

International Journal of Advances in Telecommunications Electrotechnics, Signals and Systems

a publication of the International Science and Engineering Society



**Vol. 4, No. 3
2015**

ISSN: 1805-5443

www.ijates.org

I J
A T
E S² **International Journal of**
Advances in Telecommunications
Electrotechnics, Signals and Systems

a publication of the International Science and Engineering Society

Vol. 4, No. 3, 2015

ISSN: 1805-5443

Editor-in-Chief

Jaroslav Koton, Brno University of Technology, Czech Republic

Co-Editors

Ondrej Krajsa, Brno University of Technology, Czech Republic

Norbert Herencsar, Brno University of Technology, Czech Republic

Editorial Board

Oguzhan Cicekoglu, Bogazici University, Turkey

Sergey Ryvkin, Trapeznikov Institute of Control Sciences Russian Academy of Sciences, Russian Federation

Hongyi Li, Bohai University, China

Emilia Daniela Bordencea, TU Cluj-Napoca, Romania

Albert Abilov, Izhevsk State Technical University, Russian Federation

Joze Guna, University of Ljubljana, Slovenia

Jaroslav Koton, Brno University of Technology, Czech Republic

Ondrej Krajsa, Brno University of Technology, Czech Republic

Danilo Pelusi, University of Teramo, Italy

Aims and Scope

The International Journal of Advances in Telecommunications, Electronics, Signals and Systems (IJATES²) is an all-electronic international scientific journal with the aim to bring the most recent and unpublished research and development results in the area of electronics to the scientific and technical societies, and is supported by the ISES (International Science and Engineering Society, o.s.). The journal's scope covers all the aspects of telecommunication, signal processing, theory and design of circuits and systems for electronics.

The IJATES² is ready to publish experimental and theoretical full papers and letters submitted by prospective authors. Paper submitted for publication must be written in English and must follow a prescribed format. All papers are subjected to a critical peer-review prior to publication.

The IJATES² is an open access journal which means that all content is freely available without charge to the user or his/her institution. Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the articles in this journal without asking prior permission from the publisher or the author. This journal provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge.

www.ijates.org

Copyright © 2012-2015, by ISES, o.s.

All the copyright of the present journal belongs to the International Science and Engineering Society, o.s.

CONTENTS

Vol. 4, No. 3, 2015

ISSN: 1805-5443

Test Program for the Performance Analysis of DNS64 Servers <i>Gábor Lencse</i>	60
Non Minimum Phase Channel Estimation by Blind and Adaptive algorithms <i>Zidane Mohammed Said Safi Mohamed Sabri, Ahmed Boumezzough</i>	66
Design and FPGA Implementation of Variable Cutoff Frequency Filter based on Continuously Variable Fractional Delay Structure and Interpolation Technique <i>Sumedh Dhabu, Achutavarrier Prasad Vinod</i>	72
Linear Robot as The Approach towards Individual Abdominal Ultrasound Scanning in Developing Countries? <i>Christina Pahl</i>	80
DNA Repeats Detection Using a Dedicated Dot-Plot Analysis <i>Gavril Petre Pop</i>	90

Test Program for the Performance Analysis of DNS64 Servers

Gábor Lencse

Abstract—In our earlier research papers, bash shell scripts using the `host` Linux command were applied for testing the performance and stability of different DNS64 server implementations. Because of their inefficiency, a small multi-threaded C/C++ program (named `dns64perf`) was written which can directly send DNS AAAA record queries. After the introduction to the essential theoretical background about the structure of DNS messages and TCP/IP socket interface programming, the design decisions and implementation details of our DNS64 performance test program are disclosed. The efficiency of `dns64perf` is compared to that of the old method using bash shell scripts. The result is convincing: `dns64perf` can send at least 95 times more DNS AAAA record queries per second. The source code of `dns64perf` is published under the GNU GPLv3 license to support the work of other researchers in the field of testing the performance of DNS64 servers.

Keywords—DNS64, DNS query, Linux, performance analysis, TCP/IP socket interface programming.

I. INTRODUCTION

The combination of DNS64 [1] and NAT64 [2] is the best available solution which can enable IPv6 only clients to communicate with IPv4 only servers. There are several DNS64 and NAT64 implementations exist and the network operators need to choose the one that best fits for their purposes. Performance, stability and security are all very important aspects.

Several papers were published concerning the performance of DNS64 and NAT64 implementations, e.g. [3]-[5], however their authors measured the common performance of the combination of a given DNS64 implementation and a given NAT64 implementation. We have shown in [6] that the performance of the NAT64 and DNS64 implementations should be measured independently.

In our previous works [6]-[8], we measured the performance of several DNS64 implementations using single core and also multi-core test devices. We used simple bash shell scripts for load generation (see details in section II), but they proved to be inefficient and therefore we decided to prepare a small special purpose C/C++ test program. We do not know of any other similar test program for the performance analysis of DNS64 servers, except for the general DNS server performance test program called `DNSPerf` [9] which could be used, however it uses a text file as input [10] and thus the reading and processing of the text file may be a

bottleneck. (The uncompressed size of its sample query input file [11] is 248,609,309 bytes.) As for the hardware platform of the load generator for our further DNS64 performance measurements, we plan to use a cluster of SBCs (Single Board Computer), similar to the one documented in [12] but with more powerful SBCs. (We have already benchmarked some of them [13] and we are currently designing a 16 element cluster.) The SBCs use micro SD cards as storage and they produce poor transfer rate (e.g. 10-20MB/s), thus the reading of hundreds of megabytes size files from a micro SD card may be a serious bottleneck. Therefore we considered that a small special purpose DNS64 performance test program was worth writing. In this paper, we discuss the design, implementation and testing of our DNS64 performance test program.

The remainder of this paper is organized as follows. The DNS64 server performance testing algorithm to be implemented by the test program is discussed in section II. The theoretical background including the structure of the DNS messages and the basics of TCP/IP socket interface programming is summarized in section III. The design and implementation questions are detailed in section IV. The method of the comparison of the performances of the old bash scripts and of the new test program is disclosed and the results are presented and discussed in section V. Our plans for future research are outlined in section VI.

II. ALGORITHM FOR THE TEST PROGRAM

We were satisfied with the testing method used in [6]-[8] therefore our goal was to design and implement a test program which performs very similar tests but at least one order of magnitude faster.

The original testing method was designed to eliminate the effect of caching therefore it requested name resolution for different domain names. Originally, we used the namespace of `10-{0..10}-{0..255}-{0..255}.zonat.tilb.sze.hu` which was mapped to the `10.0.0.0 - 10.10.255.255` IPv4 address only and to no IPv6 address. (It was the task of the DNS64 server to synthesize the so called *IPv4 Embedded IPv6 Addresses* [14] for these domain names.)

In order to make a well tunable load generator, we used 1-8 computers, which executed the same bash scripts. The *synchronized start* of the client scripts was done by using the “Send Input to All Sessions” function of the terminal program of KDE (called `Konsole`) in [6] and [7] and it was done by using multicast in [8].

To ensure non-overlapping namespace for each client computers, the number of the computers were included into the domain name generation: it was used as the number right

Manuscript received June 3, 2015, revised August 29, 2015.

G. Lencse is with the Department of Telecommunications, Széchenyi István University, Győr, Hungary (phone: +36-20-775-82-67, fax: +36-96-613-646, e-mail: lencse@sze.hu)

10.11601/ijates.v4i3.121

after the beginning “10-” string. The two other numbers from 0 to 255 made it possible for each clients to generate 65536 different domain names. They were grouped into 256 experiments (256 requests were sent in each of them) and the time of each experiments were measured. Two bash scripts were used. The first one was:

```
#!/bin/bash
i=`cat /etc/hostname|grep -o .$`
rm dns64-$i.txt
for b in {0..255}
do
    /usr/bin/time -f "%E" -o dns64-$i.txt \
        -a ./dns-st-c.sh $i $b
done
```

As it can be seen from the code above, the bash script performs 256 experiments and measures the execution time of each execution of the `dns-st-c.sh` script. The latter script was changed in [8] to request only the AAAA record (without the `-t AAAA` option it also requested the MX record in [6] and [7]). The change can be justified by the fact that though requesting the MX record is the default behavior of the host command, it is not need in the vast majority of the cases (e.g. when opening a web page). Thus we decided to implement the latest version of the script in our program. It is the following:

```
#!/bin/bash
for c in {0..252..4} # that is 64 iterations
do
    host -t AAAA 10-$1-$2-$c.zonat.tilb.sze.hu &
    host -t AAAA 10-$1-$2-$(($c+1)).zonat.tilb.sze.hu &
    host -t AAAA 10-$1-$2-$(($c+2)).zonat.tilb.sze.hu &
    host -t AAAA 10-$1-$2-$(($c+3)).zonat.tilb.sze.hu &
done
```

The script was designed to issue four `host` commands concurrently to utilize the computing power of all the four cores of our client computers.

Of course, the interpretation of the bash shell script and the starting of 65536 `host` commands by each clients are rather time and computing power consuming. We need a much more efficient test program to be able to measure the performance of multi-threaded DNS64 implementations executed by current multi-core computers with several CPU cores. Replacing the two bash shell scripts by a single (but multi-threaded) C/C++ program can be at least an order of magnitude faster.

III. THEORETICAL BACKGROUND

To be able to implement the presented algorithm in C/C++ (or just to follow the design steps) one need to know the structure of the DNS messages and the basics of TCP/IP socket interface programming.

A. The Structure of DNS Messages

We have given an in-depth introduction to the structure of DNS messages in [15], where we disclosed the design principles of MTD64, our new multi-threaded DNS64 implementation. Now, we give a very short summary of the information relevant for our test program on the basis of [15]. We mention only those fields of the DNS message here, which were actually used in our program. (For more

details see [15] or the specifying RFC [16].)

1) Top level structure

The DNS query and reply messages usually travel over UDP. A DNS message is built up by five sections, these are: *Header*, *Question*, *Answer*, *Authority*, *Additional* sections. Whereas the length of the *Header* section is always 12 bytes, the latter four sections have variable length and some of them may be empty.

2) Header section format

The *Header* section is built up by six 16-bit fields. The first one is *Transaction ID*, which is used by the client to identify the answer of the server for different questions. The second 16-bit field is decomposed into smaller fields and bits, out of which we had to set only the *RD* (Recursion Desired) bit to ask recursive query. (The *QR* bit should be set to 0 to specify query and the also 0 value of the *OPCODE* field means “Standard Query”, what we need.) The third 16-bit field is the *QDCOUNT* field which specifies the number of entries in the *Question* section.

3) Question section format

From among the latter 4 sections we deal with the *Question* section only. It contains *QDCOUNT* number of entries (usually 1). An entry is build up by the following three fields: The variable length *QNAME* field contains the domain name using special encoding (see: Domain name encoding). The *QTYPE* filed specifies the *RR* (Resource Record) type by 16-bit long binary vales. An AAAA record (IPv6 address) is denoted by the value of 0x1C. The *QCLASS* field contains the 0x01 16-bit binary value for denoting the *IN* (Internet) class.

4) Domain name encoding

The domain names stored in the *QNAME* field follow special encoding. A domain name is built up by so called *labels* separated from each other by “.” characters. The labels must be no longer than 63 characters. When domain names are encoded in DNS messages, the first character gives the length of the first label then the characters of the first label follow. After that, a character stands that specifies the length of the next label and the characters of the next label follow, etc. Finally, a zero character after the last label signals the end of the domain name. We illustrated this encoding with a detailed example in [15].

B. Basics of TCP/IP Socket Interface Programming

Before a socket can be used for communication, it must be created (or opened) by the `socket()` function call. Whereas programmers must use the `bind()` function call for assigning IP address and port number information to server sockets, it can be omitted for clients sockets. Functions `sendto()` and `recvfrom()` can be used for sending and receiving UDP packets, respectively. The latter one is a so-called ‘blocking’ function call. To change its behavior, timeout can be set by using the `setsockopt()` function. The return value of the socket handling functions should be checked for error and the standard `perror()` function can be used for giving appropriate error messages to the user.

As for the most important data structures, the `sendto()` function takes the IP address and port number of the destination host (together with some other pieces of information) from a structure of type `struct sockaddr`. If

IPv6 is used then the actual structure is of type **struct sockaddr_in6** (and its pointer has to be type casted to **struct sockaddr**). The same type of data structures are used by **recvfrom()** to store the parameters of the sender of the received packet.

All the above mentioned functions and data structures are well documented in the man pages of the Linux operating system.

IV. DESIGN AND IMPLEMENTATION OF THE TEST PROGRAM

A. High Level Requirements and Design Decisions

To achieve the highest possible speed, the whole functionality of the two bash shell scripts should be included into a single program. The program should be able to utilize the computing power of all the CPU cores of the computer used for its execution. Therefore, we decided that the program be multi-threaded. As for programming language, C was found to be appropriate and C++ was used only for thread handling.

The program should follow the testing method defined by the bash scripts but it should be suitable for other researchers using different test environments. Therefore an independent name space should be used for domain name resolution and the number of the threads should be a parameter.

The program should be as simple as possible for both minimizing the programming effort and making it easily understandable and modifiable by others. Therefore the program consists of a single file and takes only the essential input parameters.

Other researchers should be able to use, study its operation and modify the program as well as distribute their modified version. Therefore the test program is distributed under the GNU GPL v3 license [17].

B. Input Parameters

The program takes five input parameters:

1. The first one is the number of the executing client (stored in variable **a**; and is used in the domain name right after the beginning "10-" string).
2. The second one is the number of threads (stored in variable **n**). It MUST be a power of two (e.g. 1, 2, 4, 8, etc.), but it is not checked by the program.
3. The third one is the timeout value for the **recvfrom()** function measured in seconds (used when the test program waits for the reply of the DNS64 server).
4. The fourth one is the IPv6 address of the DNS64 server to be tested.
5. The fifth one is the port number at the server (where the DNS64 program listens). It was put to the last position because it is optional. If it is not specified then 53 is used as the default value.

C. Structure and Operation of the Test Program

The whole program contains three functions only:

- Function **main()** reads the input parameters, and executes a cycle for **b=0..255**. Each cycle contains an experiment. The execution time of each

experiment is measured (and printed out). In each experiment, **n** number of threads are started (the program code of the threads is the **dnstest()** function), and they each are to perform **N=256/n** number of DNS queries for AAAA records.

- Function **dnstest()** is executed as a thread. It takes six parameters: **a**, **b**, **c0**, **N**, the timeout value (as a **struct timeval * pointer**), the IPv6 address of the DNS64 server to be tested (in the appropriate format in a **struct addrinfo**) and the port number at the DNS64 server. It issues **N** number of AAAA queries for domain names beginning with the **10-a-b-{c0..c0+N-1}** labels.
- Function **dnsencode()** encodes the domain name with the special encoding introduced in subsection III.A.4.

D. Implementation Details

1) Name space for the experiments

An independent name space was chosen so that the program could be used anywhere by anyone. The program requests the AAAA records of the **10-a-b-c.dns64perf.test** domain names, where the values of the **a**, **b** and **c** variables are printed out in decimal format.

2) Time measurement

The standard **clock_gettime()** function is used for time measurements. It uses nanosecond resolution, however its accuracy is not necessarily 1 nanosecond. The result is printed out is milliseconds as the usual values of the execution time of one experiments is from a couple times ten to a couple time hundred milliseconds in a typical setup. The source code can be easily modified if a different (e.g. microsecond) resolution is preferred.

3) Preparation of the DNS queries

When a new DNS query is prepared, first, its whole memory area is initialized to 0x00 by the **memset()** function. One may decide to omit it as the majority of the area is overwritten with new values. However, then care must be taken to set some bytes to zero, as the current code assumes that the memory area has just been initialized to zero.

The unique value for the 16-bit *Transaction ID* is composed of the values of variables **a** and **b**. The next 16-bit field was set to ask standard recursive query. The number of questions was set to 1. The *QNAME* field of the *Question* section was encoded by the before mentioned **dnsencode()** function. Finally, the values of the *QTYPE* and *QCLASS* fields were set as we specified them in subsection III.A.3.

4) Compilation and linking

The test program was compiled and linked using the following command line:

```
g++ -std=c++11 -O4 dns64perf.cc -lrt -lpthread \
-o dns64perf
```

V. TESTING AND RESULTS

The performance of our new test program named *dns64perf* was compared to that of the original bash scripts using identical hardware and software environment.

A. Test Network

A very simple test network was used, see Fig. 1. Both the

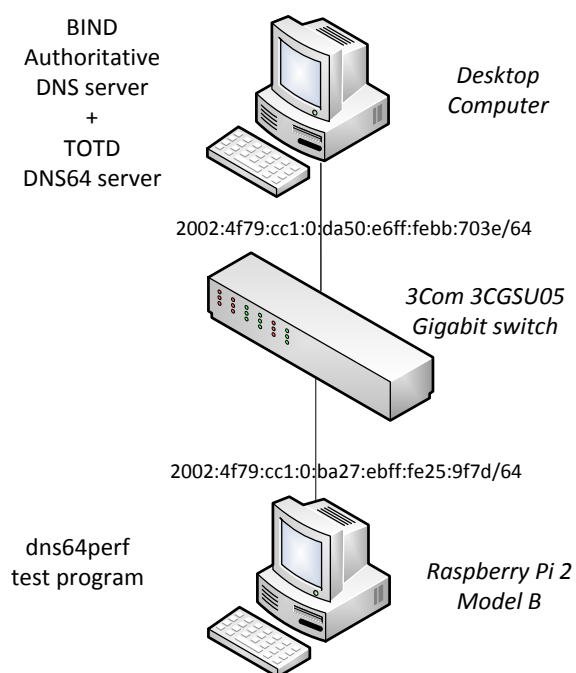


Fig. 1. DNS64 test network.

DNS64 server and authoritative DNS server were running on the same Intel desktop computer shown at the top of the figure. The test program was executed by the Raspberry Pi 2 single board computer (see at the bottom of the figure). The two computers were interconnected by a 3Com switch.

The devices had the following hardware and software configuration:

Desktop computer: 3200GHz Intel Core i5-4570 CPU (4 cores, 6MB L3 cache), 16GB 1600MHz DDR3 SDRAM, 250GB SSD; Debian GNU/Linux 7.8 operating system.

Raspberry Pi 2 Model B: 900MHz Broadcom BCM2836 CPU (4 cores, ARM Cortex-A7), 1GB RAM, 16GB microSD card (Class 10); Debian GNU/Linux 8.0 operating system.

switch: 3CGSU05 5-port 3Com Gigabit Ethernet switch

B. Testing Method

1) Authoritative DNS server

BIND 9.8.4 was used as authoritative DNS server. The zone file for the `dns64perf.test` zone was generated by the following script:

```
#!/bin/bash
cat > db.dns64perf.test << EOF

\${ORIGIN} dns64perf.test.
\${TTL} 86400
@      IN      SOA      localhost. root.localhost. (
                2015033101 ; Serial
                604800   ; Refresh
                86400    ; Retry
                2419200  ; Expire
                86400   ) ; Negative Cache TTL
;
@      IN      NS      localhost.

EOF

for a in {0..255}
do
    for b in {0..255}
```

```
do
    echo '$'GENERATE 0-255 10-$a-$b-$ IN A \
        10.$a.$b.$ >> db.dns64perf.test
done
echo "" >> db.dns64perf.test
```

Note that the `$GENERATE` directive is a shorthand. One line of this kind is equivalent to 256 traditional A record lines.

To tell BIND about the new zone, the `named.conf.local` configuration file of BIND was appended with the following lines:

```
zone "dns64perf.test" {
    type master;
    file "/etc/bind/db.dns64perf.test";
```

BIND was made to listen on port 1053 using the following line in the `named.conf.options` file:

```
listen-on port 1053 { 127.0.0.1; };
```

2) DNS64 server

TOTD 1.5.2 was used as DNS64 server. Its setup was very simple. TOTD was made known that the authoritative DNS server was listening on port 1053 of `localhost` and the prefix for the generation of IPv4 embedded IPv6 addresses was set to `dead:beef::/96` using the following two lines in the `totd.conf` file:

```
forwarder 127.0.0.1 port 1053
prefix dead:beef::
```

3) Measurement with the shell scripts

First, the `dns-st-c.sh` script was modified to use the new independent namespace:

```
#!/bin/bash
for c in {0..252..4} # that is 64 iterations
do
    host -t AAAA 10-$1-$2-$c.dns64perf.test &
    host -t AAAA 10-$1-$2-$((c+1)).dns64perf.test &
    host -t AAAA 10-$1-$2-$((c+2)).dns64perf.test &
    host -t AAAA 10-$1-$2-$((c+3)).dns64perf.test &
done
```

Then our TOTD DNS64 server was set as the primary DNS server of the Linux system. The following line was inserted as the first line of the `/etc/resolv.conf` file at the Raspberry Pi 2 computer:

```
nameserver 2002:4f79:cc1:0:da50:e6ff:febb:703e
```

After that, the DNS64 name resolution was working correctly. It was tested by the following command:

```
host -t AAAA 10-1-1-1.dns64perf.test
10-1-1-1.dns64perf.test has IPv6 address
dead:beef::a01:101
```

Finally, the bash shell scripts were executed. The results can be found in Table 1.

4) Measurement with dns64perf

The test program was executed as the “first” client ($a=1$), using all the 4 cores of the CPU and the timeout was set to 1 second using the following command line:

```
./dns64perf 1 4 1 2002:4f79:cc1:0:da50:e6ff:febb:703e
```


TABLE I
DNS64 PERFORMANCE TEST RESULTS

		bash shell		dns64perf	
		scripts	4 threads	1 thread	
Execution time of one experiment (s)	average	4.663	0.049	0.167	
	std. dev.	0.018	0.007	0.011	
	maximum	4.710	0.079	0.193	
Performance (requests/s)		55	5234	1535	

The program was also executed using only a single thread for the purpose of comparison.

C. Results and Evaluation

Table 1 shows the results of the tests. For all three tests, average and standard deviation of the execution times of the 256 experiments were calculated and also the maximum values were selected.

The performance (the number of AAAA record requests per second) was calculated according to (1).

$$N_{AAAA} = \frac{256}{t_E} \quad (1)$$

Where N_{AAAA} and t_E denote the number of AAAA record requests per second and the average execution time of an experiment (that is sending of 256 AAAA record requests), respectively.

The results show that dns64perf (using 4 threads) could perform an average of 5234 AAAA record queries per second, whereas the shell scripts could do only 55. This means that our programming efforts resulted in a 95 times speed-up.

It can also be observed that the standard deviation of the execution time was as low as 0.018s (which is less than 0.4% of the 4.663s average value) using the bash scripts. But it became relatively much higher using dns64perf: its 0.007s value is more than 14% of the 0.049s average. And also the maximum value of the execution time (0.079s) is now significantly higher than the average. We consider that it was caused by the way the threads are used: they always have to wait for the latest finishing one before the experiment can be completed. We checked our hypothesis with the results of the single thread execution. In this case, the 0.011s standard deviation value is only 6.6% of the 0.167s average. And also the maximum value of the execution time (0.193s) is much closer to the average.

D. Discussion of the Results

We have also checked the CPU utilization¹ of both the desktop computer and the Raspberry Pi 2 using the **top**² Linux command. These values were observed by human eyes only, and they were even fluctuating, therefore the following numbers are to be interpreted as order of magnitude estimations. They will be used for qualitative

¹ It was done not during the above measurements because it could have influenced the execution speed of the programs, but during repeated ones.

² This command gives both a summary of the CPU states (e.g. user, system, idle, wait, etc.) and a list of the most resource consuming processes displaying their percentage of CPU and memory consumption. However, the interpretation of 100 percent is different in the two cases. In the CPU summary, 100% means all the computing power of the existing CPU cores. In the process list, 100% means the computing power of a single core. Therefore we will always specify, how "100%" is to be interpreted.

analysis only, to find out the limits of the testing environment, that is: which component of the system limited the results?

When the bash shell scripts were used for testing, both BIND and TOTD used only about 1-2% of the computing power of a CPU core of the desktop computer, while the CPU idle time was only 5-6% (of the computing power of all 4 cores) at the Raspberry Pi 2 (that is, it was nearly fully utilized). Therefore it is clear that *the performance of the client limited the performance of the system*.

When dns64perf used 4 threads, TOTD used 90-95% of a CPU core of the desktop computer (BIND was under 40%, so it was not the bottleneck) while the Raspberry Pi 2 had more than 80% idle time. (The network could not be the bottleneck: DNS AAAA queries and answers were about 100-110 bytes and 210-220 bytes long, respectively. Let us consider the longer ones: about 5300 answers being each 220 bytes long would result in less than 10 Mbit per second whereas Raspberry Pi 2 has Fast Ethernet NIC.) Therefore, *the bottleneck was the single threaded TOTD DNS64 implementation*. We have also checked the CPU utilization value of TOTD when dns64perf used only one thread: TOTD used 45-47% of the computing power of a CPU core. (And it was 77-82% with dns64perf using 2 threads.) Thus we can clearly state that *dns64perf can perform significantly better!*

On the one hand it would be interesting to measure how many times dns64perf is faster than the method using bash shell scripts, but on the other hand the result would be useful for being proud of it only and nothing else. The performance ratio which has a real significance in DNS64 testing, is the following: *dns64perf executed by a single board computer can produce enough load to test the performance of a DNS64 server implementation executed by the CPU of a modern PC*. This result opens up the possibility of testing modern servers (e.g. with 16, 32 or even more cores) using dns64perf executed by a cluster of SBCs. This can be a very much cost effective solution.

VI. DIRECTIONS OF OUR FUTURE RESEARCH

A. Testing with an SBC cluster

Our next step is to build a 16 element cluster of ODRUID C1+ [18] SBCs for load generation. We plan to use dns64perf to test the same four DNS64 implementations which we tested in [8], but now we will be able to test their performance up to 16 cores (instead of up to 4 cores).

B. Plans for Benchmarking RFC

Benchmarking methodology for network interconnect devices was described in RFC 2544. It was expanded to address some IPv6 specificities in RFC 5180. However, it is stated there, that IPv6 transition mechanisms are outside the scope of the RFC.

Efforts were made to define benchmarking methodology for IPv6 transition technologies [19]. For more details of the measurements, see [20]. Marius Georgescu (the first author of [20]) has invited the author of this paper to take part in the further development of the draft RFC [19], to cover the testing methodology of DNS64 servers, too.

VII. CONCLUSION

We conclude that our efforts to make an efficient test program for the performance analysis of DNS64 servers were successful. Our new test program, dns64perf can perform at least 95 times more AAAA record queries per second than the old bash shell scripts could, but we have shown that its performance is even higher than that. We hope that dns64perf may be a useful testing tool for many researchers interested in the performance analysis of different DNS64 server implementations. The source code of the program is available under the GNU General Public License v3 from [21].

ACKNOWLEDGMENT

The author thanks Gábor Horváth, Dept. of Networked Systems and Services, Budapest University of Technology for lending the Raspberry Pi 2 single board computer.

REFERENCES

- [1] M. Bagnulo, A Sullivan, P. Matthews and I. Beijnum, "DNS64: DNS extensions for network address translation from IPv6 clients to IPv4 servers", IETF, April 2011. ISSN: 2070-1721 (RFC 6147)
- [2] M. Bagnulo, P. Matthews and I. Beijnum, "Stateful NAT64: Network address and protocol translation from IPv6 clients to IPv4 servers", IETF, April 2011. ISSN: 2070-1721 (RFC 6146)
- [3] K. J. O. Llanto and W. E. S. Yu, "Performance of NAT64 versus NAT44 in the context of IPv6 migration", in *Proc. International MultiConference of Engineers and Computer Scientists 2012 (IMECS 2012)*, Hong Kong, March 14-16, 2012, vol. I, pp. 638-645.
- [4] C. P. Monte et al, "Implementation and evaluation of protocols translating methods for IPv4 to IPv6 transition", *Journal of Computer Science & Technology*, ISSN: 1666-6038, vol. 12, no. 2, (August, 2012). pp. 64-70.
- [5] S. Yu, B. E. Carpenter, "Measuring IPv4 – IPv6 translation techniques", Technical Report 2012-001, Department of Computer Science, The University of Auckland, January 2012.
- [6] G. Lencse and S. Répás, "Performance analysis and comparison of different DNS64 implementations for Linux, OpenBSD and FreeBSD", in *Proc. IEEE 27th Internat. Conf. on Advanced Information Networking and Applications (AINA 2013)*, Barcelona, Spain, 2013, pp. 877–884. DOI: 10.1109/AINA.2013.80
- [7] G. Lencse and G. Takács, "Performance analysis of DNS64 and NAT64 solutions", *Infocommunications Journal*, vol. 4, no 2, pp. 29–36, Jun. 2012.
- [8] G. Lencse, S. Répás, "Performance analysis and comparison of four DNS64 implementations under different free operating systems", unpublished.
- [9] Nominum, "Measurement tools: DNSPerf and ResPerf downloads", <https://nominum.com/measurement-tools/>
- [10] die.net, "dnsperf(1) – Linux man page", <http://linux.die.net/man/1/dnsperf>
- [11] Nominum, "Sample query input file", <ftp://ftp.nominum.com/pub/nominum/dnsperf/data/queryfile-example-10million-201202.gz>
- [12] S. J. Cox, J. T. Cox, R. P. Boardman, S. J. Johnston, M. Scott, N. S. O'Brien, "Iridis-pi: a low-cost, compact demonstration cluster", *Cluster Computing*, vol. 17, no. 2, June 2014, pp. 349-358. DOI: 10.1007/s10586-013-0282-7
- [13] G. Lencse and S. Répás, "Benchmarking Single Board Computers for Building a Mini Supercomputer for Simulation", *38th Internat. Conf. Telecomm. and Signal Proc. (TSP 2015)*, Prague, Czech Republic, July 9-11, 2015, Brno University of Technology, pp. 246-251.
- [14] C. Bao, C. Huitema, M. Bagnulo, M Boucadair and X. Li, "IPv6 addressing of IPv4/IPv6 translators", IETF RFC 6052, 2010.
- [15] G. Lencse, A. G. Soós, "Design of a tiny multi-threaded DNS64 server", in *Proc. 38th Internat. Conf. Telecomm. and Signal Proc. (TSP 2015)*, Prague, Czech Republic, 2015, Brno University of Technology, pp. 27-32.
- [16] P. Mockapetris, "Domain names – implementation and specification", IETF, November 1987. (RFC 1035)
- [17] Free Software Foundation, *GNU General Public License*, Version 3, June 29, 2007, <http://www.gnu.org/licenses/gpl-3.0.en.html>
- [18] Hardkernel, *Odroid C1+*, (product description of the manufacturer), http://www.hardkernel.com/main/products/prdt_info.php?g_code=G143703355573
- [19] M. Georgescu, "Benchmarking Methodology for IPv6 Transition Technologies", IETF 93, Prague, Czech Republic, slides available: <https://www.ietf.org/proceedings/93/slides/slides-93-bmwg-6.pdf>
- [20] M. Georgescu, H. Hazeyama, Y. Kobayashi, S. Yamaguchi, "Empirical analysis of IPv6 transition technologies using the IPv6 Network Evaluation Testbed", *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol 2, no 2, e1, (2015), DOI: 10.4108/inis.2.2.e1
- [21] G. Lencse, "dns64perf source code", <http://ipv6.tilb.sze.hu/dns64perf/>



Gábor Lencse received his MSc in electrical engineering and computer systems at the Technical University of Budapest in 1994, and his PhD in 2001. He has been working for the Department of Telecommunications, Széchenyi István University in Győr since 1997. He teaches Computer networks, Computer architectures, IP-based telecommunication systems and the Linux operating system. Now, he is an Associate Professor. He is responsible for the specialization of the information and communication technology of the BSc level electrical engineering education. He is a founding member of the Multidisciplinary Doctoral School of Engineering Sciences, Széchenyi István University. The area of his research includes discrete-event simulation methodology, performance analysis of computer networks and IPv6 transition technologies. He has been working part time for the Department of Networked Systems and Services, Budapest University of Technology and Economics (the former Technical University of Budapest) since 2005. There he teaches Computer architectures and Computer networks. Dr. Lencse is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Non Minimum Phase Channel Estimation by Blind and Adaptive algorithms

Mohammed Zidane, Said Safi, Mohamed Sabri, Ahmed Boumezzough

Abstract—In this paper the problem of the Non Minimum Phase (NMP) channel identification using blind and adaptive algorithms is theoretically and numerically evaluated in noise environment case for different signal to noise ratio (SNR). For this problem, three blind algorithms based on Higher Order Cumulants (HOC) versus adaptive algorithms, i.e. Recursive Least Squares (RLS) and Least Mean Squares (LMS) are presented. In order to assess the performance of these approaches to identify the parameters of non minimum phase channels, we have selected the Macchi channel model. The simulation results in noisy environment and for different data input channel, provided to illustrate the performance of the blind approaches and compare them with adaptive algorithms.

Keywords—Higher Order Cumulants, channel identification, blind algorithms, adaptive algorithms, RLS, LMS.

I. INTRODUCTION

RECENTLY blind channel identification has drawn a great deal of attentions in the literature [1–10], [12–15]. In this work, identification of non minimum phase channel driven by a non gaussian in noisy environment is considered. The first methods are based on the autocorrelation function (second order cumulants) of the observed sequences. Thus, the autocorrelation sequence is insensitive to the phase characteristics of the system, and a non minimum phase system cannot be identified correctly using the second order statistics [2], their performances degrade when the output is noisy, because the second order cumulants for a gaussian process are not identically zero. To overcome these problems, other approaches were proposed and consist in using higher order cumulants (HOC) [1], this is because the gaussian noise will vanish in the higher order cumulants (3^{rd} and 4^{th} cumulants) domain, and the HOC preserve the phase characteristics, unlike the autocorrelation function. Further, the autocorrelation sequence fails to provide a complete statistical description for a non gaussian process, it was shown that consistent estimates of the parameters of any finite impulse response (FIR) system can be obtained by using higher order statistics or cumulants of the observed process [2–4]. Indeed, the HOS constitute a powerful tool in modelling stationary processes when the output signal of a system is corrupte with an additive gaussian noise. However, blind algorithm allows identification of non minimum phase channel with only output measurements, unlike the adaptive

algorithms exploiting the information of input and output for the impulse response channel estimation, for this reason the users periodically transmit a training sequence known by the receiver, which then estimates the impulse response channel parameters. Indeed, 20% of the bandwidth is devoted to training in GSM (up to 40% in UMTS). Blind methods are thus attractive so as to guarantee a high communication rate by eliminating (or reducing) the training sets [19]. In this paper, we address the problem of the blind and adaptive identification of non minimum phase channel. In order to evaluate the presented algorithms, we consider the Macchi channel model. The simulation results in noisy environment and for different data input channel, demonstrate that the blind and adaptive algorithms could estimate the phase and magnitude of Macchi channel with different precision.

The rest of this paper is organised as follows: In section II, the model and its assumptions, basic relationships are presented. In the following, blind and adaptive algorithms are presented to estimate the parameters of NMP channel. The simulations results showing the performance of the developed methods are presented in section VII. This paper is finally concluded in section VIII.

II. PROBLEM STATEMENT

A. Channel modeling

We consider non minimum phase channel (Fig. 1) described by the following diagram:

In noise free case:

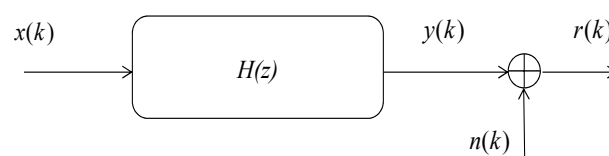


Fig. 1. Channel model

$$y(k) = \sum_{i=0}^q x(i)h(k-i) \quad (1)$$

In presence of noise:

$$r(k) = y(k) + n(k), \quad (2)$$

where $x(i)$ is the input sequence, $h(k)$ and q are the parameters and the order of linear channel, respectively, $y(k)$ represents the channel output in noiseless case and $r(k)$ is the observed channel output corrupted by additive gaussian

M. Zidane, M. Sabri are with Department of Physics, Faculty of Sciences and technology, Morocco. (e-mail: zidane.ilco@gmail.com; sipt03@yahoo.fr)

S. Safi is with the Department of Mathematics and Informatics, Polydisciplinary Faculty, Morocco. (e-mail: safi.said@gmail.com)

A. Boumezzough is with Department of Physics, Polydisciplinary Faculty, Morocco. (e-mail: ahmed.boumezzough@gmail.com)

Manuscript received August 18, 2015; revised October 4, 2015.

noise $n(k)$.

For this system we assume that:

- The input sequence, $x(i)$, is independent and identically distributed (*i.i.d*) zero mean, and non-gaussian;
- The system is supposed causal and truncated, i.e. $h(k) = 0$ for $k < 0$ and $k > q$, where $h(0) = 1$;
- The system order q is known;
- The measurement noise sequence $n(k)$ is assumed zero mean, *i.i.d*, gaussian and independent of $x(i)$ with unknown variance.

B. Basic Relationships

In this section, we present the general relationships for impulse response coefficients that constitute the basic relations in development of most linear HOC based methods proposed in the literature.

The m^{th} order cumulants of the $y(k)$ can be expressed as a function of impulse response coefficients $h(k)$ as follows [17]:

$$C_{m,y}(t_1, t_2, \dots, t_{m-1}) = \xi_{m,x} \sum_{i=0}^q h(i)h(i+t_1)\dots h(i+t_{m-1}), \quad (3)$$

where $\xi_{m,x}$ represents the m^{th} order cumulants of the excitation signal $x(i)$ at origin.

If we take $m = 2$ into Eq. (3) we obtain the second order cumulant (autocorrelation function):

$$C_{2,y}(t_1) = \xi_{2,x} \sum_{i=0}^q h(i)h(i+t_1) \quad (4)$$

For $m = 3$, Eq. (3) becomes:

$$C_{3,y}(t) = \xi_{3,x} \sum_{i=0}^q h(i)h(i+t_1)h(i+t_2) \quad (5)$$

Stogioglou and McLaughlin [18] presents the relationship between different n^{th} cumulants slices of the output signal $y(k)$, and the coefficients $h(k)$, are linked by the following relationship:

$$\begin{aligned} & \sum_{j=0}^q h(j) \left[\prod_{k=1}^r h(j+t_k) \right] C_{n,y}(\beta_1, \dots, \beta_r, j+\alpha_1, \dots, j+\alpha_{n-r-1}) \\ &= \sum_{i=0}^q h(i) \left[\prod_{k=1}^r h(i+\beta_k) \right] C_{n,y}(t_1, \dots, t_r, i+\alpha_1, \dots, i+\alpha_{n-r-1}), \end{aligned} \quad (6)$$

where $1 \leq r \leq n-2$.

If we take $n = 4$ into Eq. (6) we obtain the following equation:

$$\begin{aligned} & \sum_{j=0}^q h(j)h(j+t_1)h(j+t_2)C_{4,y}(\beta_1, \beta_2, j+\alpha_1) \\ &= \sum_{i=0}^q h(i)h(i+\beta_1)h(i+\beta_2)C_{4,y}(t_1, t_2, i+\alpha_1) \end{aligned} \quad (7)$$

III. SAFI, et al. ALGORITHM : ALGcum1

The Fourier transform of Eqs. (4) and (5) gives us the bispectra and the spectrum respectively [5]:

$$S_{3,y}(\omega_1, \omega_2) = \xi_{3,x} H(-\omega_1 - \omega_2) H(\omega_1) H(\omega_2) \quad (8)$$

$$S_{2,y}(\omega) = \xi_{2,x} H(-\omega) H(\omega) \quad (9)$$

If we suppose that $\omega = \omega_1 + \omega_2$, Eq. (9) becomes:

$$S_{2,y}(\omega_1 + \omega_2) = \xi_{2,x} H(-\omega_1 - \omega_2) H(\omega_1 + \omega_2) \quad (10)$$

Then, from Eqs. (8) and (10) we obtain the following equation:

$$S_{3,y}(\omega_1, \omega_2) H(\omega_1 + \omega_2) = \mu H(\omega_1) H(\omega_2) S_{2,y}(\omega_1 + \omega_2), \quad (11)$$

with $\mu = \frac{\xi_{3,x}}{\xi_{2,x}}$

The inverse Fourier transform of Eq. (11) demonstrates that the 3^{rd} order cumulants, the autocorrelation function and the impulse response channel parameters are combined by the following equation [5]:

$$\sum_{i=0}^q C_{3,y}(t_1-i, t_2-i) h(i) = \mu \sum_{i=0}^q h(i) h(t_2-t_1+i) C_{2,y}(t_1-i) \quad (12)$$

If we use the autocorrelation function property of the stationary process such as $C_{2,y}(t) \neq 0$ only for $-q \leq t \leq q$ and vanishes elsewhere if we take $t_1 = -q$, Eq. (12) takes the forme [5]:

$$\sum_{i=0}^q C_{3,y}(-q-i, t_2-i) h(i) = \mu h(0) h(t_2+q) C_{2,y}(-q), \quad (13)$$

else, if we suppose that $t_2 = -q$, Eq. (13) will become:

$$C_{3,y}(-q, -q) h(q) = \mu h(0) C_{2,y}(-q) \quad (14)$$

Using Eqs. (13) and (14) we obtain the following relation [5]:

$$\sum_{i=0}^q C_{3,y}(-q-i, t_2-i) h(i) = C_{3,y}(-q, -q) h(t_2+q) \quad (15)$$

The system of Eq. (15) can be written in matrix form as:

$$\begin{pmatrix} C_{3,y}(-q-1, -q-1) & \dots & C_{3,y}(-2q, -2q) \\ C_{3,y}(-q-1, -q) - \theta & \dots & C_{3,y}(-2q, -2q+1) \\ \vdots & \ddots & \vdots \\ C_{3,y}(-q-1, -1) & \dots & C_{3,y}(-2q, -q) - \theta \end{pmatrix} \times \begin{pmatrix} h(1) \\ \vdots \\ h(i) \\ \vdots \\ h(q) \end{pmatrix} = \begin{pmatrix} 0 \\ -C_{3,y}(-q, -q+1) \\ \vdots \\ -C_{3,y}(-q, 0) \end{pmatrix}, \quad (16)$$

where $\theta = C_{3,y}(-q, -q)$.

Or in more compact form, the Eq. (16) can be written as follows:

$$M h_q = d, \quad (17)$$

where M is the matrix of size $(q+1) \times (q)$ elements, h_q is a column vector constituted by the unknown impulse response parameters $h(k)$ for $k = 1, \dots, q$ and d is a column vector of size $(q+1)$ as indicated in the Eq. (17).

The least squares solution of the system of Eq. (17), permits blindly identification of the parameters $h(k)$ and without any information of the input selective channel. Thus, the solution will be written under the following form:

$$\hat{h}_q = (M^T M)^{-1} M^T d \quad (18)$$

IV. ANTARI, *et al.* ALGORITHM : *ALGcum2*

Antari, *et al.* [9] demonstrates that the coefficients $h(k)$ for an finite impulse response system can be obtained by the Eq. (12), if we take $t_2 - t_1 = q$ and the considered system is causal. The following equation describes (*ALGcum2*) algorithm:

$$\sum_{i=0}^q C_{3,y}(t_1 - i, t_1 + q - i) h(i) = \frac{\xi_{3,x}}{\xi_{2,x}} h(0) h(q) C_{2,y}(t_1), \quad (19)$$

where $t_1 = -q, \dots, q$ the Eq. (19) can be written as follows:

$$M h_e = d, \quad (20)$$

where M is the matrix of size $(2q+1) \times (q+1)$ elements, h_e is a column vector constituted by the unknown impulse response parameters $h(k) : k = 0, \dots, q$ and d is a column vector of size $(2q+1)$ as indicated in the Eq. (20). The least squares solution of the system of Eq. (20), will be written under the following form:

$$\hat{h}_e = (M^T M)^{-1} M^T d \quad (21)$$

V. ABDERRAHIM, *et al.* ALGORITHM : *ALGcum3*

Abderrahim, *et al.* [10] use the relationship (7), based on fourth order cumulants, with $\beta_1 = \beta_2 = 0$, $t_1 = q$ and $t_2 = 0$. This algorithm is given by the following relation:

$$\sum_{i=0}^q h^3(i) C_{4,y}(q, 0, i + \alpha_1) - h(q) C_{4,y}(0, 0, \alpha_1) = -C_{4,y}(q, 0, \alpha_1), \quad (22)$$

where $\alpha_1 = -q, \dots, q$. In more compact form, the system of Eq. (22) can be written in the following form:

$$M\theta = A \quad (23)$$

$\theta = [h(q) \quad h^3(1) \quad h^3(2) \dots h^3(q)]^T$ is a column vector of size $(q+1)$;

$A = [0 \dots 0 \quad -C_{4y}(q, 0, 0) \quad -C_{4y}(q, 0, 1) \dots -C_{4y}(q, 0, q)]^T$ is a vector of size $(2q+1)$;

The least squares solution of the system of Eq. (23), will be written under the following form:

$$\hat{\theta} = (M^T M)^{-1} M^T d \quad (24)$$

The parameters $h(i)$ for $k = 1, \dots, q$ are estimated from the estimated values $\hat{\theta}(i)$ using the following equation:

$$\hat{h}(i) = \sqrt[3]{\hat{\theta}(i+1)} \quad (25)$$

VI. ADAPTIVE ALGORITHMS

A. Description of the LMS Algorithm

From the method of steepest descent, the weight vector equation is given by [20]:

$$\hat{H}_{n+1} = \hat{H}_n - \frac{1}{2} \mu \nabla \bar{\varepsilon}(H)|_{\hat{H}_n}, \quad (26)$$

with $e(n) = r(n) - X^T(n)H$, $\bar{\varepsilon} = E[|e(n)|^2]$ and $H = [h(0), h(1), \dots, h(q)]^T$. μ is the step-size parameter and controls the convergence characteristics of the LMS algorithm including between 0 and 1.

The gradient vector in the above weight update equation can be computed as:

$$\nabla \bar{\varepsilon}(H)|_{\hat{H}_n} \approx \nabla |e(n)|^2 = \frac{\partial |e(n)|^2}{\partial H} |_{\hat{H}_n} \quad (27)$$

$$\nabla |e(n)|^2 = -2e(n)X(n) \quad (28)$$

The LMS algorithm is initiated with an arbitrary value $h(0)$ for the weight vector at $n = 0$. The successive corrections of the weight vector eventually leads to the minimum value of the mean squared error. Therefore the LMS algorithm can be summarized in the following equations:

$$\hat{H}(n+1) = \hat{H}(n) + \mu \hat{e}(n)X(n), \quad (29)$$

with

$$\hat{e}(n) = r(n) - \hat{H}_n X^T(n) \quad (30)$$

B. Description of the RLS Algorithm

The RLS algorithm [12] is described by the following equations (with initial conditions $\hat{H}(0) = 0$).

$Q^{-1}(0) = \delta^{-1}I$, δ is a small positive constant value.

$$k(n) = \frac{\lambda^{-1}Q^{-1}(n-1)X(n)}{1 + \lambda^{-1}X^T(n)Q^{-1}(n-1)X(n)}, \quad (31)$$

$$e(n) = r(n) - X^T(n)\hat{H}(n-1), \quad (32)$$

$$\hat{H}(n) = \hat{H}(n-1) + k(n)e(n), \quad (33)$$

$$Q^{-1}(n) = \lambda^{-1}Q^{-1}(n-1) - \lambda^{-1}k(n)X^T(n)Q^{-1}(n-1). \quad (34)$$

VII. SIMULATION RESULTS

To verify the performance of these algorithms, we have applied a Macchi channel model. To measure the strength of noise, we define the signal to noise ratio (*SNR*) as:

$$SNR = 10 \log \left[\frac{\sigma_y^2(k)}{\sigma_n^2(k)} \right] \quad (35)$$

To measure the accuracy of parameter estimation with respect to the real values, we define the mean square error (*MSE*) for each run as

$$MSE = \frac{1}{q} \sum_{i=0}^q \left[\frac{h(i) - \hat{h}(i)}{h(i)} \right]^2, \quad (36)$$

where $\hat{h}(i)$, $i = 1, \dots, q$ are the estimated parameters in each run, and $h(i)$, $i = 1, \dots, q$ are the real parameters in the model.

1) *Macchi channel*: The Macchi channel is defined by the following equation:

$$\begin{cases} y(k) = 0.8264x(k) - 0.1653x(k-1) + 0.8512x(k-2) \\ \quad + 0.1636x(k-3) + 0.8100x(k-4), \\ \text{zeros: } z_1=0.5500 + 0.9526j; z_2=0.5500 - 0.9526j; \\ z_3=-0.4500 + 0.7794j; z_4=-0.4500 - 0.7794j. \end{cases} \quad (37)$$

The Macchi channel is a non minimum phase because two of its zeros are outside of the unit circle.

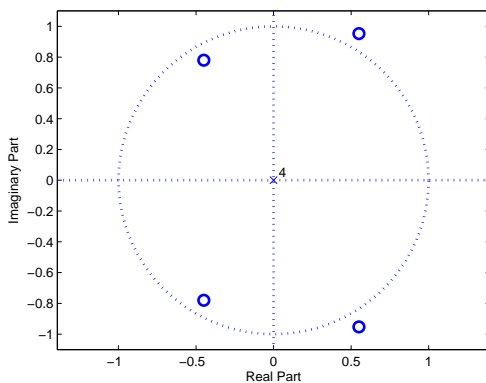


Fig. 2. The zeros of Macchi channel

2) *Impulse response parameters estimation*: The results of simulation are shown in the Table I for different values of sample sizes and for SNR=8 dB.

For all sample sizes and for a SNR=8 dB, the values, of mean square error (*MSE*), of the (*ALGcum1*), *LMS* and *RLS* algorithms are small than those obtained by the (*ALGcum2*) and (*ALGcum3*) algorithms, this implies the true parameters are near the estimates parameters.

The Fig. 3 give us a good idea about the precision of these algorithms.

In the Table II we well present the impulse response parameters estimation of the Macchi's channel for different *SNR* and data length $N = 2048$.

From the Table II we can conclude that: the *RLS*, *LMS* and (*ALGcum1*) algorithms are more effective than (*ALGcum2*) and (*ALGcum3*) algorithms in noisy environment, with an advantage of the *RLS* algorithm in terms of *MSE*. But the results given by the algorithms based on higher order cumulants, are blindly without any information about the input, comparing to adaptive algorithms exploiting the information of input and output for the estimation of the impulse response channel. For that the adaptive methods send occasionally a training sequence known to the transmitter and receiver.

The Fig. 4 give us a good idea about the precision of these algorithms for different *SNR*.

3) *Magnitude and phase estimation*: In the following figure (Fig. 5) we have presented the estimation of the magnitude and phase of the impulse response channel using the presented

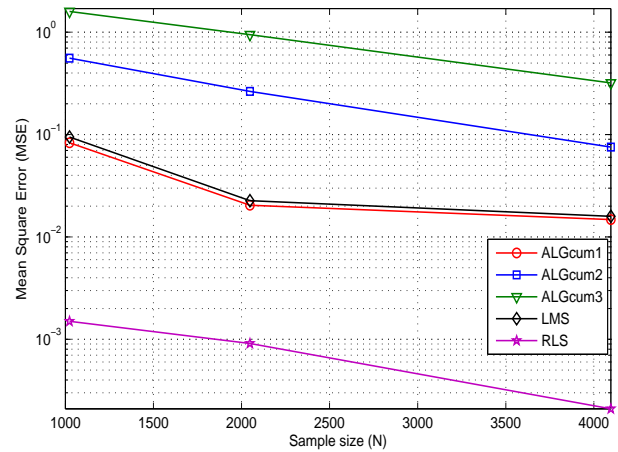


Fig. 3. Comparison of algorithms for Macchi channel for SNR=8 dB

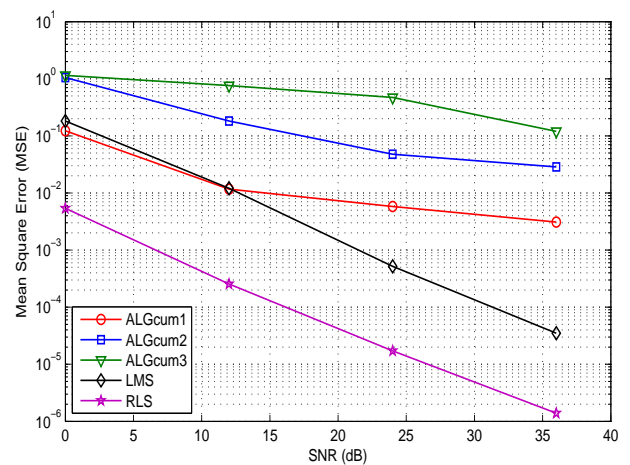


Fig. 4. Comparison of algorithms for Macchi channel for $N = 2048$

algorithms for $N = 2048$ and SNR=8 dB.

From the Fig. 5 we observe that the estimated magnitude and

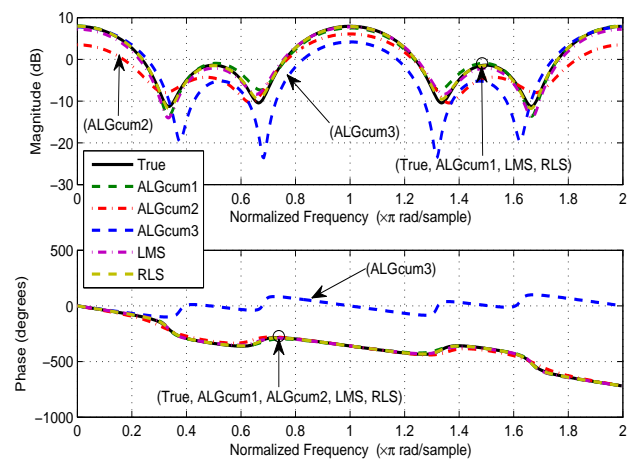


Fig. 5. Estimated magnitude and phase of the Macchi channel impulse response, for different algorithms, $N = 2048$ and SNR =8 dB.

TABLE I
Estimated parameters of Macchi channel model excited by different sample sizes and for SNR =8 dB

N	$\hat{h}(i) \pm \sigma$	$ALGcum1$	LMS	$ALGcum2$	$ALGcum3$	RLS
1024	$\hat{h}(1) \pm \sigma$	0.7896±0.1338	0.8456±0.0629	0.5607±0.2748	0.5101±0.5314	0.8114±0.0070
	$\hat{h}(2) \pm \sigma$	-0.0559±0.2056	-0.2734±0.1140	-0.4244±0.1628	0.2557±0.4035	-0.1588±0.0071
	$\hat{h}(3) \pm \sigma$	0.7857±0.1099	0.8075±0.0874	0.5550±0.3079	0.5790±0.6220	0.8536±0.0060
	$\hat{h}(4) \pm \sigma$	0.1892±0.1053	0.1886±0.0901	0.0302±0.1063	-0.1049±0.5160	0.1759±0.0079
	$\hat{h}(5) \pm \sigma$	0.6740±0.1408	0.6982±0.1735	0.8061±0.2618	0.4907±0.9437	0.8076±0.0069
	MSE	0.0831	0.0946	0.5577	1.5976	0.0015
2048	$\hat{h}(1) \pm \sigma$	0.8797±0.1168	0.8233±0.0662	0.5137±0.1847	0.6921±0.3818	0.8351±0.0048
	$\hat{h}(2) \pm \sigma$	-0.1506±0.1482	-0.2039±0.0512	-0.3257±0.0969	0.2236±0.3769	-0.1643±0.0042
	$\hat{h}(3) \pm \sigma$	0.8154±0.1087	0.8306±0.0510	0.6500±0.2366	0.7497±0.4089	0.8565±0.0067
	$\hat{h}(4) \pm \sigma$	0.2161±0.0830	0.1290±0.0557	0.0642±0.1143	0.1903±0.4751	0.1737±0.0054
	$\hat{h}(5) \pm \sigma$	0.7499±0.1224	0.7181±0.0506	0.5890±0.2197	0.5895±0.4283	0.7905±0.0063
	MSE	0.0204	0.0226	0.2641	0.9459	9.1098 × 10 ⁻⁴
4096	$\hat{h}(1) \pm \sigma$	0.7871±0.0805	0.8185±0.0297	0.5806±0.1338	0.7450±0.1372	0.8287±0.0050
	$\hat{h}(2) \pm \sigma$	-0.1977±0.1038	-0.1198±0.0236	-0.2438±0.0730	0.0105±0.3597	-0.1625±0.0041
	$\hat{h}(3) \pm \sigma$	0.8552±0.0740	0.8569±0.0277	0.6203±0.1822	0.7980±0.1381	0.8526±0.0041
	$\hat{h}(4) \pm \sigma$	0.1281±0.0526	0.1601±0.0304	0.1821±0.1147	0.0299±0.5077	0.1681±0.0050
	$\hat{h}(5) \pm \sigma$	0.8335±0.0877	0.8568±0.0296	0.6233±0.1701	0.5475±0.2014	0.8106±0.0043
	MSE	0.0148	0.0159	0.0756	0.3197	2.0973 × 10 ⁻⁴
	True parameters $h(i)$	$h(1) = 0.8264$	$h(2) = -0.1653$	$h(3) = 0.8512$	$h(4) = 0.1636$	$h(5) = 0.8100$

TABLE II
Estimated parameters of Macchi channel model for different SNR and excited by sample sizes $N = 2048$

SNR	$\hat{h}(i) \pm \sigma$	$ALGcum1$	LMS	$ALGcum2$	$ALGcum3$	RLS
0 dB	$\hat{h}(1) \pm \sigma$	0.8528±0.1256	0.8309±0.0540	-0.1310±0.9704	-0.0146±0.5764	0.8446±0.0127
	$\hat{h}(2) \pm \sigma$	-0.1157±0.1140	-0.2864±0.0839	-0.1543±0.5630	0.0316±0.4848	-0.1660±0.0148
	$\hat{h}(3) \pm \sigma$	0.7554±0.1463	0.8615±0.0602	0.1448±0.8429	0.2827±0.6526	0.8624±0.0152
	$\hat{h}(4) \pm \sigma$	0.2935±0.0777	0.0700±0.0455	-0.1602±0.2333	-0.1418±0.4816	0.1882±0.0147
	$\hat{h}(5) \pm \sigma$	0.7634±0.1416	0.6345±0.0537	0.3275±0.6321	0.2376±0.5866	0.7607±0.0125
	MSE	0.1229	0.1823	1.0514	1.1477	0.0054
12 dB	$\hat{h}(1) \pm \sigma$	0.8752±0.1081	0.8294±0.0167	0.5806±0.1957	0.7387±0.1427	0.8317±0.0037
	$\hat{h}(2) \pm \sigma$	-0.1938±0.1234	-0.1940±0.0128	-0.3110±0.1091	0.0957±0.3542	-0.1656±0.0036
	$\hat{h}(3) \pm \sigma$	0.8311±0.1096	0.8412±0.0147	0.6797±0.1813	0.8193±0.1271	0.8548±0.0044
	$\hat{h}(4) \pm \sigma$	0.1942±0.0583	0.1363±0.0115	0.1062±0.1204	-0.0655±0.5145	0.1687±0.0048
	$\hat{h}(5) \pm \sigma$	0.7810±0.1101	0.7721±0.0159	0.6071±0.1848	0.5510±0.2510	0.7975±0.0036
	MSE	0.0117	0.0121	0.1819	0.7615	2.5485 × 10 ⁻⁴
24 dB	$\hat{h}(1) \pm \sigma$	0.8534±0.0920	0.8256±0.0037	0.7273±0.1873	0.8027±0.1540	0.8277±0.0008
	$\hat{h}(2) \pm \sigma$	-0.1814±0.1497	-0.1703±0.0038	-0.2460±0.0979	0.0862±0.4024	-0.1654±0.0011
	$\hat{h}(3) \pm \sigma$	0.8767±0.0989	0.8517±0.0047	0.8098±0.1859	0.8242±0.1691	0.8519±0.0008
	$\hat{h}(4) \pm \sigma$	0.1400±0.0942	0.1572±0.0039	0.1914±0.1294	0.0515±0.5461	0.1649±0.0008
	$\hat{h}(5) \pm \sigma$	0.8486±0.1029	0.7994±0.0038	0.7681±0.2359	0.6482±0.3391	0.8067±0.0009
	MSE	58 × 10 ⁻⁴	5.1989 × 10 ⁻⁴	0.0478	0.4710	1.7192 × 10 ⁻⁵
36 dB	$\hat{h}(1) \pm \sigma$	0.8564±0.0982	0.8267±0.0012	0.7103±0.1974	0.8383±0.1537	0.8268±0.0002
	$\hat{h}(2) \pm \sigma$	-0.1728±0.1327	-0.1667±0.0008	-0.2233±0.0886	-0.0428±0.3878	-0.1653±0.0003
	$\hat{h}(3) \pm \sigma$	0.8887±0.0986	0.8509±0.0010	0.7805±0.2284	0.8728±0.1253	0.8514±0.0003
	$\hat{h}(4) \pm \sigma$	0.1552±0.0733	0.1620±0.0009	0.1872±0.1490	0.2287±0.5115	0.1640±0.0002
	$\hat{h}(5) \pm \sigma$	0.8940±0.0958	0.8079±0.0009	0.7859±0.2864	0.7197±0.3714	0.8092±0.0002
	MSE	31 × 10 ⁻⁴	3.5059 × 10 ⁻⁵	0.0286	0.1201	1.3921 × 10 ⁻⁶
	True parameters $h(i)$	$h(1) = 0.8264$	$h(2) = -0.1653$	$h(3) = 0.8512$	$h(4) = 0.1636$	$h(5) = 0.8100$

phase using ($ALGcum1$), LMS and RLS algorithms have the same allure and we have not more difference between the estimated and the true ones. The ($ALGcum2$) algorithm is able to estimate the phase response, but we have more difference between the magnitude estimated and the true ones, of the Macchi channel. The results given by ($ALGcum3$) algorithm are more different to the true ones.

VIII. CONCLUSION

In this paper, we have compared blind identification methods based on HOC, with the adaptive algorithms (RLS and LMS), to identify the parameters of the impulse response

of the frequency selective channel such as the Macchi channel. The simulation results show the efficiency of the RLS , LMS and ($ALGcum1$) algorithms, but the blind algorithm present the advantage to estimate the impulse response of the frequency selective channel blindly, comparing to adaptive algorithms. The magnitude and phase of the impulse response are estimated, using RLS , LMS and ($ALGcum1$) algorithms with a very good precision in noisy environment, but the ($ALGcum2$) and ($ALGcum3$) give us a significant difference between the estimated and the true ones.

REFERENCES

- [1] K. Abederrahim, H. Mathlouthi and F. Msahli, "FIR System Identification using Higher Order Statistics," American Control Conference Hyatt Regency Riverfront, St. Louis, MO, USA June 10–12, pp. 5598–5603, 2009.
- [2] L. Srinivas and K. V. S. Hari, "FIR System Identification Based on Subspaces of a High Order Cumulant Matrix," IEEE Transactions on signal processing, vol. 44 no. 6, pp. 1485–1491, 1996.
- [3] M. Bakrim and D. Aboutajdine, "Cumulant based identification of non gaussian moving average signals," Traitement du Signal, vol. 16 no. 3 pp. 175–186, 1999.
- [4] C.-Y. Chi and J.-Y. Kung, "A new identification algorithm for all pass systems by higher order statistics," Signal Processing vol. 41, pp. 239–256, 1995.
- [5] S. Safi, M. Frikel, M. M'Saad, and A. Zeroual, "Blind Impulse Response Identification of frequency Radio Channels : Application to Bran A Channel," Int. J. Sig. Proces, vol. 4 no. 1 pp. 201–206, 2007.
- [6] S. Safi and A. Zeroual, "Blind non minimum phase channel identification using 3rd and 4th order cumulants," Int. J. Sig. Proces., vol. 4 no. 1, pp. 158–168, 2008.
- [7] M. Zidane, S. Safi, M. Sabri and A. Boumezzough, "Impulse Response Identification of Minimum and Non Minimum Phase Channels," 4th Workshop on Codes, Cryptography and Communication Systems (WC-CCS'13) 07-08 November 2013, Meknes, Morocco.
- [8] J. Antari, A. El Khadimi, D. Mammas and A. Zeroual, "Developed Algorithm for Supervising Identification of Non Linear Systems using Higher Order Statistics: Modeling Internet Traffic," International Journal of Future Generation Communication and Networking, vol. 5 no. 4, pp. 17–28, 2012.
- [9] J. Antari, A. Zeroual and S. Safi, "Stochastic analysis and parametric identification of moving average (MA) non Gaussian signal using cumulants," International Journal of Physical and Chemical News, vol. 34, pp. 27–32, 2007.
- [10] K. Abderrahim, R. B. Abdenmour, F. Msahli, M. Ksouri, and G. Favier, "Identification of non minimum phase finite impulse response systems using the fourth order cumulants," Progress in system and robot analysis and control design, Springer, vol. 243, pp. 41–50, 1999.
- [11] K. Abderrahim, R. B. Abdenmour, G. favier, M. Ksouri and F. Msahli, "New results on FIR system identification using cumulants," APII–JESA., vol. 35 no 5, pp. 601–622, 2001.
- [12] A. Y. Kibangou, "Modèles de volterra à complexité reduite : estimation parametrique et application à l'égalisation des canaux de communication," Thèse de doctorat, Universit de Nice-Sophia Antipolis, 2005.
- [13] M. Zidane, S. Safi, M. Sabri, and A. Boumezzough, "Blind Identification of Minimum Phase Channels Based On Higher Order Cumulants," International Conference on Intelligent Information and Network Technology (IC2INT'13) 13–14 November 2013, Settat, Morocco.
- [14] S. Safi, A. Zeroual and M. M. Hassani, "Prediction of global daily solar radiation using higher order statistics," Renewable Energy, vol. 27, pp. 647–666, 2002.
- [15] S. Safi and A. Zeroual, "Blind identification in noisy environment of non minimum phase Finite Impulse Response (FIR) using higher order statistics," Int. J. Sys. Anal. Modell. Simul., Taylor Francis, vol. 43 no. 5, pp. 671–681, 2003.
- [16] J. Antari, R. Iqdour and A. Zeroual, "Forecasting the wind speed process using higher order statistics and fuzzy systems," Renewable Energy, vol. 9, 237–251, 2006.
- [17] D. Brillinger and M. Rosenblatt, "Computation and interpretation of k^{th} order spectra," In Spectral Analysis of Time Signals, NewYork: Wiley, pp. 907–938, 1967.
- [18] A. G. Stogioglou and S. McLaughlin, "MA parameter estimation and cumulant enhancement," IEEE Transactions on Signal Processing, vol. 44 no. 7, pp. 1704–1718, 1996.
- [19] D. Nion and L. De Lathauwer, "A Block Component Model Based Blind DS-CDMA Receiver," IEEE Transactions on signal processing, vol. 56 no. 11, pp. 5567–5579, 2008.
- [20] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm : analysis and simulations," IEEE Transactions on Signal Processing, vol. 45 no. 3, pp. 631–639, 1997.

Design and FPGA Implementation of Variable Cutoff Frequency Filter based on Continuously Variable Fractional Delay Structure and Interpolation Technique

Sumedh Dhabu and A. P. Vinod

Abstract—This paper presents the design and FPGA implementation of interpolated continuously variable fractional delay structure based filter (ICVFD filter) with fine control over the cutoff frequency. In the ICVFD filter, each unit delay of the prototype lowpass filter is replaced by a continuously variable fractional delay (CVFD) element proposed in this paper. The CVFD element requires the same number of multiplications as that of the second-order fractional delay structure used in the existing fractional delay structure based variable filter (FDS based filter), however it provides fractional delays corresponding to the higher-order fractional delay structures. Hence, the proposed ICVFD filter provides wider cutoff frequency range compared to the FDS based filter. The ICVFD filter is also capable of providing variable bandpass and highpass responses. We use two-stage approach for the FPGA implementation of the ICVFD filter. First, we use pipelining stages to shorten the critical path and improve the operating frequency. Then, we make use of specific hardware resource, i.e. RAM-based Shift Register (SRL) to further improve the operating frequency and resource usage.

Keywords—FPGA implementation, fractional delay structure based filter, reconfigurable digital filter, variable cutoff frequency filter.

I. INTRODUCTION

Variable finite impulse response (FIR) filters (FIR filters whose frequency response can be changed based on the desired specifications) are widely used in digital communications. The frequency response of an FIR filter can be changed by completely changing its coefficients or by modifying the impulse response using various operations. In the programmable digital filters [1]-[4], the desired frequency responses are obtained by updating all the filter coefficients which are stored in the memory. This is a very simple approach, and in general, the variable coefficient filters are optimum in a sense that the filter length for the particular frequency response specifications is the minimum. However, when the frequency response of the filter needs to be changed frequently, large number of memory access operations make updating routines of these filters time consuming. Other approaches proposed in the literature [5]-[12] modify the impulse response of the fixed-coefficient

prototype filter by controlling fewer parameters, without the need of updating all the filter coefficients.

In the interpolation approach [5], each delay of the fixed-coefficient filter structure is replaced by M delays to obtain a multiband response and then the desired band is extracted using a masking filter. In the coefficient decimation method (CDM) [6], the impulse response of the fixed-coefficient filter is modified by retaining every D^{th} coefficient of the filter and either replacing the remaining coefficients by zeros or by completely discarding them. The cutoff frequency of the coefficient decimated filter can be an integer multiple of the cutoff frequency of the prototype filter. Even though the interpolation and CDM techniques are simple to implement (as they need only multiplexers to vary M or D) and the filters realized using [5] and [6] have low complexities, they provide only coarse control over the cutoff frequency due to the discrete nature of the controlling parameters (M and D).

A very fine control over the cutoff frequency of the filter can be obtained at the cost of increase in the complexity of the filter structure. In [7], an all-pass transformation based variable filter is realized by replacing the unit delay of the prototype filter by the first- or second-order all-pass structure. Even though the prototype filter in [7] is a linear-phase filter, the resultant filter is not a linear-phase filter due to the use of all-pass transformation. As opposed to the all-pass transformations, the frequency transformation based filters preserve the linear-phase property of the prototype filter [8]. However, the transition bandwidth of the frequency transformation based filter can be significantly wider than that of the prototype filter. The spectral parameter approximation (SPA) technique [9], [10] makes use of weighted combination of the fixed-coefficient FIR sub-filters to generate the desired frequency response and provides absolute control over the cutoff frequency of the filter in the desired range. However, the complexity of the SPA technique is higher than all the other approaches. In [11], a set of fixed-coefficient filters is used where each filter takes care of only specific part of the variable frequency regions. This technique requires large number of filters when the desired cutoff frequency range is large.

In [12], the unit delay element of the filter is replaced by the fractional delay structure (FDS). In this FDS based filter [12], a single parameter (d) varies the value of the fractional delay; thereby modifying the sample values and the length of the impulse response of the filter, resulting in a variable

digital filter with fine control over the cutoff frequency. A second-order modified Farrow structure [13] is used to replace the unit delay of the prototype filter in [12]. Therefore, cutoff frequency of the filter varies according to the value of the fractional delay ($1 \leq 1+d < 2$). However, the cutoff frequency, f_c , can be varied only in the limited range given by $f_{c_mod}/2 < f_c \leq f_{c_mod} \leq 0.2$ where f_{c_mod} is the cutoff frequency of the prototype filter. (Please note that all the frequency values mentioned in this paper are normalized with respect to half the sampling frequency, i.e. π .) It is observed that the FDS based filter provides unity magnitude response and constant phase response only for the cutoff frequencies in the lower range of the Nyquist band [12]. The second-order modified Farrow structure can provide the unity magnitude response and constant phase response only for the low frequencies (approximately up to 0.2) [13, 14], which results in degradation in the response of the FDS based filter for higher cutoff frequencies. Therefore, the maximum cutoff frequency obtained from the FDS based filter can be approximately 0.2.

In [12], CDM is used to increase the cutoff frequency range of the FDS based filter. Therefore, the cutoff frequency range can be $f_{c_mod}/2 < f_c \leq 2f_{c_mod} \leq 0.2$. However, the prototype filter needs to be overdesigned, i.e. its order should be increased, in order to compensate for the passband ripple, stopband attenuation and transition bandwidth degradation which is inherent to the CDM.

In this paper we present the design and FPGA implementation of modified fractional delay structure based filter to overcome the lower limit on the cutoff frequency of the FDS based filter. The proposed interpolated continuously variable fractional delay structure based filter (ICVFD filter) uses the continuously variable fractional delay (CVFD) structure. This CVFD element provides wider delay range, equivalent to the delay range obtained from the higher-order fractional delay structure, without increasing the number of multiplications required. The ICVFD filter uses the CVFD element and the interpolation technique and provides a *continuous* control over the cutoff frequency of the filter. The ICVFD filter is capable of producing variable lowpass, bandpass and highpass filter responses.

The rest of the paper is organized as follows. Section II presents the details of the CVFD element and the ICVFD filter. A design example and comparison of the ICVFD filter with the existing variable filters is also presented in Section II. Section III presents the details of pipelining and use of specific hardware resources for implementing the ICVFD filter. The FPGA implementation results are presented in Section IV. Finally Section V concludes this paper.

II. PROPOSED ICVFD FILTER

A. CVFD Element

The CVFD element provides the delay,

$$D_p = p + 1 + d \quad (1)$$

where the fractional delay equal to $1+d$ is provided by the second-order modified Farrow structure, and the variable number of p unit delays are added using a multiplexer, as shown in Fig. 1. The fractional delay range of this CVFD element can be changed online as follows. For $p = 0$, the

fractional delay range is $1 \leq D_p < 2$, whereas the fractional delay range changes to $2 \leq D_p < 3$ for $p = 1$, and so on. For the other fractional delay structures, the multiplication complexity of the fractional delay structure (i.e. the number of multiplications required) increases with the range of fractional delay to be obtained [13-15]. However, the CVFD element provides the fractional delay same as the higher-order fractional delay structure, at the same multiplication complexity as that of the second-order fractional delay structure. Further, as the second-order modified Farrow structure has the least multiplication complexity among the second-order fractional delay structures, the proposed CVFD element is capable of providing the fractional delay equivalent to the fractional delay provided by higher-order fractional delay structures for the least multiplication complexity possible.

Another advantage of the CVFD element is that the fractional delay range can be changed online. For the modified Farrow structure based fractional delay structures [13], only the value of fractional delay can be changed online, and not the fractional delay range. The fractional delay range depends on the order of the structure, and therefore, the structures of different orders are required to change the fractional delay range. The second-order fractional delay structures can provide the fractional delay range of 1 to 2 only, and the third-order fractional delay structures can provide the delay range of 2 to 3 only. Hence, these are not suitable when the fractional delay range needs to be changed in the FDS based filter.

Similar to the proposed CVFD element, the fractional delay structure proposed in [15] is capable of changing the fractional delay range on-the-fly. However, as mentioned previously, multiplication complexity of the CVFD element is less than that of the fractional delay structure in [15]. It may be possible to use fractional delay structures based on other implementation strategies [14] which can change the fractional delay range online, but the multiplication complexity of such structures is higher, and therefore the multiplication complexity of the FDS based filter increases. In the CVFD element, the fractional delay range can be changed online without any additional multiplication complexity for the FDS based filter.

B. ICVFD Filter

The proposed ICVFD filter is the combination of the FDS based filter in which the CVFD element replaces the unit delay of the filter and the interpolation technique as shown in Fig. 2. Note that the input signal is fed to the filter and not to the CVFD element. The CVFD element is just used to replace the unit delay of the filter. The interpolation of the CVFD element by factor M results in a delay, D_c , given by

$$D_c = p + (1+d) \times M \quad (2)$$

Therefore, the cutoff frequency and the transition bandwidth of the ICVFD filter, f_{c_ICVFD} and tbw_{ICVFD} respectively, are given by

$$f_{c_ICVFD} = f_{c_mod} / D_c \quad (3)$$

$$tbw_{ICVFD} = tbw_{mod} / D_c \quad (4)$$

where f_{c_mod} is the cutoff frequency of the prototype (modal) filter and tbw_{mod} is the transition bandwidth of the prototype filter.

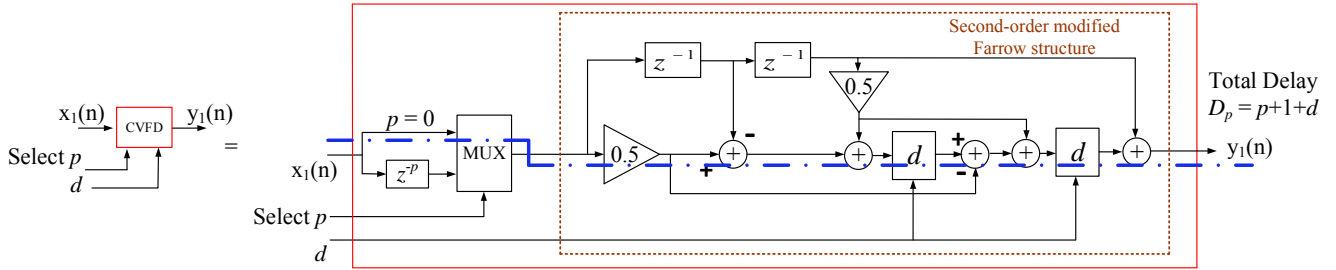


Fig. 1. CVFD element.

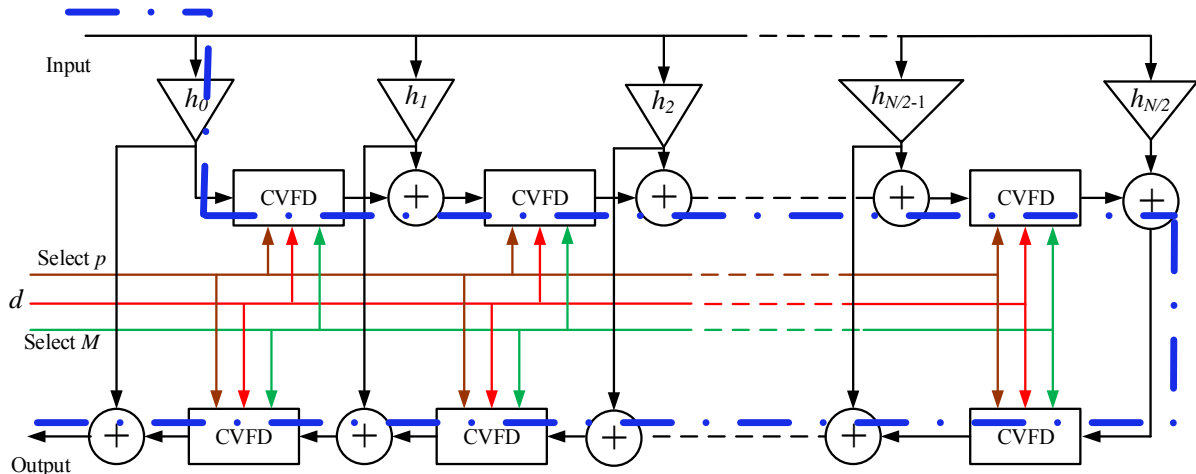


Fig. 2. Modal filter structure of the proposed ICVFD filter.

C. Variable Filter Responses obtained from ICVFD Filter

1. When $p = 0$

In the ICVFD filter, the cutoff frequency of the filter is controlled using three parameters viz. d , M and p . When $p = 0$ and $M = 1$, the ICVFD filter is equivalent to the FDS based filter. When $p = 0$ and M is varied, the ICVFD filter produces variable lowpass, bandpass and highpass filter responses. The cutoff frequencies of the bands in the multiband response are given by $f_{Ai} \pm f_{c_ICVFD}$, where f_{c_ICVFD} is defined in (3), with parameter $p = 0$ for the fractional delay D_c , i.e. the cutoff frequencies are $f_{Ai} \pm \{f_{c_mod}/(1+d)\}/M$, where f_{Ai} are the center frequencies of the bands in the multiband response given by,

$$f_{Ai} = 2i / M, \quad i = 0 \text{ to } M - \lceil M / 2 \rceil \quad (5)$$

The transition bandwidth of the bands in the multiband response is given by (4). The desired band can be extracted by using a suitable masking filter.

The variable bandwidth and variable center frequency responses obtained for the restricted range of D_c , i.e. when $p = 0$ are shown in Fig. 3. The magnitude responses obtained for $f_{c_mod} = 0.2$, $M = \{1, 2, 3, 4\}$ and $d = \{0, 0.1, 0.3, 0.7\}$ are shown in this figure. Four distinct colors and line styles are used to distinguish between the responses obtained for different values of M . The four responses of same color and line style (i.e. the responses obtained for the same value of M) correspond to the four values of the parameter d . Only 4 values of the parameter d are used in this illustration, so as to maintain the clarity of the figure and also to simultaneously show the fine variations in the cutoff frequency. The cropped version of the responses are also shown as inset to illustrate that no degradation occurs in the passband ripple of the ICVFD filter compared to the passband ripple of the modal filter. This is in contrast to the

FDS based filter where magnitude response for all the frequencies above 0.2 is distorted, even though they fall in the passband of the FDS based filter.

2. When $p \neq 0$

As can be seen from Fig. 3, the parameter d controls the fine variations in the cutoff frequency of the filter and the parameter M increases the cutoff frequency range. The new parameter p introduced in the CVFD element increases the range as well as the resolution of the cutoff frequency variation. The various lowpass filter responses that can be obtained from the ICVFD filter when p is also varied are shown in Fig. 4. The cutoff frequency of the modal filter is $f_{c_mod} = 0.2$ and the values of the interpolation factor used for this example are $M = \{1, 2\}$, and the values of the fractional delay parameter are $d = \{0, 0.1, 0.3, 0.7\}$. Only 4 values are used for the parameter d and the zoomed and cropped versions of the responses are shown so as to maintain the clarity of the figure. The values of the parameter p are 0, 1 and 2. The magnitude responses with line style

- 1) 'dash' (the responses in the blue color) are obtained for $p = 0$ and $M = 1$,
- 2) 'solid' (the responses in the red color) are obtained for $p = 0$ and $M = 2$,
- 3) 'dash and dot' (the responses in the black color) are obtained for $p > 0$.

The magnitude responses in the blue color are same as in the case of the FDS based filter. The magnitude responses in blue and red together are same as the lowpass filter responses obtained in Fig. 3 for $M = 1, 2$. As can be seen from the magnitude responses in black color in Fig. 4, when $p > 0$, the range as well as the resolution of the cutoff frequency variations is increased for the same sets of values of the parameters d and M . For instance, let $d = 0.1$. For $p = 0$ and $M = 1$, $D_c = 1.1$. For $p = 0$ and $M = 2$, $D_c = 2.2$. When

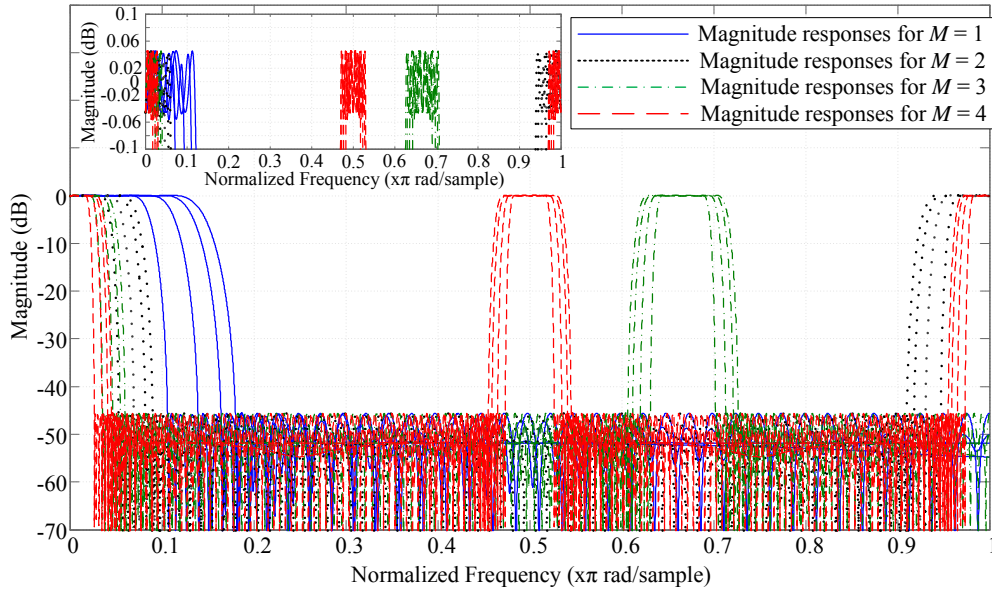


Fig. 3. Multiband responses obtained from the ICVFD filter.

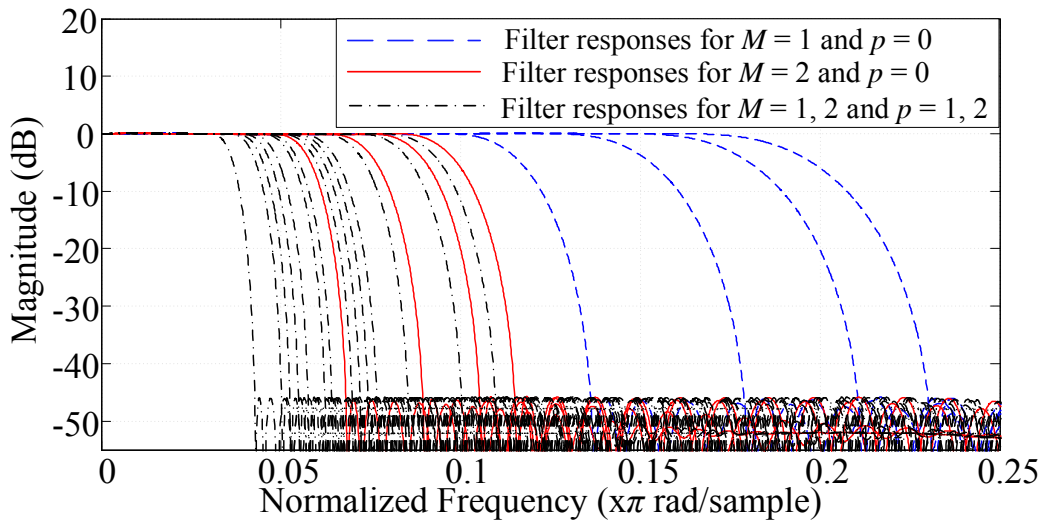


Fig. 4. Lowpass filter responses obtained from the ICVFD filter.

$p = 1$, $D_c = 2.1$ and 3.2 for $M = 1$ and 2 respectively. When $p = 2$, $D_c = 3.1$ and 4.2 for $M = 1$ and 2 respectively. Note that the fractional delay values of 2.1 , 3.2 , 3.1 , and 4.2 , and hence the corresponding cutoff frequencies are not possible in the case of the FDS based filter. The cutoff frequencies corresponding to these fractional delays can be obtained in the ICVFD filter without any increase in the multiplication complexity of the prototype filter structure compared to the FDS based filter. (As the multiplication complexity of the CVFD element is the same as that of the second-order modified Farrow structure, for the same filter order, the total number of multiplications required for the fractional delays remains the same for the ICVFD filter and the FDS based filter.)

D. Properties of ICVFD Filter

1. Passband ripple and stopband attenuation

In the ICVFD filter, all the filter coefficients of the modal filter are used for the filtering operation. Hence, unlike the CDM [6], no degradation occurs in the passband ripple or the stopband attenuation of the resultant ICVFD filter response compared to the prototype filter response. To illustrate this point, the magnitude responses of the prototype

filter and the ICVFD filters (for two different parameter settings) are shown in Fig. 5. The zoomed and cropped versions of the responses are also shown in the inset.

2. Transition bandwidth

As seen from (4), the transition bandwidth of the ICVFD filter is always less than or equal to the transition bandwidth of the prototype filter.

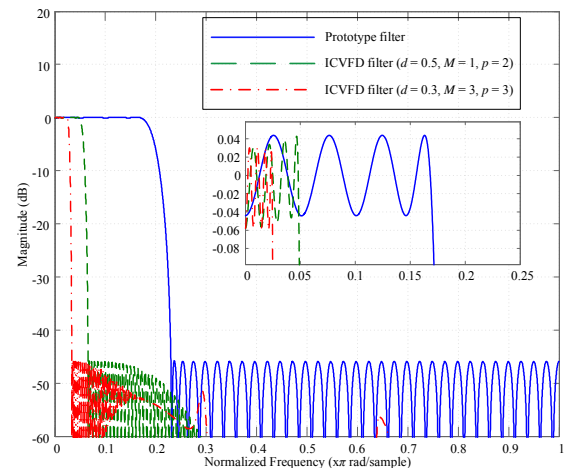


Fig. 5. Magnitude response of the prototype filter and the ICVFD filters.

3. Linear phase

Unlike the all-pass transformation based filter in [7], the ICVFD filter maintains the linear-phase property in its passband region. The magnitude-phase response plots of the ICVFD filter for two different parameter settings are shown in Fig. 6. Phase delay plots are shown in the inset of figures.

4. Cutoff frequency range

By proper choice of the fractional delay value and the interpolation factor, the cutoff frequency of the ICVFD filter can be varied anywhere below the cutoff frequency of the prototype filter (i.e. $f_{c_ICVFD} \leq f_{c_mod}$). One limitation of the proposed ICVFD filter is that $f_{c_ICVFD} \leq f_{c_mod} \leq 0.2$. This is because of the inherent limitation of the fractional delay structure that it provides unity magnitude response and constant phase delay only up to the normalized frequency of approximately 0.2 [13-15]. Beyond this range, the magnitude and the phase delay start deviating from the desired values.

E. Comparisons

The FDS based filter (for $d \geq 0.85$) as well as the ICVFD filter require a low complexity masking filter for suppressing the undesired bands in the filter response. The comparison of the ICVFD filter and the FDS based filter (without and

with CDM) is presented in Table I, for generating the lowpass filter responses. The SPA technique [9] and the technique in [11] are also considered for the comparison. Transposed direct form filter implementation is considered in each case. The desired final specifications are peak to peak passband ripple = 0.1 dB, stopband attenuation = -45 dB, and transition bandwidth = 0.1. All the filters considered for this comparison are designed to satisfy these specifications.

The ICVFD filter is designed with $M = \{1, 2\}$ and $p = \{0, 1, 2, 3\}$. The order of the modal filter is 46, and therefore, when implemented in transposed direct form, it requires 24 multipliers, 46 CVFD elements and 46 adders. Each of the CVFD elements requires 2 multipliers (for multiplication with d) and 5 adders, along with two 2:1 multiplexers for M and one 4:1 multiplexer for p . The order of the masking filter is 38, and it requires 20 multipliers and 38 adders. Therefore, the ICVFD filter requires total 136 (as $24 + 46 \times 2 + 20$) multipliers, total 314 (as $46 + 46 \times 5 + 38$) adders, 92 number of 2:1 multiplexers and 46 number of 4:1 multiplexers.

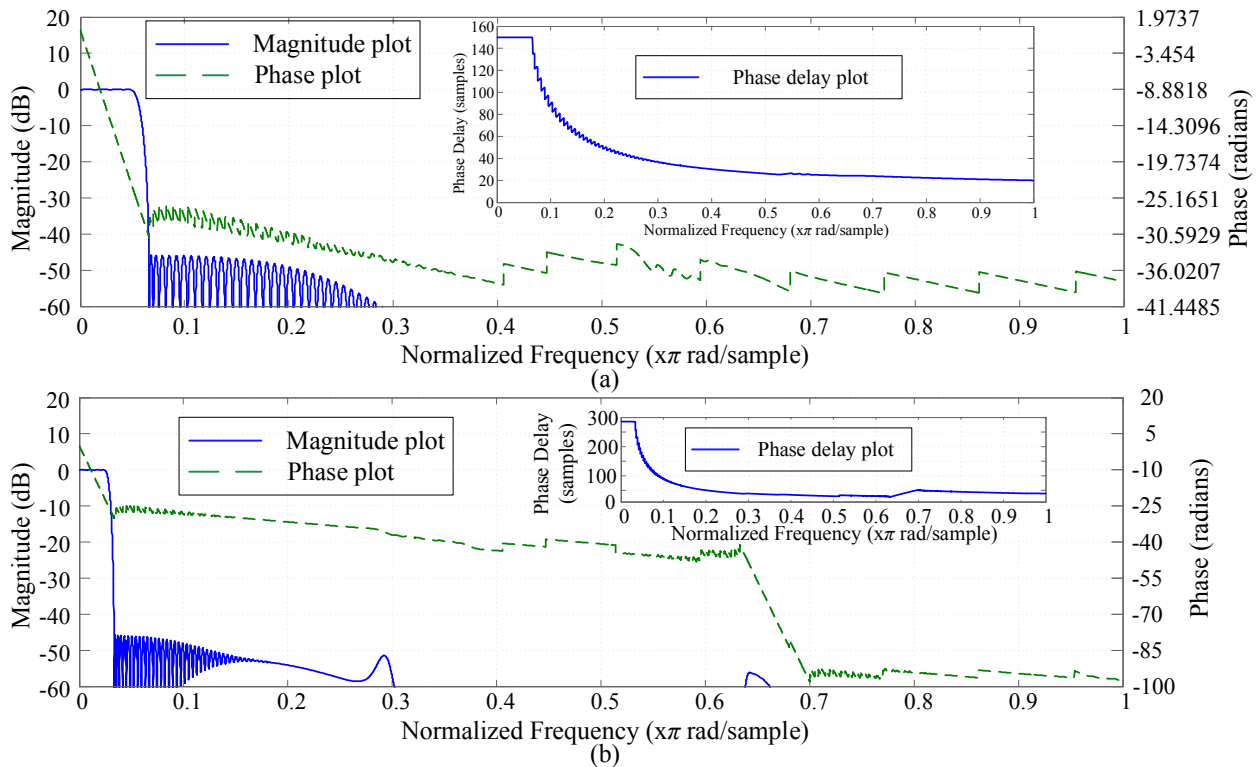


Fig. 6. Magnitude and phase response plots of ICVFD filter (a) for $d = 0.5, M = 1, p = 2$, (b) for $d = 0.3, M = 3, p = 3$.

TABLE I
COMPARISON OF RECONFIGURABLE DIGITAL FILTERS

	Number of				Total gate-count	Cutoff frequency range	Transition bandwidth
	Multipliers	Adders	2:1 mux	4:1 mux			
Proposed ICVFD filter	136	314	92	46	185008	0.0287 to 0.2	0.0143 to 0.1
FDS based filter [12]	122	286	0	0	160436 (-13%)	0.1 to 0.2	0.05 to 0.1
FDS based filter with CDM [12]	282	670	55	0	373548 (+102%)	0.05 to 0.2	0.025 to 0.1
SPA based filter [9]	>250	>500	0	0	>320500 (+73%)	0.0287 to 0.2	0.1
Filter in [11]	>250	>500	0	0	>320500 (+73%)	0.0287 to 0.2	0.1

Similarly, the total number of multipliers, adders, and multiplexers required for each of the filters considered for comparison are presented in Table I. Note that when FDS based filter is combined with the CDM technique, a higher order modal filter, and therefore, more resources are required in order to satisfy the final transition bandwidth and stopband attenuation specifications.

A 16x16 bit multiplier, a 16 bit adder, a 4:1 mux, a 2:1 mux, and a 2-input NAND gate are synthesized on a TSMC 65nm process using the Synopsis Design Compiler. The area of each component is normalized by the area of NAND gate. The total gate-count calculated from these normalized values represents the area of the filter in terms of the equivalent number of NAND gates. The number of multipliers, adders, multiplexers and the total gate-count calculated as explained above is presented in Table I for each of the filters. The ($\pm x$) values in Table I indicate the percentage increase or decrease in the total gate-count for the respective filters when compared with the ICVFD filter. As can be observed, when compared to the FDS based filter, the ICVFD filter offers wider cutoff frequency range and narrower transition bandwidth at the cost of only moderate increase in the area. Alternatively, the FDS based filter with CDM requires 102% more area when compared to the ICVFD filter, for comparable cutoff frequency range and transition bandwidth.

III. HARDWARE REALIZATION OF ICVFD FILTER

In order to realize the ICVFD filter on FPGA, we optimize the filter design in two steps, viz. use of pipelining (for improving operating speed and reducing the resource utilization) and utilization of FPGA specific feature (to improve the operating frequency further).

A. Pipelining for Hardware Implementation

The structure CVFD element is shown in Fig. 1, along with its critical path (shown as 'dash and dot' line with blue color). As can be seen, the critical path of the CVFD element extends from its input to the output. Therefore, as shown in Fig. 2 with 'dash and dot' line with blue color, the critical path of the modal filter of the ICVFD filter consists of a fixed-coefficient multiplier h_0 , N number of CVFD elements and N adders, where N is the order of the modal filter. As the ICVFD filter consists of the interpolated modal filter and a fixed-coefficient masking filter (used to extract the desired band from the multiband response), its critical path extends from its input of the modal filter to the output of the masking filter.

Such a long critical path makes the hardware implementation of the filter design infeasible without any pipelining. To improve the operating frequency, we add two levels of pipelining stages. First, in order to break the long critical path from input to output, a unit delay has been added between the interpolated modal filter structure and the masking filter.

After this first level of pipelining, the critical path is found to be from the input to the output of the interpolated modal filter. Therefore, to break this critical path, one pipelining delay can be added after each of the CVFD elements (second-level pipelining with one unit delay).

In order to shorten the critical path further, instead of

adding one unit delay after every CVFD element, two pipelining delays are added inside each of the second-order modified Farrow structure (second-level pipelining with two unit delays). As the variable multipliers, i.e. the multipliers with one input as d , are the most computationally intensive blocks, these two pipelining delays are inserted in order to separate these blocks. Additional delay elements wherever required are added in the filter structure, such that the overall filter functionality remains unaffected.

B. Hardware Realization of Variable-Length Delays

There are two variable-length delay structures in the ICVFD filter, viz. M variable delays inside the second-order modified Farrow structure (due to the interpolation) and p variable delays (as required in (1)). A straightforward way to implement such variable delays is to use multiple unit delay elements and select the appropriate number of delays using a multiplexer, with a select line with appropriate input for M or p . However, for FPGAs, multiplexers are costly in terms of both resource utilization as well as propagation delay.

As the hardware implementation of the delay element in the filter structure is done by a register, selection of variable number of delay elements (for M as well as for p) can be realized by using the addressable shift registers. We make use of Xilinx's IP core of RAM-based Shift Registers (SRLs) [16] to implement variable-length delays. The IP provides variable-length shift registers, which can be used as variable-length delay elements, with reduction in the propagation delay as well as resource requirement.

IV. IMPLEMENTATION RESULTS

The filter models were created considering the specifications mentioned in Section II-E. These filter models were created using MATLAB Simulink and Xilinx System Generator. The filters are implemented in the Xilinx Virtex 6 xc6vlx760-1ff1760 FPGA, using Xilinx ISE 14.6.

A. Effect of Pipelining

Filter implementation without any pipelining and after the first level of pipelining (i.e. separating the modal filter and masking filter) results in a very long critical path, resulting in infeasible designs. The estimated clock period after synthesis for ICVFD filter design with no pipelining and after first level of pipelining is more than 1000 ns. If one pipelining delay is added after every CVFD element (second-level pipelining with one unit delay), the ICVFD filter implementation becomes feasible with the post-place-and-route (post-PAR) maximum operating frequency of 30 MHz. Use of two pipelining delays inside each of the second-order modified Farrow structures (second-level pipelining with two unit delays) significantly improves the post-PAR maximum operating frequency to 58 MHz.

The implementation results for the interpolated modal filter structure with second-level pipelining with one unit delay and the interpolated modal filter structure with second-level pipelining with two unit delays are presented in Table II. Filter with second-level pipelining with two unit delays requires 84% more slice registers compared to the filter with second-level pipelining with one unit delay. However, due to the compact packing of the logic, it results in reducing the

requirement of LUTs and slices, and improving the post-PAR maximum operating frequency. As the overall area requirement is determined by the number of slices, use of two unit delays for pipelining actually results in reducing the area requirement by 19% and improving the maximum operating frequency by 91%.

B. Effect of SRLs

Pipelining improves the maximum operating frequency as well as reduces the number of occupied slices. To improve the operating frequency and reduce this area requirement further, variable-length delay structure can be realized using SRL instead of multiple delays and a multiplexer. Two ICVFD filter (interpolated modal filter + masking filter) models are created in MATLAB Simulink using the Xilinx System Generator block. One model utilizes multiple unit delay elements and multiplexer and the other utilizes addressable shift registers, which can then be realized as SRLs while generating Verilog implementation. The results are summarized in Table III. Use of SRLs results in reducing the requirement of slice registers by 39%. This results in compact packing of logic and better routing which improves the post-PAR maximum operating frequency by 7%. Use of SRLs also results in small (2%) improvement in overall area requirement (number of occupied slices).

C. Comparison with FDS based Filter

Similar to the pipelining of the ICVFD filter mentioned above, FDS based filter model with second-level pipelining with two unit delays was created. The FPGA implementation results of this FDS based filter are summarized in Table IV,

along with that of the ICVFD filter (for delay elements and multiplexer based design). The specifications are same as that considered for the comparison in Section II-E. The FDS based filter with CDM and filters based on the techniques in [9] and [11] are not considered for this comparison due to their high complexity. As can be observed from Table IV, the FPGA implementation results (increase in number of occupied slices) are in agreement with the theoretically estimated (increase in gate-count) results. The post-PAR minimum period of the ICVFD filter is slightly more than that of the FDS based filter, due to the more complex structure of the CVFD element used in the ICVFD filter when compared to the fractional delay structure used in the FDS based filter.

V. CONCLUSIONS

In this paper a continuously variable fractional delay (CVFD) element is proposed, which is used to replace the unit delay in the prototype filter of the proposed interpolated continuously variable fractional delay structure based filter (ICVFD filter). The CVFD element provides wide fractional delay range at the minimum complexity possible, and is capable of changing the fractional delay range on-the-fly. When compared to the existing fractional delay structure (FDS) based filter, the proposed ICVFD filter has dynamically variable, wider cutoff frequency range. It is also capable of providing variable bandpass and highpass filter responses. The ICVFD filter is suitable for obtaining the variable narrowband responses, especially in the lower

TABLE II
FPGA IMPLEMENTATION RESULTS OF INTERPOLATED MODAL FILTER OF THE ICVFD FILTER FOR DIFFERENT PIPELINING DELAYS

	Number of			Post-PAR minimum period (in ns)	Post-PAR maximum operating speed (in MHz)
	Slice registers	LUTs	Occupied slices		
Interpolated modal filter with second-level pipelining with one unit delay	9559	68800	20104	32.816	30.473
Interpolated modal filter with second-level pipelining with two unit delays	17580 (+84%)	56891 (-17%)	16231 (-19%)	17.203 (-48%)	58.192 (+91%)

TABLE III
FPGA IMPLEMENTATION RESULTS OF ICVFD FILTER FOR DELAY ELEMENTS AND MULTIPLEXER BASED DESIGN AND SRL BASED DESIGN

	Number of			Post-PAR minimum period (in ns)	Post-PAR maximum operating speed (in MHz)
	slice registers	LUTs	Occupied slices		
ICVFD filter with second-level pipelining with two unit delays – delay elements and mux based design	18545	60973	17395	16.69	59.916
ICVFD filter with second-level pipelining with two unit delays – SRL based design	11292 (-39%)	61342 (+1%)	17108 (-2%)	15.635 (-7%)	63.959 (+7%)

TABLE IV
FPGA IMPLEMENTATION RESULTS FOR ICVFD FILTER AND FDS BASED FILTER

	Number of			Post-PAR minimum period (in ns)	Post-PAR maximum operating speed (in MHz)
	Slice registers	LUTs	Occupied slices		
ICVFD filter	18545	60973	17395	16.69	59.916
FDS based filter [12]	12627 (-32%)	54795 (-10%)	15598 (-10%)	15.784 (-5%)	63.355 (+6%)

region of the frequency spectrum. Two-stage approach for the FPGA implementation of the ICVFD filter was presented. It was shown that the FPGA implementation results of are in agreement with the theoretical comparison of the ICVFD filter and the FDS based filter.

REFERENCES

- [1] Woo Jin Oh and Yong Hoon Lee, "Implementation of programmable multiplierless FIR filter with powers-of-two coefficients," *IEEE Transactions on Circuits and Systems II*, vol. 42, no. 8, pp. 553-556, Aug. 1995.
- [2] H. R. Lee, C. W. Jen and C. M. Liu, "A new hardware efficient architecture for programmable FIR filters," *IEEE Transactions on Circuits and Systems II*, vol. 43, no. 9, pp. 637-644, Sept. 1996.
- [3] T. Solla and O. Vainio, "Comparison of programmable FIR filter architectures for low power," *Proceedings of 28th European Solid-State Circuits Conference*, pp. 759-762, Italy, Sept. 2002.
- [4] K. H. Chen and T. D. Chiueh, "A low power digit-based reconfigurable FIR filter," *IEEE Transactions on Circuits and Systems II*, vol. 53, no. 8, pp. 617-621, Aug. 2006.
- [5] R. Mahesh and A. P. Vinod, "Reconfigurable frequency response masking filters for software radio channelization," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 55, issue 3, pp. 274-278, March 2008.
- [6] R. Mahesh and A. P. Vinod, "Low complexity flexible filter banks uniform and non-uniform channelization in software radios using coefficient decimation," *IET Circuits, Devices and Systems*, vol. 5, no. 3, pp. 232-242, May 2011.
- [7] A. G. Constantinides, "Spectral transformations for digital filters," in *Proceedings of the Institution of Electrical Engineers*, vol. 117, no. 8, pp. 1585-1590, Aug. 1970.
- [8] S. C. D. Roy and S. S. Ahuja, "Frequency transformations for linear-phase variable cutoff digital filters," *IEEE Transactions on Circuits and Systems*, vol. 26, no. 1, pp. 73-75, Jan. 1979.
- [9] S. C. Chan, C. K. S. Pun and K. L. Ho, "A new method for designing FIR filters with variable characteristics," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 274-277, Feb. 2004.
- [10] P. Lowenberg and H. Johansson, "Minimax design of adjustable-bandwidth linear phase FIR filters," *IEEE Transactions on Circuits and Systems I*, vol. 53, issue 2, pp. 431-439, Feb. 2006.
- [11] H. Johansson and P. Lowenberg, "On linear phase FIR filters with variable bandwidth," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 51, no. 4, pp. 181-184, Apr. 2004.
- [12] S. J. Darak, A. P. Vinod and E. M-K. Lai, "A new variable digital filter design based on fractional delays," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1629-1632, Prague, Czech Republic, May 2011.
- [13] V. Valimaki, "A new filter implementation strategy for Lagrange interpolation," *IEEE International Symposium on Circuits and Systems*, vol. 1, pp. 361-364, USA, May 1995.
- [14] T. I. Laakso, V. Valimaki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay," *IEEE Signal Processing Magazine*, vol. 13, no. 1, pp. 30-60, Jan. 1996.
- [15] C. Candan, "An efficient filtering structure for Lagrange interpolation," *IEEE Signal Processing Letters*, vol. 14, no. 1, pp. 17-19, Jan. 2007.
- [16] DS222: LogiCORE IP RAM-based Shift Register, Xilinx, 2011.

Linear Robot as The Approach towards Individual Abdominal Ultrasound Scanning in Developing Countries?

Christina Pahl

Abstract—In order to improve comfort in abdominal ultrasound scanning, reduce operator dependency and increase image resolution, a novel method for automatic ultrasound cervix scanning, especially for developing countries, has been developed. To determine the reference scanning point for customized diagnosis procedures, personalized scanning algorithms based on umbilicus detection are utilized. The setup consists of three linear motors controlled by an ATmega328 microcontroller and is based on real-time data of an IR sensor as feedback information. For testing purposes 10 subjects were scanned using a robotic platform. Robotically acquired data were compared with manually taken ultrasound images. Safety and image quality have been evaluated and approaches developed to counteract mentioned deficits. A correlation analysis as well as image quality classification between histograms and probe pressure image data was performed. Results show an automated method for personalized ultrasonography with a maximum accuracy of the defined scanning reference being 98.4 % for one dimension. It can be concluded that simple image processing techniques like 3D reconstruction and merging in combination with robotically guided ultrasound probe motion are highly efficient to counteract existing ultrasonography deficits. Future works consider the application of automated ultrasonography on related medical fields.

Keywords—linear robot, merging, personalized ultrasonography, probe pressure, safety.

I. INTRODUCTION

Referring to the most recent cancer statistics provided on the World Health Organisation (WHO) database, maintained by the International Agency of Research on Cancer (IARC), cervical cancer represents the second most common cancer in women worldwide. The international incidence rates of cervical cancer show that almost 80 % of cervix cancer cases occur in low-income countries [1]. Therefore, in developing countries like Malaysia, cervical cancer has become a serious problem and is currently part of discussions in health Politics [2]. According to WCRF International the new incidence rates worldwide of colorectal cancer show that with a number of 614,000 in the year of 2012, this cancer represents nearly 10 % of all international cancer cases [3]. It can be seen that cancer types occurring in the abdominal area are dominating in the list of all cancers in females as well as males worldwide.

This work was supported by Universiti Teknologi Malaysia under Grant 03G12. Manuscript received Nov. 12 2015, revised Dec. 18, 2015. C. Pahl is with the Biomechatronics Group, Technische Universität Ilmenau, 98693 Ilmenau, Germany (phone: +49 3677-694690; e-mail: christina.pahl@tu-ilmenau.de) as well as with IJN-UTM Cardiovascular Engineering Centre, University Technology Malaysia, Skudai, 81310, Malaysia.

In order to decrease the number of abdominal cancer cases by improving the diagnosis of these malignances, an approach on the automation of ultrasonography based on a personalized setup [4] consisting of a starting point with the lowest risk factors and costs will be presented here [5]. Although, ultrasonography is highly operator dependent and provides a low contrast resolution [6], it is used to evaluate the size and extent of tumors [7]. Among existing imaging techniques, ultrasound is preferred in the examination of the abdomen. It is affordable compared to MRI and CT [8], widely available, noninvasive, non-ionizing [9] and painless [10]. Ultrasonography is a medical imaging technique that uses high frequency sound waves [11] and their echoes [12]. The technique is similar to the echolocation used as SONAR by submarines [13].

The machine displays the distances and intensities of the echoes on the screen [14] forming a two dimensional [15] image [16]. Table I shows a basic comparison of existing technologies for abdominal cancer visualization. However, the use of ultrasound in medical examinations encompasses some limitations. A major drawback of ultrasonography is high operator dependency [17]. Since image acquisition is complex, diagnostic accuracy in ultrasound is directly related to skills, training and experience of the operator [18].

Moreover, experience is highly required in analyzing and interpreting the anatomical appearance of malign structures [19]. Each operator may interpret one and the same ultrasound image of cancer differently [20]. Concerning the limited capabilities of ultrasound in cervix imaging, it is important to develop alternative approaches to remove the current disadvantages. Autonomous systems can help in minimizing operator dependency, to increase reproducibility and thus to increase comparability of image data for further diagnosis. One approach towards this represents the usage of linear motors performing the abdominal cervix scan. Its high accuracy in moving the transducer probe and stability in keeping a desired pose [21] allow to acquire highly comparable data [22]. Currently, there are no automatic systems developed for abdominal ultrasound diagnostics like the proposed system [23].

TABLE I
COMPARISON OF CT, MRI AND ULTRASOUND (X=LOW, XX=MEDIUM, XXX=HIGH; N=NO, Y=YES)

Modality	Invasiveness	Cost	Time	Sensitivity
CT	N	XX	XX	XX
MRI	N	XXX	XXX	XXX
US	N/Y	X	X	X

A. Ultrasound Safety Considerations

Besides its apparently excellent safety record, medical ultrasonography still exposes energy in the organic tissues [24]. There appear biological effects if high power in combination with long exposure time is applied. Due to its thermal effects, safe usage of this imaging method should be taken into consideration in clinical applications. In order to correlate potential ultrasound effects with exposure, a characterization of the ultrasound beam is essential [25]. In this work, brightness mode (B-mode) imaging will be applied to acquire 2D ultrasound images of the human cervix.

B. Effects of Ultrasound Imaging

In contrast to Doppler diagnosis, B-mode imaging has not demonstrated to produce harmful temperature rises in the body [26]. However, a basic approach towards safe use of ultrasound is the as low as reasonably achievable (ALARA) principle. Two indices are introduced here: mechanical index (MI) and thermal index (TI). Whereas, MI is associated with the degree of cavitation bio-effects, TI indicates the tissue heating bio-effect. Cavitation is a non-thermal bio-effect, which takes place when small gas bubbles oscillate, increase in size, and collapse due to the influence of ultrasonic field [27]. Thermal effects indicated by TI, may result in the increase of tissue extensibility, blood flow but also reduction in joint stiffness and muscle spasm. Fig. 1 follows the guidelines for the safe use of diagnostic ultrasound equipment provided by the British Medical Ultrasound Society (BMUS) and shall give an overview about the relation between recommended scanning time and metric TI. It can be obtained that with increasing TI, the recommended scanning time decreases. At this point it can be concluded that a compromise between scanning time and TI needs to be defined in order to achieve the highest possible image quality with regards to safety requirements. Therefore, a TI value between 1.5 and 2.5 will be used.

C. Intensity and Exposure Time

A typical diagram is illustrated in Fig. 2 showing the intensity of the ultrasound beam with given frequency against the exposition time. The duration used image the inner organs of the human body is restricted by the time the ultrasound is exposed to the body.

In Fig. 2 safe, critical and unsafe area are indicated. It can be seen that with increasing scanning time, the frequency needs to be adjusted in order to stay inside the area of safe usage of the clinical ultrasound.

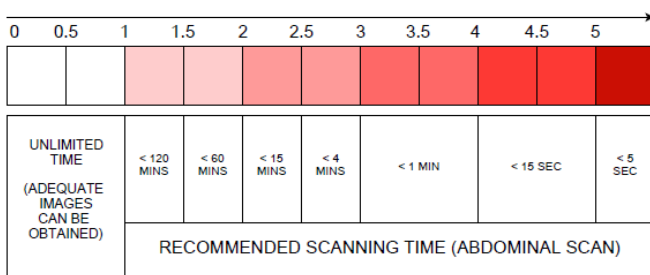


Fig. 1. Thermal Index and Recommended Scanning Time

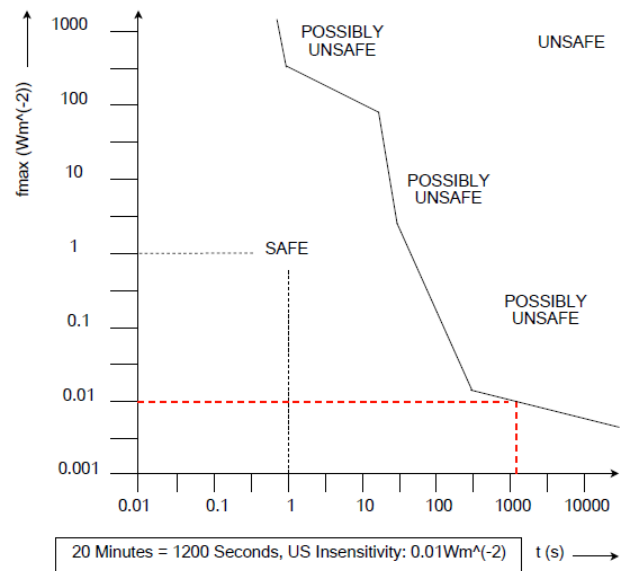


Fig. 2. Ultrasound Safety According to Frequency and Time

In Fig. 2 guidelines of the American Institute of Ultrasound in Medicine (AIUM) regarding the ultrasound intensity and allowed exposure time are shown. The beam frequency should be high for lateral resolution but depends on the depth of the organs, which need to be scanned. In order to image the cervix with the highest duration for reconstruction purposes and panoramic views, the highest scanning time possible for best possible results is required. Being here restricted to safety specifications, a compromise needs to be made. Therefore, a maximum intensity of 0.01 Wm² is chosen to be applied. This is illustrated by the dashed red line in Fig. 2.

II. DESIGN AND IMPLEMENTATION

It is important to develop alternative approaches to minimize the current disadvantages in ultrasound imaging due to the limited capabilities of ultrasound in cancer imaging. The suggested setup for automatic ultrasound scanning shall be cost-effective, stable and provide a high movement resolution. An algorithm was developed and allows operator independent detection of the umbilicus on the abdominal surface. The control scheme of the developed system can be seen in Fig. 3. It can be obtained that three motors receive data from the Atmega328 Microcontroller. Furthermore, the gear used to transmission the movement to the linear motion system and the motors used for the generation of motion proportional to the signal of the drivers. The electrical part consists of drivers, which provide voltage according to data they receive from the controller. The last part involves computing. The three motors satisfy the movements in three dimensions. These motors also move the position of the Infrared (IR) sensor.

This sensor sends data to the microcontroller, which analyzes the values according to the algorithm input from the computer and regulates subsequently the position of the sensor by correcting the movement commands of the motors [28]. The ultrasound machine used here was the TOSHIBA APLIO-MX ultrasound system.

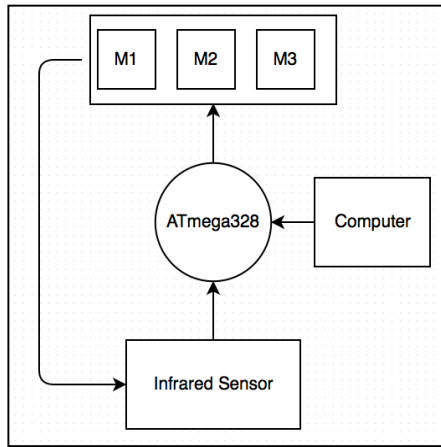


Fig. 3. Control unit of personalized ultrasound scan for automatic umbilicus localization using IR technology.

A. Characteristics for Abdominal Scanning

The IR sensor being the main part of this setup is attached to the transabdominal ultrasound probe and its mechanic holder. The setup was tested using two phantoms and 10 subjects. The characteristics of the subjects can be seen in Table II. Adult subjects, listed with the capital letter *S* and a following number, were scanned. All of the subjects were female. The youngest subject was 21 years old, whereas the oldest was documented with 36 years. The letter *A* represents the age. Here, *C* refers to the circumference of the abdomen. It is important since the abdominal detection for the umbilicus is different in the case large abdominal tissue is available or if the opposite case holds true. The letter *H* represents height. The letter *W* referring to weight can be seen as an indicator for the body mass in relation to height and therefore an index for abdominal mass. *UV* refers to the distance between umbilicus and pubic bone.

B. Umbilicus Detection

Fig. 4 displays the algorithm for the motor and driver control in order to localize the umbilicus. The algorithm consists of eight decision steps. The starting point marks the position of the three motors with X_i , Y_i and Z_i being here zero indexed.

TABLE II
CHARACTERISTICS OF SCANNED SUBJECTS

	A [YEARS]	C [CM]	H [CM]	W [KG]	UV [CM]
S1	36	79	154	59	18
S2	29	76	148	49	20
S3	28	74	152	61	19
S4	26	71	164	57	26
S5	24	69	154	52	23
S6	23	63	169	58	26
S7	23	87	158	54	19
S8	22	73	147	50	21
S9	21	73	158	58	21
S10	21	75	164	65	23

Subsequently, destination and data acquisition represents the approximation of the umbilicus based on personalized values calculated for each subject on the basis of the parameters from Table II according to (1).

$$I_{SA} = A + C + H \times W + UV^2, \quad (1)$$

where I_{SA} represents the individual scanning area defining the extreme points of the frame of probe movement as shown. Values for I_{SA} in this setup range between 7000 and 12000, where 1000 points define a 1 cm wider scan range leading to a 4 times bigger scan window for subject *S10* compared to *S2* for example. The direction and pulse are subsequently set.

Meanwhile, to the movement of the transducer and therefore the IR sensor, the values for the sensor are sent to the microcontroller and compared with experimentally predefined thresholds. Experiments show that a threshold between 360 mV and 380 mV delivers the required results for the detection of the umbilicus. In the case that the values differ from the threshold, the motor will be continuously moved until the values overlap and the motors stop automatically. The third method shows the approximation of the umbilicus by orbiting in 90° angulations defining a small circle at the starting point and resulting in bigger circles in the end. This approach showed the highest rate of success in umbilicus approximation and was therefore used in this setup.

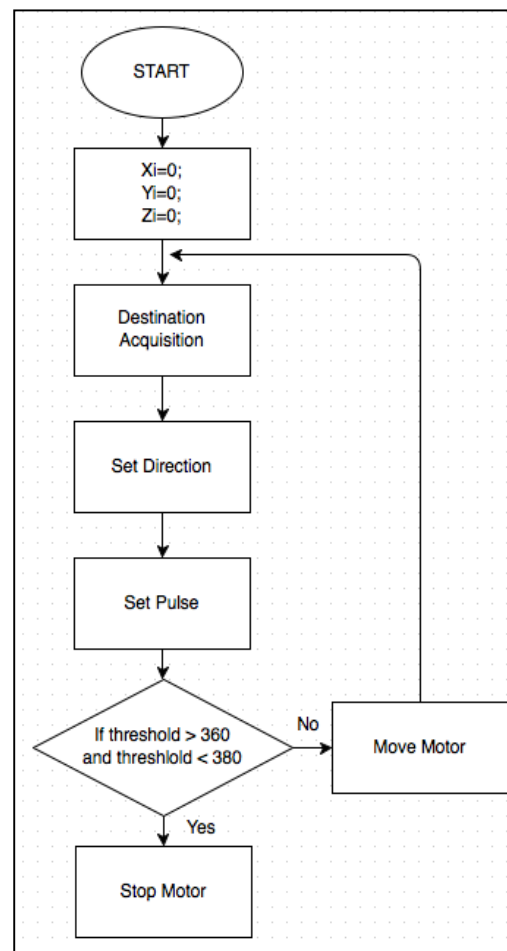


Fig. 4. Side and top view of the setup for the automatic and personalized ultrasound scans.

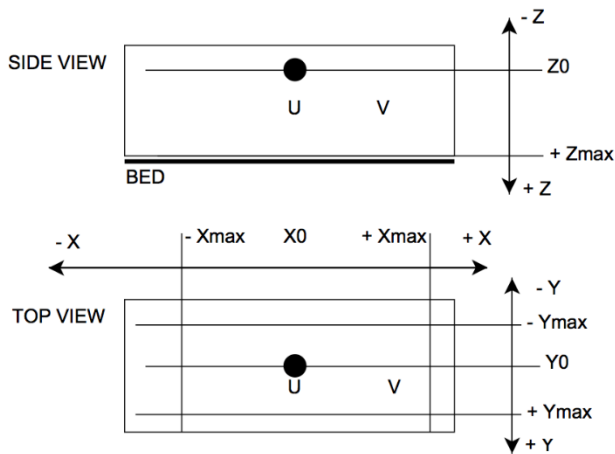


Fig. 5. Side and top view of the setup for the automatic and personalized ultrasound scans.

C. System Setup

The mechanical movement extension restricts the automatic scan [29]. Fig. 5 shows the setup for the reference scan using operator dependent parameters. There are two subfigures provided. The first one displays the side view, whereas the second one shows the top view. The allocation of the parameters is displayed. The letter *U* represents the umbilicus, whereas *V* stands for the end of the pubic bone. Moreover, Fig. 5 shows the setup for the detection of the umbilicus shown in side and front view. An IR sensor was attached to the probe holder. The voltage output of this sensor is a function of the distance between the sensor and the reflected IR waves. Because of the differing position of the sensor referring to the probe, a shift of 2.5 cm needs to be corrected in order to adjust the position of the probe for subsequent scanning. This is done after successful detection. The analog output of the IR sensor was connected to the digital converter of the μC in order to take subsequent distance measurements. On a distance between the sensor and the umbilicus ranging 4 cm to 20 cm, measurements were performed acquiring approximately 2 Million voltage values.

For this setup, the processor of the PC should be at least 2.66 GHz for the processing of the video grabber device, the Operating System (OS) should be at least Windows 2000 or any other OS like OS X 10.3 or Linux 2.6x. The hard disc should be able to store at least 5 GB, which represents 100 videos captured. The RAM should have a capacity of at least 256 MB. Out of these data, a distance between sensor and umbilicus marker of 3 cm was chosen because this value represents the highest voltage output.

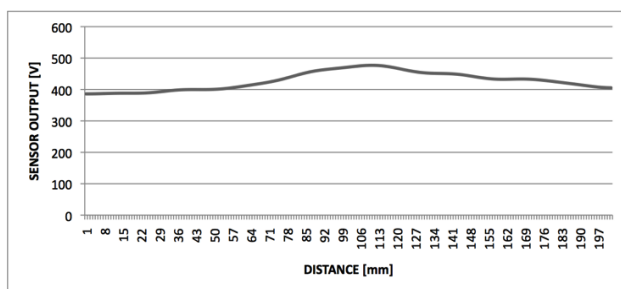


Fig. 6. Side and top view of the setup for the automatic and personalized ultrasound scans.

A squared object with an edge length of 3 cm was chosen as a marker for the umbilicus. This object was then placed above the umbilicus and used as a marker. The distance between marker and the highest point of the abdomen was set as 3 cm. The subjects were then placed on the patient bed, so that in lying position 5 linear probe motions parallel with the marker in a height of 3 cm distance from the marker could be performed. Then the subjects were scanned for 15 minutes per session. The time for umbilicus recognition was in the range of seconds. However, the duration depended on the starting point of the transducer probe and the position of the subject respectively. Humans can visually detect the umbilicus. However, in this setup human intelligence for umbilicus recognition will be substituted. Approaches towards Artificial Intelligence (AI) [30] based robotic systems may be considered for classification of data [31]. Fig. 6 shows a typical voltage output for the infrared scanning procedure above the abdomen.

III. RESULTS AND ANALYSIS

A. Scanning Safety

The feature of an emergency button is necessary, since safety for both, human and machine, is important. For the usage of the emergency button a frequency of 360,000 Hz and a resolution of 8,000 pulses per rotation was chosen.

This ensures, that the probe moves quickly from its actual position to the initial point. This is because any emergency case requires the ability for fast escape. Choosing the lowest resolution is even at the highest frequency too slow. Choosing the highest resolution is also not effective. Thus, a sufficient velocity was chosen according to the characteristics of the motor. This means, that for the actual position of the probe, which is located at the center of the linear robot, a maximum freedom of movement for the scanned subject can be ensured in emergency cases. From Table III it can be concluded, that the emergency button is working with a sufficient velocity according to human reaction capabilities [32]. The values for *X* and *Y* are therefore zero, because the emergency button only considers the *Z*-direction and no other directions. The first column is filled with zeros because to the point of time, the probe position is already in its initial point. The time required to set the probe into its initial point (2 seconds) is suitable for applications with occurring emergency cases.

B. Stop button

Coronal and sagittal images of the cervix have been taken. Moreover, the depth of the probe is displayed in Fig. 7. *Stop 1* represents soft touch. Whereas, *stop 2* refers to medium pressure, *stop 3* is already reported to be hard.

TABLE III
EMERGENCY BUTTON RESULTS

Condition	Initialization/ patient data input	Umbilicus localization	Scanning
Distance [mm]			
X	0	0	0
Y	0	0	0
Z	0	157.3	208.2
Time [s]	0	1.16	1.53

The last stop and simultaneous image capturing was performed for pain felt due to the probe depth in the lower abdomen. From image data, it can be seen that there is a correlation between probe depth and image quality. The more deep the probe is pushed into the abdomen, the better artifact causing gases can be pushed away and the closer the cervix can be examined. In order to better understand Fig. 7, the first row will be analyzed. *S1* provides images ranging a depth from 0.8 cm until 4.1 cm. The first stop indicates, that the subject experiences here soft pressure. The image quality is according to the low pressure less good compared to the following images. Moreover, the uterus and the cervix are less clear. In the next step, the stop button was pushed when medium pressure was felt. The depth has a value of 1.6 cm in this case. The structures of the uterus and consequently the cervix get more obvious. The third stop was pushed when hard pressure was felt on the abdomen. The image provides in this stage already more information and can be regarded as useful. However, the best image could be obtained with a probe depth of 4.1 cm. The details are at this stage more clearly differentiable compared to previous images. The hollow body of the uterus in this image is highly clear. *D* stands in Fig. 7 for the transducer probe depth measured from the surface of the abdomen.

C. Probe pressure and depth

Common transabdominal ultrasound scanning requires the contact pressure between the probe and the subject in order to achieve a good acoustic match. Another test was performed regarding the probe depth. Pressure must be applied evenly in order to get accurate direction of the scan. Therefore, operator dependent scanning is less suitable than automatic for better image results concerning image quality. Table IV shows the allocation of quality values and the pressure applied. The image quality parameters have been determined on the basis of anatomic landmarks used as reference points. Since the correct pressure application of probe can be used to considerably improve the image quality, probe depth was analyzed. Pressure affects the echogenicity of tissue and shortens the distance to cervical structures. All values are based on contrast properties of images by comparing their histograms. Here, it can be obtained, that in three out of ten cases, the highest pressure could not deliver best image data. Excessive pressure, however, can cause discomfort to the subject. It is assumed that too high pressure causes cramping as a protective reaction and thus influences the imaging procedure and its results. On the basis on the analysis results of an experienced sonographer, the image quality of resulting image data could be defined. This analysis considers contrast, boundary of cervix as well as artefacts. Image quality indicators from 1 to 4 were utilized. Number 4 stands for the lowest and number 1 for the highest pressure applied on the abdomen. All these values reflect subjective impressions of ultrasound probe pressure. Detail contrast can have a significant visual impact on the ultrasound image. This is because it emphasizes texture, so that the cervix appears with a higher number of details and more pronounced highlights.

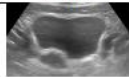
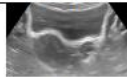
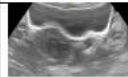
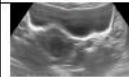
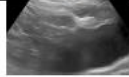

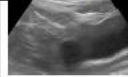

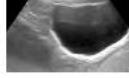
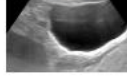
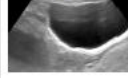
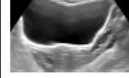
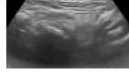




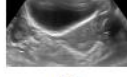
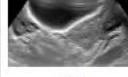
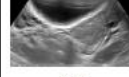








	Stop 1	Stop 2	Stop 3	Stop 4
<i>S1</i> <i>D₁</i> [cm]	 0.8	 1.6	 2.4	 4.1
<i>S2</i> <i>D₂</i> [cm]	 1.0	 2.9	 3.7	 5.2
<i>S3</i> <i>D₃</i> [cm]	 1.2	 2.2	 4.6	 6.8
<i>S4</i> <i>D₄</i> [cm]	 0.4	 1.8	 2.2	 3.9
<i>S5</i> <i>D₅</i> [cm]	 0.6	 1.8	 4.2	 3.7
<i>S6</i> <i>D₆</i> [cm]	 0.7	 1.3	 3.1	 3.8
<i>S7</i> <i>D₇</i> [cm]	 0.8	 1.7	 2.4	 3.7

Fig. 7. Stop-button Function and Depth Related Image Quality

Therefore, the histogram approach for image quality analysis was chosen. The histograms for three subjects are displayed in Fig. 8. Subject 5 and 7, show the expecting result, being higher contrast followed by higher pressure. As shown for the last case, the histogram has less contrast for the image under pain than under hard pressure. For each case, soft pressure leads to less contrast in the image, whereas medium pressure increases the contrast. The contrast quality between hard pressure and pain cannot be unambiguously correlated, see Fig. 9. This is case specific and shows the requirement of a pressure feedback using the previously introduced stop-button.

TABLE IV
IMAGE QUALITY COMPARED TO ULTRASOUND PROBE PRESSURE

Image quality \ Pressure	Pressure			
	Soft	Medium	Hard	Pain
<i>S1</i>	4	3	2	1
<i>S2</i>	4	3	2	1
<i>S3</i>	4	2	1	2
<i>S4</i>	4	3	2	1
<i>S5</i>	4	3	2	1
<i>S6</i>	4	3	2	1
<i>S7</i>	4	3	2	1
<i>S8</i>	4	3	1	2
<i>S9</i>	4	3	1	2
<i>S10</i>	4	3	2	1
Average	4	2.9	1.7	1.3

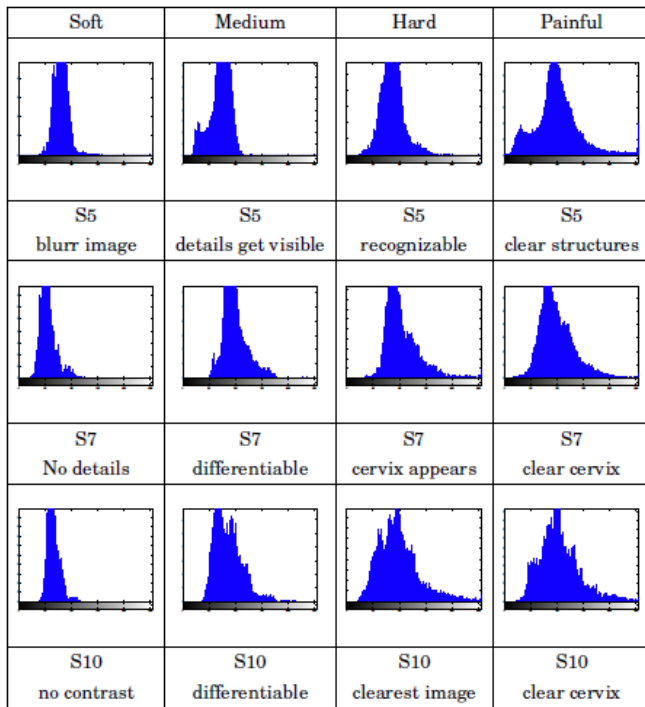


Fig. 8. Histograms of Cervix Ultrasound Images with Different Pressure

In the following, three applications shall be introduced, explaining in how far data, which are acquired by the system, can be used. First, two approaches in improving the image quality on the basis of resolution will be explained. Subsequently, the usage of the scanning system for ultrasound based panoramic views will be explained. Finally, an example using 3D reconstruction will be shown.

D. Resolution improvement

The first approach for improving the ultrasound scanning resolution is based on the idea of combining two or more scans with spatial shifts ΔX into a single scan as shown in Fig. 10. Merging ultrasound images with restricted detail information leads to the increase of information, which improves the resolution and consequently follows into a higher image quality. For this method, it is useful to have a scanning system like the one developed in this work. Such a system enables the motion of the scanning transducer probe in very small steps, so that the ultrasound probe crystals, which miss the information located in the area of their gaps can be still acquired. Moreover, for this purpose it is important to move the probe with a high accuracy, which can only be provided by machines, currently. Assuming that an ultrasound probe of a length of 6 cm is covered with 120 crystal elements, the resolution would be 0.5 mm according to (2).

$$Probe\ Resolution = \frac{60}{120} = 0.5 \quad (2)$$

Such a resolution is often not sufficient in providing reliable information on tissues, which could be affected by cancerous cells.

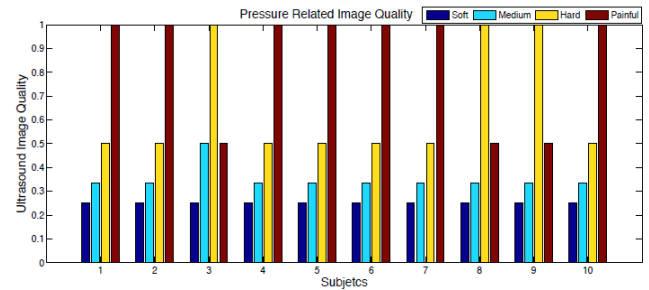


Fig. 9. Relation between Probe Pressure Classes and Image Quality

Therefore, the necessity exists to improve the resolution. Piezoelectric crystals convert electricity into sound and the other way around. The distance between each crystal element is defined as ΔX . A static scan of a phantom containing several elements of the size of the crystals and the distance between them, apparently provides only half of the information originally existing. The motion of the probe in a distance of exactly ΔX leads respectively to the result that only those phantom details are displayed, which were ignored by the beam due to the gaps. Having two scans now with a shift of the size of the crystal to crystal distance does not provide a better image quality.

Nevertheless, the combination of both images into one image leads to doubled information details. Such a merged image is of higher value for subsequent analysis procedures. Another step towards improving the resolution of ultrasound image data is based on the reconstruction of 3D data from 2D ultrasound images as shown in Fig. 11. Taking into consideration, that an ultrasound phantom consists of single structures being separated from each other in a distance of 20 μm , a static scan with the average resolution of 0.5 mm or even a dynamic scan of a scanning resolution of 20 μm and more are not sufficient to differentiate existing line structures.

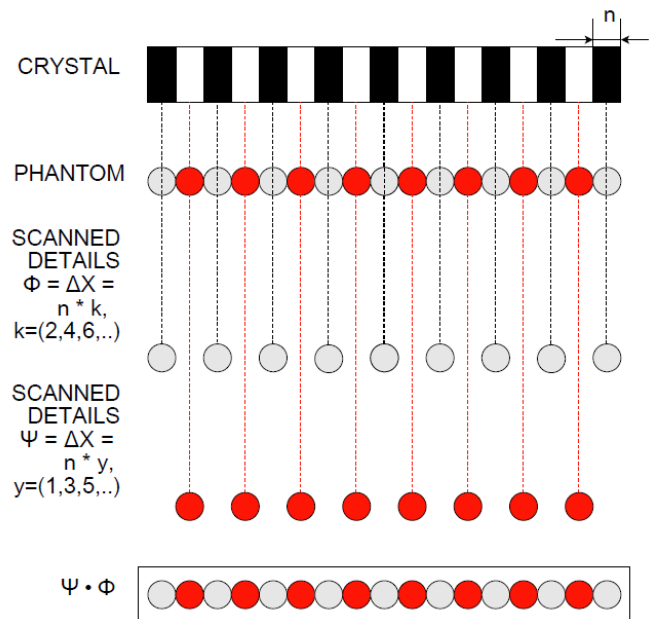


Fig. 10. Resolution Improvement by Merging

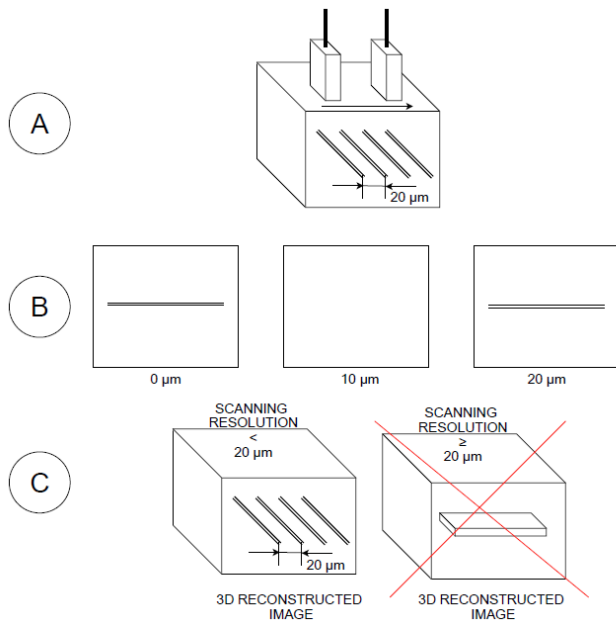


Fig. 11. Resolution Improvement by 3D Reconstruction

In order to do so, a dynamic scan with a shift of less than $20\ \mu\text{m}$ is required to obtain the single lines. Here, the figure can be divided into three parts. Part A shows the phantom, B displays single images for a scanning resolution of $10\ \mu\text{m}$ and C shows the results for a successful and a failed reconstruction. The three images displayed below the phantom, show a shift of $10\ \mu\text{m}$. It can be seen that such a shift is sufficient to display the lines clearly as single structures. After reconstructing images with a shift of $10\ \mu\text{m}$ or even lower, an accurate representation of the phantom can be achieved.

E. Panorama Scanning

In the case, that the shift is equal or bigger than the distance between the lines, the 3D representation would be similar to the crossed representation, which is a wrong representation according to the phantom. However, such a shift needs to be highly accurate, which cannot be provided by a manual scan procedure. Therefore, on the basis of the previously shown results of the system, a possible application would also be the here introduced improvement of 2D ultrasound images on the basis of shift related 3D reconstruction. High resolution ultrasound data usually provides a limited field of view.

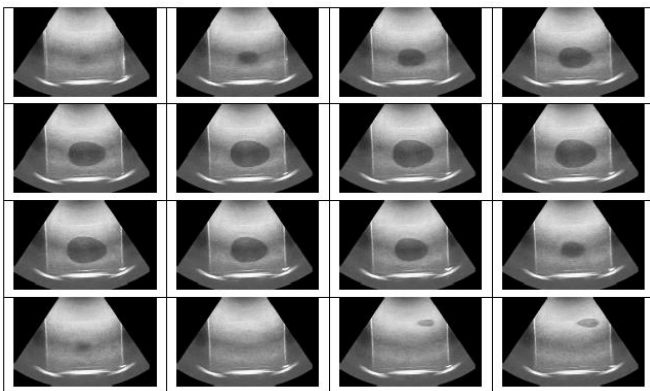


Fig. 12. Resolution Improvement by 3D Reconstruction.

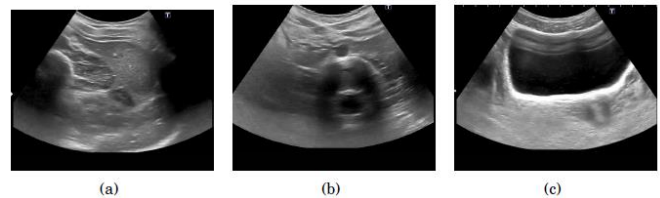


Fig. 13. Resolution Improvement by 3D Reconstruction.

For the visualization of larger volumes, a composition of the images is required. The most appropriate method to provide an extended field of view is panoramic scanning. Panoramas can be constructed by matching the images according to the location and orientation of the probe when it gets moved. The location of the probe and its motion are important parameters to subsequently display the anatomy in the same way it appears in nature. Therefore, the developed scanning system represents a suitable mechanism to provide the required accuracy in moving the probe. In order to test the performance of the system for providing panoramic scans, one phantom and a subject were scanned. For the phantom test 152 ultrasound slices have been acquired. The scanning resolution was set to $1\ \text{mm}$ for a distance of $15\ \text{cm}$. The time required was $4.05\ \text{seconds}$ with a frame rate of $37\ \text{fps}$. The phantom represents a mimicking tissue object containing an egg shaped item. In Fig. 12, the coronal slices can be obtained. The object gets scanned from positive to negative X-direction. It displays its oval shape. Moreover, a smaller object is visible in the last two frames. It can be concluded at this point, that the system is able to scan short distances with low resolution in relatively short time. For the subject scan, three sequences of each $30.000\ \text{frames}$ were acquired with a scanning resolution of $10\ \mu\text{m}$ during a scanning time of three times $14\ \text{minutes}$. The distance of the scanning area was defined as $30\ \text{cm}$. The scanned medial and lateral tissues of the abdomen can be obtained in Fig. 13.

F. 3D Reconstruction

According to previous considerations, automatic systems are highly suitable for the improvement of resolution in ultrasound images. In order to obtain the images, the software of [33] was used and modified according to the phantom image characteristics. Therefore, functions had to be adjusted.

Here, the results of two approaches, being manual and automatic scanning, are compared after 3D reconstruction based on volume rendering. This can be obtained from Fig. 14. Fig. 14(a) shows the front view on the automatically acquired slides captured with coronal probe orientation. The interpolation results of two lines are displayed after sagittal cut Fig. 14(c). Fig. 14(b) represents the front view of the manually acquired frames captured with coronal probe orientation. After interpolation, the sagittal view on one line is shown in Fig. 14(d). For this purpose, a tissue mimicking phantom for lateral scanning resolution testing was used. For both setups a total of $1021\ \text{frames}$ was captured within a time of $13.5\ \text{seconds}$ per scan. The distance of the scanning area encloses $5\ \text{cm}$. The scanning resolution of the system was set to $100\ \mu\text{m}$. In the case of automatic scanning, two lines of a diameter of $1\ \text{mm}$ are displayed in their straight appearance after 3D reconstruction. The manual scan displays a line of $1\ \text{cm}$ in diameter.

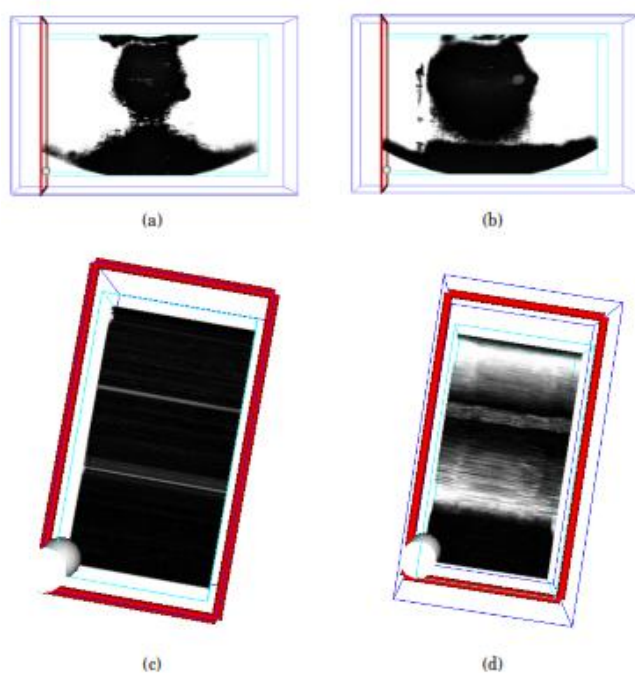


Fig. 14. Resolution Improvement by 3D Reconstruction: (a) shows the front view on the automatically acquired slides captured with coronal probe orientation, (b) represents the front view of the manually acquired frames captured with coronal probe orientation, (c) after sagittal cut of (a) and (d) sagittal cut of (b).

The white areas surrounding the line shall be ignored at this stage. Due to the fact that the data acquired by the use of freehand scanning includes variations in pressure and changing velocity, the line is shown irregularly after reconstruction. This may lead to misinterpretations of the actual structure. From these results, it can be concluded that manual scanning is highly unsuitable for accurate scan purposes and requires the substitution by a machine, like the automatic cervix scanning system.

G. Application and Testing of Umbilicus Detection

It is necessary to define an unambiguous initial point for the purpose of automatic scanning.

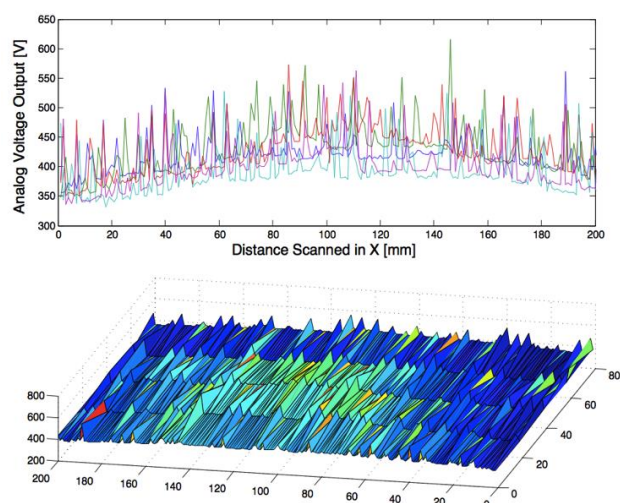


Fig. 15 2D (upper image) and 3D (lower image) plot of voltage output from IR sensor.

TABLE V
SCANNING PARAMETERS FOR SUBJECTS

	X_{1U} [cm]	X_{2U} [cm]	D_D [cm]	Total Distance [μm]	Time [s]
S1	4.5	16	0.3	205000	10.93
S2	3.2	4.6	0	078000	4.16
S3	7.7	41.9	0.5	226000	4.14
S4	1.8	18	-0.4	198000	10.65
S5	2.4	5	0.1	074000	3.95
S6	-12.1	3	-0.3	151000	8.05
S7	7	5.1	0.6	261000	13.92
S8	2.5	19.1	0.5	039000	2.18
S9	13.2	15.1	0.2	157000	8.37
S10	16.1	2.80	0.9	283000	15.10

The scan can be performed faster if the starting point is located close to the umbilicus area. Thus, the umbilicus was chosen to be the starting point of individual scan procedures. This setup resulted in a correct detection of the umbilicus in all cases with an accuracy of 95.9 % for the X -plane and 98.4 % for the Y -plane. The voltage values from the microcontroller were plotted using MATLAB R2014a.

Spikes shown in the upper image of Fig. 15 are most probably caused by noise. Here, noise represents the impact of the surrounding light information like neon lamps used in the laboratory. In order to eliminate this undesired information, averaging was performed. After replacing the acquired data vector field by its average over time, an asymptotic approximation can be obtained. In the resulting curve a voltage output between 380 and 480 for one sequence can be seen. It displays the umbilicus marker at its highest point in 3D plot in Fig. 15 representing the highest voltage output recorded. This was performed in all cases.

H. Umbilicus Detection Accuracy

The accuracy of the umbilicus detection method is shown in Table VI. The parameters X_S , X_E , Y_S and Y_E refer to the start and end points in X and Y -plane.

TABLE VI
RESULTS FOR AUTOMATIC SCANNING

	X_S [cm]	X_E [cm]	E_X [cm]	A_x [%]	Y_E [cm]	E_Y [cm]	E_Y [cm]	A_Y [%]
S1	17.3	21.2	0.9	95.9	18.2	21.7	0.5	97.6
S2	3.5	6.7	0.2	97.0	17.1	14.3	0.2	98.6
S3	8.1	4.8	0.3	93.7	22.5	25.6	0.1	99.6
S4	11.4	13.9	0.5	96.4	26.7	29.1	0.6	97.9
S5	13.2	17.3	1.1	93.6	10.3	6.4	0.9	85.9
S6	2.5	5.9	0.4	93.2	0.5	4.1	0.6	85.3
S7	1.7	5.1	0.4	92.1	8.2	10.9	0.3	97.2
S8	3.2	6.4	0.4	93.7	16.7	20.1	0.4	98.0
S9	10.3	15.1	0.2	98.6	31	34.2	0.2	99.4
S10	5.2	2.8	0.6	78.5	11	15.1	1.1	99.3
			0.5	95.9			0.49	98.4

Errors have been observed, 5 mm in X -plane and 4 mm in Y -plane. This error refers to the axial correlation between marker and skin. It is the distance of the position of the transducer to the center of the umbilicus marker. The accuracy resulting for the X -plane is 95.9 % and for the Y -plane is 98.4 %. These results show that the IR sensor approximates the umbilicus sufficiently high in order to perform abdominal scanning from this reference point.

This uncertainty of resolution in X -direction can be caused by light reflections of the skin during measurement or the fact, that human skin is close to the behavior of a black-body (Emissivity, 0.989 ± 0.01) [34] and therefore influences the axial accuracy. However, this approach can be replaced by a more accurate method in future if required. A resolution of 8000 pulses per rotation and a frequency of 50,000 Hz were set at this point. All sequences were started from the initial point of the probe with X_0 , Y_0 and Z_0 . The times, required for detecting the umbilicus, range from approximately 2 to 15 seconds, depending on the position of the subject and subsequently on the distance from the initial point to the umbilicus. The total distance in Table V refers to the distance, which is moved by the probe from the initial point to the umbilicus position. It can be obtained that currently only an unidirectional umbilicus detection method is used. The other distance measured is D_D and represents the distance moved over marker before detection. Due to the previously discussed accuracy errors of around 5 mm, the utilized method is subject to improvements.

IV. DISCUSSION

A linear robot was used to visualize the human cervix and has proved to provide advantages over manual ultrasonography. This system has been compared with two other robotic ultrasound systems. The first system compared is a 6 DOF industrial robot for automatic ultrasound system for lower limb examination [35]. The system mainly encloses a robotic arm for accurate manipulation of ultrasound probe, a scanning system for generating ultrasound images of the scanned area and a workstation that converts all position and torque data to commands and showed good results for the clinical evaluation of vessels. The performance evaluation of this system shows that a position accuracy of 0.62 ± 0.29 mm could be achieved in replay mode, which was 0.02 mm higher compared to teach mode results. The inter-distance accuracy for teach mode was measured 0.43 ± 0.32 mm, whereas for the replay mode, it was 0.42 ± 0.33 mm. Finally, the teach/replay repeatability calculated was 0.10 ± 0.22 mm.

The second system has been developed for automatic carotid artery examination. The systems main components are a master hand controller, a slave manipulator carrying the ultrasound probe and a computer control station. The operator remotely controls the probes position on the remote side. A resolution of 0.003 mm is reported here [36]. The linear robot tested in this paper allows a resolution up to 0.001 mm [37] and shows compared to both systems developed, the better performance.

Safety considerations were incorporated using initial point motions of the probe holder. However, this approach can get improved by the equipment of the holder using sensors in order to detect critical pressure.

Probe pressure must be applied evenly in order to get the accurate direction of the scan. Therefore, operator dependent scanning is less suitable than automatic for better image results concerning image quality. The pressure affects the echogenicity of the tissue and shortens the distance to the cervical structure. Moreover, image quality improvements were presented as low cost strategies in order to improve image resolution in ultrasound data. These include 3D image reconstruction using an already developed algorithm and the process of merging minimally shifted ultrasound image data. Furthermore, a precise panoramic scan of the abdomen was achieved using the linear robot in panorama scan modus offering a scanning length of 30 cm under constant pressure and with high movement resolution. This approach can be made autonomous by the usage of AI and control approaches allowing self-regulation by feedback information from sensors [38]. However, it is possible to consider the combination of the presented approaches with mobile applications [39] for procedure surveillance in terms of self-interventions and quality control for the institute carrying out the specific procedure [40]. A clinical pathway may be created for personalized abdominal scanning [41] and online analysis using open source solutions [42]. However, results concerning specifically umbilicus detection in this paper have already been presented at the International Conference on Telecommunications and Signal Processing (TSP) [28].

V. CONCLUSION

A stable and precise method enabling personalized ultrasound scanning and automatic umbilicus detection has been successfully developed. Test results show that the system is able to capture ultrasound images of phantoms and subjects with an accuracy of 95.9 % for the X -plane and 98.4 % for the Y -plane. Captured ultrasound images and positions can be used for the reconstruction of 2D and 3D panorama images for larger fields of view [43]. In order to improve the resolution of 2D panoramic view, it is recommended to use merging algorithms in combination with image registration and finite element modeling. However, it should be noted that this research does not aim at providing a ready to use device for clinical purposes since required guideline safety modalities were not considered at this point. Moreover, this study aims at improving existing ultrasound-based imaging methods by providing a motor-based scanning. Precise transducer probe movements allow accurate image capturing for subsequent processing of acquired data [44]. The analysis of ultrasound images containing abdominal structures allows a higher degree of operator independence. With this system, it is possible to perform automatic ultrasound cervix scanning with more comfort for the female patient focussing on biased cultural backgrounds as well as higher movement resolution and image quality. Furthermore, subject autonomy was achieved by defining acceptable pressure limits and the application of stop buttons for personal safety purposes [45]. This system will enable future operator independent task management [46], not only for obstetrics related diagnosis [47] but also for other internal organ imaging tools combined with E-Health platforms [48]. Therefore, it can be concluded that the presented system is a suitable low-cost approach for developing countries.

ACKNOWLEDGMENT

The author would like to appreciate material support from Universiti Teknologi Malaysia as well as structural support from Technische Universität Ilmenau. Moreover, appreciation goes to anonymous reviewers for their constructive comments.

REFERENCES

- [1] J. Ferlay, I. Soerjomataram, M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. Parkin, D. Forman, F. Bray, "Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11," *GLOBOCAN 2012 v1.0, International Agency for Research on Cancer*, 2013.
- [2] G. Lim, H. Yahaya, T. Lim, "The first report of the national cancer registry cancer incidence in Malaysia," *Kuala Lumpur: Ministry of Health*, pp. 57-58, 2003.
- [3] D. Parkin, F. Bray, J. Ferlay, P. Pisani, "Global cancer statistics," *CA: a cancer journal for clinicians*, pp.74-108, 2005.
- [4] C. Pahl, E. Supriyanto, "Personalized Cervix Ultrasound Scan Based On Robotic Arm," *PSRC-Planetary Scientific Research Center Proceedings*, 2012.
- [5] C. Pahl, E. Supriyanto, N. Mahmood, J. Yunus, "Cervix Detection Using Squared Error Subtraction," *Modelling Symposium (AMS), 2012 Sixth Asia. IEEE*, pp. 121-125, 2012.
- [6] A. I. B. Jeffree, C. Pahl, H. N. Abduljabbar, I. Ramli, N. I. B. Aziz, Y. M. Myint, E. Supriyanto, "Cervical Segmentation in Ultrasound Images using Level-set Algorithm," *WSEAS International Conference on Biomedicine and Health Engineering*, 2013.
- [7] E. Supriyanto, J. Z. E. Soh, S. M. Nooh, "Myofascial Pain Syndrome, Trigger Point Detection based on Ultrasound Image," *Proceedings of 15th WSEAS International Conference on Computers*, pp.178-183, 2011.
- [8] T. Wright Jr, M. Schiffman, D. Solomon, J. T. Cox, F. Garcia, S. Goldie, D. Saslow, "Interim guidance for the use of human papillomavirus DNA testing as an adjunct to cervical cytology for screening," *Obstetrics & Gynecology*, 2004, pp. 304-309.
- [9] G. Brix, U. Lechel, G. Glatting, S. Ziegler, W. M. nzing, S. Müller, T. Beyer, "Radiation exposure of patients undergoing whole-body dual-modality 18F-FDG PET/CT examinations," *Journal of Nuclear Medicine*, pp. 608-613, 2005.
- [10] D. Frush, L. Donnelly, N. Rosen, "Computed tomography and radiation risks: what pediatric health care providers should know," *Pediatrics*, pp. 951-957, 2003.
- [11] T. Slovis, "Children, computed tomography radiation dose, and the As Low As Reasonably Achievable (ALARA) concept," *Pediatrics*, pp.971-972, 2003.
- [12] C. Pahl, A. Jeffree, Y. Myint, E. Supriyanto, "Cervix Segmentation in Ultrasound Data Based on Numerical Technique," *International Journal of Biology and Biomedical Engineering*, Issue 4, Volume 7, pp. 157-163, 2013.
- [13] A. Fenster, D. Downey, H. Cardinal, "Three-dimensional ultrasound imaging," *Physics in medicine and biology*, 2001.
- [14] R. Prager, U. Ijaz, A. Gee, G. Treece, "Three-dimensional ultrasound imaging," *Proceedings of the Institution of Mechanical Engineers, Journal of Engineering in Medicine*, pp.193-22, 2010.
- [15] J. Szilard, "An improved three-dimensional display system," *Ultrasonics*, pp. 273-276, 1974.
- [16] K. Pandey, A. Pradhan, A. Agarwal, A. Bhagoliwal, N. Agarwal, "Fluorescence Spectroscopy: A New Approach in Cervical Cancer," *The Journal of Obstetrics and Gynecology of India*, pp. 1-5, 2012.
- [17] J. Henderson, M. Saraiya, G. Martinez, C. Harper, G. Sawaya, "Changes to cervical cancer prevention guidelines: Effects on screening among US women ages 15-29," *Preventive Medicine*, 2012.
- [18] C. Ravin, S. George, "Radiological Society of North America," *Radiology*, pp. 9-10, 2012.
- [19] J. Rouyer, P. Lasaygues, S. Mensah, "Novel Ultrasound Tomograph for Anatomical Inspection," *Acoustical Imaging*, pp. 3-10, 2012.
- [20] J. Charboneau, D. Levine, "Diagnostic ultrasound," *JAMA*, 2012.
- [21] A. Wong, J. Scharcanski, "Monte Carlo despeckling of transrectal ultrasound images of the prostate," *Digital Signal Processing*, 2012.
- [22] R. Rockland, H. Kimmel, R. Carpinelli, L. Hirsch, L. Burr-Alexander, "Medical Robotics in K-12 Education," *Robots in K-12 Education*, p. 120, 2012.
- [23] P. N. Wells, "The safety of diagnostic ultrasound: report of a British Institute of Radiology Working Group," *British Inst. of Radiology*, 1987.
- [24] W. R. Hedrick, D. L. Hykes, "Biological Effects of Ultrasound I. Specification of Intensity," *Journal of Diagnostic Medical Sonography*, 7(4), pp. 188-193, 1991.
- [25] P. N. Wells, "Ultrasonic imaging of the human body," *Reports on progress in physics*, 62(5), p. 671, 1999.
- [26] E. Kimmel, "Cavitation bioeffects," *Critical Reviews™ in Biomedical Engineering*, 34(2), 2006.
- [27] C. A. Speed, "Therapeutic ultrasound in soft tissue lesions," *Rheumatology*, 40(12), pp. 1331-1336, 2001.
- [28] C. Pahl, E. Supriyanto, "Infrared based clinical landmark determination for ultrasound image acquisition," *In Telecommunications and Signal Processing (TSP), 2015 38th International Conference on*, IEEE, pp. 236-240, 2015.
- [29] C. Pahl, E. Supriyanto, "Design of automatic transabdominal ultrasound imaging system," *In Methods and Models in Automation and Robotics (MMAR), 2015 20th International Conference on*, IEEE, pp. 435-440, 2015.
- [30] C. Pahl, T.Y. Oon, E. Supriyanto, "Fuzzy system based face tracking for head movement control in progressive health care," *In Methods and Models in Automation and Robotics (MMAR), 2015 20th International Conference on*, IEEE, pp. 880-885, 2015.
- [31] C. Pahl, E. Supriyanto, "Guide to fuzzy logic based transesophageal echocardiography motorized probe control," *In Methods and Models in Automation and Robotics (MMAR), 2015 20th International Conference on*, IEEE, pp. 892-896, 2015.
- [32] W. E. Woodson, B. Tillman, P. Tillman, "Human factors design handbook: information and guidelines for the design of systems, facilities, equipment, and products for human use," 1992.
- [33] K. W. Lai, "Foetal abnormalities detection based on three dimensional nuchal translucency," *Doctoral dissertation*, Universiti Teknologi Malaysia, 2012.
- [34] F. De Oliveira, S. Moreau, C. Gehin, A. Dittmar, "Infrared imaging analysis for thermal comfort assessment," *Engineering in Medicine and Biology Society, 29th Annual International Conference of the IEEE*, pp. 3373-3376, 2007.
- [35] M. A. Janvier, L. G. Durand, M. H. R. Cardinal, I. Renaud, B. Chayer, P. Bigras, ... & G. Cloutier, "Performance evaluation of a medical robotic 3D-ultrasound imaging system," *Medical image analysis*, 12(3), pp. 275-290, 2008.
- [36] W. H. Zhu, S. E. Salcudean, S. Bachmann, P. Abolmaesumi, "Motion/force/image control of a diagnostic ultrasound robot," *In Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, Vol. 2, pp. 1580-1585, 2000.
- [37] C. Pahl, "Novel Method for Autonomous Ultrasound Cervix Scanning," *Universiti Teknologi Malaysia*, 2012.
- [38] F. W. Lewis, S. Jagannathan, A. Yesildirak, "Neural network control of robot manipulators and non-linear systems," CRC Press, 1998.
- [39] C. Pahl, M. Zare, L. Bordey, M. H. Satria, E. Supriyanto, Y. W. Hau, "Smartphone Applications in the Support of Weight Reduction and Goal Maintenance," *Journal of Soft Computing and Decision Support Systems*, 1(1), pp. 12-16, 2014.
- [40] M. Zare, C. Pahl, M. Nilashi, N. Salim, O. Ibrahim, "A Review of Semantic Similarity Measures in Biomedical Domain Using SNOMED-CT," *Journal of Soft Computing and Decision Support Systems*, 2(6), pp. 1-13, 2015.
- [41] C. Pahl, E. Supriyanto, V. Detschew, "Evaluation of UML modeling tools for clinical pathways used in health information systems." *In Information Systems and Technologies (CISTI), 2015 10th Iberian Conference on*, IEEE, pp. 1-6, 2015.
- [42] C. Pahl, N. B. Mohammad, E. Supriyanto, "Analysis on the electroretinography response for flickering and current stimulations." *In Telecommunications and Signal Processing (TSP), 2015 38th International Conference on*, IEEE, pp. 482-485, 2015.
- [43] J. Alcázar, M. Aubá, B. Olartecoechea, "Three-dimensional ultrasound in gynecological clinical practice," *Reports in Medical Imaging*, pp.1-13, 2012.
- [44] C. Pahl, J. Lam, H. N. Abduljabbar, E. Supriyanto, Y. M. Myint, "Performance Observation of Gabor Filter for Wall Thickness Measurement of Human Colon Based on Ultrasound Image," *International Journal of Information and Electronics Engineering*, 4(2), p. 171, 2014.
- [45] K. Baker, V. Robertson, F. Duck, "A review of therapeutic ultrasound: biophysical effects," *Physical Therapy*, pp.1351-1358, 2001.
- [46] C. Pahl, H.A.B Mazle, E. Supriyanto, "Design of virtual instrument for automatic temperature visualization in magnetic fields using LabVIEW in combination with fiber-optical temperature measurement," *In Automation, Robotics and Applications*, pp. 393-399, 2015.
- [47] C. Pahl, M. Zare, M. Nilashi, M.A. de Faria Borges, D. Weingaertner, V. Detschew, O. Ibrahim, E. Supriyanto, "Role of OpenEHR as an open source solution for the regional modelling of patient data in obstetrics." *Journal of biomedical informatics*, 55, pp. 174-187, 2015.
- [48] C. Pahl, et al., "Identification of Quality Parameters for an E-Health Platform in the Federal State of Thuringia in Germany." *Journal of Soft Computing and Decision Support Systems* 1.1, pp. 17-23, 2014.

DNA Patterns Localization Using a Dedicated Dot-Plot and Image Analysis

Petre G. Pop

Abstract— DNA repeats are believed to play significant roles in genome evolution and diseases. Many of the methods for finding repeated sequences are part of the digital signal processing (DSP) field and most of these methods use distances, similarities and consensus sequences to generate candidate sequences. This paper presents results obtained using a dedicated numerical representation with a mapping algorithm (using DNA distances and consensus types) and a custom dot-plot analysis (using similarities to represent DNA patterns) combined with image processing techniques, to visual isolate the position of DNA patterns with different lengths. The final images that best put in evidence the presence of repeated sequences were obtained using weighted cosine cross-correlation, Jukes-Cantor distance and Motyka similarity.

Keywords— DNA repeats, DNA numerical representations, DNA distances, dot-plot analysis, numerical similarities.

I. INTRODUCTION

The existence of DNA repeats is a fundamental characteristic of all biological genomes. A repeat is the simplest form of regularity and the detection of repeats (pattern, length, position, repeat number) is important in biology and medicine as it can be used for phylogenetic studies and disease diagnosis. A major difficulty in identification of DNA repeats is caused by the fact that the repeat units can be of different lengths and either in tandem or dispersed or exact or imperfect. Although research in this domain is made for many years, the problem is not yet considered solved. The main approaches for repeats detection are reviewed in [1].

DNA or nucleotides sequences are represented by sequences of the characters A, T, C and G, corresponding to nucleotides A (adenine), T (thymine), C (cytosine) and G (guanine). DNA repeats are sequences that are present in more than one copy. An exact repeat is a short string of nucleotides repeated contiguously at least twice. An approximate repeat is a string of nucleotides repeated consecutively with differences between the instances (mainly due to different types of mutations). Repeats, whose copies are distant in the genome, are named distant or dispersed repeats. Among those, we can distinguish micro-satellites, mini-satellites, and satellites, depending on the

length of their repeated unit [1].

The interest in finding DNA repeats may be theoretical, technical or medical, as follows:

- Theoretical interest: related to their role in the structure and evolution of the genome.
- Technical interest: repeats can be used as polymorphic markers, either to trace the propagation of genetic traits in populations or as genetic identifiers in forensic studies.
- Medical interest: the presence of specific types of tandem repeats has been associated to different severe diseases (e.g. Huntington's disease, myotonic dystrophy). In healthy individuals, the tandem repeat size varies around a few tens of copies, while in affected individuals the number of copies at the same locus reaches hundreds or even a thousand in some cases.

The centromere of most complex eukaryotic chromosomes is a specialized locus comprised of repetitive DNA that is responsible for chromosome segregation at mitosis and meiosis. Alpha satellite DNA is composed of a tandem array of repeat units and has been identified at every human centromere. There are two major types of alpha satellite, higher-order and monomeric [2]. Higher-order alpha satellite is the predominant type in the genome and made up of ~171 bp (base pairs) monomers organized in arrays of multimeric repeat units that are highly homogeneous. Monomeric alpha satellite lies at the edges of higher-order arrays and lacks any higher-order periodicity; its monomers are only on average ~70% identical to each other [2]. Our present research was focused on determining these alpha satellites DNA.

The numerical representation of DNA sequences becomes very important as almost all DSP techniques require two parts: mapping the symbolic sequence (letters corresponding to nucleotides) into a numeric form and calculating a kind of transform of the resulting numeric sequence [3]. Most of the numerical representations associate a single numerical value to one position in the sequence using numerical values associated to each nucleotide and, finally, reflect the presence or the absence of a certain nucleotide in a specific position (e.g. indicator sequences) [3]. Another approach could be to include information about the number and type of consecutive nucleotides and to generate only one numerical value for each DNA subsequence which may be associated with a

repeat. In this regard, we have introduced a representation which considers the length of the expected repeats and the number of possible mismatches, based on polynomial-like representation [4]. This representation needs a mapping algorithm which uses distances and then evaluates a consensus sequence to generate candidates.

This paper presents results obtained using this dedicated numerical representation with associated mapping algorithm (that uses DNA distances and consensus types) combined with a custom dot-plot analysis (involving similarities to represent DNA patterns) to generate in-memory dot-plot images which are then processed to highlight the position of repeats with different lengths in DNA sequences.

II. NUMERICAL REPRESENTATION AND MAPPING ALGORITHM

In a previous work [4] we have proposed a DNA numerical representation and a mapping algorithm, which includes both the length of the DNA repeats and the number of mismatches due to point mutations. For a DNA sequence of length L , a numerical value is associated using a polynomial-like representation:

$$V = \sum_{k=0}^{L-1} V_{\alpha_k} 10^k, \quad \alpha \in \{A, G, C, T\}, \quad (1)$$

where V_{α} is the value of a single nucleotide. These coefficients should be different natural numbers such that the resulting numerical value is unique for a given DNA subsequence.

However, in the case of two DNA sequences with a high degree of similarity (which differ, for instance, by a single nucleotide) representation (1) will produce two very different values. Therefore we need an algorithm that allows finding similar sequences (considering possible mismatches), determine the associate consensus sequence and then generates a single numerical value for these similar sequences, using (1) for calculated consensus sequence.

To pass from initial DNA sequence to a single, final sequence of numerical values, we need both a type of distance and a type of consensus:

- The distance should measure the number of mismatches between DNA subsequences of the same length; if two subsequences are identical, this distance should be zero.
- Given a number DNA of subsequences of the same length, the consensus sequence is a sequence pattern derived from multiple, similar DNA sequences that represents the nucleotide most likely to occur at each position in analyzed sequences.

The proposed mapping algorithm has the following steps:

- Step-1: Consider all successive DNA subsequences

of same length L ;

- Step-2: Determine all the positions (and the associated subsequences) in the original DNA sequence for which the distance (against a subsequence from Step-1) is less or equal to the prefixed maximum mismatches allowed number M_m ;
- Step-3: Determine the consensus sequence for all (similar) subsequences from Step-2; Calculate the distance between the consensus sequence and each associated subsequence; those subsequences whose distance is greater than M_m must be reassign; Determine again the consensus sequence for remaining sequences.
- Step-4: Using (1), compute the numerical value for consensus sequence and assign this value to all starting positions of subsequences determined in Step-2.

As output, the algorithm generates a single sequence of numbers; each number is associated to a unique subsequence of length L (possible a repeat unit).

An important property of this mapping algorithm is that if the L value is a prime factor of repeated sequence length then the entire repeated sequence will be emphasized. This allows a significant reduction of the computational effort in case of long repeats. On the other hand, the final numerical values contain information about the structure of associated consensus sequence, which can be used to specify the structure of detected repeated sequences.

Finding similar sequences in Step-2 and determination of consensus sequence requires evaluating the distance between two DNA subsequences. In our experiments we used Hamming distance and Jukes-Cantor distance (an evolutionary distance).

The Hamming distance determines the number of different nucleotides between two equal length DNA sequences.

Let x and y , two DNA sequences of length n . The Jukes-Cantor distance between DNA sequences is defined by [5]:

$$d_{JC}(x, y) = -\frac{3}{4} \ln\left(1 - \frac{4}{3} \frac{\sum 1_{x_i \neq y_i}}{n}\right). \quad (2)$$

We also used these distances in Step-3 to determine the distance between the consensus sequence and candidate DNA subsequences.

Also, in Step-3 we determine the consensus sequence for all similar DNA subsequences determined in Step-2 and, on this basis, we calculate the associated numerical value (using (1)).

We used the following types of consensus:

- Most frequently occurring nucleotide (in each position or column), even if it is not the majority.
- Majority with fixed cutoff: use the fraction of nucleotides in a position to establish majority for that position, provided that the fraction is greater than the cutoff parameter.

- Majority with global appearing frequency cutoff: same as previous case but the cutoff for each nucleotide is computed as the appearing frequency in the original sequence.
- Majority with local appearing frequency cutoff: same as previous case but the cutoff for each nucleotide is computed as the appearing frequency in the analyzed similar subsequences.

For last three consensus types, if there is no nucleotide that exceeds the threshold we consider that we have no valid consensus sequence and those subsequences must be reassigned. In case that more than one nucleotide is calculated to have the same confidence, and this exceeds the consensus threshold, the nucleotides are assigned in descending order of global appearing frequency precedence [5].

III. NUMERICAL SIMILARITIES AND DOT-PLOT ANALYSIS

In case of DNA sequences, dot plots are two-dimensional representations where each axis represents a sequence (possibly the same) and the plot itself shows a comparison of analyzed sequences by a calculated score for each position of the sequences. Most of time dot plots are used to determine regions of similarity within a single DNA sequence (i.e. repeats) or between two different sequences.

Some important characteristics of patterns appearing in DNA dot plots are [6]:

- Parallels to the main diagonal indicate repeated regions on different parts of the analyzed sequences (Fig. 1-a).
- Blocks of parallel lines indicate tandem repeats in both sequences and the distance between the lines equals the distance of the repeats (Fig. 1-b).

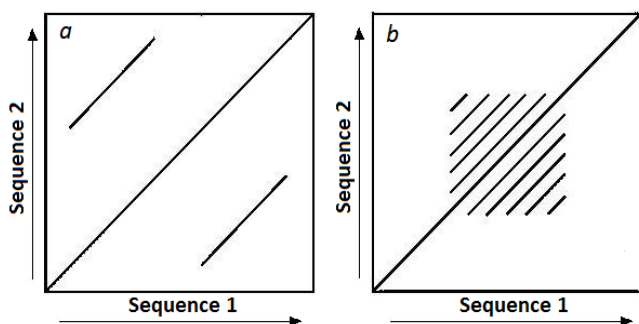


Fig. 1. Characteristic patterns appearing in dot plots.

If the length of sequences is large, windows are used to perform the analysis. In this case, if a window of fixed size on one sequence (one axis) match another window (possibly of different size) on the other sequence, a dot is drawn at the plot.

To evaluate the results obtained in the experiments we need a customized dot-plot analysis as:

- The analyzed sequence is not symbolic but a numerical one (the output of mapping algorithm).
- In most cases the length of analyzed sequence far

exceeds the number of points represented on each axis. For this reason, the dot plot analysis is done using windows.

- Due to the large number of values resulted from mapping and the different resolution on each axis, we need to determine the degree of similarity between windows of different lengths to decide if a dot will be plot or not.

To determine the degree of similarity between two numerical sequences of length m and n , a similarity coefficient is calculated for two equal sequences of length n , $(m-n)$ times, with a sliding window, then determine the average value.

In our experiments we used several similarities: cross-correlation, Motyka, Bray-Curtis, Kulczynski-1, Kulczynski-2, Ruzicka, Roberts, and Baroni [7]. The best results were obtained using: Motyka, Bray-Curtis and Kulczynski-2 similarities.

The Motyka similarity is defined by [7]:

$$s_{Mot}(x, y) = \frac{\sum \min(x_i, y_i)}{\sum (x_i + y_i)}. \quad (3)$$

Bray-Curtis similarity is defined by [7]:

$$\frac{2}{n(x+y)} \sum \min(x_i, y_i). \quad (4)$$

The definition of Kulczynski similarity 2 is [7]:

$$\frac{n}{2} \left(\frac{1}{x} + \frac{1}{y} \right) \sum \min(x_i, y_i). \quad (5)$$

The average similarity coefficient is scaled in the interval $(0, 1)$ and based on its values, a dot with a grey level value between 0 and 255 will be plot.

Finally, we obtain a gray level image to which we applied a threshold based on the mean (μ) and image variance (σ), using the formula [8]:

$$T_h = k_1 \cdot \mu + k_2 \cdot \sigma. \quad (6)$$

The constants k_1 and k_2 are image type dependent and, after some experiments, we used $k_1 = 1$ and $k_2 = 1.75$.

Following this procedure we obtain a graphical representation for each combination of parameters L (length of the searched repeated sequence), M_m (maximum number of allowed mismatches). The quality of representation varies depending on the number of similar subsequences of the same length that are found in the original sequence.

The graphical representation allows locating repeated sequences of a certain length using position of segments parallel to the main diagonal, position represented on each axis. Using a prime factor of repeated sequence length for parameter L allows highlighting of patterns in the dot-plot

image. In this case it is necessary to run the application using multiple combinations for parameters values, followed by analysis of the images obtained to determine the length of repeated sequences. Therefore, a priori information is needed about the domain values for the length of repeated sequences.

To determine the length of repeated sequences, we used another approach that allows calculation of dot-plot images, in memory, for several values of the parameter L (length of repeated sequences). Then extract some features from each image that allows the localization of repeated sequences, features that are represented in the form of intensities on a single line, for each value of L .

For this purpose we extracted next features from each in-memory dot-plot image:

- The average density for each column of the image:

$$D_i = \frac{\sum_{j=1}^N I(i, j)}{N - n^*}, i = 1, N, \quad (7)$$

where, n^* is the number of column not null values.

- Lagged autocorrelation of image values from each column of the image:

$$r_{k,i} = \frac{\sum_{j=1}^{N-k} (I(i, j) - \bar{I}_i)(I(i+k, j) - \bar{I}_i)}{\sum_{j=1}^N (I(i, j) - \bar{I}_i)^2}, i = 1, N. \quad (8)$$

- Cross-correlation between two adjacent columns of the image, expressed by the Pearson similarity coefficient:

$$r_i = \frac{\sum_{j=1}^N (I(i, j) - \bar{I}_i)(I(i+1, j) - \bar{I}_{i+1})}{\sqrt{\sum_{j=1}^N (I(i, j) - \bar{I}_i)^2 \sum_{j=1}^N (I(i+1, j) - \bar{I}_{i+1})^2}}, i = 1, N-1. \quad (9)$$

- Cross-correlation between two adjacent columns of the image, expressed by the cosine similarity coefficient:

$$s_i = \frac{\sum_{j=1}^N I(i, j)I(i+1, j)}{\sum_{j=1}^N I(i, j)^2 \sum_{j=1}^N I(i+1, j)^2}, i = 1, N-1. \quad (10)$$

Following this procedure we can represent results for a set of values of the L parameter (length of searched repeated sequences) as lines (or segmented lines) and the intensities and positions of these lines give information about the presence of repeated sequences and their positions. Our goal was to obtain a synthetic image for a

range of values (as widely as possible) of the L parameter so we can easily appreciate the presence of repeated sequences of a certain length and their position.

IV. EXPERIMENTS AND RESULTS

Our case study was the high order repeats in AC010523 from Homo sapiens chromosome 19 (GenBank) and in AC136363 from human chromosome 17 (GenBank) which contain both higher-order and monomeric DNA alpha-satellite [9], of approximately 171 bp (base pairs), arranged in tandem, in a head-to-tail fashion. High-order repeats were identified in the front domain while in the back and central domain, alpha satellite monomers were found [10] (Fig. 4, Fig. 5). Numerical representation based on (1) and associated mapping algorithms were used to obtain the associated numerical sequence.

Several experiments were performed using several combinations of parameters L and M_m , combinations of Hamming and Jukes-Cantor distances (in Step-2 and Step-3 of mapping algorithm) and different similarities (Motyka,

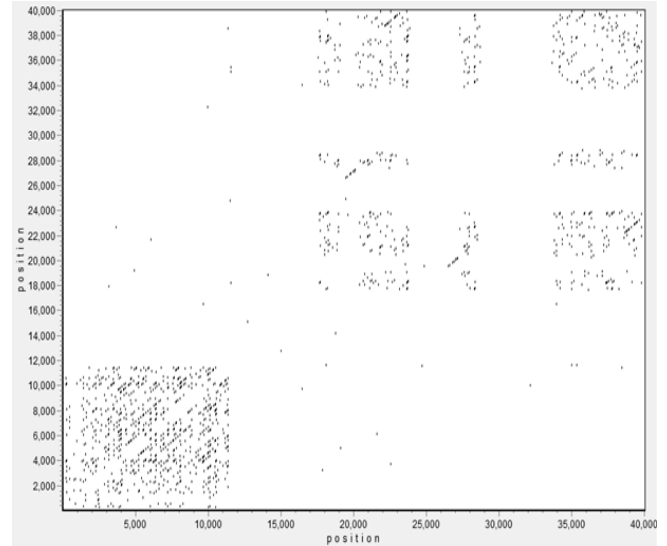


Fig. 2. Lower quality dot-plot image, with $L = 7$, $M_m = 2$, for AC010523

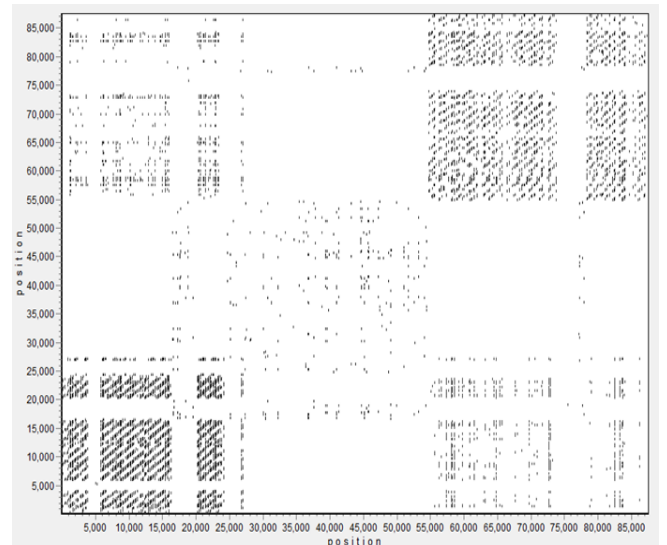


Fig. 3. Lower quality dot-plot image, with $L = 7$, $M_m = 2$, for AC136363

Bray-Curtis, Kulczynsky-2) for in-memory dot plot image construction, as they were introduced in section III.

For consensus sequence evaluation we used majority with local appearing frequency cutoff (use the fraction of nucleotides in a position to establish majority for that position, provided that the fraction is greater than a cutoff, computed as the appearing frequency in the analyzed similar DNA subsequences) [6].

Figures 2, 3, 4, 5 show dot plots results of different qualities using different values for L parameter. As one can see, the value of $L = 19$ gives better results compared to $L = 7$ as 19 is a prime factor of the length of the satellites (~171 bp). As mentioned previously, we need more pictures like this, to be able to tell what kind of patterns are present in the analyzed DNA sequence.

The following figures (Fig. 6-13) shows the best results obtained with the second approach using parameter values for L between 5 and 25 (represented on vertical axis) for

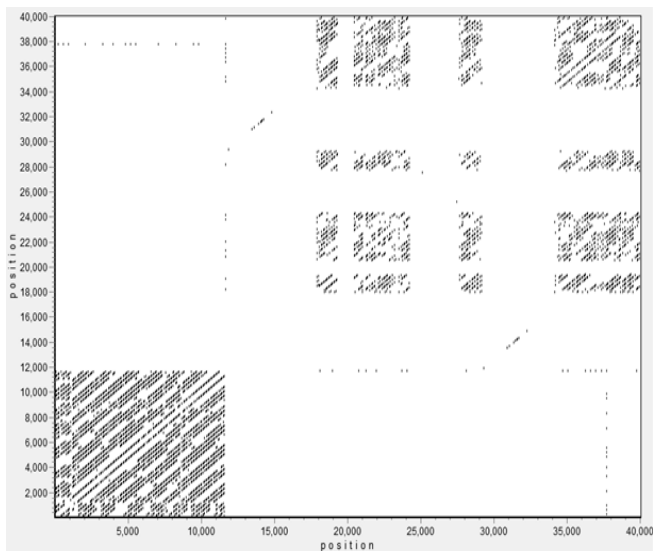


Fig. 4. Higher quality dot-plot image, with $L = 19$, $M_m = 4$, for AC010523

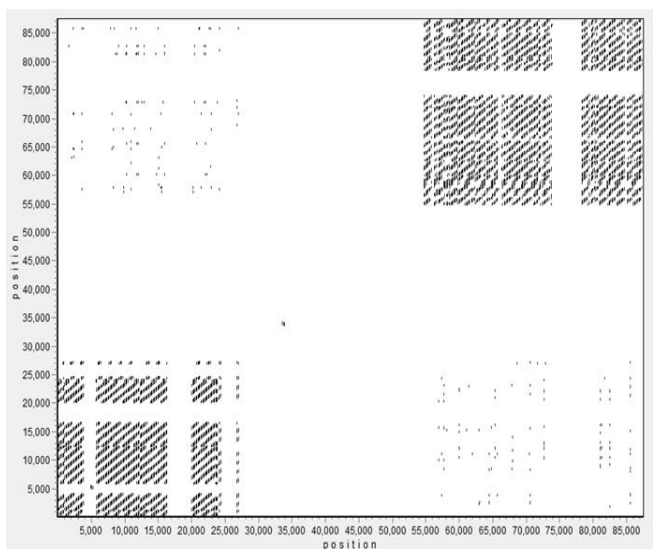


Fig. 5. Higher quality dot-plot image, with $L = 19$, $M_m = 4$, for AC136363

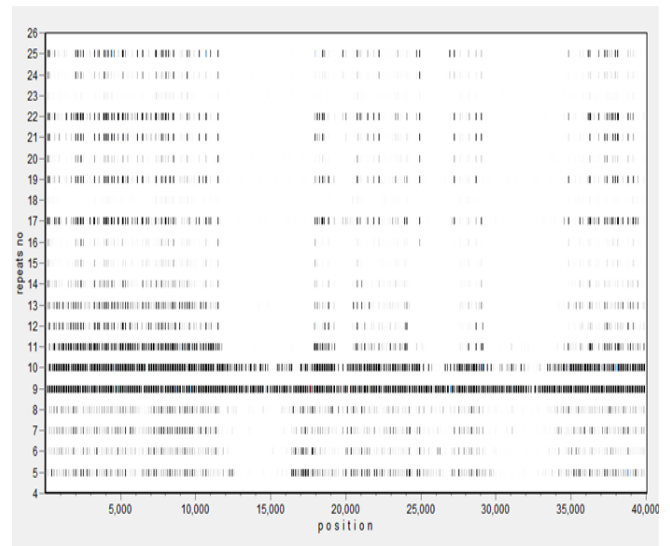


Fig. 6. Results obtained using average density, Hamming-Hamming distances, Motyka similarity (AC010523).

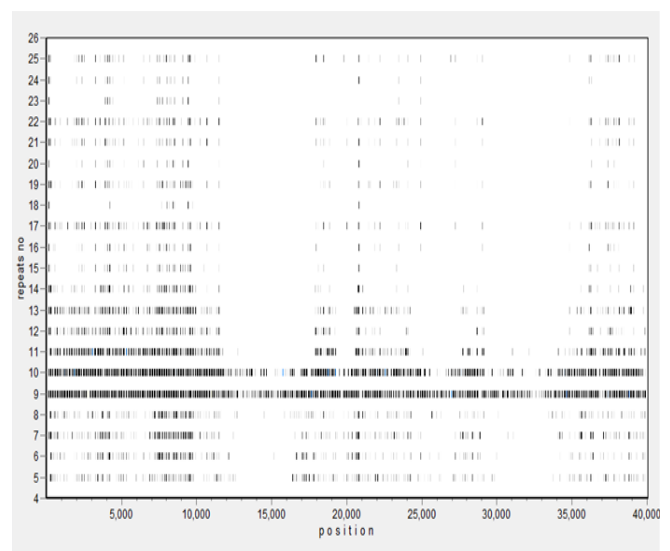


Fig. 7. Results obtained using auto-correlation, Hamming-Jukes Cantor distances, Motyka similarity (AC010523).

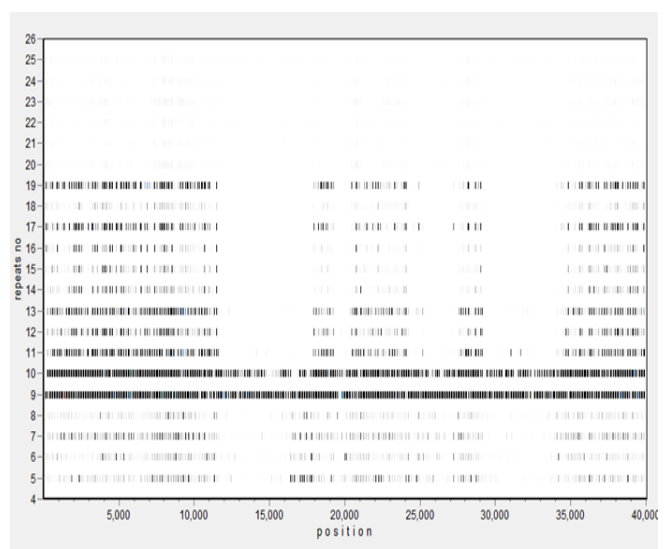


Fig. 8. Results obtained using auto-correlation, Jukes Cantor-Hamming distances, Kulczynski-2 similarity (AC010523).

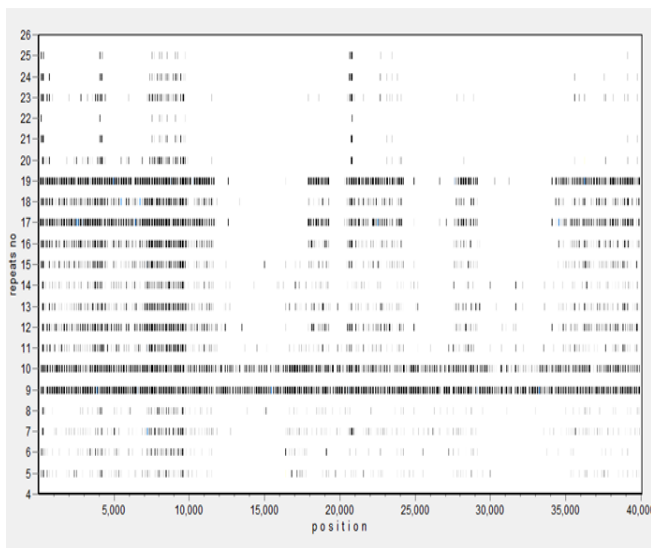


Fig. 9. Results obtained using auto-correlation, Jukes Cantor-Jukes Cantor distances, Kuczynski-2 similarity (AC010523).

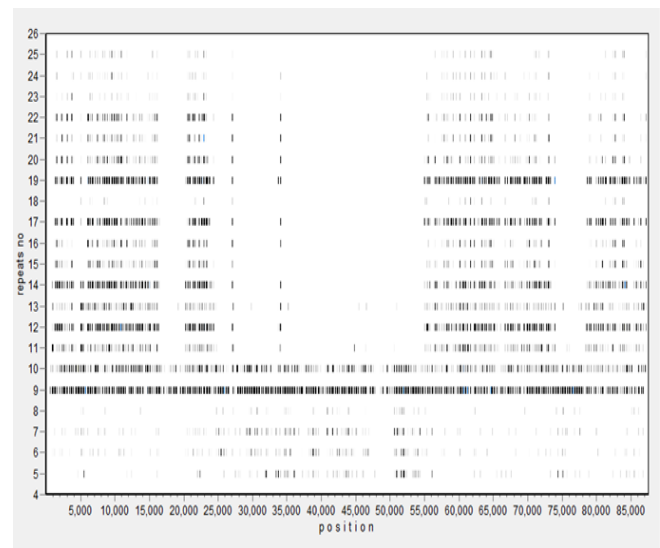


Fig. 12. Results obtained using auto-correlation, Jukes Cantor-Hamming distances, Bray Curtis similarity (AC136363).

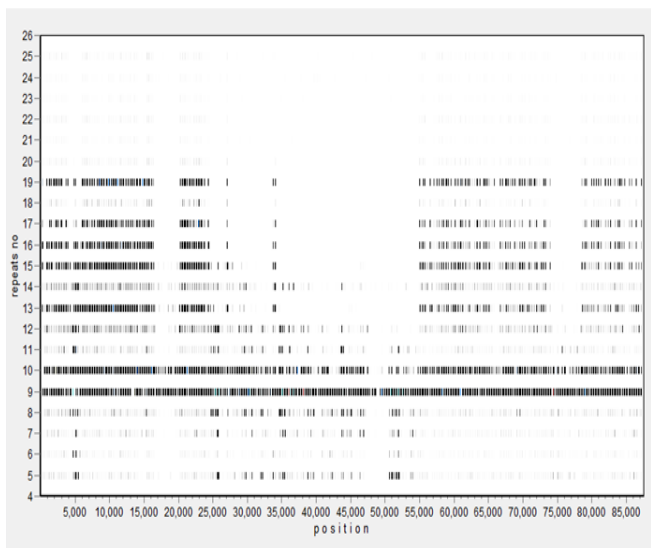


Fig. 10. Results obtained using average density, Hamming-Hamming distances, Kuczynski-2 similarity (AC136363).

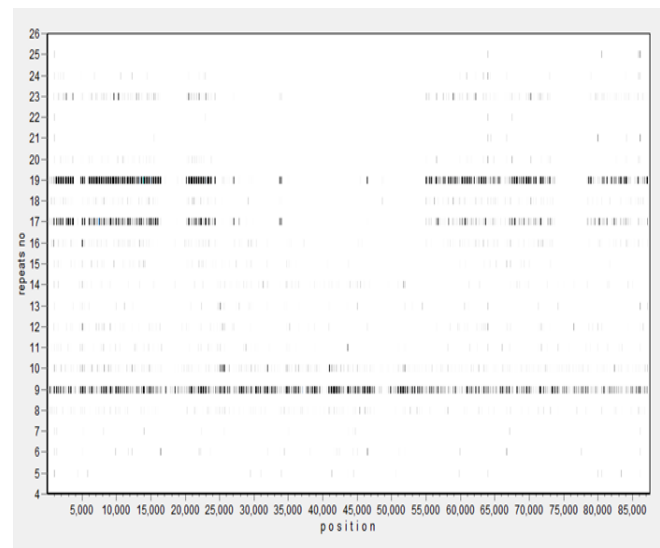


Fig. 13. Results obtained using cosine cross-correlation, Jukes Cantor-Jukes Cantor distances, Motyka similarity (AC136363).

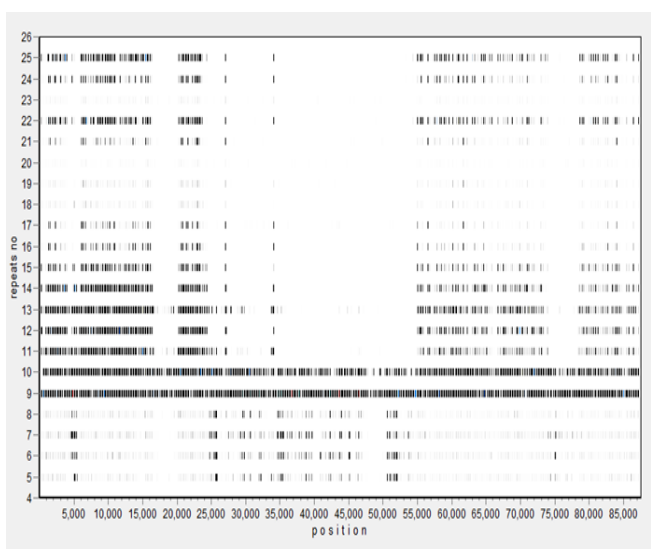


Fig. 11. Results obtained using average density, Hamming-Jukes Cantor distances, Kuczynski-2 similarity (AC136363).

our case studies (AC010523 and AC136363). As a final feature extracted from the in-memory dot-plot images we used: the average density, autocorrelation (with the target repeat length as lag) and cross-correlation (weighted Pearson and cosine similarity with the rate of common not null values), as they were introduced in section III. The images represent the best results for each type of distances combination in Step-2 and Step-3 from the mapping algorithm.

Image analysis presented in previous figures allows us to make the following statements.

When using Hamming-Hamming distances combination: (Fig. 6, Fig. 10):

- Average density allows best results combined with Motyka or Kulczynski-2 similarity.
- Values for $L = 9$, $L = 19$ are more visible, both in the front domain and in central and back domain; this can be explained by the fact that both values are the prime factors of 171 (DNA satellite

length).

- Besides the values for $L = 9$, $L = 19$, are also highlighted values for $L = 10$, $L = 17$ but at a lower intensity.

When using Hamming-Jukes Cantor distances combination: (Fig. 7, Fig. 11):

- Average density and autocorrelation allow best results combined with Motyka or Kulczynski-2 similarity.
- Values for $L = 9$, $L = 19$ are more visible, both in the front domain and in central and back domain.
- Other values, such as for $L = 10$, $L = 17$, even if they are prime factors of 170 (close to 171), are not highlighted.

When using Jukes Cantor-Hamming distances combination: (Fig. 8, Fig. 12):

- Auto-correlation allows best results combined with Bray-Curtis or Kulczynski-2 similarity.
- Values for $L = 9$, $L = 19$ are quite visible, both in the front domain and in central and back domain.
- Values for $L = 10$, $L = 17$, are also highlighted but at a lower intensity, as they are prime factors of 170 (close to 171, the DNA satellite length).

When using Jukes Cantor-Jukes Cantor distances combination: (Fig. 9, Fig. 13):

- Auto-correlation and cosine cross-correlation allow best results combined with Motyka or Kulczynski-2 similarity.
- Values for $L = 9$, $L = 19$ are more visible, both in the front domain and in central and back domain.
- Values for $L = 10$, $L = 17$, are also highlighted but at a lower intensity.
- Other values are less visible (in particular in Fig. 13).

Most images also highlights DNA satellite values in the front domain and the back domain for other values of L parameter. This suggests that there are repeated sequences of higher lengths in those areas.

If using weighted cosine similarity (Fig. 13), the values of the parameter L that are closer to the divisors of length of DNA satellite, are favored particularly those areas with a higher similarity patterns (front domain) while the other values parameter L and remaining areas are less highlighted.

V. CONCLUSION

An original DNA sequence representation (which includes information about repeats length and the number of mismatches) and a mapping algorithm are used to provide a single associated numerical sequence which is then used for a dedicated in-memory dot-plot representations of DNA patterns. These images can be used to extract some features allowing a suggestive graphical representation of the

position and length of repeated sequences. These final images provide visual and numerical information about the length of repeats and their approximate position.

We investigated the effect of several features, distances and similarities on final images. The images that best put in evidence the presence of repeated sequences were obtained using weighted cosine cross-correlation, Jukes-Cantor distance and Motyka similarity.

While final images are based on numerical features, assessment of these images is visual and, consequently, the assessment results (the presence and position of a certain length patterns based on identification of line segments more intense) are approximate. But this information can be very useful in case of long initial sequences for the use of more accurate methods is cumbersome and time consuming (because they generate very long list of candidate sequences, with associated information: score, position, difficult to interpret). This visual information can be used to dramatically narrow the search domain for more accurate methods. Looking ahead, we intend to try to automatic determine the position of those line segments whose intensity exceeds a certain threshold and extracting pattern structure information, stored in the numerical value (associated with a consensus sequence), which is represented in the in-memory images.

REFERENCES

- [1] K. G. Lim, C. K. Kwok, L. Y. Hsu, A. Wirawan, "Review of tandem repeat search tools: a systematic approach to evaluating algorithmic performance," *Brief Bioinform.*, vol. 14, no. 1, pp. 67-81, Jan. 2013.
- [2] M. K. Rudd, G. A. Wray, H. F. Willard, "The evolutionary dynamics of a-satellite," *Genome Res*, 16:88-96, 2006.
- [3] J. V. Lorenzo-Ginori, A. Rodriguez-Fuentes, R. G. Abalo, R. S. Rodrigues, "Digital signal processing in the analysis of genomic sequences," *Curr. Bioinformatics*, vol. 4, pp. 28-40, 2009.
- [4] P. G. Pop, A. Voinea, "Numerical Representations Involved in DNA Repeats Detection Using Spectral Analysis," *Studies in Informatics and Control*, vol. 20, no. 2, pp. 163-180, 2011.
- [5] D. Yankov, E. Keogh, S. Lonardi, "Dot plots for time series analysis," *Proc. 17th IEEE Int. Conf. on Tools with Artificial Intelligence*, pp. 159-168, 2005.
- [6] C. Sung-Hyuk, "Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions," *Int. J. of Mathematical Models and Methods in Applied Sciences*, vol. 1, no. 2, pp. 300-307, 2007.
- [7] T. D. Schneider, "Consensus Sequence Zen," *Applied Bioinformatics*, vol. 1, no. 3, pp. 111-119, 2002.
- [8] N. Hamadani, "Automatic target cueing in IR imagery", *Master Thesis, Air Force Institute of Technology, WPAFB, Ohio*.
- [9] C. Alkan, M. Ventura, N. Archidiacono, M. Rocchi, S. C. Sahinalp, E. E. Eichler, "Organization and evolution of primate centromeric DNA from whole-genome shotgun sequence data," *PLoS Comput Biol.*, vol. 3, no. 9, pp. 1807-1818, Sep. 2007.
- [10] V. Paar, N. Pavin, I. Basar, M. Rosandic, M. Gluncic, N. Paar, "Hierarchical structure of cascade of primary and secondary periodicities in Fourier power spectrum of aliphoid higher order repeats," *BMC Bioinformatics*, vol. 9, no. 1, 466, 2008.

ISSN 1805-5443



9 771805 544150



03 >