

(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
2018年11月8日(08.11.2018)



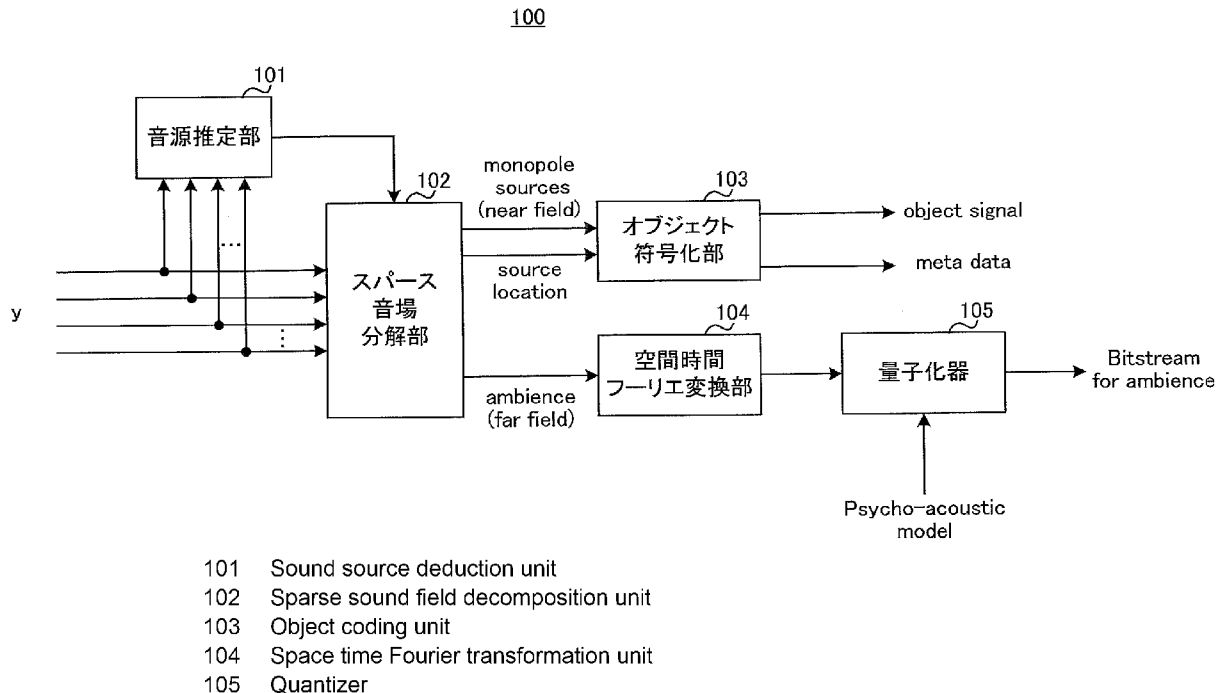
(10) 国際公開番号

WO 2018/203471 A1

- (51) 国際特許分類:
G10L 19/008 (2013.01) *G10L 25/51* (2013.01)
G10L 21/0272 (2013.01)
- (21) 国際出願番号: PCT/JP2018/015790
- (22) 国際出願日: 2018年4月17日(17.04.2018)
- (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 2017-091412 2017年5月1日(01.05.2017) JP
- (71) 出願人: 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(PANASONIC INTELLECTUAL PROPER-
TY CORPORATION OF AMERICA) [US/US];
- 90503 カリフォルニア州トーランス, ス
イート 200, マリナー アベニュー
20000 California (US).
- (72) 発明者: 江 原 宏 幸 (EHARA Hiroyuki);
〒5718501 大阪府門真市大字門真 1 0 0 6 番
地 パナソニック株式会社内 Osaka (JP). 川
村 明久(KAWAMURA Akihisa). ウ カイ(WU
Kai). ナギセティ スリカンス(NAGISETTY
Srikanth). ネオ スア ホン(NEO Sua Hong).
- (74) 代理人: 徳 田 佳 昭, 外 (TOKUDA Yoshiaki et
al.); 〒5718501 大阪府門真市大字門真 1 0 0 6
番地 パナソニック株式会社内 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,

(54) Title: CODING APPARATUS AND CODING METHOD

(54) 発明の名称: 符号化装置及び符号化方法



(57) **Abstract:** A sound source deduction unit (101) deduces an area where a sound source is present, by using a second mesh size larger than a first mesh size at a position where the sound source is assumed to be present in sparse sound field decomposition in a space for which the sparse sound field decomposition is to be carried out. A sparse sound field decomposition unit (102) carries out a sparse sound field decomposition process with the first mesh size for an acoustic signal observed by a microphone array within the area of the second mesh size in which the sound source has been deduced

BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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))

to be present in the space, and decomposes the acoustic signal into a sound source signal and an ambient noise signal.

(57) 要約: 音源推定部(101)は、スパース音場分解の対象となる空間において、スパース音場分解において音源が存在すると仮定する位置の第1の粒度よりも粗い第2の粒度で、音源が存在するエリアを推定し、スパース音場分解部(102)は、空間のうちの音源が存在すると推定された第2の粒度のエリア内において、マイクロホンアレイで観測される音響信号に対して、第1の粒度でスパース音場分解処理を行って、音響信号を音源信号と環境雑音信号とに分解する。

明 細 書

発明の名称：符号化装置及び符号化方法

技術分野

[0001] 本開示は、符号化装置及び符号化方法に関する。

背景技術

[0002] 波面合成符号化技術として、波面合成符号化を時空間周波数領域で行う方法が提案されている（例えば、特許文献1を参照）。

[0003] また、立体音響に対して主要音源成分と環境音成分とに分離符号化する高能率符号化のモデル（例えば、特許文献2を参照）を波面合成に適用し、スパース音場分解（sparse sound field decomposition）を用いて、マイクロホンアレイで観測される音響信号を、少数の点音源（monopole source）と点音源以外の残差成分とに分離して波面合成を行う方法が提案されている（例えば、特許文献3を参照）。

先行技術文献

特許文献

[0004] 特許文献1：米国特許第8, 219, 409号明細書

特許文献2：特表2015-537256号公報

特許文献3：特開2015-171111号公報

非特許文献

[0005] 非特許文献1：M. Cobos, A. Marti, and J. J. Lopez. "A modified SRP-PHAT functional for robust real-time sound source localization with scalable spatial sampling." IEEE Signal Processing Letters 18.1 (2011): 71-74

非特許文献2：Koyama, Shoichi, et al. "Analytical approach to wave field reconstruction filtering in spatio-temporal frequency domain." IEEE Transactions on Audio, Speech, and Language Processing 21.4 (2013): 685-696

発明の概要

- [0006] しかしながら、特許文献1では、音場情報を全て符号化するため、演算量が膨大となる。また、特許文献3では、スパース分解を用いて点音源を抽出する際に、分析対象となる空間内の点音源が存在し得る全ての位置（格子点（grid point））を用いた行列演算が必要となり、演算量が膨大となる。
- [0007] 本開示の一態様は、低演算量で音場のスパース分解を行うことができる符号化装置及び符号化方法の提供に資する。
- [0008] 本開示の一態様に係る符号化装置は、スパース音場分解の対象となる空間において、前記スパース音場分解において音源が存在すると仮定する位置の第1の粒度よりも粗い第2の粒度で、音源が存在するエリアを推定する推定回路と、前記空間のうちの前記音源が存在すると推定された前記第2の粒度のエリア内において、マイクロホンアレイで観測される音響信号に対して、前記第1の粒度で前記スパース音場分解処理を行って、前記音響信号を音源信号と環境雑音信号とに分解する分解回路と、を具備する構成を採る。
- [0009] 本開示の一態様に係る符号化方法は、スパース音場分解の対象となる空間において、前記スパース音場分解において音源が存在すると仮定する位置の第1の粒度よりも粗い第2の粒度で、音源が存在するエリアを推定し、前記空間のうちの前記音源が存在すると推定された前記第2の粒度のエリア内において、マイクロホンアレイで観測される音響信号に対して、前記第1の粒度で前記スパース音場分解処理を行って、前記音響信号を音源信号と環境雑音信号とに分解する。
- [0010] なお、これらの包括的または具体的な態様は、システム、方法、集積回路、コンピュータプログラム、または、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および記録媒体の任意な組み合わせで実現されてもよい。
- [0011] 本開示の一態様によれば、低演算量で音場のスパース分解を行うことができる。
- [0012] 本開示の一態様における更なる利点および効果は、明細書および図面から

明らかにされる。かかる利点および／または効果は、いくつかの実施形態並びに明細書および図面に記載された特徴によってそれぞれ提供されるが、1つまたはそれ以上の同一の特徴を得るために必ずしも全てが提供される必要はない。

図面の簡単な説明

- [0013] [図1]実施の形態1に係る符号化装置の一部の構成例を示すブロック図
- [図2]実施の形態1に係る符号化装置の構成例を示すブロック図
- [図3]実施の形態1に係る復号装置の構成例を示すブロック図
- [図4]実施の形態1に係る符号化装置の処理の流れを示すフロー図
- [図5]実施の形態1に係る音源推定処理及びスパース音場分解処理の説明に供する図
- [図6]実施の形態1に係る音源推定処理の説明に供する図
- [図7]実施の形態1に係るスパース音場分解処理の説明に供する図
- [図8]音場の空間全てに対してスパース音場分解処理を行う場合の説明に供する図
- [図9]実施の形態2に係る符号化装置の構成例を示すブロック図
- [図10]実施の形態2に係る復号装置の構成例を示すブロック図
- [図11]実施の形態3に係る符号化装置の構成例を示すブロック図
- [図12]実施の形態4の方法1に係る符号化装置の構成例を示すブロック図
- [図13]実施の形態4の方法2に係る符号化装置の構成例を示すブロック図
- [図14]実施の形態4の方法2に係る復号装置の構成例を示すブロック図

発明を実施するための形態

- [0014] 以下、本開示の実施の形態について図面を参照して詳細に説明する。
- [0015] なお、以下では、符号化装置において、スパース分解を用いて点音源を抽出する際の分析対象となる空間（音場）内の点音源が存在する可能性のある位置を表す格子点数を「N」個とする。
- [0016] また、符号化装置は、「M」個のマイクロホンを含むマイクロホンアレイを備える（図示せず）。

[0017] また、各マイクロホンで観測される音響信号を「 y 」 ($\in \mathbb{C}^M$) と表す。また、音響信号 y に含まれる、各格子点における音源信号成分（モノポール音源成分の分布）を「 x 」 ($\in \mathbb{C}^N$) と表し、音源信号成分以外の残りの成分である環境雑音信号（残差成分）を「 h 」 ($\in \mathbb{C}^M$) と表す。

[0018] すなわち、次式（１）に示すように、音響信号 y は、音源信号 x と環境雑音信号 h とで表される。すなわち、符号化装置は、スパース音場分解において、マイクロホンアレイで観測される音響信号 y を、音源信号 x と環境雑音信号 h とに分解する。

[数1]

$$y = Dx + h \quad \dots (1)$$

[0019] なお、 D ($\in \mathbb{C}^{M \times N}$) は、各マイクロホンアレイと各格子点との間の伝達関数（例えば、グリーン関数）を要素とする $M \times N$ の行列（dictionary matrix）である。行列 D は、例えば、符号化装置において、各マイクロホンと各格子点との位置関係に基づいて、少なくともスパース音場分解の前に求められていけばよい。

[0020] ここで、スパース音場分解の対象となる空間において、ほとんどの格子点における音源信号成分 x がゼロとなり、少数の格子点の音源信号成分 x が非ゼロとなる特性（スパース性。sparsity constraint）を仮定する。例えば、スパース音場分解では、スパース性を利用して、次式（２）で示される基準を満たす音源信号成分 x を得る。

[数2]

$$\min \|y - Dx\| + \lambda J_{p,q}(x)$$

$$\text{where: } J_{p,q}(x) = \sum_{n=1}^N \|x[n]\|_q^p \quad \dots (2)$$

[0021] 関数 $J_{p,q}(x)$ は、音源信号成分 x のスパース性を生じさせるためのペナルティ関数を示し、 λ は、ペナルティと近似誤差とのバランスを取るパラメータである。

[0022] なお、本開示におけるスパース音場分解の具体的な処理については、例えば、特許文献3に示された方法を用いて行われればよい。ただし、本開示において、スパース音場分解の方法は、特許文献3に示された方法に限定されず、他の方法でもよい。

[0023] ここで、スパース音場分解アルゴリズム（例えば、M-FOCUS S/G-FOCUS S又は最小ノルム解に基づく分解など）では、分析対象となる空間内の全ての格子点を用いた行列演算（逆行列など複素行列演算）が必要となるため、点音源を抽出する場合には演算量が膨大になってしまう。特に、格子点の個数Nが多くなるほど、式(1)に示す音源信号成分 x のベクトルの次元が大きくなり、演算量がより大きくなってしまう。

[0024] そこで、本開示の各実施の形態では、スパース音場分解の低演算量化を図る方法について説明する。

[0025] （実施の形態1）

〔通信システムの概要〕

本実施の形態に係る通信システムは、符号化装置（encoder）100及び復号装置（decoder）200を備える。

[0026] 図1は、本開示の各実施の形態に係る符号化装置100の一部の構成を示すブロック図である。図1に示す符号化装置100において、音源推定部101は、スパース音場分解の対象となる空間において、スパース音場分解において音源が存在すると仮定する位置の第1の粒度よりも粗い第2の粒度で、音源が存在するエリアを推定し、スパース音場分解部102は、空間のうちの音源が存在すると推定された第2の粒度のエリア内において、マイクロホンアレイで観測される音響信号に対して、第1の粒度でスパース音場分解処理を行って、音響信号を音源信号と環境雑音信号とに分解する。

[0027] 〔符号化装置の構成〕

図2は、本実施の形態に係る符号化装置100の構成例を示すブロック図である。図2において、符号化装置100は、音源推定部101と、スパース音場分解部102と、オブジェクト符号化部103と、空間時間フーリエ

変換部 104 と、量子化器 105 と、を含む構成を採る。

[0028] 図 2 において、符号化装置 100 のマイクロホンアレイ（図示せず）から音響信号 y が音源推定部 101 及びスパース音場分解部 102 に入力される。

[0029] 音源推定部 101 は、入力される音響信号 y を分析（音源推定）して、音場（分析対象となる空間）の中から音源の存在するエリア（音源が存在する確率の高いエリア）（格子点のセット）を推定する。例えば、音源推定部 101 は、非特許文献 1 に示されたビームフォーミング（BF）を用いた音源推定方法を用いてもよい。また、音源推定部 101 は、スパース音場分解の分析対象となる空間における N 個の格子点よりも粗い格子点（つまり、少ない格子点）での音源推定を行い、音源の存在する確率の高い格子点（及びその周囲）を選択する。音源推定部 101 は、推定したエリア（格子点のセット）を示す情報をスパース音場分解部 102 に出力する。

[0030] スパース音場分解部 102 は、スパース音場分解の分析対象となる空間のうち、音源推定部 101 から入力される情報に示される、音源が存在すると推定されたエリア内において、入力される音響信号に対してスパース音場分解を行って、音響信号を音源信号 x と、環境雑音信号 h とに分解する。スパース音場分解部 102 は、音源信号成分（monopole sources(near field)）をオブジェクト符号化部 103 に出力し、環境雑音信号成分（ambience(far field)）を空間時間フーリエ変換部 104 に出力する。また、スパース音場分解部 102 は、音源信号の位置（source location）を示す格子点情報をオブジェクト符号化部 103 に出力する。

[0031] オブジェクト符号化部 103 は、スパース音場分解部 102 から入力される音源信号及び格子点情報を符号化し、符号化結果をオブジェクトデータ（object signal）とメタデータのセットとして出力する。例えば、オブジェクトデータ及びメタデータは、オブジェクト符号化ビットストリーム（object bitstream）を構成する。なお、オブジェクト符号化部 103 において、音響信号成分 x の符号化には既存の音響符号化方法を用いればよい。また、メタ

データには、例えば、音源信号に対応する格子点の位置を表す格子点情報等が含まれる。

[0032] 空間時間フーリエ変換部 104 は、スパーズ音場分解部 102 から入力される環境雑音信号に対して空間時間フーリエ変換を行い、空間時間フーリエ変換後の環境雑音信号（空間時間フーリエ係数、二次元フーリエ係数）を量子化器 105 に出力する。例えば、空間時間フーリエ変換部 104 は、特許文献 1 に示された二次元フーリエ変換を用いてもよい。

[0033] 量子化器 105 は、空間時間フーリエ変換部 104 から入力される空間時間フーリエ係数を量子化及び符号化して、環境雑音符号化ビットストリーム（bitstream for ambience）として出力する。例えば、量子化器 105 において、特許文献 1 に示された量子化符号化方法（例えば、心理音響モデル（psycho-acoustic model））を用いてもよい。

[0034] なお、空間時間フーリエ変換部 104 及び量子化器 105 は、環境雑音符号化部と呼ばれてもよい。

[0035] オブジェクト符号化ビットストリーム及び環境雑音ビットストリームは、例えば、多重されて復号装置 200 へ送信される（図示せず）。

[0036] [復号装置の構成]

図 3 は、本実施の形態に係る復号装置 200 の構成を示すブロック図である。図 3 において、復号装置 200 は、オブジェクト復号部 201 と、波面合成部 202 と、環境雑音復号部（逆量子化器）203 と、波面再合成フィルタ（Wavefield reconstruction filter）204 と、逆空間時間フーリエ変換部 205 と、窓かけ部 206 と、加算部 207 と、を含む構成を採る。

[0037] 図 3 において、復号装置 200 は、複数のスピーカから構成されるスピーカアレイを備える（図示せず）。また、復号装置 200 は、図 2 に示す符号化装置 100 からの信号を受信し、受信信号をオブジェクト符号化ビットストリーム（object bitstream）と環境雑音符号化ビットストリーム（ambience bitstream）とに分離する（図示せず）。

[0038] オブジェクト復号部 201 は、入力されるオブジェクト符号化ビットスト

リームを復号して、オブジェクト信号（音源信号成分）とメタデータとに分離し、波面合成部202に出力する。なお、オブジェクト復号部201は、図2に示す符号化装置100のオブジェクト符号化部103で用いた符号化方法の逆の処理により復号処理を行えばよい。

[0039] 波面合成部202は、オブジェクト復号部201から入力されるオブジェクト信号、メタデータ、及び、別途入力又は設定されているスピーカ配置情報（loudspeaker configuration）を用いて、スピーカアレイの各スピーカからの出力信号を求め、求めた出力信号を加算器207に出力する。なお、波面合成部202における出力信号の生成方法は、例えば、特許文献3に示されている方法を用いてもよい。

[0040] 環境雑音復号部203は、環境雑音符号化ビットストリームに含まれる二次元フーリエ係数を復号して、復号された環境雑音信号成分（ambience。例えば、二次元フーリエ係数）を波面再合成フィルタ204に出力する。なお、環境雑音復号部203は、図2に示す符号化装置100の量子化器105における符号化処理と逆の処理により復号処理を行えばよい。

[0041] 波面再合成フィルタ204は、環境雑音復号部203から入力される環境雑音信号成分、及び、別途入力又は設定されているスピーカ配置情報（loudspeaker configuration）を用いて、符号化装置100のマイクロホンアレイで集音された音響信号を復号装置200のスピーカアレイから出力するべき信号に変換し、変換された信号を逆空間時間フーリエ変換部205に出力する。なお、波面再合成フィルタ204における出力信号の生成方法は、例えば、特許文献3に示されている方法を用いてもよい。

[0042] 逆空間時間フーリエ変換部205は、波面再合成フィルタ204から入力される信号に対して逆空間時間フーリエ変換（Inverse space-time Fourier transform）を行い、スピーカアレイの各スピーカから出力されるべき時間信号（環境雑音信号）に変換する。逆空間時間フーリエ変換部205は、時間信号を窓かけ部206に出力する。なお、逆空間時間フーリエ変換部205における変換処理は、例えば、特許文献1に示されている方法を用いてもよ

い。

[0043] 窓かけ部 206 は、逆空間時間フーリエ変換部 205 から入力される、各スピーカから出力されるべき時間信号（環境雑音信号）に対して窓かけ処理（Tapering windowing）を施して、フレーム間の信号をスムーズに接続する。窓かけ部 206 は、窓かけ処理後の信号を加算器 207 に出力する。

[0044] 加算器 207 は、波面合成部 202 から入力される音源信号と、窓かけ部 206 から入力される環境雑音信号とを加算し、加算信号を最終的な復号信号として各スピーカに出力する。

[0045] [符号化装置 100 の動作]

以上の構成を有する符号化装置 100 における動作について詳細に説明する。

[0046] 図 4 は、本実施の形態に係る符号化装置 100 の処理の流れを示すフロー図である。

[0047] まず、符号化装置 100 において、音源推定部 101 は、例えば、非特許文献 1 に示されたビームフォーミングに基づく方法を用いて、音場の中の音源が存在するエリアを推定する（ST101）。この際、音源推定部 101 は、スパース分解の分析対象となる空間において、スパース音場分解時に音源が存在すると仮定する格子点（位置）の粒度よりも粗い粒度で、音源が存在するエリア（coarse area）を推定（特定）する。

[0048] 図 5 は、スパース分解の分析対象となる各格子点（つまり、音源信号成分 x に対応）からなる空間 S （surveillance enclosure）（つまり、音場の観測エリア）の一例を示す。なお、図 5 では空間 S を二次元で表すが実際の空間は三次元でもよい。

[0049] スパース音場分解は、図 5 に示す各格子点を単位として音響信号 y を音源信号 x と環境雑音信号 h とに分離する。これに対して、図 5 に示すように、音源推定部 101 のビームフォーミングによる音源推定の対象となるエリア（coarse area）は、スパース分解の格子点よりも粗いエリアで表される。つまり、音源推定の対象となるエリアは、スパース音場分解の複数の格子点に

よって表される。換言すると、音源推定部 101 は、スパース音場分解部 102 が音源信号 x を抽出する粒度よりも粗い粒度で音源の存在する位置を推定する。

[0050] 図 6 は、音源推定部 101 が図 5 に示す空間 S において音源が存在するエリアとして特定したエリア (identified coarse areas) の一例を示す。図 6 では、例えば、 S_{23} 及び S_{35} のエリア (coarse area) のエネルギーが他のエリアのエネルギーよりも高くなっているとする。この場合、音源推定部 101 は、音源 (source object) が存在するエリアのセット S_{sub} として、 S_{23} 及び S_{35} を特定する。

[0051] 次に、スパース音場分解部 102 は、音源推定部 101 で音源が存在すると推定されたエリア内の格子点についてスパース音場分解を行う (ST102)。例えば、音源推定部 101 において図 6 に示すエリア ($S_{\text{sub}} = [S_{23}, S_{35}]$) が特定された場合、スパース音場分解部 102 は、図 7 に示すように、特定されたエリア ($S_{\text{sub}} = [S_{23}, S_{35}]$) 内におけるスパース音場分解の格子点についてスパース音場分解を行う。

[0052] 例えば、音場推定部 101 で特定されたエリア S_{sub} 内の複数の格子点に対応する音源信号 x を「 x_{sub} 」と表し、行列 D ($M \times N$) のうち、 S_{sub} 内の複数の格子点と符号化装置 100 の複数のマイクロホンとの関係に対応する要素からなる行列を「 D_{sub} 」と表す。

この場合、スパース音場分解部 102 は、次式 (3) のように、各マイクロホンで観測された音響信号 y を、音源信号 x_{sub} と環境雑音信号 h とに分解する。

[数3]

$$y = D_{\text{sub}} x_{\text{sub}} + h \quad \dots (3)$$

[0053] そして、符号化装置 100 (オブジェクト符号化部 103、空間時間フーリエ変換部 104、量子化部 105) は、音源信号 x_{sub} 及び環境雑音信号 h を符号化し (ST103)、得られたビットストリーム (オブジェクト符号化

ビットストリーム、環境雑音符号化ビットストリーム)を出力する(ST104)。これらの信号は復号装置200側へ送信される。

[0054] このように、本実施の形態では、符号化装置100において、音源推定部101は、スパース音場分解の対象となる空間において、スパース音場分解において音源が存在すると仮定する位置を示す格子点の粒度(第1の粒度)よりも粗い粒度(第2の粒度)で、音源が存在するエリアを推定する。そして、スパース音場分解部102は、空間のうちの音源が存在すると推定された、上記第2の粒度のエリア(coarse area)内において、マイクロホンアレイで観測される音響信号 y に対して、上記第1の粒度でスパース音場分解処理を行って、音響信号 y を音源信号 x と環境雑音信号 h とに分解する。

[0055] すなわち、符号化装置100は、音源が存在する確率の高いエリアを予備的に探索し、スパース音場分解の分析対象を、探索されたエリアに限定する。換言すると、符号化装置100は、スパース音場分解の適用範囲を、全ての格子点のうち、音源が存在する周辺の格子点に限定する。

[0056] 上述したように、音場内に存在する音源は少数であることが仮定される。これにより、符号化装置100では、スパース音場分解の分析対象のエリアがより狭いエリアに限定されるので、全ての格子点についてスパース音場分解処理を行う場合と比較して、スパース音場分解処理の演算量を大幅に削減することができる。

[0057] 例えば、図8は、全ての格子点に対してスパース音場分解を行う場合の様子を示す。図8では、図6と同様の位置に2つの音源が存在している。図8では、例えば、特許文献3に示される方法のように、スパース音場分解において、分析対象となる空間内の全ての格子点を用いた行列演算が必要となる。これに対して、図7に示すように、本実施の形態のスパース音場分解の分析対象となるエリアが S_{sub} に削減されている。このため、スパース音場分解部102において、音源信号 x_{sub} のベクトルの次元が小さくなるので、行列 D_{sub} に対する行列演算量が削減される。

[0058] よって、本実施の形態によれば、低演算量で音場のスパース分解を行うこ

とができる。

[0059] また、例えば、図 7 のように行列 D_{sub} の列数の削減によって劣決定系の条件 (under-determined condition) が緩和されるので、スパース音場分解の性能を向上させることができる。

[0060] (実施の形態 2)

[符号化装置の構成]

図 9 は、本実施の形態に係る符号化装置 300 の構成を示すブロック図である。

[0061] なお、図 9 において、実施の形態 1 (図 2) と同様の構成には同様の符号を付し、その説明を省略する。具体的には、図 9 に示す符号化装置 300 は、実施の形態 1 の構成 (図 2) に対して、ビット配分部 301 及び切替部 302 を新たに備える。

[0062] ビット配分部 301 には、音源推定部 101 から、音場内に存在すると推定される音源の数 (つまり、音源が存在すると推定されたエリア (coarse area) 数) を示す情報が入力される。

[0063] ビット配分部 301 は、音源推定部 101 で推定された音源の数に基づいて、実施の形態 1 と同様のスパース音場分解を行うモード、及び、特許文献 1 に示される時空間スペクトル符号化を行うモードの何れを適用するかを決定する。例えば、ビット配分部 301 は、推定される音源数が所定数 (閾値) 以下の場合、スパース音場分解を行うモードに決定し、推定される音源数が所定数を超える場合に、スパース音場分解を行わずに、時空間スペクトル符号化を行うモードに決定する。

[0064] ここで、所定数としては、例えば、スパース音場分解による符号化性能が十分に得られないほどの音源数 (つまり、スパース性が得られないほどの音源数) でもよい。または、所定数としては、ビットストリームのビットレートが決まっている場合には、当該ビットレートで送信可能なオブジェクトの数の上限値でもよい。

[0065] ビット配分部 301 は、決定したモードを示す切替情報 (switching infor

mation) を切替部 302、オブジェクト符号化部 303、及び、量子化部 305 に出力する。また、切替情報は、オブジェクト符号化ビットストリーム及び環境雑音符号化ビットストリームとともに、復号装置 400（後述する）へ送信される（図示せず）。

[0066] なお、切替情報は、決定したモードに限らず、オブジェクト符号化ビットストリームと、環境雑音符号化ビットストリームとのビット配分を示す情報でもよい。例えば、切替情報は、スパース音場分解を適用するモードでは、オブジェクト符号化ビットストリームに割り当てられるビット数を示し、スパース音場分解を適用しないモードでは、オブジェクト符号化ビットストリームに割り当てられるビット数がゼロであることを示してもよい。または、切替情報は、環境雑音符号化ビットストリームのビット数を示してもよい。

[0067] 切替部 302 は、ビット配分部 301 から入力される切替情報（モード情報又はビット配分情報）に応じて、符号化モードに応じた音響信号 y の出力先の切り替えを行う。具体的には、切替部 302 は、実施の形態 1 と同様のスパース音場分解を適用するモードの場合には音響信号 y をスパース音場分解部 102 に出力する。一方、切替部 302 は、時空間スペクトル符号化を行うモードの場合には音響信号 y を空間時間フーリエ変換部 304 に出力する。

[0068] オブジェクト符号化部 303 は、ビット配分部 301 から入力される切替情報に応じて、スパース音場分解を行うモードの場合（例えば、推定された音源数が閾値以下の場合）には、実施の形態 1 と同様にして音源信号に対してオブジェクト符号化を行う。一方、オブジェクト符号化部 303 は、時空間スペクトル符号化を行うモードの場合（例えば、推定された音源数が閾値を超える場合）には符号化を行わない。

[0069] 空間時間フーリエ変換部 304 は、スパース音場分解を行うモードの場合にスパース音場分解部 102 から入力される環境雑音信号 h 、又は、時空間スペクトル符号化を行うモードの場合に切替部 302 から入力される音響信号 y に対して、空間時間フーリエ変換を行い、空間時間フーリエ変換後の信

号（二次元フーリエ係数）を量子化器 305 に出力する。

[0070] 量子化器 305 は、ビット配分部 301 から入力される切替情報に応じて、スプース音場分解を行うモードの場合には、実施の形態 1 と同様にして二次元フーリエ係数の量子化符号化を行う。一方、量子化器 305 は、時空間スペクトル符号化を行うモードの場合には、特許文献 1 と同様にして二次元フーリエ係数の量子化符号化を行う。

[0071] [復号装置の構成]

図 10 は、本実施の形態に係る復号装置 400 の構成を示すブロック図である。

[0072] なお、図 10 において、実施の形態 1（図 3）と同様の構成には同様の符号を付し、その説明を省略する。具体的には、図 10 に示す復号装置 400 は、実施の形態 1 の構成（図 3）に対して、ビット配分部 401 及び分離部 402 を新たに備える。

[0073] 復号装置 400 は、図 9 に示す符号化装置 300 からの信号を受信し、切替情報（switching information）をビット配分部 401 に出力し、その他のビットストリームを分離部 402 に出力する。

[0074] ビット配分部 401 は、入力される切替情報に基づいて、受信したビットストリームにおけるオブジェクト符号化ビットストリームと環境雑音符号化ビットストリームとのビット配分を決定し、決定したビット配分情報を分離部 402 へ出力する。具体的には、ビット配分部 401 は、符号化装置 300 でスプース音場分解が行われた場合、オブジェクト符号化ビットストリーム及び環境雑音符号化ビットストリームにそれぞれ配分されているビット数を決定する。一方、ビット配分部 401 は、符号化装置 300 で時空間スペクトル符号化が行われた場合、オブジェクト符号化ビットストリームへのビットを配分せずに、環境雑音符号化ビットストリームにビットを配分する。

[0075] 分離部 402 は、ビット配分部 401 から入力されるビット配分情報に従って、入力されるビットストリームを各種パラメータのビットストリームに分離する。具体的には、分離部 402 は、符号化装置 300 においてスパー

ス音場分解が行われた場合には、実施の形態１と同様、ビットストリームを、オブジェクト符号化ビットストリームと環境雑音符号化ビットストリームとに分離し、オブジェクト復号部２０１及び環境雑音復号部２０３にそれぞれ出力する。一方、分離部４０２は、符号化装置３００において時空間スペクトル符号化が行われた場合には、入力されるビットストリームを環境雑音復号部２０３へ出力し、オブジェクト復号部２０１には何も出力しない。

[0076] このように、本実施の形態では、符号化装置３００は、音源推定部１０１において推定された音源の数に応じて、実施の形態１で説明したスパース音場分解を適用するか否かを決定する。

[0077] 上述したように、スパース音場分解では、音場における音源のスパース性を仮定しているため、音源数が多い状況は、スパース音場分解の分析モデルとして最適でない場合がある。すなわち、音源の数が多くなると、音場における音源のスパース性が低下し、スパース音場分解を適用した場合には、分析モデルの表現能力又は分解性能が低下してしまう恐れがある。

[0078] これに対して、符号化装置３００は、音場の数が多くなり（スパース性が弱くなり）、スパース音場分解によって良好な符号化性能が得られない場合には、例えば、特許文献１に示すような時空間スペクトル符号化を行う。なお、音場の数が多い場合の符号化モデルは、特許文献１に示すような時空間スペクトル符号化に限定されるものではない。

[0079] このように、本実施の形態によれば、音源の数に応じて符号化モデルを柔軟に切り替えることができるので、高能率な符号化を実現することができる。

[0080] なお、ビット配分部３０１には、音源推定部１０１から、推定された音源の位置情報が入力されてもよい。例えば、ビット配分部３０１は、音源の位置情報に基づいて、音源信号成分 x と環境雑音信号 h とのビット配分（又は、音源数の閾値）を設定してもよい。例えば、ビット配分部３０１は、音源の位置がマイクロホンアレイに対して正面の位置に近い位置であるほど、音源信号成分 x のビット配分をより多くしてもよい。

[0081] (実施の形態3)

本実施の形態に係る復号装置は、実施の形態2に係る復号装置400と基本構成が共通するので、図10を援用して説明する。

[0082] [符号化装置の構成]

図11は、本実施の形態に係る符号化装置500の構成を示すブロック図である。

[0083] なお、図11において、実施の形態2（図9）と同様の構成には同様の符号を付し、その説明を省略する。具体的には、図11に示す符号化装置500は、実施の形態2の構成（図9）に対して、選択部501を新たに備える。

[0084] 選択部501は、スパース音場分解部102から入力される音源信号 x （スパース音源）のうちの一部の主要な音源（例えば、エネルギーが大きい順に所定数の音源）を選択する。そして、選択部501は、選択した音源信号をオブジェクト信号（monopole sources）としてオブジェクト符号化部303に出力し、選択されなかった残りの音源信号を環境雑音信号（ambience）として空間時間フーリエ変換部502に出力する。

[0085] つまり、選択部501は、スパース音場分解部102で生成（抽出）された音源信号 x の一部を、環境雑音信号 h として分類し直す。

[0086] 空間時間フーリエ変換部502は、スパース音場分解が行われた場合、スパース音場分解部102から入力される環境雑音信号 h 、及び、選択部501から入力される環境雑音信号 h （分類し直された音源信号）に対して時空間スペクトル符号化を行う。

[0087] このように、本実施の形態では、符号化装置500は、スパース音場分解部102で抽出された音源信号のうち、主要な成分を選択し、オブジェクト符号化することにより、オブジェクト符号化で利用可能なビット数に限りがある場合でも、より重要なオブジェクトに対するビット配分を確保することができる。これにより、スパース音場分解による全体的な符号化性能を向上させることができる。

[0088] (実施の形態4)

本実施の形態では、スパース音場分解によって得られた音源信号 x と、環境雑音信号 h とのビット配分を当該環境雑音信号のエネルギーに応じて設定する方法について説明する。

[0089] [方法1]

本実施の形態の方法1に係る復号装置は、実施の形態2に係る復号装置400と基本構成が共通するので、図10を援用して説明する。

[0090] [符号化装置の構成]

図12は、本実施の形態の方法1に係る符号化装置600の構成を示すブロック図である。

[0091] なお、図12において、実施の形態2（図9）又は実施の形態3（図11）と同様の構成には同様の符号を付し、その説明を省略する。具体的には、図12に示す符号化装置600は、実施の形態2の構成（図9）に対して、選択部601及びビット配分更新部602を新たに備える。

[0092] 選択部601は、実施の形態3の選択部501（図11）と同様、スパース音場分解部102から入力される音源信号 x のうちの一部の主要な音源（例えば、エネルギーが大きい順に所定数の音源）を選択する。この際、選択部601は、スパース音場分解部102から入力される環境雑音信号 h のエネルギーを算出し、環境雑音信号のエネルギーが所定の閾値以下の場合には、環境雑音信号のエネルギーが所定の閾値を超える場合よりも多くの音源信号 x を、上記主要な音源としてオブジェクト符号化部303に出力する。選択部601は、音源信号 x の選択結果に応じて、ビット配分の増減を示す情報をビット配分更新部602に出力する。

[0093] ビット配分更新部602は、選択部601から入力される情報に基づいて、オブジェクト符号化部303で符号化される音源信号に割り当てるビット数と、量子化器305において量子化される環境雑音信号に割り当てるビット数との配分を決定する。すなわち、ビット配分更新部602は、ビット配分部301の切替情報（ビット配分情報）を更新する。

[0094] ビット配分更新部602は、更新後のビット配分を示す切替情報をオブジェクト符号化部303及び量子化部305に出力する。また、切替情報は、オブジェクト符号化ビットストリーム及び環境雑音符号化ビットストリームとともに、復号装置400（図10）へ多重して送信される（図示せず）。

[0095] オブジェクト符号化部303及び量子化器305は、ビット配分更新部602から入力される切替情報に示されるビット配分に従って、音源信号 x 又は環境雑音信号 h に対して符号化又は量子化をそれぞれ行う。

[0096] なお、エネルギーが小さく、ビット配分が減らされた環境雑音信号に対しては、符号化が全く行われなくてもよく、復号側で所定の閾値レベルの環境雑音として疑似的に生成されてもよい。または、エネルギーが小さい環境雑音信号に対して、エネルギー情報が符号化・伝送されてもよい。この場合、環境雑音信号に対するビット配分が必要となるが、エネルギー情報のみであれば環境雑音信号 h を含む場合と比較して少ないビット配分で済む。

[0097] [方法2]

方法2では、上述したように環境雑音信号のエネルギー情報を符号化して伝送する構成を有する符号化装置、及び、復号装置の一例について説明する。

[0098] [符号化装置の構成]

図13は、本実施の形態の方法2に係る符号化装置700の構成を示すブロック図である。

[0099] なお、図13において、実施の形態1（図2）と同様の構成には同様の符号を付し、その説明を省略する。具体的には、図13に示す符号化装置700は、実施の形態1（図2）の構成に対して、切替部701、選択部702、ビット配分部703及びエネルギー量子化符号化部704を新たに備える。

[0100] 符号化装置700において、スパース音場分解部102で得られる音源信号 x は選択部702に出力され、環境雑音信号 h は切替部701に出力される。

[0101] 切替部701は、スパース音場分解部102から入力される環境雑音信号のエネルギーを算出し、算出した環境雑音信号のエネルギーが所定の閾値を超え

るか否かを判断する。切替部 701 は、環境雑音信号のエネルギーが所定の閾値以下の場合、環境雑音信号のエネルギーを示す情報 (ambience energy) をエネルギー量子化符号化部 704 に出力する。一方、切替部 701 は、環境雑音信号のエネルギーが所定の閾値を超える場合、環境雑音信号を空間時間フーリエ変換部 104 に出力する。また、切替部 701 は、環境雑音信号のエネルギーが所定の閾値を超えたか否かを示す情報 (判断結果) を選択部 702 に出力する。

[0102] 選択部 702 は、切替部 701 から入力される情報 (環境雑音信号のエネルギーが所定の閾値を超えたか否かを示す情報) に基づいて、スパース音源分離部 102 から入力される音源信号 (スパース音源) の中から、オブジェクト符号化の対象となる音源数 (選択する音源の数) を決定する。例えば、選択部 702 は、方法 1 に係る符号化装置 600 の選択部 601 と同様、環境雑音信号のエネルギーが所定の閾値以下の場合にオブジェクト符号化対象として選択する音源数を、環境雑音信号のエネルギーが所定の閾値を超えた場合にオブジェクト符号化対象として選択する音源数よりも多く設定する。

[0103] そして、選択部 702 は、決定した数の音源成分を選択し、オブジェクト符号化部 103 に出力する。この際、選択部 702 は、例えば、主要な音源 (例えば、エネルギーが大きい順に所定数の音源) から順に選択してもよい。また、選択部 702 は、選択されなかった残りの音源信号 (monopole sources(non-dominant)) を空間時間フーリエ変換部 104 に出力する。

[0104] また、選択部 702 は、決定した音源数、及び、切替部 701 から入力される情報をビット配分部 703 に出力する。

[0105] ビット配分部 703 は、選択部 702 から入力される情報に基づいて、オブジェクト符号化部 103 で符号化される音源信号に割り当てるビット数と、量子化器 105 において量子化される環境雑音信号に割り当てるビット数との配分を設定する。ビット配分部 703 は、ビット配分を示す切替情報をオブジェクト符号化部 103 及び量子化部 105 に出力する。また、切替情報は、オブジェクト符号化ビットストリーム及び環境雑音符号化ビットスト

リームとともに、後述する復号装置 800（図 14）へ多重して送信される（図示せず）。

[0106] エネルギー量子化符号化部 704 は、切替部 701 から入力される環境雑音エネルギー情報を量子化符号化して、符号化情報（ambience energy）を出力する。符号化情報は、環境雑音エネルギー符号化ビットストリームとして、オブジェクト符号化ビットストリーム、環境雑音符号化ビットストリーム及び切替情報とともに、後述する復号装置 800（図 14）へ多重して送信される（図示せず）。

[0107] なお、符号化装置 700 は、環境雑音エネルギーが所定の閾値以下の場合には、環境雑音信号を符号化せずに、ビットレートが許容する範囲において音源信号を追加でオブジェクト符号化してもよい。

[0108] また、方法 2 に係る符号化装置は、図 13 に示す構成に加え、実施の形態 2（図 9）で説明したように音源推定部 101 で推定される音源数に応じてスパース音場分解と他の符号化モデルとを切り替える構成を備えてもよい。または、方法 2 に係る符号化装置は、図 13 に示す音源推定部 101 の構成を含まなくてもよい。

[0109] また、符号化装置 700 は、上述した環境雑音信号のエネルギーとして、全てのチャンネルのエネルギーの平均値を算出してもよく、他の方法を用いてもよい。他の方法としては、例えば、環境雑音信号のエネルギーとして、チャンネル個別の情報を用いる方法、又は、全てのチャンネルをサブグループに分け、各サブグループでの平均エネルギーを求める方法等が挙げられる。この際、符号化装置 700 は、環境雑音信号のエネルギーが閾値を超えるか否かの判断を、全てのチャンネルの平均値を用いて行ってもよく、他の方法を用いる場合には、チャンネル又はサブグループ毎に求めた環境雑音信号のエネルギーのうち最大値を用いて行ってもよい。また、符号化装置 700 は、エネルギーの量子化符号化として、全てのチャンネルの平均エネルギーを用いる場合にはスカラー量子化を適用してもよく、複数のエネルギーを符号化する場合にはスカラー量子化又はベクトル量子化を適用してもよい。また、量子化・符号化効率を向上さ

せるために、フレーム間相関を利用した予測量子化も有効である。

[0110] [復号装置の構成]

図 1 4 は、本実施の形態の方法 2 に係る復号装置 8 0 0 の構成を示すブロック図である。

[0111] なお、図 1 4 において、実施の形態 1（図 3）又は実施の形態 2（図 1 0）と同様の構成には同様の符号を付し、その説明を省略する。具体的には、図 1 4 に示す復号装置 8 0 0 は、実施の形態 2（図 1 0）の構成に対して疑似環境雑音復号部 8 0 1 を新たに備える。

[0112] 疑似環境雑音復号部 8 0 1 は、分離部 4 0 2 から入力される環境雑音エネルギー符号化ビットストリーム、及び、別途復号装置 8 0 0 が保持する疑似環境雑音源を用いて、疑似環境雑音信号を復号し、波面再合成フィルタ 2 0 4 に出力する。

[0113] なお、疑似環境雑音復号部 8 0 1 において、符号化装置 7 0 0 のマイクロホンアレイから復号装置 8 0 0 のスピーカアレイへの変換を考慮した処理を組み込んでおけば、波面再合成フィルタ 2 0 4 への出力をスキップして、逆空間時間フーリエ変換部 2 0 5 に出力するような復号処理とすることも可能である。

[0114] 以上、方法 1 及び方法 2 について説明した。

[0115] このように、本実施の形態では、符号化装置 6 0 0，7 0 0 は、環境雑音信号のエネルギーが小さい場合には、環境雑音信号を符号化するよりも、音源信号成分の符号化に可能な限り多くのビットを配分し直してオブジェクト符号化を行う。これにより、符号化装置 6 0 0，7 0 0 における符号化性能を向上させることができる。

[0116] また、本実施の形態によれば、符号化装置 7 0 0 のスパース音場分解部 1 0 2 で抽出された環境雑音信号のエネルギーの符号化情報が復号装置 8 0 0 に送信される。復号装置 8 0 0 は、環境雑音信号のエネルギーに基づいて、疑似環境雑音信号を生成する。これにより、環境雑音信号のエネルギーが小さい場合には、環境雑音信号の代わりに、少ないビット配分で済むエネルギー情報を

符号化する分、音源信号に対してより多くのビットを配分できるので、音響信号を効率良く符号化することができる。

[0117] 以上、本開示の各実施の形態について説明した。

[0118] なお、本開示はソフトウェア、ハードウェア、又は、ハードウェアと連携したソフトウェアで実現することが可能である。上記実施の形態の説明に用いた各機能ブロックは、部分的に又は全体的に、集積回路であるLSIとして実現され、上記実施の形態で説明した各プロセスは、部分的に又は全体的に、一つのLSI又はLSIの組み合わせによって制御されてもよい。LSIは個々のチップから構成されてもよいし、機能ブロックの一部または全てを含むように一つのチップから構成されてもよい。LSIはデータの入力と出力を備えてもよい。LSIは、集積度の違いにより、IC、システムLSI、スーパーLSI、ウルトラLSIと呼称されることもある。集積回路化の手法はLSIに限るものではなく、専用回路、汎用プロセッサ又は専用プロセッサで実現してもよい。また、LSI製造後に、プログラムすることが可能なFPGA (Field Programmable Gate Array) や、LSI内部の回路セルの接続や設定を再構成可能なりコンフィギュラブル・プロセッサを利用してもよい。本開示は、デジタル処理又はアナログ処理として実現されてもよい。さらには、半導体技術の進歩または派生する別技術によりLSIに置き換わる集積回路化の技術が登場すれば、当然、その技術を用いて機能ブロックの集積化を行ってもよい。バイオ技術の適用等が可能性としてありえる。

[0119] 本開示の符号化装置は、スパース音場分解の対象となる空間において、前記スパース音場分解において音源が存在すると仮定する位置の第1の粒度よりも粗い第2の粒度で、音源が存在するエリアを推定する推定回路と、前記空間のうちの前記音源が存在すると推定された前記第2の粒度のエリア内において、マイクロホンアレイで観測される音響信号に対して、前記第1の粒度で前記スパース音場分解処理を行って、前記音響信号を音源信号と環境雑音信号とに分解する分解回路と、を具備する。

[0120] 本開示の符号化装置において、前記分解回路は、前記推定回路で前記音源

が存在すると推定されたエリアの数が第 1 の閾値以下の場合に前記スパース音場分解処理を行い、前記エリアの数が前記第 1 の閾値を超える場合に前記スパース音場分解処理を行わない。

[0121] 本開示の符号化装置において、前記エリアの数が前記第 1 の閾値以下の場合に、前記音源信号を符号化する第 1 の符号化回路と、前記エリアの数が前記第 1 の閾値以下の場合に前記環境雑音信号を符号化し、前記エリアの数が前記第 1 の閾値を超える場合に前記音響信号を符号化する第 2 の符号化回路と、をさらに具備する。

[0122] 本開示の符号化装置において、前記分解回路で生成された音源信号のうちの一部をオブジェクト信号として出力し、前記分解回路で生成された音源信号のうちの残りを前記環境雑音信号として出力する選択回路、をさらに具備する。

[0123] 本開示の符号化装置において、前記分解回路で生成された前記環境雑音信号のエネルギーが第 2 の閾値以下の場合に選択される前記一部の音源信号の数は、前記環境雑音信号のエネルギーが前記第 2 の閾値を超える場合に選択される前記一部の音源信号の数よりも多い。

[0124] 本開示の符号化装置において、前記エネルギーが前記第 2 の閾値以下の場合に、当該エネルギーを示す情報を量子化符号化する量子化符号化回路、をさらに具備する。

[0125] 本開示の符号化方法は、スパース音場分解の対象となる空間において、前記スパース音場分解において音源が存在すると仮定する位置の第 1 の粒度よりも粗い第 2 の粒度で、音源が存在するエリアを推定し、前記空間のうちの前記音源が存在すると推定された前記第 2 の粒度のエリア内において、マイクロホンアレイで観測される音響信号に対して、前記第 1 の粒度で前記スパース音場分解処理を行って、前記音響信号を音源信号と環境雑音信号とに分解する。

産業上の利用可能性

[0126] 本開示の一態様は、音声通信システムに有用である。

符号の説明

- [0127] 100, 300, 500, 600, 700 符号化装置
- 101 音源推定部
- 102 スパース音場分解部
- 103, 303 オブジェクト符号化部
- 104, 304, 502 空間時間フーリエ変換部
- 105, 305 量子化器
- 200, 400, 800 復号装置
- 201 オブジェクト復号部
- 202 波面合成部
- 203 環境雑音復号部
- 204 波面再合成フィルタ
- 205 逆空間時間フーリエ変換部
- 206 窓かけ部
- 207 加算器
- 301, 401, 703 ビット配分部
- 302, 701 切替部
- 402 分離部
- 501, 601, 702 選択部
- 602 ビット配分更新部
- 704 エネルギー量子化符号化部
- 801 擬似環境雑音復号部

請求の範囲

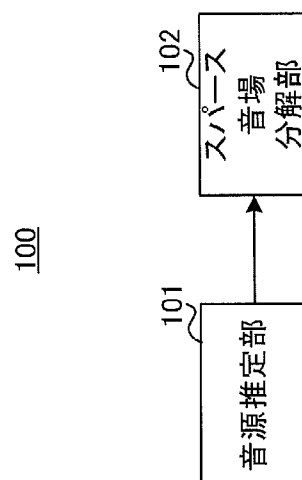
- [請求項1] スペース音場分解の対象となる空間において、前記スペース音場分解において音源が存在すると仮定する位置の第1の粒度よりも粗い第2の粒度で、音源が存在するエリアを推定する推定回路と、
- 前記空間のうちの前記音源が存在すると推定された前記第2の粒度のエリア内において、マイクロホンアレイで観測される音響信号に対して、前記第1の粒度で前記スペース音場分解処理を行って、前記音響信号を音源信号と環境雑音信号とに分解する分解回路と、
- を具備する符号化装置。
- [請求項2] 前記分解回路は、前記推定回路で前記音源が存在すると推定されたエリアの数が第1の閾値以下の場合に前記スペース音場分解処理を行い、前記エリアの数が前記第1の閾値を超える場合に前記スペース音場分解処理を行わない、
- 請求項1に記載の符号化装置。
- [請求項3] 前記エリアの数が前記第1の閾値以下の場合に、前記音源信号を符号化する第1の符号化回路と、
- 前記エリアの数が前記第1の閾値以下の場合に前記環境雑音信号を符号化し、前記エリアの数が前記第1の閾値を超える場合に前記音響信号を符号化する第2の符号化回路と、をさらに具備する、
- 請求項2に記載の符号化装置。
- [請求項4] 前記分解回路で生成された音源信号のうちの一部をオブジェクト信号として出力し、前記分解回路で生成された音源信号のうちの残りを前記環境雑音信号として出力する選択回路、をさらに具備する、
- 請求項1に記載の符号化装置。
- [請求項5] 前記分解回路で生成された前記環境雑音信号のエネルギーが第2の閾値以下の場合に選択される前記一部の音源信号の数は、前記環境雑音信号のエネルギーが前記第2の閾値を超える場合に選択される前記一部の音源信号の数よりも多い、

請求項 4 に記載の符号化装置。

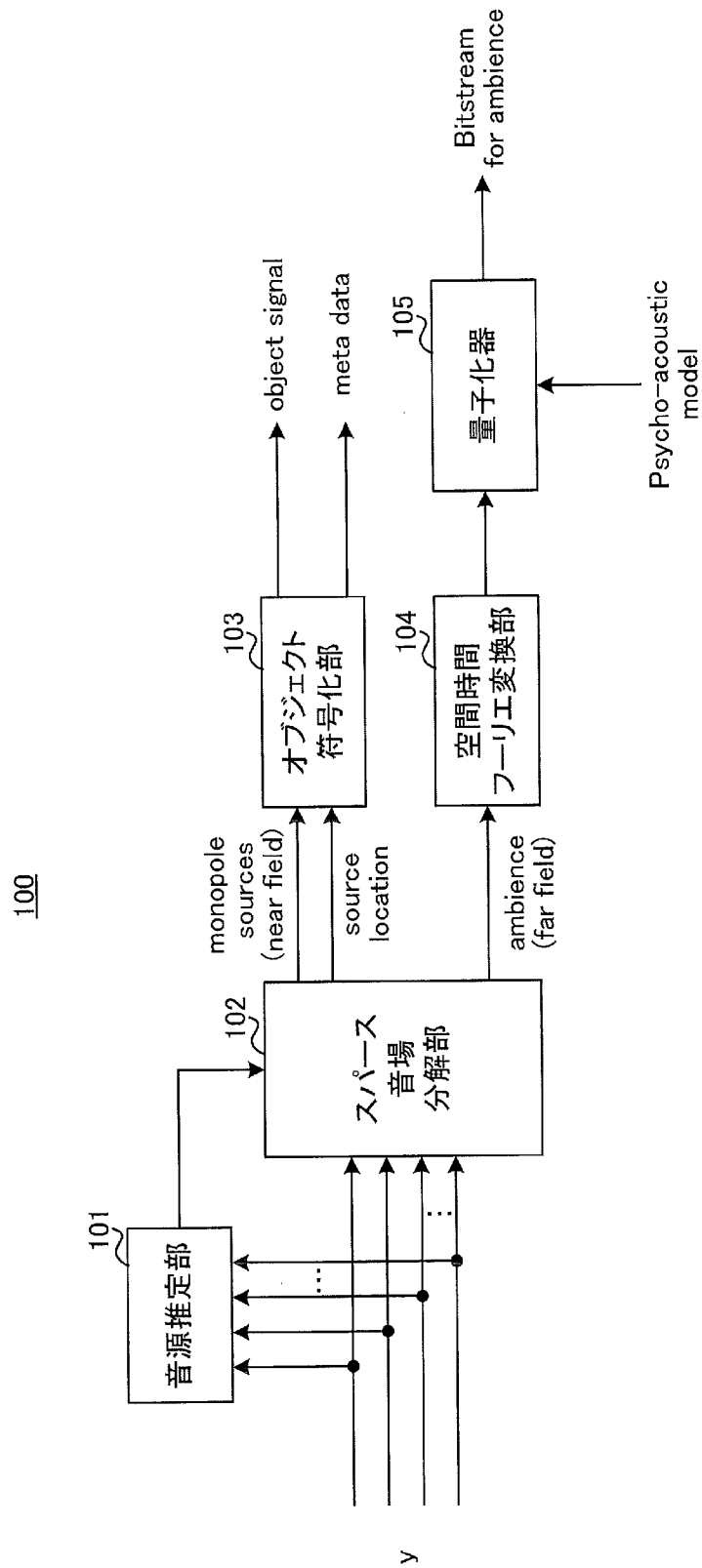
[請求項6] 前記エネルギーが前記第 2 の閾値以下の場合に、当該エネルギーを示す情報を量子化符号化する量子化符号化回路、をさらに具備する、
請求項 5 に記載の符号化装置。

[請求項7] スパース音場分解の対象となる空間において、前記スパース音場分解において音源が存在すると仮定する位置の第 1 の粒度よりも粗い第 2 の粒度で、音源が存在するエリアを推定し、
前記空間のうちの前記音源が存在すると推定された前記第 2 の粒度のエリア内において、マイクロホンアレイで観測される音響信号に対して、前記第 1 の粒度で前記スパース音場分解処理を行って、前記音響信号を音源信号と環境雑音信号とに分解する、
符号化方法。

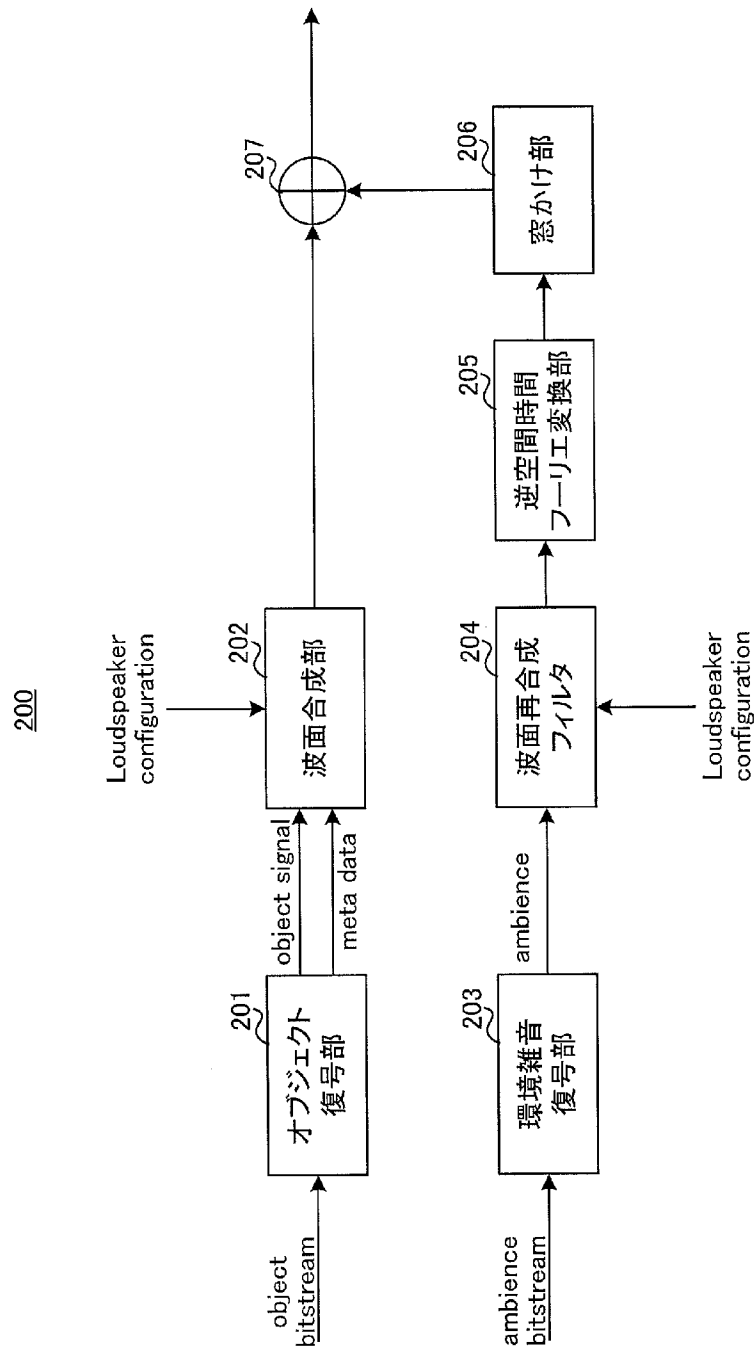
[図1]



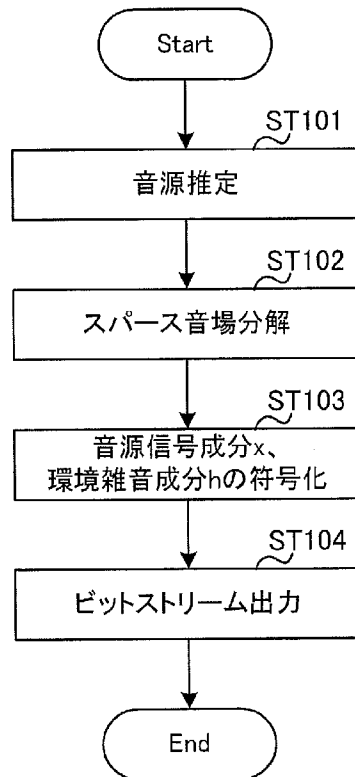
[図2]



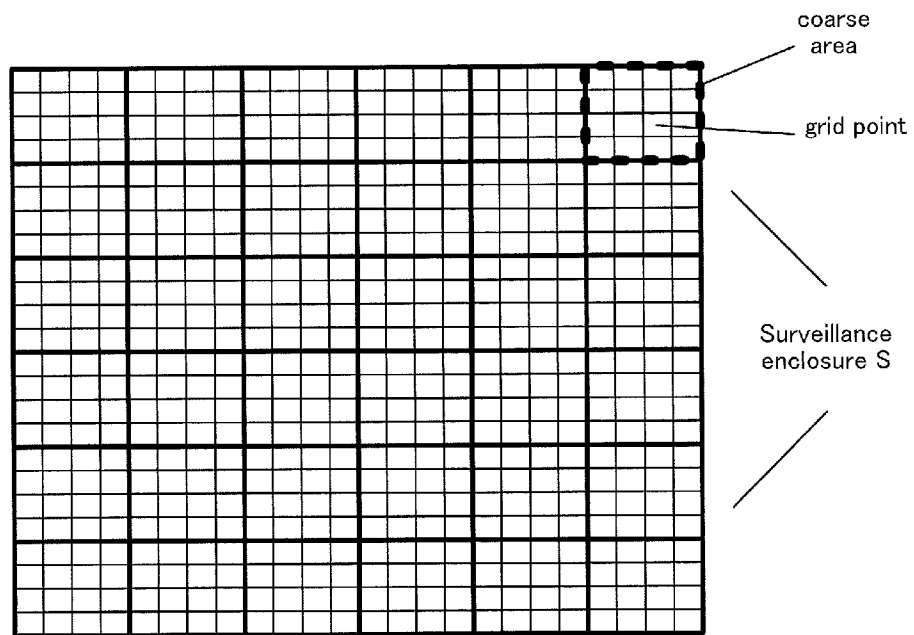
[図3]



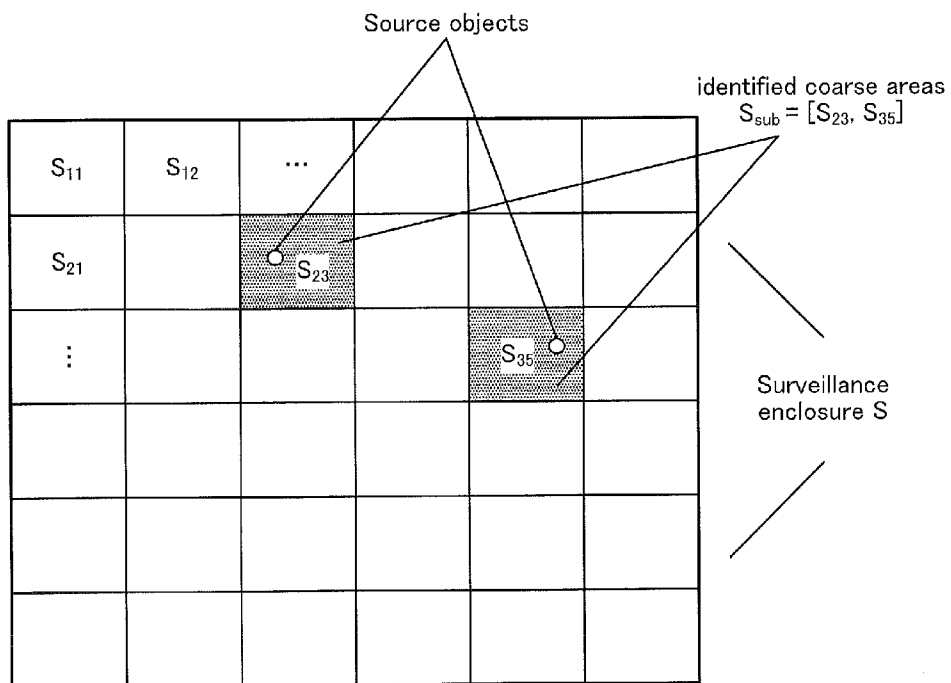
[図4]



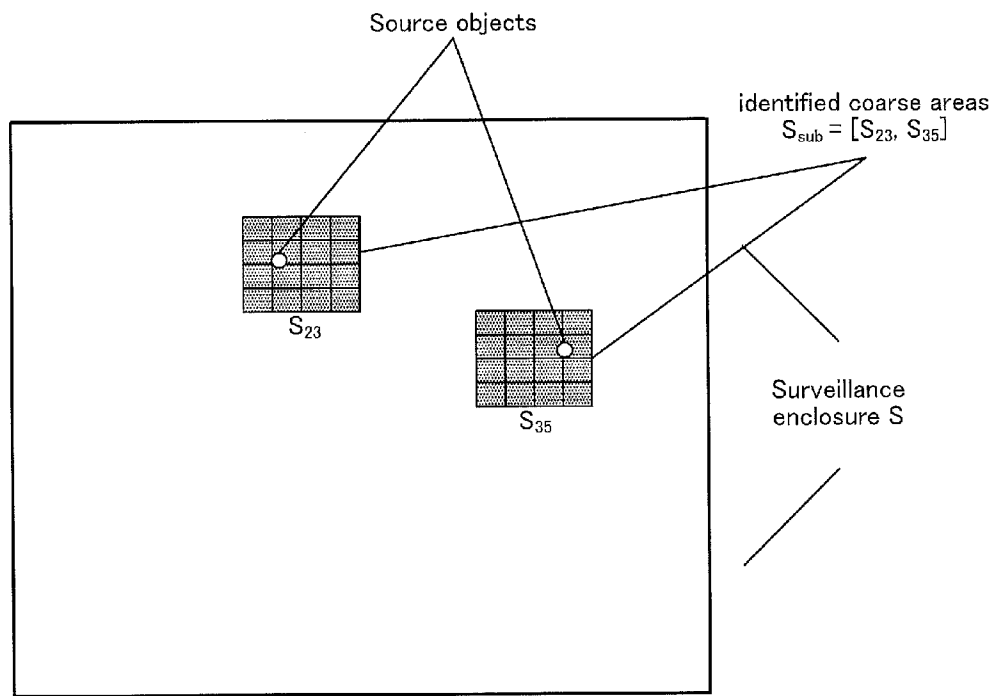
[図5]



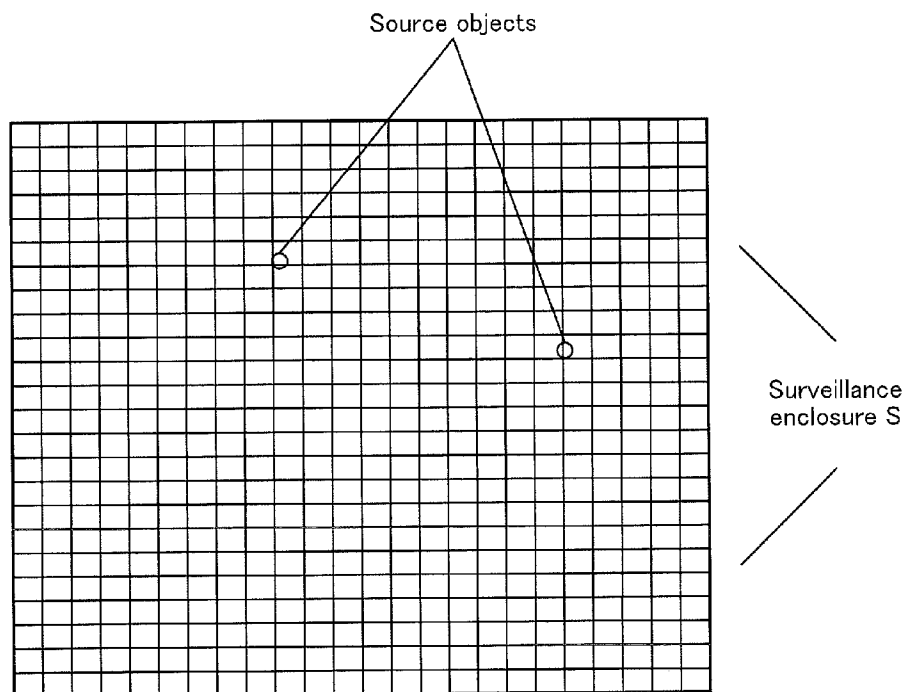
[図6]



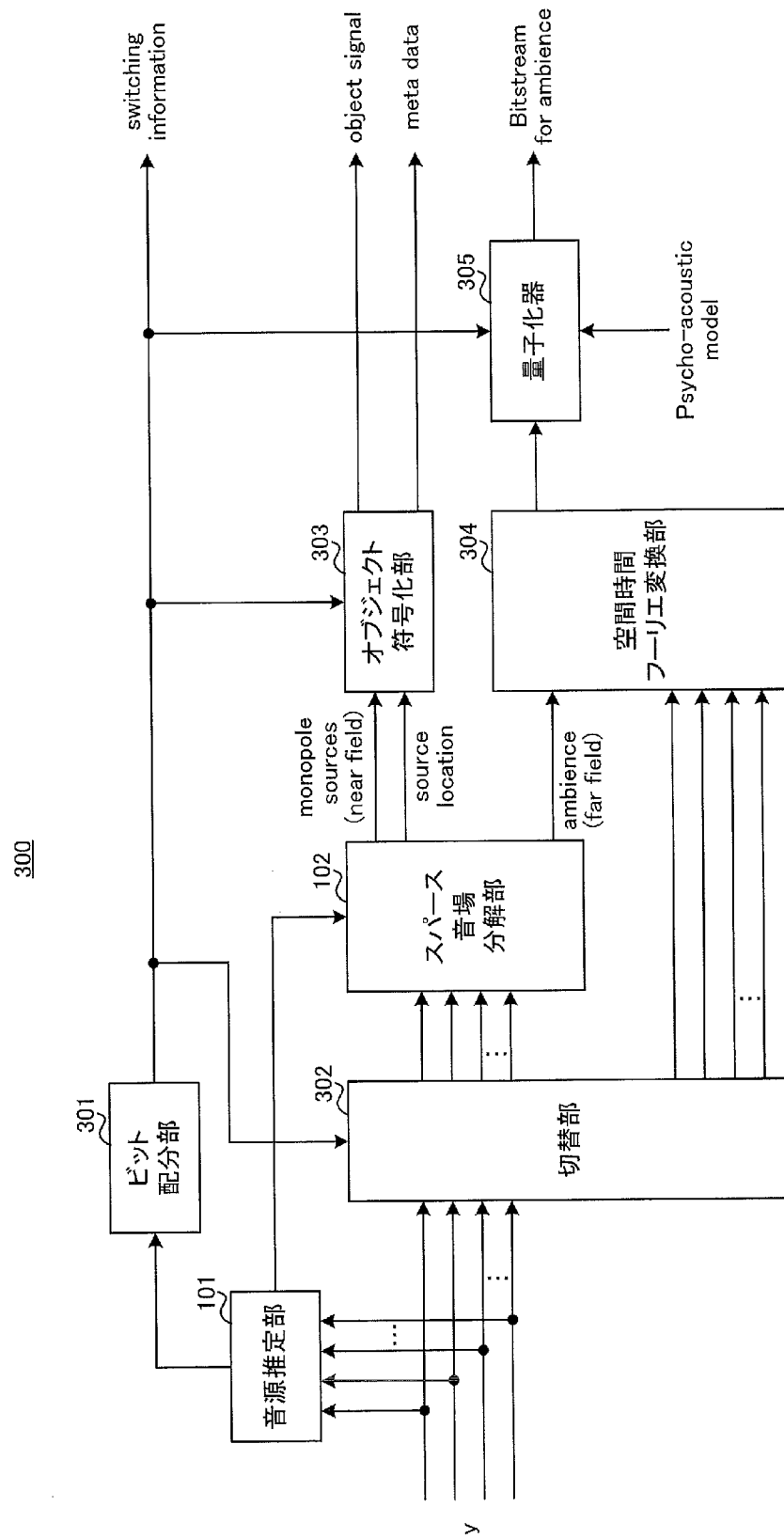
[図7]



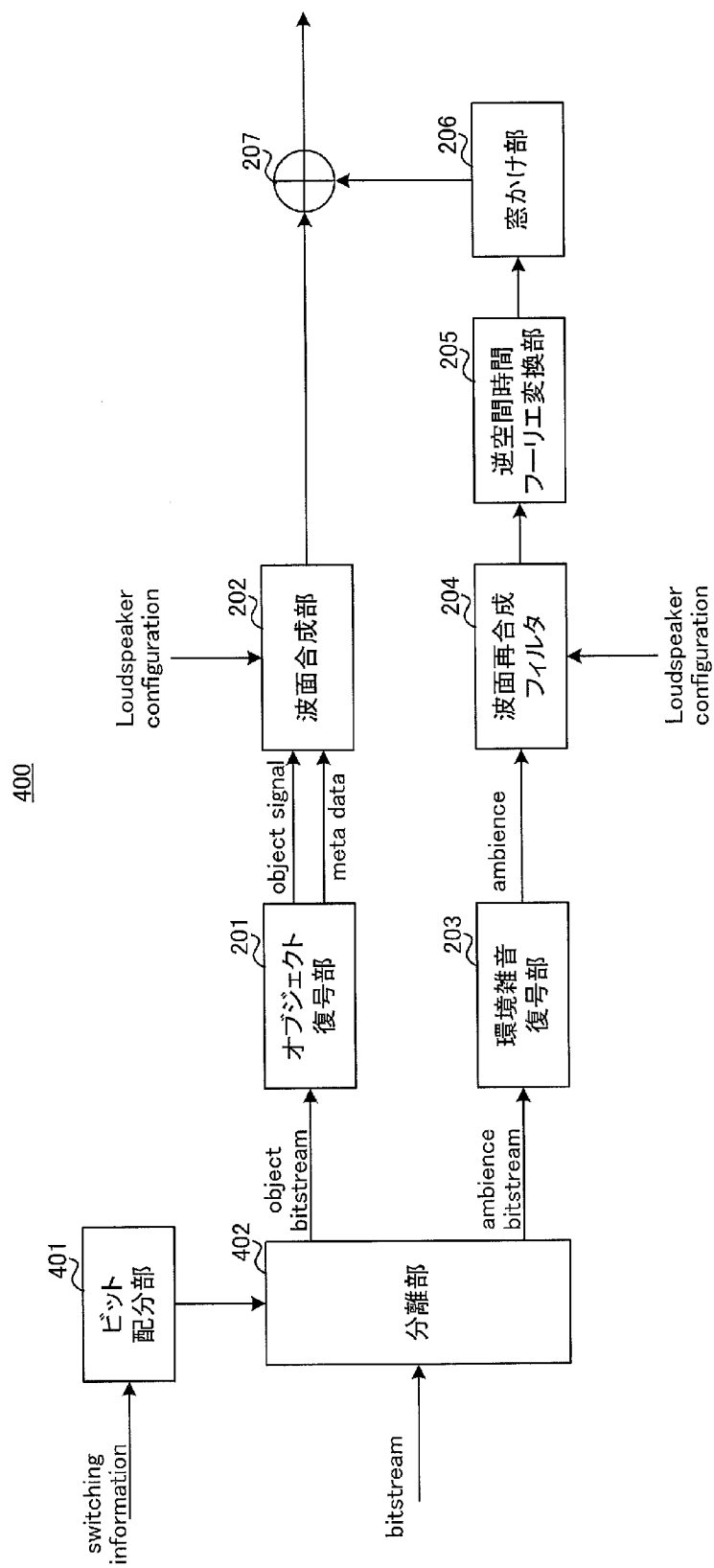
[図8]



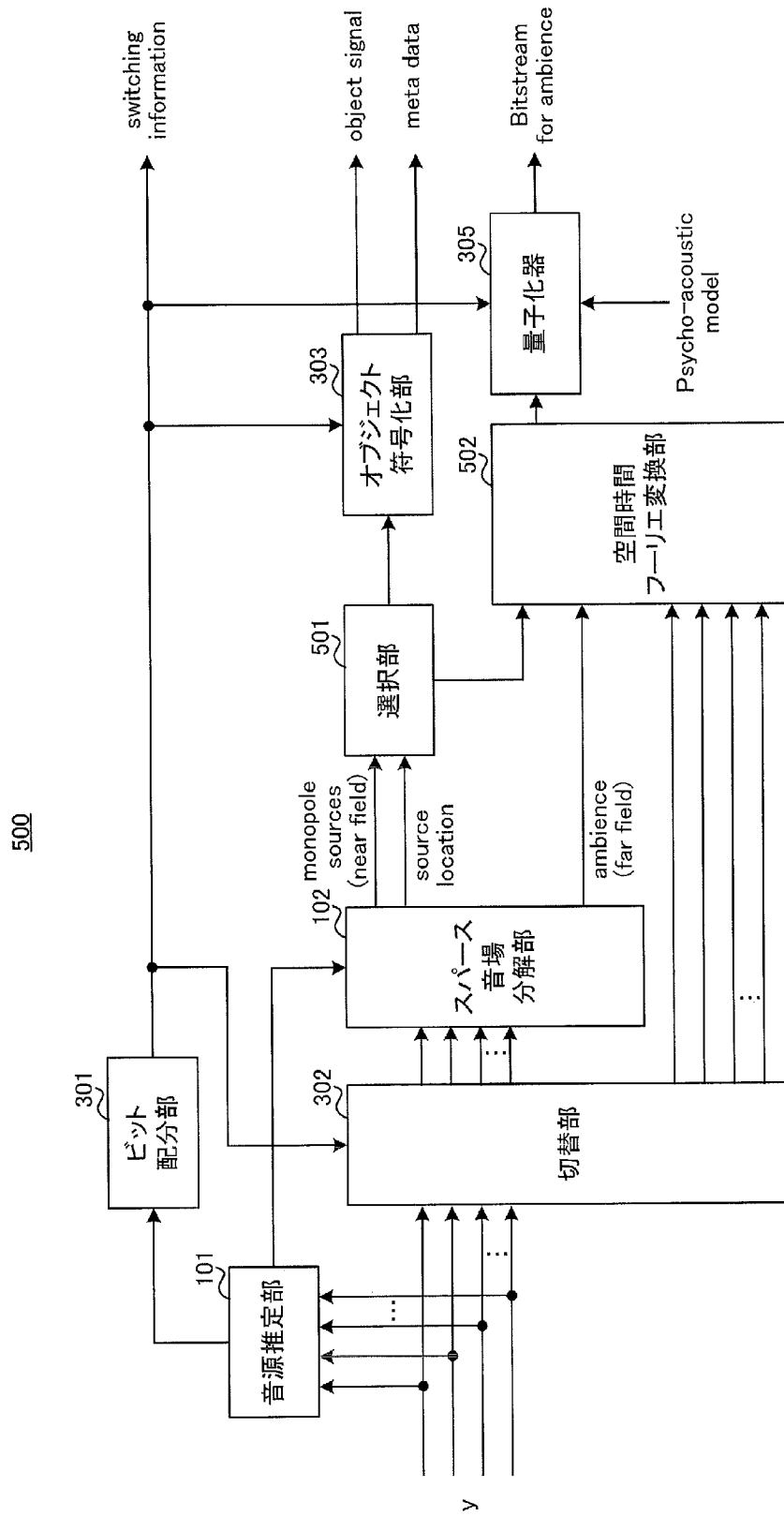
[図9]



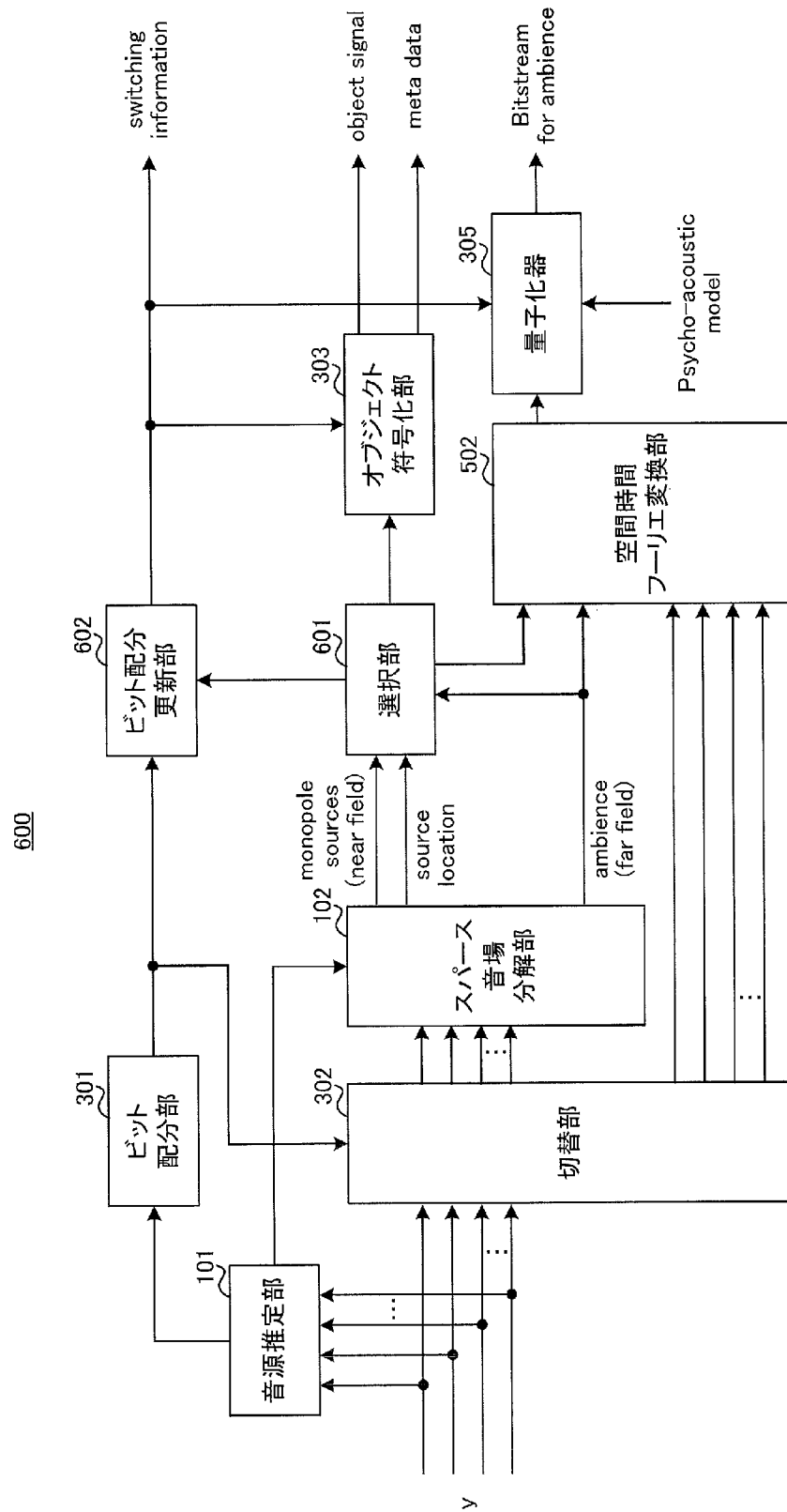
[図10]



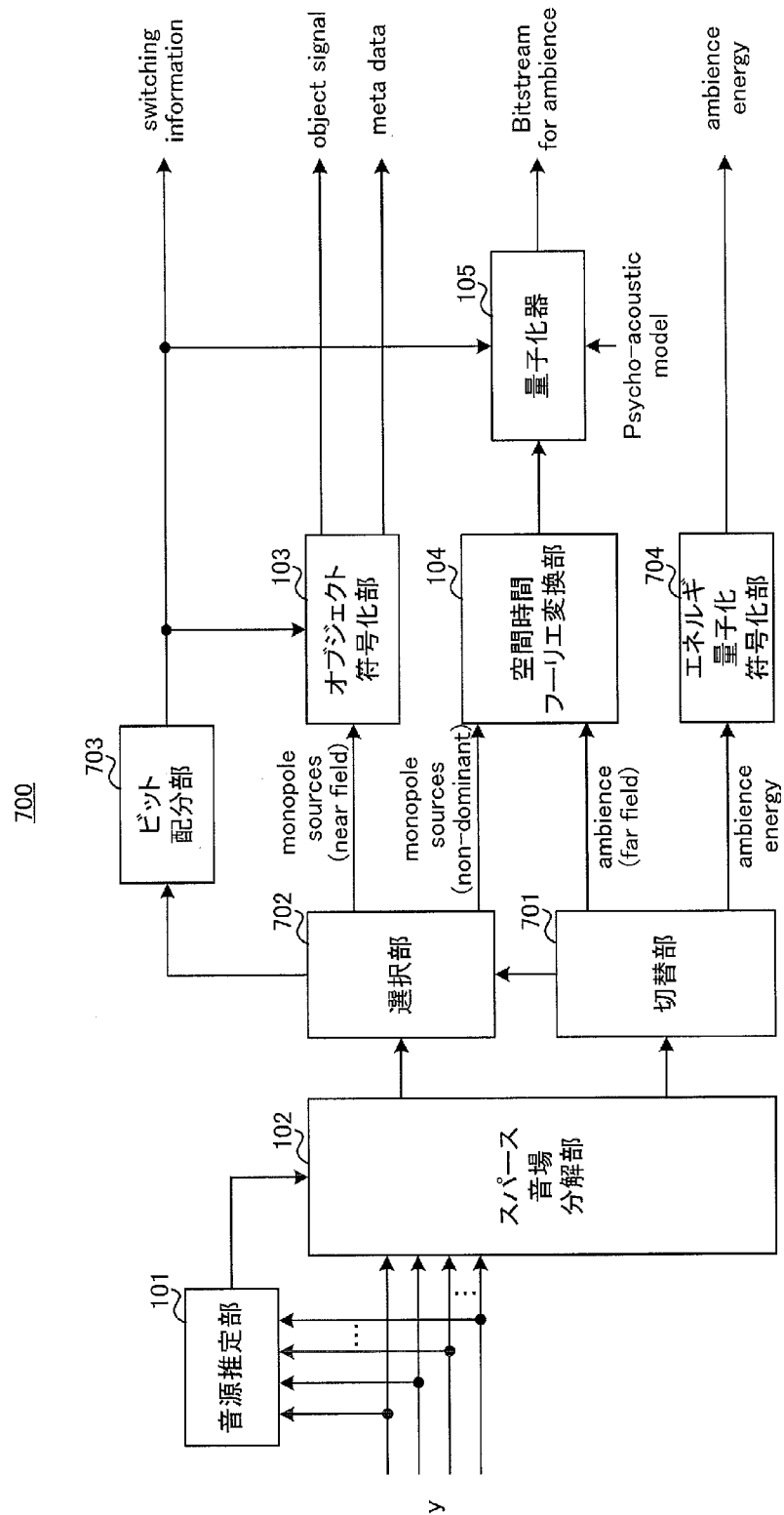
[図11]



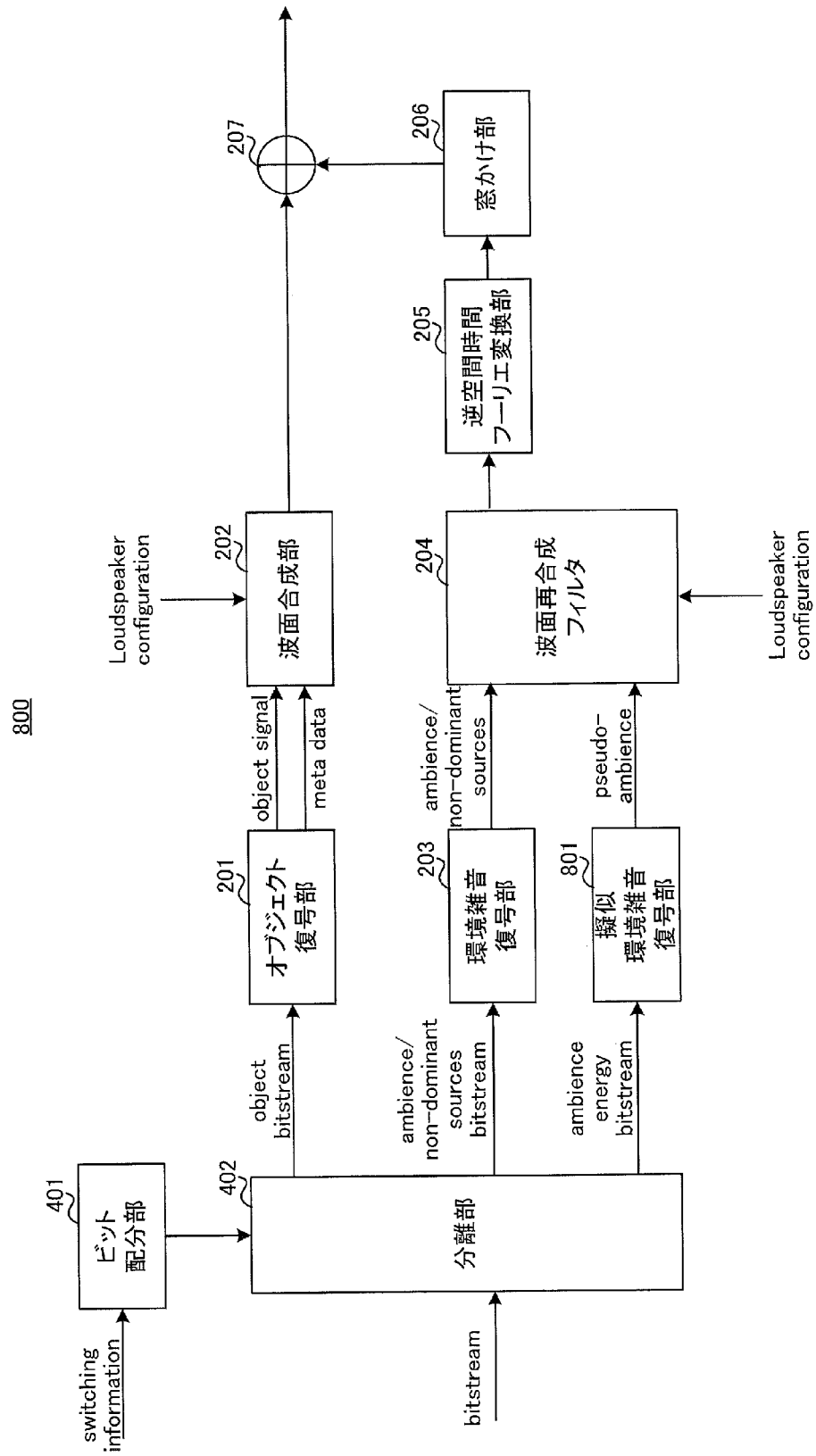
[図12]



[図13]



[図14]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2018/015790

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G10L19/008 (2013.01) i, G10L21/0272 (2013.01) i, G10L25/51 (2013.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl. G10L19/00-99/00, H04S3/00-5/02

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 1971-2018

Registered utility model specifications of Japan 1996-2018

Published registered utility model applications of Japan 1994-2018

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2016-524721 A (FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V.) 18 August 2016, entire text & US 2016/0064006 A1, entire text & EP 2804176 A1 & WO 2014/184115 A1 & KR 10-2016-0009631 A & CN 105378832 A	1-7
A	JP 2015-171111 A (NIPPON TELEGRAPH AND TELEPHONE CORP.) 28 September 2015, entire text (Family: none)	1-7



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"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)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"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

"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

"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

"&" document member of the same patent family

Date of the actual completion of the international search

26 June 2018 (26.06.2018)

Date of mailing of the international search report

10 July 2018 (10.07.2018)

Name and mailing address of the ISA/

Japan Patent Office

3-4-3, Kasumigaseki, Chiyoda-ku,

Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2018/015790

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2011/013381 A1 (PANASONIC CORP.) 03 February 2011, entire text & US 2011/0182432 A1, entire text & EP 2461321 A1 & CN 102171754 A	1-7
A	JP 2008-145610 A (THE UNIVERSITY OF TOKYO) 26 June 2008, entire text (Family: none)	1-7
A	JP 2015-516093 A (QUALCOMM INCORPORATED) 04 June 2015, entire text & US 2013/0300648 A1, entire text & EP 2847650 A1 & CN 104246878 A & KR 10-2015-0016494 A	1-7
A	WO 2016/014815 A1 (DOLBY LABORATORIES LICENSING CORPORATION) 28 January 2016, entire text & US 2017/0215019 A1 & EP 3172731 A1 & CN 105336335 A	1-7

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G10L19/008(2013.01)i, G10L21/0272(2013.01)i, G10L25/51(2013.01)i

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））

Int.Cl. G10L19/00-99/00, H04S3/00-5/02

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	1971-2018年
日本国実用新案登録公報	1996-2018年
日本国登録実用新案公報	1994-2018年

国際調査で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	JP 2016-524721 A（フラウンホーファー・ゲゼルシャフト・ツール・ フェルデルング・デル・アンゲヴァンテン・フォルシュング・アイ ンゲトラーゲネル・フェライン）2016.08.18, 全文 & US 2016/0064006 A1, 全文 & EP 2804176 A1 & WO 2014/184115 A1 & KR 10-2016-0009631 A & CN 105378832 A	1-7
A	JP 2015-171111 A（日本電信電話株式会社） 2015.09.28, 全文（ファミリーなし）	1-7

C欄の続きにも文献が列挙されている。

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

「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矛盾するものではなく、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

26.06.2018

国際調査報告の発送日

10.07.2018

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

山下 剛史

5Z

8946

電話番号 03-3581-1101 内線 3591

C (続き) . 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 2011/013381 A1 (パナソニック株式会社) 2011.02.03, 全文 & US 2011/0182432 A1, 全文 & EP 2461321 A1 & CN 102171754 A	1-7
A	JP 2008-145610 A (国立大学法人 東京大学) 2008.06.26, 全文 (ファミリーなし)	1-7
A	JP 2015-516093 A (クゥアルコム・インコーポレイテッド) 2015.06.04, 全文 & US 2013/0300648 A1, 全文 & EP 2847650 A1 & CN 104246878 A & KR 10-2015-0016494 A	1-7
A	WO 2016/014815 A1 (DOLBY LABORATORIES LICENSING CORPORATION) 2016.01.28, 全文 & US 2017/0215019 A1 & EP 3172731 A1 & CN 105336335 A	1-7