AN ILS COMPATIBLE WIDE BAND SPECTRUM ANALYSIS/PLOTTING PROGRAM (WBS)

R.A. Wills and Y.C. Tong

Department of Otolaryngology
University of Melbourne

ABSTRACT - The software tool WBS which plots spectrograms from digitally recorded ILS compatible speech files is presented. The program operation is explained and comparison shown for a speech file plotted on a postscript printer with selected frequency/time resolution and the normal spectrogram plot from a "KAY" machine.

INTRODUCTION

Speech analysis is performed by the examination of spectral and temporal information in the recorded signal, of the many ways to present this information the more easily understood presentation is in the form of a spectrogram plot. However, the spectrogram machines are traditionally analogue machines and cannot accept digitally stored input signals. Therefore a software tool which operates on digital input and integrates in to an ILS environment, has been developed. The program is called Wide Band Spectrogram (WBS) and has adjustable frequency and time resolution settings and is able to produce high quality spectrogram presentations using high resolution postscript laser printers.

THE WBS PROGRAM

WBS is executed as any other ILS command and has its own unique set of command line parameters. These parameters allow the analysis plot to be configured at run time and allow the user to specify; (i) the ILS starting frame, (ii) number of frames, (iii) bandpass filter band width, (iv) Bandpass filter centre frequency spacing, (v) time resolution, (vi) plotted time axis range, (vii) maximum plotted frequency axis range and (viii) the functional relationship between the rectified filter amplitude and the plotted gray level.

Basically the program passes the speech data file through a bank of bandpass digital filters, rectifies the output of each filter, averages the output for the time resolution selected, scales this averaged value against the previously calculated maximum value, maps the scaled value from amplitude to gray scale level and then writes postscript commands to an ascii file for output to a postscript printer.

The speech data file used is the primary data file as defined by ILS's FIL command and is read with calls to GETD, which is an ILS callable subroutine that returns data file buffers at successive context (ie. number of sample values) increments. After this buffer is returned it is passed through the ILS windowing subroutine WINDW to perform the pre-emphasis weighting if this has been set.

The filters used are fourth order Butterworth digital filters, whose coefficients are designed with the ILS program DFI at the start of the program, using the centre frequency spacing and band width specified by the user on the command line. All filters used are bandpass filters except the highest and lowest frequency bands which are high and low pass filters.
The time resolution of the plot is defined by setting the size of the buffer returned by GETD to be equal to the context, which is also specified by the user on the command line. The buffer returned by GETD is passed through the digital filter, then rectified using the ABS function and averaged to give a single data value.

After the data file has been prassed through the filter bank. The maximum averaged amplitude value in the time-frequency range to be plotted is calculated. The program then uses the maximum value found on the first pass, together with the black saturation and gray shading dB levels specified by the user, to map each amplitude value to the actual gray shade of a rectangular area drawn on the paper. The spectrogram is drawn in the form of many small rectangular areas filled with a particular gray shade. The dimensions of this rectangular area is determined by the filter bank centre frequency increment and the time resolution selected by the user.

The functional relationship between the gray scale plotted by the laser printer and the average rectified filter amplitude is;

$$\text{while } max <= x > max*10^{**}(-b)$$

$$f(x) = 0.0$$

$$\text{and while } max*10^{**}(-b) <= x > max*10^{**}(-(g+b))$$

$$f(x) = 1.0 - [\log(x) - \log(max*10^{**}(-b)) \\ /(\log(max*10^{**}(-b)) - \log(max*10^{**}(-(g+b))]$$

otherwise $f(x) = 1.0$ and is not plotted.

where: x=the current averaged rectified filter amplitude,
max=the maximum average rectified filter amplitude,
b=the user specified black saturation level (dB),
g=the user specified gray dynamic range (dB),

Note:(1) the gray level f(x) used here is between 0.0 (black) and 1.0 (white).
(2) all "log" symbols used in the above equations are logarithm to the base ten.

EXAMPLES

Figures 1 and 2 are spectrograms of the utterance "Buenos Airis". Figure 1 is a spectrogram produced on a "KAY" machine and Figure 2 is the WBS spectrogram plotted on an Apple Laser printer.

The two figures are very similar in information content. However, the recording for figure 1 was produced on a high quality NAGRA reel to reel tape recorder using a good quality microphone, while figure 2 was produced on a medium quality cassette recorder and microphone. This explains the greater amount of unwanted high frequency noise produced in the cassette in figure 2.

ACKNOWLEDGMENTS

The WBS program was developed in collaboration with Dr. J. B. Millar, of the Research School of Physical Sciences, The Australian National University.

The spectrogram in Figure 1 was produced on the KAY machine, of the School of Communication Disorders, Lincoln Institute.

REFERENCES

ILS (or Interactive Laboratory System) is the trade mark of Signal Technology Inc. California, USA.

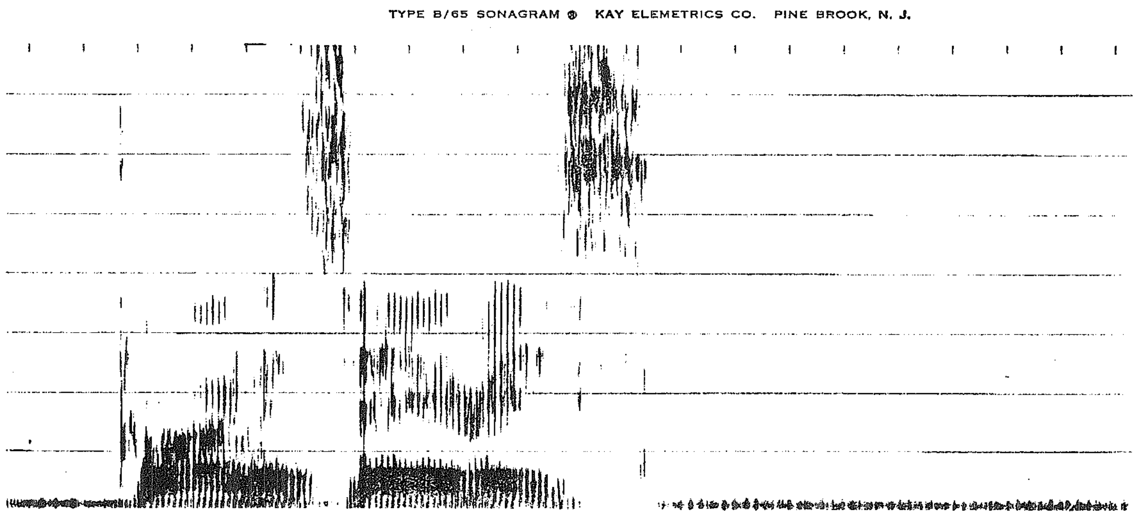

Figure 1. "KAY" spectrogram.

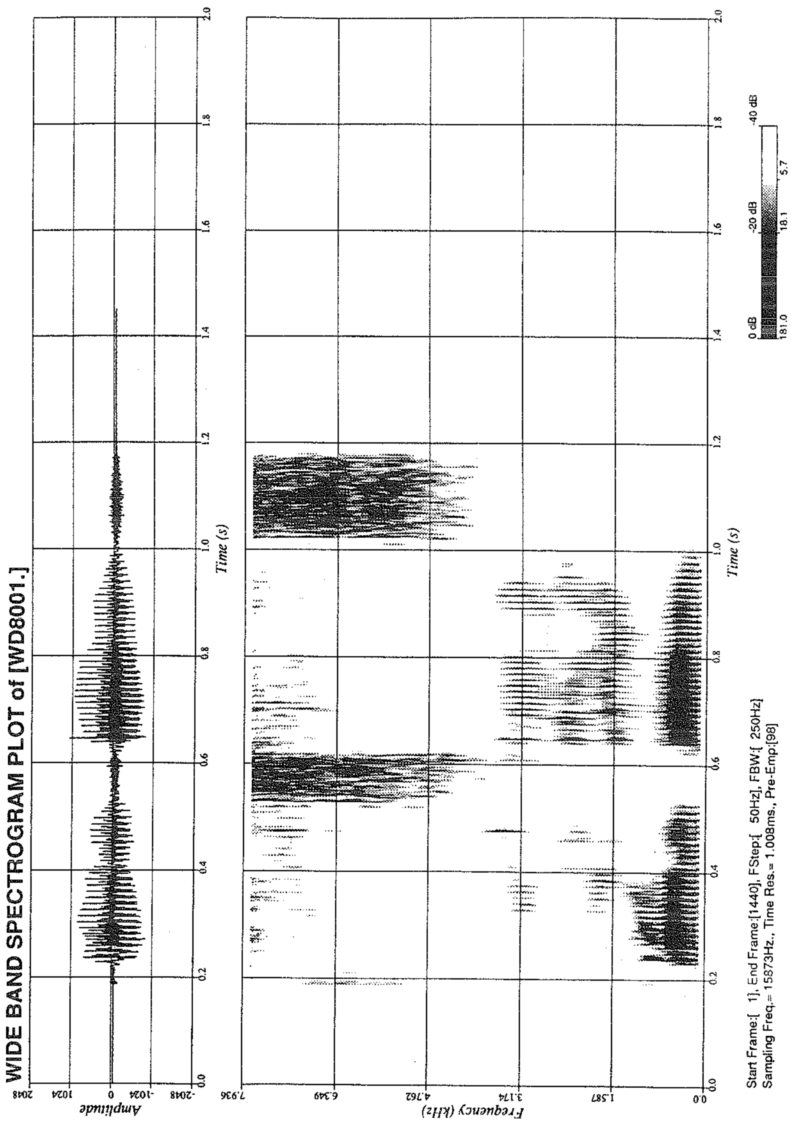

Figure 2. "WBS" spectrogram.

AN INTEGRATED AUDIO SIGNAL INTERFACE FOR USE IN THE TEACHING LABORATORY

Arthur Lagos and Michael Wagner

Department of Computer Science
University College
University of New South Wales

ABSTRACT- An audio signal interface for the IBM PC-AT is described which was specifically designed for student use in the teaching laboaratory. The interface which is implemented as an IBM PC-AT plug-in board allows the recording and playback of audio signals directly to and from disk files. All functions of signal conditioning, data conversion and the DMA bus interface are integrated on the board which provides for microphone/line inputs and headphone/line outputs.

INTRODUCTION

When a course in *Computer Speech Processing* was first planned at the Computer Science Department of UC/UNSW for 3rd- and 4th-year computer science and engineering students, the question of a suitable laboratory setup needed to be addressed. The open architecture of the IBM Personal Computer in conjunction with the program development environment provided by MS-DOS and C, the graphics facilities of the Extended Graphics Adapter and, very importantly, cost considerations led to the decision to establish a computer speech laboratory based on6 IBM PC-ATs with full analog input and output facilities for acoustic signals.

The specific requirements of a teaching laboratory called for an integrated analog interface with the following characteristics:

- the board is powered by the PC;
- all signal amplification and low-pass filtering in both input and output directions is contained on the interface board without the need for any external devices or cables;
- for analog audio input to the interface, either a microphone or a taperecorder is connected through standard audio connector plugs;
- analog audio output from the interface is provided for headphones, for a taperecorder or both and connections are made through standard audio connector plugs;
- the board has 4 connector jacks for MIKE-IN, LINE-IN, PHONES-OUT, and LINE-OUT;
- the sampling frequency is fixed at 16,000 samples/s, the cut-off frequency of the low-pass filters is 7.6 kHz and quantisation is 12 bits/sample;
- direct memory access (DMA) is used to transfer data to and from the hard disk without restricting the length of the data stream by the size of memory; and
- all amplifier and filter settings are preset using on-board potentiometers that are not accessible to the user.

THE HARDWARE

Overview

The logical functions of the audio interface board are shown in figure 1. For the input direction, the speech signal is conditioned by the input buffer and is then low-pass filtered. Analog-to-digital conversion takes place and the signal samples are transferred to memory using one of the DMA channels.

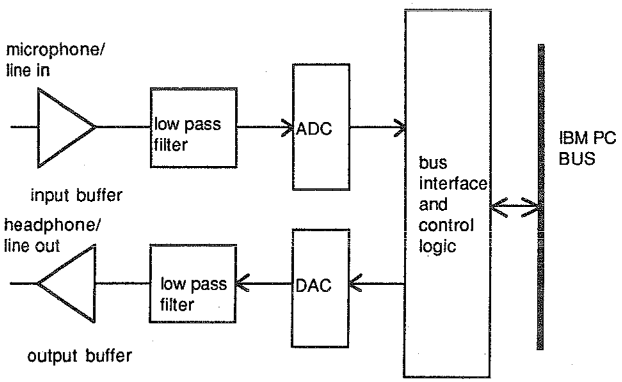

Figure 1. Block Diagram of audio interface board.

In the output direction, samples are transferred from memory to the digital-to-analog converter using DMA. The resulting analog signal is low-pass filtered and amplified to the levels required at the headphones and line jacks by the output buffer.

Input Output Buffers

The input buffer consists of a microphone amplifier with variable gain and a summing amplifier. The summing amplifier mixes the amplified microphone signal with the line-in signal and the resulting signal is fed into the filter stage. The microphone signal gain is set by an on-board potentiometer which is not accessible to the user. Under normal operating conditions, the user will connect either a microphone or a taperecorder for input to the interface. However, the input buffer could also be used in a configuration where microphone and line inputs are mixed.

The output buffer consists of two separate variable-gain amplifiers, one for the headphone signal and one for the line-out signal. The user can therefore listen to the output signal and simultaneously record it on tape. As for the input buffer, the amplifier gains are set by on-board potentiometers inaccessible to the user. A total of four operational amplifiers is used for the input and output buffers.

Low Pass Filters

The anti-aliasing and smoothing filters for the input and output streams are 8th-order Butterworth filters with a cut-off frequency of 7.6 kHz. Each filter is realised in 4 second-order stages where each stage is a biquad with 3 operational amplifiers and an RC network (Wagner & Fulcher, 1986).

Analog-to-Digital and Digital-to-Analog Converters

The converters used in the design were chosen on considerations of performance, cost and availability. For the input stream, the Analog Devices AD574 12-bit analog-to-digital converter is used.