

An Implementatin of a Multi-Channel Speech Surveillance System Over Telephone Lines

*Sung-Soo Kim, *Yong-Seok Kim and *Sung-Ho Cho

*This work is supported in part by the Electronic Materials and Computer Research Center of Hanyang University.

Abstract

This paper presents an implementation of a multi-channel speech surveillance system over telephone lines using TMS320C31 DSP chips. The incoming speech signals into each telephone line are first compressed simultaneously in real-time by the popular vector-sum excited linear predictive (VSELP) speech coding algorithm at the rate of 8 Kbps. The compressed speech bit streams are then multiplexed with those of other users. The multiplexed speech bit streams are transferred to the system storage equipments with some other required information so that a system operator can later monitor the stored speech data whenever it is necessary. The host program runs under Microsoft Windows95 for an efficient man-machine interface and a future upgrade-ability. We have confirmed that the overall 64-channel system operates satisfactorily in real-time. We also have checked approximately up to 2,880 total hours of recording capability of the system on a playback module and two removable backup drives.

1. Introduction

Recently, there have been a growing demand for the real-time multi-channel speech surveillance system over telephone lines in a wide variety of areas such as teleting businesses, banks, stock markets, police and fire stations, emergency services, and so on. This multi-purpose speech surveillance system must provide features, such as low bit-rate data compression, digital sound quality, fast information access and reliable data capture, and high storage capacity, so that sufficient reliability performance, convenience, and flexibility can be achieved.

In this paper, we present an implementation of a 64-channel speech surveillance system over telephone lines based on the TMS320C31 DSP chips [1]. The incoming voice signals into each telephone line are first compressed simultaneously in real-time by the popular VSELP speech coding algorithm at the rate of 8 Kbps [2], [3]. The compressed voice bit streams are then multiplexed with those from other telephone lines. The multiplexed voice bit streams are combined with some other required information, such as the channel ID and the calling time and date, and transferred to the storage equipments, so that a system operator can monitor the stored speech data whenever it is necessary.

The implemented voice surveillance system can record as few as eight channels, and can be easily expanded in 8-channel increment up to 64 total channels. The system includes a 6 Gbyte hard disk drive (HDD) for instant playback and two 2.6 Gbyte magneto optical drives (MOD) for the removable backup. These drives are configured to record in parallel. With the system's instant playback module and dual-mode removable storage media, the multi-channel voice information is recorded by time and date, and can be easily located and instantly played back with speed and precision. Of equal importance, the playback module also provides an extra level of redundancy when used as part of the dual MOD system.

The system is designed to maximize recording capacity without uncomfortable sacrifice of voice quality. The HDD and each MOD handle up to 1,440 and 720 hours of recording, respectively, so that 2,880 total hours of voice information can be recorded without any interrupt.

The overall system configuration is illustrated in Figure 1. A basic 8-channel DSP board is implemented to handle voice conversations simultaneously over up to eight telephone lines. This 8-channel DSP board consists of eight 1-channel DSP modules, and each 1-channel DSP module consists of a TMS320C31 DSP chip, memory devices such as SRAM and EPROM, and two input/output buffers, an automatic gain control circuit, an A/D converter, and a hook-off detector. As mentioned earlier, the speech surveillance system is capable of supporting up to eight 8-channel

* School of Electrical and Computer Engineering Hanyang University
Manuscript Received : September 7, 1998.

DSP boards.

The bit streams coming out of each 8-channel DSP board are multiplexed in the MUX system. The MUX system basically consists of a field programmable gate array (FPGA) and a buffer. The FPGA (Altera FLEX EPF10K50GC403-3) is programmed using very high-speed hardware description language (VHDL) [4]-[8] for the purpose of the address decoding, channel ON/OFF control, and buffer control functions.

The host program that runs under the Microsoft Windows95 environment is also utilized using Visual C++ for the efficient man-machine interface and the future upgrade-ability. The host program provides ease for managing the whole system, searching the stored data, and decoding the compressed speech for playback.

This paper is organized as follows. In the next section, a real-time implementation strategy of the VSELP algorithm is described. The system configurations are explained in Section III, and the concluding remarks are made in Section IV.

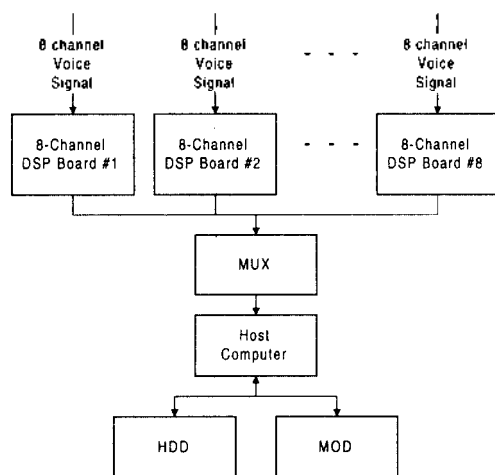


Figure 1. The block diagram of the speech surveillance system.

II. A Real-Time Implementation of the VSELP Algorithm

The VSELP speech coding algorithm has been very popular and widely recognized in a variety of applications. In particular, the 8 Kbps Motorola VSELP algorithm [2] was recently chosen by the Telecommunications Industry Association as a standard for use in North America digital mobile cellular telephone systems. The 6.7 Kbps VSELP [3] was also selected as a digital mobile cellular standard in Japan.

Our speech surveillance system follows every step of the well-known 8 Kbps VSELP speech coding algorithm

suggested in [2]. Unlike the Code-Excited Linear Predictive (CELP) speech coding algorithm having a major drawback of large computational complexities, the VSELP algorithm utilizes a codebook with a structure that allows for a very efficient search procedure. Furthermore, the algorithm was designed to accomplish the highest possible speech quality, the reasonable computational complexity, and the robustness to channel errors all simultaneously. Two VSELP excitation codebooks were used to achieve high speech quality, while maintaining reasonable complexity. A unique gain quantizer was also employed to achieve high coding efficiency, while providing robustness to channel errors. A new adaptive pre/post filter arrangement is used to enhance the reconstructed speech quality.

Table 1 shows the bit allocations for the 8 Kbps VSELP coder implemented in our system.

Table 1. Bit Allocations for 8 Kbps VSELP Algorithm [2].

Parameter	Bits / 5 msec	Bits / 20 msec
10 LPC coefficients	-	38
Average speech energy	-	5
Excitation codewords from two VSELP codebooks	14	56
Lag of pitch filter	7	28
Gain parameters	8	32
unused	-	1
Total	29	160

We first implement the 8 Kbps VSELP in C language to check the performance of the algorithm, and optimized the C source with deliberate consideration of the consequent DSP performances. Operations that require excessive computations, such as the root-square operation or the floating-point division are mostly avoided. We then translate the optimized C source into the TMS320C31 assembly code for real-time operation.

Our prime strategy for the real-time implementation of the algorithm includes 1) maximum use of the parallel instructions, 2) maximum use of the DSP internal memory, 3) minimization of the pipeline conflicts, and 4) construction of the efficient computation loops [1], [9]-[11].

The parallel instructions execute two operations per one clock cycle, and thus are the most important factor to save the overall processing time. A special care should be made occasionally, however, since parallel instructions may impose a restriction in using the DSP registers. The DSP internal memory is used basically for data that requires fast access and frequent use. Included in the 2 Kwords internal memory of the TMS320C31 in our system are two sets of the excitation codebook, speech data, 10 LPC

coefficients, and so on.

The TMS320C3x pipeline structure consists of five major units: 1) the fetch unit to fetch the instruction words from memory and update the program counter, 2) the decode unit to decode the instruction word and perform address generation, 3) the read unit to read the operand from memory if required, 4) the execute unit to read the operands from the register file, perform the necessary operation, and write results to the register file if required, and finally 5) the direct memory access (DMA) channel to read and write memory. If a pipeline conflict takes place, one instruction for one clock cycle can not be made possible.

There are three different types of the pipeline conflicts: the branch conflict, the register conflict, and the memory conflict. Among these, the register and memory conflicts are very difficult to avoid in advance. While working on the assembly code in our system, we have paid our attention mainly to minimize the branch conflict. In particular, we have tries to use the delayed branch instructions as much as possible rather than the standard branch instructions in order to keep away from the branch conflicts. The following is an example of the delayed branch instruction used in the middle of the lag search procedure:

```

;      example of delayed branch
;
      CMPI    39, R7
      BGTD   _LAG_SEA8      ; delayed branch
      STF    R1, *AR2++
      LDF    **AR0(0), R3
||     LDF    **AR0(0), R1
      LDI    @PTR_P, AR4
;
;      Branch occurs here!!!
;
      LDI    @PTR_LAG_BL, AR2
      LDI    @PTR_LAG_ZL, AR1
      ADDI   R7, AR2
      SUBI   R7, R6, R5
      LDI    R5, RC
      RPTB   _LOOP1
      ADDF   *AR1++, *AR2, R2
_LOOP1: STF   R2, *AR2++
_LAG_SEA8:
      LDI    @PTR_LAG_BL, AR2
      RPTS   39
      MPYF   *AR2++, *AR4++, R1
||     ADDF   R1, R3

```

The TMS320C31 can support looping without any overhead. There are two instructions for this purpose: RPTB for repeating a block of code and RPTS for repeating a single instruction. Using the two repeat instructions whenever they are necessary, we were able to improve

the processing time in loop computations. The following example shows a loop computation in the C source and corresponding assembly code employing the RPTS and parallel instructions:

```

;      for (i = 0; i < 170; i++, s++, d++)
;
;          *d = *s;
;
      LDF    *AR6++, R0
      RPTS   168
      LDF    *AR6++, R0      ; R0 = *s++
||     STF   R0, *AR7++      ; *sd++ = R0
      STF   R0, *AR7++

```

Figure 2 illustrates the relative execution time of the major functional modules of the 8 Kbps VSELP algorithm. As can be seen in the figure, the lag search operation requires 36% of the overall execution time. The VSELP algorithm performs lag search at every subframe (i.e., 5 msec). In order to reduce computational complexity required for lag search, we have made a full lag search in the first subframe, while in the second subframe, we make only a partial search in the neighborhood of the lag value obtained from the first subframe. The same procedure is repeated in the third and fourth subframes. By doing so, we were able to reduce approximately 6% more from the overall computations without noticeable sacrificing the voice quality.

The implemented VSELP algorithm runs at 19 MIPS, which, we believe, is enough for real-time operation on the 25 MIPS TMS320C31 single chip.

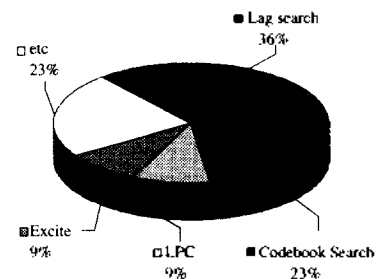


Figure 2. Relative computational complexity of the implemented VSELP algorithm.

III. System Configuration

An 8-channel DSP board consists of eight 1-channel DSP modules and is implemented to handle voice conversations simultaneously over eight telephone lines. The block diagram of the 1-channel DSP module is depicted in Figure 3.

When a hookoff status is detected, the hookoff detection

circuit sends an interrupt signal to the TMS320C31, and the TMS320C31 enables the Analog Devices AD676 16-bit linear codec. Before passing through the AD676, the speech signals on a telephone line go into the automatic gain control circuit to resolve the near-far voice level problem.

At every 20 msec, the speech data of 160 samples are stored in the internal memory of the TMS320C31 via the 4 Kwords input FIFO. The TMS320C31 produces the VSELP encoded speech data and combines them with additional header information, such as the channel ID and the calling time and date. These combined data are then transferred to the MUX system through the 4 Kwords output FIFO.

The MUX system is designed to cover up to eight 8-channel DSP boards (i.e., up to 64 channels). The Altera FLEX EPF10K50GC403-3 FPGA is employed and programmed for the purpose of the address decoding, channel ON/OFF control, and buffer control functions. The block diagram of the MUX system is illustrated in Figure 4. The address decoder sends an enable signal to the channel ON/OFF or the buffer control unit depending on the operating condition of the host computer. The channel ON/OFF control unit is used to manage the installed 8-channel DSP systems according to the operator's preference. The buffer control unit activates the multiplexed bit stream transmission from the DSP boards to the 16 Kwords FIFO in the MUX system. The buffer control unit also monitors the status of the FIFO continuously, and sends an interrupt request signal to the host computer whenever the FIFO gets full.

Once the compressed and multiplexed speech bit streams arrive at the host computer, they are first demulti-

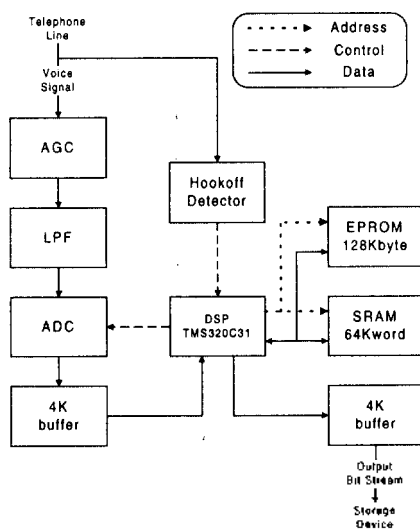


Figure 3. The block diagram of the 1-channel DSP module.

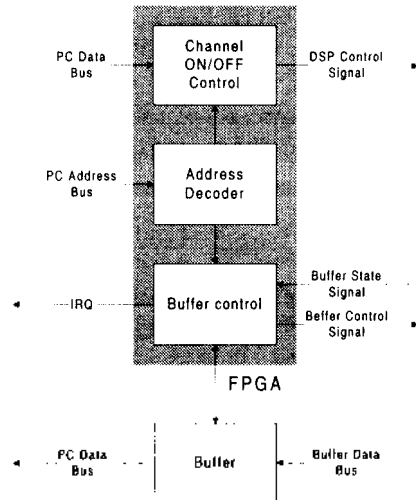


Figure 4. The block diagram of the MUX system.

plexed. The compressed speech bit streams are identified and assorted according to header information. This header information is then removed, and the speech bit streams of individual channel are finally stored in the HDD and the MOD.

IV. Concluding Remarks

In this paper, we have presented an implementation of a speech surveillance system that covers up to eight 8-channel DSP boards so that voice information of maximum 64 telephone lines can be simultaneously protected. The popular VSELP speech coding algorithm at the rate of 8 Kbps was utilized in real-time. An 8-channel DSP board based on the TMS320C31 was designed. A MUX system that can handle up to 64-channel voice data was also implemented. The system was designed to maximize recording capacity without unnecessary sacrifice of voice quality. The implemented system includes a 6GB HDD and two 2.6GB MODs, so that minimum 2,800 hours of voice information can be recorded without any interrupt.

We are currently working on implementing a more efficient speech surveillance system using a single TMS320C6201 to cover up to 16 telephone lines.

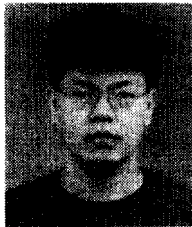
References

1. Texas Instruments, *TMS320C3x user's guide*, 1993.
2. I. A. Gerson and M. A. Jasiuk, "Vector sum excited linear prediction (VSELP) speech coding at 8 Kbps," *Proc. Int. Conf. on Acoust., Speech, and Signal Processing*, pp.461-464, Apr. 1990.
3. I. A. Gerson, "Vector sum excited linear prediction (VSELP) speech coding for Japan digital cellular," *IEICE*, pp. 35-40,

November 1990.

4. D. L. Perry, *VHDL 2nd Ed.*, R. R. Donnelley & Sons Company, 1993.
5. J. R. Armstrong and F. G. Gray, *Structured logic design with VHDL*, Prentice-Hall, 1993.
6. Zainalabedin Navabi, *VHDL - Analysis and modeling of digital systems*, McGraw-Hill, 1993.
7. Pran Kurup, Taber Abbasi, *Logic synthesis using Synopsys*, Kluwer Academic Publishers, 1995.
8. David W. Knapp, Behavioral synthesis - digital system design using the Synopsys behavioral compiler, Prentice-Hall, 1996.
9. R. Chassaing, Digital signal processing with C and the TMS320C30, John Wiley & Sons, Inc. 1992.
10. M. H. Sunwoo and S. I. Park, "Real-time implementation of the VSELP on a 16-bit DSP chip," Speech Coding Workshop, Whisler, Canada, 1991.
11. Motorola Inc., Principles of vector-sum excited linear predictive (VSELP) speech coder and its implementation on the DSP56156, Austin, Texas, 1991.

▲Sung-Soo Kim



Sung-Soo Kim was born in Seoul, Korea in 1970. He received the B.E. degree and M.S. degree in electronic engineering in 1996 and 1998, respectively, from Hanyang University, Korea. He is currently a Ph.D. student in electronic engineering at the same school. His research interests include digital communications, wireless mobile communications, digital signal processing and its applications, and image processing.

▲Yong-Soek Kim



Yong-Soek Kim was born in Seoul, Korea on November 2, 1971. He received the B.E. degree in electronic engineering in 1997 from Seoul National Polytechnic University, Korea. He is currently an M.S. student in electronic engineering at Hanyang University, Korea. His research interests include applications of digital signal processors and digital communications.

▲Sung Ho Cho



Sung Ho Cho received the B.E. degree in electronic engineering from Hanyang University, Korea, in 1982, the M.S. degree in electrical and computer engineering from the University of Iowa, Iowa City, USA, in 1984, and the Ph.D. degree in electrical engineering from the University of Utah, Salt Lake City, USA, in 1989. From August 1989 to August 1992, he was a senior member of technical staff at the Electronics and Telecommunication Research Institute (ETRI), Taejeon, Korea, developing digital communication systems. In September 1992, he joined the Department of Electronic Engineering of Hanyang University, Korea, where he is currently an Associate Professor. His current research interests include digital communications, wireless mobile communications, digital signal processing and its applications, adaptive filtering, and stochastic process. Dr. Cho is a member of Eta Kappa Nu and Tau Beta Pi.