

移動端末のマイクロフォン で情報を取得する 音波情報伝送方式

音声・音楽に情報を埋め込み、スピーカからマイクロフォンに音波で情報を伝送する技術である Acoustic OFDM を開発した。Acoustic OFDM の特徴は、OFDM 伝送信号を聴覚に不快にならないように変形し、音声・音楽に重畳することにある。これにより、従来の音響透かし技術と比較して伝送できる情報量が飛躍的に向上した。

まつおか ほうせい なかしま ゆうすけ よしむら たけし
松岡 保静 中島 悠輔 吉村 健

1. まえがき

2次元コードの普及によって、URL (Uniform Resource Locator) や電話帳などの情報を移動端末のカメラを用いて簡単に読み取ることができるようになった。広告や書籍などにも関連WebサイトのURLを含んだ2次元コードが掲載されるようになり、移動端末からWebサイトへのアクセスも容易になった。このような2次元コードに相当する情報を、テレビやラジオ放送などの音声・音楽に埋め込み、それを移動端末のマイクロフォンで取得できれば、より多くの場面でユーザが移動端末に情報を取り込むことが容易になる。

しかし、音波で情報を伝送する場合、従来の音響透かし^{*1}技術に基づいた手法では1秒間に1文字程度の情報しか送ることができない。これでは簡単なURLを伝送するのにも数十秒もかかってしまい、ユーザがストレスを感じないレスポンスタイムをはるかに超えてしまう。また超音波を用いる手法では、伝送速度は高速にできるが、市販のオーディオ機器で超音波を再生・録音できるものは少なく、さらにテレビやラジオ放送では超音波帯域の信号を送ることはできないため、利用できる場面がほとんどなくなる。

以上のような背景から、次の条件を満たす音波情報伝送技術を開発することとした。

- ・ URLや簡単なテキスト情報を1～2秒で伝送できる

^{*1} 音響透かし：音声や音楽に人間の聴覚には知覚できないように情報を埋め込む技術。

- ・可聴帯域^{*2}の音波で情報を伝送し、一般的な市販のスピーカ、移動端末のマイクロフォンで情報伝送が可能である
- ・伝送信号を聴覚に不快にならないように音声・音楽に重畳する

本稿では、従来の音波情報伝送技術の問題を述べたうえで、前述の条件を満たす新しい音波情報伝送技術であるAcoustic OFDM (Orthogonal Frequency Division Multiplexing) とその伝送実験結果について解説する。また、Acoustic OFDMを用いたプロトタイプシステムの概要と今後の展望について述べる。

2. 音波情報伝送技術

主にデジタルコンテンツの著作権保護に使用される音響透かし技術として、エコー操作方式、スペクトル拡散方式および周波数パッチワーク方式などがあるが、これらの手法を用いて情報を空中伝搬させようとすると、以下のような問題がある。

①エコー操作方式

聴覚が、短いエコーを知覚できないという特性を利用し、エコーの遅延や振幅を変化させることによって伝送ビットを識別できるようにする方法[1]。空中伝搬では反射波やスピーカの減衰振動によってさまざまなエコーが生じるため適用できない。

②スペクトル拡散方式

聴覚心理モデル^{*3}[2]を用いて周波数マスキング^{*4}のしきい値を計算し、伝送信号にPN (Pseudo-random Noise) 系列^{*5}を乗算して全周波数帯に拡散した信号を周波数マスキングのしきい値以下になるようにして重畳する方法[3]。空中伝搬や環境雑音にも強い耐性があるが、PN系列の拡散率を高くしないと信号を抽出できないため、必然的に伝送速度が低速になる (10bit/s程度)。

③周波数パッチワーク方式

任意の2つの周波数帯を選び、一方のパワーを増加、他方は減少というように統計的な偏りをつくって伝送信号を重畳する方法[4]。伝送速度は40bit/s程度が可能であり、それ以上の伝送速度を求めると音質の劣化が著しくなる。

一方、超音波を用いて情報を伝送する場合は、聴覚への影響を考慮する必要はなく、利用できる周波数帯域も可聴帯域に比べて広いため伝送速度は桁違いに高速にできるが、前述したとおり特殊な再生・録音デバイスを必要とするため、導入コストや普及の観点から考えると現実的ではない。本稿では、伝送速度、聴覚への影響、導入コスト、普及の観点から実用的なレベルを満たせる新しい音波情報伝送技術Acoustic OFDMを提案する。

3. Acoustic OFDM

Acoustic OFDMは、次世代の移動通信にも有望とされているOFDM方式を応用している。OFDM方式は、複数の狭帯域信号を周波数多重して並列伝送を行うため周波数利用効率^{*6}に優れている。また、遅延波干渉の影響に対処しやすい点も反射波の影響が強い音波通信に有効である。

Acoustic OFDMの特徴は、OFDM変調信号を聴覚に不快にならないように変形して音声・音楽に重畳することにある。一般的にフラットなパワースペクトルの音はノイズのように聞こえ、聴覚には不快に感じることが多い。一方、特定の周波数にパワーが偏っている音は、ノイズよりもトーンに近い音になる。通常のOFDM変調信号は、すべてのサブキャリア^{*7}のパワーが一樣であるためノイズのような音に聞こえるが、各サブキャリアのパワーを重畳する音声・音楽のパワースペクトルに合わせることで、不快なノイズ音を音声・音楽に調和するトーンに変換できる。

Acoustic OFDMの基本となる変調方式を図1に示す。まず、オリジナル音源 (図1①) をフーリエ変換^{*8}し、周波数スペクトルを求める。その後、低域通過フィルタ (LPF: Low Pass Filter) で高域成分を除去し、低域音響信号 (図1②) を生成する。次に、高域の周波数のサブキャリアを伝送信号で変調したOFDM変調信号 (図1③) を生成する。このOFDM変調信号をオリジナル音源のスペクトル包絡に合わせてサブキャリアのパワーを調節し、高域音響信号 (図1④) を生成する。最後に、低域音響信号と高域音響信号を合成して合成音響信号 (図1⑤) を生成し、この信号をスピーカから再生する。

一般に人間の可聴帯域は20kHzまでといわれているが、市販のスピーカや移動端末のマイクロフォンが対応している周波数特性は20kHzよりも低いものがほとんどである。

*2 可聴帯域：正常な聴覚を持つ人間が聞き取ることができる音の周波数の範囲。通常、20Hz～20kHz。
 *3 聴覚心理モデル：耳の感度やマスキング効果をモデル化した人間の聴覚特性。
 *4 周波数マスキング：近くの周波数の音が妨害しあう効果で、大きな音の周波数の近傍周波数の小さな音が知覚できないこと。

*5 PN系列：擬似ランダム雑音 (Pseudo-random Noise) のビット列。擬似ランダム雑音は、あらかじめ決められたビット列を周期的に繰り返すため、自己同期のとりやすい系列。
 *6 周波数利用効率：単位時間、単位周波数帯域当りに送信できる情報ビット数。

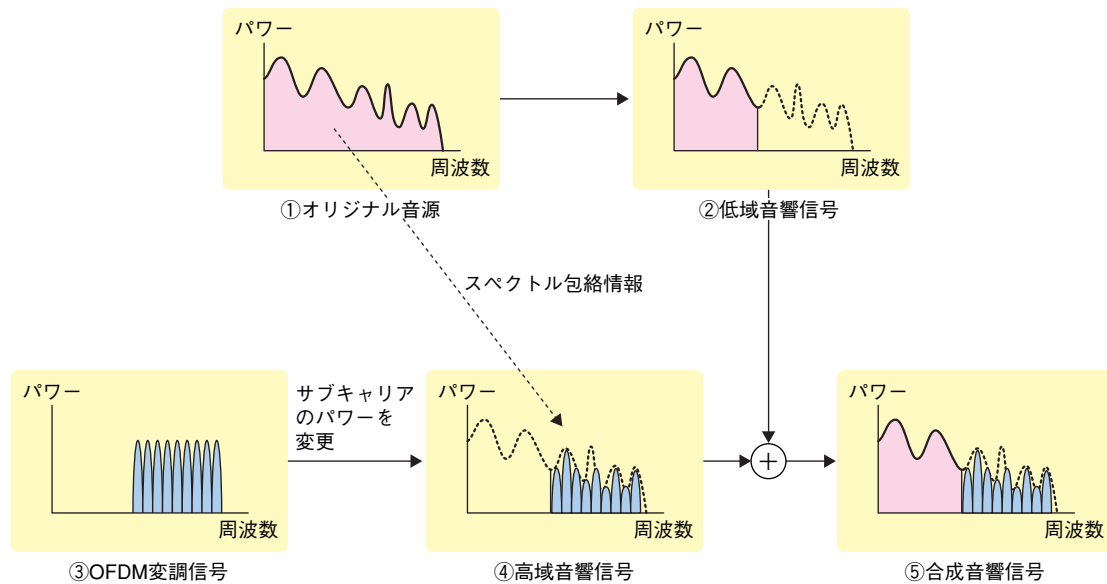


図1 Acoustic OFDM変調方式

特にマイクロフォンは10kHz付近までしか録音できないものが多い。また、オリジナル音源の低域信号に割り当てる周波数は音質劣化の観点から4kHz以上は必要であり、OFDM変調信号に割り当てられる周波数は多くても5～10kHz程度となる。しかし、5kHzの周波数帯域幅があれば、OFDMのガード時間^{*9}や誤り訂正符号などを考慮しても1kbit/s以上の情報伝送は可能であると考えられる。したがって、URLや簡単なテキスト情報であれば、1～2秒程度で伝送できる。

この基本伝送方式に、音波特有の課題を解決する手法や聴覚への影響を考慮した手法などが加わる。以下、それらについて詳しく説明する。

3.1 フレーム境界

OFDMの各サブキャリアの変調に2位相偏移変調(BPSK: Binary Phase Shift Keying)^{*10}や4位相偏移変調(QPSK: Quadrature Phase Shift Keying)^{*11}などの位相変調を用いると、OFDMフレーム間で位相が不連続になり、この部分は聴覚には耳障りな音に聞こえてしまう。そのため、OFDMフレーム間での位相不連続を軽減する措置が必要となる。通常のOFDMフレームは、データ信号部とその後方部分を前方にコピーしたガード時間で構成されるが、それに加えてフレーム境界に信号を滑らかにつなぐ区間を挿入

する。これにはフレームに対してオリジナルの高域信号に台形窓^{*12}をかけてオーバーラップさせるOverlap-Add方式(図2(a))が有効である。また、フレーム境界にトーン信号を挿入することで、不連続部分の耳障りな音を聞こえにくくする時間マスキング^{*13}方式(図2(b))も可能である。この場合、耳障りな音の代わりにトーン信号が聞こえるようになる。各フレーム境界のトーン信号の周波数を選択することでメロディを構成することもでき、このメロディを伝送信号が含まれていることを示す合図音として使うこともできる。

3.2 フレーム同期

受信側でOFDM復調を行うためには、OFDMフレームの境界を検出する必要がある。ガード時間とOFDM変調信号との相関で検出する方法もあるが、反射波などの遅延波があると精度が落ちるため、フレーム同期用の信号を追加する。このフレーム同期信号は、聴覚心理モデル[5]を利用して、音声・音楽の低域信号に周波数マスキングのしきい値以下になるようにして重畳する。Acoustic OFDMにおけるフレーム同期信号のスペクトルを図3に示す。まず、音声・音楽の低域信号の周波数マスキングのしきい値を計算する。次にPN系列で低域信号帯に拡散したフレーム同期信号を周波数マスキングのしきい値以下にレベルを調節し

*7 サブキャリア：複数の搬送波で情報ビットを並列伝送するマルチキャリア変調方式におけるそれぞれの搬送波。
 *8 フーリエ変換：信号の中に含まれる周波数成分とその割合を抽出する処理。
 *9 ガード時間：遅延波によるシンボル（*16参照）間干渉を避けるために、シンボル間に信号を挿入する一定時間分の区間。

*10 2位相偏移変調：デジタル変調方式の1つで、2つの位相にそれぞれ1つの値を割り当てることにより、同時に2値(1bit)の情報を送信可能。
 *11 4位相偏移変調：デジタル変調方式の1つで、4つの位相にそれぞれ1つの値を割り当てることにより、同時に2bitの情報を送信可能。
 *12 台形窓：ある時間ごとに区切られた信号が滑らかにつながるように信号区間の両端を減衰させる窓関数。

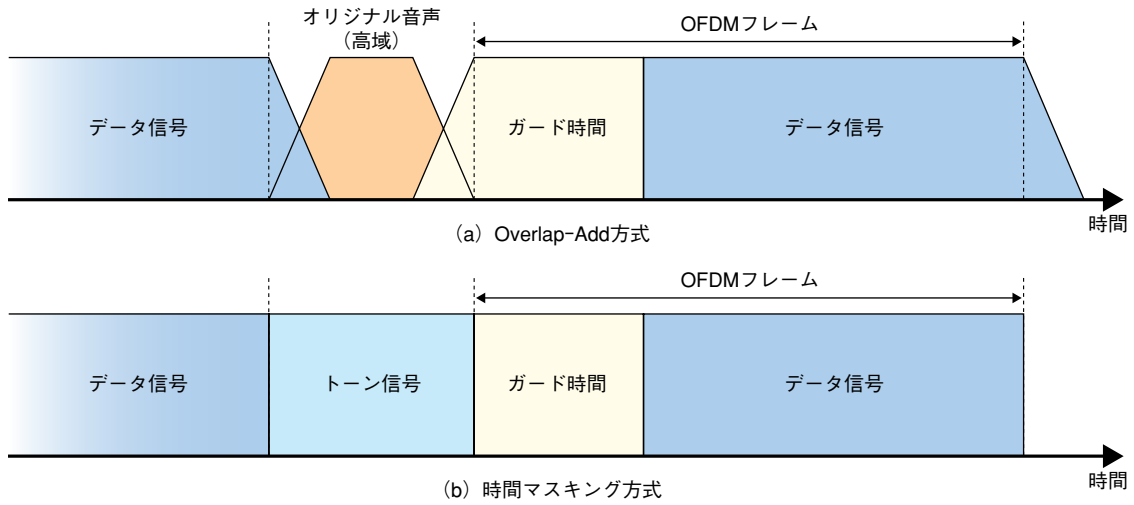


図2 Acoustic OFDM信号

て重畳する．これによりフレーム同期信号の音は聴覚には知覚できなくなる．受信側では，受信信号とPN系列で相関計算を行い，相関値の最も高くなるポイントをOFDMフレームの先頭とし，OFDM復調を行う．

3.3 ステレオ再生

一般的に音声や音楽を再生する機器には，ステレオ再生用に左右2つのスピーカが備わっている場合が多い．モノラル信号を再生する場合は，左右のスピーカから同じ信号を再生し，ステレオ信号を再生する場合は左用と右用の信号をそれぞれのスピーカから再生する．Acoustic OFDMの伝送信号を単純に2つのスピーカから再生した場合，マルチパス干渉により顕著な周波数選択性フェージング^{*14}が生じる．そのため，ステレオで再生されることを考慮し，左右それぞれのスピーカで再生される伝送信号をステレオ信号で生成しておく必要がある．一方，移動端末に搭載されているマイクロフォンはたいてい1つであり，録音信号はモノラル信号となる．したがって2スピーカ・1マイクロフォンでの送信ダイバーシチ方式が有効になる．

送信ダイバーシチには式(1)の生成行列Gで示される符号化率1の時空間ブロック符号化 (STBC: Space Time Block Code)^{*15} [6]を行う．

$$G = \begin{pmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{pmatrix} \quad (1)$$

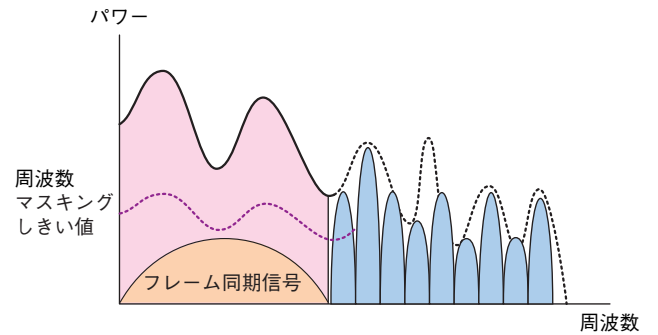


図3 Acoustic OFDM フレーム同期信号のスペクトル

ここで s^* は s の複素共役で， s_1, s_2 はそれぞれ時刻 $nT, (n+1)T$ (T はフレーム長) の送信シンボル^{*16}を表し，スピーカLからそれぞれ $s_1, -s_2^*$ を，スピーカRからそれぞれ s_2, s_1^* を送信する．受信側での信号検出は，時刻 $nT, (n+1)T$ における受信シンボル r_1, r_2 と，LRそれぞれのスピーカからの伝達関数 h_L, h_R を用いて，式(2)~(5)により検出できる．

(受信信号)

$$r_1 = h_L s_1 + h_R s_2 \quad (2)$$

$$r_2 = h_R s_1^* - h_L s_2^* \quad (3)$$

(送信ダイバーシチ復号)

$$\begin{aligned} \hat{s}_1 &= h_L^* r_1 + h_R^* r_2 \\ &= h_L^* (h_L s_1 + h_R s_2) + h_R^* (h_R s_1^* - h_L s_2^*) = (|h_L|^2 + |h_R|^2) s_1 \end{aligned} \quad (4)$$

*13 時間マスキング：大きな音の前後の音が聞こえにくくなる効果で，前方へは5～20ms程度，後方へは50～200ms程度の効果がある．

*14 周波数選択性フェージング：反射波などの遅延波による干渉によって，周波数軸上で一定周波数周期で信号電力が低下する現象．

*15 時空間ブロック符号化：送信ダイバーシチ技術に用いる符号化方式．時間と空間に相関を付加することによって空間多重された信号を分離できるという特徴がある．

*16 シンボル：本稿では，伝送するデータの最小単位で，例えばQPSKの場合，1シンボル当り2bitの情報を持つ．

$$\begin{aligned}\hat{s}_2 &= h_R^* r_1 - h_L r_2^* \\ &= h_R^* (h_L s_1 + h_R s_2) - h_L (h_R^* s_1 - h_L^* s_2) = (|h_L|^2 + |h_R|^2) s_2\end{aligned}\quad (5)$$

$\hat{s}$ は、送信ダイバーシチ復号で分離したそれぞれのシンボルであり、これにより s_1 , s_2 を抽出でき、送信ダイバーシチ効果を得ることができる。2つのスピーカから単純に同じ信号を送信した場合、干渉によって受信信号が抽出できなくなる地点が多数存在するが、送信ダイバーシチを用いる場合、どの地点においても受信信号電力を高めることができる。高周波数の音波は指向性が鋭いため、ステレオの送信ダイバーシチで再生することにより、伝送信号の届く範囲を広げることができる。

3.4 周波数オフセット・ドップラーシフト

音波通信においても無線通信と同様に送信機と受信機のクロック周波数のずれによって、周波数オフセット^{*17}が生じる。もともと音響装置の再生・録音機は通信用に設計されているわけではないので、このクロック周波数のずれは音響装置の方は大きく、機器間で5,000ppm (parts per million)^{*18}程度の周波数オフセットが生じている場合もある。また、可聴帯域の周波数でOFDM変復調する場合、サブキャリアごとにずれる周波数が顕著に異なる。例えば、送受信機間で5,000ppmのクロック周波数のずれがある場合、5kHzのサブキャリアは25Hzずれるが、10kHzのサブキャリアは50Hzずれる。したがって、無線通信で一般的に用いられているようなある一定の周波数の正弦波を乗算して周波数オフセットを補正する方法は適用できない。補正にはリサンプリングによるピッチ変換^{*19}が必要となる。また録音マイクログフォンの位置変動によるドップラーシフト^{*20}も無視できない。音速は約340m/sで、電波の速度の約百万分の一であり、わずかな位置変動でドップラーシフトが顕著に現れる。このドップラーシフトの補正と前述の周波数オフセットを同時に補正するリサンプリング方式を以下に説明する。

まず、送信側で基準となるドップラーシフト補正用のパイロット信号をOFDM変調信号の周波数帯よりも高い周波数に配置する。この周波数を f_i とし、これは受信側でも既知である。受信側では受信録音信号から高域通過フィルタ

(HPF: High Pass Filter) でドップラーシフト補正用のパイロット信号だけを抽出する。この受信パイロット信号を周波数 f_i でFM復調し周波数の時間変動を検出する。

移動端末がスピーカに対して速度 $v(t)$ で移動している場合、移動端末が検出するパイロット信号の周波数 f_o はドップラー効果によって式(6)になる。

$$f_o = \frac{V-v(t)}{V} f_i = f_i - \frac{v(t)}{V} f_i \text{ [Hz]} \quad (V \text{ は音速}) \quad (6)$$

この周波数 f_o で観測されたパイロット信号を周波数 f_i でFM復調すると、時刻 t における角周波数のずれ $z(t)$ が検出できる。

$$\text{(FM 復調)} \quad z(t) = -2\pi \frac{v(t)}{V} f_i \text{ [radian]} \quad (7)$$

この $z(t)$ に基づいて受信信号をリサンプリングしてピッチ変換することでドップラーシフトを補正できる。 $z(t)$ とサンプリング周波数 f_s から、式(8)でリサンプリングするサンプリング点を計算できる。

$$\text{(サンプリング点のずれ)} = \frac{z(t)f_s}{2\pi f_i} \text{ [sample]}^{*21} \quad (8)$$

このサンプリング点のずれに基づいてリサンプリングすることにより、ドップラーシフトおよび前述の周波数オフセットも同時に補正できる。周波数オフセットはFM復調した際の $z(t)$ の直流成分として現れてくる。

4. 音波情報伝送実験

Acoustic OFDMの基本伝送方式である5~10kHzの周波数を用いた音波OFDMの伝送実験の結果を述べる。実験に使用したスピーカとマイクログフォンの基本仕様を表1に、伝送パラメータを表2に示す。このパラメータにおいての周波数利用効率はQPSKで1.25bit/s/Hzになり、BPSKでは0.63bit/s/Hzになる。これに誤り訂正符号の符号化率をかけたものが実際に伝送できる情報量となる。

4.1 伝搬距離

伝搬距離とBER (Bit Error Rate) の関係を図4に示す。1~4mの伝搬距離においてQPSKおよびBPSKで各サブキャ

*17 周波数オフセット：本稿では、送信機と受信機の発振器のクロック周波数のずれによって生じる搬送波周波数のずれ。

*18 ppm：100万分のいくらかの割合を示す単位。主に濃度を表すために用いられるが、搬送波周波数のずれの割合を示す場合にも用いられる。

*19 ピッチ変換：音響信号の再生スピードを変えることで音程（周波数）を変化させること。

*20 ドップラーシフト：ドップラー効果によって生じる搬送波周波数のずれ。

*21 sample：デジタル化の際に、サンプリングした各時間のデータ。本稿ではフレーム長などをsample数で表記している。

表1 実験に用いたスピーカ・マイクロフォンの仕様

	スピーカ	マイクロフォン
周波数特性	20～20,000Hz	50～16,000Hz
口径	6.5cm	3.6cm

表2 伝送パラメータ

サンプリング周波数	44.1kHz
量子化ビット数	16bit
サブキャリア変調方式	BPSK/QPSK
サブキャリア間隔	21.5Hz
サブキャリア数	234 + 14 (周波数パイロット)
ガード時間	11.6ms
再生音圧	約 70dB SPL

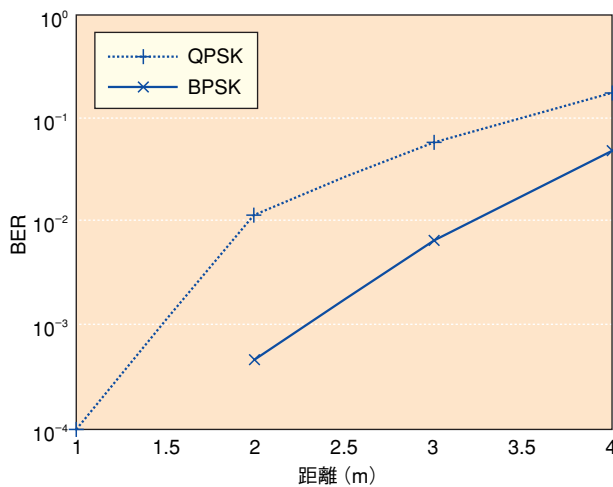


図4 伝搬距離とBER

リアを変調したときのBERを測定した。誤り訂正符号で5%程度のBERまでは訂正できるとすると、QPSKでは3m、BPSKでは4m程度の伝搬が可能である。

4.2 指向角度

音波の指向性は、周波数が高くなるにつれ、またスピーカの口径が大きくなるにつれて鋭くなる。Acoustic OFDMでは高域の周波数で信号を送送するため、伝送信号の指向性が鋭くなるという特徴がある。音波の到来方向の角度とBERの関係を図5に示す。伝搬距離2mにおける0～60°の向きでのBERを測定した。実験に用いたスピーカの口径は6.5cmである。これも5%程度のBERまで訂正できるとすると、QPSKで20°、BPSKで50°までが伝搬可能である。

4.3 周波数応答

各サブキャリアにおける振幅特性および位相特性を図6に示す。振幅特性については、スピーカとマイクロフォンの振幅特性および周波数選択性フェージングが影響し、±5dB程度の差がみられる。位相特性に関しては、周波数ごとに位相の偏位が異なっているが、反射波などの遅延波による影響が主原因であるため、位相特性はほぼ線形（群遅延が一定）になっている。そのため、ある周波数間隔でパイロット信号を入れておくことで位相の偏位は補正できる。

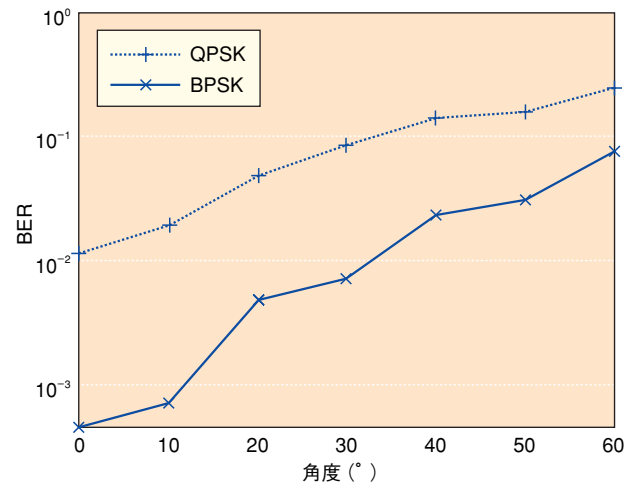
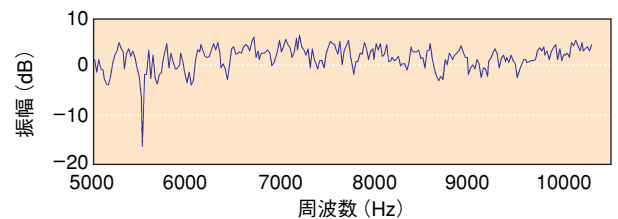
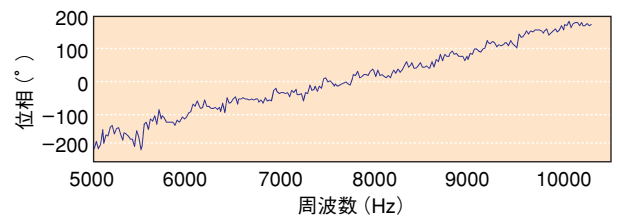


図5 到来方向の角度とBER



(a) 振幅特性



(b) 位相特性

図6 振幅特性と位相特性

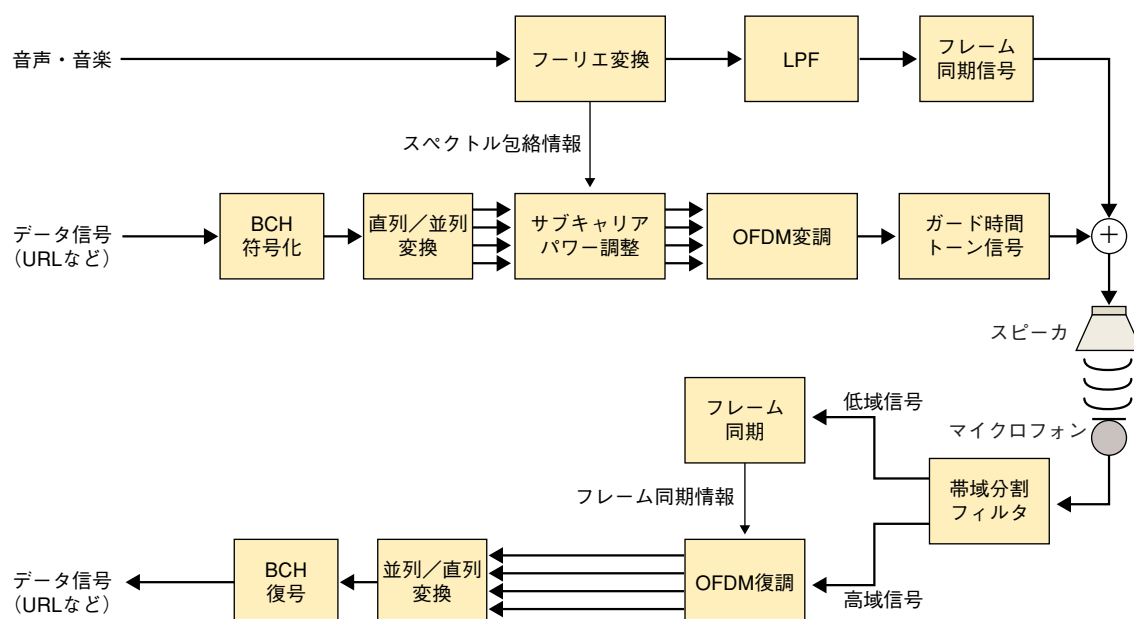


図7 システム構成

5. プロトタイプシステムの概要

Acoustic OFDMを用いたプロトタイプシステムを実装した。システムの構成を図7に示す。このシステムでは、URLなどの簡単なテキスト情報を音声・音楽に埋め込みながら繰り返しカルーセル送信^{*22}する。データ送信中は3.1節で記述した時間マスキング方式におけるトーン信号をOFDMフレーム境界に挿入し、各トーン信号の周波数を選択することで伝送情報が含まれていることを示す合図音を構成する。ユーザはこの合図音が聞こえている任意の区間を移動端末で約1.5秒間録音することでデータを抽出できる。

送信側では音声・音楽の信号とそれに重畳して送信するURLなどのデータ信号を入力し、データ信号をBCH（Bose Chaudhuri Hocquenghem）符号で符号化する。これにより、5%程度のビットエラーであれば受信側でエラー訂正して元の信号を抽出できる。この符号化したビット列を直列から並列に変換する。音声・音楽の信号は、まずフーリエ変換によって周波数スペクトルを計算し、LPFによって高域信号をカットする。ここで計算した周波数スペクトルに基づいてOFDMのサブキャリアのパワーを調節し、このサブキャリアで並列変換したデータ信号をOFDM変調し、ガード時間とトーン信号を挿入する。音声・音楽の低域信号に

はOFDMフレームの周期で高い自己相関値を持つPN系列によるフレーム同期信号を付加し、OFDM変調信号と合成してスピーカから出力する。出力された音波をマイクロフォンで録音し、その信号を帯域分割フィルタで低域信号と高域信号に分離する。低域信号とPN系列の相関をとり、高い相関値を示したところをOFDMフレームの先頭と識別し、フレーム同期をとる。これによりフレーム単位で高域信号からガード時間・トーン信号を除去し、OFDM復調する。復調された並列信号を直列に変換しBCH符号で復号してデータを抽出する。受信側ではどのフレームから受信を始めてもある一定数のフレームを受信すればデータを復号できる。BCH符号は巡回符号のため、符号語のビット列が

表3 システムパラメータ

サンプリング周波数	44.1kHz
OFDM フレーム長	2,032 sample
データ長	1,024 sample
ガード時間	600 sample
トーン信号区間	408 sample
帯域フィルタ遮断周波数	5,512.5 Hz
サブキャリア数	33+4 (周波数パイロット)
PN系列長	127
チップレート	2,756.25 Hz

* 22 カルーセル送信：同じデータを繰り返し送信すること。あるデータを送信し終わったら、また最初からデータを送信する。

巡回していても誤り訂正の復号が可能であり、符号語の巡回ビット数からデータ信号の先頭を検出できる。

プロトタイプシステムのパラメータを表3に示す。このパラメータによるシステムの受信アプリケーションを移動端末に実装し、テレビCMの音声・音楽などにデータ信号を重ねて試験したところ、スピーカから1m程度の範囲であれば90%以上の認識率で72byte程度のテキスト情報を取得できることを確認した。

6. あとがき

本稿では、伝送速度、聴覚への影響および普及の観点から実用的と考える音波情報伝送技術Acoustic OFDMを提案し、音波でのOFDM伝送実験結果およびプロトタイプシステムの概要について述べた。今後、さらなる伝送性能の向上、計算量の削減を目指す。

文 献

- [1] D. Gruhl, A. Lu and W. Bender: "Echo Hiding," Information Hiding 1996, pp. 295-315.
- [2] "Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s," ISO/IEC 11172, 1993.
- [3] L. Boney, A.H. Tewfik and K.N. Hamdy: "Digital watermarks for audio signals," IEEE Intl. Conf. on Multimedia Computing and Systems, pp. 473-480, Mar. 1996.
- [4] 中山 彰, 岩城 敏: "HyperAudio: 音を媒介にしたマンマシンイ

ンタフェースの一手法," ヒューマンインタフェース No. 088, 2000.

- [5] J. Johnston: "Transform Coding of Audio Signals Using Perceptual Noise Criteria," IEEE Journal on Selected Areas in Communications, Vol. 6, pp. 314-323, Feb. 1988.
- [6] S.M. Alamouti: "Simple Transmit Diversity Technique for Wireless Communications," IEEE J.Sel.Areas Commun., Vol. 16, pp. 1451-1458, Oct. 1998.