The Kernel Board for Vocal Signal Analysis/Synthesis Algorithms Implementation

George MAHALU
Suceava University, Romania
e-mail : mahalu@eed.usv.ro

Abstract

In this work is presented a kernel system for vocal signal analysis/synthesis algorithms implementation around of the LSI OKI Semiconductor's MSM6518 circuit. This circuit incorporates a synthesis and analyses stage structure which is based on the Adaptive Differential Pulse Code Modulation (ADPCM) method of data compression. Thanks to the specific structure of the circuit the users are enables to develop their own speech analysis and synthesis systems.

1. INTRODUCTION

In the pattern recognition and/or synthesis speech signals domain these are numerous hardware and software techniques to resolving specific questions. In a lot of cases are necessary very complex electronic equipments and/or sophisticated programme environments. This work become to suggest a more simple method of testing and to put in point for varied algorithms of implementation with a low price.

2. SCHEME DESCRIPTION

In figure 1 is presented the principle scheme of the development kernel system for some specific applications. The structure contain an ADPCM speech analysis/synthesis IC, an AD converter structure with serial data output and a logic control structure to ensure the communication protocol between the MSM5218 IC and other devices. To understand the manner working of the system is necessary to describe the exchange signals on the board.

The V_{ck} is refereed like a signal whose frequency is equal to the sampling frequency selected by S_1 and S_2 inputs.

The D_0 up to the D_3 are four data bits through is exchanged the ADPCM cod between a local processing subsystem and the memory system. When data is rather ADPCM coded through three bits the D_0 is not used.

The ANA/\SYN signal is analysis/synthesis function selector. It controls data flow direction into the I/O port . When high , data I/O are outputs and simultaneous analysis and synthesis occur. When low, data I/O are inputs and no analysis occurs.

The 4B/\3B signal specifies whether 3-bit or 4-bit ADPCM data is to be used.

The RESET is an active high input which initialises the MSM5218RS internal circuit. To be effective, must be held true for at least one V_{ck} time.

The S_1,S_2 inputs select the sampling frequency.

The SICK is a clock input for clocking in serial PCM data from an external ADC into the internal 12-bit shift register.

The ADSI signal is serial PCM data.

The \SCON output signals the start of conversion.

The SOCK is a MSM5218RS output which provides a 192 kHz signal which is synchronised with the output of the serial PCM data through the MSB/\SO pin. For all those is necessary to serial PCM data output be selected (DAS=H). Each bit of the 12-bit PCM data will be valid before the positive edge of this 192 kHz signal.

The DAS signal selected for analog signal output (DAS=L), or serial PCM data output (DAS=H).

The DAOUT is an analog output from the MSM5218RS circuit.

The MSB/\SO is the MSB of the data in the internal 10-bit DAC which will appear if analog signal output mode (DAS=L) is selected, or can be the serial PCM data clocked out when serial PCM data output mode is selected (DA=H).

In the next rows we'll refer to the internal structure of the MSM5218RS circuit with its functional features. Thus, this have a 12-bit shift register for the ADSI signal input processing, an internal oscillator coupled with a timing circuit, an ADPCM analysis stage block and an ADPCM synthesis stage block coupled with a 10-bit DAC and a 12-bit shift register in output. The ADPCM analysis stage and ADPCM synthesis stage blocks dispatches and receives the ADPCM data through a 4-bit bus.

The control logic block can be see in the figure 1. It is achieved with two flip-flop type D circuit and some logic gates NAND and NOT. It is present an asynchronous binary seven bits counter too.

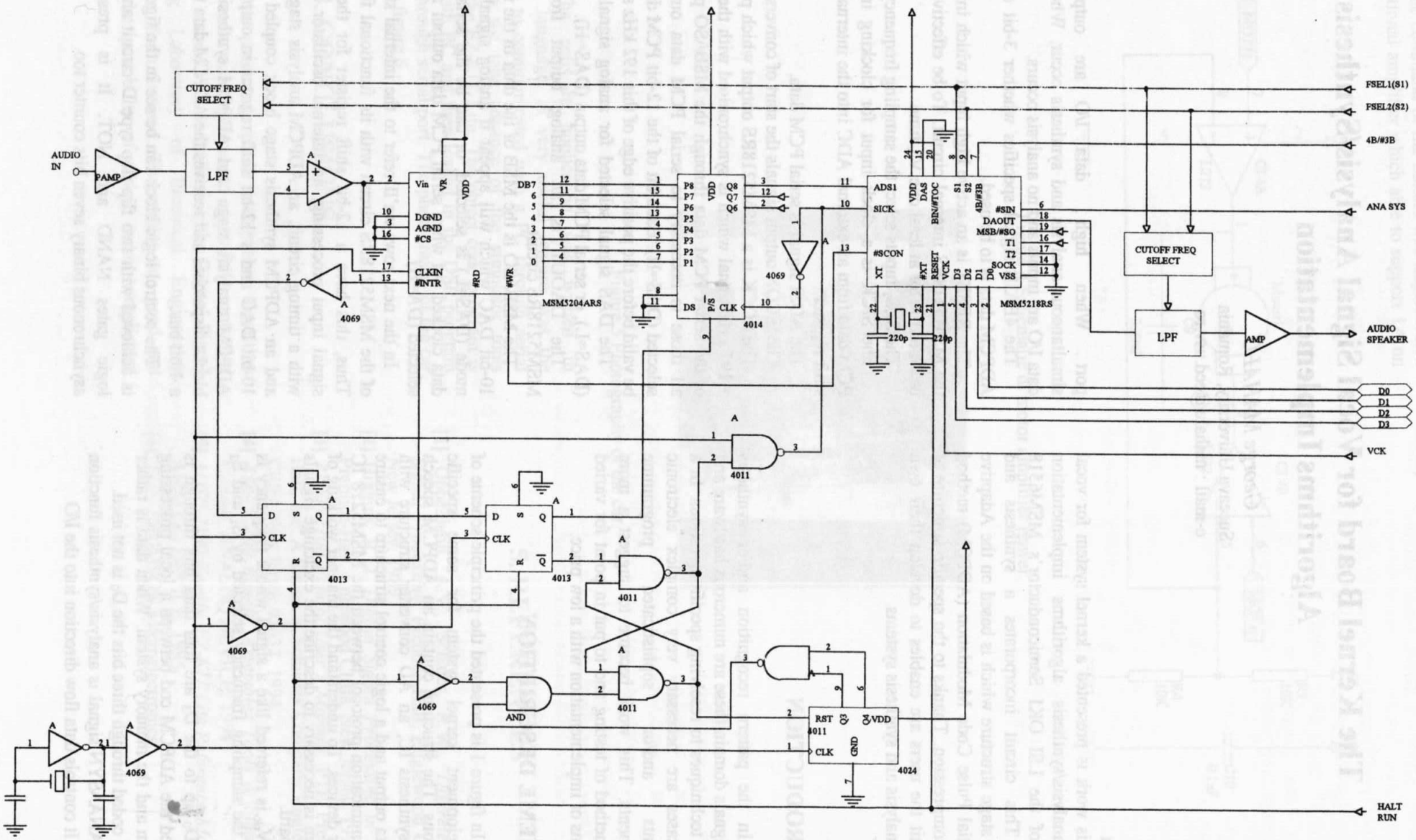
AUDIO IN
PAMP
LPF
CUTOFF FREQ SELECT
Vin
VR
VDD
DGND
AGND
#CS
CLKIN
#INTR
#RD
#WR
DB7
MSM5204ARS
P8
P7
P6
P5
P4
P3
P2
P1
Q8
Q7
Q6
DS
P/S
CLK
4014
4069
4011
4013
D
S
Q
CLK
R
AND
RST
Q3
Q4
VDD
GND
4024
ADS1
SICK
#SCON
XT
DAS
BIN/#TOC
#XT
RESET
VCK
S1
S2
4B/#3B
#SIN
DAOUT
MSB/#SO
T1
T2
SOCK
VSS
MSM5218RS
220p
FSEL1(S1)
FSEL2(S2)
4B/#3B
ANA SYS
AUDIO SPEAKER
AMP
D0
D1
D2
D3
VCK
HALT RUN

Figure 1

3. THE SYSTEM FUNCTION

Just after RESET signal applied on the HALT RUN input, a same internal RESET signal become active into the MSM5218RS circuit. The reset signal is latched within the LSI by the reset latch timing. Analysis is commenced by switching the external reset signal from H to L before a timing. The state of the reset input signal is strobed into the internal reset latch by the reset latch timing pulse. Analysis begins on the H to L transition of internal reset.

The analog speech signal is applied on the V_{in} pin of the MSM5204ARS circuit. This is an analog/digital converter with serial input/parallel output. The time of the acquisition is marked by the \SCON (Start CONversion) signal. When the ADC loaded a properly speech signal piece is generated an output signal named \INTR. This signal pass through those two flip-flop type D circuit (4013) and applies a positive pulse to the P/S input of the shift register with parallel/serial input serial output (4014). In this mode is made a parallel 8-bit load. Since that moment the data serial signal is await by the ADSI MSM5218RS input. To ensure a properly transfer of data signal between the 4014 register and the 5218 circuit will be necessary to synchronise the transfer on an appropriate clock signal. This clock signal is applied both SICK (on the MSM5218RS) and CLK (on the 4014) inputs. This clock signal is achieved like as burst of 12 pulses with 500 kHz rate by during high V_{ck} pulse. Therefore, up to 12 bits of PCM data may be strobed into the 5218 device by SICK. If more than 12 SICK pulses occur in a given V_{ck} cycle, only the last 12 are regarded valid. The cycle of SICK pulses must be completed before the next \SCON pulse.

For created that 12-bit burst is present in the structure the 4024 binary counter. After RESET signal the 4024 counts up to the 0CH code and dispatches one pulse. Thanks to this pulse the output NAND gate from the logical control subsystem is validated and supplies the SICK signal.

After all these steps, the 5218 dispatch a new \SCON signal for a new conversion command.

We can observe that first D flip-flop switches on the positive edge of the clock and thus supplies the \RD signal which command the 5204 converter to output data. The second D flip-flop switches on the next negative edge of the clock and supplies the P/S signal for the 4014 register (one positive level, therefore the active signal is P).

4. POSSIBLE ANALYSIS/SYNTHESIS SPEECH SIGNAL ALGORITHMS

After conversion of analog data signal to ADPCM code, this information can be stored into local system memory or in the PC memory through an acquisition on the parallel or serial port. In that moment can be made a properly processing after that the system can says if the speech signal is or not is recognised. Can be imagined varied appropriated algorithms. Thus, if in one memory array is stored the witness code (this code is stored at one previous moment) and in other memory array is stored the command code, the decision if this last code is or not is the correct code is done after some comparisons between the contents of those two memory arrays. Because those two codes not have often same phase and same "weight centre" will be necessary to launch more comparison session and will keep those results which exceed a defined level. The problem to establish this level is not simple but it can be resolved, in the first approximation by empirical mode. An other algorithm which will manage to resolve the decision of affiliation the command code to the witness class can analyses and compares the bit packets. For that the algorithm must does a properly choice of the bit packets following varied procedures. In this case can be involve some statistics methods for choice of the bit groups and for analyse those.

In the synthesis of the speech signal is often necessary to be used the filter rebuilding signal through the Fourier transform method. But on other hand, in used of the time analog windows methods can be utilise the Hilbert transform. This fact put the problem of use the digital filtering in rebuilding of the speech signal.

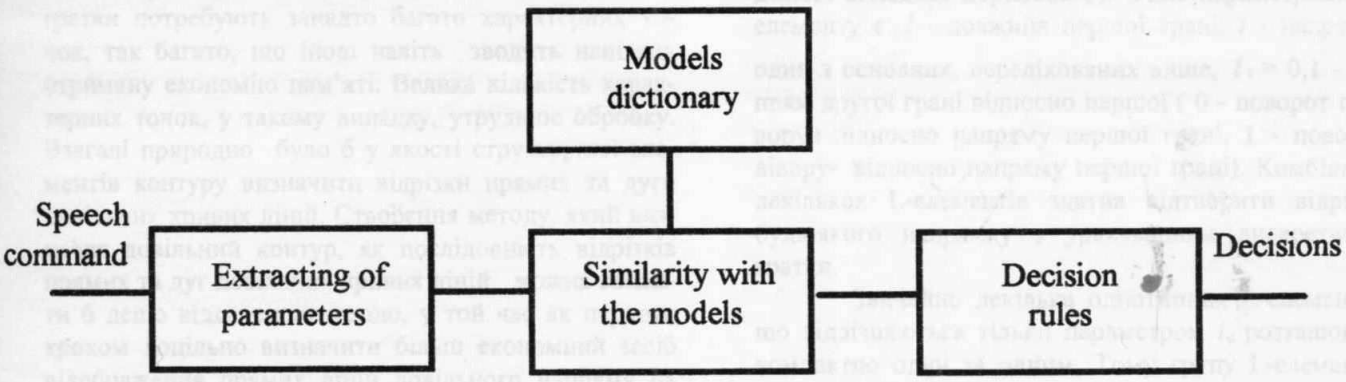

Figure 2

In the analysis domain can be involved numerous processes in understanding and recognition speech signal such as :

- the acoustic analysis, through are extracted acoustic parameters;
- the phonetic analysis, through are extracted speech sounds features;
- the prosodic analyse, through which to append the intonation information, the rhythm and accent, for purpose of the great linguistic units identification;
- the syntactic analysis, through which is testing the syntactic consistence for one hypothetically speech word comparative with the previous speech words;
- the semantic analysis, through which is check up of the understanding hypothetically sequence;
- the pragmatic analysis, through which are predicted most probable future words.

A simple structure for a determinist system of recognition can be that from the figure 2.

5. CONCLUSION

The kernel board just here presented have the advantage that has a very low cost and gives the possibility to develop a lot of software applications oriented to the pattern recognition. This system although is go like a distinct and independent structure however it can be interfacing with the PC computer and in this case it can be have all specific features. The required memory system is not so large because the achieved dates are converted on three or four bits through an ADPCM algorithm compression. On other hand this system has a very simple structure, though a high reliability. Because the work system frequency is high sufficiently we have the possibility to use it even for the vocal signal recognition, but we can use it same into the many other application such as the signs and finger print recognition (or other same). We can remark that for some applications have wrote the specific programs in C programming language with achieved and processing signal vocal purposes. Evidently, in these cases has used a PC interface structure type which running the program into the computer memory. Is possible in other cases to make a hardware link between this kernel board and an other microcontroller data processing system. In this last case the ensemble obtains a real autonomy.

REFERENCES

1. *The OKI Semiconductor Data Book - 1994/1995.*

2. D. E. Knuth - *Tratat de programare a calculatoarelor. Algoritmi seminumerici.* Editura Tehnică, Bucureşti, 1983.

3. *Data Book* - Microelectronica S.A. - 1991-1992.

4. Viorica Mărâi, Gheorghe Mărâi - *Comanda vocală a sistemelor tehnice* Editura militară, Bucureşti, 1991.