WEDNESDAY, THURSDAY AND FRIDAY / FEBRUARY 10-11-12, 1982

1982 IEEE INTERNATIONAL

SOLID-STATE CIRCUITS CONFERENCE

1982 DIGEST of TECHNICAL PAPERS

SPONSORS: IEEE SOLID-STATE CIRCUITS COUNCIL / IEEE SAN FRANCISCO SECTION / BAY AREA COUNCIL / UNIV. OF PA.

Feb. 12, 1982: 9:00 A.M.–12:15 P.M.

SESSION XIX: Speech Processing

Imperial Ballroom

ISSCC 82 Speakers, Chairman/Moderators, Committee Members


S. Magar


R. McCharles


G. McGlinchey

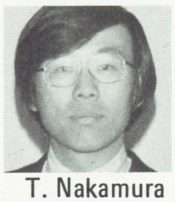

T. Macnee


K. Mathews


O. Minato


A. Mukherjee


F. Nagahisa


T. Nakamura


T. Nakashima


K. Niclas


K. Ochii


M. Odaka


T. Oura


T. Oura

SESSION XIX: SPEECH PROCESSING

FAM 19.4: A Single-Chip Sound Synthesis Microcomputer

Toshio Oura, Tomoaki Isozaki, Sachiyuki Toufuku and Hatsuhide Igarashi

Nippon Electric Co., Ltd.

Kawasaki, Japan

THIS PAPER will describe a single-chip 8b sound synthesis microcomputer which includes an on-chip 64Kb and 64-byte RAM. The SSM can synthesize high quality speech, musical sounds, imitation sounds, effective sounds and melody with a low bit rate by use of a Variable Sampling Rate Speech Segment Synthesis (VSRSSS) technique.

Figure 1 shows a block diagram of the microcomputer, which can execute powerful 182 instructions, mostly 1 machine cycle; 1.33μs. The word length of the ROM is 16b for high speed processing. In the ROM, the program for the sound synthesis and all information for sound synthesis are stored. The 32 bytes out of the 64 bytes RAM are accessible directly. They can also be used to store table addresses as 16b pair registers and as temporary memory of speech parameters. The rest of the on-chip RAM can be used as an 8-level stack register. The on-chip RAM is able to be read and written with either a 16b or 8b unit. The processing unit contains a 8b ALU, 8b accumulator, 5b multiplier register, 7b multiplicand register, 10b mode register, 7b quasi-random generator, 8b alternate register for the accumulator, 6b data pointer, 3b stack pointer, judge circuit, 1b sign register and 1b noise sign counter. In addition to general functions of an ordinary ALU, the processing unit can execute high-speed multiplication between the speech segment and envelope data. It can superpose two data with sign and execute the addition of 32-notation or 64-notation, within 1 machine cycle.

The interrupt controller contains an 8b timer and a 9b binary counter. The timer periodically generates a tone interrupt signal determined by the divide ratio. But, when the divide ratio is in the range of 0 to 7, the timer does not generate it. The minimum period of the tone interrupt is identical to 64 times of 1 machine cycle. The divide ratio is divided into the four ranges; i.e., 8 to 15, 16 to 31, 32 to 63 and 64 to 255. The period of the tone interrupt determines the sampling rate of speech segments. Two kinds of interrupt signals are generated from the binary counter; i.e., one is a noise interrupt to select any one of four periods of 32, 64, 256 or 512 times of 1 machine cycle by the mode register. The other is a time interrupt of 512 times period of 1 machine cycle. An external interrupt is generated by EXINT line, also. These interrupts are enabled by the mode register and only one of them is accepted at one time.

The current-output 9b D/A converter can drive directly a 32 to 64-ohm loudspeaker at 200mW output power. Maximum output power can be varied by the volume control line. This on-chip D/A converter can supply sound signals, with a maximum sampling rate of 23.4kHz to the loud speaker.

The multi-purpose two 8b I/O ports have four modes determined by the mode register. When the PB port can receive CS, WR and RD signals from other microcomputers, the second synthesis microcomputer can be controlled by them. When the PB port can supply WR, RD and ALE signals to peripherals by software, the microcomputer can also control them. The microcomputer can receive and transmit Byte-unit data through a PA port connected to an external data bus. It can also interface to key-matrix directly and/or drive display devices as a stand alone microcomputer. The PB port is TTL and key-matrix compatible, and the PA port is TTL compatible. Thus, the microcomputer can be applied to various systems. Figure 2 shows an example of a single-chip system. Figure 3 shows an example of a multi-chip system.

A photomicrograph of the microcomputer chip appears in Figure 4. It is fabricated by an N-channel MOS process. It contains 47,000 transistors and the die size is 4.83mm x 8.09mm (246mil square). It can operate with a single power supply over +4.5 to +6.6V at a 6MHz clock rate. In operation, the power dissipation is 450mW, but in standby, it is down to 50μW. It is housed in a 28pin plastic DIP.

Figure 5 illustrates the synthesis method. The ROM contains the software for speech synthesis, speech segments normalized for amplitude and time, envelope data, pitch data and sequence data. The Normalized Speech Segments (NSS), envelope data and pitch data are read from the ROM in order, in accordance with the sequence data. The sampling rate of the NSS is varied by the pitch data, namely the divide ratio. The processing unit multiplies the NSS data and the envelope data at sampling, and supplies multiplied data to the D/A converter. The shape of the envelope and the pitch can be freely varied against time-axis by changing each rate, and the amplitude of the envelope can be varied by peak data. The repeat numbers of the NSS can be varied freely, also. The sequence data table can contain commands of subroutine call, jump and return for the current table, as well as initial set, change and rest.

The data compression rate is higher than that of other speech synthesizers. The device is capable of synthesizing high quality 90 seconds of female speech with a bit rate of about 700b per second or 50 seconds of male speech. Speech, musical sounds, imitation sounds, effective sounds and melody by the interrupts, even during executing of the main program for control or computation can be synthesized. It can be applied not only to the VSRSSS method, which is the main purpose, but also to other speech synthesis methods, such as fixed sampling rate speech segment synthesis (FSRSSS). And it can synthesize any sounds freely by changing software. As the device does not require an on-chip digital filter, and the hardware for speech synthesis is physically very little, the performance to per cost ratio is high.

Acknowledgments

The authors thank M. Sakai, M. Endo, H. Aoyama, I. Fujitaka, T. Mukawa, E. Sugimoto, F. Tsukuda and N. Miyake for their helpful support.

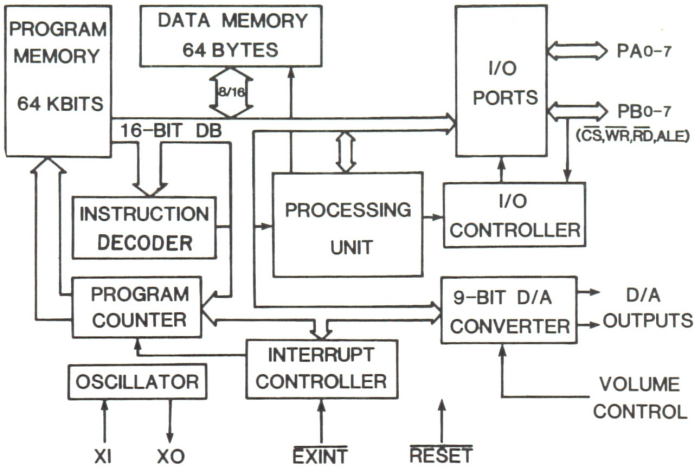


FIGURE 1—Block diagram of the sound synthesis microcomputer.

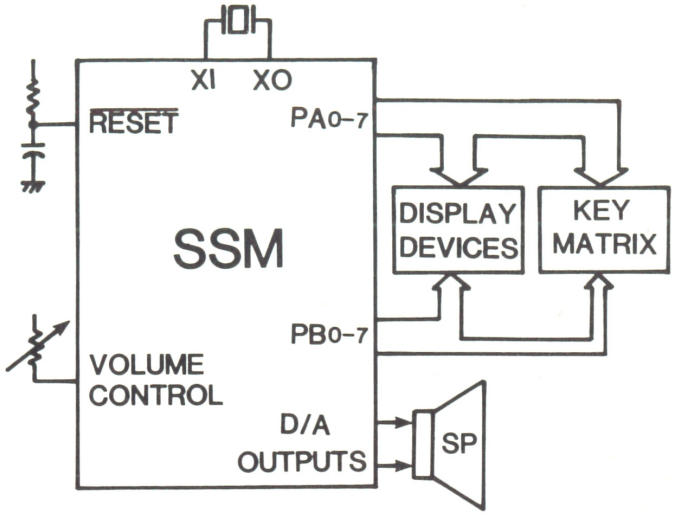


FIGURE 2—Example of single-chip system.

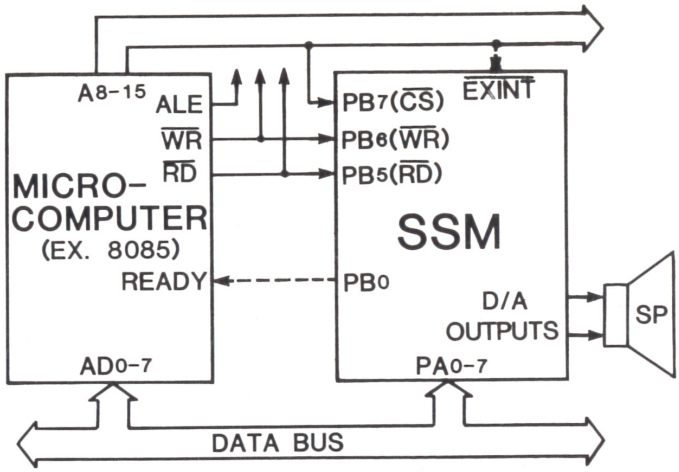


FIGURE 3—Example of multi-chip system.

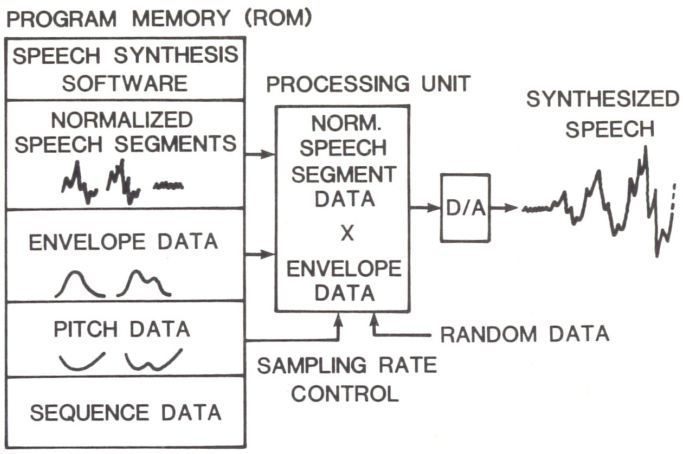


FIGURE 5—Variable sampling rate speech segment synthesis (VSRSSS) method.

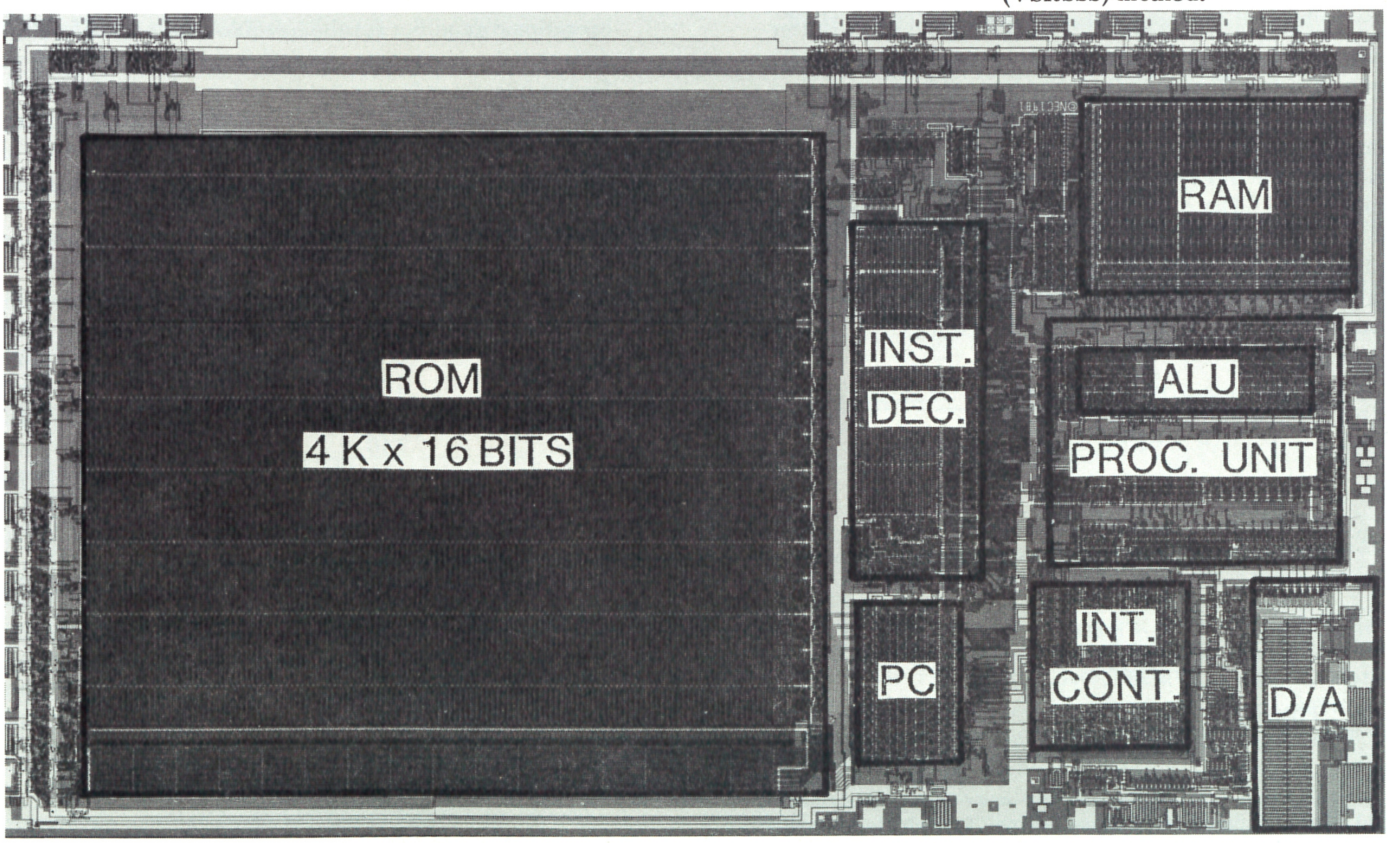


FIGURE 4—Photomicrograph of the chip.