

Computer Utilisation for Speaking Re-education

Radu V. Ciupa, Alexandra A. Ciupa, Emil Simion

Technical University of Cluj-Napoca
3400 Cluj-Napoca, Romania

Abstract. The purpose of the research was to study the problems which exists in the speech analysis domain and to work out a system through which speaking ability regained for this kind of patients was facilitated as far as speed, quality and reduction of useless staff in the educational process are concerned. The device was conceived and carried out with a double-aimed role: -a computer interface, for the analysis of the quality of uttered words and the display on a video screen of the results of comparison in graphical form proper for the usual feedback related to pronunciation correctness and, -a system of auditive feedback carried out with the same aim.

1 Introduction

The loss of speaking ability, the difficulties encountered during the speech learning of the handicapped children, mainly those having a poor hearing or with acquired deafness are a source of serious psycho-social disbalance for the patient.

The purpose of the research was to study the problems which exists in the speech analysis domain and to work out a system through which speaking ability regained for this kind of patients was facilitated as far as speed, quality and reduction of useless staff in the educational process are concerned.

The main objectives aimed during speech analysis are the following [1]: a) an efficient codification of voice signal in order to transmit or record it, b) synthesis of voice signal; there could also be synthesised vocal messages with predetermined content that have been recorded after a more or less short treatment, c) recognition of speech as an essential base (as well as the synthesis) in the development of a system of man-machine dialogue.

2 Background

In its simplest form, direct codification consists in the representation of each sample of signal or of its spectrum irrespective of all the other samples. By speech compression it may be considerably reduced the volume occupied by the spoken signal thus adapting it to the diminished perceived volume of an auditory handicapped. In the system of parameter compression, speech is submitted to an analyser in order to

extract a number of essential parameters. With their help, speech is then recognised or restored by a complementary synthesis operation.

The considerable development of micro-electronics enables now a very complex numerical treatment of the signal, even in the real time. The numerical representation can be made both in the time and frequency domain. The latter frames the sub-band coding (SBC) system too [1]. This coding system consists in dividing the signal spectrum in a certain number of sub-bands of width B_k be they equal or not. The whole made up from them reconstitutes the original spectrum, in principle. The minimal sampling frequency is $2B_k$.

The signals supplied by each filter (Fig.1) are quantified so as to equally distribute the quantifying noise as function of frequency for a certain binary rate. The quantification must be finer the lower bands in order to maintain the formants structure and fundamental component. The signals are encoded (with Adaptive Differential Pulse Code Modulation), multiplexed and transmitted.

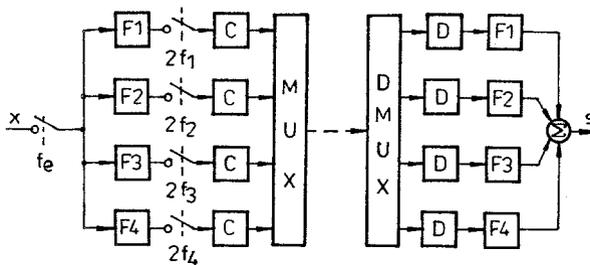


Fig. 1. The sub-band coding system

The decomposition in sub-bands is obtained with a set of filters which constitute the main system element and which can be also met in the receiver where the vice-versa process, that of de-coding takes place. As each filter presents a non-null transition zone, two consecutive bands cross each other in order to avoid lacks in spectrum recovering. A re-make of the spectrum of each sub-band after re-sampling is the consequence of this operation.

3 Development

The solution we chose takes into account the double aim of the device, i.e. that of auditive feedback or interface for systems provided with visual feedback. The block diagram is presented in Fig.2 [2]. The information received from the external loudspeaker is applied to the input amplifier I.A. (Fig.3). Diodes D_1 and D_2 are protection diodes; if the input signal amplitude exceeds 0.6V the mass short-circuiting takes place and the protection of IC_1 is the results. The potentiometer P_1 , ensures the control of the input level amplitude. R_1 limits the signal coming from the microphone. C_1 is a DC separation condenser. Condenser C_4 and resistance R_6 constitute a low pass passive filter. Condenser C_5 plays the part of a DC separator between the input amplifier and the analog/digital converter. The amplified signal is passed through a passive filter RC and reaches the input of the analog/digital converter.

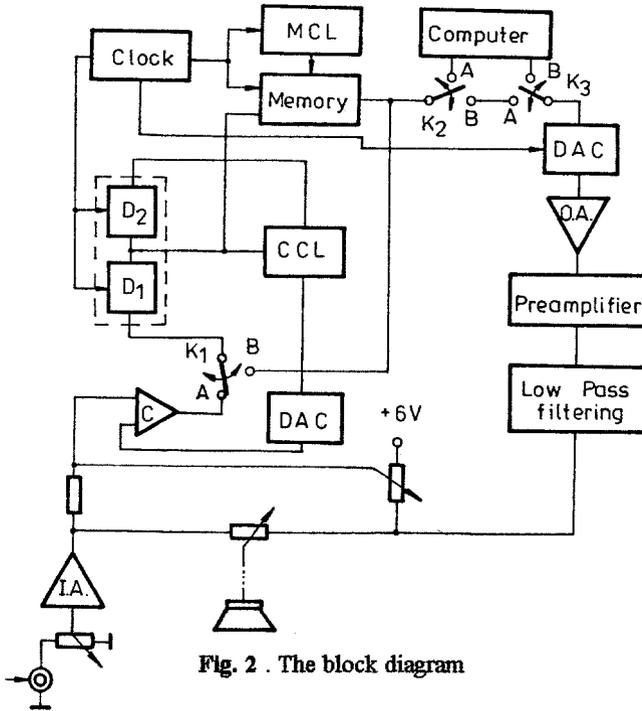


Fig. 2 . The block diagram

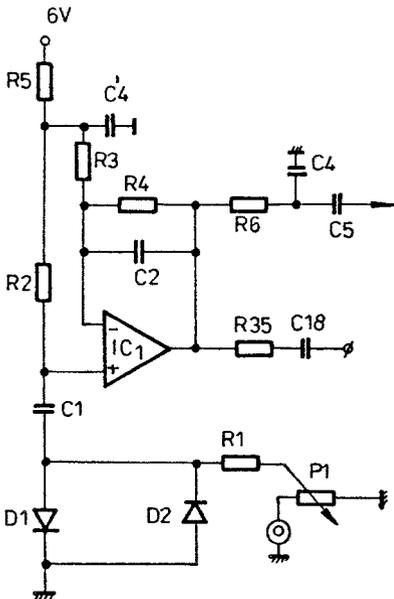


Fig. 3 . The input amplifier

The comparator C, the displacement register of 2 bits D_1 , D_2 , a digital/analog converter (DAC) and its Control Logic (CCL) are also parts of its structure. The actual bit from the latch D_1 is stored in the internal memory of the device ($M=6 \times 64$ kbits) at each clock tact.

Some technical characteristics of the memory unit components: maximum access time-150ns, read/write cycle-300ns, read/modify/write cycle-345ns. Due to the multiplexing of 16 address bits in the 8 address pins a high storing density is allowed. The two address level are ensures by the two TTL clocks: Row Address Strobe (RAS) and Column Address Strobe (CAS). After one reading or a valid read/modify/write cycle, the information is maintained on D_{out} through the CAS transition in 0. The Data Output is brought back in a state of high impedance by the returning of CAS in 1.

The circuit can maintain the information at the output by keeping CAS in 0 while RAS is used for refreshing (RAS - only cycles). The refresh takes place simultaneously with the development of RAS only cycles or ordinary cycles read/write through those 128 $A_0...A_6$ address combinations during 2ms. The memory clock generator for the memory control logic (MCL) and displacement registers of 2 bits is given by an astable provided with invertors.

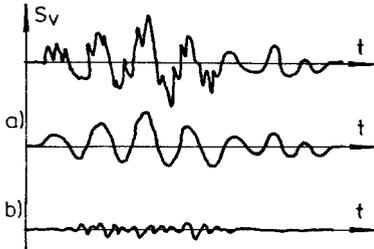


Fig. 4 . The signal oscillogram

The analysed signal oscillogram as well as the low (a) and high (b) frequency components can be seen in Fig.4. As the variation of the fundamental tone of the speaker causes a change in the spectral frequencies, when designing the filters, it has been taken into consideration a certain type of patient, having his fundamental tone round 100 Hz.

Each filter may be introduced in, or taken out from the circuit with the help of a potentiometer

which can also diminish, more or less the signal from the respective spectral region. Thus we may provide an artificially amplified or attenuated signal in certain frequency bands compared to the normal signal, a fact necessary in certain cases of poor hearing. Being assured the necessary synchronism between MUX control and scanning frequency on the horizontal, the output of each filter will be selected at a given moment and on a well established period of time within a TV line. In this case the length of the vertical bands that appear on the display is proportional to the amplitude of the corresponding spectral components. This relation of proportionality is carried out by the comparator block COMP to whom there are applied a multiplexed signal at the measure input and a saw-tooth voltage at the reference input. The syncrogenerator block provides a set of auxiliary TV signals necessary for a proper operation of the TV monitor. The complex video signal results from adding the useful signal to the auxiliary TV signals in the adder-formative block, in which the required levels of these signals are also set.

Three types of programs have been elaborated, namely: acquisition of speech parameters, learning the typical parameters by a speaker and recognition of the learned phonemes. For the duration of speech, the acquisition program undertakes a set of values of the detected voltages at the output of filters every 2ms. Then, by mediating the sets of adjacent values in time, a numeric sample of the frequency characteristic is obtained. Based on these acquired values there has been tested the codification by delta modulation. In any case, sub-band coding is more complex than the adaptive delta modulation [3].

4 Results

The device was conceived and carried out with a double-aimed role: - a computer interface, for the analysis of the quality of uttered words and the display on a video-screen of the results of comparison in graphical form proper for the usual feedback

related to pronunciation correctness and - a system of auditive feedback carried out with the same aim.

The process of speech re-education of the clinical cases showed the fact that a repeated pronunciation of model-words or words uttered by patients is beneficial due to a better concentration of the patient and his affective-volitive involvement in regaining a proper speech. More than often, especially patients of an educated extraction behaved as inhibited or refractory patients due to either their embarrassment because of an incorrect utterance in front of strangers -be they re-educators, medical staff or their comparison to little children learning to speak. The proposed system offer the patients the possibility of comparing their pronunciation with the correct one, thus eliminating the physical participation of the medical staff for longer periods of time during their logopaedic treatment.

References

1. R. Boite, M. Kunt: Traitement de la parole. Lausanne, Presses polytechniques romandes, ch.3.2, pp.55-59, ch.3.5, pp.89-100, 1987
2. R. Ciupa: Method and device for speaking re-education. Proceedings of the VI th Mediterranean Conf. on MBE, Capri, Italy, vol.II, 1297-1300, (1992)
3. J.L. Flanagan: Speech coding. IEEE Trans. Com-27, no.4, 710-736, (1979)