Audio Engineering Society
Convention Paper 6444

Presented at the 118th Convention
2005 May 28–31 Barcelona, Spain

Comparison between different surround reproduction systems: ITU 5.1 vs PanAmbio 4.1

Enrico Armelloni [1], Paolo Martignon [2], Angelo Farina [3]

[1] Department of Industrial Engineering, University of Parma, Via delle Scienze 181/A, Parma, 43100, Italy
enrico.armelloni@unipr.it

[2] Department of Industrial Engineering, University of Parma, Via delle Scienze 181/A, Parma, 43100, Italy
paolo.martignon@inwind.it

[3] Department of Industrial Engineering, University of Parma, Via delle Scienze 181/A, Parma, 43100, Italy
farina@unipr.it

ABSTRACT

Audio reproduction of a movie inside a not dedicated room is critical; setting an ITU 5.1 system at home, for example, requires to place a large number of speakers around the room. But positions of speakers are often conditioned by the furniture. Bad alignment reduces spatial performances of the system dramatically.

For circumventing the above problems, most stereo TV sets, nowadays, are equipped with some form of "virtual surround" reproduction, employing substantially the Stereo Dipole method. This provides a very good frontal sound stage, but indeed sucks regarding the emulation of virtual surround loudspeakers.

An alternative reproduction technique is PanAmbio 4.1, based on a double stereo dipole system (frontal and rear). In this work the authors propose a comparison between the standard 5-ways surround system, and the new one. Validation is performed by subjective tests inside a domestic room.

1. INTRODUCTION

Difficulty to set properly an ITU 5.1 system inside a furnished room and to reach a good spatial reproduction of the movie's audio stream is well known. Correct angles and distances between speakers are unlikely to be respected: the furniture, for example, often forces speaker position. Result of this improper placement of loudspeakers is a bad spatial sound field, with inaccurate sound image that causes incorrect perception of sound effects.

An alternative technique to reproduce multi-channel signals could be PanAmbio 4.1, proposed by Ralph Glasgal (Ambiophonics's father). This system is based on a double stereo dipole configuration, one frontal (as found in many consumer stereo TV sets) and the other behind the listener.

In the PanAmbio 4.1, the speakers are positioned according with the rules of the Stereo Dipole, where two closely loudspeakers (around +/- 10° degrees) are placed in front and behind of listeners. It means that the frontal dipole can be easily set on the left and right sides of video, while the rear dipole can be assembled as a

single box, placed on a single stand and located behind listener position.
In this work authors performed listening comparative tests between two systems, ITU 5.1 and PanAmbio 4.1, in order to evaluate their global performances, their differences and usage flexibility.
Systems were set in a listening room and were used to reproduce the audio stream from a DVD-Video player. Digital processing was performed real-time using a personal computer supplied with AudioMulch software.
Subjective results showed how PanAmbio configuration allows reaching a very large sound image compared to other configuration, sound source localization is possible over all 360 degrees horizontal plane. A better frontal/rear harmonization of the sound was also achieved.
At present a stand-alone version of PanAmbio 4.1 system is available and it is working properly in the listening room of Ambiophonics Institute (Rockleigh NJ). This system is implemented on a commercial ADSP 21161 EZ-Kit Lite platform, an evaluation board supplied from Analog Devices.

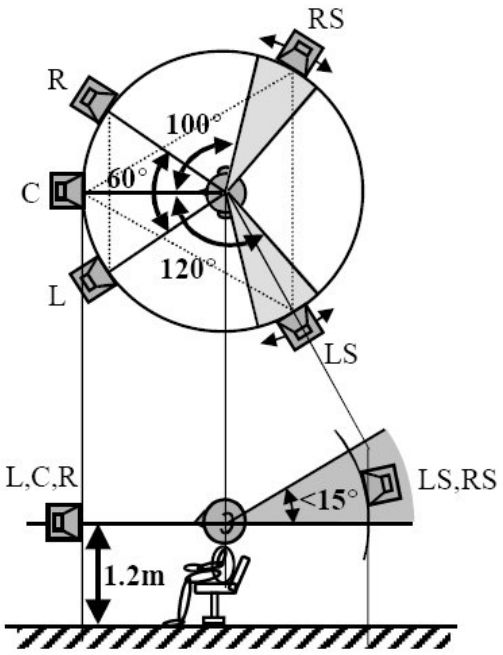

Figure 1: Speakers setup recommended by ITU-R BS. 775-1

2. ITU 5.1 SURROUND SYSTEM

2.1. Loudspeakers setup

Every multichannel "surround" system was born to create to the listener a 360° sound image. In a ITU 5.1 system five loudspeakers plus a subwoofer are employed; the best setup to listener music is depicted in Figure 1.
Optimum speakers placement is related to different aspects like dimensions and acoustic treatment of the room, distance between source and listener position etc. Furthermore character of the media (music, video) influences the reproduction system setup [1].
Angle between L and R speakers and the position of rear ones are the most critical variables as explained in ITU-R recommendations.

- **Frontal Channels L and R**, related to Figure 1, the angle of separation for L and R loudspeakers is typically 45° or 60° depending the media is played. Typically 45° is chosen for video reproduction, while 60° is used in music one.

- **Rear Channels LS and RS**, usually, depending of environments, two different placements are used: "direct surround" and "diffused surround". In the first method, Figure 2, one pair of surround speakers is aimed directly at the listening point. It is characterized by a "hot spot" listening area and used in music reproduction (ITU-R is an example). Meanwhile in the second method, "diffused surround", the absence of pin-point sound source allows to enlarge listening area. This setup is often used in theatres and cinema, Figure 3.

Direct sound method (Figure 2) is characterized by a trade-off between "surround panning" and "sense of rear stereo". The angle of separation for rear loudspeakers is critical:

- **ITU-R: 100° ÷ 200° from frontal axis -** In this configuration rear speakers are located "at the side" rather than "at the rear". It means the static phantom image is easily created by a good left/right separation. By surround panning is possible to move sound source from left to right but the image behind listener's head is blurred: the sense of depth is absent. The rear sound source movement is not spatially continuous due to the large separation of rear speakers.

- **135° from frontal axis -** This is a typical setup for speakers in domestic environment. When the angle increases, listener perceives sound more and more behind him.

- **150° from frontal axis -** This is a symmetrical configuration, frontal and rear speakers present same angle between them. By surround panning it is easier to move sound source over all 360°.

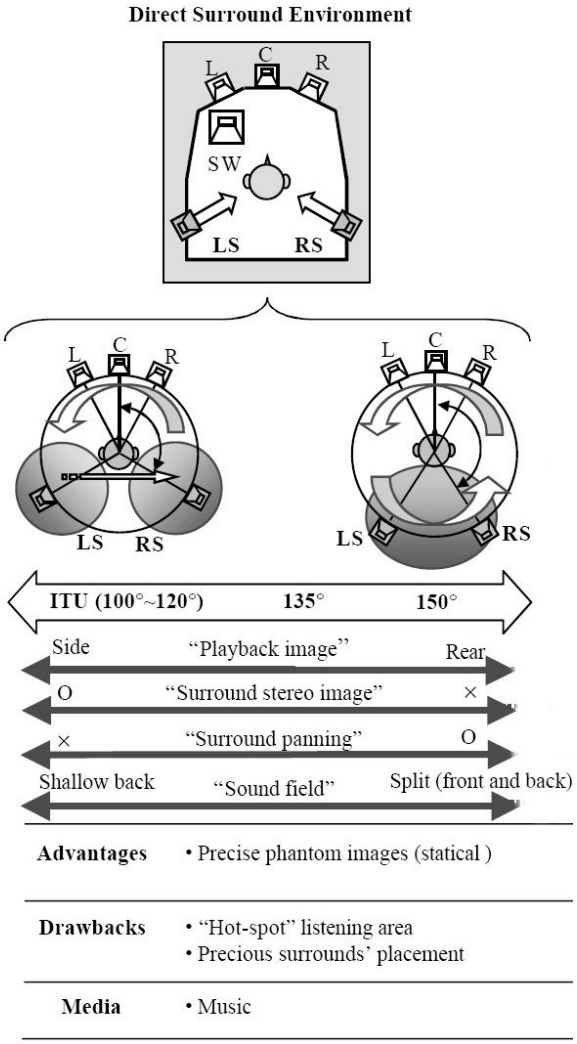

Figure 2: Direct surround environment

In diffuse surround configuration (Figure 3) rear speakers are located in both positions, side and behind. Now 360° surround panning and complete sense of rear stereo are possible.

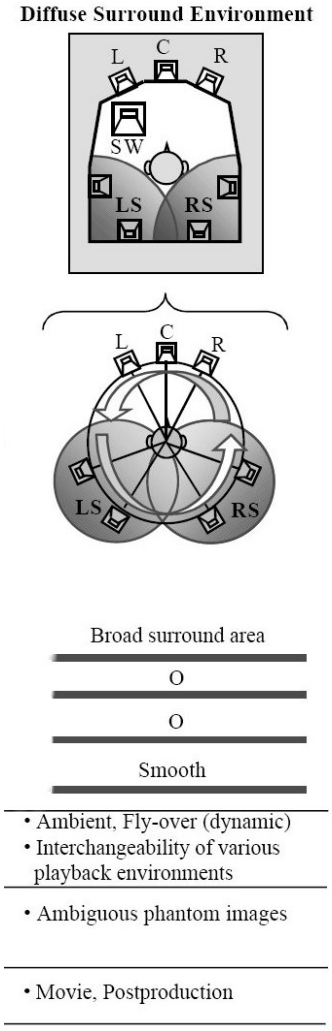

Figure 3: Diffuse surround environment

2.2. Configuration features

As described above, employing direct surround method is possible to create static phantom images in a 360° range. This configuration is often used for music reproduction. However, in order to precisely reproduce the phantom sound field, a dedicated setup of speakers is necessary for each kind of reproduced media.
Diffused surround, allowing surround panning to better create 360° audio source movements, is often used to reproduce the movie's audio stream. Using this method, compared to diffuse surround one, speakers' placement is less critical. It is easy to understand how this method can be considered efficient to play "general purpose" material.

3. CROSSTALK CANCELLATION AND STEREODIPOLE

Frontal speakers of a standard ITU 5.1 scheme are, substantially, a typical stereophonic scheme (+/- 30°) with a central speaker to reinforce the sound image. The worse defect of a system like this is the cross-talk effect at the listener ears. Cross-talk, which usually is not present in the music recording and is an artifact of a stereophonic reproduction, is a bad reproduction of sound at a location where it is not intended to be heard. For example the sound emitted from the left loudspeaker and heard at the right ear is a cross-talk. Proper cross-talk cancellation is necessary to widen the stereo sound stage; it is achieved using a stereo dipole system.

3.1. Review of stereo dipole theory

The approach employed here is derived from the formulation originally developed by Kirkeby and Nelson [2], with refinement from one of the authors [3]. Figure 4 shows the cross-talk phenomenon in the reproduction space.

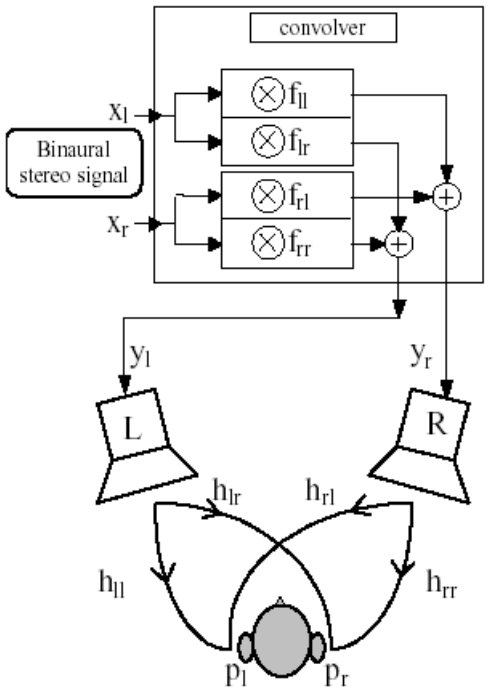

Figure 4: Cross-talk cancelling scheme

The 4 cross-talk cancelling filters *f*, which are convolved with the original binaural material, have to be designed so that the signals collected at the ears of the listener are identical to the original signals. Imposing that $p_l = x_l$ and $p_r = x_r$, a 4x4 linear equation system is obtained. Its solution yields:

$$\begin{cases} f_{ll} = (h_{rr}) \otimes InvDen \\ f_{lr} = (-h_{lr}) \otimes InvDen \\ f_{rl} = (-h_{rl}) \otimes InvDen \\ f_{rr} = (h_{ll}) \otimes InvDen \\ InvDen = InvFilter(h_{ll} \otimes h_{rr} - h_{lr} \otimes h_{rl}) \end{cases}$$

The problem is the computation of the InvFilter (denominator), as its argument is generally a mixed-phase function. In the past, the authors attempted [4] to perform such an inversion employing the approximate methods suggested by Neely & Allen [5] and Mourjopoulos [6], but now the Kirkeby-Nelson frequency-domain regularization method is preferentially employed, due to its speed and robustness. A further improvement over the original method consists in the adoption of a frequency dependent regularization parameter. In practice, the denominator is directly computed in the frequency domain, where the convolutions are simply multiplications, with the following formula:

$$C(\omega) = FFT(h_{ll}) \cdot FFT(h_{rr}) - FFT(h_{lr}) \cdot FFT(h_{lr})$$

Then, the complex inverse of it is taken, adding a small, frequency-dependent regularization parameter:

$$InvDen(\omega) = \frac{Conj[C(\omega)]}{Conj[C(\omega)] \cdot C(\omega) + \varepsilon(\omega)}$$

In practice, $\varepsilon(\omega)$ is chosen with a constant, small value in the useful frequency range of the loudspeakers employed for reproduction (100 – 20k Hz in this case), and a much larger value outside the useful range. A smooth, logarithmic transition between the two values is interpolated over a transition band of 1/3 octave.

4. PANAMBIO 4.1 SURROUND SYSTEM

4.1. System setup

PanAmbio 4.1 is a surround system where two Stereo Dipoles, are positioned around the listener, one in the front and the other in the rear [7], like showed in Figure 5. This system is suitable to reproduce movie or music audio streams.

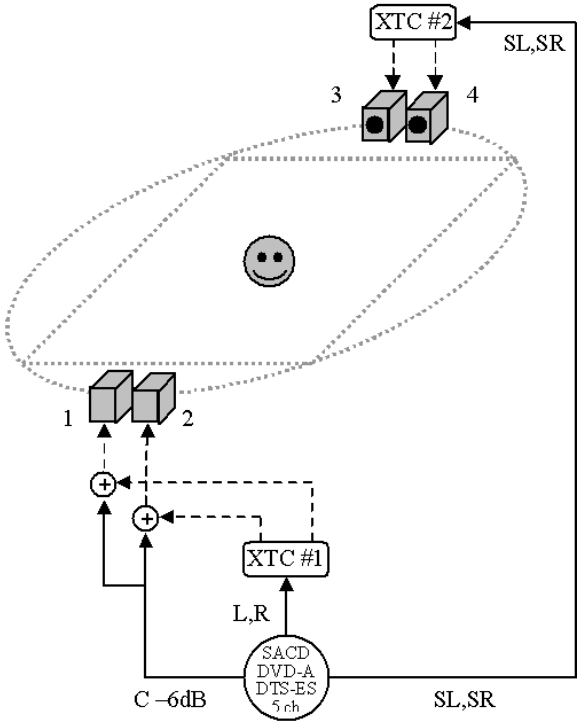

Figure 5: PanAmbio 4.1 configuration

In according with the rules of Stereo Dipole the angle between speakers is around +/- 10° degrees. It means that, in a hypothetical commercial setup, frontal couple of speakers could be easily set on the left and right sides of video. Rear dipole, instead, could be assembled as a single box, placed on a single stand and located behind listener position.

Audio signals Left, Right, Surround Left and Surround Right feed directly frontal and rear cross-talker. Therefore processed streams are sent to Stereo Dipoles.

The Central signal (C) is reduced by 5÷6 dB and added to two channels of the frontal Stereo Dipole (without passing through the cross-talk cancellation network). The choice of 5÷6 dB is due to the fact that the two signal deriving from C channel splitting over the two Stereo Dipole speakers are highly correlated at listener position (on axis). The LFE signal, instead, drives directly the subwoofer.

4.2. Implementation on a DSP board

As already explained in the introduction, nowadays a stand-alone version of PanAmbio 4.1 exists which is implemented on a commercial ADSP 21161 EZ-Kit Lite platform, an evaluation board supplied from Analog Devices. Cross-talk cancellation was obtained using filters based on a FIR structure: its coefficients was supplied by Wareing and worked into frequency range 200 ÷ 8000 Hz.

The DSP board is inserted between the multichannel outputs of the DVD player and the loudspeakers, as next Figure 6 shows. Signals are managed like explained in previous paragraph. In this configuration C signal can be added using an external mixer, since DSP board can handle only four analog input.

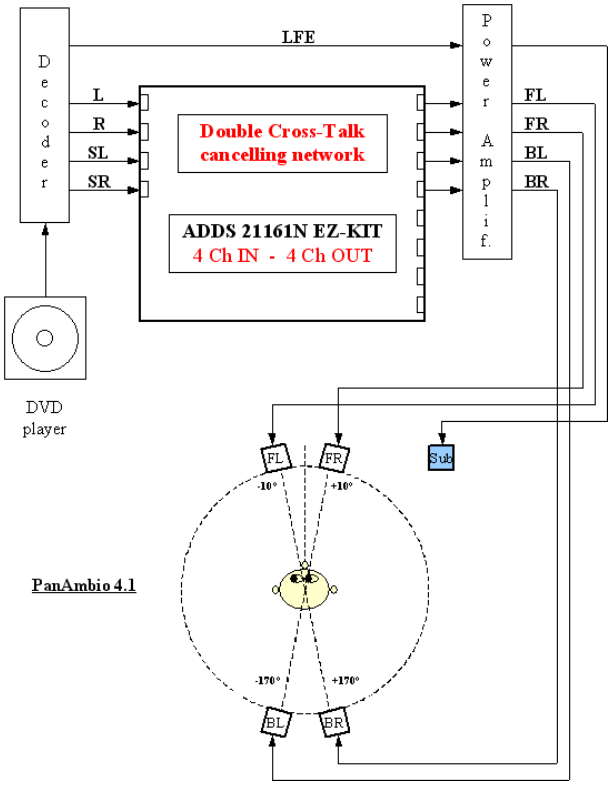

Figure 6: PanAmbio 4.1 on DSP board

PanAmbio 4.1 was tested first time, during AES 24th International Conference on Multichannel Audio (Banff, June 2003) by Ambiophonics Team (E. Armelloni, A. Farina, R. Glasgal, R. Miller, A. Torger), Figure 7;

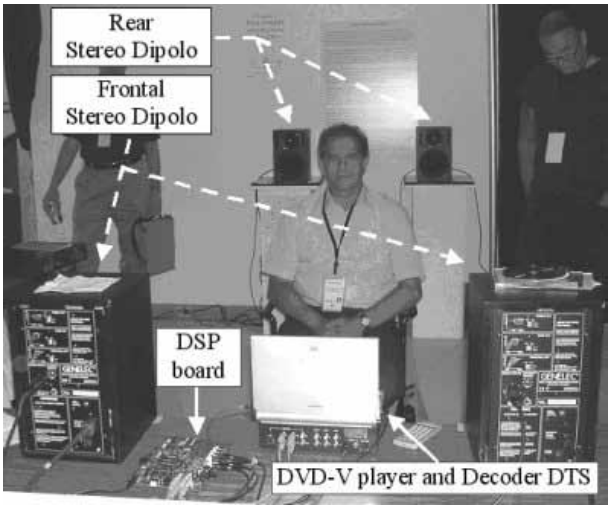

Figure 7: Demo-room at AES 24th International Conference on Multichannel Audio

Nowadays it is working properly in the two listening rooms of Ambiophonics Institute, Figure 8 and Figure 9.

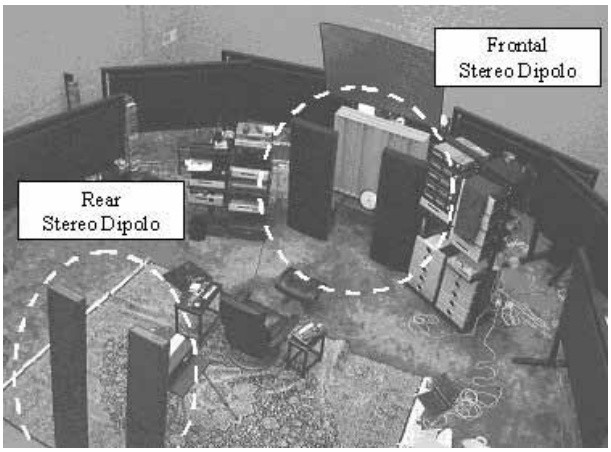

Figure 8: "Listening room 1" in Ambiophonic Institute

5. EXPERIMENTAL LISTENING TEST

5.1. Hardware and software settings

Comparison tests were performed inside a listening room of "House of Music" in Parma (Italy) where two different systems were set: ITU 5.1 and PanAmbio 4.1.
Digital output of DVD player was fed to the decoder – 5.1 amplifier connected to the ITU 5.1 system: five TurboSound Impact 50 passive speakers and a subwoofer. Environment implemented was "Direct Surround" with an angle between rear speakers and front central axis of 115°.

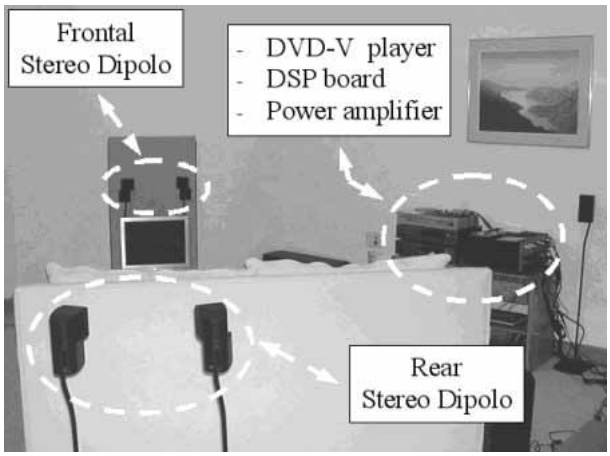

Figure 9: "Listening room 2" in Ambiophonic Institute

Analog output of DVD player, instead, was processed by a Personal Computer equipped with AudioMulch® and Woxengo Convolver plugin.
Elaborated signals drove directly the frontal Stereo Dipole, made of two self-powered Genelec S30D speakers. Signals to rear Stereo Dipole were reproduced through a further couple of TurboSound Impact 50 speakers powered by an external final amplifier.
During the test the listener could change system by a switch. Figure 10 and Figure 11 show scheme of hardware network and frontal view of system setup.

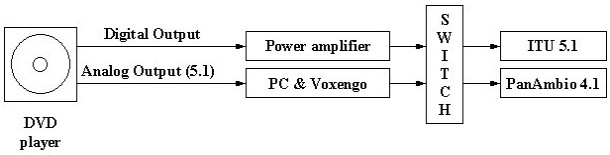

Figure 10: Scheme of hardware network.

As above mentioned, cross-talk cancellation is obtained filtering the two couple of signals (L,R and SL,SR) in real time using Woxengo Convolver, a VST plugin hosted by AudioMulch. It handles all the rooting including the center signal, as explained in paragraph 4.1. Convolver filters coefficients are the ones suggested by Wareing, which are obtained without measurements basing on a free field model. Figure 12 depicts the implemented network. Inside AudioMulch, a low-cut filter is applied to the front dipole to make the Genelecs roughly sound like the TurboSounds.

Moreover gain controls allow to set, for the two systems, same pressure level at listener position.

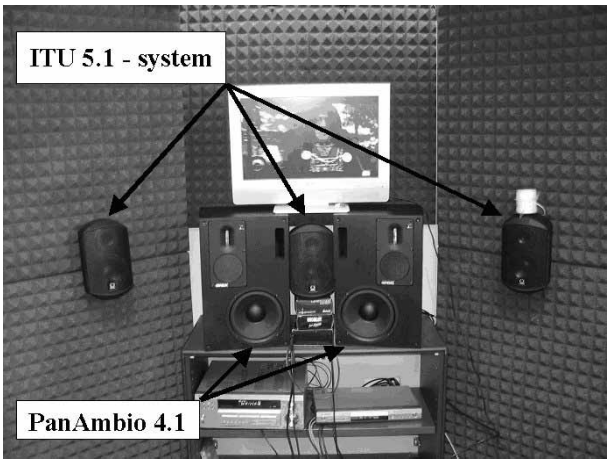

Figure 11: Frontal part of electroacoustic system.

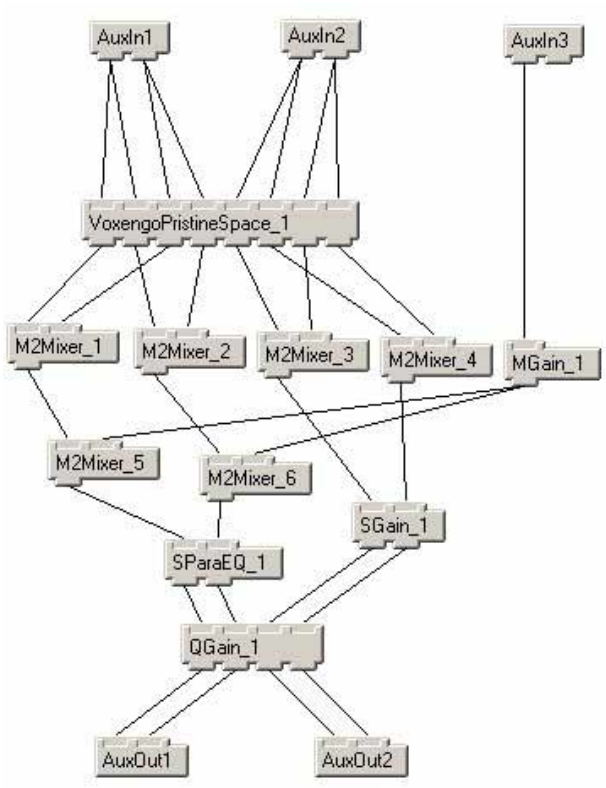

Figure 12: AudioMulch network.

5.2. Subjective results

The authors performed several subjective listening tests in order to compare the two different reproduction systems implemented. Different DTS drama and music audio/video samples were played, switching between the two systems. The impressions collected after the tests are reported below.

With ITU 5.1 reproduction system the general impression is that sound comes exactly from the 5 speakers. This is probably emphasized by the low speaker to receiver distance (1.5 m); however it is feasible that in a home cinema distances can't be much larger. The stereo dipole, on the contrary, provides for the following sound image features:

- Sound image is larger than in 5.1 case. With the former the widest angle for a virtual frontal source is 30°, coinciding with its sound coming only from one of the two front lateral speakers: On the contrary with crosstalk cancellation this case corresponds to a source placed roughly at 90°, since the opposite ear is completely shielded. A similar discussion can be made for the surround image; moreover the problem of low spatial resolution for surround sound is improved using the rear stereo dipole.

- Contributes of stereo dipole systems (deriving from 5.1 left and right front and surround channels) dramatically increase "sound depth": virtual sources distances are, so to say, shifted to infinity. Instead, the frontal contribution to sound image deriving from the original center signal, which doesn't pass into cross-cancellation, is much more present and appears to come exactly from the video. This greater differentiation of virtual sources depth and less localization of real sources is very pleasant and gives a quite superior realism sense compared to ITU 5.1.

- The spectral sound quality of stereo dipole seems to be not worsened by cross-cancellation, and no particular interference spatial artifacts due to composition of rear and front dipoles arises: that is, moving back and fro with the head, sound keeps stationary.

- Unfortunately the qualities described so far are limited to a sweet spot suitable for a single listener. Authors tried putting two chairs one beside the

other across the longitudinal central axis, hosting two listeners whose head centers resulted to be 90 cm distant. For reasons explainable looking deeper in stereo dipole theory, the left listeners could still correctly hear just the right side of the total image, and vice versa; it can be said that the sound image is half-preserved. For more off axis positions the sound image is completely lost, and it can be fairly said that this situation can't get better even changing the speaker listener distance and angle.

- Stereo dipole is in general suitable for binaural measurement, of course, but also for the coincident capsules recording method, stereo or surround. Also with artificial panning surround production the result can be fine. Problems arise when artificial delay are added or distant microphones techniques are employed; This case was experimented playing some DTS recording: in which the tracks, that sounded natural in ITU 5.1, appeared as strange artifacts when passed in a Stereo Dipole network.

6. CONCLUSION

In reason of above discussion, PanAmbio features appear to very charming in order to obtain more suggestive sound experiences with a simpler speaker setup. For this purpose PanAmbio may appear like an interesting improvement of traditional TV. In the future companies may think to integrate this system in their devices by setting speakers on the sides of video and adding output to a rear Stereo Dipole. This would mean a setup slightly more complex than today's "virtual surround" TV sets, but still quite simpler than a full discrete 5-channel systems (a single loudspeaker on his own stand located behind the listener, instead of 4 displaced around).
Since 5.1 tracks can have been made with different several methods, not all the 5.1 products, now available on the market, can yield good result with PanAmbio 4.1.

7. REFERENCES

[1] Masataka Nakahara "Multichannel Monitoring Tutorial Booklet" – Rev 2.3.0, June 2002.

[2] O. Kirkeby, P. A. Nelson, H. Hamada, "The "Stereo Dipole"-A Virtual Source Imaging System Using Two Closely Spaced Loudspeakers" – *JAES* vol. 46, n. 5, 1998 May, pp. 387-395.

[3] O.Kirkeby, P.A. Nelson, P. Rubak, A. Farina, "Design of Cross-talk Cancellation Networks by using Fast Deconvolution" - *106th AES Convention*, Munich, 8-11 May 1999.

[4] A. Farina, F. Righini, "Software implementation of an MLS analyzer, with tools for convolution, auralization and inverse filtering" - *Pre-prints of the 103rd AES Convention*, New York, 26-29 September 1997.

[5] S.T. Neely, J.B. Allen, "Invertibility of a room impulse response" - *J.A.S.A.,* vol.66, pp.165-169 (1979).

[6] J.N. Mourjopoulos, "Digital Equalization of Room Acoustics" - *JAES* vol. 42, n. 11, 1994 November, pp. 884-900.

[7] R. Glasgal, "Surround Ambiophonic Recording and Reproduction" - *24th AES International Conference on Multichannel Audio*, Banff, 26-28 June 2003.