IMPLEMENTATION OF AN LSI TONE GENERATOR CHIP FOR CHINESE SPEECH SYNTHESIS

K. C. Zhou

School of Electrical Engineering
University of Sydney

ABSTRACT - A custom LSI chip has been designed and manufactured to meet the requirements of the Chinese Text-to-Speech (CTS) system using a number of pre-settable counters. The Tone Generator of Chinese (TGC) accepts a specified 3-bit tone number code and generates tone pulse trains for the four basic and three sandhied tones. Fabricated in a 5μ nMOS process, the manufactured chip contains 1170 transistors and is 3 mm $\times$ 2.445 mm.

HOW TO DEFINE TONE PULSE TRAINS

The pitch level changes of the four basic Chinese tones as described in the Five Level Descriptions (FLD) (Chao, 1965) are shown in Figure 1(a). A transition rule has been added by the author to the FLD, assuming that the pitch period changes in a monotonic section linearly with time rather than linearly with pitch frequency as in the conventional understanding. The pitch contours of the four basic and three sandhied tones for a chosen male voice are plotted as frequency against in Figure 1(b) (Zhou, 1984; Zhou and Cole, 1984). The frequency axis is in Hertz and the time axis in sample locations. These contours have been verified by software simulation of the Chinese Text-to-Speech (CTS) system.

In the hardware design, the tone contours have been further modified as in Figure 1(c). Each contour is simply characterized by the parameters: the starting pitch period, the number of repetitions of the current pitch period, the period change direction and the location range for producing a tone pitch pulse train. Seven sets of parameters for different tone patterns are listed in Table I.

Tone	Starting Period	Repetition Number	Period Change	Location Range
1	66	1	none	0001..4096
2	80	4	down	0001..4096
3	80	1	up	0001..2048
3	102	1	down	2049..4096
4	50	1	up	0001..4096
5	80	7	down	0001..4096
6	80	2	up	0001..4096
7	61	3	up	0001..4096

Table I Simplified description of seven tone patterns

TGC DESIGN

The design was done at the laboratory for Imaging Science and Engineering (ISE), in the School of Electrical Engineering, University of Sydney. The design steps and CAD tools are: the algorithm design and

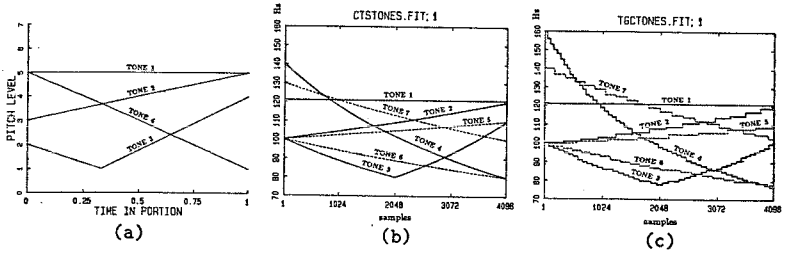

Figure 1 Tone Contours in (a) FLD; (b) Software CTS; and (c) TGC.

simulation (author's Fortron program), the logic design and simulation (ISE's Coarse Analog SIMulator - CASIM), the layout design, geometric design rule checking (Baker's DRC), circuit extraction (Berkeley's MEXTRA), electric rule checking (Berkeley's MOSERC) and simulation (CASIM) for the leaf cells, the compositions and the whole chip. The simulation results, as well as the testing results of the manufactured chip by the ISE's Chip Logic Operation Tester (CLOT) can be ploted by ISEPLOT. (Dunn A., 1984).

Following Mead and Conway's design methodology (Mead and Conway, 1980), two phase clocks are used. The clock rate is 8 kHz, the same as the sampling rate for the original speech signal.

The functional block diagram of TGC is shown in Figure 2. As soon as the TGC chip accepts the specified 3 bit tone number with a syllable synthesis start signal, it generates the tone pulse train and gives a 12 bit time scaling code at the sampling rate of 8 kHz. This time scaling code indicates each tone pulse's position in the resultant syllable for the overlap-add operation. The chip also sends out a syllable finish signal to indicate completion of the pulse train. So it has a 3 bit tone select signal, a start signal and two phase clocks as inputs; a 12 bit time scaling code, a finish signal and the tone train signal as outputs. A test bit is necessary for chip simulation and testing. This will be discussed later.

The heart of the generator is a seven bit presettable counter named the Period Counter (PC). It is preset with the 2's complement of the starting pitch period N. After N counts it produces a counter output pulse called CO_{pc}. CO_{pc} is one pulse in the expected tone train. It also acts as a reset pulse to reload PC with the next value of N from a presettable, up/down counter called Up Down Counter (UDC). CO_{pc} is also the counter input pulse CI_{ic} for the Increment Counter (IC) which has been preset with the 2's complement of the repetition number M. After M counts, IC sends an output pulse CO_{ic} which resets IC as well as acting as the input pulse CI_{udc} of UDC. That means after M equally spaced pulses in the tone train a new period, N, will be given by changing UDC's state contents by +1, −1, or 0 depending on the counting control signals of UDC. The new period will be the preset value of PC. This procedure is carried on for 4096 clock cycles to obtain an entire tone pulse train.

The total 4096 clock cycles for a complete syllable are counted by the first 12 bits of a 13 bit Length Counter (LC). LC starts counting

after receiving the signal *START*. It gives a *MID* signal from bit 11 to change the up/down control of UDC for tone 3. It also gives a signal *FINISH* while bit 12 is 1. The *FINISH* signal is used as an interrupt signal in the hardware system (Zhou, 1986).

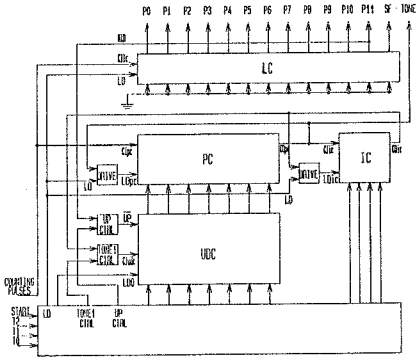

Figure 2 Functional block diagram of TGC

There are two types of counter cells in TGC: the basic counter cell and the up/down counter cell.
(1)A presettable toggle flip-flop counter has been chosen as the basic counter cell. It satisfies equations (1) and (2)

$$CO = Q_n \cdot CI \tag{1}$$
$$Q_{n+1} = Q_n \oplus CI. \tag{2}$$

The up/down counter cell satisfies the following equations.

when $\overline{UP}$ = 0 (incrementing)

$$CO = Q_n \cdot CI \tag{3}$$
$$Q_{n+1} = Q_n \oplus CI \tag{4}$$

when $\overline{UP}$ = 1 (decrementing)

$$CO = \bar{Q}_n \cdot CI \tag{5}$$
$$Q_{n+1} = Q_n \oplus CI \tag{6}$$

$\overline{UP}$ is the up/down control signal.

(2) Its state changes within a complete clock cycle when the counter is enabled, i.e. it counts clock cycles.
(2) Electrically, the cell must be able to be cascaded, in other words, its output signal *CO* must satisfy the requirements of its own input signal *CI* in order to be the input to the next stage.
(3) Geometrically, the cell should be easy to cascade to lead to a compact layout.
(4) The up/down control part is to be only an addition to the basic counter cell.

The schematics of these two cells are shown in Figure 3.

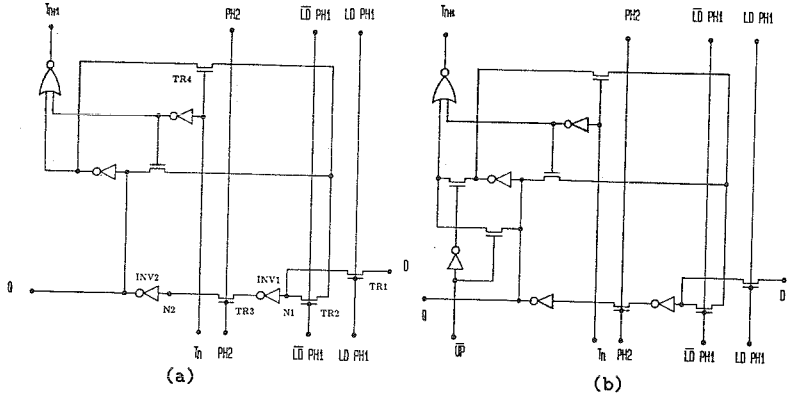

Figure 3 Design of (a) Basic Counter Cell, (b) Up/down Counter Cell.

A DRIVE cell and two interface gates LDGATE and ANDGATE are used to meet the driving and timing requirements to ensure the correct operation.

TONE 3 CONTROL

The *MID* signal acts as a tone3 control. The up/down control $\overline{UP}$ is low while the *MID* signal remains low, until LC counts to hexadecimal 00800. Then *UP* will be high to change the counting direction of the UDC while *MID* is high. This makes the pitch period increase in the first section and decreases in the second section in the tone 3 train.

A TESTING CONSIDERATION

A simulation with CASIM on the VAX 11/780 needs more than 7000 seconds of CPU time to complete 32 clock cycles for the full chip. It is unrealistic to simulate an entire pulse train (4096 cycles) which would take about 254 hours! In addition, the tester CLOT can only handle 256 samples due to its memory size. In order to tackle these problems, an additional input called *TEST* has been included in the design. While *TEST* is high, the chip is set in a testing mode which produces shorter tone trains from a second set of parameters. It is only intended to demonstrate the trend of the period changes for each tone within a limited 32 clock cycles.

THE ON CHIP CONTROLLER - THE PLA

Generated by the design tool PLAGEN (Berkeley), a Programable Logic Array (PLA) on TGC is employed as the timing controller and tone parameter generator. Table II shows the relationship between the inputs and outputs of the PLA.

SIMULATION OF THE COMPLETE TGC

The pulse trains of the four basic tones are shown in Figure 5(a), as produced from CASIM simulation under the testing mode. These pulse trains show the trend of pitch period variations to check the function

SIGNALS	T0	T1	T2	TEST	STAD	MIR	L1	L2	L21	L1D	LD0	LDTONE1	-UP	UD0	UD1	UD2	UD3	UD4	UD5	UD6	ID0
STATE	X	X	X	X	0	X	X	X	0	1	0	1	X	X	X	X	X	X	X	X	X
	X	X	X	X	1	X	1	0	1	0	1	0	X	X	X	X	X	X	X	X	X
NORMAL																					
TONE1	1	0	0	0	X	X	X	X	X	X	X	X	0	X	1	0	1	1	1	0	1
TONE2	0	1	0	0	X	X	X	X	X	X	X	X	1	0	0	0	0	0	1	1	0	0
TONE3	1	1	0	0	X	0	X	X	X	X	X	X	1	1	0	0	0	0	1	1	0	1
TONE3	1	1	0	0	X	1	X	X	X	X	X	X	1	0	0	1	0	1	1	0	0	1
TONE4	0	0	1	0	X	X	X	X	X	X	X	X	1	1	0	1	1	1	0	0	1	1
TONE5	1	0	1	0	X	X	X	X	X	X	X	X	1	0	0	0	0	0	1	1	0	1
TONE6	0	1	1	0	X	X	X	X	X	X	X	X	1	1	0	0	0	0	1	1	0	0
TONE7	1	1	1	0	X	X	X	X	X	X	X	X	1	1	1	1	0	0	0	0	1	1
TEST																					
TONE1	1	0	0	1	X	X	X	X	X	X	X	X	0	X	1	0	1	1	1	1	1	1
TONE2	0	1	0	1	X	X	X	X	X	X	X	X	1	0	1	0	0	1	1	1	1	0
TONE3	1	1	0	1	X	0	X	X	X	X	X	X	1	1	1	0	1	1	1	1	1	1
TONE3	1	1	0	1	X	1	X	X	X	X	X	X	1	0	1	0	1	1	1	1	1	1
TONE4	0	0	1	1	X	X	X	X	X	X	X	X	1	1	1	0	1	1	1	1	1	1
TONE5	1	0	1	1	X	X	X	X	X	X	X	X	1	0	1	0	0	1	1	1	1	0
TONE6	0	1	1	1	X	X	X	X	X	X	X	X	1	1	1	0	1	1	1	1	1	0
TONE7	1	1	1	1	X	X	X	X	X	X	X	X	1	1	1	0	1	1	1	1	1	0

Table II Truth Table of the PLA

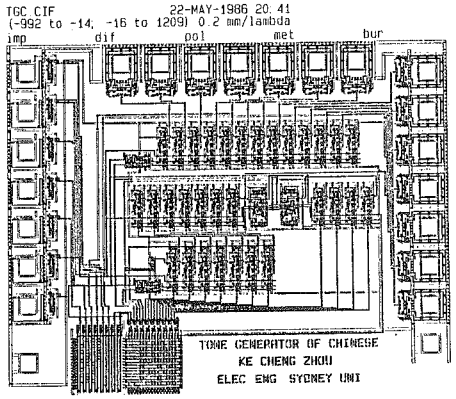

Figure 4 Layout of TGC

of the chip.

TESTING RESULTS OF THE MANUFACTURED CHIP

The fabricated chip has been tested on CLOT in the test mode. For the basic tones the real output signals are plotted in Figure 5(b). The similarity of Figure 5(a) and Figure 5(b) has proved that the architecture design was correct.

DISCUSSION AND CONCLUSIONS

The prototype TGC has been successfully tested. It will be used in the special hardware synthesizer. The key aspects of the design are:

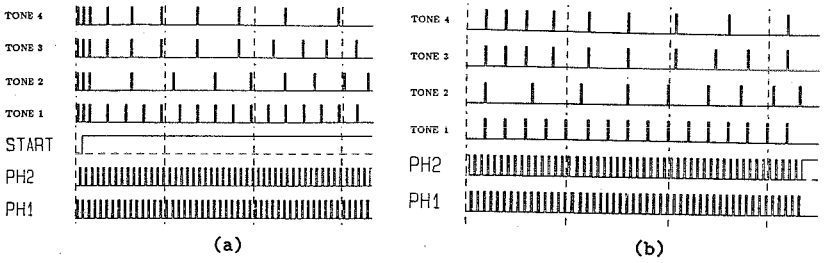

Figure 4 Tone Pulse Trains Produced by (a) CASIM and (b) CLOT

(1) A simple counter algorithm to make tone pulse trains for Chinese syllables; (2) Syllable duration can be varied by changing the LC's length; (3) Any tone pattern can be generated by varying the pattern parameters arranged in a PLA. This suggests that the method can be applied to the synthesis of any tonal language as well as Chinese, because the differences are in patterns; (4) An alternative way to specify pattern parameters may be the use of a set of programmable registers on the chip. The real time assignment of parameters could be used if the chip is designed for changeable multi-tone generation.

REFERENCES

CHAO, Y. R. (1968) "A grammar of spoken Chinese", Chapter 1, (University of California Press, Berkeley).

DUNN, A. (1984) "VLSI at ISE - the design environment", Conference Digest, VLSI-PARC, Melbourne, Australia, pp. 73-74, May 1984.

MEAD, C., CONWAY, L. (1980) "Introduction to VLSI systems", (Addison-Wesley Publishing Company, Inc.).

ZHOU, K. C. (1984) "Make a computer talk Chinese", ASAA Conference, Adelaide, Australia, May.

ZHOU, K. C., COLE, T. W. (1984) "A chip designed for Chinese text-to-speech synthesis", Journal of Electrical and Electronics Engineering, Australia, Vol.4, No.4, pp. 314-318, December.

ZHOU, K. C. (1986) "Hardware realization of Chinese text-to-speech synthesis", Proceedings of The Second Australian Computer Conference, pp. 211-220, August.