

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4347422号
(P4347422)

(45) 発行日 平成21年10月21日(2009.10.21)

(24) 登録日 平成21年7月24日(2009.7.24)

(51) Int.Cl.

H04S 3/00 (2006.01)

F 1

H04S 3/00

Z

請求項の数 8 (全 16 頁)

(21) 出願番号 特願平11-503930
 (86) (22) 出願日 平成10年6月1日(1998.6.1)
 (65) 公表番号 特表2002-505058 (P2002-505058A)
 (43) 公表日 平成14年2月12日(2002.2.12)
 (86) 國際出願番号 PCT/GB1998/001594
 (87) 國際公開番号 WO1998/058523
 (87) 國際公開日 平成10年12月23日(1998.12.23)
 審査請求日 平成17年5月26日(2005.5.26)
 (31) 優先権主張番号 97304218.7
 (32) 優先日 平成9年6月17日(1997.6.17)
 (33) 優先権主張国 歐州特許庁(EP)

(73) 特許権者
 ブリティッシュ・テレコミュニケーションズ・パブリック・リミテッド・カンパニー
 イギリス国、イーシー1エー・7エージェイ、ロンドン、ニューゲート・ストリート
 81
 (74) 代理人 弁理士 鈴江 武彦
 (74) 代理人 弁理士 村松 貞男
 (74) 代理人 弁理士 橋本 良郎
 (74) 代理人 弁理士 白根 俊郎

最終頁に続く

(54) 【発明の名称】空間形成されたオーディオの再生

(57) 【特許請求の範囲】

【請求項 1】

リスニング空間を定めるスピーカアレイから音場を生成する方法であって、スピーカの出力が結合して、バーチャルの音源空間の知覚を与え；該方法は、アレイ内の各スピーカについて、各スピーカの出力を制御するために各出力成分 P_n を生成することを含み、該出力は入力信号内で運ばれるデータから得られ、該データは和参照信号 W 、およびバーチャルの音源によって生成されるときは異なる方向をもつサウンド成分を表わす方向性サウンド成分 X 、 Y 、(Z) を含み；該方法は、各スピーカについて、各成分 P_n が同位相または逆位相で和参照信号 W に変化するか否かを知覚する段階と、前記信号が逆位相であるときはそれを修正する段階と、生成された修正成分を各スピーカへ供給する段階とを含む方法。

10

【請求項 2】

各方向性のサウンド成分が、各方向性サウンド成分の関数であるワーピング要素によって乗算されて、移動するバーチャルの音源が、リスニングフィールド内の任意の場所において聞き手によって知覚されるような滑らかな軌跡に続き、さらにリスニングフィールド内の任意の他の場所において知覚されるような滑らかな軌跡に続く請求項 1 記載の方法。

【請求項 3】

ワーピング要素が、方向性成分の 2 乗以上の偶数幕である請求項 2 記載の方法。

【請求項 4】

ワーピング要素が、方向性成分のシヌソイド関数である請求項 2 記載の方法。

20

【請求項 5】

音場を生成する装置であって：リスニング空間を定め、スピーカの出力を結合して、バーチャルの音源空間の知覚を与えるスピーカのアレイと；入力信号内を搬送されるデータを送信および処理する手段と；和参照信号Wとバーチャルの音源によって生成されるように異なる方向においてサウンドを示す方向情報要素X, Y, (Z)とを含むデータと；アレイ内の各スピーカの出力を制御する各出力成分P_nの前記データからの生成手段と；各スピーカに対して各成分P_nが和参照信号Wに対して同位相または逆位相で変化することを知覚する手段と；前記信号が逆位相のとき、それを修正する手段と、各スピーカに対して生成された修正成分を供給する手段とを含む装置。

【請求項 6】

10

各方向性成分の関数であるワーピング要素によって各方向性要素を乗算する手段をさらに含み、移動しているバーチャルの音源がリスニングフィールド内の任意の場所において聞き手によって知覚される平滑な軌跡に続き、リスニングフィールド内の他の場所において知覚される平滑な軌跡が続く請求項5記載の装置。

【請求項 7】

ワーピング要素が方向性成分の2乗以上の偶数幕である請求項6記載の装置。

【請求項 8】

ワーピング要素が方向性成分のシヌソイド関数である請求項6記載の装置。

【発明の詳細な説明】

本発明は、音響条件が理想的でない浸漬環境 (immersive environment) における空間形成されたオーディオ (spatialised audio) の再生に関する。浸漬環境は将来の通信システムの重要な要素になることが期待されている。浸漬環境とは、ユーザがシステムによって描写された環境を、テレビジョンのような通常の平面スクリーンで外部から観察するのではなく、環境内に置かれているような感覚をユーザに与える環境である。この“浸漬 (immersion)”はユーザが主題となっている素材 (subject material) とより完全に関係することができるようになる。視覚については、ユーザの全視野が3次元の印象を与える視覚表現で占領されるようにし、ユーザが複雑な幾何学的形状を知覚できるようにすることによって、浸漬環境を生成できる。

20

浸漬効果を現実的にするために、ユーザは浸漬効果に貢献する全ての感覚に対して適切な入力を受領しなければならない。とくに、オーディオとビデオとを組合せて使用することは、大抵の浸漬環境の重要な特徴事項である：例を参照されたい：

30

ANDERSON.D. & CASEY.M. “Virtual worlds-The sound dimension” IEEE Spectrum 1997, Vol.34, No 3, pp46-50:

BRAHAM.R. & COMERFORD.R. “Sharing virtual worlds” IEEE Spectrum 1997, Vol.34, No 3, pp18-20:

WATERS.R & BARRUS.J “The rise of shared virtual environments” IEEE Spectrum 1997, Vol.34, No 3, pp20-25.

空間形成されたオーディオは周知のように、2以上のスピーカを使用すると、聞き手がスピーカから離れた源から放出されていると知覚するオーディオ効果を生成する。その最も簡単な形態では、立体音響効果が数十年間にわたってオーディオシステムに使用されてきた。本明細書では“バーチャル”的音源という用語を使用して、聞き手によって知覚されるような、実際の音源、すなわちスピーカから離れている現実の音源を意味する。

40

浸漬環境は、臨場感 (Telepresence)、テレビ会議、“フライングスルー (flying through)”建築家の設計図、教育、および医療で使用するために研究されている。幅広い視界を空間のオーディオと組合せて、“そこにいる”ような感じを作り、通信プロセスを助け、さらに大きさと奥行きの感覚を付加し、効果的な共同作業の設計空間を用意することができる。

浸漬環境の幾つかの例は、D.M.Traill、J.M.Bowskill、およびP.J.Lawrenceによって、文献 (“Interactive Collaborative Media Environments” (British Telecommunications Technology Journal Vol 15 No 4 (October 1997)), pp.130ないし139) に記載されてい

50

る。浸漬環境の1つの例はBT/AEC VisionDome（前出文献の135ないし136頁および図7に記載されている）であり、視覚イメージはユーザがイメージ内にいる状態で大型凹形スクリーン上で表される（図1および2参照）。8つのスピーカを有する多チャンネルの空間形成されたオーディオシステムを使用して、オーディオ浸漬を与える。さらに詳しくは次を参照されたい：

http://www.labs.bt.com/people/walkergr/IBTE_VisionDome/index.htm

第2の例は、同文献の134ないし135頁（および図6）に記載されている“SmartSpace”的椅子であり、これは広角ビデオスクリーン、コンピュータ端末、および空間形成されたオーディオを組合せ、全てが回転椅子の回転で動作するようにされており、このシステムはブリティッシュテレコミュニケーションズ（British Telecommunications plc）によって現在開発中である。椅子の回転によって環境内のユーザの方向が変わり、それにしたがってビジュアルおよびオーディオの入力が修正される。SmartSpaceの座長は聴覚を超越した（transaural）処理を使用し、文献（“Prospects for transaural recording”, Journal of the Audio Engineering Society 1989, Vol. 37, No1/2, pp3-19）に記載されているように、ユーザの周りに“サウンドバブル（sound bubble）”を準備して、ユーザが完全にオーディオに浸漬した感じを与えながらラップアラウンドスクリーンは視覚的な浸漬を与える。

浸漬環境が対話式であるときには、イメージおよび空間形成されたサウンドは実時間で（一般的にコンピュータアニメーションとして）生成され、その一方では対話式でない材料はしばしばアンビソニックB-Formatサウンドトラックで供給される。なおアンビソニックB-Formatサウンドトラックの特徴は本明細書において後で記載することにする。アンビソニックコーディングは浸漬オーディオ環境については一般的に普及している選択肢である。その理由は3または4のみの伝送チャンネルを使用して、チャンネルを幾つでもデコードできるからである。しかしながら、アンビソニック技術は、後に記載するように臨場感（telepresence environment）のある環境で使用されるときに制限される。

ここで浸漬環境におけるサウンドの局所化に関する幾つかの問題を検討することにする。図1および2は、VisionDome（ビジョンドーム）の平面図および側断面図であり、8つのスピーカ（1, 2, 3, 4, 5, 6, 7, 8）と、ラップアラウンドスクリーンとをもち、通常のユーザ位置が記されている。多チャンネルアンビソニックオーディオトラックは通常長方形のリスニングルーム内で再生される。半球形ドーム内で再生されるときは、リスニング環境の幾何学的形状によって空間形成の品質が悪くなる。半球体内での反響はサウンドフィールド（音場）の再結合を壊すことである：これは壁表面を適切な吸収材料で処理することによって最小にできるときもあるが、常に実用的である訳ではない。リスニングルームとして固いプラスチックのドームを使用すると、複数の反響によって主として生じる多くの音響問題を生み出す。ドームの音響特性は、処理されないまま残されるときは、サウンドが多数の源から発生し、意図されたサウンドの空間形成効果が壊れるようと考えられる。1つの解決案ではドームの内側表面を、反響を低減する吸収材料で被覆する。ビデオスクリーン材料、それ自体はサウンド吸収性であるのでサウンドの反響を低減するのを助けるが、ビデオスクリーン材料はさらにスクリーンの背後に置かれたスピーカから発生するサウンドに対して高周波数を相当に減衰させる。この高周波数の減衰は、スクリーンの背後に置かれたスピーカ1,2,3,7,8へ供給される信号を等化することによって克服される。

プラスチックのドーム以外のリスニング環境は特有の音響特性を有し、大抵の場合に反響はエラーの原因となる。ドームに音響タイルを貼付して反射量を低減することによって、ユーザは一層オーディオ信号を適切に空間形成できるようになる。

大抵の投影スクリーンおよびビデオモニタは平面（またはほぼ平面）のスクリーンである。予め録音されたB-Format（Bフォーマット）サウンドトラックは動画（moving video）イメージに一致するように構成されているときは、一般的にこのような平面ビデオスクリーンを使用してスタジオ内で構成される。正確な空間知覚（知覚した音場）を与えるために、B-Formatコーディングを使用して、オーディオを平面ビデオスクリーンにマップする

10

20

30

40

50

。しかしながらVisionDomeのような多数のユーザを含む大きな環境を使用するときは、凹形スクリーン上でビデオを再生し、観察者に正確に見えるようにビデオイメージを適切に修正する。しかしながら幾何学的形状によるオーディオへの効果は最早ビデオと一致せず、知覚上の同期を回復するのに非線形のマッピングを必要とする。対話式材料の場合、B-Formatコーダは単位円の環境上にバーチャル源を置き、スクリーンの曲率をマッピングする。

聞き手のグループが小さい領域内に位置している環境では、アンビソニック再生システムは聞き手のほとんどに対して所望の聞くことに関する空間形成を生成するのに失敗しそうである。その1つの理由は、スピーカのみによって生成される種々の音場が正確に結合して、“スイートスポット”として知られるように、1つの位置における“バーチャル”的音源の所望の効果を生成できるからである。精密なスイートスポットには（多くとも）たった1人の聞き手が置ける。これは、同位相と逆位相との信号が正確に再構成して所望の信号を与える場合の真のスイートスポットが小さい領域であり、スイートスポットの外側の関係者は同位相および逆位相の信号の不正確な組合せを受取るからである。事実、半球形のスクリーンにおいて通常ビデオプロジェクタは半球体の幾何学的形状の中心にあり、アンビソニックは一般に、“スイートスポット”もスピーカアレイの幾何学的中心にあって、スクリーンと同心になるようにされている。したがって実際の“スイートスポット”はプロジェクタによって占められているので、その位置には誰も存在できない。

スイートスポットを移動して、聞き手の位置と一致させる効果は、BURRASTON, HOLLIER & HAWKSFORDによって研究してきた（文献“Limitations of dynamically controlling the listening position in a 3-D ambisonic environment” Preprint from 102nd AES Convention March 1997 Audio Engineering Society (Preprint No 4460) 参照）。これにより元のスイートスポット内に位置していなかった聞き手がアンビソニックデコード信号の正確な組合せを受領できるようになった。しかしながらこのシステムは単一ユーザ用に設計され、スイートスポットは一度に1つのみの位置に移動できる。この文献では、聞き手をスイートスポットの外側に位置付けた場合の効果（バーチャルの会合場所に聞き手のグループがいるときに発生するような効果）について記載し、非常に多くの正式なリスニング試験に基づいて、聞き手がスイートスポットに位置付けられているときのみ、聞き手はサウンドを正確に局所化できると結論している。

音源が移動し、聞き手がスイートスポット以外の位置にいるとき、興味深い効果が示される。サウンドが前方右から前方左へ移動し、聞き手が中心から外れて、前方に近付いている例について検討する。サウンドは最初に右側スピーカから放出され、しばらくそこにとどまり、次に迅速に中心を横切って、左側スピーカへ移動すると思われる。サウンドはスピーカの周りに“引掛けり(hang)”、音響的には中空の中心領域、すなわち“ホール(hole)”を生成する。スイートスポットに位置していない聞き手にとっては、どんなバーチャルの音源も概ねスピーカの1つに近付きすぎると思われる。（スイートスポットにいる聞き手によって知覚されるところによると）空間をサウンドが滑らかに移動するとすると、スイートスポットにいないユーザは1つのスピーカの位置の近くに留まっているバーチャルの源と知覚し、次に別のスピーカに突然に飛び移ることになる。

幾何学的形に協同する補正の最も簡単な方法は、スピーカの幾何学的位置をアンビソニックデコーダ内にプログラムするときにスピーカの位置を歪ませる(wrap)ことを含んでいる。デコーダは、スピーカを真の位置よりも中心に近付けた位置でプログラムされる：この結果サウンドはスクリーンの縁部で迅速に、スクリーンの中心近くでゆっくりと移動し-この結果スクリーン上のイメージに関してサウンドの知覚される線形移動を生ずる。この原理はアンビソニックデコーダのみに応用でき、アンビソニックデコーダは選択可能なスピーカの位置へ向うB-Format信号をデコードできる。すなわちこの原理は（立方体の8つの角部かまたは正方形の4つの角部のような）固定のスピーカ位置について設計されたデコーダに使用できない。

非線形のパニング(panning)方式が開発され、これは入力としてモノフォニック音源、所望のサウンド位置(x, y, z)および再生システムにおけるNスピーカの位置(x,

10

20

30

40

50

y, z) をもつ。このシステムは 1 または複数の別々の入力源を有することができ、該入力源は空間内の別々の場所で個々に局所化できる。バーチャルの音源は非線形パニングの特徴を使用して 1 つの位置から別の位置へパンされる。非線形パニングは、オーディオの“ホール”が知覚されるという上述の効果を補正する。知覚の経験が補正されて、最初の位置から最後の位置への線形オーディオの軌跡を与える。非線形パニング機構は、強度パニングに基づき、アンビソニックシステムの中でのように波面再構成には基づいていない。ワーピング (warping) は強度パニングに基づくので、他のスピーカの逆位相信号は存在せず、それ故にマルチユーザシステムでは聞き手の全ては正確に空間形成されたオーディオを経験することになる。非線形ワーピングアルゴリズムは完全なシステムであり (すなわち、信号座標を採用して、3 次元空間内に信号を位置付ける) 、したがって実時間素材のみに使用でき、アンビソニック記録のワーピングには使用できない。

本発明にしたがって、リスニング空間を定めるスピーカアレイから音場を生成する方法であって、スピーカの出力が結合して、バーチャルの音源空間の知覚を与える；該方法は、アレイ内の各スピーカについて、各スピーカの出力を制御するために各出力成分 P_n を生成することを含み、該出力は入力信号内で運ばれるデータから得られ、該データは和参照信号 W 、およびバーチャルの音源によって生成されるときは異なる方向をもつサウンド成分を表わす方向性サウンド成分 $X, Y, (Z)$ を含み；該方法は、各スピーカに対して、各成分 P_n が同位相または逆位相で和参照信号 W に変化するか否かを知覚する段階と、前記信号が逆位相であるときはそれを修正する段階と、生成された修正成分を各スピーカへ供給する段階とを含む方法を提供する。

本発明の第2の態様にしたがって、音場を生成する装置であって：リスニング空間を定め、スピーカの出力を結合して、バーチャルの音源空間の知覚を与えるスピーカのアレイと；入力信号内を搬送されるデータを送信および処理する手段と；和参照信号 W とバーチャルの音源によって生成されるように異なる方向においてサウンドを示す方向情報要素 $X, Y, (Z)$ とを含むデータと；アレイ内の各スピーカの出力を制御する各出力成分 P_n の前記データからの生成手段と；各スピーカに対して各成分 P_n が和参照信号 W に対して同位相または逆位相で変化することを知覚する手段と；前記信号が逆位相のとき、それを修正する手段と；各スピーカに対して生成された修正成分を供給する手段とを含む装置を提供する。

好みいのは各方向性サウンド成分は、各方向性サウンド成分の関数であるワーピング要素によって乗算されて、移動しているバーチャルの音源がリスニングフィールド内の任意の場所において聞き手によって知覚される平滑な軌跡に続き、さらにリスニングフィールド内の他の場所において知覚される平滑な軌跡に続く。したがってバーチャルの音源は、リスニングフィールドの中心領域よりも他の領域に現れる傾向がある。ワーピング要素は方向性サウンド成分の 2 乗以上の偶数幕か、またはシヌソイド関数であってもよい。

ここで 2 次元再生システムに関するアンビソニック B-Format コーディングおよびデコーディングの方程式を簡単に記載する。この段落ではアンビソニックの詳細な理論を記載しないが、フィールドにおける他の研究者の成果を記載する。アンビソニック理論は方向性情報をオーディオ信号にコード化するという問題に対する解決案を表している。信号は、少なくとも 4 つのスピーカ (パントホニック (pantophonic) - 水平面 - システム) または 8 つのスピーカ (無方向性 (ペリホニック) - 水平および垂直面 - システム) のアレイ上で再生されることを意図されている。信号は “B-Format” と呼ばれ、パントホニックシステム (W, X, Y) の 3 つの成分および無方向性システム (W, X, Y, Z) の 4 つの成分から成る。サラウンド - サウンドおよびアンビソニック理論の詳細な解析に関しては次の文献を参照されたい：

BAMFORD.J.& VANDERKOYJ. “Ambisonic sound for us” Preprint from 99th AES Convention October 1995 Audio Engineering Society (Preprint No4138)

BEGAULT.D. “Challenges to the successful implementation of 3-D sound” Journal of the Audio Engineering Society 1991, Vol.39, No 11, pp864-870

BURRASTON 他 (上述参照)

10

20

30

40

50

GERZON.M. "Optimum reproduction matrices for multi-speaker stereo" Journal of the Audio Engineering Society 1992, Vol.40, No 7/8, pp571-589

GERZON.M. "Surround sound psychoacoustics" Wireless World December 1974, Vol.80, p 483-485

MALHAM.D.G "Computer control of ambisonic soundfields" Preprint from 82nd AES Convention March 1987 Audio Engineering Society (Preprint No2463)

MALHAM.D.G. & CLARKE.J. "Control software for a programmable soundfield controller" Proceedings of the Institute of Acoustics Autumn Conference on Reproduced Sound 8, Windermere 1992, pp265-272

MALHAM.D.G. & MYATT.A. "3-D Sound spatialisation using ambisonic techniques" Computer Music Journal 1995, Vol.19 No4, pp58-70

POLETTI.M. "The design of encoding functions for stereophonic and polyphonic sound systems" Journal of the Audio Engineering Society 1996, Vol.44, No 11, pp948-963

VANDERKOY.J. & LIPSHITZ.S. "Anomalies of waveform reconstruction in stereo and surround-sound reproduction" Preprint from 83rd AES Convention October 1987 Audio Engineering Society (Preprint No 2554)

本明細書に記載したアンビソニックシステムは全て第1次、すなわち $m = 1$ であり、チャンネル数は2次元のシステムでは $2m + 1$ (3チャンネル: w, x, y) によって、3次元システムでは $(m + 1)^2$ に (4チャンネル: w, x, y, z) よって得られる。本明細書では、2次元システムのみを検討するが、本明細書に記載した発想は完全な3次元再生システムで使用するために直ぐにスケール合わせをすることができ、請求項の技術的範囲はこのような3次元再生システムを含む。

2次元システムを使用すると、エンコードされた空間形成されたサウンドは1つの平面、すなわち (x, y) 面内のみにある。音源が単位円内に $x^2 + y^2 = 1$ (図3参照) 位置付けられていれば仮定する。単位円上に位置付けられたモノフォニック信号については：

$$x = \cos(\theta)$$

$$y = \sin(\theta)$$

ここで、 θ は図3に規定したように音源の元の位置と所望の位置との間の角度である。

B-Format信号は3つの信号W, X, Yを含み、これらは次のように定められている(上述のMalhamおよびMyatt参照)

$$W = S \cdot 1 / \sqrt{2}$$

$$X = S \cdot \cos(\theta)$$

$$Y = S \cdot \sin(\theta)$$

ここで、Sは空間形成されたモノフォニック信号である。

バーチャルの音源が単位円上にある；すなわち $x = \cos(\theta)$ および $y = \sin(\theta)$ であるとき、xおよびyに関するW, X, Yに対する方程式を与える：

$$W = 1 / \sqrt{2} \cdot S \quad \text{周囲の信号}$$

$$X = x \cdot S \quad \text{前 - 後信号}$$

$$Y = y \cdot S \quad \text{左 - 右信号}$$

同じくMalhamおよびMyattによって記載されているように、デコーダは次のように動作する。Nスピーカの標準のアレイでは、パントホニックシステムデコーディング方程式は次の通りである：

$$P_n = (1/N) \cdot W + 2X \cos(\theta_n) + 2Y \sin(\theta_n)$$

なお、 θ_n はスピーカ "n" の方向であり(図4参照)、したがって図4に示したような標準の4つのスピーカアレイにおいて、各スピーカに供給される信号は次の通りである：

$$P_1 = \frac{1}{4} W + \frac{2X}{\sqrt{2}} + \frac{2Y}{\sqrt{2}} \quad P_2 = \frac{1}{4} W - \frac{2X}{\sqrt{2}} + \frac{2Y}{\sqrt{2}}$$

$$P_4 = \frac{1}{4} W + \frac{2X}{\sqrt{2}} - \frac{2Y}{\sqrt{2}} \quad P_3 = \frac{1}{4} W - \frac{2X}{\sqrt{2}} - \frac{2Y}{\sqrt{2}}$$

本発明の方法を使用して、B-Formatアンビソニック信号（または後で記載するワープされたB'-Format信号）をとり、逆位相の成分を低減して、非線形のパンニングタイプ信号を生成し、ユーザのグループが空間形成されたサウンドを経験できるようにする。真の波面再構成は最早達成されないので、再生成は最早アンビソニクシステムではない。デコーダ
ワーピングアルゴリズムはアンビソニックデコーダから出力をとり、出力が各再生成チャ
ンネルに供給される前に出力をワープし、したがってN出力チャネルのそれぞれに対し
てデコーダワーパ（warper）の1つの構成がある。B-FormatまたはB'-Formatデコーダ出
力の何れかからの信号の位相成分がずれているとき、その位相はW入力信号に対して逆に
され-したがってデコーダ出力をWと比較することによって、信号の位相がずれているか
否かを判断することができる。所定のデコーダ出力の位相がずれているとき、出力は減衰
要素Dによって減衰される：

$$P'_n = P_n \cdot D$$

ここで、 $\text{sign}(P_n) \text{ sign}(W)$ の場合に $0 < D < 1$ 、さもなければ $D = 1$

この簡単なアルゴリズムは、聞き手がスイートスポットから離れているとき、最も近いスピーカに対してサウンドの局所化が壊れる見込みを低減する。

B-FormatワーピングはアンビソニックのB-Format記録をとり、知覚した非線形の軌跡を補正する。システムへの入力はB-Format記録であり、出力はワープされたB-Format記録である（ここではB'-Format記録と記載される）。B'-Format記録は、既存のデコーダを使用できるB-Formatデコーダでデコードすることができる。アンビソニックシステムは、音場が正しく再構成する再生領域内に‘スイートスポット’を生成し、他の領域では聞き手は正しく空間形成されたサウンドを経験しないことになる。ワーピングアルゴリズムの目的は、 x & y 値の線形領域から非線形領域へ変更することである。サウンドが右から左へ移動する例を検討する；サウンドは最初に迅速に移動し、次にゆっくりと中心を横切って、最後に遠くの左側へ迅速に横切って、知覚を修正することが必要である。ワーピングはさらには、静止しているオブジェクトの知覚ビューにも影響を与えるが、その理由はワーピングなしでは、スイートスポットから離れている聞き手が最もバーチャルである音源が幾つかの領域に集中することを知覚することになるからであり、中心領域は一般的に密度が低く、知覚されたオーディオの“ホール”となっているのである。B-Format信号成分X, Y, およびWに関して、 x および y の元の値の推定値を判断することができ、したがって元の信号Sを再構成して、 $S' = W/2$ を求め、 $S' = W/2$ から推定値 x' および y' を得ることができる：

$$x' = \frac{X}{S'} \quad \text{および} \quad y' = \frac{Y}{S'}$$

したがって

$$x' = \frac{X}{W\sqrt{2}} \quad y' = \frac{Y}{W\sqrt{2}}$$

\hat{x}' および \hat{y}'

は、範囲 ($\pm 1, \pm 1$) 内の正規化された x および y の値を表す。一般的なワーピングアルゴリズムは次の式によって与えられる：

$$x' = x \cdot \int (\hat{x}')$$

$$\text{および} \quad y' = y \cdot \int (\hat{y}')$$

しかしながら、 x は X の関数であり、 y が Y の関数であるので、次のようになる。

$$x' = X \cdot \int (Y) \quad \text{および} \quad y' = Y \cdot \int (X)$$

10

20

30

40

50

生成された信号 X'、Y'、および W は B'-Format 信号と呼ばれることになる。

ここで 2 つの可能なワーピング関数を記載することにする。

1) 幕のワーピング

幕のワーピングを使用して、X の値は偶数の幕で幕乗された \hat{x}'

によって乗算され (X を奇数の幕で効果的に幕乗して そのサインを維持する)、Y は同じやり方でワーピングされる。

$$\int(x) = (\hat{x}')^{2i} \quad \text{および} \quad \int(y) = (\hat{y}')^{2i}$$

$$\int(x) = \left(\frac{x}{w\sqrt{2}}\right)^{2i} \quad \text{および} \quad \int(y) = \left(\frac{y}{w\sqrt{2}}\right)^{2i}$$

10

$$\text{すなわち, } x' = x \cdot \left(\frac{x}{w\sqrt{2}}\right)^{2i} \quad \text{および} \quad y' = y \cdot \left(\frac{y}{w\sqrt{2}}\right)^{2i}$$

これらの方程式では、 $i = 0$ を選択すると、ワーピングされていない構成が得られ、 $i > 0$ のときは、非線形ワーピングが生成される。

2) シヌソイドのワーピング

シヌソイドワーピングの異なる関数を使用して、異なる関数 $f(X)$ および $f(Y)$ を \hat{x}' および \hat{y}'

20

の範囲の異なる部分に使用する。シヌソイドワーピングはバーチャルの音源がその範囲の極値であり、中心領域に迅速に遷移するときに一定のレベルを与えるために行なわれる。幕乗された正弦波の半周期を使用して、値と中心領域との間に滑らかに補間する。

X において :

$$1. \quad -1 < \hat{x}' \leq x_1 \quad f(X) = \frac{1}{|\hat{x}'|}$$

$$2. \quad x_1 < \hat{x}' \leq x_2 \quad f(X) = \frac{1}{2 \cdot |\hat{x}'|} \left\{ \sin \left(\frac{(\hat{x}' + |x_1|) \cdot \pi}{|x_2 - x_1|} + \frac{\pi}{2} \right) + 1 \right\}$$

30

$$3. \quad x_2 < \hat{x}' \leq x_3 \quad f(X) = 0$$

$$4. \quad x_3 < \hat{x}' \leq x_4 \quad f(X) = \frac{1}{2 \cdot |\hat{x}'|} \left\{ \sin \left(\frac{(\hat{x}' - |x_3|) \cdot \pi}{|x_4 - x_3|} + \frac{\pi}{2} \right) - 1 \right\}$$

$$5. \quad x_4 < \hat{x}' \leq +1 \quad f(X) = \frac{1}{|\hat{x}'|}$$

Y において

40

$$\begin{aligned}
 1. \quad -1 < \hat{y}' \leq y_1 \quad f(Y) &= \frac{1}{|\hat{y}'|} \\
 2. \quad y_1 < \hat{y}' \leq y_2 \quad f(Y) &= \frac{1}{2|\hat{y}'|} \left\{ \sin \left(\frac{(\hat{y}' + |y_1|) \cdot \pi}{|y_2 - y_1|} + \frac{\pi}{2} \right) + 1 \right\} \\
 3. \quad y_2 < \hat{y}' \leq y_3 \quad f(Y) &= 0 \\
 4. \quad y_3 < \hat{y}' \leq y_4 \quad f(Y) &= \frac{1}{2 \cdot |\hat{y}'|} \left\{ \sin \left(\frac{(\hat{y}' - |y_3|) \cdot \pi}{|y_4 - y_3|} + \frac{\pi}{2} \right) - 1 \right\} \quad 10 \\
 5. \quad y_4 < \hat{y}' \leq +1 \quad f(Y) &= \frac{1}{|\hat{y}'|}
 \end{aligned}$$

定数 $x_1 \dots 4$ および $y_1 \dots 4$ に対する通常の値は次の通りである：

$$x_1 = y_1 = -0.75; \quad x_2 = y_2 = -0.25; \quad x_3 = y_3 = 0.75; \quad x_4 = y_4 = 0.75$$

ワーピングアルゴリズムへの入力としてB-Format信号を使用することは、他の方法よりも多くの長所を有する。バーチャルの会合環境において、ユーザの音声はB-Format信号でコード化することができ、次にこのB-Format信号はシステム内の全ての他のユーザへ送られる（他のユーザは世界のどこかにいると仮定する）。他のユーザが位置している物理的な環境は著しく変化することがあり、バイノーラ（2聴覚の）ルヘッドホン応用システムを使用してもよい（MOLLER.H. "Fundamentals of binaural technology" Applied Acoustics 1992, Vol. 36, pp171-218）。別の環境は、ワープされたアンビソニックを使用したビジョンドーム（Visiondome）内であってもよい。さらに別の環境は單一ユーザを使用する真のアンビソニックシステム、または聴覚を超越した2つのスピーカ再生システムであってもよく、これはCooperおよびBauckによって文献（上述参照）に記載されている。その内容は図5に示した。

ここで専用の装置（proprietary equipment）を使用して、本発明の2つの構成（一方はディジタル、他方はアナログ）を記載することにする。バーチャルの会合環境において、オーディオは実時間で処理する必要がある。ここでは全てのコーディングはアナログまたはDSP応用ハードウェアの何れかを使用して実時間で実行されると仮定する。

実際のバーチャルの会合位置は、数メートルまたは何千キロメートルも離れていてもよい。各関係者（participant）間のオーディオ接続は一般的に、ISDN、LANmまたはWANのような広帯域ディジタルネットワークを介して行なわれる。したがって有益なことはディジタル領域内でコーディングおよびデコーディングを行って、不必要なD/AおよびA/D変換段階をやめることである。コーディングは従来のB-Formatコーダを使用することによって実行され、デコーディングは修正を加えた（ワーピング）コーダによって実行される。この例外として非線形パニングの使用であり、こうして非線形パニングを使用するときは、その座標をもつモノフォニック信号か、またはNチャンネル信号を送ることが必要である。

こうして、非線形パニングを遠隔のバーチャルの会合場所を採用するシステムで使用するにはより不適切とする。

Lake HURON DSPエンジンはアンビソニックB-Format信号を生成およびデコーディングする専用の方法であり、1または複数の任意の間隔で位置付けられたスピーカで2次元および3次元の両方のオーディオをデコードすることができる。これは<http://www.lakedsp.com/index.htm>に記載されている。Huronは必要なツールを供給されて、注文製のDSPプログラムを生成し、また本明細書中に示したワーピングアルゴリズムの計算は比較的に簡単であるので、これをアンビソニックデコーダの構成に含んでもよい。この方法の主な長所は、ハードウェアが既に開発されており、システムが多数のI/Oチャンネルを処理できるこ

10

20

30

40

50

とである。

第2のデジタル構成方法は、主要なDSPチップ製造業者から販売されている多くのDSP開発システムの1つの上でDSPチップをプログラミングすることを含む。このようなシステムには2または3の入力チャンネルおよびより多くの出力チャンネル（普通は4または8）が必要である。この構成では容易に大量生産できる高度に専門化されたデコーダを生成する。

PCおよびサウンドカード技術が増大するのにしたがって、実時間のアンビソニックデコーディングおよびワーピングが現実的になり、複雑なDSPシステム設計に対する要件が低減する。

B-Formatワーピングおよびデコーダワーピングをアナログ領域においてアナログの乗算器を使用して交互に実行することができる。従来のアンビソニックデコーダを使用して、B'-Formatデコーディングを実行し、デコーダの出力をデコーダワーパハードウェアに供給する。このようなシステムは図6に示した。B-Formatワーパおよびデコーダワーパのブロック図は、それぞれ図7および8に示した。ブロック図はアナログ乗算器で使用可能な機能ブロックに対応し、一般的な種類のアナログ乗算器は://www.analog.com/products/index/12.htmlに記載されている。

ここで実時間で動作するのではなく、上述の方法を使用する多数のシミュレーションを記載することにする。実際的な実施形態に対して要求されるように、これらの例を生成するのに使用される処理は、適切なオーディオインターフェイスをもつPCを使用してオフラインで計算された。正規化された座標として各xおよびyがが-1ないし+1の値のみをとることができると仮定して、最初に単一の音源を(-1, -1)から(1, 1)へ移動する例を検討する。オーディオトラックの始めにバーチャルのサウンドを位置(-1, -1)に位置付け、トラックの終わりにバーチャルのサウンドを位置(1, 1)に位置付ける。サウンドをコード化して、開始位置から終了位置へ線形に移動する。説明を分かり易くするために、空間形成されたモノフォニック源信号は正のDC電圧にセットした。上述のB-Formatコーディング技術を使用することによって、3チャンネル信号を構成し、次にこれを上述のワーピングアルゴリズムでデコードした。

図9は、図4に示したスピーカの幾何学的形状をもつ4つの各スピーカの出力が、4チャンネルデコーダから、従来のアンビソニックB-Formatコーディングを使用して供給することを示す。最初にバーチャルの源を、完全な大きさの出力をもつスピーカ3の近くに位置付け、スピーカ1の出力は最初は逆位相であり、スピーカ2および4はWの値をもつことが分かる。バーチャルの源が中心領域を移動するとき、スピーカ1、2、3、および4のレベルは等しい。例示した軌跡の終わりにスピーカ1は高出力レベルをもち、スピーカ3は逆位相であり、スピーカ2および4は一定Wレベルを維持する。

図10は、B-Formatワーピング(B'-Format信号)を導入する効果を示す。スピーカは軌跡の最初と最後の場所で従来のB-Formatワーピングと類似のレベルをもつが、ここで経路は主として中央領域内にあり、したがって個々のスピーカ“の周りに引掛かる(hang around)”または“崩壊していく(collapsing to)”サウンドの知覚が除かれる。

図9および10に示したスピーカはアンビソニック信号を供給し、ここでは同位相または逆位相信号のベクトルの和によって正しい信号がスイートスポットで得られる。デコーダワーピングアルゴリズムは、よりコヒーレントな信号をスイートスポットに位置していない聞き手に向けて提供する逆位相の成分を減衰する。図11は、付加的にデコーダワーピングを行った基本的なアンビソニックB-Formatデコーディング(図9参照)を示す。逆位相成分の除去は、この例ではD=0のとき明白に分かる。図12は、デコーダワーピングを行ったB'-Formatデコーディング(図10参照)を示し、逆位相の減衰効果が分かる。上述の例では(-1, -1)から(1, 1)、すなわち後方左から前方右への軌跡を検討したが；次の例では(1, 1)から(-1, 1)、すなわち前方右から前方左への軌跡を検討する。図13、14、15、および16はそれぞれ、B-Formatデコーダ、B'-Formatデコーダ、デコーダワーピングをもつB-Formatデコーダ、およびデコーダワーピングをもつB'-Formatデコーダの効果を示している。この例では、逆位相信号は、選択したバーチャ

10

20

30

40

50

ル源軌跡によってより優れている。上述の例に関して、デコーダワーピング要素Dはゼロにセットされ、逆位相成分を全て取り除く。

グラフ表現を明らかにするために、ここでは正のDC電圧をバーチャルの源として使用した2つの例を記載する。しかしながら実際には正弦波および複雑な波面（実際のオーディオ信号）が使用される。デコーダアルゴリズムは複雑な波面で試験され、正確な動作を確かめた。

空間形成された最終的なアビータ（仲裁者）は聞き手である。オーディオサウンドの効果は前方右から前方左への軌跡をもつB-Format信号へコード化され、次に上述の同じ4つのでコーディングアルゴリズムを使用してデコードされた。非公式のリスニング試験をビジョンドームで行い、次のリスニング位置で聞き手によって続く観察を行った：

1. スイートスポットにおいて

・B-Format

移動する音源の知覚を与えるためにスピーカ信号を正確に結合した。しかしながらリスニング環境の幾何学的形状および音響特性のために、サウンドは線形の軌跡をもつリスニング空間を横切って移動することは考えられなかった。

・B'-Format

B-Formatの例のように、個々の音場を正確に再構成して、移動する音源の知覚を与えるようにした。バーチャルの音源は、非線形ワーピングの使用によって知覚された線形の軌跡をもつた。

・デコーダワーピングをもつB-Format

サウンドは非線形の軌跡を伴ってリスニング領域を横切って移動すると考えられた。この知覚はB-Formatの例の知覚と類似していた。

・デコーダワーピングをもつB'-Format

サウンドは線形の軌跡を伴ってリスニング領域を横切って移動すると考えられた。この知覚はB'-Format知覚の例と類似していた。

2. 前方左または前方右のスピーカへの接近（図4の位置1および4）

・B-Format

バーチャルの音源の位置は最も近いスピーカに対して“崩壊を与える（collapse）”し、このスピーカの寄与は聴覚のランドスケープ（aural landscape）を支配しており、軌跡の感覚はほとんどまたは全く得られない

・B'-Format

バーチャルの音源の位置は最も近いスピーカに対して“崩壊を与える”このスピーカの寄与は聴覚のランドスケープを支配しており、全体的な音場は、後方の逆位相のスピーカの供給から何も寄与しないからであり、軌跡の感覚がわずかに残る。

・デコーダワーピングをもつB-Format

移動の感覚は向上するが、知覚される軌跡は非線形である。

・デコーダワーピングをもつB'-Format

サウンドが1つの位置から別の位置への移動が、ほぼ線形の知覚された軌跡経路を伴う。

3. 前方左と後方左のスピーカ（4および3）間の中途または前方右と後方右のスピーカ（1および2）間の中途

・B-Format

2つの別個の軌跡が知覚される：（スピーカ4および1から）同位相信号が右から左へ移動し、逆位相信号が左から右へ移動する。2つの別個の軌跡は混乱の原因となり、軌跡が全くないときよりも紛らわしくなる。

・B'-Format

この信号の知覚はB-Format信号の知覚と類似しているが、B-Format信号よりも類似の度合いは低く、2つの別々のバーチャル源の軌跡の感覚は少なくなる。

・デコーダワーピングをもつB-Format

1つのみの軌跡が観察されたが、軌跡は明らかに非線形であった。

・デコーダワーピングをもつB'-Format

10

20

30

40

50

ここでは1つの軌跡が観察され、この軌跡は知覚された軌跡の中でB'-Format信号よりも線形であった。非線形の歪みの度合いがより大きいと局所化がより明白になる。

4. 後方左と後方右のスピーカ(3および2)の間

・B-Format

2つの主要なスピーカ源は後方のスピーカ(2および3)であるので、主要な音源は逆位相成分である。バーチャルの音源は、意図されていた方向と逆方向に移動すると考えられる。したがって、音源が浸漬環境内のビデオ源と結合されるときは危険であると推測される。サウンドおよび視界を反対方向に移動することは、明らかにモード相反という容認できない形式となる。

・B'-Format

10

観察された効果はB-Format信号と同じである。

・デコーダワーピングをもつB-Format

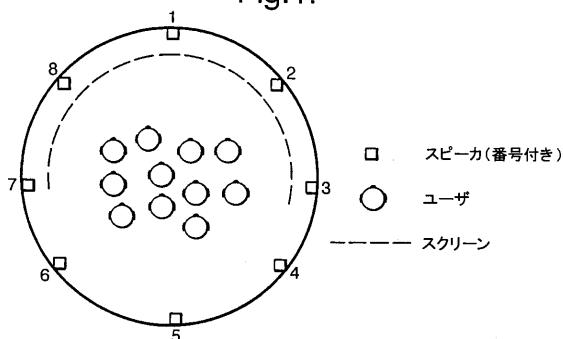
通路の軌跡は逆位相成分の除去により、非線形ではあるが、はっきりしている。

・デコーダワーピングをもつB'-Format

前方右のスピーカから前方左のスピーカへは、はっきりした線形の軌跡になる。

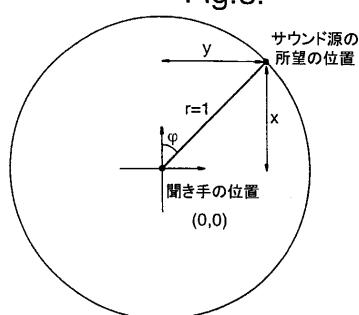
【図1】

Fig.1.



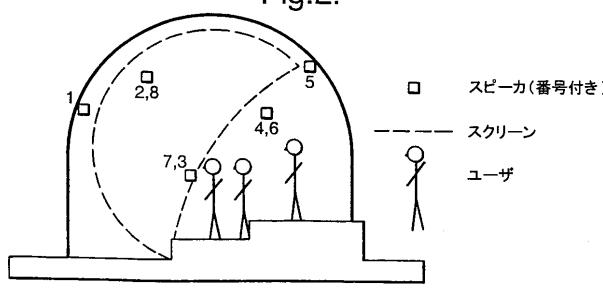
【図3】

Fig.3.



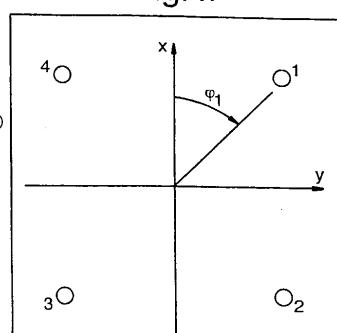
【図2】

Fig.2.

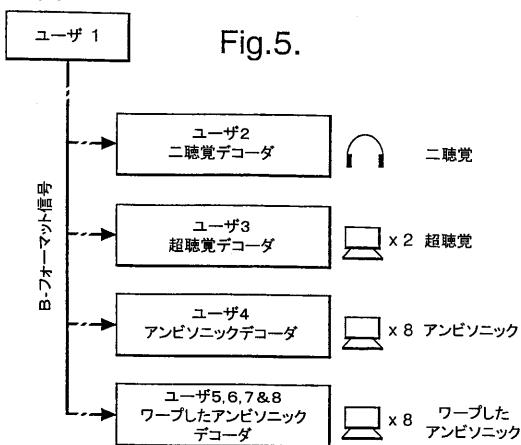


【図4】

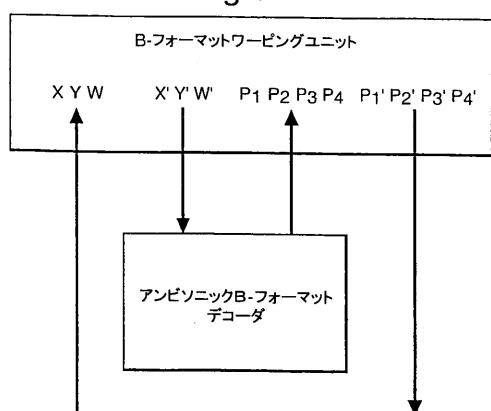
Fig.4.



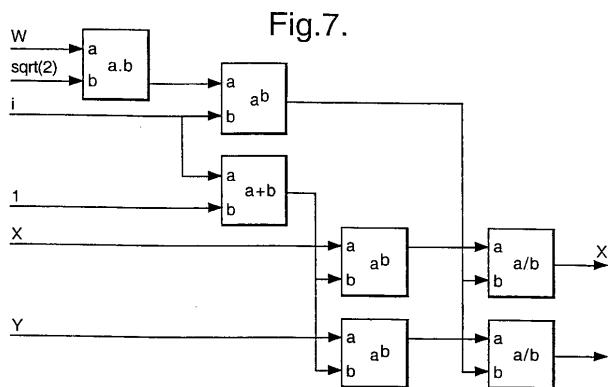
【図5】



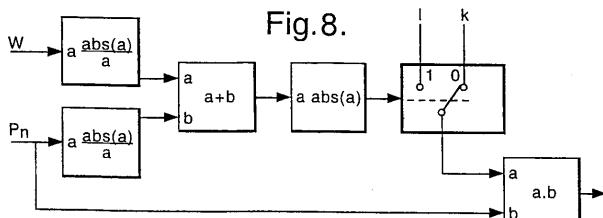
【図6】



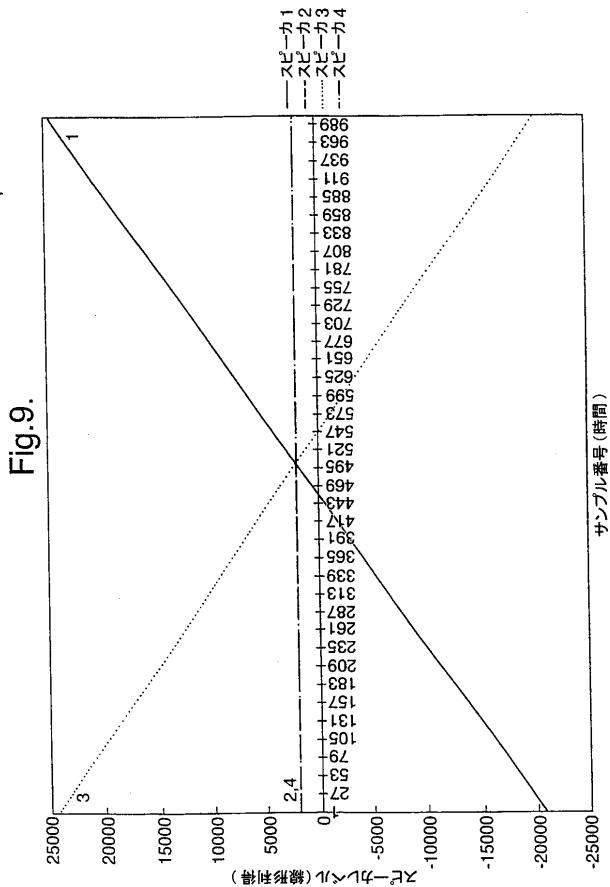
【図7】



【図8】

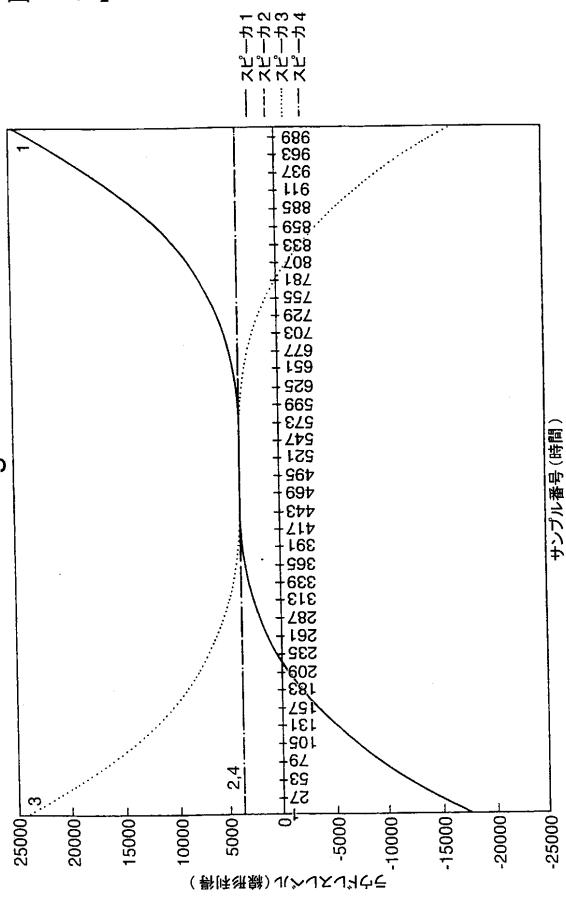


【図9】



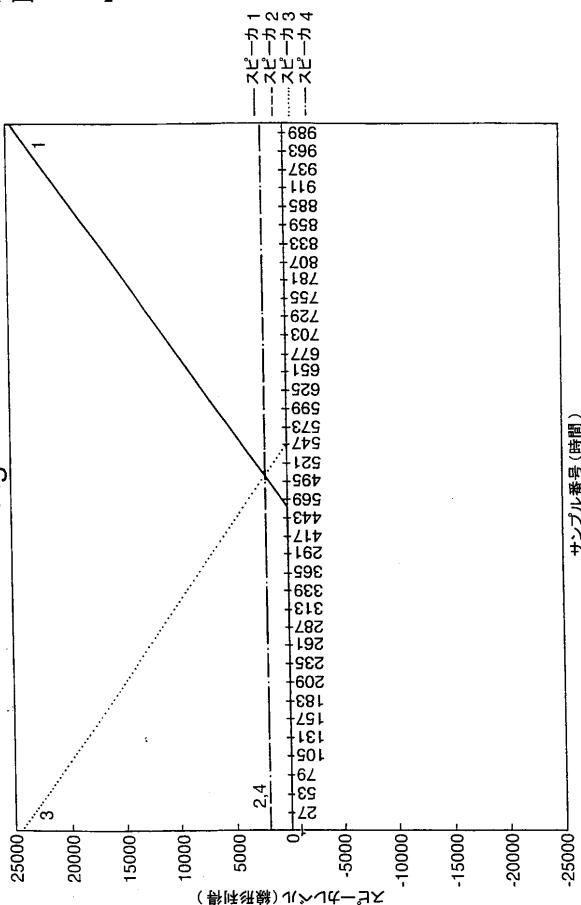
【図 10】

Fig.10.



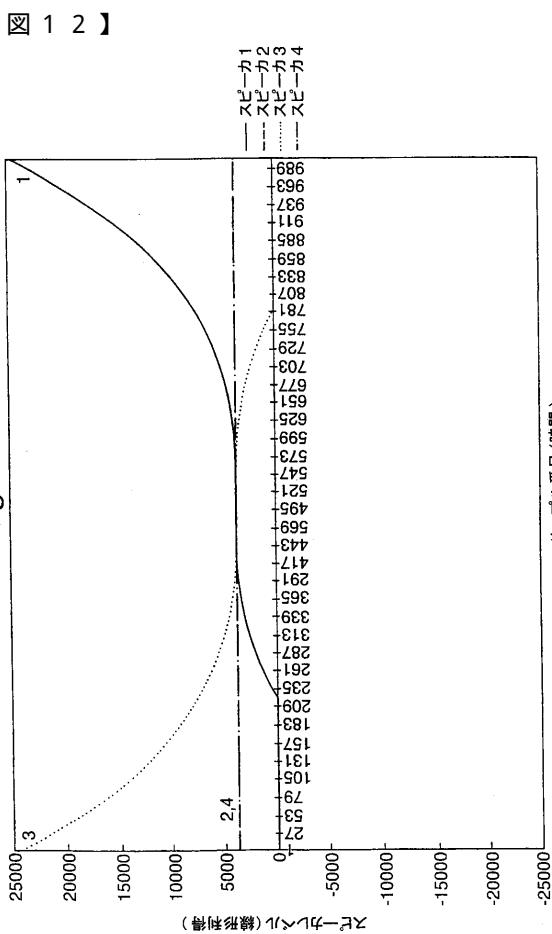
【図 11】

Fig.11.



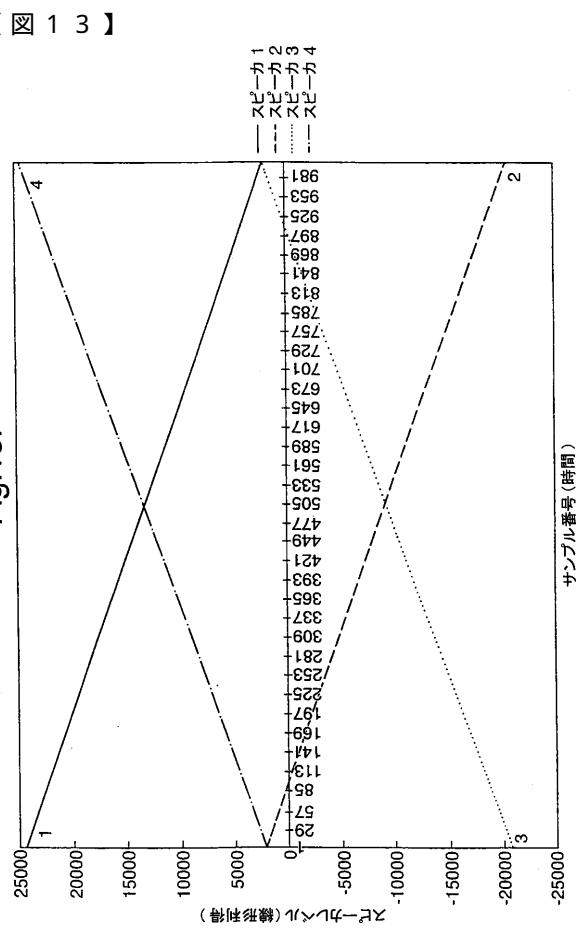
【図 12】

Fig.12.

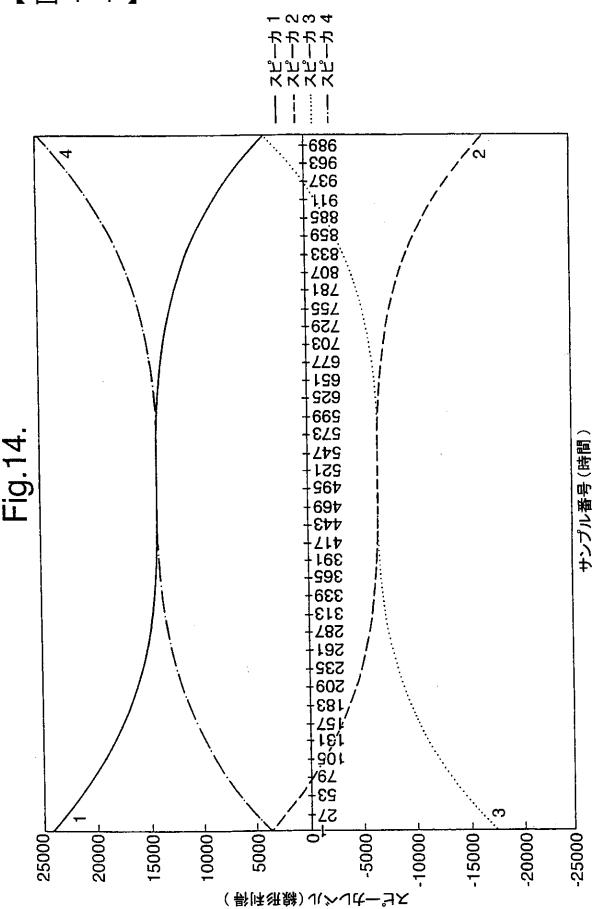


【図 13】

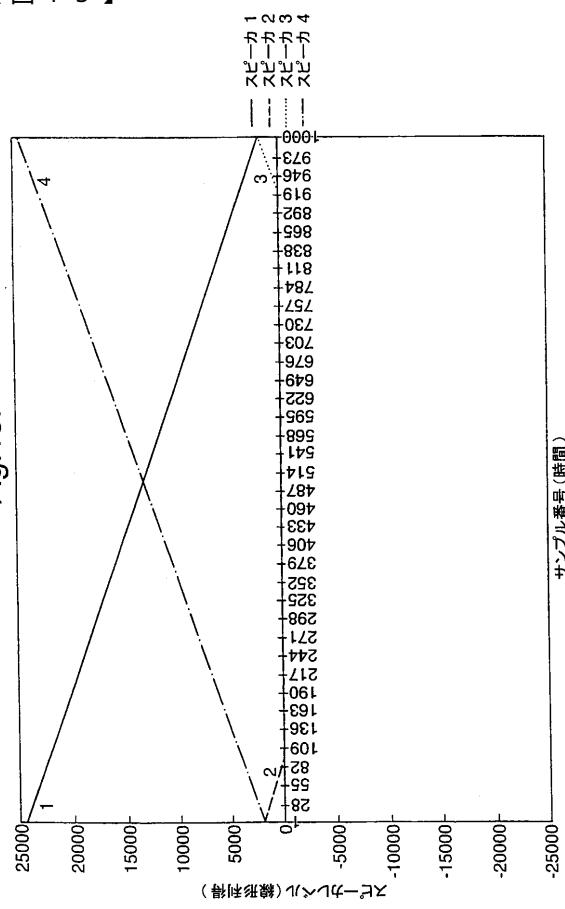
Fig.13.



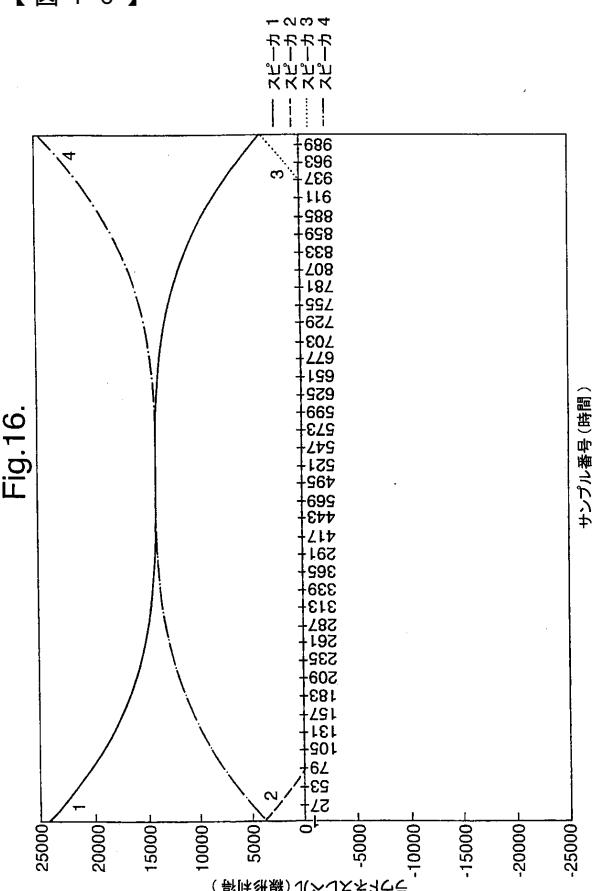
【図 1 4】



【図 1 5】



【図 1 6】



フロントページの続き

(72)発明者 リメル、アンドリュー

イギリス国、アイピー4・5ビーティー、サフォーク、イプスウィッч、アップランド・ロード
57

(72)発明者 ホーリール、マイケル・ピーター

イギリス国、アイピー5・2エックスエックス、サフォーク、イプスウィッч、ケスグレイブ、グランジェ・ファーム、ハーバート・ロード 15

審査官 井出 和水

(56)参考文献 米国特許第04392019(US, A)

米国特許第05533129(US, A)

米国特許第05307415(US, A)

MICHAEL A. GERZON, Ambisonics in Multichannel Broadcasting and Video, Journal of the Audio Engineering Society, 米国, Audio Engineering Society, 1985年11月, vol.33 no.11, pages 859-871

MICHAEL A. GERZON, Optimum Reproduction Matrices for Multispeaker Stereo, Journal of the Audio Engineering Society, 米国, Audio Engineering Society, 1992年 8月, vol.40 no.7-8, pages 571-589

(58)調査した分野(Int.Cl. , DB名)

H04S 1/00 - 7/00

H04R 3/00 - 5/04

G10K 15/00 - 15/12