Quantitative Analysis of Sound in a Short Horror Film

Nick Redfern
Independent scholar
Leeds, United Kingdom
E-mail: nickredfernres@outlook.com

Abstract:

In this article I analyse the structure of sound in a short horror film *Behold the Noose* (Jamie Brooks, 2014), applying quantitative methods in order to determine the structure of the film's soundtrack and to understand how sounds at different frequencies create an emotional experience for the viewer. Examination of the normalised aggregated power envelope shows the soundtrack is organised at different scales, from local transient events to medium scale sections easily distinguishable from those that precede and follow, and a step change between large scale segments. In this film, sections of increasing tension are associated with non-linear increases in sound energy. Using the short-time Fourier transform I show that sounds at different frequencies are associated with different emotional effects in the viewer and that the range of frequencies used increases as the emotional effect of the soundtrack intensifies.

Keywords: Computational film analysis, digital humanities, horror cinema, sound design, normalised aggregated power envelope, short-time Fourier transform, Behold the Noose

1. Introduction

Sound pays a key role in creating a frightening experience for the viewer in horror cinema (Grøn 2013). William Whittington writes that "in general, horror films use music and sound effects to stablish emotive intensity and impact far more aggressively and conceptually than any other genre" (Whittington 2007, 130); and Ivanka Pavlovic′ and Slobodan Markovic′ (2011) report that visual information in the cinema has a stronger emotional impact than auditory information, with the exception of the emotion of fear. The use of stylised sound effects makes the everyday strange, transforming the traditional role of sound effects from creating a sense of realism and coherence to one of estrangement where the audience is kept off-balance (Bullins 2013). As sound designer Gregg Rudolff states, "Even on something like a door close, I might add more low end. I may want to have a little weirdness to it. I might add a little flange or pitch bend ... it's not a traditional sound the audience is comfortable with" (Farinella 2007, 51). Localised sound events are linked to specific moments terror: acoustic blares elicit startle responses from the viewer to make them jump (Sbravatti 2019), while the slow attack of the soundtrack as it builds to a crescendo creates a feeling of apprehension (Moncrieff, Dorai, and Venkatesh 2002). Music plays a key role in shaping emotional responses, as Neil Lerner points out: "Frightening images and ideas can be made even more intense when accompanied with frightening musical sounds, and music in horror film frequently makes its audience feel threatened and uncomfortable through its sudden stinger chords and other shock effects" (Lerner 2010, ix). The result is that the sound design of a horror film is able to manifest a "direct emotion and a primary psychology" (Donnelly 2009, 121).

Despite the central role it plays in the viewer's experience of watching a film, the analysis of sound in the cinema is not yet a significant part of the digital humanities. Applications of digital humanities approaches to questions of film style and film form have focused on editing (Heftberger 2018), cinematography (Kovács 2014), scriptwriting (Murtagh, et. al. 2009), and the narrative structure of films (Cutting 2016), but not sound. This is surprising given the soundtrack to a film can be accessed directly through the use of non-linear editing software and is readily available for processing and analysis, along with the easy availability of analytical tools and methods developed in the fields of computational music analysis, sound studies, and the natural sciences (e.g. acoustic ecology, bioacoustics, etc.) that can be applied to understanding film audio.

Audio segmentation comprises a set of techniques for analysing the features of audio signals, including motion picture soundtracks. Parsing the structure of film audio allows us to identify scenes and change points, features in the audio envelope (attack, sustain, release), the distribution of sound energy, and the presence of affective events in a soundtrack. This approach has been used to detect narrative structure via audio pace (Moncrieff and Venkatesh 2007), exploring temporal coincidences between the visual and audio in montage editing (Zeppelzauer, et. al. 2011), and to identify changes in the soundscapes of motion pictures (Orio 2013). In this paper I apply audio segmentation to analyse the structure of sound in *Behold the Noose* (Jamie Brooks, 2014), a short horror film about a sheriff's deputy who arrives at a mysterious farmstead while searching for a missing girl. Applying quantitative methods, I use the short-time Fourier transform and the normalised aggregated power envelope to understand how the frequencies of the different noises used and the temporal evolution of sound energy in motion pictures create an emotional experience for the audience. Quantitative methods have previously been applied to horror film soundtracks in the context of multimedia content analysis and indexing: Moncrieff, et. al. (2002) applied quantitative methods to affective sound energy events in horror films, identifying specific patterns of sound energy dynamics correlated with particular affective events (e.g. startle, sustained alarm, building tension) in horror films and were able to automatically distinguish between horror and non-horror films based on the presence absence of these dynamic events in a film's soundtrack. This article is the first to adopt a quantitative approach to understanding how sound functions in a horror film and demonstrates how the digital humanities can be expanded to include the analysis of film soundtracks.

2. Methods

I use two methods to analyse the soundtrack of *Behold the Noose*: the short-time Fourier transform and the time contour of the normalised aggregated power envelope. These methods are implemented using Python (v.3.5.1) and the librosa package (v0.4.0; McFee, et al. 2015). The methods used in this paper are described in detail in Redfern (n.d.), which also includes examples of the code used.

I downloaded *Behold the Noose* from YouTube as an mp4 file (24 fps) and using a non-linear editing suite exported the film's audio as a mono 16-bit wave file sampled at 22.05 kHz. This audio file was then normalised to a peak volume of 0.0 dB using Audacity (v. 2.0.6) and re-exported as a mono 16-bit PCM wave file (Fig. 1. (a)).

I calculated the short-time Fourier transform (STFT) to the wave to produce a 2D time-frequency representation of the signal called a spectrogram (see Goodwin 2008). A signal as a function of time (such as the sound wave in Fig. 1. (a)) can be thought of as the linear combination of a series of sine waves. The Fourier transform is a method for decomposing a signal and expressing it in terms of its component sinusoidal waves, and is applied globally assuming the signal does not change over time (i.e. that the signal is *stationary*). This process cannot be applied directly to film soundtracks because they are non-stationary signals, their properties (amplitude, spectral composition, etc.) varying with time. The STFT is a commonly used tool for analysing such a signal, dividing the signal into windows and calculating the Fourier transform for each window. The result is a Fourier transform of the signal localised in time dependent upon the shape (rectangle, Hann, etc.), size (the number of samples within a window), and overlap of the window used. The window size also controls the trade-off between time and frequency resolution: the more precise the time resolution the less precise the frequency resolution, and *vice versa*. The spectrogram describes how the magnitudes of the individual frequencies comprising a signal vary over time, with the power spectral density indicated by colour.

Fig. 1. (b) shows the log amplitude spectrogram of the soundtrack of *Behold the Noose* using Hann windows with a length of 2048 samples and hop length of 512. The log amplitude spectrogram is used to display the STFT of the soundtrack because most of the information in the film's soundtrack is in the range 1 Hz to 3 kHz making it easier to visualise details at lower frequencies. The number of short-time spectra in Fig. 1. (b) is equal to the product of the running time of the film and the sampling rate divided by the hop length while the temporal precision is equal to the hop length divided by the sampling frequency. This results in (568 × 22050) / 512 = 24464 short-time spectra at intervals of 512 / 22050 = 0.0232s. Each short-time spectrum in Fig 1. (b) is a Fourier transform with a length equal to half the window length (2048 / 2 = 1024).

Summing the power spectral density values in each of the short-time spectra in the STFT of the signal in Fig. 1. (a) produces an aggregate power envelope, which is then normalised to a unit area and treated as a probability mass function. The resulting plot shows how the energy of a signal evolves over time. I calculated the power envelope using the original STFT rather than the log amplitude spectrogram in Fig. 1. (b) because, while summing the spectra in the log amplitude spectrogram produces the same time contour plot of the normalised aggregated power values, the resulting plot is rendered with a linear y-axis that is potentially misleading. I used LOWESS regression to fit a trendline to the envelope, and Fig. 1. (c) presents the time-contour plot of the normalised aggregated power envelope of *Behold the Noose* and its trendline.

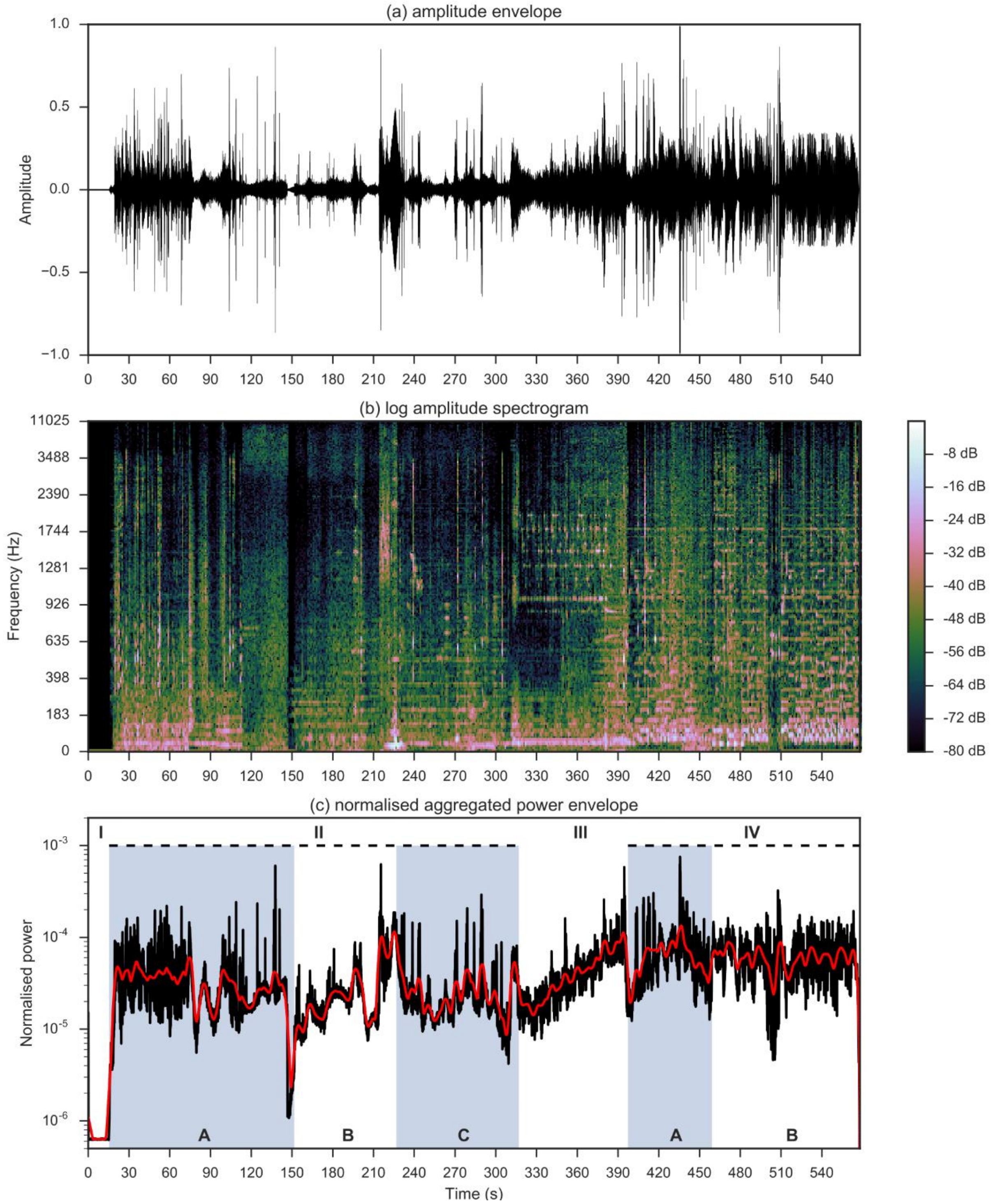

Fig. 1. The soundtrack of *Behold the Noose* (Jamie Brooks, 2014): (a) the amplitude envelope of the signal as a 16-bit mono PCM wave sampled at 22050 Hz normalised to a peak volume of 0.0 dB; (b) log amplitude spectrogram of the short-time Fourier transform of the wave form, with a Hann window of length 2048 samples and a hop length of 512; (c) the normalised aggregated power envelope with fitted LOWESS trendline.

3. Sound in *Behold the Noose*

From the short-time Fourier transform in Fig. 1. (b) and the normalised aggregated power envelope in Fig. 1. (c) we can identify a range of features in the film's soundtrack. In this section I apply these methods to the soundtrack of *Behold the Noose* to describe how the temporal evolution of sound energy in the film and the use of different sounds at different frequencies create emotional effects in the viewer.

3.1 Structure and evolution

Analysis of the normalised aggregated power envelope allows us to describe the large-scale temporal structure of *Behold the Noose* and to segment the audio and narrative structure of the film. It also allows us to understand the nature of the temporal evolution of sound energy for specific narrative features in the film.

Table 1: The structure of *Behold the Noose* based on segmentation of the normalised aggregated power envelope of the film's soundtrack

Segment	Section	Time (s)		Action
		Start	End	
I		0.0	15.4	Title card
II	A	15.4	151.7	The deputy is despatched to search for a missing girl and arrives at a farmhouse.
	B	151.7	226.9	Deputy exploring the grounds; a body in a shed is revealed to the audience.
	C	226.9	317.0	The deputy continues with his search until he arrives at the main house.
III		317.0	397.3	Deputy enters and searches the house, discovering the killer's "shrine."
IV	A	397.3	459.2	The deputy discovers the "hanging tree" and is stabbed by the killer.
	B	459.2	568.0	The deputy's body is added to the "hanging tree" and the missing girl is declared safe. Credit sequence.

There are four main segments in the film and segments II and IV can be further divided into three and two sections, respectively (see Table 1). This segmentation of the film is based on the structure of the film's audio identifiable from the LOWESS trendline in Fig. 1. (c) and the statistical summaries of the different sections derived from the normalised aggregated power envelope. The overall pattern of the evolution of the film's sound energy is a step function change from the slow build-up of tension in the second part of the film to the violence of the final segment linked by a transitional phase (segment III). There is a clear change in the level of sound energy: the median value of the normalised aggregated power envelope in Fig. 1. (c) is 2.19×10^{-5} for segment II and 5.41×10^{-5} for segment IV. The median level for segment II is consistent with the early part of segment III, while the median level of segment IV is consistent with the sustain section of the sound envelope of segment III. It is the gradual shift between these levels that marks out segment III as a transitional phase of the film's soundtrack. The spread of sound energy in segment IV (IQR = 4.19×10^{-5}) is also higher than the spread in segment II (IQR = 1.69×10^{-5}). Finally, the cumulative proportion of energy in segment II is just over a third of the film's total (39.1%) even though at 301.6 seconds this segment accounts for over half the film's running time (53.1%), whereas the cumulative proportion in segment IV (which at 170.7s is 30.1% of the total running time) is just under half the total energy (45.3%).

Segment I comprises a title attesting to the veracity of events in the movie and advising viewer discretion, but there are no sounds during this part of the film.

The second segment of the film comprises three distinct sections. Section II.A establishes the film's premise as the Sheriff's deputy is called over the radio to assist in the search for a missing girl and drives to the farmhouse. On arriving, he exits the vehicle and retrieves his shotgun before heading off to search the premises. The sounds in this part of the film include the voice of the deputy and the dispatcher on the radio, the noises of the vehicle (the engine while driving and the parking brake on arrival), while the atmosphere of the sequence is created through a combination of music and crickets chirping in the background. Throughout this section the sound energy level remains even, though there are changes in the spread when no-one is talking (e.g. from 75.0s to 100.0s).

Section II. B begins in silence as we see the deputy walking on a CCTV monitor, though there is no indication of who is watching. There is little dialogue in this part of the film (the deputy uses his radio twice), and the effects used comprise the deputy's footsteps and the chirping of crickets. Within this section there is a series of individual peaks, each associated with an increase in emotional intensity as the deputy enters the yard (163.0s), turns a corner into the unknown (183.5s), discovers the bloody sheet (196.5s), and, finally, at 216.0s the deputy opens a door to release a flock of cawing birds before the final "true shock" of this sequence is revealed – a body hanging in the shed (the peak at 225.3s). These last two features produce a "double peak" in sound energy.

The hanging body is shown to the audience, but the deputy is unaware of its presence as we move into section III. C and he continues his search of the grounds. In this section the film returns to the relatively even sound energy level of section II. A with the voice of the deputy as he calls out and the crickets in the background. There is no music in this section to speak of, with the atmosphere created by a mixture of wind effects, throbbing noises, and ghost-like noises.

Segment III begins as the deputy enters the farmhouse. The sound energy in the early part of this sequence continues at the level of section II. C, but the sounds themselves have changed. As in section II. B there is little dialogue until the deputy radios for help (384.6s) and there are few effects aside from the buzzing of flies in the kitchen, which is eventually replaced by a similar non-diegetic noise that eventually distorts. The principal musical element of the soundtrack is the rise and fall of a childish melody clearly visible by the pattern in Fig. 1. (b), and the energy of this tune also increases throughout the sequence until it starts to distort as the deputy makes his way through the farmhouse. The sound energy of this sequence increases from 327.9s as the deputy moves deeper into the house to 378.3s, when he discovers the killer's shrine to his victims – a scene that is first revealed to the audience with a shot of a skull in a jar. This increase in sound energy is related to the spatial experience of the film's "terrible place:" as the deputy moves further into the house the energy of the soundtrack increases until peaking at the moment where he is furthest from safety and confronted with his terrible discovery. The peak level is sustained until 397.3s as the deputy surveys the photographs in the shrine and begins to panic, desperately radioing for help. The sound energy then rapidly decays to the level at which this segment began after a banging noise and laughter is heard off-screen and the deputy turns to exit the building. At 395.1s the song "Witch Lynching" by Forever and Everest (2012) begins and dominates the remainder of the film's soundtrack.

The final segment of the film contains all the violence of the film and comprises two distinct audio sections. In section IV.A the deputy makes his way outside the farmhouse to discover numerous bodies hanging from a tree. The energy of the first part of this sequence (397.3s-421.5s) increases as the song "Witch Lynching" takes over the soundtrack and the deputy calls out into the darkness, reaching a peak as the deputy gasps in shock at the sight of the hanging tree. The sound energy of this feature does not immediately decay and is sustained from 421.5s until the deputy is stabbed is by the killer and his shotgun fires. From 437.0s the energy of the soundtrack decays until 459.2s as the blade is driven deeper into the deputy's body and he finally falls to the floor. The volume of the song is reduced to bring the body horror of the deputy's goring to prominence and the soundtrack emphasises the sounds of the blade twisting in the deputy's body and his final, gasping breaths. Section IV.B begins with the rapid attack of the trumpets from the song and returns to the level of section IV.A, which is then sustained until the end of the film. The action in this part of the film sees the deputy's body dragged through the words and hanged before cutting to the deputy's vehicle as the dispatcher comes over the radio to announce the missing girl is safe and sound. The sound energy drops to a low at 505.0s as the dispatcher's voice fades and the volume of the music reduces before the sudden attack of the whistling refrain of "Witch Lynching" returns as flames burst from the farmhouse's chimneys to illuminate the night sky. After this the credits roll over the remainder of the song.

The overall structure of the soundtrack is immediately clear from Fig. 1. (c) and it is evident that the sound design of the film is organised at a relatively large scale: each section of the film is associated with a particular temporal evolution of sound energy. For example, the time contour plot in Fig. 1. (c) reveals a key feature of the sound mixing in sections II. B, III, and IV. A not previously identified in studies of sound in horror cinema: in each case the attack of the sound event increases slowly in the first part of the sequence before accelerating rapidly to a climax and follows a *non-linear crescendo* pattern. The decay of the sound event in section IV.A (437.0-459.2s) is also non-linear as the deputy succumbs to his fate. Smaller-scale sound events exist within these section-level events. For example, section II. B features some local events (the individual transient peaks) within a non-linear increase in sound energy that runs for the whole length of the sequence. In this section, each peak has a higher level of sound energy than the last and so we should not see these shocks as existing isolation but as part of a single dynamic structure producing a cumulative effect in the viewer. Because the sound energy does not fall back to the level prior to the peak these small-scale events contribute to the overall non-linear increase of the whole section.

A feature of horror film soundtracks noted by several scholars is the use of "assaultive blasts that coincide with shock or revelation" (Lerner 2010, ix) or *stingers* to produce emotional effects in the audience. Stingers take the form of noises such as screams, orchestral music, or sound effects characterised by their suddenness, their stabbing shortness, and a sudden increase in volume to produce a startle effect (Hutchings 2004, 134-137; Baird 2000; Sbravatti 2019) that triggers a basic reflex of shock or surprise resulting from the collision of loud and soft sounds (Whittington 2014):

> The fact that these scenes are often equally effective underscores the significance of *contrast* between a relatively *loud* sound bursting into a relatively *quiet* scene. In gestalt theoretical terms: a figure that stands off perceptibly from a ground (Hanich 2010, 134; original emphasis).

Hanich (2010, 134) claims that the crescendo culminating in a short-sound stab is a less effective method of producing the desired audience response than a sudden, unexpected increase in volume, though he offers no empirical support for this claim and does not specify the form of the crescendo. Analysis of the dynamics of sound energy in *Behold the Noose* shows this does not adequately account for the evolution of sound energy over time. The evolution in sound energy in section II. B occurs at both the local scale of the transient events and at the medium scale of the section. Consequently, the difference between "figure" and "ground" is not clear-cut: the peaks (or "figures") associated with each event are part of the larger structure of the soundtrack evolving at a higher scale and contribute to the overall dynamic of the section as well as creating transitory moments of heightened tension. Those small-scale moments cannot be separated from the larger scale of the crescendo and so the claim that one method is more effective than another cannot be justified. When analysing horror film soundtracks it is therefore necessary to examine not only those moments of extreme horror but also their place within the larger structure of the soundtrack.

3.2 Frequency analysis

Close analysis of the short-time Fourier transform in Fig. 1. (b) reveals that sounds at different frequencies are used to create different emotional effects in the viewer and that the range of frequencies used increases as the emotional effect of the soundtrack intensifies.

From Fig. 1. (b) we see that most of the energy in the soundtrack of *Behold the Noose* is concentrated in the low frequency part of the audio spectrum. Specifically, the energy of the soundtrack is concentrated in the sub-bass (20 Hz-60 Hz), and bass (60 Hz-250 Hz) ranges. (For descriptions of frequency ranges see Owsinksi 2006: 25-26). The function of these tones is to create and maintain a persistent sense of unease and dread throughout the sequence and, although frequencies in this part of the spectrum increase in loudness during moments of intense emotion, they are not necessarily characteristic of those moments. Emotionally intense moments of horror in those scenes in which the deputy explores the "terrible place" are associated with frequencies in the mid-range (250 Hz-4 kHz) and occasionally at frequencies in the treble range (greater than 4 kHz).

For example, in section II. B sub-bass and bass range frequencies are a constant presence throughout the sequence, and though the amplitude varies these frequencies are not associated with specific actions in the film. It is the synchronisation of mid-range frequency tones with key moments of the film's action that create specific moments of heightened terror in this section. These mid-range tones arise prior to the deputy turning a corner as he explores the grounds and they occur with a shot of what appears to be a bloody sheet and with the escape of the birds that provides the "false" shock peaking at 216.0s. The peak in the normalised aggregated power of section II. B occurs when the body is revealed (225.3s) and shows increased energy in three frequency ranges: the introduction of a heartbeat effect with energy concentrated in the infrasonic and sub-bass ranges, the persistent bass hum that begins early in the film, and the expansion of the higher frequency tones to a range of 500 Hz-10 kHz. Fig. 1. (b) shows this in the increased amplitude at 216.0s and 225.3s across a much broader range of frequencies than those moments in the film immediately preceding and following.

The soundtrack in section III comprises a very low frequency hum in the range 50 Hz-80 Hz that continues throughout the sequence, intermittent low-frequency tones in the range 100 Hz-200 Hz, and a childish melody in the range 1 kHz-2 kHz. Over the course of this segment the childish melody increases in energy, becoming louder at 340.0s and again at 358.0s, eventually distorting as the deputy makes his way through the house and gets closer to the shrine. In Fig. 1. (b) we can also see harmonic tones in the range 2 kHz-7 kHz increasing in frequency until the end of this sequence and which contribute a high-pitched buzzing sound to the scene. A key moment in the sound of this sequence occurs at 378.3s when the audience is shown a close-up of a skull from the deputy's point-of-view. Here the soundtrack heightens emotion by introducing a ghost-like hollow tone that uses a much wider range of frequencies in the low end than in earlier parts of the segment (1 Hz-900 Hz) as well as increasing the amplitude of frequencies in the mid- and treble ranges. The sequence ends with a banging noise off-screen and the sound of laughter at 394.5s, also using the full range of frequencies in the sample. It is this noise that draws the deputy outside to his fate.

As noted above the energy of the soundtrack increases as the emotional effect intensifies. In part this is due to increases in the loudness of sounds present but it also due to an increase in the range of frequencies used at moments of emotional intensity. Furthermore, the greater the range of frequencies used the more intense the emotional shock. The use of sounds with frequencies greater than 4 kHz in the most emotionally intense moments of sections II. B and section III is interesting as these frequencies are classed as "presence" (4 kHz-6 kHz) and "brilliance" (6 kHz-20 kHz). "Presence" refers to the listener's experience spatial relationship with a sound source: boosting sounds in this frequency makes the source feel close by while reducing sounds at ~5 kHz makes sounds seems distant. The "brilliance" range is composed of harmonics and although we do not perceive these as individual sounds they add to the timbre of a sound event and shape our emotional involvement in a scene. For the buzzing noises in section III this creates the aural sensation of being in the room with rather than simply watching the deputy move through the house. As this scene progresses the sounds move from the diegetic (flies in the filthy conditions of the farmhouse) to the non-diegetic (simulated and distorting buzzing noises) and increase in both energy and frequency to create an intense emotional effect.

4. Conclusion

This paper demonstrates the use of quantitative methods for analysing the structure of sound in horror cinema using the short-time Fourier transform and the normalised aggregated power envelope. These methods have several features to recommend them for analysing the structure of sound in motion pictures. First, they are easy to compute using freely available software, though the computing power and the time required for analysing feature film soundtracks will be high. Second, they have descriptive power, communicating detailed information about a film's soundtrack in a straightforward manner that is easy to interpret. Third, they have analytical power supporting a bottom-up approach to analysing film style that allow the researcher to identify interesting features in a soundtrack and then to look beneath them in order to understand what is going on here.

Applying these methods to a short horror film I identified a range of features to provide a detailed understanding of how sound functions in *Behold the Noose*. The normalised aggregated power envelope allows us to segment the film and define its structure at different scales. At the micro scale individual moments of horror in the film stand out, while at the medium scale each section of the film can be easily distinguished from those that precede and follow it based on the time contour plot. A key feature revealed by this method is the non-linear sound mixing in the sequences in which the deputy searches the film's "terrible place" that builds tension to a climax to create an emotional response in the spectator. This is an original result not previously identified in studies of horror film soundtracks. This feature is not identifiable from simply listening to the film's soundtrack or from examining either the waveform or the short-time Fourier Transform of the soundtrack; but is immediately apparent in the normalised aggregated power envelope, demonstrating the value of this method. Finally, at the macro scale the evolution of the sound energy is a step function change between the set-up of the narrative and early, transient scares and the sustained emotional intensity and bloody violence of the deputy's murder.

In *Behold the Noose* emotional intensity is related not only to changes in sound energy but also to the range of frequencies used. Heightened moments of anxiety and horror in the film use a wider range of frequencies, particularly increasing the sound levels in the mid-range and above, in order to increase the audience's involvement in a scene through enhancing the presence and brilliance of the soundtrack. Using methods such as the STFT we gain a complete picture of what sounds are present in the film and thereby form an integrated understand of a film's soundtrack.

In doing so we will be better able to explain how horror films effect audiences physiologically and psychologically.

Filmography:

Behold the Noose (Jamie Brooks, USA, 2014, 9:28 min), https://www.youtube.com/watch?v=OhlDW9_Ehik, accessed 6 July 2020.

References:

Baird, Robert. "The Startle Effect: Implications for Spectator Cognition and Media Theory." *Film Quarterly* 53. 3 (2000): 12-24.

Bullins, Jeffrey. "Hearing the Game: Sound Design in the *Saw* Film Franchise." *To See the Saw Movies: Essays on Torture Porn and Post-9/11 Horror*. (Eds.) James Aston and John Wallis. Jefferson, NC: McFarland, 2013. 176-193.

Cutting, James E. (2016). "Narrative Theory and the Dynamics of Popular Movies." *Psychonomic Bulletin & Review* 23. 6 (2016): 1713–1743.

Donnelly, Kevin J. "*Saw* Heard: Musical Sound Design in Contemporary Cinema." *Film Theory and Contemporary Hollywood Movies*. (Ed.) Warren Buckland. New York: Routledge, 2009. 103-123.

Farinella, David John. "Horror Audio: Mixing Mayhem can be Fun." *Mix* (1 September 2007): 51-52. https://www.mixonline.com/sfp/horror-audio-369182, accessed 6 July 2020.

Forever and Everest. "Witch Lynching." *Filthy Songs & Dirty Stories* (2012) https://foreverandeverest.bandcamp.com/track/witch-lynching, accessed 6 July 2020.

Goodwin, Michael M. "The STFT, Sinusoidal Models, and speech Modification." *Springer Handbook of Speech Processing*. (Eds.) Jacob Benesty, M. Mohan Sondhi, and Yiteng Arden Huang. Berlin: Springer, 2008. 229-258.

Grøn, Nis. "The Sound of Horror: Silence & Sound Contrasts in Sci-Fi Horror Movies." *Tidsskrift for Medier, Erkendelse og Formidling* 1. 1 (2013): 97-109.

Hanich, Julian. *Cinematic Emotion in Horror Films and Thrillers: The Aesthetic Paradox of Pleasurable Fear*. London: Routledge, 2010.

Heftberger, Adelheid. *Digital Humanities and Film Studies: Visualising Dziga Vertov's Work*. Cham: Springer, 2018.

Hutchings, Peter. *The Horror Film*. Abingdon: Routledge, 2004.

Kovács, András Bálint. "Shot Scale Distribution: An Authorial Fingerprint or a Cognitive Pattern?" *Projections* 8. 2 (2014): 50-70.

Lerner, Neil. "Preface: what about horror's ear?" *Music in the Horror Film: Listening to Fear*. (Ed.) Neil Lerner. London: Routledge, 2010: viii-xi.

McFee, Brian, et. al. *librosa: v0.4.0*. Zenodo, (2015). http://dx.doi.org/10.5281/zenodo.18369, accessed 6 July 2020.

Moncrieff, Simon, et. al. "Determining Affective Events Through Film Audio." *Media Computing: Computational Media Aesthetics*. (Eds.) Chitra Dorai and Svetha Venkatesh. Boston, MA: Kluwer Academic Publishing, 2002: 131-157.

Moncrieff, Simon and Svetha Venkatesh. “Film Audio Pace.” *International Journal of Intelligent Systems Technologies and Applications* 3. 3-4 (2007): 296-308.

Murtagh, Fionn, , et. al. "The Structure of Narrative: The Case of Film Scripts." *Pattern Recognition* 42. 2 (2009): 302-312.

Orio, Nicola. “Soundscape Analysis as a Tool of Movie Segmentation.” *Cinergie* 3 (2013): 157-163.

Owsinski, Bobby. *The Mixing Engineer’s Handbook*, 2nd edn. Boston: Thomson Course Technology, 2006.

Pavlovic´, Ivanka and Slobodan Markovic´. “The Effect of Music Background on the Emotional Appraisal of Film Sequences.” *Psihologija* 44. 1 (2011): 71–91.

Redfern, Nick. “Computational Analysis of a Horror Film Trailer Soundtrack with Python.” (n.d.), https://www.academia.edu/43289938/Computational_analysis_of_a_horror_film_trailer_soundtrack_with_Python, accessed 6 July 2020.

Sbravatti, Valerio. “Acoustic Startles in Horror Films.” *Projections* 13. 1 (2019): 45–66.

Whittington, William. *Sound Design in Science Fiction Films*. Austin, TX: University of Texas Press, 2007.

Whittington, William. “Horror Sound Design.” *A Companion to the Horror Film*. (Ed.) Harry M. Benshoff. Chichester: John Wiley & Sons, 2014: 168-185.

Zeppelzauer, Matthias, et. al. “Cross-modal Analysis of Audio-Visual Film Montage.” *Proceedings of 20th International Conference on Computer Communications and Networks*. Red Hook, NY: Curran Associates, Inc, 2011. 1-6.