# A Formulation of Ambisonics in Unconventional Geometries

| | |
|---|---|
| 著者 | Trevino Jorge |
| 号 | 19 |
| 学位授与機関 | Tohoku University |
| 学位授与番号 | 情博第575号 |
| URL | http://hdl.handle.net/10097/58154 |

| | |
|---|---|
| 氏名（本籍地） | ﾄﾚﾋﾞｰﾆｮ　ホルヘ<br>TREVINO　JORGE |
| 学位の種類 | 博　士（情報科学） |
| 学位記番号 | 情　博　第　575　号 |
| 学位授与年月日 | 平成２６年　３月２６日 |
| 学位授与の要件 | 学位規則第４条第１項該当 |
| 研究科、専攻 | 東北大学大学院情報科学研究科（博士課程）システム情報科学専攻 |
| 学位論文題目 | A Formulation of Ambisonics in Unconventional Geometries<br>（非球形のアレイ形状を持つ Ambisonics の定式化） |
| 論文審査委員 | （主査）東北大学教　授　鈴木　陽一<br>東北大学教　授　菅沼　拓夫　　東北大学教　授　伊藤　彰則（工学研究科）<br>東北大学准教授　坂本　修一 |

# 論 文 内 容 の 要 旨

## 第１章　Introduction

There is a growing demand for systems capable of realistically conveying the illusion of being immersed in a different, real or virtual, environment. Applications in areas such as entertainment, communications, telepresence and psychophysical research require the faithful presentation of sensorial information to one or multiple users.

Humans are sensitive to the spatial features of sound fields. We can accurately determine the direction from which sound reaches our ears, judge the distance to a given source, perceive the effects of obstacles within the sound propagation path, assess the properties of listening rooms through reverberation, and, to some extent, estimate the size and shape of sound sources. A system intended to present a realistic auditory scene to its users must be able to convey all these information.

This dissertation contributes towards the realistic presentation of auditory information by formulating new methods for the synthesis and reproduction of sound fields. A sound field reproduction system attempts to re-create the complete sound field across an extended region surrounding the listener. If the acoustic intensity around the user could be re-created, the resulting experience would be indistinguishable from that of actually being at the presented scene.

## 第２章　Mathematical modeling of sound fields

Sound field reproduction makes extensive use of the Kirchhoff integral theorem. Using this theorem, it is possible to recover the complete sound field within a source-less region from acoustic intensity measurements on the boundary enclosing it. The result reduces the dimensionality of the problem, greatly decreasing the amount of measurements and processing required.

Unfortunately, the dimensionality reduction gained from the Kirchhoff integral theorem is not enough to design a practicable system. Sound fields carry too much information for exact reproduction. A compromise must be made between the reproduction system's complexity and its accuracy. Since different applications might demand varying degrees of precision, a scalable format for the characterization of sound fields is required.

First attempts to formulate such an scalable encoding format led to a technique known as Ambisonics. In recent years, research surrounding an extension to this format, based on the spherical harmonic functions, has been very active. This is the so-called High-Order Ambisonics (HOA) encoding.

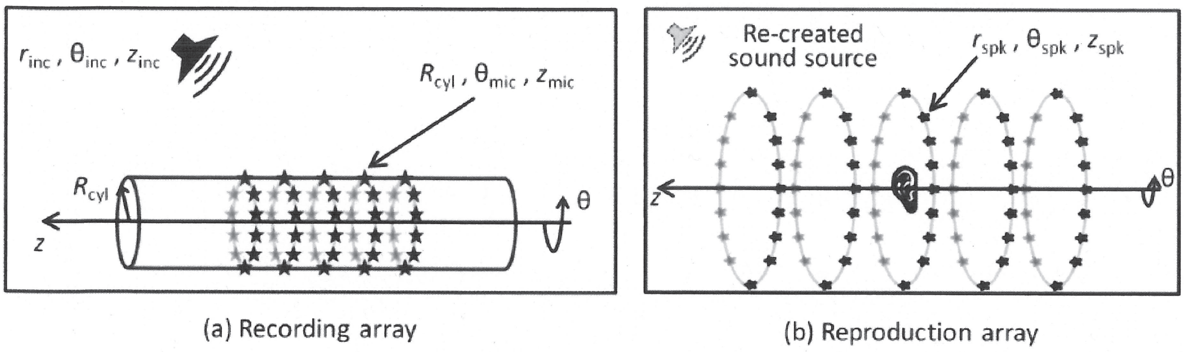



Figure 1: Layout of cylindrical recording and reproduction arrays. The black stars represent the transducers. The cylindrical geometry defines a privileged axis which can be made parallel to the horizontal plane for accurate sound field presentation while limiting the number of transducers.

## 第 3 章　Plane wave decomposition of cylindrical sound pressure distributions

The formulation of HOA is very simple and elegant. Sound pressure on a boundary surface is either calculated in a computer simulation or measured using a microphone array. The result is then expressed as a weighted-sum of spherical harmonics. The expansion coefficients, truncated at some degree, are the HOA encoding. Reliance on the spherical harmonics, however, leads to some constrains on this technique. The boundary conditions must be sampled regularly in terms of solid angles. Furthermore, distance is not part of the encoding, thus requiring the boundary to be spherical with all samples taken at an equal distance from a fixed observation point.

This chapter proposes the use of cylindrical geometries in the design of microphone and loudspeaker arrays. This approach is illustrated in Fig. 1. To use these kind of arrays, the solutions to the Helmholtz equation are derived in cylindrical coordinates. A significant result found in these solutions is the set of functions known as the cylindrical harmonics.

The cylindrical harmonic functions will eventually become a crucial building block for the encoding of sound fields recorded with cylindrical microphone arrays. In this sense, they are similar in importance to the spherical harmonics which lie at the core of HOA.

A least-squares approach to encode the sound fields recorded by cylindrical microphone arrays is presented in this chapter. This method works only in the far field; it does not consider distance information. However, the mathematical techniques used to define this encoding, known as the plane wave expansion, can be applied to more complex encodings which include distance information. This will be one of the main topics of this dissertation and is elaborated in the following chapters.

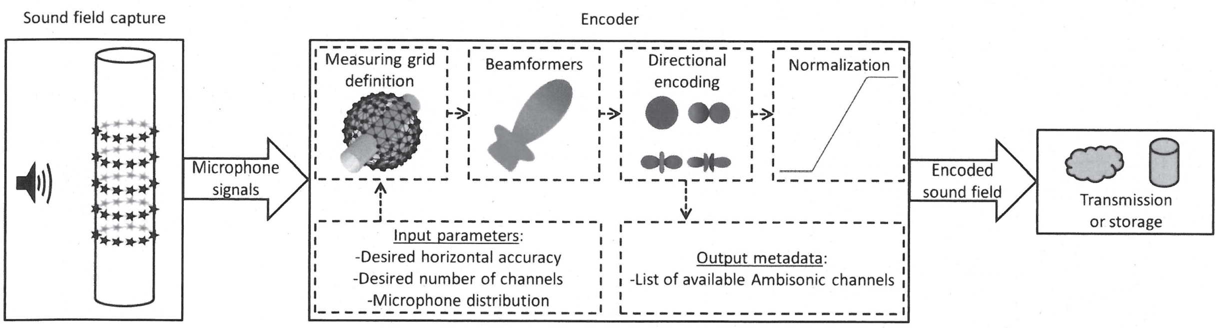



Figure 2: Block diagram for the proposed system's encoding stage. A cylindrical microphone array is used to sample the sound field over a spherical measuring grid. The result is encoded using the spherical harmonic functions. The microphone array has different resolutions for azimuth and elevation. The result is a Mixed-Order Ambisonics encoding of the sound field.

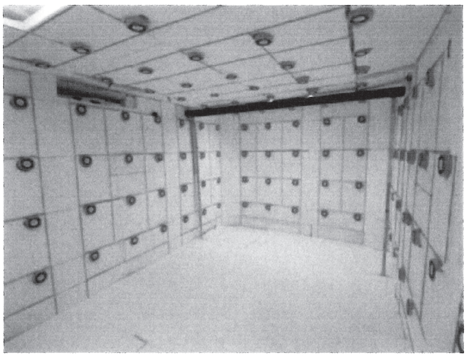


Figure 3: Photograph of a 157-channel, irregular loudspeaker array at the Research Institute of Electrical Communication. Its construction is simpler than that of a spherical array; however, conventional HOA decoders cannot be effectively applied to it.

## 第 4 章　Mixed-Order Ambisonics for cylindrical microphone arrays

Human auditory perception does not exhibit spherical symmetry. We show greater spatial accuracy when localizing sound sources in the horizontal plane. It is, thus, useful to characterize sound fields with a greater emphasis on their horizontal features. Several existing systems attempt to do this for both sound field recording and reproduction. They can be classified into three categories: plane-wave decomposition, wave field reconstruction and mixed-order Ambisonics. Methods based on the plane-wave decomposition can be applied to virtually any geometry by expressing a given sound field as a sum of plane waves. However, this leads to the loss of distance information; what is known as the evanescent field. Attempts to encode the evanescent field are, at the moment, limited to spherical harmonic expansions. The wave field reconstruction approach is system-dependent since it takes the form of a set of filters that matches given recording and reproduction arrays without an intermediate encoding.

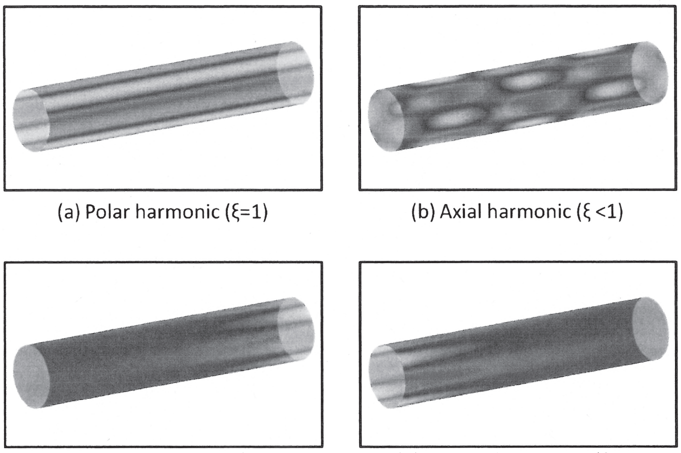
In this chapter, a mixed-order Ambisonics method is used to increase horizontal accuracy by using a large set of horizontally-oriented spherical harmonics. This approach normally requires a spherical boundary and considers a single observation point. In this dissertation, the alternative recording system shown in Fig. 2 is introduced.

The new proposal results in a mixed-order Ambisonics encoder and its output is system-independent. This was achieved by defining a measuring grid which does not necessarily uses the same distribution as the microphone array. The measuring grid is defined using two parameters provided by the user, leading to a scalable encoding using considerably less channels than previous methods, such as those based on the helical wave spectrum. More specifically, the proposed method results in $M^2 + 2N + 1$ channels, where the linear dependency N sets the system accuracy in the horizontal plane. The presented system is considered to be future-proof since it is compatible with the Ambisonic B-format with little modifications.

## 第 5 章　Decoding generalized Ambisonics for arbitrary loudspeaker configurations

The spherical geometry used in HOA leads to simple encoding and decoding equations. However, it poses difficulties for its practical use. Spherical microphone arrays can be easily build and transported, however, the spherical geometry is difficult to preserve when building a loudspeaker array. A more realistic loudspeaker array is shown in Fig. 3. The loudspeakers in this system cover the walls and ceiling of a rectangular room. The loudspeaker separation is constant; therefore, it does not offer a regular sampling of the sphere.

This chapter focused on sound field reproduction, specifically that of HOA-encoded sound fields. Conventional HOA encodings and their decoders do not consider distance explicitly since they assume that the far-field approximation holds. Furthermore, transducers in both the recording and reproduction arrays are considered to be equidistant from the observation point.



(a) Polar harmonic (ξ=1) (b) Axial harmonic (ξ <1) (c) Damped harmonic (ξ >1, positive orientation) (d) Damped harmonic (ξ >1, negative orientation)

Figure 4: The cylindrical harmonic functions of fourth polar order. Three types of cylindrical harmonics exist. The polar harmonics are constant along the axial coordinate. The axial harmonics present an oscillatory behavior along both coordinates. The damped harmonics show an exponential tendency in the axial coordinate.

A new HOA decoding method was introduced by taking into consideration the difficulty of building a spherical loudspeaker array, as well as the desire for a larger listening region to accommodate the listener(s). Its main difference when compared with a conventional decoder is the introduction of a new error metric which penalizes large variations in the reconstruction error as the listener moves away from the center. This is mathematically expressed as the minimization of the radial derivative of the reconstruction error.

The performance of the proposal was compared with that of a conventional HOA decoder. Not only was the accuracy in the reconstruction of the field considered for evaluation, but also the interaural cues that each decoding method conveys to the listener.

## 第 6 章　3D Cylindrical Ambisonics

When considering a cylindrical geometry, the present research introduces a new format called 3D Cylindrical Ambisonics (3DCA). The encoding is based on the cylindrical harmonic functions; these are illustrated in Fig. 4. The proposal can use different resolutions to encode sound field information in the horizontal plane and elevation angles.

The most salient advantage of the encoding methods being proposed in this dissertation is that sound field information can be broadcast to multiple users who use different reproduction systems. On the listener's end, a decoder is used to generate suitable signals for virtually any specific loudspeaker system. A user with an advanced loudspeaker array can get a very realistic experience, while others can still decode and experience the same contents within the limitations of their specific reproduction systems.

Spherical or cylindrical arrays of loudspeakers make it straightforward to decode HOA or 3DCA. However, they are difficult to build and use. As part of this dissertation, loudspeaker arrays of arbitrary geometries are considered and a new decoding approach is advanced.

## 第 7 章　Conclusions

This dissertation outlines the basis for sound field recording, analysis and reproduction using non-spherical arrays and in particular a cylindrical microphone array. It is hoped that the proposal presented here, in particular 3D Cylindrical Ambisonics can become an important step towards the realization of ultra-realistic spatial audio systems.

## 論文審査結果の要旨

高臨場感情報通信の高度化を実現するうえで，3次元音空間情報の集音・再生技術は重要な基盤である。特にキルヒホッフ・ヘルムホルツ積分方程式に基づいて音場を再現する技術は，音場の精密な再現が可能であることから原理的に優れている。そのような技術の中で，集音系と再生系を完全に独立に取り扱えるという性質を持つAmbisonics法はネットワークを介して音場情報を共有する上で高い有効性を持っている。しかし，従来のAmbisonics法は音場を球座標上で調和解析する球面調和解析法を用いているため，再現される音場の精度が球状に均一となってしまうだけでなく，球上ほぼ均等にスピーカを配置しないと音場の再現領域が狭くなるという問題点を有していた。著者は，Ambisonics法の有効性に着目し，その問題点の解決と高度化を図るための研究を進めてきた。本論文は，その成果をまとめたもので，全編7章からなる。

第1章は序論である。

第2章では，音場の球面調和解析を基本原理とする従来のAmbisonics法について述べている。

第3章では，円筒座標系を用いれば軸方向と半径方向で独立に再現音場の精度を制御可能であり，したがって人間の音像定位能力が上下よりも水平方向の方が優れているという特性を生かしうる集音・再生が可能になると発想して，球座標系に代えて円筒座標系調和解析を用い，集音系と再生系を独立に設計できる新しい音場情報集音・再生法の理論式の定式化を行っている。これは独創的で，高い有効性を持つ成果と高く評価できる。

第4章では，円筒形のマイクロフォンアレイを用いて集音した音場情報を，従来のAmbisonicsを用いつつ，円筒座標系に合わせて水平方向と上下方向で異なる精度で再生するアルゴリズムを与えている。これは，自らの提案の応用範囲を大きく広げる成果として高く評価できる。第5章では，任意のスピーカ配置を用いて高精度の音場再生を可能とする新しいAmbisonics 再生アルゴリズムを与えている。これは提案法のみならず従来のAmbisonicsの応用範囲をも大きく広げる優れた提案と評価できる。

第6章では，第3章の成果に基づき，3D Cylindrical Ambisonicsと呼ぶ実用的な音場情報集音・再生信号処理体系の定式化と定量評価結果を与えている。これは，円筒座標系調和解析法を任意のスピーカ配置に適用して，システム規模に応じた精度と音場再生領域を実現する実用的な音場情報集音・再生信号処理体系であり，有用性に富む成果として高く評価できる。

第7章は結論である。

以上要するに本論文は，高度の臨場感を持つ音場情報を高精度に集音し，集音系とは独立に設計された再生系において高精度音場再生を実現する新しい音信号処理方式を定式化したもので，システム情報科学ならびに音情報科学の発展に寄与するところが少なくない。よって，本論文は，博士（情報科学）の学位論文として合格と認める。