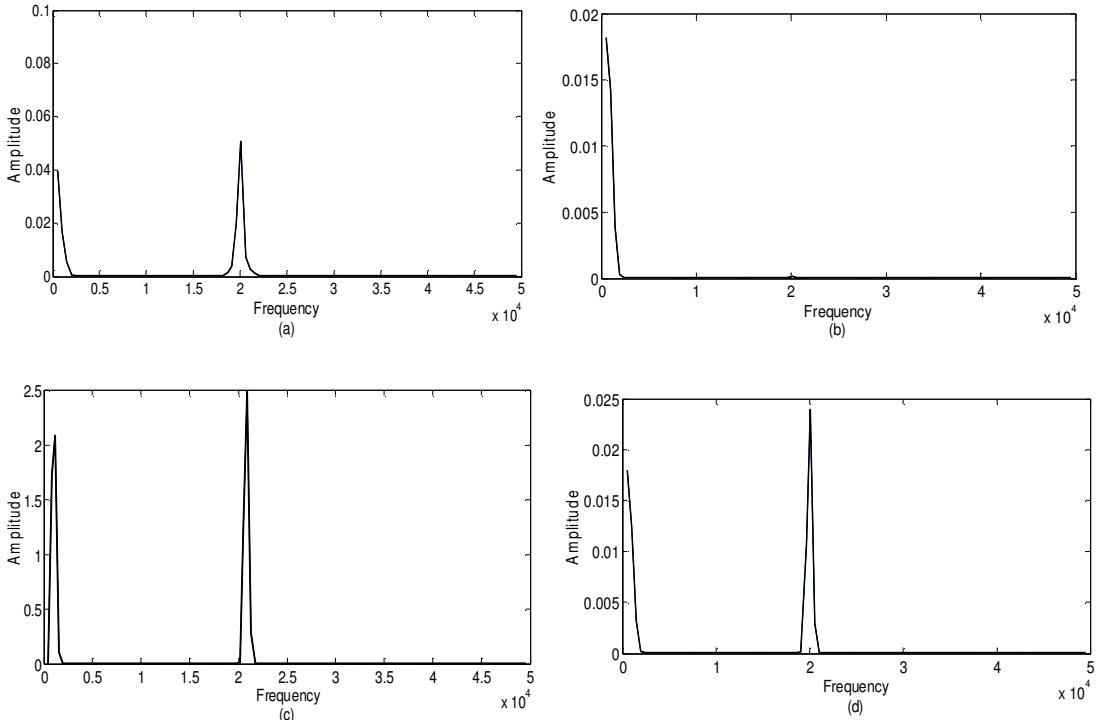


FIGURE 5: Amplitude spectrum of the output signal of (a) PD in proposed PLL, (b) LPF in proposed PLL, (c) PD in conventional PLL, (d) LPF in conventional PLL.

We also performed comparison in case of input signal having both frequency and phase difference. Fig. 6 illustrates the response of both architectures at $\omega = 2\pi \times 10100$ rad/s and $\theta(t) = \pi / 2$ in (1).

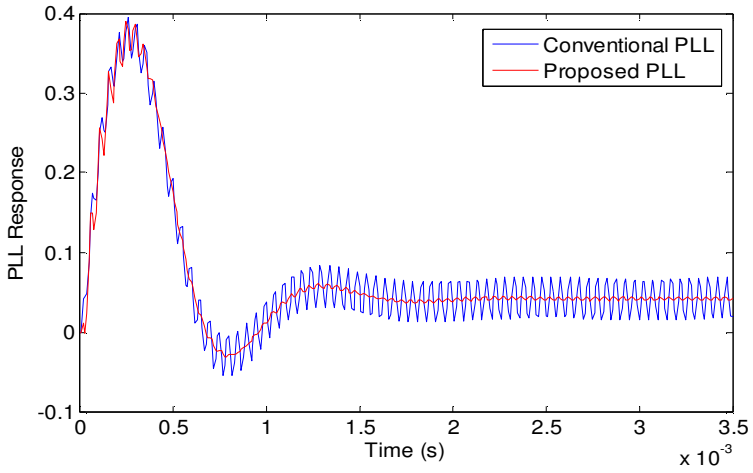


FIGURE 6: Response in case of phase difference

We also performed comparison of the two structures in the presence of AWGN. When the input signal is combined with AWGN at 35 dB SNR, the response of both PLLs is plotted in Fig. 7.

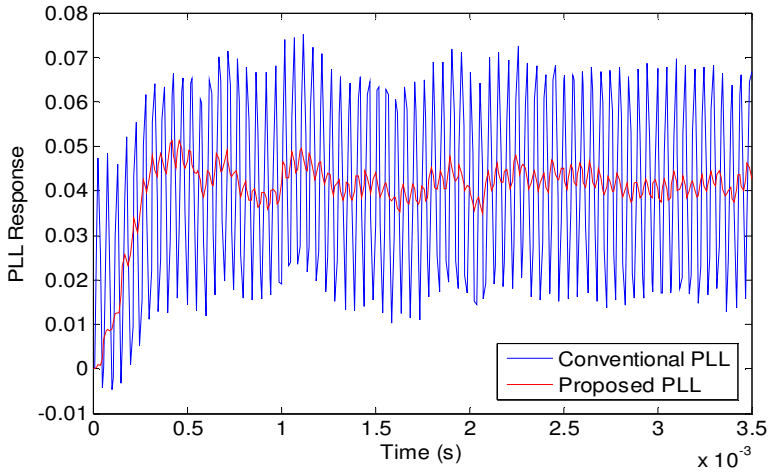


FIGURE 7: Response in the presence of AWGN at 35 dB SNR

4.2 Proposed PLL versus Higher Order Conventional PLL

In this group of simulations, we have used a first order LPF in the proposed PLL while second order LPF is used in conventional one. Fig. 8 shows both responses when the frequency of input signal is changed and the phase stays fixed, i.e, $\omega = 2\pi \times 10100$ rad/s and $\theta(t) = 0$.

Proposed PLL locked to the input signal while conventional PLL did not achieve locking although noise level is reduced. As mentioned before, increasing the order of LPF in the loop increases instability of the system. Therefore, the parameters of conventional PLL must be adjusted carefully to control the location of poles and zeros to achieve stability of higher order PLLs.

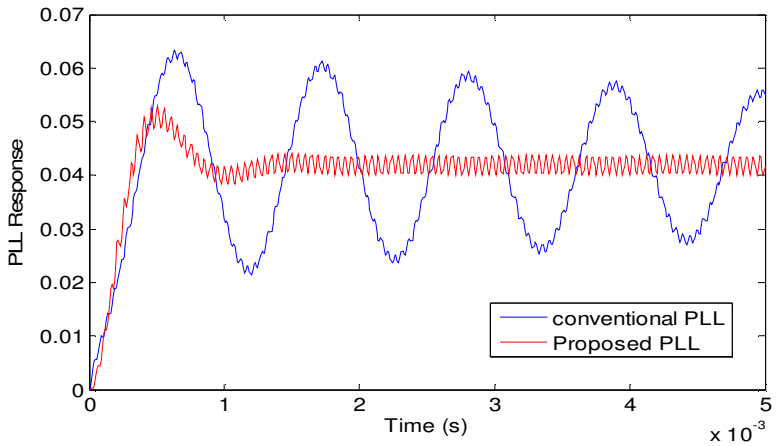


FIGURE 8: Response when conventional PLL has higher order LPF than proposed one

The above simulations clarify and explain operation of the proposed architecture through a comparison with conventional one in different cases such as a frequency difference or phase difference in the input signal as shown in both Fig. 4 and Fig. 6. Furthermore simulations show us the signals' amplitudes in different stages inside both architectures as in Fig. 5. The situation in case the received signal has AWGN is clarified in Fig. 7. In case the order of LPF in conventional PLL is increased, Fig. 8 shows that the noise is suppressed while the PLL loses stability and

unlocked the input signal while proposed one is stable. All previous simulations clearly demonstrate that the proposed PLL has low ripple and is more stable than the conventional one.

5. Hardware Implementation of Proposed PLL

Since Digital Phase Locked Loops (DPLL) find applications in most of the state of the art equipment, we discuss hardware implementation of the digital version of our proposed architecture in this section. Hardware modeling is performed using VHDL language [11,12]. In Fig. 3, the input signal is 8 bits-length, comes from Analog to digital converter (ADC) and generates 8 bits as an output of DPLL architecture. The following sub blocks describe the operation of each part:

5.1 Phase detector (PD)

The Phase Detector detects the phase error between input signal and generated signal from NCO. This operation is done using a multiplier and a register as shown in Fig. 9. In the VHDL model, Booth's multiplication algorithm [13] is used to achieve smaller area and higher speed of multiplication.

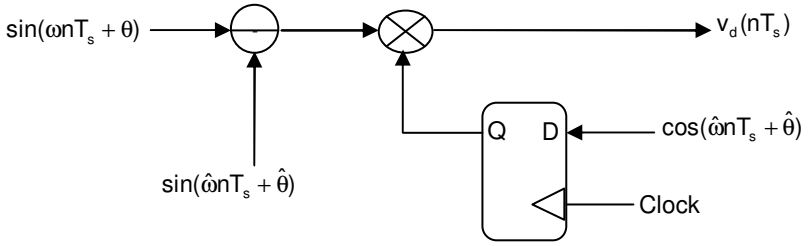


FIGURE 9: Phase Detector module

5.2 Loop Filter

The loop filter is a digital filter responsible for eliminating high frequency term from the output of phase detector unit. The filter is a first order IIR low pass filter. In designing filter it is important to choose the pole, which controls cut-off frequency, to be inside the unit circle to provide stability to the filter. It cannot be equal to 1 because the filter in this case becomes an integrator [14,15]. The transfer function of the filter in z-domain for an input $s_d(nT_s)$ and output $s_f(nT_s)$ is

$$H(z) = \frac{s_f(z)}{s_d(z)} = \frac{kz^{-1}}{1 - pz^{-1}} \tag{15}$$

where k, and p are coefficients of the filter. The structure of digital filter is illustrated in Fig. 10.

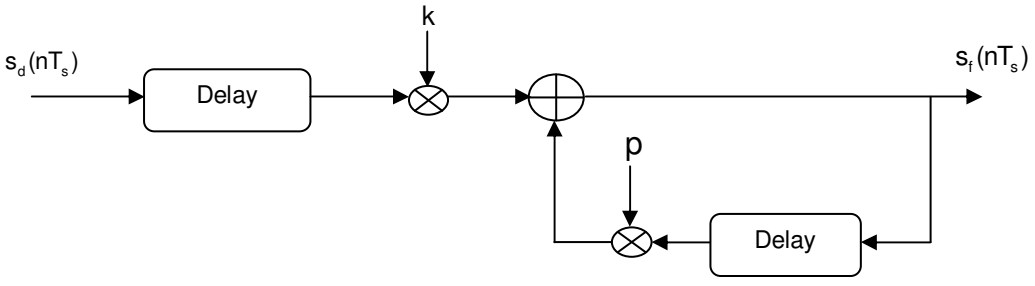


FIGURE 10: First order digital low pass filter

5.3 Quadrature NCO

NCO is an important component in DPLL because it determines power consumption of the entire system. The generation of analog sinusoidal waveform in digital domain requires storing the amplitudes of analog waveform in a read only memory (ROM). The structure of ROM in digital implementation causes high power consumption and slow operation of the circuit. These drawbacks of ROM block limit the use of DPLL in portable applications which require low power consumption [16].

Since it is desirable to have a large number of bits to achieve fine frequency tuning, several techniques have been invented to limit the ROM size while maintaining sufficient performance. One method exploits the quarter wave symmetry of the sine function to reduce by four the number of angles for which a sine amplitude is required. Truncating the phase accumulator output (eliminating a number of most significant bits (MSB) of the output) is another common method, although it introduces spurious harmonics in the generated waveform [17].

Various angular methods have been proposed to reduce memory size [18-19]. They consist of splitting the ROM into a number of smaller units each addressed by a portion of the truncated phase accumulator output bits. Data retrieved from each small ROM is added to yield a sinusoidal approximation.

In this paper instead of using ROM in the quadrature NCO, piecewise linear approximation for the first quarter of the sine waveform is employed. From the first quarter of sine wave a complete waveform for sine and cosine can be generated due to symmetry of both functions. The first quarter of sine waveform which lies between $(0,\pi/2)$ is represented by eight linear equations. Slopes and constants for these linear equations are chosen according to minimum mean square error (MMSE) criterion between the approximated and ideal sine wave.

$$\sin(t) \approx a_i t + b_i, \qquad \frac{i}{16}\pi \le t < \frac{i+1}{16}\pi, \qquad i = 0,1....7 \tag{16}$$

where a_i is segment slope and b_i is segment constant.

The quadrature NCO block diagram is shown in Fig.11. It consists of 3 blocks, the first one is the accumulator, which receives input signal $s_f(nT_s)$ corresponding to phase difference between θ and $\hat{\theta}$. The accumulator width is 16 bits.

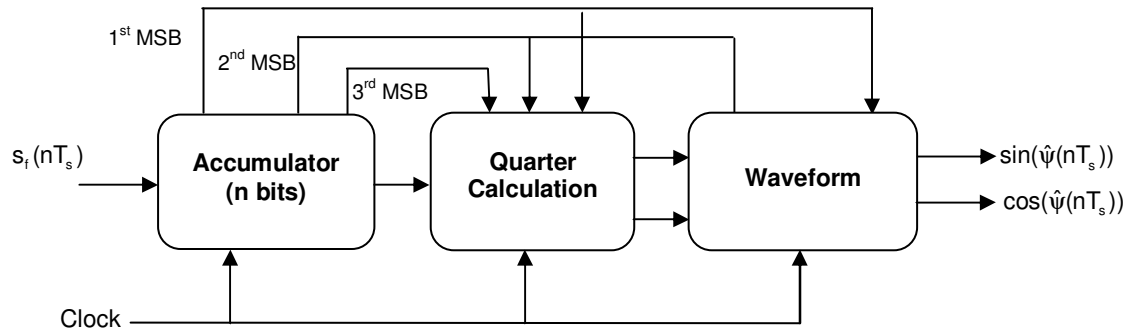


FIGURE 11: Structure of modified NCO

The accumulator works as a circular counter, a complete rotation of accumulator represents one cycle of output waveform. All bits leaving accumulator are directed to Quarter calculation block. The first three MSBs are used to choose between slopes and constants for every linear segment. After calculation of first quarter, the results are directed to waveform block. This block forms

complete sine and cosine waveforms according to the 1st MSB and 2nd MSB of accumulator block. The frequency of generated waveforms f_{nco} is

$$f_{nco} = (\frac{s_f + \omega_c}{2^n}) \times f_{clk}, \quad (17)$$

where f_{nco} is the generated frequency, s_f are the input binary bits to the NCO, ω_c is a constant value which represents the free running frequency of NCO, n is number of bits or width of accumulator which is 16 bits and f_{clk} is the clock frequency which is 50 MHz. When input signal s_f is zero, the architecture generates free running frequency.

Spurious free dynamic range (SFDR) is a measure of spectral purity of the waveform generated by NCO. It represents the ratio between amplitude of the fundamental generated frequency and the amplitude of the largest spur in the dynamic range of NCO $(0 : f_s/2)$ [20,21]. For our proposed NCO, SFDR is 60 dBc which is sufficient not to introduce ripples to DPLL.

Fig. 12 shows the output spectrum for input word of value 1317 representing $v_f(n)$, at a clock frequency of 50 MHz, ω_c=1317 and accumulator width n = 16. The fundamental frequency is approximately 2 MHz with -30.057 dB and the spurious appears at 14.46 MHz with -89.925 dB, so SFDR = 59.868 dBc.

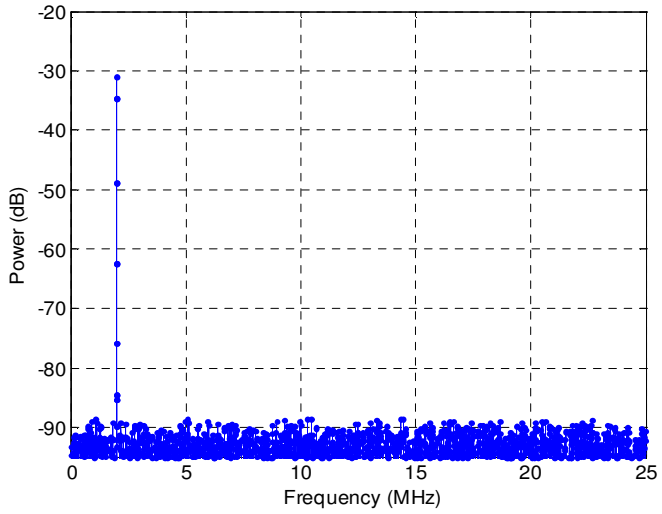


FIGURE 12: SFDR in dBc of fundamental frequency 2 MHz

6. FPGA Implementation Results

The proposed DPLL architecture is written in VHDL and simulated using Modelsim program. The simulation is done at clock frequency 50 MHz, the input is sinusoidal waveform of frequency 1 MHz. Fig. 13 shows the simulation waveforms for proposed PLL. The input signal is multiplied by the modified NCO signal after subtracting the other sinusoidal output from NCO and the resultant is subjected to digital filter. The final output shows that digital simulation agrees with the expected waveform. Fig. 14 shows the response of conventional PLL for the same input signal. It is clear that the proposed PLL suppresses noise more effectively than the conventional one.

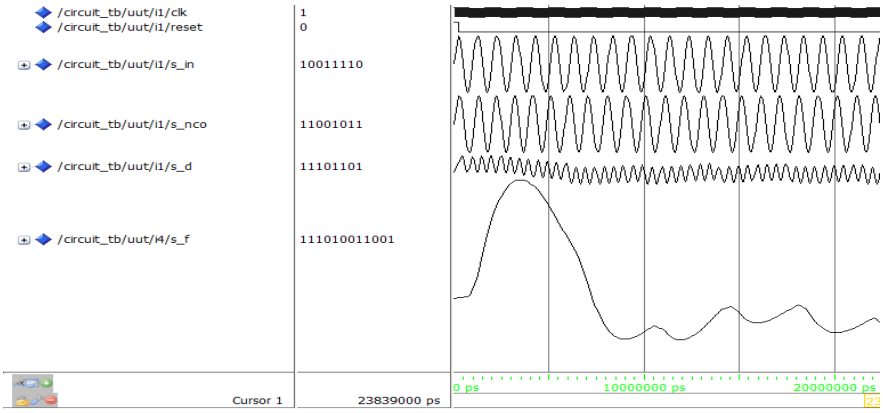


FIGURE 13: Proposed DPLL waveforms for input sinusoidal of 1 MHz

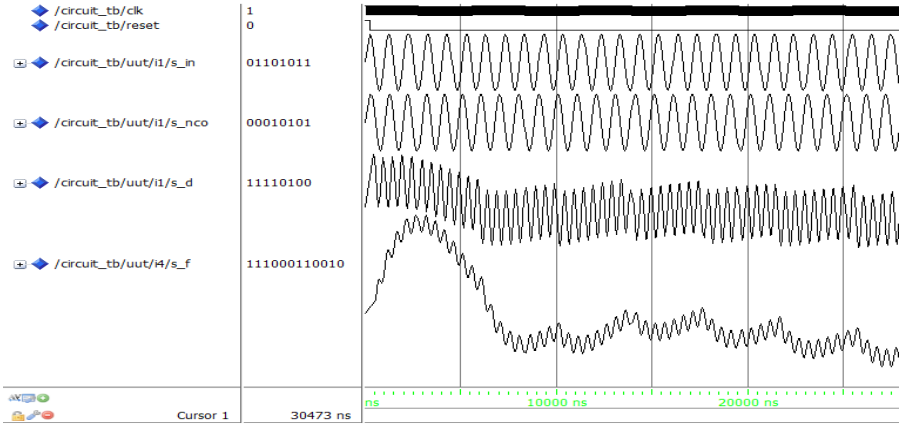


FIGURE 14: Conventional DPLL waveforms for input sinusoidal of 1 MHz

A comparison between conventional DPLL (in which the phase detector is only a multiplier and NCO uses ROM) and proposed DPLL is done by implementing both architectures on the same FPGA device (cyclone II, EP2C35F67C). The result is illustrated in TABLE 1. It is clear that the proposed PLL saves area, reduces power consumption and works at higher frequency than the conventional one.

Architecture	Conventional DPLL	Proposed DPLL
Total logic elements	762	655
Logic registers	106	78
Memory bits	2048	0
Maximum clock	115.02 MHz	155.8 MHz
Core dynamic power consumption(100 MHz)	19.54 mW	15.44 mW

TABLE 1: DPLL implementation comparison

7. CONSLUSION

An improved PLL design was presented. The signal estimated by the VCO was subtracted from the signal input to the PLL, before passing the resultant to the phase detector. This resulted in

eliminating the noise present because of the double frequency ripple, without increasing order of the LPF. Computer simulations performed using Matlab showed significant performance improvements in the case of changing frequency, phase and both frequency and phase. A digital version of the improved PLL was also proposed using VHDL and then implemented through FPGA. High power consumption and low operation speed in conventional DPLL results because of the NCO in which the generation of sinusoidal waveform depends on ROM. As accuracy of the generated signal increases, the size of ROM is increased. Proposed structure replaces the traditional NCO with another one, which depends on piecewise linear approximation to the sine function thus ROM is not needed in this case. The proposed architecture reduces noise, power consumption, area consumption and works at higher frequency than the conventional DPLL.

8. Acknowledgement

This work was supported in part by Grand-in-Aid for Young Scientists of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), no. 20700210 and the Japan Society for the Promotion of Science (JSPS) grant no. 20560381.
.

9. REFERENCES

1. Roland E. Best. “Phase-Locked Loops: Design, Simulation and Application”, 6th ed, McGraw-Hill, (2007)

2. M. Pardoen, J. Gerrits, V. von Kaenel. “A 0.9 V 1.2mA 200 MHz BiCMOS single chip narrow band FM receiver”. In Proceeding of the Solid-State Circuits Conference, CA,USA,1996

3. A. Sempel, H. van Nieuwenbrg. “A fully integrated HIFI PLL FM demodulator”. In Proceeding of the Solid-State Circuits Conference, CA, USA,1990

4. M. Hagiwara, M. Nakagawa .“Digital Signal Processing type stereo FM receiver”. IEEE Trans. Cons. Elect.,32(1):37-43,1986

5. N. Boutin. “An Arctangent type wideband PM/FM demodulator with with improved performance., IEEE Trans. Cons. Elect.,38(1): p 5,1992

6. G. Hsieh and C. Hung. “Phase-locked loop techniques - A survey”. IEEE Trans. Ind. Elect., 43(6): 609-615,1996

7. W. C. Lindsey and C. M. Chie. “A survey of digital phase-locked loops”. Proc. IEEE, 69(4): 410-431,1981

8. J. C. Lee and C. K. Un. “Performance analysis of digital tanlock loop”. IEEE Trans. Com.,30(10):2398-2411,1982

9. D. Abramovitch. “Phase-Locked Loops: A Control Centric Tutorial”. In Proceedings of the American Control Conference, USA, 2002

10. S. Goldman. “Phase Locked-Loop Engineering Hand Book of Integrated Circuit”, Artech House Publishers, (2007)

11. P. Pong Chu. “RTL Hardware Design Using VHDL: Coding for Efficiency, Portability and Scalability”. Wiley-IEEE Press, (2006)

12. U. Meyer-Baese. “Digital Signal Processing with Field Programmable Gate Arrays (Signals and Communication Technology)”, Springer, (2007)

13. T. Tokumaru, Hidechika. “Booth’s multiplier”. U.S. Patent 4807175, Feb. 21,1989.

14. J. Certic, L Milic. “Signal Processor Implementation of a Low-Pass/High-Pass IIR Digital Filter with Variable Cutoff Frequency”. In Proceeding of The International Conference on computer As a Tool, Belgeade, 2005

15. D. Tufts, J. Francis. “Designing digital low pass filters comparison of some methods and criteria”. IEEE Trans Audio & Electro.,18(4):p487, 1970

16. S. Kadam, D. Sasidaran, A. Awawdeh, L. Johnson, M. Soderstand_ “Comparison of various numerically controlled oscillators”. symposium circuits and systems, 2002

17. H. T. Nicholas, H. Samueli. “An Analysis of the Output Spectrum of Direct Digital Frequency Synthesizers in the Presence of Phase Accumulator Truncation”. In Proceeding of 41st Annu. Frequency Control Symp,1987

18. B. Yang. “Direct digital frequency synthesizer using a new ROM compression method”. In proceeding of 27th European Solid-State Circuit Conference, 2001

19. P. W. Ruben, E. F. Heimbecher. “Dilley, Reduced Size Phase-to- Amplitude Converter in a Numerically Controlled Oscillator”. U. S. Patent 4855946, Aug. 8, 1989

20. S. Tze-Yun, K. Lyu-Ting, H His-chin. “Low power and high-SFDR direct digital frequency synthesizer based on hybrid CORDIC algorithm”. IEEE symposium circuit and systems, Taipei, 2009

21. K. Zhang, X. Huang. “A high SFDR direct digital synthesizer with frequency error free output”. IEEE symposium circuit and systems, WA, USA, 2008

Performance Evaluation of CE-OFDM in PLC Channel

El Ghzaoui Mohammed elghzaoui.mohammed@gmail.com
Laboratory of Transmission and Data processing, University Sidi Mohammed Ben Abdellah, Fez,

Belkadid Jamal belkadid@gmail.com
Laboratory of Transmission and Data processing, University Sidi Mohammed Ben Abdellah, Fez,

Benbassou Ali alibenbassou@gmail.com
Laboratory of Transmission and Data processing, University sidi Mohammed ben Abdellah, Fez,

Abstract

One major drawback associated with an OFDM system is that the transmitter's output signal may have a high peak-to-average ratio (PAPR). High levels of PAR may be a limiting factor for power line communication (PLC) where regulatory bodies have fixed the maximum amount of transmit power. To overcome this problem, many approaches have been presented in the literature. One potential solution for reducing the peak-to-average power ratio (PAPR) in an OFDM system is to utilize a constant envelope OFDM (CE-OFDM) system. This paper describes a CE-OFDM based modem for Power Line Communications (PLC) over the low voltage distribution network. The impact of the electrical appliances on the signal transmission is investigated. The good performances of the BER have been checked by the simulation platform of real PLC channel using Matlab. Finally, CE-OFDM-CPM is compared with conventional OFDM under HomePlug AV..

Keywords: OFDM, CE-OFDM, CPM, BER, PAPR

1. INTRODUCTION

In home networks advanced communication technologies has allowed the Power Line Communication (PLC) channel to be a transmission medium that enables the transferring of high-speed digital data over the classical indoor electrical wires. Power Line Communications [1, 2, 3] have been the subject of an important research work. At the same time, the growing demand for multimedia communications provides a good prospect for PLC as a promising transmission technique for the "last mile" access network. However, the main characteristics of the power line channel seem quite unfavorable: multipath reflections induced by impedance mismatches, changes in the channel transfer function [4] due to the switching of electrical devices, a highly diverse noise environment, and so forth. Therefore, an efficient transmission scheme has to include robust techniques in order to face the difficulties of such a medium and to get reliable and

spectrally efficient transmissions. In this study, we investigate the influences of some switching power devices on PLC adapters.

Orthogonal frequency division multiplexing (OFDM) [5, 6, 7] modulation using orthogonal subcarriers reduces the delay spread by increasing robustness to multipath fading and can use overlapped bandwidth due to orthogonality on frequency domain. Thus, OFDM has been adopted for high speed data transmission of multimedia traffic such as HomePlug A/V and Mobile WiMax. However, OFDM also has a drawback of a high PAPR (peak-to-average-power-ratio). This high PAPR takes place due to parallel processing of a number of data at once using a fast Fourier transform (FFT) processor. Implementation of multicarrier system needs more precision Digital-to-Analog Converter (DAC) in Transmitter and more precision Analog-to-Digital Converter (ADC) in Receiver. The DAC clips all samples that exceed certain maximum amplitude, the clip level. Setting this level is a compromise between clipping probability and quantization noise level: decreasing the chip level will increase the average clip noise but decrease the quantization noise. It is usually set to that the total Signal-to-Noise Ratio (SNR) is minimized. A lower PAR will increase the SNR or allow for a DAC with lower resolution to be used. In the literature, many PAR reduction schemes have been studied, such as block coding [8], clipping [9], trellis shaping [10] and selected mapping (SLM) [11].

To alleviate this problem, constant envelope OFDM (CE-OFDM) signal has been introduced in [12, 13, 14], which combines orthogonal frequency division multiplexing and phase modulation or frequency modulation. Furthermore, by utilizing continuous phase modulation (CPM) in a CE-OFDM system, the PAPR can be effectively reduced to 0 dB. Although the CPM has low spectral efficiency, it features low system complexity and favorable performance due to low PAR and robustness to amplitude variation and impulsive noise [15]. The CPM decreases the side lobe of the power spectrum by means of continuously connecting the phase that contains the information.

This paper is organized as follows. Next section describes the PAPR of OFDM system. In section 3, we give a brief insight on CPM modulation. Then the CE-OFDM-CPM signal is introduced in section 4. Experimental and Simulation results are presented in Section 5 and Section 6 concludes the paper.

2. PAPR OF OFDM SYSTEM

The OFDM baseband waveform can be represented by

$$m(t) = \sum_{k=0}^{N-1} I_k e^{j2\pi \frac{k}{T_B} t} \qquad 0 \le t \le T_B \qquad (1)$$

N is the number of subcarriers, T_B is the signaling interval, and data symbol I_k modulates the k[th] subcarrier $e^{j\pi \frac{k}{T_B} t}$. The data symbols are chosen from a complex set defined by an M-point signal constellation such as PSK or QAM. Consider sampling $m(t)$ without cyclic prefix, at the sampling rate $f_{sa} = \frac{JN}{T_B}$ samp/s, where $J \ge 1$ is the oversampling factor. The signal samples are:

$$m[n] = \sum_{k=0}^{N-1} I_k e^{j2\pi \frac{k}{N_B J} n} \qquad n = \{0,1,.....,N_B - 1\} \qquad (2)$$

It can be seen that the sequence $\{m[n]\}$ can be interpreted as the inverse discrete Fourier transform (IDFT) of the OFDM data block *I* with $(J-1)N$ zero padding. It is well known that PAPR of the continuous-time OFDM signal cannot be obtained precisely by the use of Nyquist rate sampling, which corresponds to the case of $J = 1$. It is shown in [16] that $J = 4$ can provide sufficiently accurate PAPR results. The PAPR computed from the J-times oversampled time-domain signal samples is given by

$$PAPR_s = \frac{\max_{0 \le n \le JN} |m(n)|^2}{E_s\left\{|m[n]|^2\right\}} \qquad (3)$$

Where E{.} denotes expectation.

The instantaneous signal power, $|s(t)|^2 = \Re^2\{s(t)\} + \Im^2\{s(t)\}$, with the peak and average signal power, are plotted in Figure 1. For this example the peak-to-average power ration is more than 7.9 dB.

There have been many schemes proposed in the research literature aimed at reducing the impact of the PAPR problem. All PAPR reduction techniques have some advantages and disadvantages. These PAPR reduction techniques should be chosen carefully for getting the desirable minimum PAPR. In the following section we introduce an improved PAPR reduction scheme using constant envelope modulation.

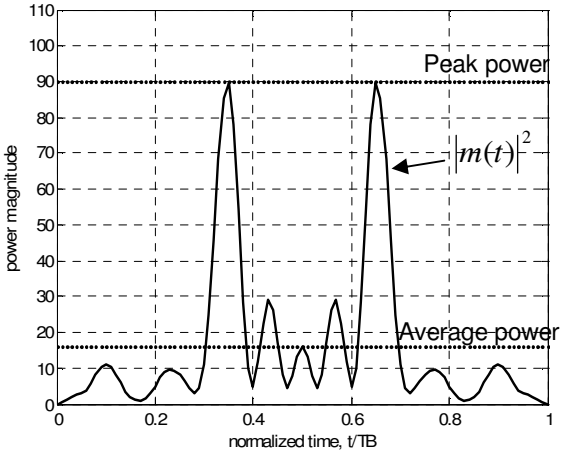


FIGURE 1: Instantaneous signal power of OFDM signal, also showing average and peak powers. N=16, $I_k \in \{\pm 1\}$

3. CONTINUOUS PHASE MODULATION

Continuous phase modulation (CPM) is a good example of nonlinear modulation with memory. Signal dependence is introduced purposely to shape the spectrum of the transmitted signal to achieve better spectral efficiency. As a result, CPM modulation is a nonlinear modulation scheme with memory [17]. Furthermore, CPM is typically implemented as a constant-envelope waveform. The CPM baseband signal is given by:

$$x(t) = Ae^{j\phi(t,I)} \qquad (4)$$

where A is the signal energy and the phase is given by

$$\phi(t,I) = \phi_0 + 2\pi \sum_{k=-\infty}^{n} h_k I_k q_k(t - kT), \quad nT \le t < (n+1)T \qquad (5)$$

where ϕ_0 is the phase memory, I_k is the sequence of M-ary information symbols, h_k is a sequence of modulation indices and the waveform $q(t)$ can be represented as,

$$q(t) = \int_0^t g(\tau)d\tau \qquad (6)$$

Normally, the function $g(t)$ is a smooth pulse shape over a finite time interval $0 \le t \le LT$ and zero outside. Thus, L is the pulse length and T is the symbol period. From the definition of the above class of constant amplitude modulation schemes the pulse $g(t)$ is defined in instantaneous frequency and its integral $q(t)$ is the phase response [18].

4. CE-OFDM-CPM SIGNAL DESCRIPTION

4.1. CE-OFDM Signal Définition

Consider the baseband OFDM waveform:

$$m(t)=\sum_{i}\sum_{k=1}^{N} I_{i,k}\,q_k(t-iT_B) \quad (7)$$

Where $\{I_{i,k}\}$ are the data symbols and $\{q_k(t)\}$ are the orthogonal subcarriers. For conventional OFDM the baseband signal is up-converted to band-pass as:

$$\begin{aligned} y(t) &= \Re\left\{m(t)e^{j2\pi f_c t}\right\} \\ &= A_m(t)\cos\left[2\pi f_c t+\phi_m(t)\right] \end{aligned} \quad (8)$$

f_c : Carrier frequency

Where $A_m(t)=|m(t)|$ and $\phi_m(t)=\arg[m(t)]$. For real-valued $m(t)$, $\phi_m(t)=0$ and $y(t)$ is simply an amplitude modulated signal. For CE-OFDM, $m(t)$ is passed through a phase modulator prior to up-conversion. The baseband signal is $s(t)=Ae^{j\alpha m(t)}$ where α is a constant. The band-pass signal For real-valued *m*(t) is ,

$$y(t)=\cos\left[2\pi f_c t+\alpha\, m(t)\right] \quad (9)$$

Therefore $y(t)$ is a phase modulated signal.

The OFDM signal is phase modulated onto a carrier signal to obtain a constant envelope signal with 0dB PAPR. CE-OFDM requires a real-valued OFDM [19] message signal, that is, $\phi_m(t)=0$. Therefore the data symbols in (7) are real-valued, this one dimensional constellation is known as pulse-amplitude modulation (PAM). Thus the data symbols are selected from an M-PAM set. The subcarriers $q_k(t)$ must also be real-valued, and may be expressed as:

$$q(t)=\begin{cases}\cos(2\pi kt/T_B) & 0\le t<T_B \quad k<N,\\ \sin(2\pi(k-N/2)t/T_B) & 0\le t<T_B, \quad k>N/2\\ 0 & \text{otherwise.}\end{cases} \quad (10)$$

For k = 1, 2, . . . , N,
The subcarrier orthogonality condition holds:

$$\int_{iT_B}^{(i+1)T_B} q_{k1}(t-iT_B)\,q_{k2}(t-iT_B)\,dt=\begin{cases}E_q & k_1=k_2\\ 0 & k_1\neq k_2\end{cases} \quad (11)$$

Where $E_q=T_B/2$
The baseband CE-OFDM signal is,

$$s(t)=Ae^{j\phi(t)} \quad (12)$$

Where A is the signal amplitude.

The phase signal during the i^{th} block is written as:

$$\phi(t)=\theta_i+2\pi hc_N\sum_{k=1}^{N} I_{i,k}\,q_k(t-iT_B), \quad iT_B\le t<(i+1)T_B \quad (13)$$

Where

$$\theta(t)=\phi(iT-\varepsilon)-\phi(iT+\varepsilon), \quad \varepsilon\to 0$$

The phase memory θ_i may be used in conjunction with a phase unwrapper at the receiver to ensure a continuous phase at the symbol boundaries and hence better spectral containment [20,

12]. Here *h* refers to modulation index; N is the number of sub-carriers $I_{n,k}$ represents M-PAM data symbols; T_B is the l^{th} block interval, and $q_k(t)$ represents the set of subcarrier waveforms. The normalizing constant, c_N, is set to $c_N = \sqrt{\frac{2}{N\sigma_I^2}}$ where σ_I is the variance of the data symbols, and consequently the variance of the phase signal will be $\sigma_\phi^2 = (2\pi h)^2$. Assuming that the data is independent and identically distributed, it follows that $\sigma_I^2 = \frac{M^2-1}{3}$.

To guarantee continuous phase, the memory terms set to

$$\theta_i = K\sum_{l=0}^{\infty}\sum_{k=1}^{N}[I_{i-l,k}\, A_b(k) - I_{i-1-l,k}\, A_e(k) \qquad (14)$$

Where

$K = 2\pi h c_N$, $A_b(k) = q_k(0)$, $A_e(k) = q_k(T_B - \varepsilon)$, $\varepsilon \to 0$.

The benefit of continuous phase CE-OFDM-CPM is a more compact signal Spectrum [20].

4.2. Spectral Efficiency

The effective double-sided bandwidth, defined as the twice the highest frequency subcarrier, of *m*(t) is:

$$W = 2\frac{N}{2T_B} = \frac{N}{T_B} \qquad (15)$$

The bandwidth of $s(t)$ is at least *W*, and depending on the modulation index the effective bandwidth can be greater than *W*. The RMS (root-mean-square) bandwidth is obtained by borrowing a result from analog angle modulation [21], [22],

$$B_{rms} = \sigma_\phi W = 2\pi h W \qquad (16)$$

As defined in (16), the RMS bandwidth can be less than *W*. A more suitable bandwidth is thus,

$$B_s = \max(2\pi h, 1)W \qquad (17)$$

The power density spectrum can be easily estimated by the Welch method [23] of periodogram averaging. Figure 2 shows Power Spectrum of CE-OFDM for different value of modulation index.

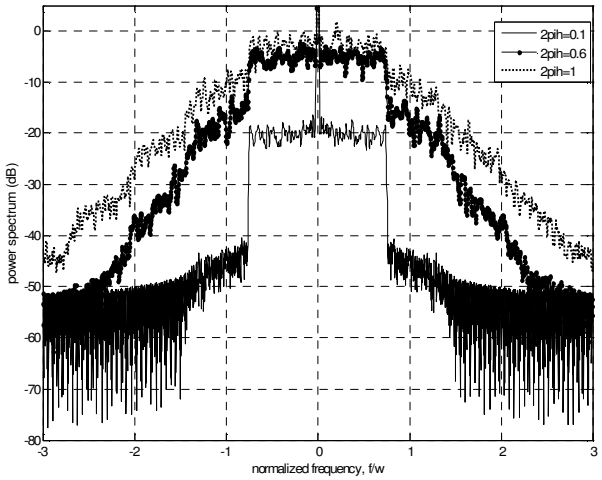


FIGURE 2: Power Spectrum of CE-OFDM-CPM. N=512

Since the modulation index controls the CE-OFDM spectral containment, smaller h can be used if a tighter spectrum is required. The tradeoff is that smaller h results in worse performance. Therefore, the system designer can trade performance for spectral containment, and visa versa, consequently the modulation index also controls the system performance. Figure 2 shows that the power spectrum of the CE-OFDM signal is nearly constant over the range $\left|\frac{f}{W}\right| < 0.8$. To calculate the spectral efficiency versus performance, the data rate must be defined, which for uncoded CE-OFDM is:

$$R = \frac{N \log_2(M)}{T_B} \qquad (18)$$

Using (17) as the effective signal bandwidth, the spectral efficiency is,

$$\eta = \frac{R}{B_s} = \frac{\log_2(M)}{\max(2\pi h, 1)} \quad b/s.Hz \qquad (19)$$

5. MEASUREMENTS AND SIMULATIONS RESULTS

We have made frequency response measurements for in-building power line channels in the frequency range of (1–100 MHz). As can be seen, the frequency response exhibits considerable frequency dependent variation, due to the specific wiring configurations encountered. The frequency dependent channel fading is the result of reflections and multipath propagation. One main source for reflection and multipath propagation is impedance discontinuity. There are many possible reasons for impedance discontinuity, such as change of gauge of wires connected to each other, connected loads or branch wires, etc.

5.1. Individual Appliances Connection

Electrical appliances connected or not connected to the network at anytime generate changes on network characteristics. We consider the topology in Figure 3. The electrical appliances at point B were varied as personal computer charger, cell phone charger, heater, and without electrical appliances. Figure 4, we plot the realization of a channel frequency response magnitude measured by network analyzer in downlink from a household for different electrical appliances. Measurements demonstrate that the impedance varies significantly from one electrical appliance to another.

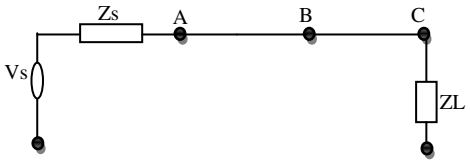
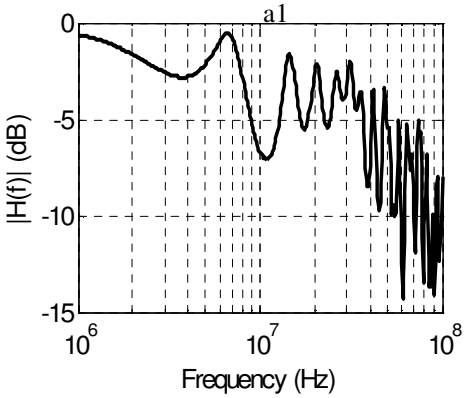
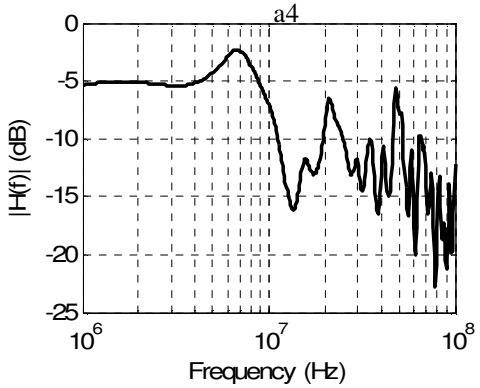


FIGURE 3: Power line network between sending and receiving ends.





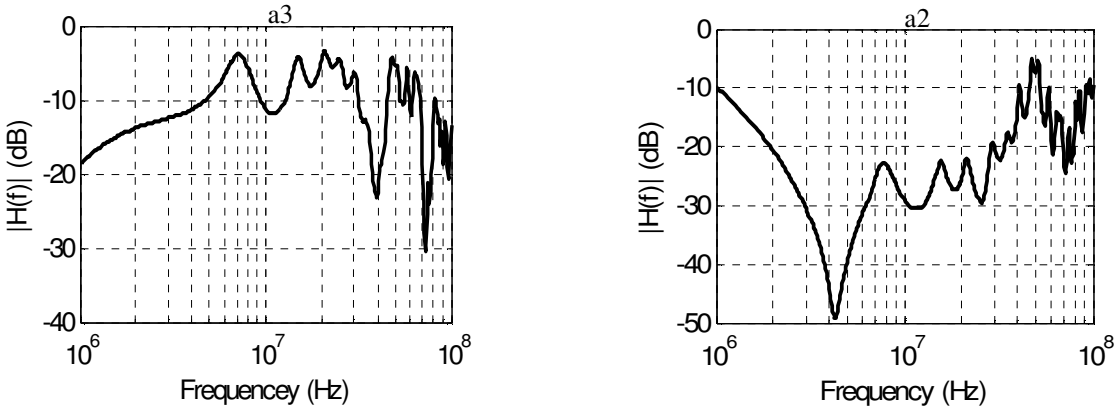


FIGURE 4: Experimental measurements of frequency response magnitude for different Electrical appliances (a1) without Electrical appliances (a2) heater, (a3) personal computer charger, (a4) cell phone charger.

5.2. Combinations of Appliance Connection

In Figure 5 the measurements results regarding the simultaneous connection of two outlet cable with a length of 1.5m, heater, personal computer charger, cell phone charger and arbitrary generator are recorded.

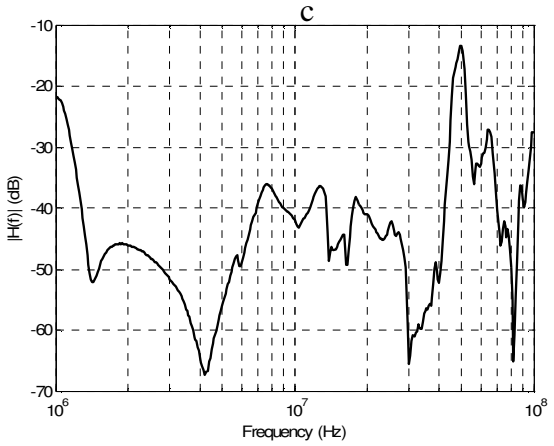


FIGURE 5: Experimental measurements of frequency response magnitude of arbitrary combinations of appliance connection.

5.3. NUMERICAL RESULTS

In this section, the bit error rate (BER) performance of the proposed CE-OFDM-CPM PLC system is investigated. The signal-to-noise ratio (SNR) is defined as E_b/N_0, where E_b is the received energy per information bit and N_0 is the power spectral density of the noise. The BER performance of CE-OFDM-CPM over PLC channel is evaluated using computer simulation. First let us see the influence of different electrical appliances on transition on PLC channel.

Figure 6 shows the performance of the CE-OFDM-CPM system for different electrical appliances cases. A good channel performance is seen for without electrical appliances with the bit error probability of 3.10^{-7} at a E_b/N_0 per bit of 20 dB. The power is 26 dB, 28 dB and 39 dB, for heater, personal computer charger, and cell phone charger, respectively. When the cell phone is connected the power loss is 39 dB indicating degraded performance. The results in Figure 6 show the dramatic performance degradation as a consequence of the severe frequency selectivity. The analysis of the results leads to a classification of the electrical appliances, according to their influence on the transmitted PLC signal in the narrow-band frequency range. Appliances such as cell phone charger cause a severe attenuation on the transmitted signal. For

example, Lin et al. showed that cell phone chargers caused serious degradation of the bandwidth of the PLC adapters [24]. The severe attenuation is related to the presence of the compensation capacitors of the devices. The signal levels are also affected by the transmitter and the distance of the connection point from the transmitter.
Figure 7 shows the performance of the CE-OFDM-CPM system for channel (c) (Combinations of appliances connection). A good channel performance is seen for without electrical appliances with the bit error probability of 10^{-7} at a E_b/N_0 per bit of 20 dB, but if the simultaneous connection of electrical appliances the power loss is 59 dB indicating degraded performance (the performance is severely degraded due to the several frequency selectivity).

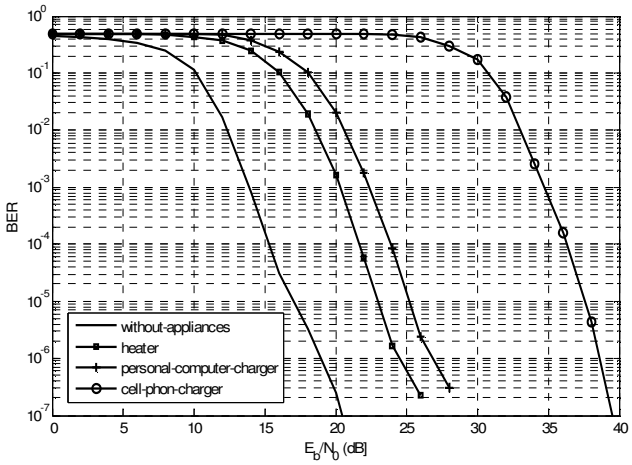


FIGURE 6: Performance of CE-OFDM in PLC channel for different Electrical appliances, (M=4, N=512, 2πh=1, J=4, MMSE).

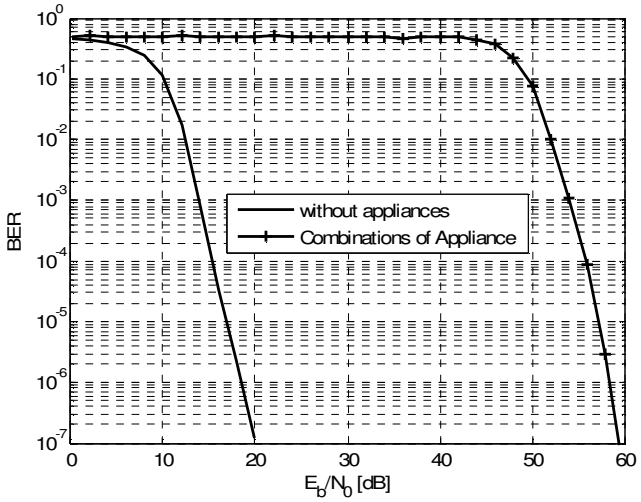


FIGURE 7: Performance of CE-OFDM in PLC channel for simultaneous connection of electrical appliances. (M=4, N=512, 2πh=1, J=4, MMSE).

In figure 8, the BER performances of CE-OFDM-CPM have been simulated and compared with OFDM system under *HomePlug AV*. In this case 2πh=0.8 and M=16. CE-OFDM-CPM is shown to outperform OFDM at high bit energy-to-noise. CE-OFDM-CPM is shown to outperform OFDM at high E_b/N_0. Alternatively, if operation at low E_b/N_0 is important, then CE-OFDM-CPM may not be well suited due to the threshold effect.

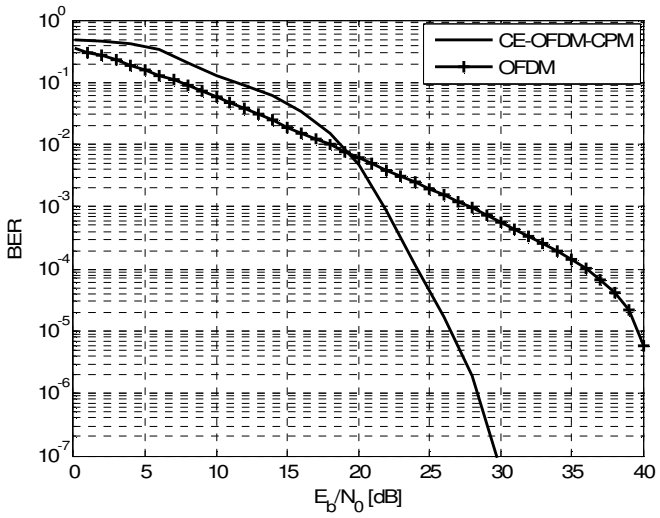


FIGURE 8: BER performances comparison of CE-OFDM, and OFDM systems under *HomePlug AV*

6. CONCLUSION

In this paper the performance of power line channel has been investigated. The propagation of high frequency signals in an indoor power network is strongly influenced by the connection of various electrical appliances, by the complexity and by the varying star topology of the power distribution installation. Therefore, a PLC application must be carefully designed and must be combined with proper experimentation, regarding signal transmission characteristics. Together with the PLC transmission environment, a possible modulation technique is introduced in order to create a complete picture of the PLC transmission. Design parameters for a CE-OFDM-CPM communication system for high rate data transmission over power lines are specified. OFDM-CPM is compared with conventional OFDM under HomePlug AV. CE-OFDM-CPM is shown to outperform OFDM at high E_b/N_0. However, at low E_b/N_0 the CE-OFDM-CPM phase demodulator receiver suffers from a threshold effect.

7. REFERENCES

1. Pavlidou N, Han Vinck A J, Yazdani J, Honary B. Power line communications: State of the art and future trends. IEEE Communications Magazine, 2003, 41(4): 34-40.

2. Gotz M, Rapp M, Dostert K. Power line channel characteristics and their effect on communication system design. IEEE Communications Magazine, 2004, 42(4): 78-86.

3. A. B. Gutierrez, A. Darmand, V. Watt, L. Ngalamou, “Design of an Analog Electronic Interface for a Power Line Based Telephony System”, IEEE-ISPLC 2006, Florida, USA, March 2006, pp. 232-238.

4. C. Papaleonidopoulos, C. G. Karagiannopoulos, D. P. Agoris, P. D. Bourkas, N. J. Theodorou. HF Signal Transmission over Power Lines and Transfer Function Measurement. Proc. of the Sixth IASTED International Conference. July 3-6, 2001, Rhodes, Greece. (502-505).

5. N. Sood, A. K Sharma, M. Uddin, “On Channel Estimation of OFDM-BPSK and –QPSK over Nakagami-m Fading Channels,” Signal Processing, An International Journal (SPIJ), Volume (4): Issue (4). 2010,pp. 239 – 246.

6. R. Prasad, OFDM for Wireless Communications Systems, Boston, MA, Artech House, 2004.

7. Ma Y H, So P L, Gunawan E. Performance analysis of OFDM systems for broadband power line communications under impulsive noise and multipath effects. IEEE Trans-actions on Power Delivery, 2005, 20(2): 674-682.

8. D. Wulich and L. Goldfeld, "Reduction of peak factor in orthogonal multicarrier modulation by amplitude limiting and coding," IEEE Trans. Commununication, vol. 47, no. 1, pp. 18–21, January 1999.

9. L. Wang and C. Tellambura, "A simplified clipping and filtering techniques for PAR reduction in OFDM systems," IEEE Signal Process. Lett., vol. 12, no. 6, pp. 453–456, June 2005.

10. W. Henkel and B. Wagner, "Another application for trellis shaping: PAR reduction for DMT (OFDM)," IEEE Trans. Commununication, vol. 48, no. 9, pp. 1471–1476, September 2000.

11. J. Sun Lee; H. Oh; J. Kim; J. Y. Kim;" Performance of scaled SLM for PAPR reduction of OFDM signal in PLC channels" Power Line Communications and Its Applications, IEEE International Symposium 2009,.

12. Y. Tsai and G. Zhang, "Orthogonal Frequency Division Multiplexingwith Phase Modulation and Constant Envelope Design," in Proc. of IEEE Milcom 2005, Atlantic City, NJ, Oct. 2005.

13. S.C. Thompson, A.U. Ahmed, J.G. Proakis, and J.R. Zeidler, "Constant Envelope OFDM Phase Modulation: Spectral Containment, Signal Space Properties and Performance," in Proc. of IEEE Milcom 2004, Monterey, CA, Nov. 2004.

14. S.C. Thompson, J.G. Proakis, and J.R. Zeidler, "Binary OFDM Phase Modulation," in Proc. Of IEEE Milcom 2003, Boston, MA, Oct. 2003.

15. K. Dostert, Powerline communications, Prentice-Hall, 2001.

16. C. Tellambura, "Computation of the continuous-time PAR of an OFDM signal with BPSK subcarriers" IEEE Commun. Lett., vol. 5, pp.185–187,May2001.

17. Tor Aulin John B. ANderson and Carl-Erik Sundberg, Digital Phase Modu-

18. lation, Kluwer Academic Publishers, 1986;

19. T. Aulin, N. Rydbeck, and C. Sundberg, "Continuous phase modulation- Part I and II," IEEE Trans. on Comm., vol. 29, no. 3, pp. 196–225, 1981.

20. S. C. Thompson, "Generating Real-Valued OFDM Signals with the Discrete Fourier Transform." [Online]. Available: http://zeidler.ucsd.edu/~sct/pubs/t5.pdf.

21. S. C. Thompson, A. U. Ahmed, J. G. Proakis, and J. R. Zeidler, "Constant Envelope OFDM—Part II: Spectral Properties," IEEE Trans. Commun., 2006.

22. J. G. Proakis and M. Salehi, communication systems Energineering. New Jersey: Prentice Hall, 1994.

23. H. Roberts, Angle Modulation. London: Peter Peregrinus Ltd., 1977.

24. "Spectral Estimation of Digital Signaling Using The Welch Method". [Online]. Available: http://zeidler.ucsd.edu/~sct/holdings/welch/

25. C. K. Lin, S.C. Yeh, and H. H. Chen, “Bandwidth estimation of in-home power line networks,” in Proceedings of the 2007 IEEE International Symposium on Power line Communications and Its Applications, Pisa, Italy, Mar. 2007,pp.413–418.

Classification of Electroencephalograph (EEG) Signals Using Quantum Neural Network

Ibtisam A. Aljazaery ibtisamalasay@googlemail.com
College Of Engineering/ Elect. Eng. Dep.
Basrah University
Hilla, Iraq

Abduladhem A. Ali abduladem1@yahoo.com
College Of Engineering/ Comp. Eng. Dep.
Basrah University
Basrah, Iraq

Hayder M. Abdulridha drenghaider@uobabylon.edu.iq
College Of Engineering/ Elect. Eng. Dep.
Babylon University
Hilla, Iraq

Abstract

In this paper, quantum neural network (QNN), which is a class of feedforward neural networks (FFNN's), is used to recognize (EEG) signals. For this purpose ,independent component analysis (ICA), wavelet transform (WT) and Fourier transform (FT) are used as a feature extraction after normalization of these signals. The architecture of (QNN's) have inherently built in fuzzy. The hidden units of these networks develop quantized representations of the sample information provided by the training data set in various graded levels of certainty. Experimental results presented here show that (QNN's) are capable of recognizing structures in data, a property that conventional (FFNN's) with sigmoidal hidden units lack . Finally, (QNN) gave us kind of fast and realistic results compared with the (FFNN). Simulation results show that a total classification of 81.33% for (ICA), 76.67% for (WT) and 67.33% for (FT).

Keywords: Quantum Neural Network, EEG, ICA, Wavelet

1. INTRODUCTION

Brain is the center of central processing of Physical and mental activities, which is mostly affected by the Physical performance.

Neurons, or nerve cells, are electrically active cells which are primarily responsible for carrying out the brain's functions. Neurons create action potentials, which are discrete electrical signals that travel down axons and cause the release of chemical neurotransmitters at the synapse, which is an area of near contact between two neurons [1,2].

An electroencephalograph (EEG) is the measurement of electrical activity generated by the brain. First measured in humans by Hans Berger in 1929 [3].

In general, EEG is obtained using electrodes placed on the scalp with a conductive gel. In 1998, Rodrigo Q. Q. described and extended two new approaches that started to be applied to (EEG) signals (a) the time-frequency methods, and (b) the methods based on Chaos theory [4].

Quantum neural network (QNN's) is a promising area in the field of quantum computation and quantum information. In 1996, Lov K. Grover, proposed a method can speed up a range of search applications over unsorted data using Quantum mechanics [5]. Several models have been proposed in the literature but for most of them need a clear hardware requirements to implement such models. One of the most exciting emerging technologies is quantum computation, which attempts to overcome limitations of classical computers by employing phenomena unique to quantum-level events, such as nonlocal entanglement and superposition. It is therefore not surprising that many researchers have conjectured that quantum effects in the brain are crucial for explaining psychological phenomena, including consciousness [6]. Jarernsri. L. Mitrpanont, Ph. D. Ananta Srisuphab, presented the approach of the quantum complex-valued backpropagation neural network or QCBPN. The challenge of their research is the expected results from the development of the quantum neural network using complex-valued backpropagation learning algorithm to solve classification problems [7].

Independent component analysis (ICA) is essentially a method for extracting useful information from data. It separates a set of *signal mixtures* into a corresponding set of statistically independent component signals or *source signals*. ICA belongs to a class of *blind source separation* (BSS) methods for separating data into underlying informational components [8]. The mixtures can be sounds, electrical signals, e.g., electroencephalographic (EEG) signals, or images (e.g., faces, fMRI data). The defining feature of the extracted signals is that each extracted signal is statistically independent of all the other extracted signals[9].

The basis signal, or wavelet, used to decompose a signal does not produce information about “frequency” in the traditional sense, but rather a distribution of time and scale is created. A change in scale represents stretching or compressing the wavelet by a factor of two. It is therefore possible to reconstruct any signal using one wavelet as the basis and placing as many wavelets as are needed at different times with different amplitudes and scales[10].

Fourier Transform (FT) which transforms a signal (function) that exists in the time (or space) domain to the frequency domain. The FT accomplishes this task through a kernel composed by sine and cosine waveforms. This is the origin of the main disadvantage of FT for signal analysis [11].

The FFT (Fast Fourier Transform) can be computed relatively quickly, at or around real-time. The FFT does have its disadvantages, however. The frequencies used to decompose a signal are a function of the sampling frequency of the signal and the number of frequency bins desired. Without modifying these two parameters, these frequencies are not selectable. A simple sine wave whose frequency does not fall on one of the frequencies of the transform will produce a spectrum with energy spread to many frequencies[10].

2. METHODOLOGY

In (ICA), each signal is described as a scalar variable, and a set of signals as a vector of variables, and the process of obtaining signal mixtures from signal sources using a set of mixing coefficients [9]. ICA showing us how a set of source signals can be represented as a scattergram in which each point corresponds to the values of the signals at one time, and that a set of mixing coefficients can be used to implement a geometric transformation of each point.

$$x1 = as1 + bs2 \quad (1)$$

$$x2 = cs1 + ds2 \quad (2)$$

where (a, b, c, d), a set of mixing coefficients.

The resultant set of “mixture” points can be transformed back to the original set of “source signal” points using a set of unmixing coefficients, which reverse the effects of the original geometric transformation from source signals to signal mixtures.

s1 = αx1 + βx2 .....(3)

s2 = γ x1 + δx2.....(4)

where (α, β, γ, δ), a set of unmixing coefficients and (s1, s2) are the original signals.

The desecrate wavelet (DWT) of a signal *x* is calculated by passing it through a series of filters. First the samples are passed through a low pass filter with impulse response *g* resulting in a convolution of the two:

$$y[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k]g[n-k] \quad \text{.....(5)}$$

The signal is also decomposed simultaneously using a high-pass filter *h*. The outputs giving the detail coefficients (from the high-pass filter) and approximation coefficients (from the low-pass). It is important that the two filters are related to each other and they are known as a quadrature mirror filter. However, since half the frequencies of the signal have now been removed, half the samples can be discarded according to Nyquist's rule. The filter outputs are then subsampled by 2 (g- high pass and h- low pass):

$$y_{low}[n] = \sum_{k=-\infty}^{+\infty} x[k]g[2n-k] \quad \text{.....(6)}$$

$$y_{high}[n] = \sum_{k=-\infty}^{+\infty} x[k]h[2n+1-k] \quad \text{.....(7)}$$

This decomposition has halved the time resolution since only half of each filter output characterises the signal. However, each output has half the frequency band of the input so the frequency resolution has been doubled.

With the downsampling operator $\downarrow$

$$(y \downarrow k)[n] = y[kn] \quad \text{.....(8)}$$

the above summation can be written more concisely

$$y_{low} = (x * g) \downarrow 2 \quad \text{.....(9)}$$

$$y_{high} = (x * h) \downarrow 2 \quad \text{.....(10)}$$

However computing a complete convolution *x* * *g* with subsequent downsampling would waste computation time. The Lifting scheme is an optimization where these two computations are interleaved [12,13].

The fast Fourier transform (FFT) is a discrete Fourier transform algorithm which reduces the number of computations needed for Npoints from $2N^2$to $2N\lg N$, where lg is the base-2 logarithm. If the function to be transformed is not harmonically related to the sampling frequency, the response of an FFT looks like a sinc function. Discrete Fourier transform can be computed using an FFT by means of the Danielson-Lanczos lemma if the number of points $N$is a power of two. If the number of points $N$is not a power of two, a transform can be performed on sets of points corresponding to the prime factors of Nwhich is slightly degraded in speed. Prime factorization is slow when the factors are large, but discrete Fourier transforms can be made fast for $N = 2$, 3, 4, 5, 7, 8, 11, 13, and 16 using the Winograd transform algorithm [14].

Fast Fourier transform algorithms generally fall into two classes: decimation in time, and decimation in frequency. The Cooley-Tukey FFT algorithm first rearranges the input elements in bit-reversed order, then builds the output transform (decimation in time). The basic idea is to break up a transform of length N into two transforms of length N/2 using the identity [15].

$$X(k) = \sum_{n=0}^{N-1} x(n)\omega_N^{nk}, \qquad k = 0,..., N-1 \quad \text{.....(11)}$$

$$x(n) = \frac{1}{N}\sum_{n=0}^{N-1} X(k)\omega_N^{-nk}, \qquad k = 0,..., N-1 \quad \text{.....(12)}$$

where,

$$\omega_N = \exp(\frac{-2\pi \quad j}{N}) \quad \text{.....(13)}$$

x(n) is the signal in time domain.

X(k) is the signal in the frequency domain.

To selecting the best features from the signal which is dealt with (ICA), (WT) and (FFT), classification method was used for this purpose.

The individual within-class scatter matrix and the total within-class scatter matrix is defined by

$$S_i = \sum_{x \in \chi_i} (x - \mu_i)(x - \mu_i)^T \qquad i = 1,2 \quad \text{.....(14)}$$

$$S_w = S_1 + S_2 \quad \text{.....(15)}$$

where, **μi** is the mean vector of the class **i** .

We can obtain the transform vector *w* with maximal between class distance and minimal within class variance by Fisher criterion function and Lagrange multiplier method: [16]

$$w = S_w^{-1}(\mu_1 - \mu_2) \quad \text{.....(16)}$$

The QNN consists of n_i inputs, one hidden layer of n_h nodes, each hidden node represents a multilevel function (Eq. 17), and n_o output units. The output units are sigmoidal [17].

The equation of the output of hidden layer can be written as:

$$\tilde{h}_{j,k} = \frac{1}{n_s}\sum_{r=1}^{n_s} h_{j,k}^{r} = \frac{1}{n_s}\sum_{r=1}^{n_s} \mathrm{sgm}\left(\beta_h\left(\bar{h}_{j,k} - \theta_j^{r}\right)\right) \quad \text{.....(17)}$$

$$\mathrm{sgm}(\tau) = 1/(1+\exp(-\tau))$$

Where: βh is a slope factor, θ_j^{r}'s define the jump positions in the transfer function, and n_s is the number of levels or sigmoids in the hidden unit.

3. RESULTS AND DISCUSSION

This section presents experimental classification results on the (EEG) data set which is used in (QNN). The results were obtained by using the (ICA), (WT) and (FFT) are from two different electrodes of the scalp hat, as in tables 1,2 and 3:

	(QNN) electrode no.1					(QNN) electrode no.5				
	TP	FN	FP	Se	PP	TP	FN	FP	Se	PP
CLASS 1 (baseline)	25	1	4	.96	.86	25	0	5	1	.83
CLASS 2 (rotation)	25	2	3	.926	.89	24	0	6	1	.8
CLASS 3 (multiplication)	24	2	4	.923	.857	26	0	4	1	.867
CLASS 4 (counting)	26	2	2	.929	.929	19	1	10	.95	.655
CLASS 5 (letter comp.)	22	4	4	.846	.846	22	0	8	1	.733
TCA	81.33%					77.33%				

TABLE 1: Classification Results With QNN By Using (ICA).

	(QNN) electrode no.1					(QNN) electrode no.5				
	TP	FN	FP	Se	PP	TP	FN	FP	Se	PP
CLASS 1 (baseline)	25	2	3	.926	.89	25	0	5	1	.833
CLASS 2 (rotation)	22	2	6	.916	.79	23	3	4	.8846	.85
CLASS 3 (multiplication)	23	3	4	.88	.85	22	7	1	.758	.956
CLASS 4 (counting)	24	2	4	.923	.86	22	5	3	.814	.88
CLASS 5 (letter comp.)	21	3	6	.875	.78	19	8	3	.703	.86
TCA	76.67%					74%				

TABLE 2: Classification Results With QNN By Using WT(db1).

	(QNN) electrode no.1					(QNN) electrode no.5				
	TP	FN	FP	Se	PP	TP	FN	FP	Se	PP
CLASS 1 (baseline)	21	6	3	.78	.875	20	2	8	.9	.714
CLASS 2 (rotation)	20	4	6	.83	.77	20	6	4	.769	.83
CLASS 3 (multiplication)	19	6	5	.76	.79	18	3	9	.857	.67
CLASS 4 (counting)	20	3	7	.87	.74	17	5	8	.77	.68
CLASS 5 (letter comp.)	21	4	5	.84	.807	22	0	8	1	.73
TCA	67.33%					64.67%				

TABLE 3: Classification Results With QNN By Using (FFT).

Where,

(TP_i) is (true positive) classification for class i.

(FN_i) is (false negative) classification for class i.

(FP_i) is (false positive) classification for class i.

$$Sensitivit\ y(Se) = \frac{TP_i}{TP_i + FN_i} \quad \text{.....(18)}$$

$$Positive\ \ \Pr edictivity(PP) = \frac{TP_i}{TP_i + FP_i} \quad \text{.....(19)}$$

From tables(1,2 and 3), we can notice that, different results were obtained from different electrodes and that means (EEG) signals are different from electrode to another according to mental tasks.

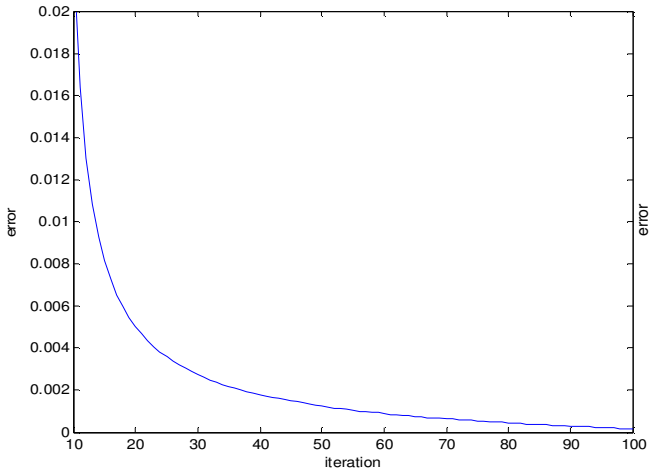


FIGURE 1: Error By Using ICA (1rst.electrode).

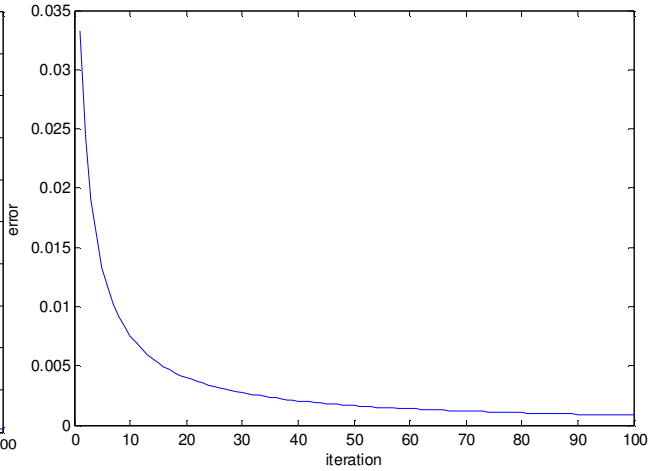


FIGURE 2: Error By Using ICA (5th.electrode).

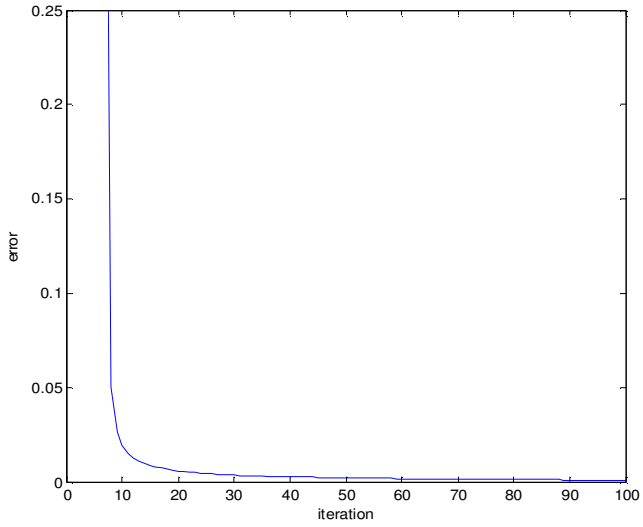


FIGURE 3: Error By Using WT(db1) (1rst.electrode).

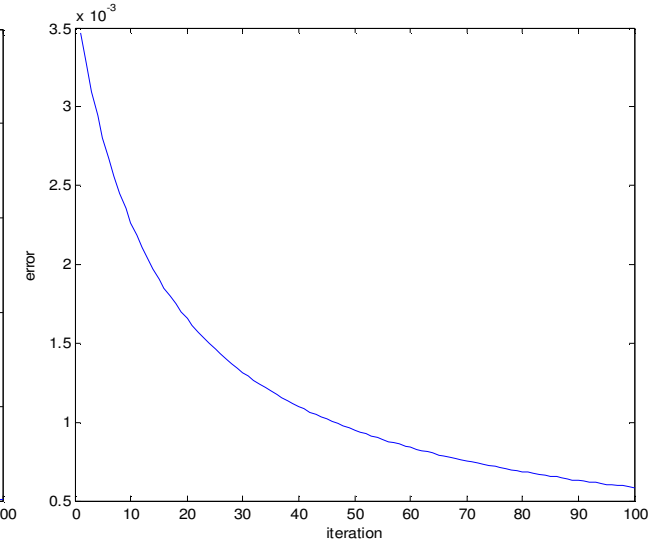


FIGURE 4: Error By Using WT(db1) (5th.electrode).

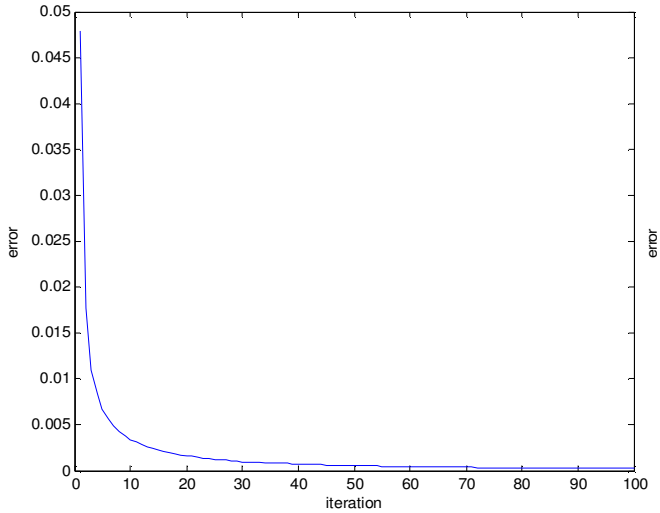


FIGURE 5: Error By Using FFT (1rst.electrode).

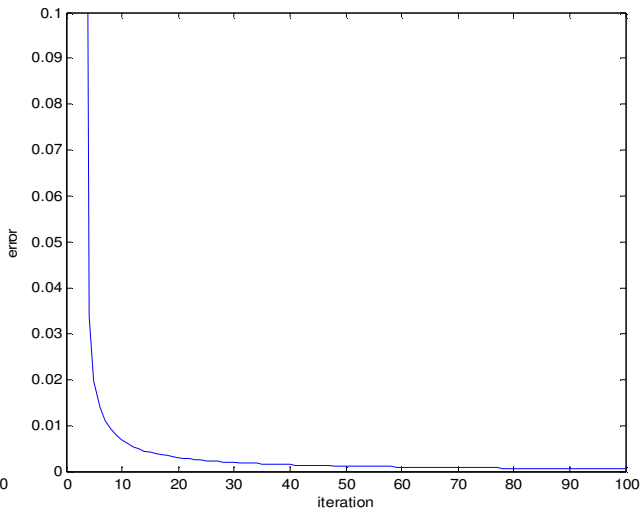


FIGURE 6: Error By Using FFT (5th.electrode).

The error diagrams for the first and fifth electrode, (Fig.1 to Fig.6), tell us that, the value of error is decrease whenever the training is progress until convenient weights and biases are obtained.

4. CONCLUSION

As resultant for this research, different results were obtained from different electrodes. Because of different mental tasks for different positions of brain.

Generally, the results were obtained from first electrode best than the fifth electrode's results. And the classification of (EEG) signals by using (ICA) best than by using (WT and FFT), and the classification by using (WT) best than by using (FFT). Besides, classification of class(4) from the first electrode is best than fifth electrode. It means the mental tasks which read from the first electrode as (EEG) signals ,may be they were specialist with routine actions (like counting) more than the others.

5. REFERENCES

[1] Shimogawara Masahiro, Tanaka Hiroaki, Kazumi Kunio, Haruta Yasuhiro, *"MEGvision Magnetoencephalograph System and Its Applications"*, Yokogawa Technical Report English Edition, No. 38, pp. 23-27, 2004.

[2] Gaetano Di Chiara, M.D. "*The Principles of Nerve Cell Communication*", Vol. 21, No. 2, pp. 107-108, 1997.

[3] William O. Tatum, IV. Aatif M. Husain, Selim R. Benbadis, Peter W. Kaplan, "*Handbook of EEG interpretation*", Made in the United State of America, pp. 1-7, 2008.

[4] Rodrigo Quian Quiroga, "*Quantitative analysis of EEG signals: Time-frequency methods and Chaos theory*", Institute of Physiology-Medical University Lubeck and Institute of Signal Processing-Medical University Lubeck, 1998.

[5] Lov K. Grover, "*Quantum Mechanics Helps in Searching for a Needle in a Haystack*", Physical View Letters, Vol. 79, No. 2, 14 July 1997.

[6] Abninder Litt, Chris Eliasmith, Frederick W. Kroon, Steven Weinstein, Paul Thagard, "*Is the Brain a Quantum Computer?*", Cognitive Science Society, University of Waterloo, No.30, pp. 593–603, 2006.

[7] Jarernsri. L. Mitrpanont, Ph. D., Ananta Srisuphab**,** "*The Realization of Quantum Complex-Valued Backpropagation Neural Network in Pattern Recognition Problem*", The 9th. International Conference on Neural Information Processing(ICONIP'OZ)Vol. 1, 2003.

[8] Isabelle Guyon, Steve Gunn, Masoud Nikravesh, Lotfi A. Zadeh, "*Feature Extraction Foundations and Applications*", USA and United Kingdom, 2006.

[9] James V. Stone, "*Independent component analysis A Tutorial Introduction*", A Bradford Book The MIT Press Cambridge, Massachusetts London, England, 2004.

[10] Matthew Van Dyke Kotvis, "*An Adaptive Time-Frequency Distribution with Applications for Audio Signal Separation*", Coral Gables, Florida, April 25, 1997.

[11] Daniela Do Amaral, Gilson A. Giraldi, "*Continuous and Discrete Wavelet Transform and Multiscale Analysis*", LNCC–National Laboratory for Scientific Computing, Av. Getulio Vargas, Petr´opolis, RJ, Brazil, 2009.

[12] Paulo M. Silveira, Paulo F. Ribeiro, "*Using Wavelet for Visualization and Understanding of Time-Varying Waveform Distortions in Power Systems*", 2007.

[13] Stéphane Mallat, "*A Wavelet Tour of Signal Processing*", 2007.

[14] Bracewell, R. "*The Fourier Transform and Its Applications"*, 3rd ed. New York, McGraw-Hill, 1999.

[15] Nussbaumer, H. J. "*Fast Fourier Transform and Convolution Algorithms"*, 2nd ed. New York, Springer-Verlag, 1982.

[16] Aihua Zhang, Bin Yang, Ling Huang, "*Feature Extraction of EEG Signals Using Power Spectral Entropy*", International Conference on Biomedical Engineering and Informatics, Lanzhou University of Technology, 2008.

[17] Gopathy Purushothaman, Nicolaos B. Karayiannis, "*Quantum Neural Networks (QNN's): Inherently Fuzzy Feedforward Neural Networks*", IEEE Transaction on Neural Networks Vol. 8, No. 3, pp. 679-693, May 1997.

New Data Association Technique for Target Tracking in Dense Clutter Environment Using Filtered Gate Structure

E. M.Saad elsayedmos@hotmail.com
Department of Electronics and Communication
Engineering Helwan University
Cairo, Egypt

El. Bardawiny bardawiny@hotmail.com
Department of Radar
MTC College
Cairo, Egypt

H.I.ALI hosam_ib_ali@yahoo.com
Department of Electronics and Communication
Engineering Helwan University
Cairo, Egypt

N. M. Shawky negmshawky@gmail.com
Department of Electronics and Communication
Engineering Helwan University
Cairo, Egypt

Abstract

Improving data association process by increasing the probability of detecting valid data points (measurements obtained from radar/sonar system) in the presence of noise for target tracking are discussed in this paper. We develop a novel algorithm by filtering gate for target tracking in dense clutter environment. This algorithm is less sensitive to false alarm (clutter) in gate size than conventional approaches as probabilistic data association filter (PDAF) which has data association algorithm that begin to fail due to the increase in the false alarm rate or low probability of target detection. This new selection filtered gate method combines a conventional threshold based algorithm with geometric metric measure based on one type of the filtering methods that depends on the idea of adaptive clutter suppression methods. An adaptive search based on the distance threshold measure is then used to detect valid filtered data point for target tracking. Simulation results demonstrate the effectiveness and better performance when compared to conventional algorithm.

Keywords: Target Tracking, Data Association, Probabilistic Data Association Algorithm, Kalman Filter.

1. INTRODUCTION

Real-world sensors often report more than one measurement that may be from a given target. These may be either measurements of the desired target or "clutter" measurements. Clutter refers to detections or returns from nearby objects, clouds, electromagnetic interference, acoustic

anomalies, false alarms, etc. Data association algorithms allow the use of the Kalman filter (KF) structure for estimation in the presence of clutter. But data association can be the source of both track loss and computational complexity issues. Two most popular KF-based algorithms for single-target tracking in clutter using data association as probabilistic data association filter (PDAF) and nearest neighbor kalman filter (NNKF) used in comparative evaluation in simulation. Failing algorithm to track a target (track loss) may be from increasing the range of clutter density. Filtering methods for tracking targets in noise are well established [1], [2]. The way to separate signals from clutter in target tracking is to use a distance measure from the predicted target positions. Since it is computationally expensive to evaluate this for all measurements, we require a gating process in order to reduce the number of candidate measurements to be considered. The gating technique in tracking a maneuvering target in clutter is essential to make the subsequent algorithm efficient but it suffers from problems since the gate size itself determines the number of valid included measurements. If we choose a too small gate size, we can miss target-originated measurements on the other hand, if we choose a gate with too large size, we will obtain many unwanted non-target measurements, giving rise to increased computational complexity and decreased performance. To find a gate volume in which we regard measurements as valid is an important consideration. There have been many types of gating techniques studied. First of all, previous approaches have used constant parameters to determine the gate size [3]-[6]. Recently, adaptive and (locally) optimal approaches to estimate gate size have also been proposed under more restricted assumptions [7]-[14]. However, this estimation is often computationally intensive. Data association is responsible for deciding which of the existing multiple measurements in gate of the predicted position should update with a tracking target. Some data techniques use a unique to update a track; i.e. at most one observation is used to update a track. PDAF is An alternative approach to use all of the validated measurements with different weights (probabilities). Due to increase in the false alarm rate or low probability of target detection (target in dense clutter environment), most of the data association algorithms begin to fail. We propose here an algorithm which is less sensitivity to false alarm targets in the gate region size than PDA and NNKF algorithms. This proposed algorithm reduces the number of candidate measurements in the gate by a filtering method that compares the measurement in the gate at the prediction step with the current measurement in the same gate at the update step and then avoids any measurement in the current gate less than the threshold value due to comparison. This is called filtering gate method which is similar to an idea taken from adaptive clutter suppression filtering methods used in radar signal processing [15,16]. The filtering gate algorithm is combined with PDA algorithm to apply the proposed algorithm in tracking targets in presence of various clutter densities. Simulation results showed better performance when compared to the conventional PDA algorithm.

2. BACKGROUND

State Space Model

In a dynamic state space model, the observed signals (observation/measurements) are associated with a state and measurement noise. Let the unobserved signal (hidden states) $\{x_t : t \in N\}, x_t \in X$ be modeled as a Markov process of initial distribution $p(x_0)$ and transition probability $p(x_t \mid x_{t-1})$. The observations $\{z_t : t \in N\}, z_t \in Z$ are assumed to be conditionally independent given the process $\{x_t : t \in N\}$ and of the marginal distribution $p(z_t \mid x_t)$.We have the following state and measurement space models at time *t*:

$$x_t = A \quad x_{t-1} + w_t$$
$$z_t = H \quad x_t + v_t \qquad (1)$$

where w_t and v_t are white Gaussian noise with zero mean and covariance *Q* and *R* respectively. *A* and *H* are matrices with appropriate sizes. The initial distribution is defined by $P(x_0) = N(x_0 \mid m_0, p_0)$ where the initial configuration defined by parameters m_0 and

p_0 is assumed to be known. In linear systems, the state space model is optimally addressed by the Kalman filter [14],[17]. The functioning of the Kalman filter consists of two recursive steps: prediction and update.

Filtered Gate Method

In the prediction step, Let $Z_{t-1} = \{z_{1,t-1}, z_{2,t-1}, \cdots z_{i,t-1}, \cdots z_{w_n,t-1}\}$ be a set of points in the 2-D Euclidean space at time *t*-1 where w_n is the number of points at time scan Δt and let $\hat{z}_t$ be a predicted position of the tracked target at time *t*. according to distance metric measure and gate size, let $\breve{z}_{t-1} = \{z_{1,t-1}, \cdots, z_{j,t-1}, \cdots z_{m,t-1}\}$ be a set of the candidate points detected in the gate of predicted position $\hat{z}_t$ whose elements are a subset from the set Z_{t-1} where *j* =1 to *m* (number of detected points in gate at time *t*-1) and $\breve{z}_{t-1}$ be a set of all valid points $z_{j,t-1}$ that satisfy the distance measure condition $\left| z_{i,t-1} - \hat{z}_t \right|$ < W where W is threshold value that determines the gate size and *i* =1 to w_n , *j* =1 to m, i. e $j = j+1$ after each valid point is detected up to m points. We consider each point $z_{j,t-1}$ in the gate is a center of very small square gate g_j its length is small δ where each value in the small gate g_j is approximately equal to $z_{j,t-1}$ i.e. $z_{j,t-1} \approx z_{j,t-1} - \delta/2$ to $z_{j,t-1} + \delta/2$.

In the updating step, let $Z_t = \{z_{1,t}, z_{2,t}, \cdots z_{w_k,t}\}$ be a set of points in the 2-D Euclidean space at time *t* where w_k is the number of points at time scan Δt. The candidate points detected in the same gate of predicted position $\hat{z}_t$ be a subset $\breve{z}_t = \{z_{1,t}, \cdots, z_{k,t}, \cdots z_{n,t}\}$ from the set Z_t where *k* =1 to *n* (number of detected points in gate at time *t*) and $\breve{z}_t$ be a set of all valid points $z_{k,t}$ that satisfy the distance measure condition $\left| z_{i,t} - \hat{z}_t \right|$ < W where *i* =1 to w_k , *k*=1 to n for *k*=*k*+1 after each valid point is detected. After receiving the measurement Z_t and detecting the valid measurements $\breve{z}_t$ in the gate , each point from $\breve{z}_t$ in the specified gate at time *t* is compared with the previous points $\breve{z}_{t-1}$ in the gate at time *t*-1 to detect the invalid points when $\left| z_{k,t} - z_{j,t-1} \right| < \delta$ and then exclude the point $z_{j,t-1}$ from the set $\breve{z}_{t-1}$ in the next iteration of comparison as shown in Fig. 1.

Finally, we obtain the reduced number of valid points in the gate while the other invalid points is not including in the data association process.

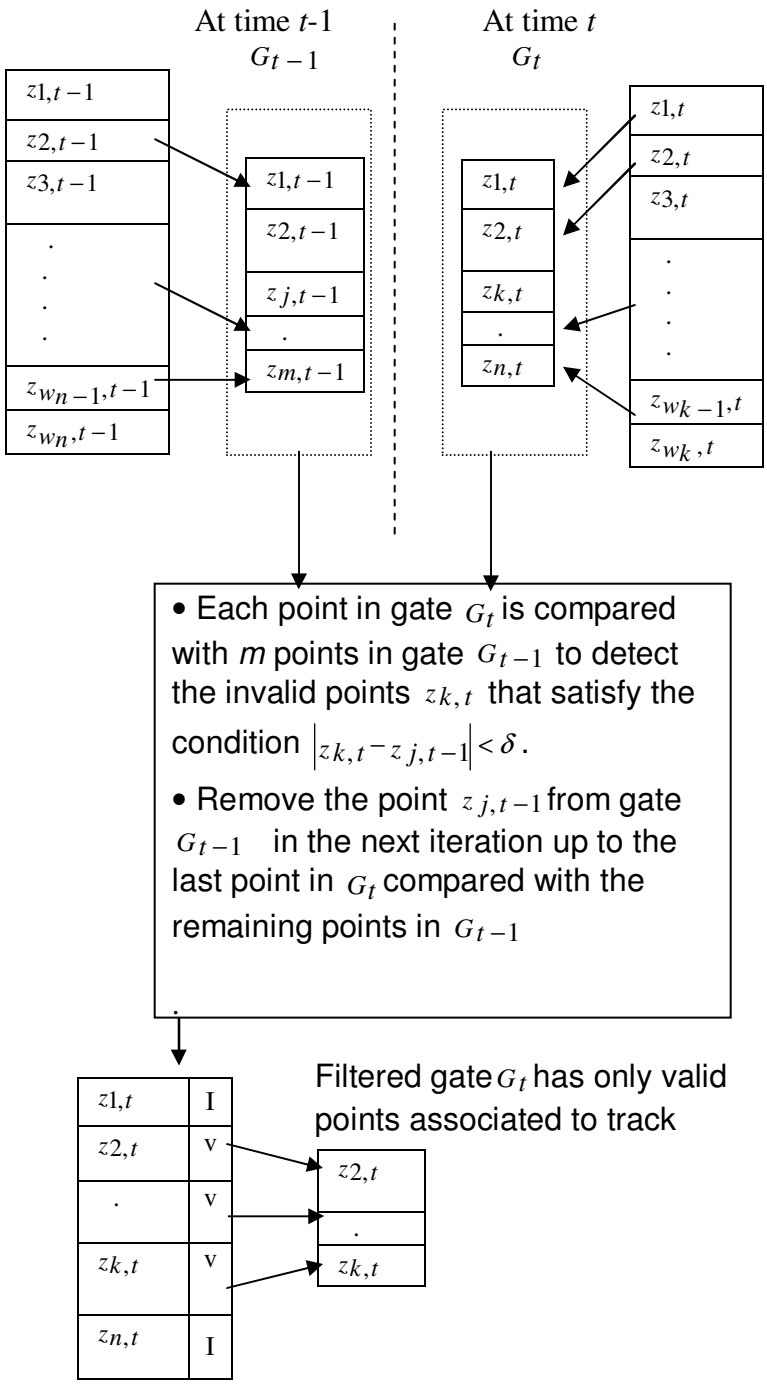


FIGURE 1: Filtered Gate scheme.

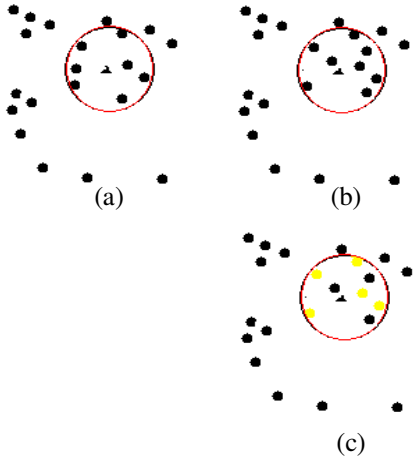


FIGURE 2: Gated measurements given an identical threshold for conventional PDA and filtered gate based PDA (FG-PDA): (a) measurements at previous scan. (b) Measurements at current scan. (c) Filtered gate based approach at the scan of (b).

3. INTEGRATION BETWEEN DATA ASSOCIATION AND FILTERED GATE

We propose an algorithm which depends on the history of observation for one scan and use a fixed threshold but operates similar to an adaptive estimator. In conventional data association approaches with a fixed threshold, all observations lying inside the reconstructed gate are considered in association. The gate may has a large number of observations due to heavy clutter, this leading to; increasing in association process since the probability of error to associate target-originated measurements my be increased. In our proposed algorithm a filtered gate structure is used to provide the possibility to decrease the number of observations in the gate by dividing the state of observations into valid and invalid that only the valid are considered in association. The proposed algorithm can be applied to all gate based approaches, including tracking and clustering. See Fig. 2. Fig. 2(a) and 2(b) show the candidates for association in both conventional probabilistic data association (PDA) and our proposed filtered gate based probabilistic data association (FG-PDA). Red circles represent the gated measurements. Our approach has measurements as well as ones inside the validated region but is divided into two states valid and invalid, yellow points represent invalid points as shown in Fig. 2(c).

4. IMPLEMENTATION OF PROBABILISTIC DATA ASSOCIATION FILTER USING FILTERED GATE METHOD

The Probabilistic Data Association (PDA) algorithm is used to calculate the probability that each validated measurements is attributable to the target in cluttered environments [1],[2],[4],[14],[18-19]. This algorithm is useful in tracking a single object (target) in clutter and referred to as the PDA filter (PDAF).

Notation for PDAF Approach

The PDAF calculates the associated probability of each element of the set of validated measurements at time t, denoted as $Z_t = \{z_t^i\}_{i=1:c_t}$ where z_t^i is the ith validated measurement and c_t is the number of measurements in the validation region at time t. under the Gaussian assumption for the prediction kernel $p(x_t \mid z_{1:t-1})$, the validation region is commonly taken to be the elliptical region

$$V_t = \left\{ Z : (z_t^i - H\ \bar{m}_t)^T S_t^{-1} (z_t^i - H\ \bar{m}_t) \le \gamma \right\} \quad (2)$$

Where γ is a given threshold and the covariance is defined by $S_t = H\ \bar{p}_t H^T + R$. We define the accumulation of validated measurements is $Z_{1:t} = \left\{ Z_j, \quad for \quad j \in \{1,....,t\} \right\}$.

Prediction step in PDAF Approach

We define the posterior distribution of x_t given the past sequence of observations $Z_{1:t-1}$ in the prediction step, i.e., $p\ (x_t \mid Z_{1:t-1})$. This process is equivalent to the prediction step of standard Kalman filter. The prediction distribution is defined by

$p\ (x_t \mid Z_{1:t-1}) = N\ (x_t; \bar{m}_t, \bar{p}_t)$,where

$\bar{m}_t = A\ m_{t-1}$ and $\bar{p}_t = A\ p_{t-1} A^T + Q$

Update Step in PDAF Approach

As mentioned in section 2.1, the hidden variables of the state space model are recursively estimated by the prediction and updating steps. The PDAF can be modeled as a state space model which can also be estimated using these recursive operations. First of all, the update step in PDAF approach is as

$$p\ (x_t \mid Z_{1:t}) = \sum_{i=0}^{c_t} N\ (x_t; m_t^i, p_t^i)\, \beta_{t,i} \quad (3)$$

where the $P\ (i \mid Z_{1:t}) = \beta_{t,i}$ is association probability and

$$p_t^i = \left[\bar{p}_t^{-1} + H^T R^{-1} H \right]^{-1}$$
$$m_t^i = p_t^i \left[H^T R^{-1} z_t^i + \bar{p}_t^{-1} \bar{m}_t \right] \quad \text{and} \quad (4)$$

for $i \in \{ 0,....,c_t \}$. In addition, we have $m_t^0 = \bar{m}_t$ and $p_t^0 = \bar{p}_t$ for i = 0 where there is no target-originated measurement (i.e., z_t^0 = nil).

Estimating Conditional Probability in PDAF Approach

In order to obtain the filtering density, we require an estimate for the parameter $\beta_{t,i}$ for $i \in \{ 0,....,c_t \}$. Under the assumption of a poisson clutter model, the association probability $\beta_{t,i}$ can be estimated as [20-21]

$$\beta_{t,i} = P\ (i \mid Z_{1:t}) = \begin{cases} \dfrac{e_i}{b + \sum_{j=1}^{c_t} e_j}, & i = 1,.....c_t \\ \dfrac{b}{b + \sum_{j=1}^{c_t} e_j}, & i = 0 \end{cases} \quad (5)$$

where

$$
\begin{aligned}
e_i &= \exp\{D\} \\
b &= \lambda \left|2\pi S_t\right|^{1/2} \frac{1-p_D p_G}{P_D} \\
D &= -\frac{1}{2}(z_t^i - H\ \overline{m}_t)^{iT}\ S_t^{-1}(z_t^i - H\ \overline{m}_t)^i \\
&= -\frac{1}{2}(v_t)^{iT}\ S_t^{-1}(v_t)^i
\end{aligned}
\tag{6}
$$

and λ is a spatial density parameter. The functions p_D and p_G denote the probability of detection and Gaussian validation. The proposed filtering gate based PDAF is represented in algorithm 1. The algorithm is divided into four major parts: prediction, finding validated regions, estimating conditional probability and finally an update step. Since only finding validated regions component is fundamentally different from the conventional PDAF, we look at this in more detail.

Algorithm 1 PDAF using filtered gate

---

1. **for** t = 1 to T **do**
2. **Do prediction step**,

$$x_{t|t-1} \sim P\ (x_t \mid Z_{1:t-1}) = N\ (x_t \mid \overline{m}_t, \overline{p}_t)$$

where $\left\{ \begin{matrix} \overline{m}_t = A\ \ m_{t-1} \\ \overline{p}_t = A\ \ p_{t-1} A^T + Q \end{matrix} \right\}$

3. **Finding validated region according to Algorithm 2**.
4. **Estimating conditional probability**, $\beta_{t,i}$ for

for $i \in \left\{\ 0,....,c_t\ \right\}$,

$$\beta_{t,i} = \left\{ \begin{matrix} \dfrac{e_i}{b + \sum_{j=1}^{c_t} e_j}, & i = 1,.....c_t \\ \dfrac{b}{b + \sum_{j=1}^{c_t} e_j}, & i = 0 \end{matrix} \right\}$$

where

$$e_i = \exp\left\{-\frac{1}{2}(v_t)^{iT}\ S_t^{-1}(v_t)^i\right\}$$

$$b = \lambda \left|2\pi S_t\right|^{1/2} \frac{1-p_D p_G}{P_D}$$

5. **Do update step**,
6. Calculate the distribution of the missing observation $P\ (x_t \mid Z_{1:t-1})$ which is for *i* = 0,

$$m_t^0 = \overline{m}_t, \qquad p_t^0 = \overline{p}_t$$

7. Calculate the distribution of the associated observation, $P\ (x_t \mid z_t^i, Z_{1:t-1}) = N\ (x_t; m_t^i, p_t^i)$ for $i = \{\ 1,....,c_t\ \}$

$$p_t^i = \left[\ \bar{p}_t^{-1} + H^T R^{-1} H\ \right]^{-1}$$

$$m_t^i = p_t^i \left[H^T R^{-1} z_t^i + \bar{p}_t^{-1} \bar{m}_t \right]$$

8. Calculate marginalized probability using Gaussian approximation, $P\ (x_t \mid Z_{1:t}) = N\ (x_t \mid m_t, p_t)$ *where* .

$$m_t = \sum_{i=0}^{c_t} \beta_{t,i} m_t^i$$

$$p_t = \sum_{i=0}^{c_t} \beta_{t,i} \left[p_t^i + (m_t^i - m_t)(m_t^i - m_t)^T \right] \quad (7)$$

9. end for

For finding the validated region, the filtered gate FG_PDAF after the prediction step checks the number of measurements z_{t-1}^i at time *t*-1 that lying inside the gate that determined by the same way of PDAF, then in the update step at time *t* also checks the number of measurements z_t^i that lying in the same gate. If any measurement in the current gate has approximately the same weight (position) to any measurement detected in the previous frame for the same gate within tolerance value with very small threshold δ as mentioned before, we consider this measurement be invalid in the gate and not taken in consideration to data association process.

Filtering the Validation Region to valid/invalid Observations

Intuitively, we find measurements in the gate with fixed size which are associated to the predicted position of the existing target before receiving new measurements. To update the predicted position, the new measurements in the gate is compared with the detected previous measurements in the same gate and avoid these new measurements which have approximately the same weight from data association process as described in algorithm 2.

Algorithm 2 Finding Validated Region of Filtered Gate based PDAF

1. **Find validated region for measurements at time *t*-1**:

$$Z_{t-1} = \{z_{t-1}^i\}, \quad i = 1,.....,m$$

By accepting only those measurements that lie inside the gate:

$$z_{t-1} = \left\{ z_{t-1} : (\ z_{t-1}^i - H\ \ \bar{m}_t)^T S_t^{-1} (\ z_{t-1}^i - H\ \ \bar{m}_t)\ \le \gamma \right\}$$

2. **Find validated region for measurements at time *t***:

$$Z_t = \{z_t^i\}, \quad i = 1,.....,n$$

By accepting only those measurements that lie inside the gate

$$z_t = \left\{ z_t : (z_t^i - H\ \bar{m}_t)^T S_t^{-1} (z_t^i - H\ \bar{m}_t) \le \gamma \right\}$$

where $S_t = H\ \bar{p}_t H^T + R$

3. for $i = 1$ to n do

4. If $\left| z_t^i - z_{t-1}^j \right| < \delta \quad j = 1,.....m$

   Set z_t^i *to* I,

   Remove z_{t-1}^j from the set Z_{t-1}, and set $m = m - 1$

5. Else

   Set z_t^i *to* V

6. End if
7. End for
8. **Obtain valid** (V) **measurements** c_t are included for data association process where the invalid (I) measurements c_f are excluded, i.e.:

   Z_t *be a set of all measurements* $\{z_t^i\} = V$,

   $i = 1$ *to* c_t *where* $c_f = n - c_t$

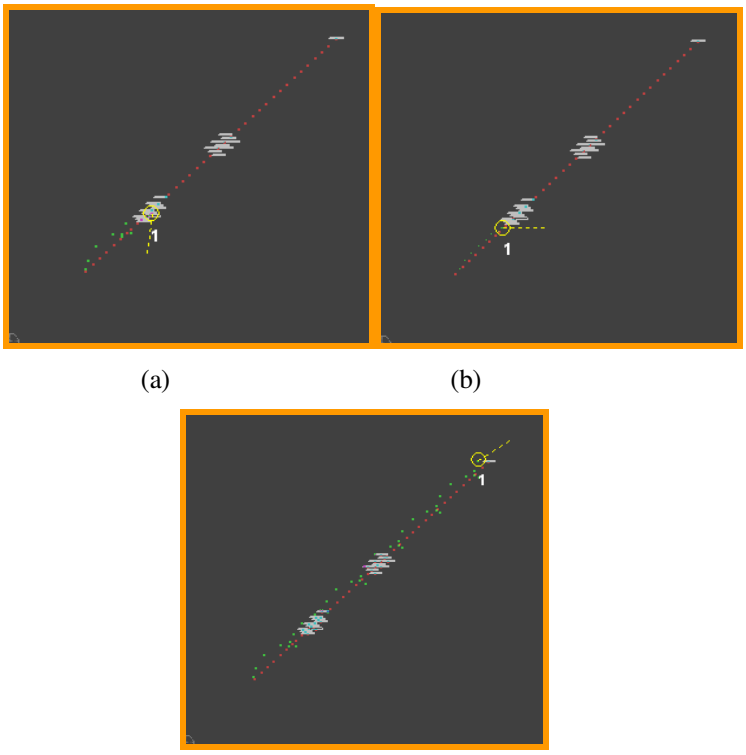
5. SIMULATION RESULTS

We used a synthetic dataset to highlight the performance of the proposed algorithm. The performance of the FG-PDAF is compared with a conventional PDAF and nearest neighbor kalman filter (NNKF) [22] . The synthetic data has a single track which continues from the first frame to the last frame. The mean and covariance for the initial distribution $p(x_0)$ is set to m_0 = [12, 15, 0, 0] and p_0 = diag ([400, 400, 100, 100]). The row and column sizes of the volume (V= $s_W \times s_H$). We initiate the other parameters as: $\tau = 148$,V=26x26 , λ = 0.001 , Δt = 4 , p_D =0.99, p_G =0.8, in addition, we also set the matrices of (1) as

$$A = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, Q = G\ G^T, \quad R = \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}, \quad G = \begin{bmatrix} \frac{\Delta t^2}{2} & 0 \\ 0 & \frac{\Delta t^2}{2} \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix}$$

Given a fixed threshold ($\gamma = 10^{-5}$), we showed that the proposed FG-PDAF succeeded to track a target in dense clutter environment while the others conventional PDAF, NNKF failed to track a target as shown in Fig. 3. We obtain trajectories for X and Y components as shown in Fig. 4(a),(b). In this figure, the blue line represents the underlying truth target of the trajectory. Our proposed algorithm (green line) detects and associates the proper sequence of observation very well compared to PDAF (red line) and NNKF (yellow line).It is clear that the values obtained from using the proposal FG-PDA algorithm, is approximately attached to the values of the true target up to the processing of the last frame number while the values obtained from the conventional two algorithms(PDA and NNKF) is started to far from the true target values after 11 number of frames due to failing in tracking process with existing more numbers of false targets . We also compared error value and root mean square (RMSE) for different approaches as shown in Fig. 5,6. It is also noted that the absolute tracking error is very successful especially after 11 number of frame using the proposed algorithm that has far lower error, RMSE values than either PDAF or NNKF over frame numbers.

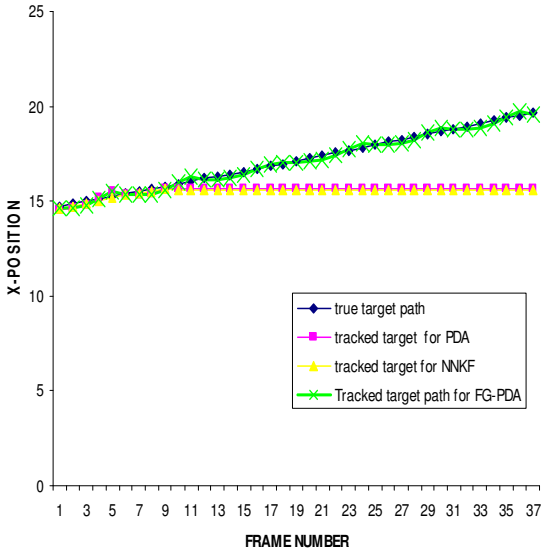
From results, Simulation have been achieved that the proposed algorithm improves the conventional PDA algorithm to be able to, continue tracking without losing the true target in heavy clutter environment, decreases the number of valid measurements region by avoiding the measurements that represent the false targets and thus, the performance of the data association process is increased.



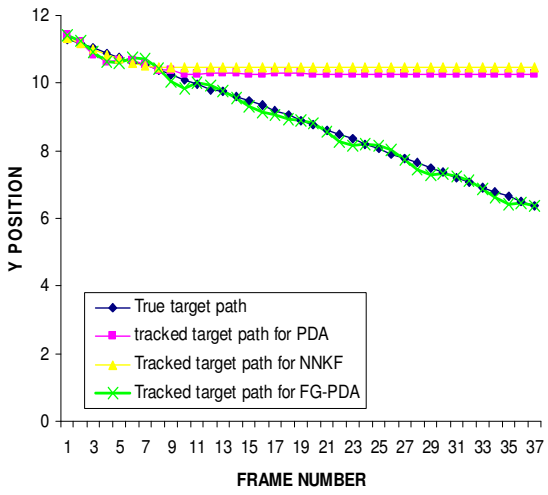
(a) (b)

(c)

FIGURE 3: The state of tracking a single target moving in heavy clutter using 3 approaches algorithm (a) PDAF failed to track (b) NNKF failed to track (c) FG-PDAF succeeded to track

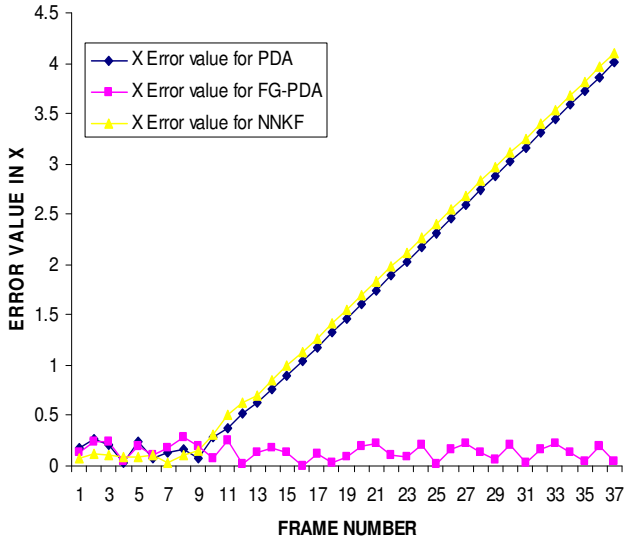


(a)

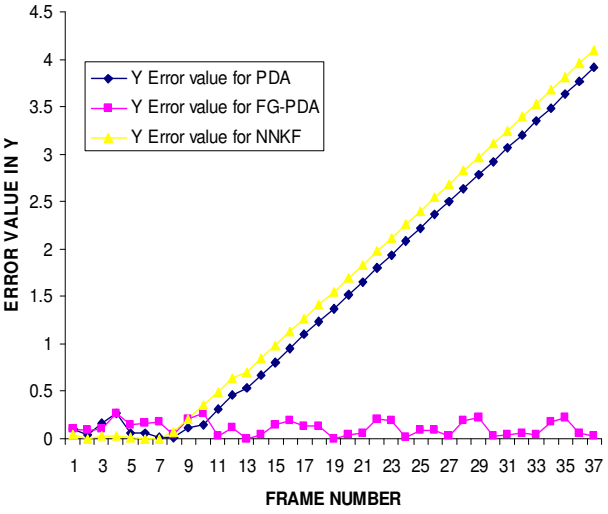


(b)

FIGURE 4: Trajectory for X and Y components for the 3 approaches algorithm used in tracking a target in dense clutter and the true target path (as presented in Fig. 3). (a) Trajectory for X (b) Trajectory for Y



(a)



(b)

FIGURE 5: Absolute error value for the 3 approaches algorithm in X- and Y-components where the error for FG-PDA is minimum compared to the increased error for PDA and NNKF due to failing in tracking (a) absolute error for X component (b) absolute error for Y component

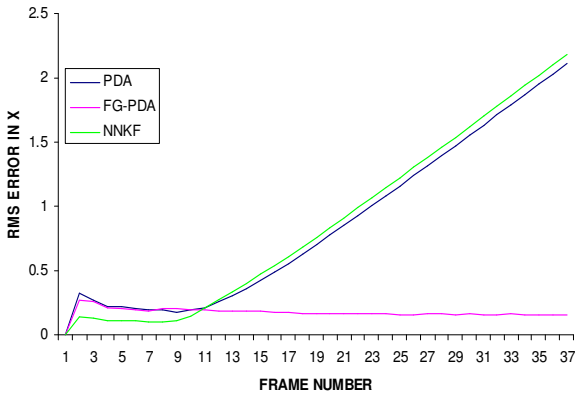


FIGURE 6: The root mean square error [RMSE] over frame number (each frame take 4 sec / one scan) for the 3 approaches algorithm and the RMSE is maintained minimum for the proposed FG-PDA and less sensitivity to dense clutter.

6. CONCLUSION

We have showed that the probabilistic data association filter (PDAF) is improved by avoiding the false targets from the valid based measurement region using a filtering method in dense clutter environment. This approach can be used to overcome the clutter of gate based approaches in tracking. With even high threshold values for gate size, we can obtain smaller validated measurement regions with improving data association Process which have been shown to give targets the ability to continue tracking in dense clutter.

7. REEFRENCES

[1] Y. Bar-Shalom and W. D. Blair." *Multitarget Multisensor Tracking:Applications and Advances"*, volume III. Archtech House, Norwood, MS, 2000.

[2] Y. Bar-Shalom and T. E. Fortmann. "*Tracking and Data Association"*. Academic Press, 1988.

[3] R. A. Singer, R. G. Sea, and K. B. Housewright." *Derivation and evaluation of improved tracking filters for use in dense multitarget environment*". IEEE Transactions on Information Theory, 20:423–432, 1974.

[4] Y. Bar-Shalom and E. Tse." *Tracking in a cluttered environment with probabilistic data-association*". Automatica, 1975.

[5] Y. Bar-Shalom. "*Tracking methods in a multitarget environment*". IEEE Transactions on Automated control, 23:618–626, 1978.

[6] D. B. Reid. "*An algorithm for tracking multiple targets*". IEEE Transactions on Automatic Control, 24:843–854, 1979.

[7] J. B. Collins and J. K. Uhlmann. "*Efficient gating in data association with multivariate Gaussian distribution states*". IEEE Transactions on Aerospace and Electronic Systems, 28(3):909–916, July 1992.

[8] S. S. Blackman and R. Popoli. "*Design and Analysis of Modern Tracking Systems".* Artech House, 1999.

[9] F. J. Breidt and A. L. Carriquiry." *Highest density gates for target tracking*". IEEE Transactions on Aerospace and Electronic Systems,36(1):47–55, Jan. 2000.

[10] X. Wang, S. Challa, and R. Evans." *Gating techniques for maneuvering target tracking in clutter*". IEEE Transactions on Aerospace and Electronic Systems, 38(3):1087–1097, July 2002.

[11] Y. Kosuge and T. Matsuzaki. "*The optimum gate shape and threshold for target tracking*". In SICE Annual Conference, 2003.

[12] D. Musicki and M. R. Morelande. "*Gate Volume Estimation for Target Tracking*". In International Conference on Information Fusion, 2004.

[13] M. Wang, Q. Wan, and Z. You." *A gate size estimation algorithm for data association filters*". Science in China, 51(4):425–432, April 2008.

[14] Ji Won Yoon and Stephen J .Roberts *"Robust Measurement Validation in Target Tracking using Geometric Structure"* IEEE Signal Pocessing Letters,17(5):493-496,May 2010

[15] Simon Haykin. "*Radar Signal Processing*"*, .* IEEE ASSP MAGAZINE*,* April 1985*.*

[16] G.Richard Curry. '*Radar System Performance Modeling*". Artich House, 2rd ed. Edition,2005.

[17] R. F. Stengel. "*Optimal Control and Estimation".* Dover Publications, 1994.

[18] G. W. Pulford and R. J. Evans. "*Probabilistic data association for systems with multiple simultaneous measurements*". Automatica, 32(9):1311– 1316, Set. 1996.

[19] X. R. Li and Y. Bar-Shalom. "*Stability evaluation and track life of the PDAF for tracking in clutter*". IEEE Transactions on Automatic Control, 36(5):588–602, May 1991.

[20] T. Kirubarajan and Y. Bar-Shalom." *Probabilistic Data Association Techniques for Target Tracking in Clutter*". Proceedings of the IEEE, 92(3):536–557, Mar. 2004.

[21] Gad.A,Majdi. F and Farooq. M. *"A Comparison of Data Association Techniques for Target Tracking in Clutter"* Proceedings of information Fusion, of the Fifth international conference vol 2:1126-1133,Nov 2002.

[22] V P S Naidu. "*Data Association and Fusion Algorithms for Tacking in Presence of Measurement Loss*". IE(I) Journal -AS, Vol 86, pages 17–28, May 2005.

F_0 Contour Modeling for Arabic Text-to-Speech Synthesis Using Fujisaki Parameters and Neural Networks

Zied Mnasri zied.mnasri@gmail.com
Ecole Nationale d'Ingénieurs de Tunis
Electrical Engineering Department
Signal, Image and Pattern Recognition Research Unit
University Tunis El Manar
Tunis, 1002, Tunisia

Fatouma Boukadida fatoumaboukadida@yahoo.fr
Institut Supérieur des Technologies Médicales
Electrical Engineering Department
University Tunis El Manar
Tunis, 1002, Tunisia

Noureddine Ellouze noureddine.ellouze@enit.rnu.tn
Ecole Nationale d'Ingénieurs de Tunis
Electrical Engineering Department
Signal, Image and Pattern Recognition Research Unit
University Tunis El Manar
Tunis, 1002, Tunisia

Abstract

Speech synthesis quality depends on its naturalness and intelligibility. These abstract concepts are the concern of phonology. In terms of phonetics, they are transmitted by prosodic components, mainly the fundamental frequency (F_0) contour. F_0 contour modeling is performed either by setting rules or by investigating databases, with or without parameters and following a timely sequential path or a parallel and super-positional scheme. In this study, we opted to model the F_0 contour for Arabic using the Fujisaki parameters to be trained by neural networks. Statistical evaluation was carried out to measure the predicted parameters accuracy and the synthesized F_0 contour closeness to the natural one. Findings concerning the adoption of Fujisaki parameters to Arabic F_0 contour modeling for text-to-speech synthesis were discussed.

Keywords: F_0 Contour, Arabic TTS, Fujisaki Parameters, Neural Networks, Phrase Command, Accent Command.

1. INTRODUCTION

TTS systems have known much improvement with a variety of techniques. However, naturalness is still a troublesome aspect, which needs to be looked after. In fact, naturalness is too large as a concept; it may be related to the speech synthesizer, which is required to produce an acoustic signal matching as closely as possible to the natural waveform, or to the listeners, who react perceptually to the sound they hear [1].

In both cases, it's prosody which is responsible of the naturalness quality. Prosody includes the underlying features spanning the speech segments. Again, there is a twofold definition of prosody, according to the adopted viewpoint. Thus, phonologically speaking, prosody stands for stress, rhythm and intonation. These terms describe the cognitive side of speech. Translated into phonetics, these abstract definitions are quantified by the signal's amplitude, duration and F_0 contour [1].

The latter feature, i.e. F_0 contour, is the physical transmitter of the acoustic information, whether linguistic, para-linguistic or non-linguistic. Linguistic information includes the lexical, syntactic and semantic data present in the text, while para-linguistic data describe the speaker's intention, attitude and style. Finally, non-linguistic information is related to his physical and emotional state [2].

All these sides, acting simultaneously to produce speech, need a model able to re-synthesize the F_0 contour transmitting them. Nevertheless, modeling F_0 contour is subject to many constraints, according to the approach to be used. Thus the way these data are dealt with is responsible of the modeling strategy, either in a timely manner, i.e. sequentially, or in a parallel and super-positional manner.

Besides, the presence/absence of parameters in the adopted model is a key index for its applicability. In fact, non-parametric modeling, in spite of its simplicity, is of lesser relevance than parametric modeling, which provides the opportunity of multi-language applicability.

For instance, the Fujisaki model is a super-positional model, inspired from the early works of Ohman [4], and developed to provide an analytical description of the phonatory control mechanism, through the introduction of three basic concepts:

1. The baseline frequency F_b
2. The phrase command
3. The accent command

Then the overall F_0 contour is calculated in the logarithmic domain as the superposition of the aforementioned concepts.

Since Fujisaki model is parametric, the main task is to measure its parameters. This can be done either by approximation, using the analysis-by-synthesis technique, [5] and [6], or by prediction, [22] and [23].

Amongst the prediction techniques, neural networks are famous for their generalization power, through capturing the latent relationship between an input set and the matching outputs, to be able to guess the value of any new coming sample. Nevertheless, supervised learning, i.e. specific and separate input and output sets, is highly recommended to ensure a faster convergence of the neural networks [7].

In the framework of Arabic TTS synthesis, we opted for a parametric tool, i.e. the Fujisaki model, to generate synthesized F_0 contours after the analysis of a phonetically balanced Arabic speech corpus.

Hence, we started by extracting the Fujisaki parameters from our corpus, using Mixdorff's tool [8]. Then neural networks were used to train a learning set, covering 80% of the corpus to predict the parameters related to the test set.

In this paper, we start by defining the different levels of intonation modeling, to locate the Fujisaki model and describe its components. Then, after a short description of our corpus and the extraction method, the selected neural architecture is explicitly shown, with the various involved phonological, contextual and linguistic features. Finally, synthesized F_0 contours and original ones are compared using statistical coefficients, and the synthetic parameters are discussed.

These instructions are for authors of submitting the research papers to the International Journal of Computer Science and Security (IJCSS). IJCSS is seeking research papers, technical reports, dissertation, letter etc for these interdisciplinary areas. The goal of the IJCSS is to publish the most recent results in the development of information technology.

2. INTONATION MODELING

2.1 Phonological vs. Phonetic Models

Phonological models attempt to explain the intonation system of a language through a set of principles of intonation and the relationship between intonation processes and the linguistic units, mainly syllables. This implies that phonological models are not directly related to the physical waveform, they are rather a symbolic representation.

In contrast, phonetic models focus more on the alignment of intonation movements with phonetic segments and their temporal location, which is a physical representation [1].

For example, the Tilt model is a phonetic intonation model using a continuous description of F_0 movements based on acoustic F_0 data [28] while ToBI model [3], is based on a linguistic survey which divides the speech tones into phonological categories having each its own linguistic function [9].

Then, F_0 pattern is inferred from the relationship between the phonological sets ad their assigned linguistic functions. However, phonological categories may linguistically interfere, causing mutual interaction between the linguistic functions. This phenomenon may seriously affect the model's results [10].

2.2 Rule-based vs. Data-driven Models

The rule-based models emphasize on the representations which capture maximal generality and focus on symbolic characterizations. This is completely compatible with the human way to produce speech, which variability depends on these abstract representations. But linguistics focus on the cognitive aspect of speech at the expense of data itself, which is, actually, the visible aspect.

On the opposite side, speech data-driven modeling doesn't require a close interaction with linguistics. Certainly some linguistic rules have to be considered or used to extract meaningful data, but do not interfere in processing, and do not impose major constraints on the output. This approach has proved that such models can explore areas that linguistics cannot reach, and give answers that they haven't found, thanks to its high-level computational processing, and also because of the difficulty to simulate the phonological and cognitive rules related to speech production.

Although these constraints can be modeled by using uncertainty, in the form of probability characterizations, it is still a wide and deep area to be explored, looking to the various and complicated interactions lying in [1].

In the case of intonation modeling, the rule-based modeling was used to generate F_0 by targeted interpolation [3]. Actually, linguistic rules are first set to define the target functions of every syllable structure. This function allows them to place the pitch target in the right range, between the top and base values [11].

Data-driven models are based on a phonetic and prosodic segmentation and labeling of the speech corpus. This data is used to predict either F_0 movements or F_0 values. For example the Tilt model is used for English pitch modeling using accent marks [11]. Besides, data-driven methods don't need a deep linguistic exploration and therefore are more adapted for statistical learning. Then they can be used to predict the pitch pattern using either a parametric representation or not.

2.3 Superpositional vs. Sequential Models

Superpositional models are built upon the idea that F_0 contour can be seen as the result of the interaction of many factors at different levels of the utterance, such as the phoneme, the syllable, the word...etc. Thus instead of processing the F_0 contour as a whole, the study can be split into many sub-models dealing each with a particular level, to be combined later to generate the desired F_0 contour.

Sequential models stand on the other edge. They aim to generate F_0 values or movements either directly or by means of parameters, but in both cases, they rely on a sole model moving from the

beginning to the end of the utterance. Hence, the components of the F_0 contour are generated together at any particular instant of speech [9].

2.4 Parametric vs. Non Parametric Models

In intonation modeling, parameterization consists in transforming the original F_0 values into some parametric forms. Hence, instead of predicting the F_0 values, it would be enough to predict the values of its parameters, to re-synthesize the F_0 contour.

In contrast, the non-parametric approach consists in estimating the F_0 values directly from a set of features. Though its simplicity and its direct scheme, the latter method provides equivalent results when compared to the first one. Actually, the F_0 values are considered as meaningful and intrinsic linguistic parameters, and thus, predictable from linguistic features.

However, non-parametric modeling proceeds directly, discarding the hierarchy and especially the interactions of the various input features. Hence, the components of the F_0 contour are ignored and its movements are neglected at the expense of its values. So forth, a post-processing is required to ensure the smoothness of the estimated F_0 contour.

This post-processing action is not required while parametric modeling, as the prediction of each parameter is made in a local level, e.g. the syllable, then smoothness is inherently processed. Nevertheless, care should be taken in the parameterization process, as too many parameters may provide better prediction accuracy, but on the other hand, may cause the loss of linguistic meaning [12].

3. FUJISAKI INTONATION MODEL

This is a phonetic data-driven superpositional and parametric model for intonation. It starts from a physiological interpretation of the intonation to provide an analytical description in the logarithmic domain through the introduction of some theoretical concepts such as the accent and the phrase commands [2].

Thus, the F_0 contour is considered as the response of the mechanism of the vocal cord vibration to the accent and phrase commands. So forth, the Fujisaki model gives a twofold description of the intonation:

1. A physiological and physical description of the phonatory control mechanisms through the shapes of phrase and accent components.
2. An analytical description through the magnitude, the timing and the superposition of the aforementioned commands.

3.1 Physiological Description

Fujisaki applied the Buchthal & Kaiser formulation of the relationship between the tension, *T*, and the elongation of skeletal muscles, *x*, to the vocal muscle

$$\begin{aligned} T &= a.(e^{bx} - 1) \\ &= a.e^{bx} \quad if \quad e^{bx} >> 1 \end{aligned} \tag{1}$$

and the relationship between the vibration frequency of elastic membranes and their tension

$$F_0 = C_0.T^{\frac{1}{2}} \tag{2}$$

to yield

$$Ln(F_0) = \frac{b}{2x} + ln(C_0.a^{\frac{1}{2}}) \tag{3}$$

Actually, the passage to the logarithmic domain is helpful to achieve a linear superposition, and so forth, a decomposition of the model into accent and phrase components. The constant $(C_0.a^{1/2})$ refers then to the baseline frequency $Ln(F_b)$ which is constant during each utterance.

As the elongation of the vocal muscle, x, is associated to the tension of the glottis, a further decomposition of the movement of the glottis into a rotation around the cricothyroid joint and a translation of the thyroid against the cricoids allows introducing the concept of the accent

component which refers to the rotation of the glottis, and the phrase component describing its translation [13].

3.2 Analytical Description

The Fujisaki model describes F_0 as a function of time in the logarithmic domain by achieving a linear superposition between:

1. The baseline frequency, which doesn't alter along the sentence
2. The phrase component
3. The accent component

The phrase and accent components are the outputs of 2 second-order linear systems, called the phrase and the accent commands [14]:

$$Ln(F_0(t)) = Ln(F_b) + \Sigma_{i=1}^{I} A_{pi} G_p(t - T_{0i}) + \Sigma_{j=1}^{J} A_{aj} [G_a(t - T_{1j}) - G_a(t - T_{2j})] \quad (4)$$

Where

$$G_p(t) = \begin{cases} \alpha^2 t e^{-\alpha t} & if \quad t > 0 \\ 0 & else \end{cases} \quad (5)$$

$$G_a(t) = \begin{cases} min(1 - (1 + \beta t).e^{\beta t}, \gamma) & if \quad t > 0 \\ 0 & else \end{cases} \quad (6)$$

The parameters A_p, T_0, A_a, T_1, T_2, α, β and γ are called the Fujisaki parameters.
As inferred by the formulation of $Ln(F_0)$, F_b denotes the asymptotic value of F_0 in absence of accent commands. Furthermore, it is proved that F_b is highly correlated to the mode of the sentence. It has higher values in direct Yes/No questions than in declarative statements [15].

1. **The Phrase Component**
   The phrase control mechanism is a second-order linear system whose impulse response is stated in (5).Then the output impulses are defined by their magnitude Ap and onset time T0. The parameter α is constant during an utterance. Hence, Ap describes the declination degree in an utterance, and therefore cannot be subject of comparison between different types of utterances.

2. **The Accent Component**
   The accent command is also a second-order linear system whose step-response is stated in (6). Then, the accent command introduces a magnitude Aa, an onset time T1 and an offset time T2. Besides, the parameter β is constant during an utterance. The same for γ, which is fixed to a ceiling value of 0.9 to ensure that the accent component will converge to its maximum in a finite delay. As T2 is usually higher than T1, the variation of F0 is proportional to the accent component magnitude Aa, which was extended to the negative domain to be able to apply the model to many other languages [15].

3. **F0 Contour Analysis**
   In order to obtain an optimal approximation of the F0 contour, the analysis by synthesis of the natural F0 contour is applied [6]. This is done by modifying the input commands of the model until:
   - The F0 contour is approximated
   - The result is linguistically interpretable

These constraints are either linguistic, describing the relationship between linguistic units and structures, or paralinguistic, dealing with phrase and accent components.

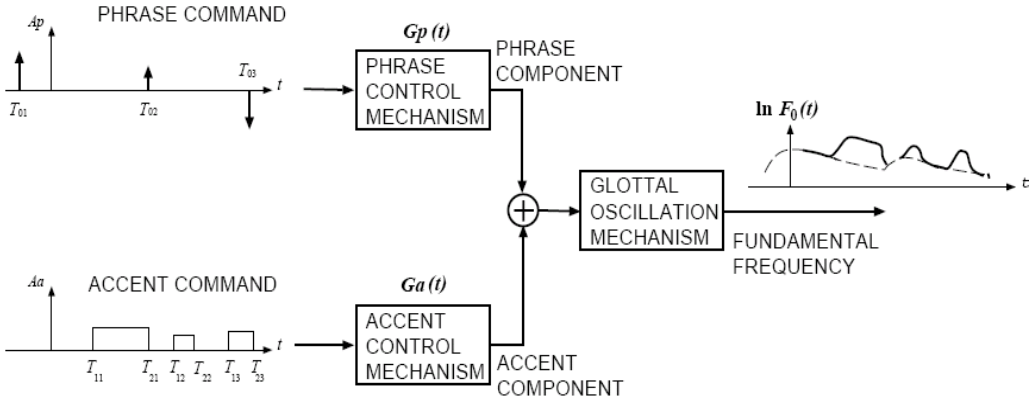


FIGURE 1: Fujisaki model components [4]

3.3 Motivation to use the Fujisaki model for Arabic

The Fujisaki model sound physiological background has been a great asset to its application to other languages. In fact, the model is based on the analytical formulation of the larynx movements, which are more related to the phonatory control system than to the linguistic rules. For instance, a polyglot speaker is able to produce, with his own larynx, the same intonation patterns as a native speaker [14].

Furthermore, the model can be modified to meet the specifications of the language to model. In fact, after its success in modeling, first Japanese [16], then English F_0 contours [36], it was adapted to many languages of different etymologies, Nordic (German and Swedish), Latin (Spanish and Portuguese) or south-east Asian (Chinese and Thai). This large use of the model brought many benefits to its primary formulation. For example, the falling accents can be modeled by introducing negative accent commands A_a.

Hence, the modified characteristics reveal the specific prosodic constraints of the language, while the invariable properties are mainly related to the speaker, regardless his language. Therefore, the Fujisaki model is potentially able to model the Arabic intonation.

4. SPEECH MATERIAL

4.1 Speech corpus

For this survey, we used a 200-Arabic-sentence corpus recorded by a male voice, with a 16-Khz sampling rate and 16-bit encoding, including the entire Arabic alphabet, composed by 28 consonants, 3 short vowels and 3 long vowels. In addition, amongst the 6 types of Arabic syllables, the most used ones are present in the corpus, i.e. /CV/, /CVV/, /CVC/ and /CVVC/ [29].

Corpus level	Type	Quantity
Sentence	Declarative	160
	Interrogative	20
	Imperative	20
Syllable	Opened and short /CV/	721
	Opened and long /CVV/	315
	Closed ans long /CVC/ and /CVVC/	519
Phoneme	Vowels	43%
	Consonants	57%

TABLE 1: Balanced Arabic corpus hierarchy and composition.

This corpus was first translated into phonetics, then segmented and labeled using spectrogram and waveform tools. The segmented data was stored in a database containing two levels: the predictors, i.e. the input features and the observations, i.e. the actual segmented durations. Then the main task while shaping the input space consists in classifying these features. Therefore, a twofold classification was suggested. The first part is linguistic, where segmented data are divided according to their contextual, positional and phonological aspects, and the second is statistical, as input data can be categorical or continuous. This classification generates a 2-dimension array where every factor is described according to its linguistic and numerical classes.

4.2 Fujisaki Parameters Extraction

Fujisaki constants, α, β and γ of the recorded voice were set at, respectively, 2/s, 20/s and 0.9 [17]. The Fujisaki parameters were obtained by Mixdorff's tool [18] which applies a multi-stage process called 'Analysis-by-Synthesis'. This process allows extracting the baseline frequency F_b, the phrase and the accent commands parameters through the minimization of the minimum square error between the optimal synthetic F_0 contour and the natural F_0 contour [6].The first step consists in quadratic stylization using the MOMEL algorithm [19] to interpolate the unvoiced segments and the short pauses within the F_0 curve, and to smooth the microprosodic variations due to sharp noises. Then, a high-pass filter is used to separate the phrase and the accent components through the subtraction of the filter output from the interpolated contour. This yields a low frequency contour containing the sum of phrase components and F_b. The third step consists in initializing the command parameters, i.e. A_p, T_0, A_a, T_1 and T_2. Finally, the synthesized contour is optimized, considering the interpolated contour as a target and the mean square error minimization as a criterion [18].

Input features types	Accent command input features
Phonological	• Sentence mode • Syllable type • Syllable accent level • Nucleus weight
Positional	• Accent command rank in sentence • Number of accent commands in sentence • Accented syllable position in sentence • Number of syllables in sentence • Nucleus position in accented syllable • Number of phonemes in accented syllable • Nucleus position in sentence • Number of phonemes in sentence
Contextual	• Nucleus duration • Previous phoneme's duration in accented syllable • Accented syllable's duration • Sentence duration • F_0 in beginning of accented syllable • F_0 at end of accented syllable • F0 movement in accented syllable
Extra features for A_a	• Predicted accent command duration (T_2-T_1)

TABLE 2: Accent command's input features

Input features types	Phrase command input features
Phonological	• Sentence mode • Phrase command syllable type • Phrase command syllable accent level • Nucleus weight
Positional	• Phrase command syllable position in sentence • Number of syllables in sentence • Nucleus position in phrase command syllable • Number of phonemes in phrase command's syllable • Nucleus position in sentence • Number of phonemes in sentence
Contextual	• Nucleus duration • Phrase command syllable duration • Sentence duration • Utterance's baseline frequency (F_b)
Extra features for T_0	• Ap predicted for phrase command

TABLE 3: Phrase command's input features

5. MODELING FUJISAKI PARAMETERS WITH NEURAL NETWORKS

5.1 General Neural Scheme

Neural networks are famous for their large ability to link the input to the output through a functional relationship. Therefore, they have been used in several intonation models, not only to predict F_0 values [20] or F_0 movements [21], but also F_0 parameters [22] and [23].

Though the modeling goal may differ, the approach is always the same. Neural networks are used to map a learning set, representing intonation-related features to a target set. Once learning is achieved, the model becomes able to predict the output from the test set inputs.
In most cases, the learning set includes 80% of the samples whereas the test set covers the remainder. The input features consist of a variety of speech characteristics, describing different aspects. However, these features should be selected looking to their correlation with the intonation. According to a previous survey we made about segmental duration using neural networks [30], the introduction of some features may give better prediction accuracy, whereas it's best to discard other features. Though the targets are different, as we are aiming to predict the Fujisaki parameters, it is still important to study the relevance of every feature or class of features.
Besides, the variety of the targets requires different implementation schemes for the neural networks. Hence, looking to our corpus composed of 200 separate sentences, we decided to assign a single phrase command for every sentence and at most one accent command for every syllable.
Then the task is to predict Fujisaki Parameters related to each aspect. For the phrase command, we need to predict A_p and T_0, and for the accent command, we need to predict A_a, T_1 and T_2. It looks obvious that we are looking for different types of targets at different levels, i.e. amplitudes and temporal locations for sentences and syllables. Therefore, the neural processing should be carried out in a parallel and distributed way to predict every single target on its own. Furthermore, we have noted that the targets themselves are well correlated with each other at each level, i.e. A_p and T_0, A_a and (T_2-T_1). Therefore we set a double tier strategy where, at every level, i.e. the sentence or the syllable level, the predicted parameters are used as inputs to predict the other ones. This strategy has been helpful to give better results, as it captures the latent relationship between the amplitude and the location of phrase and accent commands.
In addition, we opted for a 4-layer feed-forward neural network to model each parameter, i.e. using 2 hidden layers and a single-node output layer. Actually, it's been proved that a 2-hidden-layer neural network is able to model any continuous vector-valued function [24]. The first hidden layer is used to capture local features from the input, while the second hidden layer is required to capture the global features.
For the activation functions, we used a linear function at the output whereas they were non linear at the hidden layers, respectively the logistic sigmoid function and the tangent hyperbolic function.

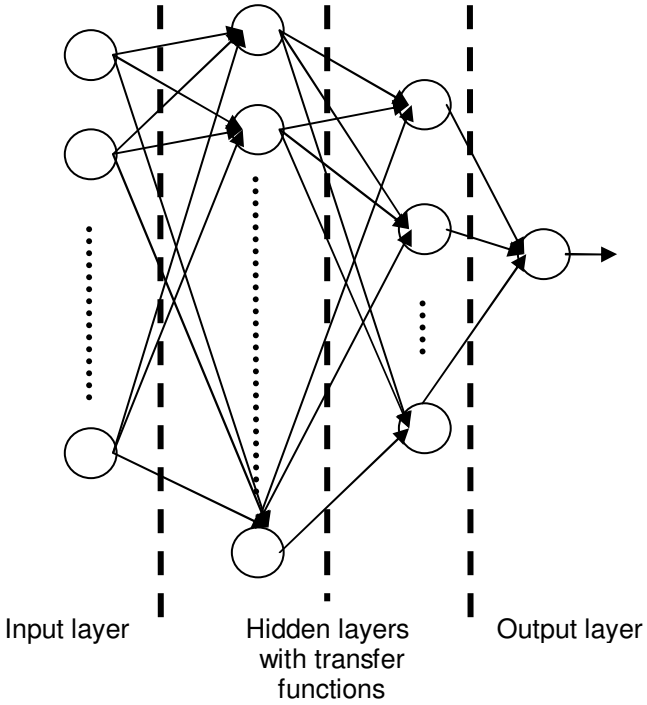


FIGURE 2: General neural network's scheme : a 2-hidden-layer FFNN with respectively, sigmoid and hyperbolic tangent transfer functions.

5.2 Modeling Phrase Command

In order to predict the phrase command parameters, i.e. A_p and T_0, we used an input set covering the contextual, phonological and positional features extracted from the corpus.

It's necessary to stress that the sentences of the corpus are totally distinct, which implies that contextual features are strictly internal to the sentence. Therefore we don't encounter any feature like 'previous phrase command amplitude' or 'previous phrase command timing' in the learning set.

However, looking to the high correlation between A_p and T_0, we start by predicting the A_p to be used later as input for T_0 prediction, but again, only within the same sentence. Also, to normalize the output, we opted to use the logarithm of the squared values for the temporal locations of the phrase commands.

5.3 Modeling Accent Command

In order to build the learning set to be mapped to the extracted targets, a certain number of constraints have to be considered:

1. Only accented syllables have accent commands, and one syllable have at most one accent command.
2. Each accent command is defined by its amplitude A_a, its onset time T_1 and offset time T_2. If the accented syllable lies in the beginning of the sentence, the accent command can start before the accent group, but in any case, the accent command cannot end beyond the accent group [25].

In addition to these constraints, we adopted 2 major modifications to the Fujisaki model, which were suggested to allow its extension to other languages, i.e. German and Swedish:

1. For the interrogative sentences, a final rise component is added. It's represented by an extra accent command [26].
2. Accent command magnitude, A_a, can be negative [27].

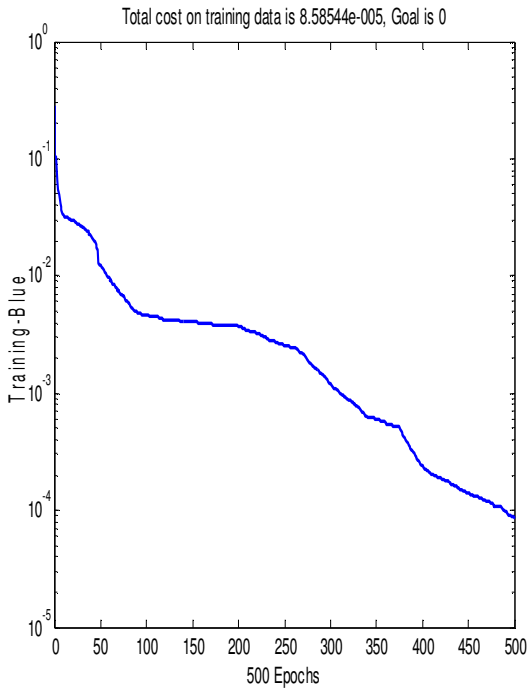


FIGURE 3: A_p training error

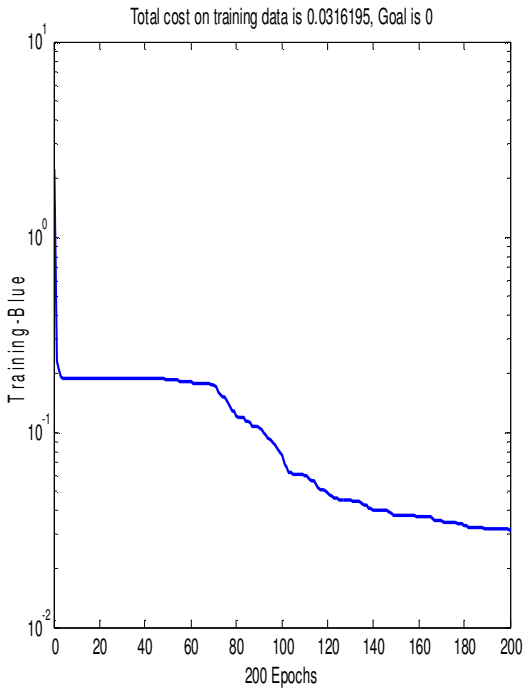


FIGURE 4: T_0 training error

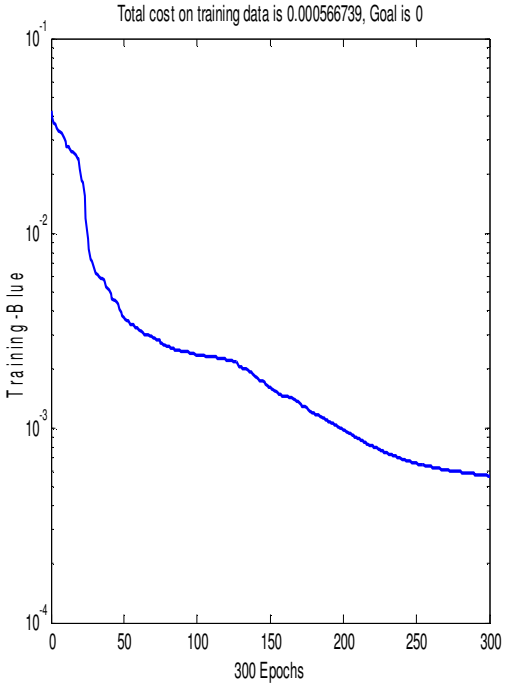


FIGURE 5: A_a training error

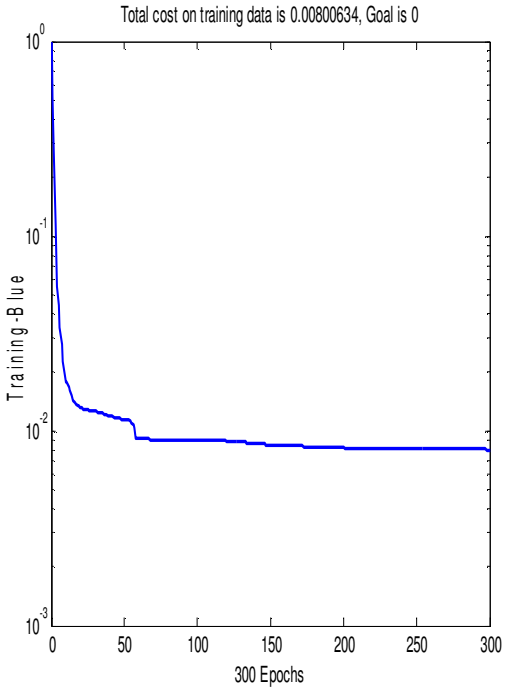


FIGURE 6: (T_1,T_2) training error

5.4 Selection of Input Features

For both levels, phrase and accent commands, the input features are treated according to their hierarchical class, i.e. the phoneme, the syllable or the sentence. It's important to note that for Arabic, the word is the acoustic unit between 2 successive pauses. However, the phonetic transcription doesn't always match with the text, because of the frequent presence of word-

boundary-linking syllables. To cope with this problem, we opted for the syllable as the acoustic unit.
The other criteria of features selection are firstly, the type of data values, whether discrete or continuous, and secondly their classes, i.e. contextual, phonological or positional.
Actually, such a broad classification is widely used while dealing with neural networks. As learning is supervised, these classes have to be defined to reduce the scarcity of data incurring a high-dimension input space, and to get rid of unnecessary data which may reduce the learning performance. This pre-processing is also useful to avoid the over-learning problem. Actually, too many inputs may yield generalizing the learning exceptions.

6. EVALUATION AND DISCUSSION

After training the neural networks for each parameter on its own, the test phase is carried out jointly with statistical evaluation. Thus we used the following statistical coefficients to measure the accuracy of the model:

- Mean absolute error

$$\mu = \frac{\sum_i | x_i - y_i |}{N} \quad (7)$$

- Standard deviation

$$\sigma = \sqrt{\frac{\sum_i d_i^2}{N}}, \; d_i = e_i - e_{mean}, \; e_i = x_i - y_i \quad (8)$$

- Correlation coefficient

$$\gamma_{X,Y} = \frac{V_{X,Y}}{\sigma_X . \sigma_Y} \quad (9)$$

$$V_{X,Y} = \frac{\sum_i |(x_i - x_{mean})| . |(y_i - y_{mean})|}{N} \quad (10)$$

X and Y are the actual and the predicted F0 values.

Then we used the test results to build synthetic F_0 contours to be compared with the original ones, which voiced parts were extracted by SHR algorithm [31] whereas we used Praat software [32] to locate the voiced and unvoiced parts of speech.

Statistical coefficients	With extracted F_0 contour by SHR
Mean absolute error	38.59 (29.75%)
Correlation in voiced parts	0.78
Standard deviation	43.12

TABLE 4: Statistical evaluation of synthesized F0 contours with extracted F0 contours in the test set.

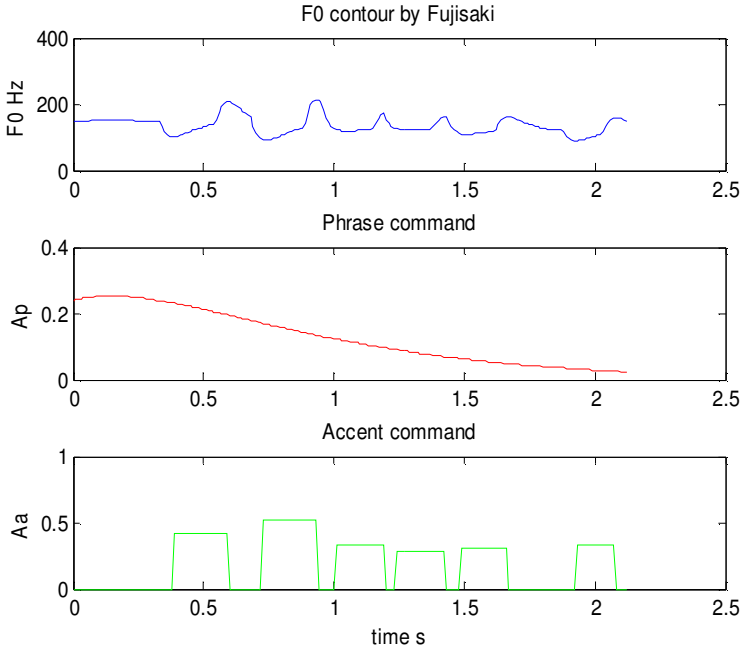


FIGURE 7: Synthesized F_0 contour, AC and PC of the Arabic sentence``hal ka:na juqabilukuma: ?" ( ``Was he meeting with both of you?")

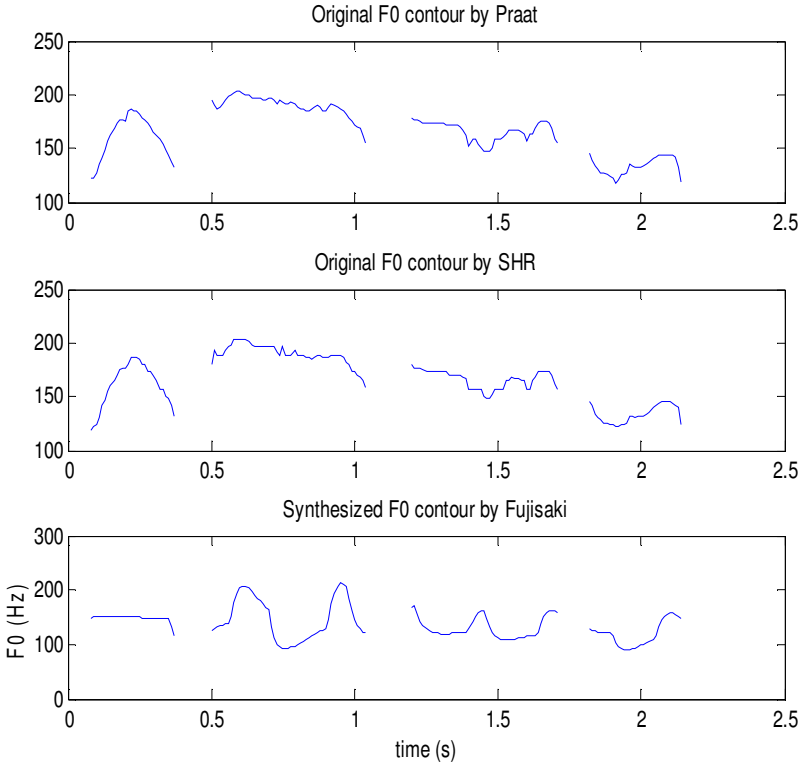


FIGURE 8: Extracted and synthesized F_0 contours in voiced parts of the Arabic sentence ``hal ka:na juqabilukuma: ?" ( ``Was he meeting with both of you?")

Though we opted to process the Fujisaki parameters separately, we have noticed a certain number of dependencies between them, especially within each level, i.e. the phrase and the accent groups. Actually, the single-node-output-layer scheme has been advantageous since it allowed the use of some already predicted parameters as inputs for the others.
Although it's still difficult to appreciate the quality of speech relatively to the synthesized F_0 contour before perceptual tests, the correlation between the natural and synthetic F_0 contour can tell about the accuracy of the model.
In addition to the statistical evaluation, these experiments revealed a certain number of notes about, not only the output parameters, but also about the predictors, i.e. the input features:

6.1 The Phrase Command's Evaluation

- **The Phrase Command' Amplitude A_p:** It's highly correlated with the duration of the sentence. It has higher values for longer sentences

- **The Phrase Command's Timing T_0:** This parameter is particularly critical, since it describes the very beginning of the phrase component, which precedes the onset of the utterance. Looking to our corpus, which is composed of separate sentences, T0 should usually be negative, unless there is a silence in the beginning of the utterance.

Generally, the phrase command is higher, only in the beginning, to decay along the sentence, and therefore the accent command is more relevant in the global shape of F_0 contour.

Phrase command parameter	Mean absolute error	Mean Value	Correlation
A_p	0.279	0.779	0.913
T_0	0.072	-0.409	0.508

TABLE 5: Phrase command parameters statistical evaluation in the test set

6.2 The Accent Command's Evaluation

- Accent command's onset and offset (T_1, T_2): Unlike the phrase component, the accent component is characterized by a variable range of temporal locations. If we consider (T2-T1) as the accent group duration, and according to a previous study we made about the segmental durations [30], then the accent command onset T_1 and offset T_2 are mainly related to the phonological, contextual and positional features of the speech. Yet, a good result of the accent group prediction has been helpful to predict its amplitude A_a.

- Accent command's magnitude A_a: A good prediction of A_a requires, not only information about the speech characteristics, but also about the neighboring accent groups, i.e. A_a and (T2-T1) of the previous accent group, and (T2-T1) of the actual accent group. In fact, the accent component is responsible of the overall shape of the F_0 contour after the decay of the phrase component, and so forth, the amplitude, the temporal locations of the previous accent group, coupled with the F_0 movement, provide a relevant indication about the accent command amplitude.

The assumption made in the beginning, stating that A_a values can be expanded to the negative domain, was verified in the test set. However, synthetic and natural A_a values, for the same accent group, don't have always the same sign. This is due, from one side, to the prediction accuracy, and from the other side, to the interaction of the phrase and accent commands, which are superposed to generate a natural looking F_0 contour. Besides, the contextual input features, such as (T2-T1), A_a and F_0 movement of the previous accent group, act jointly as correction agents to keep watching the fall/rise evolution of the synthesized contour.

Accent command parameter	Mean absolute error	Mean Value	Correlation
A_a	0.205	0.438	0.595
T_1	0.136	1.134	0.903
T_2	0.141	1.332	0.918

TABLE 6: Accent command parameters statistical evaluation in the test set

6.3 Discussion

- **Fujisaki Constants (α, β):** In our model, we used average values of the angular frequencies α=2/s and β=20/s. The variation of these factors doesn't alter significantly the results. Actually, the tool we used to extract the Fujisaki parameters, [8], allows changing their values, but without a great influence on the resulting contour. This confirms many previous researches concluding that α and β are mainly speaker-dependent, such as English [36], German [13] and Japanese [4] since they characterize the dynamic properties of the glottal control mechanism. [4].

- **Input Features:** A good approximation of F_0 contour needs, not only a good model, but mainly a representative learning set, able to describe as closely as possible, the various interactions and synergies between different contributory types of data. Hence, after the feature selection phase, pre-processing is required either to normalize input data or to broad-classifying them. In fact, a heterogeneous input set, including different data classes with different value ranges may fall in over-fitting, where exceptions could be generalized, or may require a high calculation time before the training error falls down to an acceptable minimum. Back to the obtained results, the pre-processing phase was of great help to tune the output values, improve accuracy and minimize training error. It also allowed capturing the inherent relationship between some input features and their relative outputs, and therefore guided us to build a personalized predictors set for each following parameter.

7. CONCLUSION

In the general framework of developing an Arabic TTS system, spotlight was focused on modeling F_0 contour. This task was performed using the Fujisaki model, which parameters were extracted by Mixdorff's analysis-by-synthesis-based tool and trained by neural networks.
Therefore, a phonetically-balanced Arabic corpus was analyzed, firstly to extract the Fujisaki parameters and secondly to select and pre-process the input features used as predictors for the neural network. Neural networks were hired for their large ability to capture the hidden functional mapping between input and output sets. Many neural schemes were suggested during the elaboration of this work, to select those which performed best in the try-and-error test. Then, statistical coefficients were calculated between actual and predicted Fujisaki parameters.
This study revealed also that Fujisaki parameters are dependent at each of the phrase and accent levels. Therefore, some of them were used as inputs to predict the other ones. In fact, an interaction was noted between the phrase and accent component to keep the overall shape of F_0 contour. Thus, after the decay of the phrase command, the accent command rises, and so forth, the F_0 contour becomes more sensitive to the accent variations. This note was checked out by the introduction of correction agents to the input feature while training. Furthermore, some of Fujisaki's assumptions were verified in this study. Thus negative accent commands were necessary to model the variations of the accent group; and the variation of some parameter

values such as α and β didn't have a relevant impact on the results, confirming that they are rather speaker-dependant.
As a future projection, this model can be used along with our previous study on segmental duration modeling [30], to build an integrated model of Arabic prosody, able to generate automatically the duration and the F_0 contour of an input text. Also, this study can be expanded to a paragraph-composed corpus in different Arabic dialects.

8. REFERENCES

1. M. Tatham, K. Morton, *"Developments in speech synthesis"*, John Wiley & Sons Inc. (2005)

2. H. Fujisaki, *"Prosody, information and modeling with emphasis on tonal features of speech",* in Proceedings of Workshop on spoken language processing, ISCA-supported event, Mumbai, India, January 9-11, 2003

3. J. B. Pierrehumbert, *"The phonology and phonetics of English intonation",* Ph. D. Thesis, MIT, Cambridge, 1980

4. H. Fujisaki, *"Dynamic characteristics of voice fundamental frequency in speech and singing. Acoustical analysis and physiological interpretations"*.STL-QPSR,1981, Vol. 22(1), pp 1-20, KTH, Sweden

5. S. Narusawa, N. Minematsu, K. Hirose and H. Fujisaki, *"Automatic extraction of model parameters from fundamental frequency contours of English utterances",* in Proceedings of ICSP'2000, pp 1725-1728, Denver, Colorado, USA

6. H. Mixdorff, H. Fujisaki, G. P. Chen and Y. Hu, *"Towards the automatic extraction of Fujisaki model parameters for Mandarin"*, in Proceedings of Eurospeech'03, pp 873-976, Geneva, 2003

7. M.Vainio, *"Artificial Neural networks based prosody models for Finnish text-to-speech synthesis",* PhD. Thesis, Helsinki University of Technology, Finland, 2001

8. H.J. Mixdorff, *"FujiParaEditor program"*, Available at http://www.tfh-berlin.de/~mixdorff/

9. J. Buhmann, H. Vereecken, J. Fackrell, J. P. Martens and B. Van Coile, *"Data driven intonation modeling of 6 languages"*, in Proceedings of International conference on spoken language processing, October 2000, Beijing, China, Vol. 3, pp 179-183

10. K. S. Rao and B. Yegnanarayana, *"Intonation modeling for Indian languages"*, Computer speech and language Journal, Volume 23, pp 240-256, Elsevier, 2009

11. G. P. Giannopoulos and A. E. Chalamandaris, *"An innovative F0 modeling approach for emphatic affirmative speech, applied to the Greek language"*, in Speech Prosody 2006, Dresden, Germany

12. X. Sun, *"F0 Generation for speech synthesis using a multi-tier approach",* in *Proceedings of ICSLP'02*, Denver, 2002, pp 2077-2080

13. H. Mixdorff, *"An integrated approach to modeling German prosody",* Habilitation Thesis, Technical University of Dresden, Germany, 2002

14. H. Fujisaki and S. Ohno, *"Prosodic parameterization of spoken Japanese based on a model of the generation process of F_0 contours",* in Proceedings of ICSLP'96, vol 4, pp 2439-2442, Philadelphia, PA, USA, Oct. 1996

15. B. Moebius, *"Synthesizing German F0 contours",* in J. Van Santen, R. Spraot, J. Olive and J. Hirschberg, *Progressin speech synthesis,* Chapter 32, pp 401-416, Springer Verlag, New York, 1997

16. H. Fujisaki and K. Hirose, *"Analysis of voice fundamental frequency contours for declarative sentences of Japanese"*, in Journal of the acoustic society of Japan (E), 5(4), pp 233-241, 1984

17. H. Mixdorff and O. Jokisch, *"Building an integrated prosodic model of German",* in Proceedings of Eurospeech 2001, Aaloborg, Denmark, vo2, pp 947-950

18. H. Mixdorff and O. Jokisch, *"Evaluating the quality of an integrated model of German prosody",* International journal of speech technology, Vol 6, pp 45-55, 2003

19. D. Hirst, A. Di Cristo and R. Espesser , *"Levels of representation and levels of analysis for inronation in M.Horne, Prosody: Theory and experiment",* Kluwer editions, Dortrecht, 2000

20. K. S. Rao and B. Yegnanarayana, *"Intonation modeling for Indian languages"*, in Proccedings of Interspeech'04,Jeju Island, K0rea, 4-8 October 2004, pp733-736

21. G. Sonntag, T. Portele and B. Heuft, *"Prosody generation with a neural network: Weighing the importance of input parameters"*, in Proceedings of ICASSP, pp 931-934, Munich, Germany, April 1997

22. J. P. Teixieira, D. Freitas and H. Fujisaki, *"Prediction of Fujisaki model's phrase commands"*, in Proceedings of Eurospeech 2003, Geneva, pp 397-400

23. J. P. Teixiera, D. Freitas and H. Fujisaki, *"Prediction of accent commands for the Fujisaki intonation model"*, in Proceeding of Speech Prosody 2004, Nara, Japan, March 23-26, 2004, pp 451-454

24. J. J. Hopfield, *"Neural networks and physical systems with emergent collective computational abilities"*, Proceedings of the National Academy of Sciences of the USA, vol. 79 no. 8 pp. 2554-2558, April 1982

25. E. Navas, I. Hernaez, A. Armenta, B. Etxebarria and J. Salaberria, *"Modeling Basque intonation using Fujisaki's model and CARTS"* , in Proccedings of ICSLP 2002, Denver, USA, pp 2409-2412

26. H. Mixdorff, *"Intonation patterns of German-model-based quantitative analysis and synthesis of F_0 contours"*, Ph. D. Thesis, TU Dresden, 1998

27. H. Fujisaki, S. Ohno and S. Luksaneeyanawin, *"Analysis and synthesis of F_0 contours of Thai utterances based on the command-response model"*, in Proceeding of 15th ICPhS, Barcelona, Spain, 2003, pp 1129- 1132

28. P. Taylor, *"Analysis and synthesis of intonation using the Tilt model",* Journal of Accoustic society of America, No 107, pp 1697-1714, 2000

29. F. Boukadida, *"Etude de la prosodie pour un système de synthèse de la parole Arabe standard à partir du texte",* Thèse de doctorat, Université Tunis El Manar, 2006

30. Z. Mnasri, F. Boukadida and N. Ellouze, *"Modelling segmen1tal durations by statistical learning for an Arabic TTS system",* International Revue on Computer and Software, September 2009

31. X. Sun, *"SHR program"* , available at http://mel.speech.nwu.edu/sunxj/pda.htm, Copyrightht © 2001, X.Sun, Department of communication sciences and disorders, Northwestern University, USA

32. P. Boersma and D. Weenink, *"Praat: Doing phonetics by computer, version 4.4",* available at http://www.praat.org

33. A. Black and A. Hunt, *"Generating F0 contours from ToBI labels using linear regression"* ,in Proceedings of ICSLP, Philadelphia, Pennsylvania, 1996

34. K. Dusterhoff, A. Black and P. Taylor, *"Using decision trees within the tilt intonation model to predict F0 contours",* in Proceedings of Eurospeech, Budapest, Hungary, 1999

35. S. Sakai and J. Glass, *"Fundamental frequency modeling for corpus-based speech synthesis based on a statistical learning technique",* in Proceedings of IEEE ASRU 2003, Nov. 30-Dec. 4, 2003, St. Thomas, US Virgin Islands, pp 712-717

36. H. Fujisaki and S. Ohno, *"Analysis and modelling of fundamental frequency contours of English utterances"*, in Proceedings of Eurospeech'95, pp 985-988, Madrid, Sep. 1995

CALL FOR PAPERS

Journal: Signal Processing: An International Journal (SPIJ)
Volume: 5 **Issue:** 1
ISSN:1985-2339
URL: http://www.cscjournals.org/csc/description.php?JCode=SPIJ

About SPIJ

The International Journal of Signal Processing (SPIJ) lays emphasis on all aspects of the theory and practice of signal processing (analogue and digital) in new and emerging technologies. It features original research work, review articles, and accounts of practical developments. It is intended for a rapid dissemination of knowledge and experience to engineers and scientists working in the research, development, practical application or design and analysis of signal processing, algorithms and architecture performance analysis (including measurement, modeling, and simulation) of signal processing systems.

As SPIJ is directed as much at the practicing engineer as at the academic researcher, we encourage practicing electronic, electrical, mechanical, systems, sensor, instrumentation, chemical engineers, researchers in advanced control systems and signal processing, applied mathematicians, computer scientists among others, to express their views and ideas on the current trends, challenges, implementation problems and state of the art technologies.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for SPIJ.

SPIJ List of Topics

The realm of International Journal of Signal Processing (SPIJ) extends, but not limited, to the following:

- Biomedical Signal Processing
- Communication Signal Processing
- Detection and Estimation
- Earth Resources Signal Processing
- Industrial Applications
- Optical Signal Processing
- Radar Signal Processing
- Acoustic and Vibration Signal Processing
- Data Processing
- Digital Signal Processing
- Geophysical and Astrophysical Signal Processing
- Multi-dimensional Signal Processing
- Pattern Recognition
- Remote Sensing

- Signal Filtering
- Signal Processing Technology
- Software Developments
- Spectral Analysis
- Stochastic Processes
- Signal Processing Systems
- Signal Theory
- Sonar Signal Processing
- Speech Processing

IMPORTANT DATES

Volume: 5
Issue: 1
Paper Submission: January 31, 2011
Author Notification: March 01, 2011
Issue Publication: March /April 2011

CALL FOR EDITORS/REVIEWERS

CSC Journals is in process of appointing Editorial Board Members for ***Signal Processing: An International Journal (SPIJ)***. CSC Journals would like to invite interested candidates to join **SPIJ** network of professionals/researchers for the positions of Editor-in-Chief, Associate Editor-in-Chief, Editorial Board Members and Reviewers.

The invitation encourages interested professionals to contribute into CSC research network by joining as a part of editorial board members and reviewers for scientific peer-reviewed journals. All journals use an online, electronic submission process. The Editor is responsible for the timely and substantive output of the journal, including the solicitation of manuscripts, supervision of the peer review process and the final selection of articles for publication. Responsibilities also include implementing the journal's editorial policies, maintaining high professional standards for published content, ensuring the integrity of the journal, guiding manuscripts through the review process, overseeing revisions, and planning special issues along with the editorial team.

A complete list of journals can be found at http://www.cscjournals.org/csc/byjournal.php. Interested candidates may apply for the following positions through http://www.cscjournals.org/csc/login.php.

Please remember that it is through the effort of volunteers such as yourself that CSC Journals continues to grow and flourish. Your help with reviewing the issues written by prospective authors would be very much appreciated.

Feel free to contact us at coordinator@cscjournals.org if you have any queries.

Contact Information

Computer Science Journals Sdn BhD
M-3-19, Plaza Damas Sri Hartamas
50480, Kuala Lumpur MALAYSIA

Phone: +603 6207 1607
+603 2782 6991
Fax: +603 6207 1697

BRANCH OFFICE 1
Suite 5.04 Level 5, 365 Little Collins Street,
MELBOURNE 3000, Victoria, AUSTRALIA

Fax: +613 8677 1132

BRANCH OFFICE 2
Office no. 8, Saad Arcad, DHA Main Bulevard
Lahore, PAKISTAN

EMAIL SUPPORT
Head CSC Press: coordinator@cscjournals.org
CSC Press: cscpress@cscjournals.org
Info: info@cscjournals.org

COMPUTER SCIENCE JOURNALS SDN BHD
M-3-19, PLAZA DAMAS
SRI HARTAMAS
50480, KUALA LUMPUR
MALAYSIA