5章 アンビソニックス

西村 竜一[†]

キーワード：自由視線、音場回転、マイクロホン、指向特性、球調和関数

1. まえがき

立体音響技術には、種々の方法が存在する。その中にあってアンビソニックスの特徴は、空間や聴取者の頭が回転した際の音場の変化を、ソフトウェア的に容易に再現できる点にある。この特性は、例えば、ストリートビューにおいて、インタラクティブに視線（視点ではない）を変えた映像と組合せて利用する場合に、魅力的に感じられるであろう。あるいは、図1のようなスポーツ観戦において、ある観客目線での映像と音響を伝送して再現したい場合、視点の平行移動はあまり起こらないが、視線の方向は自由に

図1 スポーツ観戦中の頭の自由度

動く可能性がある。アンビソニックスは、このような状況下での立体音の生成に、特に有効な手法である。

アンビソニックスは、英語では Ambisonics と表記され、形容詞的に用いられる場合には、アンビソニック (Ambisonic) と表現される。最初は、レコーディングのテクニックとして、現場におけるオーディオエンジニアにより開発された。その後、2000年頃から球調和関数展開による音場再現や数学的解析の研究が盛んに行われ、その一部に位置付けられる手法として再認識された。球調和関数展開は、方向性パターンに対するフーリエ変換に相当し、球面上で異なる周期を持った直交する成分に分解して解析する手法である¹⁾。アンビソニックスは、球調和関数の0次と1次のみに着目した手法とみなされる。一方、2次以上の高次の球調和関数までを考慮した手法は、Higher Order Ambisonics (HOA) と呼ばれ、従来のアンビソニックスと区別することができる。高い次数までを考慮することで、単に音の到来方向を再現するばかりでなく、究極的には広い領域で原音場を再現することが可能となる。

アンビソニックスによる音場収録・再生は、概ね、図2に示すような構成となる。本稿では、この構成に基づき、アンビソニックスによる収録と音場の操作、ならびに再生

図2 アンビソニックス収録・再生の基本構成

[†] 独立行政法人情報通信研究機構
"Ambisonics" by Ryouichi Nishimura (National Institute of Information and Communications Technology (NICT), Kyoto)

について、基本的な考え方を解説する。より詳細な技術内容については、文献2)や文献3)等を参照頂きたい。

2. 収音方法

アンビソニックスにおける収音には、アンビソニックマイクロホンと呼ばれる特殊なマイクが用いられる。図3に示すように、さまざまな種類のもが市販されているが、基本的な構造はどれも同じであり、図3(b)、図3(d)も、メッシュでカバーされているだけで、内部には図3(a)、図3(c)と同様な配置で4つのマイクロホンカプセルが存在している。正四面体(給食の牛乳のテトラパックの形状)の各面に、外向きにマイクロホンが付いている状況を想像すると、その配置がイメージできるであろう。これらの4つのマイクロホンカプセルで収音された信号は、まとめてA-フォーマットと呼ばれる。アンビソニックスでは、この信号をさらに、B-フォーマットと呼ばれる図4に示す収音指向特性で収音した場合の信号表現に変換して使用する。0次は無指向性のマイクロホンでの収音信号に対応し、1次は互いに空間的に直交する双指向性 (figure-of-eight) のマイクロホンで収音した場合に対応する。したがって、アンビソニックマイクでは、音圧と音圧傾度(ベクトル)に相当する量が

得られる。

A-フォーマットからB-フォーマットへの変換は、次のようにして行われる。図5、図6において、1つのアンビソ

図4 B-フォーマットでの収音指向特性

(a) Core Sound: TetraMic
(引用: <http://www.core-sound.com/TetraMic/1.php>)

(c) SoundField: SPS200
(引用: <http://www.tslproducts.com/soundfield/soundfieldsps200-software-controlled-microphone/>)

(b) Embrace Cinema Gear: Brahma
(引用: <http://www.embracecinemagear.com/brhma-sound-field-microphone-tetra-mic-head-ambisonics-design-1.html>)

(d) SoundField: MKV Microphone

図3 市販の種々のアビソニックマイクロホン

図5 水平面の指向特性の合成（上側から見た図）

図6 垂直面の指向特性の合成（左側から見た図）

ニックマイクに、4つのマイクロホンカプセルが使用されており、 L_F は左前方の上向き、 L_B は左後方の下向き、 R_F は右前方の下向き、 R_B は右後方の上向きに配置されているものとする。それぞれのマイクロホンは、基本的には単一指向性を有しており、図中の小さい円（実線と2点鎖線）がそれらを示している。これらを合成して、より大きな円（破線と一点鎖線）で示された特性での收音信号を取得したい。それには、

$$X = L_F - R_B + R_F - L_B \quad (1)$$

$$Y = L_F - R_B - (R_F - L_B) = L_F - R_B - R_F + L_B \quad (2)$$

$$Z = L_F - L_B + R_B - R_F \quad (3)$$

$$W = L_B - L_F + R_F - R_B \quad (4)$$

の計算をすればよい。ここで、 X 、 Y 、 Z はそれぞれの軸方向の指向性成分に、 W は無指向性成分に対応する。このように、各マイクロホン信号で得られた信号の単純な加減算で、これらの信号変換を実現することができる⁴⁾⁵⁾。

任意に配置した多くのマイクロホンを用いて收音した信号からでも、デジタル信号処理を用いれば、同様の收音特性の信号を取得することは可能である。しかし、その場合には、遅延や有限長のフィルタを用いることに起因する歪みが收音信号に重畳することになる。これに対し、最小限のマイクロホン数で、かつ、極めて単純な計算でこれらの信号を生成しているところが、アンビソニックスの特質である。

このような信号表現を用いることで、ポストプロダクションにおいて、音場全体を回転する操作も容易にできるようになる。例えば、視聴者が顔の向きを水平面上で ϕ だけ変えた時に聞こえる音場を再現したい時には、

$$\begin{bmatrix} W' \\ X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \phi & -\sin \phi & 0 \\ 0 & \sin \phi & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} W \\ X \\ Y \\ Z \end{bmatrix} \quad (5)$$

のような3次元の座標変換の行列をB-フォーマット信号に作用させることで、容易に生成することができる。これは、アンビソニックマイクロホンが收音時に ϕ だけ回転していた時に收音されたはずの信号を、後から生成することができるということでもある。図2に示した基本構成の中に描いたように、B-フォーマット信号の合成の際の係数の選び方次第では、異なる指向特性のマイクロホンで收音した時の信号を合成することも可能である。このような音場や收音特性の操作を行う製品は、すでに市販されている。また、ソフトウェアで実現できることから、種々の Virtual Studio Technology (VST) プラグインも存在している。従来のステレオマイクロホンで收音した信号に対して、後から空間情報を操作するのは、とても難しい処理であり、多くの場合、歪みを伴う。なお、アンビソニックス関連の製品やソフトウェアについての情報は、文献6)に記したウェブサイトから得ることができる。

3. 再生方法

立体音を再現するには、ラウドスピーカの配置に応じて、各ラウドスピーカを適切な信号で駆動しなければならない

図7 ラウドスピーカを四隅に配置した場合

い。これらの信号は、B-フォーマットの各信号に対して、適当な係数を掛けて足し合わせることで合成される。例えば、図7に示すような、4つのラウドスピーカが水平面上等間隔に、聴取者を取り囲むように配置されている場合を考える。聴取者の左前方のラウドスピーカを L_F 、右前方を R_F 、左後方を L_B 、右後方を R_B とする。この時、各ラウドスピーカの駆動信号は、B-フォーマット信号を元にして、

$$L_F = W + 0.707(X + Y) \quad (6)$$

$$R_F = W + 0.707(X - Y) \quad (7)$$

$$L_B = W + 0.707(-X + Y) \quad (8)$$

$$R_B = W + 0.707(-X - Y) \quad (9)$$

により合成される⁷⁾。さらに高さ方向も再現したい場合には、図8のように部屋と天井の四隅にラウドスピーカを配置して、

$$L_{FU} = W + 0.707(X + Y + Z) \quad (10)$$

$$R_{FU} = W + 0.707(X - Y + Z) \quad (11)$$

$$L_{BU} = W + 0.707(-X + Y + Z) \quad (12)$$

$$R_{BU} = W + 0.707(-X - Y + Z) \quad (13)$$

$$L_{FD} = W + 0.707(X + Y - Z) \quad (14)$$

$$R_{FD} = W + 0.707(X - Y - Z) \quad (15)$$

$$L_{BD} = W + 0.707(-X + Y - Z) \quad (16)$$

$$R_{BD} = W + 0.707(-X - Y - Z) \quad (17)$$

により駆動信号を合成できる⁷⁾。厳密な話をすると、壁からの反射の影響も考慮する必要があるが、ここでは基本的な話にとどめることにする。

図8 ラウドスピーカを部屋と天井の4隅に配置した場合

先にも述べたように、さまざまな指向特性のマイクロホンで収音した場合の信号を合成できることから、1つのB-フォーマット信号を元にして、5.1chサラウンドの各信号への変換も可能である⁸⁾。さらにヘッドホン再生の場合も、アンビソニックデコーダを用いてラウドスピーカ用の信号に一度変換してから、各ラウドスピーカの位置に対応する頭部伝達関数を畳み込んで足し合わせることで、バイノーラル信号を生成することが可能である⁹⁾。

4. 今後の展望

アンビソニックスで使用されるB-フォーマット信号は、マイクロホンの位置における音圧を球調和関数展開した時の、0次と1次の信号成分に相当する。1次の信号成分が表す音圧傾度は、音圧の空間的な微分(勾配)であることから、音源の数が1つで反射波もない場合には、音の到来方向が再現できると考えられる。しかし、音源が複数ある場合には、それらの音波が重畳されて音場が形成されるため、単純に一点における音圧傾度だけからでは、音源の方向は判断できない。そのような場合でも正しく音場を再現するには、より高次の球調和関数成分までを再現する必要がある。そのような手法がHigher Order Ambisonicsである。高い次数まで再現することで、収音時と同じ音場が再現される空間領域が拡大するが、高い次数まで観測するには多くのマイクロホンカプセルを集積配置する必要がある。その実現の困難さに起因する誤差が精度の低下を招く。そのため、市販されているアンビソニックマイクロホンは通常1次までの信号を観測するものが一般的であり、大学等における研究目的でも、5次程度までの信号成分を再現できるシステムまでしか開発されていないのが現状である¹⁰⁾。

アンビソニックデコーダの役割は、複数のラウドスピーカで再生した音場を、聴取者の位置にアンビソニックマイクロホン置いて収録した時に、原音場で収録した時と同じB-フォーマット信号が得られるようにラウドスピーカの駆動信号を決めることである。B-フォーマットが4チャンネルの信号であることから、使用するラウドスピーカの数4を超える場合には、劣決定系の問題となり、一般には複数の解が存在する。前述のHigher Order Ambisonicsにおいて次数を上げるということは、この自由度を減らし、原音場により近い音場に同定する作業に他ならない。次数を上げることで自由度を減らす以外にも、音場に何らかの仮定を設け、事前知識による制約を加えることで間違っただけに陥らないようにするアプローチも考えられる。これらについては、今後の研究課題となろう。

5. むすび

アンビソニックスは、収録に際してマイクロホンを分散配置することなく、音の空間情報を取得することができる。また、音場全体を回転操作するような音の空間情報に対する後処理も容易に行える。これらの特性は、アンビソニックスを実用的な手法として、魅力的なものにしている。すでに市販のプロダクトも存在して利用できるが、より精度の高い音空間の再現技術に発展するポテンシャルを有した手法である。そのためには、マイクロホンカプセルの小型集積化を進めたり、学習や知識付与による制約を用いるなど、新しい研究開発も必要となろう。これらの進展により、より臨場感のある映像・音響メディアが実現できるようにすることを願っている。

(2014年4月25日受付)

〔文 献〕

- 1) E.G.Williams: "Fourier Acoustics -Sound Radiation and Nearfield Acoustical Holography-", Academic Press (1999)
- 2) R.K. Furness: "Ambisonics - An Overview", AES 8th International Conference: the Sound of Audio, pp.181-190 (May 1990)
- 3) 安藤彰男: "高臨場感音響技術とその理論", IEICE Fundamentals Review, 3, 4, pp.33-46 (Apr. 2010)
- 4) K. Farrar: "Soundfield microphone", Wireless World, 85, pp.48-51 (Oct. 1979)
- 5) K. Farrar: "Soundfield microphone - 2", Wireless World, 85, pp.99-103 (Nov. 1979)
- 6) <http://www.ambisonic.net/>
- 7) D.G. Malham and A. Myatt: "3-D Sound Spatialization using Ambisonic Techniques", Computer Music Journal, 19, 4, pp.58-70 (1995)
- 8) SoundField SP451 Surround Processor, User Guide, Ver. 1.02
- 9) M. Noisternig and A. Sontacchi and T. Musil and R. Höldrich: "A 3D Ambisonic Based Binaural Sound Reproduction System", AES 24th International Conference: Multichannel Audio, the New Reality (June 2003)
- 10) T. Okamoto, Z.-L. Cui, Y. Iwaya and Y. Suzuki: "Implementation of a high-definition 3D audiovisual display based on higher-order Ambisonics using a 157-loudspeaker array combined with a 3D projection display", International Conference on Network Infrastructure and Digital Content (ICNIDC), pp.179-183 (Sep. 2010)

にしむら りょういち
西村 竜一

1998年、東北大学大学院情報科学研究科博士後期課程修了。同年、(株)ATR知能映像通信研究所客員研究員。2000年、東北大学電気通信研究所助手。2014年、同大学助教授。2006年より、(独)情報通信研究機構にて、超臨場感音響の研究に従事。博士(情報科学)。