Design of LSI speech analyzer using switched capacitor filter techniques

DESIGN OF LSI SPEECH SPECTRUM ANALYZER USING SWITCHED CAPACITOR FILTER TECHNIQUES

Kenji Nakayama, Yutaka Ishikawa and Yoshiaki Kuraishi

Transmission Div., Nippon Electric Co., Ltd.
Kawasaki-City, 211 Japan

ABSTRACT

This paper presents a design approach to an LSI speech spectrum analyzer, which constructs one board speech recognition systems with other LSIs [4]. Particularly, the purpose of this LSI speech spectrum analyzer is to achieve high performances, such as a high resolution filter bank, powerful interface to CPU and variable spectrum integration intervals. In order to realize the above spectrum analyzer on one chip, switched capacitor filter (SCF) techniques are applied to filter bank synthesis. Design techniques to considerably reduce a SCF's area are introduced. A breadboard model constructed with discrete components shows good filter performances and high speech recognition rates. It is recognized through LSI design that this LSI can be realized on a 42 mm^2 chip with power dissipation of about 300 mW using the latest CMOS technology.

INTRODUCTION

The development of MOS IC technology has been promoting to realize speech signal processing systems on LSIs. It has been possible to synthesis speech recognition systems on one board using LSI chip sets [1]-[4].

This paper presents a design approach to realize a speech spectrum analyzer on an LSI chip. Particularly, the purpose of this LSI speech spectrum analyzer is to achieve high performances, such as a high resolution filter bank, powerful interface to CPU and variable spectrum integration intervals. The analyzer can be applied to both word and monosyllable recognition systems [5], [6]. In that case, however, a filter bank requires a very high filter order as a whole, and a very large chip area results. How to reduce the filter bank area is a main design problem for the above LSI.

For this purpose, switched capacitor filter (SCF) techniques are employed, which is suitable for monolithic IC filter reslizations and for mixed analog and digital circuit LSIs. Furthermore, design techniques to reduce an area of SCFs are introduced. Designed SCF characteristics and measured performances of a breadboard model constructed with discrete components are reported.

SYNTHESIS OF SPEECH SPECTRUM ANALYZER

Figure 1 shows a functional block diagram of a speech spectrum analyzer to be realized on an LSI chip. LPF1 is an anti-aliasing filter synthesized with a 2nd-order RC active filter. All filters after LPF1 are constructed with SCF techniques. An auto gain controller (AGC) is under the CPU control and adjusts the input signal level in a dynamic range of 40dB by 0.625dB step. Sample and hold (SH), and AGC circuits are synthesized with switched capacitor (SC) circuits, and are implemented with a sampling frequency of 200 kHz. LPF2 is a band limiting lowpass filter with 6th-order for sampling frequency alternation from 200 kHz to 18.18 kHz. EQL is used to emphasize speech spectrums in a high frequency band. A number of channels of a filter bank can be selected out of 16 and 20. A 20-channel filter bank is mainly required for monosyllable (alphabet and Japanese Kana) recognition systems [6]. BPFi(i=1~16 or 20) are bandpass filters having different 6th-order transfer functions for 16ch- and 20ch- filter banks. Using these filter banks, high resolution for speech spectrums can be achieved. Rect is a full wave rectifier. LPF3 is a 4th-order lowpass filter used to integrate dc component of a full wave rectified signal, that is the speech spectrums. LPF3 can realize six kinds of cut-off frequencies from 12.5 Hz to 400 Hz. One of these cut-off frequencies is selected through the CPU control in order to obtain the optimum spectrum integration interval. In the filter bank, two sets of 11 channels in low and high frequency bands are synthesized with time division multiplexed SCFs. Each channel is implemented with a sampling frequency of 18.18 kHz, and a clock rate for multiplexed 11 channels is 200 kHz. Each group includes one dummy channel which is used to observe dc offset voltage. The output of LPF3 is decimated and the sampling frequency becomes 2.02 kHz per channel. The decimated outputs of the low and high frequency groups are multiplexed and the sampling frequency becomes 1.01 kHz per channel. Therefore, the minimum spectrum integration interval becomes 1 m second. Since 22 channels are multiplexed in an interval of 1 m second, conversion time of 45 μ second per channel is required for an analog-to-digital (A/D) convertor. The output of the A/D convertor are stored in a buffer memory. Spectrums of 20 channels stored in the buffer memory can be asynchronously read by CPU within the above minimum frame of 1 m second. The spectrum integration intervals can be selected as a multiple of 1 m second.

CH1841-6/83/0000-0515 $1.00

As mentioned previously, the LPF3 cut-off frequency is determined according as this interval of reading the spectrums from the buffer memory.

SWITCHED CAPACITOR FILTER DESIGN

As described in the previous section, the filter bank requires a very high filter order as a whole, that is (BPFi:6th + LPF3:4th)x22ch=220th-order. Therefore, it becomes a very important design problem how to reduce the SCF area. This section introduces several design techniques to achieve considerable reductions in the SCF area.

Time Division Multiplexed SCF

In order to reduce numbers of operational amplifiers and capacitors, time division multiplexed SCFs [7] are employed. Figure 2(a) shows a circuit diagram for a time division multiplexed 2nd-order SCF having a stray capacitor insensitive structure. When zeros of the transfer-function locate on the unit circle, C_{13} is not necessary. High-order SCFs are constructed as a cascade form of the 2nd-order sections. Since the capacitors C_{10}, C_{13}, C_{20} and C_{23} do not discharge, they are individually required for all channels. On the other hand, the capacitors C_{11}, C_{12}, C_{21} and C_{22} discharge in a half clock interval, then these capacitors can be utilized by all channels in common. The latter capacitor values are controlled so as to monotonously increase or decrease in ascending order of the BPFi center frequencies. Figure 2(b) shows control signals applied to switches in the time division multiplexed SCF. The switches take ON and OFF states with high and low level gate signals, respectively.

Transfer Function Approximation

Transfer functions for BPFi(i=1~16 or 20) are approximated so to have Gaussian amplitude responses in the passband and stopband attenuation of more than 40dB through an iterative approximation method. The task of LPF3 is to integrate the spectrum in the given time interval. An impulse response of LPF3 is a weighting function for the spectrum integration, and is required to have a symmetrical form. For this reason, Bessel transfer functions having zeros in a finite frequency range are employed for LPF3. In order to optimize the integration interval, six kinds of cut-off frequencies from 12.5Hz to 400Hz are prepared. One of them is selected through the CPU control.

Sampling Frequency Alternation

It is well known that element values of RC active filters are vary inversely as cut-off frequencies, and a dynamic range of capacitance ratios in SCFs is proportional to a ratio of a sampling frequency to a bandwidth. Therefore, high and low sampling frequencies are required for the SH circuit and the filter bank, respectively, in order to reduce areas of both analog and sampled data filters. LPF2 is used for this purpose, and suppresses high frequency spectrums. A sampling frequency for the LPF2 output signal is alternated from 200 kHz to 18.18 kHz per channel. Both BPFi and LPF3 are implemented with sampling frequencies of 18.18 kHz per channel and 200 kHz per set of 11 channels multiplexed.

LPF3 takes a cascade form of 4th-order LPF31 and LPF32 which realize high and low cut-off frequencies, respectively. LPF31 utilizes the sampling frequency of 18.18 kHz per channel. LPF32 is implemented with an alternated sampling frequency, that is 2.02 kHz per channel. Tasks of LPF31 include band limitation for sampling frequency alternation. A decimator is utilized between LPF31 and LPF32, whose circuit diagram is described later.

Resistive Divider

Although the sampling frequencies are optimized, the ratios of sampling frequencies to bandwidths still take large values in some cases, because frequency ranges to be covered are wide. Voltage division techniques are employed in order to scale up small capacitances. There exist two kinds of voltage division methods, that is to use a capacitive divider [8] and to use a resistive divider [9]. The capacitive divider is not suited to scale up C_{13} and C_{23} in Fig. 2(a). Because these capacitors always hold charge, and the capacitive dividers are individually required for all multiplexed channels. For this reason, the resistive divider method is applied. The capacitors to be scaled up are C_{23} in BPFi, and C_{11}, C_{21} and C_{23} in LPF3.

Discrete Value Capacitance Optimization

It is well known that characteristics of SCFs are determined by capacitance ratios. Furthermore, it is desirable to realize capacitors using unit capacitors having the same dimension in order to achieve high precision in capacitance ratios with small unit capacitors on MOS ICs. Small unit capacitors make it possible to reduce the SCF area and to lighten capacitive loads for operational amplifiers. However, it is difficult to obtain sufficient filter responses after only rounding off. Therefore, discrete value capacitance optimization is inherently required and becomes a very important design process for SCFs. This discrete optimization has been well discussed in digital filters [10].

In this paper, a discrete optimization method developed for SCFs is employed [11].

Scaling capacitance: Since SCFs take the same characteristics after scaling values of capacitors connected to the same operational amplifier input, the capacitances included in the above group are scaled so as to minimize roundoff errors. These scaled and rounded off capacitances are used as initial guess in the discrete optimization.

Normalized sensitivity: Normalized sensitivity means sensitivity, at each specified frequency point, normalized by a absolute sum of sensitivities at all specified frequency points. Useful capacitors are selected based on the above normalized sensitivity, which optimize SCF responses in the specific frequency band where the maximum deviation appears.

This approach is to minimize the maximum deviation through varying discrete capacitor values. Examples of capacitance ratios for BPF8 in the 16-ch filter bank are listed in Table 1. Data in this table mean the numbers of unit capacitors. Total number of unit capacitors for BPFi(i=1~16) required

in the time division multiplexed SCFs shown in Fig. 2(b) becomes 3024.

Designed filter responses are illustrated in Figs. 4 and 5 with solid lines.

Decimator Circuit

Figure 3 shows a decimator circuit employed to decimate the LPF31 output signal. Since spectrum transitions in a time domain are important informations for speech recognition systems, differences of spectrum detecting points on a time axis among channels multiplexed are required to be small. For this purpose, the proposed decimator samples spectrums of all multiplexed channels by clocks ϕ_{31}~ϕ_{3m} within a 55 μ second interval at the LPF31 output. An interval of 55 μ second means one frame for a sampling frequency of 18.18 kHz per channel. The output signals from this decimator are held during 45 μ second per channel and 495 μ second per set of 11 channels multiplexed.

BREADBOARD MODEL IMPLEMENTATION

A breadboard model was constructed with discrete components for the case of the 16 channel filter bank using the results obtained through the design methods described in this paper.

Filter Responses

Figures 4 and 5 illustrate filter responses for BPFi and LPF3, respectively. Solid and dashed lines show designed and measured responses, respectively. Figure 5(b) only shows a designed impulse response.

The measured data for BPFi responses deviated in the direction of gain in the adjoining channel passbands. These deviations are due to cross talks between the adjoining channels occur through stray capacitors of wires in capacitor arrays [7], [12]. Therefore, layout of SCFs must be carried out so as to minimize the wire lengths in the capacitor arrays. The measured data for LPS3 responses are almost the same as the designed.

Speech Recognition Rates

The breadboard of the developed analyzer is connected with the existing speech recognition system [5]. The following recognition rates are obtained, that is 99.0% for words(a hundred of Japanese station name), 99.9% for numbers (0~9) and 93.0% for alphabet (A~Z).

From these results, it is concluded that the developed SCF speech spectrum analyzer can be successfully applied to speech recognition systems.

LSI DESIGN

Requirements for operational amplifiers utilized in BPFi are as follows: 60 dB dc gain, 7kHz band (-3 dB), and 2 μ second settling time. This operational amplifier can be realized with a 54×10^{-3} mm^2 area and power dissipation of 4mW using the latest CMOS technology. Since LPF3 has very low cut-off frequencies, requirements for the operational amplifiers can be considerablly relaxed from that for BPFi.

All capacitors are constructed using unit capacitors having the same dimension, high precision in capacitance ratios can be achieved even though the small unit capacitor is employed. An unit capacitance of 0.2 pF assures sufficient capacitance ratio accuracy for the developed analyzer. An area of 0.2 pF is about 0.3×10^{-3} mm^2.

Through LSI design, it is recognized that the developed speech spectrum analyzer can be realized on a 42 mm^2 chip and operates with a +5 v power supply. The power dissipation is about 300 mW.

CONCLUSION

Design and breadboard model implementation for a speech spectrum analyzer to be realized on an LSI chip are presented in this paper. Particularly, design methods to reduce an area of the high resolusion filter bank synthesized with SCFs and to achieve variable spectrum integration intervals are provided. Through the introduced techniques, it is confirmed that the high performance speech spectrum analyzer can be realized on an LSI chip.

ACKNOWLEDGMENT

The authors would like to thank K. Ochiai, M. Nakajima, M. Hibino, S. Tsuruta and M. Ohki for their encouragement and helpful discussion.

REFERENCES

[1] L.T.Liu et al, "A monolithic audio spectrum analyzer for speech recognition systems," in ISSCC Dig. Tech. Papers, pp.272-273, Feb. 1982.

[2] H.Ohga et al, "A Walsh-Hadamard transform LSI for speech recognition," IEEE Trans. Consumer Electronics, vol.CE-28, pp.263-270, Aug. 1982.

[3] S.Tsuruta et al, "An isolated spoken word recognition LSI chip set," Tech. Group Meeting Speech, Acoust. Soc. Japan, Oct. 1982.

[4] T. Iwata et al, "A speech recognition processor," in ISSCC Dig. Tech. Papers, Feb. 1983.

[5] Saitoh et al,"SR-100 speech input terminal," IECE of Japan, National Convention Record, p.1397, 1982.

[6] O. Izeki et al, "Speech input terminal for Japanese monosyllable," IECE of Japan, National Convention Record, p. 2310, 1981.

[7] P.W.Bosshart, "A multiplexed switched capacitor filter bank," IEEE J. Solid-State Circuits, vol. SC-15, pp.939-945, Dec. 1980.

[8] T. Hui and D.J.Allstot, "MOS switched capacitor highpass/notch ladder filters," Proc. ISCAS, pp.309-312, 1980.

[9] Y. Kuraishi, T. Makabe and K. Nakayama, "A single-chip NMOS analog front-end LSI for MODEMS," IEEE J. Solid-State Circuits, vol. SC-17, Dec. 1982.

[10] E. Avenhaus, "On the design of digital filters with coefficients of limited word length," IEEE Trans. Audio Electroacoust., vol. AU-20, pp. 206-212, Aug. 1972.

[11] S. Fushimi and K. Nakayama, "A discrete optimization method of switched capacitor filters," IECE of Japan, Rept. of Tech. Group on Circuits and Syst., vol. CAS82-107, pp.33-40, Nov. 1982.
[12] T. Kaneko, H. Kikuchi and A. Iwata, "A multiplexed switched capacitor filter," IECE of Japan, Rept. of Tech. Group on Circuits and Syst., vol. CAS81-15, pp.63-70, 1981.

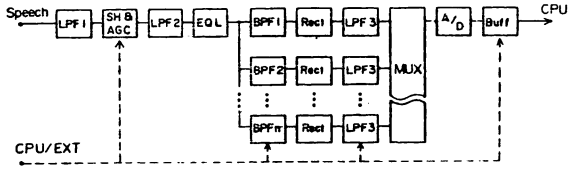

Fig. 1. Functional block diagram for speech spectrum analyzer.

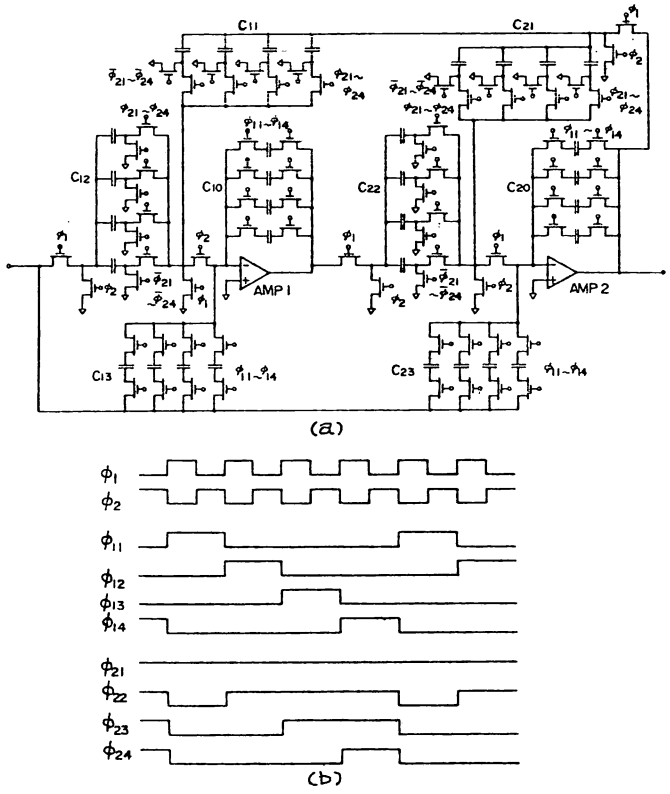

Fig. 2. (a) Circuit schematic for time division multiplexed 2nd-order SCF with 4 channels. (b) Timing chart for switch control signals.

Table 1. Example of capacitance ratios for BPF8 in 16 channel filter bank.

	Sec. 1	Sec. 2	Sec. 3
C_{10}	27	11	67
C_{11}	15	6	40
C_{12}	2	2	3
C_{13}	0	0	0
C_{20}	29	36	39
C_{21}	2	2	2
C_{22}	18	27	19
C_{23}	5	7	9

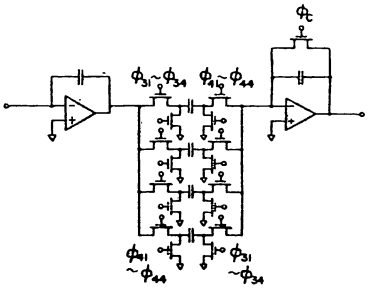

Fig. 3. Decimator circuit with 4 channels multiplexed.

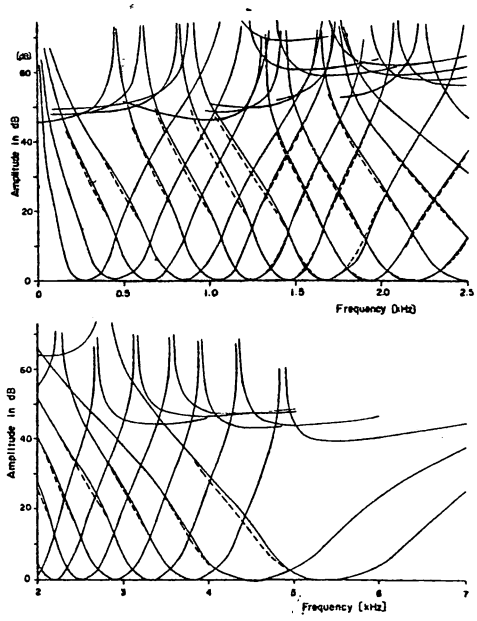

Fig. 4. Amplitude responses in dB for BPFi(i=1~16). Solid and dashed lines indicate calculated and measured responses, respectively.

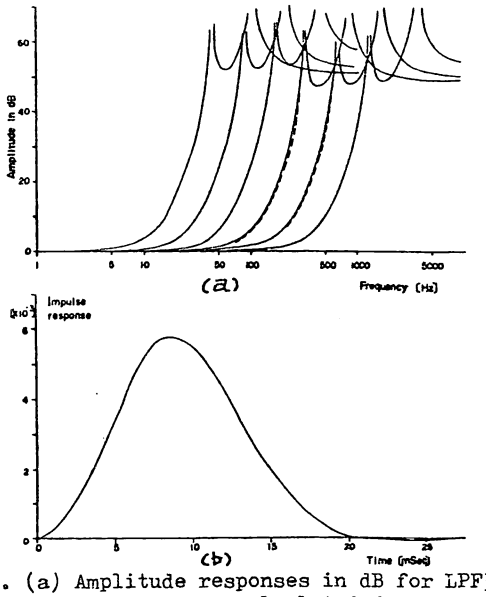

Fig. 5. (a) Amplitude responses in dB for LPF3. (b) Impulse response calculated for cut-off frequency of 50 Hz.