Sonification of Hyperspectral Image Data

M. Bernhardt, C.E. Cowell and W J. Oxford
Waterfall Solutions Limited
Parklands, Guildford, Surrey, GU2 9JX

Abstract

There are many reconnaissance tasks which involve an image analyst viewing data from hyperspectral imaging systems and attempting to interpret it. Hyperspectral image data is intrinsically hard to understand, even when armed with mathematical knowledge and a range of current processing algorithms. This research is motivated by the search for new ways to convey information about the spectral content of imagery to people. In order to develop and assess the novel algorithms proposed, we have developed a tool for transforming different aspects of spectral imagery into sounds that an analyst can hear. Trials have been conducted which show that the use of these sonic mappings can assist a user in tasks such as rejecting false alarms generated by automatic detection algorithms. This paper describes some of the techniques used and reports on the results of user trials.

Keywords: Sonic transformation, hyperspectral data, sonification

Introduction

The use of hyperspectral imagery for either reconnaissance information or strategic intelligence gathering relies on the use of human-based analysis and automated processing. People, and in particular image analysts, have a very highly developed ability to recognise things spatially. However, comprehending high-dimensional data sets is something that people, even with experience, are rather poor at. Because of this, any tools that can genuinely assist in the comprehension of hyperspectral data will assist an analyst in his or her task.

This paper describes a novel tool to assist human operators in their interpretation of hyperspectral data via real-time sonic transformations. Previous work was conducted [1] under this project which was based upon the notions of consonance (where simultaneous sounds are pleasant to the ear) and dissonance (where simultaneous sounds are harsh and unpleasant). Initial experiments showed sonification to be a very useful tool in assisting the analysis of hyperspectral data. The experiments showed that it is possible to out-perform the industry standard algorithm (RX) in terms of differentiating between genuine anomalies and false alarms. The work presented here follows on from these preliminary findings. More sonification mappings were developed and used together with the dissonance mapping to perform more extensive trials.

This paper briefly covers the sonification tool used to trial the different sonic mappings, a summary of some of the mappings used, describes experiments run and presents a precise of the results and conclusions.

Sonification Tool

In order to perform user trials a software tool was created which implements the sonic mapping of choice and allows a user to listen to a hyperspectral image (HSI). The tool allows a user to read in a hyperspectral cube in any of the common formats, or any general format specified by the dimensions of the cube, and the orientation of band data. The cube is then displayed in a window, allowing visual inspection of any chosen band or a number of user-specified bands to produce a colour image. Controls also allow the user to choose the particular mapping of interest and to determine the basic waveform for use with that mapping. If the mouse is moved over the image display then the pixels over which it moves are rendered into sound in real-time according to the selected mapping. Figure 1 shows an example screen-shot of the tool.

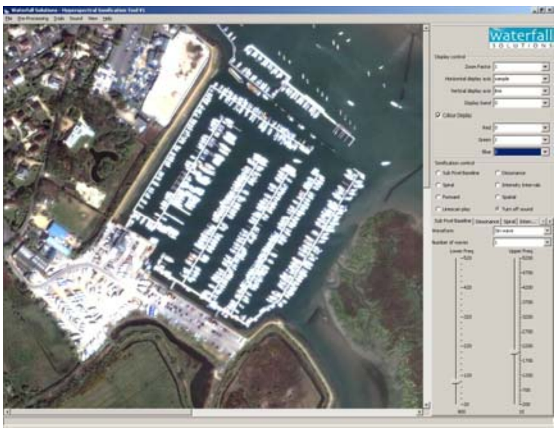

Figure 1 *Main screen of the HSI Sonification Tool.*

For initial tests, an option to read ground truth data associated with a cube together with algorithm detection data was implemented. This data was represented as a set of red dots on the screen at the locations of targets and false alarms. The user was then invited to listen to the data and to click on any red dots that were believed to be genuine anomalies. When the test was complete, the program generated the user's score in the form of a confusion matrix which shows the true class (anomaly or false alarm) in each row versus the chosen class in each column.

Formant Mapping

A formant is a peak in the frequency spectrum of any acoustic system which arises from resonance on the frequency source. Formants are the distinguishing frequency components of human speech and are particularly important in vowel formation. The ability to distinguish vowel sounds can be attributed to the differences in the first three formants. Due to their resonant origin, formants tend to stay the same even when the frequency of the fundamental waveform is changed. The formants make each vowel easily identifiable for the brain - even in foreign speech people are able to recognise vowels. The brain seems to have 'hard wired' pre-processing to deal with formants. This ability of the brain suggests that formants may lend themselves to make a humanly-interpretable acoustic mapping. Figures 2 to 4 show the frequency spectrum for three different vowels and illustrate the differences between the first three formants. Table 1 shows typical frequencies for the first three formants of three different vowels.

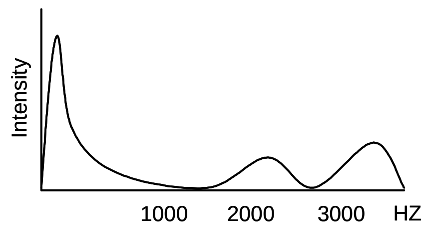

Figure 2 *Frequency spectrum for the vowel EE*

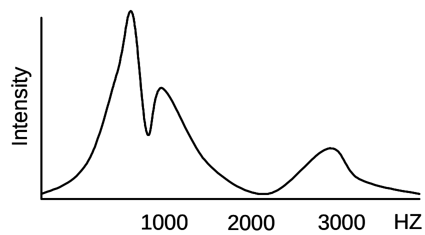

Figure 3 *Frequency spectrum for the vowel AH*

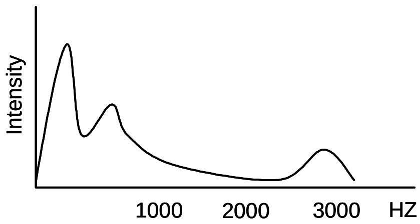

Figure 4 *Frequency spectrum for the vowel OO*

First three vowel formant frequencies in Hz			
Vowel	F_1	F_2	F_3
/i/	250	2290	3010
/a/	730	1090	2440
/u/	570	850	2410

Table 1 *Table showing the first three formant frequencies for three different vowels.*

All of this would suggest that the use of formants would be a natural choice for a sonic mapping. The formant mapping is produced by taking a sound wave of some form (e.g. a saw-tooth wave or sine wave) and filtering it with the different frequency spectra of a vowel. The level of mixing is determined by the level of the wavebands of the image and different wavebands control the levels of mixing of the different vowels. Results reported here are for the three vowels contained in Table 1.

Spiral Mapping

Both the dissonance and the formant mappings consider only spectral information. For different tasks it may be desirable to take into account the spatial information that is also available from an image. A spiral mapping was therefore produced in order to give the user the option of an acoustical mapping which contained both spatial and spectral information.

In order to produce sound, a time-series is required that represents the amplitude of the sound at each discrete time-step. The sonification tool uses pairs of time-series so that the sounds can be in stereo, and the standard (CD audio quality) sampling rate of 44.1KHz is used. The sonic mapping generates a time-series from HSI and the sound produced by this mapping is related to both the spatial and spectral content near the mouse position in a very direct way.

Spatial Mapping

The spatial mapping was created for data without any spectral content and instead concentrates on the spatial characteristics of the image. The mapping was created by measuring spatial characteristics of the data over a user-defined window. The horizontal and vertical differences and centre surround contrast are used to determine the deviation from the core notes of a chord. If the window selected covers a region that is bland then the chord will sound pure. If the window selected covers a region which has more structure and variation the chord will be much less harmonious.

Hyperspectral Analysis Experiment

The aim of the experiment in the previous work [1] was to determine if the dissonance and baseline acoustic mappings contained any useful information, and to test if the dissonance mapping was of more benefit than a baseline mapping. Given the proof of the utility of an acoustic mapping, the experiment reported here was designed to establish whether there is a particular sonic mapping which gives better results for all candidates, or whether different people have different preferences for mappings.

The concept upon which the trial was based is that another algorithm (for example, RX) would be used initially to screen a hyperspectral data cube to select a number of pixels or areas of interest. These would then be interrogated by the user using sonification in order to determine a priority order or remove false alarms.

The experiment was run giving the user 20 selected pixels, 10 of which were false alarms selected using RX and 10 of which were target spectra embedded at a range of levels. The images were presented in a random order with a different sound playing each time: first spiral, then dissonance, and then formant. Each of the targets could then be listened to with the sound mappings and the subject was asked to identify which pixels were targets.

Table 2 shows the results in the form of a so-called confusion matrix for the spiral mapping. An ideal result indicating perfect performance would have 100% in the top left and bottom right, and 0% elsewhere. Table 3 gives the dissonance mapping results and Table 4 reports on the formant mapping results. Note that these results are for targets at a range of different embedding strengths. It is interesting to observe that the clear trend apparent in these averaged results was also closely followed in the individual results, proving that averaging was justified to give a good overall summary.

All users were able to distinguish targets from false alarms using all of the different mappings after a little practice. There was a clear order of merit for the three mappings used for this experiment, with the spiral mapping being the preferred method.

		Declared as	
		RX False Alarm	Anomaly
Truth	RX False Alarm	87%	13%
	Anomaly	6%	94%

Table 2 *Confusion matrix for spiral-based sonification algorithm averaged over all embedding levels.*

		Declared as	
		RX False Alarm	Anomaly
Truth	RX False Alarm	77%	23%
	Anomaly	29%	71%

Table 3 *Confusion matrix for dissonance-based sonification algorithm averaged over all embedding levels.*

		Declared as	
		RX False Alarm	Anomaly
Truth	RX False Alarm	64%	36%
	Anomaly	30%	70%

Table 4 *Confusion matrix for formant-based sonification algorithm averaged over all embedding levels.*

Search Experiment

The aim of this experiment was to determine whether sonification can aid an analyst searching for geometric, spatial targets in a single band image. The concept on which this experiment was based is that image analysts often have a large image in which they have to search for targets whilst minimising false alarms.

A number of 500-by-500 pixel, synthetic single band images with geometric targets in realistic spatially correlated clutter were generated and a number of targets placed in the image, with different levels of noise added after target insertion. An example is shown in Figure 5. These images were generated to have power-law spatial autocorrelation functions such as those found in natural scenes.

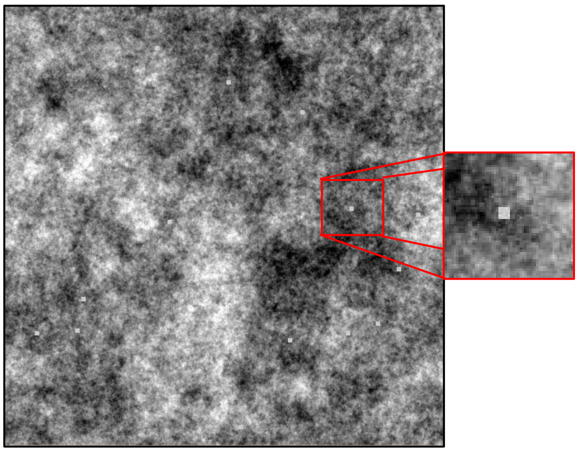

Figure 5 *Example of the synthetic images generated for search experiment with a close up view of one target*

The subject was presented with the different scenes in a random order and asked to locate the targets in the image. A time limit of 2 minutes was given per image in order to try to represent the fact that this task would normally be done with under a time constraint. The first image was presented without any sound, the second with the spiral mapping producing sound, and the third with the spatial mapping. This continued until a number of different noise levels had been heard for each sound type.

Figure 6 gives an example of results achievable with sonification. In this case the individual under test clearly improved their detection performance and false-alarm rejection abilities using the spiral mapping. In the figure a graph of probability of detection versus Signal-to-Noise-plus-Clutter Ratio (SNCR) is given and Table 5 reports the false-alarm rate results achieved.

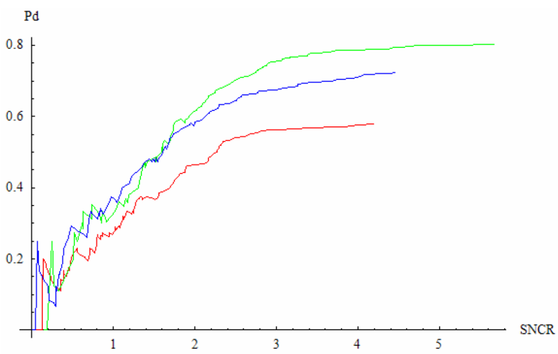

Figure 6 *Example detection result from trials. Red Curve: No Sound; Green Curve: Spiral Mapping; Blue curve: Spatial Mapping.*

Mapping	False alarms	False alarms standard deviation
No sound	63	7.8
Spiral	33	4.8
Spatial window	31	7.2

Table 5 *False alarm rate results for same subject reported in Figure 6.*

Matched Filter Experiment

The trials conducted in the first phase of this work had been completed with a limited number of people having completed the experiment. An extended trial was therefore held in which many more people took part.

In these subsequent tests, a similar experiment to the hyperspectral analysis was conducted (with a small number of changes). Due to feedback from the users of the tool, a target panel was added so that individual pixels identified as false alarms could be played by pressing the button associated with the pixel. The algorithm used to identify the false alarms was also changed to be a spectral filter (the adaptive cosine estimator, ACE). The user was given identified false alarms and targets at different embedding levels as before, and once again the target spectra were embedded at 100%.

The later results show that the spiral mapping adds considerable benefit, with the maximum percentage of targets correctly identified being 80%. Figure 7 shows the bar chart of the percentage of correct and incorrect targets identified over the different embedding levels.

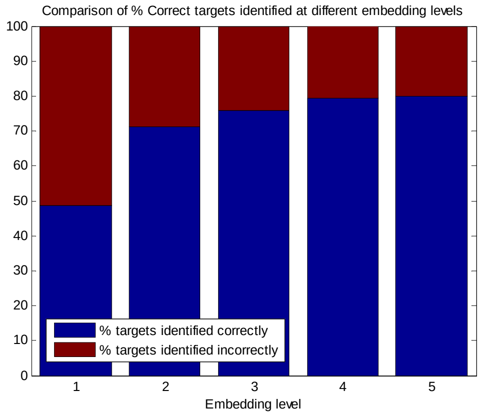

Figure 7 *Targets identified correctly and incorrectly for the different embedding levels with the spiral.*

However, results were not so encouraging for the dissonance mapping in this application. The main difference between this experiment and the hyperspectral analysis experiment reported earlier in this paper was that the false alarms generated by the matched filter resemble the target spectra, whereas those generated by anomaly detection have no particular relationship to the targets. The dissonance mapping was designed for anomaly detection where anomalous pixels which are not targets sound quite dissimilar from true targets. This is the origin of the superior performance of this mapping in the anomaly detection task. The spiral does not make such an explicit assumption – and this appears to make it much more robust, having the best performance in both applications.

Summary and Conclusions

It is clear from these experiments that sonification can be a very useful tool to assist the analysis of hyperspectral data. These results and more extensive tests have given weight to earlier findings that the novel method of using sonification to deliver hyperspectral data to a user has significant merit.

The hyperspectral analysis experiment showed that all of the sonic mappings described in this paper have utility in false alarm rejection and that users showed a clear order of preference, with the spiral mapping being their first choice.

The search experiment taught that subjects have to view potential targets with their eyes before sonification provides benefit (by confirming true targets and rejecting false alarms).

For the matched filter experiment, the spiral mapping substantially increased the level of correct targets and false alarms identified, while the dissonance mapping did not perform so well.

References

1. Bernhardt, M; Heather, JP; Edmunson, PF; 'Strategic Hyperspectral Detection through Sonification', 3rd EMRS DTC Technical Conference, Edinburgh, 2006

Acknowledgements

The work reported in this paper was funded by the Electro-Magnetic Remote Sensing (EMRS) Defence Technology Centre (DTC), established by the UK Ministry of Defence and run by a consortium including SELEX Sensors and Airborne Systems, Thales Defence, Roke Manor Research and Filtronic.