# CHORAL MUSIC RECORDING WITH ACOUSTIC QUADRAPHONY

**G. Cengarle, D. Bonsi, D. Gonzalez, D. Stanzial**
*The Musical and Architectural Acoustics Laboratory FSSG-CNR, Fondazione G. Cini, Isola di San Giorgio Maggiore, I-30124 Venezia, Italy,*
*cengarleg@libero.it*
*davide.bonsi@cini.ve.cnr.it*
*diego.gonzalez@cini.ve.cnr.i*
*domenico.stanzial@cini.ve.cnr.it*

## Abstract

Sound recording from spatially extended sources is an interesting acoustical problem that only in recent years, thanks to the advent of multichannel audio, has started to be successfully dealt with. In this respect the choral music is indeed a challenging audio test-bed both in the recording and playback phase of sound reproduction technology. A new application of the intensimetry to the audio field, named "Acoustic Quadraphony" (AQ), recently developed at FSSG-CNR Lab in Venice within the frame of the IP-Racine European project is certainly a promising technique towards the real acoustical solution of this problem. The present paper illustrates how the AQ methodology has been applied for recording a choir source of about thirty singers in the classical 4 voices, Basso-Tenore-Contralto-Soprano, composition. Once recorded, the pseudo-anechoic signals corresponding to each voice go through a multiple quad-convolution process and then the final spatially and temporally superimposed quad-signals are reproduced through a common 5.1 surround system.

## INTRODUCTION

"Acoustic Quadraphony" is a recently developed application of sound intensimetry to audio technology [Stanzial et al., 2003]. This application is based on the fact that any sound field can be identified by means of a set of 4 scalar fields: the sound pressure and the three components of the air particle velocity [Stanzial et al., 2002]; these signals contain all the necessary and sufficient information for the complete description of the sound field, including its energetic and spatial properties. Moreover, it has been shown that, in the case of a point-like source, the pressure and velocity fields measured at an arbitrary point in space can be expressed as the time convolution between a single sound pressure signal generated by the source in an anechoic environment and the set of pressure and velocity responses (or simply "Quadraphonic impulse response") obtained from an impulse-like excitation generated at the source position. A custom-made software has been developed for measuring the room quadraphonic impulse response, implementing also all the functions for the full energetic analysis of the acoustic environment in the specific source-receiver configuration. A quad-recording process can thus be executed in principle either with a direct recording of 4 tracks (i.e. recording the pressure and 3 component of air particle velocity signals) or it can be post-processed by means of a quad-convolution between the directly recorded pseudo-anechoic pressure signal with a measured quadraphonic impulse response. However, since this method relies on a point-like approximation, some relevant spatial properties of the source itself, namely its extension and directivity, cannot be accounted for.

The simplest approach which can be adopted when dealing with a spatially extended source is to approximate the sound field as a linear superposition of contributions generated by a set of point-like terms. For a given source, the degree of accuracy provided by this method depends on the choice of the decomposition, that is the number of point sources and their positioning.

In order to test this technique with a musical sound source and a suitable acoustic environment, we chose to study the rendering of a four-section choir in a church. The singers, arranged to form a semicircle surrounding the conductor, located in the centre, are gathered into four groups according to the classic left to right disposition of Sopranos, Altos, Tenors and Basses. When all the singers in a section perform the same part, each of these four sections is suitable to be associated, in a first approximation, to a single point sound source.

The canon “Alleluia, Kanon im Einklang” (KV553), composed by Wolfgang Amadeus Mozart (ed. Schott Musik International, Mainz) was chosen for the performance, since its musical structure allows the listener to easily distinguish and follow the different parts.

Usually, choral music is performed in churches or large auditoria; in the case of a sacred repertoire, a church is obviously the most appropriate environment because it provides the right amount of reverberation and a sense of “depth” that fit the music and support the performance offering a natural sustain to the sound. The environment we have chosen for the impulse response measurement is the church of San Giorgio Maggiore in Venice which, thanks to its wooden choir stalls dating back to the 16$^{th}$ century, its large dimensions and long reverberation time (RT30 $\cong$ 3 s) is suitable for enhancing the sound of a choral performance.

## MEASUREMENT OF QUADRAPHONIC IMPULSE RESPONSES

In order to measure the impulse responses of pressure and the three components of the velocity vector, we used a custom setup implementing the *Swept Sine excitation* technique [Bonsi et al., 2005]. A LOOKLINE mod. 301 dodecahedron loudspeaker was used as an omnidirectional sound source for the playback of the excitation signal. The pressure and velocity responses to the excitation signal were simultaneously acquired by means of a Microflown® USP sound intensity probe. A laptop PC equipped with a MOTU® mod. *Traveler* firewire audio interface carried out the playback and recording of the signals. The generation of the excitation signal, the calculation of the impulse response and the data analysis were performed by a software implemented on *Matlab* platform.

The choice of the source-receiver configurations was based on the positions that the singers and the choirmaster would occupy in the case of a real musical performance inside the church. The measurements were performed in the apsidal choir of the church, where the singers would sing in a semicircle, with sopranos on the left, altos half-left, tenors half-right and basses on the right, as seen from the conductor standpoint. As shown in figure 1, four positions were chosen for the source, corresponding roughly to the middle point of each voice section. An additional central position was used for the comparison with the previous single point-source approach. The probe was positioned in the centre of the semi-circle, reflecting the position of the conductor, and 20 meters away from the choir for the simulation of the sound field perceived by the audience.

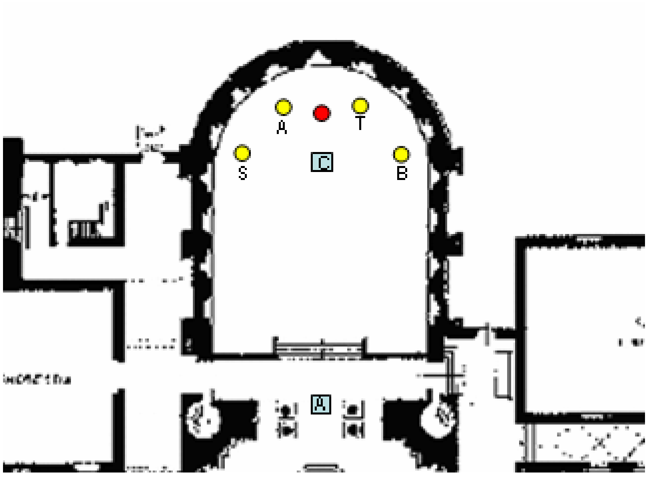


*Figure 1: Source-receiver configurations. The yellow spots S, A, T, B indicate the position of the source corresponding to the four sections of the choir; the red spot is the single source position representing the whole choir. The squares represent the positions of the receiver: C is the position of the conductor and A represents the audience.*

## ANECHOIC RECORDING OF THE CHOIR

The source contribution to the sound is represented by its pressure signal. In our approximation, the extended source was decomposed into four point-like sources representing the four sections of the choir. Obtaining the four pressure signals corresponding to these four voice sections implies recording the performance in a reflection-free environment, capturing the sound of the separate sections with closely placed pressure microphones. Ideally, the recording should be made inside an anechoic chamber, to ensure the absence of any room contribution to the recorded sound. Due to the unavailability of a suitable anechoic chamber, with enough volume to host an entire choir, the free-field condition was obtained recording the performance in a large outdoor space, i.e. a grass lawn in the country, as shown in figure 2.

The choir was arranged to mirror the placement that would have been chosen for performing in the church, with each section occupying the position where the source was placed during the measurements.

A pressure microphone was placed in front of each section, slightly above head height and approximately 1 m from the first row, to ensure a uniform pickup of all the singers in a section and to minimize the leakage from the other sections. The four pressure transducers are a matched pair of EARTHWORKS mod. *QTC40* microphones and two AKG mod. *C480 B* microphones with *CK-62* capsules. An additional omnidirectional microphone was placed above the conductor, to provide a single pressure signal of the whole choir to be compared to the multichannel approach. The signals were sent to microphone preamplifiers with matched gains feeding the inputs of a 24 bit digital multitrack recorder mod. ALESIS *HD24*, where the five signals were recorded on separate tracks running in sync.

*Figure 2: Left: omnidirectional loudspeaker and sound intensity probe in St. George church. Right: recording the choir in an open field.*

Since two different microphone pairs were used on the choir sections, a digital gain compensation was applied to the lower sensitivity pair at the post-processing stage to ensure the level matching of the four channels.

## MULTICHANNEL QUADRAPHONIC CONVOLUTION

For a single point-like source, the sound field in position $\vec{r}_M$ generated by the sound pressure source $S_i(t)$ located at $\vec{r}_i$ is represented by the four signals $\left\{p^i(t), v_x{}^i(t), v_y{}^i(t), v_z{}^i(t)\right\}$ obtained by the convolution between the source signal and the four impulse responses (pressure and velocity) corresponding to the same source-receiver configuration, as expressed by the set of equations (1):

$$\begin{cases} p^i(\vec{r}_i, \vec{r}_M, t) = \int_{-\infty}^{t} g_p(\vec{r}_i, \vec{r}_M, t-\tau) S_i(\tau) d\tau \\ v_x{}^i(\vec{r}_i, \vec{r}_M, t) = \int_{-\infty}^{t} g_{vx}(\vec{r}_i, \vec{r}_M, t-\tau) S_i(\tau) d\tau \\ v_y{}^i(\vec{r}_i, \vec{r}_M, t) = \int_{-\infty}^{t} g_{vy}(\vec{r}_i, \vec{r}_M, t-\tau) S_i(\tau) d\tau \\ v_z{}^i(\vec{r}_i, \vec{r}_M, t) = \int_{-\infty}^{t} g_{vz}(\vec{r}_i, \vec{r}_M, t-\tau) S_i(\tau) d\tau \end{cases}, \quad (1)$$

In the case of $N$ sound sources, the resulting total field is the sum of the contributions of the single sources:

$$\begin{cases} p(t) = \sum_{i=1}^{N} p^i(t) \\ v_x(t) = \sum_{i=1}^{N} v_x{}^i(t) \\ v_y(t) = \sum_{i=1}^{N} v_y{}^i(t) \\ v_z(t) = \sum_{i=1}^{N} v_z{}^i(t) \end{cases} . \tag{2}$$

Following this procedure, we must first separately determine the four sets of quadraphonic data generated by each section by means of a convolution between the anechoic signal recorded by the microphone in front of a given section and the corresponding quadraphonic impulse response. The pressure and velocity components of the four fields must then be summed to obtain the complete sound field.

After the storage of the audio files and the impulse responses as Broadcast Wave files, the convolution operation was performed by means of *SIR*, a convolution reverberator available as a VST plug-in, running on a PC equipped with Adobe Audition software. The sum was performed with Adobe Audition using the *Mixdown* feature.

## RESULTS

The increase of accuracy attainable in the reproduction of the spatial properties when using such a multi-source approach may be verified analysing the instantaneous sound intensity vector:

$$\vec{i}(t) = p(t)\vec{v}(t) . \tag{3}$$

This quantity expresses the amount of energy flowing per unit time and surface area at any time $t$ along the vector direction. Plotting the sound intensity vector as a function of time for the whole duration of a sound event gives us information about the direction of arrival and the relative magnitude of the direct sound and the subsequent reflections. An effective way of displaying the intensity vector is through a polar plot on a plane. Starting with the wave files obtained by the aforementioned convolution, the intensity vector can be calculated and plotted by means of a *Matlab* based routine. We here consider the projection of sound intensity on the $xy$ plane, which in our case is the horizontal plane containing the head of the listener. Plotting the intensity vector for the whole duration of a musical phrase performed either by the complete choir or by a single section, as in the case of the beginning and the ending of the canon, sung by sopranos and basses alone respectively, we can point out the different spatial properties of the field. Figure 3 shows the results obtained by the convolution with the impulse responses measured at the conductor's position. The plots show that a multi-source approach yields a sound field capable of preserving the position and the broadening of the sound source, while the single point-source approach causes the shrinking of the source in a central position, as expected. Moving further away from the choir, one expects the spatial extension of the source to become less perceivable, because two factors come into play: first of all, the source is seen under a narrower angle, sounding smaller and behaving as a point-like one; moreover, as one moves outside the critical distance, the reflections which form the diffuse sound field become predominant making the localization of the sound sources difficult. Plotting the intensity vector of the reconstructed sound field at the audience position, shown in figure 4, it is clear that it is not possible to discriminate the left and right location of the voice sections, although it remains evident that the source is located in the front direction.

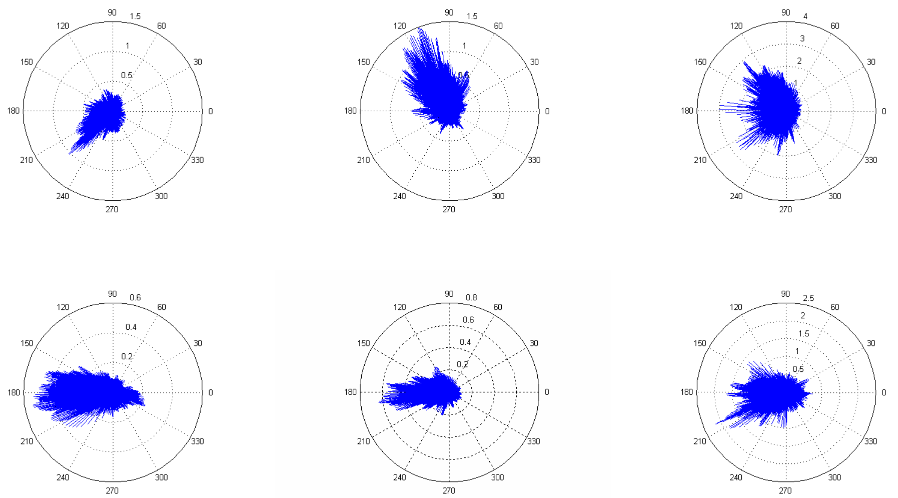

*Figure 3: Sound intensity vector polar plot in the conductor's position. Upper row: results obtained with multi-source approach; left to right: sopranos only, basses only, whole choir. Lower row: results obtained with single point-like source approach; left to right: sopranos only, basses only, whole choir.*

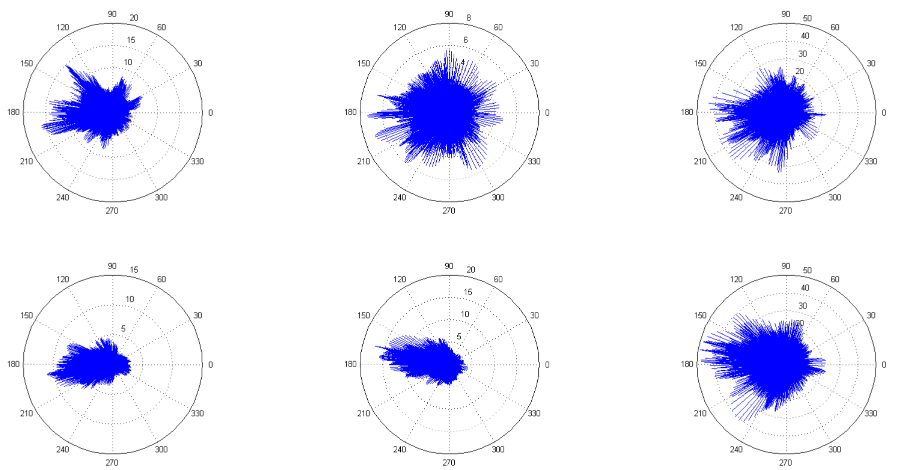

*Figure 4: Sound intensity vector polar plot in the audience position. Upper row: results obtained with multi-source approach; left to right: sopranos only, basses only, whole choir. Lower row: results obtained with single point-like source approach; left to right: sopranos only, basses only, whole choir.*

Beside the signal analysis, a listening test with direct comparison of the two techniques can clearly show the advantages and improvements conveyed by this new approach. The quadraphonic acoustic field created by convolution can be rendered in surround sound by means of a suitable *transcoding* process [Bonsi et al., 2005], which allows the signals feeding a given configuration of loudspeakers to be derived by linear combinations of pressure and velocity. Our choice has fallen on the 5.1 surround configuration [ITU-R, 1992], which employs five loudspeakers in a circle around the listener, at angles of 0°, ±30° and ±110°, plus a subwoofer for the low frequency enhancement. For the playback a DVD-Audio was authored, with the six channels of linear PCM audio feeding a surround loudspeaker system setup in FSSG-CNR anechoic chamber. The listening test confirmed that the multi-source convolution renders a sound image that is closer to the real event perceived by the conductor, where both the choir sections positions can be identified and the reverberation of the church is reproduced in a natural and convincing way, while in the case of the single point source the ability to discriminate the placement of the various sections is lost.

## CONCLUSIONS

The multi-source approach presented in this paper represents a step forward if compared with the single point-like source method of rendering the sound field created by a spatially extended source in an acoustic environment. Both measurements and listening tests have outlined the improvement in retaining the spatial characteristics of the sound source, thus offering a more accurate representation of the real event. This method can successfully be applied to any extended sound source such as music groups and orchestras, provided that a reasonable choice is done for the decomposition of the source into a set of discrete point-like sources. It is worth noting that this kind of approach, apart from its benefits in retaining the direction of arrival of sound at the listener's perspective, can also be used for rendering the field depth, that is the correct distance of the various elements of the source. Thanks to the compatibility between Acoustic Quadraphony and the industry standard formats for audio, such as stereo and 5.1 surround, the results can be implemented within the standard workflow of music production and cinema post-production.

### Acknowledgments

The authors would like to thank the choir "Coro Sante Sabide" from Goricizza (Udine – Italy) and the choirmaster m° Cristian Cozzutti for their kind collaboration and logistical support during the recording, the Benedictine monks community of the church of San Giorgio Maggiore in Venice for their kind hospitality and Mr. Alessandro Ceccato for the support during the measurement session.

## REFERENCES

Stanzial, D.; Bonsi, D.; Schiffrer, G. (**2003**) "Equipment for collection and measurement of quadraphonic sound data and metadata and relative procedure", European Patent n. 04425909.1, registered: Dec 2, 2003.

Stanzial, D.; Bonsi, D.; Schiffrer, G. (**2002**). "Four-dimensional treatment of linear acoustic fields and radiation pressure", Acta Acustica united with Acustica 89, pp. 213-224

Bonsi, D.; Gonzalez, D.; Stanzial, D. (**2005**). "Quadraphonic impulse response for acoustic enhancement of audio tracks: measurement and analysis", *Proceedings of Forum Acusticum*, Budapest, p. 335-340.

Bonsi, D.; Fontana, F.; Gonzalez, D.; Schiffrer, G.; Stanzial, D. (**2005**). "Report on technology for the development of intensimetric control of sound fields", IP-RACINE deliverable D 7.5.1 IST-2-511316

ITU-R Rec. BS.775-1 (**1992-94**) Multichannel Stereophonic Sound System with and without Accompanying Picture, International Telecommunications Union, Geneva.