



REVERSE ENGINEERING EMOTIONS IN AN IMMERSIVE AUDIO MIX FORMAT

Sreejesh Nair

Solutions Specialist, Audio, Avid
Re-Recording Mixer

ABSTRACT

The idea of immersive mixing is not new. Yet, the concept of adapting it to achieve an emotional story telling and audience control is still something that can be pushed forward. Looking at the developments happening in the world of cinema and home audio, the idea of bringing the spectator to engage with the story is now a reality. The Idea is to propose various techniques of mixing and audio processing along with the emotional impact into this world so that the artistic intention of the creator with the viewer is bridged and rather than having an audience that sees a movie, it is time to have the audience experience the story. It will span from the art of psychological placements of sounds to designed and interpreted spaces that replicate a human emotion when replayed.

INTRODUCTION

A lot has been written about the techniques and methods employed in creating a stereo mix or mix for music. There is very little literature about the actual process of a Film Mix or Immersive mix. One of the reasons is probably because this is a relatively new format and the artistic and engineering capabilities are still being explored. In this paper, I will attempt to deconstruct the Emotional Aspect of Story telling and how this can be applied to an Immersive format.

It always has been a greater exploration to find the right balance between sounds so as to get the listener to be a part of the Visual Story telling that is happening on screen. TO this extent, I would say that there are 3 perspectives in this too; First Person, Second Person and Third Person. The balance and the weaving of the perspective between these bring what I would like to call Emotional Dynamics in the context. The biggest advantage that Immersive mixing has brought to the forefront is the ability of enabling these very changes by the positioning of sound.

EMOTIONAL UNDERSTANDING

As a listener, it is very important to understand that sound has very little emotional relation or meaning to the person experiencing it unless there is a context to which it can be associated. For example, the sound of a crow in a movie may not indicate much and could be taken as part of an ambience, but when it is shown in the context of lets say witchcraft or so, the meaning associated with it changes. On the other hand, there are sounds that are conditioned into us. An example is the sound of a wolf howling. This creates an eerie atmosphere only because the sound has been used multiple times for this very purpose.

P.N. Juslin [1] has presented a study on how the brain creates an emotional response from sound with six psychological mechanisms:

- *Brain Stem Reflex*: The importance or Urgency is implied by the acoustic characteristic like loudness of the sound.
- *Evaluative Conditioning*: Anchoring an emotional event with a stimulus like using a particular score or sound for an emotion in film.
- *Emotional Contagion*: Causing the listener to reciprocate or mimic the emotion created with the sound.
- *Visual Imagery*: Sound can create a visual image or idea in relation to the emotion presented.
- *Episodic Memory*: Triggering an event or memory by sound.
- *Expectancy*: Conditioning the listener to a sequence or frequency of sounds that is expected for a particular emotion. When this is broken, tension can be created.

Ekman and Kajastila [2] have presented a paper where they presented a study on the sound source direction and width influence on the perception of scariness. They have demonstrated the effect of manipulating special width and position to fine-tune the sound and its emotional impact. On the basis of these, it is clear that there are directional influences that can impact our perception of the emotional value of sound.

POSITIONING AND RESPONSE

The way we accept and emotional response to the sound is based a lot on the position of it and more importantly the proximity. On a usual question to students and professionals on what they perceive is more intimidating; is it a sound of a twig break or a tiger growl. Usually the response is twig break. This is because of the amount of imagination we put into the context of the twig break and the reason we create for it based on the environment we are in at that moment. We can identify a tiger growl. But since the twig break can have multiple causes for it, we recognise fear more. The second question I posed was where would be scarier - in front or in the surrounds. The majority of the response was the surround. Again the reason I would put for this is based on episodic memory because a sound without a known source would be more intimidating than one we know.

The second point is dynamics. Dynamics are a very important part of story telling. Looking back at the Brain Stem Reflex reasoning provided, the loudness and the attack of the sound also signifies urgency or evokes a response. A variation in this aspect can break what is expected, thereby creating the change needed to remove the listener from monotony and breaking the emotional barrier. There are 3 ways to create the dynamics:

1. Volume
2. Frequency
3. Position



We are all well versed with the first 2 methods that are usually employed as a technique in film and music mixing to create excitement and movement. What I am exploring now is the Position based dynamics to create the emotional response.

Creating the dynamics

One of the recent films I mixed that employed these techniques is a film called Bombay Velvet [3]. The story of the film revolves around a time period that spans 20 years. A method we approached was to have the sound evolve from Mono to LCR, to 5.1, to 9.1 and finally Atmos. What this allowed us to get was the change in dynamics by varying the audience from becoming First Person to second to third and once establishing a pattern, we broke it thereby creating the emotional response needed without much need of volume or amplitude based changes. The change in perspective was created with positional changes thereby engaging the audience much more deeper into the story emotionally.

EMOTIONAL STORYTELLING

There are multiple ways of artistically using the ability of an immersive mix technique for storytelling. I will outline some of them here that help with this.

Sense of Position

One of the ways of achieving proximity or help deciding if the audience is first person or third person is the proximity of them in relation to the story on screen. The closer and more realistic the positioning is, the more effective it would become in putting the audience in the space of the character. This helps in making the audience identify with the character or situation much more easily. For example, constructing a street or restaurant that is hyper real to then taking it out of context for an emotional sound design need.

Sense of Proximity

There are 2 main factors that affect the sense of proximity to a sound. First one is the Space, the second being the attack. Space is something that can be manipulated using reverbs and EQ. Space also provides the listener with the distance that one has from the character in the given scene of the film. This means that it can be effectively used to determine what amount of engagement the audience has at that given point in the story and whether it should be broken for dynamics. Proximity in Immersive mix formats allow for a much higher accuracy because of the Height, depth as well as positional accuracy (in Object based panning systems). A combination of these three provides the listener with the current space they need to experience.

The Attack of a sound also determines how close it is to us. An object based panning system can provide us with much more accurate representation of this. This technique is something quite often used in music and film where a compression on the attack of the given sound to reduce its proximity, for example on the tambourines or snares. This can also be used in Foley sounds in film for footsteps or other sounds that can be very close or far.

Music and placement

Music is one component that has a huge emotional impact on the listener. There have been various studies regarding the emotion of instruments. [4 – 7]. Ronald Mo, Bin Wu

and Andrew Horner have also mentioned that the trumpet sounds happier in character compared to the horn. Although the study of that is not part of the scope of this paper, the placement is something we can explore. The purpose of music in films can serve multiple purposes from emotional engagement to time lapses and montages. In addition to this, there have been studies by various people, one notable being by Rob Parke, Elaine Chew, and Chris Kyriakakis [8] on whether music affects the emotional perception of a film. While there is very little debate in that space, the placement and movement of sound is something that can and is experimented in context of the story. Places where the audience needs to reflect and concentrate on the screen itself, the score can be placed very near to the character on screen. This brings a form of personalization and identification to the character. Of course for this to work correctly, the music must be in the right genre that has been established. Else, the audience will dissociate themselves from the piece and the emotional context of the music. Using this technique, it is more easy to have the instruments “Speak” to the audience and have the score compliment the shots on screen, like if it is a close shot, or a sweeping pan shot etc.

Sense of Reality and Dissociation

One of the best advantages of the Immersive format is constructing realistic scenes and then using this to bring the audience in and out of the story telling as an art. For example, creating a realistic scene of a street in India that has a character walking through it and slowly shifting it to a character perspective or POV. This is something that gives immense scope for the positional dynamics that I earlier mentioned. So, it is very easy to craft an environment that can be extremely engaging and slowly change or remove elements that influence the listener thereby shifting the focus to the screen or the character or the location as needed.

CONCLUSION

In conclusion, I would like to bring out the emotional impact of the way the above techniques can be used. It can be expanded to a wider form with a better use of these as sound design / music when it comes to storytelling in cinema. The advances happening in Reactive Experience to sound in some cases like the Virtual Reality Headsets, and the techniques being employed there, it can only be closer to achieve the emotional response we want from the audience. This new field has the advantage of being reactive unlike cinema, which is more of a guided experience. That was one of the primary reasons I feel that Immersive audio was created for film. Being able to do this and reversing the way we emotionally react to sound and its placement is something that has a lot of scope for development in the technical as well as artistic field. It has also demonstrated and shown various studies and the interpretation of that for mixing sounds that provide the listener with the experience and more importantly, the creator with the tools and understanding for better storytelling. In being able to use Primed association of audio, along with Positional dynamics, getting the audience to enjoy the creators original intent is a closer reality now than it was before.

REFERENCES

[1] P. N. Juslin, “Seven Ways in which the Brain Can Evoke Emotions from Sounds”, Sound, Mind and Emotion, The Sound Environment Centre at Lund University, Report no. 8, Sweden, 2009.

[2] I. Ekman and R Kajastila “Localisation Cues Affect Emotional Judgments – Results from a User Study on Scary Sound” Proc. AES 35th Conference on Audio for Games, London, UK (2009)

[3] <http://www.imdb.com/title/tt2979920/>

[4] B. Wu, C. Lee, and A. Horner “The Correspondence of Music Emotion and Timbre in Sustained Musical Instrument Tones,” *J. Audio Eng. Soc.*, vol. 62, pp. 663–675 (2014 Oct.), <http://dx.doi.org/10.17743/jaes.2014.0037>.

[5] B. Wu, A. Horner, and C. Lee “Musical Timbre and Emotion: The Identification of Salient Timbral Features in Sustained Musical Instrument Tones Equalized in Attack Time and Spectral Centroid,” *International Computer Music Conference (ICMC), Athens, Greece*, pp. 928–934 (14–20 Sept 2014).

[6] T. Eerola, R. Ferrer, and V. Allure “Timbre and Affect Dimensions: Evidence from Affect and Similarity Ratings and Acoustic Correlates of Isolated Instrument Sounds,” *Music Perception: An Interdisciplinary J.*, vol. 30, no. 1, pp. 49–70 (2012) <http://dx.doi.org/10.1525/mp.2012.30.1.49>.

[7] The Effects of Reverberation on the Emotional Characteristics of Musical Instruments by Ronald Mo, Bin Wu and Andrew Horner

[8] Multiple Regression Modeling of the Emotional Content of Film and Music by Rob Parke, Elaine Chew, and Chris Kyriakakis.

ACKNOWLEDGEMENTS

The Author would like to thank his colleagues and family for the help provided in the preparation and publishing of this work. In addition, he would like to thank Avid Technology and Dolby Laboratories, for the time and help provided in the completion of this paper. He would also like to thank the International Broadcasting Convention for permission to publish this paper.