2章 マルチチャネルオーディオ

[正会員] 小野 一穂[†]

キーワード：マルチチャネル、オーディオ、22.2、立体音響、3次元音響

1. まえがき

マルチチャネル音響は、空間内の1点を基準に、方向の異なる複数のチャンネルを設置し、その各チャンネルが音の到来方向と1対1に対応する方式である。

空間内でのスピーカ配置は離散的になり、再生する音場の物理的な再現精度を高めることはできないため、スピーカの方向に対しては心理音響効果により音の方向感をコントロールする。このため、心理音響モデル方式と呼ばれることもある。

マルチチャネル音響の基本は2chステレオ技術にある。そこで本稿では、2chステレオの原理について述べ、デジタル放送やパッケージメディアで広く用いられている5.1chサラウンド、そして5.1chサラウンドを超える立体音響方式について解説する。

2. 2chステレオ

2.1 2chステレオの原理

2chステレオは、一般に60°の間隔をおいておかれた2台のスピーカを用いて音の方向や広がり表現する方式である。離散的に置かれたスピーカであるが、2台のスピーカに同じ信号を振幅差をつけて再生することにより、2台のスピーカの間に、音像定位が生じる、すなわち実際の音源があるかのような方向感を与えることができる。

2.2 振幅パンニング

人間の音像定位方向の知覚は、左右の耳における音の振幅差、位相差、周波数特性の違いなどによって決まる。このうち、水平方向の音像定位知覚に特に重要なものは、両耳間振幅差 (ILD: Interaural Level Difference) と両耳間時間差 (ITD: Interaural Time Difference) である¹⁾。両耳間時間差の代わりに両耳間位相差で表現することもある。単一の実音源が耳の高さの水平面内を左右方向に移動する場合、例えば正面方向では振幅差も時間差もゼロであるが、実音源が左右に移動すると、音源からの距離の差が生じて

図1 2chステレオの原理
 g_L と g_R は左右のスピーカに与える振幅ゲイン。

両耳間時間差が生じると同時に、音源の方向と反対側の耳では頭によって音が遮蔽されるためレベルが減衰し、両耳間振幅差が生ずる。

2chステレオは、図1に示すように耳の高さの水平面内に左右2台のスピーカを配置し、左右のスピーカに振幅差を与えて音の聞こえる方向、すなわち音像の方向を制御する。2台のスピーカから同じ信号を同レベルで再生すると音像は正面に定位するが、振幅差をつけるにしたがって振幅の大きいスピーカの側に音像が移動し、最終的に片側のスピーカ上に音像が定位する。このように、2台のスピーカに振幅差をつけて再生することを、一般に振幅パンニングという。

図1において、 θ_0 は正面から各スピーカまでの角度、 θ は正面から仮想音像までの角度を示す。振幅差と方向を関係づけるパンニングの式は複数提案されているが、よく知られたものに、サイン則とタンジェント則²⁾がある。サイン則では、角度 θ_0 に音を定位させるために、左と右のスピーカに与えるゲイン g_L と g_R を

[†] 一般財団法人NHKエンジニアリングシステム
"Multichannel Audio" by Kazuho Ono (NHK Engineering System Inc., Tokyo)

$$\frac{g_L - g_R}{g_L + g_R} = \frac{\sin \theta}{\sin \theta_0} \quad (1)$$

とする．一方タンジェント則では，ゲインを，式

$$\frac{g_L - g_R}{g_L + g_R} = \frac{\tan \theta}{\tan \theta_0} \quad (2)$$

によって定める．

サイン則，タンジェント則ともに，両耳間時間（位相）差が実音源に近づくように算出された係数であり，サイン則は，頭を正面位置に固定した場合のパンニング則，タンジェント則は頭を仮想音像の方向に向けた場合のパンニング則と言われている²⁾．振幅パンニングによる再生音は，実音源の両耳間時間差と両耳間振幅差をよく模擬することが知られている．

ところで，実音源が左右方向に移動すると，両耳で振幅差とともに位相差（時間差）が生ずる．一方，2chステレオでは左右のスピーカの信号に加えるのは振幅のゲインだけであるが，左右それぞれのスピーカから再生される音が各々両耳で振幅差と位相差を有しているため，左右のスピーカに与えるゲインを変えるだけで実音源と同様に振幅差と位相差を生じさせることができる．

なお，2chステレオでは，振幅パンニングによる仮想音像定位が成立するのは，スピーカから等距離の位置のみである．聴取位置からスピーカまでの距離が異なる場合は，近い方のスピーカによる両耳間時間差が優勢となり，近い方のスピーカ付近に音像が定位することが問題となる．

3. 5.1 ch サラウンド

5.1 ch サラウンドは，ITU-R 勧告 BS.775 に規定された音響システムである³⁾．図2に示すスピーカ配置による前方3ch，後方2ch，そのほかに低域専用1チャンネルを有する．

5.1 ch サラウンドにおいては，振幅パンニングにより横や後ろの仮想音像を再生すると同時に，聴取者が左右前後から音に包み込まれる効果が得られ，臨場感や迫力が増大する．このため，5.1 ch サラウンドは，2chステレオに続く標準的なフォーマットとして，デジタル放送，DVDなどのパッケージメディア，ゲームなどで，幅広く採用されている．

4. 立体音響方式

4.1 動 向

技術の発展に伴い，5.1 ch サラウンドを超える立体音響方式も提案されてきた．これらは，次の三つに大別される．

- (1) チャンネルベース方式：2chステレオや5.1chサラウンドを拡張した，チャンネル配置が予め定められた方式
- (2) オブジェクトベース方式：素材音（オブジェクト）とその位置や動きの情報を分けて記録・伝送し，再生環境に合わせて構成する方式
- (3) シーンベース方式：音場全体の物理的な情報を記

図2 5.1 ch サラウンド

録・伝送する方式

これらの方式を網羅的に規定する国際規格が2014年2月にITU-R 勧告 BS.2051 として発行された⁴⁾．マルチチャンネル音響は主としてチャンネルベース方式に含まれ，ITU-R 勧告 BS.2051 においては，チャンネルの位置を左右方向の方向角 (Azimuth) と高さ方向の仰角 (Elevation) により規定し，垂直方向に3層以下のチャンネルを有する方式について記述している．例として，5.1 ch サラウンドの水平面内にチャンネルを追加した7.1 ch 方式，5.1 ch サラウンドを高さ方向に2層配置した10.2 ch 方式，8K スーパーハイビジョンの音響方式であり，高さ方向に3層のチャンネルを配置した22.2 マルチチャンネル音響（以下，22.2 ch 音響）などが述べられている．また，Dolby Atmos (ドルビーアトモス)⁵⁾ で用いられているチャンネルベース方式とオブジェクトベース方式を組合せた方式についても述べられている．なお，22.2 ch 音響と Dolby Atmos については後節で詳しく紹介する．

4.2 3次元パンニング

5.1 ch チャンネルを超える方式のうち，スピーカを高さ方向にも配置する方式では，振幅パンニング技術も3次元的に考慮する必要がある．このために，2chステレオのパンニング技術を拡張した手法が考案されている．ここでは，ベクトル合成に基づくVBAP (Vector Base Amplitude Panning)⁶⁾を中心に紹介する．

VBAPは，仮想音像を，近接した方向の三つのスピーカから再生する際，聴取位置からスピーカを向く3方向のベクトルを，スピーカに与える利得で重みづけして加算した合成ベクトルの向きが，仮想音像の方向と一致するように利得を定める方式である．詳細は以下の通りである．

図3に示すように，仮想音像を向く単位ベクトルを $\mathbf{p} = (p_1 \ p_2 \ p_3)^T$ ，聴取位置から3個のスピーカを向く単位長さのベクトルを $\mathbf{l}_1 = (l_{11} \ l_{12} \ l_{13})^T$ ， $\mathbf{l}_2 = (l_{21} \ l_{22} \ l_{23})^T$ ， $\mathbf{l}_3 = (l_{31} \ l_{32} \ l_{33})^T$ とする．ここで， T は行列またはベクトルの転置を表す．各スピーカに与えるゲインを g_1, g_2, g_3 とする．VBAPでは

図3 VBAF (Vector Base Amplitude Panning)

まず、仮想音像の方向 $\mathbf{p}$ が与えられたとき、 $\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3$ の作る立体角内に $\mathbf{p}$ が含まれるように、3つのスピーカーを選択する。次に、

$$\mathbf{p} = g_1 \mathbf{l}_1 + g_2 \mathbf{l}_2 + g_3 \mathbf{l}_3 \quad (3)$$

すなわち、 $\mathbf{p}$ が3つのスピーカーによる合成ベクトルの方向と一致するよう、ゲイン g_1, g_2, g_3 を、

$$\mathbf{g} = \begin{bmatrix} g_1 & g_2 & g_3 \end{bmatrix} = \mathbf{p}^T \mathbf{L}_{123}^{-1} \quad (4)$$

として求める。ここで、

$$\mathbf{L}_{123} = \begin{pmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \quad (5)$$

である。 $\mathbf{L}_{123}$ が逆行列を持つには、 $\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3$ が線形独立である必要があるが、これは、3つのスピーカーの位置が聴取位置から見て同一直線上にないことと等価である。なお、実際の振幅パンニング係数として用いる場合は、例えば

$$\mathbf{g}^{\text{scaled}} \equiv \frac{\mathbf{g}}{\sqrt{g_1^2 + g_2^2 + g_3^2}} \quad (6)$$

として、振幅を正規化して用いることが必要である。

Ando⁷⁾ は、特に22.2ch音響への応用を目的として、音響インテンシティが実音源と仮想音源との間で一致する条件をもとにした振幅パンニング式を提案している。音響インテンシティを規定することで、VBAFの場合に必要な振幅の正規化を行わなくても、そのままパンニング係数として用いることができる利点を有する。

4.3 22.2 ch音響

22.2ch音響は、8Kスーパーハイビジョンの音響方式として開発されたものである。HDTVの16倍の画素数をもつ $7,680 \times 4,320$ の8K映像と22.2ch音響による8Kスーパーハ

図4 22.2マルチチャンネル音響

イビジョンは、まさにその場にいるような高臨場感を実現できる。日本では2016年に試験放送、2020年に本放送を開始することが国のロードマップに示されており、NHKはそれに向けて研究開発および設備整備を進めている。22.2ch音響は、ITU-R勧告BS.2051において高さ方向に3層のチャンネルを有する方式の一つとして規定されている。

22.2ch音響は、ITU-R勧告BS.1909で規定される要求条件である、

- (1) 聴取位置を取り囲む全方向からの音の到来が再現可能なこと
- (2) 高品質な3次元音響空間印象（包み込まれ感）を再現できること
- (3) スクリーン上任意の位置での映像と音像の方向が一致すること
- (4) 広い視聴範囲をもつこと
- (5) 既存のマルチチャンネル音響方式との後方互換性を有すること
- (6) 多様な再生システムに合わせて高品質な3次元音響を再現できること
- (7) ライブ收音およびライブ伝送ができること

を満たす方式として開発された（図4）。

22.2ch音響のスピーカー配置の根拠は、音による包み込まれ感が得られるスピーカー間隔の上限が、水平面内では約 45° である実験結果に基づいている。この実験は、水平面内に等間隔にスピーカーを設置し、その各スピーカーから無相関な白色雑音を再生した際、これ以上スピーカー数を増やしても差が検知できないスピーカー数が8であることによる⁸⁾。さらに同様の実験を上半球面内の垂直面内で行った結果、これ以上スピーカーを増やしても差が検知できない条件はスピーカーを 45° から 60° 間隔で設置した場合であった⁹⁾。これより、垂直方向にも 45° 間隔でチャンネルを配置した。以上の条件より、中層に8ch、上層に中層のスピーカーの真上に

8chに加え、仰角90°に対応する聴取位置中心の真上に1chの計17chを配置した。

さらに、スクリーンの映像の方向に音が定位する条件を満たすために、特に重要な中層には、前方右、前方左と前方中央の間にそれぞれ1chの計2chを追加、さらに、スクリーン下半分の音像定位を向上させるために、前方下層に3chを追加し、合計5chを追加した。

上記の音の包み込まれ感に基づく17chと、スクリーン方向の定位を向上させる5chに、低音効果用の2chを加えて22.2chを構成した。

22.2ch音響は、収音マイクロホンから、ミキシング卓などの制作装置、符号化・伝送装置、家庭再生方式に至るまで2016年の試験放送、2020年の本放送に向けて研究開発を進めている。2012年のロンドン五輪では、8Kの映像とともに競技会場からライブ伝送を行い、イギリス国内、アメリカ、日本でパブリックビューイングを行った。また、バイノーラル技術やトランスオーラル技術¹⁰⁾を組合せた制作機器や家庭再生技術の研究も進めている¹¹⁾。

4.4 Dolby Atmos

Dolby Atmosは、映画の音響方式として、チャンネルベース方式を拡張し、オブジェクトベース方式も取り入れた方式である。Dolby Atmosは、Bedsと呼ばれるチャンネルベースの7.1chなどの方向が固定したチャンネルと、Objectと呼ばれる複数の音源と、それらのレベルや方向等を記述するメタデータからなる。Objectsは、ダイアログ、効果音などを、Bedsは主としてダイアログ、音楽、効果音のうち背景音などに属するものなどを担当する。Beds、Objectsのいずれも、方向、レベルなどを記述したメタデータを有し、BedsやObjectsの音信号とともに記録される。メタデータを有したこれらの音信号は各映画館に配信され、映画館では専用のプロセッサによってスピーカ用の信号を再構築する。Objectsを用いることにより、各映画館ごとに異なるスピーカ配置に適応した再構築が可能である。

Dolby Atmosでは、デジタルオーディオワークステーション用のプラグインソフト、映画館用のプロセッサをは

じめ、制作から配信までの機器がすでに開発されており、実際に映画館に導入され始めている。

5. むすび

マルチチャンネル音響は、従来の2chステレオの原理を拡張した技術である。既存の技術との親和性も高く、22.2ch音響や、チャンネルベース方式とオブジェクトベース方式を組合せたDolby Atmosなど、実用化に向けた研究開発が最も進展している高臨場感音響技術として、早期の普及が期待される。

(2014年6月19日受付)

〔文 献〕

- 1) 森本政之ほか、空間音響、鹿島出版会 (1990)
- 2) B. Bernferd: "Attempts for better understanding of the directional stereophonic listening mechanism", 44th AES Convention (1973)
- 3) ITU-R BS.775-2: "Multi-channel stereophonic sound system with or without accompanying picture", ITU (2006)
- 4) ITU-R BS.2051-0: "Advanced sound system for programme production", ITU (2014)
- 5) Dolby® Atmos® Next-Generation Audio for Cinema, <http://www.dolby.com/uploadedFiles/Assets/US/Doc/Professional/Dolby-Atmos-Next-Generation-Audio-for-Cinema.pdf>
- 6) V. Pulkki: "Virtual sound source positioning using vector base amplitude panning", J. Audio Eng. Soc., 45, 6, pp.456-466 (1997)
- 7) A. Ando: "Conversion of Multichannel Sound Signal Maintaining Physical Properties of Sound in Reproduced Sound Field", IEEE Trans. Audio, Speech, & Lang. Processing, 19, 5 (2011)
- 8) K. Hiyama, et al: "The minimum number of loudspeakers and its arrangement for reproducing", 113th AES Convention, Preprint 5674 (2002)
- 9) 大出ほか: "垂直方向に配置したスピーカによる空間的な音のつながりに関する検討", 音響学会講演集, pp.601-602 (2011)
- 10) 猿渡: "逆フィルタによる音場再現技術", 映情学誌, 68, 8, pp. 612-615 (2014)
- 11) K. Matsui: "Binaural Reproduction of 22.2 Multichannel Sound over Frontal Loudspeakers", 5th International Conference on 3D Systems and Applications, S10-1 (2013)

小野 一穂 1991年、東京大学工学部計数工学(修士)卒業。同年、NHKに入局、放送技術研究所に勤務。スピーカアレイによる音場制御の研究、音響トランスデューサ、スーパーハイビジョン音響の研究に従事。2013年より、(一財)NHKエンジニアリングシステム先端開発研究部チーフエンジニア。正会員。