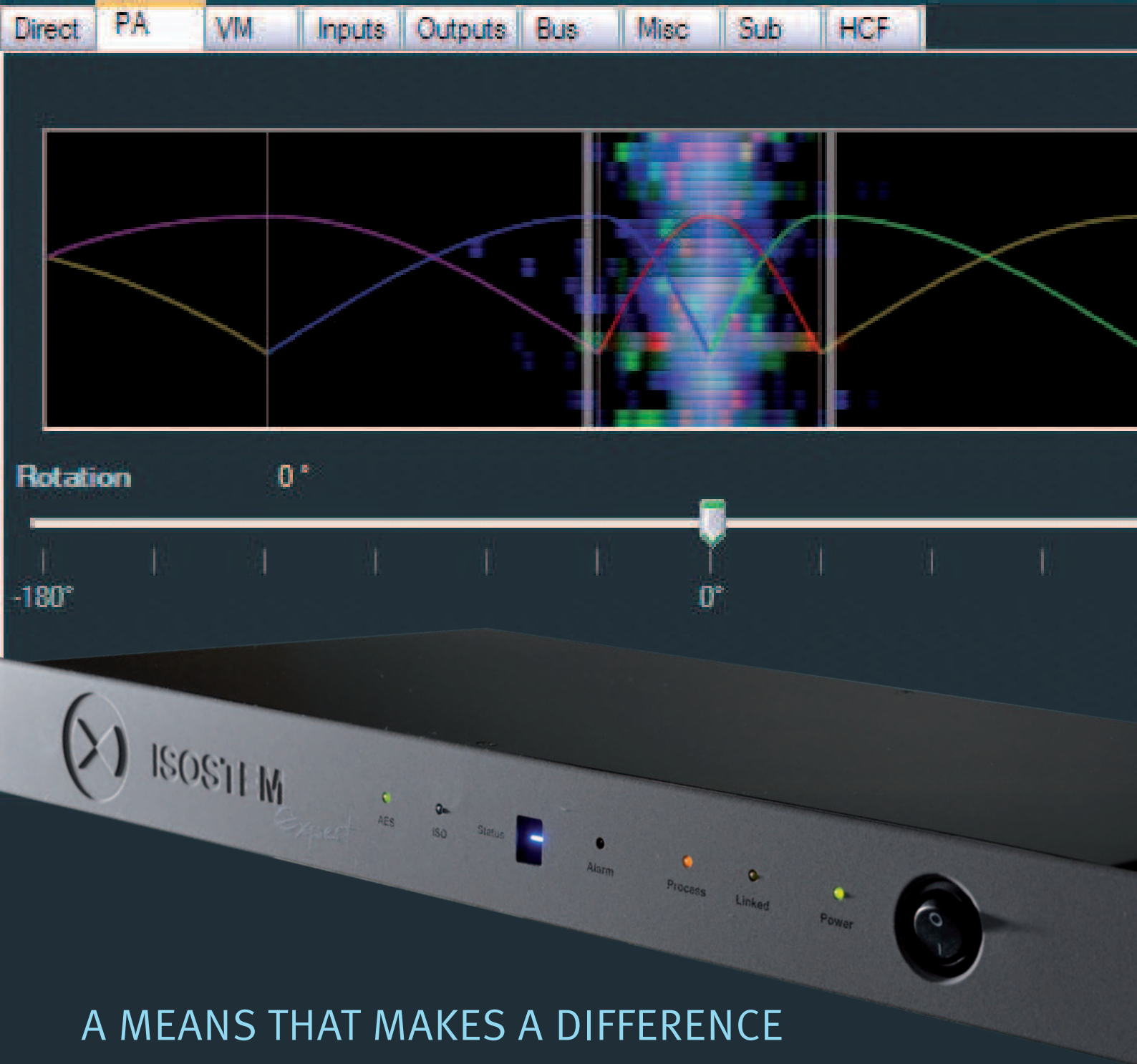
FACHZEITSCHRIFT FÜR PROFESSIONELLE AUDIOTECHNIK 05 2012

studio magazin

TEST: DSPECIALISTS ISOSTEM

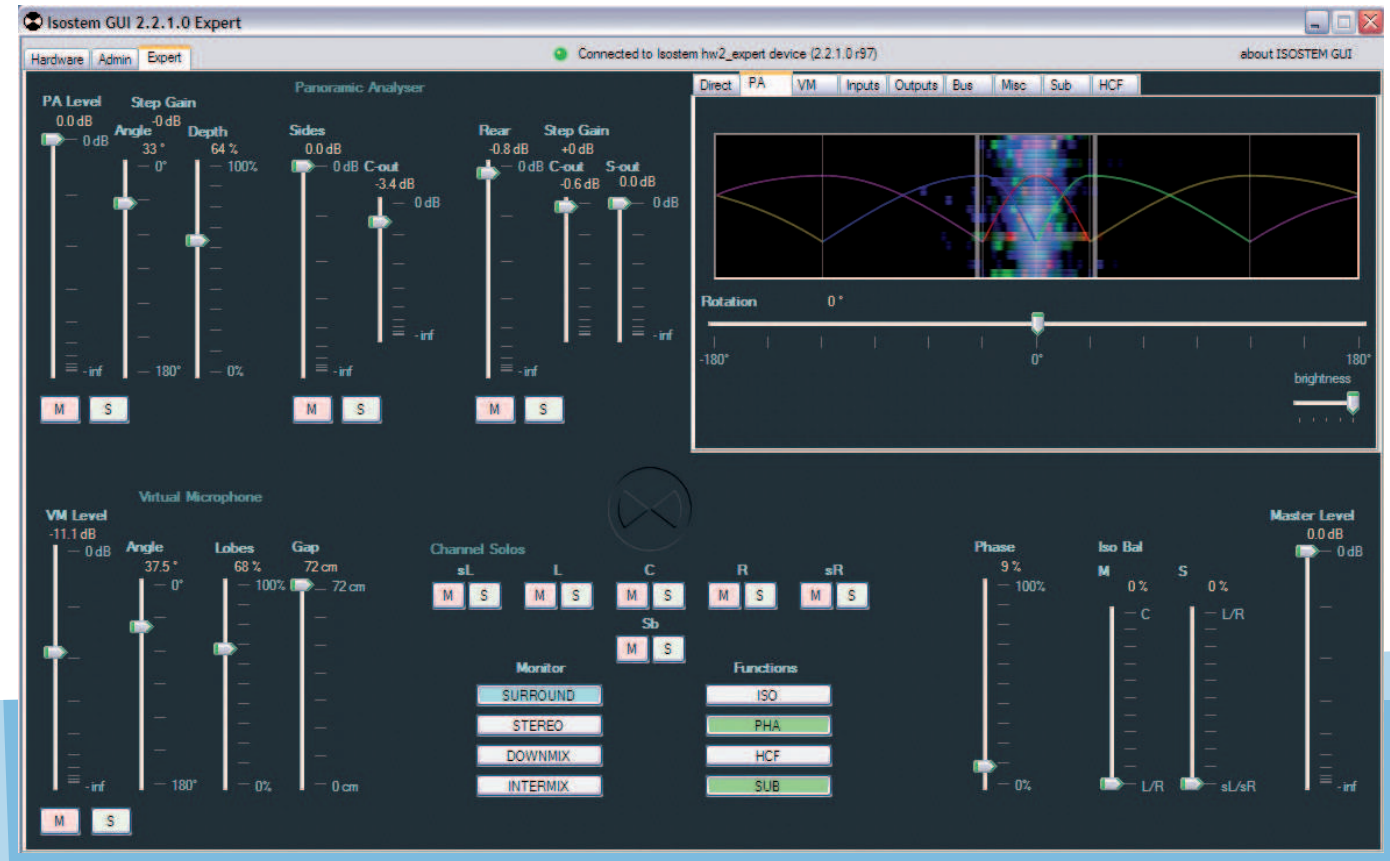
REPRINT



A MEANS THAT MAKES A DIFFERENCE

DSpecialists Isostem **Fritz Fey, Screenshots: Fritz Fey**

Today, 5.1 surround is quite common to most of us – even though the format has not yet made its well-deserved breakthrough in many areas and the majority of (mainly music producing) recording studios is still based on stereo. And while even more complex multichannel formats featuring 11 or more channels are getting trendy, 5.1 is gaining acceptance in broadcast – slowly yet steadily. Thanks to DVD and Blu-Ray, surround-enabled home theaters are becoming more and more common; HDTV and huge full-HD-enabled screens have their share as well. When 5.1 surround was introduced, I was confident that the listening experience when compared directly to stereo would make the audience quickly want to listen to nothing else but surround sound. While this hope (which I have expressed on these pages, too) has not fulfilled as quickly as I had expected, the consumers have actually taken note of the difference between stereo and surround and have already started appreciating it. The main reason why surround is not yet consistently used in broadcast is that the archives are filled with stereo programs, and the majority of recordings are still done in the stereo. There are but a few studios specialized in creating stereo and surround mixes in parallel – and thus at a reasonable cost. Today, the principal that producing in surround is expansive still holds true. And actually it sometimes is – for example, the expense of real-time surround recording of location sound would not be justifiable.



The user interface of the Expert version shows all relevant parameters. The spectrogram in the multifunctional window currently shows the filter effect and the acceptance angle.

As stations cannot keep switching between stereo and surround material, and because it would not be possible to create 5.1 mixes of existing stereo programs due to the lack of source material, broadcasters and producers alike would definitely welcome a technical approach of lifting original stereo material to the surround level in a convincing manner. The solution comes from DSpecialists and is named “Isostem”. Founded in 2003, Berlin-based DSpecialists specialize in developing products for digital audio processing and metering. Using my TC System 6000 and the Unwrap software running on it, I have myself already converted tons of live-music recordings from stereo to “fake” surround – music that is sold today on “surround DVD” in record stores and online shops. My customer was so happy with the results that he even preferred a synthesized surround version to a surround mix made entirely from scratch (where the original multitrack source material still existed). However, that was all about life performances – not studio productions.

With an approach based on spectrum analysis of pressure gradients and the detection and distribution of direct and diffuse sound portions, Isostem produces results that without a doubt must be referred to as stunning. Admittedly, the extremely complex software user interface for controlling the DSP hardware and concepts hardly any tonmeister will be familiar with clearly indicate that getting satisfying results is definitely not a piece of cake; however, thanks to hardware components and storable presets, the system smoothly integrates into the broadcast chain. What is more, settings can be adapted to various contents using GPIO in order to yield optimum results.

Overview

In 2005, Antoine Hurtado, honorary professor at the National Conservatory in Paris, filed a patent for his invention, the Isostem system, and presented its basic concepts at the AES Convention in New York. Hurtado assumed that dominance sources within a multichannel signal could be detected using spectrum analysis. This would allow for broadly separating direct and diffuse sound portions from each other using filters – similar perhaps to separating mid and side portions of a stereo signal. Simply said, the so-called “Panoramic Analyzer” identifies the distribution of sound energy in space, thus detecting signal portions containing direct or diffuse sound. Those sound portions can then be isolated and distributed to the appropriate surround channels with regard to artistic or acoustic aspects. This is where the ear of the experienced tonmeister comes into play – after all, the

system's "technological intelligence" provides the tools but no turnkey solutions. The method bases on filter banks. It produces no additional signal portions but just redistributes the existing ones. One might refer to this as a reorganization or segmentation of the stereo signal. In addition to the Panoramic Analyzer, there is the Virtual Microphone – another tool that can be used either separately or in combination with the PA method. The Virtual Microphone is based on an acoustic model that applies the existing information onto a virtual 5.0 microphone array. In practice, this is similar to playing back the stereo signal with two speakers and recording it with that array – but without the drawbacks of speaker playback. The emulation actually ignores the LFE signal – definitely a realistic approach given typical recording scenarios. Using these two processing components, which as mentioned can be used either separately or in combination, one can convert even highly complex source material into an excellent surround signal. Each module features its own mixer, so you can create complex mixes in the surround domain. My experience shows that following this approach is actually necessary in almost every scenario; but it leads to sometimes truly spectacular results that exceed all expectations.

The system comprises a DSP hardware unit in a 19" 1U enclosure and control software. The software incorporates a complex signal router, a mixer, and a "processing stage" for critical parameters and is currently available for Windows operating systems only. Signal processing is computed on a Sharc processor by Analog Devices. This ultra-low-power unit (3 watts) requires no active cooling. Audio processing occurs at a 32-bit depth, and dynamic filtering is based on 64-bit processing at a fixed rate of 48 KHz.

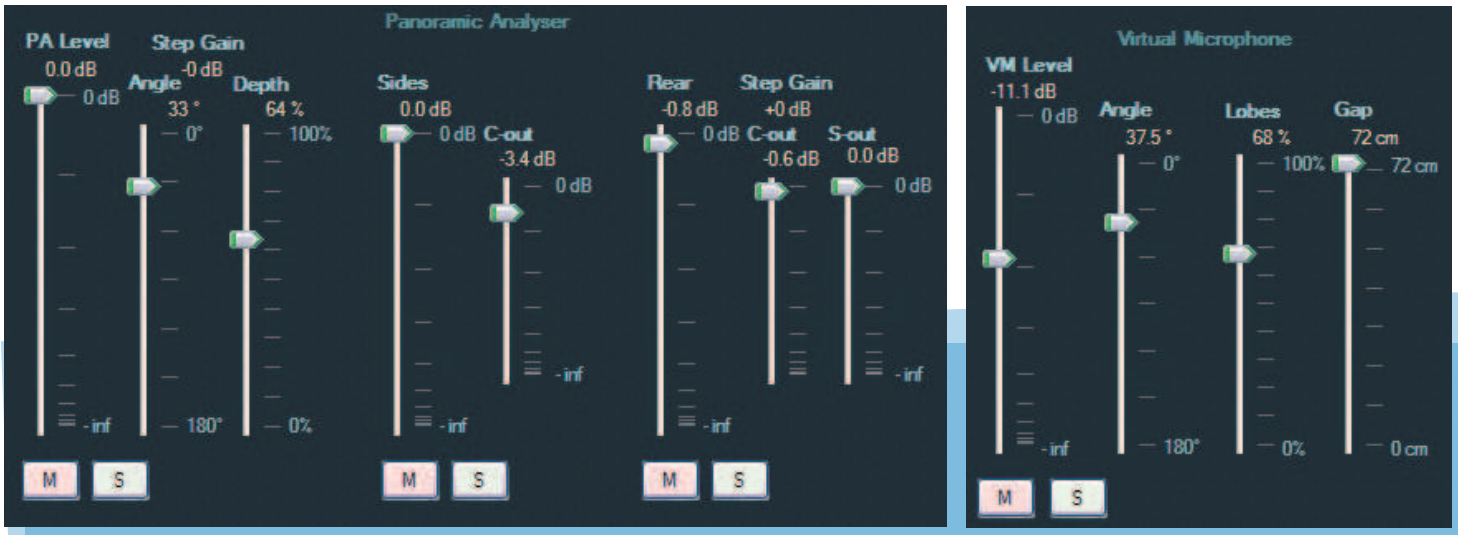
At first, the user interface looks quite straightforward. It features virtual sliders, which are very useful for mouse operation, and comes without those notorious encoders and flip switches that normally interfere with the workflow and just serve as eye catchers. Communication occurs over a serial interface. As these are rare on today's motherboards, a USB-to-serial converter is used for generating a virtual COM port. The implementation of that converter went absolutely smoothly even on the somewhat outdated Windows XP computer in our studio. Clocking options include a separate wordclock line, the AES signal, or an internal generator. When using the Isostem, keep in mind that it uses a fixed sample rate of 48 kHz. After a successful configuration of the system, all indicators on the front panel of the DSP unit will light.

There are no physical controls on the hardware unit – all parameters are controlled by the software. Two system configurations are available at different prices. The Expert version provides access to all system parameters while the Live version is a hardware-only version: It provides no parameter access but operates on the basis of presets created using the Expert version. This is useful, in particular, in broadcast scenarios where immediate parameter changes are not desirable. Analysis and filtering cause a basic latency of 75 ms; however, that can be synchronized to standard bit rates. The port configuration is Tascam-compliant with 25-pin D-sub terminals, each carrying eight AES inputs and eight AES outputs.

The User Interface

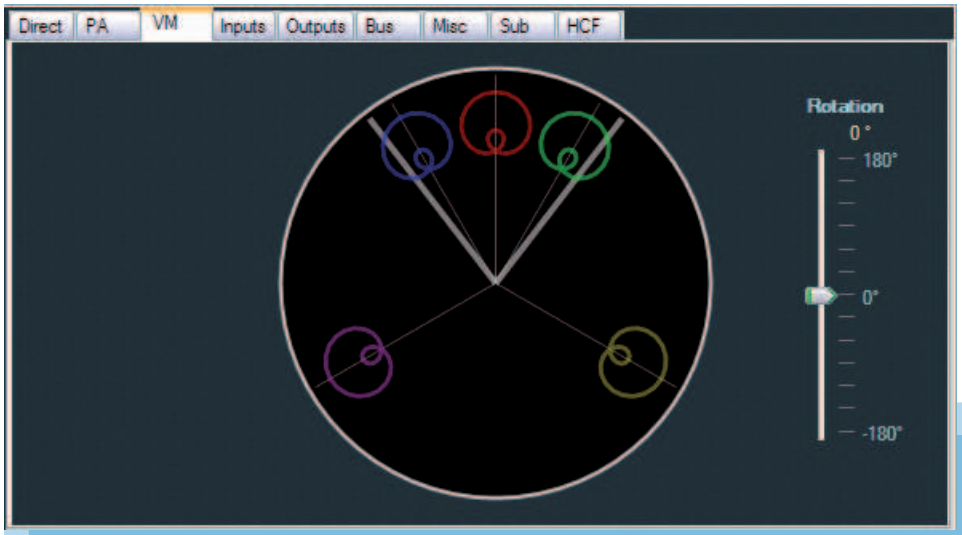
There is no need to install the software: You can start it right away by launching an EXE file. When the setup is complete, even experienced tonmeisters will frown and start scratching their heads. Looking rather simple at first sight, the parameter set unfolds to a powerful tool structured with countless submenus and tabs. The main window named Current is sub-

Our test setup. The USB-to-serial converter was used to generate the COM 5 port. The software uses that port for communicating with the hardware.

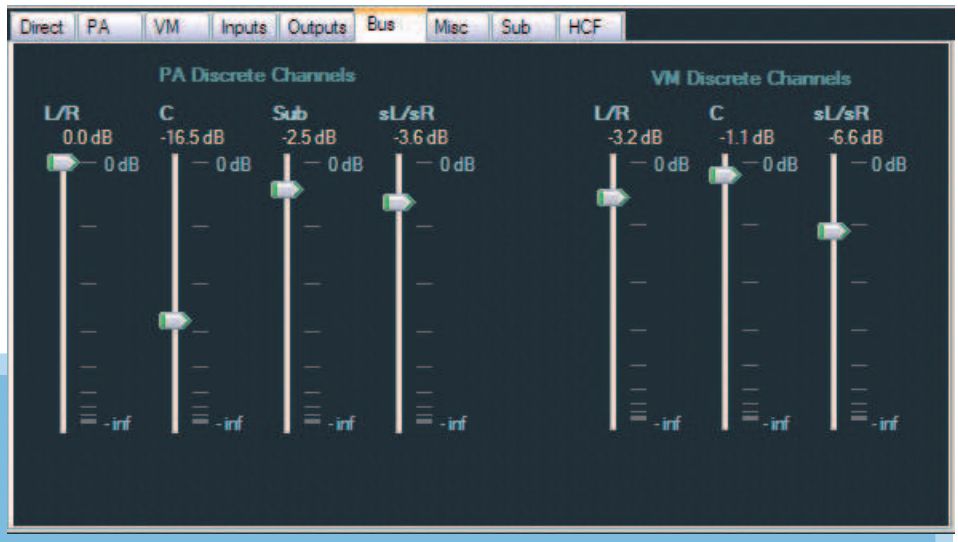


PA Mixer and VM Mixer: The Panoramic Analyzer and Virtual Microphone modules each have their own parameter mixer.

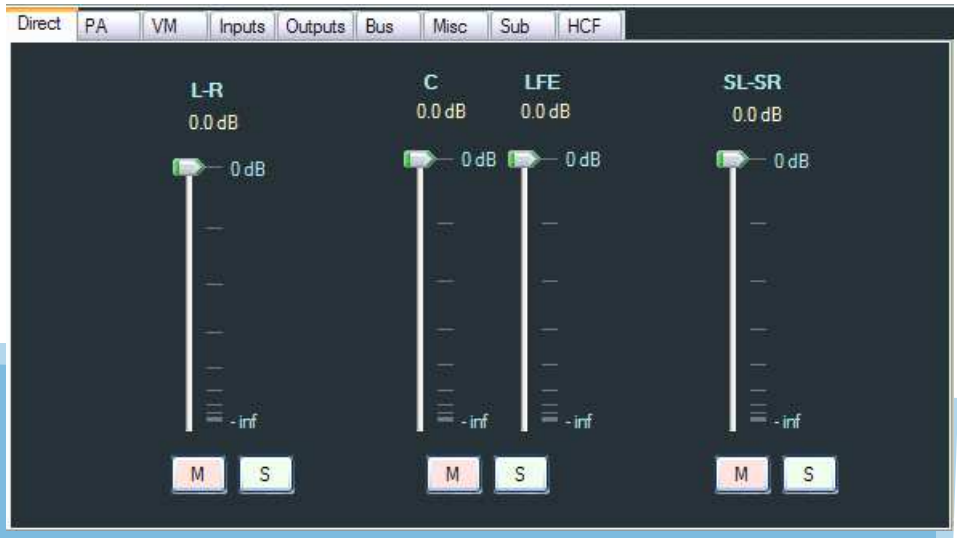
divided into four “action areas”. One of those areas allows for selecting between nine sub windows that give access to signal routing, graphical views, the bus mixer, and filters. You will soon become familiar with that structure and find out quickly, which windows should be permanently displayed and which ones need to be open only now and then. The action areas include the Panoramic Analyzer (or, more precisely, its “mixing console”); the Virtual Microphone with its own parameter set; the multifunctional window (router, bus mixer, filters, etc.); and the Master area, which provides solo/mute functions, the master level, a phase control, and the ISO Balance parameter. (More information about the latter two is to follow in a moment.) A total of eight controls is available for operating the Panoramic Analyzer. The Analyzer not only analyzes but also controls the underlying dynamic filter banks. First you might want to open the PA tab in the multifunctional window where you will find a graphical view of the filter effect plus another slider used for controlling the rotation of the overall simulation. This is useful, for example, when correcting the orientation of an unbalanced stereo panning. As you can see from the illustration, there are multiple filters operating in overlapping ranges. The acceptance angle can be set using the Angle slider. Two delimiting lines represent the acceptance angle between 0 and 180°. Between those lines, the constantly varying input stereo signal is displayed as a “spectrogram”. Using the Depth control, you can manipulate the filter depth. Used in combination with the master fader, this control forms the basis for separating direct and diffuse (reverberant) sound portions with more or less high accuracy. There is no rule of thumb: which depth setting is appropriate depends on the character of the source material, or, in the end of the day, on your aural perception and taste. The Sides and Rear slider sets have a subtractive characteristic in order to precisely separate direct and diffuse sound portions from each other. In other words, adding the Sides signal (L/R) creates the reverberant counterpart to the direct signal on the left and right front channels of the surround signal. The same is true for the Rear slider set: here, too, purposeful subtraction leads to a more or less accurate direct-sound imaging on the rear channels. The Sides and Rear sets each include a Cen-



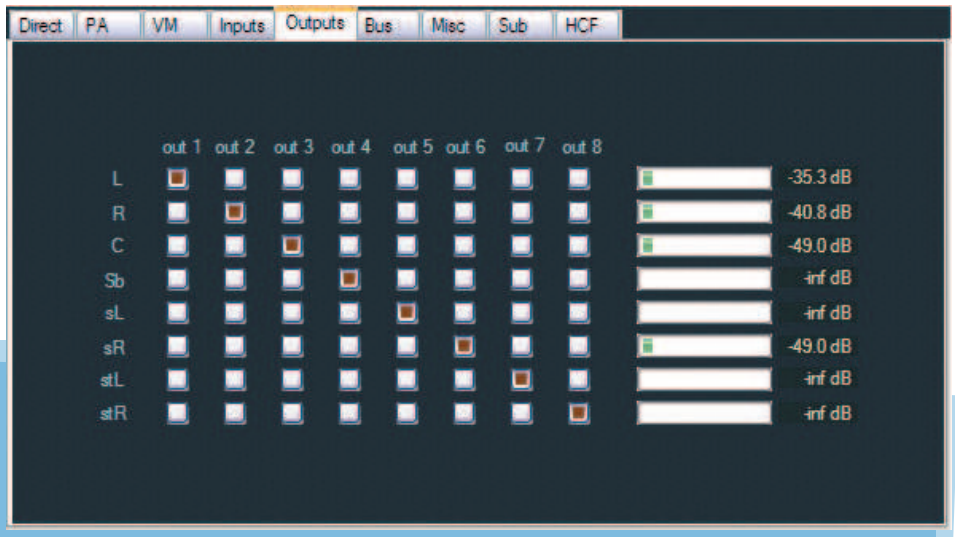
The virtual microphones are shown on a surround circle. You can adjust the acceptance angle, the polar pattern, and the distance between the microphones.



The bus mixer allows for flexibly mixing the two emulation models.



The input mixer used for integrating ready-to-use surround material



The output matrix with the meters

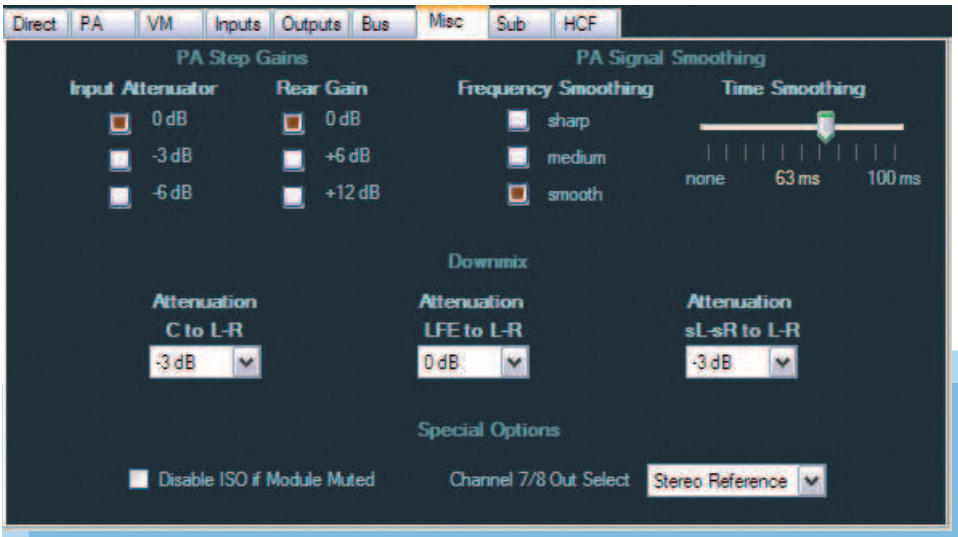
ter Out slider that causes level-dependent subtraction, too. This allows for selectively reducing or even fully eliminating the direct sound portion of the stereo signal's phantom center from the front and rear channels of the surround signal. In addition, the Rear section also includes a Sides Out slider, which can be used for adjusting the diffuse sounds portions on the two front channels in relation to the direct sound on the rear channels. All these controls interact with each other to a certain extent, so when changing one setting, it might be necessary to readjust other parameters as well.

As already said, the Virtual Microphone is a different emulation model. When using it, you should first open the corresponding graphical view in the multifunctional window. It shows the classic surround circle with five microphone positions. Unlike with the Panoramic Analyzer, the Virtual Microphone emulates a recording scenario using a 5.0 microphone array rather than applying subtractive techniques. A look at the figure illustrates the situation. The four sliders are used for defining the properties of the microphone array: the overall level, the acceptance angle, the polar pattern, and the distances between the microphones (0-72 cm). The polar pattern is set for all virtual microphones. It can be adjusted continuously from omnidirectional to cardioid to bidirectional. As you can use the Virtual Microphone in parallel to the Panoramic Analyzer and each module features its own surround mixer, you can combine any signal portions from the two emulation models with each other. The master section provides functions for phase, ISO balance, and master level. The conversion from stereo to surround introduces a negative phase correlation between the front and rear surround channels and the diffuse portions extracted there because the correlated portion of the original stereo signal may "pile up" in the center. This is compensated using the Phase control: a single parameter reintroduces a phase shift into the emulated model, thus restoring positive correlation between front and rear.

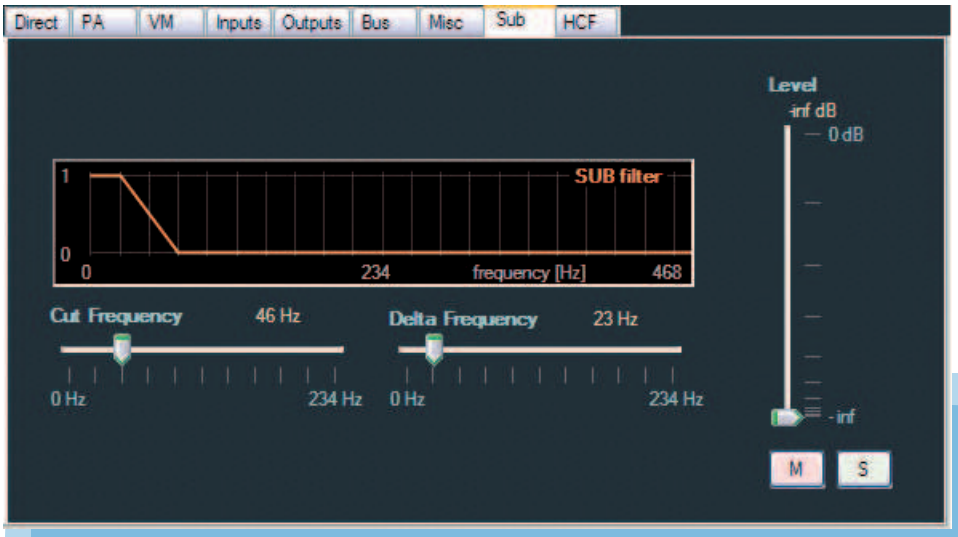
The ISO function provides settings for en-

suring full compatibility between the downmix made from the generated signal and the original stereo material. The so-called “intermix” – the differential signal from the downmix and the stereo source – feeds the surround model, making the downmix identical to the stereo signal on the mathematical level. This amount of compatibility may not be needed in all situations; you are free to go without it in order to purposefully redesign the surround stage. Finally, we will describe the monitoring and functional sections before we turn to the multifunctional window and its parameters. (Some elements of that window will anyway be referenced below.) The monitoring section provides a full set of solo and mute keys for these six surround channels plus a matrix allowing for monitoring either the emulated surround signal, the original stereo signal, the stereo downmix, or the intermix. The function block includes buttons for the ISO, Phase, HCF, and Sub functions, which you can enable or disable separately in any combination. HCF is a slope filter for high and mid-range frequencies. The filter depth and rate are continuously adjustable. Use that filter for “designing” a pleasant sonic image on the surround rear channels by cutting the high and mid frequency ranges.

This is where you process the diffuse and/or direct sound portions on the rear channels by attenuating their definition or presence on the frequency domain. The Sub function is a kind of bass management: Here, you set the separation frequency and the filter slope; frequencies below that setting will go to the LFE channel. Regardless of the setting, however, the five surround channels will produce the full signal range. If you disable the Sub function, the generated LFE component will be rerouted to the two surround front channels. The bus mixer accesses all signals coming from the emulation modules and allows for freely adjust the level relations. Using the Direct module, you can feed an external surround source to the internal bus and mixer. The fact that a total of eight inputs is available indicates that the unit is capable of processing a mix

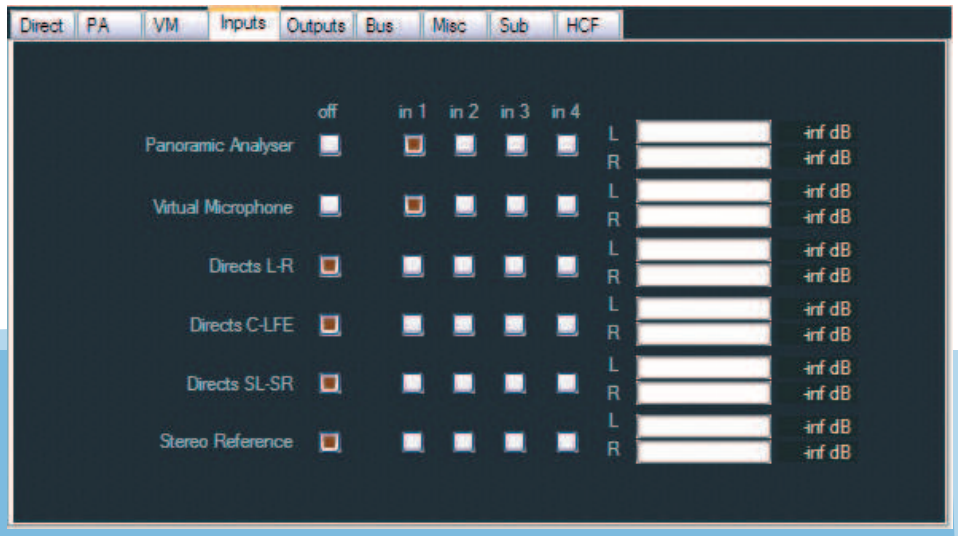


Level, smoothing, and downmix parameters in the same window



LFE extract filter

Design filter for rear surround channels



Input matrix for emulated and genuine surround material

of native surround signals and stereo material as well – for example, when processing a playlist with mixed formats in automated GPIO operation. As already mentioned, the purpose of the design is to create a consistent surround-program format from surround as well as stereo sources. To achieve this, the Expert version allows for storing presets to the computers disk; the operator can subsequently load and process those presets in the broadcast environments using the Live version of the system.

The Multifunctional Window

The nine tabs in this window provide essential setup, operating, and feedback functions. Let's start with the Direct window, which we already described above. Basically, the presence of eight AES digital inputs allows for connecting one surround and one stereo source at the same time. Using the I/O matrix, you can route the signals on the system. Direct is an input mixer that directly accesses the internal bus mixer. There are four sliders (L/R, C, LFE, and LS/RS), each with dedicated mute and solo buttons. If you have set up the system properly, there will always be a surround output signal – no matter whether the source has been a routed-through surround signal or one generated from a stereo signal.

"PA" is a visual utility that shows the filter's acceptance angle and depth and also provides a rotation control used for correcting the orientation of the emulated surround image. "VM" displays the placement of the virtual surround microphones, their front acceptance angles, and their polar patterns. "Inputs" in an input matrix that allows for routing all relevant signals from the AES inputs: PA, VM, direct to L/R, direct to C/LFE, direct to LS/RS, and "Stereo Reference." Its counterpart is the "Outputs" tab that allows for routing the L, R, C, LFE, LS, RS, Lt, and Rt to the eight available AES outputs. "Bus" is the internal bus mixer for discreet PA and VM outputs (PA: L/R, C, LFE, LS/RS; VM L/R, C, and LS/RS, as virtualization is limited to five channels). The Sub tab shows the filter cutoff frequency and slope as a filter-curve diagram and also includes all necessary sliders. "Misc" provides a small set of parameters for level adjustment and the downmix parameters. This is where you get a separately enabled input-level attenuator (0, –3, or –6 dB) plus an amplifier section for the rear surround channels (0, +6, or +12 dB). "PA Signal Smoothing" is an important parameter set allowing for smoothing the filter setting (three settings are available) and for adjusting the integration of the dynamic filter effect (0…100 ms), which corresponds to smoothing the emulation process on the time domain.
Finally, this is also where you set the downmix attenuation for specific signal path: C to L/R, LFE to L/R, and LS/RS to L/R. "HCF" controls the already mentioned slope filter that allows for attenuating high and middle frequencies on the LS/RS channels to your taste. Here again, a filter-curve diagram and frequency, slope, and depth controls are available.

In Practice

I had spent quite a few hours of bewilderment and joy with the Isostem before I started slowly getting to the point and creating satisfying results. Each time, you have to keep in mind what you are actually using the system for: Do you want to create a generic preset for a particular kind of audio program – for example, live reporting, sports broadcasts, specific musical genres, game shows, or talk shows? Or would you rather produce an amazing surround signal from a stereo mix and optimize its acoustic features? As I am talking to pros here, it goes without saying that approaching your objective purely intuitively and without knowing the key parameters does not really make sense: The system is just too complex, and the effects of the parameters are to "powerful"; and what is more, many parameters even interact with each other. Only when you have understood the functions of the effects of all parameters, you will be able to act purposefully. If you consider this bad news, rest assured: With a little practice, you will be able to produce extremely convincing results. However, be sure to create an accurate balance using the Panoramic Analyzer and/or Virtual Microphone mixers. Sometimes, considerable differences in level can occur in the resulting signal when fully removing the direct sound portions from the channel. As this can be done even on the L/R front channels, the center channel may become the only one to carry direct sound, thus creating an unproportional overall program. During my tests, I

learnt how to create a well-balanced relation between the direct and diffuse portions on the surround front channels while producing a surround envelope by using only generated diffuse sound on the rear channels. This way, I came up with clear front-channel localization plus an amazing spatial impression. Experimenting with direct sound portions sometimes leads to beautiful localization effects of signals that exist on the outer sides of the stereo panorama. Of course, this requires monitoring the signal in a precisely configured studio environment – a setup you typically cannot configure in your home theater even if you try to position your speakers as accurately as possible. So creating such amazing phenomena seems to be at least partly random; however, this is also true for generic surround mixes. Therefore, in my opinion, a virtual surround mix that is to function everywhere should always incorporate some kind of a "safety aspect." I also believe that sometimes it might be desirable to remaster a surround program created with the Isostem system, in particular, when it is a "production" rather than a continuous broadcast format. Using EQ and dynamics, one could certainly make even more of it. The parameters of the system respond with mathematical precision – one actually gets what one expects. It goes without saying that there is a major focus on removing direct sound portions located in the phantom center from the rear surround channels (unless you do not care about localization fuzziness). Of course, a generic mix offers entirely different signal-distribution and signal-design options; but nevertheless, Isostem is a fantastic system for converting a stereo program to an absolutely convincing surround program. It's virtually magic – I was extremely surprised what I could achieve using this emulation package.

Summary

When it comes to converting stereo material of any kind to surround, Isostem offers fantastic possibilities. As I know the daily tasks in broadcast or the features of a uniform multichannel broadcast format only in theory, I can only guess how valuable such a system might be in that context. However, one thing is clear: The investment into an Isostem system will pay off in virtually no time because of considerable cost savings on the production side. So if future viewers or listeners will enjoy the enveloping surround sound of live reports in newscasts, or if sports broadcasts (though conventionally produced in stereo) give them the impression of being right in the middle of it, they will not feel any difference between the original stereo or surround signals. That system presents genuine added value! When it comes to production (where I am better qualified), the parameter set of the Isostem system forms a reliable basis as well as an experimental playground for creating convincing surround productions even if the signal source comes from the stereo-recording back catalogue of the music industry. Without a doubt, the emulation of a surround field produced with the Isostem is definitely an upgrade to the source material. From already exquisite stereo sources, I created a number of amazing surround tracks, all of which I liked much more than the respective original recordings. The way there, however, is child's play only to those who are using the Live version of the system and just need to press a button for recalling a preset. An interesting finding that impressed me a lot was that a set of parameter settings configured for a specific program was suitable for any material from the same genre. So if you create presets such as "sports broadcast" or "music played on acoustic instruments", you will be able to rely on them in the future. According to DSpecialists, the Expert version providing full system access will be 6900 while the "preset-playing" Live version, which is suitable for automated broadcast operation, will be 4400 (all plus VAT/sales tax). While the price level will certainly deter the majority of studio owners at first, the Isostem system can definitely be considered quite affordable for broadcasters due to its realistic cost/benefit ratio. And when the books are well filled, even recording studios might consider purchasing such a system. As far as I can see, there is a plethora of stereo programs waiting to be converted to surround that cannot be mixed again because the multitrack source material is not available. Therefore, the good news is that DSpecialists is currently preparing Isostem software plug-ins for various platforms – and certainly at a much lower cost. Isostem is a professional tool of technical excellence and has the potential of making 5.1 surround a considerable success at last in both broadcast and production. In the next years, surround sound in combination with HD/3D flat screens will become a matter of course (because it has to). In this context, however, high-quality stereo productions certainly will (and shall) not become obsolete. But I'm convinced that thanks to Isostem this surely won't happen …

DSPECIALISTS
Digitale Audio- und Messsysteme GmbH
Helmholtzstraße 2-9 L · 10587 Berlin
Tel. +49-30-467 805-0 · Fax +49-30-467 805-99
info@dspecialists.de · www.isostem.de