ADX TRAX

Insights on Audio Source Separation

Audionamix has served the professional audio community for years, lending expertise and creative solutions to individuals and companies at the forefront of the music, film, and broadcast industries. Delivering pristine voice isolations for everything from blockbuster movies and 3-D animated shorts to virtual duets by legendary artists, and exciting new remixes by top DJs and music producers, the company is positioned as the global leader in audio source separation. Audionamix now provides some insight into the world of audio source separation through the lens of its new software release, ADX TRAX.

By
Rick Silva
(United States)

ADX TRAX is designed to help producers, DJs, mash-up artists, and creative audio engineers isolate vocal or melodic motifs to create new hooks, innovative remixes, or virtual duets. The software also enables users to raise or lower the vocal track or a solo instrument in a mix during the mastering process without requiring access to the multi-track master sessions. ADX TRAX enables users to "duck" vocals or melodic instruments within a mixed track to keep the audience focused on the dialogue, voice-over talent in a commercial, or to meet profanity requirements for radio mixes. ADX TRAX is used by the production specialists at Audionamix to deliver its world renowned vocal isolations.

Melodic Audio Source Separation

Automatic voice activity detection (AVAD) and melodic audio source separation are the core components of TRAX software. Upon importing an audio file, TRAX quickly analyzes the content and sends all relevant information to Audionamix's cloud-based servers where the AVAD process begins. Once the process is complete, the servers return an estimated "pitch guide." What appears in the interface is a zoomed in, spectral view of the melodic content the AVAD process was tracking to perform the automatic separation process. Audionamix calls this view its "pitchogram" overview. The TRAX pitchogram shows the main melodic content's fundamental frequencies along with the estimated pitch guide used to separate the melody from the music (see **Figure 1**).

Following the AVAD process, the estimated pitch guide is sent through Audionamix's Vocal Extraction algorithm (VEX), then TRAX generates a separation. As a result, you get the first separation almost instantly (see **Figure 2**).

It's important to note that the AVAD process only creates an estimation of the main melodic content. Although it saves time by calculating the pitch guide estimation, AVAD cannot define the melodic content's nuances nearly as accurately as you can. This is where the manual refinement process begins.

Once the initial pitch guide appears in the pitchogram, you can see where it missed some of the desired melodic content. To correct these imperfections, choose from the standard editing tools at the bottom of the interface to create a customized pitch guide to zero in on the desired melodic content for separation. In places where the melodic content's amplitude was too low for the AVAD algorithm to track, you can use the piano

roll editor to audition the melody's pitch that you are trying to separate. Then, apply a marquee tool selection to force VEX to separate melodic content within a specifically defined frequency range and time duration. Unlike editing within a full range spectrogram, you are only required to make your pitch guide align with the fundamental frequency of the melodic content you are trying to separate. Then TRAX automatically extrapolates the upper and lower harmonic partials during the separation process. Once all your refinements are complete and you've created a custom pitch guide, click the "Separate" button located at the bottom of the Separate screen and a refined separation is created and listed in your separation bin (see **Figure 3**).

Nonlinear Editing, Spectral Editing, and Audio Source Separation

Since the release of TRAX, there have been numerous requests for spectral editing features. However, TRAX is source separation software, not a spectral editor. Although they may appear to be very similar, they are indeed very different.

A waveform overview display of audio as shown in **Figure 2** represents the amplitude of the audio over time. This is a great view for basic cut, copy and pasting techniques to remove or rearrange audio. By zooming in to the sample level, you can even make minor corrections and fix clicks or pops with a pencil tool. Regardless, any edits made to waveform overviews are going to affect the entire frequency range of the audio, and there is no way to edit individual frequencies or even visualize the melody with just this waveform overview.

On the other hand, with a spectral editing software's spectrogram view, you can visualize and edit frequency specific content over the full frequency range of audio because it displays time on the X-axis and frequency on the Y-axis. Amplitude is usually displayed by brightness, with louder sounds being brighter and softer sounds being darker. Having this view of the audio is incredibly useful for carefully chipping away from
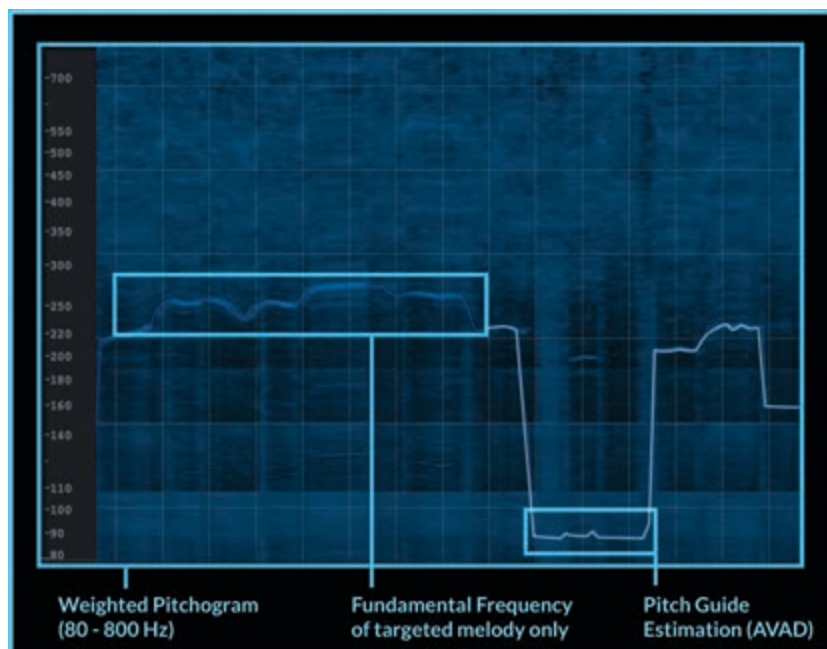

Figure 1: The melodic content is shown within an 80-to-800-Hz frequency range. You can see the automatic estimation of the pitch guide traced over the melodic content.

the content you are trying to isolate (e.g., isolating specific melodic content or vocal from a mastered recording). With a spectral view of the audio, you can almost immediately identify melodic content, including its fundamental frequency and most of its harmonic partials, as well as all of the other content a spectral view of the audio reveals (see **Figure 4**).

Through the use of spectral editing tools, you can directly edit by time and frequency, applying gain to specific sounds, attenuate others, or remove them entirely from the frequency spectrum. You can also apply spectral repair, decrackling, and many other useful audio restoration techniques. Using a common spectral editor and applying a lot of time, effort, and tedious audio manipulation, you can achieve a fairly decent level of isolation.

The process of Melodic Audio Source Separation in digital signal processing aims to identify a specific
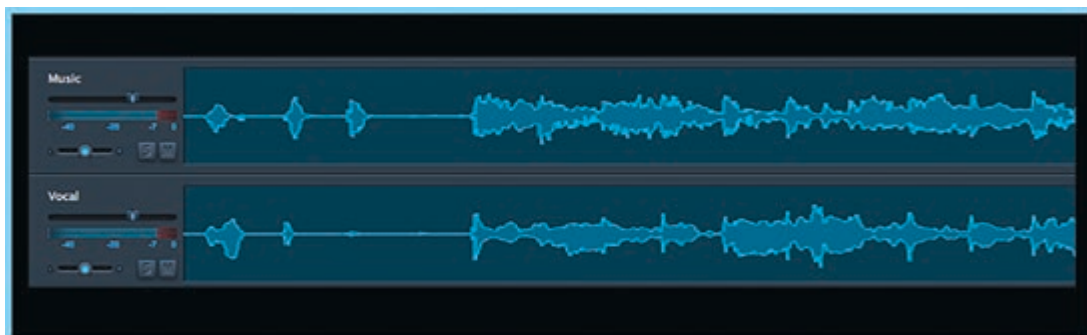

Figure 2: The automatic separation is displayed as typical waveform overviews—a music file and a vocal file.

Figure 3: The Separations Window shows the editing tools, Guide Tone controls, Transport, Separate Button, Separation Bins, and separated music and vocal files.
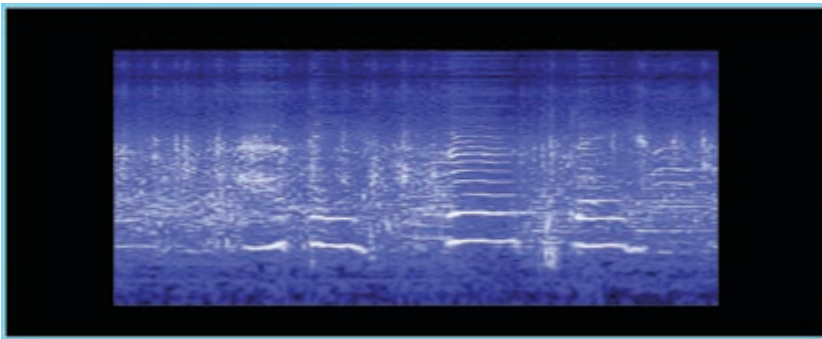

Figure 4: A full mix of music in a standard spectogram shows the fundamental frequencies of melodic content, all its harmonic partials, and a visual representation of everything else in the audio file.
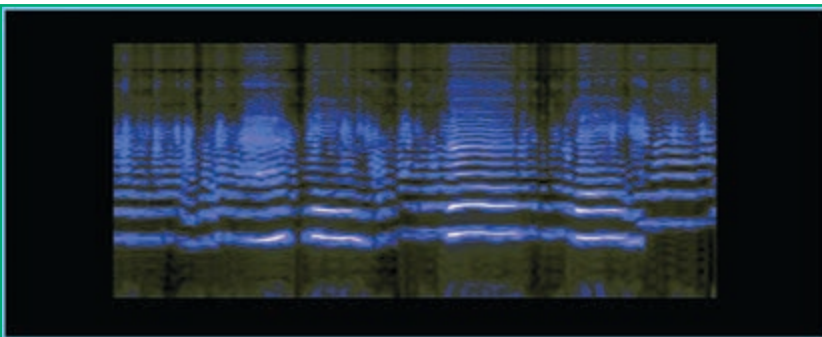

Figure 5: The spectogram view of an isolated vocal is shown after an ADX TRAX automatic separation is complete and before it is refined.

signal from a blend of many different signals that have been mixed together and then separate it from the mixture. For this reason, pure audio source separation is often referred to as "the holy grail" to almost every professional audio engineer. Audio engineers have known for more than 100 years that once signals are mixed together and combined into one master recording, you cannot go back and "un-mix" that master recording. That is why audio engineers strive to capture a great, live performance of a band in one take. Otherwise multi-track recorders are needed to access the separate audio "stems" and re-record any undesirable performances or re-mix unbalanced levels among the instruments or the voices.

"Un-mixing" audio is what Audionamix is known for and ADX TRAX clearly demonstrates the ultimate first step of the "un-mixing" process. With an understanding that TRAX is a melodic audio source separation program and not a spectral editor, you'll realize that once the TRAX separation and refinement process is complete it is much easier to go through the final steps of audio clean up using spectral editing tools to achieve a stunning vocal isolation. **Figure 4** showed a significant amount of information in the spectrogram when importing a full mix. Now compare the same audio clip after it is run through ADX TRAX (see **Figure 5**).

Notice how the melody is much more isolated from the rest of the instrumentation in the audio file. The results shown in **Figure 5** occurred just after the ADX TRAX automatic separation process, prior to refinements, and only took a matter of minutes. Simply starting from the results of TRAX in a spectral editor is an incredible time saver when trying to isolate a vocal.

To add some perspective, if your goal was to reach this amount of isolation with a typical spectral editor it could take anywhere from two to 10 times the amount of manual effort than using TRAX and more often than not, the result would be completely different from one user to the next. Remember, an incredibly high skill set is required to not damage the file in a typical spectral editor. Additionally, if any of the vocal audio is deleted from the spectral work you've done, there is no way to retrieve it—the only thing that remains is the isolated audio. The rest of the audio has been deleted from your project.

Although fundamental frequencies may be easy to identify using a spectrogram view, the target vocal contains many harmonic partials. These harmonic partials are crucial for maintaining the timbre, clarity, and integrity of the voice. When using a spectral editor, most of the upper partials

are difficult to see and isolate. By simply refining the main pitch guide-line using Audionamix's "pitchogram," and its user-friendly editing tools, TRAX automatically calculates the harmonic partials of the melody and separates as much of the original voice as possible. When TRAX isolates a vocal, it transfers everything else to a second, complementary file (the music file). These files, when combined, will always equal the sum of the original track. This is one of the key differentiators with the ADX approach.

This unique process has immediate every-day use for digital audio enthusiasts. You can raise or lower vocals from a master recording 6 to 9 dB and adjust pan positioning in the stereo field without requiring the multi-track master sessions. This is great for touch-ups in the mastering or re-mastering process, creating profanity law remixes, or lowering melodies in expensive sync music so it doesn't pull focus of the center channel dialog in a television or movie mix. Vocalists, guitarists, brass, and woodwind players can even use this software to play along with an actual band. To take melodic isolation and instrumental creation to the next level, Audionamix unveils its unique approach to spectral editing with the release of ADX TRAX 2.0 Spectral.

ADX TRAX 2.0 Spectral

While the TRAX audio source separation process does most of the heavy lifting when trying to isolate vocals or melodic content, spectral editing tools are perfect for removing unwanted content in your TRAX separations. But remember, with traditional spectral editors, once the information is removed, it's gone for good. So what if you want to put the remaining music from your vocal track, back into your music track to get a better instrumental backing track? And what if there is a bit of missing vocal still in the music track? How can you get that back?

ADX TRAX 2.0 Spectral offers a unique approach to spectral editing software that is seamlessly integrated into the familiar interface and workflow of the original TRAX software. Now, you can separate the vocal from the music track, refine and re-process your results, create a comp track to get the "best of" separation, and use spectral editing tools.

ADX TRAX 2.0 Spectral includes the most important standard spectral editing tools, but the result is not standard at all. ADX TRAX 2.0 Spectral provides a workflow that remains non-destructive and continues to work between two files (vocal and music). This means any spectral changes you make to one file are automatically reflected in the other.

For example, after processing with TRAX you might notice that part of a word has been left out of the vocal file. You can enter the spectral editing view, look at the music file, and identify the missing vocal part. Use a lasso tool to select it, and then use the cut function. You'll see it disappear from the music track, and automatically reappear in the vocal track. You can instantly switch back and forth between tracks to hear the results. This straightforward, non-destructive approach to spectral editing is the type of innovation that has kept high-profile clients requesting professional services from Audionamix for years. With the release of ADX TRAX 2.0 Spectral, Audionamix is finally starting to share their advanced technology and creative workflow with the world.

ADX TRAX 2.0 Spectral will be showcased at the 137th Audio Engineering Society (AES) Convention from October 9–12, 2014 at the Los Angeles Convention Center in Los Angeles, CA. During the convention, ADX TRAX experts will demonstrate ADX TRAX 2.0 Spectral, the all-in-one audio software for creating vocal isolations, instrument samples and instrumental tracks for a multitude of creative uses. ax

Author's Note: As a result of the continued support from audioXpress, *Audionamix is offering a special 25% off discount on all ADX TRAX plans to all* audioXpress *subscribers through October 31, 2014. Visit www.audionamix.com/shop and enter the promo code AXpress25 at checkout through the end of October to receive this special offer. We hope you enjoy using ADX TRAX and look forward to providing new and exciting technology to the pro audio community.*

About the Author

Rick Silva is the Vice President of Production and Product Management for Audionamix, which is a global leader in audio separation. Based on years of audio signal processing research, the company developed its patented ADX Technology. Rick is also a studio owner, published author, and video artist for Hal Leonard and Alfred Publishing. Rick works as an audio engineering instructor at the Musicians Institute in Hollywood, CA. He is also a musician/guitar player and a recording, mixing, and mastering engineer.