

Multichannel quadrature coherent demodulator for a portable time-domain OCT device

P. Guerra Gutierrez^{1,2}, J.J. Gómez Valverde¹, A. Martín Gómez³,
M.J. Ledesma Carbayo^{1,2}, J.L. Rubio Guivernau³, A. Santos Lleó^{1,2}

¹ Biomedical Image Technologies, Universidad Politécnica de Madrid, Madrid, Spain, pedro.guerra@upm.es

² CIBER-BBN, Madrid, Spain

³ MedLumics S.L. Tres Cantos, Spain, jlrbio@medlumics.com

Abstract

The high demands on energy consumption and data processing that face medical devices have boosted significant advances in real-time signal processing, filtering and massive parallel computing. In this context, we propose a novel design of a multichannel quadrature coherent demodulator included in the development of an innovative portable time-domain optical coherence tomography (OCT) device. Two aspects of the design are brought into consideration: the transformation of a conventional signal processing algorithm into an equivalent version that is suitable for hardware implementation and the use for high speed data transmission of on-chip modules originally developed for mass-electronics applications. The results show an adequate implementation for the demanding design constraints in terms of noise rejection, numerical accuracy and band selectivity.

1. Introduction

Medical devices have to become more powerful, energy efficient, faster and smarter, with the goal of providing newer functionalities that help saving money and improving patient care. Current advances in medical imaging for diagnosis and therapeutic applications are based on new instrumentation imaging technology and intensive image processing [1]. This applies to static and mobile medical devices such as compact ultrasound devices, complex tomographic imaging devices, medical tablets, personal computers (PC) or patient data monitoring systems. In many of these applications, it is required to acquire, process, move and display or store huge amounts of data in real time. At the same time, devices need to become more compact, energy efficient and cost effective in order to succeed in an increasingly mobile and wider field of application. Constraints are so demanding that usually dedicated electronic systems must be designed for the purpose.

In this work, we review the process of transforming a signal processing algorithm with application in optical coherence tomography into an architecture that is suitable for hardware implementation. We also discuss the architectural elements of modern programmable devices that allow high speed data streaming.

2. Optical Coherence Tomography

Optical Coherence Tomography (OCT) is a non-invasive imaging technique that falls in between ultrasound and

microscopy, with an axial resolutions ranging from 1 to 15 μ m, and penetration depth in tissue around 1.5 mm, limited by light scattering. OCT is a technique that can be integrated with a wide range of instruments such as endoscopes, catheters, laparoscopes, or needles, which enable internal body imaging [2].

OCT performs cross-sectional images (B-scans) of the tissue by measuring the magnitude and echo time of backscattered light along predefined lines (A-scan) using low-coherence interferometry, or white light interferometry. OCT provides information about the tissue microstructure, but it is also possible to extract information about movements, which is the case of blood flow. This functional extension of OCT has become known as optical Doppler tomography, colour Doppler OCT, or simply Doppler OCT (D-OCT) [3].

3. Architecture of a hand-held OCT probe

The system under consideration aims at developing a cost effective, reliable and compact optical biopsy tool for dermatology and point-of-care diagnosis. The system includes a pen-like OCT probe and an associate console on which the acquired images are to be displayed and analyzed. In this probe, the splitter and the reference arms have been implemented using integrated optics, while the external source and detectors are currently coupled with optical fibers. Thanks to the integration, the varying group delay in the reference arm of the Michelson interferometer is obtained by means of the thermo-optic effect of silicon [4]. This approach reduces the complexity, footprint and accurate mechanical adjustments needed in conventional time-domain OCT implementations, which resort to a tilting mirror and a grating to implement the delay line.

In our particular implementation of the delay line, several waveguides are integrated on the substrate to combine the reference light with the backscattered light and thus resolve contributions from a number of points at the sample [5]. The retrieval of the OCT information requires simultaneously demodulating the different contributions to reconstruct the image by sequentially scanning line after line. Currently, the integrated optics device enables the concurrent exploration of four different channels. The proposed architecture for OCT data retrieval, shown in Figure 1, includes one analog to digital converter (ADC),

four quadrature channels to demodulate the coherent signal and build the A-Scan line. Images are streamed to an external platform by a hardware implementation of the User Datagram Protocol (UDP), which is part of the internet protocol (IP) stack.

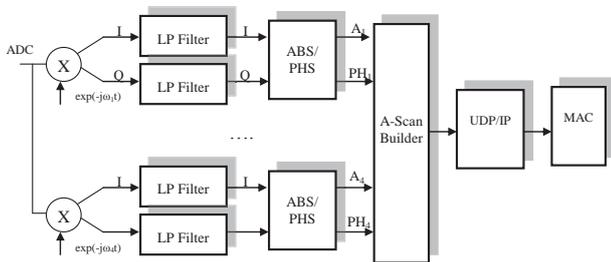


Figure 1: Proposed architecture for OCT data retrieval. Different channels are demodulated to build the A-Scan line. Images are streamed by a hardware implementation of the UDP/IP protocol.

Signal demodulation is achieved by direct down conversion of the OCT signal sections, taking into account the central frequency ω_i of the OCT signal, low-pass (LP) filtering of the inphase (I) and quadrature (Q) components to reject adjacent information and finally computing the amplitude (ABS) and phase (PHS) of the I/Q signals. Moreover, the high amount of data being acquired and the unidirectional direction of the flow justify the development of hardware accelerated implementation of the UDP/IP protocol stack. This HW-based processing module is able to sustain high data rates, above 800 Mbps, while keeping the overall system complexity moderate, avoiding the need of an additional microprocessor.

4. Implementation of the digital quadrature demodulation

The Digital signal processing (DSP) algorithms used to demodulate and build the OCT image are typically specified as infinite precision operations. Their actual implementation on a field programmable gate array (FPGA) relies on fixed-point approximations. One essential step of a top-down design flow is to determine the numerical accuracy at every signal node, namely the word-length, truncation mode and overflow mode, need to fulfil the constraints of the application. This is commonly referred to as floating-point to fixed-point conversion (FFC) problem. In the process of implementing a given algorithm into HW following a top-down design flow, we usually face two well defined stages: transforming the ideal algorithm into a functionally equivalent version that can be efficiently implemented (i.e. that only uses integer values to represent numbers), and determining the architecture, to maximize parallelism.

4.1. Floating to fixed point conversion rules

In the case of dedicated electronics, implemented with Application Specific Integrated Circuits (ASICs) or FPGAs, the implementation of a given algorithm with integer fixed-point arithmetic enables the implementation of highly parallel processing pipelines capable of massive

data processing. In order to represent a real number with fixed point, the number of fractional bits n and integer bits m has to be specified. This results in the well-known Q-notation where the integer value X with representation $Q(m_1, n_1)$ is used to represent the real number $Z = X * 2^{-n_1}$, being the total word length $w_i = n_1 + m_1 + 1$, where X is an integer represented in two's complement.

From a practical point of view, we have to point out that when implementing a signal processing algorithm in fixed point the $Q(m_1, n_1)$ representation is implicit, in the sense that the number of bits devoted to represent the fractional and integer part of a given variable are not automatically supported by the tool. That implies that, in order to perform arithmetic computations with numbers in fixed-point certain rules have to be taken into account manually to properly track the evolution of the fixed point representation after each operation.

To exemplify the fixed point conversion process, we will describe the development of an infinite impulse response (IIR) filter from specification to implementation.

4.2. Quadrature Receiver Implementation

The mixer is implemented as two tables, to store sine and cosine values, which are indexed using the phase value and a pair of multipliers. The phase value is incremented every clock cycle by a fixed amount to generate the oscillating signal with the desired frequency.

Multiple alternatives are possible to the implementation of a narrowband low pass or band pass filter. However, the filter order needed to achieve high frequency selectivity with a finite impulse response (FIR) approach is so high that other architectures, such as cascaded integrator-com (CIC) [6] filters or infinite impulse response (IIR) are chosen to provide a reasonable resource/performance trade-off. For a given frequency response it is well known that IIR filters can do with much less hardware resources than FIR filters but at the expense of a nonlinear phase response. Jackson et.al [7] showed that second-order sections (SOS) optimized in isolation and connected in cascade, called sectional optimal structures, perform very close to block optimal cascades.

4.3. Quadrature Receiver: Fixed point conversion

The ultimate goal is to implement a parameterized HDL block for IIR low pass filtering, which may be automatically generated with a high level tool, such as Matlab (The Mathworks, Natick, MA, USA) and synthesized into an FPGA. After considering alternative architectures, the decision was to implement the IIR structure as a cascade of biquadratic or second-order sections (SOS). The biquad implementation is particularly useful for fixed point implementations, as the effects of quantization and numerical stability are of no great concern. However, the overall success of any biquad implementation is dependent upon the available number precision, which must be sufficient enough in order to ensure that the quantized poles are always inside the unit circle.

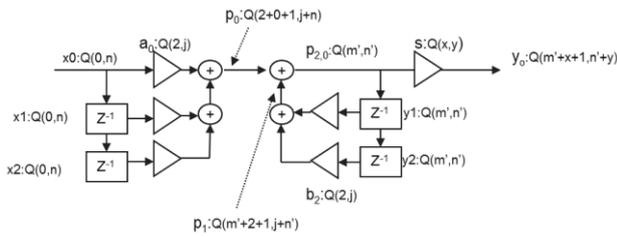


Figure 3: Fixed point analysis of the module. For the sake of clarity, we will assume that the scale factor and the representation after scaling is the same for all SOS modules. In our case, we have to specify the representation of the biquad output y_i , the internal representation $p_{2,i}$ as well as that of the scaling s and filter coefficients.

Next we perform the optimization analysis of an 8th-order Butterworth low-pass filter implementation with $W_n=0.1$. To evaluate the approximation goodness, the max relative error between the floating point implementation and fixed point implementation are compared at different points of the algorithm. Once the accuracy representation is determined, the analysis of the dynamic range at internal nodes helps us finding the number of integer bits that are required. After several simulations, we decided to constrain our internal representations to:

Node Name	Fixed Point Requirements
x_0	Q(0,12)
a_i	Q(2,11)
b_j	Q(2,11)
p_0	Q(3,23)
p_1	Q(12,15)
p_2	Q(9,4)
y_0	Q(0,8)

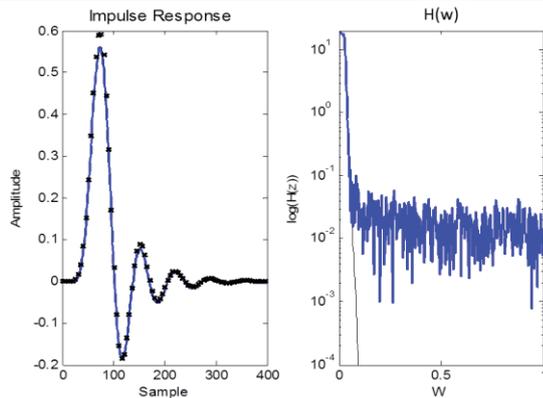


Figure 4: Floating point impulse response (black) vs fixed point counterpart (blue) in time (Left) and frequency (Right) domain in log-scale.

With these representations, the partial time response after each SOS shows no clear distortion due to quantization, and the global time and frequency responses, shown in Figure 4, shows that the final fixed-point implementation is very close to that of the ideal one.

4.4. Quadrature Receiver: Algorithm pipeline

The pipelining of the SOS module, shown in Figure 5, is straightforward: registers are inserted at regular points as shown in the following figure, which does not show fixed-point word length adjustment blocks. These intermediate registers reduce the combinational complexity between registers without affecting the numerical result and eases the automatic generation of high order filters as a cascade of this building block, at the cost of higher latency. According to synthesis results with an Artix-7 FPGA (Xilinx, Logic Drive, San Jose CA, USA), this architecture would support data processing with sampling rates up to 90 MHz, independently of the filter order thanks to pipelining.

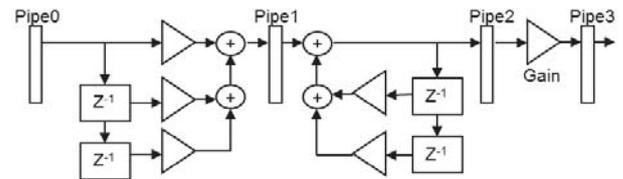


Figure 5: Block diagram of the pipelined biquad structure that is used as building block of the filter 8th order IIR low-pass filter used to select the OCT channel of interest.

4.5. Implementation of a UDP/IP HW module for high speed data streaming

In this context, a dedicated module has been implemented to send a continuous data flow over an Ethernet network using off-the-shelf components. Data is encapsulated as UDP datagrams, which are built by a hardware module in the FPGA and sent to user configurable destination ports.

This communication module connects to an on-chip medium access controller, that handles the interface with an external gigabit transceiver (88E1111, Marvell, Hamilton Bermuda). Additional connectivity to a microprocessor (uP) is provided to enable bidirectional communication with the acquisition system via software. Moreover, a light weight custom implementation of the ARP level is used by the module to determine the MAC address of the destination.

Input data, organized as lines, are buffered in a first-first out (FIFO) buffer. Whenever a full line is waiting for transmission, a finite state machine builds the datagram, including cyclic redundancy checks, transfers it to a second FIFO and request the MAC to start the transmission and sends data to the external gigabit transceiver. It has been measured that the module transfers the input data without any loss up to 800 Mbit/s.

5. Results

The implemented architecture is configured for four OCT demodulation channels, with central frequencies at 1.5, 1.7, 2.0, 2.3 MHz, spectral bandwidths of 90 kHz and a sampling rate of 12.5MHz. Low pass filters are designed to guarantee adjacent-channel rejection higher than 100 dB. Dimensioning the word sizes is carried out with this constraint in mind. As an example, Figure 6 shows the generation of the A-scan for the fourth reception channel,

including the evolution of the digital signals at different stages of the demodulation process.

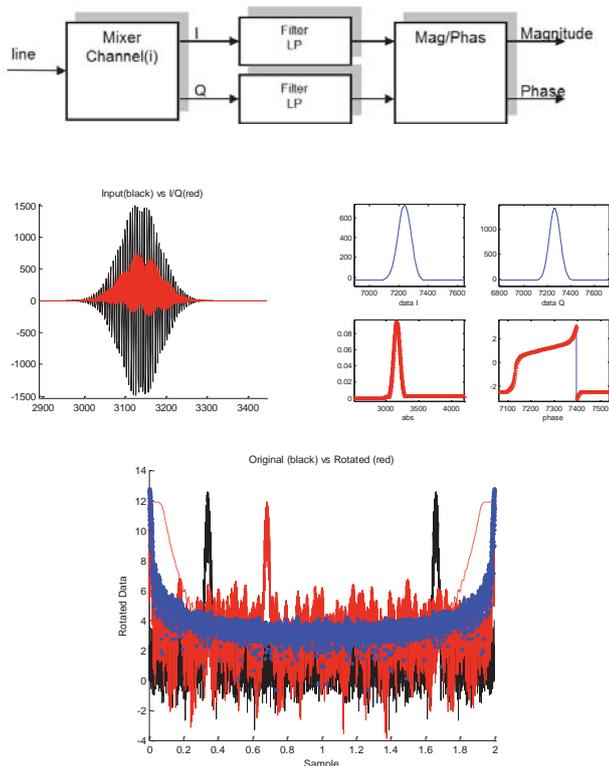


Figure 6: Diagram showing the generation of the A-scan for the 4th reception channel. These data are obtained simulating the behavior of synthesizable description of the hardware that is to be implemented in the FPGA.

In the previous diagram the demodulator steps needed to produce A-scan line are shown. The information from each channel, properly multiplexed and synchronized in time with the exploration control signals, enable the composition of a B-scan image, as shown in Figure 7. Currently, the system is configured for the generation of 850 lines per second, yielding 1 B-scan image per second, with size 850x512 pixels.

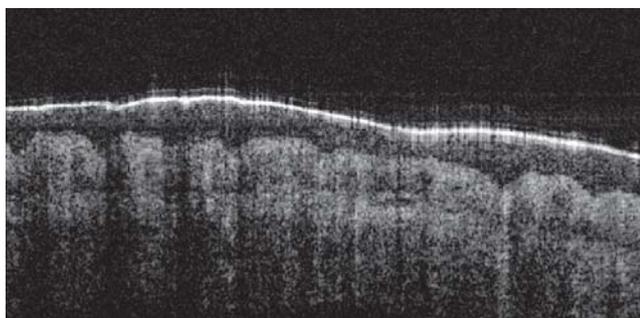


Figure 7: B-scan raw OCT image of the skin, resulting from the direct composition of A-scan lines without further post-processing.

6. Conclusions

The design of a portable time-domain optical coherence tomography (OCT) probe has been made possible by the development of an innovative technology of integrated optics. As part of this compact OCT system, a module for multi-channel coherent demodulation of the interferometric signals has been specified and implemented. Significant effort has been needed in the front-end processing to meet the demanding design constraints in terms of noise rejection, numerical accuracy and band selectivity. The processing hardware core has been highly parameterized, so that the complete module can be automatically generated with the aid of a scripting tool. Currently, the probe already produces high quality images of the skin. However future generations of the integrated optical device will demand for even sharper filters that may require new processing architectures. In this line, multirate techniques may provide a viable approach to further narrow the pass band without compromising filter stability.

Acknowledgments

This work was supported by the European Union through the FP7 project BiopsyPen and the European Fund for Regional Development (FEDER). We also acknowledge the Spanish Ministry of Science and Innovation, currently integrated in the Ministry of Economy, for their support under grants IPT-300000-2010-003 and the Comunidad de Madrid for their grant TOPUS P2013/MIT-3024

References

- [1] T. M. Deserno, H. Handels, K. H. Maier-Hein, S. Mersmann, C. Palm, T. Tolxdorff, G. Wagenknecht, and T. Wittenberg, "Viewpoints on Medical Image Processing: From Science to Application," *Current medical imaging reviews*, vol. 9, no. 2, pp. 79, 2013.
- [2] W. Drexler, and J. G. Fujimoto, *Optical coherence tomography: technology and applications*: Springer, 2008.
- [3] V. X. Yang, I. A. Vitkin, E. Regar, T. G. v. Leeuwen, and P. W. Serruys, "Principles of Doppler OCT," *Optical Coherence Tomography in Cardiovascular Research*, I. Healthcare, ed., 2007.
- [4] M. Geljon, E. Margallo-Balbas, G. Pandraud, D. H. B. Wicaksono, and P. J. French, "Time-domain Optical Coherence Tomography system with integrated delay line for surgical guidance applications." pp. 3017-3020.
- [5] J. Rubio-Guivernau, and E. Margallo-balbas, *Patent application 20130201485 "Flexible Waveguides for Optical Coherence Tomography"*, to Medlumics SL, 2013.
- [6] E. Hogenauer, "An economical class of digital filters for decimation and interpolation," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 29, no. 2, pp. 155-162, 1981.
- [7] L. B. Jackson, "Roundoff noise bounds derived from coefficient sensitivities for digital filters," *Circuits and Systems, IEEE Transactions on*, vol. 23, no. 8, pp. 481-485, 1976.