MULTISOURCE MONITOR MIXING

TIME ALIGNMENT, YES OR NO?

Otto Kroymann, Adlib Audio

ABSTRACT

In today's world on stage the focus is more and more shifting to an all-out performance. The days where a musician was a static fixture on stage delivering a musical performance with a little help from lighting and special effects are long gone.
With the heightened mobility on stage, the work of a monitor engineer has undergone radical changes. The old concept of single source monitor mixing has long reached its limits, as the artists are no longer in a small defined space but take the whole stage as their performance space. Therefore, a monitor engineer has to provide a multi-source monitor mix for the artists in order to be able to reach the whole performance area.
One of the biggest problems of multi-source monitor mixing is sound field coherence. In a setup with centre wedge speaker and side fill speaker the sheer audio path difference is a potential disaster. For any monitor engineer the fundamental question in is: Time alignment or not?
The paper will analyse psychoacoustic effects of multisource monitor mixing and discuss the pros and cons of time alignment versus unaligned sources and the impact on the artist's experience

CONTENTS

1. INTRODUCTION

In a live musical performance the stage monitor sound is a vital element for the performance. A good monitor mix helps the artist to create the ultimate show either by using In Ear Monitors (IEM) or 'traditional' stage monitors. Today's performer is no longer 'glued' to one spot on the stage as in times when their microphone was cable bound but they use the whole stage and therefore, a single source monitor mix is no longer an option. With using IEMs for every performer on stage free movement is no problem at all but IEMs are criticized for giving the performer a sense of isolation and disconnection from fellow musicians as well as from the audience. For these reasons many performers prefer stage monitors. Not compromising in the artist's performance and quality of the mix the set-up and mixing of a sophisticated multisource monitor system is a challenge.
A show is more than hearing it is feeling the music. In general monitor mixing is as much, if not more, about being able to translate human emotions and the performer's feelings into a sound image as it is about technical expertise and skills.
This paper is a hands-on guide addressed to researchers in order to show them what is important from a sound engineer's point of view. It is also addressed to practitioners to give them insight in the theoretical background of their work and to strengthen discussion about different ways to create a multisource monitor setup and mixing on it.

2. THE AIM OF A MONITOR MIX

The aim of a monitor mix is to create the performers preferred space. “A good monitor mix is like being snuggled up in your favourite blanket and told that everything is going to be OK”[1]. This statement is monitor mixing in a nutshell. As a monitor engineer you have to give your performer the confidence and the security that everything is going to be fine, that he can concentrate solely and exclusively on his audience and his performance.
In that aim the monitor engineer has to have a multifold approach. On the one hand it is essential to have a close relationship with the artist and to understand his personality and his needs, on the other hand the engineer has to know the all aspects and possibilities of their technical tools. In creating a monitor mix that exceeds the coverage area of a single source there are two fundamental acoustic phenomena the engineer will have to deal with on any given stage.
These two are comb filtering and the Haas effect.

3. COMBINATION PHENOMENA OF ACOUSTIC SOURCES.

a. Comb Filtering

$$|H(\omega)| = \sqrt{1 + 2g\cos(2\pi f D) + g^2}$$

Figure 1: Equation for the calculation of comb-filtering

The combination of two or more sources leads to a direct interaction of the merging signals. Depending on the phase relationship between the signals comb filtering occurs. Parts of the signal will be attenuated while other frequency ranges can be emphasised. In the practical approach this can lead to the loss of audibility of individual instruments or the loss of distinction between instruments[2] (two things sounding similar…)

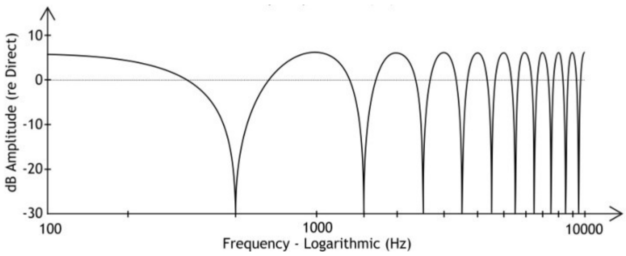

Figure 2: Comb-filtering

b. Doubling of signals

According to the theory of Helmut Haas the human ear will not recognise two signals of similar content as individual signals if the time of arrival difference is less than 10 – 20ms. If two signals are within this ‘Haas window’ the directionality cue will be taken from the first arriving wave front. The secondary is, regardless of the real direction, perceived as being the same signal coming from the same source. Any time of arrival differences outside the ‘Haas – window’ are perceived by our ears as two discrete signals either from different locations or even from the same source.

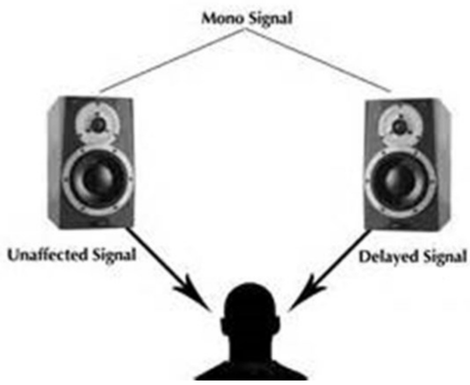

Figure 3: The Haas-Effect[1]

[1] B. Cowan
http://www.lsionline.co.uk/video/Meyer%20Sound/Mixing%20Metallica%3A%20Backstage%20Video%20Interviews%20with%20Band's%20Dual%20Monitoring%20Team/8eps1a, 20.10.2015

[2] http://recordingology.com/in-the-studio/delay/comb-filter-calculations/

4. Multiple Sources

Looking closely at the following example setup, certain problematic areas can be easily identified. The most obvious is the apparent path difference and therefore inherent time of arrival difference (Δtoa) between the centre wedge speaker right in front of the performer and the side fill speakers at either side of the stage.

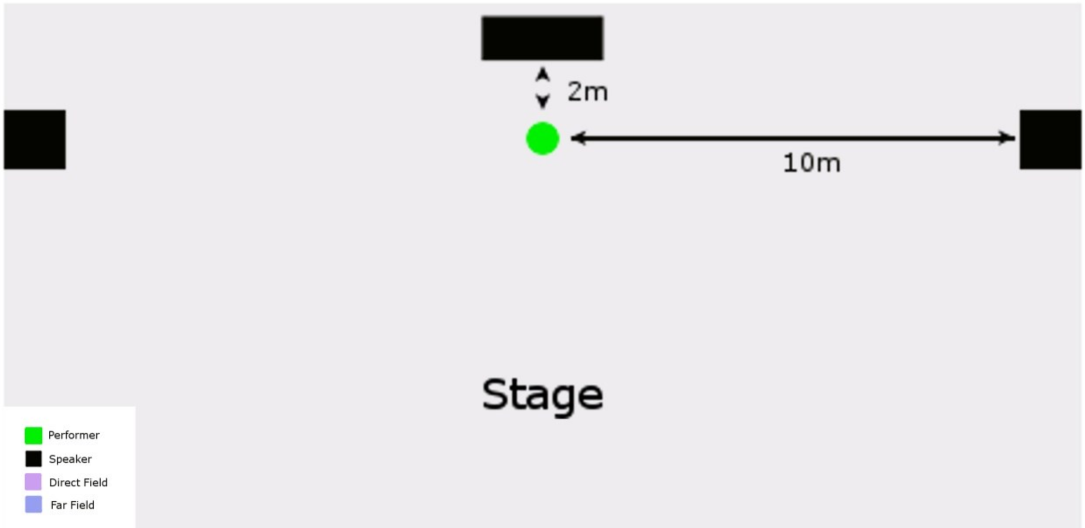

Figure 4: Basic Setup

a. Unaligned sources

Left without time alignment of the various sources, the setup is as shown below. Looking at the graph, the previously determined problematic areas are even more obvious. The performing position at centre stage, even though it is in the direct sound field (light red) of the centre wedge speaker, is still possibly prone to far field (light blue) interference stemming from either side fill speaker.

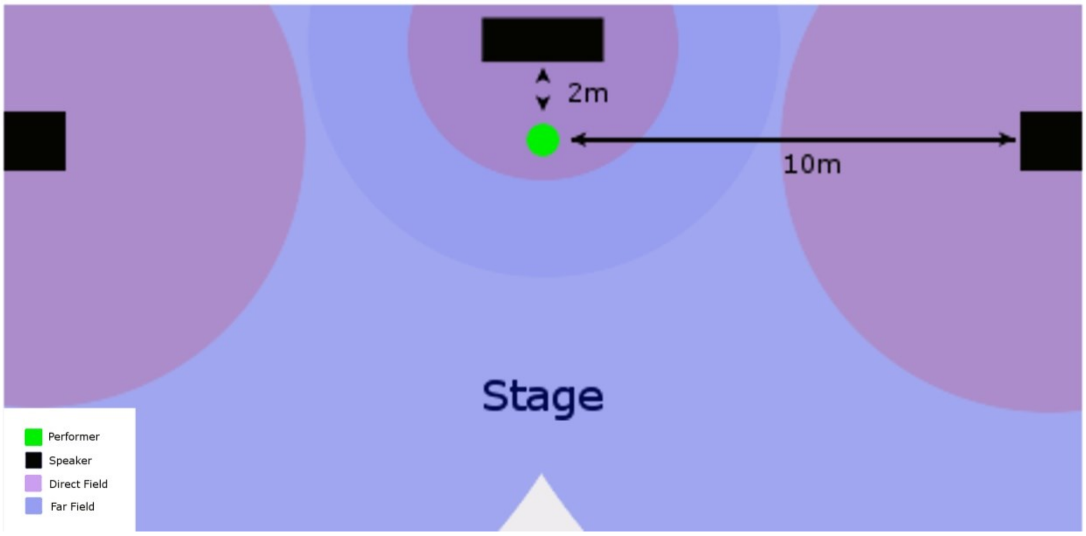

Figure 5: Basic Setup unaligned sources

The main advantage of the unaligned setup is that the performer will have a very direct and immediate sound field. The system latency is kept to a bare minimum.

The latency can be described as

Latency = microphone path + inherent system latency + system delay + speaker path

In this, the microphone and the speaker path is the physical distance between the transducer and the acoustic source or receiver, respectively. These factors can only be influenced by a physical change and, in the case of a monitor speaker, by a reduction of usable performance space. The inherent system latency results from the electrical pathway between microphone and speaker. Today one of the predominant factors are the conversion processes in and out of the digital domain. These conversions are with the use of digital consoles unavoidable. Therefore, the only factor that could be changed is the remaining system delay. In the case of unaligned sources this is at zero.

Psychoacousticly, here the performer will be at an advantage as he will be able to exploit the effects of visually aided hearing[3]. This means that the human ear is capable of filtering signals from the background noise according to the theory of the 'Cocktailparty – Effect' even under conditions where general signal to noise ratio is not favourable. This is achieved by visually 'anchoring' the signal stream to a distinct source.

In the example the stage setup is fairly standard for a decent size concert. The stage is 20 meters wide. For the centre position of the performer this means, that the side fill speakers are 10 meters

[3] E. Picou, T. Ricketts, B. Hornsby
How hearing aids, background noise, and visual cues influence objective listening effort.
Ear Hear. 2013 Sep; 34(5):e52-64.

(or 30.3ms) away from the performer. The centre wedge speaker in front of him is 2 meters (or 6.06ms) away from his ears. At his position the time of arrival difference between the different sources is 24.24ms. Looking at the phase relationship between the two signals at 200Hz results that, according to $\varphi = \Delta t * f(Hz) * 360°$, the two signals are 1745.28° out of phase. This means the side fill signal is 305,28° out of phase with the 4th wave front of the centre wedge speaker.

At this position due to the level difference between the two sources (in this example we are assuming equal sound pressure levels from both sources) - resulting of the proximity to the centre wedge speaker - the interference signal from the side fill speakers will not result in destructive comb filtering. The same is true if the performer is moving to the respective position in front of the side fill speakers.

Radically different on the other hand is the situation moving out of the direct sound field of both sources. This area is depicted by the dark blue area in Figure 6. In this dark blue area the two sources are at similar sound pressure levels and are therefore directly interacting.

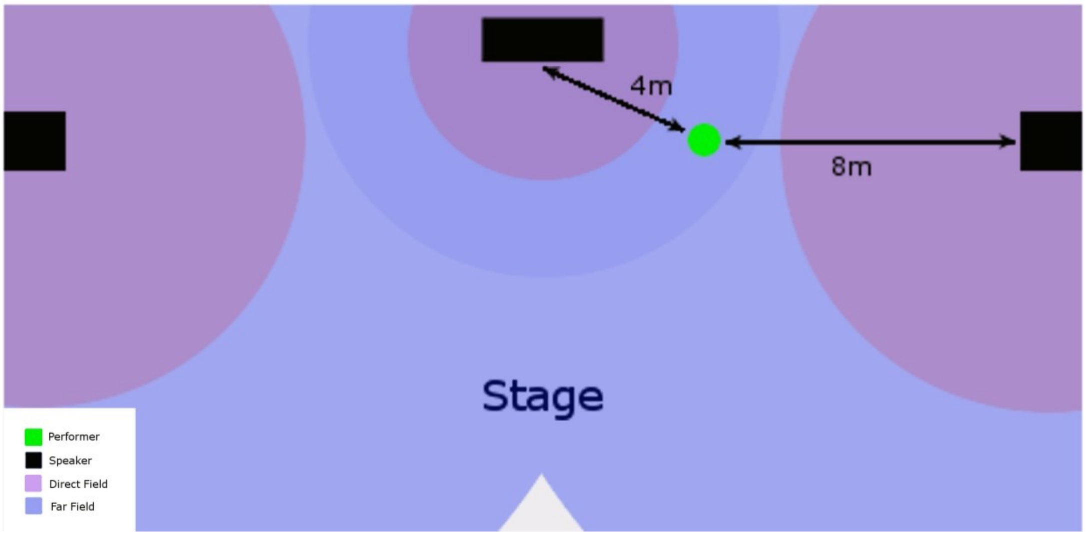

Figure 6: Transition area for unaligned sources

For a performer in this area this means that he is 4 meters (or 12.12ms) from the centre wedge speaker and 8 meters (or 24.24ms) from the side fill speaker. Therefore the ΔTOA is 12.12ms. According to $\varphi = \sqrt{1 + 2g\cos(2\pi f D) + g^2}$ this results in destructive comb filtering. For example at 125Hz a cancellation would occur ($f(Hz) = \frac{\Delta t * 360°}{\varphi}$).
Furthermore, as ΔTOA > 10ms, the signals might be already outside the 'Haas-window' and therefore the performer might hear two discrete signals which will lead to musical timing difficulties. This is especially valid in the case of predominantly percussive signal components.

b. Time aligned sources:

In light of these issues, the other option is the time alignment of the various monitor sources. In our example the physical setup is still the same as before. The changing factor is that now the system delay in the equation

Latency = microphone path + inherent system latency + system delay + speaker path

with regards to the centre wedge speaker will not be zero. By adding electronic delay the centre wedge speaker, whilst physically not moving, is now virtually further away from the performer.

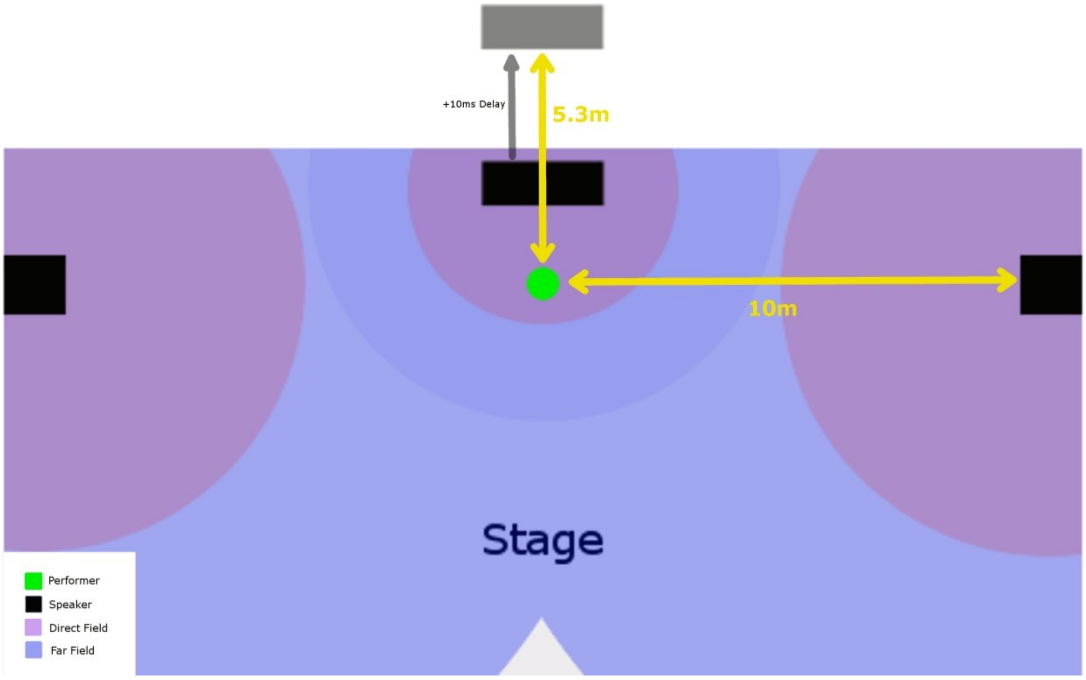

Figure 7: Basic Setup time aligned sources

The advantage of the virtual move rather than the physical one is that the direct sound field is still reaching the performance area, same as before. The required delay time for the example stage is 18.18ms. Unfortunately, this delay time is resulting in an overall latency in a region that musical timing is becoming progressively difficult. For a professional system the added delay time should not exceed 8 – 10ms[4]. In professional monitoring systems this is resulting in an overall system latency of less than 13 – 15ms.

[4] This is coming from personal experience of the author.

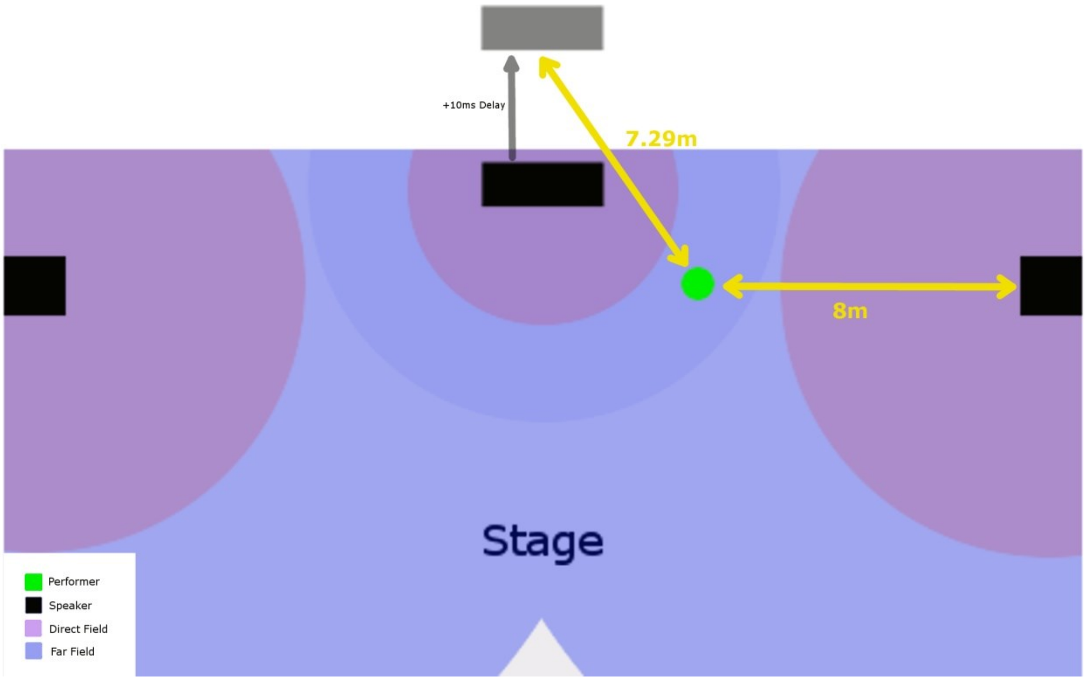

Figure 8: Transition area for time aligned sources

The result of the correct application of added delay to the centre wedge speaker is perceived by the performer as an emersion in a coherent sound field. This coherence leads to a less destructive transition area between the multiple sources. The aligned sources are combining across the stage. The most destructive combination areas are now well within the relative direct sound fields of the respective speakers and are therefore less predominant and less audible. The performer has now a more acoustically unified performance area rather than 'islands' of usable performance space throughout the stage.

Unfortunately, this approach to multisource monitor mixing does not come without its unique set of drawbacks. The predominant one being the fact that the above mentioned effect of visually aided hearing does not apply in this setup. In the creation of the coherent sound field the human ear is now losing the help of the visual anchor to the source as the perceived acoustical source does no longer coincide with the visual. This does not allow the brain to apply the above mentioned effects. Most often this will lead to the monitor mix having to be a bit louder than before.

5. Conclusion

Looking at the two different approaches it becomes apparent that in light of the respective advantages and also the drawbacks of each one, these two approaches both have their time and their place. It depends very much on the overall situation and also the preference of the performer, which approach is the one to be applied. As a tendency, taking specifically the drawbacks into account, it seems that the first approach using unaligned sources might be more useful on louder stages, whereas the use of time aligned sources might work a bit better on a quieter stage focussing on one main performer.

6. References

G. M. Ballou (ed.)
Handbook for Sound Engineers, The New Audio Cyclopedia, 2nd Ed.
Focal Press;1998.

W. Ahnert, F. Steffen
Beschallungstechnik, Grundlagen und Praxis
S. Hirzel; 1993.

B. Cowan
http://www.lsionline.co.uk/video/Meyer%20Sound/Mixing%20Metallica%3A%20Backstage%20Video%20Interviews%20with%20Band's%20Dual%20Monitoring%20Team/8eps1a, [02.10.2015]

D. Davis, C. Davis
Sound System Engineering, 2nd Ed.
Focal Press; 1997.

J. Eargle, C. Foreman
Audio Engineering for Sound Reinforcement
Hal LeonardCorp.; 2000.

J. Eargle
The Microphone Book
Focal Press; 2001.

T. Görne
Mikrofone in Theorie und Praxis, 2nd Ed.
Elektor-Verlag GmbH; 1996.

E. Picou, T. Ricketts, B. Hornsby
How hearing aids, background noise, and visual cues influence objective listening effort.
Ear Hear. 2013 Sep; 34(5):e52-64.