Automatic Monitor Mixing for Live Musical Performance*

MICHAEL J. TERRELL, *AES Student Member*, **AND JOSHUA D. REISS,** *AES Member*
(michael.terrell@elec.qmul.ac.uk) (josh.reiss@elec.qmul.ac.uk)

Queen Mary University of London, London E14NS, UK

A model has been developed which describes the monitor mix experienced by each participant in a musical performance. Using this model, and the mix requirements of each performer, an objective function was defined. This function was then minimized to find the monitor mix settings that best replicate the target monitor mix of each performer, subject to the side constraints of maximum and minimum allowable sound pressure level and the prevention of acoustic feedback.

0 INTRODUCTION

A monitor mix is required on stage during a musical performance. This mix, referred to as the stage sound, will differ markedly from the mix that is heard by the audience, referred to as the front-of-house sound. Emphasis is placed on the front-of-house sound to ensure that the music is recreated in a way that satisfies the artist, yet the stage sound is also important. At a minimum it is vital that the different musicians be able to hear themselves and one another. As the stage sound is improved, the musicians will hear a more natural mix of their music, and as a result will produce a more proficient and expressive performance. This is highlighted in [1].

The front sound and the stage sound are not independent. On-stage sources such as guitar amplifiers, or a naturally acoustic instrument such as a drum kit, will be audible both on stage and out front. Such sources will attenuate and disperse based on their individual characteristics. Monitor loudspeakers are used on stage to provide a mix of all instruments. These tend to be facing away from the audience and so have minimal effect on the front-of-house sound. The front-of-house sound generally has minimal impact on the stage sound as the main public-address loudspeakers are in front of the stage. Standard practice at a concert is to set up the front-of-house sound, and then optimize the stage sound by adapting those elements of the stage sound that will not affect the front-of-house sound, namely, the mix from the monitor loudspeakers.

Each musician will have different requirements in terms of the optimal mix. There are limitations on the variations in monitor mix that can be given to each musician. It is a function of the number of monitor loudspeakers, the number of independent monitor mixes, the location of the performers, and the location and amplitude of the live sources. There is also a minimum and maximum threshold on the overall monitor level; too quiet and the musicians cannot hear themselves/each other, too loud and it may affect the sound out front. Musical performances in television studios for live broadcast are particularly susceptible to the level of the stage sound, as stated in [1].

A number of approaches to automatic mixing have been presented. Dugan [2], [3] introduced the concept of automatic microphone mixing. His work was geared toward conferencing applications. Real-time automatic mixing techniques for musical applications have been presented by Gonzalez and Reiss [4]–[6]. In [4] an automatic gain normalization technique is presented that will prevent feedback, enabling the gain and equalization parameters of the mixing desk to be adjusted freely. In [5] an algorithm is presented that will automatically pan a multi-input musical source based on the spectral content. In [6] the authors investigate the effects of masking on multichannel musical sources and present a method to reduce the masking effects by analyzing the frequency content of each source. Kolasinski [7] used a genetic algorithm to extract track gain settings in multitrack musical recordings. Reed [8] presents a system that can learn the EQ adjustments needed to affect a certain type of perceptual change (that is, make the sound brighter) based on the type of input. To the best of the authors' knowledge, however, no research has been undertaken in the field of automatic monitor mixing.

*Manuscript received 2009 January; revised 2009 June 1 and August 27.

Acoustic feedback is a common problem during musical performances. The mechanism of acoustic feedback is described in [9], [10]. At a musical performance, feedback is particularly prevalent when the input signal to a microphone is low, and hence large amounts of gain must be applied, such as is often the case with vocals. Techniques for suppressing feedback exist, but they are generally unsuitable for musical applications as they require changes to be made to the spectral envelope [11] or they introduce unwanted audio artifacts [12]. The automatic system presented by Gonzalez and Reiss [4] analyzes the feedback loop transfer function and applies a broad-band gain to ensure that the peak of this transfer function is below 0 dB (or a some lower threshold, which acts as a buffer). This is analogous to the process the engineer will go through when setting up the monitor mix. At all frequencies the feedback loop transfer gain must be below unity.

In this paper a model is developed to describe the monitor mix at each listener location on stage. Constraints are then placed on the model based on the requirements of each performer, overall constraints on the mix levels, and the requirement that feedback be prevented. The set of equations generated are then solved to identify the mix settings that best satisfy the constraints.

1 THEORY

The following variables are used to describe the monitor mix.

1.1 Monitor Mix Model

1) Direct sound sources S, which does not include monitor loudspeakers, such as vocalist, guitar amplifiers, or drum kit (although each of the direct sources can pass through the monitors). The SPL of a direct source is a function of frequency and is measured in dB.

2) Monitor speakers M.

3) Total sound pressure level (SPL) L, experienced by each listener, as a function of frequency and measured in dB.

4) Transfer functions, describing each stage in the signal path:

- The combined effect of acoustic attenuation, source dispersion, and room effect on sound emanating from a direct source G_{D} or a monitor speaker G_{MA} and traveling to a listener.
- The gain and equalization settings applied to the indirect signal passing through the mixing desk, G_{I}.
- The gain and equalization settings on the monitor loudspeaker, G_{M}.
- The feedback loop transfer function between each microphone and reenforecement speaker, F.

5) The SPL A experienced by each listener from a particular source, as a function of frequency and measured in dB.

One element of a stage setup that is overlooked here is the conversion of acoustic to electric energy by the microphones used on the direct sources, and the conversion of electric energy to acoustic energy by the monitor loudspeakers. It is assumed that the system is normalized before the application of any mixer or monitor gain, that is, with 0-dB gain settings in the mixer and monitor loudspeakers, the SPL of the signal coming out of the source is equal to the SPL of the signal coming out of the monitor loudspeakers. This could be incorporated into the model easily, as a transfer function for each microphone–monitor loudspeaker combination, but was omitted here for simplicity.

1.2 Model Simplifications

It is our purpose to develop a framework under which the monitor mix can be optimized, subject to a set of requirements. There are a large number of variables in the model proposed in Section 1.1. A significant study into each aspect would be required to get a fully functioning model of the monitor mix. For this reason a number of simplifying assumptions are made.

1.2.1 Frequency Domain

The variables that describe the monitor mix are frequency dependent. In order to simplify the model in this developmental stage the frequency component is omitted. The level of the sources and monitors as well as the transfer functions are described by a scaler gain. It is the intention of the authors to address the frequency dependence in future work.

1.2.2 Acoustic Attenuation

Energy is absorbed by any medium through which a sound wave propagates. This attenuation is a function of the properties of the medium, the temperature, and the frequency of the signal. Higher frequency signals are attenuated to a greater extent. Losses in signal strength caused by acoustics attenuation are omitted in this study.

1.2.3 Source Dispersion

Using an idealized point source to represent each instrument and loudspeaker on stage is inaccurate, as the dispersion pattern will differ from one source to the next, and will be affected by boundary conditions, that is, the interaction of the floor and the wedge monitor loudspeakers. The dispersion pattern is also a function of frequency, but as already discussed, the frequency dependence is omitted in this paper. Approximate dispersion patterns which are independent of frequency will be used, and will be discussed later in this paper. This also applies to microphone polar attenuation patterns.

1.2.4 Room Effect

There is a significant body of work in the literature on room acoustics. Kuttruff's [13] is a standard text on the subject. The effect of an enclosure on the sound at a given location is caused by reflections of the sound source. The reenforcement effect of early reflections on the perceived level of the monitor mix will be significant, particularly in small enclosures. If early reflections reach the listener within 20–30 ms of the direct sound, the brain will process them as a single instance of the sound. This is a well-known psychoacoustic effect, known as the Haas effect, details of which can be found in [14]. It would be possible to track early reflections within this time frame using the image source method, and to combine them with the direct sound. This should yield a more accurate prediction of the SPL at a given location. Given that in 20 ms the sound waves will travel about 7 m, the number of early reflections that must be tracked in order to estimate the reenforcement effect should not prove too computationally intense. The modification and incorporation of the image source method to investigate this reduced time frame is a significant undertaking, and in order to ensure that the issues involved are studied in detail this will be reserved for future work.

1.3 Amplitude Matrices

A performance will involve one or more listeners and one or more sources. The amplitude matrix A_{ij} contains the SPL experienced by each listener from each source. The total SPL vector L_i contains the total SPL of all sources combined for each listener. Where used, all matrices and vectors follow the convention that i is the listener index and j is the source index. The index k is used as the monitor loudspeaker index.

Each element of the amplitude matrix will contain a contribution of the direct sound source (that which travels straight from the source to the listener) and the indirect sound source (that which travels from the source through the mixer desk and out through the loudspeakers, and onto the listener). The direct amplitude matrix A_{D} and the indirect amplitude matrix A_{I} hold the two components of the amplitude matrix. The direct and indirect amplitude matrices define the connectivity between listeners and sources.

The direct amplitude matrix is a function of the source SPL and the acoustic attenuation and dispersion from the source to the listener,

$$A_{\mathrm{D}ij} = S_j + G_{\mathrm{D}ij}. \tag{1}$$

The indirect amplitude matrix is a function of the source SPL, the monitor mix gain applied on the mixing desk, the monitor loudspeaker gain, and the attenuation and dispersion from the loudspeaker to the listener. For each listener–source combination the contribution of all monitor loudspeakers must be included,

$$A_{\mathrm{I}ij} = S_j + 20\log_{10}\left[\sum_{k=1}^{N_k} 10^{\frac{G_{\mathrm{I}kj}+G_{\mathrm{M}k}+G_{\mathrm{MA}ik}}{20}}\right]. \tag{2}$$

Assuming that the two sources are incoherent, the direct and indirect matrices are combined as follows:

$$A_{ij} = 20\log_{10}\left[10^{\frac{A_{\mathrm{D}ij}}{20}} + 10^{\frac{A_{\mathrm{I}ij}}{20}}\right]. \tag{3}$$

From this the total SPL at listener location i can be determined,

$$L_i = 20\log_{10}\left[\sum_{j=1}^{N_i} 10^{\frac{A_{ij}}{20}}\right]. \tag{4}$$

The feedback loop gain for each microphone must be determined. F'_{km} identifies the feedback loop gain of microphone m with monitor loudspeaker k. Unwanted feedback will most often occur with vocal microphones. It is therefore assumed that the position of the microphone will coincide with a listener and source location; hence m can be mapped to a specific combination of i and j. Using this mapping, the feedback loop gain can be calculated by

$$F''_{km} = 20\log_{10}(G_{\mathrm{I}jk} + G_{\mathrm{M}k} + G_{\mathrm{MA}ik} + G_{\mathrm{P}km}) \tag{5}$$

where $G_{\mathrm{P}km}$ reflects the polar attenuation of the microphone in relation to the monitor loudspeaker. The total feedback for each microphone is then given by

$$F'_m = 20\log_{10}\left[\sum_{k=1}^{N_k} 10^{\frac{F''_{km}}{20}}\right] \tag{6}$$

and the feedback loop gain for the entire system is

$$F = 20\log_{10}\left[\sum_{m=1}^{N_m} 10^{\frac{F'_m}{20}}\right]. \tag{7}$$

1.4 Constraint Equations

The purpose of the monitor mix is to give each listener the stage mix that best matches his/her optimal mix. Due to limitations in the number of monitor speakers, the number of independent mixes, and the levels of the direct sources it is unlikely that all requirements will be met exactly. Constraints are imposed on the elements of amplitude matrix A and total SPL vector L. The former represents the desired mix that each listener wants, and the latter represents limitations on the total SPL at each listener location.

1.4.1 Mix Constraints

The mix constraints are used to form the objective function. Each listener defines a specific mix that he or she wants. Consider a band consisting of vocals, guitar, bass, and drums. A set of constraint equations is generated for each listener. Each constraint equation within this set will correspond to a specific listener–source combination and will be measured as a relative value from the listener's own source. For example, the set of constraint equations for the vocalist will express the desired mix of the guitar, bass, and drum levels relative to the vocals. This will generate three constraint equations. Once all listeners are considered, there will be 12 constraint equations in total.

The mix matrix M_{ij} contains the relative gain settings, where i is the listener index and j is the source index for which the relative level is being set. Assuming that the orders of listener and source indices are vocalists and vocals, guitarist and guitar amplifier, bassist and bass amplifier, and finally drummer and drum kit, the example mix matrix given by

$$M = \begin{bmatrix} 0 & -2 & -4 & -2 \\ 0 & 0 & -3 & -2 \\ 0 & -3 & 0 & 0 \\ 0 & -4 & -1 & 0 \end{bmatrix} \tag{8}$$

implies that the vocalist requires the vocals to be 2 dB higher than the guitar, 4 dB higher than the bass, and 2 dB higher than the drums. Constraint equations are formed using the equation

$$A_{ii} + M_{ij} = A_{ij}. \tag{9}$$

In the instance where $i = j$ the constraint is redundant (as $M_{ii} = 0$).

1.4.2 Total SPL Constraints

The total SPL constraints are there to impose an absolute limit on the maximum and minimum total SPLs at each listener location. This will ensure that each listener can hear the performance clearly, but that it is not so loud that the stage sound will affect the front-of-house sound. The latter issue is common at venues where there is a total SPL limit, such as venues in residential areas. The following criteria must be met for all listeners,

$$\mathrm{SPL}_{\mathrm{min}} \leq L_{\mathrm{i}} \leq \mathrm{SPL}_{\mathrm{max}}. \tag{10}$$

1.4.3 Feedback Prevention Constraints

The feedback constraints are used to prevent any of the microphones from developing uncontrolled feedback. Feedback will grow exponentially if the gain loop between a microphone and the monitor loudspeakers is above 0 dB. At levels close to 0 dB unsustained feedback artifacts can also occur. The following feedback prevention criteria must be met,

$$F < F_{\mathrm{limit}} \tag{11}$$

where $F_{\mathrm{limit}} < 0$ and represents the headroom given to the feedback loop.

1.4.4 Optimization Algorithm

The Matlab function fmincon is used to perform the optimization. This is a minimization subject to a set of side constraints. A line-search method was used [15]. The constraints are enforced using Langrangian functions [16]. The function argument is

$$x = f\,\mathrm{mincon}[f(x), x_0, B, C, D, E, \mathrm{LB}, \mathrm{UB}, P, Q, \mathrm{OPT}] \tag{12}$$

where $f(x)$ is the objective function to be minimized. The solution vector x contains all elements of the indirect and monitor gain matrices G_{I} and G_{M}. It is derived from the mix constraints. B and C contain the linear inequalities, and D and E contain the linear equalities. The linear equalities can be used to constrain elements of the solution vector x to equal one another if there are a limited number of monitor mixes available,

$$Bx \leq C \tag{13}$$

$$Dx = E. \tag{14}$$

LB and UB contain the lower and upper bounds of the solution vector x and P and Q are the nonlinear inequalities and equalities. $Q(x)$ contains the total SPL and feedback loop constraints that must be met,

$$P(x) = 0 \tag{15}$$

$$Q(x) \leq 0. \tag{16}$$

OPT contains the input options.

2 TYPICAL BAND CASE STUDY

In this section the automatic monitor mixing will be applied to a typical band case study. The band will comprise vocals, guitar, bass, and drums. The stage layout is shown in Fig. 1.

2.1 Formulation of the Constraint Equations

Fig. 2 shows the direct signal path of the guitar amplifier (source 2), and Fig. 3 shows the indirect signal path of all sources passing through monitor loudspeaker 1. There will be an equivalent set of parameters for all other monitor loudspeakers. These are the parameters to be solved by the optimization algorithm.

Figs. 2 and 3 show the signal path traveling to the listener from a direct source and from monitor loudspeakers. The level to which the signal gain is reduced in this path will be a function of the individual source or loudspeaker dispersion. These characteristics will vary widely from one source to another and will have a strong influence on the final monitor mix. Dispersion characteristics of specific loudspeakers and amplifiers are not readily available, so an assumed dispersion function is

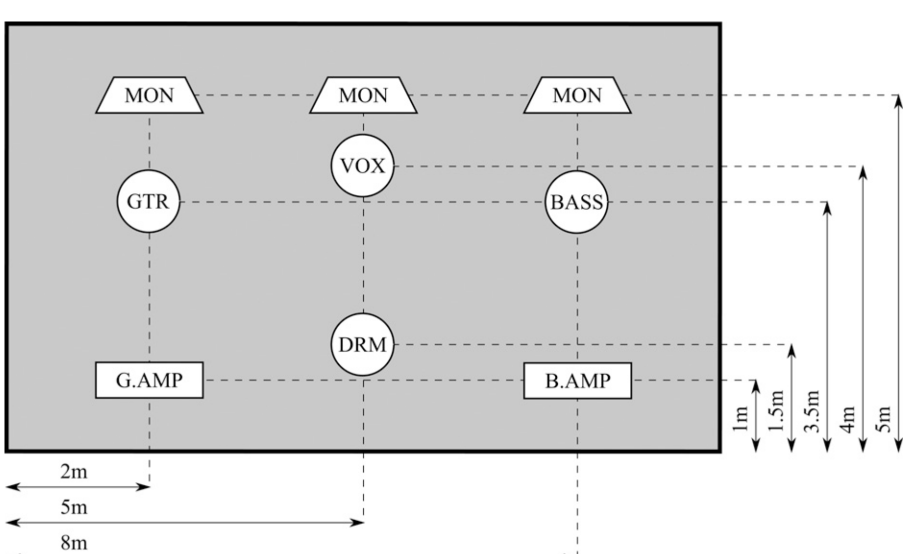


Fig. 1. Stage setup for typical four-piece band.

used. This does not affect the validity of the model presented as a more accurate dispersion and attenuation profile can be substituted into the algorithm as and when it becomes available.

The source dispersion profile used is shown in Fig. 4. The gain values shown reflect the reduction in signal level. Directly in front of the source the gain is 0 dB, directly behind the source the gain is −6 dB. The attenuation is assumed to be −6 dB for every doubling of the distance from a reference of 1 m. These are approximations to the profile given by the Mapp system [17].

The SPL levels shown in Table 1 are used for the four sources.

Using the stage dimensions, the loudspeaker orientation shown in Fig. 1, and the dispersion and attenuation model presented, the direct source and monitor attenuation matrices are as follows:

$$G_{\mathrm{D}} = \begin{bmatrix} 0 & -12.72 & -12.72 & -7.93 \\ -10.75 & -7.93 & -16.68 & -11.43 \\ -10.75 & -16.68 & -7.93 & -11.43 \\ -13.89 & -10.33 & -10.33 & 0 \end{bmatrix} \tag{17}$$

$$G_{\mathrm{MA}} = \begin{bmatrix} -10.51 & 0 & -10.51 \\ -3.51 & -10.90 & -16.39 \\ -16.39 & -10.90 & -3.51 \\ -13.40 & -10.84 & -13.40 \end{bmatrix}. \tag{18}$$

The constraint equations are extracted from the user-defined mix matrix M. The elements of M have been tabulated for clarity. (See Table 2.) Limits are placed on the total SPL at each listener location,

$$\mathrm{SPL}_{\max} = 105\,\mathrm{dB}$$
$$\mathrm{SPL}_{\min} = 95\,\mathrm{dB}.$$

Weighting is applied to the constraint equations to reflect the relative importance of each listener's monitor mix. The weighting matrix is

$$W = \mathrm{diag}(10, 4, 2, 2). \tag{19}$$

The directivity of the microphone is based on the Shure SM58 polar pattern. It is an approximation to the pattern at 1 kHz and is shown in Fig. 5. It is assumed that the only microphone under consideration is the vocal microphone, which corresponds to listener 1 and source 1. The microphone polar attenuation is given by

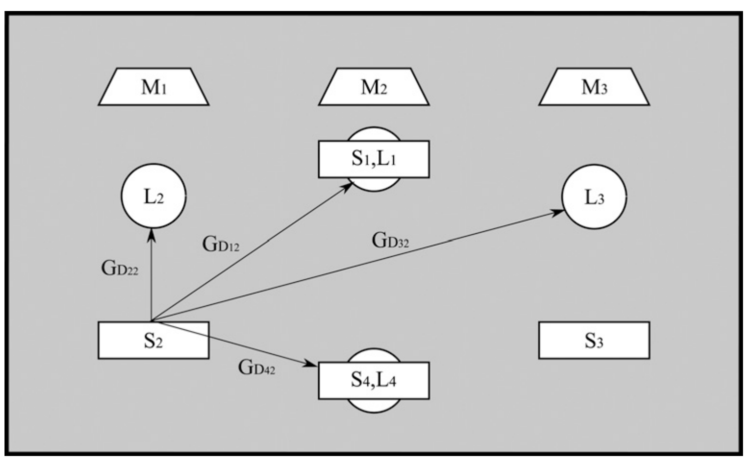


Fig. 2. Signal path of direct source 2.

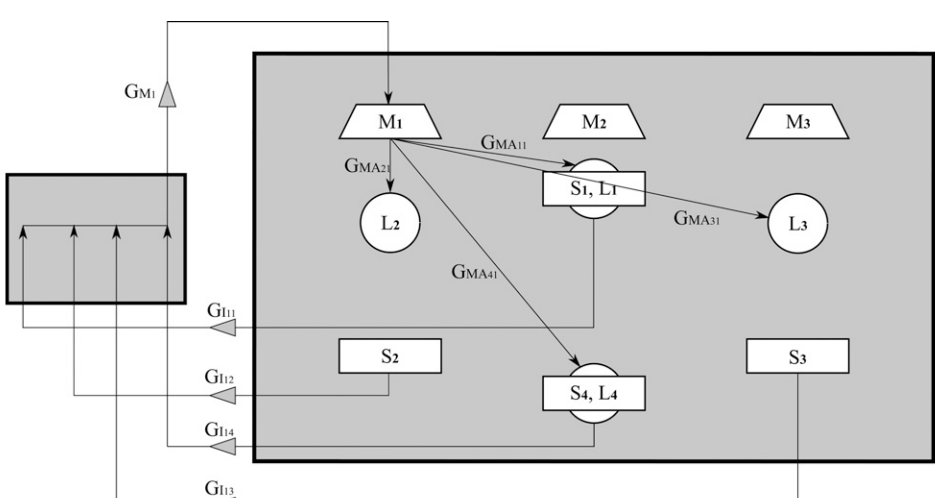


Fig. 3. Magnetic field of dual-coil motor design in Fig. 2. Units are in tesla.

$$G_P = \begin{bmatrix} -8.46 \\ -20.00 \\ -8.46 \end{bmatrix} \tag{20}$$

where G_{Pk} corresponds to the attenuation of monitor loudspeaker k. The headroom for the feedback loop is set at 3 dB. Therefore the feedback constraint is set such that $F < -3$ dB.

3 RESULTS

The mixer and monitor gain matrices that best satisfy the constraint equations are

$$G_I = \begin{bmatrix} 18.00 & -40.00 & 6.41 & 5.64 \\ 18.00 & -17.623 & -40.00 & -40.00 \\ -40.00 & 18.00 & -40.00 & 3.30 \end{bmatrix} \tag{21}$$

Table 1. Direct source sound pressure levels.

Source Number	Source Name	SPL (dB)
1	Vocals	70
2	Guitar	90
3	Bass	95
4	Drums	92

where the row index reflects the monitor k and the column index reflects the source j, and

$$G_M = \begin{bmatrix} -21.21 \\ -2.01 \\ -21.09 \end{bmatrix}. \tag{22}$$

The resultant amplitude matrix is

$$A = \begin{bmatrix} 87.51 & 84.72 & 84.36 & 85.93 \\ 78.12 & 84.61 & 83.64 & 84.06 \\ 76.82 & 86.18 & 87.70 & 83.63 \\ 76.68 & 83.70 & 85.78 & 92.55 \end{bmatrix}. \tag{23}$$

The total SPL at each listener location is

$$L = [97.76 \quad 95.00 \quad 96.49 \quad 98.49] \tag{24}$$

and the feedback loop gain is

$$F = -3.00\,\text{dB}. \tag{25}$$

Table 2 contains the target relative monitor mix levels. For example, the vocalist requires the vocal SPL to be 3, 3, and 2 dB higher than the guitar, bass, and drum SPL, respectively. The target mix has been defined based on the author's experience but by no means represents a subjective study into typical mix requirements. The actual relative monitor mix levels are shown in Table 3.

Table 2. Desired relative monitor mix levels (dB).

Listener Name	Vocal Mix	Guitar Mix	Bass Mix	Drum Mix
Vocalist	0	−3	−3	−2
Guitarist	−1	0	−3	−1
Bassist	−1	−3	0	−1
Drummer	−1	−3	−1	0

Table 3. Actual relative monitor mix levels (dB).

Listener Name	Vocal Mix	Guitar Mix	Bass Mix	Drum Mix
Vocalist	0	−2.79	−3.15	−1.58
Guitarist	−6.49	0	−0.97	−0.55
Bassist	−10.89	−1.53	0	−4.07
Drummer	−15.87	−8.85	−6.77	0

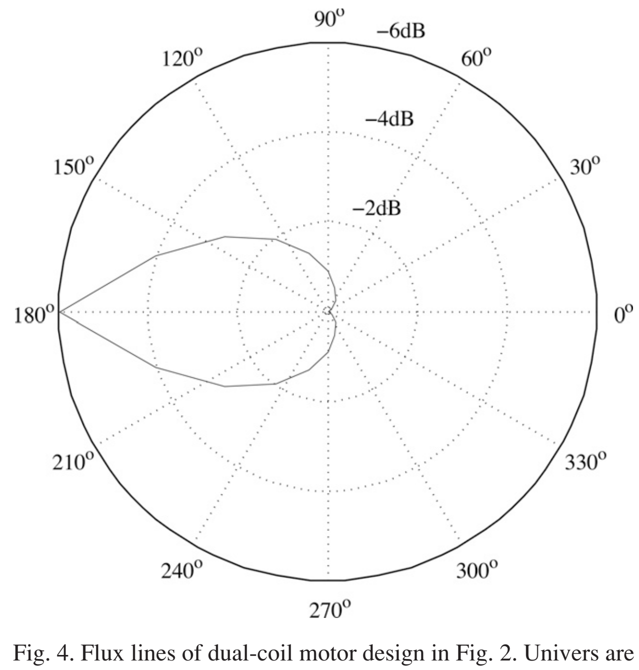


Fig. 4. Flux lines of dual-coil motor design in Fig. 2. Univers are in volt-seconds.

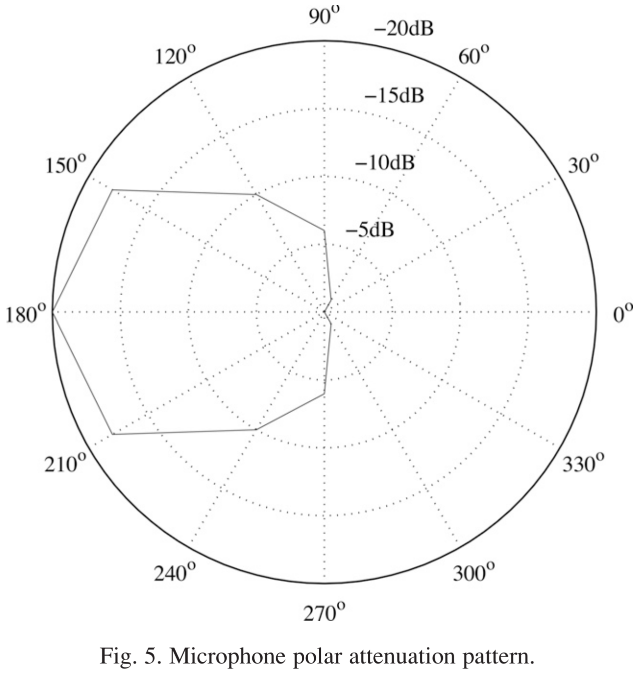


Fig. 5. Microphone polar attenuation pattern.

The differences between the target and the actual relative monitor mix levels are shown in Table 4.

4 DISCUSSION

Table 3 shows the relative monitor mix levels that best fit the desired mix given in Table 2. Table 4 shows the differences in the actual and target mix levels and the rms error for each listener. The weighting matrix [Eq. (19)] heavily favored the vocalist's requirements. It can be seen that the vocalist's requirements have been met to an extent, but the requirements for the other performers, particularly the drummer, are not met. The SPL and feedback constraints define a subspace in which the solution must exist. The size of this subspace will affect how well the mix requirements are met. Eq. (24) shows that the total SPL levels are toward the lower end of the allowable range, with the total guitarist level at the minimum. Eq. (25) shows that the feedback loop gain is at the maximum allowable level. The solution identified in this example is therefore on the limits of the solution subspace.

An investigation of the size of the solution subspace reveals features of the stage setup and the likelihood that certain mixes can be obtained. The range of possible solutions is found by evaluating the amplitude matrix using the maximum and minimum parameter values, which are 18 and −40 dB in the examples presented here. In reality a mixer channel or monitor loudspeaker can have a gain of negative infinity. If this is the case then $A_{\mathrm{I}} = [0]$ and $A = A_{\mathrm{D}}$. A finite bound was placed on the minimum gain for use in the optimization algorithm as the error function becomes very insensitive to changes in the gain parameters below 40 dB. The amplitude matrix and total SPL vector at the upper limit are

$$A_{\max} = \begin{bmatrix} 110.15 & 130.08 & 135.08 & 132.10 \\ 106.90 & 126.91 & 131.88 & 128.89 \\ 106.90 & 126.88 & 131.91 & 128.89 \\ 103.12 & 123.14 & 128.14 & 125.27 \end{bmatrix} \tag{26}$$

$$L_{\max} = [\ 142.42 \quad 139.23 \quad 139.23 \quad 135.52\ ]. \tag{27}$$

It can be seen that for the guitar, bass, and drums there is sufficient head room in the system to raise the SPL at each listener location well above the maximum SPL limit. This is not the case for the vocals, and so the vocal level will impose a limit on the solution space.

Table 4. Error in relative monitor mix levels (dB).

Listener Name	Vocal Mix	Guitar Mix	Bass Mix	Drum Mix	Rms Error
Vocalist	0	0.21	−0.15	−0.42	0.50
Guitarist	−5.49	0	2.03	0.45	5.87
Bassist	−9.87	1.47	0	−3.07	10.46
Drummer	−14.87	−5.85	−5.77	0	16.99

The amplitude matrix and the total SPL vector at the lower limit are

$$A_{\min} = \begin{bmatrix} 70.00 & 77.29 & 82.29 & 84.07 \\ 59.25 & 82.07 & 78.32 & 80.57 \\ 59.25 & 73.32 & 87.07 & 80.57 \\ 56.12 & 79.67 & 84.67 & 92.00 \end{bmatrix} \tag{28}$$

$$L_{\min} = [\ 91.93 \quad 90.25 \quad 91.78 \quad 96.55\ ]. \tag{29}$$

It can be seen that the total SPL experienced by the drummer is already above the minimum SPL when no amplification has been applied. There is a limit to the SPL that can be sent to the drummer before the maximum SPL limit is reached.

Eq. (20) shows the polar attenuation of each monitor loudspeaker with respect to the vocal microphone. This is combined with the loudspeaker dispersion to give a feedback loop gain when the mixer and monitor loudspeakers have a gain of 0 dB. This is given by G_{F},

$$G_{\mathrm{F}} = \begin{bmatrix} -18.97 \\ -20.00 \\ -18.97 \end{bmatrix} \tag{30}$$

where element k of G_{F} corresponds to the feedback loop gain of the microphone with monitor loudspeaker k. When considering each monitor loudspeaker in isolation, the maximum gain that can be applied to the microphone input signal (the vocals) is equal to the magnitude of the corresponding element of G_{F} minus the feedback limit [Eq. (25)]. When the feedback loops for all monitor loudspeakers are combined, the allowable amount of gain is reduced. In order to maximize the SPL of the vocals at the location of the microphone (which coincides with the location of the vocalist), the vocal signal is amplified by monitor loudspeaker 2 only as this has the lowest feedback loop gain. It will be equal to

$$A_{(1,1)\max} = 20\log_{10}\left(10^{\frac{S_1}{20}} + 10^{\frac{S_1-(G_{\mathrm{F2}}-F)}{20}}\right) \tag{31}$$

which includes the contribution of the direct source and equals 88.15 dB. Inspection of Eqs. (21) and (22) shows that the monitor mix identified has almost all of the vocal signal coming from monitor 2. Element (2, 1) of G_{I} is added to element 2 of G_{M} to give a gain of 16 dB, compared to the maximum of 17 dB. The corresponding vocal levels at each of the other listener locations for the monitor mix identified are 78.12, 76.82, and 76.68 dB. Inspection of Eq. (28) shows that even if the other sources are not amplified, the vocal SPL for all other listeners will be lower than the SPL of all other sources (with the exception of the guitar SPL at the bassist location, element (3, 2) of $A_{\min}$). This reduces the solution subspace significantly.

The primary requirements of the monitor mix are those of the vocalist. These state that the vocal SPL should be 3, 3, and 2 dB higher than the guitar, drums, and bass guitar, respectively. While these requirements have been satisfied to an extent, the limit to the vocal level clearly has a detrimental effect on all other mix requirements. Raising the vocal level in the monitor mix to the desired level before the onset of acoustic feedback during a musical performance is a common problem, particularly if the vocalist has a weak voice. The model can potentially be used to identify areas where the stage setup can be improved. The ideal scenario would be to increase the input SPL of the vocals, but this is outside the control of the engineer. Alternatively the levels of the other instruments can be lowered. It is difficult to do this with the drums (unless the drummer uses "hot rods" instead of wooden drum sticks), but the levels of the guitar and bass can be reduced by turning down the amplifiers. While this may not be ideal from the perspective of the musicians (as the tonal quality, especially of distortion, can be compromised), it is an option worth considering. It would also be possible to move and/or rotate the guitar and bass amplifiers. The effect of the feedback constraint can be lessened by moving monitor loudspeakers further from the microphone, and by rotating them so that they are facing away from the microphone.

The monitor mix is again optimized, but the SPL of the guitar and bass amplifiers, and the positions and angle of the guitar and bass amplifiers and the monitor loudspeakers are included as parameters in the optimization algorithm. Limits are placed on maximum displacements of 1 m in the x direction, 0.5 m in the y direction, and 90° rotation. No limits are placed on the guitar and bass amplifier levels. The resultant guitar and bass amplifier SPLs are 93.4 and 95.1 dB, respectively, and the optimized stage layout is shown in Fig. 6. The resultant errors in the monitor mix levels are shown in Table 5.

A comparison of Tables 5 and 4 shows that a significant improvement in the match with the target mix (Table 2) is found when the stage layout parameters are included in the optimization algorithm.

The optimized stage layout shows some interesting and unexpected results. The guitar and bass amplifier SPLs have both been increased, although the bass increase is by a minimal margin. It was expected that these SPLs would be reduced to increase the relative SPL of the vocals at the guitarist and bassist locations. Instead, the two amplifiers have been moved further away from guitarist and bassist, and have been rotated inward to face the drummer. This has the advantage of decreasing both the guitar and bass SPLs at the guitarist and bassist locations and increasing both of these SPLs at the drummer location, where initially there was very little of either in the mix.

Monitor loudspeakers 1 and 3 have been rotated away from the vocal microphone and moved further away from the stage front, which will decrease the gain in the feedback loop. Monitor loudspeaker 2 has been rotated toward the guitarist, which will increase the SPL of the vocals at the guitarist location. The guitarist is favored over the bassist as the guitarist requirements are weighted more heavily [Eq. (19)]. Monitor loudspeaker 2 has also been moved further away from the vocalist. This will reduce the dispersion angle of the vocal signal coming from monitor loudspeaker 2, and so will increase the vocal level experienced by all listeners.

While the approximation of the desired mix has been improved, it is still not possible to raise the vocal level significantly at the guitarist, bassist, and particularly the drummer locations. In order for this happen it is likely that an additional monitor loudspeaker would be required. There is, however, no guarantee that this will increase the vocal SPL without violating the feedback loop gain limit.

5 FUTURE WORK

In the examples considered in this paper an arbitrarily defined, desired monitor mix has been used. This enabled

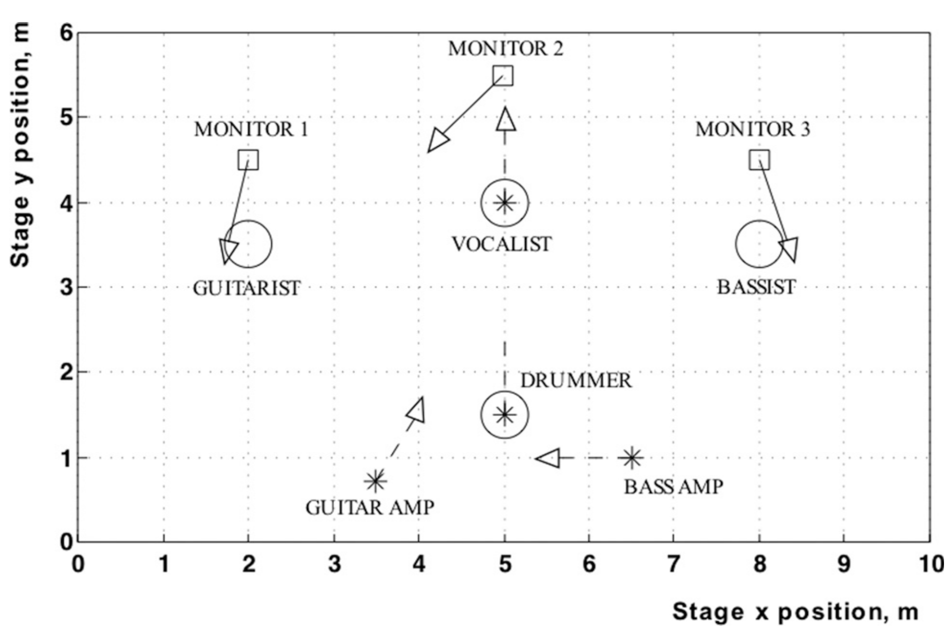


Fig. 6. Optimized stage layout.

the accuracy of the optimized monitor mix to be defined objectively. In order for the model to be used in a practical environment, consideration of the subjective requirements of a monitor mix must be included. The study of typical monitor mix requirements is a topic of future work. It may be possible to set up the monitor mix at a given venue in the traditional way and then extract the desired monitor mix. The desired mix could then be used at subsequent venues.

As discussed in Section 1.2, a number of model simplifications have been made. In order for the model to be used in a practical environment these must be addressed. It is the intention of the authors to extend the method to include frequency-dependent transfer functions. This is critical in assessing feedback accurately, and will also impact on the monitor mix requirements since masking between two sources is far more pronounced when the frequency separation is low. The incorporation of real acoustic effects, such as the room effect and the true dispersion from different types of sources, is also essential.

6 CONCLUSIONS

A model has been presented which describes the monitor mix experienced by each musician in a live musical performance. The framework of equations developed has been used to optimize the monitor mix experienced by each musician, subject to their individual requirements. It has been applied and evaluated for simple examples in a static setting. While the monitor mix could be optimized by a skilled engineer, it is not a trivial task since each listener's mix is highly coupled. The model presented enables all listener requirements to be considered simultaneously while adhering to other limitations, such as the maximum and minimum allowable SPL on stage and the prevention of acoustic feedback. It also incorporates the relative importance of the listener requirements. It has been shown that these additional constraints, in particular the prevention of acoustic feedback, reduce the size of the solution space and make it more difficult to match the target mix. It has also been shown that the position of some sources and monitor loudspeakers can be included in the optimization algorithm, enabling an improved stage setup to be found.

Table 5. Error in relative monitor mix levels for optimized stage setup (dB).

Listener Name	Vocal Mix	Guitar Mix	Bass Mix	Drum Mix	Rms Error
Vocalist	0	−0.00	−0.02	−0.31	0.31
Guitarist	−2.95	0	0.93	0.83	3.20
Bassist	−4.11	1.30	0	0.94	4.42
Drummer	−12.37	−0.04	−0.23	0	12.38

7 REFERENCES

[1] J. Gale, "Chris Trimby: Monitor Engineer," *Sound on Sound Mag.* (2006 Mar.).

[2] D. Dugan, "Applications of Automatic Mixing Techniques to Audio Consoles," presented at the 87th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 37, p. 1058 (1989 Dec.), preprint 2835.

[3] D. Dugan, "Automatic Microphone Mixing," presented at the 51st Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 23, p. 492 (1975 July/Aug.), preprint 1034.

[4] E. P. Gonzalez and J. Reiss, "An Automatic Maximum Gain Normalization Technique with Applications to Audio Mixing," presented at the 124th Convention of the Audio Engineering Society, (Abstracts) www.aes.org/events/124/124thWrapUp.pdf, (2008 May), convention paper 7411.

[5] E. P. Gonzalez and J. Reiss, "Automatic Mixing: Live Downmixing Stereo Panner," in *Proc. 10th DAFx Conf.* (Bordeaux, France, 2007 Sept.).

[6] E. P. Gonzalez and J. Reiss, "Improved Control for the Selective Minimization of Masking Using Interchannel Dependancy Effects," in *Proc. 11th DAFx Conf.* (Espoo, Finland, 2008 Sept.).

[7] B. Kolasinski, "A Framework for Automatic Mixing Using Timbral Similarity Measures and Genetic Optimization," presented at the 124th Convention of the Audio Engineering Society, (Abstracts) www.aes.org/events/124/124thWrapUp.pdf, (2008 May), convention paper 7496.

[8] D. Reed, "A Perceptual Assistant to Do Sound Equalization," in *Proc. 5th Conf. on Intelligent Used Interfaces* (New Orleans, LA, 2000).

[9] S. H. Antman, "Extension to the Theory of Howlback in Reverberant Rooms," *J. Acoust. Soc. Am.*, vol. 2, no. 2 (1965).

[10] D. Davies and E. Patronis, Jr., *Sound System Engineering*, 3rd ed. (Elsevier, Oxford, UK), pp. 402–404.

[11] D. Troxel, *Understanding Acoustic Feedback and Suppressors* (Rane Corp., Arlington, VA, 2005).

[12] T. Waterschoot and M. Moonen, "Comparative Evaluation of Howling Detection Criteria in Notch-Filter-Based Howling Suppression," presented at the 126th Convention of the Audio Engineering Society, (Abstracts) www.aes.org/events/126/126thWrapUp.pdf, (2009 May), convention paper 7752.

[13] H. Kuttruff, *Room Acoustics*, 4th ed. (Elsevier, London, UK).

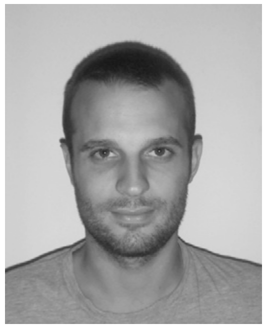
[14] D. M. Howard and J. A. S. Angus, *Acoustics and Psychoacoustics* (Reed Educational and Professional Publ., Elsevier, Oxford, UK, 2001).

[15] M. A. Wolfe, *Numerical Methods for Unconstrained Optimization* (Van Nostrand Reinhold, New York, 1978), pp. 53–73.

[16] P. E. Gill and W. Murray, *Numerical Methods for Constrained Optimization* (Academic Press, London, 1974), pp. 21–28.

[17] Mapp Online Pro, http://www.meyersound.com/products/mapponline/pro.

THE AUTHORS



M. Terrell

J. Reiss

Michael Terrell received his first Ph.D. degree in aeronautical engineering from the University of Bristol, Bristol, UK, specializing in structural dynamics. After working in the finance sector for a number of years he returned to academia in the field of digital audio and is presently a postgraduate student within the Centre for Digital Music at Queen Mary University, London, UK. He is an experienced drummer and music producer, and the focus of his Ph.D. research is the development of intelligent mixing systems. This includes tools to be used to assist the setup up and design of performance spaces, and algorithms that can be used to automate aspects of musical production.

●

Josh Reiss received a Ph.D. degree in physics from the Georgia Institute of Technology, Atlanta, GA, specializing in analysis of nonlinear systems. He is a senior lecturer with the Centre for Digital Music at Queen Mary University of London, UK. He made the transition to audio and musical signal processing through his work on sigma–delta modulators, which led to patents and a nomination for a best paper award from the IEEE. He has investigated music retrieval systems, time scaling and pitch shifting techniques, polyphonic music transcription, loudspeaker design, automatic mixing for live sound, and digital audio effects. His primary focus of research, which ties together many of these topics, is on the use of state-of-the-art signal processing techniques for professional sound engineering.

Dr. Reiss has published over 70 scientific papers and serves on several steering and technical committees. As coordinator of the EASAIER project, he led an international consortium of seven partners working to improve access to sound archives in museums, libraries, and cultural heritage institutions. His services to the Audio Engineering Society include being a member of the Review Board for the *Journal*, a member of the Technical Council, and vice chair of the Technical Committee on High-Resolution Audio. He has written numerous papers for the AES *Journal*, conferences, and conventions, as well as participating in and organizing workshops and tutorials. In 2007 he was chair of the 31st AES International Conference, "New Directions in High Resolution Audio," and in 2009 he was general secretary of the 35th AES International Conference, "Audio for Games."