世界知的所有権機関
国際事務局

PCT

特許協力条約に基づいて公開された国際出願

(51) 国際特許分類 5 H04N 7/14, 7/13, H04R 3/00 H04R 3/320	A1	(11) 国際公開番号	WO 94/06246
		(43) 国際公開日	1994年3月17日 (17.03.1994)

(21) 国際出願番号 PCT/JP93/01213
(22) 国際出願日 1993年8月27日(27. 08. 93)

(30) 優先権データ
特願平4/228572 1992年8月27日(27. 08. 92) JP

(71) 出願人(米国を除くすべての指定国について)
株式会社 東芝(KABUSHIKI KAISHA TOSHIBA)〔JP/JP〕
〒210 神奈川県川崎市幸区堀川町72番地 Kanagawa, (JP)
(72) 発明者;および
(75) 発明者/出願人(米国についてのみ)
岡田 理(OKADA, Osamu)〔JP/JP〕
〒165 東京都中野区新井2-8-1-201 Tokyo, (JP)
南 重信(MINAMI, Shigenobu)〔JP/JP〕
〒252 神奈川県綾瀬市綾西4-10-15 Kanagawa, (JP)
山崎彰一郎(YAMASAKI, Shoichiro)〔JP/JP〕
〒154 東京都世田谷区池尻3-1-1-908 Tokyo, (JP)
(74) 代理人
弁理士 鈴江武彦, 外(SUZUYE, Takehiko et al.)
〒100 東京都千代田区霞が関3丁目7番2号
鈴榮内外國特許事務所内 Tokyo, (JP)

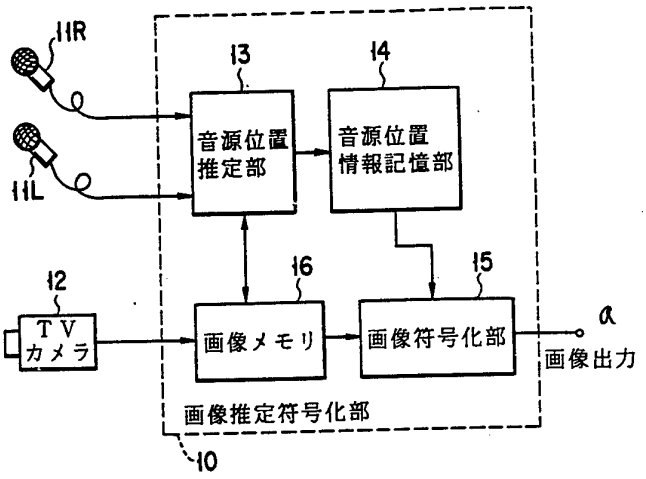
(81) 指定国
CA, US, 欧州特許(DE, FR, GB).

添付公開書類 国際調査報告書
補正書

(54) Title : MOVING PICTURE ENCODER

(54) 発明の名称 動画像符号化装置

10 ... image estimation encoding unit
12 ... television camera
13 ... sound source position estimation unit
14 ... sound source position information storage
15 ... image encoding unit
16 ... image memory
a ... image output



(57) Abstract

An encoder comprising a television camera (12) which images an object and generates corresponding signals; plural microphones (11L and 11R) which are apart from each other to collect sound waves of voice of the object imaged by the television camera and output voice signals; presuming circuit (13) which presumes the position of the sound source based on the voice signals obtained from the microphones; and an encoding circuit which encodes the image signals in an image area of a given extent whose center is in the position of the sound source presumed by the presuming circuit by assigning them codes of an amount slightly greater than those assigned to image signals in the other imaging areas so that the resolution of the image of the image area of the given extent may be higher than that of the other areas.

(57) 要約

被写体を撮像して画像信号化するテレビカメラ１２と、互いに離間して配置され、テレビカメラにより撮像される被写体の音声を収音して音声信号を出力する複数のマイクロフォン１１Ｌ，１１Ｒと、これらマイクロフォンから得られた音声信号から音源位置を推定する推定回路１３と、推定回路により推定された音源位置を中心に所定の範囲の画像領域を他の画像領域よりも高い解像度となるように所定範囲の画像領域の映像信号を他の画像領域の映像信号よりも多めに符号量を割り当てて符号化する符号化回路とにより構成する。

情報としての用途のみ
PCTに基づいて公開される国際出願のパンフレット第1頁にPCT加盟国を同定するために使用されるコード

AT オーストリア
AU オーストラリア
BB バルバードス
BE ベルギー
BF ブルキナ・ファソ
BG ブルガリア
BJ ベナン
BR ブラジル
BY ベラルーシ
CA カナダ
CF 中央アフリカ共和国
CG コンゴー
CH スイス
CI コート・ジボアール
CM カメルーン
CN 中国
CS チェッコスロヴァキア
CZ チェッコ共和国
DE ドイツ
DK デンマーク
ES スペイン
FI フィンランド
FR フランス
GA ガボン
GB イギリス
GN ギニア
GR ギリシャ
HU ハンガリー
IE アイルランド
IT イタリー
JP 日本
KP 朝鮮民主主義人民共和国
KR 大韓民国
KZ カザフスタン
LI リヒテンシュタイン
LK スリランカ
LU ルクセンブルグ
LV ラトヴィア
MC モナコ
MG マダガスカル
ML マリ
MN モンゴル
MR モーリタニア
MW マラウイ
NE ニジェール
NL オランダ
NO ノルウェー
NZ ニュー・ジーランド
PL ポーランド
PT ポルトガル
RO ルーマニア
RU ロシア連邦
SD スーダン
SE スウェーデン
SI スロヴェニア
SK スロヴァキア共和国
SN セネガル
TD チャード
TG トーゴ
UA ウクライナ
US 米国
UZ ウズベキスタン共和国
VN ヴェトナム

明　　細　　書

動画像符号化装置

技術分野

この発明は、映像信号を符号化するための符号化装置に係わり、特に画像の注目部分を、映像信号とともに得られる音声信号により特定し、その特定した画像領域については割り当て符号量を増大して画像符号化を行う動画像符号化装置に関する。

背景技術

近年、通信技術の進歩に伴い、離れた場所にいながら会議のできる遠隔会議システム（テレビ会議システム）や、個人でも利用できるテレビ電話システムなどが実用に供されるようになっている。

このようなシステムにおいては、電話回線等の通信回線を使用して映像と音声を伝送するが、そのためにチャネル当たりの伝送可能な符号量が制限されるために、この制限される上限の符号量以内に映像信号のデータ量を抑えるべく、符号化して画像情報を伝送する。

単位時間当たりの伝送可能な符号量は十分なものではないから、動画像を伝送する場合には、自然な動きを確保するために１コマ当たりの画像の符号量は伝送レートで決まってしまう。

一般的には、画面全体が均一の解像度になるように符号化を行うが、そのために、相手の顔が不鮮明であったりする弊害がある。通常、人間の感覚では画面全体に注意が行き届くわけではなく、画面中の注目部分に神経が集中する傾向がある。従って、注目部分の画質を向上させれば、他の部分は多少解像度が悪くとも気にならない。

このような観点から、主観的な画質を向上させることを目的として、より重要な情報源である人物の顔領域を他の領域よりも鮮明に表示する符号化方式の検討が行われており、その中にフレーム間差分画像を用いた手法（文献「上野他、“カラー動画ＴＶ電話における顔領域検出方式の一検討”，1989年電子情報通信学会春期全国大会D-92」）等が提案されている。

このシステムは、テレビカメラで通話者を撮影し、これにより得られた映像信号から、画像上の動きのある部分を検出し、検出した領域から発言者の顔領域を推定し、この推定した顔領域について符号量を多めに割り当て、他の領域には割り当て符号量を少なくする符号化処理を施すことにより、人物の顔領域を他の領域よりも鮮明に表示できるようにしている。

このような動画ＴＶ電話における顔領域検出方式を会議システムに応用する場合に、人物以外に動きのあるものが写ってしまったり、あるいは複数の人物が写っていて、それぞれの表情に動きがあるような場合などでは発言者の顔領域の推定が困難になってしまう。

このように、複数の人物が写っていたり、人物以外に動きのあるものが写っている場合に、動画像の顔領域を検出する方式で最も重要である発言者の顔領域のみを抽出することができないと云う欠点がある。

そこで、この発明の目的は、映像信号中の発言者の位置を精度良く推定することができて、画面中の発言者の領域を適確に抽出でき、従って、発言者の写っている領域を鮮明に表示することができる動画像符号化装置を提供することにある。

発明の開示

この発明によると、映像信号を符号化して伝送する画像伝送装置において、被写体を撮像し、映像信号を発生するテレビカメラと、このテレビカメラにより撮像される被写体の音声を収音し、音声信号を出力し、互いに離間して配置された複数のマイクロフォンと、これら複数のマイクロフォンから得られた音声信号から音源位置を推定する音源位置推定回路と、この音源位置推定回路により推定された音源位置を中心に所定の範囲の画像領域を高い解像度となるようにこの画像領域に対応する映像信号を、他の画像領域より多めに割り当てた符号量(coded bit rate)で符号化する符号化回路とより構成される動画像符号化装置が提供される。

上記構成の動画像符号化装置によると、テレビカメラが被写体を撮像し、映像信号を出力する。一方、被写体の前に互いに離間して配置された複数のマイクロフォンが音声を収音し、音源位置推定回路は収音した複数チャンネル分の音声信

号に基づいて被写体中の音源位置を推定する。符号化回路はテレビカメラから出力される映像信号を符号化するに当たり、推定回路により推定された音源位置を中心に所定の範囲の画像領域が高い解像度となるように他の画像領域よりも多めに割り当てた符号量で当該画像領域の映像信号が符号化される。

この結果、画面の音源位置周辺を重点的に高解像度で符号化することができるので、発言者をより鮮明に映し出されるように映像信号を符号化できる動画像符号化が実現できる。特に、推定された音源位置を中心に所定の範囲の画像領域を画面中の被写体顔領域の範囲に合わせておくことにより、発言者の顔領域が高い解像度となるように映像信号を符号化できる。

図面の簡単な説明

図１は、本発明の一実施例に係わるテレビ会議システムの画像符号化部の構成例を示すブロック図。

図２は、本発明の実施例を説明するための図であって、本発明に係わるテレビ会議システムの会議室の構成を示す図。

図３は、図１に示す音源位置推定部の構成を示すブロック図。

図４Ａおよび図Ｂは、図３に示す音源位置推定回路の構成を示す回路図。

図５は、図３に示す音源位置推定回路の推定方法を説明するための図。

図６は、図１に示す画像符号化部の重点符号化領域の決定方法を説明するための図。

図７は、図１に示す画像符号化部のブロック回路図。

発明を実施するための最良の態様

以下、この発明の一実施例について図面を参照しながら説明する。この発明は、複数チャンネルの音声信号から音源位置を推定し、この推定した音源位置周辺を重点的に画像を符号化することにより、発言者をより鮮明なるように符号化する動画像符号化方式を採用した画像符号化装置を提供する。

図２は、この発明の画像符号化装置を有するテレビ会議システムの会議室の概略構成を示しており、この図では、１台のテレビカメラが３名の会議出席者を捉えている。

図２に示すように、会議出席者Ａ１ないしＡ３が座っている机９上には、出席者の音声を取り込むように左右に一定の間隔で２個のマイクロフォン（感音手段）１１Ｒおよび１１Ｌが置かれている。また、机９の前方には、テレビカメラ１２が設けられ、机９に並んで座っている会議出席者Ａ１ないしＡ３の像を捉えて映像信号を発生する。

右用および左用のマイクロフォン１１Ｒおよび１１Ｌによって入力された音声信号およびテレビカメラ１２によって入力された映像信号は、画像処理系である図１に示す画像推定符号化部１０に入力され、ここで１画面当たり所定の符号量に収まるように符号化される。なお、音声信号は図示しない音声信号処理系にも与えられ、ここでディジタル信号に変

換され、符号化された映像信号と共に伝送路へと送り出され、相手側に伝送される。

画像処理系である画像推定符号化部１０は、テレビカメラ１２が捉えた会議出席者Ａ１ないし Ａ３の像のうち、発言者の顔領域の位置を推定してこの推定位置の領域の映像信号を他の領域の映像信号より多めに割り当てた符号量Ｍ（ｉ）で符号化し、他の領域をその残りの符号量Ｍ（０）で符号化する。すなわち、１画面当たりの総符号量Ｍ(total) は決められており、この決められた符号量を推定位置の領域に割り当てる符号量（Ｍ（ｉ））と、その他の領域に割り当てる符号量（Ｍ（０））とに配分される。即ち、Ｍ(total) ＝Ｍ（ｉ）＋Ｍ（０）となる。

この画像推定符号化部１０は図１に示すように、音源位置推定部１３と音源位置情報記憶部１４と画像符号化部１５および画像メモリ１６とからなる。画像メモリ１６はテレビカメラ１２から得られる映像信号をディジタル変換して得た画像データを、画面単位で一時保持するメモリであり、画像を扱うために複数枚分の画像を記憶する容量を有し、画像データを逐次更新記憶する。音源位置推定部１３は音源位置を推定する。即ち、この推定部１３は発言者の位置をマイクロフォン１１Ｒおよび１１Ｌの音声信号出力から推定すると共に、画像メモリ１６に保持された画像データから左右のマイクロフォン１１Ｌおよび１１Ｒ位置から画像上の音源位置、すなわち、発言者領域を推定する。音源位置情報記憶部１４は音源位置推定部１３にて推定された音源位置の情報と、推

定を行った時刻の情報とを記憶する。このとき、時刻情報は外部から与えるか、若しくは画像推定符号化部１０に時計回路を設けてこれより得るようにする。

画像符号化部１５は音源位置情報記憶部１４の情報を用いて画像メモリ１６に保持された画像データを符号化して出力する。即ち、これは、映像信号のうち、発言者位置を中心とした領域をより鮮明に表示するような符号化を行う。そのために、画像符号化部１５は、音源位置情報記憶部１４に記憶された発言者位置情報をもとに、その発言者位置の画像上の領域を重点符号化領域として決定し、この重点符号化領域の映像信号に対して符号量Ｍ（ｉ）を割り当て、他の領域の映像信号に対して符号量Ｍ（０）を割り当てて、その割り当てられた範囲内に収まるように、それぞれの領域の映像信号を符号化する。

音源位置推定部１３は図３に示すように、遅延回路３１、推定回路３２、減算回路３３および音源位置推定回路３４とから構成されている。これらのうち、遅延回路３１は左マイクロフォン１１Ｌで得た左チャンネルの音声入力信号を遅延し、推定回路３２はこの遅延回路３１から出力する遅延左チャンネルの音声入力信号と右マイクロフォン１１Ｒで得た右チャンネル音声信号とから左チャンネル音声信号を推定する。減算回路３３は遅延回路３１から出力する遅延左チャンネル音声信号と推定回路３２から出力する推定左チャンネル音声信号を入力とし、左チャンネル音声信号から推定左チャンネル音声信号を差し引いてその差信号を得る回路である。

この差信号が推定回路３２にフィードバックされることにより、推定回路３２は当該差信号が零になるような推定左チャンネル音声信号を推定して出力し、それにより、推定回路３２は遅延左チャンネルの音声入力信号を参照して右マイクロフォン１１Ｒで得た右チャンネル音声信号から左チャンネル音声信号を推定インパルス応答系列Ｈ（ｋ）として推定できることになる。音源位置推定回路３４は推定回路３２の求めた推定インパルス応答系列Ｈ（ｋ）を用いて音源位置を推定する。

上記のような構成において、テレビカメラ１２により会議出席者を撮影し、同時に机９上のマイクロフォン１１Ｒおよび１１Ｌにより音声を収音する。テレビカメラ１２からの映像信号は画像符号化部１５に送られ、マイクロフォン１１Ｒおよび１１Ｌからの音声信号は音源位置推定部１３に送られる。音源位置推定部１３はこの音声信号をもとに、音源の位置を推定し、推定結果は音源位置情報記憶部１４に記憶される。

画像符号化部１５はこの音源位置情報記憶部１４に記憶されている最新の音源位置情報を利用してテレビ画像における画面上の音源位置対応領域を特定し、その領域を予め設定した符号量Ｍ（ｉ）で、また、その他の領域は符号量Ｍ（０）で符号化し、伝送する。これにより、テレビ会議に出席している人のうち、発言者が受信側のモニタ（図示せず）に高解像度で表示できる。

発言者の特定についてもう少し具体的に説明する。

図３において、発言者Ａ１が発した音声をＸ（ω）とすると、この音声Ｘ（ω）はマイクロフォン１１Ｒおよび１１Ｌに収音される。いま、この音声Ｘ（ω）を発することにより右側のマイクロフォン１１Ｒの入力音声信号をＹＲ（ω）および左側のマイクロフォン１１Ｌの入力音声信号をＹＬＯ（ω）とすると、これら入力音声信号ＹＲ（ω）およびＹＬＯ（ω）は発生源からマイクロフォンまでの音声の伝播遅延並びに室内の音声特性で決まる伝達関数ＦＲ（ω）およびＧＬ（ω）により以下のように表わされる。ただし、ωは角周波数である。

$$YR(\omega) = FR(\omega)\,X(\omega) \qquad \cdots (1)$$

$$YLO(\omega) = GL(\omega)\,X(\omega) \qquad \cdots (2)$$

さらに、左チャンネル入力音声信号ＹＬＯ（ω）は、推定回路３２における因果律を保証するための遅延回路３１によってＣ（ω）なるフラットな遅延を受ける。これによって左チャンネル入力音声信号ＹＬＯ（ω）は、遅延回路３１まで含めた伝達関数ＦＬ（ω）により以下のＹＬ（ω）のように表わすことができる。

$$YL(\omega) = C(\omega)\,GL(\omega)\,X(\omega)$$

$$= FL(\omega)\,X(\omega) \qquad \cdots (3)$$

この左チャンネル入力音声信号ＹＬ（ω）は、減算回路３３に入力される。推定回路３２は、右チャンネルの音声信号ＹＲ（ω）および左チャンネルの音声信号ＹＬ（ω）を用い、右チャンネル音声信号ＹＲ（ω）より左チャンネル音声信号ＹＬ（ω）を得るための伝達関数Ｇ（ω）を次式（４）

に基づき推定し、この伝達関数Ｇ（ω）から推定伝達関数Ｇｐ（ω）を生成する。

$$G(\omega) = GL(\omega) / FR(\omega) \quad \cdots (4)$$

上記伝達関数Ｇ（ω）に対する推定伝達関数Ｇｐ（ω）の生成は、具体的には次のようにして行う。

推定回路３２は、まず右チャンネルの音声信号ＹＲ（ω）を用いて時間領域の推定左チャンネル音声信号ｙｐ（ω）を算出する。推定回路３２は、図４Ａに示すような時間領域の推定左チャンネル音声信号ｙｐ（ｋ）を算出する適応トランスバーサルフィルタ３２ａと、図４Ｂに示すような伝達関数Ｇ（ω）の推定インパルス応答系列Ｈｐ（ｋ）を逐次更新する修正回路３２ｂとにより構成されている。適応トランスバーサルフィルタ３２ａと修正回路３２ｂは図示しないクロック発生源から与えられるシステムクロックに同期して動作する。適応トランスバーサルフィルタ３２ａは、入力音声信号ＹＲ（ω）を順次送って右チャンネル音声信号ｘ（ｋ）ないしｘ（ｋ－ｎ＋１）を各時間成分毎の値に変換するｎタップのシフトレジスタ41_1ないし41_{n-1}と、修正回路３２ｂで修正された各時間成分毎の推定インパルス応答ｈｐ１（ｋ）ないしｈｐｎ（ｋ）と上記シフトレジスタ４１１ないし41_{n-1}を経て得られる右チャンネル音声信号ｘ（ｋ）ないしｘ（ｋ－ｎ－１）との各成分毎の乗算を行う乗算器42_1ないし42_nと、この乗算結果の総和（Σ）を求めて推定左チャンネル音声入力信号ｙｐ（ｋ）を得る加算器４３とより構成されている。

すなわち、修正回路３２ｂは後述する（１０）式の演算を行って推定インパルス応答系列ｈｐ１（ｋ）ないしｈｐｎ（ｋ）を求め、それらを時間成分別に分けて適応トランスバーサルフィルタ３２ａの対応する乗算器42_1ないし42_nに与える。乗算器42_1ないし42_nは推定インパルス応答系列ｈｐ１（ｋ）ないしｈｐｎ（ｋ）とシフトレジスタ41_1ないし41_{n-1}を経て得られる右チャンネル音声信号ｘ（ｋ）ないしｘ（ｋ－ｎ＋１）とを各成分毎に乗算し、時間成分別の推定左チャンネル音声信号を得る。これら時間成分別推定左チャンネル音声信号を加算器４３が加算することによって推定左チャンネル音声信号ｙｐ（ｋ）を求める。

このような推定回路３２において、まず右チャンネル音声信号ｘ（ｋ）は一段当たり１サンプル時間分の遅延時間を有するｎ段のシフトレジスタ41_1ないし41_{n-1}に入力され、次式（５）に示すような時系列ベクトルが生成される。

$$X(k)=(x(k),\ x(k-1),\ \cdots,\ x(k-n+1))^T \quad \cdots(5)$$

但し、$(\ )^T$は転置ベクトルを示す。

一方、推定伝達関数Ｇｐ（ω）に時間領域において近似した推定インパルス応答系列Ｈｐ（ｋ）を次式（６）により表すと、

$$Hp(k)=(hp1(k),\ hp2(k),\ \cdots,\ hpn(k))^T \quad \cdots(6)$$

次式（７）により左チャンネル音声信号ｙ（ｋ）の推定値である推定左チャンネル音声信号ｙｐ（ｋ）を得ることができ

る。

$$yp(k) = Hp(k)^T \cdot X(k) \quad \cdots (7)$$

このとき、伝達関数G（ω）のインパルス応答系列Hが次式（8）により表されると（但し、nは任意の整数）、伝達関数の推定が良好に行われることになる。

$$H = (h1, h2, \cdots, hn)^T \quad \cdots (8)$$

従って、推定インパルス応答系列Hp（k）が

$$Hp(k) = H \quad \cdots (9)$$

となったときに、推定左チャンネル音声信号yp（k）が実際の左チャンネル音声信号y（k）にかなり近似する。

従って、（9）式の関係が得られる伝達関数G（ω）になるような推定伝達関数Gp（ω）を探せば良いわけであり、それには推定伝達関数Gp（ω）が伝達関数G（ω）になるような推定インパルス応答系列Hp（k）を推定できれば良いことになる。

ここで、推定回路32における推定インパルス応答系列Hp（k）の推定は、適応トランスバーサルフィルタ32aにおいて、n段のシフトレジスタ41_1ないし41_{n-1}の入出力として得られる時系列ベクトルx（k）ないしx（k－n＋1）を用いて、修正回路32bにより逐次的に例えば、以下の演算を行うことにより達成される。

$$Hp(k+1) = Hp(k) + \alpha \cdot e(k) \cdot X(k) / \|X(k)\|^2$$

ただし、Hp（0）＝0

このアルゴリズムは、公知の学習同定法である。なお、

（１０）式において、ｅ（ｋ）は、図３の減算回路３３の出力であり、この出力ｅ（ｋ）は推定左チャンネル音声信号をｙｐ（ｋ）とすると、つぎの（１１）式の関係を有している。

$$e(k) = y(k) - yp(k) \quad \cdots (11)$$

従って、減算回路３３の出力ｅ（ｋ）は左チャンネル音声信号ｙ（ｋ）に対する推定左チャンネル音声信号ｙｐ（ｋ）との差信号と云うことになる。また、（１０）式において、αは（１０）式の収束速度および安定性を決定する係数となり、音源５１から左右のマイクロフォン１１Ｌおよび１１Ｒ位置までの距離差を表す。

従って、画像推定符号化部１０では、画像メモリ１６に保持された画像データから左右のマイクロフォン１１Ｌおよび１１Ｒ位置を調べ、距離差αを求め、これと減算回路３３の出力ｅ（ｋ）を利用して修正回路３２ｂは（１０）式の演算を行うことにより、推定インパルス応答系列Ｈｐ（ｋ）の推定を行うことができる。

以上の処理により求められた推定インパルス応答系列Ｈｐ（ｋ）から、音源位置推定回路３４により音源位置が推定される。この推定はつぎのようにして行う。

推定インパルス応答系列Ｈｐ（ｋ）の係数のうち、最大値をとる項をM_xとする。このとき、サンプリング周期をＴ（sec）、音速をｖ（m/sec）、タップ数をｎとすると、音源から左右のマイクロフォン１１Ｌおよび１１Ｒ位置までの距離差αは次式（１２）により推定できる。

$$\alpha = v \cdot T (M_x - N/2) \quad \cdots (12)$$

ここで、図５に示すように、左右のマイクロフォン１１Ｌおよび１１Ｒを直線５２で結び、この直線５２に平行な直線５３を想定し、音源５１の位置が左右のマイクロフォン１１Ｌおよび１１Ｒから一定距離離れた直線５３上に存在すると仮定する。このとき、直線５２における左右のマイクロフォン１１Ｌおよび１１Ｒの中心点位置Ｐｏを通り、直線５２に垂直な線５４と直線５３との交点から音源５１までの距離をａ、右マイクロフォン１１Ｒから音源５１までの直線距離をｂ、音源５１を通る直線５３とマイクロフォン１１Ｌおよび１１Ｒを通る直線５２との間の垂線の長さをｃ，マイクロフォン１１Ｌおよび１１Ｒ間の距離を２ｄとすると、次の連立方程式が成り立つ。

$$(b+a)^2 = (d+a)^2 + c^2$$

$$b^2 = (d-a)^2 + c^2 \quad \cdots (13)$$

この連立方程式からｂを消去してａを解くことにより、音源位置Ｐａを推定することができる。

上記のように推定された音源位置Ｐａのデータが音源位置情報記憶部１４を介して画像符号化部１５に入力されると、音源位置を中心とする画像領域が重点符号化領域としてこの領域に対応する画像データが他の領域の画像データよりも多い符号化量で符号化される。この符号化について詳しく説明する。

画像メモリ１６には、図６に示されるように１フレームの画像データが、例えば、１ブロックを８画素×８ラインとして４４×３６ブロックに分けて記憶されている。この画像メ

モリ１６に記憶された画像データはブロック単位で順次に画像符号化部１５に送られる。画像符号化部１５は、図７に示されるように画像メモリ１６の読み出し端子に接続される直交変換（ＤＣＴ）回路７１と、ＤＣＴ回路７１の出力端子に接続される量子化回路７２と、量子化回路７２の出力端子に接続される可変長符号化回路７３と、量子化回路７２の制御端子に接続される量子化ステップサイズ決定回路７４とにより構成される。この画像符号化部１５には、更にマーカ認識回路７５および重点符号化領域決定回路７６が含まれている。マーカ認識回路７５は画像メモリ１６から読み出された画像データから左右マイクロフォン１１Ｌおよび１１Ｒの位置に対応付けて設けられた２つのマーカ６１ａおよび６１ｂを認識し、画面上でのマイクロフォン１１Ｌと１１Ｒとの間の距離２ｄ’を求める。なお、マーカはマイクロフォンを会議室に配置したときにオペレータによって装置に入力される。

求められた距離２ｄ’の情報が重点符号化領域決定回路７６に入力されることにより、この回路７６はこの距離（２ｄ’）情報と音源位置情報記憶部１４から読み出される音源位置情報とから距離２ｄ’の中心から発言者の位置６２までの距離ａ’を次式１４により求める。

$$a' = a \cdot d' / d \qquad \cdots (14)$$

更に、重点符号化領域決定回路７６は発言者位置６２を中心として予め設定した幅２ｗ´の領域６３を重点符号化領域と決定する。この重点符号化領域に関する情報がステップサイズ決定回路７４に入力されると、このステップサイズ決定回

路７４は重点符号化領域の画像データを、他の領域の画像データよりも多くの符号量で符号化するためのステップサイズを決定する。決定されたステップサイズの情報が量子化回路７２に入力されると、量子化回路７２は画像メモリ１６から読み出され、ＤＣＴ回路７１により直交変換された画像データを決定されたステップサイズで、即ち符号量で量子化する。この場合、重点符号化領域６３に対応する画像データが量子化回路７２に入力されたときに決定されたステップサイズで量子化されるが、他の領域の画像データは領域６３の画像データに対するステップサイズよりも荒いステップサイズで量子化される。量子化された画像データは可変長符号化回路７３により可変長符号化され、符号化画像データとして出力される。

上記のように符号化された画像データが受信側に送られ、受信モニタに表示されると、発言者の画像が他の画像よりも高解像度で表示される。

上記実施例においては、音源位置情報記憶部１４には音源情報だけ記憶されるとして説明したが次のように時間情報をも記憶しても良い。

即ち、音源位置推定部１３は音源位置推定回路３４において、推定インパルス応答系列Ｈｐ（ｋ）の係数のうち、最大値をとる項をもとに音源位置Ｐａを推定するが、この音源位置推定部１３にて推定された音源位置Ｐａの情報と、推定を行った時刻を、図示しない制御装置の制御のもとに音源位置情報記憶部１４に記憶する。このとき、現在よりｔ時刻だけ

過去の音源位置Ｐａ（ｔ）が、最新の音源位置Ｐａから左右にそれぞれ一定の幅ｗ以内にある場合には過去の音源位置Ｐａ（ｔ）の記憶情報を音源位置情報記憶部１４から消去するように、制御装置によって記憶部１４が制御される。これによって、音源位置情報記憶部１４には現在の発言者の発言位置と、過去に発言した者（Ｎ人）の最後の発言位置の情報が以下のように記憶される。

Ｔ（１），Ｌ（１）

Ｔ（２），Ｌ（２）

：　　：

Ｔ（Ｎ），Ｌ（Ｎ）

ただし、

Ｔ（１）＜Ｔ（２）＜　…　＜Ｔ（Ｎ）　…（１５）

ここで、Ｔ（ｉ）は発言者ｉが最後に発言してから経過した時間、Ｌ（ｉ）は発言者ｉが最後に発言した位置を示すデータである。また、Ｔ（１）は現在の発言者の音声サンプリングにより上記演算処理した時点での時間、Ｌ（１）は現在の発言者の発言した位置を示すデータである。

画像符号化部１５では、音源位置情報記憶部１４に記憶された最新の発言者の位置Ｌ（１）の情報をもとに、上述したように画像の符号化を行う。

いま、画面全体の符号量をＭ、画面全体の幅をＷとし、発言者ｉの重点符号化領域の重要度をＲ（ｉ）、重点符号化領域以外の領域の重要度をＲ（０）とする。このとき、重要度Ｒ（ｉ）およびＲ（０）は自由に設定できるが、より最近

に発言した人に高い重要度を与えるとすると、

$$R(1) > R(2) > \cdots > R(N) > R(0) \quad \cdots (16)$$

となるように設定できる。

このとき、最新の発言者の重点符号化領域（最新の発言者の画像領域）の符号量M（ｉ）、この重点符号化領域以外の領域の符号量M（０）は、

$$M(i) = M \cdot w' \cdot R(i) / RT$$

$$M(0) = M \cdot (Wl - N \cdot w') R(0) / RT$$

となるように割り当てる。ただし、ＲＴは

$$RT = w' (R(1) + R(2) + \cdots + R(N)) + (Wl - N \cdot w') R(0) \quad \cdots (17)$$

従って、このように発言者ｉの重点符号化領域に多めの符号量M（ｉ）を割り当て、他の領域に残りの符号量M（０）を割り当ててその割り当て範囲内で符号化を行うことにより、発言者位置を中心とした領域をより鮮明に表示するような符号化を行うことができるようになるため、画面当たりの総符号量は従来と変わりないが、画面全体の主観的な画質の向上を計ることができるようになる。

以上のように、異なる位置に配置した複数のマイクロフォンでそれぞれ収音した複数チャンネルの音声信号と、マイクロフォンおよび発言者を含めた映像画面上のマイクロフォン位置とから音源位置を推定するようにしたことにより、映像画面上の発言者の画像領域を適確に抽出できるようになり、発言者の画像領域について符号量を多めに割り当てて符号化

するようにしたことにより、発言者の像領域について鮮明に表示することが可能な動画像符号化方式が得られる。

なお、この発明は、上述した実施例に限定されるものではなく、その要旨を変更しない範囲内で適宜変形して実施し得る。

例えば、上述した実施例における音源位置推定部１３の推定回路３２では、時間領域の適応トランスバーサルフィルタを用いているが、周波数領域の適応トランスバーサルフィルタなどの他の回路構成であっても良い。また、推定アルゴリズムとして学習同定法を例にして説明したが、最急降下法などの他の学習アルゴリズムを用いることもできる。

さらに、音源位置推定回路３４において、推定インパルス応答系列Ｈｐ（ｋ）の係数のうち、最大値をとる項をもとに音源位置を推定したが、他の方法を用いても良い。

また、画像符号化部１５の重点符号化領域の決定方法も、上述した方法に限らず、重点符号化領域６３において顔領域検出を行うなどの他の方法であっても良い。また、画像符号化部１５における重要度の設定方法についても、現在までに発言した時間に応じて重要度を設定する方法、最後に発言してからの時間と現在までに発言した時間との両方を加味して重要度を設定する方法などの他の手法であっても構わない。

また、テレビ会議システムでは、被写体はほぼ定位置固定であり、テレビ画面は被写体に対して同一視野角を維持するために、画面上の被写体は被写体自身が移動しない限り、位置に変化はないから、画像符号化部１５における重要度の設

定や重点符号化領域の設定を外部から行うことにより、例えばＶＩＰは常に高精細に符号化を行うようにすることもできる。更には、画面と被写体との関係が変わらないことから、発言者の画像領域ではなく、発言者の顔部分の領域を特定することは容易であるから、この特定した顔領域について解像度を高めるように符号量を割り当てる構成とすることもできる。

さらに、画像符号化部１５の符号化方式についても、上述の実施例において、各フレーム毎に重点符号化領域６３に多くの符号量を与えて精細に符号化を行う手法について述べたが、重点符号化領域６３以外の部分をコマ落としの状態にすることによって重点符号化領域６３に多くの符号量を与えて精細に符号化を行っても良い。また、発言者の履歴に対応して最新の発言者ほど高解像度にし、発言の古い者ほど低解像度にする等の、発言順位等に対応させた重み付けにより解像度を変えるなどしても良い。

なお、上述した実施例では、音声入力を２チャンネルとしたが、３チャンネル以上あっても良い。この場合、マイクロフォンの配置に上下方向の高低差を与えることにより、音源位置の２次元の推定が可能となり、この場合には音源として画面上の一点を推定することができるようになり、より高精度の音源位置推定が可能となる。

産業上の利用可能性

上述した発明によれば、複数チャンネルの音声信号から音

源位置を推定し、音源位置周辺を重点的に符号化することにより、発言者をより鮮明なるように符号化する動画像符号化方式を提供できる。

請　求　の　範　囲

1．音声を発する少なくとも1つの被写体を撮像して映像信号を出力する撮像手段と、

互いに離間して配置配され、前記撮像手段により撮像される被写体の音声を収音して音声信号を出力する複数の感音手段と、

前記複数の感音手段から出力される音声信号から音源位置を推定する推定手段と、

前記推定手段より推定された音源位置を中心に所定の範囲の画像領域に対応する映像信号を、他の画像領域に対応する映像信号よりも多く符号量を割当て符号化する符号化手段と、

により構成される動画像符号化装置。

2．前記感音手段は、複数の被写体に対して左右に配置され、左右チャンネルの音声信号を発生する左右マイクロフォンにより構成され、前記推定手段は、前記左マイクロフォンから出力される左チャンネル音声信号を遅延する遅延回路と、前記遅延回路から出力する遅延左チャンネル音声信号と前記右マイクロフォンから出力される右チャンネル音声信号とから左チャンネル音声信号を推定する推定回路と、前記遅延回路から出力する遅延左チャンネル音声信号と前記推定回路から出力する推定左チャンネル音声信号との差信号を得る減算回路と、前記推定回路が前記差信号をフィードバックにより受けたとき前記差信号が零になるような推定左チャンネル音声

信号を推定し、前記推定回路から出力される推定インパルス応答系列を用いて音源位置を推定する音源位置推定回路とにより構成される請求項１記載の動画像符号化装置。

３．前記推定回路は、時間領域の推定左チャンネル音声信号を算出する適応トランスバーサルフィルタと、推定インパルス応答系列を逐次更新する修正回路とにより構成される請求項２記載の動画像符号化装置。

４．前記適応トランスバーサルフィルタは、右チャンネル音声信号を順次送って右チャンネル音声信号を各時間成分毎の値に変換するｎタップのシフトレジスタと、前記修正回路により修正された各時間成分毎の推定インパルス応答と前記シフトレジスタを経て得られる右チャンネル音声信号との各成分毎の乗算を行う乗算器と、乗算結果の総和を求めて推定左チャンネル音声入力信号を得る加算器とより構成される請求項３記載の動画像符号化装置。

５．前記修正回路は、推定インパルス応答系列を求め、それらを時間成分別に分けて前記適応トランスバーサルフィルタの対応する乗算器に与える回路手段を有し、前記適応トランスバーサルフィルタの前記乗算器は推定インパルス応答系列と前記シフトレジスタを経て得られる右チャンネル音声信号とを各成分毎に乗算し、時間成分別の推定左チャンネル音声信号を出力し、これら時間成分別推定左チャンネル音声信号を前記加算器が加算することによって推定左チャンネル音声信号を求める請求項４記載の動画像符号化装置。

６．音声を発する少なくとも１つの被写体を撮像して映像信

号を出力する撮像手段と、

互いに離間して配置され、前記撮像手段により撮像される被写体の音声を収音して音声信号を出力する複数の感音手段と、

前記複数の感音手段から得られた音声信号から音源位置を推定する推定手段と、

前記推定手段により推定された現在及び過去の音源位置情報の履歴を記憶する音源位置記憶手段と、

前記音源位置記憶手段に記憶された音源位置情報及び過去の音源位置情報の履歴を利用して位置に応じた符号量で映像信号を符号化する符号化手段と、

により構成される動画像符号化装置。

7．前記画像符号化手段は、前記音源位置記憶手段に記憶された少なくとも1つの音源位置とその近傍を高画質領域として各々の画質レベルを設定し、他の領域より前記画質レベルに応じて高画質になるように符号量を割り当てて映像信号を符号化する請求項6記載の動画像符号化装置。

8．前記画像符号化手段は、外部より高画質領域と画質レベルを設定し、他の領域より高画質になるように符号量を割り当てて映像信号を符号化する機能を有する請求項6記載の動画像符号化装置。

9．前記音源位置推定手段は、前記複数チャンネルの音声信号の遅延差と位相差、レベル差の少なくとも一方を利用して検出する請求項6記載の動画像符号化装置。

10．前記画像符号化手段は、音源位置の出現頻度に応じて

画質レベルを設定する請求項８記載の動画像符号化装置。

１１．前記感音手段は、複数の被写体に対して左右に配置され、左右チャンネルの音声信号を発生する左右マイクロフォンにより構成され、前記推定手段は、前記左マイクロフォンから出力される左チャンネル音声信号を遅延する遅延回路と、前記遅延回路から出力する遅延左チャンネル音声信号と前記右マイクロフォンから出力される右チャンネル音声信号とから左チャンネル音声信号を推定する推定回路と、前記遅延回路から出力する遅延左チャンネル音声信号と前記推定回路から出力する推定左チャンネル音声信号との差信号を得る減算回路と、前記推定回路が前記差信号をフィードバックにより受けたとき前記差信号が零になるような推定左チャンネル音声信号を推定し、前記推定回路から出力される推定インパルス応答系列を用いて音源位置を推定する音源位置推定回路とにより構成される請求項６記載の動画像符号化装置。

１２．前記推定回路は、時間領域の推定左チャンネル音声信号を算出する適応トランスバーサルフィルタと、推定インパルス応答系列を逐次更新する修正回路とにより構成される請求項１１記載の動画像符号化装置。

１３．前記適応トランスバーサルフィルタは、右チャンネル音声信号を順次送って右チャンネル音声信号を各時間成分毎の値に変換するｎタップのシフトレジスタと、前記修正回路により修正された各時間成分毎の推定インパルス応答と前記シフトレジスタを経て得られる右チャンネル音声信号との各成分毎の乗算を行う乗算器と、乗算結果の総和を求めて推定

左チャンネル音声入力信号を得る加算器とより構成される請求項１２記載の動画像符号化装置。

１４．前記修正回路は、推定インパルス応答系列を求め、それらを時間成分別に分けて前記適応トランスバーサルフィルタの対応する乗算器に与える回路手段を有し、前記適応トランスバーサルフィルタの前記乗算器は推定インパルス応答系列と前記シフトレジスタを経て得られる右チャンネル音声信号とを各成分毎に乗算し、時間成分別の推定左チャンネル音声信号を出力し、これら時間成分別推定左チャンネル音声信号を前記加算器が加算することによって推定左チャンネル音声信号を求める請求項１３記載の動画像符号化装置。

補正された請求の範囲
[1994年1月6日(06.01.94)国際事務局受理;出願当初の請求の範囲1、2および6は補正された;他の請求の範囲は変更なし。(5頁)]

1. （補正後）音声を発する少なくとも1つの被写体を撮像して映像信号を出力する撮像手段と、

複数の被写体に対して左右に配置される左右マイクロフォンにより構成され、前記撮像手段により撮像される被写体の音声を収音して、左右チャンネルの音声信号を出力する左右チャンネル感音手段と、

前記左マイクロフォンから出力される左チャンネル音声信号を遅延する遅延回路と、前記遅延回路から出力する遅延左チャンネル音声信号と前記右マイクロフォンから出力される右チャンネル音声信号とから左チャンネル音声信号を推定する推定回路とを含み、前記左右チャンネル感音手段から出力される左右チャンネル音声信号から音源位置を推定する推定手段と、

前記推定手段より推定された音源位置を中心に所定の範囲の画像領域に対応する映像信号を、他の画像領域に対応する映像信号よりも多く符号量を割当てて符号化する符号化手段と、

により構成される動画像符号化装置。

2. （補正後）前記推定手段は、前記遅延回路から出力する遅延左チャンネル音声信号と前記推定回路から出力する推定左チャンネル音声信号との差信号を得る減算回路と、前記推定回路が前記差信号をフィードバックにより受けたとき前記差信号が零になるような推定左チャンネル音声信号を推定し、

前記推定回路から出力される推定インパルス応答系列を用いて音源位置を推定する音源位置推定回路とにより構成される請求項1記載の動画像符号化装置。

3. 前記推定回路は、時間領域の推定左チャンネル音声信号を算出する適応トランスバーサルフィルタと、推定インパルス応答系列を逐次更新する修正回路とにより構成される請求項2記載の動画像符号化装置。

4. 前記適応トランスバーサルフィルタは、右チャンネル音声信号を順次送って右チャンネル音声信号を各時間成分毎の値に変換するnタップのシフトレジスタと、前記修正回路により修正された各時間成分毎の推定インパルス応答と前記シフトレジスタを経て得られる右チャンネル音声信号との各成分毎の乗算を行う乗算器と、乗算結果の総和を求めて推定左チャンネル音声入力信号を得る加算器とより構成される請求項3記載の動画像符号化装置。

5. 前記修正回路は、推定インパルス応答系列を求め、それらを時間成分別に分けて前記適応トランスバーサルフィルタの対応する乗算器に与える回路手段を有し、前記適応トランスバーサルフィルタの前記乗算器は推定インパルス応答系列と前記シフトレジスタを経て得られる右チャンネル音声信号とを各成分毎に乗算し、時間成分別の推定左チャンネル音声信号を出力し、これら時間成分別推定左チャンネル音声信号を前記加算器が加算することによって推定左チャンネル音声信号を求める請求項4記載の動画像符号化装置。

6. （補正後）音声を発する少なくとも1つの被写体を撮像

して映像信号を出力する撮像手段と、

互いに離間して配置され、前記撮像手段により撮像される被写体の音声を収音して音声信号を出力する複数の感音手段と、

前記複数の感音手段から得られた音声信号から音源位置を推定する推定手段と、

前記推定手段により推定された現在及び過去の音源位置情報の履歴を記憶する音源位置記憶手段と、

前記音源位置記憶手段に記憶された現在の音源位置情報及び過去の音源位置情報の履歴を利用して位置に応じた符号量で映像信号を符号化する符号化手段と、

により構成される動画像符号化装置。

7. 前記画像符号化手段は、前記音源位置記憶手段に記憶された少なくとも1つの音源位置とその近傍を高画質領域として各々の画質レベルを設定し、他の領域より前記画質レベルに応じて高画質になるように符号量を割り当てて映像信号を符号化する請求項6記載の動画像符号化装置。

8. 前記画像符号化手段は、外部より高画質領域と画質レベルを設定し、他の領域より高画質になるように符号量を割り当てて映像信号を符号化する機能を有する請求項6記載の動画像符号化装置。

9. 前記音源位置推定手段は、前記複数チャンネルの音声信号の遅延差と位相差、レベル差の少なくとも一方を利用して検出する請求項6記載の動画像符号化装置。

10. 前記画像符号化手段は、音源位置の出現頻度に応じて

画質レベルを設定する請求項８記載の動画像符号化装置。

１１．前記感音手段は、複数の被写体に対して左右に配置され、左右チャンネルの音声信号を発生する左右マイクロフォンにより構成され、前記推定手段は、前記左マイクロフォンから出力される左チャンネル音声信号を遅延する遅延回路と、前記遅延回路から出力する遅延左チャンネル音声信号と前記右マイクロフォンから出力される右チャンネル音声信号とから左チャンネル音声信号を推定する推定回路と、前記遅延回路から出力する遅延左チャンネル音声信号と前記推定回路から出力する推定左チャンネル音声信号との差信号を得る減算回路と、前記推定回路が前記差信号をフィードバックにより受けたとき前記差信号が零になるような推定左チャンネル音声信号を推定し、前記推定回路から出力される推定インパルス応答系列を用いて音源位置を推定する音源位置推定回路とにより構成される請求項６記載の動画像符号化装置。

１２．前記推定回路は、時間領域の推定左チャンネル音声信号を算出する適応トランスバーサルフィルタと、推定インパルス応答系列を逐次更新する修正回路とにより構成される請求項１１記載の動画像符号化装置。

１３．前記適応トランスバーサルフィルタは、右チャンネル音声信号を順次送って右チャンネル音声信号を各時間成分毎の値に変換するｎタップのシフトレジスタと、前記修正回路により修正された各時間成分毎の推定インパルス応答と前記シフトレジスタを経て得られる右チャンネル音声信号との各成分毎の乗算を行う乗算器と、乗算結果の総和を求めて推定

左チャンネル音声入力信号を得る加算器とより構成される請求項１２記載の動画像符号化装置。

１４．前記修正回路は、推定インパルス応答系列を求め、それらを時間成分別に分けて前記適応トランスバーサルフィルタの対応する乗算器に与える回路手段を有し、前記適応トランスバーサルフィルタの前記乗算器は推定インパルス応答系列と前記シフトレジスタを経て得られる右チャンネル音声信号とを各成分毎に乗算し、時間成分別の推定左チャンネル音声信号を出力し、これら時間成分別推定左チャンネル音声信号を前記加算器が加算することによって推定左チャンネル音声信号を求める請求項１３記載の動画像符号化装置。

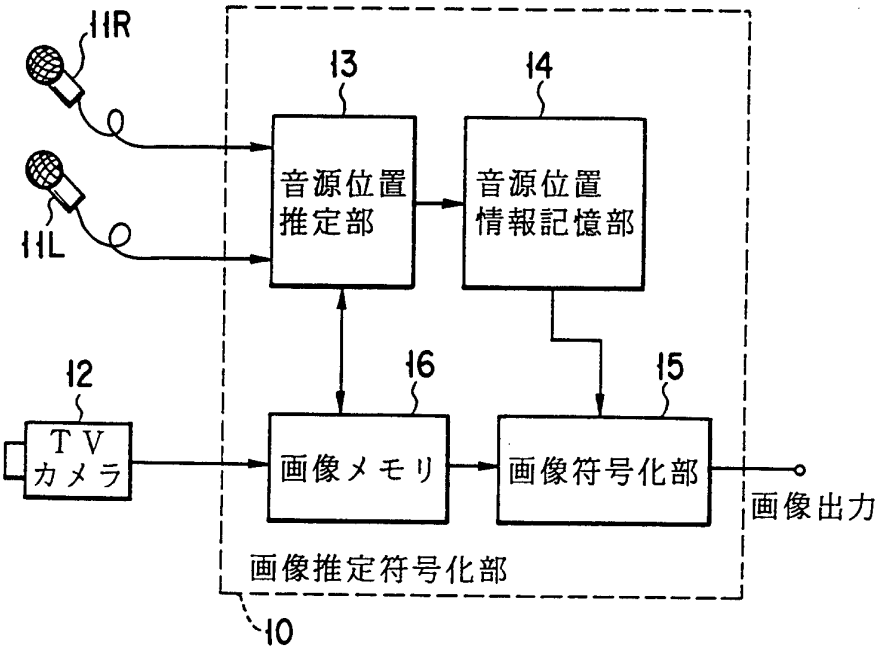

11R
11L
13
14
音源位置
推定部
音源位置
情報記憶部
12
TV
カメラ
16
15
画像メモリ
画像符号化部
画像出力
画像推定符号化部
10

FIG. 1

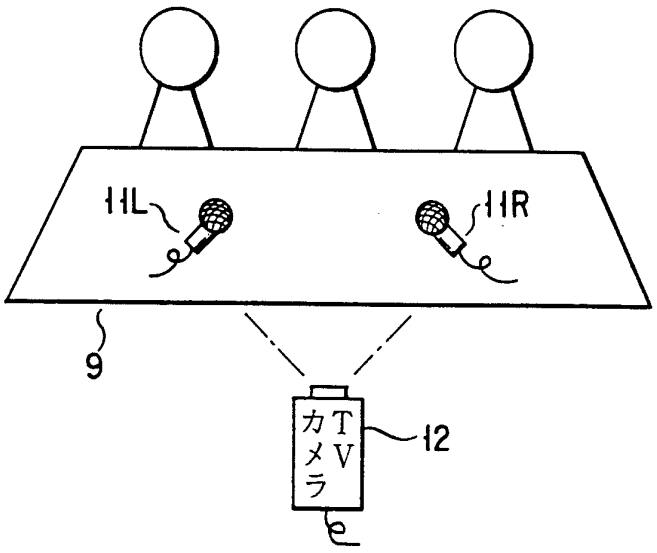

11L
11R
9
TV
カメラ
12

FIG. 2

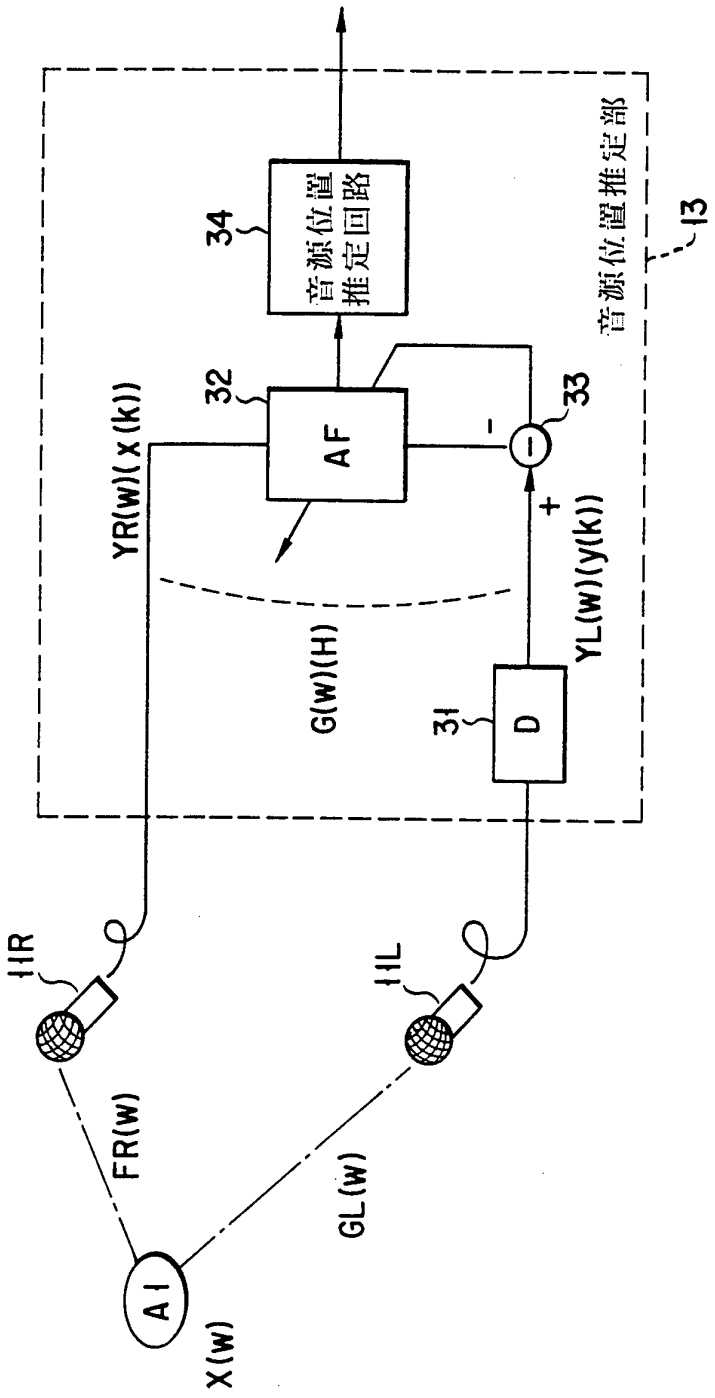

音源位置推定部
13
音源位置推定回路
34
AF
32
33
-
+
YR(w)(x(k))
YL(w)(y(k))
G(w)(H)
D
31
11R
11L
FR(w)
GL(w)
A1
X(w)

FIG. 3

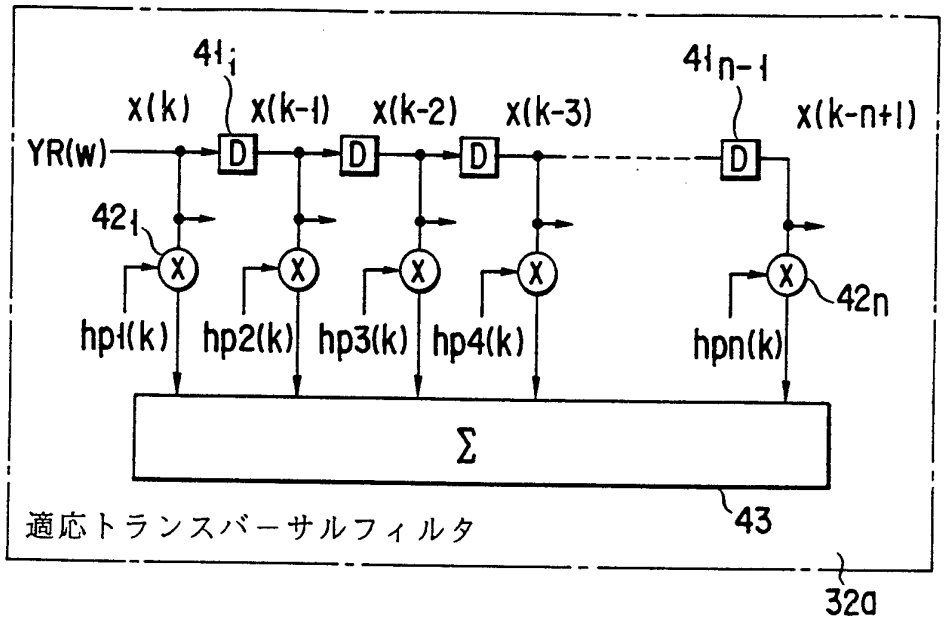

41i
41n-1
x(k)
x(k-1)
x(k-2)
x(k-3)
x(k-n+1)
YR(w)
D
42i
42n
hp1(k)
hp2(k)
hp3(k)
hp4(k)
hpn(k)
Σ
43
適応トランスバーサルフィルタ
32a

FIG. 4A

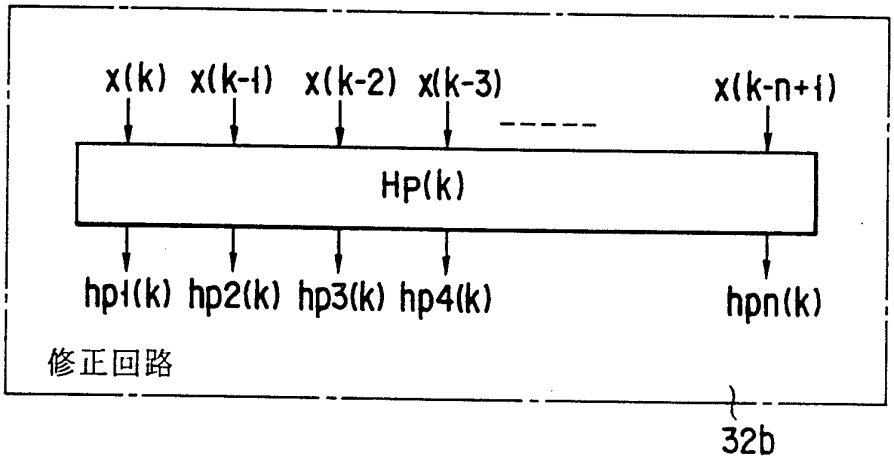

x(k)
x(k-1)
x(k-2)
x(k-3)
x(k-n+1)
Hp(k)
hp1(k)
hp2(k)
hp3(k)
hp4(k)
hpn(k)
修正回路
32b

FIG. 4B

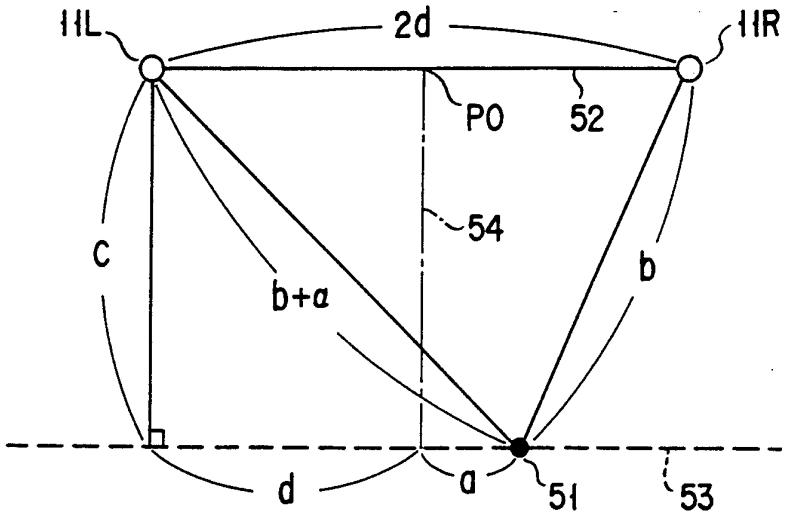


FIG. 5

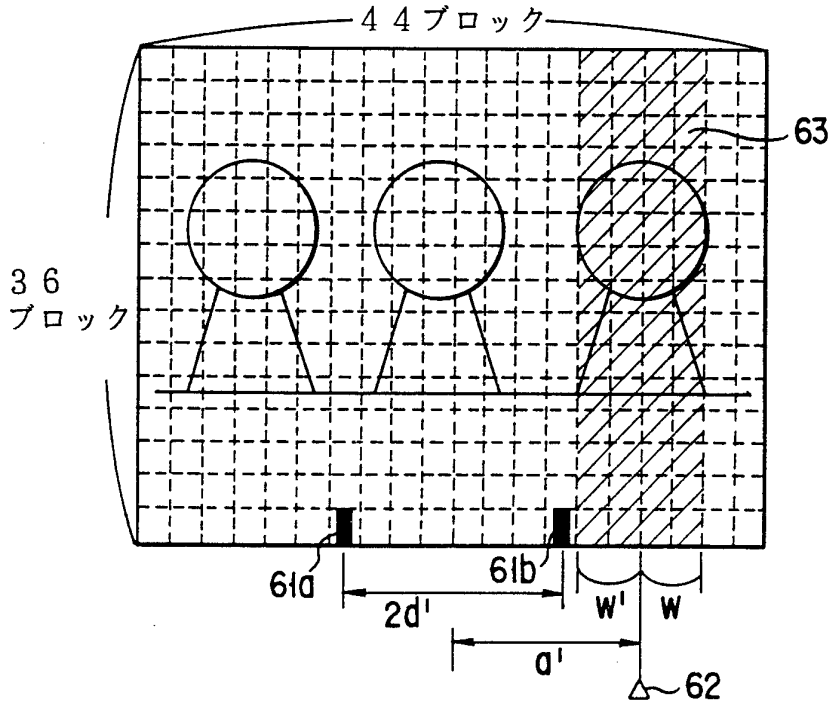


FIG. 6

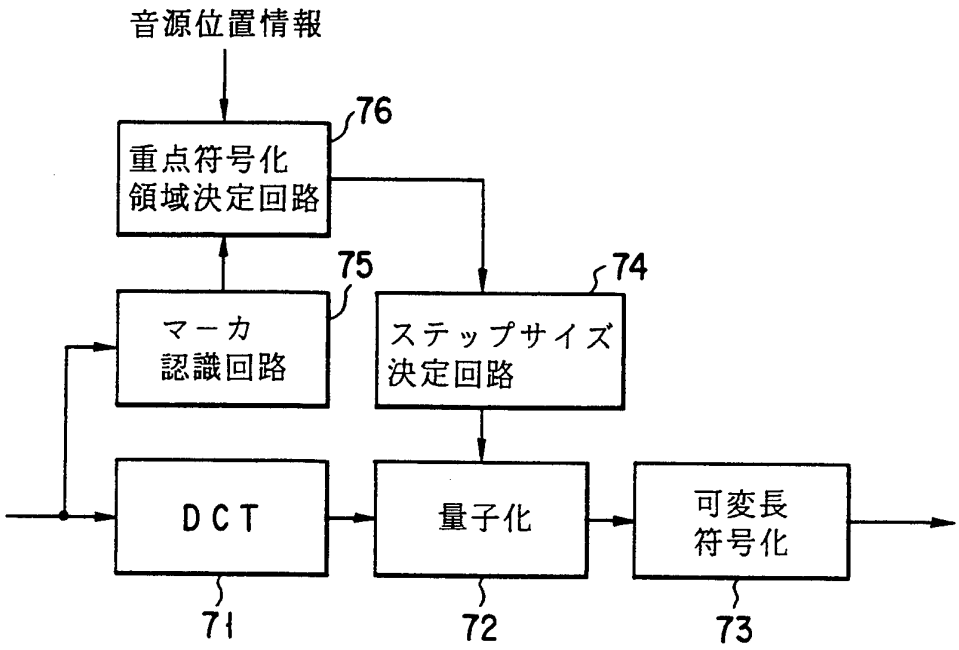

音源位置情報
76
重点符号化
領域決定回路
75
74
マーカ
認識回路
ステップサイズ
決定回路
DCT
量子化
可変長
符号化
71
72
73

FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP93/01213

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl^5 H04N7/14, H04N7/13, H04R3/00, 320

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^5 H04N7/14, H04N7/13, H04R3/00, 320

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1926 - 1993
Kokai Jitsuyo Shinan Koho 1971 - 1993

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.
X A	JP, A, 59-21186 (Westarn Electric Co., Inc.), February 3, 1984 (03. 02. 84), & US, A, 4494144 & DE, A, 3322413 & FR, A, 2529422	1 2-14
A	JP, A2, 3-96999 (Aishin Seiki Co., Ltd., Shin Sangyo Kaihatsu K.K.), April 22, 1991 (22. 04. 91), (Family: none)	2-5, 7-14
A	JP, B2, 63-64120 (Hitachi, Ltd.), December 9, 1988 (09. 12. 88), (Family: none)	2-5, 7-14
A	JP, B2, 61-29163 (Mitsubishi Electric Corp.), July 4, 1986 (04. 07. 86), (Family: none)	3, 12

[X] 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 document 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	Date of mailing of the international search report
October 22, 1993 (22. 10. 93)	November 16, 1993 (16. 11. 93)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

Form PCT/ISA/210 (second sheet) (July 1992)

国際調査報告

国際出願番号 PCT/JP 93/01213

A. 発明の属する分野の分類（国際特許分類（IPC））

Int. Cl⁵ H04N7/14, H04N7/13, H04R3/00, 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））

Int. Cl⁵ H04N7/14, H04N7/13, H04R3/00, 320

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 1926－1993年
日本国公開実用新案公報 1971－1993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	関連する請求の範囲の番号
X A	JP, A, 59－21186（ウエスターン・エレクトリック・カンパニー・インコーポレーテッド）, 3. 2月. 1984（03. 02. 84） &US, A, 4494144&DE, A, 3322413 &FR, A, 2529422	1 2－14
A	JP, A 2, 3－96999（アイシン精機株式会社，株式会社 新産業開発）, 22. 4月. 1991（22. 04. 91）（ファミリーなし）	2－5， 7－14

☑ 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」先行文献ではあるが、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
22. 10. 93

国際調査報告の発送日
16.11.93

名称及びあて先
日本国特許庁（ISA/JP）
郵便番号100
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）
佐藤聡史 印
5C 8943

電話番号 03－3581－1101 内線 3541

様式PCT/ISA/210（第2ページ）（1992年7月）

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	関連する請求の範囲の番号
A	JP，B2，63－64120（株式会社　日立製作所），9．12月．1988（09．12．88）（ファミリーなし）	2－5，7－14
A	JP，B2，61－29163（三菱電機株式会社），4．7月．1986（04．07．86）（ファミリーなし）	3，12