THE SPEECH LAB

Jan Sedivy, Jan Uhlir

Czech Technical University, Prague
Department of Circuit Theory

Introduction.

This contribution describes a simple programming environment called Speech Lab (SL). The system is designed for personal computers operating under CP/M-80 or MS-DOS.

The SL was designed at the Department of Circuit Theory of Czech Technical University in Prague for the speech processing. The scope of task being solved is very large. It includes the basic signal processing algorithms, wave-shape coding methodes, LPC analysis etc. Recently a simple knowledge base was added to support works in speech synthesis.

SL is used in all speech oriented works in our department. It is used by students, by after graduate students and by staff of the department. It enables easy way of data and program exchange.

The SL structure.

The SL is structured programing environment. It consists of the following building blocks:

    User interface
    Data Acquisition system
    Data Processing Package
    Graphics Package

The SL is controlled trough the User Interface. It is made up from powerful commands. Commands can be divided in following groups, which correspond with the building blocks:

    Database Commands
    Data Acquisition Commands
    Data Processing
    Graphics Commands
    External Procedures
    Help Menu System

The first set of commands are database commands. They are used to display the database records, to retrieve, erase and update records. Other commands are devoted to get and put the data on the disk. Special command serves to import and export ASCII files.

The Data Acquisition Commands cooperate with the D/A and D/A convertors. They assure a fast data acquisition and the immediate check of processed speech. A simple command "SPEAK" is very useful in many speech applications.

For simple data processing are available many commands. Some of them are build-in and some are external commands. They are used for signal processing.

A very important feature of the SL is the simple ability to make up own procedures. The whole SL is written in TURBO Pascal [2]. This implementation of Pascal programing language is very popular in the PC compatibles and CP/M-80 world. The user created procedures written in TURBO Pascal can be called from SL with single command. They can use the same data as SL. It is an easy task to go back to SL after processing the data in a common block. This is accomplished using the TURBO Pascal command chain. In this way the user can use all the SL commands and create his own procedures. The burden of all data housekeeping and many processing is minimized.

User interface is equipped with large help menu system. This help can be called any time.

The SL uses a 20k or 32k long buffer for the data storage depending on the particular implementation. This buffer is used as above mentioned common block which is used by both SL and user-written procedures. To simplify the orientation in data, the buffer is divided in particular number of 128 bytes long sectors. Every byte in any sector is user accessible. The same structure of the data pointer is used in all commands. The data pointer is composed from two numbers. The first number is used for a sector and the second for a byte in the sector. We found this type of pointers useful for the sort of applications the SL was designed for.

The data structure.

The data in the buffer are speech samples or any other type of data. For example the LPC coefficients, spectrum, histogram etc. This data can be processed and then graphically displayed on the screen. There are available different graph shapes for the spectrum, histogram or the speech time series. User can chose to display the data in any shape. The graphs can be print out. Simultaneous display of different records is possible.

The file structure.

The SL record consists from two files a data file and a dictionary file. The dictionary file carries the information about the data like the sampling frequency, the date of creating the record, remarks etc. The data file contains pure data. Both files are automatically maintained using the SL commands. The dictionary file is short and it is presented in the SL on line to speed up the directory operation. The data file is updated after any changes has taken place.

A/D D/A converters.

The SL system enables an easy installation of A/D and D/A drivers. The drivers must be written in assembly language and must be patched in the SL. The SL supports any sampling frequency for data acquisition. We usually use double sampling frequency then required, to decrease the analog filter requirements. The final data are the digitally filtered with FIR filter. Simultaneously is the signal decimated in frequency. In this way the linear phase response can be preserved.

Graphics.

The SL uses Graphical Commands to display the contents of the BUFFER on the screen. The way how it is displayed depends on the hardware abilities. The graphical system is a separate part of the SL. It is written for different graphical systems. The CP/M versions are not so rich in graphical abilities compared to MS-DOS version. Both versions give the same type of the printer output.

The user can chose a proper shape of the graph different types. Shapes are designed to display a speech wave-shape, the power spectra, the histogram, the autocorrelation function etc. In the graphical part of the SL are also included commands to display a difference of two signals.

Graphs can be send to the printer or plotter.

Data processing.

The SL has build-in the most important signal processing procedures. They can be applied to active buffer. These procedures are:

    Histogram
    Autocorrelation
    Windowing
    FFT

External procedures are in the processing library. It contains a large set of algorithms for spectral and LPC analysis.

Different algorithms for speech wave-shape coding can be called. For example different types of PCM and DPCM quantization, adaptive versions of PCM and DPCM etc.

We also developed a package of clustering algorithms for the purpose of vector quantization of descriptors or wave-shape of speech.

For the purpose of synthesis by rule we developed a special package with a simple knowledge database. It is used for development of synthesizers with limited number of words and high quality of speech. The knowledge database simplifies the construction of similar sounds.

User written-subroutines.

A very important feature of the SL is the ability to include a user-written subroutine and to use the facilities of the system too. The whole system is written in Turbo PASCAL. The data buffer can be equally accessed from the main SL system and from the user written routine. The user-written routine can be debugged separately in Turbo Pascal environment and then translated as a CHN file and executed from SL. This simplifies the development and debugging very much. The user-written procedure is invoked from the system command line.

The system is designed to receive different data in the form of ASCII string. This way of transfer represents an easy link to large set of other programs written in different languages. These commands are used to transfer data from other programs. For example with SL we use the standard library of signal processing programs for digital filtering, FFT etc. [1].

Conclusion.

The SL system was successfully used in our department to solve various tasks. The SL is a very simple system but gives the user many capabilities and simplifies the

overall development.

References.

[1] Digital Signal Processing Committee of
    IEEE Acoustics, Speech, and Signal
    Processing Society, "Programs for di-
    gital Signal Processing", IEEE Press,
    1979.
[2] TURBO Pascal Reference Manual Version
    3.0. Borland International Scotts
    Valley, California, 1985.