Variation in Multitrack Mixes: Analysis of Low-level Audio Signal Features

ALEX WILSON, *AES Student Member,* **AND BRUNO M. FAZENDA,** *AES Member*

(a.wilson1@edu.salford.ac.uk)   (b.m.fazenda@salford.ac.uk)

Acoustics Research Centre, University of Salford, Greater Manchester, M5 4WT, UK

To further the development of intelligent music production tools towards generating mixes that would realistically be created by a human mix-engineer, it is important to understand what kind of mixes can be created, and are typically created, by human mix-engineers. This paper presents an analysis of 1501 mixes, over 10 different songs, created by mix-engineers. The primary dimensions of variation in the full dataset of mixes were "amplitude," "brightness," "bass," and "width" as determined by feature-extraction and subsequent principal component analysis. The distribution of representative features approximated a normal distribution and this is then used to obtain general trends and tolerance bounds for these features. The results presented here are useful as parametric guidance for intelligent music production systems.

0 INTRODUCTION

There are a number of stages in the music production process from the initial composition to the final distribution. Central to this process is the creation of the mix, when the recorded audio is assembled into the arrangement and sound for which the song will become recognized. While the recording engineer may capture a great number of individual and group performances, it is the mix engineer who is tasked with the challenge of combining all of these elements into one mix; a challenge that is often both highly creative and highly technical. The task of creating a mix from multitrack audio can be considered an optimization problem, albeit one with a large amount of variables and a target that is not well defined. Studies have investigated mix-diversity by compiling best-practice behaviors for the art of multitrack mixing, either by interviewing professional mix engineers [1] or from the analysis of subjective ratings and comments in reviews of mixes by students on music technology related subjects [2, 3]. Consequently, many of the "best-practice" techniques in mix-engineering are anecdotal and limited in generality. Material available for education in mix-engineering is typically based on the experience of a small number of professionals who have each produced a large number of mixes over their careers [4–6]. Due to the proliferation of the digital audio workstation as a low-cost audio production platform and the distribution of software, audio and educational materials via the internet, it is possible to reverse this paradigm, and study the actions of a large number of engineers on a small number of music productions. This allows both quantitative and qualitative study of mixing practices—the dimensions of mixing, and the variation along these dimensions, can be investigated. For a human mix-engineer it is of course important to treat each song individually and create the optimal mix, even if based on general rules that have been learned by the engineer. For the development of automated/intelligent music production systems, the study of alternate mixes by many mix-engineers may allow for an insight into human decision-making in mixing that has not previously been exploited.

The authors' previous work [7] demonstrated that, with some subjective rating, one can learn which features might be correlated to the perception of quality. Here, the focus lies in defining what trends might exist across mixes of a song and in general for many songs. Perceptual rating is implicit in the choices made by each mixer as they strive to achieve the best mix from their own viewpoint. Arguably, the fact that each song has an associated variance for each feature is evidence that there is a subjective/perceptual aspect at play and that no "perfect" mix exists.

1 METHODOLOGY

The data used in this study was collected directly from Cambridge Multitracks (http://www.cambridge-mt.com), which hosts multitrack content along with a forum where members can publicly post their mixes of that content. The database categorizes multitrack content by genre and of the ten most mixed sessions, eight belong to the "Rock/Punk/Metal" category. The songs that have attracted the most mixes (as of Nov. 2015) were specifically favored. Due to the "Rock/Punk/Metal" category being preferred,

this study focusses on these genres—often-mixed songs from other categories are omitted in place of slightly less-often mixed songs from within this category. This allows the creation of a dataset that contains a consistent selection of instruments and sounds, including, but not limited to, drums, electric bass, guitars, and vocals.

1.1 Pre-Processing

The majority of the mixes were only available in MP3 format at bit-rates between 128 kbps and 320 kbps. All downloaded files were converted to .wav format, at a sampling rate of 44.1 kHz and a bit-depth of 16 bits. While lossy encoding would have an effect on certain objective measures of the signal, such as reducing the value of *Spectral Centroid* and *Rolloff* features, this effect can be demonstrated to be negligible. For a given song, each mix was of a different length, due to varying amounts of silence at the start and end of each file and also various acts of re-arrangement such as the removal or duplication of certain bars. This made it difficult to use the entire audio in the analysis. To normalize the choice of audio segment, the audio was cut to short segments containing the second chorus of the song. Each of these segments was then time-aligned, which was achieved by determining the peak in the cross-correlation vector when comparing one mix to all others. All of the mixes but one were zero-padded to align the files accordingly. Each mix was then trimmed to a 30-second length containing the chorus. This ensures that feature extraction tasks can be performed fairly on all mixes. This process was applied to each batch of mixes of each song. It assumes that tempo does not vary across mixes of the same song, which is demonstrated to be true in this dataset.

1.2 Feature-Extraction

As many established audio signal features have been designed for Music Information Retrieval (MIR) tasks such as instrument recognition or genre classification, it is not widely understood which features would be best suited to categorizing mixes of a given song. Features relating to the perception of polyphonic timbre were thought to be important based on earlier work [8] and so the sub-band spectral flux was determined [9]. The statistical moments of the sample amplitude probability mass function (PMF) have been shown to categorize different types of distortion in mixing and mastering processes [10] and so these features are also used. Spatial features were derived from the stereo panning spectrogram (SPS) [11]. Table 1 contains a full list of features. At this stage, features related to rhythm are not included since the structure, form, and meter of varying mixes should be identical. Further discussion of rhythm can be found in Sec. 3.

1.3 Research Questions

Subjective appraisal of these mixes, in the conventional sense of controlled listening tests, is not included in this paper due to the overwhelming size of the dataset. However, as all mixes were created in real-world conditions, we assume each engineer produced their mix to the best of their

Table 1. Audio signal features used in analysis. Features with KMO <0.6, marked with an asterix, are not included in the PCA.

Feature	Label	Ref.	KMO
Spectral Centroid	SpecCent	[12]	0.758
Spectral Spread	SpecSpr	[12]	0.797
Spectral Skew	SpecSkew	[12]	0.851
Spectral Flatness	SpecFlat	[12]	0.898
Spectral Kurtosis	SpecKurt	[12]	0.852
Spectral Entropy	SpecEnt	[12]	0.639
Crest Factor	CF		0.967
LoudnessITU	LoudITU	[13]	0.834
Top1dB	Top1dB	[10]	0.900
Harsh	Harsh	[14]	0.633
LF Energy	LF	[14]	0.631
Rolloff85	RO85	[15]	0.819
Rolloff95	RO95	[15]	0.677
Gauss	Gauss	[14]	0.965
PMF Centroid	PMFcent	[10]	0.938
PMF Spread	PMFspr	[10]	0.890
PMF Skew		[10]	0.534*
PMF Flatness	PMFflat	[10]	0.962
PMF Kurtosis	PMFkurt	[10]	0.907
Width (all)	W.all	[11, 8]	0.966
Width (band)		[11, 8]	0.591*
Width (low)	W.low	[11, 8]	0.778
Width (mid)		[11, 8]	0.540*
Width (high)		[11, 8]	0.567*
Sides/Mid ratio			0.593*
LR imbalance		[16]	0.518*
Spectral Flux	sbflux1–10	[9]	All >0.8

abilities and towards their desired target. In this sense, subjective evaluation is implicit in the data itself. This dataset of mixes can be used to address a variety of challenges, a number of which are explored herein.

1. Which features vary most across mixes?
2. What are the dimensions of mix-engineering practice, across all songs and for a particular song?
3. How are the values of low-level features distributed in the dataset? What are their typical means and variance?

2 ANALYSIS OF MIX DATASET

Outlier detection was performed in the 36-dimensional feature-space (see Table 1). The Z-score of each point was determined by the Euclidean distance to the three nearest neighbors. Thirty-five samples where $Z > 2.5$ were deemed outliers, leaving 1466 audio samples remaining.

2.1 Principal Component Analysis

In order to reduce the dimensions of the feature-space, Principal Component Analysis (PCA) was used. The appropriateness of PCA was tested as follows using R [17]. Using Bartlett's test of sphericity (using the psych package [18]), the null hypothesis that the correlation matrix of the data is equivalent to an identity matrix was rejected ($\chi^2(630, N = 1466) = 97162.75$, $p < 0.001$). This indicated that factor analysis was a suitable analysis method. The Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) was
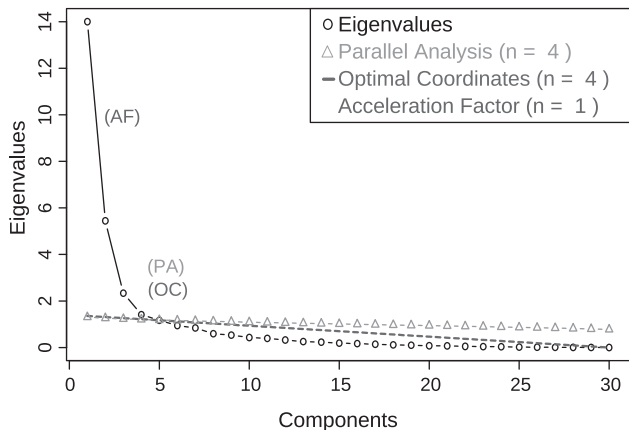
Fig. 1. Scree plot for initial PCA.

Table 2. Eigenvalues of revised PCA.

|	1st	2nd	3rd	4th
Eigenvalue	14.00	5.45	2.34	1.42
% var	46.68	18.15	7.80	4.72
Cuml. % var	46.68	64.83	72.62	77.34

evaluated. KMO for the full set of variables was 0.845, above the recommended value of 0.6 [19], suggesting that factor analysis would be useful. KMO for each individual variable was determined and any individual variables with a value less than 0.6 were excluded from analysis (see Table 1). Consequently, PCA was conducted with the remaining 30 variables. Each variable was standardized prior to PCA, i.e., mean $\mu = 0$ and standard deviation $\sigma = 1$. This initial PCA is unrotated and there was no limit on the number of components. The plot of eigenvalues is shown in Fig. 1.

Using the `nFactors` package [20] a variety of methods were employed in order to determine the number of dimensions to keep in further analysis, shown in Fig. 1. Kaiser's rule [21] suggests retaining those dimensions with eigenvalues greater than 1, which in this case was the first five components. The acceleration factor (AF) [20] determines the knee in the plot by examining the second derivative—this method would retain only the first dimension but is known to underestimate [22]. The optimal coordinates (OC) method [20] suggested that the first four dimensions be kept. Parallel analysis (PA) [23] also suggested that the first four dimensions were suitable to retain. Based on agreement suggested by three of the four methods, four dimensions were kept for the subsequent analysis.

As before, 30 variables were used for a revised PCA, now limited to four dimensions and rotated using the varimax method [24]. Rotation was applied so that the resultant factors were easier to interpret, by ensuring variables had high loading on one dimension and low loading on those remaining. The eigenvalues of this PCA are shown in Table 2, four dimensions accounting for ≈77% of the variance. The following is an interpretation of each of the first four dimen-
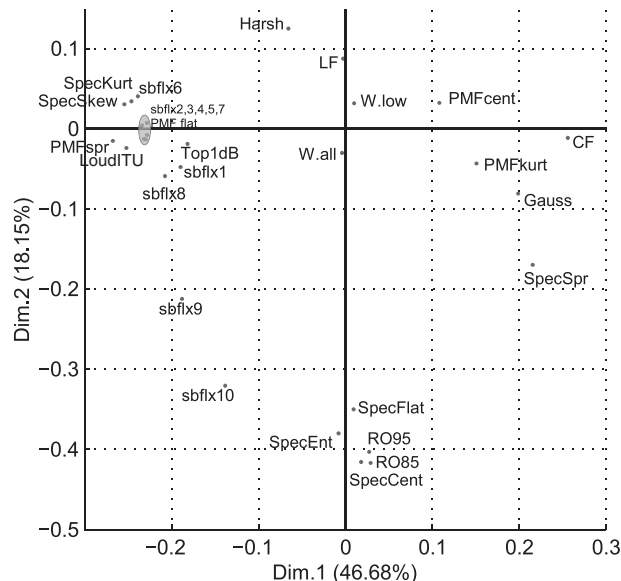
(a) Dimension 1 relates to mostly amplitude features and dimension 2 to mostly high-frequency spectral features.
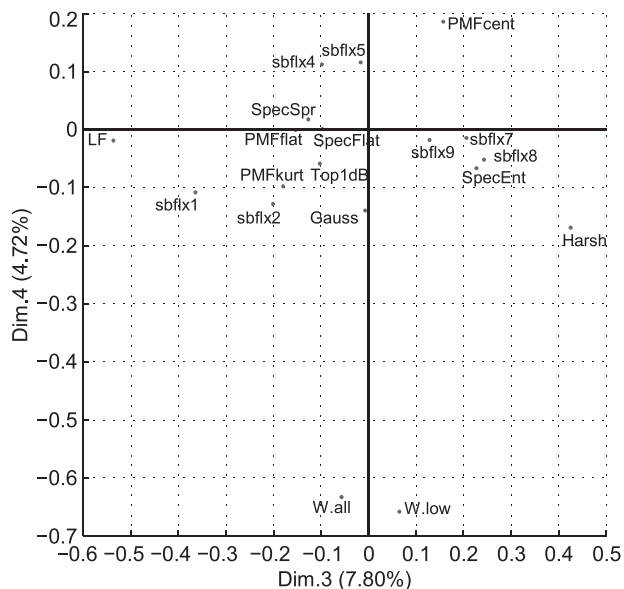

(b) Dimension 3 relates mostly to either low or high-frequency features and dimension 4 to spatial features. Loadings < 0.1 are removed for clarity.
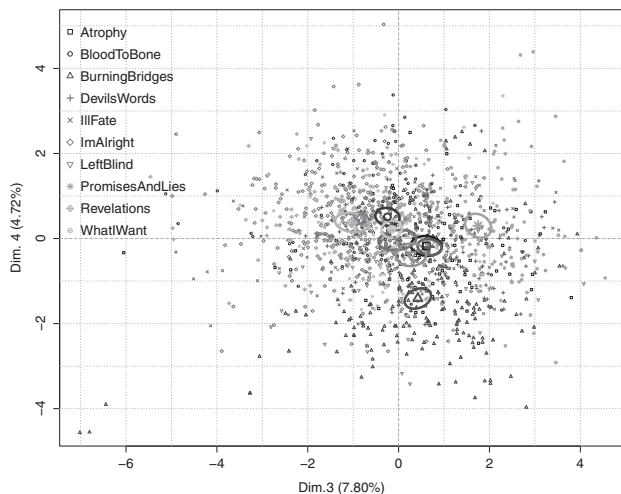
Fig. 2. Results of PCA for 1466 audio samples. The variables factor maps, shown in (a) and (b), indicate loadings of variables on the varimax-rotated principal components.

sions, based on the loadings of the individual features, as shown in Figs. 2a and 2b. This addresses research questions 1 and 2 from Sec. 1.3.

1. Many of the input variables associated with signal amplitude, dynamic range, and loudness are strongly correlated with the first principal component. Negative values indicate high amplitude mixes (see Fig 2a).

2. The second dimension can be described by the many strong correlations to spectral features with negative
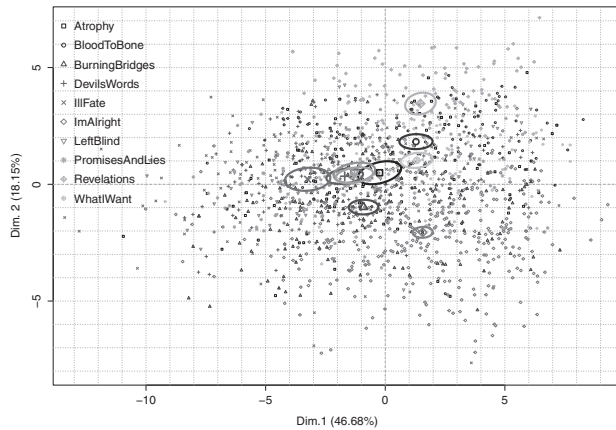
(a) Mixes of a song vary more in dim.1 than dim.2, while songs differ from one another more along dim.2 than dim.1. The mixes of all songs overlap greatly in this feature-reduced space.

(b) There is great overlap in this space, yet the central value of certain songs differ from others. Mixes of three specific songs stand out in the upper-left, right and bottom of the plot.

Fig. 3. Results of PCA for 1466 audio samples. The individual factor maps, shown in (a) and (b), display the placement of each audio sample in the space, grouped by song. The centroid of each group is marked by thick markers and the ellipses represent regions of 95% confidence in the population centroid of that group. From this result it can be seen that, while clustering is evident, songs are not easily categorized by the features used.

values denoting mixes that have a greater proportion of energy in higher frequencies (see Fig 2a).

3. Features associated with low frequencies are more strongly loaded onto dimension 3 in the negative direction, while treble range features are loaded with positive values (see Fig 2b).

4. Dimension 4 can be explained by the correlation of the spatial features to this dimension. As the value of this dimension decreases, the perceived width of the stereo image increases (see Fig 2b).

Figs. 3a and 3b show the dataset of mixes placed in the varimax-rotated PCA space. Each point represents a mix of a song, where the song is coded by a unique color and symbol combination. We can see significant overlap

between the range of mixes for all 10 songs. The estimated centroid of each group, and the 95% confidence ellipse of that centroid estimation, are also indicated in Figs. 3a and 3b. There is an indication that some songs, and their range of mixes, might form clusters for given dimensions, suggesting that there are central tendencies in mixing when these dimensions are considered (see Sec. 3).

2.2 Distribution of Audio Signal Features

The density of each extracted feature was estimated using the `density` function in R with a Gaussian smoothing kernel. Fig. 4 shows the estimated density of four of the features extracted, considered representative of the principal components due to their high loadings. The plots indicate that the distribution of features shows central tendency, while some curves display additional modes. A Shapiro-Wilk test of normality was carried out [25]. As this test is known to be biased for large sample sizes, the test was carried out not only on the raw data for each song but also the smoothed distributions shown in Fig. 4. The majority of these distributions tested were determined to be significantly different from a normal distribution.

A Gaussian Mixture Model (GMM) was used to determine how well the distribution over all mixes could be characterized by a sum of normal distributions. This was implemented using the `mixtools` package [26]. The model parameters are shown in Table 3 and Fig. 5, where λ_n is the mixing proportion (thus summing to 1), μ_n is the mean, and σ_n is the standard deviation of each of the n Gaussian functions in the model. The coefficient of determination, R^2, is shown in Table 3, indicating the proportion of the estimated density that can be explained by the model where n = 2. As this value is close to 1 in all cases it can be said that the sum of just two Gaussian functions well-approximates the estimated densities.

3 DISCUSSION

Hitherto, there have not been any studies looking at feature variance over such a large number of alternative mixes of the same song. In this study, the features extracted were amplitude-based, spectrum-based or spatial features. Over all 10 songs considered, the dimensions of variation revealed by the PCA were described as "amplitude," "brightness," "bass," and "width" in order of variance explained. Equivalent descriptions of the four dimensions were found in an earlier study that used a subset of the dataset [7]—the dimensions of "brightness," "bass," and "width" were found to be related to the perception of mix quality. Additionally, the description of the first two principal components is equivalent to those found in a related study on popular music, using a similar set of features [8]. This shows that all songs, within their range of mixes, varied in terms of their perceived loudness and dynamics. Fig. 3a shows certain songs with distinct dynamic range values when compared to other songs—the lowest values of dimension 1 (loud, low dynamic range) apply to songs in hard rock or metal styles, whereas the soft rock styles attain higher values along this dimension.
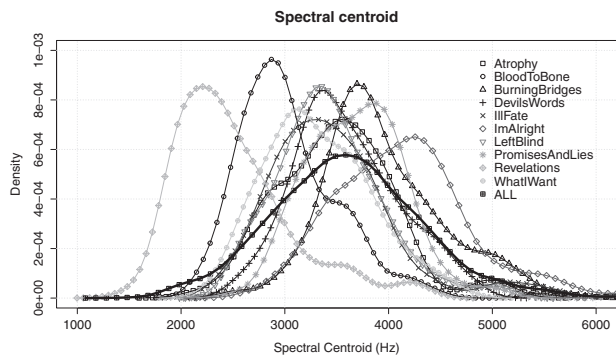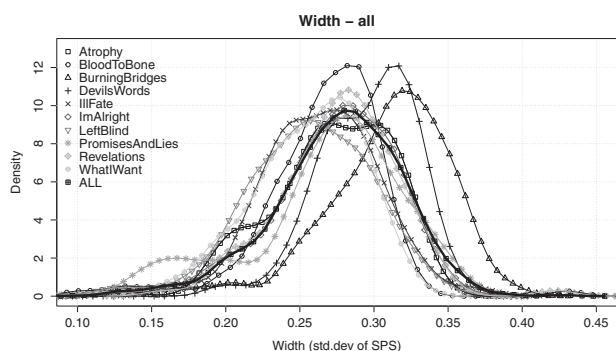
(a) The distributions of *spectral centroid* shows distinct variation from song to song.
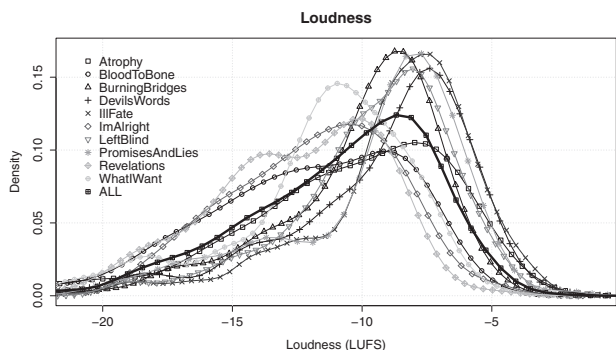
(b) Many mixes were subject to mastering-style processing, resulting in high values of perceived loudness.
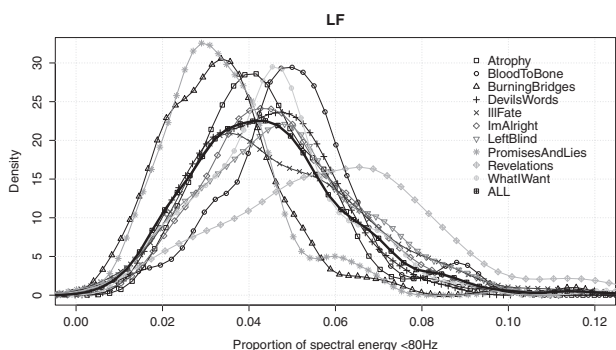
(c) Notable inter-song differences in *LF energy*

(d) Most mixes occupy a narrow range of *width* values. Here the feature used is the value of *width* over all frequencies. Note that a value of 0 represents a mono mix.

Fig. 4. Kernel Density Estimation (KDE) for four of the signal features, shown for all 1501 mixes and also for all mixes of each song. The distributions are typically multi-modal but dominated by one mode.

As the data points in Fig. 3a are spread out over the space, and not definitively grouped by song, it is observed that any one song can be mixed with the overall loudness/dynamics or brightness of any other song. Despite this, trends are apparent. The song "Revelations" had the highest average value of *dim.2*, meaning the least amount of brightness. This may be due to the fact that the multitrack content was recorded in 1975, sourced from an analogue tape. While little is known about the precise recording conditions, it is likely the reduced high-frequency content in mixes of this song was due to the limitations of the recording technology used at the time or the use of era-specific mixing techniques by the mix engineers. The song with the lowest values of *dim.2* (the brightest mixes) is "I'm Alright," which features acoustic guitars and shakers, instruments with emphasis on high frequencies. *Dim.3* is difficult to interpret as it represents emphasis on bass or treble frequencies depending on the value, and there is little inter-song difference. Mixes of the song "Promises and Lies" tended to have a higher concentration of spectral energy between 2 kHz and 5 kHz than other songs, or a lack of spectral energy below 80 Hz. There is little observed difference in the group centroids along *dim.4*, which represents stereo width, particularly at low frequencies, as expected.

Feature distributions in Fig. 4 suggest multi-modal behavior, often dominated by one specific mode, which is dependent on the song. This distribution holds well for the songs considered, providing evidence for central tendency or even "optimal" values. In Fig. 4a, typical values of *Spectral Centroid* differ from song to song, suggesting each song has a range of possible values that can be tolerated, based on the arrangement, instrument timbre, key, etc. The distribution of *Loudness* values in Fig. 4b is quite similar from song to song. This is a possible side effect of the fact that many mixes were subjected to mastering-style processing, particularly heavy dynamic range processing. Fig. 4c indicates that the proportion of spectral energy below 80 Hz is reasonably consistent from song to song, with some variation. This is possibly dependent on the key of the song, the precise arrangement and the relationship between bass guitar and kick drum performances. *Width* distributions shown in Fig. 4d are similar for each song, occupying a narrow range of values. We find songs being mixed with a very wide range of panning conditions, from mono to wide stereo. However, central tendencies can be observed with clear distributions around them. This result indicates that panning conventions are applied similarly in all songs, restricted by the medium of two-channel stereo reproduction, and that a central tendency is observed.

3.1 Implications for Intelligent Music Production

By examining a large dataset of mixes, from hundreds of individual mix-engineers of varying skill levels, the results here indicate the dimensions over which mixes vary and the amounts by which they vary in these dimensions. This could help to inform targets and bounds for
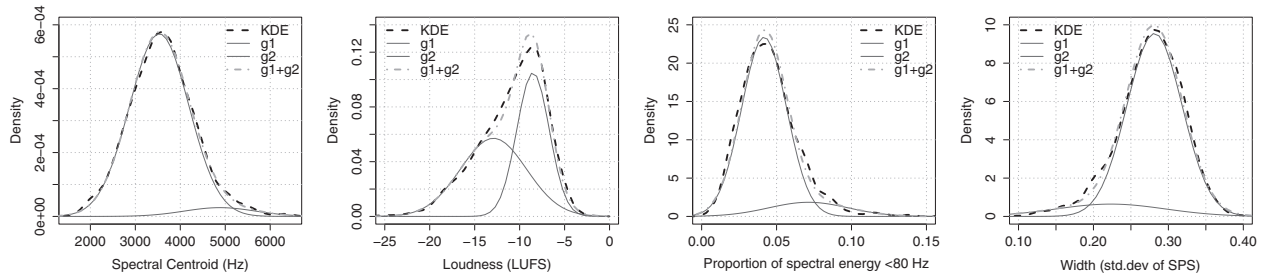
Fig. 5. GMM parameters from Table 3. The dashed curve represents the estimated density and the solid curves represent the GMM. While *Loudness* shows a bi-modal distribution, *Spectral Centroid*, *LF Energy,* and *Width* are well characterized by a single Gaussian function.

Table 3. GMM parameters for distributions of all 1501 mixes. R^2 is the coefficient of determination describing the fit of $(g1 + g2)$ to the KDE curve.

Feature	λ_1	λ_2	μ_1	μ_2	σ_1	σ_2	R^2
SpecCent	0.945	0.055	3532	4880	659	794	0.998
LoudITU	0.526	0.474	−12.910	−8.511	3.672	1.801	0.993
LF	0.877	0.123	0.042	0.071	0.015	0.026	0.989
Width	0.118	0.882	0.223	0.281	0.073	0.037	0.995

intelligent mixing tools. For example, Fig. 5 and Table 3 suggest that values of *Spectral Centroid* are normally distributed with a mean of ≈ 3.5 kHz and standard deviation of ≈ 660 Hz. Consequently, and also shown by Fig. 4a, few rock mixes would have a *Spectral Centroid* value below 2 kHz, although there may exist specific, context-dependent productions where this is possible, such as when analogue recording media are utilized. The results in Table 3 could inform a system that monitors the mix, in an automatic or human-operated system, and offers advice when the values of certain features deviate strongly from expected values.

3.2 Implications for Music Information Retrieval

In a number of tasks in Music Information Retrieval (MIR), feature-extraction is used as a means of characterizing audio data, so that each data point, representing a song or instrument, can be described in a meaningful way. For example, when attempting to train a classifier to perform genre prediction, each song is labelled as belonging to a specific genre and features are extracted from each song. The assumption is that the features can be used to represent useful attributes of that song, and thus, its genre. However, perhaps the features only represent attributes of the recording of the song and not the song itself.

In this study, where there are hundreds of alternate mixes of a given song, we can see that these features do not clearly distinguish between songs. What are the implications then for tasks such as genre prediction? If a classifier was developed with α songs in genre *A* and β songs in genre *B*, how would the performance of the classifier change if alternate mixes were substituted for all $\alpha + \beta$ songs, or for all

possible permutations of classifier that could be made from hundreds of alternative mixes?

Of course, this problem is simplified should estimated tempo be included, as the tempo of a song does not typically change with mix. However, the perception of a song's rhythm can change when instruments are presented at different volumes. Consequently, a detailed study on rhythm in multitrack mixes would be useful in furthering our perception of why certain music mixes are created.

4 CONCLUSIONS

A dataset was prepared containing 1501 audio files representing the mixes of 10 songs. The number of mixes of each song ranged from 97 to 373. A variety of objective signal features were extracted and principal component analysis was performed, revealing four dimensions of mix-variation for this collection of songs, which can be described as "amplitude," "brightness," "bass," and "width." Feature distribution suggests multi-modal behavior dominated by one specific mode. This distribution appears to be robust to the choice of song, with variation in modal parameters. This has provided insight into the creative decision making processes of mix engineers.

Suggested further work is to obtain subjective quality ratings from a subsection of this dataset in order to examine the relationship between audio signal features and the perception of audio quality and mix-preference. Also, as the study presented here only considered features relating to amplitude, spectrum, and stereo panning, an in-depth study using rhythmic and metrical features is planned. It is anticipated that this dataset can be used to test the robustness of algorithms used in MIR, for tasks such as tempo estimation, genre prediction, and music structure analysis.

We are conscious that furthering the understanding of these concepts will be necessary for the design of future intelligent/automated music production systems. However, this incipient study shows that measures of central tendency and distribution are useful targets for such systems. Under higher level human supervision, this concept could be used to achieve sonic qualities that approximate current accepted practices, or as a creative contrast, to challenge current trends and exploit results that may lie at the boundaries of the feature spaces studied.

5 REFERENCES

[1] P. Pestana and J. D Reiss, "Intelligent Audio Production Strategies Informed by Best Practices," presented at the *AES 53rd International Conference: Semantic Audio* (2014 Jan.), conference paper S2-2.

[2] B. De Man, M. Boerum, B. Leonard, R. King, G. Massenburg, and J. D. Reiss, "Perceptual Evaluation of Music Mixing Practices," presented at the *138th Convention of the Audio Engineering Society* (2015 May), convention paper 9235.

[3] B. De Man and J. D. Reiss, "Analysis of Peer Reviews in Music Production," *J. Art of Record Production*, vol. 10 (2015 July).

[4] A. Case, *Mix Smart: Professional Techniques for the Home Studio* (Focal Press, 2011).

[5] B. Owsinski, *The Mixing Engineer's Handbook* (Delmar, 2013).

[6] M. Senior, *Mixing Secrets for the Small Studio* (Taylor & Francis, 2011).

[7] A. Wilson and B. M. Fazenda, "101 Mixes: A Statistical Analysis of Mix-Variation in a Dataset of Multitrack Music Mixes," presented at the *139th Convention of the Audio Engineering Society* (2015 Oct.), convention paper 9398.

[8] A. Wilson and B. M. Fazenda, "Perception of Audio Quality in Productions of Popular Music," *J. Audio Eng. Soc.*, vol. 64, pp. 23–34, (2016 Jan./Feb.), http://dx.doi.org/10.17743/jaes.2015.0090.

[9] V. Alluri and P. Toiviainen, "Exploring Perceptual and Acoustical Correlates of Polyphonic Timbre," *Music Perception*, vol. 27, no. 3, pp. 223–242 (2010), http://dx.doi.org/10.1525/mp.2010.27.3.223.

[10] A. Wilson and B. Fazenda, "Characterization of Distortion Profiles in Relation to Audio Quality," in *Proc. of the 17th Int. Conference on Digital Audio Effects (DAFx-14)*, Erlangen, Germany (2014), pp. 1–8.

[11] G. Tzanetakis, R. Jones, and K. McNally, "Stereo Panning Features for Classifying Recording Production Style," *ISMIR* (2007).

[12] O. Lartillot and P. Toiviainen, "A Matlab Toolbox for Musical Feature Extraction from Audio," *Proc. of the 10th Int. Conference on Digital Audio Effects (DAFx-07)*, pp. 1–8 (2007).

[13] ITU, "ITU-R BS.1770-3 Algorithms to Measure Audio Programme Loudness and True-Peak Audio Level" (2012).

[14] A. Wilson and B. Fazenda, "Perception & Evaluation of Audio Quality in Music Production," in *Proc. of the 16th Int. Conference on Digital Audio Effects (DAFx-13)*, Maynooth, Ireland (2013), pp. 1–6.

[15] G. Tzanetakis and P. Cook, "Musical Genre Classification of Audio Signals," *IEEE Speech Audio Process.*, vol. 10, no. 5, pp. 293–302 (2002), http://dx.doi.org/10.1109/TSA.2002.800560.

[16] B. De Man, B. Leonard, R. King, and J. D. Reiss, "An Analysis and Evaluation of Audio Features for Multitrack Music Mixtures," *ISMIR*, pp. 137–142 (2014).

[17] R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2015).

[18] W. Revelle, *Psych: Procedures for Psychological, Psychometric, and Personality Research* (Northwestern University, Evanston, IL, 2015), R package version 1.5.8.

[19] G. D. Hutcheson and N. Sofroniou, *The Multivariate Social Scientist: Introductory Statistics Using Generalized Linear Models* (Sage, 1999).

[20] G. Raîche, T. A. Walls, D. Magis, M. Riopel, and J.-G. Blais, "Non-Graphical Solutions for Cattell's Scree Test," *Methodology: European J. Research Methods for the Behavioral and Social Sciences*, vol. 9, no. 1, pp. 23 (2013).

[21] H. F. Kaiser, "The Application of Electronic Computers to Factor Analysis," *Educational and Psychological Measurement* (1960).

[22] J. Ruscio and B. Roche, "Determining the Number of Factors to Retain in an Exploratory Factor Analysis Using Comparison Data of Known Factorial Structure," *Psychological Assessment*, vol. 24, no. 2, pp. 282 (2012), http://dx.doi.org/10.1037/a0025697.

[23] J. L. Horn, "A Rationale and Test for the Number of Factors in Factor Analysis," *Psychometrika*, vol. 30, no. 2, pp. 179–185 (1965), http://dx.doi.org/10.1007/BF02289447.

[24] H. F. Kaiser, "The Varimax Criterion for Analytic Rotation in Factor Analysis," *Psychometrika*, vol. 23, no. 3, pp. 187–200 (1958).

[25] S. S. Shapiro and M. B. Wilk, "An Analysis of Variance Test for Normality (Complete Samples)," *Biometrika*, vol. 52, no. 3-4, pp. 591–611 (1965 Dec.).

[26] T. Benaglia, D. Chauveau, D. R. Hunter, and D. Young, "Mixtools: An R Package for Analyzing Finite Mixture Models," *J. Stat. Softw.*, vol. 32, no. 6, pp. 1–29 (2009).

THE AUTHORS

Alex Wilson

Bruno Fazenda

Alex Wilson is currently a Ph.D. student at the University of Salford, investigating the perception of audio quality in sound recordings with a focus on music productions. He received a B.Sc. in experimental physics from NUI Maynooth in 2008 and a B.Eng. in audio technology from University of Salford in 2013, which included a year of industrial experience in studio monitor design. He maintains interests in digital audio processing, music psychology, and the art of record production.

Bruno Fazenda is a senior lecturer and researcher at the Acoustics Research Centre, University of Salford. His research interests span room acoustics, sound reproduction, and psychoacoustics, in particular, the assessment of how an acoustic environment, technology or psychological state impacts on perception of sound quality. He is a researcher in a number of research council funded projects. He is also a keen student on aspects of human evolution, perception, and brain function.