SoundID

A Revolutionary Sound Recognition System

www.soundid.net

Table of Contents

Introduction

This manual is organised in the order of the modules that the reader should be familiarised with first. The module "LPC and GD" is the one that should be mastered before proceeding too far, because it embodies in the most straightforward manner the principles that the other modules are based on. This module compares individual cut WAV sounds and introduces the Geometric Distance concept. This module is not only the best one to clarify the concepts, but it is the one that is important for studying and verifying the reference library.

We have been running training courses for some years now and recently the course has been extended to one week. It is probable that the manual is sufficient for users to grasp the essential concepts without that training, but a fuller understanding of the concepts, many of which are entirely new, cannot be conveyed in any manual. Once grasped, the concepts are found to be intuitive, but it may require some effort initially.

The software mostly follows standard Windows formats, but there are a few custom controls that need to be mastered.

The target is to get better than 95% recognition accuracy. This is comparable to the accuracy of a human expert. If the user cannot achieve that then most likely the software is not being applied properly. At the end of this manual there is a checklist of things that can cause problems. Also, make sure you understand Chapters 6, 7 and 9 which describe how to get the most out of your software.

There was a problem with signals that had gliding frequencies. This problem has been addressed in the new 64 bit release which will be a Beta version initially. This type of signal is known as a Bart's Head as seen below with the Eastern Whip-bird call. This problem is due to the compression of the call into a single time-slot which tends to turn any wide-band signal into a Bart's Head. Thus it is difficult, in this way, to distinguish one wide-band signal from another. This is not an inherent problem with the technique as long as the signal is processed differently. The multiple peaks are the result of a rising or falling pitch. By choosing smaller frame lengths the Bart's Head becomes a Matterhorn.
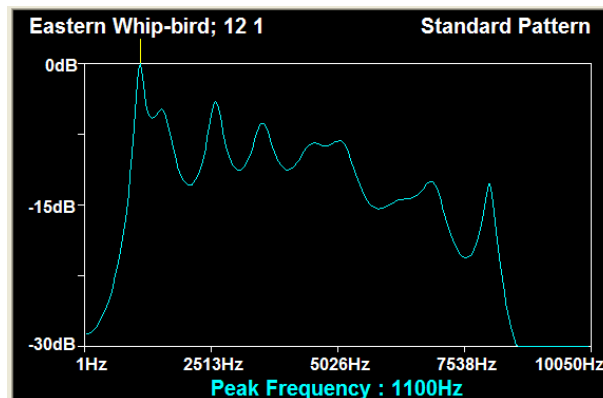
Bart's Head   Bart's Head Signal

The new 64 bit software processes the signal differently as seen in the 64 bit signal image. This overcomes the Bart's Head limitations.
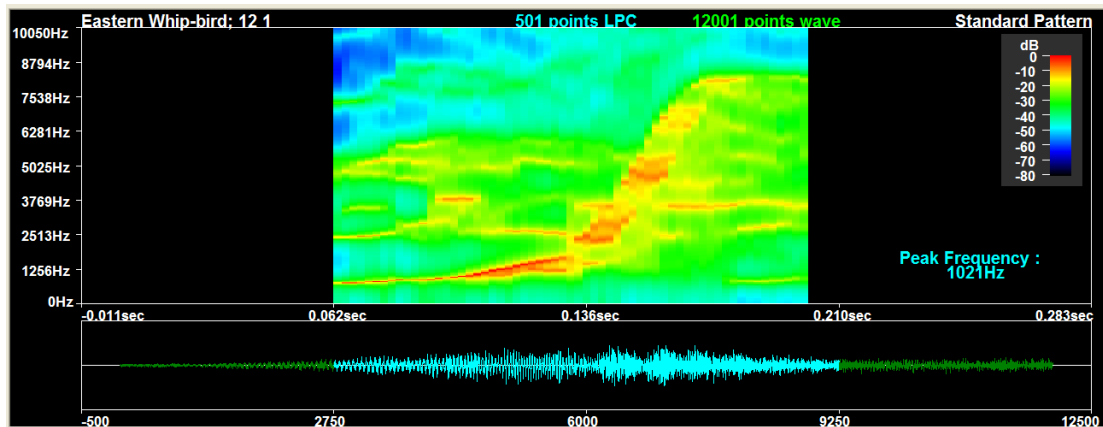
The same signal processed in the time domain as well as the frequency domain is far more distinctive.

Installing the SoundID Software
The software should auto-install once you place the CD in the CD drive. If it does not, click on the "Setup" icon on the CD.

If downloaded from the internet, then click on the setup icon.

Your Product

The trial version of the software has full functionality for 30 days. Licencing is done on the internet (www.soundid.net) and involves you providing us with your registration key. This process will ordinarily happen within 12 hours but we do recommend you get you licence key a few days before the expiry to ensure uninterrupted use.

Note carefully that the SoundID Professional is a different product to SoundID but the licencing key will work for the respective products. So the download is the same installation file for both products and the product can be upgraded with a new licence key.
It is important to note that the licence only covers one PC (although it may be transferred between PCs). If you need multiple PC licencing contact us first.

Sound Samples
Some sound samples of the Double Eyed Fig-parrot are included to make it easier to get started on this program. They have been provided by David Stewart. (http://www.naturesound.com.au/).

Parrot Icon

The lovely red and green Australian King Parrot that we use as an icon (and as seen at the top of this document) is a regular visitor, and has given permission to SoundID for the use of his images.

Author
Neil J Boucher: SoundID, Maleny, Queensland, Australia

Inventor and Patent Holder
Michihiro Jinnai: Nagoya Women's University, Japan

December 2013

Glossary

Geometric Distance

The concept of Geometric Distance (GD), a mathematical method of finding the similarity between two patterns, is at the heart of SoundID. The calculation of the Geometric Distance is computationally expensive, but once you get used to interpreting it, you will agree that it is worthwhile. The geometric distance, measured in degrees, varies between 0 (for a perfect match) and 90 (if there are no discernible similarities; and here, 90 degrees means a right angle).

The value of the GD for matching is relative. So if the signal is viewed at the default settings a good match is typically 3-6 degrees. If you zoom in on the signal by narrowing the bandwidth and/or the LPC dB (noise floor) much higher GDs may be needed to indicate a match (sometimes 20 degrees or more). We have introduced the GD Normalisation concept to allow for this relativity.

Notice that GD and Weighting Vector are interactive. The GD measures the difference between two sounds, but the weighting vector changes what is measured (and hence the same two sounds will have a different GD when measured with different weighting vectors).

GD Trim

If a signal found at any point in time matches to more than one reference, then a decision must be made as to which one matches best. If the GD trim is set to zero them the best match only (lowest GD) is reported. If you need to see all reasonable matches then set the trim according. For example a GD Trim of 2 will leave all time co-incident matches that are within degrees of the best match.

Frame Width

The Frame width is the number of points that are used for the calculation of the LPC. It is equivalent to the "bin Size" in the FFT.

Linear Predictive Coding (LPC)

LPC is the name of a mathematical transform which, in our implementation, for practical purposes, is much like the FFT (Fast Fourier Transform) in that it splits a given signal into its frequency components. The LPC, however, is more robust than the FFT and also more accurate for vocalisation. The disadvantage is that the LPC is computationally more demanding than the FFT, and so runs slower. There may also be some non-vocal sounds for which the FFT might be more appropriate.

Normalised Geometric Distance

The Geometric Distance is a relative and not an absolute measure. Thus, for one group of signals that are alike the GD between members of the group may be, for example, 3 degrees whereas for another group it may be 7 degrees. So if a sound matches as 2.5 degrees to the first group, but at 5 degrees to the second group, it is not immediately obvious which is the closer match. The normalised GD of the first match is $2.5/3 = 0.83$ and to the second group is $5/7 = 0.71$. So it is a closer match to the second group. When running mixed templates that have

different reference GDs the normalised mode should be used in Recognition and the Batch Mode.

Template

A template is a group of reference calls that are similar enough to be processed with a single set of parameters. For example these might be advertising calls of a particular parrot.

Wave Normalisation

This sets the wave file maximum value to 0dBm0. It is equivalent to setting the maximum value recorded to the maximum value that could be recorded. In this way it is possible to compare signals that are recorded at different levels.

Weighting Vector

The weighting vector is a signal distortion vector that adds pre-emphasis to the signal in a way that adds weight to the stronger parts of the signal and de-emphasis to the low energy parts.

Weighting Vector Spread

The weighting vector factor spreads the weighting vector (which is a Gaussian distribution) in a way that diminishes the effect of the vector as the spread is increased.

Chapter 1

Comparing WAV Files

Introduction

SoundID is a powerful sound recognition system which uses entirely new concepts. Since the best way to get an understanding of it is to use it, this paper will describe each of the programs in the SoundID suite, showing examples for each one.

The method is essentially image match of either the full high resolution LPC spectrum (2-d) or of the frequency components only of that spectrum (1-d)

SoundID cannot only detect differences in sounds, but using the Geometric Distance it can *measure* the difference. This is an important advance in understanding the differences in sounds. In fact SoundID is better described as a *similarity* recogniser rather than as a difference recogniser. SoundID in fact looks for similarity, not difference!

SoundID can process continuous files, but it is easier to understand how it works if you start with single calls or sounds. For tutorial purposes, we have included a few files that can be used to get started. These files are in the same directory tree where SoundID was installed (typically `C:\Program Files\SoundID`).

Be aware that the concepts used by SoundID are different from those normally encountered in acoustics, and the reader may need to spend some time to become familiar with them. But a little effort will pay big dividends, and the techniques are powerful and useful. More details can be found on the SoundID website in the "research papers" section www.soundid.net .

Getting Started

Having installed the software, you will find that it placed a King Parrot icon on the desk-top:

Double-clicking the icon will bring up the SoundID *Main Menu* showing the individual programs that make up the SoundID suite:
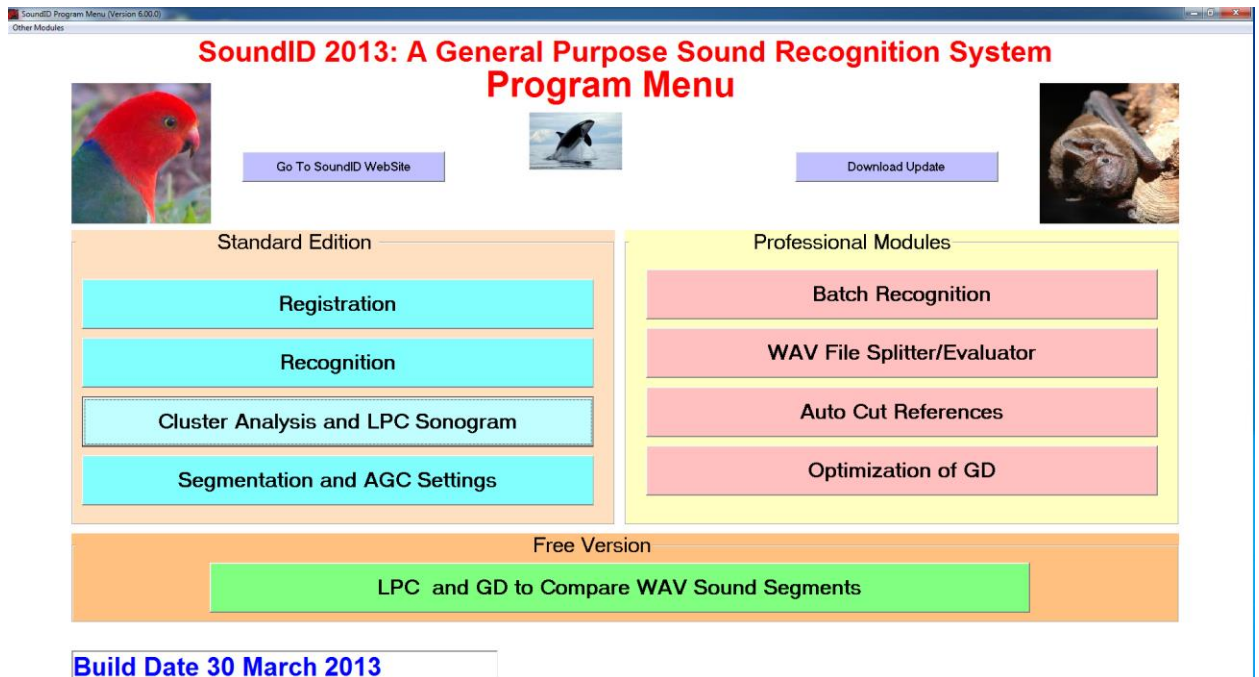
Pressing a button on the main menu causes the associated SoundID program to be launched. These programs are described in detail in the following sections of the document.

You will notice that the menu is divided into three sections. The first section is the free version that will run indefinitely. The standard and professional modules will run during the trial period, but after the trial period they will need to be licenced to continue to run.

LPC and GD with Noise

The program that handles individual calls is called LPC and GD (Linear Predictive Coding and Geometric Distance). Selecting this program on the Main Menu will cause the following screen to be displayed. It can also be used as an editor for the reference files.
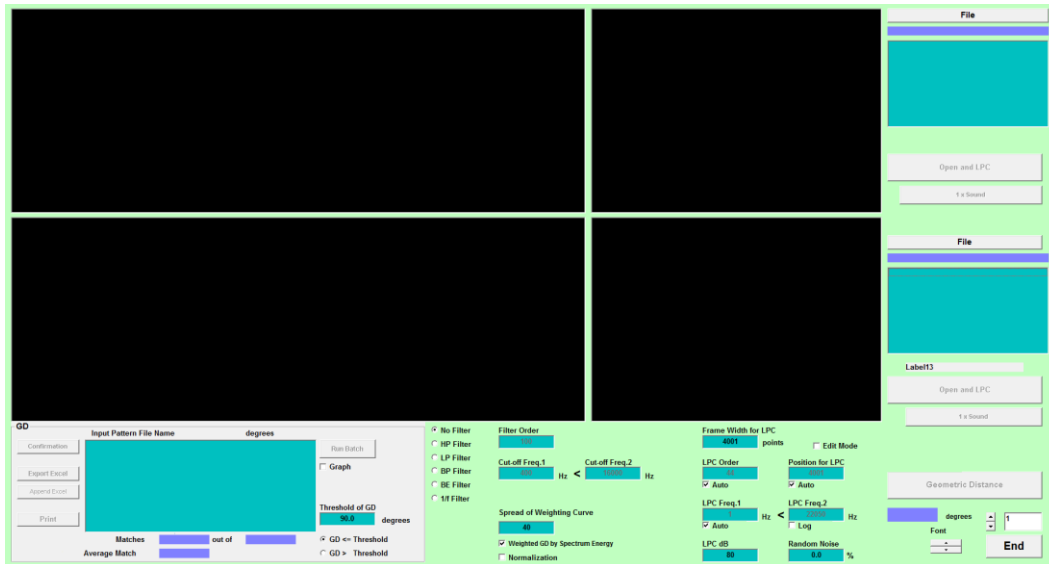
This screen contains four distinct areas:

- Upper left: Waveform display windows
- Upper right: File selection and processing
- Lower left: GD controls and results list
- Lower right: Filter and algorithm settings

Defaults

The program comes with a setup page that presets the default values. You should not change these until you have an understanding of how they work. Be aware also that these values change and this can cause the results to differ from those in the examples.

Waveform display windows

These are the windows in which the 'raw' or WAV sounds and the derived LPC power spectrums are both displayed (the 'raw' sounds appear in the wider windows).

The upper windows show what is referred to as the 'Standard Pattern', while the lower windows show the 'Input Pattern'. This naming reflects the primary purpose of SoundID: to search a set of recorded calls (the 'Input Patterns') for any that resemble the reference call (the 'Standard Pattern').

File selection and processing

To the right of each of the waveform display windows is an area containing some buttons and file list boxes. These buttons are used to load the sound files you wish to work with, and then to calculate the LPC power spectrum for those sounds.

File

Clicking this button causes a file selection dialogue box to appear. Navigate to the folder that contains the sound files you wish to work with, and select any file in that folder. The program will search the folder for all *wav* files, and will display a

sorted list of their names in the box below the button. Although this is not a standard way for Windows directories to behave it is most convenient.

Open and LPC

Clicking this button tells the program to process the file that is selected in the list box, causing it to populate the adjacent display windows with both the raw input waveform and the derived LPC power spectrum.

Double-clicking a file name in the list box will give the same result.

Play Sound

Clicking this button will cause the program to play the sound associated with the currently-highlighted item in the list box. Note that this is *not necessarily the same sound whose waveforms are displayed, as it will usually be somewhat longer, and including parts that were not processed in the image.* This button may even be used before any sounds have been processed.
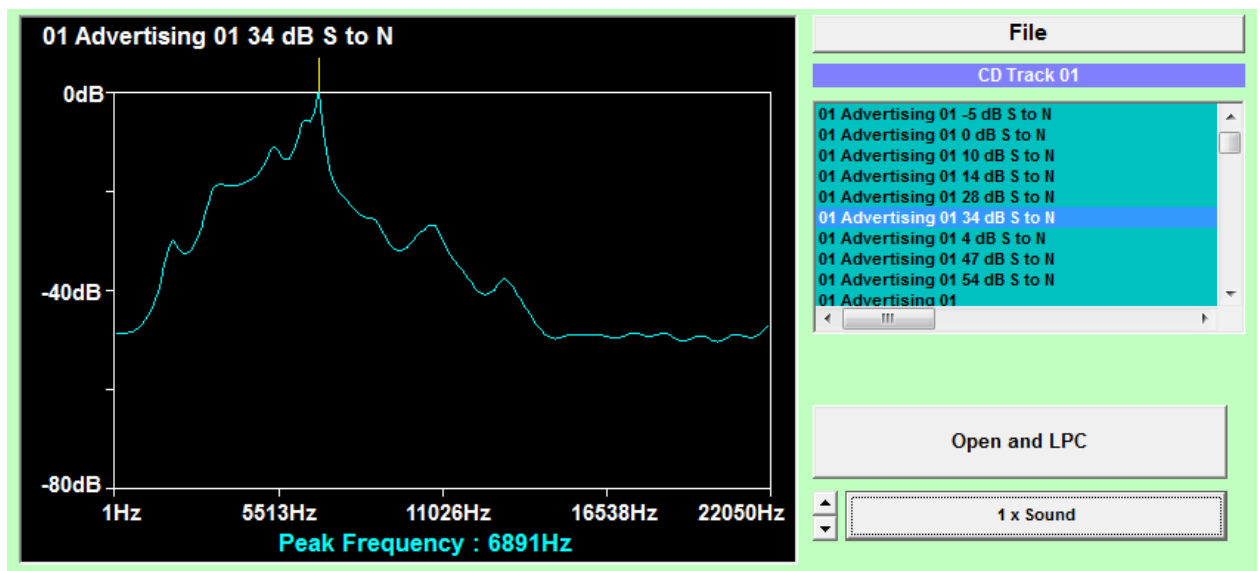
The button named 1 x Sound enables the sound to be played. The scroll bars are used to scale down or scale up the frequency by a given factor. This is useful for very low or very high frequencies as it makes them more audible.

No Graph

Just below the second list box is a selector that allows you to suppress the real-time display in the lower waveform windows. This is simply an aid to boost performance when large numbers of files are being processed. See the *GD controls* section below for more information about processing multiple files.

Geometric Distance

When the program has processed a pair of sound files and calculated their LPC power spectrums, this button causes the spectrums to be compared. The result, the *Geometric Distance*, is a measure of how similar the two sounds are considered to be, and has a value between 0 (a perfect match) and 90 (no discernible similarities).

GD controls and Results

The panel in the lower left-hand part of the screen contains a list box, a series of buttons, and an input field. Between them, these controls allow you to compare multiple files, vary the result threshold, and export the results. The panel is reproduced here for ease of reference.

Confirmation

This button plays the sound segment that was analysed to allow the user to confirm correct recognition.

Export Excel

This button causes a dialogue box to open up, allowing you to specify a file to which the current batch results will be exported. The results are in the form of simple text, with each line containing a set of comma-separated values suitable for input to the spreadsheet program of your choice. The first line of the output contains column headers.

Append Excel

The function of this button is similar to that of the *Export Excel* button, except that it will append the current results to an existing output file.

List box

This box contains a list of all the input files that match the standard pattern according to the specified threshold value.

Run Batch

This button causes the program to compare the 'Standard Pattern' of interest (the one currently selected in the list box in the upper right-hand part of the screen) with all of the 'Input Patterns' (those listed in the lower list box).

Threshold of GD
As defined elsewhere, the GD (*Geometric Distance*) is a measure of how closely one sound resembles another – the lower the number, the more similar are the two sounds. This input area allows you to specify a number, between 0.00 and 90.00, which is the value used by the program to determine whether or not to display a result.

If you enter a number less than 0, it will be interpreted (without warning) as 0. If you enter a number greater than 90, it will be interpreted (without warning) as 90. If you enter something other than a number, an error message will be displayed.

GD Threshold radio buttons
Two radio buttons are provided, allowing you to choose between selecting results that are less than or equal to the threshold value, or results that are greater than the threshold. By default, the program displays results that are less than or equal to the threshold.

Matches
A pair of output fields, *Matches* and *out of,* shows the number of matches determined by the program, taking into account the threshold settings.

Average Match
For the sounds that resulted in a match, this field displays the arithmetic average of their GDs.

Filter and algorithm settings

This panel, occupying the lower right-hand part of the screen, contains controls that allow you to affect how the waveforms are processed. The panel is reproduced here for ease of reference.

Filter selection radio buttons
The left-hand side of the panel contains a series of controls for various filters that can be applied to the waveforms.

You may select:
• No Filter – the waveforms are processed unchanged

- HP Filter – a high-pass filter is applied to the waveforms
- LP Filter – a low-pass filter is applied to the waveforms
- BP Filter – a bandpass filter is applied to the waveforms
- BE Filter – a band elimination filter is applied to the waveform

Filter Order

This is a measure of how 'steep' the filter cut-off is. Lower numbers mean the filter takes effect more slowly as the frequency increases. In general an order of 20 is a good compromise.

Cut-off Freq 1 and 2

For high-pass and low-pass filters, *Cut-off Freq 1* is the frequency above or below which the desired filtering is to take effect.

For bandpass and band elimination filters, both frequencies are used to define the filter limits.

Using these filters is easier if you understand that the algorithm is "sort of" comparing the same images that you see on the screen as the LPC. If the image is nicely centred then all is probably OK. However if the image is crunched into one end as seen in the frog call below then the matching is all about matching the long noise tails.
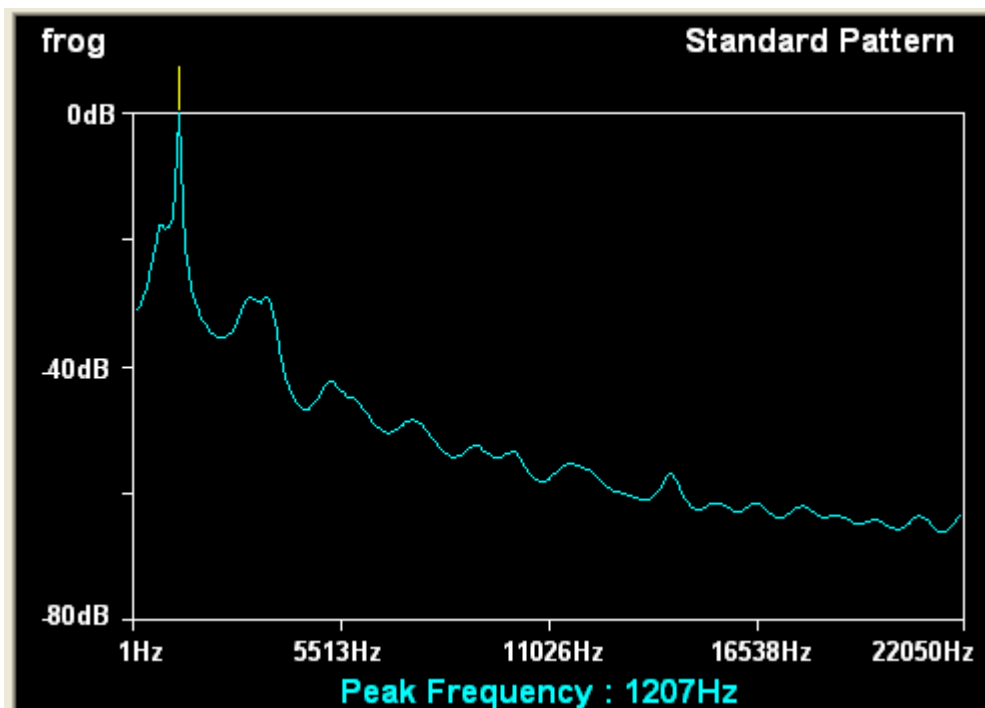
Frog recorded at 44.1 kHz
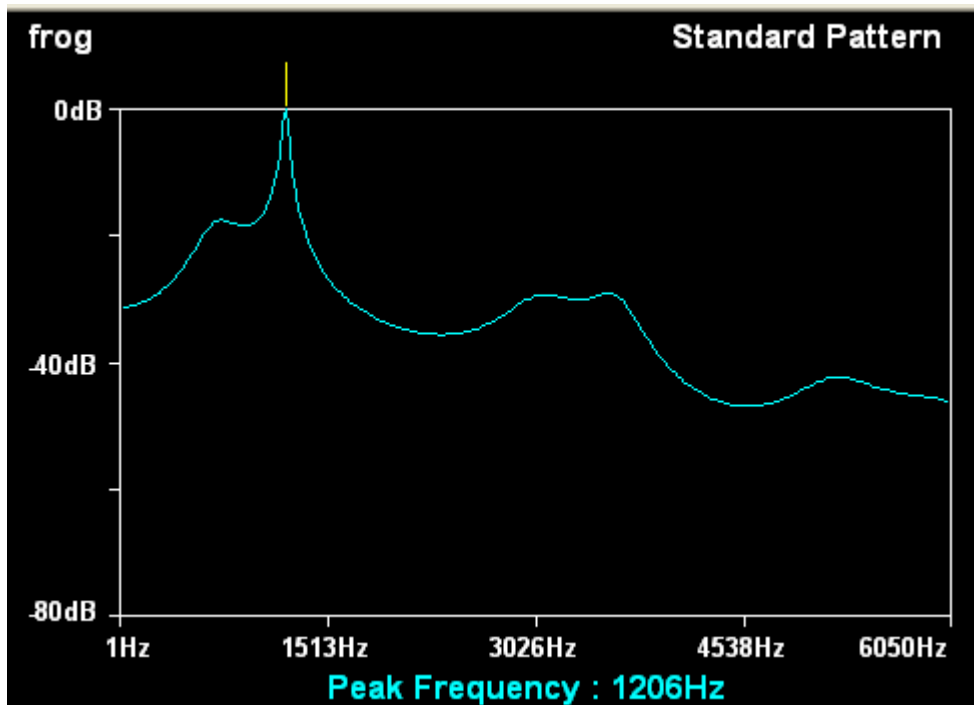
Frog call set to a more realistic frequency range.

Notice that if this is done it will change the GD and improve the recognition, but it is most important that this is done identically to the reference call and the recordings being examined (there is no automatic check on this).

Other Controls

Frame Width

The software will look for the highest energy in the cut call, which is one frame width long. The frame width is measured in samples. This means that if the sample rate is 44,100 Hz, then a sample width of 4410 is equal to one tenth of a second. It is most important that you set this value so that it will be equal to the regular length of the call. The effect of varying this value can be seen by looking at the waveform and noting the area of that waveform that is in blue. The blue is the part of the call that the software is looking at. The green is totally ignored

LPC Order

You will notice that the LPC order and the next four controls are greyed out. This is done because it will be rarely necessary to change these. However should you need to change them you first need to un-tick the auto box above. If you increase the order the LPC will be done in finer detail and this will generally be seen as more detail showing up in the LPC.

Position for LPC

It is unlikely that you will need to change this. However if you do, you must first un-tick the auto box and then change the LPC position. The LPC position is the location of the middle of the frame that the software is looking at (the blue portion of the signal). The location of the signal as a function of the position is displayed at the bottom of the WAV graph.

LPC Freq1 and Freq2

These frequencies are the frequencies displayed in the graphics. For some signals it is better to restrict the waveform frequency range when the target signal is narrow band. An example would be bats, which have high frequency calls but output nothing in the low end of the spectrum. So for the case of a bat call at 16 kHz you may set Freq.1=10000 and Freq.2=22000. To change the frequency you must first un-tick the auto checkbox below LPC Freq.1.

LPC dB

The LPC dB is the depth to which the signal will be examined. For example if your recording is near a busy road and the background noise is 60 dB below the 0 dB level, then you might set this to 60 dB as anything below that value will be predominately noise and will only obscure the recognition.

Random Noise

This adds random noise to your signals so that you can see the effect of noise on the recognition capabilities. This should always be set at zero when you are processing calls.

<u>**Examples**</u>

Having described the *LPC and GD* program in some detail, now we can try it out using the supplied sample sound files. These files may be found in the `Recognition` subfolder where the SoundID program was installed.

1. From the Main Menu, start the *LPC and GD* program.
2. Click the upper *File* button and navigate to the `Recognition` folder.
3. Navigate to the `Double Eyed Fig-parrot Advertising` folder.
4. Select any file in that directory and click *Open*.
5. Repeat steps 2, 3, and 4 for the lower *File* button.

6. Click the buttons labelled *LPC and GD* below each of the list boxes, and the screen will display as:
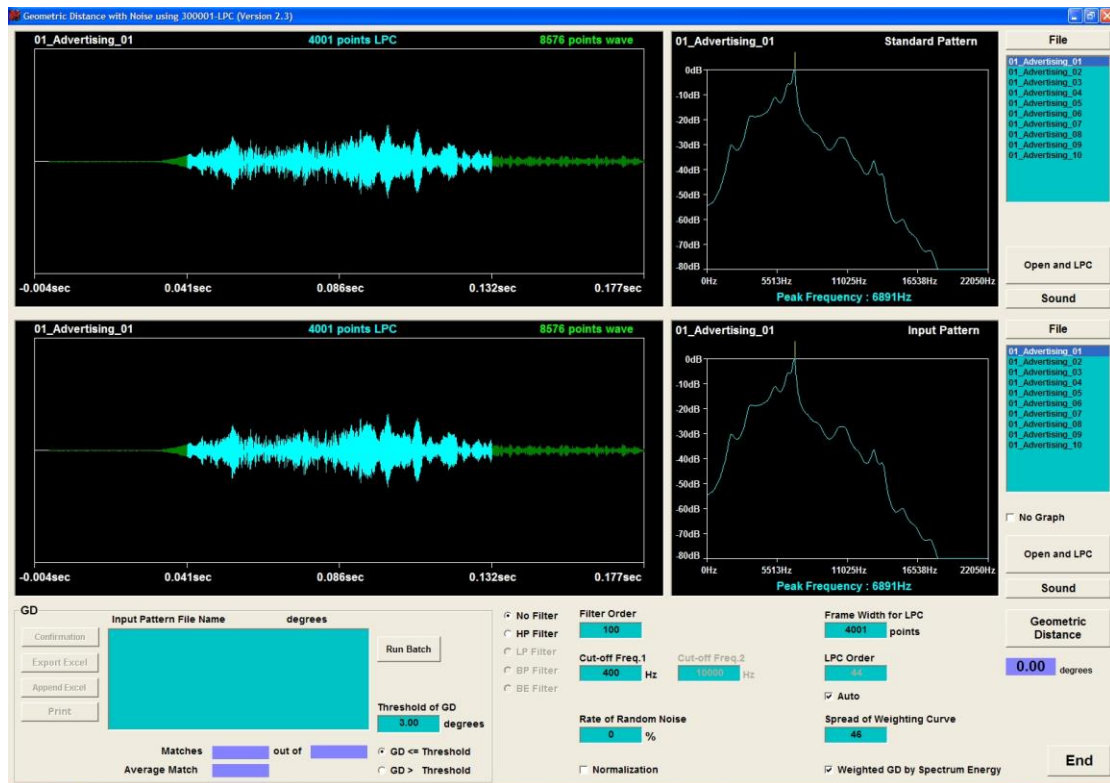
The waveform of call `01_Advertising_01` is displayed on the left, and its LPC appears to the right of it. At this stage we are comparing `01_Advertising_01` with itself.

Click the button labelled *Geometric Distance* (lower right), and you will see its value displayed as 0.00 degrees, since, by definition, two identical patterns have a geometric distance of 0.

In the lower file list box, double-click on `01_Advertising_10`. This will cause the program to load that file, calculate its LPC, and display the new geometric distance:
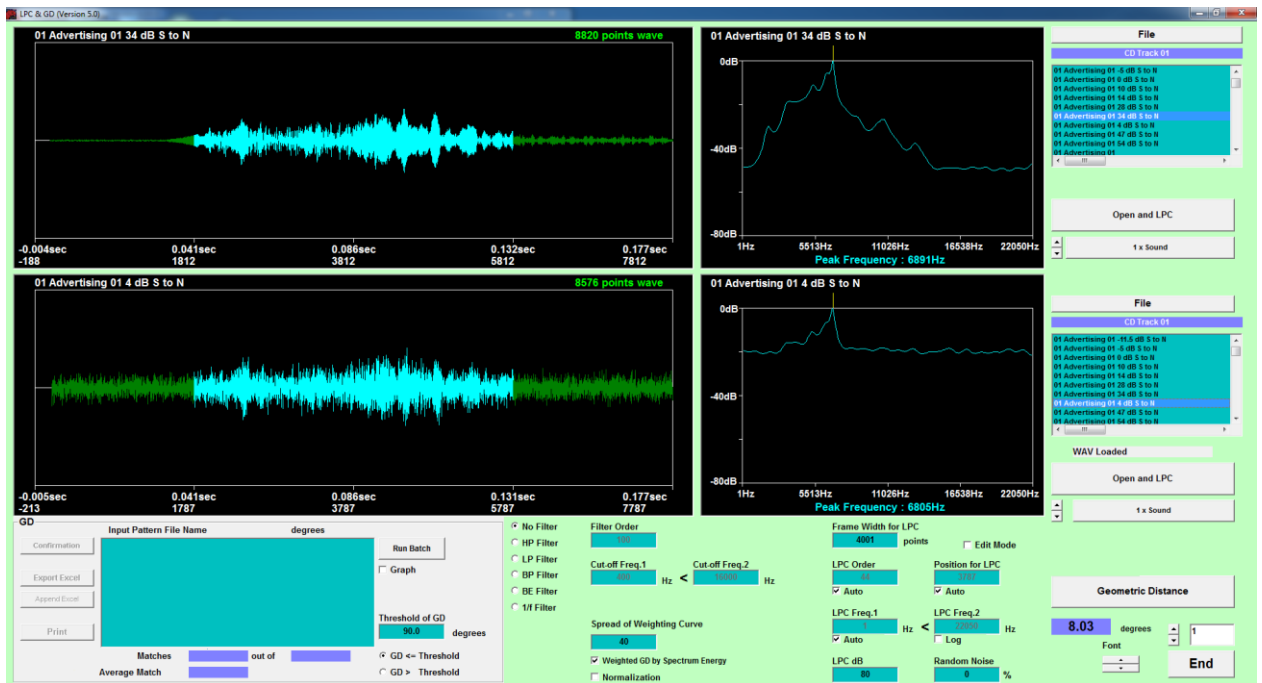
You can also press either of the *Sound* buttons to hear the sound that is currently displayed. Notice that the part of the call highlighted in blue is the part that has been compared. However the sound button plays the *whole* of the cut segment of the WAV file.

Next, click on the *Run Batch* button in the *GD* panel:
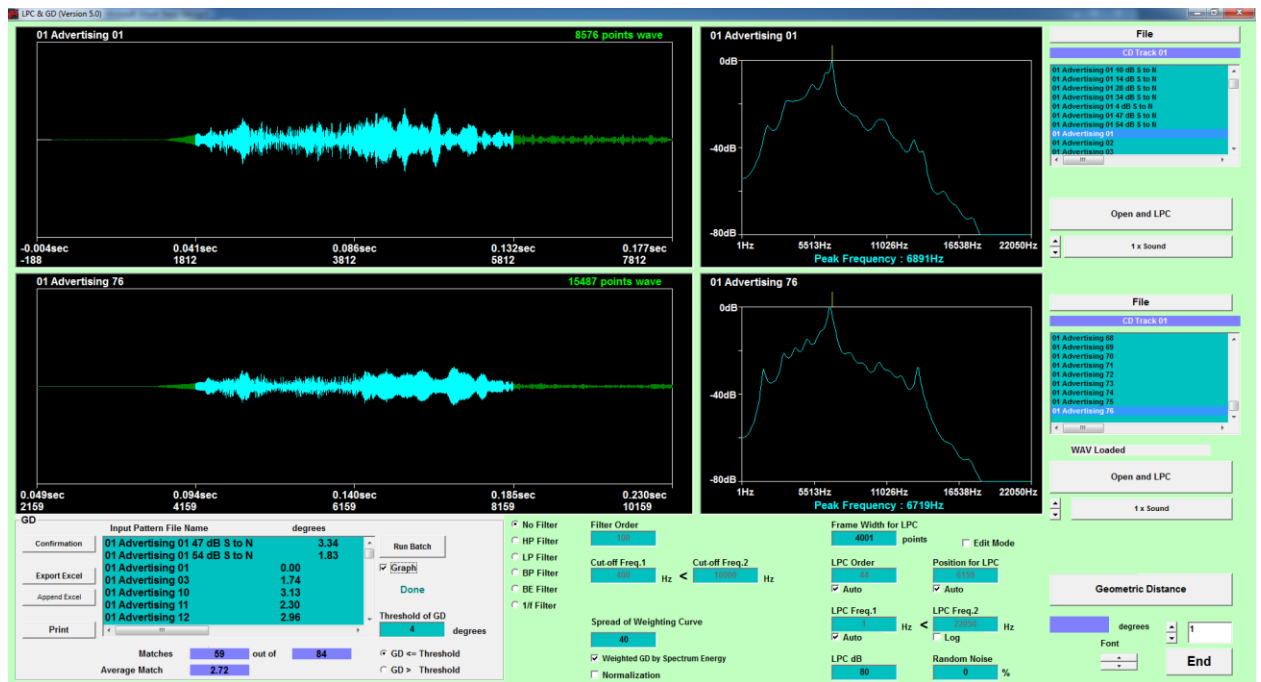
The program will compare the selected call in the top box, `01_Advertising_01`, with each of the calls in the bottom box, totalling 59 calls out of 84 compared. We find also that it matches only with the least noisy versions of itself.

Increase the matching level (threshold) that we will accept, by changing its value from 3 to 90, and press the *Run Batch* button again:
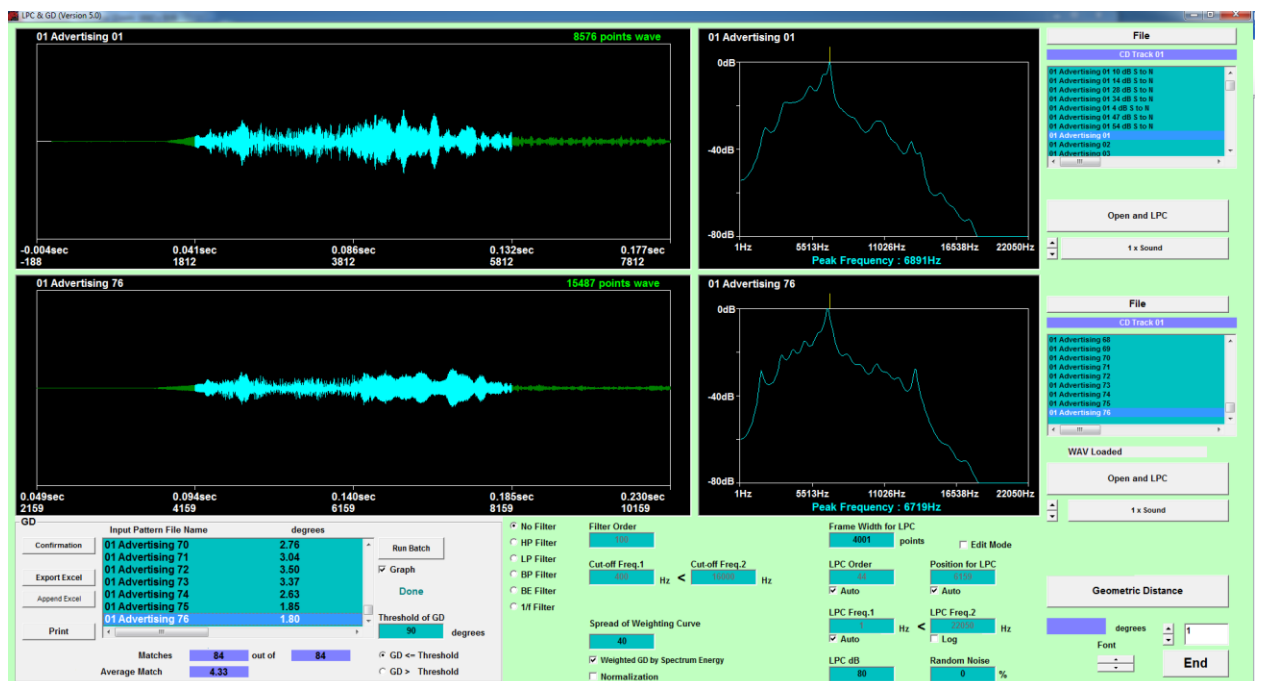
Because we set the highest threshold, all calls now match to `01_Advertising_01`, with an average match of GD = 4.33 degrees.

If you click on the respective *Sound* buttons, you will hear that the calls are indeed similar. In general, for parrots, we have found that a GD of 5 or less signifies a good match, but matching can occur up to a GD of 10.

You should note that the GD for matching depends on the sound you are examining, and on how close, for your purposes, you need the calls to be to accept them as a match.

Speeding up the Process

When you press the *Run Batch* button you will only see the graphs being drawn in real time if you tick the *No Graph* tick box (lower right). The default mode will switch off the graphic display and considerably decrease the processing time.

Digital Filter

If you tick the digital filter check box, a filter will be applied to the sounds. This will remove most wind noise and a lot of ambient noise. It will also increase the amount of processing that needs to be done, and will slow down the recognition process. Take care that the cut-off frequency is not set so low as to cause the loss of any part of the target signal.

Running a non-matching file

Before we leave this section, try running a file that does not match.

1. Click the lower *File* button and navigate to the `Recognition` folder. From there, navigate to the `Double Eyed Fig-parrot Contact` folder.
2. Select any file in that directory and click *Open*.
3. Click the buttons labelled *LPC and GD* below each of the list boxes, and then select *Run Batch*:
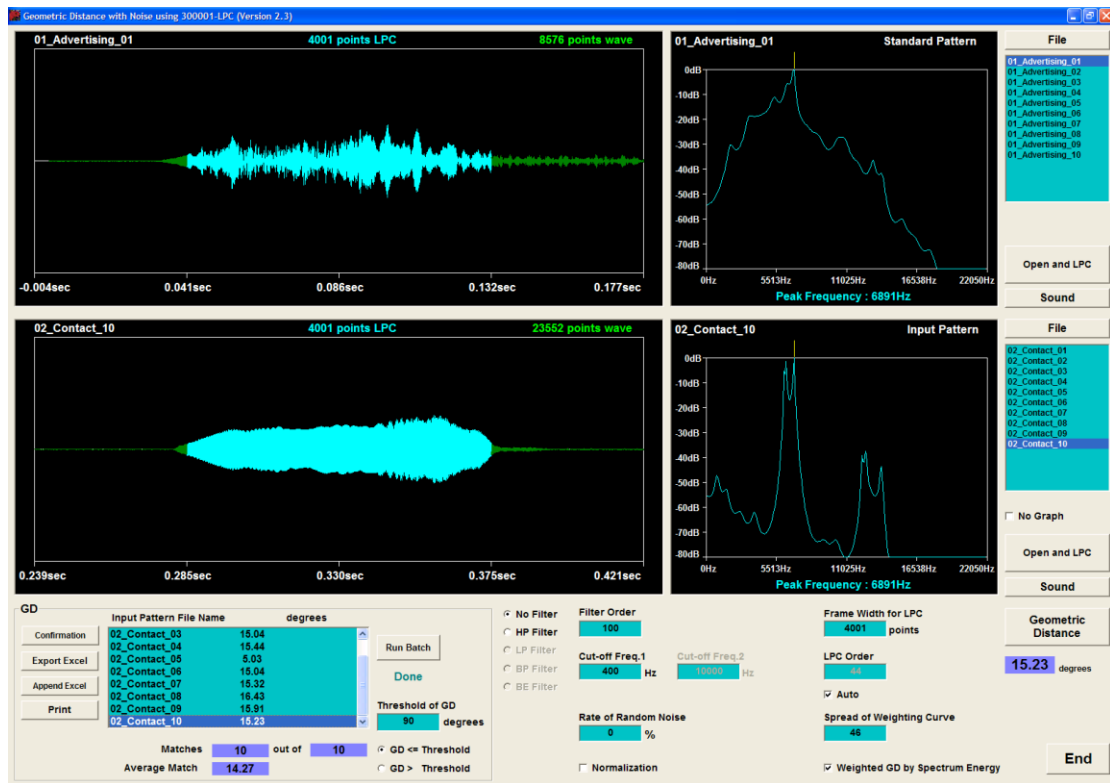
Notice that the average match is now 14.27, indicating that the calls are very different. If you press the respective *Sound* buttons, or look at the LPC waveforms, you will agree that they are indeed different. However, there is a suspicious outlier at `02_Contact_05`, which has a GD of only 5.03.

Double click on `02_Contact_05` in any of the list boxes where it appears, to get this screen display:
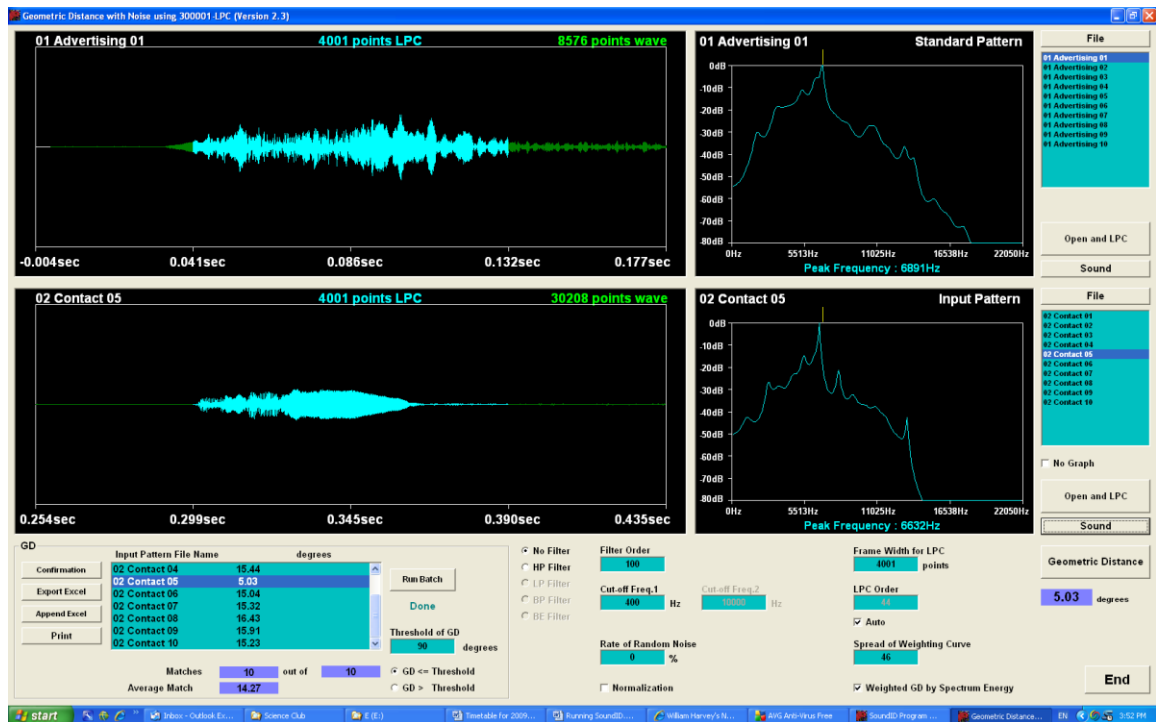
Notice that the LPCs look similar, and playing the sounds will confirm this. In fact, **02_Contact_05** turns out to be an advertising call that was wrongly categorised (the mismatch was discovered at the time of writing this paper, but was left in place to illustrate how the software works).

Running Your Own Files

There is a formal and precise way to set up the matching criteria, which we will cover later but for those who are getting anxious, here is a rough and ready way to get started quickly with your own sound catalog. There are two ways to do this.

Manually
 To run you own reference files you first need to cut them (use Audition or similar) into sound samples. Do not cut them precisely, but leave about 30% either side of the portion of interest and let the software find the highest energy portion.

Auto Cut

The auto-cut module is part of the professional version of SoundID (although you can still run it in the trial version). This allows you to cut a continuous file into chunks suitable for use as references with a few key strokes (see chapter 10).

Running the References
Start with a batch of sounds which are similar (the same call or sound type) but are variations of a theme. Add a few calls that should not be classified as matches. Now run the LPC and GD module again in batch mode and vary the GD until you find the

range of values of GD that cause all the calls that should match to do so, while leaving all that should not match higher than the GD threshold and hence classified as non-matches. The GD that you end up with should be in the range of about 2.0 (sounds that are highly uniform like machines) and 4.5-5.0 sounds that have a lot of inherent variability.

Using LPC & GD as an Editor

If the Edit Mode check box in the lower right hand corner is ticked a "Delete Sound" button become visible. With this ticked you can load a group of reference calls into LPC and GD and see visually how it will be as a reference. If you want to delete it from the reference calls just click the delete button. Be aware that this will permanently delete that file from your references.
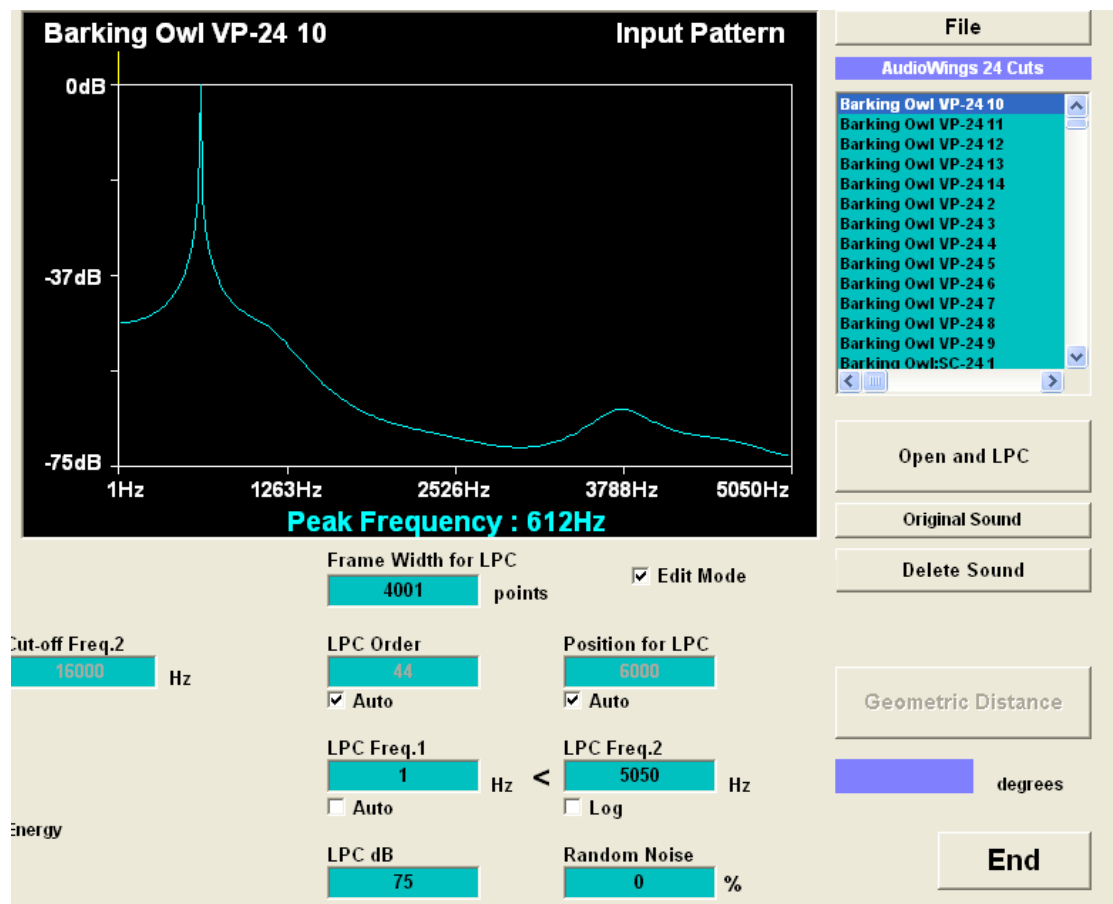
Below is a wider screen shot of the same process.
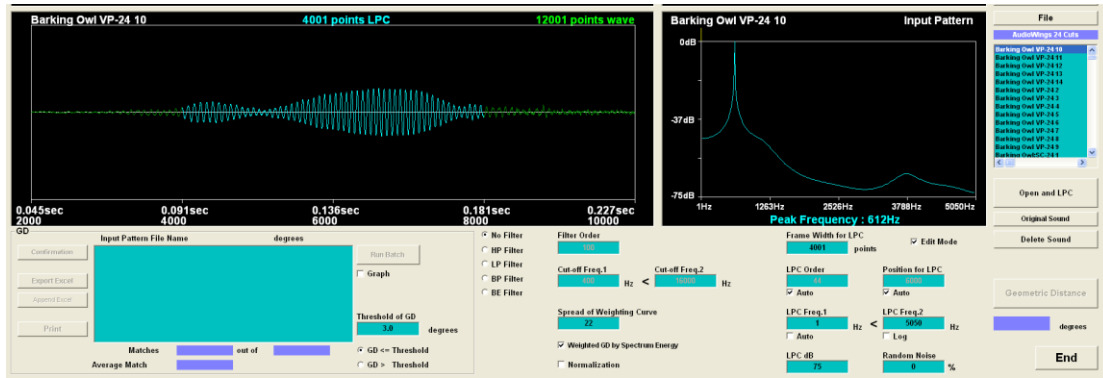
In order to properly compare sounds it may be necessary to stretch high frequency calls so that the peak frequency is closer to the optimal human hearing range or to scale up the peak for very low frequency sounds. Either way this is easy to do. In the screen-shot below you can see the up/down buttons. These can scale up the frequency from 1x to 32 or from 1 to 1/32 (stretch). The screen shot shows a sound being stretched by a factor of 6x (you hear it at 1/6 of the true frequency).

Chapter 2
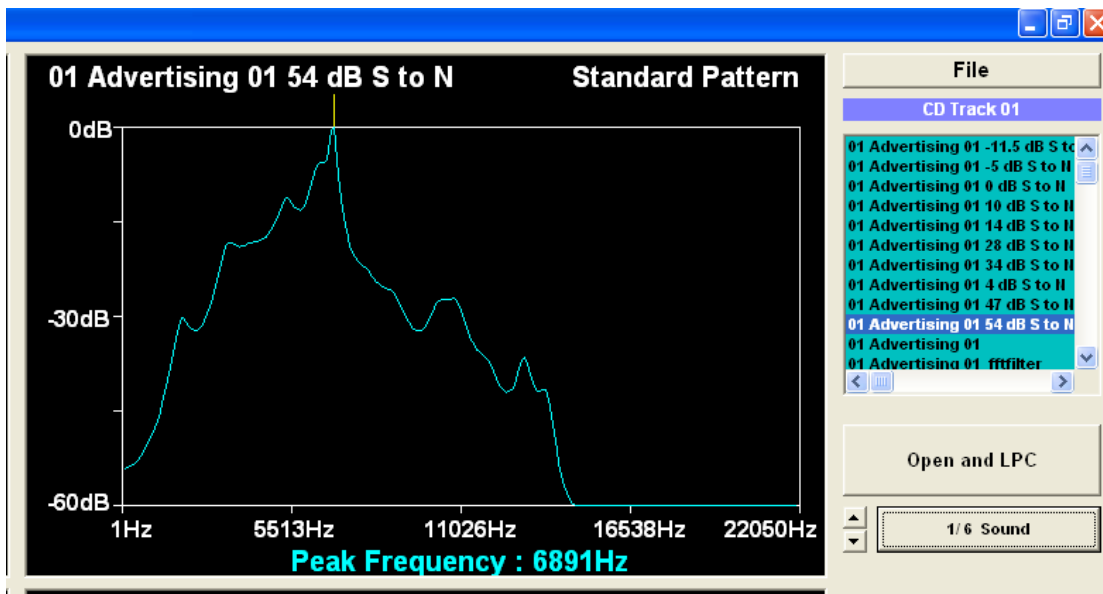
Registering Your References

Registration

The reference file is a library of calls that can be searched to look for matches in a long WAV file recording (s).

SoundID uses your cut WAV file samples to form its library. However to expedite processing the references are not stored as the original WAVs, but rather as pre-processed mathematical images of those files. Having cut a set of reference files it is good practice to run them through the LPC and GD module to confirm that there are sufficiently good examples and that they represent the target well.

Care here! If you have only a few examples of a call from a species that makes dozens or even hundreds of different calls you will have little reason to believe that the software will likely find your target. Also keep in mind that commercial recordings are often made in one location and at one time. There may be dozens of calls but they may be representative of only a few distinct call types.

There is no real limit to the number of calls that can be added to the references and in general the more examples, the better the prospects of finding the target. However noisy references and ones corrupted with other sounds are likely to lead to false positives and should be weeded out. In fact once you have a reference file it is a good idea to run it over a known recording and look for false positives. Invariably, references that consistently produce false positives will be found to have been corrupted.
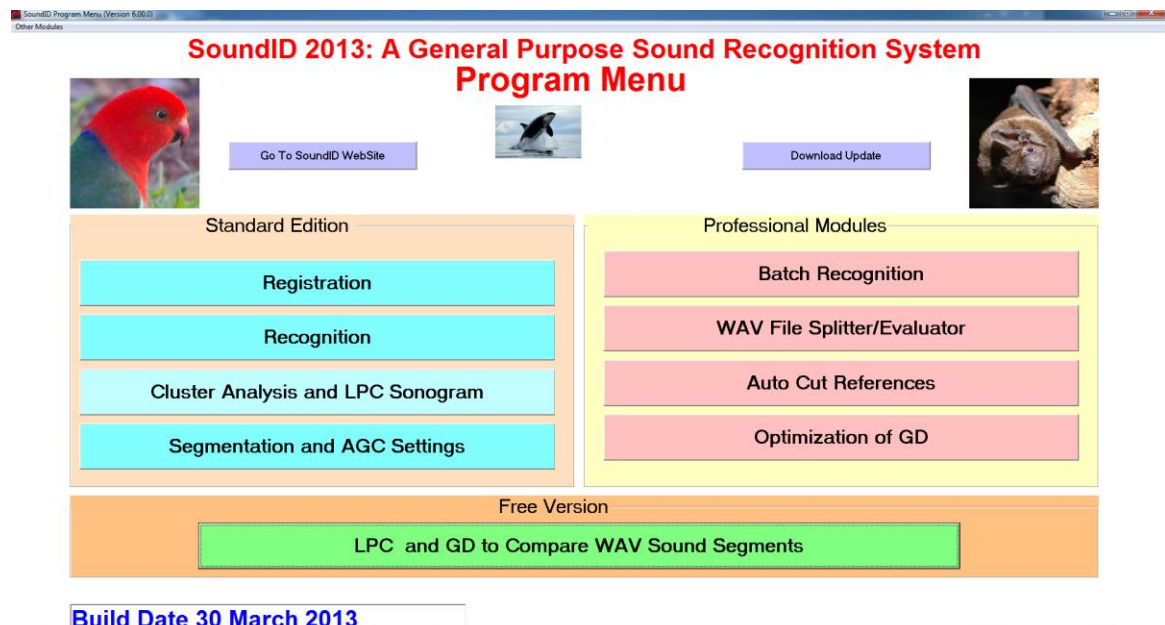
It is in the Registration window that you determine the settings that will be used in subsequent processing. The settings in the Recognition windows should not be different from the original Registration settings (and in most cases cannot be different). If you want to try different settings for different outcomes, do it in this window.

To begin click on registration and the following screen will appear.

Registration Screen
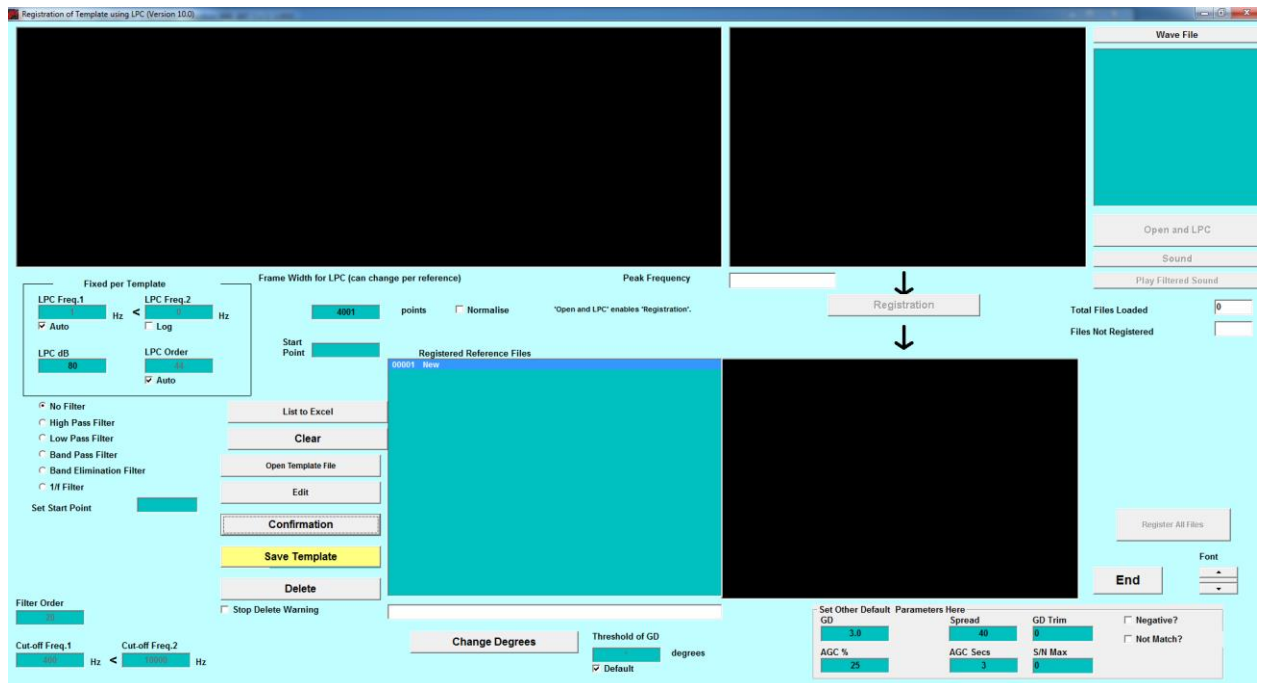
At the top right there is a button called Wave file. Click on this and follow the Windows dialog box to locate your cut reference files. Click on any one of them and then click OPEN in the dialog box.
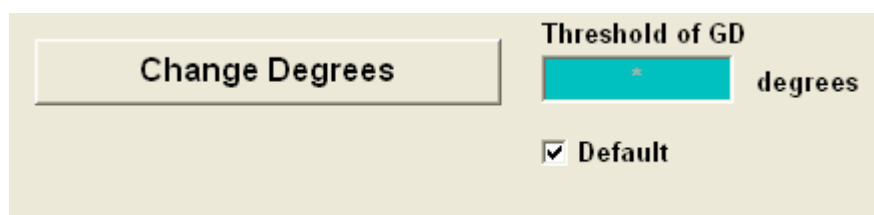
Click "Open and LPC". Files can be registered individually by clicking the "registration" box, or collectively by clicking "Register All Files".

Threshold of GD

The threshold of GD is the value of Geometric Distance in degrees that is considered to be small enough so that calls are considered the same. The threshold of GD will ordinarily be set to the default value which will be set in later programs, but there are times when it needs to be altered at registration. This is particularly true for calls that have been set to narrow bandwidths and so have large GD values for matching. The correct value for matching can be determined by using the Cluster Analysis module. In other cases there may be a need to set the GD to rather extreme values. Examples are calls or references that are rare and of great interest. In these cases it may be advisable to accept a larger (say 4 degrees) than normal GD, so that any matches no matter how tentative will be brought to attention. Notice of course that this invites false positives. Alternatively, there may be a recording of a reference that is of poor quality, but nevertheless it is desirable to include it. In this case it may be a good idea to set the GD smaller (say to 2 degrees), so that only very good matches to it are noted.

Change Degrees

There are two ways to change the degrees of matching of a particular call. The first is to change the value in the box below "Threshold of GD". This will cause any new file to be registered at the set number of degrees.

Alternatively, if the file has already been registered and it is desired to change its degree value, highlight it in the list box, change the number of degrees and then click "Change Degrees".

File Manipulation
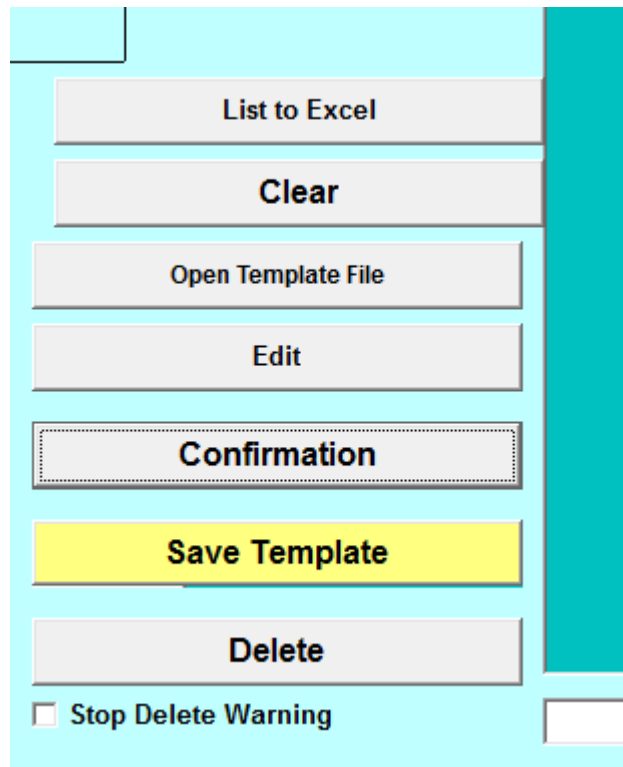
The reference file can be manipulated with the buttons shown

Clear

This will clear everything in the list box.

Open Template

If it is required to append something extra to an existing reference file then you can click on the "Open Template" and use the Windows dialog box to navigate to the file. Click open and it will load. Then whatever WAV files are processed will be appended to that file.
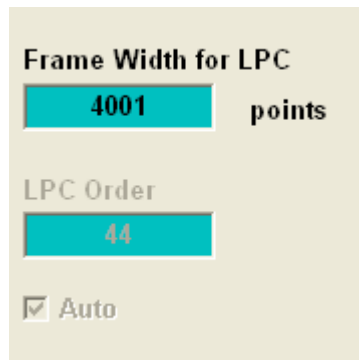
Confirmation

Plays the WAV.

Delete

Deletes the highlighted referenced file from the list box.

Save Template

This will save the whole of the reference template.

Frame Width

Frame Width for LPC

4001 points
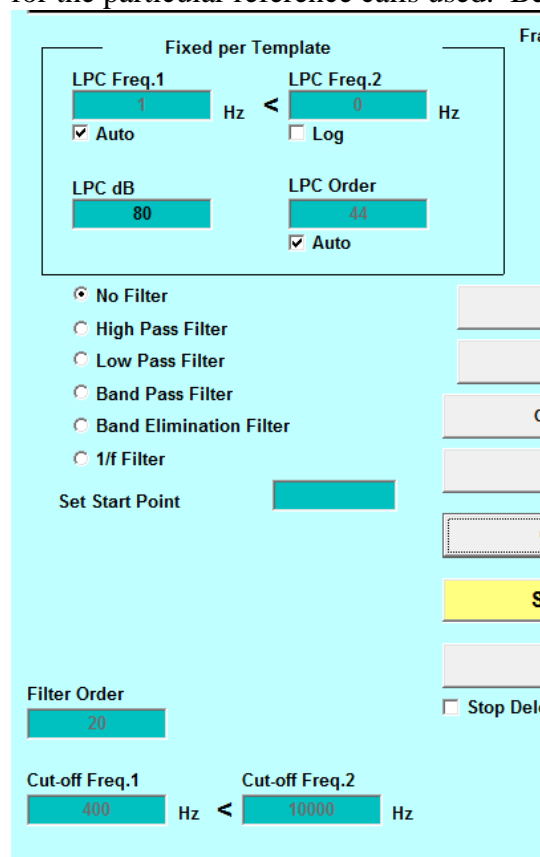
LPC Order

44

☑ Auto

As before the frame width sets the size of the reference selection. So, 4410 points on a 44,100 kHz recording represents 1/10 second. It is most important to get this right as its value determines the proportion of the call that is processed.

The frame width may be optimally different for different calls. Different lengths can be registered on a per call basis, but two things need to be kept in mind. Firstly longer calls take longer to process and secondly if there is more than one call length then the runs are repeated for each call length listed, so keep the number of different lengths to a minimum.

For most survey work 1001 to 2001 points would be ideal, but do try variations outside this range for special sounds.

Other Settings.

The Registration window has a number of settings that can be stored and optimised for the particular reference calls used. Below you will see two of the most important ones. The **LPC window** sets the frequency range. For a parrot with a peak frequency of 6 kHz this Window might be 1000 Hz to 11,000; the selection being based on including most of the signal and excluding most of the noise.

Fixed per Template

LPC Freq.1   LPC Freq.2
1   Hz   **<**   0   Hz
☑ Auto   ☐ Log

LPC dB   LPC Order
80   44
    ☑ Auto

⦿ No Filter
○ High Pass Filter
○ Low Pass Filter
○ Band Pass Filter
○ Band Elimination Filter
○ 1/f Filter

Set Start Point

Filter Order
20

Cut-off Freq.1   Cut-off Freq.2
400   Hz   **<**   10000   Hz

These frequencies are the frequencies displayed in the graphics. For some signals it is better to restrict the waveform frequency range when the target signal is narrow band. An example would be bats, which have high frequency calls but output nothing in the low spectrum. So for the case of a bat call at 16 kHz you may set Freq.1=10000 and Freq.2=22000.

☐ Stop Del

31

The **LPC dB** sets the noise floor that you want to work to. For most real world recordings this should be from 15-30 dB. Setting it higher will force the software to only match with low noise examples of the references (and you may want to do this deliberately). However for survey work you want to match even very noisy calls so set it low.

These two settings are the most important ones to get right if you are processing noisy sounds.

Filters

The filters should be set as required. It needs to be noted that filtering a signal will distort it and will also significantly increase the processing time. It can increase the processing time by double or even more. The filter should only be used if necessary. Indicators for its use are wind noise, extraneous noise outside the target's frequency range and noise enhancement for narrow band targets.

The filter order can be increased to sharpen its effect, but this will significantly increase processing time, in general for very small gains. The filter cut-off points should be set as desired.

The filter characteristics used in the reference files should match the settings in the Recognition window or there will be no perfect matches with the originals (a filtered signal is different to the unfiltered one).

The filter and the LPC frequency windows do much the same thing. However for really noisy signals the filter will be sharper and so filter out more noise than using the LPC settings. The filter however consumes more processor time than using the LPC settings so it's a bit of a trade-off.

Three more settings need attention as shown below.

Set Other Default Parameters Here			
GD	Spread	GD Trim	☐ Negative?
3.0	40	0	☐ Not Match?
AGC %	AGC Secs	S/N Max	
25	3	0	

The default should ordinarily be set high as this will determine the widest matching that you will be able to study. A good default is 10. The spread set to about 40 will suit most signals.

If the negative box is ticked, then the software will look for the reference, identify it, but never acknowledge it.

The default of 4001 points for the frame width shows the part of the signal that is processed highlighted in blue in the top left hand screen.

Example of Making a Reference File from Cut Calls

We can make a reference file from call snippets cut using a sound editing program (Cool Edit or Audition etc). Note that when you use an editing program to cut reference calls cut them at least 60% longer than the actual call itself. The software will find the most energetic part of the call to use for a reference automatically. If you cut the calls too short compared to the frame-width they may not register properly.

We already have some cut calls, that come installed with this software and we can use these to make the reference file.

Double click on the Desk Top King Parrot icon and then click on the "Registration" button. You should have the screen below.
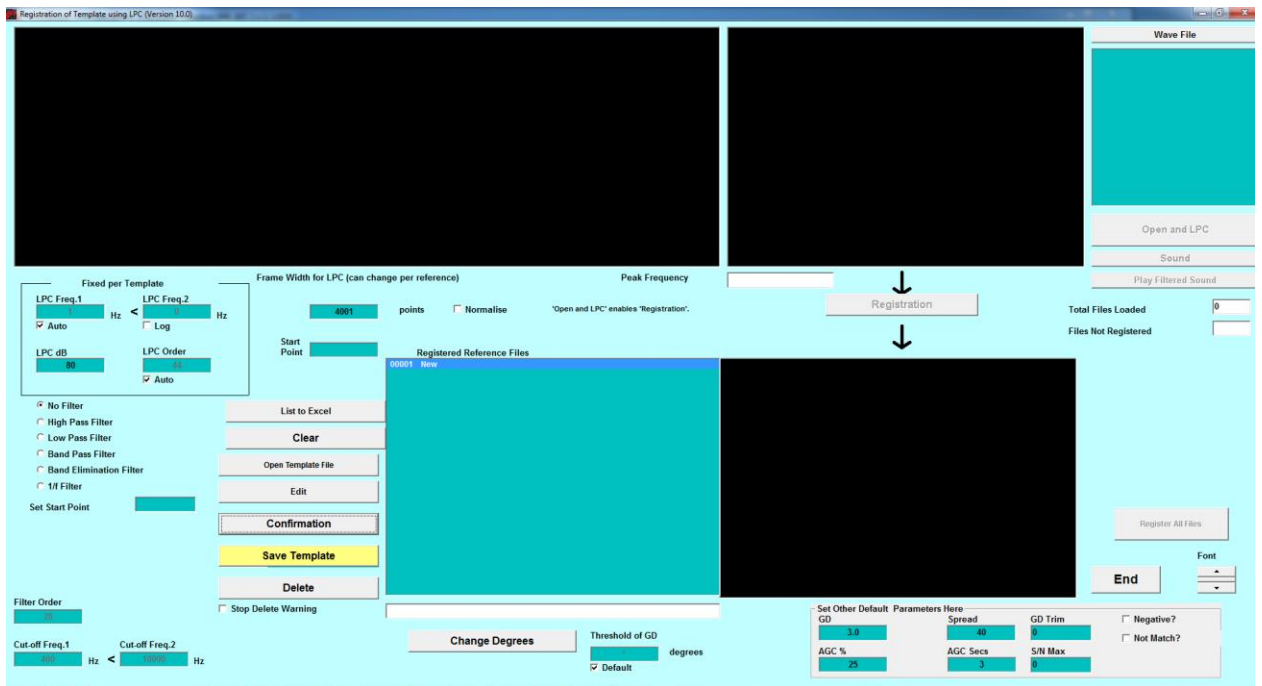
Now click on the "File" button on the top right and open C:\Program Files\SoundID\Recognition\Double Eyed Fig-parrot\01 Advertising 01 and when you have done so, click on the "Open and LPC" button just below. You should now see this screen.
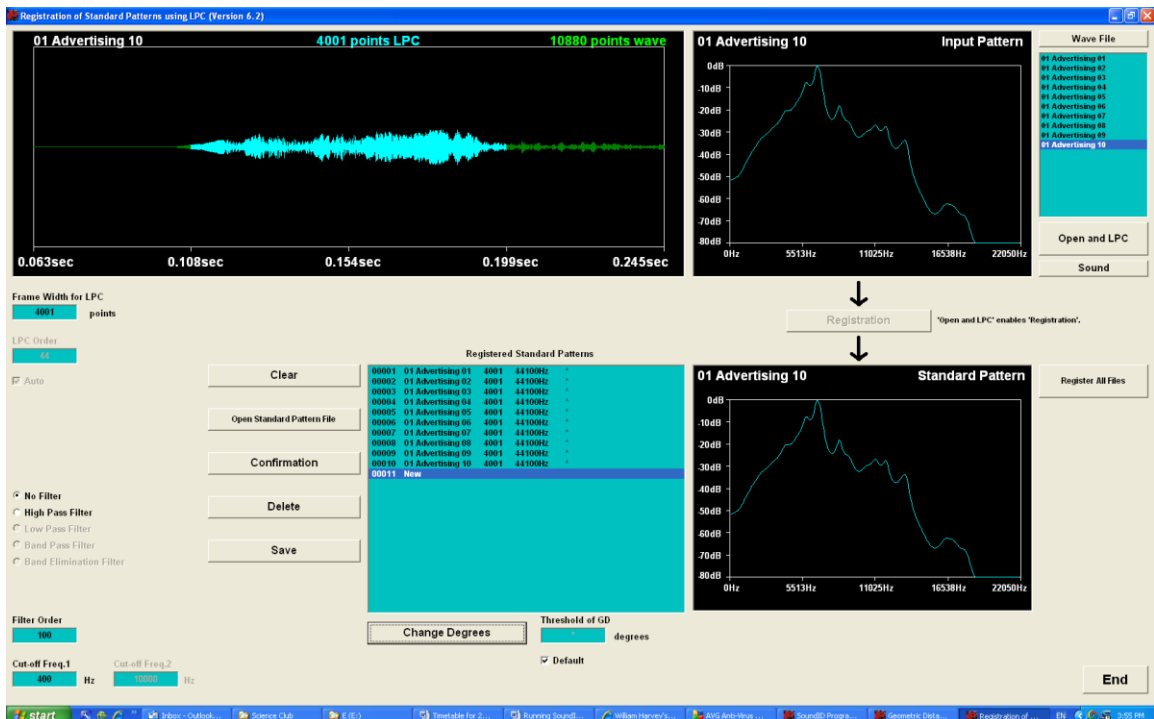
At this stage you can register one file or all of them at once. Let's register the lot by clicking the button "Register All Files", (on the middle right of the screen). This will cause the "Registered Standard Patterns" list box to fill out like this.

Now click save and give the file a name as below in the "Patterns" folder.
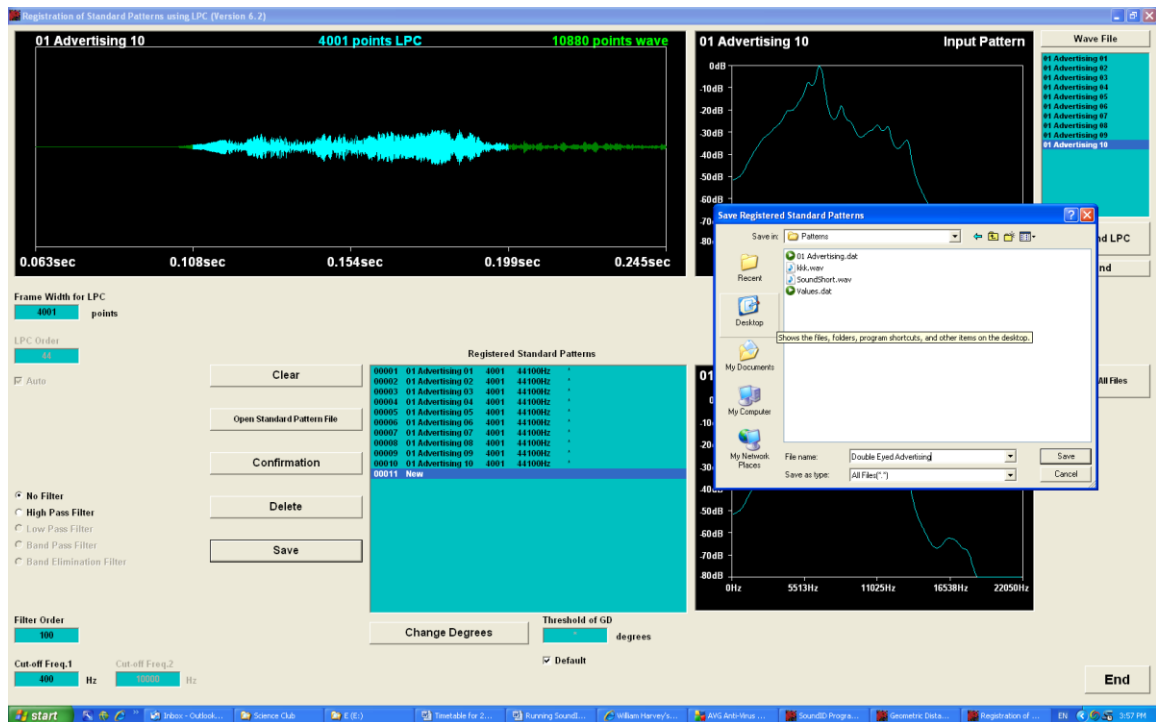
   You now have a reference file called "Double Eyed Fig-parrot Advertising" that holds examples of all of these calls.

What is saved in the Registration process?

You can see what is saved in the Registration process by clicking on the Menu item "Other Modules" (top left). Next click on template file header and use the Windows explorer to find the templates you have saved. This will display the information that is saved in the header of the template. All except the last two items in the list "Low Freq" and High Freq" are details that appear on the Registration screen. During the process of saving the templates the software will look at the peak frequency of each of the sounds in the template and record the highest and lowest frequency that was found.

New Version of Registration

The new version of Registration now saves all the settings that it displays and automatically uses these settings when running either Recognition or the Batch Recognition. Below you will see a screen shot with the parameters that are now automatically sent to the Recognition modules from the .dat file.

In the past this was not done and so a lot of parameters (like preferred GD) had to be set to one fixed value per run. With the new system the values can be set on a per .dat file basis and so the run is much better optimised.

If you have older .dat files that did not save all of these parameters you can either untick the check boxes in the red circles above and set the values manually as before or you can update the .dat files as seen below.
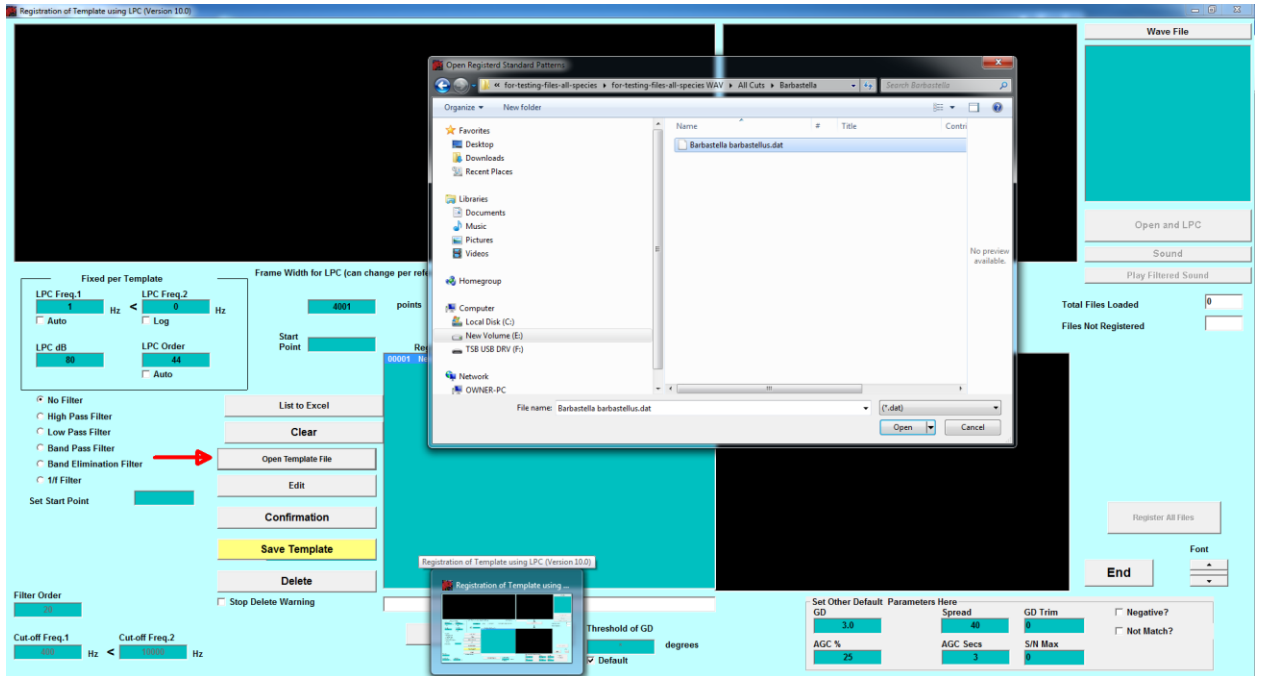
Click on the "Open Template File" button and it will load your original Template file.
Now set the parameters that you need to set and then press the yellow "Save
Template" button to save them, along with the original data.

Chapter 3

Using Recognition for Long Files

SoundID has been designed to search massive volumes of recordings for matches. To run this program you need first to create a reference file (see Chapter 2), but we have included a brief reference file for those too impatient to make the reference.

From the menu below choose "Recognition"
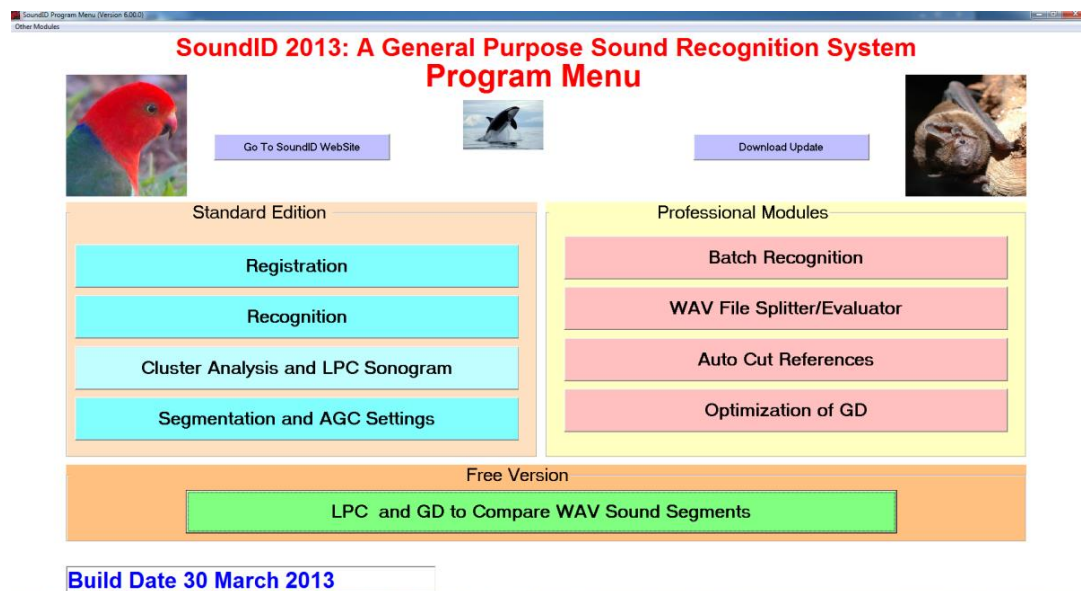
You will now have a screen like this.
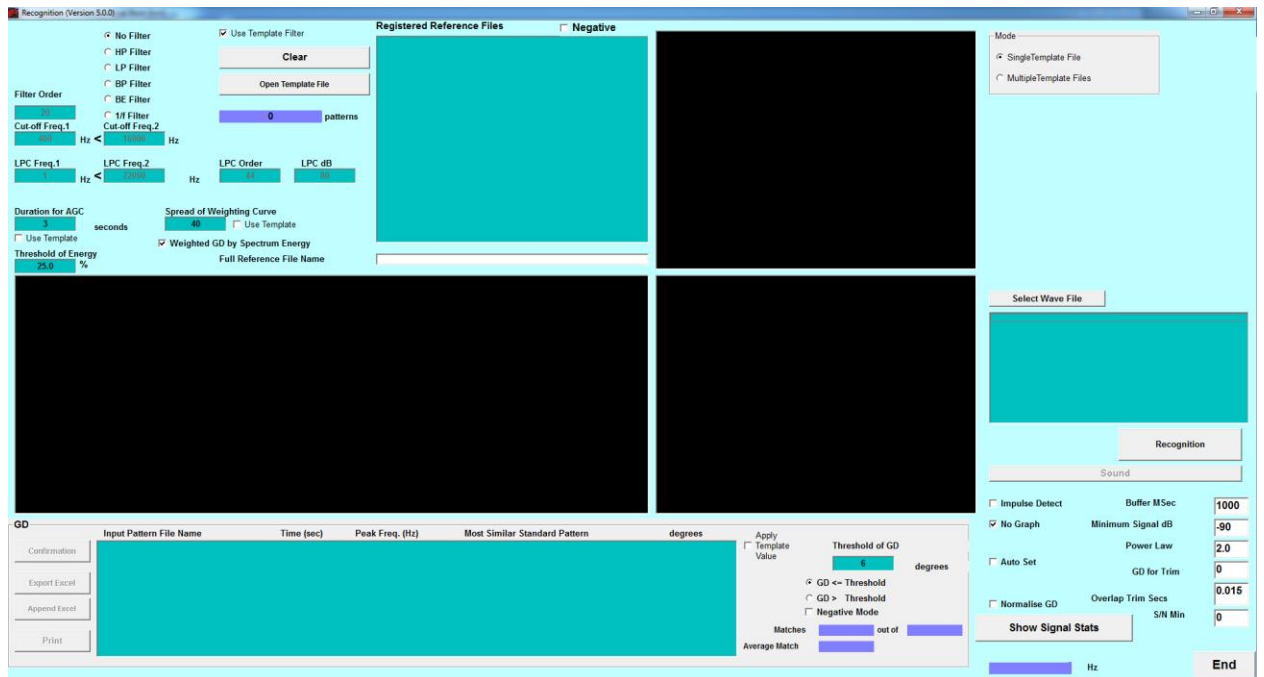
To load the reference file click on the "Open" button (top left) and if you want to use the reference file included it is located at C:\Program Files\SoundID\Patterns\Double Eyed Fig-parrot Advertising.
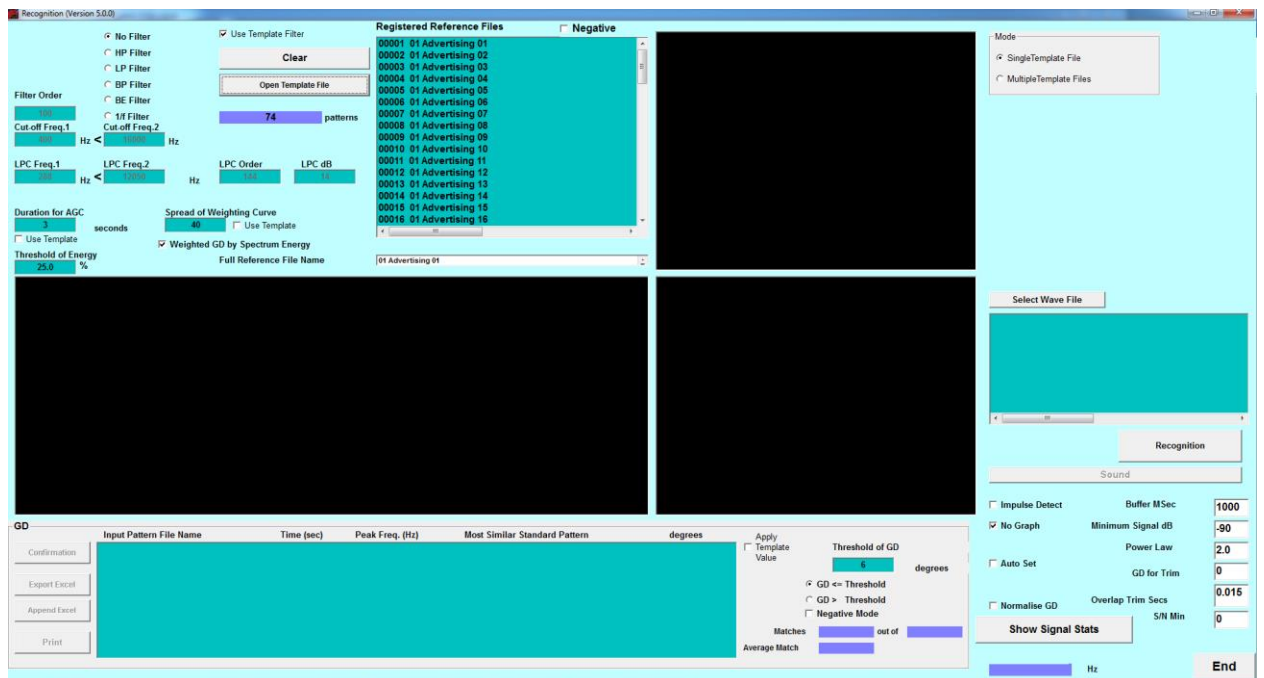
Next we need to open the file that we want to scan. An example will be found at C:\Program Files\ SoundID\ Recognition\Recorded WAV files\ 01_____AUDIOTRACK_01
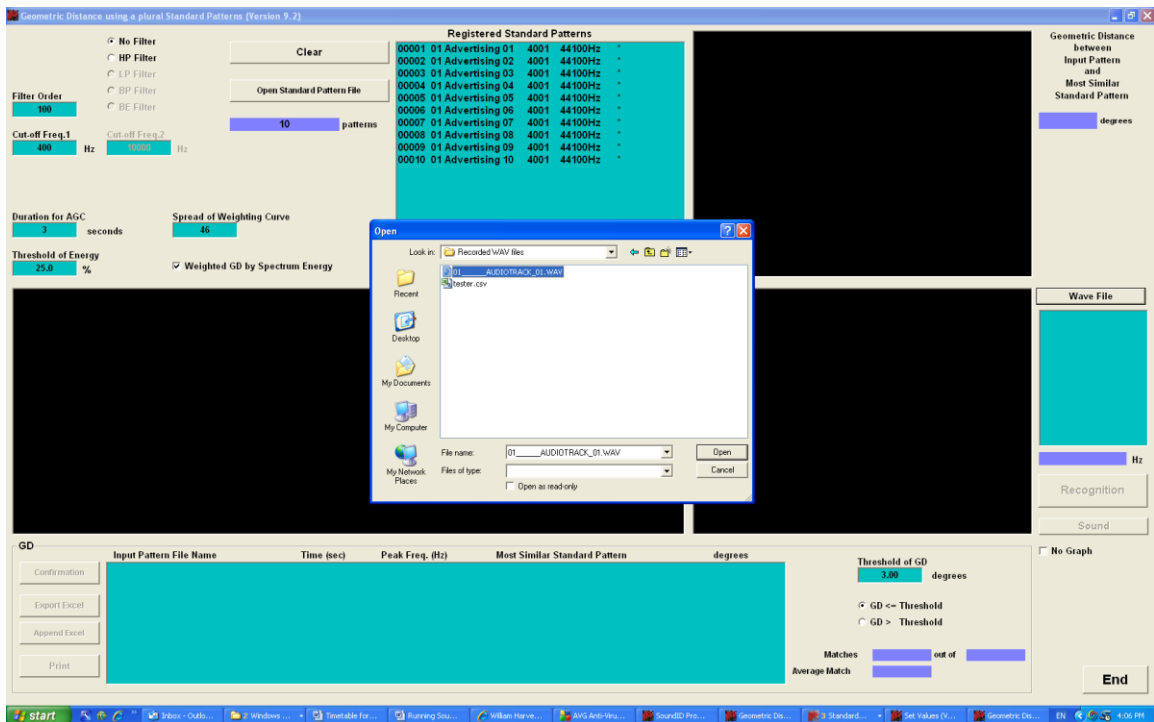
And then "Open" the file.
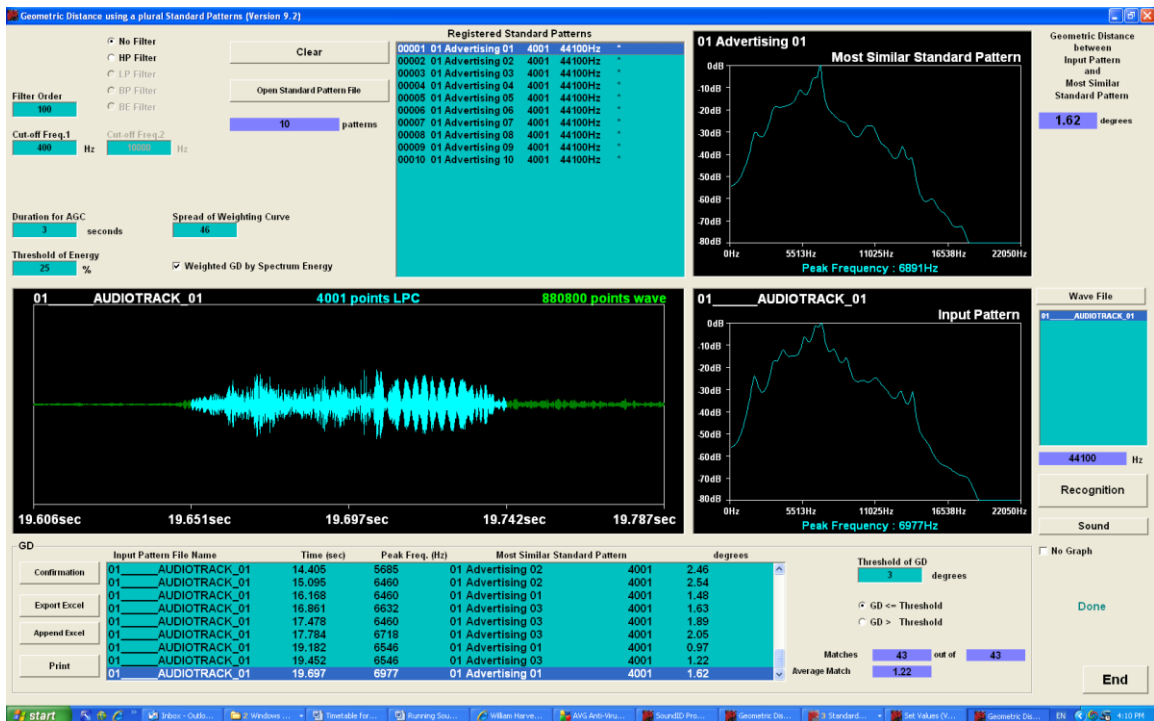
A Word of Caution
The default in the Registration and Recognition programs is no HP filter. If you use the filter on one of the programs and not the other, there will be a slight mismatch between the reference calls and the recognition. There is nothing to warn you of that. A good way to keep track is to add the filter bandwidth to the name of the reference file e.g. "Double Eyed Fig-parrot HP400".

This run shows an average match of GD=1.38 (which is close) over the whole of the recorded file. The list box in the bottom left lists all matches that are closer than GD = 3.0 degrees and there are 23 of them out of 61 calls detected on that file. The average match refers to the average match of the 23 calls (and not to the whole 61).

Notice that the first 10 matches have GD = 0. This is because the first 10 reference calls were cut from the *same* file, namely "01 AUDIOTRACK 01"

Again we can turn off the graphics to significantly increase the processing speed. We also have a high pass filter option if needed.

Recognition with Multiple Reference Files

The earlier versions of the software permitted a few parameters to vary on a per-sound basis. The new version permits any number of references to be used in a single study, each having its own optimised settings. Note that while this dramatically increases the accuracy (particularly for things like the dawn chorus) it does come at the expense of increased computing time. The extra computing time is roughly the time for one reference file multiplied by the number of reference files used. Should two reference files use identical settings then this will be detected by the software and they will be merged into a single larger reference (and this will run the merged files almost as fast as a single reference file).
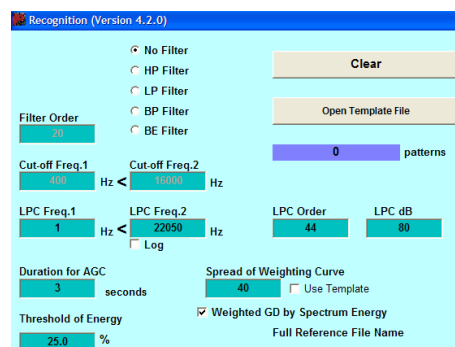
Notice that all of the settings that *need* to be tied to the original Registered templates settings will be set accordingly at run-time. Some settings are still optional and in particular the *spread* can be set for a particular run or it can be set to the value at the time the template was made. The reason for this is that some original references might have an optimum spread and this can be added as an option at the time of running the Registration. You will see in the screen-shot opposite that you may use the Registration value by ticking the "Use Template" check box.

The new screen, shown below offers the choice of multiple reference files.
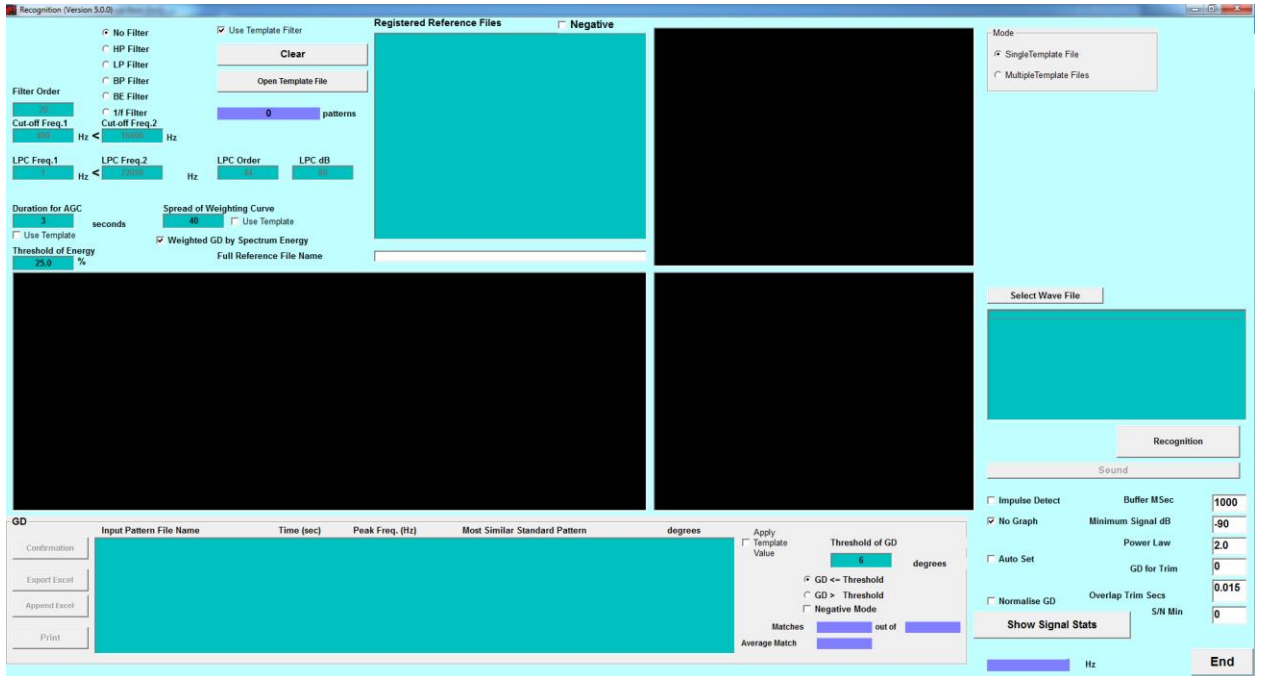
If we click on the "Multiple Templates Files" the screen will appear as below. Here
we have clicked on "Select References" and added a batch of ".dat" files to run. Then
we have clicked on to the "Load References" button. This causes the first of the files
to load into the Registered Standard Patterns" list-Box and you will note it has
replaced the default settings with the ones that are associated with the Sedge Frog
reference pattern.

Next we load the WAV files to be run using the "Select WAV File" button as before.
If we now press the "Run All References" button the software will begin to run the
references and their respective settings sequentially.
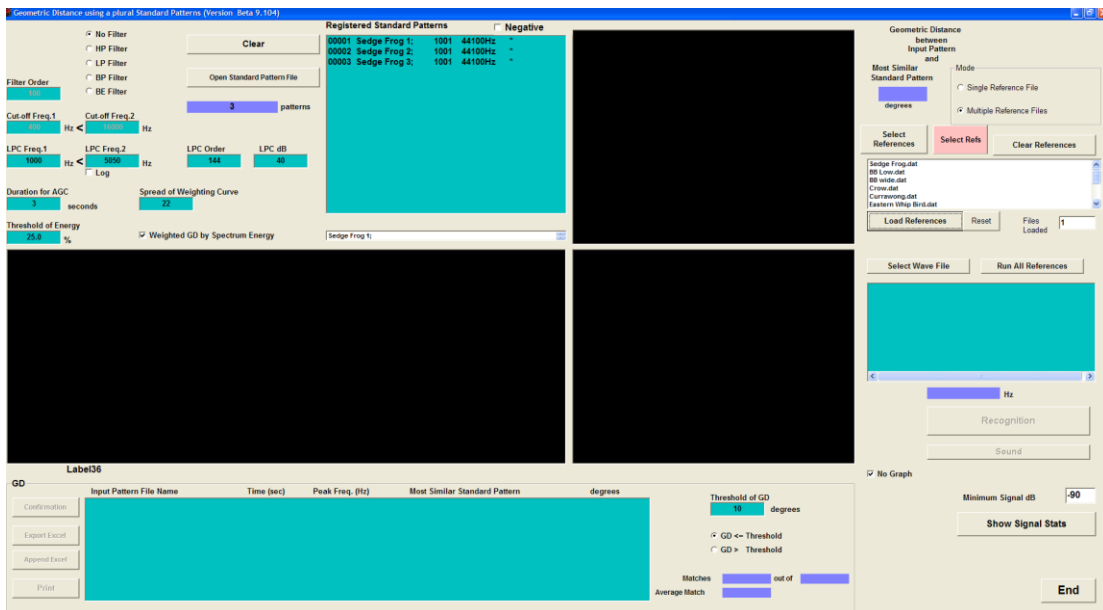
When all of the references have been run a new screen as seen below will pop up. These references and the recording being studied are those of the Dawn Chorus in Maleny, Queensland, Australia.
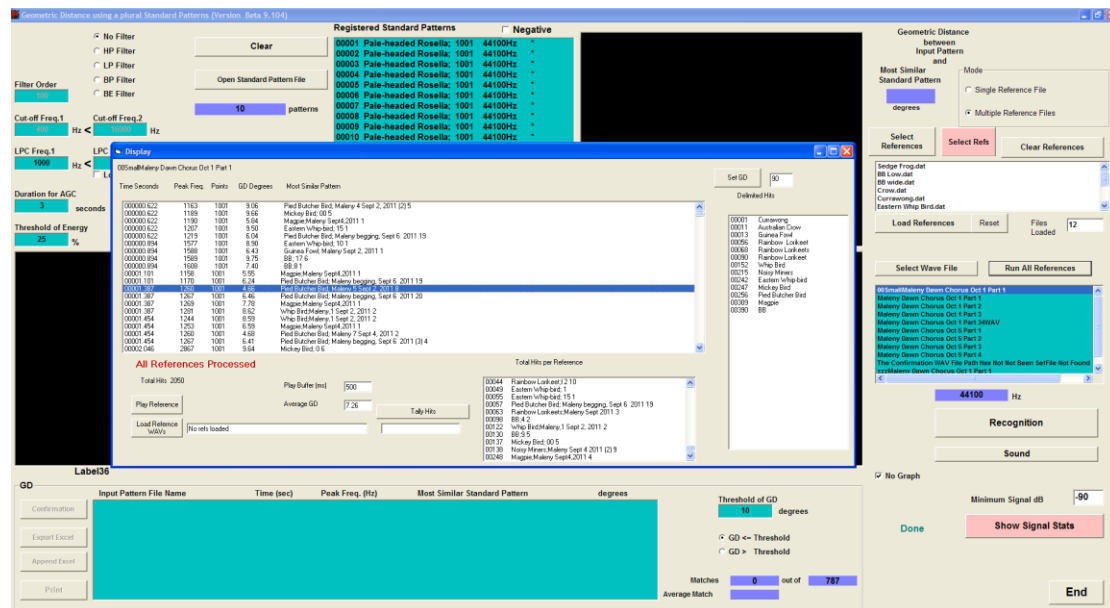
Name Length

The first thing to notice is that the new list has the "long" names, whereas the recognition screen does not. The original recognition screen had names limited to 20 characters. The new version has a limit of 128 characters.

There is a complete list of all the "hits" and you can confirm these by clicking on them to play the respective sound. You can also play the "hit" against the original reference if you press the "Load Reference WAVs" and use the resultant search window to locate the original WAV files that the references were cut from (they should all be contained in one folder).

The window will also present the user with a complete summary of how many of each species were found (in the right hand list box see that the BB (Butcher Bird) was found 390 times and the magpie 309 times. The adjacent lower box lists the number of times that any particular reference was matched. Be suspicious of any particular reference matching too often as bad (and noisy) references tend to match simply noise with noise.

Species List

A new feature here is the species list. Using a naming convention with a semi-colon delimiter, you can get the species list (top right). So for example the reference file name might be

Butcher Bird, Maleny, April, 2011
If you rename it to

Butcher Bird; Maleny, April, 2011

The software will treat everything to the left of the semi-colon as the species and everything after it as metadata.

Powerful Options

	Recognition	
	Sound	
☐ Impulse Detect	**Buffer MSec**	1000
☑ No Graph	**Minimum Signal dB**	-90
	Power Law	2.0
☐ Auto Set	**GD for Trim**	0
	Overlap Trim Secs	0.015
☐ Normalise GD	**S/N Min**	0
Show Signal Stats		
	Hz	**End**

In the illustration above we see a set of powerful new options for 2013. The default values can be used for most applications but there will be times when it is best to customise some of these.

Buffer mSec

This set a sound buffer before and after the segment that was matched to be used when playing the sound. The 1 second (1000 milli-second) buffer makes it easy for the user to confirm that the identification is correct. The time resolution of the human ear is about 100 milliseconds and sometimes the ID will be based on time frames as small as or even smaller than this making human verification virtually impossible.

Minimum Signal dB

This sets the minimum signal measured as dBm0 (level in dB below 0 dBm on the recording). At -90 dB there is virtually no signal affected, but were it to be set at say 40 dB all low level signals would simply be ignored. This is particularly useful if you are using field recordings to extract references as it can be used to ignore all but the best signals.

Power Law

Traditionally the most popular method for signal detection is to measure the signal power at any point in the recording and then to use variations in the power level to determine if a signal is present. This of course occurs before the recognition phase as for now we are only concerned with determining if there is a signal that might be worth looking at.
Other people have found different measures of the signal from the entropy to powers as high as 5.5 of the amplitude work best for their particular signals. The useful range for the power law value is about 0.5 to 6.0. If you find the software is failing to detect some signals it might be worth looking at this value. It is best to see its effect in the "Auto Cut References" module.

Multiple Coincident Calls

This multiple reference software can discriminate multiple overlapping calls. The calls in the above list are sorted according to their time and if multiple calls overlap, they may be listed as occurring at the same time.

Overlap Trim Secs

There are two ways to address overlapping calls. First it is possible (probable) that not all templates in the templates list will have the same frame width. This means that calls that are coincident in time may overlap a little. By setting the "Overlap Time Trim" (default 0.015 seconds) this will be taken care of. Setting its value to "0.0" will disable this function.

GD for Trim

Next we have "GD for Trim". Set to default of zero this disables the function but set to any other value it defines how much the coincident match has to exceed the best match before it is deleted. So for example if two coincident signals have a GD or 3 and 4 respectively and the "GD for Trim" is 2 then both signals will be displayed. If the "GD for Trim" were set to 0.5 then only the match of GD=3 will be displayed.

Set GD

A new and powerful feature is the button at the bottom right. (Apply template Value to GD). Because the software is now trading processor time for accuracy we want to limit the number of times anything has to be run. This can be done by setting a high GD initially and in the example it was set to 10. Then if you want to see the result with any other GD just reset the value and it will recompute and relist the values.

Chapter 4

Batch Recognition

Batch recognition is designed to allow the searching of large collections of recordings. It can search folders of even whole HDDs and has been routinely used on terabyte HDDs holding thousands of hours of half hour recordings in multiple folders. It can take some time to run such large volumes of recordings and it can be advantageous to divide the recordings into groups of roughly equal size and run multiple instances of the "batch recognition" module on PCs with multiple processors. Note there is not much advantage in running more instances than there are processors in the PC. A bit of experimentation will find the best number of instances. Also if the PC is to be used for other purposes at the same time and the idea is to run the search in the background, the number of instances should be one less than the number of processors, otherwise the batch program will have all the PC resources tied up and make it run very slowly.

> **Caution:** This module uses the results from Recognition and *assumes* that 3 or more consecutive spaces delimits the data. Therefore a file name with 3 or more consecutive spaces will be mis-read as two separate variables, and will cause problems. This applies both to the original recordings and the Reference file names.

Multiple Processors

The software has been designed to run with multiple processors, and will put the results in different folders for each processor. The results of the runs are stored in a folder called tmpResults {sequence number}. The sequence number is allocated uniquely to each processor. Note the prefix "tmp" signifies that these folders are temporary and will be overwritten each time the PC is rebooted. So if you want to keep the results, make sure to save them before shutting down the PC.

Be careful not to select the same folders for two different instances of the program to scan. Conflicts can occur if a file is already opened and another attempt is made to open it, which can cause the run to crash.

The data stored in the folders is shown below. This is the same format as the output of Recognition. The space separated values are "file name", "time" (in seconds to hit), "peak frequency", "nearest matching file", "file length", "GD". The addresses of these files are shown in the "Results Saved in" text box.

```
"D:\AAAAAUP Full Wave\"
"Average Match  0.00     Matches  137     Total Hits  137"
"01_____AUDIOTRACK_01        1.964      6891      01 Advertising 01        4001      0.00"
"01_____AUDIOTRACK_01        2.142      6461      01 Advertising 02        4001      0.00"
"01_____AUDIOTRACK_01        2.303      6805      01 Advertising 03        4001      0.00"
"01_____AUDIOTRACK_01        2.527      5599      01 Advertising 04        4001      0.00"
"01_____AUDIOTRACK_01        2.716      5513      01 Advertising 05        4001      0.00"
"01_____AUDIOTRACK_01        2.901      5858      01 Advertising 06        4001      0.00"
"01_____AUDIOTRACK_01        3.110      5772      01 Advertising 07        4001      0.00"
"01_____AUDIOTRACK_01        3.650      5772      01 Advertising 08        4001      0.00"
"01_____AUDIOTRACK_01        3.847      5858      01 Advertising 09        4001      0.00"
```

|						
"01_____AUDIOTRACK_01	4.612	6633	01 Advertising 10	4001	0.00"
"01_____AUDIOTRACK_01	4.863	6891	01 Advertising 11	4001	0.00"
"01_____AUDIOTRACK_01	5.265	6891	01 Advertising 12	4001	0.00"
"01_____AUDIOTRACK_01	6.283	6461	01 Advertising 13	4001	0.00"

Note the reference file name " 01 Advertising 13" has single spaces, which is OK, and even two consecutive spaces would not be a problem. But three or more spaces will cause the program to read this string incorrectly.

Batch Recognition

Click "Batch Recognition" from the menu and you will notice that two forms are loaded as seen in the screenshot below.
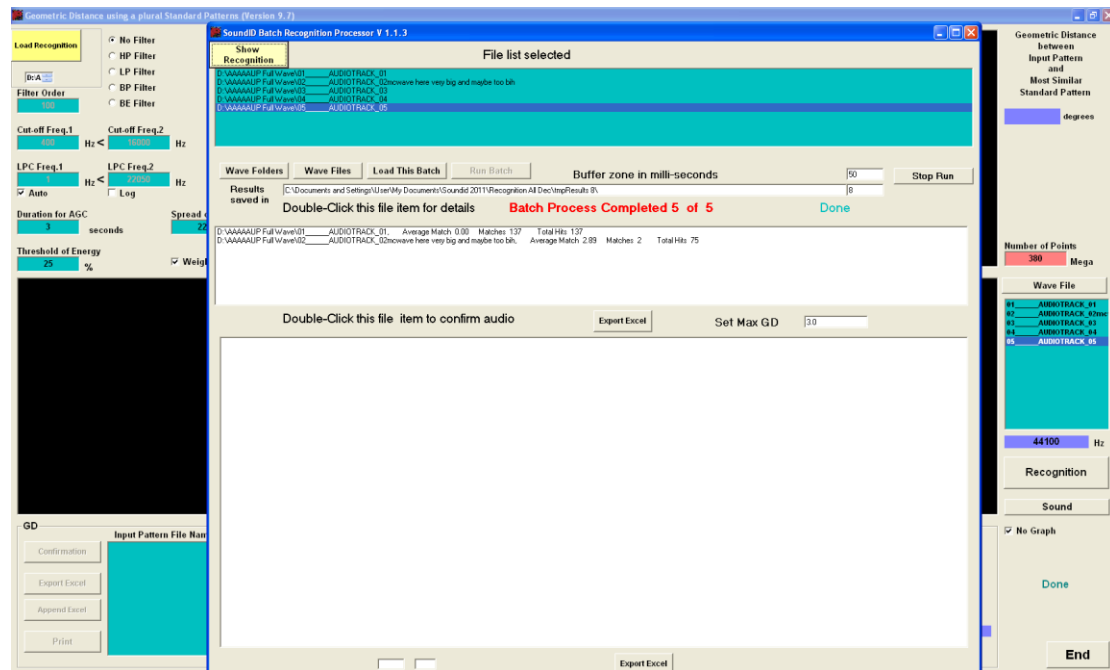

The two yellow buttons enable the user to toggle between these screens. However if the user has two screens it will be best to drag the batch processor to the second screen and click the maximiser button so that it fills the second screen.
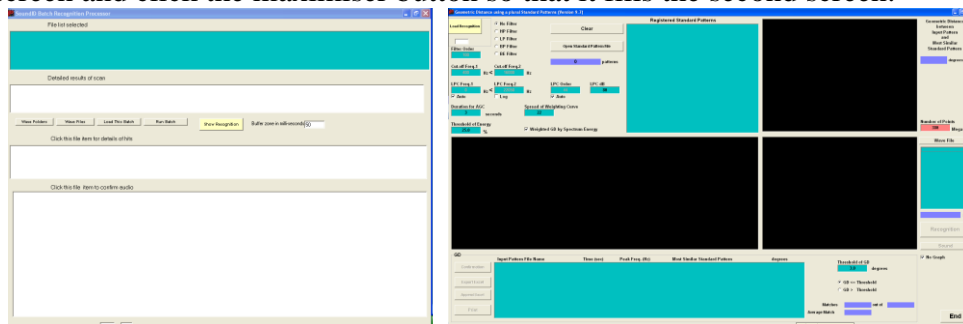
With two screens we have the display shown above. The form on the right is the familiar Recognition module and it works in much the same way as the single file recognition. Keep in mind that the Batch form on the right will load in the files to be processed as required, but that the processing is done on a file by file basis in Recognition. The results of the processing are sent back to the batch form.

Consequently it is necessary to set any needed parameters in the Recognition Window. However the files to be examined are loaded in from the batch processor.
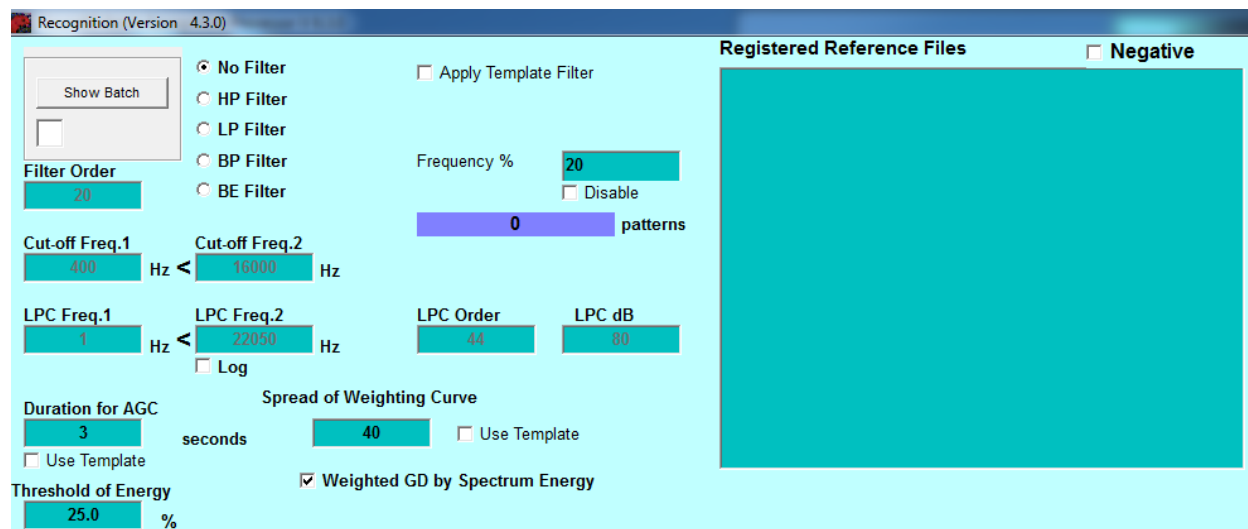
There are a number of things that cannot be set in this window including all of the text boxes that are "greyed out". These are set automatically to match the template currently being run.

You can set the filter, as required (even if it is originally "greyed out"), but remember this will slow down the processing significantly *and* it will interfere with the matching if the filter settings differ in any way from the settings used to create the template. For really noisy recordings this is sometimes worth doing anyhow.

Notice also (see diagram above) that you can automatically apply the same filter (if any) that was used in making the templates. This is done by clicking on the "Apply Template Filter" check box.

Template Bandwidth Check

You have an option to *disable* the frequency test. Each Template holds information about the maximum and minimum peak frequencies of the signals that make it up. By default the minimum will be decreased by 20% and the maximum increased by 20 %. In this mode the software automatically does *not* process any templates whose frequency max/min is outside the bandwidth of the target.

For example if the software looks at the target and finds its peak frequency is 300 Hz and the template peaks are 5000 Hz maximum and the minum is 1000 Hz (bandwidth 800 Hz to 6000 Hz once the 20% is factored in) then it will not process this signal because it is out of band. Instead it will go on to the next template and repeat the process until it finds a template that has 300 Hz within its bandwidth.

This process can significantly increase the processing speed when there are a large number of templates. However should it prove problematic you can disable it by clicking on the "Disable" checkbox.

The "duration of AGC" and "threshold" can be set maunally or linked to the value set when the template was made. The spread of the weighting curve can likewise be set or locked. Note that if these values were not purpose set when the template was made they will contain the default values in the Registration module.

As an example we first load the reference files in the Recognition and set the parameters to use as in the screen below. Note that no files, other than the registered patterns, are added directly into this Window.
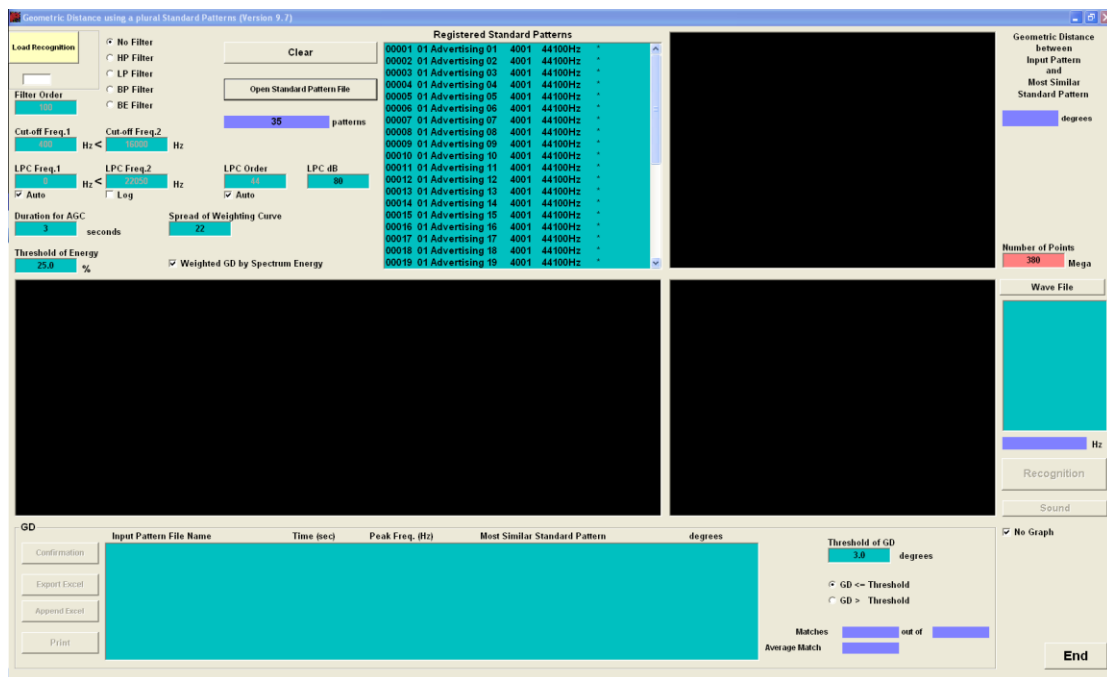

Next, from the Batch Window, click either the "Wave Folders" to load folders or drives, or "Wave Files" to load a group of files. The result is seen below.
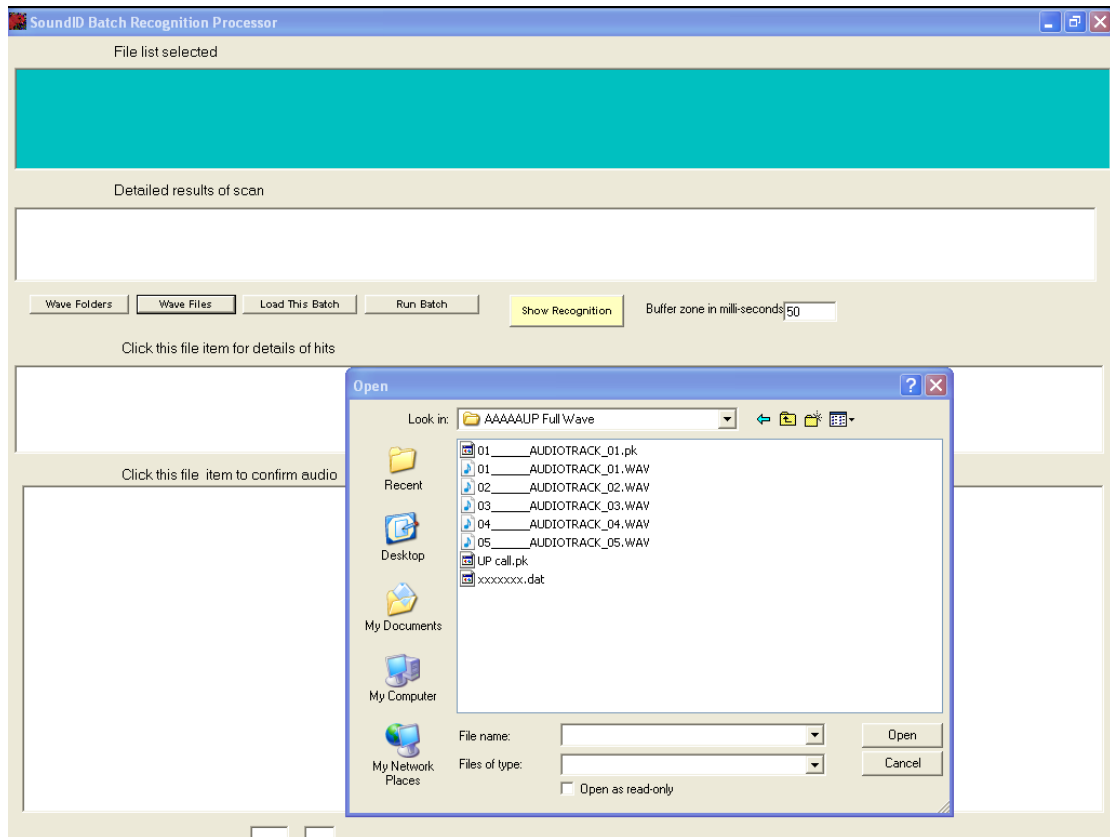
Once the files are selected we find the list of files loaded into the Batch Window. Click the "Load This Batch" button to send these files to the Recognition Window.

Now press the "Run Batch" button and you will see that Recognition runs all of these files sequentially. However the results are first transferred to the Batch list box. The results are shown below. Notice that for the truly impatient it is possible to confirm results, even while the program is still running during the Recognition cycle. This is not possible during the "Opening" and "Done" cycles (making this possible would slow the program execution time unnecessarily).

You will notice that in the example below the matches on all of these are "0.00", which is to be expected as the searched files were the same as the ones from which the references were cut. Also the best matches are sequentially in the file indicating the way the references were cut.
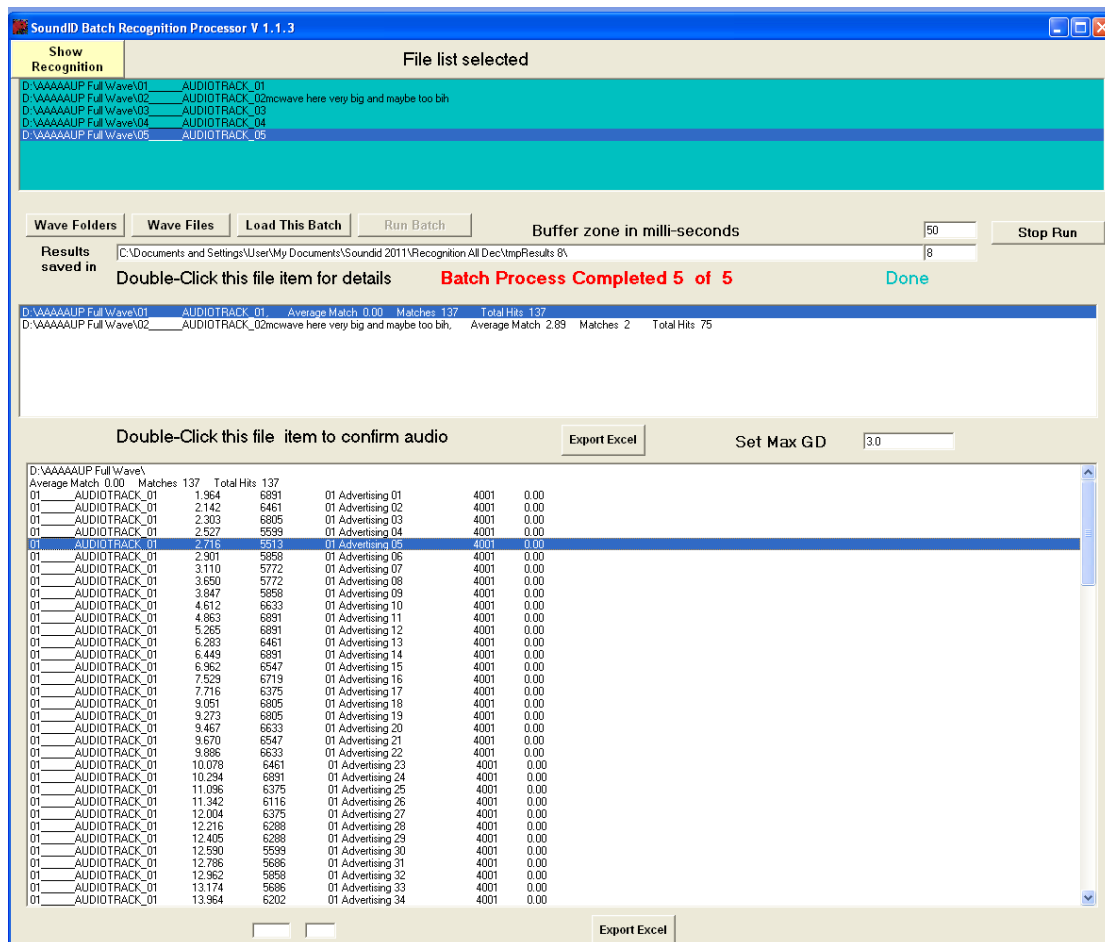
Setting the Recognition Bandwidth

At the top left you will find a few options. The first option "Apply a template filter, will apply the SAME filter that was used in the original template made at the time of Registration.

Below that you will see the "Frequency %" and its default value of 20%. This will set the pre-filter, which, during a batch run, has the effect of first looking at the target frequency and then looking at the highest and lowest frequency in that template. If the target frequency is either 20% lower than the lowest frequency or 20% higher than the highest frequency it will assume that there is no possible match to the target in the current template and cancel the scan for that template. This will speed up processing by about 10% to 200% depending on how many templates still need to be run. In the simple case of two templates which do not have overlapping frequency ranges, it is easy to see that the speed improvement will approach double. There may be circumstances when you do not wish to run this option and the disable check box will turn the option off.
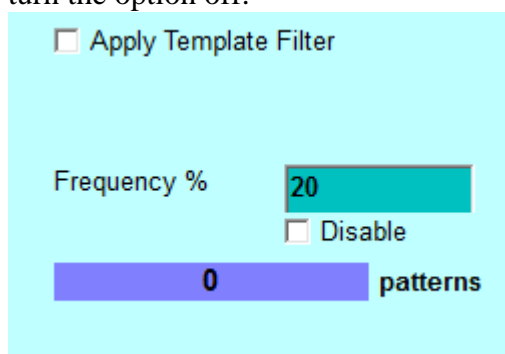

Note also that the frequency bandwidth information was first added in December of 2012 and templates made before this time cannot run this option. The software will automatically identify the older templates and so will disable the option for them. See page 32 for more information on this.

51

Setting GD

The GD is set in the Recognition module. However there is also a setting for GD in the batch module. This can save a lot of time. If for example the Recognition is set to 5.0 it may well find too many hits. So in the batch you can refine the search by setting a lower GD (say 2.8). This will load into confirmation only those matches that are 2.8 or less. Because this does not involve the actual calculation of the GD it is much faster, and allows the user to refine different GDs to suit different targets.

Loading Folders

Loading folders is a non-standard Windows operation so we will cover it in a bit of detail. On the upper left hand side there are two buttons, one called "Wave Files" and the other "Wave Folders". The "Wave Files" button will load files as does the "Recognition" module and will process them also in the same way.
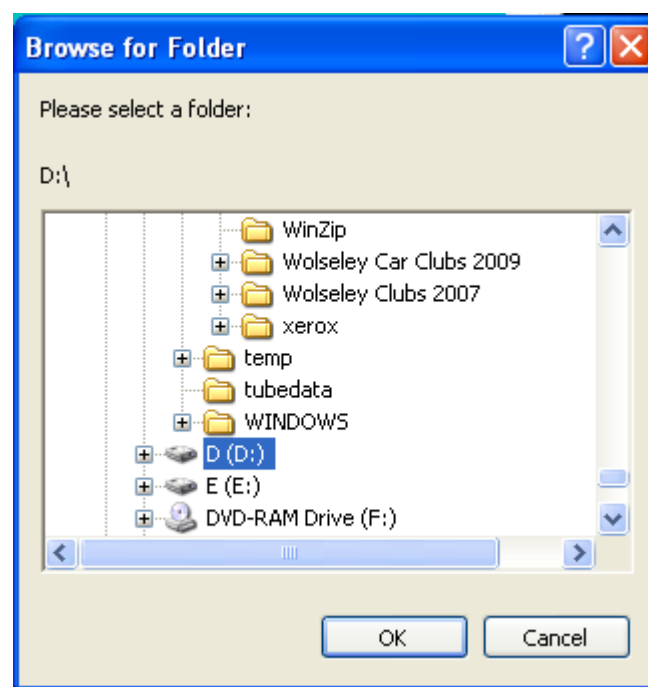
The "Wave Folders" button will open a Dialog box like the one above and list all the folders and HDDs on the PC. You can select a folder or a HDD and press OK to download all WAV files on that HDD. Alternatively you can open the HDD or any folder and select its subfolders. In any case, whatever is selected the module will open that folder (or HDD) and all its subfolders and load the WAV files into the list box.

> **Caution** The Batch mode uses the absolute addresses of the original files as they were run in Recognition. So it is essential that the drives not be removed after the usage. If for example the processing was done from drive "G:\", then the Batch software will look for the file on drive "G:\". If it is absent or if some other drive is in its place, the Batch mode will fail to find the file.

Confirmation

To confirm a result, double click on the file in the third list box, and the results from that file will be displayed in the bottom list box. Confirmation is done by double clicking on the items in the bottom list box. When double clicked the original sound from the recording will be played. Note that beside the buttons (below second list box) there is a buffer zone variable. This is a measure in milliseconds that determines the time before and after the recognised call that the wav is played. For most calls the 50 milliseconds default should be fine. It can be adjusted on a call by call basis.

| Confirmation buffer zone in milli-seconds	50

To confirm the recorded sound against the reference, it is first necessary to inform the software where the original WAV files can be found. This is best done using the "Locate Confirm Files" button, seen at the left and middle of the screen shot below. Once the WAV file location is found and clicked, it will enter the folder address into the text box adjacent tot the "Locate Confirm Files" button. Alternatively the address can be entered directly into the text box. Once this has been done it is possible to play the original reference from the "Confirm" button and play the recorded match by double clicking its list box line.

Note these WAV files need to have exactly the same name as the name used in the reference files. It does not need to be exactly the same as the reference file content as the software searches the folder that the WAVs are in and looks for a matching name. Thus this directory could be a master directory of original WAVs that includes many more than the current run. If a file is not found a confirmation cannot be run, but the program will continue to run normally.
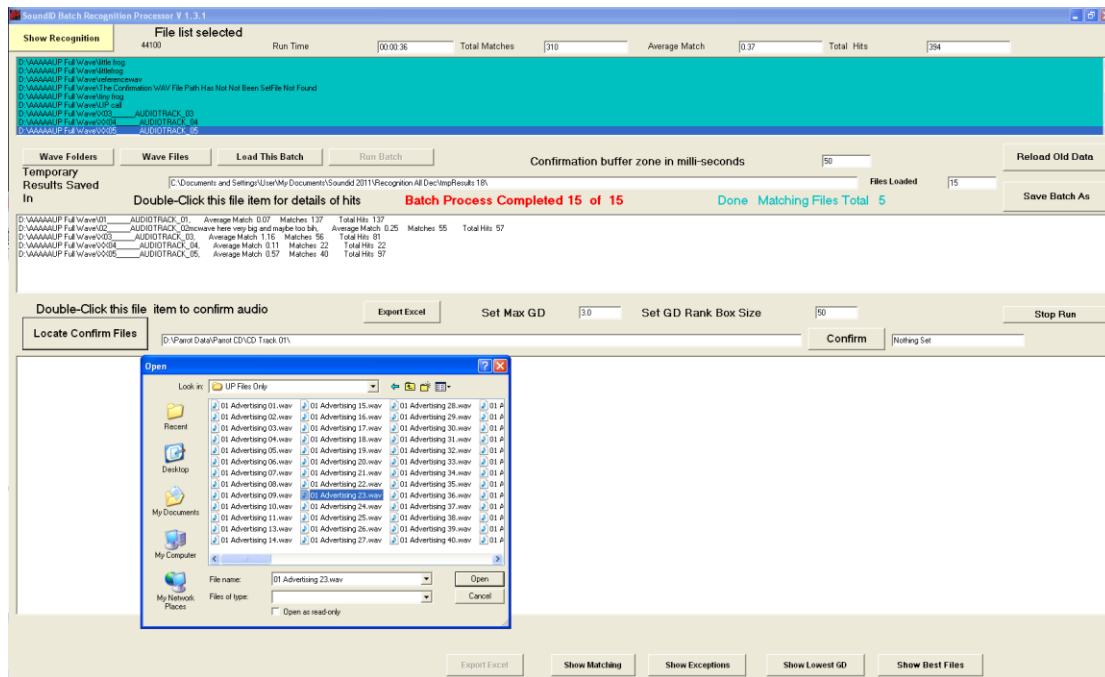
Post-Processing

There may be many thousands of matches as the result of some batch runs. In order to expedite processing of these some post-processing modules have been added. At the bottom of the Batch Window you will find the four buttons below.
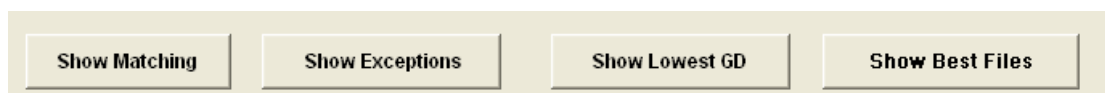

The Show Matching will list all the references that matched in the current run, sorted by order of how often they were matched. Clicking on any one of the listing will cause its original sound to be played and its image to be displayed (only if the WAV file location has been set). This enables a check on the quality of the call and its relevance (often reference calls from CDs have other calls mixed in with them).

Poor quality calls (often with a lot of noise) will sometimes reveal themselves as frequently matching signals. This is because there is a lot of noise out there that they can match with. Always check the higher matches for quality and correctness (stray calls often get mixed in with references).

The maximum level of the call is reported as the maximum averaged over a specified number of samples (set at the top right and defaulting to 1000). This value should ideally be -25 dB or higher. The minimum is likewise reported. The saturation count should be zero, as it measures the number of time the recorder was overloaded by the signal. However a small count here can be OK.

You can also zoom in on this graphic by left clicking and dragging the mouse to lasso the area you want to zoom into. When you release the mouse it will zoom on the lassoed area. You can restore the original graphic with a right click.
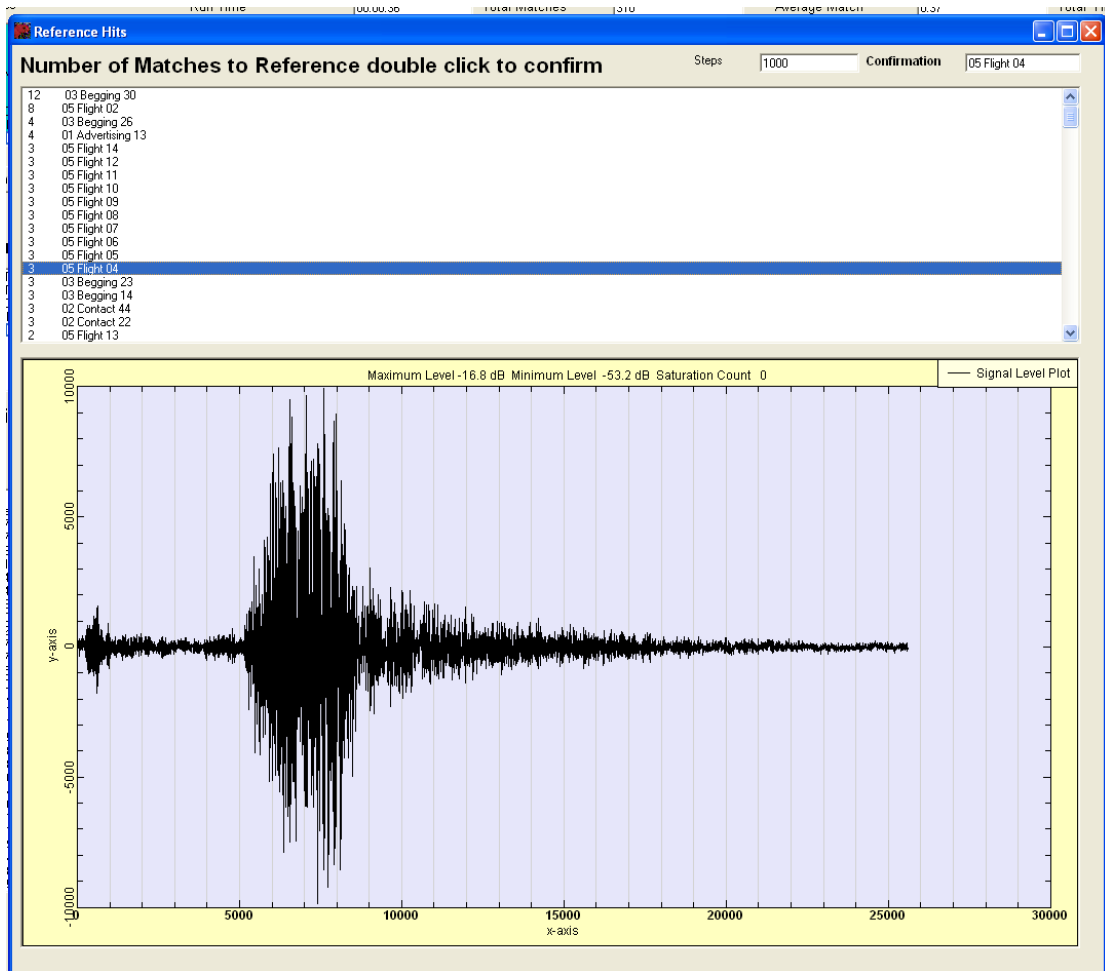
The Show Exceptions button reports on any files that were not processed and the reason that they were not.
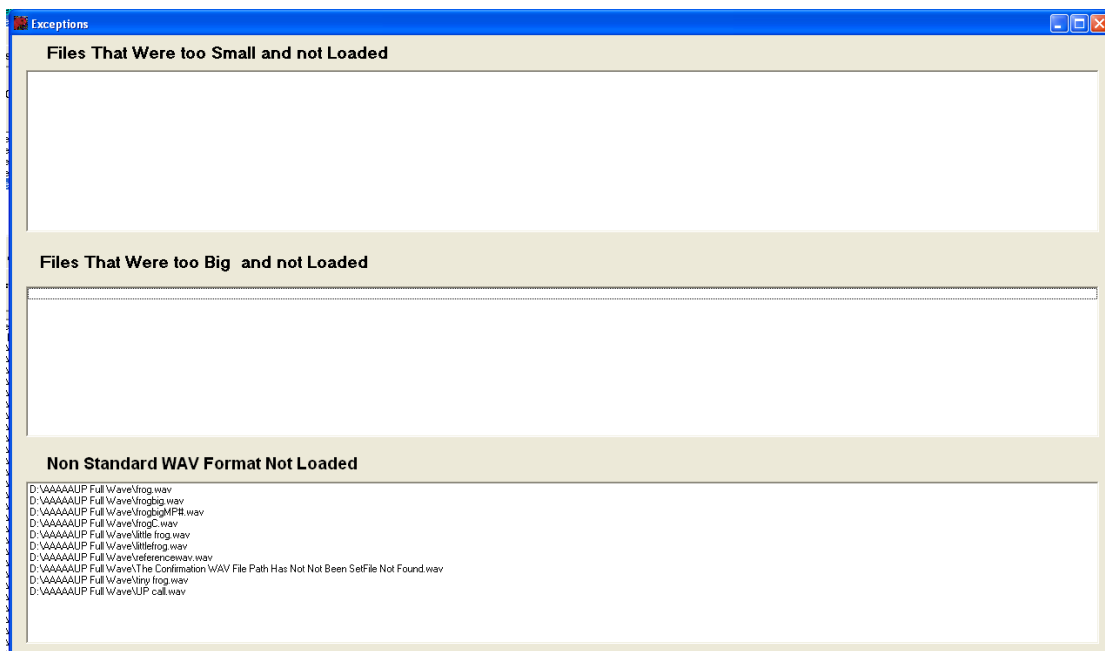
The Show Lowest GD button will show a list of all of the matches sorted in order of the best matches first. The number of matches in the list will be determined by the value in the "Set GD Rank Box Size". This value defaults to 50, but can be set as high as 32,000. Again in this report it is possible to hear the original sound by double clicking the listing, and to confirm it against its reference match by clicking the confirm button as seen below.
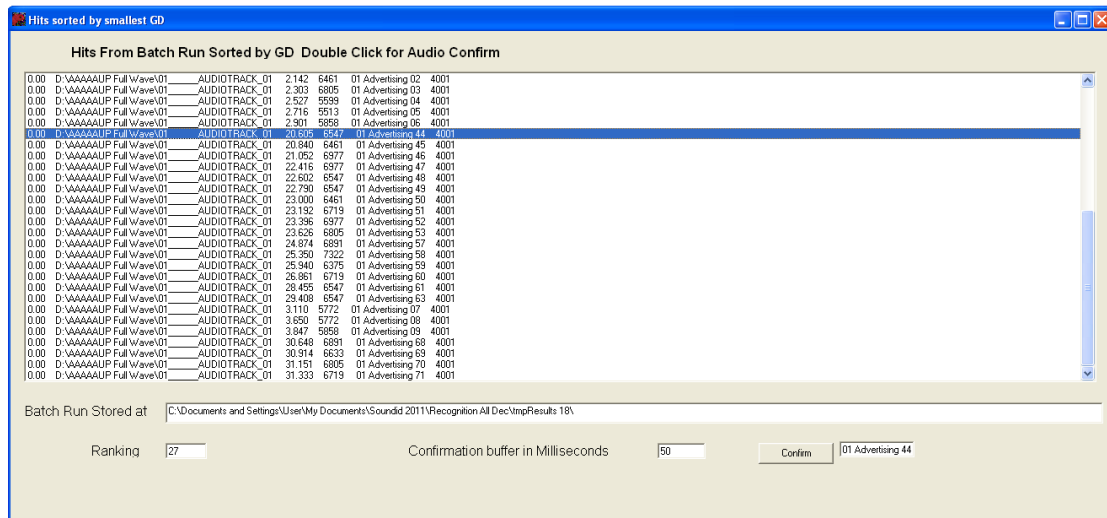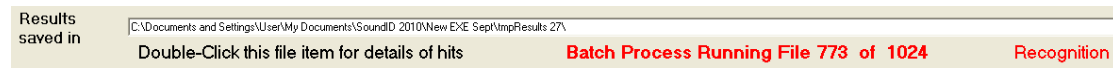

The button on the right selects the files with the most promising number and quality of matches. A number equal to {The average value of matching GD}/{The number of matches in the file} is calculated. The smaller this number the more promising is the match. When you select this view the files that were examined are listed in a ranking based on this number. Double click on the file name to reveal the actual matches in that file. If the actual matches are in turn double clicked the original recording will be played. Confirmation against the reference file can also be done by pressing the confirm button.
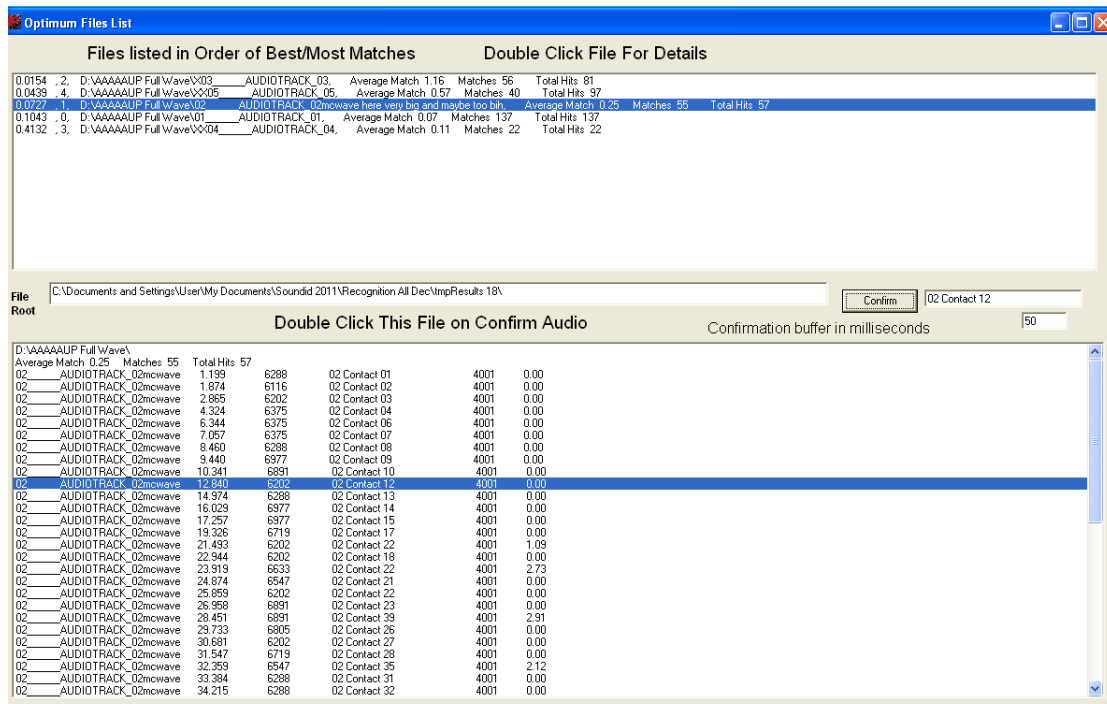
Notice that closing any of these reports can cause a loss of the data that is in them and so a message box will prompt you to be aware of this in case you accidently hit the close Window button. If all the results have been saved in the Batch Window then they can be recovered.

Saving a Batch Run

A Batch run can take a lot of time and so it may be a good idea to save the results. While a Batch run is underway the progressive results are saved in a temporary folder as seen below.

Results saved in	C:\Documents and Settings\User\My Documents\SoundID 2010\New EXE Sept\tmpResults 27\	
	Double-Click this file item for details of hits	**Batch Process Running File 773 of 1024**	Recognition

This folder is a temporary folder in that it may be overwritten on a subsequent run. However it can be copied and saved like any other folder or this can be done from the software using the two buttons below. The file structure is such that folder "tmpResults 1" holds the results of the first WAV file examined, in full detail in file "0 File " (file number zero). The other WAV files are in "tmpResults 2" "tmpResults 3" ..... "tmpResults n" respectively. tmpResults1, 2 …. n reside in the same directory as the SoundID program.
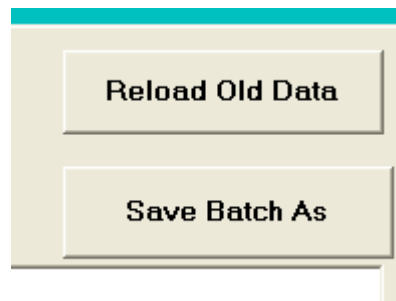
While running, these buttons are greyed-out and cannot be used. However once the run is completed the "Save Batch As' button will prompt you to save the results. These can later be retrieved under the 'Reload Old Data" button.

Stop Run

There is a Stop Run button in the Batch mode as seen below that will stop the run during the Recognition phase. It will not stop instantly but will let the current Recognition phase complete first. This is a useful alternative to closing the Window and losing all data. Once stopped all the grey-out buttons will become visible and the results can be viewed and processed as they would in a normally completed run.

You may have to press this button a number of times to get the run to stop as there are no interrupts allowed during a calculation (this would seriously slow the processing).

Example: Searching 1024 half-hour Recordings for the Coxen's Fig Parrot

Using the batch mode it is possible to download a whole HDD or any folder (with sub-folders) to be examined. In this study 1024 half-hour files (512 hours of recordings) are studied.

The objective of the study is to find the rare Coxen's Fig Parrot and the files are searched for matches to it. The recordings are from Mary Cairncross Reserve, a rainforest park in Maleny, Queensland. In the same park there are numerous other parrots and these are searched for separately as a calibration.

We use the 195 reference calls that we have of the related species the Double Eyed Fig Parrot, with the GD set at 3.5. From previous studies with this target species, we have found that a GD of 3.0 is sufficient we can set it a bit higher to see what is in the GD=3.0-3.5 zone, as seen in Figure 1 below..
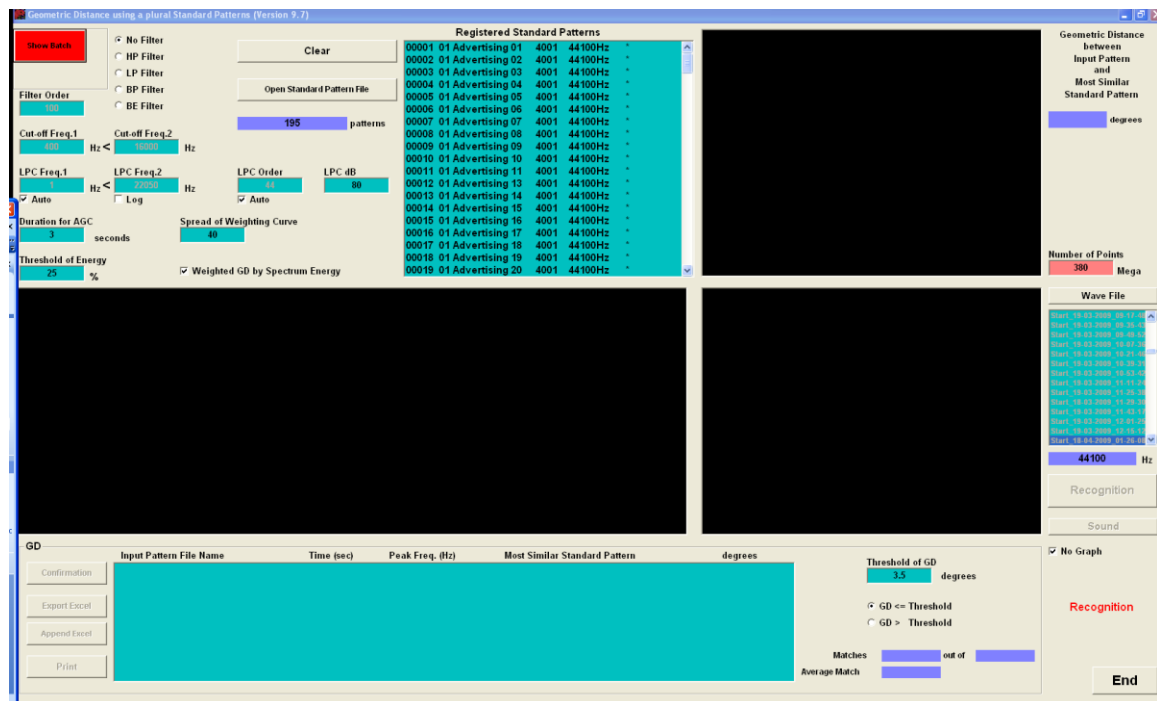

Figure 1. The setting of the recognition module

We then loaded the 1024 WAV files into the Batch module and set it running as seen in Figure 2. The results are summarised across the top. The run took 15 hours and 48 minutes, there were 1.332,314 sounds examined. There was a lot of rain in the period of the recoding and a lot of the sounds that would have been examined were the sounds of the rain. The rain noise therefore slows up the recognition process, by doubling the number of sounds that need to be examined.

 At a GD of 3.5 only two matches were found and both were poorer matches that the 3.0 we would have hoped for. It is possible from the Batch mode to click on any one

of the results and hear a playback of the call, for audio confirmation. One of these is definitely a parrot, with a lot of other birds calling on top, the second is less certain as there are many other birds calling. If we declare both to be false positives, we have a false positive rate of 2/1332314 x 100 = 0.00015%..
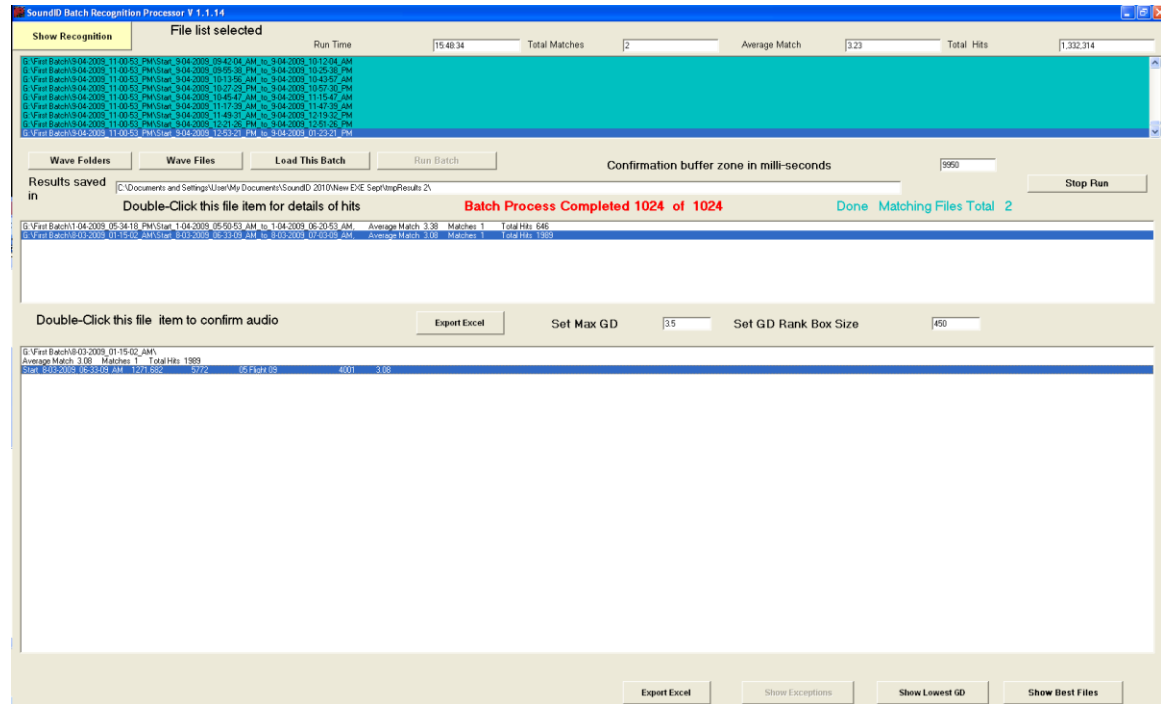

Figure 2. The Batch run for the Coxen's Fig Parrot.

The next run contained the non-Coxen's parrots, which included the Musk Lorikeet, Little Lorikeet, Scaly-breasted Lorikeet, Rainbow Lorikeet and Australian King Parrot. There were 364 reference calls in total for these birds. All of these birds are likely to be in the same general area as the Coxen's. Figure 3 below shows the batch run about ¾ of the way through. At the end of the run there were 2414 matches, most of which were to the musk Lorikeet flying/feeding calls. Of these, 286 have a GD of 3.00 or less. However most of the ones with GD up to 3.5 clearly have a parrot call in them.

The rainforest is a very special environment, not only for the inhabitants, but for the propagation of sound. High frequencies will be attenuated more than the lows, and this will distort the sound heard as though it had been through a filter. Interestingly a lot of the matches are with the calls that look as though they have already been filtered by the forest as they show attenuation in the higher frequencies.
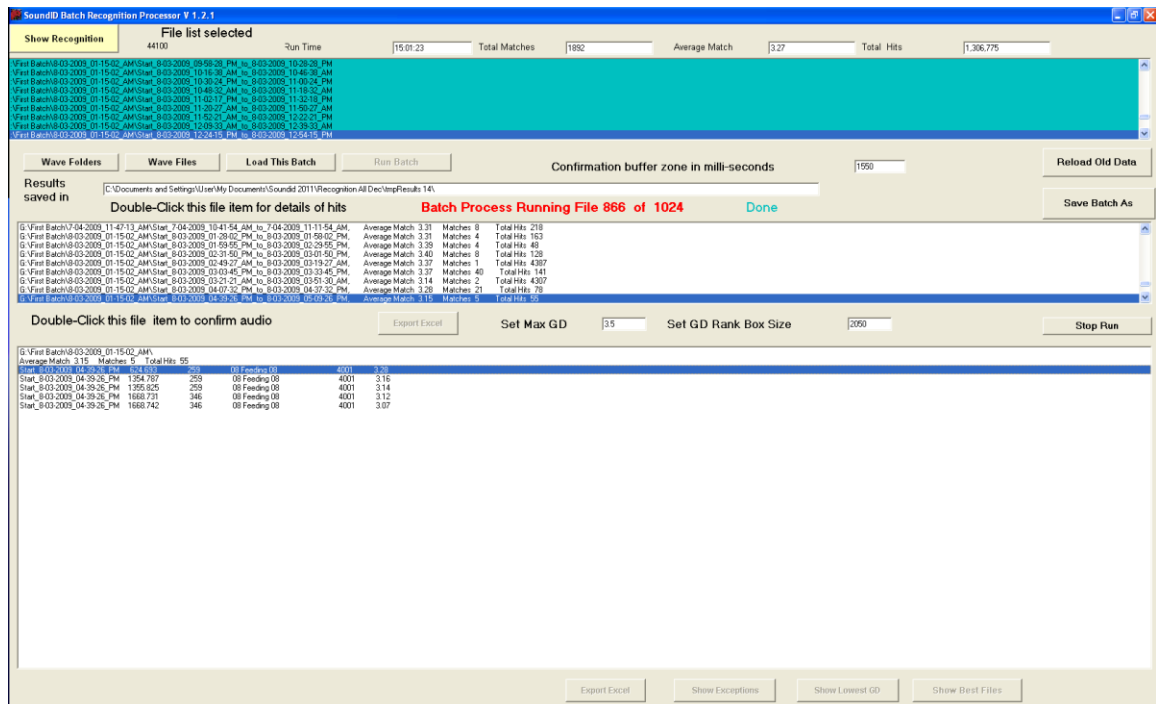
Figure 3 The non-Coxen's Fig Parrot detection

The reader might notice that the Non-Coxen's run was done on a later version of software. The later version will not change the results, but it does have the added ability to store all the results of the Batch run for later retrieval, without the need to run the recognition again.

A new reference file was then run with three different reference file sizes. The first run was done with all reference files of 4001 points. When there are multiple reference lengths a separate run is done for each length, and so the execution time will be extended. However as the lengths represent more truly the actual lengths of the calls, the run should also be more accurate.

As can be seen below this new run found six matches (where the original run found only 2), but the original two matches now have a larger GD (meaning they are less significant. None of the matches are less than GD=3.0. However the matches such as they are, should perhaps be referred for an expert opinion as the Coxen's is so rare than any possible match would be a significant thing.

This run took 26 hours 55 minutes and found 1,592,725 sound events to examine. This is 70% longer than the case when all the calls were of the same length.
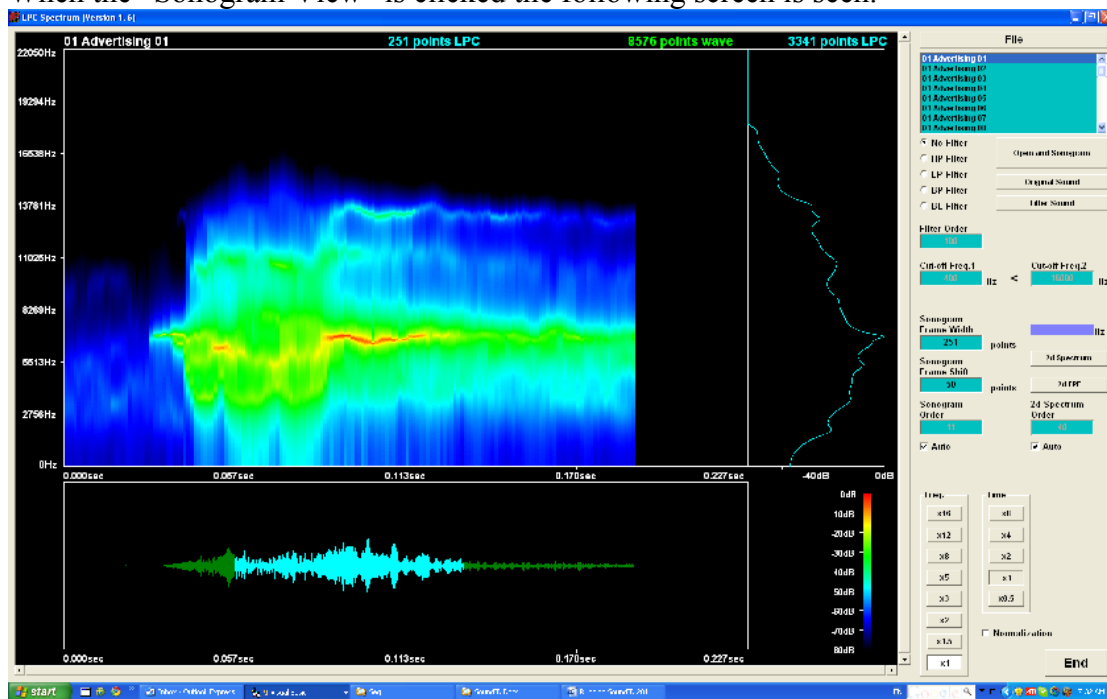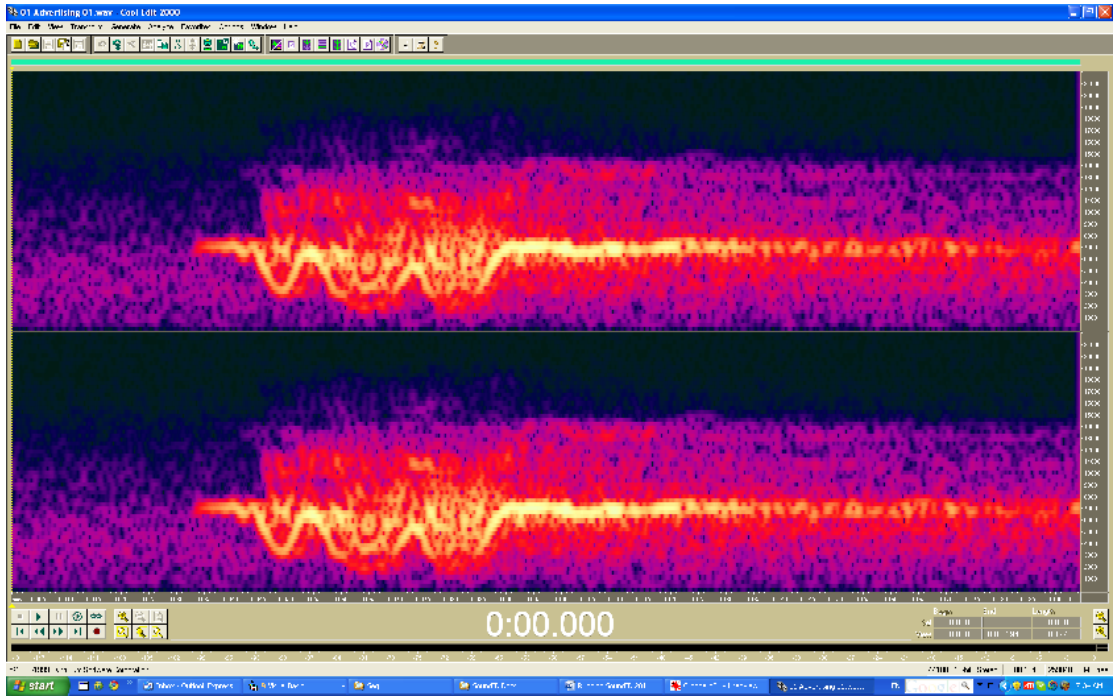
Chapter 5

Sonogram View

The whole of this software is based on the LPC transform and this transform can be used in the same manner as the FFT to produce a sonogram. This module produces a sonogram view based on the LPC. Notice that it is slow and its speed will be addressed in subsequent releases.

It may be better to use the Cluster Analysis module, which has the same capabilities but is faster until the Sonogram view is fixed.

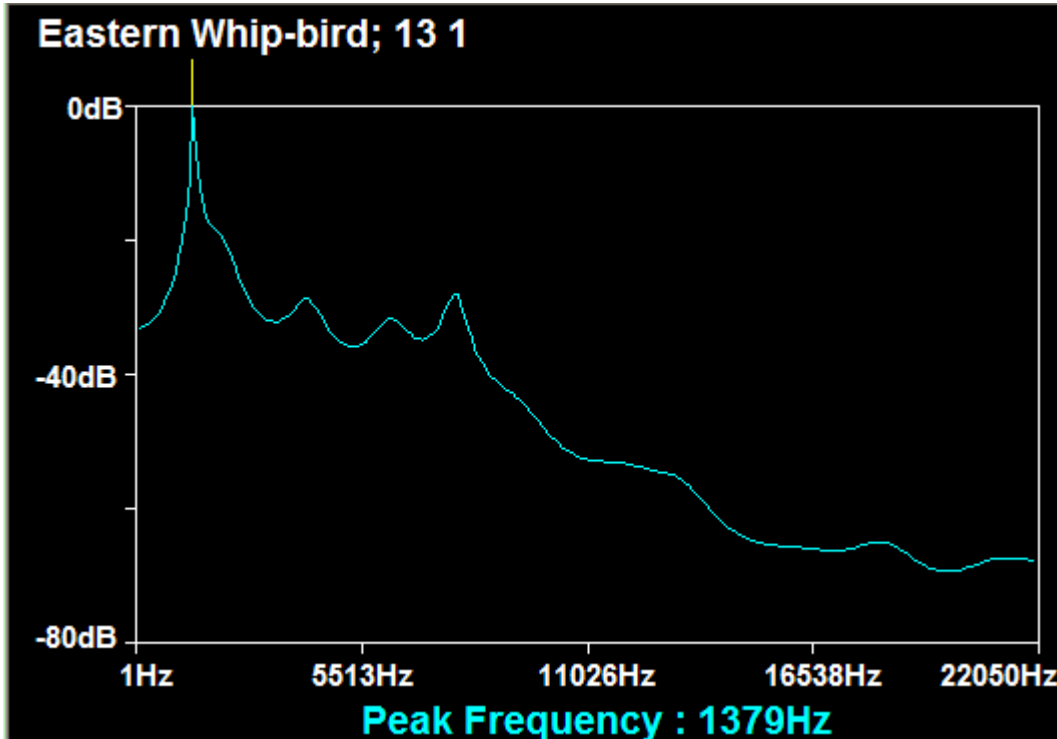When the "Sonogram View" is clicked the following screen is seen.

The advantage of this view over the FFT can be seen by looking at the FFT version below generated by Cool Edit.

Chapter 6

Windowing to Get High Accuracy

To get the full potential out of SoundID you need to understand our windowing concept. Consider the LPC image below.

Now consider the proposition (which is only approximately true) that SoundID will generate an image like this from any file that you are sampling and base its match on how similar the sample image is to this image. Now look more closely at the image (which is of an Eastern Whip-bird) and you will see that it is mostly noise above about 10 kHz. So when it is looking at any other image, part of what it will be looking for is essentially noise above 10 kHz. So we don't really want that. We can Window out everything above 10 kHz without losing much information about our call. To do this we start in the Registration Window and un-tick the auto-frequency and then change the default frequencies.

to
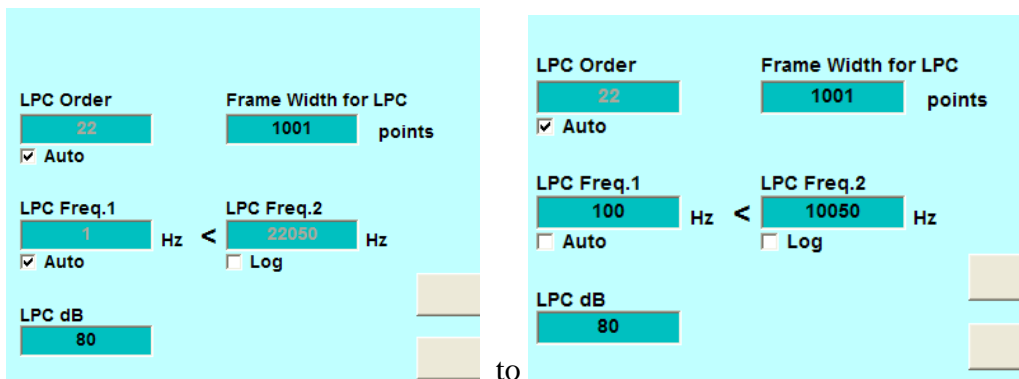
Now the image looks like this. Notice we have also clipped the first 100 Hz as this is basically just adding a wind filter.
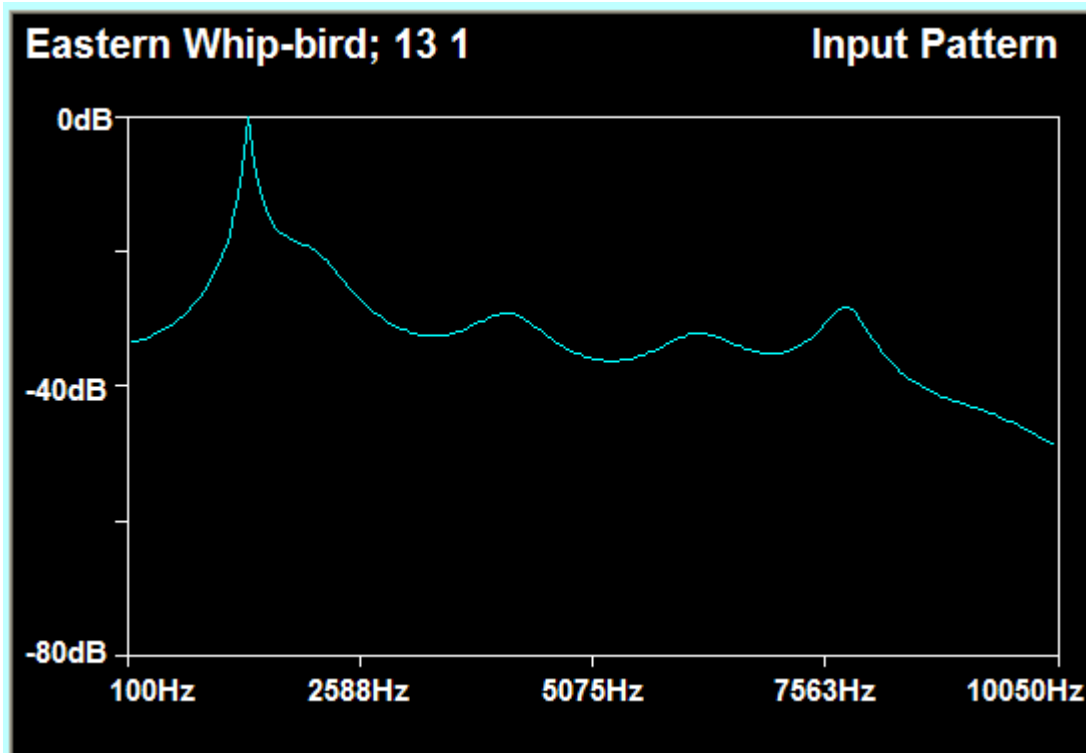
No we see that there is a huge area of the diagram below the noise floor of about 40 dB. So we change the LPC dB to 40 dB, and while we are at it lets cut the frequency off at 5 kHz.
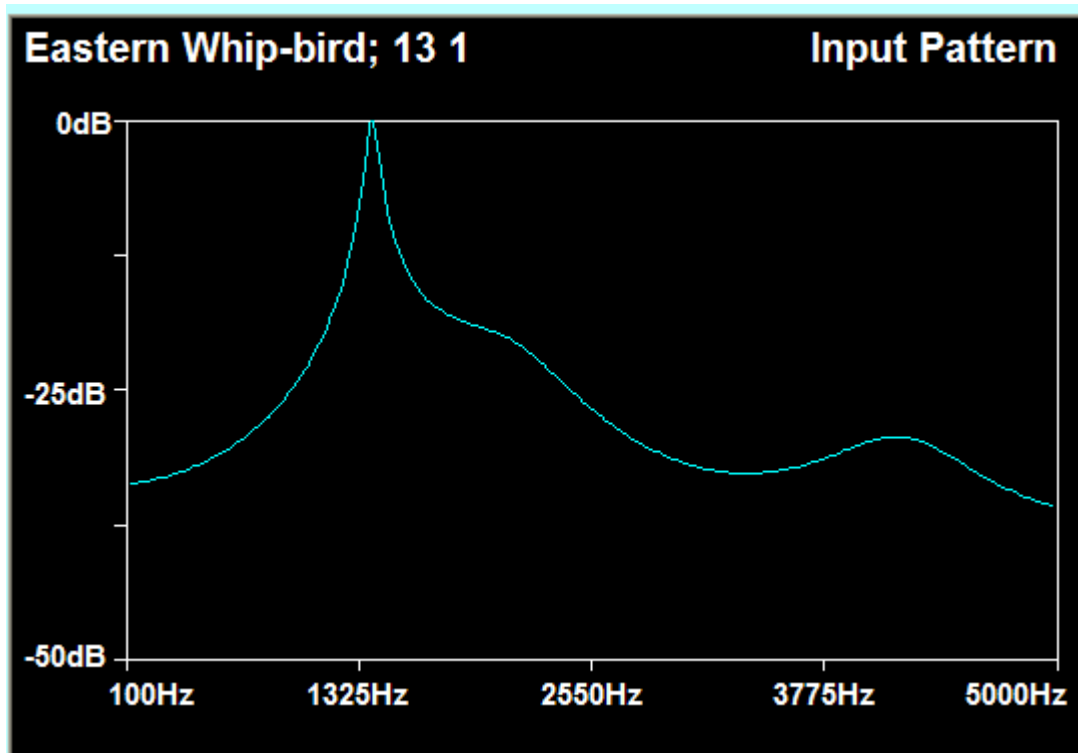
Now the signal is mostly filling the screen and maybe there is still a bit too much noise. However we notice also that the frequency mix looks a bit nondescript and maybe a higher LPC order would improve things a bit. Take note however that a higher LPC order, while giving a better frequency description will come at a relatively high CPU time cost. But let's do it anyway!

So now the settings are.
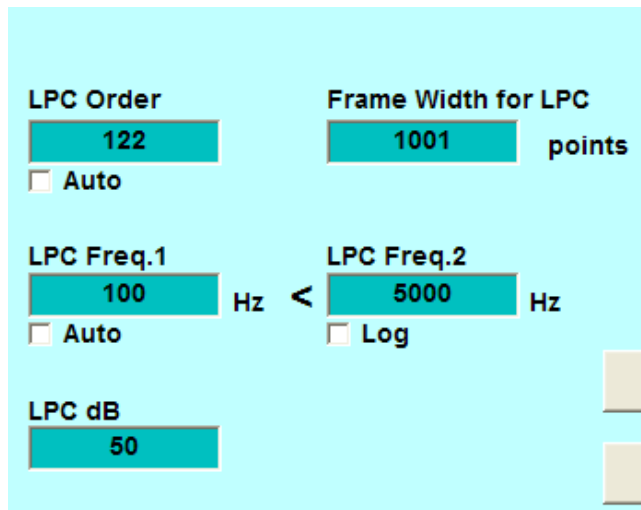
And our image is.
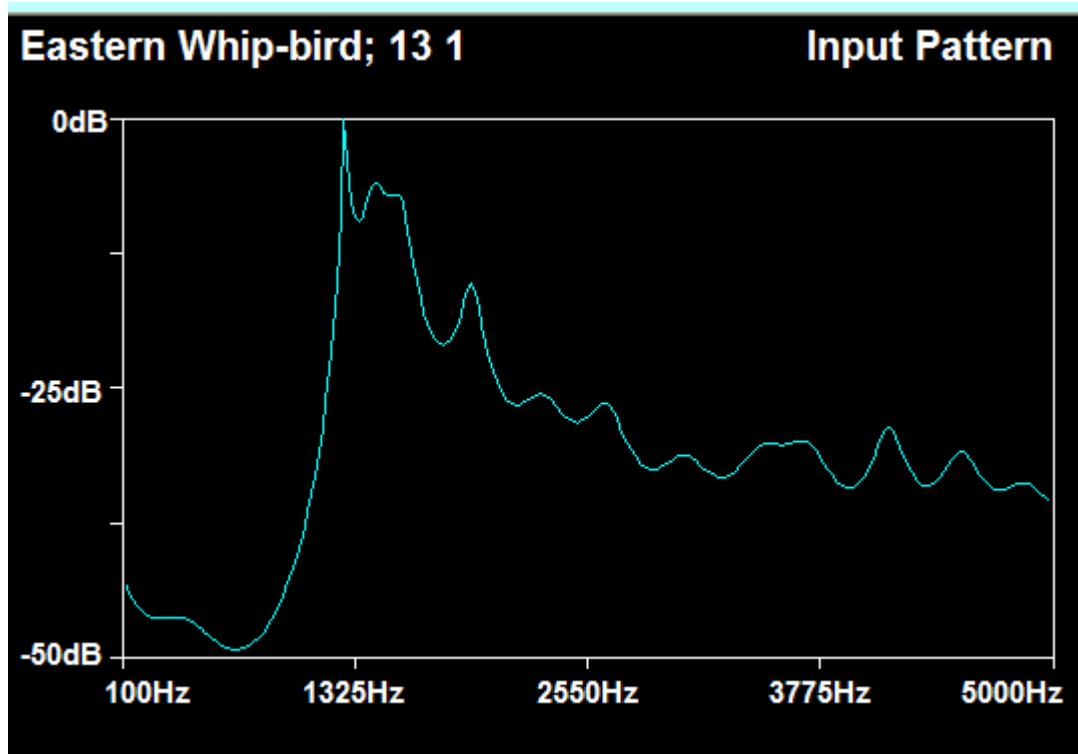
This is a much more distinctive image to match, and when we match with it will mostly be matching signal with signal (instead of signal + noise + a lot of empty

image). In short we have zoomed in on the important part of the signal and that is what we will match with.

Now, we do not have to do this on an image by image basis. And in fact we can't. Once you register and image with the parameters set as above they will grey out and you can't change them anymore. So the whole of the registered files need to be set to one consistent Window. With a bit of practice this process will prove to be simple enough.

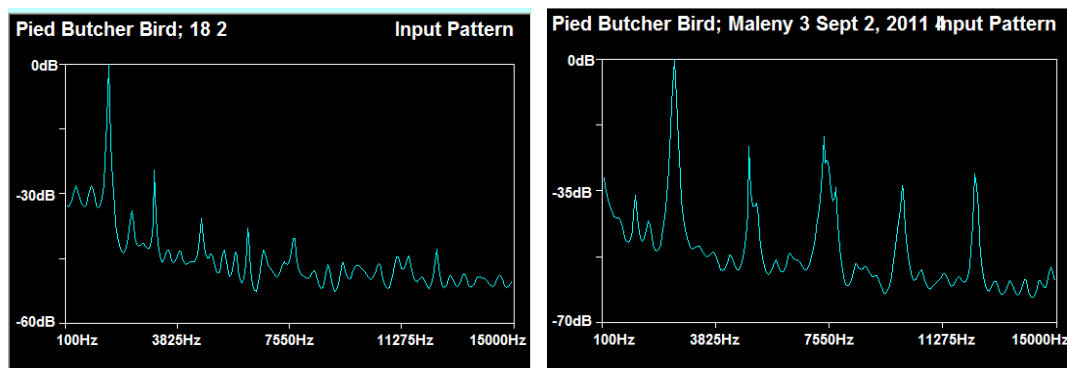But what if the caller has two or more very different kinds of calls, like for example the Butcherbird below?

Here we have one call where all the energy of consequence is below 5 kHz and above -60 dB and another where the bandwidth is at least 10 kHz and the floor is more like -70 dB. Easy. Make two templates one for the first type of call and another for the second. In both Recognition and the Batch Mode any number of templates can be run together.

The Number of Points……And Most Important for Bats

One of the parameters that can be set above is the Frame Width. It is the only parameter in this group that can be set on a call by call basis (all of the others are fixed for the template). However, each time you add a template with a different frame width within the template you will invoke and extra pass of the software to compute it. So two different frame widths will take nearly twice as long as one to execute, and so on. Use them sparingly.   Now here we have to distinguish between bats and birds (and possibly other sounds.

Birds

Birds mostly string together a collection of syllables and these syllables are of short duration (typically 0.02-0.1 seconds). The best way to recognise birds is by the syllables. This is because if they change the song they often change the order of the syllables rather than the syllables themselves. Also because the syllables are short they do not put too much of a load on the CPU time. In a nut-shell for birds you the shorter the frame the better. However be careful of getting too short. As you reduce the size of the frame you will begin to notice a change in the LPC spectrum. If it gets too small you will have only part of a syllable and this which change thing

dramatically. For most birds, a frame width of 801 to 1501 points (at 44.1 kHz sampling) seems to work, but be prepared to experiment outside those ranges.

Bats

Now bats don't seem to deal in syllables, but rather they deal in pulses. For bats you need to sacrifice some CPU time to accommodate the temporal range of the pulses. Some short pulses are incorporated in the long ones (more or less) and some are not. To get good recognition for most bats we probably need to have about 4 frame widths and make sure that the pulses that are used for reference fit wholly within the frame. So this is OK.
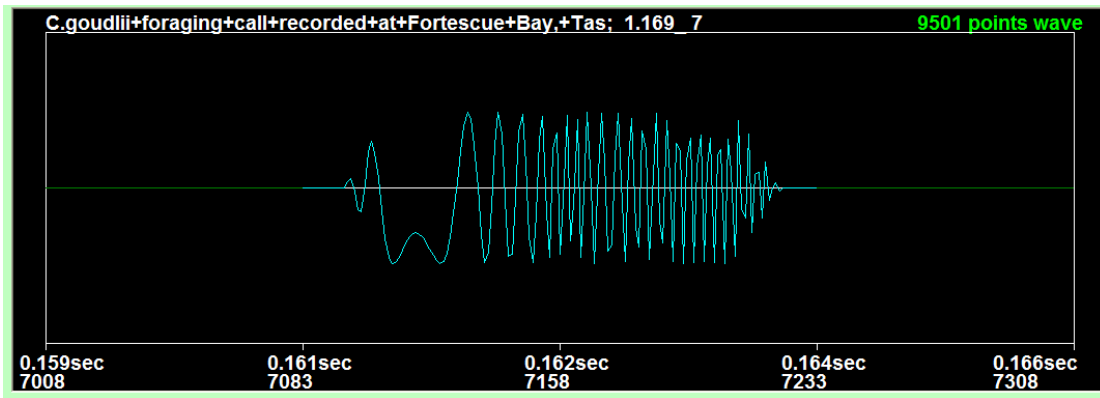
Because the Frame (the blue bit) fully contains the pulse. But this is not because the frame is too short. The green bits are outside the frame and will not be processed.

Chapter 7
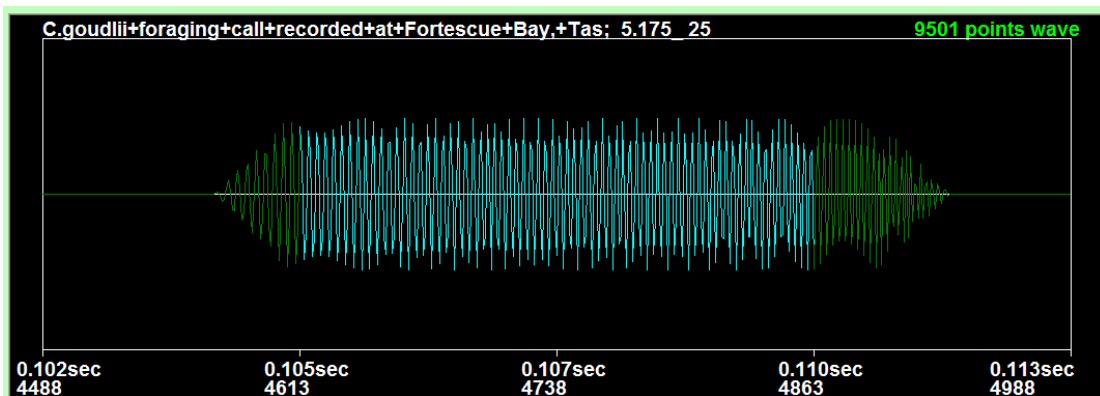
Working with Noisy Files

If you are searching only for nice clean copies of files that are like your references (and there may be many reasons why you are not interested in noisy examples) then use the default settings that include the full frequency range of the recording (i.e. a 44100 Hz recording will default to 22,050 Hz bandwidth as seen below).

SoundID can recognise sounds in noisy recordings at least as accurately as a human expert. However you need to understand how noisy recordings differ from studio recordings and how to select the settings to work with noisy signals (note the settings used for noisy signals will also work for clean signals but not vice-versa).

Technically, SoundID can recognise signals at a S/N ratio of 3 dB or less, with good accuracy. At this level it is difficult for a human expert to identify the signals.

It is not a good idea to use noisy references to find noisy signals. Any signals that are recorded in noisy areas will contain local noises that will not relate to the target signal. In fact it is possible that noisy signals will match other noisy signals on the basis on the "same" noise alone. With poor settings for example it is possible to match dissimilar noises only on the basis that they both have similar wind noise.
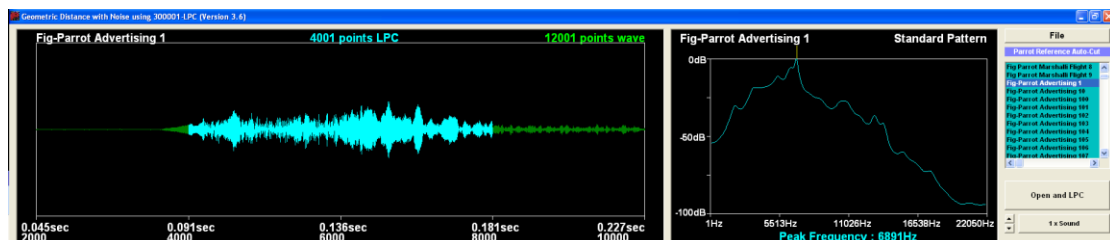
If the recordings you are using have come from the field they will not be nice and clean like this one. More likely they will look like this below. This is the same signal with a S/N of 10 dB.
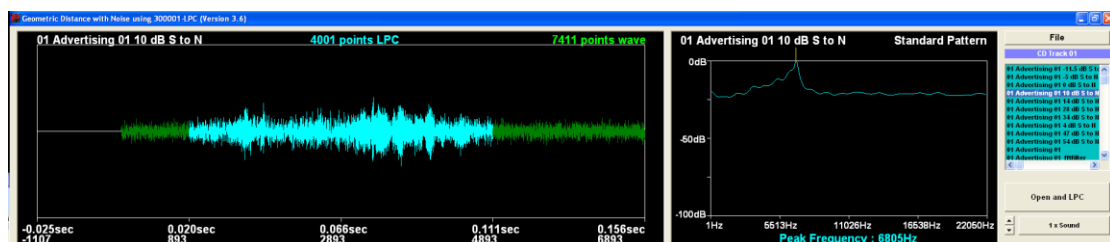
Notice, that not far from the signal peak, the pattern is largely feature-less and flat. Also this signal no longer looks much like the "clean" version and may be missed in a

search if the clean version is the reference. In fact the GD between these to signals is 10.67!

What we need to do is to limit the search to the area that still has some recognisable features. If we look at the noisy signal, it is distinguished from the noise in the bandwidth of about 3,000 Hz to 8,000 Hz.

Also note that the noisy call has a depth of only about 25 dB, compared to 100 dB for the clean signal.

So if we set the bandwidth and depth accordingly as below.
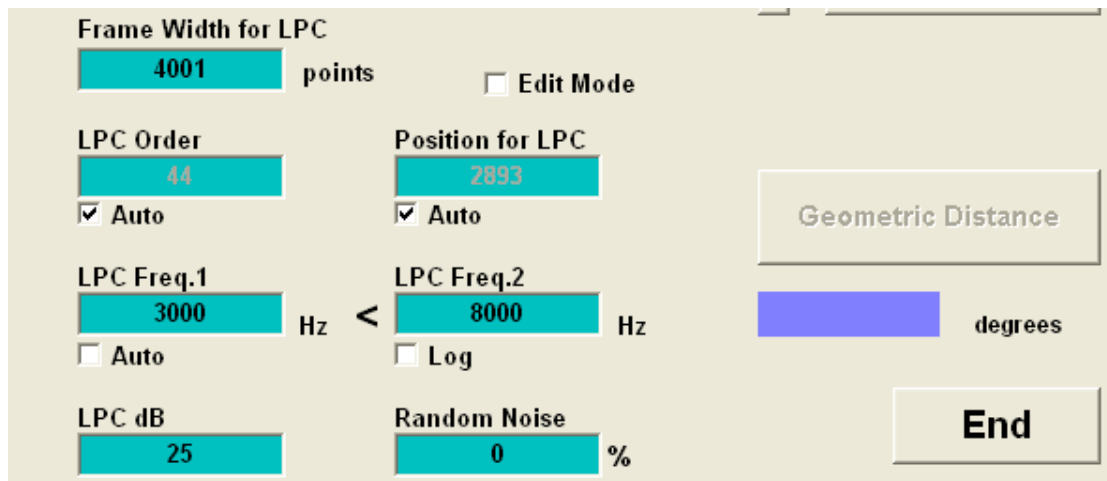
This results in the clean signal looking as shown below (notice the cropping of the signal at the lower end of the frequency spectrum. Matching could be improved by narrowing the bandwidth even more from.
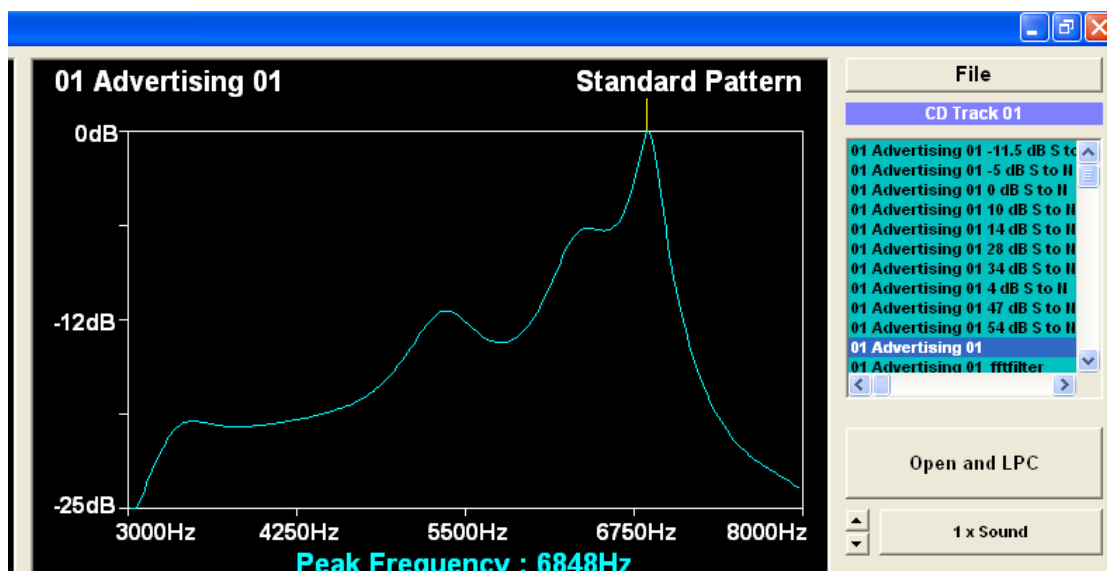
If the noisy signal is cropped accordingly then the similarity now becomes clear by looking at the two images. Below we have the noisy signal and it is image. The GD is 5.09. Now when working with noisy signals and cropped references you need to understand one very important thing, The matching GD range will be more like 6-13 dB than the 3 dB default for clean signals using the full default recording bandwidth.. As a rule when working with noisy signals and cropping the signal as shown above set the "Threshold of GD" to about 15 dB for a match.
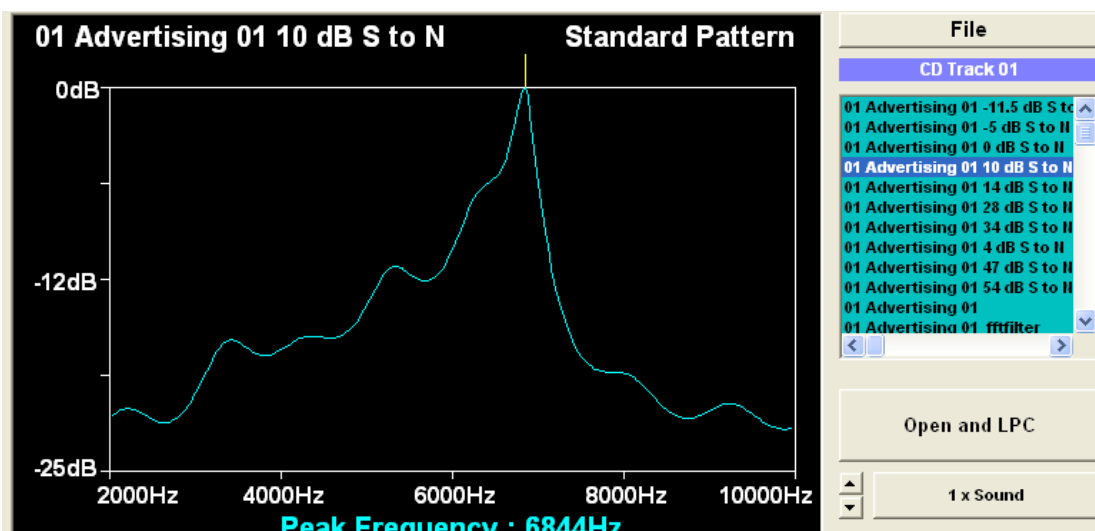

Other Settings

When you zoom in on a narrow frequency band as above it is a good idea to increase the LPC order to about 150. Also if the files are noisy and you are not interested in the very low frequencies, add the filter. A word about the filter; it is SLOW. It works fine but it is not optimised for speed. This is on our to-do list, but for now be aware that putting the filter on, will significantly slow the processing. We hope to have it fixed soon.

Example Study of the Dawn Chorus

The dawn chorus is the time, just before day-break and some time after it, that all the birds come out to sing. It has been suggested that this would make an ideal test for Sound ID. So on the first weekend of October 2011we set out to see how the SoundID software would handle this task.

The first thing to do was to get some recordings, so an Olympus LS-11 recorder was left out from 5.30 to about 8.00 a.m. A total of 160 minutes of recordings were made. From some previous recordings in the same area and also by picking out the best of the weekend's recordings, a reference file was made using the Auto-cut SoundID module, with the minimum energy left at the default -50 dB (this prevents low signal

levels, which are likely to be noisy from becoming references). This resulted in about 650 reference files.

The species definitively identified by us were

Crow
Currawong
Eastern Whip Bird
Grey Shrike-Thrush
Guinea Fowl
Magpie
Noisy Miner
Pale Headed Rosella
Pied Butcher Bird
Plover
Rainbow Lorikeet

Plus the **Sedge Frog**

The reference files were then opened in Windows and played (by double clicking) in turn on Windows media player. Any that were not clear or contained clashed calls (more than one call at a time) were deleted. This left 575 useful calls. Now, while this might sound a lot, for some species it was adequate being > 50 calls, but for others it was only a few calls and thus inadequate.

There were some other species recorded, but since we could not identify them they were not used in the search.

The testing was done in the SoundID batch mode and these settings were used. Notice that the "LPC dB" is set here at 55 dB which is much higher than the 25 dB used earlier in this chapter. This is because the noise level in these recording was only moderate.
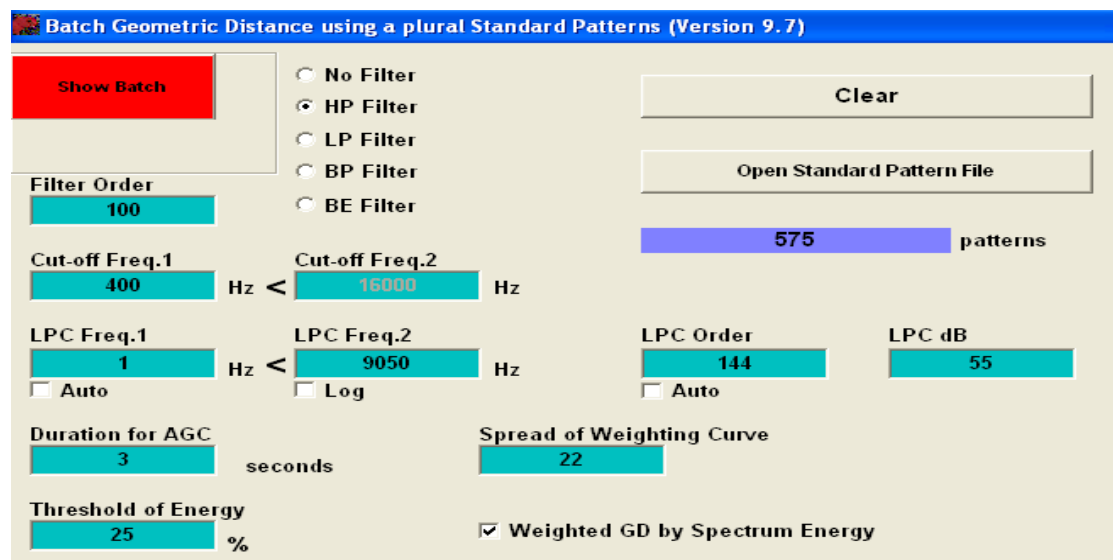
Next the Batch Window was used to try various settings of the GD and check the sounds (by double clicking on them as seen below.

The results of one test are as shown below

GD=3   66 calls identified False positive =0
GD=3.5 70 calls identified False positive =0
GD=4   73 calls identified False positive =0
GD=4.5 78 calls identified False positive =1
GD=5   88 calls identified False positive =6
GD=6   148 calls identified False positive =23

And the detailed results are

File	Time	Freq	Species	Points	GD
Maleny Dawn Chorus	40.709	4031	Rainbow Lorikeet; 9	2001	2.66
Maleny Dawn Chorus	246.423	3324	Mickey Bird; 4 15	2001	2.44
Maleny Dawn Chorus	256.669	3324	Mickey Bird; 4 33	2001	2.19
Maleny Dawn Chorus	259.932	3394	Mickey Bird; 4 40	2001	2.9
Maleny Dawn Chorus	260.445	3713	Mickey Bird; 4 41	2001	2.07
Maleny Dawn Chorus	294.624	2829	Pale-headed Rosella;	2001	1.97
Maleny Dawn Chorus	295.846	2299	Pale-headed Rosella;	2001	2.45
Maleny Dawn Chorus	296.064	2829	Pale-headed Rosella;	2001	1.89
Maleny Dawn Chorus	297.008	2263	Pale-headed Rosella;	2001	2.25
Maleny Dawn Chorus	297.996	2758	Pale-headed Rosella;	2001	2.34
Maleny Dawn Chorus	298.282	2687	Pale-headed Rosella;	2001	2.37
Maleny Dawn Chorus	299.098	2263	Pale-headed Rosella;	2001	2.03

Maleny Dawn Chorus   299.321 1380 Pale-headed Rosella; 2001 2.07
Maleny Dawn Chorus   299.471 2405 Pale-headed Rosella; 2001 2.81
Maleny Dawn Chorus   483.502 6399 Mickey Bird; 7 18 2001 2.41
Maleny Dawn Chorus   484.614 6434 Mickey Bird; 7 20 2001 2.43
Maleny Dawn Chorus   486.236 6505 Mickey Bird; 7 23 2001 2.22
Maleny Dawn Chorus   488.445 6187 Mickey Bird; 7 28 2001 2.08
Maleny Dawn Chorus   489.192 3359 Mickey Bird; 7 30 2001 2.25
Maleny Dawn Chorus   625.715 5409 Mickey Bird; 8 9 2001 2.21
Maleny Dawn Chorus   626.074 3253 Mickey Bird; 8 10 2001 3.08
Maleny Dawn Chorus   626.805 6116 Mickey Bird; 8 11 2001 2.12
Maleny Dawn Chorus   627.941 5939 Mickey Bird; 8 13 2001 2.39
Maleny Dawn Chorus   628.474 6505 Mickey Bird; 8 14 2001 2.4
Maleny Dawn Chorus   632.478 5056 Mickey Bird; 8 21 2001 2.67
Maleny Dawn Chorus   633.775 3041 Mickey Bird; 8 23 2001 2.33
Maleny Dawn Chorus   653.892 1945 Currawong; 1 1 2001 2.15
Maleny Dawn Chorus   690.139 2970 Mickey Bird; 9 2 2001 2.15
Maleny Dawn Chorus   690.394 2970 Mickey Bird; 9 3 2001 2.27
Maleny Dawn Chorus   799.159 1415 Grey Shrike-Thrush; 2001 2.54
Maleny Dawn Chorus   890.338 3677 Rainbow Lorikeet; 2 2001 2.57
Maleny Dawn Chorus   890.761 4490 Rainbow Lorikeet; 2 2001 2.28
Maleny Dawn Chorus   Maleny Dawn Chorus   1326.540 3889 Sedge Frog 3; 2001 2.27
Maleny Dawn Chorus   1327.372 3889 Sedge Frog 2; 2001 3.95
Maleny Dawn Chorus   1377.053 1768 Crow; 1 2001 1.86
Maleny Dawn Chorus   1377.488 1203 Crow; 2 2001 2.23
Maleny Dawn Chorus   1378.097 1344 Crow; 3 2001 2.59
Maleny Dawn Chorus   1378.509 1203 Crow; 4 2001 2.56
Maleny Dawn Chorus   1583.504 3854 Rainbow Lorikeet; 2 2001 2.3
Maleny Dawn Chorus   1620.583 1380 Magpie; 1 1 2001 2.28
Maleny Dawn Chorus   1621.007 1238 Magpie; 1 2 2001 2.96
Maleny Dawn Chorus   1621.979 1167 Magpie; 3 2001 2.43
Maleny Dawn Chorus   1621.990 1167 Magpie; 3 2001 2.88
Maleny Dawn Chorus   1667.551 1415 Currawong; 10 1 2001 2.26
Maleny Dawn Chorus   1674.493 1450 Currawong; 11 1 2001 1.73
Maleny Dawn Chorus   1680.056 1450 BB; 12 2 2001 2.34
Maleny Dawn Chorus   1683.977 1698 Currawong; 12 1 2001 1.84

Maleny Dawn Chorus   1683.977 1698 Currawong; 12 1 2001 1.84
Maleny Dawn Chorus   1689.100 1380 Currawong; 7 1 2001 1.95
Maleny Dawn Chorus   1695.214 1380 Currawong; 14 1 2001 2.06
Maleny Dawn Chorus   1700.140 4278 Rainbow Lorikeet; 2 2001 2
.Maleny Dawn Chorus   1710.942 1592 Currawong; 16 1 2001 1.93
Maleny Dawn Chorus   1729.266 1486 Currawong; 17 1 2001 1.43
Maleny Dawn Chorus   1735.055 1698 Currawong; 18 1 2001 1.87
Maleny Dawn Chorus   1775.345 1592 Grey Shrike-Thrush; 2001 3.13
Maleny Dawn Chorus   1776.048 1415 BB; 16 2 2001 2.4
Maleny Dawn Chorus   1879.517 4278 Rainbow Lorikeet; 2 2001 1.47
Maleny Dawn Chorus   1886.468 4172 Rainbow Lorikeet; 2 2001 2.77
Maleny Dawn Chorus   1888.038 4349 Rainbow Lorikeet; 2 2001 2.43
Maleny Dawn Chorus   2042.364 3713 Rainbow Lorikeet; 2 2001 2.19
Maleny Dawn Chorus   2103.667 4419 Rainbow Lorikeet; 3 2001 2.09
Maleny Dawn Chorus   2104.941 4349 Rainbow Lorikeet; 3 2001 1.24
Maleny Dawn Chorus   2113.992 4313 Rainbow Lorikeet; 3 2001 2.19
Maleny Dawn Chorus   2114.153 4172 Rainbow Lorikeet; 3 2001 2.65
Maleny Dawn Chorus   2115.043 4419 Rainbow Lorikeet; 3 2001 2.57
Maleny Dawn Chorus   2115.501 4313 Rainbow Lorikeet; 3 2001 2.28

Maleny Dawn Chorus   929.338 4137 Rainbow Lorikeet;l 2 2001 3.25
Maleny Dawn Chorus   1064.209 1415 Pied Butcher Bird; M 2001 4.04
Maleny Dawn Chorus   1277.309 3819 Sedge Frog 1; 2001 4.42
Maleny Dawn Chorus   1287.720 3571 Sedge Frog 1; 2001 4.37

Maleny Dawn Chorus   1292.451 3571 Sedge Frog 1; 2001 3.91
Maleny Dawn Chorus   1306.099 3536 Sedge Frog 1; 2001 3.88
Maleny Dawn Chorus   1307.079 3182 Sedge Frog 1; 2001 4.1
Maleny Dawn Chorus   1317.397 3925 Sedge Frog 2; 2001 4.39
Maleny Dawn Chorus   1325.858 3571 Sedge Frog 1; 2001 2.22
Maleny Dawn Chorus   1326.176 3677 Sedge Frog 2; 2001 2.22

It was later realised that in this test from which the above list was derived the Whip Bird reference files had been let out and this being one of the more common birds increased the count by about 40 recognitions in the same time frame on a subsequent run.

Chapter 8

Multiple References and Run-Time

The Recognition and Batch Recognition modules can both use multiple templates. Multiple templates significantly increase the accuracy of the system, but they do slow in down roughly in proportion to the number of *unique* templates. So templates that have identical settings (F1, F2, LPC depth etc) can be run as a single large template and this will run almost twice as fast as two different templates.
The templates should be optimised on a species by species (or sound type by sound type) basis, and then checked in two ways.

Optimise Run Time

Firstly depending on the settings used the processing time for the template can vary over at least an order of magnitude. For example a file with a large number of points and a high LPC order will run particularly slowly. You can use Recognition and a small continuous WAV file to check the run times. In the figure below we have loaded a number of templates into recognition and then ticked the check box on the bottom left "Check References". When checked this will cause the software to report the time (in seconds) that each template was run for.

We see on the results page that the run time for the Rosella was 0.5 seconds whereas the Grey Shrike Thrush and Guinea Fowl took 0.9 seconds. These particular templates have had their settings modified for processing time. The original settings had run-times with some of them taking up to 10 x more than others. A ratio of 2:1 is mostly not a problem.
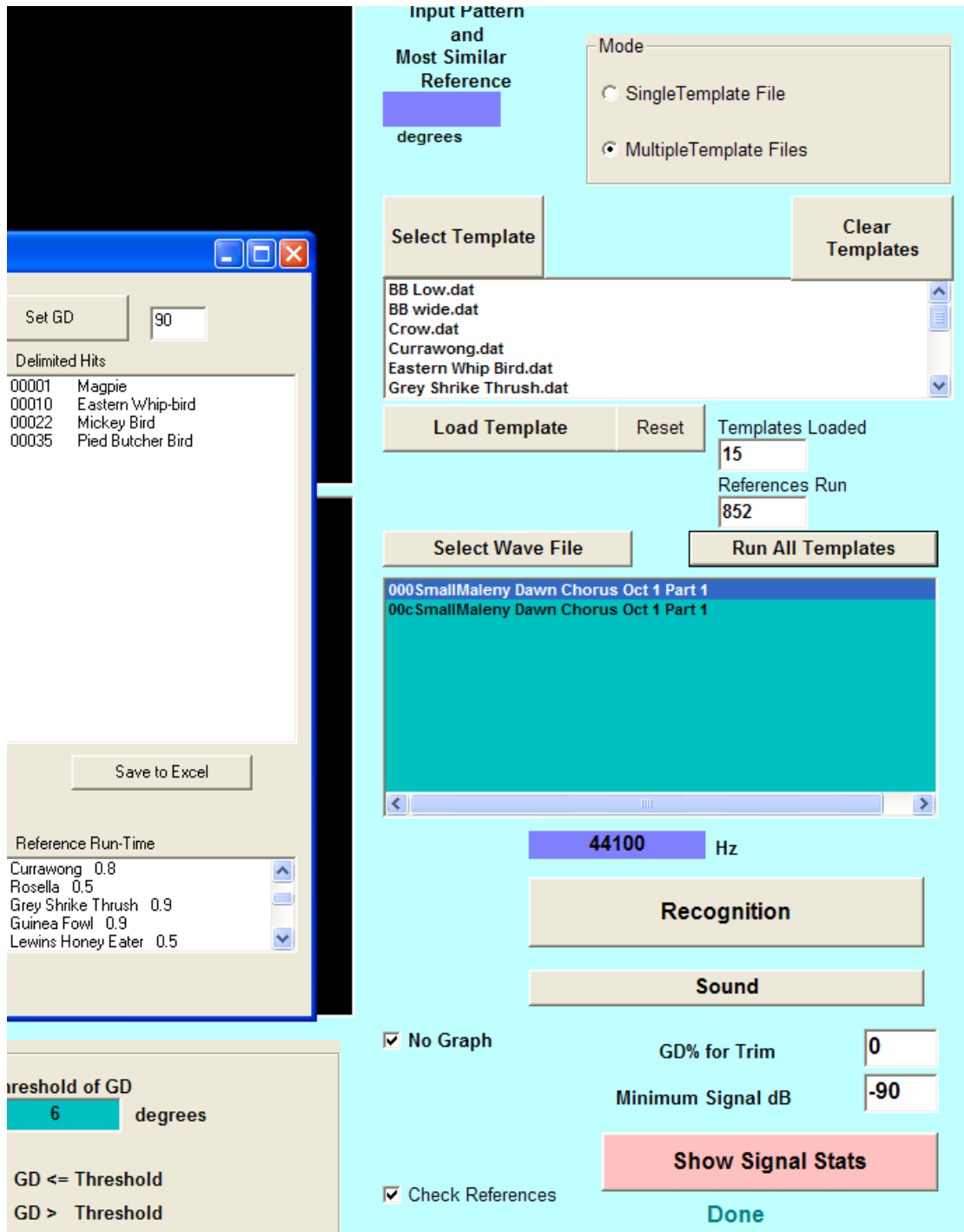
Optimise (minumise) the Number of Unique References

When you have a large number of templates there is a module to enable you to manage them. By loading the "Template File Header" from the "Other Modules" menu (top right) you will get the list of settings for each of the templates. Look for ones that are most similar and consider making them identical so that they will run faster.

Read Reference Header Files

Load Templates

Species	LPC Order	Freq F1	Freq F2	LPC dB	GD	Spread	Points	Is Neg=1
BB Low	44	200	5050	60	10	40	5990053384	0
BB wide	44	200	15050	65	10	22	0	0
Crow	82	200	10050	50	10	22	3703301005	0
Currawong	82	200	10050	50	10	22	7138081236	0
Eastern Whip Bird	44	400	14000	60	10	22	1001	0
Grey Shrike Thrush	82	200	8050	50	10	22	7138081236	0
Guinea Fowl	76	1000	7550	45	10	22	3211010244	0
Lewins Honey Eater	44	1	7050	60	10	40	1500646549	0
Magpie	54	200	7050	60	10	40	3349734687	0
Mickey Bird	22	700	10050	50	10	22	5718795527	0
OldSedge	42	200	10050	50	3	40	1001	0
Plover	22	1	7050	40	10	22	3767814725	0
Rainbow Lorikeet	44	1	12050	60	10	40	7489540551	0
Rosella	44	400	14000	60	10	22	1001	0
Sedge Frog	60	200	10000	60	3	22	1001	0

Chapter 9

Segmentation and AGC Settings

Segmentation

In the Segmentation module you will find the "Optimisation" but which can be used to get the detection right. *Detection* comes before recognition. So the software first scans the file to find a signal of sufficient magnitude to decide whether or not to stop and look at it. *If you have the AGC settings wrong you may not get sufficient detections and so the software is not presented with anything to recognise.*

The segmentation module shows the effect of varying the AGC level and the threshold. The waveform that is detected as "signal" for any setting of threshold and AGC is seen in *blue*. The signal in *green* is ignored as noise. It can also be used to optimise the AGC settings.

If you select Segmentation from the menu and load a file using the "File" button (top right) to import a WAV file, and this use the fast forward buttons (bottom left) you will see the way a long file is scanned for various settings. The part of the file in blue is the part that will be used to try and find the match. An example is seen below.
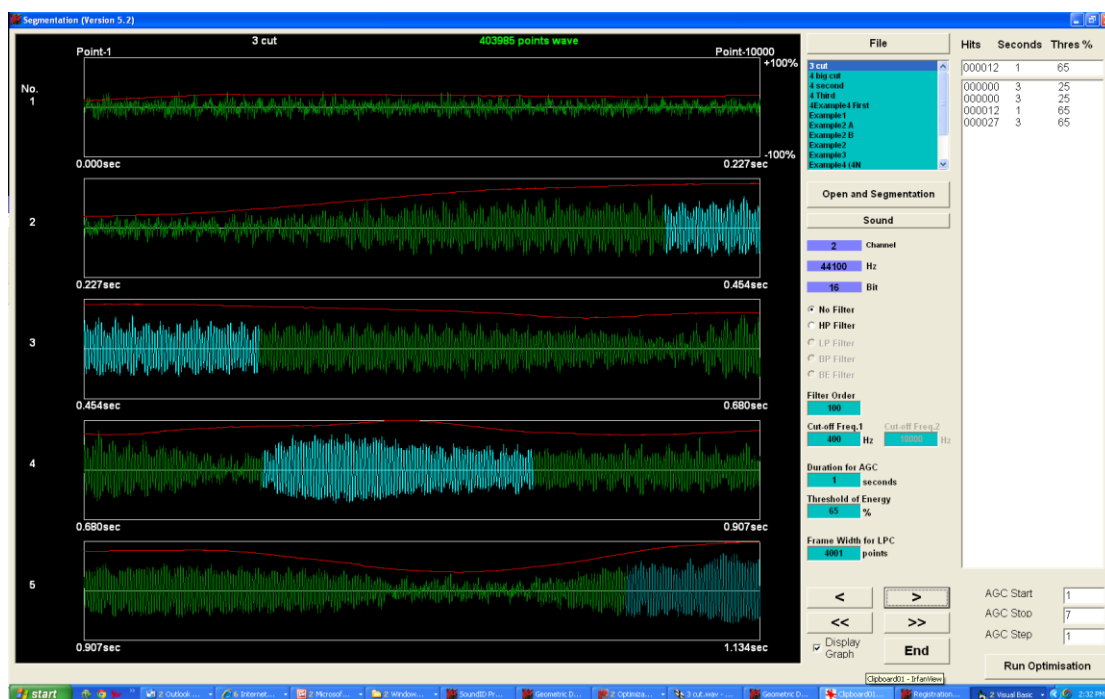
Segmentation (scanning of a long File)

To find the settings that best suit your particular recordings, first select a typical recording segment of a few minutes and load it in using the "File" button on the top right. Note that if you use a long file (instead of just a part of that file), because this program detects the signal for hundreds of different AGC settings, it can take a very long time to complete.

It will load in all WAV files from the selected directory. On the bottom right you will see the AGC scan settings for the AGC. You can change the default values, but it is probably a good idea to run them first. The values all must be positive. Note that fractional values like 0.1 can be used for the AGC step size.

When you click the "Run Optimisation" button the software will scan the file for all values of the AGC time (as set) and for all values of Energy Threshold from 2% to 99%. Notice that since a full scan of the selected file takes place, the choice of a long file may mean that the run-time is very long.

The results will appear in the list box in ascending order. In general you will want the settings that give the highest number of hits and these will be at the bottom of the list box. The results are sorted as they are produced so that the top of the list box will usually have a "000000" listing. Above the list box is the value of the last setting calculated. This can be useful particularly if the values are being entered manually.

Note however that it is possible to set the AGC to an ultra-sensitive setting that results in too many detections. Mostly the settings that give the greatest number of detections (at the bottom of the list box) will be ideal, but it does pay to confirm the results by listening.

Chapter 10

Auto-Cut References

The Auto-Cut References module (in the professional edition only) has the Segmentation and AGC Settings as described in the previous chapter, but it can also automatically cut reference files from a long recording, for example a CD of the bird calls you may be interested in. In this example we will automatically cut the calls that are used as examples, with the software. A novel use of this module as discovered by one of our users was to set it to examine mostly quiet recordings to find and cut sounds in the desert. By suitably selecting the "Minimum Neg Energy" setting (about -75 dB seemed to work) the module could be made to scan 20 minute segments of recordings and extract the rare, but important, calls that were detected. This saved our friend 20 minutes of tedious listening per recording segment. The source of the files used in this example is from a CD by David Stewart.

First, using the "File" button at the top right, we load in the long files of interest, as seen below. Typically these might be things like tracks from an audio CD (after converting it to WAV format) of target species of interest. You should have previously looked at these files to determine the correct parameters for cutting them (e.g. how many points and the frequency range).

The "Min Energy Neg dB", which defaults to -40dB sets the level of the signal that will be added to reference file. This stops low energy signals (which will have a lot of embedded noise) ending up as references. As the value is negative, the higher the value the lower the signal levels that will be cut as references. The default value will ordinarily suffice for most situations.

You will notice that the "Display Graph" check-box at the bottom right is unchecked. This turns off the graphical display of the segmentation process (and which if it was left on would slow down the PC). Also you can set the number of buffer points (this is the number of points before and after the target cut. So with the above setting the 4001 point reference will be cut with 4000 additional points of the recording at either end of its peak energy range. This makes it easier for the software to handle and easier for a human to verify the sound acoustically for short files.

To save the segments as files that can be used to make a reference file, click the RED button and it will open the dialog box seen above. At this point add a new folder to save the segments in (in this case we have used "01 Advertising" and we have called the file name the same thing (it need not be the same). Having done this click the save button and all the WAV segments will be saved as seen below.
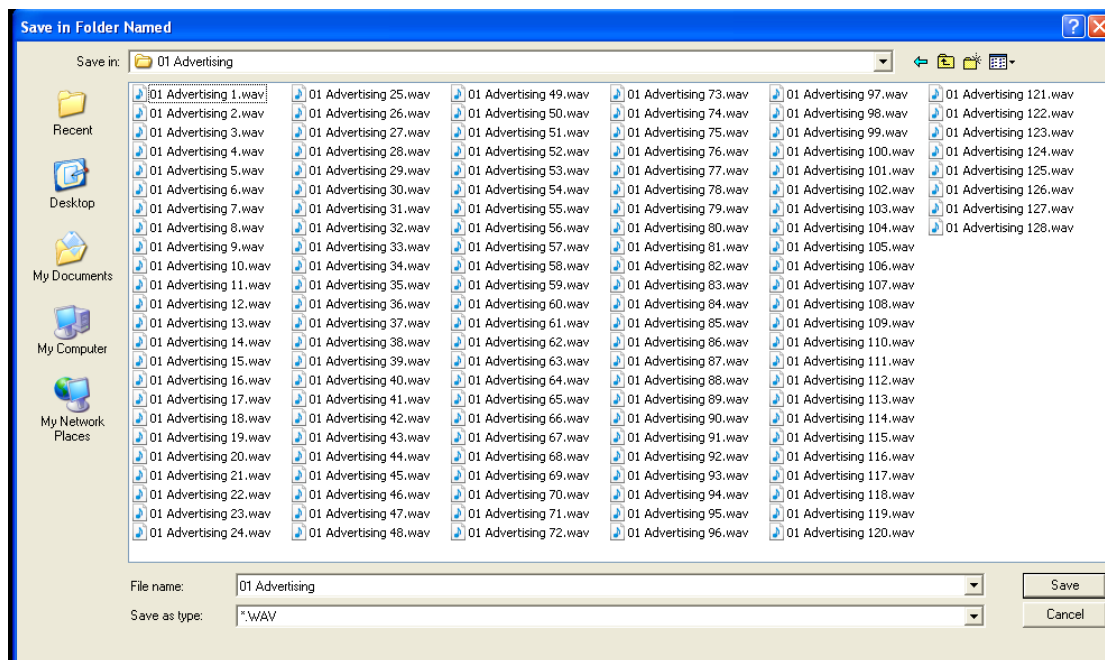
Here we see that the software has cut the long file into 128 segments ready for use as reference files. Note that the buffer defaults to 4000 points. This is the portion of the file that is added to the cuts above the Frame Width (4001 points in this example). So in this instance 4000 additional points from the original WAV are included before the segment and another 4000 points at the end.

To cut these files manually would take probably a whole morning, whereas the Auto-Cut can do this in seconds. This way you can build a very large library in a short time.

Auto-Cut Batch Files

Auto cutting a batch of files is just as easy and the only difference is that the "Cut All Waves" button to the left of the red button is selected and then the software will cut references from all the files in the top right hand list box.

AGC Settings

This screen also allows the optimisation of the AGC settings for the files that you are using as a reference. For details on how to run this, see the previous chapter "Segmentation and Settings" as it runs exactly the same way.

Chapter 11

Optimisation

The optimisation of the values for Geometric Distance and Weighting Vector can be found using this module. Be aware that different sound types can be optimally search for with different GDs and weighting vectors.

This module works by comparing a batch of reference WAV files that should match, with another batch that should not match. The non-matching batch of WAV sounds should preferably be those that are likely to be mistaken for the target or those that are prevalent in the vicinity of the target sounds.

One way to do this is to divide the reference calls into two halves and run that against an unwanted set of calls. In this example we have divided the reference calls call "01Advertising XX" into two batches and compared them to the call of a different parrot "01 Advertising XX"
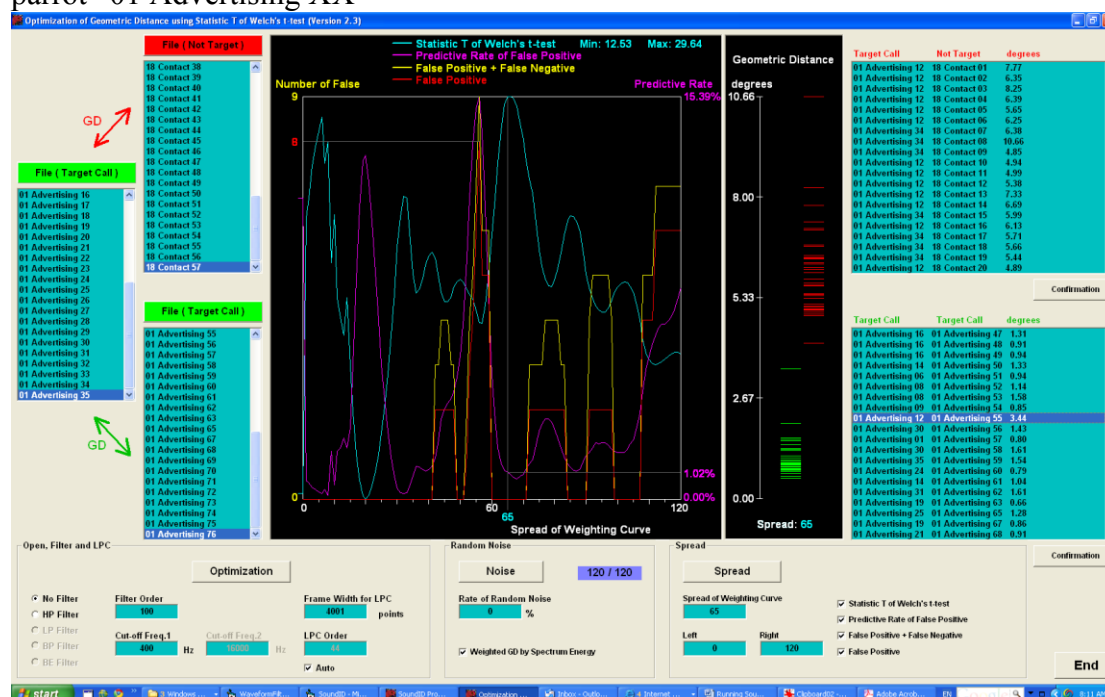
The first half of the "01 Advertising XX" calls are loaded into the first target call box and the second half into the second target list box (both green). Then the "18 Contact XX" batch will be loaded into the non-target box.

Click on the optimisation button and the files are compared. The output on the far right of the display shows the target call values (green) and the non-target values (red) do not overlap (there will be no false positives). This plot does show however that one of the reference calls stands out as having a GD very far from the main batch. A

closer inspection of this from the list box of target calls (bottom right), reveals that the worst match is between call 12 and call 55. Inspection reveals that call 55 has a lot of noise and that it probably ought to be removed from the target call collection.
The module recommends a weighting vector spread of 65.

Chapter 12

Cluster Analysis

So What is Cluster Analysis?

Cluster analysis is a fancy term for sorting things into similar groups. Figure 1 shows a group of farm animals that are sorted into like pens. All that is required to do this, is to follow the simple rule "put animals of the same species in unique pens".

Figure 1. A simple cluster analysis; animals sorted by species.
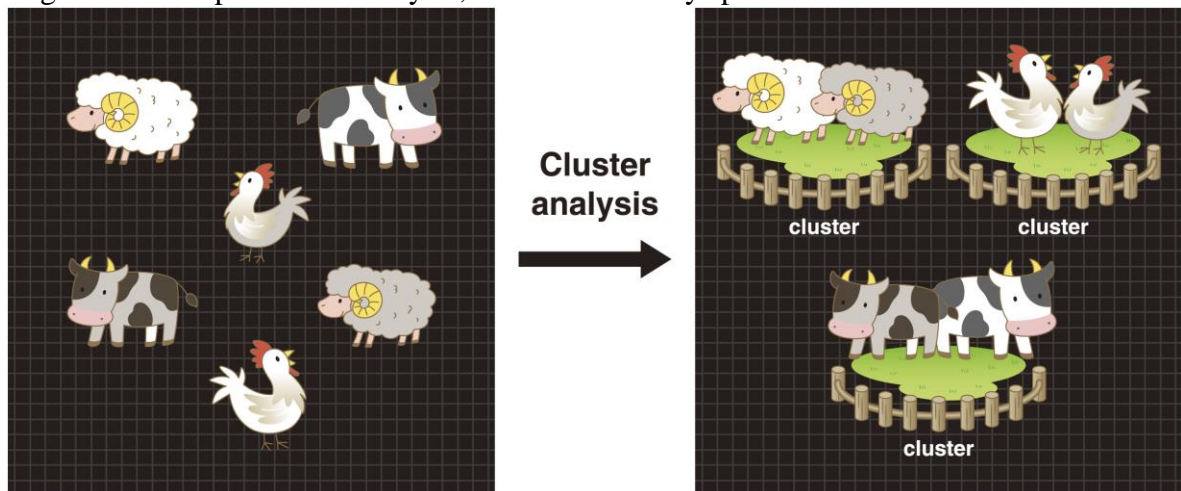
The analysis can be more complex than this. For example, what happens if the number of pens is less than the number of species? In this case we might group the animals into "most closely related species". While that is fair enough we need another rule to decide what constitutes a close relationship.

The main point here is that the clustering can be done for any number of items which results in them being sorted according to a set of rules. In our case, we sort calls into calls that are similar as measured by their Geometric Distance (a mathematical measure of similarity). The Geometric Distance is an angular measurement in degrees (as in 90 degrees is a right angle), and is such that a distance of zero degrees means "no difference" or a perfect match. Sounds that are similar to the human ear are typically less than 3.5 degrees apart. It is worth noting here that the human ear has rather poor discrimination compared to that of a bird and so similarity to the human ear is not necessarily a sufficient measure.

Running the Cluster Analysis

To begin click on the Cluster Analysis and load the "Wave Files" as below

These files may be found in the **Recognition** subfolder where the SoundID program was installed.

1. From the Main Menu, start the *Cluster Analysis* program.
2. Click the upper *File* button and navigate to the **Recognition** folder.
3. Navigate to the **Double Eyed Fig-parrot Advertising** folder.
4. Select any file in that directory and click *Open*.
5. This will load the selected WAV files into both boxes.
6. Double click on the second item in the lower list box, and the screen will display as:
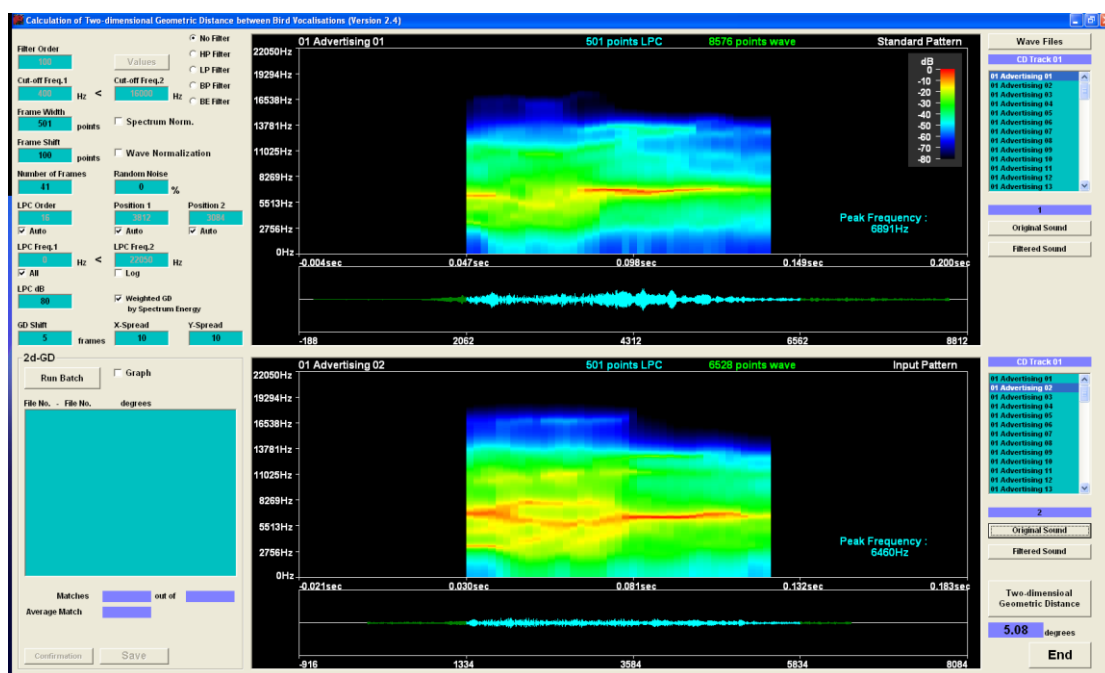
What we have here is a spectrogram view of the calls based on the LPC transform. You will find that this view reveals at lot more detail of the call than any FFT based spectrogram. Furthermore you can zoom in on any part of it by changing the LPC frequency range. Below we, un-ticked the check box below "LPC Freq.1, and have restricted the frequency range from 100 Hz to 12050 Hz.
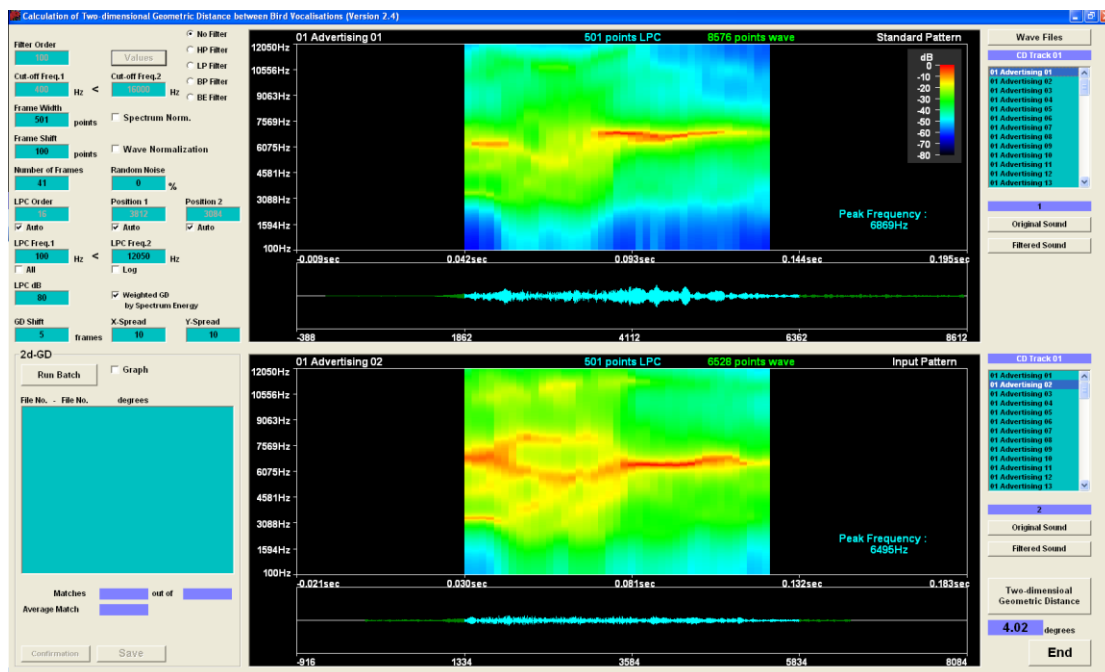
You might have noticed that restricting the frequency range, or zooming, has changed the GD from 5.08 to 4.02. This is because the software actually uses the image as displayed to work out the similarity and there is more similarity revealed in the zoomed version. Part of this might be attributed to the low signal level of **01 Advertising 02,** so that the zooming has removed some of the noise.

The filter is available as before and it needs to be noted that as before the use of filtering will significantly slow the processing *and* it will distort the signals so changing the GD.

Frame Width

The frame width refers to the number of points in the frame that is used to analyse the signal. The frame shift (100 points) and the number of frames (41) determine how much of the signal is considered (the blue area in the WAV view. In points this is

ViewPoints = (n-1)* (Frame Shift) + (Frame Width)

Where n=number of frames

In this example this is (41-1)*100+501=4501 points or about ½ the signal.

LPC Order

The LPC order defaults to a calculated value. The LPC order can be increased to improve the spectrum resolution, but if it is increase too much the LPC transform can become unstable. Additionally the calculation time increases with the LPC order quite significantly. In general, for most sounds, an LPC order of 100 or less will be stable. To see the transform go unstable set the LPC to a high number (if you select anything higher than 499, it will default to 499) and the screen shot below shows the resultant instability.
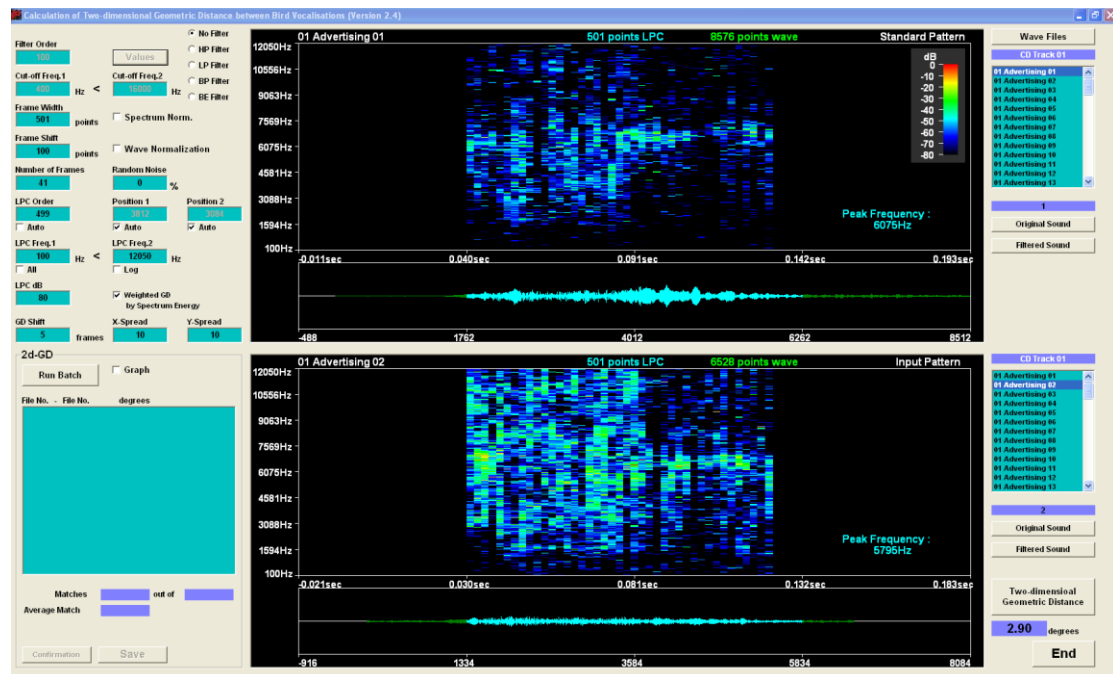
You should experiment with different values and see the effect.

LPC dB

The default value of 80 dB the LPC dB refers to the dynamic range of the recorder with 80 dB, being typical for most good quality recorders operating in the open air environment (is a studio you can do a lot better). There will rarely be a need to change this except perhaps to decrease it for very noisy recordings.

X Spread and Y Spread

The X and Y spreads refers to the weighting vector of the time axis and the frequency axis respectively. The maximum values for these are 20 for the X spread and 100 for the Y spread (if you input higher numbers than these the maximum defaults will be used. Changing these values has the same effect as changing the spread of weighting vectors in the one dimensional version (LPC & GD).

Frame Shift

To look for a match the GD image is shuffled along to find the best match. As seen in the figure below where the GD of the two images is calculated as the bottom image is shuffled along.
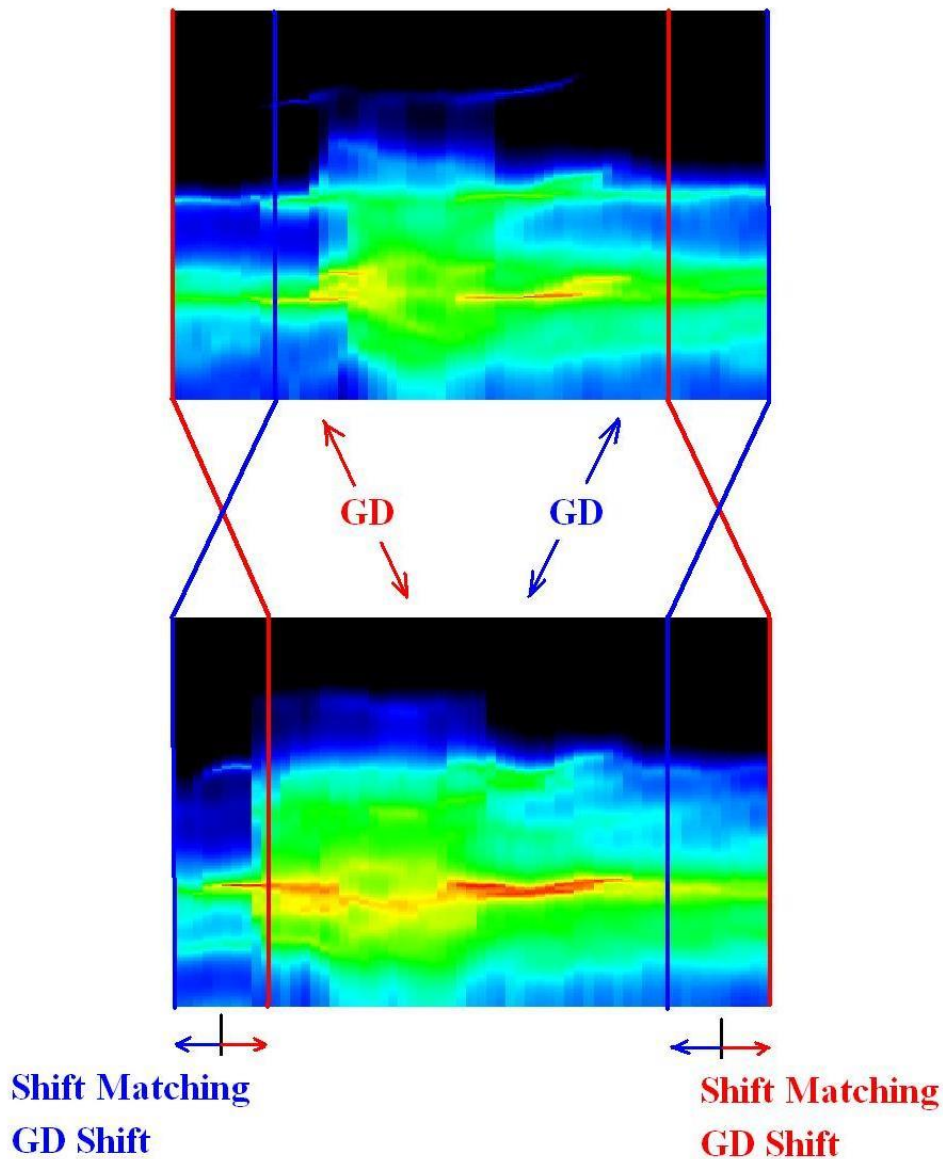
Running in Cluster Mode

Running in cluster mode is a two stage process. First run the software by clicking on "Run Batch". This will cause the software to calculate the GD in the dimensions of time and frequency, between all the pairs of WAV files as shown below.

2d-GD		
Run Batch	☐ Graph	
		Done
File No. - File No.	**degrees**	
1 - 2	1.84	
1 - 3	0.82	
1 - 4	2.11	
1 - 5	1.66	
1 - 6	1.50	
1 - 7	1.33	
1 - 8	1.55	
1 - 9	1.09	
1 - 10	1.27	
1 - 11	1.01	
1 - 12	0.73	
1 - 13	1.09	
1 - 14	2.09	
1 - 15	0.76	

Matches 2701 out of 2701

Average Match 1.79

Confirmation   Save

To begin the cluster analysis you need to save this result, by pressing the ""Save" button.

If you now run the "Cluster Sort" program, and click on "Open" to open the recently saved file you will get this screen shot.

By clicking on the buttons marked "Mean Equation", "Maximum" or "Minimum" you will get the corresponding clusters, initially divided into the 10 default clusters. You can change the number of clusters or you can sort based on the minimum GD between the clusters (which will result in whatever number of clusters is necessary to meet this criterion).

The most meaningful distance measure is the "Mean Equation" and that should be used unless there is a good reason to use the Maximum or Minimum clusters. What follows is an extract from Wikipedia on the clustering technique we use.

Hint

The Frame width when changed will automatically change the Frame Shift which is set to be 1/5 of the frame width. If you have a really large batch of data and want to process it quickly you can set the Frame Shift to a higher number (but not higher than the Frame Length) and it will run significantly faster. However you will get a bit less precision in the LPC transform. Decreasing the Frame Shift (as a proportion of the Frame Width) does increase precision a little but at the cost of a significant increase in processing time.
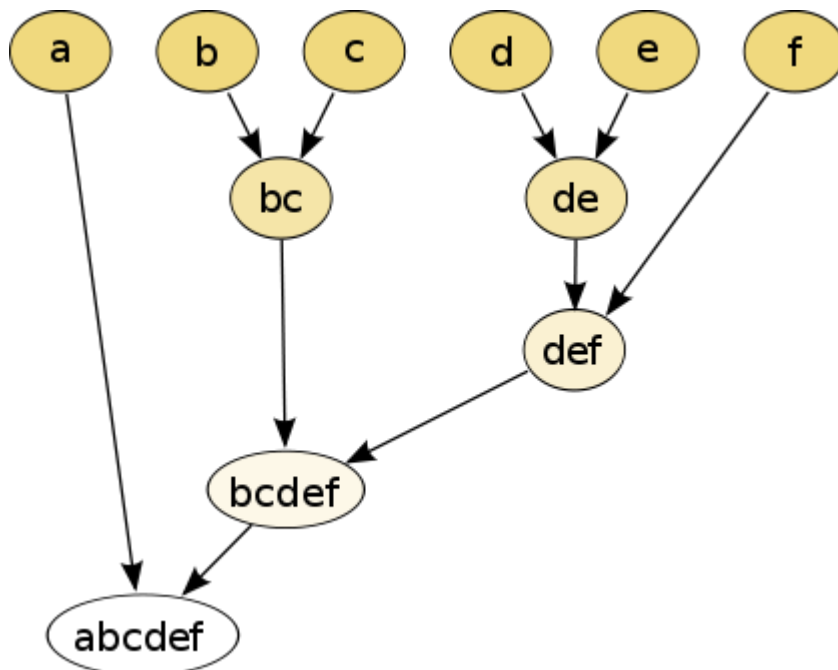
Agglomerative hierarchical clustering

For example, suppose this data is to be clustered, and the Euclidean distance is the distance metric.
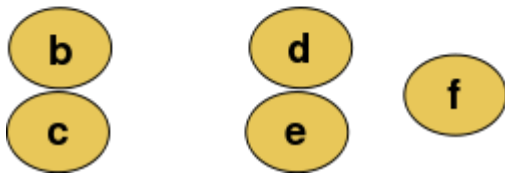
Raw data

The hierarchical clustering dendrogram would be as such:

Traditional representation

This method builds the hierarchy from the individual elements by progressively merging clusters. In our example, we have six elements {a} {b} {c} {d} {e} and {f}. The first step is to determine which elements to merge in a cluster. Usually, we want to take the two closest elements, according to the chosen distance.

Optionally, one can also construct a distance matrix at this stage, where the number in the i-th row j-th column is the distance between the i-th and j-th elements. Then, as clustering progresses, rows and columns are merged as the clusters are merged and the distances updated. This is a common way to implement this type of clustering, and has the benefit of caching distances between clusters. A simple agglomerative clustering algorithm is described in the single-linkage clustering page; it can easily be adapted to different types of linkage (see below).

Suppose we have merged the two closest elements b and c, we now have the following clusters $\{a\}$, $\{b, c\}$, $\{d\}$, $\{e\}$ and $\{f\}$, and want to merge them further. To do that, we need to take the distance between $\{a\}$ and $\{b\,c\}$, and therefore define the distance between two clusters. Usually the distance between two clusters $\mathcal{A}$ and $\mathcal{B}$ is one of the following:

- The maximum distance between elements of each cluster (also called complete linkage clustering):

$$\max\{\, d(x, y) : x \in \mathcal{A}, y \in \mathcal{B}\,\}.$$

- The minimum distance between elements of each cluster (also called single-linkage clustering):

$$\min\{\, d(x, y) : x \in \mathcal{A}, y \in \mathcal{B}\,\}.$$

- The mean distance between elements of each cluster (also called average linkage clustering, used e.g. in UPGMA):

$$\frac{1}{|\mathcal{A}| \cdot |\mathcal{B}|} \sum_{x \in \mathcal{A}} \sum_{y \in \mathcal{B}} d(x, y).$$

- The sum of all intra-cluster variance.
- The increase in variance for the cluster being merged (Ward's criterion).
- The probability that candidate clusters spawn from the same distribution function (V-linkage).

Each agglomeration occurs at a greater distance between clusters than the previous agglomeration, and one can decide to stop clustering either when the clusters are too far apart to be merged (distance criterion) or when there is a sufficiently small number of clusters (number criterion).

Chapter 13

File Cutter and Evaluator

Because SoundID does all its work in RAM (as that is the fastest way to do it), very long files can easily cause the PC to run out of memory. The file cutter will automatically slice long files into more digestible pieces and can do this in a batch mode. The file evaluator can open any file (and in particular pre-cut files that are being considered as reference files) and rate them according to the signal to noise level in the file. This can be used to determine not only the quality of the references, but it can be used to calibrate a recorder and pick when it is performing below par.

When you first click on this module you will get this page.
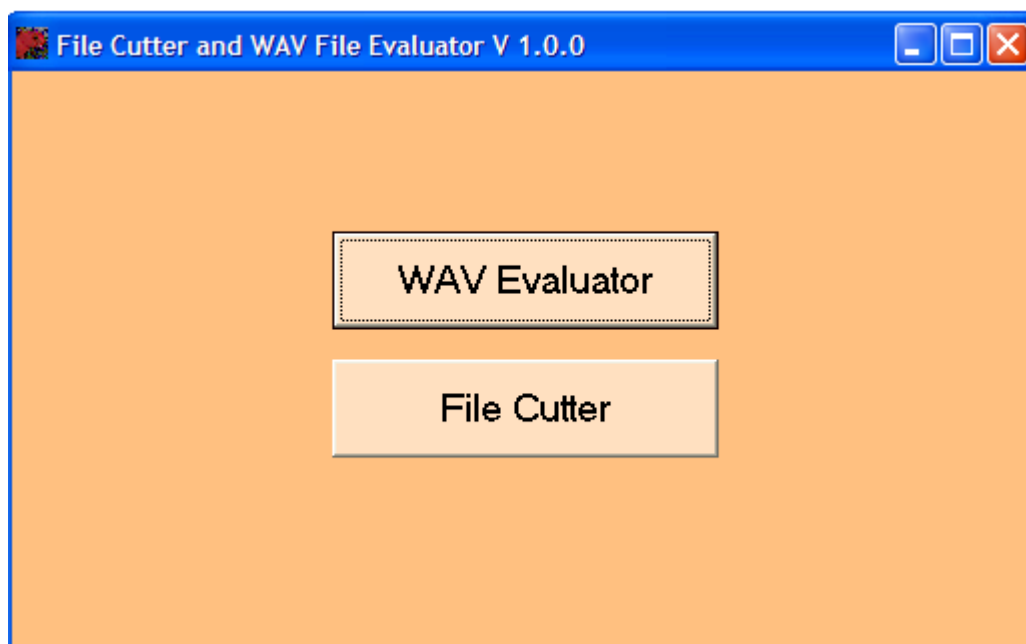
The File Cutter
First we look at the file cutter module. Quite a few recorders on the market today record continuous files in 2 GB chunks. This is far too much to load into RAM and as there is no option to save the files in smaller chunks (at least on many of the commercially available recorders), this could present some problems. But the file cutter can fix this. Click on the "File Cutter" button as seen above to get to this page.

You could alternatively use a good WAV file editor and likewise run a batch cut.
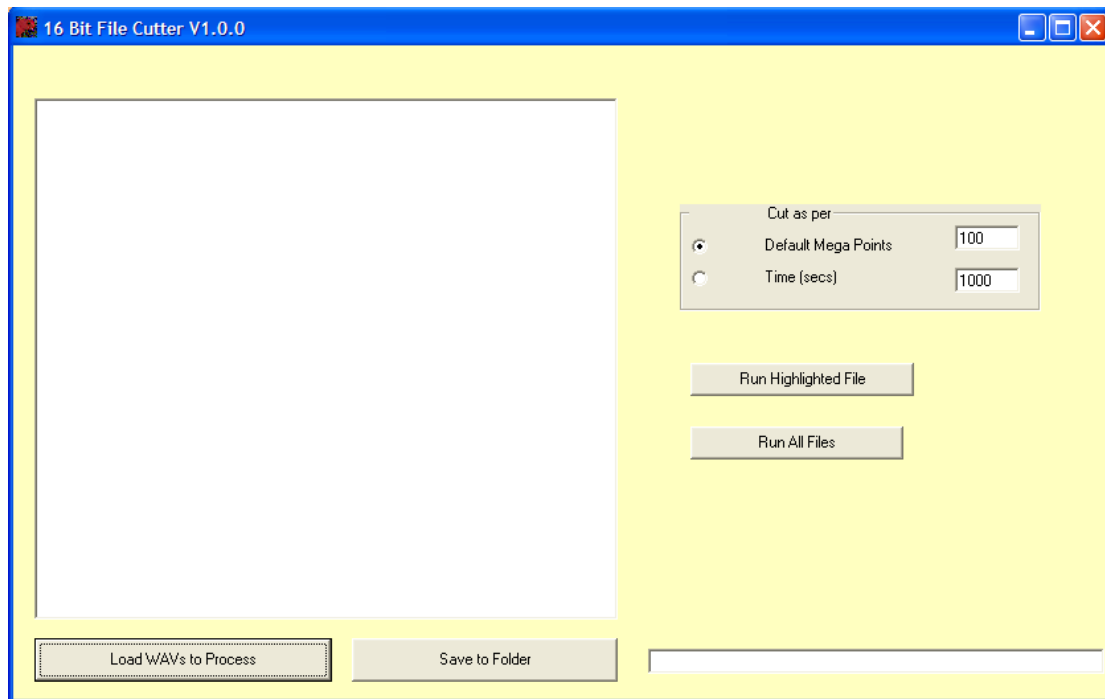
Now click on "Load WAVs to Process" to select the long WAV files that you want to cut.
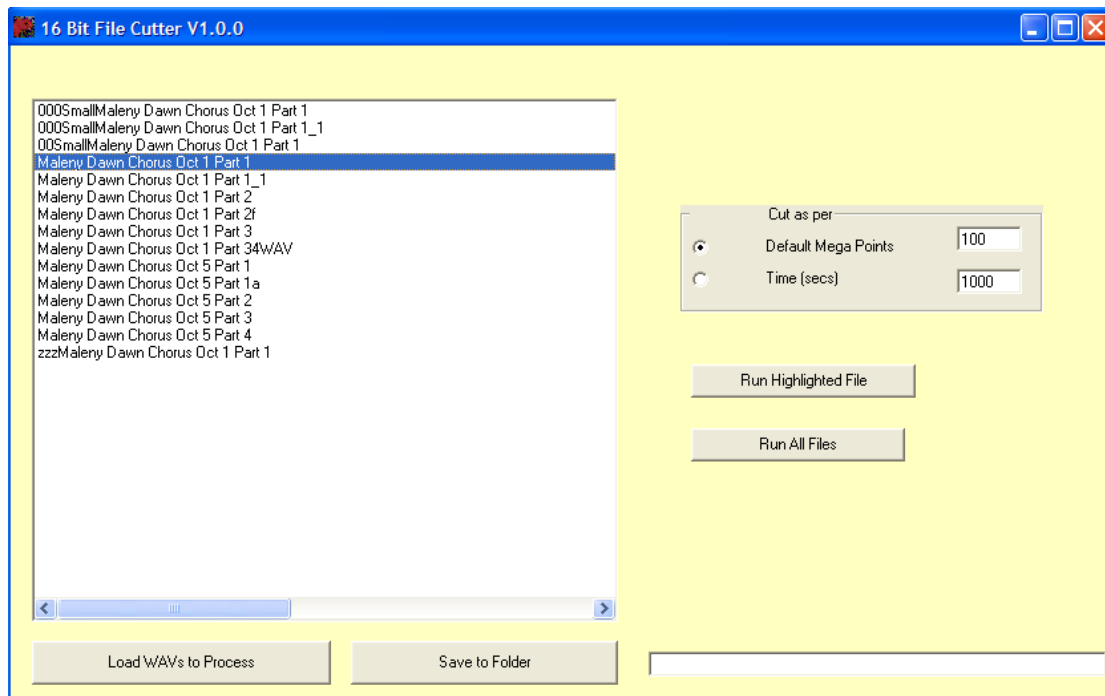
You now can choose one of the default cut sizes or enter a custom one. Having selected the cut wav file size you can choose just to cut the highlighted file or to cut the whole batch. When you select either option you will be asked to nominate of folder to save the cuts into. The cut files will then be saved with the original file name with _1, _2,……_n extensions with the extension indicating the order in which the files were cut. The highlighted file above was cut to the default 100 Mega points, as seen in the illustration below.
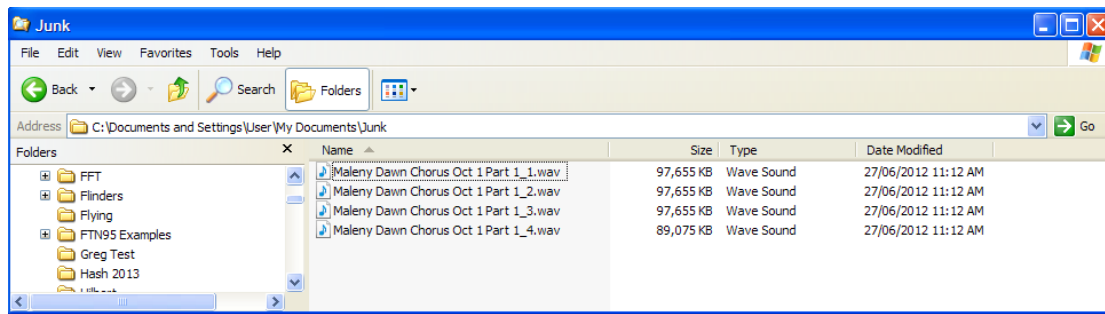
The WAV File Evaluator

It is often required to make decisions about which cut references to use and which to throw out. The references must of course be confirmed by listening to include only examples of the target species. If you include a duck call in with your frog references you should not be too surprised when the software incorrectly identifies the duck as your frog. But particularly if you have used the auto-cut module, there may be a lot of reference calls to check out. You will see that the File Evaluator can quickly identify the noisiest files which you can eliminate early to save to much manual work. Additionally you will see that the File Evaluator can be used to calibrate your recorder and to compare the results of different recorders and different recording techniques.

So let's look first at evaluation of some actual reference cuts.

First we click on the "WAV Evaluator Button to see the page below.
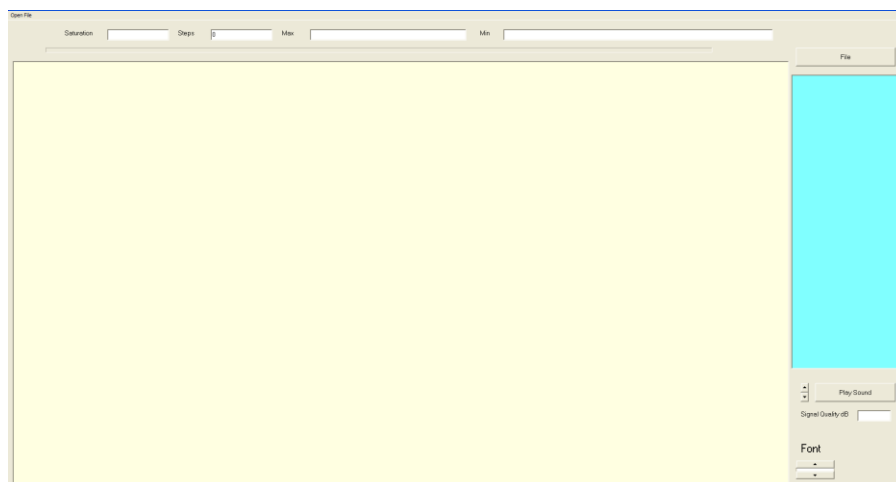
Now click on the "File" button and load in the reference WAVs as you would do in any other modules. Double click on one of the references to get the signal evaluation.

In this case I have selected a sample of Currawong calls and have clicked on the first one in the list. It reports back a signal quality (loosely a Signal to Noise value over 1000 points) and declares it to be excellent. The actual signal is also depicted graphically. The level is reported in linear amplitude points as well.

As with the other modules you can listen to the sound and stretch or compress it as required. You also can scroll the fonts to male them larger if you happen to be working with a smallish screen.
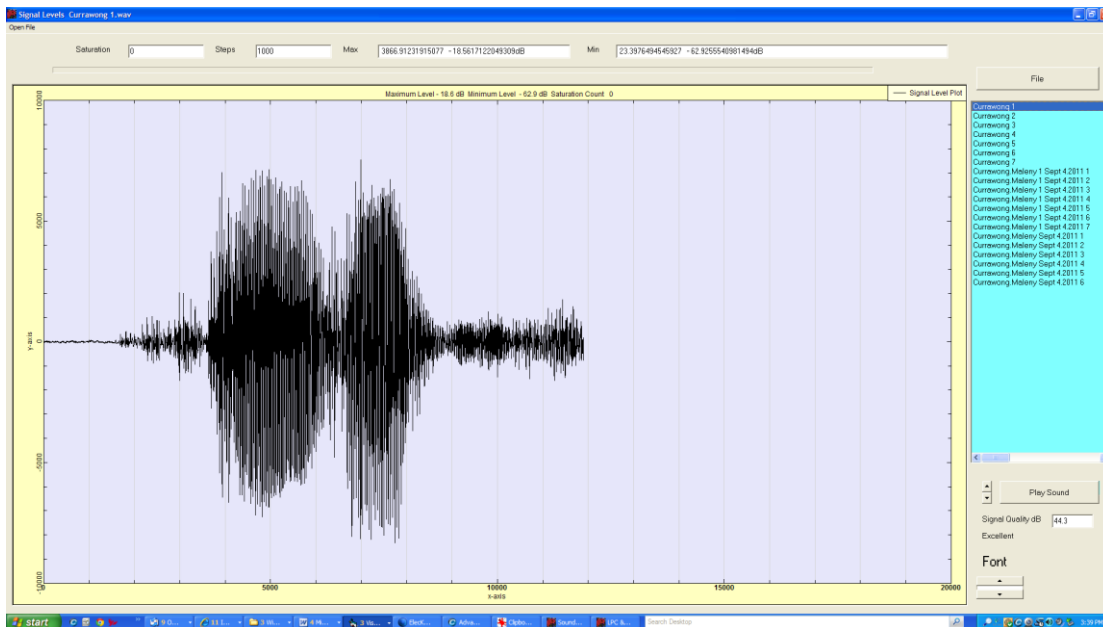
Next we look at a not-so-good signal which just happens to be the last one in that list.
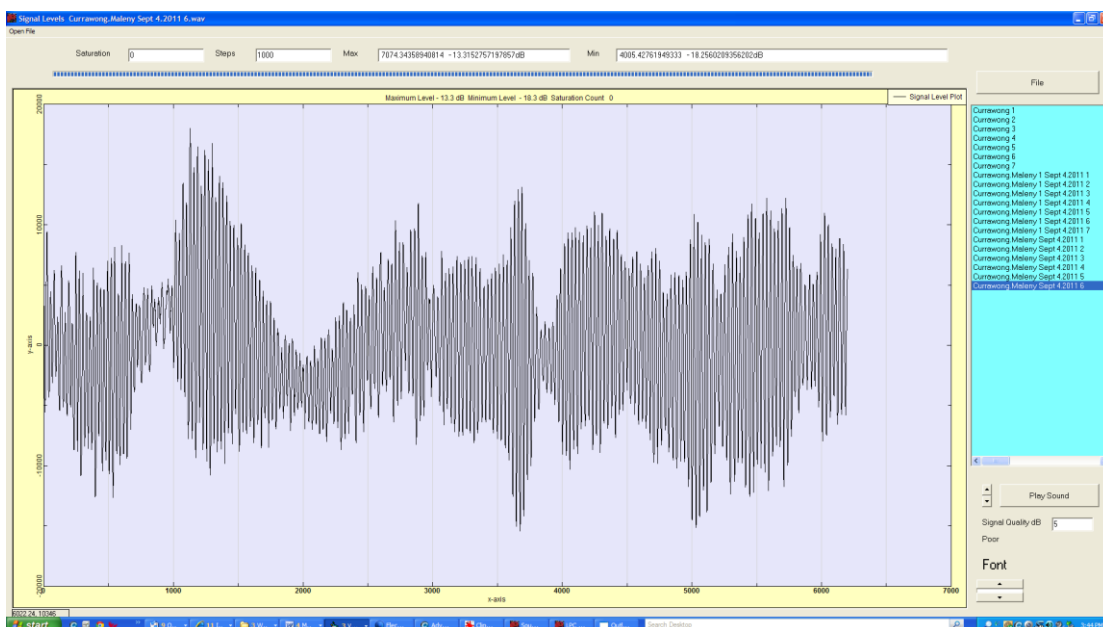
Notice that now the software reports the signal to be only 5 dB (S/N) and of poor quality. This particular signal has a lot of wind noise with it, and it is possible that it could be cleaned up with a filter.

Testing Recorders

This same module can be used to test your recorder. Find a "standard" sound source and a quiet room. At a pinch (but it is a bit rough) you can simply count to ten as the "standard" sound source. Much better however would be to purchase a standard sound source (see our web site) or improvise one using something like a small electric organ (remembering to always use the same key in the test). Then allow a time frame of about the same length of time as the "standard" source had of silence. The

resultant WAV file can then be put into this module, which will report the difference between the peak and minimum signal levels, which gives you a good idea of the recorders performance. Most important however is that the "standard" sound source is always at the same distance to the microphone (1 metre is the scientifically correct distance).

Recorders that are left outdoors for long periods of time will deteriorate. This is particularly true of the microphone. It is a very good idea to test each recorder before field deployment and here you have an easy way to do it.

Chapter 14

How to Transfer a SoundID Licence to a New PC

1. First go to the program folder and then to the SoundID folder. In the SoundID folder you will find a folder called Util0 as below.
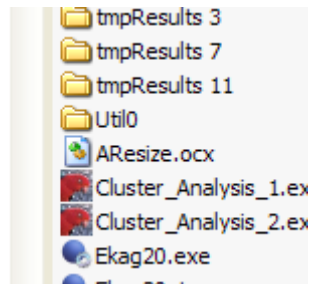
2.Open that folder and click the Transfer.exe application.
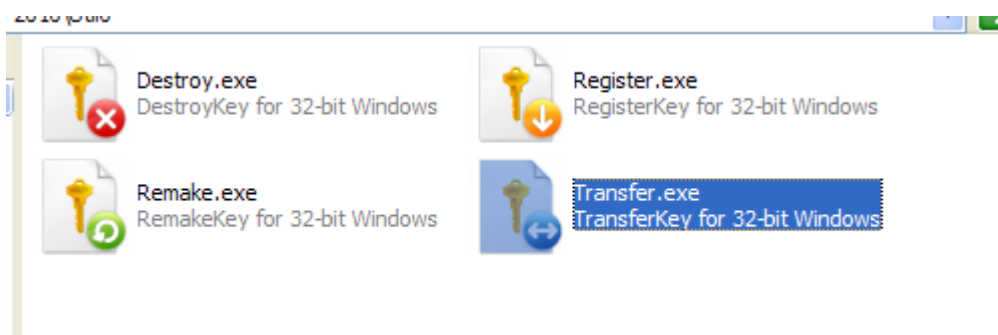
3. You will then see this screen.
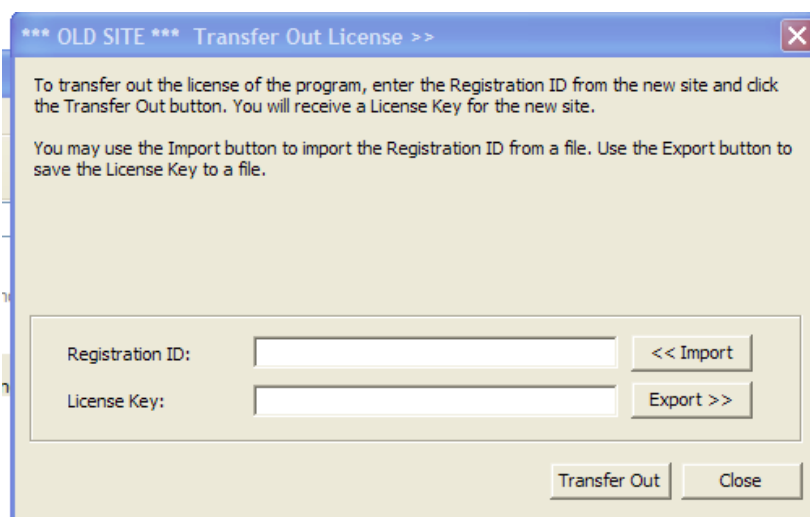
4. Now start SoundID on the new machine and cut and paste its Registration number into the text box above. Click transfer out and it will generate the key for you new PC.

Other Modules

There are a number of other modules accessible from the software these are accessed from the menu as shown below. These are modules that you will probably only need to access occasionally. However they can be useful and you are advised to become familiar with them.

The next few chapters will show how to use these modules.

Chapter 15

Digital Filter

The Digital Filter program will show you the effect of using the filter on a noisy signal. Select "High Pass Filter" from the drop down menu and you get this screen, after you click on the button marked "BF" and then click on the FIR Digital Filter button. If you adjust the parameters it will show you the effect of so doing. Notice however that this program is demonstrative only and does not have any direct effect on the running of any program.
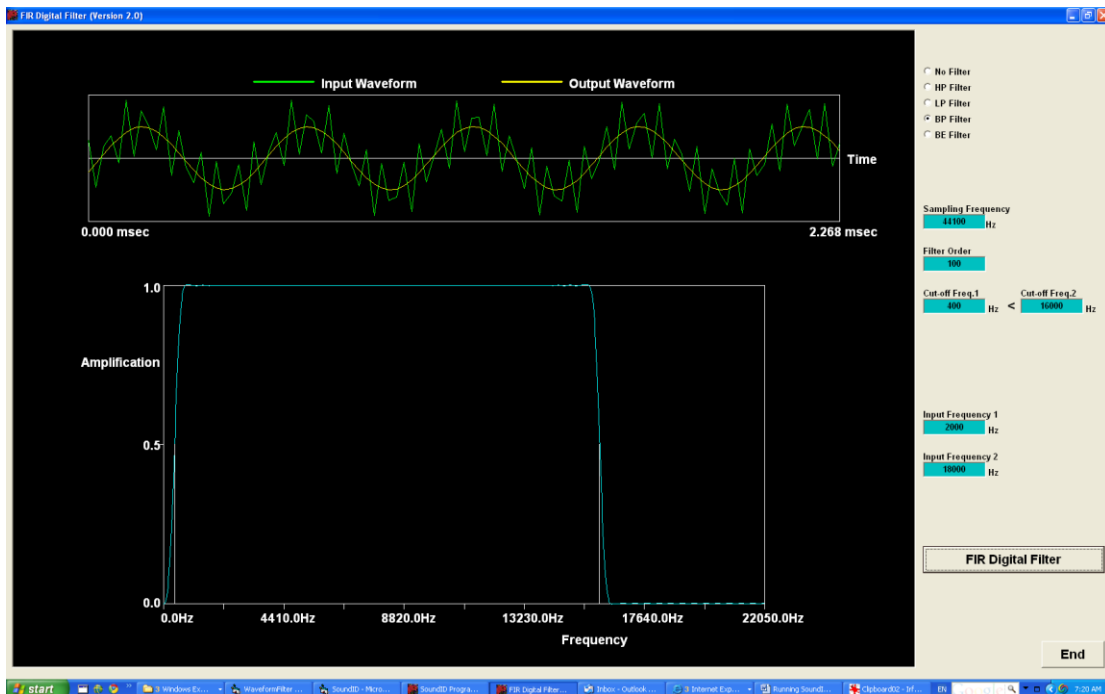
High Pass Filter

Chapter 16

PC Speed Test

The performance of SoundID in terms of time taken depends very much on how the templates are set up. But it also depends on the PC that you run it on. In particular, the software being intensely mathematical does an awful lot of number crunching.

How fast is your PC? The best measure for SoundID purposes is the flop rate, which is usually measured in Gigaflops on today's PCs. A reasonably fast Notebook is about 15 Gigaflops, while a moderately paced one is more like 5-10 Gigaflops. The 3770K, top end i7 (2012) is about 140 Gigaflops. The time to process a given workload using SoundID is inversely proportional to the Gigaflop rating (to a very good approximation). So the higher the Gigaflop rating is, the better.

To measure your PC performance, go to http://www.vtaskstudio.com/support.php and download "QwikMark". This is a very small (64 kB) .exe which you need to click on to make it run. You can also get to this site from the Menu page by clicking on the "Other Modules" menu item on the top left as seen below

When you run it you will see a screen shot like this.
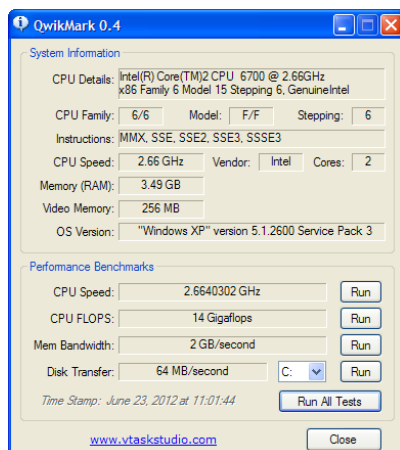
Notice the box near the bottom "CPU Flops" reads 14 Gigaflops. This tells you that this desktop PC (which is about 5 years old) runs at a modest 14 Gigaflops and is comparable in speed to a modern middle of the range Notebook like an i5.

But Wait, There is More
Having determined that you PC is fast (or not) you need to understand one more thing. What the

104

software measures is NOT the Megaflops of one CPU, but the total throughput of the machine. So now look towards the top right of the software and see that the machine has 2 cores. To get the actual per CPU speed we therefore need to divide by 2 (and get 7 Gigaflops/CPU).

Why do we need to do this? Because the current version of SoundID is 32 bit and when it runs it uses only one core.

However the batch mode of SoundID has been optimised to run *multiple instances* so that if you have a really big job to run, you can break it down into chunks and run them concurrently. When this is done you can run as many instances as there are cores to maximise you're processing.

Beware however that even here there is a trap for young players. Many people, including many PC shop vendors, confuse "cores" and "threads". The confusion is that most modern PCs have 2 threads per core and this is often *wrongly* equated to the PC having twice as many cores. *Wrong, wrong, wrong!!!!* For the purposes of CPU intensive software like SoundID, invoking the threads may *slow down* the processing by as much as 50%. If you want maximum throughput do not run more instances than there are cores. However, depending on the processing load vs the I/O load it may be that running a few extra threads may actually be advantageous. So feel free to experiment and use what is best for your data.

 If you want to use the PC for anything else while it is running a large SoundID job, use a maximum of cores -1. This will leave one core free for other tasks. If you fail to do this and use all of the cores, you will still be able to use the PC for other tasks, but you will be interrupting the processors when you run something else on the PC as it shares a core (using 2 threads), between SoundID and your other task.

Chapter 17

Template Header

If you have a large group of templates, you might want to know what settings each of them have. When you click on the "Other Modules" menu item as below you will see this screen.
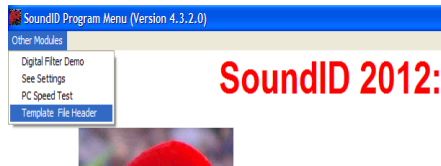
Now click on the "Template File Header" and then click the "Load Templates" button to select the templates of interest. They will all be added into the list box below and you can see immediately what they have been set to.

Species	LPC Order	Freq F1	Freq F2	LPC dB	GD	Spread	Points	Is Neg=1	Low Freq	High Freq	Filter Pos	Filter F1	Filter F2	Filter Order	Is Log=1	AGC %	AGC Secs	M Val	F power
Barbastella barbastellus	25	5000	100000	50	15	22	1301	0	0	0	0	400	16000	100	0	2	3	0.8	1.7
Eptesicus nilssoni	44	5000	70000	60	5	22	4001	0	0	0	0	400	16000	100	0	2	3	0.8	1.7
Hypsugo savii	38	5000	70000	60	5	22	3001	0	0	0	0	400	16000	100	0	2	3	0.8	1.7
Miniopterus schreibersii	22	2500	100000	60	9	22	1001	0	51631	59629	0	400	16000	100	0	2	3	0.8	1.7
Myotis alcathoe;	22	2500	100000	50	8	22	1001	0	0	48789	0	400	16000	100	0	2	3	0.8	1.7
Myotis blythii	22	2500	100000	50	12	22	1001	0	0	0	0	400	16000	100	0	2	3	0.8	1.7
Myotis brandtii	25	30000	80000	35	14	22	1301	0	0	0	0	400	16000	100	0	2	3	0.8	1.7
Myotis capaccinic	25	30000	80000	50	10	22	1301	0	41328	53633	0	400	16000	100	0	2	3	0.8	1.7
Myotis daubentonii	25	3000	100000	50	11	22	1301	0	38996	60973	0	400	16000	100	0	2	3	0.8	1.7
Myotis emarginatus	16	1	100000	60	5	22	501	0	10938	77735	0	400	10000	20	0	2	3	0.8	1.7
Myotis myotis	25	1200	100000	55	10	22	2001	0	8919	44039	0	400	16000	100	0	2	3	0.8	1.7
Myotis mystacinus	25	1200	100000	55	8	22	1001	0	42109	49056	0	400	16000	100	0	2	3	0.8	1.7
Myotis nattereri	25	1200	100000	55	9	22	1001	0	24742	62950	0	400	16000	100	0	2	3	0.8	1.7
Nyctalus leisleri	25	1200	50000	55	9	22	5001	0	9588	28078	0	400	16000	100	0	2	3	0.8	1.7
Nyctalus noctula	25	1200	50000	65	6	22	5001	0	19500	25028	0	400	16000	100	0	2	3	0.8	1.7
Pipistrellus kuhlii	25	1200	80000	65	6	22	2001	0	36906	39984	0	400	16000	100	0	2	3	0.8	1.7
Pipistrellus pipistrellus;	25	12000	80000	65	7	22	2001	0	38031	49719	0	400	16000	100	0	2	3	0.8	1.7
Pipistrellus pygmaeus;	25	30000	80000	60	6	22	2001	0	51875	57734	0	400	16000	100	0	2	3	0.8	1.7
Plecotus auritus	16	5000	100000	50	4	22	501	0	20900	25863	0	400	16000	100	0	2	3	0.8	1.7
Plecotus austriacus	31	5000	100000	55	6	22	2001	0	0	0	0	400	16000	100	0	2	3	0.8	1.7
Rhinolophus ferrumequinum	70	15000	150000	50	10	22	10001	0	29766	81973	0	400	16000	100	0	2	3	0.8	1.7
Tadarida teniotis	70	1500	20000	80	7	22	10001	0	9160	12773	0	400	16000	100	0	2	3	0.8	1.7
Vespertilio murinus	70	1500	40000	70	5	22	10001	0	20900	25863	0	400	16000	100	0	2	3	0.8	1.7

If you have old templates then expect to find "rubbish" numbers in some of the cells. These old files can be rejuvenated as explained in the Registration chapter.

Chapter 19

WAV File Header

It is not unusual for a WAV file to get corrupted. Also there are a lot of WAV files that are non-standard and cannot easily be read. If you are having trouble with a particular WAV file then this module will read the header information in the header file and it also displays the meaning of the data. A screen shot of this is seen below.
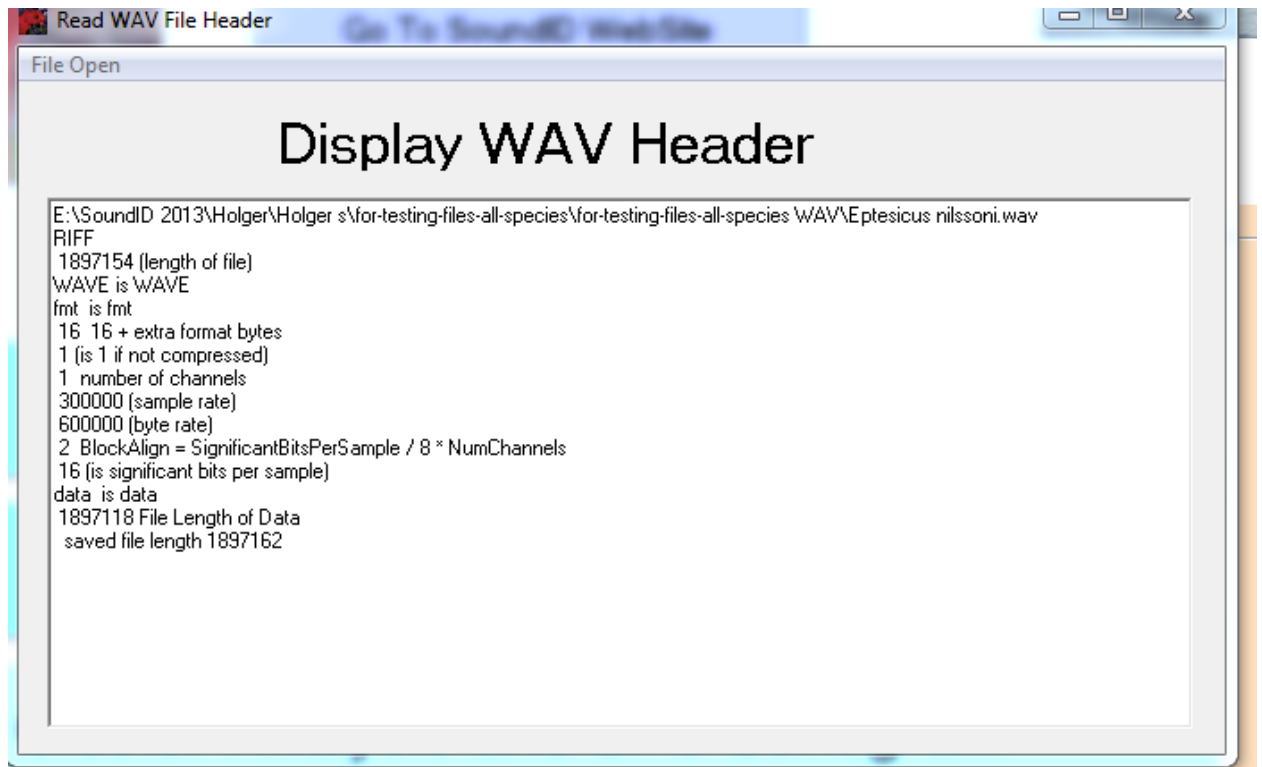

If the file is empty, this Window will often return the file address and nothing else. Such a file will be found to contain 0 bytes and so has not information in it. The file should be deleted.

Chapter 20

An Advanced Example of Running SoundID

In this section we consider how to use all the modules that are available in SoundID to generate a template for automated recognition. The objective will be to produce an optimised Template.

We begin by starting with our reference recordings and building a template from it. We use the example of the recording. We start with the recording of the Double Eyed Fig Parrot. These files are in the same directory tree where SoundID was installed (typically `C:\Program Files\SoundID`). The first thing is to load the recording into the Auto-cutter and cut the recording into its component call parts as below.
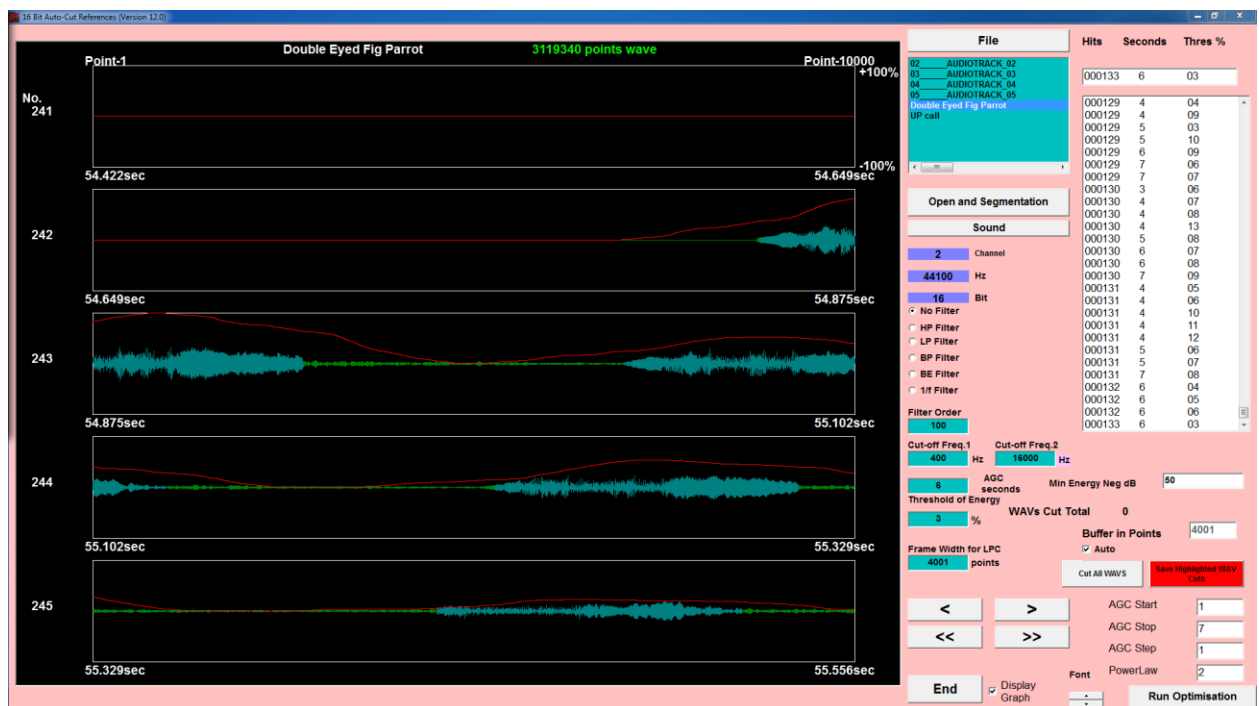
First look at thesettings on this screen. The filter is off and this should be the default with the filter only being used when it is necessary. Use of the filter will slow up the recognition process and will cause some loss of signal detail. If the filter needs to be used the next step will reveal this,

The default settings for the AGC is 3 seconds and 25% threshold. This may or may not be suitable for your recording and you can test this by running the "Run Optimisation" (the button at the bottom right). Be aware that this process can be quite slow for long recodings and you may prefer to cut the recording down to a 30-60 second sample for this process, the speed it up.

In the default mode the graphic display is turned off (as this seriously slows the processing if the graph is turned on). Once the run is completed, by scrolling to the

bottom of the list box (top right) you will find the values for the AGC that finds most of the significant pulses.

Other Factors
The next setting is the 4001 points. This number reflects the fact that the original software was designed to find parrots and for the particular parrot that was originally sought the typical call length was about 4001 points long. This can be observed in the graphic above where the call detected is seen displayed in blue. If your signals are significantly longer or short than this then you should adjust this value accordingly.

The minimum energy default is -50 dB (bottom right). This ignores low level signals which are generally not desirable. However if you calls are all low level and that is all you have, then you may need to increase this value (remember the value is negative).

The buffer is automatically set and adds to the selected signal that part of the signal that is a buffer number of points before and after the cut.

When the optimum setting has been found then press the red button below to cut the selections and to save them. The files will be saved as WAV files.

It is also of course possible to cut these files out manually and that is exactly what was done in the early days of SoundID development. The same files cut manually took about 4 hours to cut and 76 of them were found compared to the 133 found with an optimised automatic cut.

The next step is to examine and check the cuts. There are several ways to do this but the following is recommended. Load the cuts into the Cluster Analysis module as seen below.
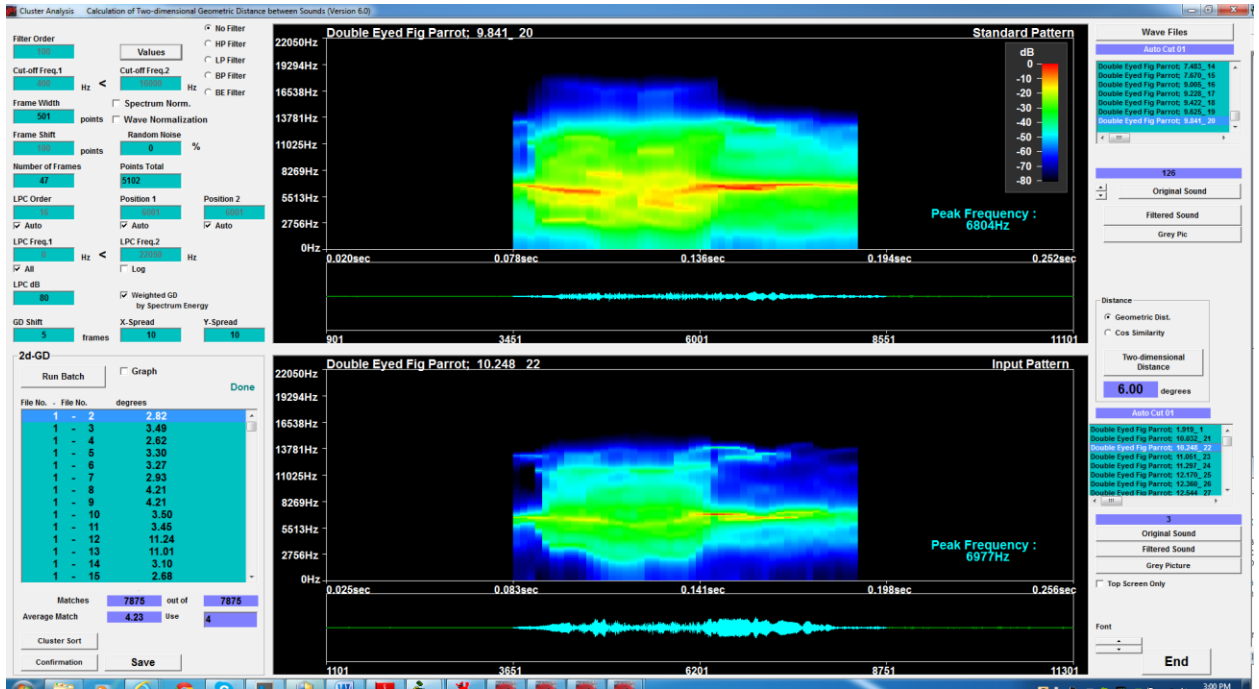
Once the cuts have been selected and the appropriate settings applied then press the "Run Batch" button (middle left). This will run a comparison between each cut and every other one.

Note that in the bottom left there are four text boxes that show how many matches were considered, the average a match and a "Use" value. The use value is the *suggested* GD to match to these signals. Notice that this should be treated as a starting point only as in the next step we will see how to look more deeply into these cuts. To do so press the "Save" button (lower left). Once you have done so the following screen will appear.
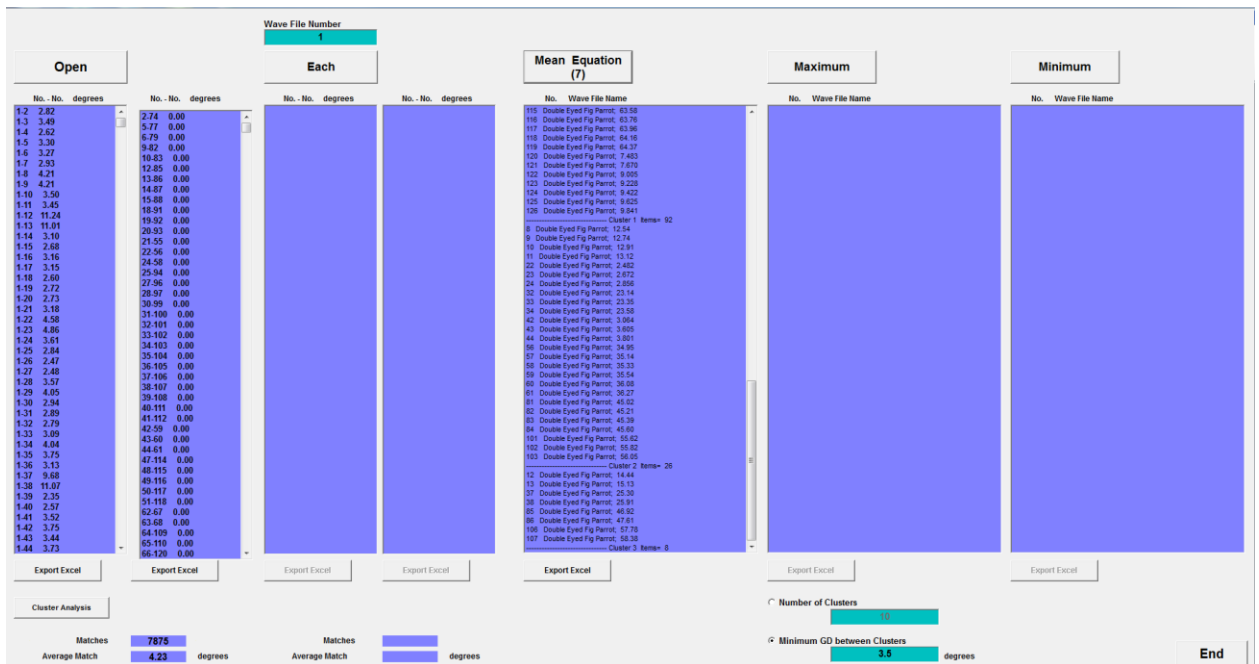
Set the radio buttons at the bottom to "Minimum GD between Clusters" and we find that there are only three clusters. This suggests that the files are quite good. At this point if there were signals that should not be there (for example a different bird that was intruding on the call) it should appear in a small or perhaps single cluster of its own. Any small clusters at this stage should be checked to see if they are spurious signals or simply rare wanted calls.

Interlude for very large reference files

You may or may not want to do this next step. As there are only three closely related clusters, if you have a very large reference set >1000 you might want to trim it down a bit and remove some of the examples from each cluster that are similar to one another,

Here we will also have a closer look at the signals within and between the clusters. Let's take the first and last in the first cluster (nos 115 and 126) as seen below.
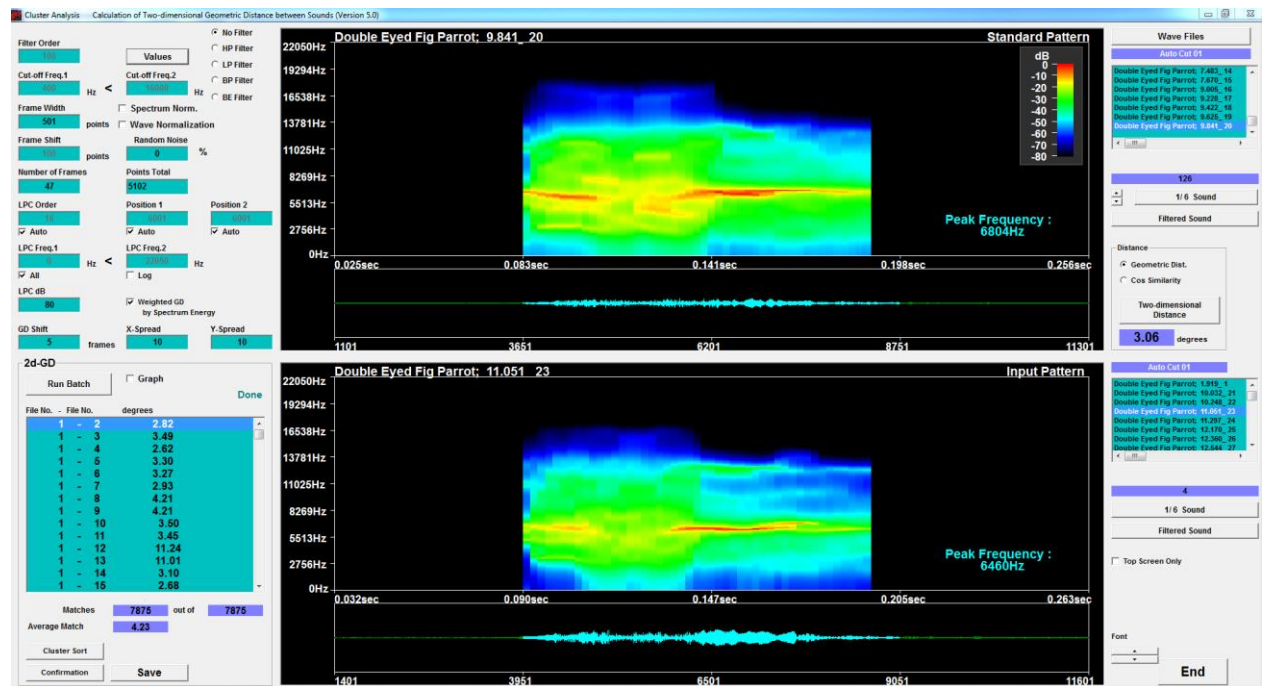
Notice that on the left there is a button reading 1/6 sound. This has been selected to slow down the sound to make it easier for a human to judge the similarity. The GD between the sounds is 3.06 which suggests they are very similar and the listening test confirms this. If your reference set is smaller than 1000 then it would probably pay not to trim any examples.

Now lets look at the first call (no 115) in the first cluster and the last call (107) in the last cluster.
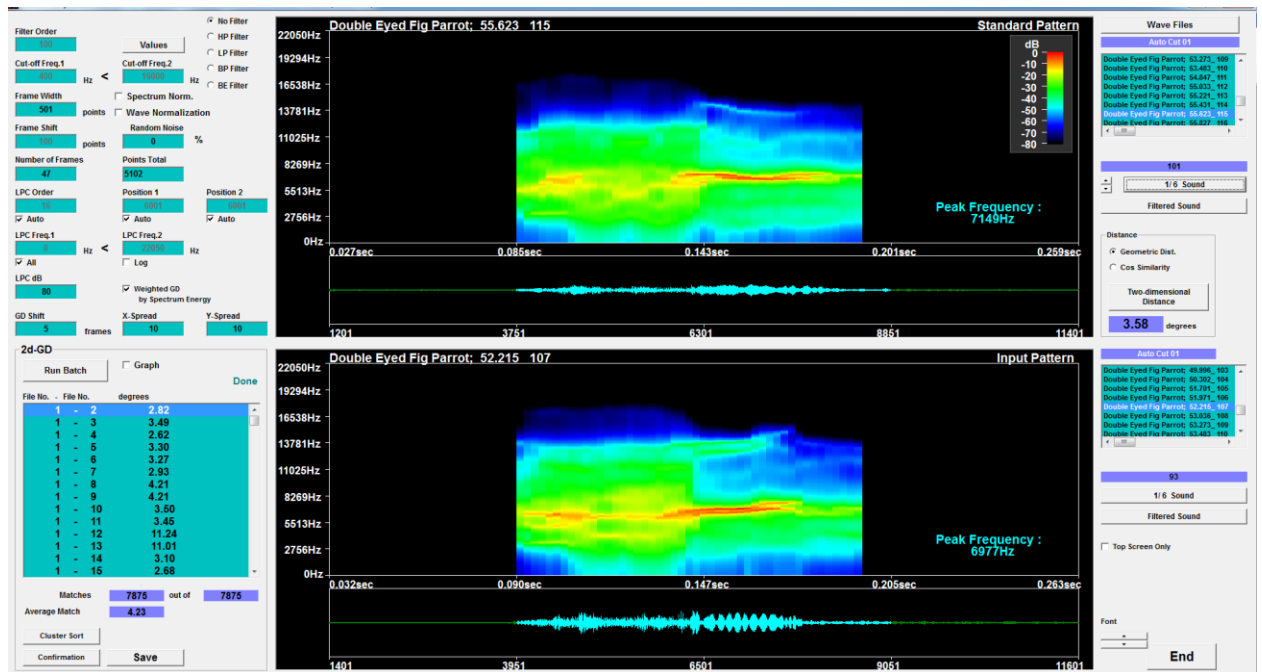
These calls are not too disimilar in the spectral view, and have a GD of 3.58. They do sound distinctly different and so it is a different version of a call type.

End of Interlude

So now we have 133 calls but only three distinct call types. They appear to be separated by about a GD of 3, which is not too far from the originally suggested GD of 4. For this particular species we do have other four other call types and in a real search application we would use all of the calls (about 500 in all) in the search.

If you have the patience and the skill, it would be a good idea to listen to each of these calls to ensure that the set consists only of the wanted call types. In the rea world, recording anything without getting the occasional intrusion from non-target species is very difficult and regardless of your sources you should be on the look-out for such intrusion. In particular, in early runs it is well to be on the look-out for strange matches which might indicate that an unwanted sound has crept into the reference list.

Making the Template

Now we are ready to make the template so we use the Registration module to do this. When we run this module we get the screen shot below. Notice that the area top right encircled by a white circle has an area of the signal that has "bottomed out. We could prevent this by increasing the LPC depth.
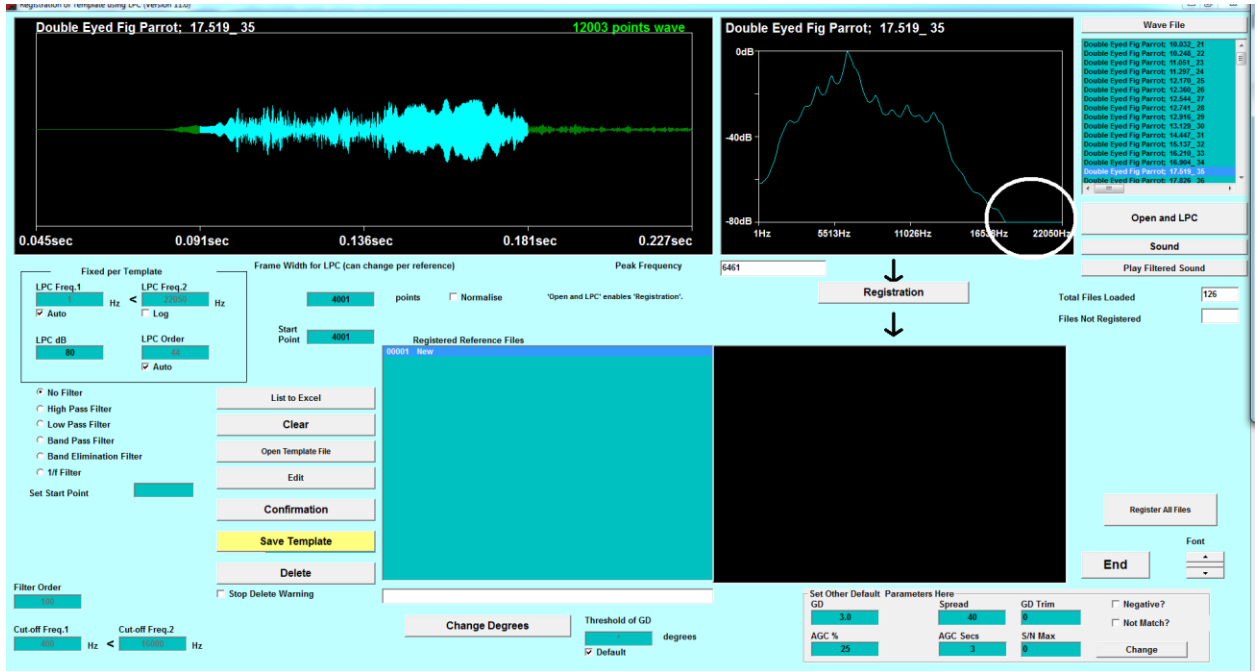
Noting that birds mostly cannot hear above about 12 kHz and that the signal is most likely to be noise below 500 Hz, we could reset Freq 1 and Freq 2 (top left) to narrow the band width. Now we have the screen-shot below.


Having done that we now see (white arrow above) that the signal is now well above the noise floor. So we can set if for 60 dB (instead of the default of 80 dB) and the graphic now fills the screen quite well. If we now press the "Register All Files" on the bottom left we can watch the image for each reference and confirm that the "bottoming out" does not occur.

Next we run the cluster analysis again with those settings and get the screen shot below.
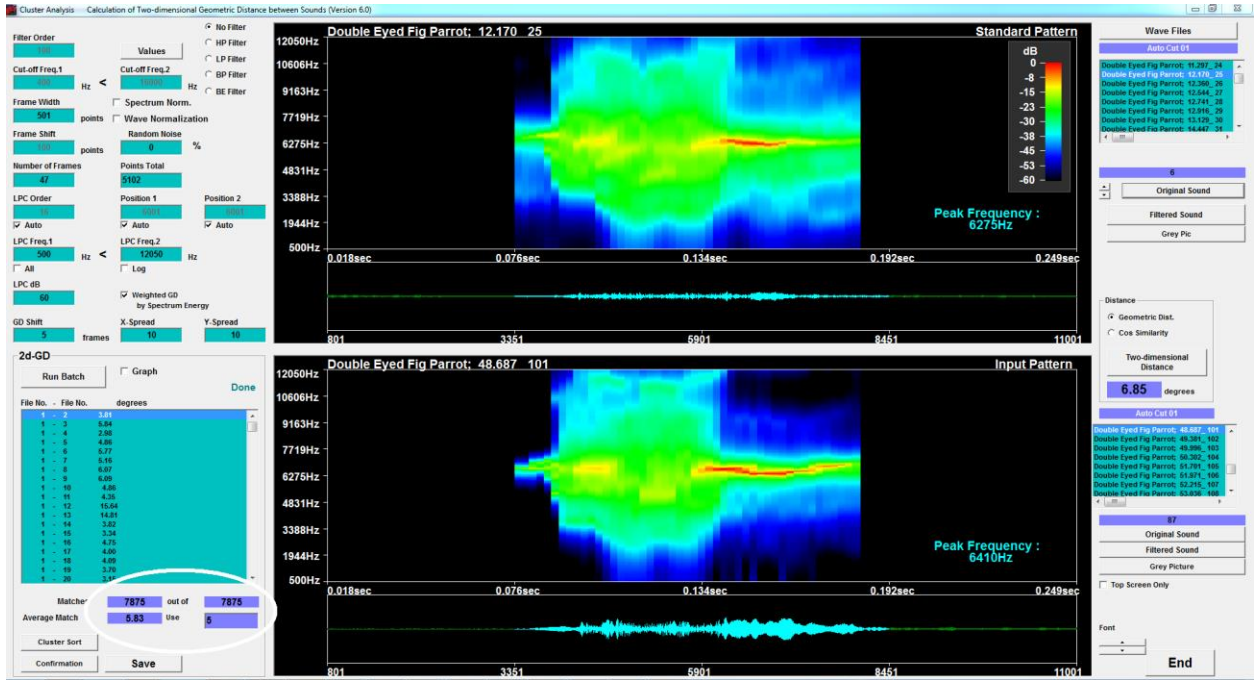
Notice that the suggested GD to use (bottom left) is now 5 (from the previous value of 4). If we set the cluster minimum GD to 5 we now get 4 clusters as seen below.
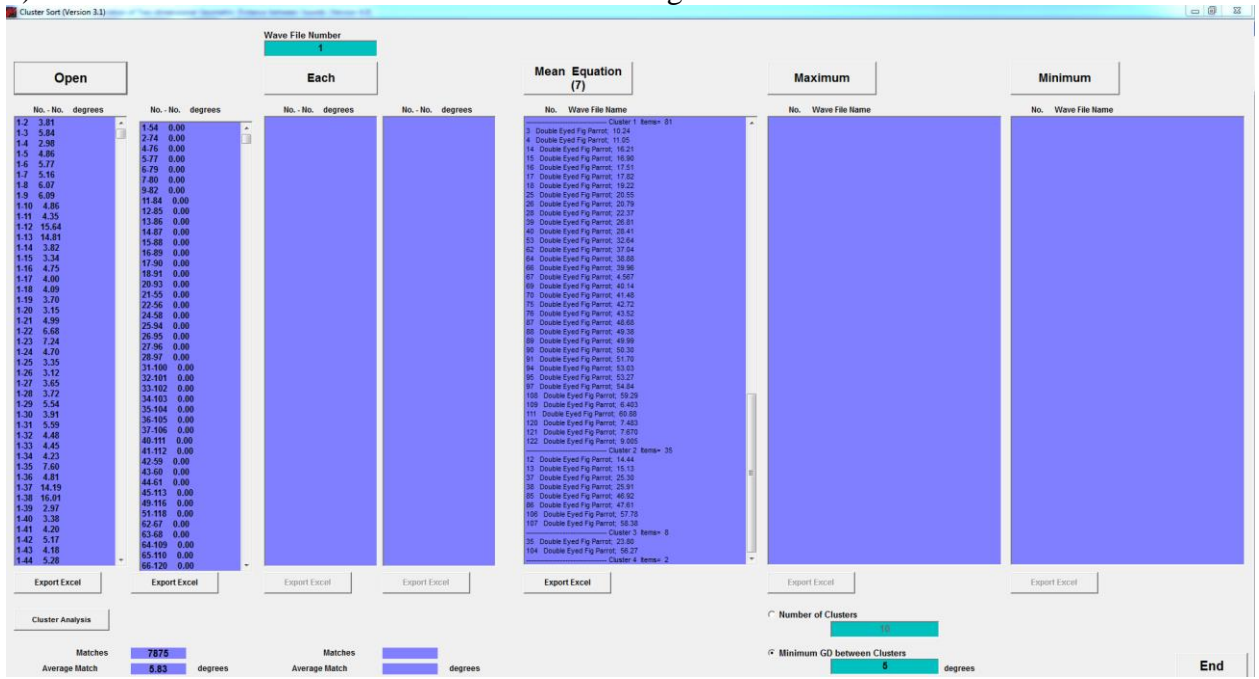
In general, if we zoom in by using values of F1 and F2 less than the default values and/or lower values of LPC dB, then we should expect to get a higher matching value of GD.

Setting the AGC

The AGC setting that was applied to the reference calls, will probably not work particularly well for the field calls, which will be much noisier. To determine the AGC for the field calls cut a few representative examples to lengths of about 1 minute and run them as you di with the reference calls through the Auto-cutter and optimise for the best number of detections. Let us assume that the field calls were optimum with AGC=10% and 4 seconds. You would now enter these values into the Recognition module and run the "Register all Files" to get the screen below.

Save this Template and you now have your optimised Template made.

Testing

The next thing is to test this template and it is probably best to do this in Recognition. First run a test with the file that the Template was cut from and confirm that it returns the expected number of matches. In this case we started with 126 files and when we run in Recognition as seen below we get 137 matches. The difference is due to the AGC settings (which you will recall were in this case arbitrary).
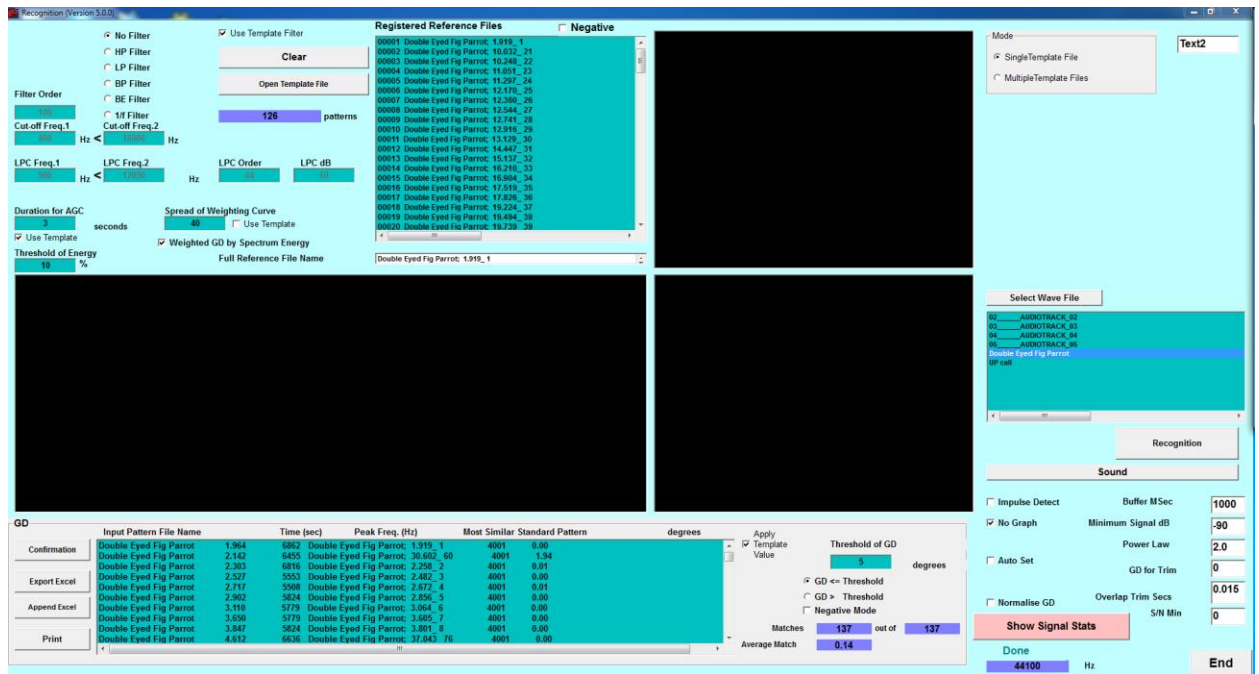
If you have some more known long recordings that should match, run them also. Remember that regional accents and the limited number of examples (there were only 4 similar clusters in this group) may result in non-matching and this may mean that you must expand your template(s).

Having passed the easy test we should now apply some more difficult tests to see if we can avoid false positives. This species often reportedly is miss-indentified by humans with the Scaly Breasted Lorrikeet and the Musk/Little Lorrikeet. So we can run a few examples of that species by the software. In the original development of the software to recognise th Coxen's Fig Parrot we used a collection of recordings from Dave Stuart to make sure that we would be able to recognise the target parrot when other parrots where known to be present in the same area. So we ran all of these parrots past the software and indeed the Musk/Little Lorikeet (mixed), did give the false positives as below.
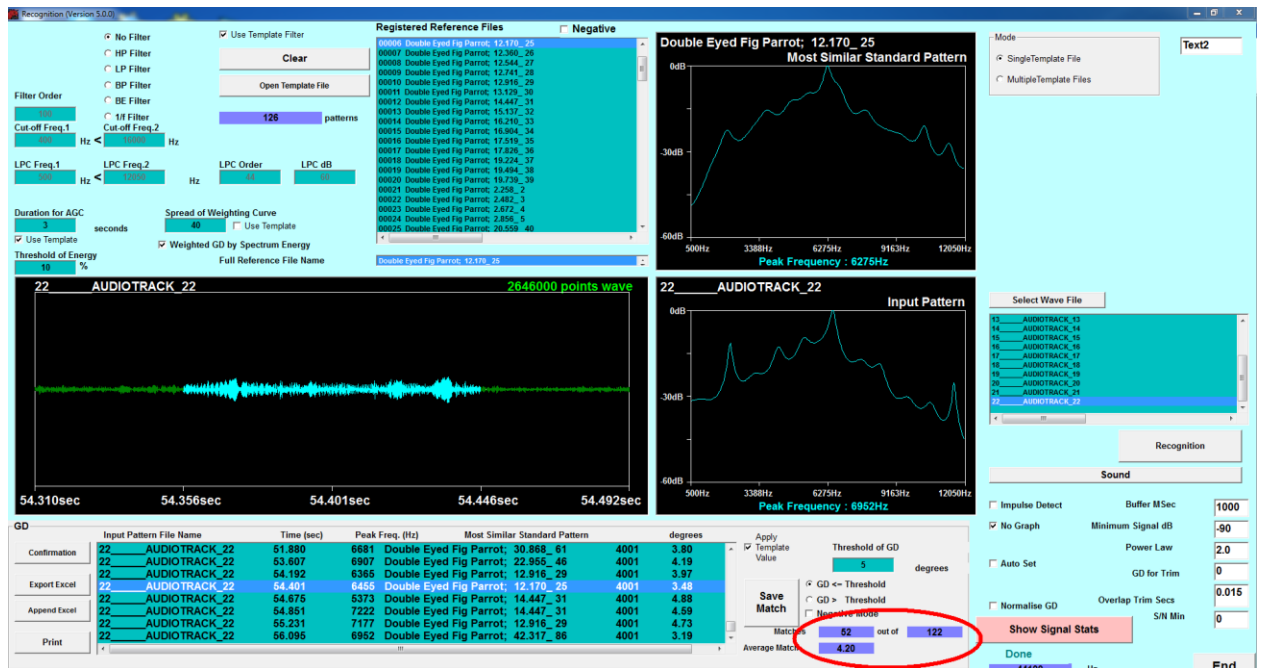
Now we find that there are a lot of matches that are false positives. The circled area shows 52 matches out of 122 hits. This is good, not bad!

If we now scroll through the results we find that the "best match" is GD=3.14 (just above the red ellipse). So if we set the GD to 3.00 we would have NO false positives. And that is what we should do. To do this simply import the Template into the Registration file, but clicking "Clear" and then "Open Template File" and change the GD to 3.00. Now save the template and it is done.
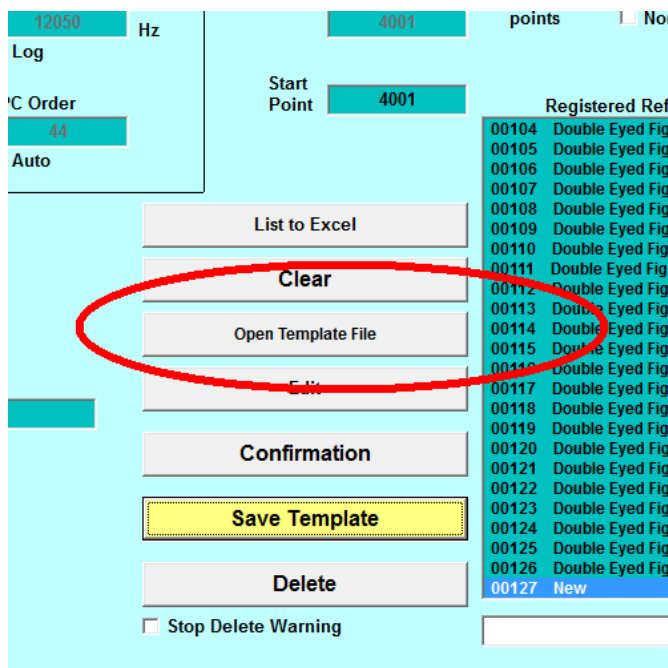
Research Possibilities

Distinguishing Subtle Sound Differences

Because SoundID can distinguish subtle differences in sound, better than can the human ear, there is the possibility of studying small differences in sounds. The ideal module for such a study is the LPC & GD. Examples of such studies are regional differences in bird or other animal calls, running engine and other mechanical sounds. Most importantly of course this same module can be used to examine the reference calls to ensure that a sufficient cross section of target calls are included.

Differences of about 2 degrees of GD are on the limit of the ability of a human to discriminate. Although a human may detect a slight difference if these are played sequentially, such small differences would not ordinarily be noted. However for SoundID, 2 degrees is quite a substantial difference and it can return a metric of such differences.

Searching for Rare Sounds

Although SoundID is mostly intended to search for similarity, it may be used for the exact opposite; that is to seek out exceptions. By setting the search to >GD SoundID will find all the sounds that do not match the references. This can have many uses and one of the most obvious is when extracting references from a long recording, the references can then be run in the exception mode to reveal sounds that were not selected. This way the reference file can be improved, with software finding the excluded sounds that should be part of the reference file.

Other Recommended Software

Although we have tried to make the software as self-contained as possible, there are a few other programs that may be useful and can be used in addition to SoundID. The first is an excellent and free image editing software called InfanView, http://www.irfanview.com/. This is very useful for capturing images and for passing SoundID results directly into other documents. Highly recommended.

SoundID (professional) can be run entirely without a sound editor, but in many cases it is nice to have one. The standard edition of SoundID does not have the Auto-cut for the references and so will need a sound editor to support it. Cool Edit is widely used and works well, but it is no longer supported and this is beginning to show, with some strange behaviour on some of the latest Windows operating platforms. Its successor Audition is a very capable editor, but it is a bit more complex and expensive than it needs to be. However for serious users we do recommend it. Most people have their preferred sound editor and as it is only an adjunct to SoundID, if the user is happy with their existing editor then it should be fine.

Recommended Hardware

The SoundID software will run on *any* Windows PC. However it was developed on high end machines with 1080 screens and sometimes small screens can be problematic for the way they display. If you are an intensive user here is what we would recommend.

If you do have a small screen, look for the "Font" scroll bars and adjust them for the best results for your screen.

If you are on a budget

The highest clock speed you can get for the price.
A 1080 screen (or an external screen with 1080 specs if you have a lap top).
4 GB or more of RAM

For those with an unlimited budget

An Intel i7 4770K processor or higher (the earlier 3770K will be fine also and is negligibly slower). The higher the clock speed the better (this processor is 3.9 GHz).

16 GB or more of RAM
A video card capable of driving three or more screens.

Three or more 27 inch screens.

Windows Ultimate (one of the best features of Windows Ultimate is its excellent multi-screen capabilities).

Checklist

Not Detecting Signals

A simple check once you have made your reference (.dat) files is to run that .dat file, with the original WAV file that the references were cut from, in Recognition or Batch and you should get 100% of the original cuts in the .dat files matching at GD=0.0. If you do not get that, it is most likely that you have not correctly set the detection parameters (AGC seconds/threshold). Run Segmentation and then Optimisation to find the values that give full detection (AGC seconds/threshold). Reset those values and re-run.

False positives >5%

The false positives should be well below 5%. First check that you have sufficient references (more than 20) and that those references are of good quality (noisy references are always a problem). If you are still getting false positives check the Windowing. The call should be Windowed (the settings of the number of points, LPC F1 and F2, LPC dB, and LPC Order) to best typify that call.

Trap

Avoid the twisted thinking that the real world is noisy and so noisy references are more realistic in the real world. This simplistic thinking leads to problems. Noisy references have noise that is characteristic of where and when they were recorded. It is most unlikely that the reference noises will be the same as the real world noises. By using noisy references you are setting the software the task of finding the (target call+ the background reference noise), which is probably not what you really want.

Filtering
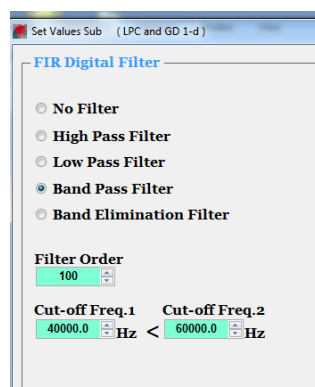
Filtering needs to be used with considerable caution not only because it slow the processing, but because the filtering itself can cause things to "match" that should not. Below we see two bat calls the Myotis nattereri and the Myotis emarginatus with no filtering and a GD of 22.41 correctly indicating "no match". However is we apply a filter as seen opposite we find that it distorts the image so much that the same two signals now match with a GD of 3.67. The filter, wrongly applied can cause this kind of problem with all kinds of signals!

Two bat calls with no filtering that do not match,

The same two calls with the filter set between 40 KHz and 60 KHz causes false matching!
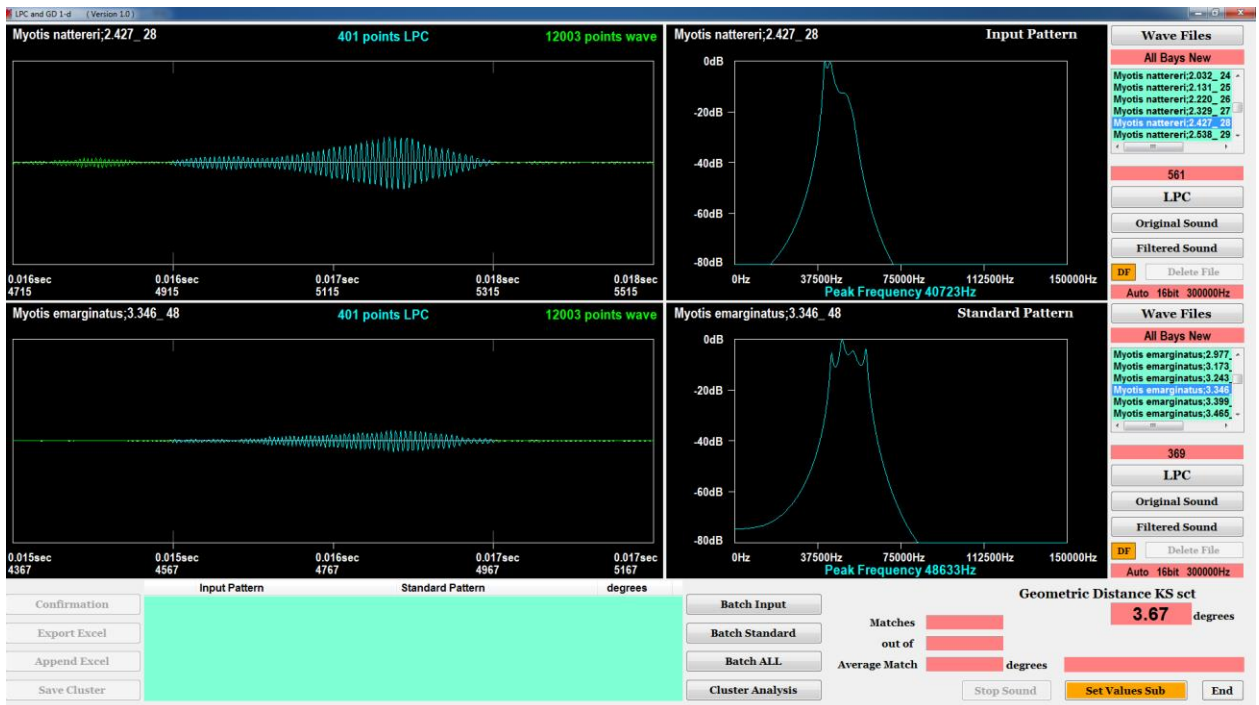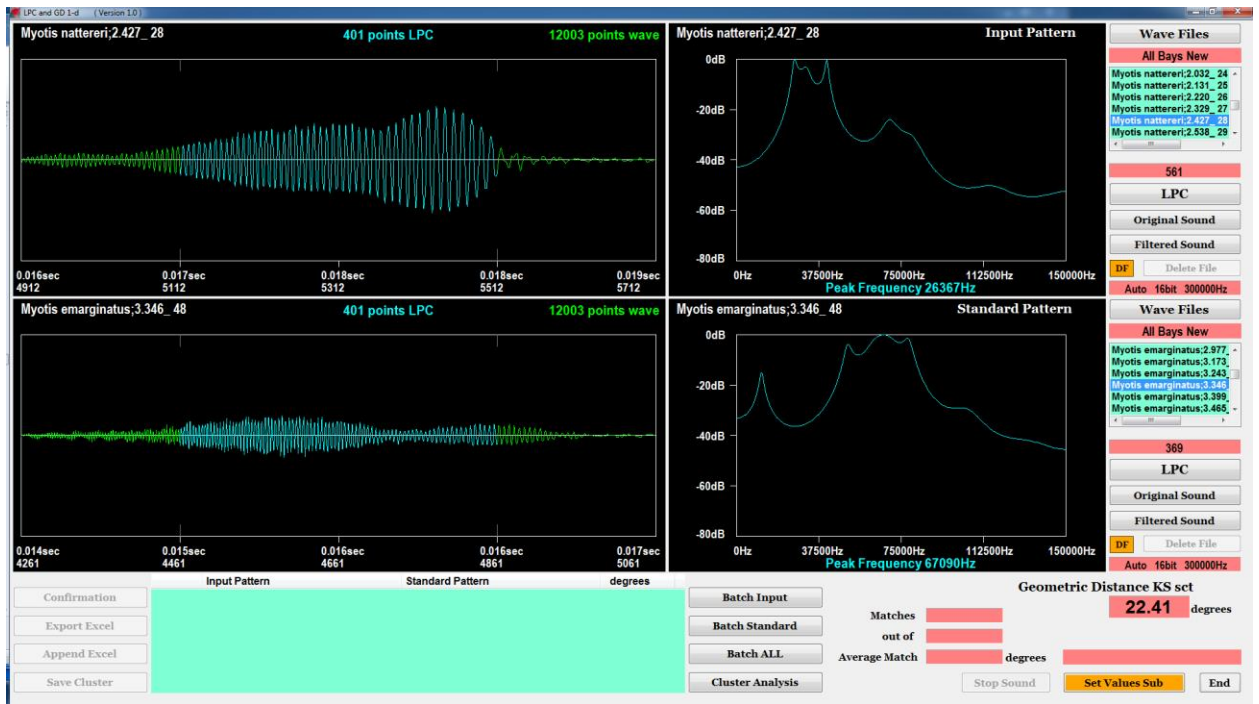
Important Hints

Sufficient References

Always make sure that there are enough samples in the registered files to include a good representation of the calls of the target. The software relies on a good range of sample calls. In general any reference file of less than 20 original WAVs is unlikely to be representative. Remember also that there is no real upper limit to the number of references that can be used and that even large numbers of references do not slow the processing up significantly. For example on tests done with 50 reference calls compared to 1000, the extra processing time will be about 5%.

Regional accents

If the target is animal species, be aware of regional accents and make sure the reference file contains examples of the local versions.

Use of Filters

The default in the Registration and Recognition programs is no HP filter. If you use the filter on one of the programs and not the other, there will be a slight mismatch between the reference calls and the recognition. A good way to keep track is to add the filter bandwidth to the name of the reference file e.g. "Double Eyed Fig-parrot HP400".

Minimise Frame-width Variations

In theory every reference call can be cut to a different frame-width (number of points). This however is not a good idea, as the software treats each batch of reference calls of a given frame-width as separate reference blocks. This in turn means that a reference file with two different frame-widths will take about twice as long to process as it would had all the frame-widths be the same (remember that the total number of references in a given frame-width does not matter very much in regard to processing time). The more frame-widths that you have in the reference file however the slower will the processing be. To a good approximation it will slow down in direct proportion to the number of different frame-widths. This of course should not stop users from using multiple frame-widths where appropriate, but one should be aware of the trade-off in processing time.

Long Frame Widths

Long frame-widths can be necessary for things like whale calls. It should be noted that the processing time increases roughly as the square of the frame-width and for very long calls this may be an issue.

File Size

Be aware that the software loads all files processed into memory and this limits the length (or more strictly the size) of files that can be processed. If an "out of memory" message comes up this is most likely the issue. It is recommended that PCs used for this software have 4 Gig of RAM fitted. With 4 Gig of RAM it is possible to run 3 hour files recorded at 44,100 kHz. If multiple processors are available and used to run multiple instances, then the total file size (size x number of instances) should be kept below 3 hours for 4 G of RAM). In Windows it is not possible to say definitively how much of the RAM is available for running applications. Windows uses RAM for all sorts of background tasks. Accordingly for PCs with less than 4 G of RAM, experimentation will reveal how large the files can be.

As the memory size is based on both the sample rate and the length of the recording, the above rule should be modified accordingly for different file recording rates.

Frequency

You can change the frequency band that is considered. The default frequency band is 1 Hz to 22,000 Hz. If you have narrow band signals or high frequency signals, you will get better resolution if the restrict the frequency band so that it does not look outside your band of interest. For example parrots typically call at 5 kHz (but not all parrots) and rarely make sounds below 800 Hz. So for a parrot you could restrict the bandwidth to be between 800 Hz to 12,000 Hz.

INDEX