Aurora listens to the traces of pyramid power

Ramsete is a new pyramid-tracing computer program, a variation on ray or beam tracing, which provides 3-D simulation of indoor and outdoor building acoustics. For the first time, pyramids are in competition with cones as the basic geometry for ray-tracing noise models. Aurora is a new auralization program which is being used to test and improve Ramsete. Both were devised by Dr Angelo Farina of the University of Parma in Italy, and the judging panel are students at the university. Here he describes how the two programs work.

Ramsete simulates sound propagation in large rooms and outdoors taking into account source directivity and extension, specular and diffuse reflections, angle and frequency dependent absorption, excess attenuation and screen shielding. The results can be used for auralization via Aurora.

Ramsete contains a source manager for source directivities and sound power data files (figure 1). It can automatically define directivity balloons from experimental measurements conducted according to ISO codes 3744 and 3746, with direct reading of the most common RTA file formats. The program also reads loudspeaker directivity data stored in the Bose Modeler format.

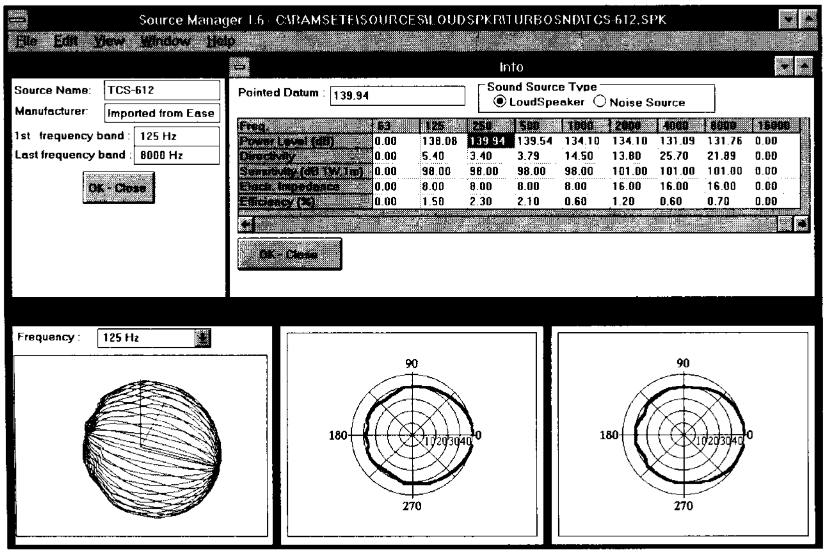

Fig. 1: Ramsete's Source Manager tackles directivity (lower left window) and plot angles (lower centre and right windows).

The package includes a large material-properties database, including absorption coefficient and sound reduction index data. The program has its own Windows-based 3-D CAD. It is compatible with Autocad and reads and writes DXF files. The post-processing tools include calculation of standard acoustic quantities (sound level, reverberation time, centre time, clarity, STI, LE, LF), with graphical and tabular representation, and a 3-D subprogram that can be used for generating contour maps.

Sound level maps in complex geometries can be generated very rapidly. Usually the work is completed in a couple of minutes for each sound source (using a 66 MHz 486 PC), and only in rare cases is computation time as long as 20 minutes.

Ramsete uses a completely new system of pyramid tracing. Pyramid tracing triangles perfectly cover the source sphere and so avoid the problems presented by conical beam tracing of multiple detection of the same path and underestimation of the reverberant queue. Beam tracing cones must be overlapped to fully cover the source sphere surface. An algorithm is required to avoid multiple detections or to weight the energy so that, on average, the multiple contributions produce the correct sound level.

In Ramsete, the surface sphere is subdivided on the basis of octants. The number of pyramids generated can be any power of 2, and all of them have almost the same base area, giving a nearly isotropic sound source.

For the Sabinian sound field, Ramsete contains a multiplicative correction factor so that there is no sudden jump in the detection density. The correction is computed separately for each receiver based on the statistics (mean free path) of that particular receiver.

Alternatively, Ramsete can work as a purely deterministic model. If the number of pyramids is large enough, it is possible to obtain "exact" values for the more sensitive acoustic parameters. In practice, only a small number of pyramids is needed to obtain reasonably accurate estimates of the sound level and reverberation time, providing that the time resolution selected for the impulse response is not too small.

For non-Sabinian sound fields, the use of a very large number of pyramids is ideal to ensure that no queue correction is required. The correction applied to Sabinian sound fields does not work for non-Sabinian fields because the absorption distribution is non-uniform and reflection paths assume values dispersed from the mean free path on account of the larger number of obstacles. The correction can be adjusted but this must be done according to the individual geometry of the site.

Tests and improvements

The numerically calculated impulse responses can be compared with experimental impulse response measurements taken in the spaces being modelled. Another way to compare the numerical and experimental results is by listening to musical pieces, spoken text, or other signals converted to pressure impulse responses by the Aurora program. Aurora is equipped with fast convolution software, which runs on a PC and does not require a DSP board. It can process standard WAV files of unlimited length with binaural impulse responses up to 200 000 taps. These can be either experimental or numerical. A standard 16-bit PC audio board and high quality headphones are then used to submit the convoluted signal to a judging panel.

Some of the panel were able to distinguish between signals convoluted with experimental and numerical impulse responses, but they gave broadly similar subjective judgements regarding reverberation, bass and treble balance and spatial impression. Thus the simulated responses are not perfectly identical to the experimental ones, but they reproduce correctly the most important acoustic effects.

The research is now continuing to obtain reasonably good reproductions through loudspeakers. In addition, the Ramsete pyramid tracing algorithm is being updated, taking into account diffusing surfaces and extended sources. The convolver is being improved towards the goal of real-time computation.

Conversion procedure

Ramsete produces energy impulse responses of unlimited length, sampled at intervals typically of 10, 5 or 1 ms. A separate response is computed for each octave band (31.5 Hz to 16 kHz). To obtain reasonably good binaural simulations, two receivers must be located at the sides of a sound diffracting sphere, approximating two ears. In this way, the response at one receiving point is a couple of arrays, each of them with 10 columns (the frequency bands) and some hundreds or thousands of rows (the time intervals).

It is necessary to convert each of these two arrays to a single pressure response, with a bandwidth of 20 Hz to 20 kHz, sampled at 44.1 kHz, or preferably 48 kHz. Usually these impulse responses are stored in standard WAV format, or in proprietary impulse response formats such as the TIM format of the MLSSA system. Powerful software tools enable any kind of mathematical manipulation of these pressure response files.

The conversion process begins with the generation of a sample of white noise with proper length and sampling rate, covering the impulse response duration. Ten copies of this white noise signal are made, and each of them is modulated with an amplitude envelope obtained by the square root of the energy impulse response in the corresponding frequency band. Each modulated white signal is passed through a 6-pole IIR digital octave filter, centred on the corresponding frequency. The ten filtered signals are then summed together, obtaining a wide-band pressure response. Eventually, the result is normalized and converted to a 16-bit integer.

Although this process can be criticized in many ways, it is very fast and produces impulse responses that exhibit the same energy/time curves, the same reverberation times, and almost exactly the same objective acoustical parameters as the original energy responses. Obviously this is not enough to be accepted as realistic for audible simulations. A subjective evaluation needs to be made for this.

There is a problem regarding the overall frequency response. Often Ramsete simulations make use of a theoretical omnidirectional source (OMNI), having a sound power level of 110 dB in each octave band. It is also possible to use sound source files emulating many loudspeakers with proper directivity balloons. In this case the computation is conducted as if the sound source was fed with pink noise, because this simplifies the presentation of the resulting spectra in terms of overall frequency response. However, a pressure impulse response unfiltered in any frequency is inherently white, not pink. The pink-to-white conversion is automatically accomplished by the procedure described above.

Some Ramsete simulations are conducted with sound source data that are already white. For example, if the sound source file was built by MLSSA measurements on a loudspeaker, the above procedure introduces a 3 dB/octave increase of the response with frequency. This can be removed by setting an appropriate flag during the conversion process.

Another case is a numerical simulation conducted with an OMNI source, that is to be compared with experimental measurements ▶

obtained with a sound source not perfectly flat in the whole frequency range. The conversion software allows for this, enabling an octave-band equalizer to adjust each octave band prior to the summing. If the equalizer is adjusted by an octave analysis of the experimental impulse response (probably limited to the direct wave), this already takes into account the "whiteness" of the spectrum. Therefore, it is necessary to set the "pink" flag again in order to avoid a 3 dB/octave positive slope in the computed response.

Fast convolution

Although specialized hardware tools which perform continuous delayed convolution have been available for some years, they are expensive and do not allow for easy digital transfer of the input and output signals on a PC. These devices use frequency domain processing with large blocks of data, resulting in a delay that is three to four times longer than the impulse response length. While time domain processing DSP boards conduct time domain convolution with no delay, they are limited to a few thousand filter taps in the better cases.

The Aurora system employs a different approach. Both the input and output data files are stored on the hard disk in standard WAV format and, once convolution is performed, comparative tests can easily be conducted with just a point-and-click delay.

Convolution is performed through the well known "select-save" algorithm. Now the program has been extended to longer impulse responses (up to 200 000 taps) and has been speeded up.

The computing times for the implementation of Aurora on a 486 DX2-66 PC, with a mono input signal of three different lengths and two (binaural) impulse responses of five increasing lengths, are given in table 1.

Usually there is no problem in employing samples over 30 s for comparison tests. The program takes 10–20 s to initialize, performing the FFTs on the impulse responses; then the select-save loop begins. The I/O time due to disk access is always less than 10% of the total computation time. The table shows that the effective CPU time spent in the select-save loop ranges from 5 to 15 times the input sample duration. Thus, a large speed

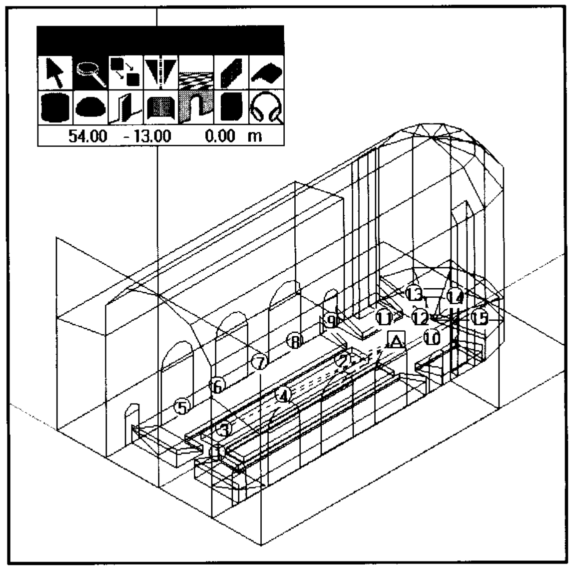

Fig. 2: Microphone placements in the Church of Santa Lucia in Foligno.

improvement of the system would be required for real-time delayed operation.

Some tests have shown a speed improvement of a factor of five using a Pentium-90. Real-time delayed processing can actually be obtained on a Pentium-90 for impulse responses of 16 kpoints at 44 kHz, or 32 kpoints at 22.05 kHz. A special multi-thread version of the software is under development now in order to test it on multi-processor Dual Pentium machines.

Subjective tests

The proper working of the convolution software was previously verified, comparing convoluted signals with direct recordings taken during loudspeaker playback of anechoic music in theatres. Ramsete was also tested and showed good correspondence between the computed objective energy impulse parameters and the experimental ones. Subjective tests were thus required to validate the Aurora's energy-to-pressure conversion algorithm, which is the most delicate point of a computerized audible simulation based on geometrical acoustic assumptions.

Two enclosed spaces were employed for this test. The first was a church in Foligno (figures 2 and 3) and the second a large sports arena in Modena. The same sites were used for the comparison of objective parameters. Two receivers were chosen in the church, one very near the source (with a high direct/reverberant ratio) and one in the far field. A single receiver was chosen in the sports arena, as the sound field was very similar everywhere.

Two anechoic samples were used for the test. The first was a piece of music by Mozart, the overture from "The Marriage of Figaro", bars 1–

Table 1: Aurora computing times (in s)

	Impulse response length (kpoints)				
	16	32	64	128	179.5
input 10 s, 44.1 kHz	61.85	79.0	108.2	160.4	209.6
input 20 s, 44.1 kHz	108.4	125.1	171.6	300.7	344.8
input 30 s, 44.1 kHz	156.3	183.4	244.6	359.7	488.8

18. The second was a poem by Leopardi, *Il Sabato del Villaggio*. The samples were digitally transferred from CD to hard disk through a Toshiba CD-ROM drive.

They were convoluted both with the numerically simulated impulse responses and with the experimental ones. Thus six pairs of samples were obtained, the presentation order of which was randomly shuffled. To obtain comparison data, two control groups each of other six pairs of samples were mixed with the "true comparisons" set: six "really equal" pairs (obtained playing the same sample twice) and six "really different" pairs (obtained with random associations of samples coming from different rooms).

Fourteen subjects were asked to listen to the 18 pairs and to answer a series of questions for each pair. They were asked if the two samples A and B were equal or not. If not, they were asked to grade the difference between each sound pair numerically as a lot (–2), slight (–1), no difference (0), slight (+1), or a lot (+2). These gradings were applied to various sound qualities: more reverberant; more pronounced bass; more pronounced treble; and wide spatial impression.

The results of the first question on equality, were 23.0% for the Ramsete/experimental comparison, 71.8% for the truly equal samples, and 2.5% for the truly different samples. These percentages show that Ramsete's simulations were fooling one subject out of four. Table 2 shows the results of the sound quality perceptions reported by the students.

Surprisingly the Ramsete/Aurora simulations give an almost correct impression of the four subjective qualities. For average differences, they are judged nearly equal to the experimental responses. They diverge, as expected, from the "truly different" control group, while the "truly equal" group shows average values which are close to the numerical/experimental comparison. This result is probably due to the fact that differences in "truly equal" responses are found by the most unreliable subjects, while differences between numerical and experimental samples are found by the most sharp-eared listeners. The standard deviation of the numerical/experimental comparison is greater than that of the "truly equal" pairs, while the standard deviation of the "truly different" group is greater still.

The subjective results show that the simulated responses are not perfectly indistinguishable from the experimental ones: a trained listener can easily identify the difference in a comparison test. On the other hand, these differences are actually very little, as the main subjective parameters are judged equal. Asked to specify the true nature of the difference, most subjects said they could distinguish a slight difference, but that they could not identify its cause.

Table 2: Average differences (a) and standard deviations (b)

	Ramsete/Exp. comparison		Truly equal samples		Truly different samples	
	(a)	(b)	(a)	(b)	(a)	(b)
more reverberant	–0.06	±0.92	–0.03	±0.58	–0.36	±1.38
pronounced bass	–0.06	±0.77	–0.10	±0.41	+0.24	±1.30
pronounced treble	0.00	±0.72	+0.04	±0.41	+0.35	±1.07
wide spatial impression	+0.24	±0.87	+0.06	±0.40	+0.64	±1.22

Probably a wider subjective experiment is needed, with a larger number of subjects and including some other rooms with different acoustic behaviour. A better questionnaire can now be arranged based on the improvements suggested by many of the subjects.

This test shows that the Aurora system has the capability of producing realistic binaural simulations, that make it possible to appreciate even slight modifications in the numerical model of a room. Usually the goal of an auralization system is to provide an audible feedback to the designer, or to demonstrate the effects of proposed acoustic treatments of the room. These capabilities are now available using a low-cost PC and no additional specialized hardware, with computation times short enough to perform dozens of consecutive tests, and with no practical limitation on the sampling frequency and the length of both the convoluted signals and the impulse responses. ■

Contact: Angelo Farina, Department of Industrial Engineering, University of Parma, Via delle Scienze, 43100 Parma, Italy. Tel: +39 521 905854. Fax: +39 521 905705. E-mail: farina @aida.eng.unipr.it.

Fig. 3: Noise contour map of Santa Lucia.