

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関  
国際事務局

(43) 国際公開日  
2024年10月17日(17.10.2024)



(10) 国際公開番号

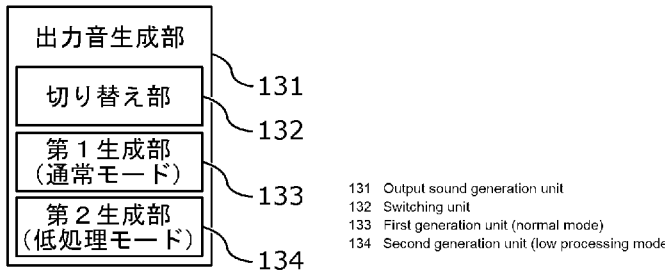
WO 2024/214799 A1

- (51) 国際特許分類:  
H04S 7/00 (2006.01)
- (21) 国際出願番号: PCT/JP2024/014744
- (22) 国際出願日: 2024年4月11日(11.04.2024)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:  
特願 2023-066552 2023年4月14日(14.04.2023) JP
- (71) 出願人: パナソニックホールディングス株式会社 (PANASONIC HOLDINGS CORPORATION) [JP/JP]; 〒5718501 大阪府門真市大字門真1006番地 Osaka (JP). 公立大学法人秋田県立大学 (AKITA PREFECTURAL UNIVERSITY) [JP/JP]; 〒0100195 秋田県秋田市下新城野字街道端西241-438 Akita (JP).
- (72) 発明者: 榎本 成悟 (ENOMOTO, Seigo). 宇佐見 陽 (USAMI, Hikaru). 中橋 康太 (NAKHASHI, Kota). 石川 智一 (ISHIKAWA, Tomokazu). 西口 正之 (NISHIGUCHI, Masayuki).
- (74) 代理人: 新居 広守 (NII, Hiromori); 〒5320011 大阪府大阪市淀川区西中島5丁目3番10号イトーピア新大阪ビル6階新居国際特許事務所内 Osaka (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,

(54) Title: INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 情報処理装置、情報処理方法、及び、プログラム

[図4]



(57) Abstract: An information processing device (101) comprises: an acquisition unit (111) that acquires sound information including a sound signal and information relating to the position of a sound source object in a three-dimensional sound field; a first generation unit (133) that generates an output sound signal using a head-related transfer function corresponding to an arrival direction based on the position of the sound source object and the position of a user in the three-dimensional sound field, and the sound signal; and a second generation unit (134) that generates an output sound signal using a head-related transfer function corresponding to a representative direction based on the position of a representative point set in the three-dimensional sound field and the position of the user, and the sound signal.

(57) 要約: 情報処理装置(101)は、音声信号と、三次元音場内の音源オブジェクトの位置の情報とを含む音情報を取得する取得部(111)と、音源オブジェクトの位置と三次元音場内のユーザの位置とに基づく到来方向に応じた頭部伝達関数と、音声信号とを用いて、出力音信号を生成する第1生成部(133)と、三次元音場内に設定された代表点の位置とユーザの位置とに基づく代表方向に応じた頭部伝達関数と、音声信号とを用いて、出力音信号を生成する第2生成部(134)と、を備える。

PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK,  
SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA,  
UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保  
護が可能): ARIPO (BW, CV, GH, GM, KE, LR, LS,  
MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG,  
ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU,  
TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ,  
DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS,  
IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT,  
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF,  
CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE,  
SN, TD, TG).

添付公開書類 :

一 国際調査報告 (条約第21条(3))

## 明 細 書

**発明の名称**：情報処理装置、情報処理方法、及び、プログラム  
**技術分野**

[0001] 本開示は、情報処理装置、情報処理方法、及び、プログラムに関する。

### 背景技術

[0002] 従来、仮想的な三次元空間内で、立体的な音をユーザに知覚させるための音響再生に関する技術が知られている（例えば、特許文献1参照）。また、このような三次元空間内で音源オブジェクトからユーザへと到来するように音を知覚させるためには、元となる音情報から出力音情報を生成する処理が必要となる。特に、仮想空間内でユーザの身体の動きに応じた立体的な音を再生するためには膨大な処理が必要になる。特にコンピュータグラフィックス（CG）の発展により視覚的に複雑な仮想環境を比較的容易に構築することが可能になり、対応する聴覚情報を実現する技術が重要となっている。加えて、音情報から出力音情報を生成するまでの処理を事前に行う場合には、事前に計算した処理結果を保存する大きな記憶領域が必要になる。また、そのような大きな処理結果のデータを伝送する場合には広い通信帯域が必要となる場合がある。

[0003] より現実に近い音環境を実現するため、仮想的な三次元空間内で音を出すオブジェクトの数が増えたり、反射音や回折音や残響などの音響効果が増えたり、さらにユーザの動きに対してこれら音響効果を適切に変化させる必要があり、大きな処理量が要求される。そこで、このような大きな処理量を削減するという目的で、三次元空間内の音をあらかじめ三次元空間内に設定された、いくつかの代表点からの音によって表現するパニング処理と呼ばれる変換技術が知られている。

### 先行技術文献

#### 特許文献

[0004] 特許文献1：特開2020-18620号公報

## 発明の概要

### 発明が解決しようとする課題

[0005] ただし、パニング処理のような変換処理では、処理量の削減に効果を成さない場合がある。そこで、本開示では、変換処理を効果的に適用するための情報処理装置などを提供することを目的とする。

### 課題を解決するための手段

[0006] 本開示の一態様に係る情報処理装置は、音声信号と、三次元音場内の音源オブジェクトの位置の情報とを含む音情報を取得する取得部と、前記音源オブジェクトの位置と前記三次元音場内のユーザの位置とに基づく到来方向に応じた頭部伝達関数と、前記音声信号とを用いて、出力音信号を生成する第1生成部と、前記三次元音場内に設定された代表点の位置と前記ユーザの位置とに基づく代表方向に応じた頭部伝達関数と、前記音声信号とを用いて、出力音信号を生成する第2生成部と、を備える。

[0007] また、本開示の別の一態様に係る情報処理装置は、複数の方向のそれぞれと時間シフト調整量及びゲイン調整量とを対応付けて記憶する記憶部と、音声信号と、三次元音場内の音源オブジェクトの位置の情報とを取得する取得部と、前記音声信号と、前記音源オブジェクトの位置と前記三次元音場内のユーザの位置とに基づく第1方向に対応する前記時間シフト調整量及びゲイン調整量とを用いて、第2方向から前記ユーザの位置に到来する音として出力音信号を生成する第2生成部と、を備える。

[0008] また、本開示の一態様に係る情報処理方法は、音情報を処理することにより、仮想的な三次元音場内の音源オブジェクトから到来する音として出力音信号を生成する、コンピュータによって実行される情報処理方法であって、前記音源オブジェクトの位置、及び、音声信号であって、当該音声信号により前記音源オブジェクトにおいて発せられる再生音を含む音声信号を取得するステップと、前記三次元音場内のユーザの位置を取得するステップと、前記音源オブジェクトの位置から前記ユーザの位置に到来する前記再生音の到来方向を算出するステップと、算出した前記到来方向に応じた頭部伝達関数

と前記再生音とを用いて前記出力音信号を生成するステップと、前記三次元音場内に設定された代表点の位置と前記ユーザの位置とに基づく代表方向に応じた頭部伝達関数と、前記音声信号とを用いて、前記出力音信号を生成するステップと、を含む。

[0009] また、本開示の一態様は、上記に記載の情報処理方法をコンピュータに実行させるためのプログラムとして実現することもできる。

[0010] なお、これらの包括的又は具体的な態様は、システム、装置、方法、集積回路、コンピュータプログラム、又は、コンピュータ読み取り可能なCD-ROMなどの非一時的な記録媒体で実現されてもよく、システム、装置、方法、集積回路、コンピュータプログラム、及び、記録媒体の任意な組み合わせで実現されてもよい。

### 発明の効果

[0011] 本開示によれば、変換処理を効果的に適用することが可能となる。

### 図面の簡単な説明

[0012] [図1]図1は、実施の形態に係る音響再生システムの使用事例を示す概略図である。

[図2]図2は、実施の形態に係る音響再生システムの機能構成を示すブロック図である。

[図3]図3は、実施の形態に係る取得部の機能構成を示すブロック図である。

[図4]図4は、実施の形態に係る出力音生成部の機能構成を示すブロック図である。

[図5]図5は、実施の形態に係る情報処理装置の第1の動作例を示すフローチャートである。

[図6]図6は、実施の形態に係る情報処理装置の第2の動作例を示すフローチャートである。

[図7]図7は、実施の形態に係るパニング処理の処理対象について説明するための図である。

[図8]図8は、実施の形態に係る情報処理装置の第3の動作例を示すフローチャートである。

ャートである。

## 発明を実施するための形態

[0013] (開示の基礎となった知見)

従来、仮想的な三次元空間内（以下、三次元音場という場合がある）で、立体的な音をユーザに知覚させるための音響再生に関する技術が知られている（例えば、特許文献1参照）。この技術を用いることで、ユーザは仮想空間内の所定位置に音源オブジェクトが存在し、その方向から音が到来するかのごとく、この音を知覚することができる。このように仮想的な三次元空間内の所定位置に音像を定位させるには、例えば、音源オブジェクトが鳴らしている音の信号（音源オブジェクトにおいて発せられる音、又は、再生音ともいう）に対して、立体的な音として知覚されるような両耳間での音の到来時間差、及び、両耳間での音のレベル差（又は音圧差）などを生じさせる計算処理が必要となる。このような計算処理は、立体音響フィルタを適用することによって行われる。立体音響フィルタは、元の音情報に対して、当該フィルタを適用した後の出力音信号が再生されると、音の方向や距離などの位置や音源の大きさ、空間の広さなどが立体感をもって知覚されるようになる情報処理用のフィルタである。

[0014] このような立体音響フィルタの適用の計算処理の一例として、所定方向から到来する音として知覚させるための頭部伝達関数を目的の音の信号に対して畳み込む処理が知られている。この頭部伝達関数の畳み込みの処理を、音源オブジェクトの位置からユーザの位置までの再生音の到来方向に対して、十分に細かい角度で実施することで、ユーザが体感する臨場感が向上される。

[0015] また、近年、仮想現実（VR: Virtual Reality）に関する技術の開発が盛んに行われている。仮想現実では、ユーザの動きに対して仮想的な三次元空間内の音オブジェクトの位置が適切に変化し、あたかもユーザが仮想空間内を移動しているように体感できることが主眼に置かれている。このためには、ユーザの動きに対して、仮想空間内の音像の定位位置を相対的に移動させる

必要が生じる。このような処理は、元の音情報に対して、上記の頭部伝達関数のような立体音響フィルタを適用することで行われてきた。ただし、三次元空間内でユーザが移動する場合などには、音の反響及び干渉など、音源オブジェクトとユーザとの位置関係ごとに、音の伝達経路が時々刻々と変化する。そうすると、その都度、音源オブジェクトとユーザとの位置関係をもとに、音源オブジェクトからの音の伝達経路を決定し、音の反響及び干渉などを考慮して伝達関数を畳み込んでいては、情報処理が膨大となり、大規模な処理装置がなければ、臨場感の向上が望めないことがある。

[0016] そこで、このような膨大化する処理量を削減するという目的で、再生音にパニング処理を適用して、頭部伝達関数の畳み込み量を削減するという試みが行われている。具体的には、三次元空間内に、いくつもある音源オブジェクトのそれぞれについて、再生音に頭部伝達関数を畳み込むのではなく、音源オブジェクトからの再生音を、三次元空間内にあらかじめ設定された、いくつかの代表点からの音（代表音）によって表現しなおす。そして、代表音に代表点からのユーザの位置までの頭部伝達関数を畳み込むだけで、ユーザに遜色ない立体音を知覚させることが可能となる。代表点が元の音源オブジェクトの数よりも少なければ、当然、頭部伝達関数の畳み込みを行う対象も少なくなるため、処理量の観点で有利となる。

[0017] 一方、このようなパニング処理を適用する場合、元の音源オブジェクトが少ないなど、その他いくつかの条件においては、パニング処理そのものの処理量の増大分もあるため、全体としての処理量の削減効果が得られないことが起こる。そこで、本開示では、パニング処理を適用する場合と、適用しない場合との両方が可能なように、2通りの出力音信号生成のための処理部を備える情報処理装置を提供する。これにより、パニング処理が処理量の削減に効果的であれば、パニング処理を適用した出力音信号の生成を行い、そうでなければ、パニング処理を適用することなく出力音信号の生成を行うことが可能となる。つまり、パニング処理のような変換処理を効果的に適用することが可能となる。

- [0018] より具体的な本開示の概要は、以下の通りである。
- [0019] 本開示の第1態様に係る情報処理装置は、音声信号と、三次元音場内の音源オブジェクトの位置の情報とを含む音情報を取得する取得部と、音源オブジェクトの位置と三次元音場内のユーザの位置とに基づく到来方向に応じた頭部伝達関数と、音声信号とを用いて、出力音信号を生成する第1生成部と、三次元音場内に設定された代表点の位置とユーザの位置とに基づく代表方向に応じた頭部伝達関数と、音声信号とを用いて、出力音信号を生成する第2生成部と、を備える。
- [0020] このような情報処理装置によれば、第1生成部を用いて算出した到来方向に応じた頭部伝達関数により出力音信号を生成することと、第2生成部を用いて、代表方向に応じた頭部伝達関数により出力音信号を生成することとを行うことができる。例えば、第2生成部を用いることが処理量の削減に効果的に作用する場合は、第2生成部を用いるようにし、そうでなければ、第1生成部を用いるようにすることができる。つまり、処理量の観点で、例えば条件分けなどを行うことにより、変換処理を効果的に適用することが可能となる。
- [0021] また、第2態様に係る情報処理装置は、第1生成部が、音声信号によって音源オブジェクトにおいて発せられる再生音に、到来方向に応じた頭部伝達関数を畳み込むことで出力音信号を生成し、第2生成部が、再生音を代表点から到来する代表音に変換する変換処理を実行し、代表方向に応じた頭部伝達関数を畳み込むことで出力音信号を生成する、第1態様に記載の情報処理装置である。
- [0022] これによれば、第1生成部が、到来方向に応じた頭部伝達関数を再生音に畳み込むことで出力音信号を生成し、第2生成部が、パニング処理などの変換処理により、到来方向からの音を、三次元音場内に設定された代表点のそれぞれから到来する代表音によって表現する出力音信号を生成することができる。例えば、変換処理を適用することが処理量の削減に効果的に作用する場合は、第2生成部を用いるようにし、そうでなければ、第1生成部を用い

ることができる。つまり、処理量の観点で、例えば条件分けなどを行うことにより、変換処理を効果的に適用することが可能となる。

[0023] また、第3態様に係る情報処理装置は、変換処理では、再生音に対して時間シフト調整とゲイン調整とを適用して代表音に変換する、第2態様に記載の情報処理装置である。

[0024] これによれば、変換処理において、再生音に対して時間シフト調整とゲイン調整とを適用して代表音に変換することができる。この結果、変換処理を適用しても違和感が低減され、より臨場感の高い出力音信号を生成することができる。

[0025] また、第4態様に係る情報処理装置は、音情報が、複数の音源オブジェクトのそれぞれの位置及び音声信号によって複数の音源オブジェクトのそれぞれにおいて発せられる再生音を含み、代表点の数が、音源オブジェクトの数に基づいて決定される、第1～第3態様のいずれか1態様に記載の情報処理装置である。

[0026] これによれば、音源オブジェクトの数に基づいて、代表点の数を動的に変化させ、都度適切な変換処理を行うことができる。

[0027] また、第5態様に係る情報処理装置は、代表点の数が、音源オブジェクトの数より少ない数である、第4態様に記載の情報処理装置である。

[0028] これによれば、音源オブジェクトの数に基づいて、代表点の数を動的に変化させ、都度適切な変換処理を行うことができる。特に、音源オブジェクトの数に対して少ない数の代表点とすることができるので、変換処理による処理量の削減効果を高めやすいという利点がある。

[0029] また、第6態様に係る情報処理装置は、変換処理の時間シフト調整では、再生音に対して、到来方向に応じた頭部伝達関数と代表方向に応じた頭部伝達関数との相互相関が最大になるように算出された時間シフト、又は、当該時間シフトに負号を付した時間シフトを行う、第3態様に記載の情報処理装置である。

[0030] これによれば、到来方向の頭部伝達関数と代表方向の頭部伝達関数との相

相互相関が最大になるように算出された時間シフト、又は、当該時間シフトに負号を付した時間シフトを行うことで再生音に対してする時間シフト調整を行うことができる。

[0031] また、第7態様に係る情報処理装置は、変換処理では、時間シフト調整及びゲイン調整の少なくとも一方が、周波数軸上の重み付けフィルタをかけてから相互相関が最大になるように算出された時間シフト、又は、当該時間シフトに負号を付した時間シフトを行う、第6態様に記載の情報処理装置である。

[0032] これによれば、周波数軸上の重み付けフィルタをかけてから相互相関が最大になるように算出された時間シフト、又は、当該時間シフトに負号を付した時間シフトを行うことで、時間シフト調整及びゲイン調整の少なくとも一方を行うことができる。

[0033] また、第8態様に係る情報処理装置は、変換処理では、2以上の代表点のそれぞれについて、時間シフトした再生音に、再生音及び代表方向毎に設定されたゲインをかける、第6態様に記載の情報処理装置である。

[0034] これによれば、2以上の代表点のそれぞれについて、時間シフトした再生音に、再生音到来方向及び代表方向毎に設定されたゲインをかけて変換処理を行うことができる。

[0035] また、第9態様に係る情報処理装置は、変換処理では、代表方向に応じた頭部伝達関数ベクトルの和で到来方向に応じた頭部伝達関数ベクトルを合成する際に、合成された頭部伝達関数ベクトルと到来方向に応じた頭部伝達関数ベクトルとの誤差信号ベクトルが代表方向に応じた頭部伝達関数ベクトルと直交するようにして算出したゲインを用いる、第8態様に記載の情報処理装置である。

[0036] これによれば、代表方向の頭部伝達関数ベクトルの和で到来方向の頭部伝達関数ベクトルを合成する際に、合成された頭部伝達関数ベクトルと到来方向の頭部伝達関数ベクトルとの誤差信号ベクトルが代表方向の頭部伝達関数ベクトルと直交するようにして算出したゲインを用いて変換処理を行うこと

ができる。

[0037] また、第10態様に係る情報処理装置は、変換処理では、合成された頭部伝達関数ベクトルと到来方向に応じた頭部伝達関数ベクトルとの誤差信号ベクトルのエネルギー又はL2ノルムを最小化するようにして算出されたゲインを用いる、第8態様に記載の情報処理装置である。

[0038] これによれば、合成された頭部伝達関数ベクトルと到来方向の頭部伝達関数ベクトルとの誤差信号ベクトルのエネルギー又はL2ノルムを最小化するようにして算出されたゲインを用いて変換処理を行うことができる。

[0039] また、第11態様に係る情報処理装置は、誤差信号ベクトルには、周波数軸上の重み付けフィルタをかけたものを用いる、第10態様に記載の情報処理装置である。

[0040] これによれば、誤差信号ベクトルとして、周波数軸上の重み付けフィルタをかけたものを用いることができる。

[0041] また、ある態様に係る情報処理装置は、情報処理装置が、頭部伝達関数を記憶するための記憶部に記憶されていない、新たな頭部伝達関数を読み込んだ場合に、当該新たな頭部伝達関数に対して、変換処理に用いられる時間シフト調整及びゲイン調整における調整量を決定し、読み込んだ新たな頭部伝達関数と、決定した調整量とを紐づけてデータベースに記憶させ、変換処理では、再生音に対して、記憶部に記憶された新たな頭部伝達関数に紐づけられた調整量で時間シフト調整とゲイン調整とを適用して代表音に変換する、第3態様に記載の情報処理装置である。

[0042] これによれば、頭部伝達関数を記憶するため記憶部に記憶されていない、新たな頭部伝達関数を読み込んだ場合に、当該新たな頭部伝達関数に対して、変換処理に用いられる時間シフト調整及びゲイン調整における調整量を決定し、読み込んだ新たな頭部伝達関数と、決定した調整量とを紐づけて記憶部に記憶させ、これを変換処理に用いることができる。新たな頭部伝達関数には、当該頭部伝達関数に適した調整量があり、そのような調整量を変換処理を開始する前に（例えば、音信号のデコード時、音響再生システムの電源

投入時、又は、音響再生システムの初期化時等)、決定しておくことで、適切な調整量での変換処理を、処理量の増大を抑制しながら行うことができる。

[0043] また、第12態様に係る情報処理装置は、初期化時に代表方向の頭部伝達関数と、変換処理に用いられる時間シフト調整及びゲイン調整における調整量とを頭部伝達関数の各方向ごとに紐づけた調整量テーブルを記憶部に記憶させ、変換処理では、再生音に対して、記憶部に記憶された調整量テーブルにおいて、代表方向に応じた頭部伝達関数の各方向に紐づけられた調整量で時間シフト調整とゲイン調整とを適用して代表音に変換する、第3態様に記載の情報処理装置である。

[0044] これによれば、変換処理において、初期化時に記憶部に記憶された調整量テーブルから、代表方向に応じた頭部伝達関数の各方向に紐づけられた調整量で時間シフト調整とゲイン調整とを適用して代表音に変換することができる。

[0045] また、第13態様に係る情報処理装置は、初期化時に、複数の代表方向を決定し、調整量テーブルは、決定された複数の代表方向の頭部伝達関数に基づいて作成される、第12態様に記載の情報処理装置である。

[0046] これによれば、変換処理において、決定された複数の代表方向の頭部伝達関数に基づいて作成された頭部伝達関数の各方向に紐づけられた調整量で時間シフト調整とゲイン調整とを適用して代表音に変換することができる。

[0047] また、第14態様に係る情報処理装置は、音情報が、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを指定するフラグを含み、情報処理装置が、第1生成部又は第2生成部のうち、取得した音情報に含まれるフラグにおいて指定された一方を用いて出力音信号を生成する、第1～第13態様のいずれか1態様に記載の情報処理装置である。

[0048] これによれば、音情報に含まれているフラグによって、第1生成部又は第2生成部のうち指定された一方を用いて出力音信号を生成することができる。

。つまり、第1生成部又は第2生成部のうち、いずれを用いるかをフラグによって指定することができる。

[0049] また、第15態様に係る情報処理装置は、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替える切り替え部を備える、第1～第14態様のいずれか1態様に記載の情報処理装置である。

[0050] これによれば、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替えることができる。

[0051] また、第16態様に係る情報処理装置は、切り替え部が、音情報に含まれる音源オブジェクトの数と、三次元音場内に設定された代表点の数とを比較し、比較結果に応じて、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替える、第15態様に記載の情報処理装置である。

[0052] これによれば、切り替え部が音情報に含まれる音源オブジェクトの数と、三次元音場内に設定された代表点の数とを比較して、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを適切に切り替えることができる。

[0053] また、第17態様に係る情報処理装置は、切り替え部が、頭部伝達関数を記憶するための記憶部内に記憶されている頭部伝達関数が所定の条件を満たさない場合に、第1生成部を用いて出力音信号を生成するように切り替える、第15態様に記載の情報処理装置である。

[0054] これによれば、切り替え部が記憶部内の頭部伝達関数が所定の条件を満たさない場合に、第1生成部を用いて出力音信号を生成するように切り替えることができる。

[0055] また、第18態様に係る情報処理装置は、音声信号によって音源オブジェクトにおいて発せられる再生音の伝播経路を算出し、算出した再生音の伝播経路に応じた再生音の間接的な伝播によりユーザの位置に到来する合成音及び当該合成音の到来方向を算出する経路算出部を備える、第1～17態様の

いずれか1態様に記載の情報処理装置である。

[0056] これによれば、経路算出部により、音源オブジェクトからの再生音の伝播経路を算出し、算出した、再生音の伝播経路に応じた、再生音の間接的な伝播によりユーザの位置に到来する合成音及び当該合成音の到来方向を算出することができる。

[0057] また、第19態様に係る情報処理装置は、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替える切り替え部を備え、切り替え部が、再生音及び合成音のそれぞれについて個別に、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替える、第18態様に記載の情報処理方法である。

[0058] これによれば、再生音及び合成音のそれぞれについて個別に、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替えることができる。

[0059] また、第20態様に係る情報処理装置は、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替える切り替え部を備え、経路算出部が、互いに異なる間接的な伝播によりユーザの位置に到来する2以上の合成音及び当該2以上の合成音のそれぞれの到来方向を算出し、切り替え部が、2以上の合成音のそれぞれについて個別に、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替える、第18態様に記載の情報処理方法である。

[0060] これによれば、経路算出部が互いに異なる間接的な伝播によりユーザの位置に到来する2以上の合成音及び当該2以上の合成音のそれぞれの到来方向を算出し、この2以上の合成音のそれぞれについて個別に、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替えることができる。

[0061] また、第21態様に係る情報処理装置は、第1生成部を用いて出力音信号

を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替える切り替え部を備え、切り替え部が、再生音及び合成音の合計数と、三次元音場内に設定された代表点の数とを比較し、比較結果に応じて、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替える、第18態様に記載の情報処理方法である。

[0062] これによれば、再生音及び合成音の合計数と、三次元音場内に設定された代表点の数とを比較して、第1生成部を用いて出力音信号を生成するか、又は、第2生成部を用いて出力音信号を生成するかを切り替えることができる。

[0063] また、第22態様に係る情報処理方法は、音情報を処理することにより、仮想的な三次元音場内の音源オブジェクトから到来する音として出力音信号を生成する、コンピュータによって実行される情報処理方法であって、音源オブジェクトの位置、及び、音声信号であって、当該音声信号により音源オブジェクトにおいて発せられる再生音を含む音声信号を取得するステップと、三次元音場内のユーザの位置を取得するステップと、音源オブジェクトの位置からユーザの位置に到来する再生音の到来方向を算出するステップと、算出した到来方向に応じた頭部伝達関数と再生音とを用いて出力音信号を生成するステップと、三次元音場内に設定された代表点の位置とユーザの位置とに基づく代表方向に応じた頭部伝達関数と、音声信号とを用いて、出力音信号を生成するステップと、を含む。

[0064] これによれば、上記に記載の情報処理装置と同様の効果を奏することができる。

[0065] また、第23態様に係るプログラムは、上記に記載の情報処理方法をコンピュータに実行させるためのプログラムである。

[0066] これによれば、コンピュータを用いて上記に記載の情報処理方法と同様の効果を奏することができる。

[0067] また、別の態様に係る情報処理装置は、頭部伝達関数を用いて音情報を処理することにより、仮想的な三次元音場内の音源オブジェクトから到来する

音として出力音信号を生成する情報処理装置であって、音源オブジェクトの位置、及び、音源オブジェクトにおいて発せられる再生音を含む音情報を取得する音取得部と、三次元音場内のユーザの位置を取得する位置取得部と、音源オブジェクトの位置からユーザの位置に到来する再生音の相対的な到来方向を算出する到来方向算出部と、第3生成部と、を備え、頭部伝達関数を記憶するための記憶部に記憶されていない、新たな頭部伝達関数を読み込んだ場合に、記憶部に記憶する前に、当該新たな頭部伝達関数に対して、変換処理に用いられる時間シフト調整及びゲイン調整における調整量を決定し、読み込んだ新たな頭部伝達関数と、決定した調整量とを紐づけて記憶部に記憶させ、第3生成部が、再生音に対して、記憶部に記憶された新たな頭部伝達関数に紐づけられた調整量で時間シフト調整とゲイン調整とを適用して代表音に変換し、代表点のそれぞれの位置からユーザの位置に向かう代表方向に応じた頭部伝達関数を代表音に畳み込むことで出力音信号を生成する。

[0068] これによれば、頭部伝達関数を記憶するため記憶部に記憶されていない、新たな頭部伝達関数を読み込んだ場合に、当該新たな頭部伝達関数に対して、パニング処理等の変換処理に用いられる時間シフト調整及びゲイン調整における調整量を決定し、読み込んだ新たな頭部伝達関数と、決定した調整量とを紐づけて記憶部に記憶させ、これを変換処理に用いることができる。新たな頭部伝達関数には、当該頭部伝達関数に適した調整量があり、そのような調整量を変換処理を開始する前に（例えば、音信号のデコード時、音響再生システムの電源投入時、又は、音響再生システムの初期化時等）決定しておくことで、適切な調整量での変換処理を、処理量の増大を抑制しながら行うことができる。

[0069] また、第24状態に係る情報処理装置は、複数の方向のそれぞれと時間シフト調整量及びゲイン調整量とを対応付けて記憶する記憶部と、音声信号と、三次元音場内の音源オブジェクトの位置の情報とを取得する取得部と、音声信号と、音源オブジェクトの位置と三次元音場内のユーザの位置とに基づく第1方向に対応する時間シフト調整量及びゲイン調整量とを用いて、第2

方向からユーザの位置に到来する音として出力音信号を生成する第2生成部と、を備える情報処理装置である。

[0070] これによれば、複数の方向のそれぞれと時間シフト調整量及びゲイン調整量とを対応付けて記憶しておくことで、取得した音声信号と、三次元音場内の音源オブジェクトの位置とユーザの位置とに基づく第1方向に対応する時間シフト調整量及びゲイン調整量とを用いて、適切な調整量で処理量の増大を抑制しながら第2方向からユーザの位置に到来する音として出力音信号を生成することができる。

[0071] また、第25状態に係る情報処理装置は、記憶部が、さらに第2方向に対応する頭部伝達関数を記憶しており、第2生成部が、音声信号と、第1方向に対応する時間シフト調整量及びゲイン調整量と、第2方向に対応する頭部伝達関数とを用いて、第2方向からユーザの位置に到来する音として出力音信号を生成する、第24状態に記載の情報処理装置である。

[0072] これによれば、さらに第2方向に対応する頭部伝達関数を含む補助記憶部等を保持して、このような情報を記憶しておくことで、第1方向に対応する時間シフト調整量及びゲイン調整量と、第2方向に対応する頭部伝達関数とを用いて、第2方向からユーザの位置に到来する音として出力音信号を生成することができる。

[0073] また、第26状態に係る情報処理装置は、記憶部が、さらに第2方向及び第2方向以外の方向に対応する頭部伝達関数を記憶しており、第2生成部が、音声信号と、第1方向に対応する時間シフト調整量及びゲイン調整量と、第2方向に対応する頭部伝達関数とを用いて、第2方向からユーザの位置に到来する音として出力音信号を生成し、情報処理装置が、さらに、第1生成部を備え、第1生成部が、音声信号と、第1方向に対応する頭部伝達関数とを用いて、第1方向からユーザの位置に到来する音として音声信号を生成する、第24状態に記載の情報処理装置である。

[0074] これによれば、記憶部が、さらに第2方向及び第2方向以外の方向に対応する頭部伝達関数を含む補助記憶部等を保持して、このような情報を記憶し

ておくことで、第2生成部が、音声信号と、第1方向に対応する時間シフト調整量及びゲイン調整量と、第2方向に対応する頭部伝達関数とを用いて、第2方向からユーザの位置に到来する音として出力音信号を生成し、さらに情報処理装置が第1生成部を備え、第1生成部が、音声信号と、第1方向に対応する頭部伝達関数とを用いて、第1方向からユーザの位置に到来する音として音声信号を生成することができる。例えば、時間シフト調整量及びゲイン調整量を用いて処理を実行することが処理量の削減に効果的に作用する場合は、第2生成部を用いるようにし、そうでなければ、第1生成部を用いるようにすることができる。つまり、処理量の観点で、例えば条件分けなどを行うことにより、変換処理を効果的に適用することが可能となる。

[0075] また、さらに別の様態に係る情報処理方法は、複数の方向のそれぞれと時間シフト調整量及びゲイン調整量とを対応付けて記憶する補助記憶部を保持し、音声信号と、三次元音場内の音源オブジェクトの位置の情報とを取得し、音声信号と、音源オブジェクトの位置とユーザの位置とに基づく第1方向に対応する時間シフト調整量及びゲイン調整量とを用いて、第2方向からユーザの位置に到来する音として出力音信号を生成する、情報処理方法である。

[0076] これによれば、第2 3 様態に記載の情報処理装置と同様の効果を奏することができる。

[0077] また、さらに別の態様に係るプログラムは、上記のさらに別の態様に記載の情報処理方法をコンピュータに実行させるためのプログラムである。

[0078] これによれば、コンピュータを用いて上記のさらに別の態様に記載の情報処理方法と同様の効果を奏することができる。

[0079] さらに、これらの包括的又は具体的な態様は、システム、装置、方法、集積回路、コンピュータプログラム、又は、コンピュータ読み取り可能なCD-ROMなどの非一時的な記録媒体で実現されてもよく、システム、装置、方法、集積回路、コンピュータプログラム、及び、記録媒体の任意な組み合わせで実現されてもよい。

[0080] 以下、実施の形態について、図面を参照しながら具体的に説明する。なお、以下で説明する実施の形態は、いずれも包括的又は具体的な例を示すものである。以下の実施の形態で示される数値、形状、材料、構成要素、構成要素の配置位置及び接続形態、ステップ、ステップの順序などは、一例であり、本開示を限定する主旨ではない。また、以下の実施の形態における構成要素のうち、独立請求項に記載されていない構成要素については、任意の構成要素として説明される。なお、各図は模式図であり、必ずしも厳密に図示されたものではない。また、各図において、実質的に同一の構成に対しては同一の符号を付し、重複する説明は省略又は簡略化される場合がある。

[0081] また、以下の説明において、第1、第2及び第3等の序数が要素に付けられている場合がある。これらの序数は、要素を識別するため、要素に付けられており、意味のある順序に必ずしも対応しない。これらの序数は、適宜、入れ替えられてもよいし、新たに付与されてもよいし、取り除かれてもよい。

[0082] (実施の形態)

[概要]

はじめに、実施の形態に係る音響再生システムの概要について説明する。

図1は、実施の形態に係る音響再生システムの使用事例を示す概略図である。図1では、音響再生システム100を使用するユーザ99が示されている。

[0083] 図1に示す音響再生システム100は、立体映像再生装置300と同時に使用されている。立体的な画像及び立体的な音を同時に視聴することで、画像が聴覚的な臨場感を、音が視覚的な臨場感をそれぞれ高め合い、画像及び音が撮られた現場に居るかのように体感することができる。例えば、人が会話をする画像（動画像）が表示されている場合に、会話音の音像（音源オブジェクト）の定位が当該人の口元とずれている場合にも、ユーザ99が、当該人の口から発せられた会話音として知覚することが知られている。このように視覚情報によって、音像の位置が補正されるなど、画像と音とが併せら

れることで臨場感が高められることがある。

[0084] 立体映像再生装置300は、ユーザ99の頭部に装着される画像表示デバイスである。したがって、立体映像再生装置300は、ユーザ99の頭部と一体的に移動する。例えば、立体映像再生装置300は、図示するように、ユーザ99の耳と鼻とで支持するメガネ型のデバイスである。

[0085] 立体映像再生装置300は、ユーザ99の頭部の動きに応じて表示する画像を変化させることで、ユーザ99が三次元画像空間内で頭部を動かしているように知覚させる。つまり、ユーザ99の正面に三次元画像空間内の物体が位置しているときに、ユーザ99が右を向くと当該物体がユーザ99の左方向に移動し、ユーザ99が左を向くと当該物体がユーザ99の右方向に移動する。このように、立体映像再生装置300は、ユーザ99の動きに対して、三次元画像空間をユーザ99の動きとは逆方向に移動させる。

[0086] 立体映像再生装置300は、ユーザ99の左右の目それぞれに視差分のずれが生じた2つの画像をそれぞれ表示する。ユーザ99は、表示される画像の視差分のずれに基づき、画像上の物体の三次元的な位置を知覚することができる。なお、音響再生システム100を睡眠誘導用のヒーリング音の再生に使用する等、ユーザ99が目を閉じて使用する場合等には、立体映像再生装置300が同時に使用される必要はない。つまり、立体映像再生装置300は、本開示の必須の構成要素ではない。立体映像再生装置300としては、専用の映像表示デバイスの他にも、ユーザ99が所有するスマートフォン、タブレット装置など、汎用の携帯端末が用いられる場合もある。

[0087] このような汎用の携帯端末には、映像を表示するためのディスプレイの他に、端末の姿勢や動きを検知するための各種のセンサが搭載されている。さらには、情報処理用のプロセッサも搭載され、ネットワークに接続してクラウドサーバなどのサーバ装置と情報の送受信が可能になっている。つまり、立体映像再生装置300及び音響再生システム100をスマートフォンと、情報処理機能のない汎用のヘッドフォン等との組み合わせによって実現することもできる。

[0088] この例のように、頭部の動きを検知する機能、映像の提示機能、提示用の映像情報処理機能、音の提示機能、及び、提示用の音情報処理機能を1以上の装置に適切に配置して立体映像再生装置300及び音響再生システム100を実現してもよい。立体映像再生装置300が不要である場合には、頭部の動きを検知する機能、音の提示機能、及び、提示用の音情報処理機能を1以上の装置に適切に配置できればよく、例えば、提示用の音情報処理機能を有するコンピュータ又はスマートフォンなどの処理装置と、頭部の動きを検知する機能及び音の提示機能を有するヘッドホン等によって音響再生システム100を実現することもできる。

[0089] 音響再生システム100は、ユーザ99の頭部に装着される音提示デバイスである。したがって、音響再生システム100は、ユーザ99の頭部と一体的に移動する。例えば、本実施の形態における音響再生システム100は、いわゆるオーバーイヤーヘッドホン型のデバイスである。なお、音響再生システム100の形態に特に限定はなく、例えば、ユーザ99の左右の耳にそれぞれ独立して装着される2つの耳栓型のデバイスであってもよい。

[0090] 音響再生システム100は、ユーザ99の頭部の動きに応じて提示する音を変化させることで、ユーザ99が三次元音場内で頭部を動かしているようにユーザ99に知覚させる。このため、上記したように、音響再生システム100は、ユーザ99の動きに対して三次元音場をユーザ99の動きとは逆方向に移動させる。

[0091] ここで、ユーザ99が三次元音場内を移動する場合、ユーザ99の三次元音場内の位置に対する相対的な音源オブジェクトの位置が変化する。そうすると、ユーザ99が移動する度に音源オブジェクトとユーザ99との位置に基づく計算処理を行って再生用の出力音信号を生成する必要がある。通常このような処理は処理量が膨大になるため、本開示では、処理量の削減の観点で変換処理の1つとしてのパニング処理を適用して再生音を代表点からの代表音によって表現する。この結果、代表音に頭部伝達関数を畳み込むだけで、音源オブジェクトからの再生音をユーザ99に知覚させることが可能とな

る。以下、本実施の形態では、変換処理の一例としてパニング処理を用いる場合を説明するが、変換処理としては、パニング処理に限られず、条件によって、その変換で処理量の削減が見込まれる変換処理であれば、あらゆる変換処理が適用されうる。

[0092] 三次元音場内にあらかじめ設定された代表点が、音源オブジェクトの数よりも少なければ、頭部伝達関数の畳み込みの処理が少なくなるので、処理量の削減に寄与できる。ただし、パニング処理にはそれ自体に、元の再生音に頭部伝達関数を畳み込む場合にはない処理が要求されるので、処理量の削減効果が得られるのは、音源オブジェクトの数が代表点よりも数倍多いときに限られる。その他、処理量の削減効果を得るには、いくつかの条件があるため、本開示では、処理量の削減が見込まれない場合には、再生音に頭部伝達関数を畳み込む通常モードでの出力音信号の生成処理を行う。

[0093] [構成]

次に、図2を参照して、本実施の形態に係る音響再生システム100の構成について説明する。図2は、実施の形態に係る音響再生システムの機能構成を示すブロック図である。

[0094] 図2に示すように、本実施の形態に係る音響再生システム100は、情報処理装置101と、通信モジュール102と、検知器103と、ドライバ104と、を備える。

[0095] 情報処理装置101は、音響再生システム100における各種の信号処理を行うための演算装置である、情報処理装置101は、例えば、コンピュータなどの、プロセッサとメモリとを備え、メモリに記憶されたプログラムがプロセッサによって実行される形で実現される。このプログラムの実行によって、以下で説明する各機能部に関する機能が発揮される。

[0096] 情報処理装置101は、取得部111、経路算出部121、出力音生成部131、及び、信号出力部141、記憶部105を有する。情報処理装置101が有する各機能部の詳細は、情報処理装置101以外の構成の詳細と併せて以下に説明する。

[0097] 通信モジュール102は、音響再生システム100への音情報の入力を受け付けるためのインタフェース装置である。通信モジュール102は、例えば、アンテナと信号変換器とを備え、無線通信により外部の装置から音情報を受信する。通信モジュール102は、S O F Aファイルなどの頭部伝達関数のセットを外部の装置から受信してもよい。より詳しくは、通信モジュール102は、無線通信のための形式に変換された音情報を示す無線信号を、アンテナを用いて受波し、信号変換器により無線信号から音情報への再変換を行う。これにより、音響再生システム100は、外部の装置から無線通信により音情報及び頭部伝達関数のセットを取得する。通信モジュール102によって取得された音情報及び頭部伝達関数のセットは、取得部111によって取得される。このように、取得部111は、音取得部の一例である。音情報は、以上のようにして情報処理装置101に入力される。なお、音響再生システム100と外部の装置との通信は、有線通信によって行われてもよい。

[0098] 音響再生システム100が取得する音情報は、音響再生システム100によって再生される再生音についての情報（音信号）と、当該音の音像を三次元音場内において所定位置に定位させる（つまり所定方向から到来する音として知覚させる）際の定位位置に関する情報とで構成される。再生される再生音についての情報は、例えば、MPEG-H 3D Audio（ISO / IEC 23008-3）等の所定の形式で符号化された音信号や、符号化されていないPCM信号であってもよい。定位位置に関する情報は、音源オブジェクトに関する情報と読み替えることもできる。つまり、音情報には、音源オブジェクトの三次元音場内における位置と、音源オブジェクトが鳴らす音とを含んでいる。また、音情報には、パニング処理を適用するかしないかを決定するためのフラグが含まれている場合がある。このフラグについては後述する。

[0099] 音情報は、上記のように入力データとして得られ、再生音についての情報である音声信号（音響信号）と、その他の情報である音源オブジェクトの三

次元音場内位置の情報とを含んでいる。その他の情報には、他に、三次元音場を定義するための情報が含まれる場合がある。そのため、その他の情報を包括して音源オブジェクトの位置の情報及び三次元音場を定義するための情報等を含む、空間に関する情報（空間情報）という場合がある。音声信号を主体として見る場合には、入力データは、音声信号にその他の情報（メタデータ）が付帯する音情報であるといえる。また、空間情報を主体として見る場合には、入力データは、空間情報に音声信号が付帯する情報であるといえる。あるいは、このような入力データの両側面を有することから、入力データを音空間情報と考えてもよい。

[0100] 一具体例として、音情報には第1の再生音及び第2の再生音を含む複数の音に関する情報が含まれ、それぞれの音が再生された際の音像を三次元音場内における異なる位置から到来する音として知覚させるように定位させる。そのため第1の再生音の音源オブジェクトは、三次元音場内における第1の位置に、第2の再生音の音源オブジェクトは、三次元音場内における第2の位置に定位される。音情報には、このように、複数の音が含まれていることがある。

[0101] 立体的な音によって、例えば、立体映像再生装置300を用いて視認される画像と併せて、視聴されるコンテンツなどの臨場感を向上することができる。なお、音情報には、再生音についての情報のみが含まれていてもよい。この場合、所定位置に関する情報を別途取得してもよい。また、上記したように、音情報は、第1の再生音に関する第1音情報、及び、第2の再生音に関する第2音情報を含むが、これらを別個に含む複数の音情報をそれぞれ取得し、同時に再生することで三次元音場内における異なる位置に音像を定位させてもよい。このように、入力される音情報の形態に特に限定はなく、音響再生システム100に各種の形態の音情報に応じた取得部111が備えられればよい。

[0102] なお、取得した直後の音情報には、直接音に関する音声信号が含まれており、副次音を計算する変換処理によって残響音、1次反射音、回折音などの

それぞれの音声信号を含む音情報へと変換される。あるいは、直接音に関する音声信号を含む音情報の他に、このような副次音に関する音声信号が含まれる音情報が取得されてもよい。副次音を計算によって音情報に付加する変換処理には、三次元音場の空間環境の条件（例えば、三次元音場内のオブジェクトの位置、反射、回折特性等）の情報が用いられる。このように、副次音は、1つの再生音に関する音情報から、三次元音場の空間環境の条件によって計算的に生成される。1つの副次音からは、その副次音の伝搬によってさらに別の副次音が生じることもある。なお、空間環境の条件の情報は、空間情報の一部であり、入力された音情報によって、音声信号とともに取得される。

[0103] 副次音の到来方向には、反射音の場合のどのようなオブジェクトで反射するか、及び、その反射時の減衰率はどの程度かなどの付加情報を含む。付加情報は、入力された音情報によって計算された副次音の到来方向に含まれている。つまり、付加情報は、音情報から計算的に生成し取得される。

[0104] 空間情報について整理すると、空間情報には、空間（三次元音場）における音源オブジェクトの空間位置（音源オブジェクトの位置の情報）、当該音源オブジェクトにおける音の反射、回折特性（併せて、空間環境の条件の情報）、及び、三次元音場の広さなどのさらなる情報を含んでいる。経路算出部 121 が空間情報に基づいて、再生音がどの音源オブジェクトで反射又は回折するかによって副次音を生成し、その副次音の到来方向と、副次音が反射又は回折によって減衰した後の音量などを付加情報として算出する。音情報（入力データ）は、音声信号と付帯するメタデータの形で空間情報を含んでおり、その空間情報には、上記したように、音声信号以外の情報として、音を立体音にして三次元音場内に音源オブジェクトを位置させるようにするために必要な情報、及び／又は、音を立体音にして三次元音場内に音源オブジェクトを位置させるようにするために必要な情報を計算するのに用いられる情報を含んでいる。

[0105] ここで、取得部 111 の一例を、図 3 を用いて説明する。取得部 111 は

、出力音生成に必要な情報を取得する処理部であって、出力音生成に必要な情報は、音情報及び頭部伝達関数のセット、センシング情報などを含む。図3は、実施の形態に係る取得部の機能構成を示すブロック図である。図3に示すように、本実施の形態における取得部111は、例えば、エンコード音情報入力部112、デコード処理部113、及び、センシング情報入力部114を備える。

[0106] エンコード音情報入力部112は、取得部111が取得した、符号化された（言い換えるとエンコードされている）音情報が入力される処理部である。符号化された音情報とは、例えばMPEG-H 3D Audio (ISO/IEC 23008-3)等の所定の形式で符号化された音信号を含む。エンコード音情報入力部112は、入力された音情報をデコード処理部113へと出力する。デコード処理部113は、エンコード音情報入力部112から出力された音情報を復号する（言い換えるとデコードする）ことにより音情報に含まれる再生音（音信号）と、音源オブジェクトの位置と、フラグとを、以降の処理に用いられる形式で生成する処理部である。センシング情報入力部114については、検知器103の機能とともに、以下に説明する。

[0107] なお、エンコード音情報入力部112及びデコード処理部113で行われる処理は、情報処理装置101の外部の装置で実行されてもよい。つまり取得部111は、音情報を取得すればよく、外部の装置でデコード処理が行われた音情報を、通信モジュール102を通じて取得してもよい。また、音情報が符号化されている例を説明したが、音情報は符号化されていなくてもよい。例えば、再生音の情報は、PCM信号など符号化されていない音信号として取得されてもよい。

[0108] 音情報に含まれる音信号と空間情報は、それぞれ別のストリームやファイルで取得されてもよいし、同一のストリームやファイルで取得されてもよい。

[0109] また、取得部111は図示しない頭部伝達関数入力部を備えていてもよく

、通信モジュール102を通じて外部から取得された頭部伝達関数のセットを取得し、記憶部105に出力してもよい。

[0110] 検知器103は、ユーザ99の頭部の動き速度を検知するための装置である。検知器103は、ジャイロセンサ、加速度センサなど動きの検知に使用される各種のセンサを組み合わせて構成される。本実施の形態では、検知器103は、音響再生システム100に内蔵されているが、例えば、音響再生システム100と同様にユーザ99の頭部の動きに応じて動作する立体映像再生装置300等、外部の装置に内蔵されていてもよい。この場合、検知器103は、音響再生システム100に含まれなくてもよい。また、検知器103として、外部の撮像装置などを用いて、ユーザ99の頭部の動きを撮像し、撮像された画像を処理することでユーザ99の動きを検知してもよい。

[0111] 検知器103は、例えば、音響再生システム100の筐体に一体的に固定され、筐体の動きの速度を検知する。上記の筐体を含む音響再生システム100は、ユーザ99が装着した後、ユーザ99の頭部と一体的に移動するため、検知器103は、結果としてユーザ99の頭部の動きの速度を検知することができる。

[0112] 検知器103は、例えば、ユーザ99の頭部の動きの量として、三次元空間内で互いに直交する3軸の少なくとも一つを回転軸とする回転量を検知してもよいし、上記3軸の少なくとも一つを変位方向とする変位量を検知してもよい。また、検知器103は、ユーザ99の頭部の動きの量として、回転量及び変位量の両方を検知してもよい。

[0113] センシング情報入力部114は、検知器103からユーザ99の頭部の動き速度を取得する。より具体的には、センシング情報入力部114は、単位時間あたりに検知器103が検知したユーザ99の頭部の動きの量を動きの速度として取得する。このようにしてセンシング情報入力部114は、検知器103から回転速度及び変位速度の少なくとも一方を取得する。ここで取得されるユーザ99の頭部の動きの量は、三次元音場内のユーザ99の位置及び姿勢（言い換えると座標及び向き）を決定するために用いられる。その

ため、取得部 111 は、センシング情報入力部 114 によって位置取得部としても機能する。音響再生システム 100 では、決定されたユーザ 99 の座標及び向きに基づいて、音像オブジェクトのユーザ 99 に対する相対的な位置を決定して音が再生される。具体的には、経路算出部 121、出力音生成部 131 によって、上記の機能が実現されている。

[0114] 経路算出部 121 は、決定されたユーザ 99 の座標及び向きに基づいて、再生音について、音源オブジェクトの位置からユーザ 99 の位置に到来する相対的な到来方向を算出する到来方向算出機能と、音源オブジェクトからの伝播経路を算出し、算出した再生音の伝播経路に応じた再生音の間接的な伝播によりユーザ 99 の位置に到来する合成音及び当該合成音の到来方向を算出する合成音算出機能とを含んでいる。つまり、経路算出部 121 は、到来方向算出部の一例でもある。

[0115] 経路算出部 121 は、再生音が直接音としてユーザに届く際の再生音の到来方向を算出することと、再生音の間接的な伝播によりユーザ 99 の位置に到来する合成音（例えば、反射音、回折音及び残響音など）をその到来方向とともに算出することができれば、どのような処理によって実現されてもよい。経路算出部 121 は、再生音及び合成音について三次元音場内のいずれの方向から到来する音としてユーザ 99 に知覚させるかを上記のユーザ 99 の座標及び向きに基づいて決定し、出力音信号が再生された場合に、そのような音として知覚されるように、音情報を処理する。

[0116] 出力音生成部 131 は、音情報に含まれる再生音に関する情報を処理することにより、出力音信号を生成する処理部である。

[0117] ここで、出力音生成部 131 の一例を、図 4 を用いて説明する。図 4 は、実施の形態に係る出力音生成部の機能構成を示すブロック図である。図 4 に示すように、本実施の形態における出力音生成部 131 は、例えば、切り替え部 132、第 1 生成部 133、及び第 2 生成部 134 を備える。切り替え部 132 は出力音信号を生成する際に、第 1 生成部 133 を用いるか、第 2 生成部 134 を用いるかを切り替えるための処理部である。そのため、切り

替え部 1 3 2 には、第 1 生成部 1 3 3 を用いるか又は第 2 生成部 1 3 4 を用いるかを判定するための情報を取得する機能を有する。

[0118] 第 1 生成部 1 3 3 は、パニング処理を適用せずに、再生音に対してそのまま頭部伝達関数が畳み込まれる場合に用いられる処理部である。第 1 生成部 1 3 3 は、いわゆる「通常モード」で出力音信号を生成する場合に用いられる処理部である。第 1 生成部 1 3 3 は、再生音と、再生音の到来方向に対応する頭部伝達関数とを取得して、再生音に対して取得した頭部伝達関数の畳み込み処理を行い、出力音信号を生成する。

[0119] 第 2 生成部 1 3 4 は、パニング処理を適用して、再生音を代表音に変換する変換処理を行った後に、変換後の代表音に対して頭部伝達関数を畳み込む場合に用いられる処理部である。第 2 生成部 1 3 4 は、いわゆる「低処理モード」で出力音信号を生成する場合に用いられる処理部である。第 2 生成部 1 3 4 は、再生音及び代表点の位置を取得して、再生音を代表点からの音によって再現するための代表音への変換処理を行う。

[0120] 例えば、音源オブジェクトが 2 つの代表点の中間に位置する場合に、再生音と同じ音を 2 つの代表点のそれぞれから鳴らすように音を生成する。そして、音源オブジェクトの位置に合うように、生成した音のゲイン調整を行うことで、代表音を生成することができる。再生音からの代表音への変換は、このような例に限られない。例えば、後述するように時間シフト調整とゲイン調整とを行うことで再生音から代表音への変換を行ってもよいし、その他の、既存のあらゆる変換によって、再生音を代表点からの音によって再現するための代表音への変換を行うことができればよい。時間シフト調整とゲイン調整とを行う変換の例については後述する。第 2 生成部 1 3 4 は、変換によって得られた代表点の数と同じだけの代表音と、各代表点からユーザ 9 9 の位置までの代表方向に対応する頭部伝達関数とを取得して、代表音に対して取得した頭部伝達関数の畳み込み処理を行い、出力音信号を生成する。

[0121] 図 2 を再び参照する。出力音生成部 1 3 1 は、出力音信号生成のために用いる頭部伝達関数を記憶部 1 0 5 から取得する。記憶部 1 0 5 は情報を記憶

するための記憶装置としての機能と、記憶された情報を読み出して、情報処理装置に含まれる他の処理部へ出力する記憶コントローラとしての機能とを併せ持つ情報記憶装置である。記憶部105は情報処理装置101に備わるメモリと読み替えてもよい。記憶部105には、取得部111で取得した頭部伝達関数がユーザ99への到来方向ごとに記憶されている。記憶部105に含まれる頭部伝達関数は、万人に用いることができる汎用の頭部伝達関数のセット、又は、ユーザ99個人に最適化された頭部伝達関数のセット、又は、一般に公開されている頭部伝達関数のセットである。記憶部105は、出力音生成部131から、到来方向をクエリとした問い合わせを受け、その到来方向に対応する頭部伝達関数を出力音生成部131へと出力する。また、出力音生成部131は、切り替え部132からの問い合わせを受けて、頭部伝達関数のセットをすべて出力したり、頭部伝達関数のセット自体の特性などを出力したりする場合もある。頭部伝達関数のセットは、取得部111において、例えばS O F Aファイルの形式で外部から取得され、その後記憶部105で保存されてもよい。

[0122] 信号出力部141は、生成された出力音信号をドライバ104へと出力する機能部である。信号出力部141は、出力音信号に基づいてデジタル信号からアナログ信号への信号変換などを行うことで、波形信号を生成し、波形信号に基づいてドライバ104に音波を発生させ、ユーザ99に音を提示する。ドライバ104は、例えば、振動板とマグネット及びボイスコイルなどの駆動機構とを有する。ドライバ104は、波形信号に応じて駆動機構を動作させ、駆動機構によって振動板を振動させる。このようにして、ドライバ104は、出力音信号に応じた振動板の振動により、音波を発生させ（出力音信号を「再生」することを意味する、すなわち、ユーザ99が知覚することは「再生」の意味には含まれない）、音波が空気を伝播してユーザ99の耳に伝達し、ユーザ99が音を知覚する。

[0123] [動作]

次に、図5～図8を参照して、上記に説明した音響再生システム100の

特に情報処理装置101の動作について説明する。

- [0124] まず、図5は、実施の形態に係る情報処理装置の第1の動作例を示すフローチャートである。情報処理装置101の第1の動作例では、取得部111が通信モジュール102を介して音情報を取得する（ステップS11）。音情報は、デコード処理部113によって再生音に関する情報と、音源オブジェクトの位置に関する情報とフラグとにデコードされ、出力音信号の生成が開始される。
- [0125] センシング情報入力部114は、ユーザ99の位置に関する情報を取得する（ステップS12）。経路算出部121は、音源オブジェクトの位置及びユーザ99の位置に基づいて再生音の到来方向を算出する（ステップS13）。ここで、音情報に含まれるフラグは、音情報を作成した際に、作成者によって付されるフラグである。そしてこのフラグは、第1生成部133によって出力音信号を生成させるか、第2生成部134によって出力音信号を生成させるかを指定するためのフラグである。作成者は、元の音情報にどのような音源オブジェクトが含まれているかを把握しているので、例えば、音情報に含まれている音源オブジェクトの数がかなり少ないなどの理由から第1生成部133によって出力音信号を生成させるフラグを付すことが可能である。
- [0126] あるいは、作成者は、音情報に含まれている音源オブジェクトの数がかなり多いなどの理由から第2生成部134によって出力音信号を生成させるフラグを付すことが可能である。フラグは、第1生成部133によって出力音信号を生成させることを指定していれば、第2生成部134によって出力音信号を生成させないことを指定しているものと同等に扱われてもよい。また、フラグは、第2生成部134によって出力音信号を生成させることを指定していれば、第1生成部133によって出力音信号を生成させないことを指定しているものと同等に扱われてもよい。
- [0127] 図5の動作例では、フラグに第1生成部133を用いることが指定されているか否かの判定が行われる（ステップS14）。そして、フラグに第1生

成部133を用いることが指定されていれば（S14でYes）、第1生成部133により出力音信号を生成する（ステップS15）。一方で、フラグに第1生成部133を用いることが指定されていなければ（S14でNo）、第2生成部134により出力音信号を生成する（ステップS16）。なお、出力音生成部131は、切り替え部132を用いて、ステップS14の判定及び、第1生成部133を用いて出力音信号を生成させるか、又は、第2生成部134を用いて出力音信号を生成させるかの切り替えを行ってもよいし、図示しないフラグ判定部がステップS14の判定を行い、判定結果に応じて、取得部111が直接第1生成部133に音情報を入力したり、第2生成部134に音情報を入力したりしてもよい。つまり、第1生成部133を用いて出力音信号を生成させるか、又は、第2生成部134を用いて出力音信号を生成させるかの切り替えは、必須ではない。

[0128] 次に、図6は、実施の形態に係る情報処理装置の第2の動作例を示すフローチャートである。図6に示す動作例では、ステップS14の代わりにステップS24が実行される他は、図5と同様であるため説明を省略する。図6の動作例では、ステップS13の後に、音情報に含まれる音源オブジェクトの数と、三次元音場内に設定された代表点の数とを比較し、比較結果が所定の条件を満たすか否かに応じて、ステップS15を実行するか、ステップS16を実行するかが切り替わる。具体的には、切り替え部132は、音情報を取得して、音源オブジェクトの数を計数する。

[0129] また、切り替え部132は、三次元音場内に設定された代表点の数を取得する（代表点の数は、図示しない記憶部などに設定情報として記憶されている）。そして、切り替え部132は、音源オブジェクトの数と代表点の数とを比較する。切り替え部132は比較結果が、例えば、所定の条件として音源オブジェクトの数が代表点の数係数倍未満であるか否かにより、比較結果が所定の条件を満たすか否かを判定する（ステップS24）。所定の条件を満たす（音源オブジェクトの数が代表点の数係数倍未満である）場合は（S24でYes）、切り替え部132は、ステップS15が実行されるように

切り替える。また、所定の条件を満たさない（音源オブジェクトの数が代表点の数係数倍以上である）場合は（S 24でNo）、切り替え部132は、ステップS 16が実行されるように切り替える。ここでの係数倍は、すなわち、通常モードでの出力音信号の生成が、低処理モードでの出力音信号の生成よりも処理量の観点で同等又は有利である場合を想定して設定される。上記に説明したように、パニング処理は、それ自体の処理量を有するので、係数は、実装されるパニング処理に応じて、1倍、3倍、5倍などの数倍から、10倍、30倍、50倍などの数十倍程度の範囲内で変化する。つまり、係数としては、パニング処理の態様に応じた数値が適宜設定されればよい。

[0130] また、図7に示すように、三次元音場（図中の最外部の矩形）において、白抜き円形で示す音源オブジェクトから再生音が、黒塗り円形で示すユーザ99の位置に到来するとき、再生音が直接到来する直接音の他に、間接的な伝播によって生成される反射音や回折音（逆三角形で示す空間内のオブジェクトの影響）、又は、残響音（不図示）などが併せて発生する。このとき、反射音、回折音及び残響音など、互いに異なる間接的な伝播によって、再生音があらゆる方向から到来することになる。そのため、再生音に対して、出力音信号を生成すると不自然に聞こえる場合があるため、本実施の形態では、間接的な伝播によって到来する再生音を合成音として生成する。このような合成音も適切な到来方向からの音として知覚される必要があり、音源オブジェクトからの再生音と同様に出力音信号に含まれる必要がある。つまり、再生音と合成音との組、又は、互いに異なる間接的な伝播による合成音と合成音との組について、パニング処理をすべきか否か判断すべきである。そこで、図6に戻り、ステップS 24で代表点の数と比較する対象として、再生音と合成音との合計数を用いてもよい。この場合も、係数倍などの所定の条件を満たすか否かの判定によって、ステップS 15を実行するか、ステップS 16を実行するかが切り替わる。

[0131] 次に、図8は、実施の形態に係る情報処理装置の第3の動作例を示すフローチャートである。図8に示す動作例では、ステップS 14の代わりにステ

ップS 3 4 が実行される他は、図 5 と同様であるため説明を省略する。図 8 の動作例では、ステップ S 1 3 の後に、記憶部 1 0 5 に含まれる頭部伝達関数がパニング処理の適用による処理量削減効果を十分に発揮できる程度に密になっているか否かに応じて、ステップ S 1 5 を実行するか、ステップ S 1 6 を実行するかが切り替わる。具体的には、切り替え部 1 3 2 は、記憶部 1 0 5 への問い合わせを行い、頭部伝達関数のセットを読み出すか、又は、頭部伝達関数の疎密に関する特性情報を読み出す。そして、切り替え部 1 3 2 は、記憶部 1 0 5 に含まれる頭部伝達関数が、予め設定された閾値よりも疎であるか、密であるかを判定する。

[0132] つまり、切り替え部 1 3 2 は、頭部伝達関数の疎密に関する特性が、疎密に関する閾値よりも密であるか否かにより、所定の条件を満たすか否かを判定する（ステップ S 3 4）。所定の条件を満たさない（頭部伝達関数の疎密に関する特性が閾値よりも疎である）場合は（S 3 4 で Y e s）、切り替え部 1 3 2 は、ステップ S 1 5 が実行されるように切り替える。また、所定の条件を満たす（頭部伝達関数の疎密に関する特性が閾値よりも密である）場合は（S 3 4 で N o）、切り替え部 1 3 2 は、ステップ S 1 6 が実行されるように切り替える。この疎密に関する閾値は、例えば、水平方向に 5 度、1 0 度、1 5 度、及び、垂直方向に 5 度、1 0 度、1 5 度などの少なくとも一方の方向に対する刻み角よりも密で頭部伝達関数が含まれているか否かなどのようにして設定される。疎密に関する閾値は、音情報に含まれる音源オブジェクトからの再生音の到来方向にも依存するし、設定された代表点からの代表方向にも依存する。そのため、音情報に含まれる音源オブジェクトからの再生音の到来方向及び代表点からの代表方向に応じて適宜設定されればよい。

[0133] [パニング処理の具体例]

パニング処理では、複数個の音源オブジェクトからの再生音を、複数の代表方向からの代表音によって表現する。この代表方向には、例えば、2～3 方向を用いることが可能である。具体的には、パニング処理では、音源オブ

ジェクトの個数より少ない個数の代表点にまとめ、この代表点に対する代表方向の頭部伝達関数のみで再生音を到来方向からの音として知覚させることが可能である。また、パニング処理は、再生音を代表点（代表方向）に分配する処理であると読み換えてもよい。具体的には、それぞれの音源オブジェクトの位置に紐づけられた再生音の音信号を、代表点の位置に分配し、代表点（代表方向）からリスナへ到来する代表音を生成する。ここで代表方向は、リスナの頭部方向と代表点の位置の関係で定められる方向である。例えば、リスナの正面から見た代表点の方向のことを指す。また、リスナの顔の正面が向いている方向を基準としたときの代表点の方向、リスナの目から見た代表点の方向と言い換えてもよい。

[0134] この際、パニング処理では、音源オブジェクトからの到来方向の頭部伝達関数と代表方向の頭部伝達関数との相互相関が最大になる時間シフト（ディレイ、時間遅延）を算出する。ここで得られた時間シフト、又はこの時間シフトに負号を付した時間シフトを音源オブジェクトの再生音に付与した、時間シフト後の信号が代表方向にあるものとして、以降の処理を行う。

[0135] この時間シフトは、サンプリング周期より短い時間での時間シフト（サンプル位置が小数で示されるシフト。以下、「小数シフト」という。）も許容してもよい。この小数シフトは、オーバーサンプリングにより行うことが可能である。

[0136] ここで、パニング処理では、音源オブジェクトの再生音を時間シフトした代表方向の信号にゲインをかけて、代表点毎に算出されたそれらの信号に各代表点における頭部伝達関数を畳み込んだものの和を算出することで、音源オブジェクトの再生音に到来方向の頭部伝達関数を畳み込んだものと等価な信号を合成する。

[0137] 一方、パニング処理では、代表方向の頭部伝達関数（ベクトル）の和で到来方向の頭部伝達関数（ベクトル）を合成する際、合成された頭部伝達関数（ベクトル）と到来方向の頭部伝達関数（ベクトル）の誤差信号ベクトルが代表方向の頭部伝達関数（ベクトル）と直交させるようにして、ゲインを算

出してもよい。なお、頭部伝達関数（ベクトル）とは頭部伝達関数の時間領域での表現である頭部インパルス応答の時間波形をベクトルと見立てたものである。以下、この頭部伝達関数（ベクトル）を、単に「頭部伝達関数ベクトル」とも記載する。

[0138] パニング処理では、このゲインについて、音源オブジェクトの位置からのユーザ99の左右の耳までの頭部伝達関数のエネルギーバランスが、パニング処理により複数の代表点からの頭部伝達関数で合成された頭部伝達関数でも維持されるように補正する。すなわち、パニング処理では、音源オブジェクトによるユーザ99の左右の耳の頭部伝達関数のエネルギーバランスが、パニング処理により合成された頭部伝達関数でも維持されるようにゲインを補正してもよい。

[0139] 本実施形態においては、パニング処理では、音源オブジェクトの各到来方向について、代表方向の頭部伝達関数に乗ずるゲイン値と、代表方向の頭部伝達関数に施す時間シフト値とを算出して、後述するテーブルデータ（頭部伝達関数テーブルまたは調整量テーブル）に格納しておくことが可能である。

[0140] この上で、パニング処理では、各音源オブジェクトの到来方向に対応する時間シフト値及びゲイン値で、各音源オブジェクトの時間シフトを行い、ゲインをかけて、これの和をとって和信号とする。パニング処理では、この和信号が代表点の位置に存在するものとして扱う。パニング処理では、この和信号に、代表点の方向の頭部伝達関数を畳み込んで、ユーザ99の耳元の信号を生成することが可能である。

[0141] また、パニング処理には、合成されたHRIRベクトルと音源方向のHRIRベクトルとの誤差信号ベクトルのエネルギー又はL2ノルムを最小化するようにして算出されたゲインを用いてもよい。HRIRベクトルは、頭部伝達関数の時間領域の波形を48kHzのサンプリング周波数でサンプルした値を要素に持つベクトルである。

[0142] また、上記においては、相互相関を最大化する時間シフト及びゲインの算

出時に、頭部伝達関数そのものを用いている例について記載した。一方で、時間シフト及び／又はゲインは、周波数軸上の重み付けフィルタをかけてから相互相関が算出されたものを用いてもよい。

[0143] すなわち、相互相関を最大化する時間シフトおよびゲインの算出時に、周波数軸上の重み付けフィルタ（以下、「周波数重み付けフィルタ」ともいう。）をかけたものを用いることが可能である。

[0144] この周波数重み付けフィルタは、ヒトの聴感の感度が高い周波数帯域近傍かそれよりやや高い周波数をカットオフ周波数として、それより高い帯域、すなわちヒトの聴感の感度が低くなっていく帯域を減衰させるようなフィルタを用いることが好適である。たとえば、カットオフ周波数を3000Hz～6000Hz、6db/oct（オクターブ）～12db/oct程度のローパスフィルタ（LPF）を用いることが好適である。

[0145] なお、パニング処理においては、記憶部105に含まれる頭部伝達関数のセットに応じて、時間シフト調整及びゲイン調整における調整量を決定し、再生音に対して、決定した調整量で時間シフト調整とゲイン調整とを適用して代表音に変換してもよい。頭部伝達関数に応じて、パニング処理に用いられる時間シフト調整及びゲイン調整における調整量は、最適な値が変化するため、まず、SOF Aファイルなどの頭部伝達関数のセットを取得した際、または記憶部105に含まれている頭部伝達関数のセットを読み出した際に、頭部伝達関数のセットに含まれる頭部伝達関数それぞれに合わせた時間シフト調整及びゲイン調整における調整量を決定することで、以降、この頭部伝達関数のセットを用いる限りにおいて、同じ調整量を流用できるので、処理量の観点で有利である。より具体的には、パニング処理に例えば3方向の代表方向を用いる場合、まず、頭部伝達関数のセットを取得したとき（例えば初期化時）に、頭部伝達関数のセットに含まれる全天球の方向から複数の代表方向候補（例えば8方向）を選択する。次に、頭部伝達関数のセットに含まれる全天球の頭部伝達関数それぞれについて、複数の代表方向候補のうちいずれの3方向を代表方向として用いるかを決定する。次に、頭部伝達関

数のセットに含まれる全天球方向それぞれについて、特定した3つの代表方向に信号を振り分けるための時間シフト調整とゲイン調整における調整量を算出する。そして、算出した調整量を、頭部伝達関数のセットに含まれる全天球方向のそれぞれに紐づく調整量として決定する。

[0146] 頭部伝達関数テーブルは、記憶部105に記憶された、頭部伝達関数を含むテーブルデータの一例であるが、頭部伝達関数テーブルには、頭部伝達関数とともに、その頭部伝達関数に応じて決定した時間シフト調整及びゲイン調整における調整量が互いに紐づけられて格納されている。つまり、記憶部105に含まれる頭部伝達関数ごとに、時間シフト調整とゲイン調整における調整量を予め算出して頭部伝達関数テーブルを構築しておいてもよい。このように、それぞれの頭部伝達関数と当該調整量とを紐づけた頭部伝達関数テーブルのテーブルデータを、記憶部105に記憶しておいてもよい。なお、頭部伝達関数ごとの調整量の算出は、第2生成部134、またはデコード処理部113で行ってもよい。または、外部の装置で調整量の算出を行い、外部の装置のメモリに記憶しておいてもよい。この場合は、外部の装置のメモリが記憶部の一例に相当する。

[0147] また、時間シフト調整とゲイン調整における調整量を、頭部伝達関数のセットに含まれる複数の頭部伝達関数のそれぞれの方向ごとに予め算出して、複数の代表方向のそれぞれと、頭部伝達関数のセットに含まれる複数の頭部伝達関数のそれぞれの方向ごとの調整量とを紐づけた調整量テーブルを構築し記憶部105に記憶しておいてもよい。このとき、複数の代表方向は複数の代表方向候補（例えば8方向）から選択された代表方向（例えば3方向）であることから、調整量テーブルには、頭部伝達関数のセットに含まれる全天球の方向ごとに、複数の代表方向候補（例えば8方向）からいずれの代表方向（例えば3方向）を選択したかの情報を含む。調整量テーブルは複数の代表方向それぞれの頭部伝達関数と、頭部伝達関数のセットに含まれる複数の頭部伝達関数のそれぞれの方向ごとの時間シフト調整とゲイン調整における調整量とを紐づけたテーブルデータを含んでいてもよいし、複数の代表方

向それぞれの頭部伝達関数は、予め取得し記憶部105に記憶している全天球（複数方向）の頭部伝達関数のセットからレンダリング時またはシステムの初期化時に抽出してもよい。また、システムの初期化時に外部から頭部伝達関数のセットを取得し、初期化時に調整量テーブルを構築した後、記憶部105に記憶しておいてもよい。その場合、音声信号の出力処理時に記憶部105に記憶した調整量テーブルを読みだして使用してもよい。

[0148] なお、本開示における初期化時及び出力処理時について、以下に説明する。

[0149] 例えば、実施の形態において、空間情報の更新処理（情報更新スレッド）と、音響処理を付加した音声信号の出力処理（音声スレッド）とは、1つのスレッドで実行されてもよく、異なるスレッドで実行されてもよい。これらの2つの処理を異なるスレッドで実行する場合、スレッドの起動頻度を個々に設定してもよいし、並行して処理を実行してもよい。

[0150] 特に、これらの2つの処理を異なる独立したスレッドで実行する場合、音響処理を付加した音声信号の出力処理に優先的に演算資源を割り当てることが可能である。これにより、僅かな遅延も許容されないような、例えば、1サンプル（0.02 msec）の遅延でプチっというノイズが発生するような出音処理を安全に実行することが可能である。

[0151] その際、空間情報の更新処理には、演算資源の割り当てが制限される。しかし、空間情報の更新は、音声信号の出力処理と比較して、低頻度の処理（例えば、受聴者の顔の向きへの更新のような処理）であるため、音声信号の出力処理のように必ずしも遅延のないように略リアルタイムで行われなくてもよい。したがって、演算資源の割り当てが制限されても、音響的な品質に大きな影響はない。

[0152] 空間情報の更新は、予め設定された時間又は期間毎に定期的に行われてもよいし、予め設定された条件が満たされた場合に実行されてもよい。また、空間情報の更新は、受聴者又は音空間の管理者によって手動で実行されてもよいし、外部システムの変化をトリガとして実行されてもよい。

- [0153] 例えば、受聴者によってコントローラが操作され、受聴者自身のアバターの立ち位置が瞬間的にワープしたり、時刻が瞬時に進められたり戻されたりした場合に、空間情報が更新されてもよい。あるいは、仮想空間の管理者によって、突如、場の環境を変更するような演出が施された場合、空間情報が更新されてもよい。これらの場合において、空間情報の更新処理のためのスレッドが、定期的な起動に加えて、単発的な割り込み処理として起動されてもよい。
- [0154] 例えば、空間情報の更新処理は、仮想空間作成時（ソフトウェアの作成時）、仮想空間の情報（シーン情報）の読み込み時、仮想空間の処理の開始時（ソフトウェアの起動時又はレンダリング開始時）、又は、仮想空間の処理において定期的に発生する情報更新スレッドが発生したタイミング等で行われてもよい。また、仮想空間の作成タイミングは、音響処理の開始前に仮想空間を構築するタイミングであってもよいし、仮想空間の情報（空間情報）の取得時であってもよいし、ソフトウェアの取得時であってもよい。
- [0155] このように、本開示においては、イレギュラーに発生する処理スレッド、受聴者の顔の向きの変更などの低頻度で定期的に発生する処理スレッド、及び、出音処理などの高頻度で定期的に発生する処理スレッドのように異なる頻度で発生する3つの処理スレッド（言い換えるとワークフロー）が存在する。本開示における初期化時の処理は、上記の内、イレギュラーに発生する処理スレッドに該当する。
- [0156] また、調整量テーブルは、全天球の頭部伝達関数のセットの各頭部伝達関数の方向から受聴者の位置に到来する音信号に対して、例えば複数の代表方向のうちいずれの代表方向にその信号を分配するかの情報と、分配する際の代表方向毎に音声信号に乗じる、時間シフト調整量とゲイン調整量の情報とを含むテーブルであってもよい。
- [0157] 頭部伝達関数の畳み込みの処理を行う際には、記憶部105に記憶された調整量テーブルを参照し、適用する方向の頭部伝達関数に紐づけられた、時間シフト調整とゲイン調整における調整量を用いることで、畳み込みの処理

ごとに調整量を算出する必要がなく、処理量の削減に寄与できる。

[0158] なお、本発明の実施形態は記憶部105に含まれない新たな頭部伝達関数のセット（例えば、S O F Aファイル）に適用することも可能である。音信号のデコード時や音響再生システム100の電源投入時、あるいは音響再生システム100の初期化時に、三次元音場全体の頭部伝達関数を新たに読み込み、本実施形態で開示した手法または別の手法で、改めて代表方向を決定し、新たな頭部伝達関数のセットに含まれる頭部伝達関数ごとの調整量の算出を行ってもよい。その場合、頭部伝達関数と当該調整量とを紐づけたテーブルデータを、記憶部105に記憶しておいてもよい。または、外部の装置で調整量の算出を行い、外部の装置のメモリに記憶しておいてもよい。頭部伝達関数の読み込みの処理を行う際には、適用する頭部伝達関数に紐づけられた時間シフト調整とゲイン調整における調整量を参照することで、読み込みの処理ごとに調整量を算出する必要がなく、処理量の削減に寄与できる。

[0159] このように、記憶部105に記憶されていない新たな頭部伝達関数のセットを読み込んだ場合に、記憶部105に記憶する前に当該新たな頭部伝達関数のセットに対して、パニング処理に用いられる時間シフト調整及びゲイン調整における調整量を決定し、新たな頭部伝達関数のセットと、決定した調整量とを紐づけて頭部伝達関数テーブルを構築し、頭部伝達関数テーブルを記憶部105に記憶させてもよい。そして、パニング処理を行う際には、記憶部105からこの調整量が読み出され、その調整量によって、シフト調整及びゲイン調整が適用される。なお、新たな頭部伝達関数は、それ以前に記憶部105に記憶されていたものが、音信号のデコード時、音響再生システム100の電源投入時、又は、音響再生システム100の初期化時等に一時的に記憶部105から取り除かれ、再び記憶部105に記憶しなおされたものであってもよい。または、新たな頭部伝達関数のセット用のテーブルデータと、それ以前に記憶部105に記憶されていたテーブルデータのそれぞれを、別の頭部伝達関数のセットに対応するテーブルデータとして記憶部105に記憶しておいてもよい。なお、ここで記憶部105に記憶するテーブル

データは頭部伝達関数テーブルでもよいし、頭部伝達関数テーブルの一部である調整量テーブルであってもよいことは言うまでもない。また、このような調整量の決定は、パニング処理を行うか否かの切り替えを行わずとも効果的である。すなわち、第1生成部133及び第2生成部134とは異なる第3生成部であって、再生音に対して、記憶部105に記憶された新たな頭部伝達関数に紐づけられた調整量で時間シフト調整とゲイン調整とを適用して代表音に変換し、代表点のそれぞれの位置からユーザの位置に向かう代表方向に応じた頭部伝達関数を代表音に畳み込むことで出力音信号を生成する第3生成部を、第1生成部133及び第2生成部134の代わりに備えてもよい。

[0160] (その他の実施の形態)

以上、実施の形態について説明したが、本開示は、上記の実施の形態に限定されるものではない。

[0161] 例えば、上記の実施の形態に説明した音響再生システムは、構成要素をすべて備える一つの装置として実現されてもよいし、複数の装置に各機能が割り振られ、この複数の装置が連携することで実現されてもよい。後者の場合には、情報処理装置に該当する装置として、スマートフォン、タブレット端末、又は、PCなどの情報処理装置が用いられてもよい。例えば、音響効果を付加した音響信号を生成するレンダラとしての機能を有する音響再生システム100において、レンダラの機能のすべて又は一部をサーバが担ってもよい。つまり、取得部111、経路算出部121、出力音生成部131、信号出力部141のすべて又は一部は、図示しないサーバに存在してもよい。その場合、音響再生システム100は、例えば、コンピュータ又はスマートフォンなどの情報処理装置と、ユーザ99に装着されるヘッドマウントディスプレイ(HMD)やイヤホンなどの音提示デバイスと、図示しないサーバとを組み合わせ実現される。なお、コンピュータと音提示デバイスとサーバとが同一のネットワークで通信可能に接続されていてもよいし、異なるネットワークで接続されていてもよい。異なるネットワークで接続されている

場合、通信に遅延が発生する可能性が高くなるため、コンピュータと音提示デバイスとサーバとが同一ネットワークで通信可能に接続されている場合にのみサーバでの処理を許可してもよい。また、音響再生システム100が受け付けるビットストリームのデータ量に応じて、レンドラのすべて又は一部の機能をサーバが担うか否かを決定してもよい。

[0162] また、本開示の音響再生システムは、ドライバのみを備える再生装置に接続され、当該再生装置に対して、取得した音情報に基づいて生成された出力音信号を再生するのみの情報処理装置として実現することもできる。この場合、情報処理装置は、専用の回路を備えるハードウェアとして実現してもよいし、汎用のプロセッサに特定の処理を実行させるためのソフトウェアとして実現してもよい。

[0163] また、上記の実施の形態において、特定の処理部が実行する処理を別の処理部が実行してもよい。また、複数の処理の順序が変更されてもよいし、複数の処理が並行して実行されてもよい。

[0164] また、上記の実施の形態において、各構成要素は、各構成要素に適したソフトウェアプログラムを実行することによって実現されてもよい。各構成要素は、CPU又はプロセッサなどのプログラム実行部が、ハードディスク又は半導体メモリなどの記録媒体に記録されたソフトウェアプログラムを読み出して実行することによって実現されてもよい。

[0165] また、各構成要素は、ハードウェアによって実現されてもよい。例えば、各構成要素は、回路（又は集積回路）でもよい。これらの回路は、全体として1つの回路を構成してもよいし、それぞれ別々の回路でもよい。また、これらの回路は、それぞれ、汎用的な回路でもよいし、専用の回路でもよい。

[0166] また、本開示の全般的又は具体的な態様は、装置、装置、方法、集積回路、コンピュータプログラム又はコンピュータ読み取り可能なCD-ROMなどの記録媒体で実現されてもよい。また、本開示の全般的又は具体的な態様は、装置、装置、方法、集積回路、コンピュータプログラム及び記録媒体の任意な組み合わせで実現されてもよい。

[0167] 例えば、本開示は、コンピュータによって実行される音声信号再生方法として実現されてもよいし、音声信号再生方法コンピュータに実行させるためのプログラムとして実現されてもよい。本開示は、このようなプログラムが記録されたコンピュータ読み取り可能な非一時的な記録媒体として実現されてもよい。

[0168] その他、各実施の形態に対して当業者が思いつく各種変形を施して得られる形態、又は、本開示の趣旨を逸脱しない範囲で各実施の形態における構成要素及び機能を任意に組み合わせることで実現される形態も本開示に含まれる。

[0169] なお、本開示における符号化された音情報は、音響再生システム100によって再生される所定音についての情報である音信号及び、当該所定音の音像を三次元音場内において所定位置に定位させる際の定位位置に関する情報であるメタデータを含むビットストリームと言い換えることができる。例えばMPEG-H 3D Audio (ISO/IEC 23008-3)等の所定の形式で符号化されたビットストリームとして音情報が音響再生システム100に取得されてもよい。一例として、符号化された音信号は、音響再生システム100によって再生される所定音についての情報を含む。ここでいう所定音は、三次元音場に存在する音源オブジェクトが発する音又は自然環境音であって、例えば、機械音、又は人を含む動物の音声等を含み得る。なお、三次元音場に音源オブジェクトが複数存在する場合、音響再生システム100は、複数の音源オブジェクトにそれぞれ対応する複数の音信号を取得することになる。

[0170] 一方、メタデータとは、例えば、音響再生システム100において音信号に対する音響処理を制御するために用いられる情報である。メタデータは、仮想空間（三次元音場）で表現されるシーンを記述するために用いられる情報であってもよい。ここでシーンとは、メタデータを用いて、音響再生システム100でモデリングされる、仮想空間における三次元映像及び音響イベントを表す全ての要素の集合体を指す用語である。つまり、ここでいうメタ

データとは、音響処理を制御する情報だけでなく、映像処理を制御する情報も含んでいてもよい。もちろん、メタデータには、音響処理と映像処理とのいずれか一方だけを制御する情報が含まれていてもよいし、両方の制御に用いられる情報が含まれていてもよい。本開示において音響再生システム100が取得するビットストリームには、このようなメタデータが含まれている場合がある。あるいは、音響再生システム100は、後述するようにビットストリームとは別に、メタデータを単体で取得してもよい。

[0171] 音響再生システム100は、ビットストリームに含まれるメタデータ、及び追加で取得されるインタラクティブなユーザ99の位置情報等を用いて、音信号に音響処理を行うことで、仮想的な音響効果を生成する。例えば、初期反射音生成、後期残響音生成、回折音生成、距離減衰効果、ローカリゼーション、音像定位処理、又はドップラー効果等の音響効果が付加されることが考えられる。また、音響効果の全て又は一部のオンオフを切り替える情報がメタデータとして付加されてもよい。

[0172] なお、全てのメタデータ又は一部のメタデータは、音情報のビットストリーム以外から取得されてもよい。例えば、音響を制御するメタデータと映像を制御するメタデータとのいずれかがビットストリーム以外から取得されてもよいし、両方のメタデータがビットストリーム以外から取得されてもよい。

[0173] また、映像を制御するメタデータが音響再生システム100で取得されるビットストリームに含まれる場合は、音響再生システム100は映像の制御に用いることができるメタデータを、画像を表示する表示装置、又は立体映像を再生する立体映像再生装置に対して出力する機能を備えていてもよい。

[0174] また、一例として、符号化されたメタデータは、音を発する音源オブジェクト、及び障害物オブジェクトを含む三次元音場に関する情報と、当該音の音像を三次元音場内において所定位置に定位させる（つまり、所定方向から到達する音として知覚させる）際の定位位置に関する情報、すなわち所定方向に関する情報とを含む。ここで、障害物オブジェクトは、音源オブジェク

トが発する音がユーザ99へと到達するまでの間において、例えば音を遮ったり、音を反射したりして、ユーザ99が知覚する音に影響を及ぼし得るオブジェクトである。障害物オブジェクトは、静止物体の他に、人等の動物、又は機械等の動体を含み得る。また、三次元音場に複数の音源オブジェクトが存在する場合、任意の音源オブジェクトにとっては、他の音源オブジェクトは障害物オブジェクトとなり得る。また、建材又は無生物等の非発音源オブジェクトも、音を発する音源オブジェクトも、いずれも障害物オブジェクトとなり得る。

[0175] メタデータを構成する空間情報として、三次元音場の形状だけでなく、三次元音場に存在する障害物オブジェクトの形状及び位置と、三次元音場に存在する音源オブジェクトの形状及び位置とをそれぞれ表す情報が含まれていてもよい。三次元音場は、閉空間又は開空間のいずれであってもよく、メタデータには、例えば床、壁、又は天井等の三次元音場において音を反射し得る構造物の反射率、及び三次元音場に存在する障害物オブジェクトの反射率を表す情報が含まれる。ここで、反射率は、入射音に対する反射音のエネルギーの比であって、音の周波数帯域ごとに設定されている。もちろん、反射率は、音の周波数帯域に依らず、一律に設定されていてもよい。また、三次元音場が開空間の場合は、例えば一律で設定された減衰率、回折音、又は初期反射音等のパラメータが用いられてもよい。

[0176] 上記説明では、メタデータに含まれる障害物オブジェクト又は音源オブジェクトに関するパラメータとして反射率が挙げられたが、メタデータは、反射率以外の情報を含んでいてもよい。例えば、音源オブジェクト及び非発音源オブジェクトの両方に関わるメタデータとして、オブジェクトの素材に関する情報が含まれていてもよい。具体的には、メタデータは、拡散率、透過率、又は吸音率等のパラメータを含んでいてもよい。

[0177] 音源オブジェクトに関する情報として、音量、放射特性（指向性）、再生条件、ひとつのオブジェクトから発せられる音源の数及び種類、又はオブジェクトにおける音源領域を指定する情報等が含まれてもよい。再生条件では

、例えば、継続的に流れ続ける音なのかイベント発動する音なのかが定められてもよい。オブジェクトにおける音源領域は、ユーザ99の位置とオブジェクトの位置との相対的な関係で定められてもよいし、オブジェクトを基準として定められてもよい。ユーザ99の位置とオブジェクトの位置との相対的な関係で定められる場合、ユーザ99がオブジェクトを見ている面を基準とし、ユーザ99から見てオブジェクトの右側からは音X、左側からは音Yが発せられているようにユーザ99に知覚させることができる。オブジェクトを基準として定められる場合、ユーザ99の見ている方向に関わらず、オブジェクトのどの領域からどの音を出すかは固定にすることができる。例えばオブジェクトを正面から見たときの右側からは高い音、左側からは低い音が流れているようにユーザ99に知覚させることができる。この場合、ユーザ99がオブジェクトの背面に回り込むと、背面から見て右側からは低い音、左側からは高い音が流れているようにユーザ99に知覚させることができる。

[0178] 空間に関するメタデータとして、初期反射音までの時間、残響時間、又は直接音と拡散音との比率等を含めることができる。直接音と拡散音との比率がゼロの場合、直接音のみをユーザ99に知覚させることができる。

[0179] また、三次元音場におけるユーザ99の位置及び向きを示す情報が初期設定として予めメタデータとしてビットストリームに含まれていてもよいし、ビットストリームに含まれていなくてもよい。ユーザ99の位置及び向きを示す情報がビットストリームに含まれていない場合、ユーザ99の位置及び向きを示す情報はビットストリーム以外の情報から取得される。例えば、VR空間におけるユーザ99の位置情報であれば、VRコンテンツを提供するアプリから取得されてもよいし、ARとして音を提示するためのユーザ99の位置情報であれば、例えば携帯端末がGPS、カメラ、又はLiDAR (Laser Imaging Detection and Ranging) 等を用いて自己位置推定を実施して得られた位置情報が用いられてもよい。なお、音信号とメタデータとは、一つのビットストリームに格納されて

いてもよいし、複数のビットストリームに別々に格納されていてもよい。同様に、音信号とメタデータとは、一つのファイルに格納されていてもよいし、複数のファイルに別々に格納されていてもよい。

[0180] 音信号とメタデータとが複数のビットストリームに別々に格納されている場合、関連する他のビットストリームを示す情報が、音信号とメタデータとが格納された複数のビットストリームのうちの一つ又は一部のビットストリームに含まれていてもよい。また、関連する他のビットストリームを示す情報が、音信号とメタデータとが格納された複数のビットストリームの各ビットストリームのメタデータ又は制御情報に含まれていてもよい。音信号とメタデータとが複数のファイルに別々に格納されている場合、関連する他のビットストリーム又はファイルを示す情報が、音信号とメタデータとが格納された複数のファイルのうちの一つ又は一部のファイルに含まれていてもよい。また、関連する他のビットストリーム又はファイルを示す情報が、音信号とメタデータとが格納された複数のビットストリームの各ビットストリームのメタデータ又は制御情報に含まれていてもよい。

[0181] ここで、関連するビットストリーム又はファイルはそれぞれ、例えば、音響処理の際に同時に用いられる可能性のあるビットストリーム又はファイルである。また、関連する他のビットストリームを示す情報は、音信号とメタデータとを格納した複数のビットストリームのうちの一つのビットストリームのメタデータ又は制御情報にまとめて記述されていてもよいし、音信号とメタデータとを格納した複数のビットストリームのうちの二以上のビットストリームのメタデータ又は制御情報に分割して記述されていてもよい。同様に、関連する他のビットストリーム又はファイルを示す情報は、音信号とメタデータとを格納した複数のファイルのうちの一つのファイルのメタデータ又は制御情報にまとめて記述されていてもよいし、音信号とメタデータとを格納した複数のファイルのうちの二以上のファイルのメタデータ又は制御情報に分割して記述されていてもよい。また、関連する他のビットストリーム又はファイルを示す情報を、まとめて記述した制御ファイルが音信号とメタ

データとを格納した複数のファイルとは別に生成されてもよい。このとき、制御ファイルは音信号とメタデータとを格納していなくてもよい。

[0182] ここで、関連する他のビットストリーム又はファイルを示す情報とは、例えば当該他のビットストリームを示す識別子、他のファイルを示すファイル名、URL (Uniform Resource Locator)、又はURI (Uniform Resource Identifier) 等である。この場合、取得部111は、関連する他のビットストリーム又はファイルを示す情報に基づいて、ビットストリーム又はファイルを特定又は取得する。また、関連する他のビットストリームを示す情報が音信号とメタデータとを格納した複数のビットストリームのうちの少なくとも一部のビットストリームのメタデータ又は制御情報に含まれていると共に、関連する他のファイルを示す情報が音信号とメタデータとを格納した複数のファイルのうちの少なくとも一部のファイルのメタデータ又は制御情報に含まれていてもよい。ここで、関連するビットストリーム又はファイルを示す情報を含むファイルとは、例えばコンテンツの配信に用いられるマニフェストファイル等の制御ファイルであってもよい。

### 産業上の利用可能性

[0183] 本開示は、立体的な音をユーザに知覚させる等の音響再生の際に有用である。

### 符号の説明

[0184]        99 ユーザ  
              100 音響再生システム  
              101 情報処理装置  
              102 通信モジュール  
              103 検知器  
              104 ドライバ  
              105 記憶部  
              111 取得部

- 1 1 2 エンコード音情報入力部
- 1 1 3 デコード処理部
- 1 1 4 センシング情報入力部
- 1 2 1 経路算出部
- 1 3 1 出力音生成部
- 1 3 2 切り替え部
- 1 3 3 第1生成部
- 1 3 4 第2生成部
- 1 4 1 信号出力部
- 3 0 0 立体映像再生装置

## 請求の範囲

- [請求項1] 音声信号と、三次元音場内の音源オブジェクトの位置の情報とを含む音情報を取得する取得部と、
- 前記音源オブジェクトの位置と前記三次元音場内のユーザの位置とに基づき到来方向に応じた頭部伝達関数と、前記音声信号とを用いて、出力音信号を生成する第1生成部と、
- 前記三次元音場内に設定された代表点の位置と前記ユーザの位置とに基づき代表方向に応じた頭部伝達関数と、前記音声信号とを用いて、出力音信号を生成する第2生成部と、を備える
- 情報処理装置。
- [請求項2] 前記第1生成部は、前記音声信号によって前記音源オブジェクトにおいて発せられる再生音に、前記到来方向に応じた頭部伝達関数を畳み込むことで前記出力音信号を生成し、
- 前記第2生成部は、前記再生音を前記代表点から到来する代表音に変換する変換処理を実行し、前記代表方向に応じた頭部伝達関数を畳み込むことで前記出力音信号を生成する
- 請求項1に記載の情報処理装置。
- [請求項3] 前記変換処理では、前記再生音に対して時間シフト調整とゲイン調整とを適用して前記代表音に変換する
- 請求項2に記載の情報処理装置。
- [請求項4] 前記音情報は、複数の前記音源オブジェクトのそれぞれの位置及び前記音声信号によって複数の前記音源オブジェクトのそれぞれにおいて発せられる再生音を含み、
- 前記代表点の数は、前記音源オブジェクトの数に基づいて決定される
- 請求項1に記載の情報処理装置。
- [請求項5] 前記代表点は、前記音源オブジェクトの数より少ない数である
- 請求項4に記載の情報処理装置。

- [請求項6] 前記変換処理の時間シフト調整では、前記再生音に対して、前記到来方向に応じた頭部伝達関数と前記代表方向に応じた頭部伝達関数との相互相関が最大になるように算出された時間シフト、又は、当該時間シフトに負号を付した時間シフトを行う  
請求項3に記載の情報処理装置。
- [請求項7] 前記変換処理では、時間シフト調整及びゲイン調整の少なくとも一方は、周波数軸上の重み付けフィルタをかけてから前記相互相関が最大になるように算出された時間シフト、又は、当該時間シフトに負号を付した時間シフトを行う  
請求項6に記載の情報処理装置。
- [請求項8] 前記変換処理では、2以上の前記代表点のそれぞれについて、時間シフトした前記再生音に、前記再生音及び前記代表方向毎に設定されたゲインをかける  
請求項6に記載の情報処理装置。
- [請求項9] 前記変換処理では、前記代表方向に応じた頭部伝達関数ベクトルの和で前記到来方向に応じた頭部伝達関数ベクトルを合成する際に、合成された頭部伝達関数ベクトルと前記到来方向に応じた頭部伝達関数ベクトルとの誤差信号ベクトルが前記代表方向に応じた頭部伝達関数ベクトルと直交するようにして算出したゲインを用いる  
請求項8に記載の情報処理装置。
- [請求項10] 前記変換処理では、合成された頭部伝達関数ベクトルと前記到来方向に応じた頭部伝達関数ベクトルとの誤差信号ベクトルのエネルギー又はL2ノルムを最小化するようにして算出されたゲインを用いる  
請求項8に記載の情報処理装置。
- [請求項11] 前記誤差信号ベクトルには、周波数軸上の重み付けフィルタをかけたものを用いる  
請求項10に記載の情報処理装置。
- [請求項12] 前記情報処理装置は、初期化時に、代表方向の頭部伝達関数と、前

記変換処理に用いられる時間シフト調整及びゲイン調整における調整量とを頭部伝達関数の各方向ごとに紐づけた調整量テーブルを記憶部に記憶させ、

前記変換処理では、

前記再生音に対して、前記記憶部に記憶された前記調整量テーブルにおいて、前記代表方向に応じた頭部伝達関数の各方向に紐づけられた調整量で時間シフト調整とゲイン調整とを適用して前記代表音に変換する

請求項 3 に記載の情報処理装置。

[請求項13] 前記情報処理装置は、前記初期化時に、複数の前記代表方向を決定し、

前記調整量テーブルは、決定された複数の前記代表方向の頭部伝達関数に基づいて作成される

請求項 1 2 に記載の情報処理装置。

[請求項14] 前記音情報は、前記第 1 生成部を用いて前記出力音信号を生成するか、又は、前記第 2 生成部を用いて前記出力音信号を生成するかを指定するフラグを含み、

前記情報処理装置は、前記第 1 生成部又は前記第 2 生成部のうち、取得した前記音情報に含まれるフラグにおいて指定された一方を用いて前記出力音信号を生成する

請求項 1 に記載の情報処理装置。

[請求項15] 前記第 1 生成部を用いて前記出力音信号を生成するか、又は、前記第 2 生成部を用いて前記出力音信号を生成するかを切り替える切り替え部を備える

請求項 1 に記載の情報処理装置。

[請求項16] 前記切り替え部は、

前記音情報に含まれる前記音源オブジェクトの数と、前記三次元音場内に設定された前記代表点の数とを比較し、

比較結果に応じて、前記第1生成部を用いて前記出力音信号を生成するか、又は、前記第2生成部を用いて前記出力音信号を生成するかを切り替える

請求項15に記載の情報処理装置。

[請求項17] 前記切り替え部は、頭部伝達関数を記憶するための記憶部内に記憶されている頭部伝達関数が所定の条件を満たさない場合に、前記第1生成部を用いて前記出力音信号を生成するように切り替える

請求項15に記載の情報処理装置。

[請求項18] 前記音声信号によって前記音源オブジェクトにおいて発せられる再生音の伝播経路を算出し、算出した前記再生音の伝播経路に応じた前記再生音の間接的な伝播により前記ユーザの位置に到来する合成音及び当該合成音の到来方向を算出する経路算出部を備える

請求項1～17のいずれか1項に記載の情報処理装置。

[請求項19] 前記第1生成部を用いて前記出力音信号を生成するか、又は、前記第2生成部を用いて前記出力音信号を生成するかを切り替える切り替え部を備え、

前記切り替え部は、前記再生音及び前記合成音のそれぞれについて個別に、前記第1生成部を用いて前記出力音信号を生成するか、又は、前記第2生成部を用いて前記出力音信号を生成するかを切り替える

請求項18に記載の情報処理装置。

[請求項20] 前記第1生成部を用いて前記出力音信号を生成するか、又は、前記第2生成部を用いて前記出力音信号を生成するかを切り替える切り替え部を備え、

前記経路算出部は、互いに異なる間接的な伝播により前記ユーザの位置に到来する2以上の前記合成音及び当該2以上の前記合成音のそれぞれの到来方向を算出し、

前記切り替え部は、前記2以上の前記合成音のそれぞれについて個別に、前記第1生成部を用いて前記出力音信号を生成するか、又は、

前記第2生成部を用いて前記出力音信号を生成するかを切り替える  
請求項18に記載の情報処理装置。

[請求項21] 前記第1生成部を用いて前記出力音信号を生成するか、又は、前記第2生成部を用いて前記出力音信号を生成するかを切り替える切り替え部を備え、

前記切り替え部は、前記再生音及び前記合成音の合計数と、前記三次元音場内に設定された前記代表点の数とを比較し、

比較結果に応じて、前記第1生成部を用いて前記出力音信号を生成するか、又は、前記第2生成部を用いて前記出力音信号を生成するかを切り替える

請求項18に記載の情報処理装置。

[請求項22] 音情報を処理することにより、仮想的な三次元音場内の音源オブジェクトから到来する音として出力音信号を生成する、コンピュータによって実行される情報処理方法であって、

前記音源オブジェクトの位置、及び、音声信号であって、当該音声信号により前記音源オブジェクトにおいて発せられる再生音を含む音声信号を取得するステップと、

前記三次元音場内のユーザの位置を取得するステップと、

前記音源オブジェクトの位置から前記ユーザの位置に到来する前記再生音の到来方向を算出するステップと、

算出した前記到来方向に応じた頭部伝達関数と前記再生音とを用いて前記出力音信号を生成するステップと、

前記三次元音場内に設定された代表点の位置と前記ユーザの位置とに基づく代表方向に応じた頭部伝達関数と、前記音声信号とを用いて、前記出力音信号を生成するステップと、を含む

情報処理方法。

[請求項23] 請求項22に記載の情報処理方法をコンピュータに実行させるための

プログラム。

[請求項24] 複数の方向のそれぞれと時間シフト調整量及びゲイン調整量とを対応付けて記憶する記憶部と、

音声信号と、三次元音場内の音源オブジェクトの位置の情報とを取得する取得部と、

前記音声信号と、前記音源オブジェクトの位置と前記三次元音場内のユーザの位置とに基づく第1方向に対応する前記時間シフト調整量及びゲイン調整量とを用いて、第2方向から前記ユーザの位置に到来する音として出力音信号を生成する第2生成部と、を備える

情報処理装置。

[請求項25] 前記記憶部は、さらに前記第2方向に対応する頭部伝達関数を記憶しており、

前記第2生成部は、前記音声信号と、前記第1方向に対応する前記時間シフト調整量及びゲイン調整量と、前記第2方向に対応する頭部伝達関数とを用いて、前記第2方向から前記ユーザの位置に到来する音として出力音信号を生成する

請求項24に記載の情報処理装置。

[請求項26] 前記記憶部は、さらに前記第2方向及び前記第2方向以外の方向に対応する頭部伝達関数を記憶しており、

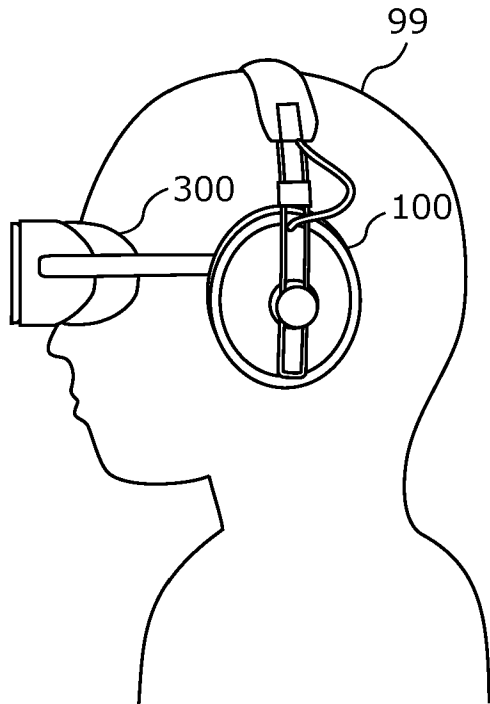
前記第2生成部は、前記音声信号と、前記第1方向に対応する前記時間シフト調整量及びゲイン調整量と、前記第2方向に対応する頭部伝達関数とを用いて、前記第2方向から前記ユーザの位置に到来する音として出力音信号を生成し、

前記情報処理装置は、さらに、第1生成部を備え、

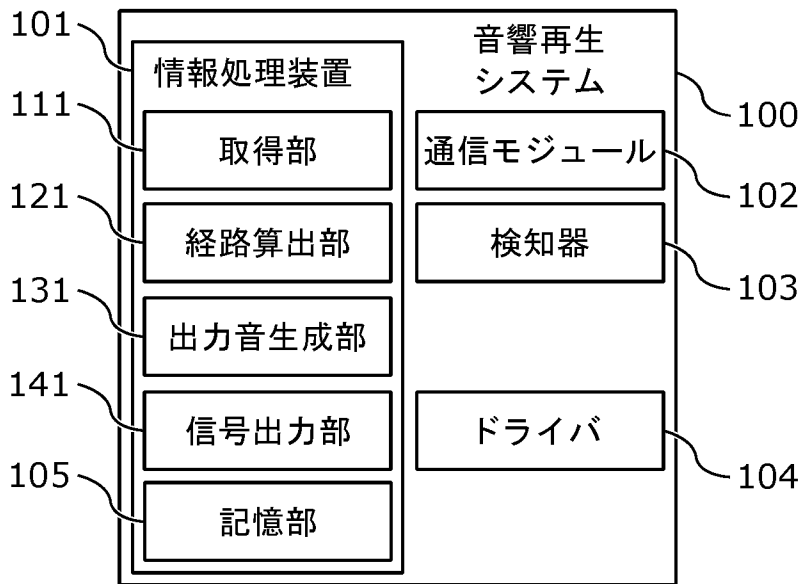
前記第1生成部は、前記音声信号と、前記第1方向に対応する頭部伝達関数とを用いて、前記第1方向から前記ユーザの位置に到来する音として音声信号を生成する

請求項24に記載の情報処理装置。

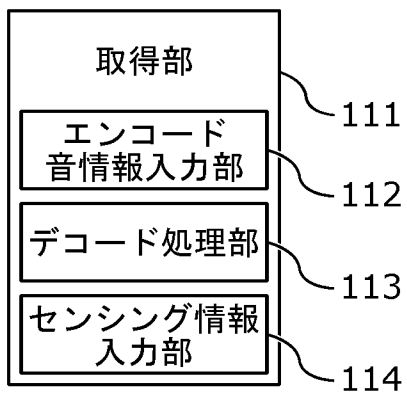
[図1]



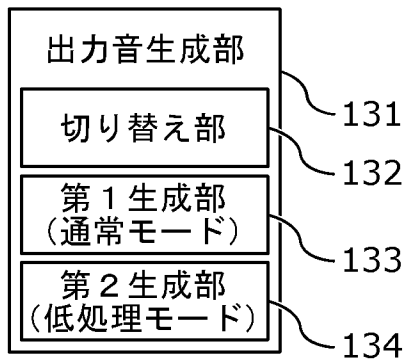
[図2]



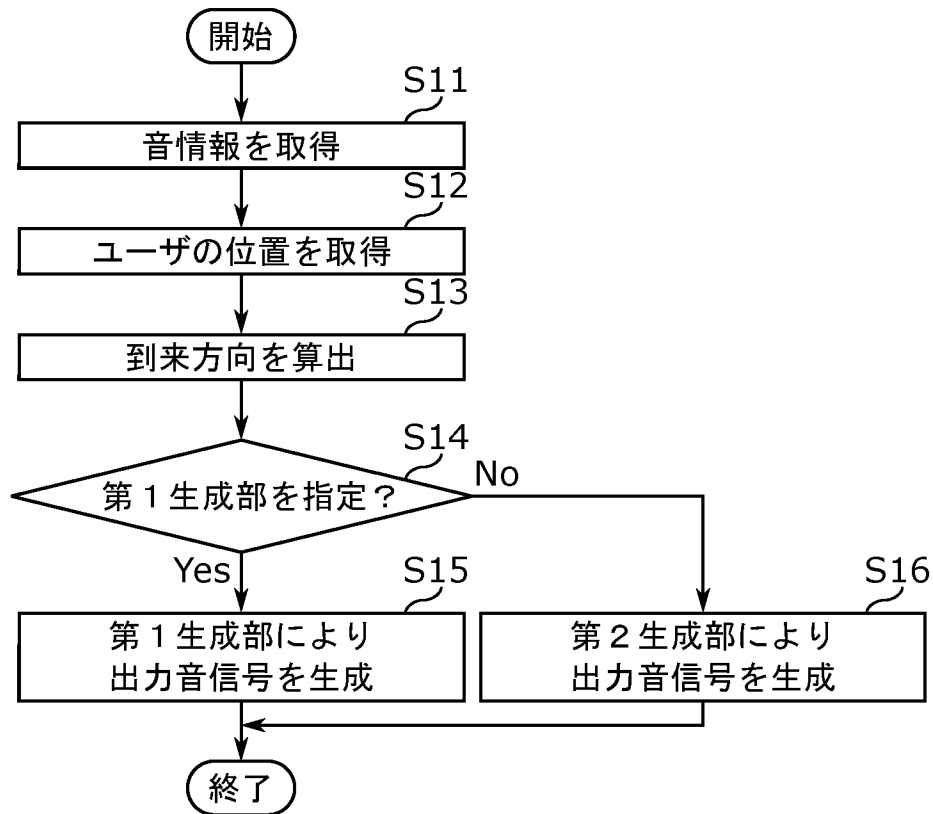
[図3]



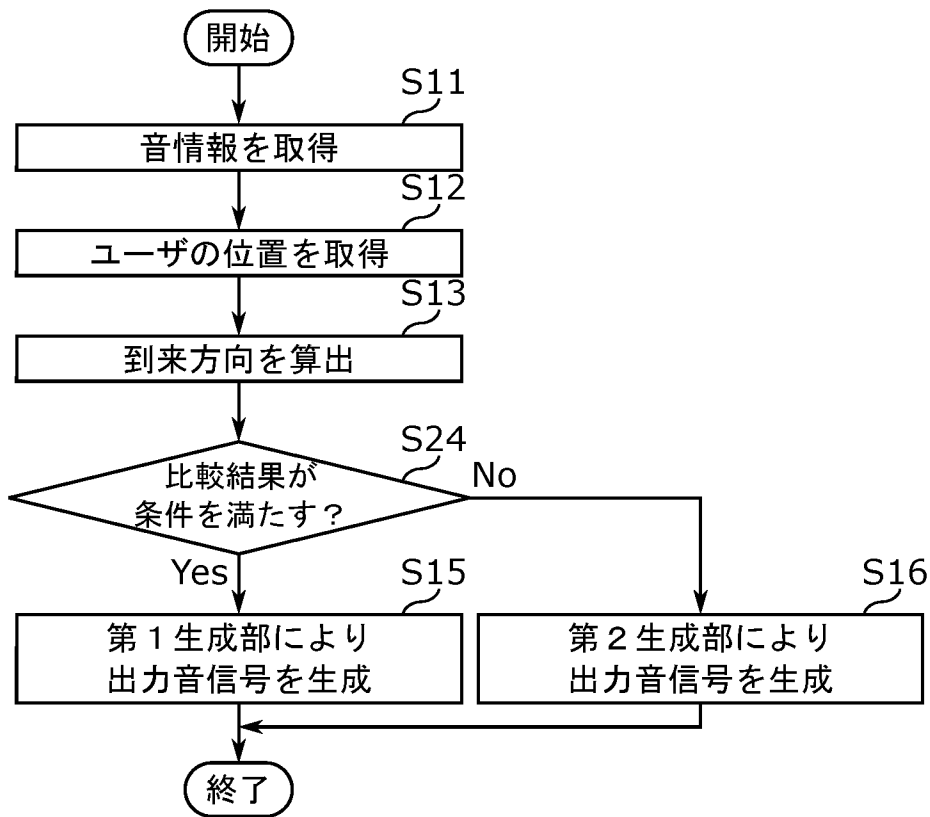
[図4]



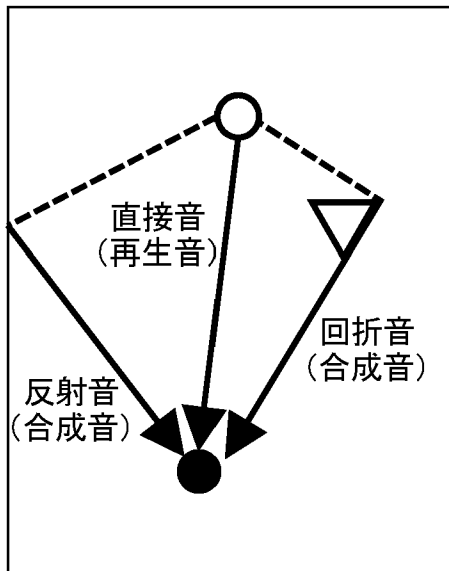
[図5]



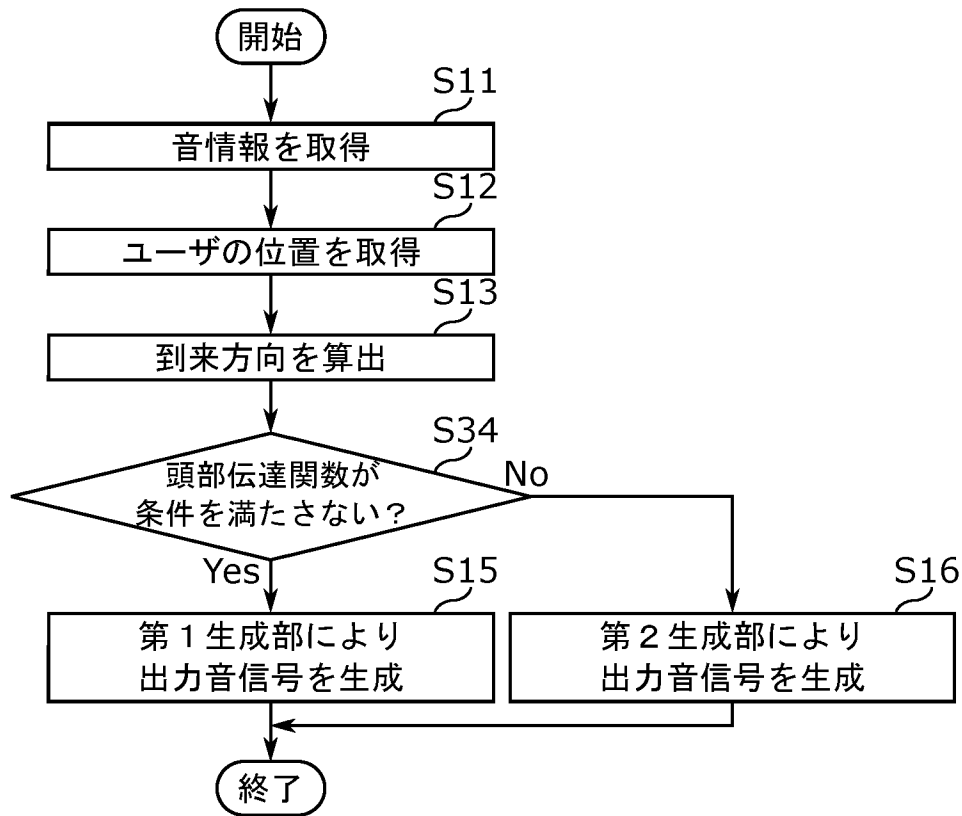
[図6]



[図7]



[図8]



## INTERNATIONAL SEARCH REPORT

International application No.

**PCT/JP2024/014744**

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
<i>H04S 7/00</i> (2006.01) FI: H04S7/00 320		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) H04S7/00		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Published examined utility model applications of Japan 1922-1996 Published unexamined utility model applications of Japan 1971-2024 Registered utility model specifications of Japan 1996-2024 Published registered utility model applications of Japan 1994-2024		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2022/038929 A1 (PANASONIC INTELLECTUAL PROPERTY CORPORATION OF AMERICA) 24 February 2022 (2022-02-24) entire text, all drawings	1-26
A	WO 2019/116890 A1 (SONY CORPORATION) 20 June 2019 (2019-06-20) entire text, all drawings	1-26
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“D” document cited by the applicant in the international application</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> <p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&amp;” document member of the same patent family</p>		
Date of the actual completion of the international search <b>17 June 2024</b>		Date of mailing of the international search report <b>02 July 2024</b>
Name and mailing address of the ISA/JP <b>Japan Patent Office (ISA/JP) 3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915 Japan</b>		Authorized officer  Telephone No.

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No. <b>PCT/JP2024/014744</b>
---

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
WO 2022/038929 A1	24 February 2022	US 2023/0179938 A1 entire text, all drawings EP 4203520 A1 CN 116018824 A	
-----			
WO 2019/116890 A1	20 June 2019	US 2021/0168548 A1 entire text, all drawings EP 3726859 A1 CN 111434126 A KR 10-2020-0096508 A	
-----			

A. 発明の属する分野の分類（国際特許分類（IPC）） H04S 7/00(2006.01)i FI: H04S7/00 320		
B. 調査を行った分野 調査を行った最小限資料（国際特許分類（IPC）） H04S7/00 最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2024年 日本国実用新案登録公報 1996-2024年 日本国登録実用新案公報 1994-2024年		
国際調査でを使用した電子データベース（データベースの名称、調査に使用した用語）		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	WO 2022/038929 A1 (パナソニック インテレクチュアル プロパティ コーポレーション オブ アメリカ) 24.02.2022 (2022-02-24) 全文全図	1-26
A	WO 2019/116890 A1 (ソニー株式会社) 20.06.2019 (2019-06-20) 全文全図	1-26
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input checked="" type="checkbox"/> パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー “A” 特に関連のある文献ではなく、一般的技術水準を示すもの “D” 国際出願で出願人が先行技術文献として記載した文献 “E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの “L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す） “O” 口頭による開示、使用、展示等に言及する文献 “P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献 “T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの “X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの “Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの “&” 同一パテントファミリー文献		
国際調査を完了した日 17.06.2024	国際調査報告の発送日 02.07.2024	
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 大石 剛 5Z 4882 電話番号 03-3581-1101 内線 3549	

国際調査報告  
 パテントファミリーに関する情報

国際出願番号

PCT/JP2024/014744

引用文献			公表日	パテントファミリー文献			公表日
WO	2022/038929	A1	24.02.2022	US	2023/0179938	A1	
				全文全図			
				EP	4203520	A1	
				CN	116018824	A	
-----							
WO	2019/116890	A1	20.06.2019	US	2021/0168548	A1	
				全文全図			
				EP	3726859	A1	
				CN	111434126	A	
				KR	10-2020-0096508	A	
-----							