

SPEECH COMMUNICATION HARDWARE

Badin P. Degryse D.

Laboratoire de la Communication Parlée
E.N.S.E.R.G. 23, rue des Martyrs
38031 GRENoble Cedex FRANCE

ABSTRACT

This paper describes a fast digital signal processor working on 16 bits data words. Unlike the other systems this processor don't use bit slice components keeping high performance (5 Million instructions per second) and programmation simplicity with a microprocessor-like instruction set. This processor is generally connected to a general computer but its input-output system simplifies real time applications. Experimental results with a formant synthesizer and a linear prediction vocoder are demonstrated.

INTRODUCTION

Most digital speech signal processing applications needs fast, flexible and easy to program processor. The "Laboratoire de la Communication Parlée de Grenoble" is engaged in the development of such processors since 1970 for real time applications (1). Actually most realizations associate bit sliced microcomputer with micro-programming techniques. They lack easy programmation due to large instruction words and they need fine knowledge of hardware structure. It seems to us very important the development of high speed signal processors with like standard computer instruction set first at all to diminish programming efforts and second to simplify compiler generation for a high-level language. Under a work supported by the "Direction des Recherches, Etudes et Techniques" we have developed a such signal computer.

The task assigned to us was to achieve the following goals. Firstly the processor's primary mission is to serve as an experimental instrument in a research laboratory. It must perform primarily digital speech signal processing such as real time Linear Prediction Vcoders. Secondly, its instruction set must be like a standard computer to facilitate programming. Thirdly the processor must be inexpensive; therefore it was constructed using commercialy available components.

With these goals in mind we have decided to build the processor using uptodate components discarding microprogramming and systematic use of bit slice elements. Specifically a transparent three level pipeline, TTL-LS components and the TRW Multiplier Accumulator were chosen to design the Speech Communication Hardware (SCH). A first unit

has been designed, constructed and tested in less than a year, and actually several units have been commercialized.

Its description in this paper is as follows. First the general architecture is summarized with a presentation of its instruction set. Next we present its environment with its input-output system, host-computer communication and software support. Finally experimental results are presented for formant synthesis and linear prediction vocoder.

ARCHITECTURE

Structure

SCH is a like microprocessor programmable 16 bits, fixed point signal processor. It is designed with TTL-LS family circuits, HMOS-Memories and a LSI Multiplier. The entire processor with basic analog input-output system occupies a 2" rack module and comprises approximatively 250 chips. The basic cycle time for the processor is 200ns.

The 5 Million instructions per second rate is achieved by using :

- a three level pipeline structure:
 - instruction fetch
 - instruction decode-address computation and read cycle
 - execution and write cycle;
- separate memories:
 - program memory of 1 K-words of 16 bits
 - data memory of 1 or 4 K-words of 16 bits.

For high efficiency on signal processing algorithms the execution level includes a TRW 16x16 bits Multiplier-Accumulator.

Instruction set

The principal goal of the SCH design was to obtain a like microprocessor instruction set. Six registers can be accessed by program:

- A register (16 bits) :primary data accumulator also used as first operand for multiplications
- B register (16 bits) :secondary data accum.
- X register (16 bits) :data memory pointer or index register for addressing
- P register (16 bits) :Most Significant Bits of accumulation-multiplication computations
- MLSB register (16 bits) :Less Significant Bits of same computation

- PC register (10 bits) :program counter

The execution level also includes a four bits condition code which can be affected by execution of arithmetic operations and tested by control flow instructions.

All the instruction set is coded on 16 bits and can be decomposed as:

- control flow instructions which permit conditional or unconditional Jumps, subroutine Calls and Returns, with a 4 max level of subroutines nesting.
- thirty Memory-Registers instructions with direct addressing of the 256 first memory locations, or indirect addressing using the X register content and a 8 bits offset. A pseudo-immediate mode is provided by assembler. The Memory-Register operation set includes standard operations as load, store, add, subtract, logical... and a multiplication and accumulation operation using TRM Multiplier-Accumulator.
- thirty-one Register-Register instructions offering a large set of inter-register transferts and shift-operations.

Generally the three level pipeline is transparent to the programmer excepted only for some well known configurations of instructions and addressing modes.

ENVIRONMENT

I/O subsystem

SCH provides a flexible digital and analog input-output system. Modularity is achieved using an unique scheme, needing 20 lines, for module connection. For each interconnection one module is called the data acceptor and the other the source. The 20 lines are decomposed as:

- sixteen lines of data (from source)
- one Ask line (from acceptor)
- one Acknowledge line (from receptor)
- one Init line (bidirectional)
- one Error line (bidirectional).

The standard I/O board provides:

- three sources of data
- three acceptors of data
- two sampling rate programmable generators for analog conversion modules. These two generators are independant and be cascaded for subrate generation.

Analog conversion is provided by a separate module with buffering by 256 words standard First In-First Out memory modules.

This modular concept permits for example, easy connection between two SCH without buffering or with buffering by a FIFO module, using wire to wire liaison.

Host-computer communication

Connection with the host-computer is achieved through a 40 wire ribbon cable designed to be independant of some specific host-computer. The host-computer has the ability to:

- reset, run, halt the SCH
- single-step the SCH on the 200ns cycle basis
- load and examine data and program memories in halt-state and also in running-state with a cycle stealing mechanism.

A hardware flag which can be set, reset and tested by the host and the SCH permits to implement synchronization between tasks in the host and in the SCH.

Actually specific interfaces have been developed for the PDP-11 and LSI-11 of Digital Equipment.

Software support

Since there is a limited memory in the SCH, the program development is performed by using a cross-macroassembler running on a PDP-11 or LSI-11. Program debugging is supported by a DEBUG program. A subroutine library permits control of the SCH for high-level programs in the PDP.

APPLICATIONS

Linear predictive vocoder

The analysis method used for the evaluation of LPC is the autocorrelation method(10 predictors) with LEROUX algorithm (2) for linear equations solution. The pitch detection is performed by the SIFT algorithm. The synthesis method is the ladder filter one. The signal is sampled at a frequency of 8 kHz.

The timing is shared out as follows:

ANALYSIS (frames of 25ms, 200 points)	
- acquisition,preemphasis	: 0.5 ms
- Hamming's ponderation	: 0.32 ms
- correlation (10 pts cor., 200 pts sig.)	: 2.0 ms
- solution of LEROUX algorithm (10 coefficients)	: 0.38 ms
- coding	: 0.1 ms
total for analysis	: 3.3 ms

.PITCH DETECTION (frames of 40ms, 320 pts)

- downsampling to 2 kHz with low-pass filtering	: 0.5 ms
-Hamming's ponderation	: 0.13 ms
- correlation (4 pts cor., 80 pts sig.)	: 0.4 ms
- solution of system (4 coefficients)	: 0.1 ms

- inverse filtering	: 0.67 ms
- autocorrelation of error signal	: 2.5 ms
- decision Voiced/Unvoiced	: 0.5 ms
total for the SIFT	: 4.8 ms
 .SYNTHESIS	: 1.5 ms .

We notice that the complete cycle takes 9.6 ms, which allows a repetition rate between frames of 10 ms. Globally that application needs about 800 instructions.

As indicated before, there is the ability to change any value in the data or program memory while the program is running. So we can change, in real time, for example the voiced detection threshold, the number of bits used in coding or the repetition rate between frames. It makes easy the research on the effect of different parameters on the analysis-synthesis quality. We can also change instructions in real time: for example we can remove the call to the pitch detection and make a synthesis with constant pitch frequency.

Formant synthesizer

We have also developed a formant synthesizer running in real time at 16.4 kHz sampling frequency. We use a structure with two canals:

- vocal/nasal canal with 5 formants and a pole/zero filter
- noise canal with 2 formants,

and two excitation sources (voiced and noise source). This structure was inspired by the works of GOLD and RABINER (4,5), and KLATT (6).

The calculation time of a sample for a single formant is 3.6 μ s.

Each new set of command parameters (F0, AV, AB, F1, F2 ...) is furnished by the host-computer, and the filter coefficients are computed by the SCH. Therefore if we chose a sampling period of 20 ms for the command parameters, the program will run in real time with a time margin equivalent to an additional formant. We can also run a more sophisticated synthesizer with a lower signal sampling frequency.

Other applications

Thanks to the SCH programming easiness, it is envisaged to develop other programs:

- analog input-output system with digital filtering and down- or over-sampling to avoid analog filters
- general signal processing (FFT, correlation...)
- display processing (CRT driving)
- music synthesizer (digital oscillators and filters) .

CONCLUSION

We have built and developed a fast digital signal processor which is able to perform 5 million instructions per second. We have seen that its structure is relatively simple, and that standard components were used, so that the device's cost is low. Several important applications are already running on this processor, and the programming easiness allows us to envisage many other applications in various domains.

REFERENCES

- (1) D. DEGRYSE, J.F. SERIGNAT, R. CARRE
"A special speech processor", Conference on Speech Communication and Processing, BOSTON 1972.
- (2) J. LEROUX, C. GUEGUEN
"A fixed point computation of partial correlation coefficients in linear prediction", IEEE International conference on Acoustics, Speech and Signal Processing, HARTFORD, CONNECTICUT 1977.
- (3) J.D. MARKEL
"The SIFT algorithm for fundamental frequency estimation", IEEE trans. on AU, Vol. 20, N°5, pp. 367-377: 1972.
- (4) B. GOLD, L.R. RABINER
"Analysis of digital and analog Formants-Synthesizers", IEEE trans. on AU, Vol. 16, N°1, pp. 84-94: 1968.
- (5) L.R. RABINER
"Digital formant synthesizer for speech synthesis studies", J. Acoust. Soc. Amer., Vol. 43, N°4, pp. 822-828: 1968.
- (6) D.H. KLATT
"Software for a cascade/parallel formant synthesizer", J. Acoust. Soc. Amer., Vol. 67, N° 3, pp. 971-995: 1980.