Implementing a Noise Cancellation System with the TMS320C31

Authors: V. Davidek, J. Sika, J. Stusak

ESIEE, Paris

September 1996
SPRA335

TEXAS INSTRUMENTS

IMPORTANT NOTICE

Texas Instruments (TI™) reserves the right to make changes to its products or to discontinue any semiconductor product or service without notice, and advises its customers to obtain the latest version of relevant information to verify, before placing orders, that the information being relied on is current.

TI warrants performance of its semiconductor products and related software to the specifications applicable at the time of sale in accordance with TI's standard warranty. Testing and other quality control techniques are utilized to the extent TI deems necessary to support this warranty. Specific testing of all parameters of each device is not necessarily performed, except those mandated by government requirements.

Certain application using semiconductor products may involve potential risks of death, personal injury, or severe property or environmental damage ("Critical Applications").

TI SEMICONDUCTOR PRODUCTS ARE NOT DESIGNED, INTENDED, AUTHORIZED, OR WARRANTED TO BE SUITABLE FOR USE IN LIFE-SUPPORT APPLICATIONS, DEVICES OR SYSTEMS OR OTHER CRITICAL APPLICATIONS.

Inclusion of TI products in such applications is understood to be fully at the risk of the customer. Use of TI products in such applications requires the written approval of an appropriate TI officer. Questions concerning potential risk applications should be directed to TI through a local SC sales office.

In order to minimize risks associated with the customer's applications, adequate design and operating safeguards should be provided by the customer to minimize inherent or procedural hazards.

TI assumes no liability for applications assistance, customer product design, software performance, or infringement of patents or services described herein. Nor does TI warrant or represent that any license, either express or implied, is granted under any patent right, copyright, mask work right, or other intellectual property right of TI covering or relating to any combination, machine, or process in which such semiconductor products or services might be or are used.

TRADEMARKS

TI is a trademark of Texas Instruments Incorporated.

Other brands and names are the property of their respective owners.

CONTACT INFORMATION

US TMS320 HOTLINE	(281) 274-2320
US TMS320 FAX	(281) 274-2324
US TMS320 BBS	(281) 274-2323
US TMS320 email	dsph@ti.com

Contents

Figures

Tables

Implementing a Noise Cancellation System with the TMS320C31

Abstract

This paper is concerned with the digital implementation of spectral subtraction modified algorithm on a Texas Instruments (TI™) TMS320C31 Digital Signal Processor (DSP). The algorithm was developed to suppress a car noise in speech picked up by one microphone in a running car for application in mobil hand-free telephony. The real-time implementation on floating-point processor is described. The speech activity detector is a limiting part of one channel spectral subtraction algorithm. In implemented version of program the cepstral speech/pause detector was used.

This document was part of the first European DSP Education and Research Conference that took place September 26 and 27, 1996 in Paris. For information on how TI encourages students from around the world to find innovative ways to use DSPs, see TI's World Wide Web site at www.ti.com

Product Support on the World Wide Web

Our World Wide Web site at www.ti.com contains the most up to date product information, revisions, and additions. Users registering with TI&ME can build custom information pages and receive new product updates automatically via email.

Introduction

Noise cancellation in a speech signals degraded by additive noise is an important and common problem in signal processing especially in telecommunication systems. Adaptive methods offer the powerful tool for solving this problem because the both desired and undesired signals are rarely stationary in the real applications. One of them is very often used spectral subtraction method [1],[2]. We were interested in application this method for the hand-free mobile car telephony. The aim was to suppress a car noise in speech picked up by one microphone in a running car and to implement the improved algorithm [3],[4] on DSP. But this method requires reliable speech/pause detector [5],[6] because incorrect speech activity detection introduces errors in output enhanced speech signal. These errors are significant for lower signal to noise ratio. This paper describes the real-time implementation and results of modified spectral subtraction method used under real car noisy conditions.

Method Description

The block diagram of the implemented algorithm is shown in Figure 1. The simply form of modified spectral subtraction method [5] was used. The modification of this algorithm consists in the repetition of spectral subtraction with full wave rectification and in the frequency splitting which is omitted in this case.

In algorithm it is assumed that short-time stationary speech signal *s[n]* has been degraded by an uncorrelated additive noise *d[n].* The input signal *x[n]* is then

$$x[n] = s[n] + d[n] \tag{1}$$

In frequency domain taken by discrete Fourier transform

$$X[k] = S[k] + D[k] \tag{2}$$

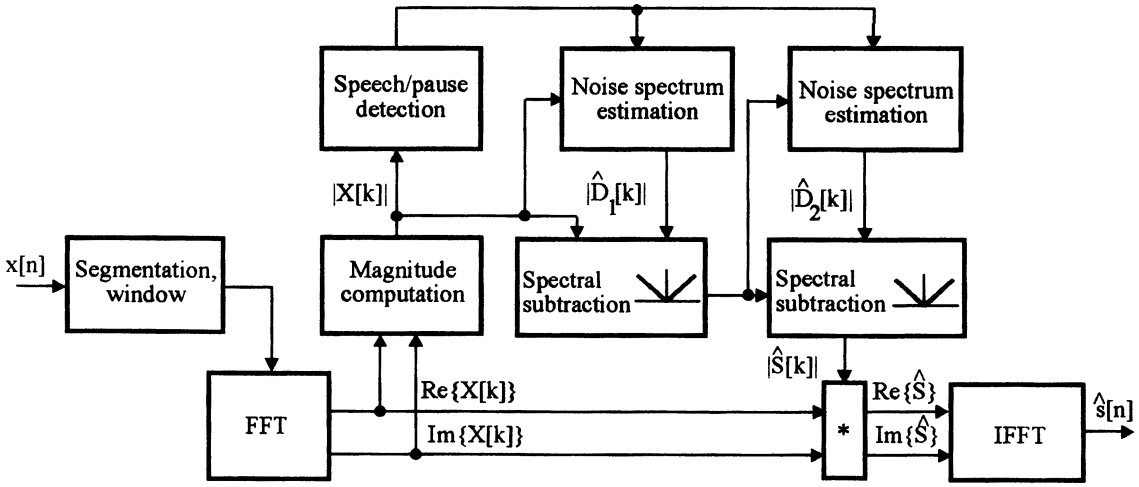
For speech enhancement based on short-time spectral amplitude the objective is to obtain an estimate $|\hat{S}[k]|$ of $|S[k]|$ and from this an estimate $\hat{s}$*[n]* of *s[n].* From the estimate $\hat{s}$*[n]* output speech signal can be generated in a variety of different ways. One approach is to use an analysis window function *w[n]* that generates s*[n]* when all frames of *s[n]* are overlapped and added with the appropriate time shifting, see the section, *Input-Output Data Communication*.

Various speech enhancement techniques differ primarily in how $|S[k]|$ is estimated from noisy speech. In the spectral subtraction technique is $|S[k]|$ estimated [1] by

$$|\hat{S}[k]|^a = |X[k]|^a - E\{|D[n]|^a\}, \tag{3}$$

where $E\{|D[n]|^a\}$ denotes the average of the noise magnitude spectra during nonspeech activity segments. For our algorithm was used the constant *a* = 1. $E\{|D[n]|^a\}$ is obtained from actual background noise in the interval where speech is not presented. The estimate (3) is not guaranteed to be nonnegative. The consequence of this negative components is the presence of a residual noise which can be heard as short tones with random frequencies (musical tones) in an output signal. To make the value on the right-hand of (3) positive is used the rectification. Full-wave rectification causes smaller musical tones than half-wave rectification but at the

Figure 1. Block Diagram of the Modified Spectral Subtraction



same time lower noise reduction. The full-wave rectification was implemented in our algorithm. Reasons for this choice are given in the references [5].

It is assumed that the short-time phase information is for the human ear not importent. Therefore each short-time speech segment $\hat{s}[n]$ is estimated by utilizing the phase associated with the input noisy speech segment [1],[2].

The magnitude spectrum of enhanced speech is estimated by repeated subtraction the estimated noise magnitude spectrum from the noisy speech magnitude spectrum.

This method requires reliable speech activity detector. We investigated some types of detectors, the energy detector and the cepstral detectors computed using FFT or using autoregressive method, coherence detector etc [5],[4],[7]. Energy and cepstral speech activity detectors were implemented and their performance in real-time was tested. The comparison of the detectors is included in the references [5],[7]. For implemented algorithm was chosen cepstral speech activity detector whose coefficients were computed using FFT.[6],[7]

Cepstral Detector

This speech activity detector uses cepstral coefficients

$$c[n] = \frac{1}{2}\frac{1}{2\pi j}\oint_c \ln\left[X\left(\frac{1}{z}\right)X(z)\right]z^{n-1}dz \qquad (4)$$

which contain the information not only about the energy of signal but also about its spectral content and that is why cepstral detector performance is significantly better than energy detector performance. Two types of cepstral detectors were tested - FFT based cepstral detector and autoregressive (AR) one.

Both algorithms can be described as follows:

1) Firstly, 11 coefficients *c[n],* for *n* = 0,..., 10 are evaluated
   a) either using the FFT followed by the IFFT of logarithm of magnitude spectrum
   b) or using Burg method followed by Levinson recursion and AR to cepstral coefficients transformation

   to form vector $c_i = [c[0], c[1], \ldots, c[10]]_i^T$, where *i* describes i-th signal segment and *T* denotes transposition. The decision results are better for the AR based detector. As compromise between decision accuracy and computation time requirements was implemented the cepstral FFT detector.

2) Second, the difference between the i-th cepstral vector c_i and the last estimation of noise background cepstral vector c_i' is computed as

$$d_i \approx 4.34\sqrt{\left(c_i[0]-c_i'[0]\right)^2 + 2\sum_{n=1}^{10}\left(c_i[n]-c_i'[n]\right)^2} \quad (5)$$

3) Finally, the distance d_i is compared to the threshold *thr* obtained using the estimation of mean value and standard deviation of d_i in the speech pauses as

$$thr = mean(d) + \lambda \bullet std(d). \quad (6)$$

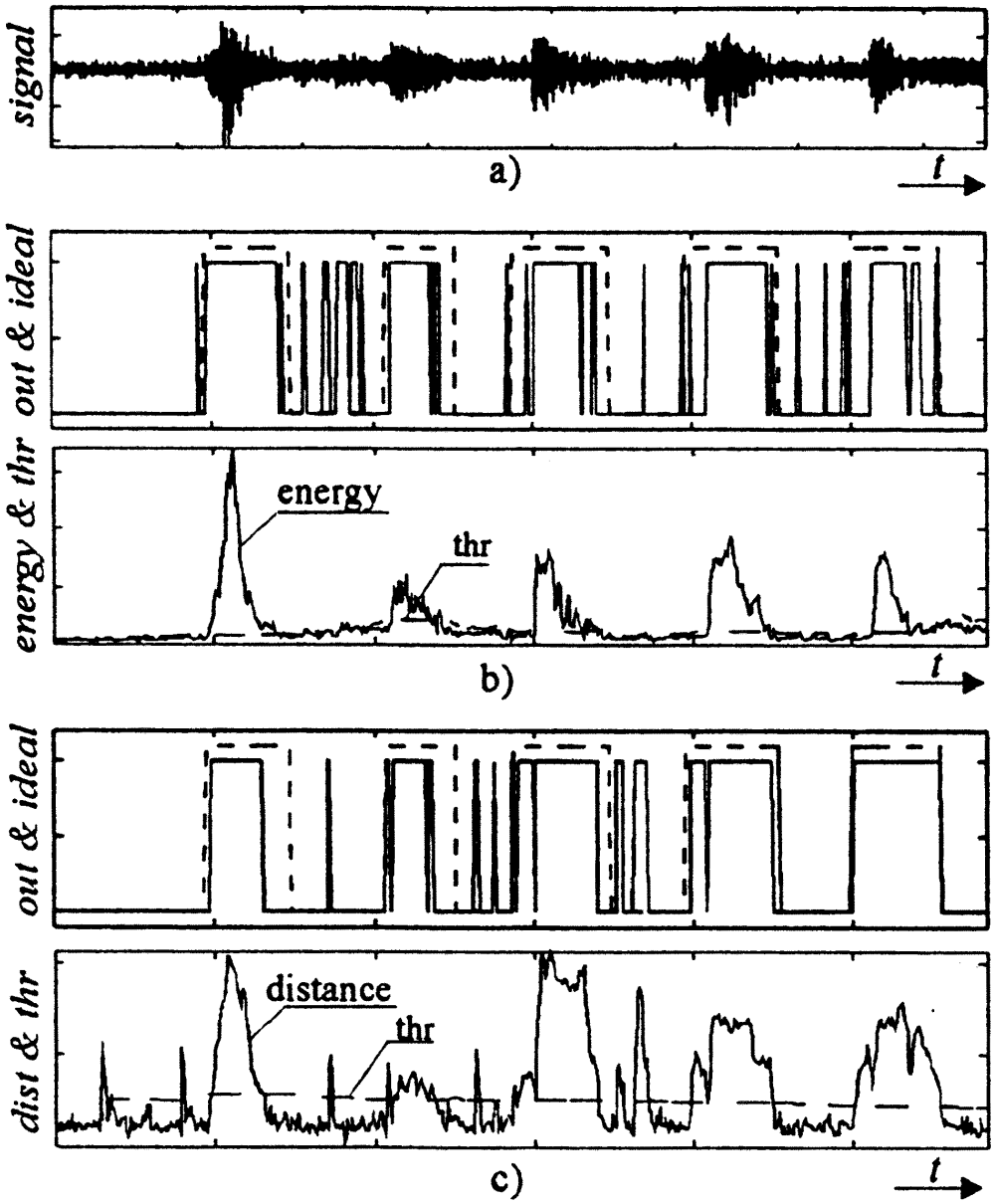
Analysis showed and real-time tests confirmed that the optimal value of λ is in the range from 1 to 2 depending on signal to noise ratio of noisy speech signal.

The coefficients of background noise cepstrum are obtained by exponential averaging in the speech pauses according to

$$c_{i+1}' = (1-\alpha)\bullet c_i' + \alpha \bullet c_i \quad (7)$$

where the optimal value of α was found in the range from 0.05 to 0.3.

Figure 2. Comparison of Isolated Word Detection a) Processed Signal b) Energy Detector c) FFT Cepstral Detector



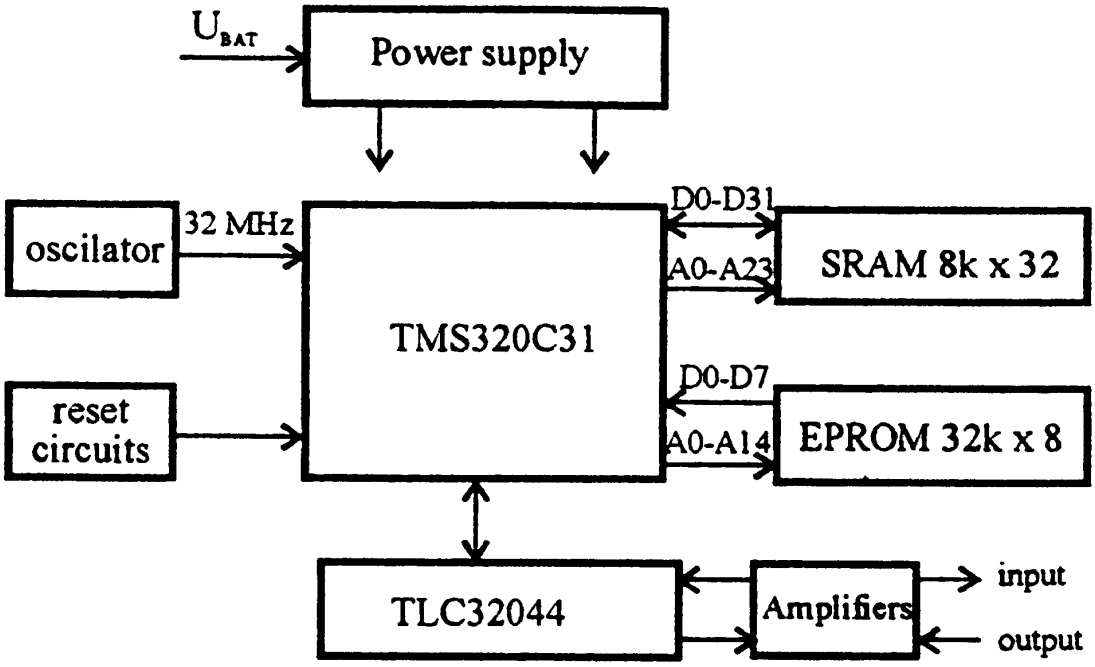
In Figure 2 is shown the operation of energy detector b) and implemented cepstral FFT based detector c). From the differences between ideal and real outputs follows our preference of cepstral detector.

Algorithm Implementation

The objective of our effort was to develop and implement computationally efficient and robust algorithm of spectral subtraction and test it in real-time on real speech signals, degraded by a car noise. The noise suppresser was implemented on floating-point signal processor TMS320C31. It was developed and realised a special board with DSP and supporting devices. The structure of this board is shown on the block schematic in Figure 3. The core of the system is DSP TMS320C31. The source code is stored off-chip in slower EPROM and by system reset is using boot loader loaded to off-chip RAM from which is executed. Analog interface circuit TLC32044 is connected to serial port of DSP and is applicated for speech signal input and output. Input clock for DSP was generated by on-board crystal oscillator operating at 32 MHz.

System needs two supply voltages, +5V is generated by voltage regulator 7805 and -5V by circuits LT1054. The voltage U_{BAT} can be from the range of 8V to 16V.

Figure 3. Block Structure of Developed Noise Suppressor

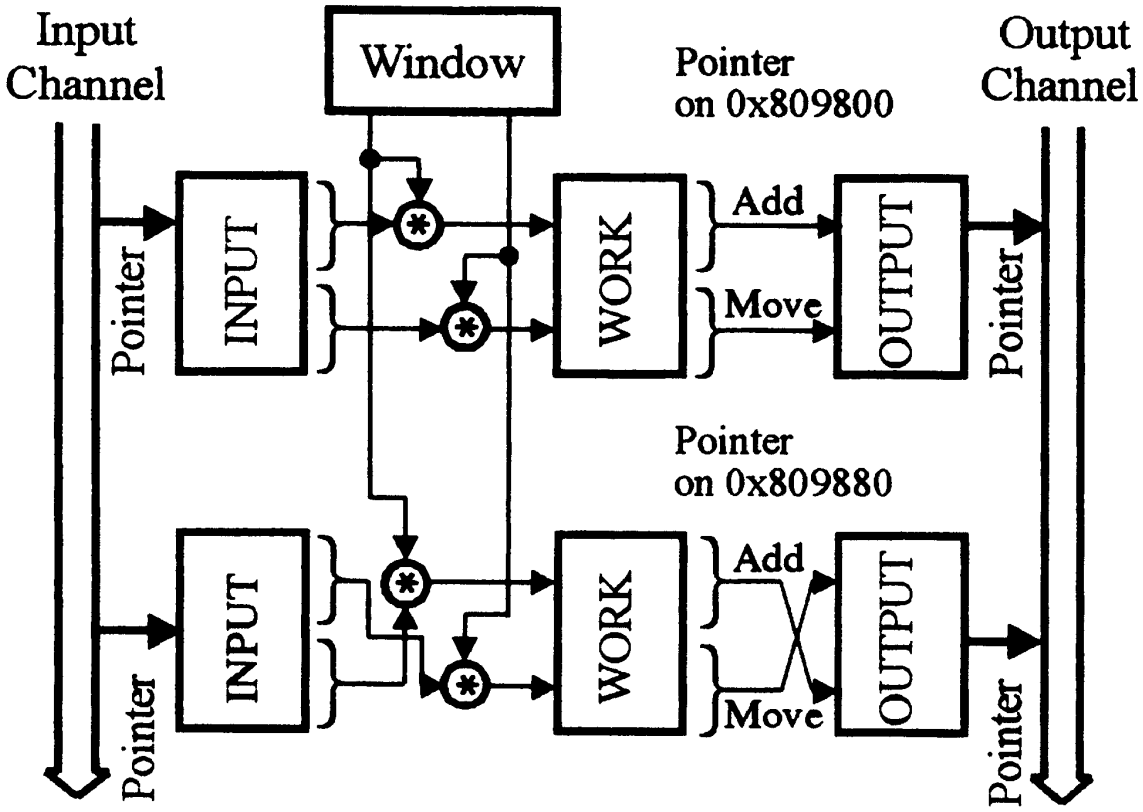


For the program debugging was used an evaluation module with TMS320C30. Following sections present specifications required for the implementation of the spectral subtraction system.

Input-Output Data Communication

Speech signal is sampled and converted by 14-bit A/D converter using IC TLC32044 with sampling rate setting in the range of 8 to 19 kHz. In implemented system was used sampling of 8 kHz. The communication between the converter and the signal processor is performed by the serial port. By every new sample from A/D converter is invoked the interrupt procedure which converts an integer sample to floating- point representation and saves it to the buffer INPUT in the internal RAM. The organization of data transfer and data buffering is shown in Figure 4. The integer value of processed sample is sent from the OUTPUT buffer to the D/A converter. The actual data positions in arrays INPUT and OUTPUT are the same and are activated by the same pointer.

Figure 4. Data Transfer Between A/D, D/A and Signal Processor (OLA Algorithm)



The data are segmented with the length of 256 samples and 50% (128 samples) overlap. We was tested also algorithm with 75% (192 samples) overlap. The output signal was similar, but the time for computing was two times shorter. Before FFT is each segment of the input buffer multiplied by the Hamming window and moved to the WORK array. The size of the FFT is the same as the length of windowed segments. All computations are performed on data in WORK sector of RAM.

Frequency Transformation

The complex spectrum $R_e\{X[k]\} + j\,\mathrm{Im}\{X[k]\}$ is computed for each data segment. The magnitude spectrum is computed in the effective way according to

$$|X[k] = |X[k]|^2 / \sqrt{|X[k]|^2} \ .$$

This equation was used because implemented algorithm computes the inverse value of the square root. The real and imaginary parts of $X[k]$ contain information about phase and they are saved for the output signal reconstruction.

For the frequency transformation the radix-2 FFT in-place algorithm was implemented. Since input data are real, the size of the FFT can be one half of the number of input samples and then the reduction of storage requirements can be achieved.

After the spectral subtraction procedure the module and phase of enhanced speech waveform are decomposed to the real and imaginary part of complex spectrum and transformed to the time domain using inverse FFT.

Speech Activity Detection

The correct decision of nonspeech segments influences the quality of noise suppression and the intelligibility of enhanced speech. In our implementation was used cepstral FFT based detector described in the *Cepstral Detector section.*

The cepstral coefficients are computed by DSP and stored in the WORK buffer. For the computation the magnitude of $X[k]$ is used the equation

$$c[n] = IFFT\{\log\{|X[k]|\}\}$$

The function *log*{} is approximated by using polynomials. Vector of 11 coefficients is used for binary decision: speech - nonspeech. It is made in three steps, described in the *Cepstral Detector* section: difference distance *d* computation, test *d <> thr,* speech/pause decision. The procedure is followed by an exponential averaging of background noise.

Computational requirements for the cepstral FFT based detector are shown in Table 1.

Noise Spectrum Estimation

The estimate of a noise magnitude spectrum is approximated by the exponential averaging of the signal magnitude spectrum $|X[k]|$ during nonspeech activity segments. The implemented algorithm requires twice repeated background noise estimation $|\hat{D}_1[k]|$ and $|\hat{D}_2[k]|$ as shown in Figure 1.

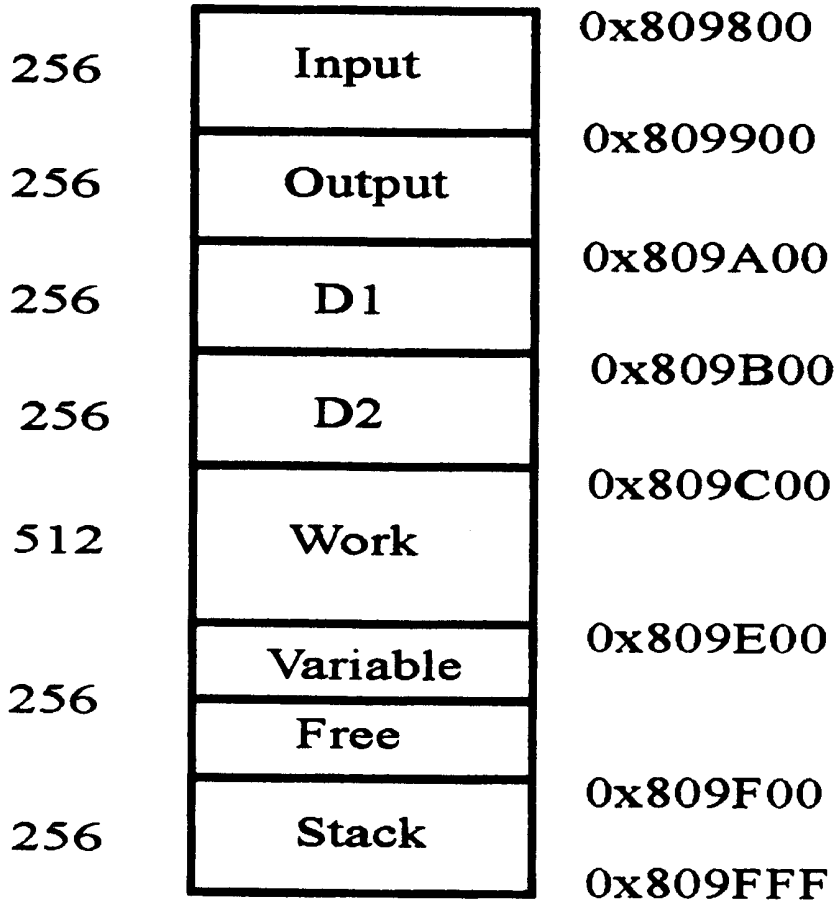
Unfortunately, this approach suffers from the fact, that the estimation of the noise spectrum differs from the instantaneous noise spectrum.

Internal Data Memory

The internal two blocks of 1k words RAM were divided in buffers for input data, output data and WORK registers. The memory map of this internal memory is shown in Figure 5.

To the sector INPUT are saved samples from A/D interface when the interrupt routine is activated. The data from this sector are moved to the sector WORK in the complete 256-sample segments, as shown in Figure 4. The data from the sector WORK are consequently processed segment by segment by the implemented algorithm. The processed segments are transformed to the time domain and transferred to the sector OUTPUT with 128 samples overlap. The movement of this data to D/A converter is realised by the interrupt procedure. The memory sectors D1 and D2 are used for the saving of noise spectra estimates.

Figure 5. Memory Map of the Internal RAM



Program Memory

The results of the program memory requirements are shown in Table 1. The implemented program code is stored in external EPROM type 27C256 organized 32 k x 8 bits. Using boot loader operations of TMS320C31 is source program loaded from EPROM in to external SRAM. With respect to selected EPROM configuration was specified 8-bits data transfer width format. For activation the boot loader function by processor reseting is driven the *MCBL/$\overline{MP}$* pin high.

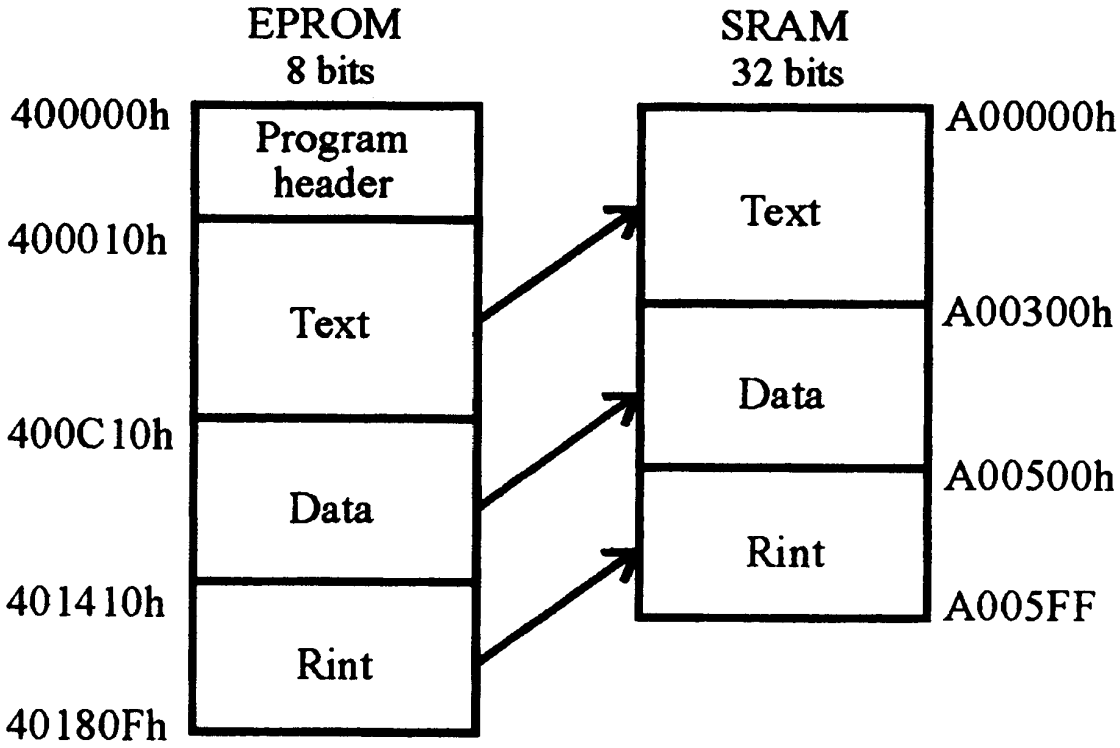
Every 32-bits word in SRAM is composed using four 8-bits words located in EPROM subsequently. On realised board is EPROM memory mapped starting at the address 0x400000h (external interrupt pin INT1 = 0 by reset). Address line A22 was used as output enable input of EPROM. A loader destination block SRAM begins at the address 0xA00000h. SRAM is composed by 4 devices HM65764 with access time t_{ac} = 63 ns organized 8 k x 8 bits.

At the addresses 0x400000h - 0x40000Fh is resided a loader header that defines

- memory width communication = 8 bits
- wait states = 4 (EPROM t_{ac} = 120 ns)
- program size = 0x7FFh
- program load starting address = 0xA00000h.

The program transfer from EPROM to SRAM with real addressed blocks is shown in Figure 6.

Figure 6. Memory Map of EPROM and SRAM



Computational Time Requirements

The instruction cycle time of an used signal processor is 63 ns for the clock frequency of 32 MHz. Between two successive data segments is the time gap of 16 ms for the sampling rate of 8 kHz. In this time gap all procedures have to be computed.

In Table 1 are presented results of signal processor timing. Computation times were measured on real signal using oscilloscope. The time interval of executed procedure was measured using I/O flag register XF1, which was seted by the operation starting and reseted when the operation was finished.

Because the implementation of spectral subtraction [4] works on signal segments of 256 samples, so the speech/pause detection is always required in the processed frame only.

Conclusions

In paper has been demonstrated that processor TMS320C31 can be effective used for noise suppression system based on spectral subtraction method. The efficiency of the instruction set coupled with speed of operation allow to implement the relatively simple suppression system working in real-time. The developed board has been constructed as universal DSP system applicable for the other algorithms in acoustic frequency range.

The modified algorithm of spectral subtraction for the reduction of acoustically added noise in speech was implemented. The algorithm was tested on actual speech corrupted by a car noise. The noise suppression of implemented system is about 5 to 8 dB depending on input noise and speech characteristic.

The implementation confirmed that the modified algorithm improves the reduction of musical tones and at the same time leads to the high noise reduction comparable with other used methods of spectral subtraction.

Table 1. Programme and Computational Time Requirements of an Implemented Spectral Subtraction Algorithm

Operation	**Progr. Memory**	**Operation Time [ms]**
System initialization	98	0.50
I/O data transfer and windowing	293	0.16
FFT, magnitude and phase computation	495	1.08
Energy detector	37	0.038
Cepstral detector FFT based	727	3.1
Spectral subtraction	65	1.04
Actual noise estimate	83	1.29
Re and Im part computation, IFFT	508	1.12
Interrupt procedure	34	0.003
The window with energy detector	1115	3.7
The window with cepstral detector	1293	6.6

Many types of speech activity detectors was tested under different input noisy conditions. The energy detector is very simple but it can be used only for nearly stationary or slowly varying nonstationary background noises. The cepstral detectors, FFT based and AR based, consumes more computational time, but better reacts to the changes of a background noise.

System on TMS320C31 is the first step in implementation. The more effective version would be a suppression system realized on TMS320C32 with 5 kwords internal RAM because off-chip RAM in this case does not necessary to use.

Future Developments

One channel methods of noise cancellation are simple to implement and in some conditions have good performance. The performance of these methods has strong limitations when non-stationary noises are met, nevertheless some noise suppression can be achieved even if for real car noises. The research works show that it is impossible by one channel method to achieve a marked improvement of noise cancellation.

Our research is now oriented on more channel methods, two channel method based on coherrence function and four channel method.

In the near future will be at our department implemented two new algorithms, echo-canceller algorithm for suppression of signal from loudspeaker picked up by microphones array and simply speech recognition algorithm for voice control dial.

References

[1] Lim, J.S. - Oppenheim, A.V.: Enhancement and Bandwidth Compression of Noisy Speech. Proceedings of the IEEE, vol.67, No.12, December 1979, pp 1586-1604.

[2] Boll, S.F.: Suppression of Acoustics Noise in Speech Using Spectral Subtraction. ASSP-27, No.2, April 1979, pp 113-121.

[3] Pollák, P. - Sovka, P. - Uhlír, J.: The Noise Suppression System for a Car. In. proc. of the 3rd European Conference on Speech Communication and Technology, EUROSPEECH '93, Berlin, Sept. 1994, pp 1073-1076.

[4] Davídek, V. - Sovka, P. - Šika, J.: Real-Time Implementation of Spectral Subtraction Algorithm for Suppression of Acoustic Noise in Speech, In: proc. of the 4rd European Conference on Speech Communication and Technology, EUROSPEECH'95, Madrid, Sept. 1995, pp 141-144.

[5] Pollák, P. - Sovka, P.: Spectral Subtraction and Speech/Pause Detection. Research Report, CTU Faculty of Electrical Engineering, Prague, 1995.

[6] Haig, J.A. - Mason, J.S.: A Voice Activity Detector Based on Cepstral Analysis, In: proc. of the 3rd European Conference on Speech Communication and Technology, EUROSPEECH'93, Berlin, Sept.1994, pp 1103-1106.

[7] Sovka, P. - Davídek, V. - Pollák, P. - Uhlír, J.:Speech/Pause Detection for Real-Time Implementation of Spectral Subtraction Algorithm, In: ICSPAT 95. The International Conference on Signal Processing Applications and Technology. Boston : DSP. 1995. pp 1955 - 1958.