PERCEPTION AND DETECTION OF AUDITORY OFFSETS WITH SINGLE SIMPLE MUSICAL STIMULI IN A REVERBERANT ENVIRONMENT

RUSSELL MASON AND STEPHEN HARRINGTON

Institute of Sound Recording, University of Surrey, Guildford, UK
r.mason@surrey.ac.uk

It is apparent that little research has been undertaken into the perception and automated detection of auditory offsets compared to auditory onsets. A study was undertaken which took a perceptually motivated approach to the detection of auditory offsets. Firstly, a subjective experiment was completed that investigated the effect of: the sound source temporal properties; the presence or absence of reverberation; the direct to reverberant level; and the presence or absence of binaural cues on the perceived auditory offset time. It was found in this case that: the sound source temporal properties had a small effect; the presence of reverberation caused the perceived auditory offset to be later in most cases; the direct to reverberant ratio had no significant effect; and the binaural cues had no significant effect on the perceived offset times. Measurements were conducted which showed that the -30dB threshold below the peak level of the slowest decaying frequency bands could be used as a reasonable predictor of the subjective results.

INTRODUCTION

One of the major challenges remaining in the analysis of audio signals is the division of a binaural sound field into separate 'sound objects'. The human perceptual system can separate a collection of audio signals arriving at two points in space into a meaningful scene in which separate sound sources and the effects of the acoustical environment can be deciphered and labelled. There are many cues that can be used to achieve this separation, most of which are summarised by Bregman [1]. The relatively new field of research that involves using computational methods to automatically separate and label sound components has been termed Computational Auditory Scene Analysis (CASA).

Two important cues that can be used to separate a sound field into separate objects are the starts and ends of sounds. A large amount of research has been conducted into detecting auditory onsets (when a sound starts), such as the work of Scheirer [2], Rodet and Jaillet [3], Bello and Sandler [4], Smith and Fraser [5] and Klapuri [6]. However, relatively little research has been undertaken into detecting auditory offsets (when a sound ends). It is possible that the onset detection techniques may be adapted for use in detecting auditory offsets, but as this paper discusses, there are a number of complications related to auditory offsets that may mean that onset detectors are unsuitable for this purpose.

Bello *et al.* summarise and compare a number of onset detection techniques [7]. They also clearly define what they mean by the term 'onset'. Their definition is that an auditory onset is a single moment in time selected to mark a longer transient at the beginning of an auditory event. Based on this, an auditory offset can be defined as a single moment in time selected to marker a longer variation at the end of an auditory event.

This paper takes a perceptually motivated approach to Computational Auditory Scene Analysis – as such, the primary interest is in predicting perceived effects, such as perceived onsets and perceived offsets, rather than attempting to base the analysis on objective criteria or to attempt to surpass the limited analytical capabilities of the human perceptual system. On this basis, the definition of a perceptual onset is the moment at which an auditory event is perceived to start, and the definition of a perceptual offset is the moment at which an auditory event is perceived to end. In the context of a musical signal in a typical acoustical environment, the perceived onset is the moment at which the note is perceived to start, and the perceived offset is the moment at which the note is perceived to end, leaving only reverberation.

There are a number of reasons why it is useful to be able to automatically detect the perceived offset of an auditory event. It is useful in the automated analysis of the spatial attributes of a signal – as discussed by Griesinger, the division of source-related and environment-related aspects needs to be undertaken in a perceptually relevant manner, rather than on the basis of an impulse response or a single time division [8]. Also, the automated separation of source signals and reverberation can assist in a number of applications, such as enhancing speech signals for the hard of hearing and pre-processing signals for speech recognition.

Therefore, this paper investigates perceived offsets in more detail. Firstly, comparisons are made between onsets and offsets, in order to judge the similarities and differences between these two types of events so that an evaluation can be made about whether similar methods can be used for each. Secondly, a subjective experiment is described that investigated the perceived offset points for a number of stimuli. Finally, comparisons are made between the subjective results and objective measurements so that an automated method of predicting perceived offsets can be developed.

1 AUDITORY ONSETS AND OFFSETS

As already mentioned above, the majority of work in this area has focused on the detection of auditory onsets. This is likely to be due to the fact that auditory onsets are more perceptually salient, and that they are easier to detect – both perceptually and using computational methods.

Considering the range of possible attack durations created by musical instruments and the voice it may be considered no easy task to automatically detect auditory onsets using computational methods. Luce and Clark investigated the duration of the attack transient of non-percussive instruments and found that it varied significantly between instruments, players, and the pitch of the note played [9]. An onset detector would not only have to cope with these variables but also with the wide orchestral dynamic range and the rapid succession or simultaneity of onsets.

A number of strategies for detecting auditory onsets have been proposed, a number of which were reviewed by Collins [10]. These include detection of level variations (e.g. [2]), detection of spectrum variations (e.g. [3]), and detection of phase variations (e.g. [4]). These are usually followed by some form of simplification and selection algorithm based on probability or psychoacoustic modelling (e.g. [6], [11], [12]).

Compared to onset detection, successful offset detection is arguably more difficult due to the following factors. Firstly, the range of decay times that are present in musical signals cover a much wider range than the attack times. A plucked or struck instrument may create a sound that decays over a large number of seconds compared to a heavily damped instrument whose sound may decay within milliseconds. This is significantly larger than the range of attack durations from 14 to 85 ms identified by Luce and Clark [9]. Secondly, musical signals may contain multiple decay segments that may confuse an offset detector. This is a particular problem with instruments that contain transient attacks, such as percussion, piano and plucked strings, where the initial transient has significantly higher level than the sustained tonal portion of the sounds. In these cases it is possible that the initial decay from the transient may be detected as an offset, unless appropriate detection suppression is included. Thirdly, there is a greater chance of offsets being obscured by other source sounds. As noted by McManus et al., the majority of musical notes contain most energy close to the onsets, and least at the end [13]. This means that there is a greater chance of an auditory offset than an auditory onset being obscured by other sound, such as noise or another note, due to the relatively lower level of the former. Finally, and perhaps most significantly, in most acoustical environments reflected energy (i.e. reverberation) will act to obscure the decay of the sound source. Depending on the characteristics of the source signal, the acoustical environment, and the relative positions of the source and receiver (or listener), there may be no variation in level, spectrum or phase at the receiver when the source stops producing sound, as the reverberation will continue.

The final factor listed above means that it is difficult, if not impossible, to use similar methods to onset detection to automatically determine the offset of a musical note. However, it also raises the question of how the auditory offset is perceived in this situation. It seems that little research has been undertaken into this topic, and understanding of the perception of auditory offsets in reverberation is still relatively basic.

An additional factor that may be utilised in the perception of auditory offsets is the change in the binaural cues caused by the direct sound ceasing to arrive at the receiver. For a musical note, the interaural cross-correlation usually settles at a relatively constant value, but then varies rapidly once the direct sound ceases, as discussed in [14]. Research has not yet been conducted into the perceptual relevance of this cue, and it may be a useful additional parameter in the detection of auditory offsets.

This section has identified a number of areas where further research is required into the perception and detection of auditory offsets. Based on this, a subjective experiment was conducted into the perception of auditory offsets in reverberant environments for musical signals. Objective analyses were then conducted of the stimuli used in the subjective experiment, to determine the characteristics of the signals that most closely matched the offset judgements.

2 SUBJECTIVE EXPERIMENT

A subjective experiment was conducted to investigate the effect of a number of factors on the perceived offset time of a number of stimuli, and the consistency of the perceived offset judgements. The factors considered were: presence or absence of reverberation; the source-receiver distance in an acoustical environment (and therefore the relative levels of the direct and reverberant sound); and the presence, absence or artificial manipulation of binaural cues.

2.1 Experiment stimuli

Three musical sounds were chosen for use in the experiment – a conga hit, a plucked guitar chord, and a cornet note. The conga hit was selected to represent a short transient signal with a rapid decay. The guitar chord was chosen to represent an initial transient attack followed by a gradual, almost exponential, decay, ending with the strings being muted. The cornet note has a slower attack, and was chosen for its more continuous level over the duration, with a gradual decay. These stimuli were anechoic recordings, obtained from the Archimedes Project CD of anechoic recordings as described in [15].

The direct anechoic stimuli were reproduced as diotic signals (i.e. identical in both channels over headphones). These were used as a comparison with the reverberant stimuli, and enabled analysis of the accuracy of the judgements of the perceived offset. The reverberant stimuli were generated by reproducing the anechoic signals over a loudspeaker in a reverberant environment and capturing the result using a Cortex MK2 head and torso simulator. The reverberant environment was a large recording studio with dimensions of approximately 14 x 17 x 8m, and a reverberation time of approximately 1.2 seconds. The capture was undertaken at two distances from the sound source – 1.5m and 5m – in order to investigate the effect of varying the direct-to-reverberant ratio.

It is possible that the difference between the diotic and binaural representation used for the direct and reverberant versions of the stimuli may have had an effect on the results. For this reason, and to investigate the effect of binaural cues on the perception of auditory offsets, the stimuli were also manipulated spatially. Three versions were made of the reverberant stimuli: diotic, binaural and exaggerated. The diotic reverberant stimuli were created by feeding the left and right channels similarly to both output channels. The binaural stimuli were unprocessed, with the left and right channels fed to the left and right output channels respectively. The exaggerated stimuli were spatially manipulated by increasing the correlation of the left and right channels during the segment of time when the direct sound was arriving at the receiver, and then reverting to the original binaural signal when the sound consisted of solely reflected energy. This had the effect of narrowing the image during the period that the direct sound was reaching the receiver, and widening the image during the reverberant (i.e. non-direct) sound segment. A summary of the conditions is contained in Table 1 below.

Number	Acoustical environment	Source / receiver distance	Binaural condition
1	Anechoic	N/A	Diotic
2	Reverberant	1.5m	Diotic
3	Reverberant	5m	Diotic
4	Reverberant	1.5m	Binaural
5	Reverberant	5m	Binaural
6	Reverberant	1.5m	Exaggerated
7	Reverberant	5m	Exaggerated

Table 1: Summary of the conditions generated for each musical extract, including presence or absence of acoustical environment, source/receiver distance, and binaural condition.

The combination of 7 conditions for each of the three musical extracts resulted in 21 stimuli for the experiment in total.

2.2 Experiment procedure

The experiment was conducted using a method-of-adjustment technique [16], where the subjects were asked to adjust the timing of a click so that it coincided with the perceived offset of each stimulus. The stimuli were presented individually, and were looped with a short pause between each repetition. The subjects could listen to the stimuli as often as they required, and it was up to them to decide when a match was good enough and that they were ready to move to the next stimulus.

The user interface was configured so that the subjects could adjust the time of the click relative to each test stimulus by adjusting a slider on the screen using the mouse. In order to avoid biases caused by order effects, the computer presented the stimuli in a random order. This also rendered the test double-blind (i.e. neither the subjects nor the experimenter knew which stimulus was which). In order to limit the chance of the subjects learning the duration of the stimuli and making their judgements based on the slider position rather than the sound, the delay time of the click after the start of each stimulus was modified by a random number, and this value was then removed prior to analysis of the data.

The method of adjustment procedure was chosen based on its time efficiency [17], high intra-subject repeatability [17], and the lower chance of boredom affecting the results due to the active role the subject has in the experiment [18]. By introducing a random modifier in the delay times, the chance of bias due to habituation and expectation was reduced [16]. As this method requires the subject to decide that the match is sufficiently accurate – a weakness of this method compared to others [19] – only subjects experienced in critical listening were used.

Twenty-five subjects undertook the experiment; most of these were final year undergraduate students in the Institute of Sound Recording at the University of Surrey.

3 SUBJECTIVE RESULTS

The results from the experiment were in the form of delay times after the start of each stimulus that the subjects judged to match the perceived offset time. These were then modified so that they were relative to the actual offset time of each stimulus; the actual offset time being judged as the point at which the anechoic stimulus decayed to the noise floor of the recording. The means and associated 95% confidence intervals calculated from these subjective data are shown in Figure 1. The results were examined for changes in the time of the perceived auditory offsets caused by the different experiment conditions, as well as changes in the variance in the results which may indicate the salience or the ambiguity of the perceived offsets.

3.1 Effect of the presence or absence of reverberation on the perceived offset

The stimuli consisted of anechoic recordings that were reproduced in a reverberant environment and captured using a head and torso simulator. Both the anechoic and reverberant stimuli were used in the experiment to compare the perceived offsets with and without reverberation.

For the bongo and cornet stimuli, the presence or absence of the reverberation had a statistically significant effect on the judgements of the perceived offset time. For both of these instruments, the addition of the reverberation caused subjects to judge the offset time to be significantly later than the actual offset time, regardless of the binaural condition or the direct to reverberant ratio. For the anechoic stimuli, the results were not statistically different from the actual offset time, and the mean values were within a few milliseconds of the actual offset time. For the reverberant stimuli, the perceived offset time had a mean value in the range of 90 to 170 ms after the actual offset of the note, which indicates that the reverberation obscures the actual note offsets and effectively lengthens the perceived note.

Contrary to this, the addition of the reverberation to the guitar chord caused a much smaller variation in the

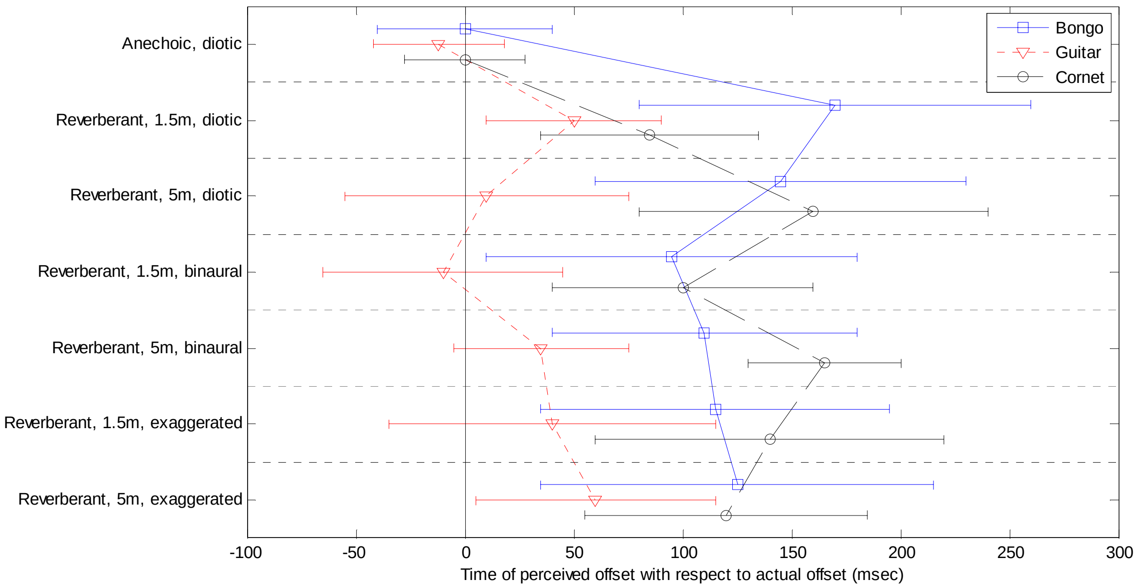

Figure 1: Plot of the means and associated 95% confidence intervals of the perceived auditory offsets for each stimulus with respect to the actual offset time.

perceived offset time. For the anechoic guitar stimulus, the perceived offset times were again not statistically significantly different from the actual offset time, and the mean value was within 15ms of the actual offset time. For the majority of the guitar stimuli with reverberation, the perceived offset times were still not statistically significant different from the actual offset time. Only two of the stimuli showed a statistically significant difference: the diotic reverberant stimulus with a 1.5m source-receiver distance; and the spatially exaggerated reverberant stimulus with a 5m source-receiver distance.

In all cases, the addition of the reverberation increased the variance in the offset judgements compared to the anechoic version. This indicates that the reverberation made the offset time more difficult to judge.

3.2 Effect of the direct to reverberant ratio on the perceived offset

The direct to reverberant ratio of the stimuli was varied by using two different source-receiver distances for all the reverberant stimuli. Comparing the results for the two source-receiver distances in each pair of otherwise similar stimuli, it is apparent that there were no statistically significant differences between the results for the two source-receiver distances. Therefore it appears that in this case changing the direct to reverberant ratio did not alter the time of the perceived offsets. There were also no consistent changes in the variance of the subjective results caused by changing the direct to reverberant ratio, which in this case indicates that altering the direct to reverberant ratio did not significantly enhance or degrade the perceived offset cues.

3.3 Effect of the binaural condition on the perceived offset

The reverberant stimuli had three different binaural conditions: diotic (the same signal in each ear); binaural (unprocessed recording from the head and torso simulator); and exaggerated (processed so that the image changes from artificially narrow to wide at the time of the actual offset). Comparing the results for the three binaural conditions in each triad of otherwise similar stimuli indicates that there were no statistically significant differences or consistent trends caused by changing the binaural conditions. Therefore it appears that in this case changing the binaural conditions did not alter the time of the perceived offsets. There were also no consistent changes in the variance of the subjective results caused by altering the binaural condition, which indicates that in this case the offset cues were not significantly degraded or enhanced by altering the binaural conditions.

3.4 The effect of the programme material on the perceived offset

There were three items of programme material used in the experiment: a bongo hit; a cornet note; and a guitar chord. As mentioned above, the results showed that the offset judgements of the guitar stimuli were less affected by the addition of reverberation than the bongo and cornet stimuli. This indicates that the perceived offset of the guitar chord is less obscured by the addition of reverberation than the other two musical signals. Objective analysis of the stimuli is required to investigate how the different items of programme material interact with the reverberant environment to give rise to this result.

4 OBJECTIVE MEASUREMENTS

In order to investigate the cues that were used by the subjects to judge the auditory offset times, a number of objective analyses were conducted of the stimuli that were used in the experiment. The comparison was simplified by taking the mean subjective results for all the reverberant conditions for each musical extract together, as there were no instances where these were statistically significantly different from each other. The mean subjective data used for the comparison are shown in Table 2, with the perceived offset times shown with respect to the measured onset time.

Stimulus	Mean perceived offset time (secs)
Bongo anechoic	0.16
Bongo reverberant	0.30
Guitar anechoic	2.65
Guitar reverberant	2.69
Cornet anechoic	3.44
Cornet reverberant	3.57

Table 2: List of the subjective data used for comparison with the objective measurements for the simplified stimulus set, with the perceived offset time taken as the mean value with respect to the measured onset time.

The objective measurements were simple time-energy-frequency analyses [20], and involved dividing the stimuli into three frequency ranges (100-400 Hz, 400-1600 Hz and 1600-4800 Hz), and measuring the signal level over time. The resulting signals were full-wave rectified, smoothed with a 10ms averaging filter for display, and normalised so that the measurements in each frequency band peaked at 0 dB for each stimulus. The plots from this analysis are shown in Figure 2 to Figure 4, overlaid with the means and associated 95% confidence interval data from the subjective results.

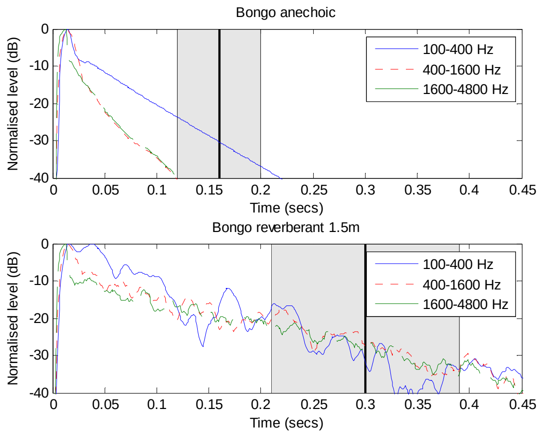

Figure 2: Simple time-energy-frequency analysis of the bongo stimuli, showing the normalised level of the three frequency bands of the anechoic version in the upper plot and the selected reverberant version in the lower plot, over time on the x-axis. The means of the subjective results are shown on each plot as a thick vertical black line, and the associated 95% confidence intervals are shown as a grey area.

For the bongo stimuli (shown in Figure 2), it seems that the perceived offset times from the experiment match closely with a point where the level in the low frequency range drops below approximately -30dB. It can be seen from the anechoic stimulus that the low frequency signal drops below this threshold at approximately 0.15s, and for the reverberant stimulus this threshold is crossed a number of times between 0.28 and 0.32s. For both stimuli, these objective results are close to the mean of the subjective results, and fall within the 95% confidence interval limits.

The results in the other frequency bands also show a delay in the offset time caused by the addition of reverberation, but the values in these cases range from 0.07s for the anechoic condition, to 0.32s for the reverberant condition. Therefore it can be seen that although the trend is the same, the match with the subjective data is inferior to the objective results derived from the low frequency range.

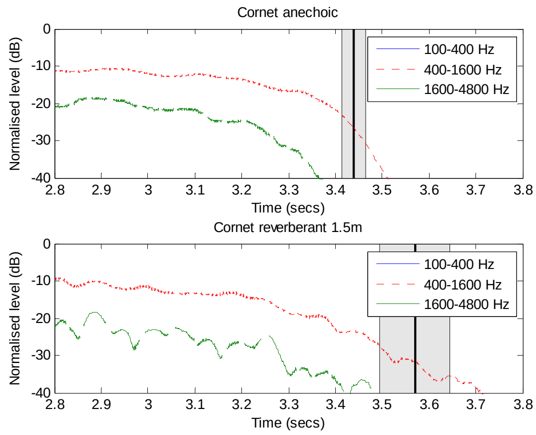

Figure 3: Simple time-energy-frequency analysis of the cornet stimuli, showing the normalised level of the three frequency bands of the anechoic version in the upper plot and the selected reverberant version in the lower plot, over time on the x-axis, scaled to focus on the note offset. The means of the subjective results are shown on each plot as a thick vertical black line, and the associated 95% confidence intervals are shown as a grey area.

The results from the cornet stimuli (shown in Figure 3) again follow a similar trend, in that a threshold of -30dB is a reasonable match to the subjective offset results. In this case, however, it is the mid frequency range that shows the best match to the subjective data, and the common trend with the previous results is that it is the slowest decaying frequency band (i.e. the last to drop below the threshold). The match is not as accurate as for the bongo stimuli, with the threshold measurement for the anechoic stimulus predicting an offset at 3.46s (compared to a mean of 3.44s for the subjective results) and the threshold measurement for the reverberant stimulus predicting an offset between 3.51 and 3.55s (compared to a mean of 3.57s for the subjective results). However, in all cases the prediction derived from the -30dB threshold falls within the 95% confident intervals of the subjective results.

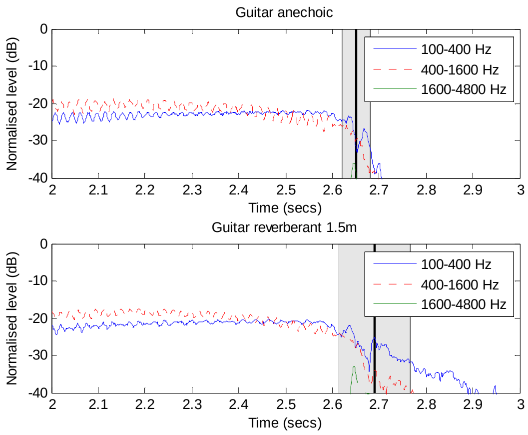

Figure 4: Simple time-energy-frequency analysis of the guitar stimuli, showing the normalised level of the three frequency bands of the anechoic version in the upper plot and the selected reverberant version in the lower plot, over time on the x-axis, scaled to focus on the chord offset. The means of the subjective results are shown on each plot as a thick vertical black line, and the associated 95% confidence intervals are shown as a grey area.

For the guitar stimuli (shown in Figure 4), the -30dB threshold is again a reasonable predictor for the perceived offset results, though it is a little more ambiguous compared to the bongo and cornet stimuli. For both the anechoic and reverberant conditions, the threshold is crossed a number of times: in the anechoic case this covers a period from 2.645s to 2.675s (compared to a mean of 2.65s for the subjective results) and in the reverberant case this covers a period from 2.67s to 2.77s (compared to a mean of 2.69s for the subjective results). Whilst the -30dB threshold results encompass the subjective mean values in both cases, for the reverberant guitar stimuli the range of objective results spread beyond the 95% confidence interval range of the subjective results. The match between the objective and subjective results could be improved by further smoothing or interpolation of the measured data, though this would be at the expense of temporal detail in the results.

A potential additional auditory offset cue can also be seen in the measurements of the guitar stimuli. The action of damping the strings at the end of the chord caused a click which can be seen as a transient peak in the high frequency band of the measurements of both the anechoic and reverberant stimuli. In both cases this peak falls within the 95% confidence intervals of the subjective results, and it is possible that this was used by some of the subjects to judge the offset time of these stimuli, which could explain why the offset times for the reverberant stimuli were not statistically significantly different from either the actual offset time or the subjective results for the anechoic stimulus.

5 DISCUSSION

This paper described a subjective experiment that was conducted to investigate the perceived offset time of three diverse musical signals, with the following conditions:

- presence or absence of reverberation
- variation of direct to reverberant ratio
- presence, absence or exaggeration of binaural cues.

It was found that the perceived offset time of the anechoic stimuli matched the actual offset time, and that for two of the three musical signals the presence of the reverberation caused the perceived offset time to be delayed, effectively lengthening the notes. The variations in the direct to reverberant ratio and the binaural conditions were not found to make statistically significant differences in this experiment. It is likely that if the direct to reverberant ratio is decreased further there will become a point where the offset will be difficult to perceive based on variations in level. In this situation, it is still possible that the variation in the interaural cross-correlation will act as useful binaural cue for perceiving the offset. Further research is required to investigate this.

Simple time-energy-frequency analysis was undertaken of the experiment stimuli, and the results of this were compared with the subjective results. It was found that the perceived offset judgements from the subjective experiment correlated with the time at which the slowest decaying frequency range of each stimulus dropped 30dB below its peak level. It is recognised that this is a relatively simplistic analysis method that appears to be effective for simple single stimuli with a low background noise level. Further research is required to develop and test this method using more complex situations such as differing levels of background noise or polyphonic musical extracts. In particular, it would be useful to conduct a detailed investigation of the properties of signals that act to obscure offsets of various stimuli, as this would help to more precisely uncover the cues that are used by the auditory system in to detect offsets in difficult listening situations.

REFERENCES

[1] A. S. Bregman, *Auditory Scene Analysis: the perceptual organisation of sound*, MIT Press (1990)

[2] E. D. Scheirer, “Tempo and beat analysis of

acoustic musical signals", *Journal of the Acoustical Society of America*, vol. 103, pp. 588-601 (1998).

[3] X. Rodet and F. Jaillet, "Detection and modelling of fast attack transients", *Proceedings of the International Computer Music Conference*, pp. 30-33 (2001).

[4] J. P. Bello and M. Sandler, "Phase-based note onset detection for music signals", *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-03)*, pp. 275-280 (2003).

[5] L. S. Smith and D. S. Fraser, "Robust sound onset detection using leaky integrate-and-fire neurons with depressing synapses", *IEEE Transactions on Neural Networks*, vol. 15, pp. 1125-1134 (2004).

[6] A. Klapuri, "Sound onset detection by applying psychoacoustic knowledge", *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Proceesing (ICASSP-99),* pp. 115-118 (1999).

[7] J. P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. B. Sandler, "A Tutorial on Onset Detection in Music Signals", *IEEE Transactions on Speech and Audio Processing*, vol. 13, pp. 1035-1047 (2005).

[8] D. Griesinger, "The psychoacoustics of apparent source width, spaciousness and envelopment in performance spaces", *Acustica*, vol. 83, pp. 721-731 (1997).

[9] D. Luce and M. Clark, "Duration of attack transients of nonpercussive orchestral instruments", *Journal of the Audio Engineering Society*, vol. 13, pp. 194-199 (1965).

[10] N. Collins, "A comparison of sound onset detection algorithms with emphasis on psychoacoustically motivated detection functions", *118th Audio Engineering Society Convention*, preprint no. 6363 (2005).

[11] H. Thornburg and F. Gouyon, "A flexible analysis-synthesis method for transients", *Proceedings of the International Computer Music Conference*, pp. 400-403 (2000).

[12] S. A. Abdallah, and M. D. Plumbley, "Probability as metadata: event detection in music using ICA as a conditional density model", *Proceedings of the 4th International Symposium on Independent Component Analysis and Signal Separation*, pp. 233-238 (2003).

[13] J. A. McManus, C. Evans and P. W. Mitchell, "The dynamics of recorded music", *95th Audio Engineering Society Convention*, preprint no. 3701 (1993).

[14] R. Mason, T. Brookes and F. Rumsey, "The effect of various source signal properties on measurements of the interaural cross-correlation coefficient", *Acoustical Science and Technology*, vol. 26, pp. 102-113 (2005).

[15] V. Hansen and G. Munch, "Making recordings for simulation tests in the Archimedes project", *Journal of the Audio Engineering Society*, vol. 39, pp. 768-774 (1991).

[16] G. A. Gescheider, *Psychophysics, the fundamentals*, Erlbaum (1997).

[17] A. Hesse, "Comparison of several psychophysical procedures with respect to threshold estimates, reproducibility and efficiency", *Acustica*, vol. 59, pp. 263-273 (1986).

[18] B. L. Cardozo, "Adjusting the method of adjustment: SD vs DL", *Journal of the Acoustical Society of America*, vol. 37, pp. 786-792 (1965).

[19] C. C. Wier, W. Jesteadt and D. M. Green, "A comparison of method-of-adjustment and forced-choice procedures in frequency discrimination", *Perception and Psychophysics*, vol. 19, pp. 75–79 (1976).

[20] A. Mertins, *Signal analysis: wavelets, filter banks, time-frequency transforms, and applications*, Wiley (1999).