Updating the Reading EPG

Simon Arnfield (S.C.Arnfield@Reading.ac.uk),
Wilf Jones

Department of Linguistic Science, The University of Reading
Whiteknights, Reading, UK. RG6 6AA

ABSTRACT

The Reading EPG[1, 2] (Electropalatograph) was developed in the Speech Research Laboratory at The University of Reading as a device to allow the number, shape and position of contacts between the tongue and the roof of the mouth to be measured and recorded. This device is used both as a research tool[3, 4] and in speech therapy as a bio-feedback device.

The preferred equipment used by major speech research laboratories has shifted in recent years from discrete devices and PCs to more powerful Unix workstations running speech and signal processing packages such as Entropic's ESPS/Waves+.

This paper describes the design and implementation of an update to the Reading EPG 3: hardware (called the EPG 3+) has been developed to acquire data from the EPG 3 direct-to-disk on a Unix workstation and software has been developed to integrate EPG into ESPS/Waves+, to calculate a variety of analyses and to provide real–time display of EPG patterns.

1. INTRODUCTION

Equipment used for speech research has shifted from custom built and specialist devices to the use of computer systems with data capture hardware and signal processing packages. Speech research often pushes the limits of what is possible with personal computers so many institutions now use Unix workstations with powerful speech and signal processing software such as Entropic's ESPS/Waves+ package. The greater power and flexibility of this combination made it desirable to sidestep the existing PC based EPG system and acquire EPG data directly onto the Unix workstation.

1.1. EPG 3 Description

The PC version of the Reading EPG (the *EPG 3*) takes the form of an IBM compatible PC equipped with a specially designed data acquisition card and an external control box connected to which is the user's artificial palate – via a multiplexor unit – and a microphone. Custom built software allows the user either to display palate contacts in real time or to acquire a small amount of speech. Hardware restrictions have meant that this is limited to a duration of 10 seconds. The acoustic speech signal is also band limited to a 10,000Hz sampling rate. Maintaining and updating the custom built software is a time consuming and costly task.

Palates Artificial palates are constructed by an orthodontist from a cast of an impression taken of the subjects hard palate. The artificial palate has a grid of 62 electrodes set into the surface in eight rows of eight (the frontmost row only has 6 electrodes). Sixty-two fine wires connect from the palate to the multiplexor unit and subsequently to the EPG control unit. An additional hand held contact from the control unit completes the circuit. As the multiplexor scans the contacts bits are collected and transmitted to the host computer's data acquisition card to signify which contacts the tongue is touching.

Communication In practice bits are transmitted as a series of nibbles (4 bit words) corresponding to the 4 contacts on the left side of the palate followed by the 4 contacts on the right side of the palate for each row in turn. Handshaking between the computer and the EPG control box thus communicates 16 four bit words for each palate frame. Each four bit word is held steady for $180\mu s$ and as there are 100 frames per second this means a gap of $7.12ms$ between frames.

Analogue Signals Provision is also made for the acquisition of two analogue signals: acoustic signal from the microphone and an auxiliary signal (for example laryngograph). In the new system these signals are handled separately by the data acquisition card.

2. EPG 3+ HARDWARE DESIGN

The update to the above mechanism for Unix workstations is called the EPG 3+. To minimise development time the update was designed to work with the existing EPG 3 external control box and multiplexor and any computer with a data acquisition card capable of 2 channel, 16,000Hz, unfiltered, DC-coupled data capture. In brief the interface works by encoding the digital data from the EPG 3 control box as an analogue signal which is then sampled and decoded by software on the workstation.

Each of the 16 four bit words from the EPG control unit is encoded by a digital to analogue device (the EPG3+ unit) which is then sampled by the computer's data acquisition card. A sample rate of 16,000Hz was chosen as the minimum commonly used sample rate that could capture the values for each nibble. Each nibble is held for $180\mu s$ equivalent to a frequency of 5555.5Hz. So the absolute minimum sample rate would be 11,111.1Hz in order to get at least 2 samples per nibble. In practice 16,000Hz is a widely used and convenient rate.

2.1. Digital to Analogue Encoding

Analogue encoding of the 16 four bit words is accomplished using the following scheme. This scheme, though apparently complex, is relatively easily achieved by a simple electronic circuit and has advantages that allow for automatic synchronisation and calibration.

Nibble values are encoded on a voltage range of -5v to +5v. This is an arbitrary range picked to suit the data acquisition card used. The voltage range is divided into two halves with the lower range -5v to 0v being used for nibbles associated with the left half of the palate and the upper range 0v to +5v for the right half of the palate.

Within each half the four bits of each nibble are treated as an integer (note bits are numbered from left to right) in the range 0 to 15 which is scaled by $\frac{5}{16}$volts and added to the base value – either -5v or 0v. These values can vary from the ideal as a calibration phase in the software can estimate voltage range and DC offset.

0111	1110
1100	0111
1000	0011
1000	0001
1000	0001
1000	0001
1000	0001
1000	0001

Table 1: Binary representation of EPG frame.

To encode the first row of the frame shown in table 1 the following calculations are made. The left nibble has a value of 0111 = 14 and the right a value of 1110 = 7. Voltage levels for these two values would be $14 \times \frac{5}{16} - 5 = -0.62$volts and $7 \times \frac{5}{16} = 2.19$volts. The thick line in figure 1 represents an idealisation of the analogue signal – the boxes show the value encoded at each point (the figure only shows the signal for the first 6 rows of the frame). Following the last nibble the voltage level remains unchanged until the first nibble of the next frame.

The resulting signal is a varying amplitude square wave. This is the source of the requirement for DC-coupled, unfiltered sampling. If a filter were used the effect would be to round-off the corners of the square wave. If AC-coupling were used DC components of the square wave would be affected. In addition the sharp edges in the signal can cause a filter to ring.

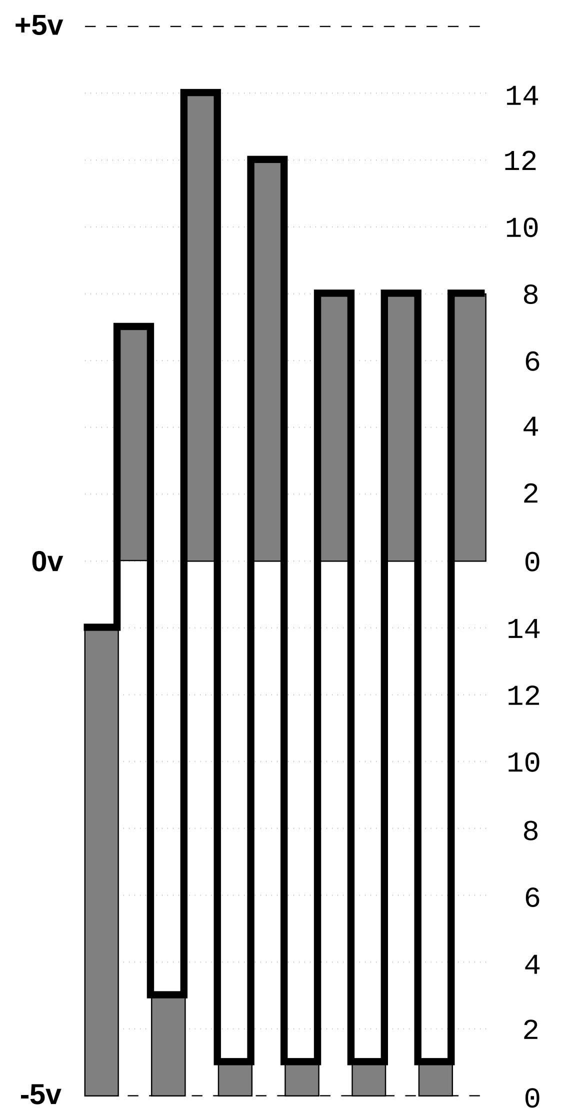

Figure 1: Analogue Encoding of part of an EPG frame.

3. SOFTWARE

There are several programs that support this system which fall into three categories:

- Acquisition and Decoding Data
- Data Analysis
- ESPS/Waves+ Support

Acquisition software is hardware dependent and interfaces with system software provided with the data acquisition card used, the Analyx AD1416–166; and therefore forms the most machine dependent part of the software. Other A/D cards could be used but no others are currently supported.

Data analysis software is designed to be system independent and will process EPG data files from either EPG system.

ESPS support software converts signals acquired from the acquisition software into ESPS format files and allows for display and analysis of EPG data.

3.1. Acquisition and Decoding

Analyx A/D Card Data Capture The Analyx AD1416–166 multichannel data acquisition SBUS card is capable of sampling from 1 to 16 channels with an aggregate sample rate of up to 160,000Hz. Its DC–coupling and unfiltered input make it suitable to sample many types of signal; this is in contrast to similar devices that are primarily designed to capture acoustic or acoustic–like signals.

To sample acoustic signals it is important to use a filter where aliasing might occur. We use a Kemo 0.01Hz – 100kHz dual variable rate filter (type VBF/8) and a Wavetek Rockland dual hi/lo filter (model 852) although the majority of our input sources (such as the Laryngograph and Rothenburg mask) have built–in filters.

For capture of EPG and acoustic data a microphone signal is low pass filtered down to around 8000Hz and connected to channel 1 of the data acquisition card's plug box and the analogue signal from the EPG 3+ is connected to channel 2. The EPG 3+ is also plugged into the EPG 3 control box with a multiplexor unit and palate connected. The microphone signal from the EPG 3 control box is not used in this instance since it is low pass filtered to less than a 5000Hz and a better recording can be made with the 16,000Hz sample rate used to capture EPG data.

A modified version of software provided with the Analyx card is used to sample data from these two channels at 16,000Hz per channel. The input channels are multiplexed and sampled into two buffers on the card (8192 samples per buffer). When one buffer is full data is written directly to disk whilst the other buffer is being filled. As the second buffer is written to disk the first buffer is overwritten with new samples. This continues until the require data capture duration is reached. The file stored on disk contains the multiplexed channel samples; that is the first value is a sample from channel 1, the second value is from channel 2, the third value channel 1 again etc. Multiplexed files of this kind are not very useful so a post–sampling program demultiplexes each channel into a separate file.

EPG Analogue Decoder After sampling a program is used to decode the digitised EPG signal back into binary frame patterns (8 bytes per frame).

A first step is to scale down the quantisation resolution by a factor derived from the signal voltage level and the quantisation resolution of the A/D converter card and taking into account any DC offset. These values are estimated from the default null input signal (ie when no palate is connected which is equivalent to no contacts) in a calibration phase. In the case of the Reading setup, which uses the Analyx AD1416–166 multichannel analogue to digital card, the scaling factor is approximately $\frac{1}{2000}$. This brings the -32768 to 32767 (16 bit) quantisation range down to -16 to 16. These values directly correspond to the 4 bits of each encoded nibble (negative values have 16 added to them first).

However, it must be realised that several samples are taken per nibble and it is possible for a sample to be taken on an edge of the square wave which could result in any value. To overcome this the number of repetitions of each scaled down value is counted. Values that occur once only are discarded (these correspond to edges). The result is 16 values per frame, the first 15 of which last for around $180\mu s$ and the 16^{th} of which lasts for around $7.12ms$. By locating this long lasting value it is possible to synchronise decoding to the start of frames. In addition checks are made for "corrupt frames" by checking the number of values and that the sign of the values alternate between negative and positive. Pairs of nibble values are combined into bytes and written to a file in the appropriate order. The start time of the first frame relative to the start time of the acoustic signal is calculated for complete accuracy as the EPG unit works in free–run mode and is not started and stopped in synchrony with data capture. The first frame may well be chopped in half, rendering it corrupt. The start time offset is usually less than $10ms$.

Real Time Frame Display By combining data capture with the analogue decoder into one program with X-windows display code for frames the real time display program can give a bio–feedback display of palate contacts. It is not possible due to speed and hardware constraints to achieve a real time display of 100 frames per second. This is above human perception levels anyway but it is quite possible to display 10 frames per second (ie every 10^{th} frame) on a Sparc–Station 5. An update rate of 20 frames per second is possible on this system but requires higher process priorities or else the X-windows code does not have enough processor time to complete before the next frame is due to be displayed.

The A/D card is instructed to acquire $50ms$ (or $100ms$) of data from the EPG input channel before signaling the program it is ready. It then continues to aquire the next $50ms$ (or $100ms$) worth of data. Meanwhile the first sample of data is scanned looking for the long stretch of values that corresponds to the last nibble value of a frame. The next 16 decoded values are taken as an EPG frame. The rest of the sampled data is discarded. The frame is displayed on an 8×8 grid by some simple X-window code before the cycle repeats using the data captured while decoding and display of the previous frame was taking place; and so on.

Since Unix is a time–sharing operating system there are no guarantees that other system or user processes will not hog processor time. In addition X-windowing systems do not necessarily perform the tasks requested of them immediately. If the decoding and X-window code does not complete in time to signal the A/D card that it has finished with the first batch of data this system breaks down. In practice this can be avoided by either running the software in single user mode or by increasing the process priority.

3.2. Data Analysis

Data analysis software currently consists of fairly trivial programs to calculate contact totals and frequencies.

Contact Totals The first analysis function sums the total number of contacts over a given time range for each contact on the palate.

Area Contact Rates The second analysis function produces contact frequency graphs over a given time range for three zones on the palate: velar (front 3 rows), palatal (middle 2 rows), and alveolar (rear 3 rows).

Frame Display To support further analysis software frames (or ranges of frames) can be output in ASCII format as a sequence of 8 bit binary numbers.

3.3. ESPS/Waves+ Support

Support software is in ongoing development. At present software exists to import and view data in Waves+, to label data and to produce postscript printouts.

Data Conversion In order to use signal files acquired with this system with Waves+ they must first be converted to ESPS format. The data capture software has a graphical user interface that allows the user to specify which data sources to use by choosing any combination from (in the current Reading setup) acoustic signal, laryngograph, oral and nasal airflow and EPG. The user interface provides information to the user on how to connect the various sources to the computer and initiates data capture. As noted earlier, multiple data sources are multiplexed into one file — the user interface software automatically demultiplexes the data and converts it to the appropriate ESPS file formats. This is largely achieved using ESPS support programs.

Frame Display Support of displays of EPG frames linked to Waves+ signal displays is provided by a Tcl/Tk script (see figure 2). Communication between the script and Waves+ allows the EPG frame to update automatically in response to the position of the cursor in the Waves+ signal display or alternatively to playback frames from a selected range of speech in the acoustic signal.

Labeling and Printout An extension to the menu of the Waves+ display allows a labelling program to run which inserts data from an EPG frame into the label file at a specified point. This is coupled with another program which produces postscript output of the label file including EPG frame images.

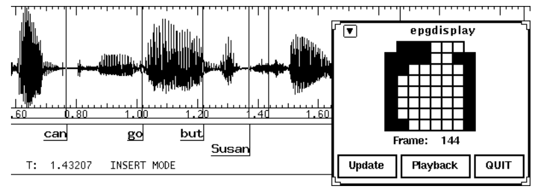

Figure 2: EPG display linked with Waves+

4. FUTURE DEVELOPMENTS

Support and analysis software is in ongoing development as time and needs demand. Negotiations are currently underway to market EPG3+.

For further details see our world wide web site. The URL is: `http://midwich.rdg.ac.uk/research/speechlab/epg/`.

5. REFERENCES

1. P.J. Roach, and W.J. Hardcastle. A computer system for the processing of electropalatographic and other articulatory data. *Proceedings of Phonetics Symposium*, Univ. of Essex Language Centre Occasional Papers, 17:127–142, 1976.
2. W.R. Jones. Electropalatograph hardware. *Phonetics Laboratory of University of Reading, Work in Progress*, 1:7–13, 1977.
3. W. Hardcastle, W. Jones, C. Knight, A. Trudgeon, and G. Calder. New developments in electropalatography: A state–of–the–art report. *Clinical Linguistics and Phonetics*, 3:1–38, 1989.
4. P.J. Roach. Notes on the processing of Electropalatographic data with a small computer. *Phonetics Laboratory of University of Reading, Work in Progress*, 1:14–20, 1977.