PROTOTYPE OF A SOFTWARE-DEFINED BROADCAST MEDIA INDEXING ENGINE

Olivier Romain[1], *Bruce Denby*[1,2]

[1]Laboratoire des Instruments et Systèmes Ile-de-France, Université Pierre et Marie Curie – Paris VI,
BC-252, 4 place Jussieu, 75252 Paris Cedex 05, France
[2]Laboratoire d'Electronique, Ecole Supérieure de Physique et Chimie Industrielles de la Ville de Paris (ESPCI-Paristech),
10 rue Vauquelin, 75005 Paris, France

ABSTRACT

The paper describes the prototype of a broadcast media indexing engine based on a 64-point weighted-overlap-add DFT filter bank implemented in an FPGA. In its present configuration, the device simultaneously demodulates and displays a speech/music flag for 5 FM radio stations, with audio output routed to a loudspeaker *via* a user-operated switch. Extension to the full FM radio band will be possible using a larger FPGA.

Index Terms— software defined radio, multimedia indexing, broadcast monitoring, audio processing, FPGA

1. INTRODUCTION

In emerging radio systems, improved ADC sampling rates and DSP clock speeds are pushing digital signal processors progressively closer to the antenna, allowing tasks previously performed in fixed analog hardware to be carried out on programmable or reconfigurable processors – the so-called *Software Defined Radio* (SDR) approach. While much of the initial interest in SDR was focussed on personal communications systems, there has been growing interest in extending the advantages of SDR to broadcast media as well [1-8]. SDR in such a context allows to imagine digitizing an entire broadcast band, decomposing it into its component channels, demodulating these to produce parallel real-time audio streams, and applying indexing or other audio processing algorithms, all in a single programmable device. The broadcast bands are thus transformed into a content-rich 'Hertzian Internet,' to which standard audio processing algorithms can be applied.

The ability to access to large corpi of current broadcast media content in digital format is of interest both from a *broadcast monitoring* standpoint – for questions of intellectual property and the like [9-11], and for *multimedia indexing* applications, such as audio record and rewind functions, music clip ID, keyword searches, speech/music detection, etc. [11-14]. Many of the possible applications will require embedded (e.g., automobile) or portable operation – cellphones, MP3 players, PDA, etc.

We describe here a prototype of a broadcast media indexing engine for the FM radio band, based on software radio, which targets future consumer applications in such low-resource, low-power settings. The following section provides a description of the overall system architecture, while in section 3, a presentation of the prototype performance is detailed. Conclusions and perspectives for future developments appear in the last section.

2. SYSTEM ARCHITECTURE

As shown in figure 1, the prototype is composed of an RF front end, an FPGA development board, a user interface containing LED's and switches, also on the FPGA board, and a loudspeaker. The conceptual layout of the system appears in block form in figure 2.

Figure 1. Photograph of the indexing engine prototype, showing RF front end, FPGA card with user interface, and loudspeaker.

2.1. RF Front End

The front end, providing amplification and bandpass filtering, consists of a commercial amplified indoor 75 Ω

FM radio antenna with a quoted frequency range of 88 to 108 MHz and variable gain of up to 30 dB [15].

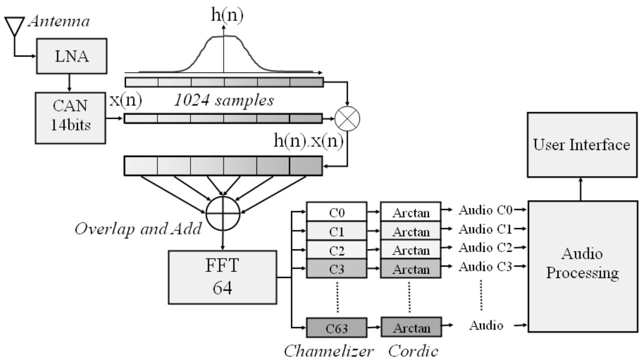

Figure 2. Block diagram of the indexing engine prototype showing RF front end, bandpass sampling, (N,K,M)=(64,16,64) WOLA FFT block, CORDIC arctan FM demodulation block, and audio/user interface block. N, K, M are explained in section 2.2.2.

2.2. FPGA Board

All post-antenna processing is performed on an Altera Stratix II EPS125 FPGA development board [16], featuring a 25000 Logic Element FPGA, 2 ADC's with rates of up to 125MSPS, 2 DAC's which furnish up to 165MSPS, as well as SRAM memory and user interfacing.

2.2.1. Bandpass Sampling

The output of the antenna, which at this point is still at RF, is bandpass sampled with 14 bit precision at 43.9 MHz, then immediately upsampled to 51.2 MHz. The choice of 43.9 MHz as the initial sampling frequency insures that the channels in the 88-108 MHz FM radio band will appear in monotonically ascending order and without aliasing at input to the filter bank (figure 3), and also maximizes the overlap between the carrier frequencies (after bandpass sampling) of local FM radio stations and the center frequencies of the filter bank. The upsampling to 51.2 MHz is necessary for the correct operation of the DFT filter bank, as described below.

2.2.2. 64-point WOLA Filterbank

A weighted-overlap-add (WOLA) DFT algorithm [17] was used to implement the filterbank. WOLA is frequently invoked as a technique for fast filter bank execution in portable, low-power systems [18, 19], as will be those required for consumer multimedia indexing applications. WOLA is based upon two key observations. The first is that an N-point FFT is equivalent to N complex digital mixers, allowing to use a single low pass filter in conjunction with the FFT 'carriers' to implement the filter bank. The second observation is that the periodicity of the FFT complex exponentials allows to time-alias the signal after multiplication with the LPF impulse response (figure 2), via a K-fold 'overlap-add,' for a dramatic reduction in the size of the FFT block required.

In the present application, N = 64 was chosen, dictated by the limited available FPGA resources. With a sampling frequency of 51.2 MHz, the filter bank will then consist of 32 useful channels spaced at 800 kHz intervals. As European FM radio stations are typically spaced at 400 kHz intervals, our system can only be sensitive to about half of these. In reality, station offsets which are multiples of 100 kHz can exist in normal operation. In its present configuration, the prototype captures only the 5 Paris area FM radio stations which lie precisely on filter bank center frequencies (figure 3). The exact set of channels selected can be varied by modifying slightly the initial bandpass sampling frequency. A larger FPGA would allow a straightforward expansion of the system to the full FM band.

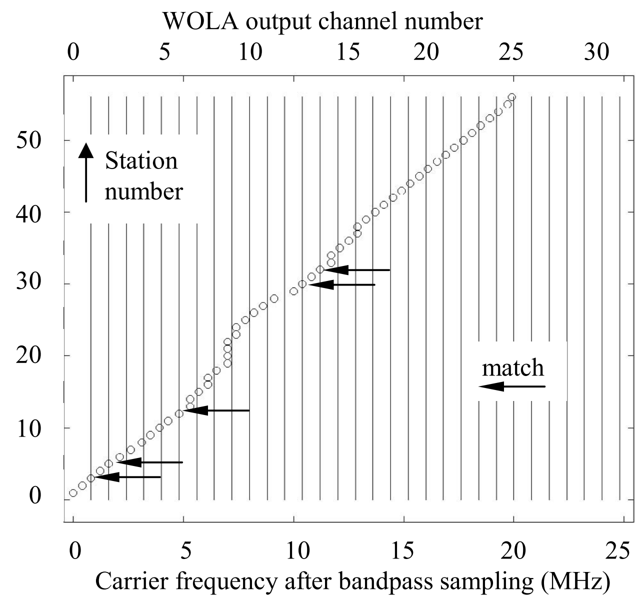

Figure 3. WOLA output channels (vertical lines, top scale) and Paris area radio stations (circles) in order of their appearance in the FM band. Carrier frequency after bandpass sampling is indicated on the lower scale. Arrows point to the 5 stations which exactly match WOLA output channels.

A 200 kHz LPF with 30 dB stopband rejection was used. The sampled signal (x(n) in figure 2) is first multiplied by the 1024 element impulse response of the LPF (h(n) in the figure). The result is broken into K = 16 blocks of 64 samples which are then pointwise summed – the overlap-add or time aliasing step – before being input to a standard 64-point FFT core. WOLA also allows to specify a decimation factor, M, independent of the time aliasing factor K. In our case, we chose M = 64, giving an output sampling frequency of 800 kHz for the 64 output streams. In a typical WOLA application, the output of the FFT must be multiplied by a complex exponential to correct for the time-advancing analysis window. With K=M, however, this step is not necessary, and thus does not appear in figure 2.

2.2.3. CORDIC arctan *Demodulation*

The baseband signals furnished by the FFT block must now be demodulated. As we are dealing with FM radio stations, the required processing amounts to simple extraction of the complex phase of the signal and temporal differentiation to obtain the instantaneous frequency. Phase extraction is accomplished using the CORDIC *arctan* algorithm [20, 21]. CORDIC is an iterative algorithm for carrying out the rotation of a two-dimensional vector using only add and shift operations, which is widely used in FPGA implementations. As shown in figure 2, channelizer logic routes the outputs of the FFT to identical copies of the CORDIC *arctan* calculation circuitry (one for each of the 5 captured stations; see discussion of chip resources below).

2.2.4. Audio Processing and User Interface

The signals exiting the *arctan*/differentiation block contain the FM radio baseband signal L+R, L-R, and RDS components. In the current configuration of the prototype, these signals are simply applied to the DAC on the Stratix board, which in turn drives the loudspeaker, the channel to be passed to the loudspeaker being selected with a DIP switch. Stereo and RDS processing are not carried out.

In addition to the possibility of selectable loudspeaker output, each channel is processed by a speech/music discriminator which sets an LED ‘on’ in the case of speech and ‘off’ for music. The discriminator algorithm, described in more detail in [7] and in [12], exploits the natural 4 Hz pause frequency characteristic of human speech, and is based upon a measurement of the fraction of time, within a sliding window, during which the audio signal remains below a fixed threshold. This algorithm, though not very sophisticated, is known to give an accuracy of some 90%, which is sufficient for a demonstrator of this type.

2.2.5. Utilization of chip resources

With the 64-point FFT, 5 copies of the CORDIC *arctan*, and 5 copies of the speech/music discriminator, 67% of the current FPGA chip resources are used. To implement the *arctan* and discriminator on *all* channels, an FPGA roughly three times as large is required, for example the Stratix II EP2S90, with 90000 logic elements.

3. PERFORMANCE OF THE PROTOTYPE

The signal amplitude at the output of the amplified antenna was about 250 mV RMS and was seen to be reasonably well band limited to the commercial FM band. The least count of the ADC is 500 μV. A summary of the results of the test session discussed in this article is given in Table I.

Carrier frequency f_p MHz	f_p after bandpass sampling MHz	WOLA FFT block output channel number	Name of radio station captured
88.6	0.8	1	Radio 88.6
89.4	1.6	2	Radio Libertaire
92.6	4.8	6	TROPICAL
98.2	10.4	13	RADIO FG
99.0	11.2	14	LATINA

Table I. Summary of the results of the test session discussed in the article. The set of stations captured can be modified by varying the initial bandpass sampling frequency. The technique can be extended to the entire FM band by using a larger FPGA.

All of the captured stations could be clearly heard and understood at a distance of several meters from the loudspeaker. No crosstalk between stations was audible, and WOLA outputs not corresponding to FM stations produced only static, indicating that the LPF used is sufficiently selective, at least for a prototype. The overall signal quality was somewhat inferior to that obtained on a standard FM radio, exhibiting static. A scope-shot of the spectrum of one channel is shown in figure 4. The the L+R band and 19 kHz pilot tone are clearly visible.

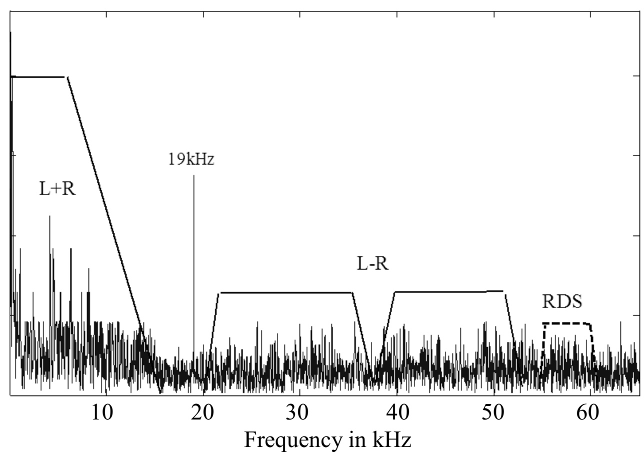

Figure 4. Scope-shot of output signal spectrum for one channel. The 19 kHz pilot and L+R band are visible; positions of other bands are included as a guide.

The speech/music discriminator LED remained “on” or “off” for passages of pure speech and music, respectively, but hovered at intermediate values during advertising passages containing music – which in fact make up a substantial fraction of FM radio air time – and for songs containing spoken lyrics.

4. CONCLUSIONS AND PERSPECTIVES

We have described a prototype broadcast media indexing engine based on SDR principles and centered around a 64-point weighted-overlap-add DFT filter bank implemented in an FPGA. The system is capable of transposing to baseband and demodulating up to 32 FM radio stations and applying audio indexing algorithms to the resulting multimedia streams. In its present configuration, our prototype performs well on the 5 commercial FM radio stations whose frequencies match the rather coarsely spaced carriers (800 kHz) imposed by limited chip resources. With a larger FPGA, the technique can be applied to the entire FM band.

In the future, in addition to increasing chip resources to cover more channels, it will be interesting to improve the signal to noise ratio, add RDS capability, and experiment other audio indexing algorithms, such as an advertising detector or search capability. The approach could of course also be applied to other multichannel radio systems, such as DAB. The ultimate goal will be to provide a simple, portable, and completely reconfigurable broadcast media indexing engine, whose precise characteristics can be tailored to the requirements of the user application at hand.

5. ACKNOWLEDGEMENTS

The authors wish to recognize the contributions of numerous students and interns who have contributed to this project over the past few years. Their work was supported in part by the French *Ministère d'Education Nationale, de la Recherche, et de la Technologie*.

6. REFERENCES

[1] http://www.vanu.com/

[2] Henrique C. Miranda, Filipe M. M. da Silva e Sousa, Rui Manuel F. S. M. Oliveira, "softFM: An all-Digital Stereo FM Transmitter with RDS based on Software Radio Architectures," ConfTele - National Telecommunications Conference, Figeira da Foz, Portugal, April 23 - 24, 2001.

[3] L. Doyle, P. Mackenzie, D. O'Mahony, K. Nolan, D. Flood, "A General Purpose Processor Component based Software Radio Engine," *Proceedings of the Second European Colloquium on Reconfigurable Radio*, Athens Greece, June 20th-22nd, 2002.

[4] Bruce Denby and Sid Ahmed Hariti, "Towards a Software Radio Enabled Broadcast Media Navigator," *EC-VIP-MC2003*, Zagreb, Croatia, July 2003.

[5] D. Flood, L. Doyle, P. Mackenzie, K. Nolan, D. O'Mahony, "Exploiting The Dynamic Flexibility Of Software Radio In FM Broadcast Receivers," *ARRL and TAPR 22nd Digital Communications Conference*, Hartford, CT, USA, Sept. 2003.

[6] Katsushi Sanda, Nobuo Itoh, Noriyoshi Suzuki, Kenji Ito, "Multi-Service Receiver for Automobiles," *R&D Review of Toyota CRDL*, Vol.38 No.4, pp. 35-42, December 2003.

[7] Bruce Denby, Olivier Romain, Sid-Ahmed Hariti, "A Software Radio Approach to Commercial FM Content Indexing," *11th International Workshop on Systems, Signals and Image Processing, IWSSIP'04 Ambient Multimedia*, September 13-15, 2004, Poznań, Poland.

[8] Olivier Romain, Bruce Denby and Sid Ahmed Hariti, "A Real Time Software Radio for the AM Long Wave Band," ibid.

[9] http://www.confirmedia.com/

[10] http://www.cyberalert.com/

[11] http://www.shazam.com/

[12] R. Jarina, N. Murphy, N. O'Connor, S. Marlow, "Speech-Music Discrimination from MPEG-1 Bitstream," *WSES/SSIP01*, Malta, September 1-6, 2001.

[13] W. Chou and L. Gu, "Robust Singing Detection in Speech/Music Discriminator Design," *ICASSP01*, Salt Lake City, May 7-11, 2001.

[14] J. Haitsma, T. Kalker and J. Oostveen, "Robust Audio Hashing for Content Identification," *CBMI01*, University of Brescia, Italy September 19-21, 2001.

[15] EuropSonic SA, 7-9 rue Cuvier, BP 80, 42302 Roanne Cedex, France.

[16] Altera Corporation, 101 Innovation Drive, San Jose, California 95134, USA.

[17] R.E. Crochière, "A Weighted Overlap-add Method of Short Time Fourier Analysis/Synthesis," *IEEE Trans. Acoust. Speech Signal Processing*, ASSP-28 (1980) 99-102.

[18] Andreas Drollinger, Dequn Sun, Christophe Waelchli, Alexandre Heubi, Peter Balsiger, Fausto Pellandini, "Efficient Sub-band Coder Implementation for Portable Low-power Applications," *International COST 254 Workshop*, Neuchâtel, Switzerland, May 5-7, 1999.

[19] Todd Schneider, Robert Brennan, Peter Balsiger and Alexandre Heubi, "An Ultra Low-power Programmable DSP System for Hearing Aids and Other Audio Applications," *ICSPAT'99 Proceedings*, November 1-4, 1999, Orlando, FL.

[20] J.E. Volder, "The CORDIC Trigonometric Computing Technique," *IRE Trans. Electronic Computers*, vol. EC-8, no. 3, 1959, pp. 330–334.

[21] U. Meyer-Base, A. Meyer-Base, and W. Hilberg, "COordinate Rotation DIgital Computer (CORDIC) Synthesis for FPGA," 4th International Workshop on Field Programmable Logic and Applications (FPL'94), Prague, Czech Republic, 7–9 Sept., 1994.