DEVELOPMENT OF A SWITCHED CAPACITOR SPEECH SPECTRUM ANALYZER[1] : SYSTEM DESIGN

J.S. Chang and Y.C. Tong

Department of Otolaryngology
Royal Victorian Eye and Ear Hospital
University of Melbourne

ABSTRACT - The design of a novel Low Power Monolithic Time-Multiplexed Switched Capacitor Speech Spectrum Analyzer is described. Essential features are specified with comments on the reasons for the design decisions. An experimental four channel spectrum analyzer has been fabricated and measurements on prototypes show that the design specifications are satisfied.

I. INTRODUCTION

There exist a number of speech analysis methods, including the filter bank spectrum analyzer, Linear Predictive Coefficients (LPC), LPC cepstrum, Partial Correlation Coefficients (PARCOR) techniques, Fast-Fourier Transform (FFT), etc. [Rabiner and Schafer, 1978; Markel and Gray, 1976]. The classical filter bank spectrum analyzer approach which provides a frequency domain representation of speech information is chosen for several reasons. First, the ear is known to process speech using a structure similar to the filter bank. Second, previous research has shown that the performance of a filter bank speech recognition system achieves recognition accuracy similar to a more complicated LPC based system on some standard word vocabularies [White and Neely, 1976]. In fact, a vast majority of commercial speech recognition systems employ filter bank analysis as the front-end processing for recognition [Dautrich, Rabiner and Martin, 1983; Kuraishi et. al., 1984, Lin et. al., 1983]. Third, a Switched Capacitor filter bank speech spectrum analyzer can easily be implemented in a single integrated circuit.

The development of a Low Power Monolithic Switched Capacitor Speech Spectrum Analyzer is described in this paper. Essential features are specified with comments on the reasons for the design decisions. Several novel design techniques are used to reduce hardware requirements for the realization of the spectrum analyzer. An experimental four channel version has been fabricated using a 5 micron double polysilicon Complementary-Metal-Oxide Semiconductor (CMOS) process. Some measurements from prototypes are reported.

II. FILTER BANK SPEECH SPECTRUM ANALYZER SPECIFICATIONS [Tong et. al., 1986-1988]

The speech spectrum analyzer, depicted in figure 1, comprises 24 channels covering the frequency range between 200 - 4000 Hz (to be extended to 7 KHz). Each filter channel is made up of a Bandpass (BP) filter, Full-Wave Rectifier (FWR), and Lowpass (LP) filter. The BP filter separates the input speech spectrum into a bandlimited frequency band. The bandlimited energy is extracted by the FWR and LP filter. A filter bank spectrum analyzer therefore provides a running spectrum of the speech signal.

From a system viewpoint, it is desirable that the speech spectrum analyzer possesses maximum frequency and temporal resolutions. However, as these specifications are conflicting parameters, a minimum time window of 20 milliseconds is an appropriate compromise. This time window is comparable to the shortest speech sounds, hence providing an appropriate short-time spectrum analysis.

From an integration viewpoint, it is desirable that the speech spectrum analyzer be implemented with minimum chip area and dissipate minimum power. These features are necessary for a practical portable system, for example the front-end processor of a multichannel cochlear implant system [Tong et. al., 1986-1988].

The principal parameters of the spectrum analyzer are :

BANDPASS FILTER BANK

a. 24 Channel
b. Critical Band Spacing
c. Transitional Butterworth-Bessel Approximation
d. Fourth Order per Channel
e. Composite Flat Magnitude Response
f. Time Window ≤ 20 milliseconds
g. Dynamic Range ≥ 60 dB

NON-LINEARITY FUNCTION

Full-Wave Rectifier is chosen instead of half wave in order to utilize as much signal as possible, hence improving dynamic range.

LOWPASS FILTER BANK

a. 24 Channel
b. 35 Hz Cutoff
c. Bessel Approximation
d. Third Order per Channel
e. Time Window ≤ 20 milliseconds
f. Dynamic Range ≥ 60 dB

III. DESIGN AND FABRICATION TECHNOLOGY

Switched Capacitor (SC) circuits are sampled data analog systems. They offer several advantages over fully analog circuits as the latter has mostly been fabricated using discrete or hybrid forms. SC filters, on the other hand, are realized as integrated circuits, hence compact, reliable and inexpensive for large volume applications. Their frequency responses are precise as they are specified by an accurate crystal controlled clock, capacitor ratios and filter topology. Capacitor ratios are held to 0.01% typically. Compared to an equivalent digital filter, a SC realization usually require a less complicated structure and hence much less chip area and power dissipation. Hence, a SC realization is attractive.

In view of the integration specifications, a double polysilicon CMOS process is chosen as the fabrication technology. A double polysilicon process provides for high quality capacitors with low dissipation factors and good temperature stability. A CMOS process permits low power circuit designs.

IV. DESIGN METHODOLOGY

On account of the large amount of signal processing (168 poles and 24 FWRs), the following design methodologies have been used to achieve a hardware efficient realization :

a. Pole-Zero Placement : Appropriate pole-zero pairing selection in a high order cascade synthesis reduces the total hardware requirements such as reduced operational amplifier (op amp) loading, smaller capacitors, etc.;
b. Transformations : Use of generic forms [Fleischer and Laker, 1979] other than the usual bilinear transformed numerator can provide a more economical design such as reduced number of capacitors and their values;
c. Filter Topology : Only micropower compatible filter subcircuits [Peteghem and Sansen, 1986] are used as these configurations place less op amp speed requirements which in turn reduce power dissipation;
d. Time-Division-Multiplexing : An attractive feature of SC filters is that op amps and some capacitors can be time-multiplexed [Chang and Tong, 1988a; Kuraishi *et. al.*, 1984; Bosshart, 1980] so that they are shared by more than one filter. This therefore significantly reduces the total component count which in turn achieves chip area saving;
e. Parasitic-Insensitivity : Parasitic Insensitive designs permit the use of small capacitor values which not only reduces op amp loading (low power dissipation) but also uses less chip area;
f. Low Clock Rates and multirate sampling : A low clock rate results in longer integration times, hence reduced power dissipation, and multirate sampling provides smaller capacitor values for reduced op amp loading and chip area;
g. Clock Generators : Simple clock generators require less chip area;
h. Op Amps : Dynamic Biased op amps utilize currents efficiently.

In addition, several novel design techniques have been employed :

i. Capacitor Sharing Technique : This new technique results in nearly 20% chip area saving for realizing the Time-Multiplexed biquadratic filters [Chang and Tong, 1988b];
j. Time-Multiplexed Biquadratic Filters : DC Offsets are usually a problem in Time-Multiplexed BP biquadratic filters [Lin *et. al.*, 1983; Kuraishi *et. al.*, 1984]. By careful design, these biquadratic filters are designed such that DC offsets are not a problem [Chang and Tong, 1988b] and allows for hardware efficient BP filter bank realizations;
k. Time-Multiplexed FWR : A novel SC Time-Multiplexed FWR features parasitic insensitivity, high sensitivity, and does not require a sample-and-hold input. It is also jitter-free (delay-free), DC offset compensated and employs micropower compatible subcircuits only. Furthermore, all its components are shared in a Time-Multiplexed application, thus achieving substantial chip area saving [Chang and Tong, 1988c].

V. EXPERIMENTAL RESULTS

An experimental four channel Switched Capacitor Speech Spectrum Analyzer has been fabricated using a 5 micron double polysilicon CMOS process. A microphotograph of this integrated circuit is shown in figure 2. The active area of the speech spectrum analyzer is approximately 1.4 mm x 3 mm. It can be seen that the layout is very regular where the area allocated for each filter channel is equal. The modularity of the design reduces the overall design time and allows easy extension to include more channels if desired.

The theoretical and measured frequency responses of the BP filter bank are depicted in figure 3 (a) (one channel is zero input to observe DC offsets) which shows good agreement. The DC offsets between different BP filter bank channels is typically 8 mV, a negligible error in this application.

Using a single 5V supply, the maximum swing at 1 % Total-Harmonic-Distortion is 3V peak-peak, and a dynamic range of 73 dB is achieved. With a single 3V supply, a dynamic range of 64 dB is attained which satisfies the specification in Sec. II. Such low voltage operation is particularly attractive for biomedical applications.

The impulse response of the BP filter bank is depicted in figure 3(b). The Time-Multiplexed outputs of 3 channels with unit impulse inputs and 1 zero input channel are shown as the middle trace. The top trace is the demultiplexed output of channel 1. It can be seen that the filter responses are settled within 20 ms, hence satisfying the time window specifications.

Figure 4 depicts the output of the Time-Multiplexed BP filter bank (middle trace) and Time-Multiplexed FWR (top trace) for an input at the passband frequency of channel 1 (bottom trace).

The LP frequency response is shown in figure 5 where the measured response agrees well with the theoretical response.

The current drawn by the speech spectrum analyzer is 500 uA for a single 3V supply. The op amps were deliberately overdesigned in this implementation for measurement purposes. The final design comprising 24 channels is expected to dissipate 2.5 mW and take up an active area of 4.65 mm x 3 mm using a 5 micron process. The power dissipation and chip area figures can be significantly improved if a smaller feature technology is available.

VI. CONCLUSIONS

The design of a novel Low Power Time-Multiplexed Switched Capacitor Audio Spectrum Analyzer for short-time speech spectrum analysis has been described. Essential features of this design has been specified with comments on the reasons for the design decisions. Several new design techniques have been used to reduce the hardware in the speech spectrum analyzer realization.

Measurements from prototypes have verified the new methodologies employed and show that a low power monolithic realization is feasible.

REFERENCES

1. Bosshart P.W. (1980), 'A Multiplexed Switched Capacitor Filter Bank' IEEE J. Solid-State Circuits 15, 939-945.
2. Chang J.S. and Tong Y.C (1988a), 'A Switched Capacitor Time-Division-Multiplexed Pole Sharing Technique for Linear Phase Bandpass Filter Banks', IEEE Int. Symp. Circuits and Systems (Espoo, Finland), 1245-1248.
3. Chang J.S. and Tong Y.C. (1988b), 'A New Filter Bank Design for a Low Power Monolithic SC Speech Spectrum Analyzer', Submitted to IEEE Int. Symp. Circuits and Systems 1989.
4. Chang J.S. and Tong Y.C. (1988c), 'A Low Power SC Speech Spectrum Analyzer', Submitted to IEEE J. Solid-State Circuits.
5. Dautrich B.A., Rabiner L.R. and Martin T.B. (1983), 'On Effects of Varying Filter Bank Parameters on Isolated Word Recognition', IEEE Acoust. Speech and Signal Proc. 31, 793-807.
6. Flanagan J.L. (1972), Speech Analysis Synthesis and Perception, 2 nd Edn., Springler-Verlag.
7. Fleischer P.E. and Laker K.R. (1979), 'A Family of Switched Capacitor Biquad Building Blocks', Bell Sys. Tech. J. 58, 2235-2269.
8. Kuraishi Y., Nakayama K., Miyadera K., et. al. (1984), 'A Single-Chip 20 Channel Speech Spectrum Analyzer Using a Multiplexed Switched Capacitor Filter Bank', IEEE J. Solid-State Circuits 19, 964-970.
9. Lin L.T., Tseng H.F., Cox D.B., et. al. (1983), 'A Monolithic Audio Spectrum Analyzer', IEEE Acoust. Speech and Signal Proc. 31, 288-293.
10. Markel J.D. and Gray Jr. A.H (1976), Linear Prediction of Speech, Springler-Verlag.
11. Rabiner L.R. and Schafer R.W. (1978), Digital Processing of Speech Signals, Prentice-Hall.
12. Tong Y.C., Chang J.S., Harrison J.M., et. al. (1986-1988), National Institutes of Health (USA) Contract No. 1-NS-5-2388 : 'Speech Processors for Auditory Prosthesis', Quarterly Progress Reports 1986-1988.
13. Van Peteghem P.M.V. and Sansen W.M.C. (1986), 'Power Consumption versus Filter Topology in SC Filters', IEEE J. Solid-State Circuits 21, 40-47.

14. White G.M. and Neely R.B. (1976), 'Speech Recognition Experiments with Linear Prediction, Bandpass Filtering and Dynamic Programming' IEEE Trans. Acoust. Speech and Signal Proc. 24, 183-188.
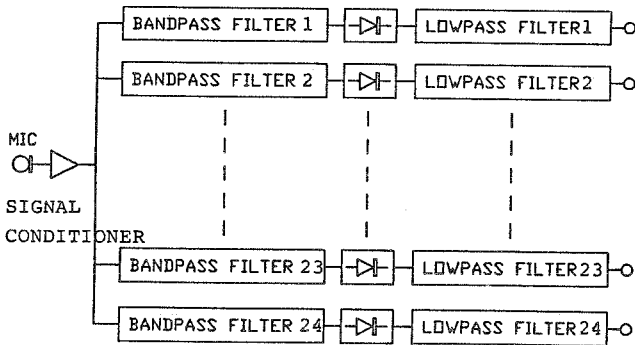
DIAGRAMS

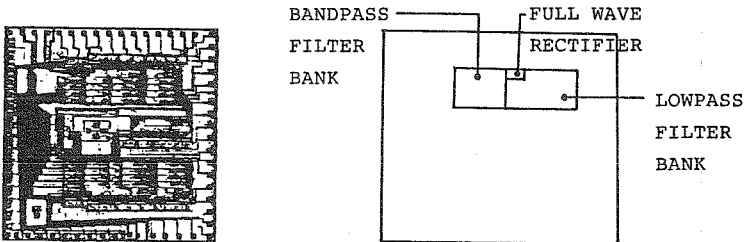FIG 1 : SPEECH SPECTRUM ANALYZER

FIG 2 : MICROPHOTOGRAPH
        OF EXPERIMENTAL SPEECH SPECTRUM ANALYZER
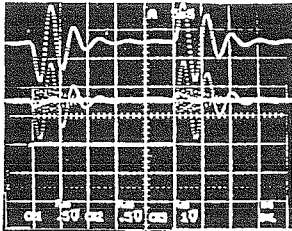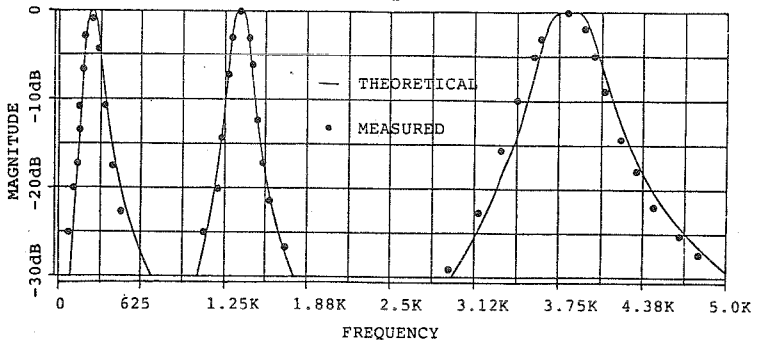
FIG 3(a): BANDPASS FILTER BANK FREQUENCY RESPONSE

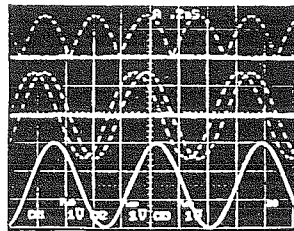FIG 3(b) : IMPULSE RESPONSE
OF BP FILTER BANK

FIG 4 : BP FILTERBANK and
FWR Output

LOWPASS FILTER FREQUENCY RESPONSE

FIG 5