SPEECH ENCODING AND ENCRYPTION IN VLSI

K.Kalyan Chakravarthy
Dept of Electronics, Cochin University of Science and Technology, Cochin, India 682022
Tel: +91-484-532161
Fax: +91-484-532800
Email: kkckals@rediffmail.com

M.B.Srinivas
VLSI & Embedded Systems Research Center
International Institute of Information Technology, Hyderabad, India 500019
Tel: +91-40-3001969, Fax: +91-40-3001413
Email: srinivas@iiit.net

Abstract-In this work, an attempt has been made to design and synthesize speech encoding and encryption as a system-on-chip. The novelty of this design is that it uses wavelet decomposition for data compression and perpetual audio masking to keep quantization noise level to a minimum. The encryption is done by implementing RSA algorithm in hardware.

1. INTRODUCTION

Present day speech communication demands high quality [1], high resolution, low power consumption, portable and secured message transfer. The use of digital speech has become widespread and has enormous advantages over analog voice signal. This paper presents a design concept and includes all the operations to be performed on a chip. The design is intended to be implemented [2] as an ASIC.

2. SYSTEM ARCHITECTURE

"Fig.1" shows a general block diagram for speech encoding and encryption. The input to the design is 16 bit linear PCM speech data. The PCM data is applied simultaneously to both subband analysis block and FFT block, which are followed by psychoacoustic model and quantizer. The encryption block consists of an RSA encoder and the output block is a Viterbi encoder to cater to the channel coding for countering the channel noise.

3 SUB BAND FILTER ANALYSIS

The unique feature of coding for sub-band filter is discrete wavelet transform based decomposition [3] [4] [5]. The novelty of the design lies in the use of wavelets to decompose the speech signal into frequency components . The frequency components which meet the threshold criteria only are transmitted and others are discarded, thus only the components required to convey the speech are transmitted. This decomposition enables a first level of speech encryption, that is, unless the coefficients of the filters are known signal cannot be decrypted "Fig.2" shows the internal architecture of sub band filter analysis. The wavelet chosen for the analysis is db2, which is to have more time-frequency resolution and to achieve data compression.

Input samples are stored in FIFO prior to being applied to sub band filter bank via buffer. The speech packet consists of 12 samples and each sample is represented in excess-7, 16 bit floating point representation [6] as follows.

Sign	15th bit
Exponent	14-11th (4 bits)
Mantissa	10-0th (11bits)

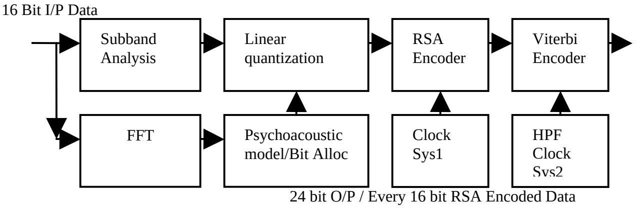

Fig 1 Block Diagram of Speech Encoding and Encryption SOC

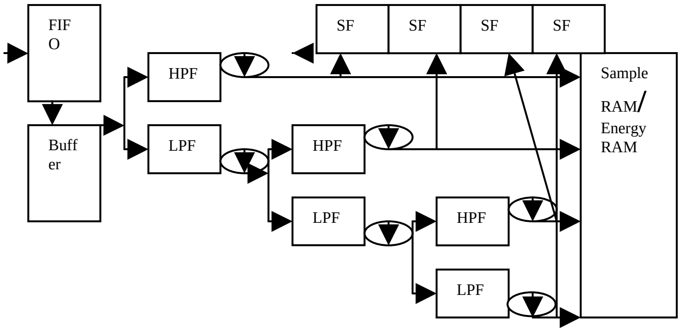

Fig 2 Internal architecture of subband analysis filter bank.

The speech packet is divided in to three slices and one slice is applied at a time to filter bank. The filter bank is based on db2 wavelet decomposition, which is a three-stage decomposition using FIR low pass and high pass filters [7]. This generates three detailed and one approximate coefficient and each of them is represented as three points. Thus a speech packet of 12 samples generates 9 samples for each sub band at the filter bank output. The points are stored and energy of these points is also calculated.

The scale factor is calculated for each subband, which is the max value of the subband under consideration (each subband consists of 9 points), which will be used later in psychoacoustic model. The sub-sampled domain corresponds approximately to perpetual domain.

2. FAST FOURIER TRANSFORM

Input samples are also applied in parallel to the FFT block. Prior to application as shown in "Fig.3", input samples are stored in a 16 X 16 RAM with an offset of 2 and centered; thus the same speech packet of 12 samples is stored. FFT [8] is performed on 16 floating point samples in parallel in 3 stages using in place computation, which is controlled by the stage controller. Energy spectrum of data segment is calculated and stored, which is used to generate a mask so that the quantization noise is confined inside this mask.

5. PSYCHOACOUSTIC MODEL

The system takes advantage of the inability of human auditory system to hear quantization noise under conditions of auditory masking. This masking is a perceptual property of the human

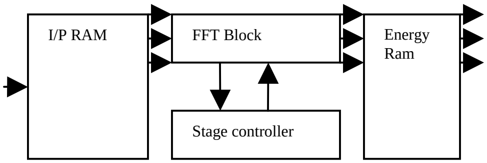

Fig 3 FFT Block Diagram

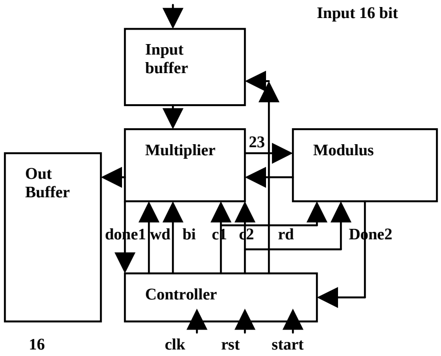

Fig 4 Block Diagram of RSA Encoder

auditory system that occurs whenever the presence of a strong speech signal makes the temporal or spectral neighborhood of weaker audio imperceptible. The psychoacoustic model analyzes the speech signal and computes the amount of masking available as a function of given signal component depends on its freq position and its loudness. Psychoacoustic model calculates individual masking threshold for each sub band, overall masking threshold, determines max masking level and calculates signal to masking ratio. The encoder uses this information to decide how best to represent the input speech samples with its limited number of code bits

6. Quantization / Bit Allocation/ Bit Packing

There is a multitude of quantizers thus allowing selecting the required quantizer. The threshold mask decides the number of bits that are to be allotted to each block. If there is not sufficient energy in a subband, the number of bits that are to be allotted might be zero. Depending on the bit allocation, one of the four quantizers is selected. Linear quantization is performed to the sub band samples and samples are scaled by a scale factor. Scale factor and quantizer number used are also packed with quantized sample data prior to being used by RSA unit.

7. RSA Encoder

RSA algorithm [9] is used for encryption of quantized data from quantizer. The encoder as shown in "Fig.4" consists of 16 bit multiplier, modulus, controller, input buffer and output buffer. Clocking system for RSA encoder is derived from system clock.

8. Viterbi Encoder

As the noise in the channel is inevitable, it can play havoc with total system performance. Obviously the need arises for system to counter the noise introduced by channel perturbations. Thus system includes viterbi encoder to counter channel noise. Viterbi encoder illustrated in the "Fig.5" receives the input from RSA encoder and produces 3 bits for every 2 bits. For K=2, k=2, n=3 the generator vectors are g1= [1111], g2= [0110], g3= [1101]

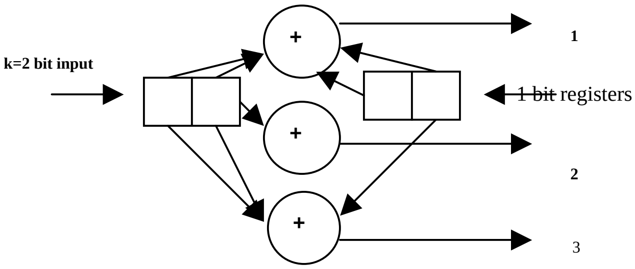

Fig 5 Block Diagram of Viterbi encoder

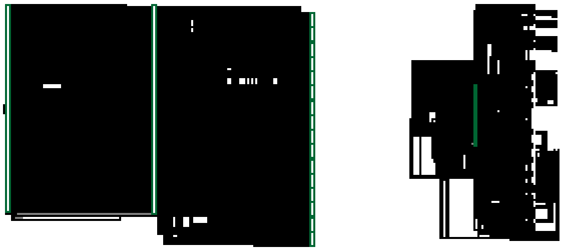

Fig 6 RTL and Technology Schematics of FFT Block

TABLE I:

DEVICE UTILIZATION FOR 2V10000ff1517

****

Resource	Used	Avail	Utilization
IOs	275	1108	24.82%
Function Generators	155211	122880	126.31%
CLB Slices	77606	61440	126.31%
Dffs or Latches	2893	126204	2.29%

9. Results And Conclusion

The system has been synthesized and verified using FPGA tools from Mentor Graphics . The synthesis results show that complete design cannot be accommodated on an FPGA. As seen from Table I even the FFT block could not be accommodated in largest FPGA that's is Xilinx Virtex II. The synthesized RTL and Technology Schematic of the FFT block are shown in "Fig.6". Efforts are being made to synthesize the entire design on a single chip using ASIC tools.

References

[1] MARVIN E. FRERKING., Digital Signal Processing in Communication Systems, Chapman Hall, ITP.

[2] JAMES R. ARMSTRONG, F.GAIL GRAY., VHDL Design Representation and Synthesis, PH PTR NJ

[3] Jaideva C. Goswami, Andrew. Chan., Fundamentals of Wavelets Theory, Algorithms, and Applications. WI. INC.

[4] RAGHUVEER M.RAO, AJIT S. BOPARDIKAR., Wavelet Transforms introduction to Theory and Applications, Addison Wesley.

[5] Robert D.Turney, Chris Dick, and Ali M.Reza., Multirate Filters and Wavelets: From Theory to Implementation, Xilinx INC, CA.

[6] Vijay K. Madisetti., Digital Signal Processors., IEEE press.

[7] David W. Knapp., BEHAVIORAL SYNTHESIS Digital System Design Using the Synopsys Behavioral Compiler, PHI NJ.

[8] JIRI JAN ., Digital Signal Filtering, Analysis and Restoration, IEE.

[9] William Stallings., Cryptography and Network Security principles and practice, PHI.