GOAL-ORIENTED MIXING

Emmanuel Deruty

Sony CSL Paris
{derutycsl}@gmail.com

ABSTRACT

In the field of music production, mixing is the process that aims at converting multi-tracks to music. Automatic mixing is a field of research that aims at performing this task without human intervention. The automatic mixing community is currently focused on producing technically correct mixes. Our position is that the mixing process serves many purposes, and that technical correctness is but one of them. We call this approach *goal-oriented mixing*.

1. INTRODUCTION

The state of the art in automatic mixing provides knowledge and tools to perform a technically correct mix. This is useful, insofar as a mixing engine should be able to produce a technically correct rough mix, leaving the interesting, creative stuff for the mixer [1]. We must look beyond correctness, and one way to do so is through goal-oriented mixing. We provide a non-exhaustive list of such goals, illustrated by real-life examples. In many cases, the goal can be the expression of an emotion. The mix can therefore be linked to BRECVEMA mechanisms [2-6]. This paper refers to audio examples. To listen to the examples please go to http://csl.sony.fr/music/wimp2016.

2. GOAL: THE MIX MUST BE TRANSPARENT (EXAMPLES 1-6)

In this section, we list cases in which either the mix must not be heard as such, or most of the mixing process consists in hiding itself. This aspect of the mixing process is reminiscent of the assumption that all sources should be heard as well as possible, be it by way of equal loudness between tracks [7-9] or masking minimization [7, 10-11]. Note that an absence of mix doesn't necessarily result in a realistic or transparent mix.

3. GOAL: THE MIX MUST CONVEY A PARTICULAR FEELING (EXAMPLES 7-19)

The mixing process may highlight or create feelings that are conveyed to the listener. Several BRECVEMA mechanisms may be concerned. In a number of examples, we observe that equalization and dynamic compression may help activate the BRECVEMA mechanism "emotional contagion". In other examples, reverberations seem to activate the mechanism "visual imagery". In yet another example, volume automation appears to activate the mechanism "brain stem reflex".

4. GOAL: THE MIX MUST FOLLOW TRENDS (EXAMPLES 20-27)

Trends apply to the mix as well. Particular trends may be expressed in terms of spectrum [12] or dynamics [13]. Revivals of vintage sounds are trends - in which case the BRECVEMA mechanism "episodic memory" may be activated. If the trend is expressed in terms of spectrum, the spectral profile may be understood as a *target* in the process of target mixing as defined by [14-16].

5. GOAL: THE MIX MUST HELP CREATE ABSTRACT STRUCTURES (EXAMPLES 28-47)

The mixing process is able to create or underline a variety of abstract structures. Such structures involve: sound scenes as defined by [17]; abstract movements of a given track; space sequences involving a group of tracks [18]; auditory illusions as defined by [19]; articulations of vocals, in relation to the use of pitch-correction plug-ins [20]; and musical structure as defined by [21]. The listener's appreciation of such abstract structures may relate to the BRECVEMA mechanism "aesthetic judgment".

6. GOAL: THE MIX MUST ENHANCE / CORRECT (EXAMPLES 48-60)

Enhancements performed during the mixing process can be diverse. They include: creating punctuations; making tracks more lush; making the music more "dancy", in which case the BRECVEMA mechanism "rhythmic entertainment" may be activated; getting tracks closer to what's perceived as musical perfection, using pitch-correction plug-ins in particular.

7. GOAL: THE MIX MUST PRIORITIZE TRACKS (EXAMPLES 61-63)

The practice of track prioritization contradicts the principle that all tracks should be made as audible as possible [10-11, 22-23]. As storage price is low, musicians may postpone decisions to the last moment. As a result, the mixer receives sessions with an inordinate number of tracks. The mixer has to tone down some tracks or if possible, remove them [24]. Track prioritization can also originate from social conventions: on Céline Dion's records, the mixer is required to mix her voice significantly louder than the customary level [25].

8. CONCLUSION

There are many goals that mixing can reach beside technical correctness:

- The mix may pretend to be transparent.
- The mix may help create / convey a particular feeling.
- The mix may follow trends (current or past).
- The mix may help create abstract structures.
- The mix may enhance and correct sounds.
- The mix may prioritize tracks.

Automatic mixing currently addresses only a few of these goals. We wish that the community would investigate the technical means to reach more goals, so that automated mixing may provide more perspectives and stand a chance to emulate a human mixer.

9. REFERENCES

[1] D. Reed. "A Perceptual Assistant to do Sound Equalization," 5th Int. Conf. on Intelligent User Interfaces, New Orleans, LA, USA, 2000.

[2] P. N. Juslin and D. Vastfjall. "Emotional responses to music: The need to consider underlying mechanisms." *Behavioral and Brain Sciences*, vol. 31, no. 5, pp. 559–575, 2008.

[3] P.N. Juslin, S. Liljestrom, D. Vastfjall and L.O. Lundqvist. "How does music evoke emotions? Exploring the underlying mechanisms," in *Handbook of music and emotion*, Oxford Press, 2010, pp. 605–642.

[4] P.N. Juslin. "From everyday emotions to aesthetic emotions: towards a unified theory of musical emotions." *Physics of life reviews*, vol. 10, no. 3, pp. 235–266, 2013.

[5] P.N. Juslin, L. Harmat and T. Eerola. "What makes music emotionally significant? Exploring the underlying mechanisms." *Psychology of Music*, vol. 42, no. 4, pp. 599-623, Jul. 2014.

[6] D. Ronan, J. Reiss and H. Gunes. "An empirical approach to the relationship between emotion and music production quality." Currently in review. Submitted to IEEE, Jul. 2016.

[7] D.Ward, J.D. Reiss and C. Athwal. "Multitrack Mixing Using a Model of Loudness and Partial Loudness." AES 133rd Conv., San Francisco, CA, USA, 2012.

[8] R.B. Dannenberg. "An Intelligent Multi Track Audio Editor." *Proc. of the 2007 Int. Computer Music Conf.*, San Francisco, CA, USA, 2007.

[9] S. Mansbridge *et al.* "Implementation and Evaluation of Autonomous Multi-Track Fader Control." AES 132nd Conv., Budapest, Hungary, Apr. 2012.

[10] S Hafezi and J.D. Reiss. "Autonomous multitrack equalization based on masking reduction." *J. Audio Eng. Soc.*, vol. 63, pp. 312-323, May 2015.

[11] E. Perez-Gonzales and J.D. Reiss. "Improved control for selective minimization of masking using interchannel dependency effects." *Proc. of the 11th Int. Conf. on Digital Audio Effects (DAFX-08)*, 2008, pp. 75–81.

[12] P.D. Pestana, Z. Ma, J.D. Reiss, Á. Barbosa and D.A.A. Black. "Spectral Characteristics of Popular Commercial Recordings 1950-2010." AES 135th Conv., New York, USA, 2013.

[13] E. Deruty and F. Pachet. "The MIR Perspective on the Evolution of Dynamics In Mainstream Music." *Proc. of the 16th Int. Society for Music Information Retrieval Conf. (ISMIR 2015)*, Oct. 2015, pp. 722-727.

[14] Z. Ma, J.D. Reiss and D.A.A. Black. "Implementation of an intelligent equalization tool using Yule-Walker for music mixing and mastering." AES 134th Conv., Rome, Italy, 2013.

[15] J. Scott, M. Prockup, E. M. Schmidt and Y. E. Kim. "Automatic multi-track mixing using linear dynamical systems." 8th Sound and Music Computing Conf., Jul. 2011.

[16] D. Barchiesi and J.D. Reiss. "Automatic target mixing using least-squares optimization of gains and equalization settings." 12th Int. Conf. on Digital Audio Effects (DAFx-09), Como, Italy, 2009.

[17] W. Moylan. *Understanding and Crafting the Mix, the Art of Recording*. Focal Press, 2007.

[18] E. Deruty and F. Pachet. ISMIR 2016 tutorial, "Why is studio production interesting?" Columbia University, New York, NY, USA, August 6th, 2016.

[19] J.C. Risset. "Pitch Control and Pitch Paradoxes Demonstrated with Computer-Synthesized Sounds." *J. Acoust. Soc. Amer.,* Jul.1969.

[20] P. White. "Shaped by technology." *Sound on Sound* (Jan. 2012).

[21] F. Bimbot, E. Deruty, G. Sargent and E. Vincent. "Methodology and Resources for the Structural Segmentation of Music Pieces into Autonomous and Comparable Blocks." *Proc. of the 12th Int. Soc. for Music Information Retrieval Conf. (ISMIR 2011)*. Miami, FL, USA, Oct. 2011, pp. 287-292.

[22] E. Perez-Gonzales and J. D. Reiss. "Automatic Equalization of Multichannel Audio Using Cross-Adaptive Methods." AES 127th Conv., 2009.

[23] P. Aichinger, A. Sontacchi and B. Schneider-Stickler. "Describing the transparency of mixdowns: The masked-to-unmasked-ratio." AES 130th Conv., May 2011.

[24] S. 'Alf' Briat. Private correspondence with the author (Jul. 2016).

[25] Y. Macé. Private correspondence with the author (Aug. 2016).