Advances in Computer Science and Engineering
Volume 7, Number 1, 2011, Pages 91-97
Published Online: August 3, 2011
This paper is available online at http://pphmj.com/journals/acse.htm

THE METHOD OF A FAST FILTERING FOR NOISE REDUCTION IN AUTOMATIC SPEECH RECOGNITION SYSTEMS

N. M. BYKOV, A. RAIMY, K. KONATE and A. E. SEGEDA

Department of Computer Control Systems
Vinnitsa National Technical University
Vinnitsa, Ukraine
e-mail: nbbykov@mail.ru

Laboratory of Telecommunication
Faculty of Sciences and Technics
University Cheikh Anta Diop of Dakar
Dakar, Senegal
e-mail: araimy16@yahoo.fr

Laboratory of Telecommunications and Network
Faculty of Sciences and Technics
University Cheikh Anta Diop of Dakar
Dakar, Senegal
e-mail: kkonate911@yahoo.fr

Department of Computer Control Systems
Vinnitsa, Ukraine
e-mail: fate2107@yandex.ru

Abstract

The noise influence of hardware and environment on signals in information channel reduces the recognition accuracy of patterns, represented by these signals. The main methods of noise influence

Keywords and phrases: filtering, signal noise, approximation, the coefficients of the filter, computational complexity, optimized frequency response.

Received March 19, 2011

elimination on recognition accuracy are: the selection of invariant for noise features and also by filtration. Usage of the second method very much often results to unjustifiable computing complexity, as the efficiency of filtering for given pattern arrangement in feature space is not considered. That is why the application in many recognition tasks of not complicated filters can increase the recognition accuracy in the same measure, as filters with considerable computing complexity (for example, Kalman filters). In this paper, the simple method of a fast signals filtration which minimizes computing complexity is reviewed.

Introduction

The analysis given in the previous authors' works [1, 2] has shown that the filtration of noise in a speech can result in a considerable increase of the recognition accuracy under definite conditions. According to this, there is a vital problem of looking for the filtration methods which require hardware with minimum computing recourses or digital devices constructed on the basis of foolproof with high rate. In the interval of time, the filtration can be evaluated by the convolution of consistent signal samples with factors of a digital filter.

Mathematical Grounds of Proposed Method

Known classical methods of smoothing, which are approaching the signal using the set of functions [2] in speech recognition systems, cannot be applied for several reasons. First, the appearance of a curve, which should perform the approximation, is not known so far, and second, such a smoothing requiring information on all samples of discrete signal in real time is impossible. Under such conditions, rational is the construction of smoothing filter, which is approximated by a polynomial of the first degree that uses the sliding average. Number of samples in the "sliding" group is defined by the requirements for precision of smoothing and the complexity of the filter. Optimal smoothing as a criterion can be used, for example, a minimum sum of squared deviations of the real signal samples from the smoothing samples:

$$\sum_{i=1}^{m}(y_i - y_i^*)^2 = \min, \tag{1}$$

where y_i is the current value of the signal at the filter input in the *i*th point in time; y_i^* are smoothed current signal samples, and m is the number of samples.

Smoothing with a "sliding" group of three points from the condition (1) yields the following expression for the samples of smoothed signal:

$$y_i^* = \frac{1}{3} y_{i-1} + \frac{1}{3} y_i + \frac{1}{3} y_{i+1}, \tag{2}$$

where y_i^* are adjusted current signal samples and y_{i-1}, y_i, y_{i+1} are three consecutive time signal samples at the filter input.

However, the smoothing algorithm implementation under this formula is not optimal in terms of computational cost, since division operations require major time and hardware costs. Besides this, it is not optimized frequency response of the smoothing filter because it suppresses useful components that lie below the sampling frequency of signal. We show this for what we find *z*-transform expression (2), and then the filter transfer function:

$$Y(z) \cdot z^{-1} = \frac{1}{3} X(z) \cdot (z^{-2} + z^{-1} + z), \tag{3}$$

$$W^*(z) = \frac{Y(z)}{X(z)} = \frac{1}{3}(z^{-1} + z + 1). \tag{4}$$

Frequency response of this filter is

$$H^*(j\omega) = \frac{1}{3}(1 + e^{-i\omega t} + e^{i\omega t}) = \frac{1}{3}(1 + 2\cos\omega t). \tag{5}$$

Figure 1, where curve 1 represents the module of the filter frequency response, shows that the suppression of useful signal components is carried out in the frequency range of $\left[\frac{2\pi}{3}, \pi\right]$. Generally, the frequency response of device that realizes smoothing with "sliding" group of three points is as follows:

$$H(j\omega) = (a_0 + 2a_1 \cos\omega t). \tag{6}$$

Set the requirement such that at zero frequency of signal, the module of transfer function is equal to 1, and the highest frequency is equal to 0:

$$\begin{cases} a_0 + 2a_1 \cos\omega T = 1, & \text{if } f_c = 0; \\ a_0 + 2a_1 \cos\omega T = 1, & \text{if } f_c = \frac{1}{2} f_d = \frac{1}{2T}, \end{cases} \tag{7}$$

where a_0, a_1 are coefficients of smoothing filter, ω is angular frequency, T is sampling period, f_c is signal frequency, and f_d is sampling frequency.

In system (7), we find that $a_0 = \frac{1}{2}$, $a_1 = \frac{1}{4}$. Substituting these values in (6) gives the expression of transfer function filter:

$$H^*(j\omega) = \frac{1}{2} + 2 \cdot \frac{1}{4} \cos \omega t. \tag{8}$$

The module of the obtained filter frequency response is depicted with curve 2 in Figure 1.

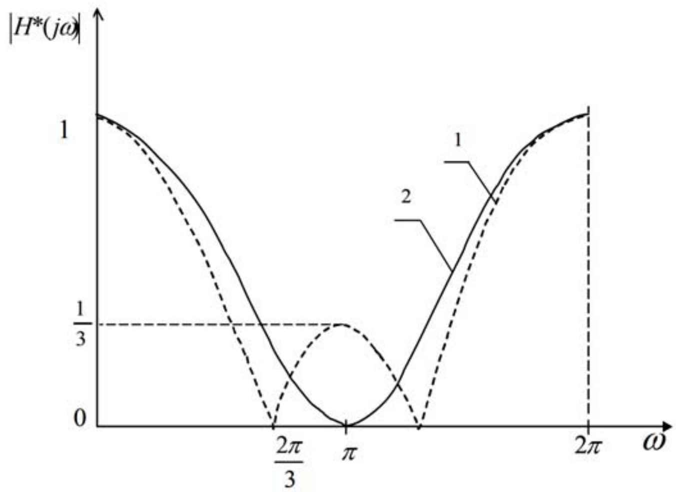

Figure 1

Filter's transfer function corresponding to this signal processing algorithm is as follows:

$$W(z) = \frac{1}{4} z^{-1} + \frac{1}{2} z^0 + \frac{1}{4} z^{+1}. \tag{9}$$

In the time domain, filter can be represented as a convolution of successive signal samples with digital filter coefficients:

$$y_i^* = \left(y_i + \frac{y_{i-1} + y_{i+1}}{2} \right) / 2 = \frac{1}{4} y_{i-1} + \frac{1}{2} y_i + \frac{1}{4} y_{i+1}. \tag{10}$$

Let us compare the rate and effectiveness of the proposed filter with the same characteristics of optimal three-point filter. Here the optimality is estimated due to criterion (1).

Filter noise reduction is determined by the sum of squares of its coefficients, therefore the loss of efficiency filtration E_f in filter (10) compared with the filter (2) determines the expression:

$$E_f = \frac{\sum_{i=1}^{m} k_{ip}}{\sum_{i=1}^{m} k_{io}} = \frac{(1/2)^2 + 2 \times (1/4)^2}{3 \times (1/3)^2} = 1{,}125. \tag{11}$$

where k_{ip} are coefficients of the proposed filter and k_{io} are coefficients of the optimal filter.

Filter (10) implements the calculation of coefficients by successive shifts of the signal samples and thus excludes the operations of multiplication and division required for the filter (2). Simple calculations show that the filtration rate is increasing in the order and above. Thus, our proposed filter has an order of magnitude higher performance compared with the optimal filter with very little loss of efficiency filtration. Such a non-recursive filter can be implemented in hardware as a chain of buffer registers, whose outputs are connected to the input of the adder (Figure 2).

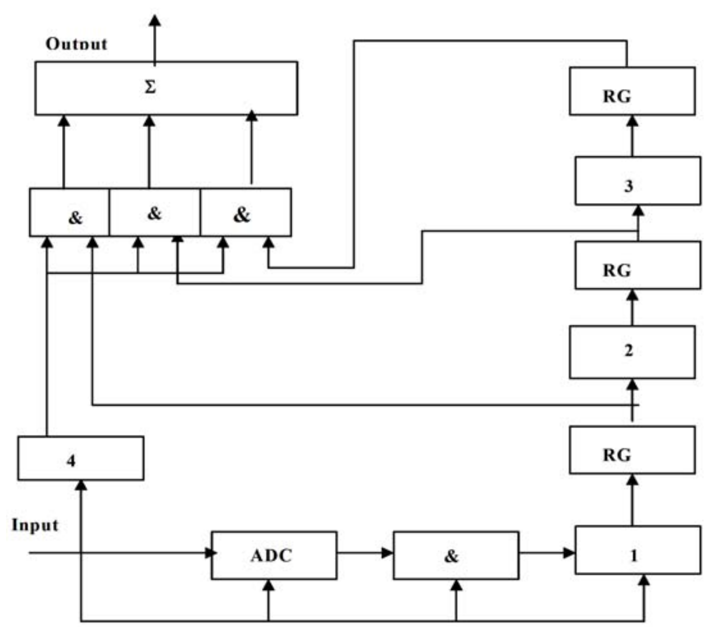

Figure 2

Following designations are taken in Figure 2: ADC - analog to digital converter, RG - registers, 1 - the chain of binary word shifting two bits to right, 2 - chain shift one bit to the left, 3 - chain shift one bit to the right, 4 - block for permit of withdrawal information. Operation of device is as follows:

Values of the signal, measured in discrete time, are converted by the ADC in a digital binary code (binary word). Also, at the same time points, ADC produces clock pulses to synchronize the entire device. Of binary words, that define the signal samples, from the register to register, they are shifted by shifting chains so that they are equal to $y_{i-1}/4$ and $y_{i+1}/4$, respectively, and the contents of the second register is equal to $y_i/2$.

Smoothed signal samples are obtained on the adder output by summation of the contents of three registers in each of clock pulses. Output permission unit 4 in the initial time prevents reading the contents of registers prior to completion of all three registers. Offered filter hardware realizes calculations of coefficients by shifting consistent signal samples already during their transmission from the register to register. Because of these operations of multiplication and divisions, indispensable for the filter (2) is eliminated. Since there is no need of performing of division operations, the arithmetic device is considerably simplified compared to the arithmetic unit to perform calculations for smoothing formula (2), because its composition does not need to introduce the device to perform division.

Winning the performance of the proposed smoothing filter versus filter (2) can be calculated by the following formula:

$$e_w = \frac{k_d(m_d - n_d)T}{T}, \tag{12}$$

where k_d is number of microoperations in firmware, which implements the algorithm of division, m_d is number of bits in dividend, n_d is number of bits in divisor, T is clock pulse duration in microoperation. Let $m_d = 8$, $k_d = 6$, $n_d = 2$ (bit of number 3), we obtain $e_w = 6(8 - 2) = 36$.

Thus, the filter is obtained by taking into account conditions for optimizing the frequency response having an order of magnitude higher performance than known. This allows us to use it to smooth the signal in real time.

On the basis of the theory of function approximation [3] and theory of digital filters [4], we can show that, by serial and parallel connection of such basic filter, fast filter of desired characteristics can be constructed. The evidence of this position leaves for frameworks of this activity.

References

[1] N. M. Bykov, I. V. Kuzmin and L. V. Proshenko, Mathematical model of the influence of noises on the precision in speech recognition, In the book: ICIM-2002, International Conference on Inductive Modeling, Lvov, 20-25 May 2002 T1, Vol. 2, Lvov, 2002, pp. 287-292.

[2] N. M. Bykov, Methods and tools for information measuring and processing in computerized speech recognition systems, Thesis for the Grade of Candidate in Technical Sciences, Vinnitsa, 1985, p. 243.

[3] L. R. Rabiner, B. Gold and C. A. McGonegal, An approach to the approximation problem for nonrecursive digital filters, IEEE Trans. Audio and Electroacoustics AU-18(2) (1970), 83-106.

[4] A. V. Oppenheim and R. W. Schafer, Digital Signal Processing, Prentice-Hall, Inc., Englewood Cliffs, New York, 1975.