

## 音楽電子透かし技術を用いたモバイル端末位置情報の検出手法

茂出木 敏雄<sup>†</sup>

<sup>†</sup> 大日本印刷株式会社 情報コミュニケーション研究開発センター  
〒141-8001 東京都品川区西五反田 3-5-20

あらまし 筆者らは、音源に携帯電話を向けるだけで情報が抽出できる音楽電子透かし技術の開発を進めてきた。音脈分凝という聴覚心理特性を活用し、ヒトの聴感特性が鋭敏な電話回線全帯域に対してデータを埋め込んでも、埋め込みノイズを抑圧し、3G携帯電話で安定した抽出を実現させることができた。これを応用して、空間上に異なる符号が埋め込まれた多数の音源を配置させ、抽出される符号により端末位置を検出する手法を開発している。本稿では、2チャンネル・ステレオ音楽の左右チャンネルに異なる符号を埋め込み、端末位置が左右のスピーカ近辺では各々埋め込まれた符号が抽出されるとともに、左右が混合された中央近辺では第3の符号の抽出を可能にする改良手法を提案する。

## Detection Method of Mobile Terminal Using Audio Watermark Technique

Toshio MODEGI<sup>†</sup>

<sup>†</sup> Media Technology Research Center, Dai Nippon Printing Co., Ltd.  
3-5-20, Nishigotanda, Shinagawa-ku, Tokyo 141-8001, Japan  
E-mail: <sup>†</sup> Modegi-T@mail.dnp.co.jp

**Abstract** We have been developing audio watermarking techniques which enable extraction of embedded data by cell phones. Applying an auditory stream segregation phenomenon, we could embed data onto full phone-line frequency ranges, where our auditory response is prominent, with suppressing play-back audible noises in order to make possible to extract data with a 3-G mobile phone. Applying this technique, we are developing a mobile terminal location detection method against multiple sound sources with unique codes embedded. In this paper, we propose an improved method, which enables to extract another third code at the center location between left and right loudspeakers by embedding two different codes to each channel of stereo signals.

### 1. まえがき

筆者らは空間に分布する音源の情報を携帯端末等で非接触に抽出し、音源に関連するサイバー空間と連動できるユビキタス音響空間を提案している[4]。それを実現する方法として、音楽電子透かし手法に注目し、携帯電話を音源に近づけるだけで、音響信号に埋め込まれた情報を抽出できる新規な音楽電子透かし手法の開発を進めてきた[3]-[6]。既提案の音楽著作権保護を目的とした電子透かし手法[1]では、埋め込みノイズが聞こえないようにヒトの聴感特性が鈍感な音域を改変し、音質を余り劣化させない程度の各種圧縮・変調処理などに対する耐性を備えていれば良かった。

これに対して、現行の3G携帯電話では電話回線帯域と同様に4kHzを超える音は収録できず、更に録音された音声データは3GPP準拠圧縮によ

り振幅側も顕著な歪みが発生する。一方、電子透かしが埋め込まれたソース音源は、アナログ/デジタル放送やストリーミングで配信される場合もあり、同様に4kHzを超える周波数成分は変調や圧縮による劣化が加わることも考慮する必要がある。従って、開発すべき電子透かし技術において、4kHz未満の聴感特性が最も鋭敏な音域に、振幅歪みがあっても抽出可能なように顕著な改変を加え、かつ再生時にノイズが聴取されないようにするという矛盾する課題を解決する必要がある。

そこで、筆者らは2チャンネル・ステレオ再生を応用して、データを埋め込んだL側スピーカで顕著に発生するノイズをR側スピーカで相殺する手法を提案した[3]。特に時分割法を用いれば、携帯電話による非接触抽出が可能であることを確認

した[3]。更に、左右双方のスピーカからも抽出可能な抽出自由度の高い埋め込み方式も考案し、埋め込み周波数帯を電話回線帯域の高域側(1.7kHz~3.4kHz)に変更すれば、最大 80 [bps]のレートで埋め込めることも確認している[4][5]。

しかし、これらの既提案手法はモノラル再生やヘッドホン再生環境ではノイズが目立ち運用が難しいこと、受信端末を再生音源のL側スピーカから遠ざけたり、室内残響が豊かであるとR側信号の影響を受けやすく、抽出精度が低下するという問題を抱えていた。そこで、ヒトの聴覚心理現象である音脈分凝[2]に着目し、データ埋め込みにより改変された信号パターンが、視聴者に対して音脈分凝[2]を引き起こすように工夫し、モノラル再生により生成されるノイズがヒトに対してのみ特異的に抑圧されるようにすることができた。更に、この音脈分凝が二重に引き起こされるように拡張することにより、再生ノイズを抑圧しながら、埋め込み周波数領域を電話回線帯域(300Hz~3.4kHz)の全体まで拡大させ、抽出感度を向上させることもできた[6]。

この音脈分凝を活用した方式をステレオ音源に適用すると、左右チャンネル別に異なる符号を埋め込み、左右のスピーカに端末を近づければ異なる符号を抽出させることができる。しかし、左右チャンネルの信号が混合する中央部では、データの抽出が不能になるという問題があった。

本稿では、この音脈分凝を活用したステレオ埋め込み方式において、左右に埋め込む符号の埋め込みレベルを各々ビット単位に2段階に変位させる拡張手法を提案する。これにより、左右スピーカ近辺では各々のチャンネルに埋め込まれた符号が抽出できるとともに、左右チャンネルの信号が混合する中央部では、左右に埋め込まれた符号と異なる第3の符号が安定して抽出されることが可能になる。即ち、ステレオ再生空間において、左右、中央3箇所端末位置識別が可能になる。

## 2. 先提案の音脈分凝を活用したモノラル音源に対する電子透かし埋め込み方式

図1は筆者らが提案した音脈分凝を活用したモノラル信号への電子透かし埋め込みの概念を示している。横軸が時間軸で後述する窓関数を用いてフレーム分割し、更に各フレームを2分割して埋め込み領域を形成し、各フレームに1ビットの情報を埋め込む。フレーム長は50 [msec]に設定す

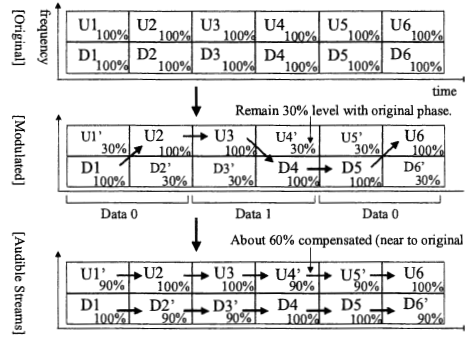


図1 音脈分凝を活用したモノラル信号への電子透かし埋め込みの概念

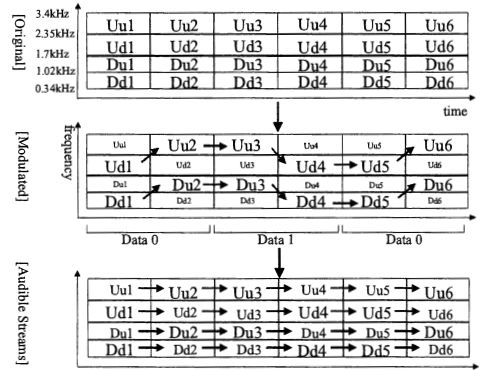


図2 二重音脈分凝を活用した広帯域電子透かし埋め込みの概念

ると、20 [bps]の埋め込みレートになり、図1の例では3フレームに3ビットの情報が埋め込まれる。縦軸は周波数で埋め込み周波数帯として、電話回線の高域部 1.7kHz~3.4kHzを設定し、この帯域を2分割して埋め込み領域を形成し、時間軸方向と合わせて、1フレームあたり4領域で形成される。

図1上段は原音信号で、3フレームの U1~U6 および D1~D6 の信号成分は全て均等 100%と仮定する。図1中段は3ビットの埋め込み処理が行われた結果を示し、ビット0を埋め込む場合は、各フレーム4領域の左上と右下成分のレベルを30%に落とし、ビット1を埋め込む場合は、左下と右上成分のレベルを30%に落としたものである。これにより、図中の矢印で示しているように、各時間領域において上下いずれかの音成分しか存在

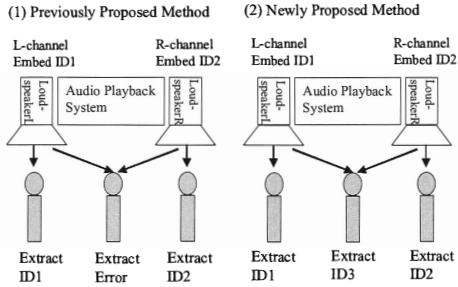


図3 位置検出を可能にするステレオ信号への電子透かし埋め込みの概念

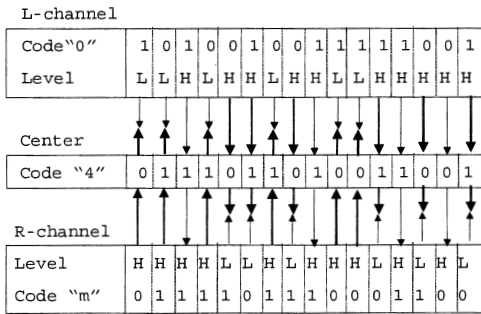


図4 左右チャンネル混合中央位置における第3の符号抽出の原理

しない、単旋律の音脈になる。

図1中段の信号が再生されている音をマイクロフォンでサンプリングし、各フレームの2対の対角成分合算値の比率でいずれのビットが埋め込まれているかを検出できる。一方、図1下段は同再生音をヒトが聴取した場合のパターンを示しており、音脈分凝[2]の効果が働き、上下2つの複旋律の音脈に聞こえてしまい、失われた U1, D2, D3, U4, U5, D6 の信号レベルは90%程度まで補間されてしまい原音との差異が殆どなくなる。

図1中段では、D1, U2, U3, D4, D5, U6 の信号レベルには改変を加えていないが、その後の研究により、逆に170%に増加させてもヒトの聞こえ方には変化が無いことが判明した。そのため、抽出感度向上のため、全信号レベルを増減させるようにしている。

図2は埋め込み周波数帯を電話回線帯域300Hz

～3.4kHzの全域に拡大する方法を示している。1.7kHz以下の低域側0.34kHz～1.7kHzにもう1セットのデータ埋め込み領域を設定し、上下で同一ビットパターンを重複して埋め込むようにすると、二重に音脈分凝が働くようになり、ノイズの増大を伴わずに抽出感度を向上させることができた。

### 3. 本提案のステレオ音源に対する位置検出可能な電子透かし埋め込み方式

図3-(2)は本稿で提案するステレオ信号への電子透かし埋め込みの概念を示している。L側チャンネルにはID1を埋め込み、R側チャンネルにはID2を埋め込むと、L側スピーカ近辺からはID1が抽出され、R側スピーカ近辺からはID2が抽出されるとともに、双方のスピーカ音が混合した中央部ではID3が抽出されるようにする。

図4は、これを実現する原理を示す。後述する表1の16ビット非組織的のハミング符号で表現されたASCIIコード"0"と"m"を各々左右チャンネルに埋め込み、ビットごとに埋め込みレベルをLとHの2段階に変位させる。中央部では左右いずれかの埋め込みレベルが強い方のチャンネルのビットを選択して第3のコード"4"を生成する。この時、第3のコード"4"も16ビット非組織的のハミング符号を満足すれば、中央部でも安定して抽出することができる。

左右チャンネルが同一ビットの場合は、左右チャンネルとも同一の埋め込みレベルHで埋め込めば良いが、左右チャンネルが異なるビットの場合は、採用するビットのチャンネル側の埋め込みレベルをHに反対側をLに設定する。埋め込みレベ

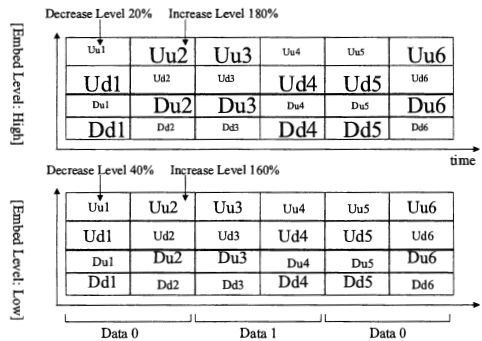


図5 二重音脈分凝手法において2種類の埋め込みレベルを設定する具体例

ルが強い音源は弱い音源に比べ、抽出可能な距離が長いため、双方の音源までの距離が等位置にある中央部では埋め込みレベルが強い音源のビットが抽出されやすい。

埋め込みレベルを変化させる方法として、図5に示し、これは図2に示した埋め込み手法に対して、この2種類の埋め込みレベルを実現する一例である。

#### 4. 改良型電子透かし埋め込み方式の詳細

本稿で提案する手法は文献[6]に記載の電子透かし埋め込みアルゴリズムにおいて、埋め込みレベルをビットごとに2段階に変位させる機能を追加するだけで実現できる。抽出側のアルゴリズムは変更不要であるが、本稿では繰り返し埋め込まれた符号を迅速に抽出できる巡回ビットパターン照合機能について併せて言及する。

##### 4.1. 電子透かし埋め込みアルゴリズム

###### 4.1.1. 音響信号の周波数次元変換

本節では以下、ソースがステレオ信号の左右いずれかのチャンネルのモノラル信号であることを前提に説明する。サンプリング周波数  $f_s$  のソース音響信号から窓幅  $N$  のフレーム  $X(i)$  ( $i=0, \dots, N-1$ ) を  $N/2$  間隔で抽出し、奇数フレームと偶数フレームをセットにして1ビットを埋め込むものとする。 $f_s=44.1$  [kHz] のとき、 $N$  は4096 (10bps), 2048 (20bps), 1024 (40bps), 512 (80bps) のいずれかを与える。図6に示される2種類の窓関数  $W_k(i)$  ( $k=1, 2$ ) を用意し、奇数フレームに対して、第1番目の窓関数  $W_1(i)$  を適用し、偶数フレームに対しては第2番目の窓関数  $W_2(i)$  を用いて離散フーリエ変換を施し、変換後データの実部を  $A_k(j)$  ( $j=0, \dots, N-1$ ,  $k=1, 2$ ), 虚部を  $B_k(j)$  ( $j=0, \dots, N-1$ ,  $k=1, 2$ ) として、以下のように与えられる。

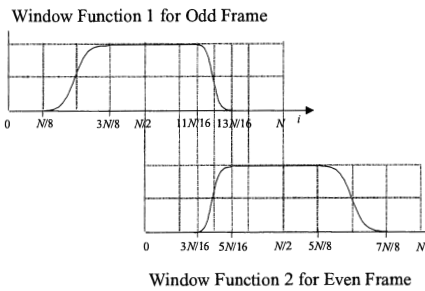


図6 提案する2つの窓関数

$$\begin{aligned} A_k(j) &= \sum_{i=0, N-1} W_k(i) \{X(i)+R(i)\} \cos(2\pi ij/M). \\ B_k(j) &= \sum_{i=0, N-1} W_k(i) \{X(i)+R(i)\} \sin(2\pi ij/M). \\ &(j=0, \dots, N-1, k=1, 2) \end{aligned} \quad (1)$$

ここで、 $R(i)$  はソース音響信号に付加する白色雑音成分で、ソース音響信号が音声信号の場合など無音区間が多く存在する場合、データの埋め込みが適切に行えないため、必要に応じて付加する。ソース音響信号が16ビットの場合、 $R(i)$  として  $\pm 32$  以内の範囲で正負に変化する一様乱数を与えれば、音質にあまり影響を与えず、無音部にも適切にデータを埋め込むことができる。

2種の窓関数  $W_k(i)$  ( $k=1, 2$ ) は、以下の通り、互いに  $N/2$  ずらして重ねると相補的な関係になるように、ハニング窓関数  $W(i)=0.5-0.5\cos(2\pi i/M)$  を時間軸で非対称に変形させたものである。

$$\begin{aligned} W_1(i) &= 0.0 \quad (i \leq N/8). \\ W_1(i) &= 0.5 - 0.5\cos(4\pi(i-N/8)/N) \quad (N/8 < i \leq 3N/8). \\ W_1(i) &= 1.0 \quad (3N/8 < i \leq 11N/16). \\ W_1(i) &= 0.5 + 0.5\cos(8\pi(i-11N/16)/N) \\ &\quad (11N/16 < i \leq 13N/16). \\ W_1(i) &= 0.0 \quad (13N/16 < i). \\ W_2(i) &= 0.0 \quad (i \leq 3N/16). \\ W_2(i) &= 0.5 - 0.5\cos(8\pi(i-3N/16)/N) \\ &\quad (3N/16 < i \leq 5N/16). \\ W_2(i) &= 1.0 \quad (5N/16 < i \leq 5N/8). \\ W_2(i) &= 0.5 + 0.5\cos(4\pi(i-5N/8)/N) \\ &\quad (5N/8 < i \leq 7N/8). \\ W_2(i) &= 0.0 \quad (7N/8 < i). \end{aligned} \quad (2)$$

###### 4.1.2. 最適埋め込み割合の決定

奇数フレーム ( $k=1$ ) および偶数フレーム ( $k=2$ ) における周波数4分割バンドのエネルギー総和値  $E_{k1}$ ,  $E_{k2}$ ,  $E_{k3}$ ,  $E_{k4}$  を計算する。 $mI$ ,  $MI$  と  $m\hat{h}$ ,  $M\hat{h}$  を下位と上位の埋め込み周波数領域の各々下限値と上限値とし、 $pI$ ,  $p\hat{h}$  を各々の領域の2分割バンドとする。例えば、 $N=4096$  の場合、 $mI=32$  (0.34kHz),  $MI=m\hat{h}=160$  (1.7kHz),  $M\hat{h}=320$  (3.4kHz) に設定し、 $pI=(MI-mI)/2=64$  (0.68kHz),  $p\hat{h}=(M\hat{h}-m\hat{h})/2=160$  (1.7kHz) に設定する。

$$\begin{aligned} E_{k1} &= \sum_{j=mI, mI+pI-1} \{A_k(j)^2 + B_k(j)^2\}. \\ E_{k2} &= \sum_{j=mI+pI, mI+2pI-1} \{A_k(j)^2 + B_k(j)^2\}. \\ E_{k3} &= \sum_{j=m\hat{h}, m\hat{h}+p\hat{h}-1} \{A_k(j)^2 + B_k(j)^2\}. \\ E_{k4} &= \sum_{j=m\hat{h}+p\hat{h}, m\hat{h}+2p\hat{h}-1} \{A_k(j)^2 + B_k(j)^2\}. \end{aligned} \quad (k=1, 2) \quad (3)$$

続いて  $\gamma = (E_{11}E_{22}E_{13}E_{24}) / (E_{12}E_{21}E_{14}E_{23})$  を算出し、埋め込みビットが 1 の場合は、 $\gamma = 1 / \gamma$  に変換する。ここで前節で述べた埋め込みレベルが H の場合、即ち埋め込むビットと中央位置で抽出されるビットが同一の場合は、埋め込み割合  $\alpha$ 、 $\beta$  の初期値を 0.72、0.28 とし、 $\gamma$  が 1.0 未満の場合、 $\alpha = 0.72\gamma^{-1/4}$ 、 $\beta = 0.28\gamma^{1/4}$  (4-1) に設定する。ただし、パルス性ノイズの発生を抑えるため、 $\alpha \leq 7.2$ 、 $\beta \geq 0.028$  に制限する。

一方、埋め込みレベルが L の場合、即ち埋め込むビットと中央位置で抽出されるビットが異なる場合は、 $\alpha$ 、 $\beta$  の初期値を 0.66、0.33 とし、 $\gamma$  が 1.0 未満の場合、

$$\alpha = 0.66\gamma^{-1/4}、\beta = 0.33\gamma^{1/4} \quad (4-2)$$

に設定する。ただし、パルス性ノイズの発生を抑えるため、 $\alpha \leq 6.6$ 、 $\beta \geq 0.033$  に制限する。

#### 4.1.3. 周波数成分の変換

上記算出した埋め込み割合  $\alpha$ 、 $\beta$  を用いて、 $A_1(j)$ 、 $A_2(j)$ 、 $B_1(j)$ 、 $B_2(j)$  の各周波数成分に対して、以下 (5) 式に基づき改変を行なう。

$$\begin{aligned} A_k'(j) &= A_k(j) \delta / |A_k(j)^2 + B_k(j)^2|^{1/2} \\ B_k'(j) &= B_k(j) \delta / |A_k(j)^2 + B_k(j)^2|^{1/2} \end{aligned} \quad (5)$$

1) ビット 0 を埋め込む場合：

$j = m/1, \dots, m/1 + P/1 - 1$  および  $j = mh, \dots, mh + Ph - 1$  の範囲で、 $k=1$ 、 $\delta = \alpha$  とし (5) 式を実行し、 $k=2$ 、 $\delta = \beta$  とし (5) 式を実行し、

$j = m/1 + P/1, \dots, m/1 + 2P/1 - 1$  および  $j = mh + Ph, \dots, mh + 2Ph - 1$  の範囲で、 $k=1$ 、 $\delta = \beta$  とし (5) 式を実行し、 $k=2$ 、 $\delta = \alpha$  とし (5) 式を実行する。

2) ビット 1 を埋め込む場合：

$j = m/1, \dots, m/1 + P/1 - 1$  および  $j = mh, \dots, mh + Ph - 1$  の範囲で、 $k=1$ 、 $\delta = \beta$  とし (5) 式を実行し、 $k=2$ 、 $\delta = \alpha$  とし (5) 式を実行し、

$j = m/1 + P/1, \dots, m/1 + 2P/1 - 1$  および  $j = mh + Ph, \dots, mh + 2Ph - 1$  の範囲で、 $k=1$ 、 $\delta = \alpha$  とし (5) 式を実行し、 $k=2$ 、 $\delta = \beta$  とし (5) 式を実行する。

#### 4.1.4. 音響信号の時間次元逆変換

更新されるフレームのデータ  $X'(i)$  は、直前の処理で逆変換されたフレームのデータ  $X(i)$  に対して  $N/2$  だけオフセットして加算される。奇数フレームの場合  $k=1$ 、偶数フレームの場合  $k=2$  とし以下式を実行する。

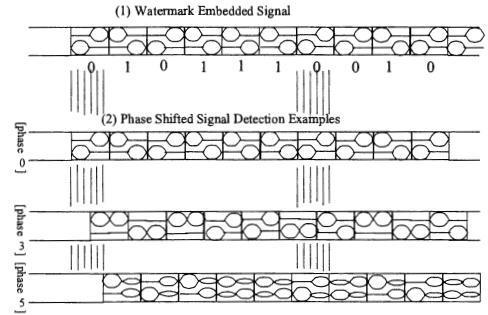


図 7 電子透かし抽出時におけるフレーム位相補正処理

$$\begin{aligned} X'(i) &= 1/N \{ \sum_{j=0, N-1} A_k'(j) \cos(2\pi ij/M) \\ &\quad - \sum_{j=0, N-1} B_k'(j) \sin(2\pi ij/M) \\ &\quad + [1 - W_k(i)] X(i + N/2) \} \end{aligned} \quad (6)$$

## 4.2. 電子透かし抽出アルゴリズム

### 4.2.1. 音響信号の周波数次元変換

電子透かしが埋め込まれた音響信号に対して、音源がステレオの場合はそれらを合成した信号に対して、モノラル信号として周波数  $f_s$  でサンプリングを行ない、4.1.1 節と同様に、窓幅  $N$  のフレーム  $X(i)$  ( $i=0, \dots, N-1$ ) を  $N/2$  間隔で抽出する。この時、埋め込み時の奇数フレームと偶数フレームに対応する区間が未知であるため、図 7 に示されるように位相  $h$  を  $H$  通り (図 7 例では  $H=6$ ) にずらしながら、 $H$  通りの奇数フレーム ( $k=1$ ) および偶数フレーム ( $k=2$ ) のセットを用意する。そして (1) 式と同様に 2 つの窓関数  $W_k(i)$  ( $k=1, 2$ ) を用いて、各位相  $h$  別に離散フーリエ変換を施し、変換後データの実部を  $A_k(j, h)$ 、虚部を  $B_k(j, h)$  ( $j=0, \dots, N-1$ 、 $k=1, 2$ 、 $h=0, \dots, H-1$ ) とする。

$$\begin{aligned} A_k(j, h) &= \sum_{i=0, N-1} W_k(i) X(i + hN/H) \cos(2\pi ij/N) \\ B_k(j, h) &= \sum_{i=0, N-1} W_k(i) X(i + hN/H) \sin(2\pi ij/N) \end{aligned} \quad (7)$$

### 4.2.2. 音響信号の仮想振幅変換

後述するビットデータの判定において隣接フレーム間での振幅大小比較を行なうにあたり、原信号の振幅変動成分を除外するため平坦化する前処理を行う。 $H$  通りの奇数フレームおよび偶数フレームにおけるエネルギー総和値を計算し、各々が既定値  $S_0$  になるようにスケール変換倍率  $S_1(h)$  と

$S_2(h)$ を決定する。

$$S_k(h) = S_0 / [\sum_{j=m, M-1} \{A_k(j, h)^2 + B_k(j, h)^2\}]^{1/2}. \quad (8)$$

( $k=1, 2, h=0, \dots, H-1$ )

$j=m, \dots, M-1$  の範囲で、 $k=1, 2$  について  $A_k(j, h)$ ,  $B_k(j, h)$  の各要素に  $S_k(h)$  を乗じるスケール変換を施す。以下 (9) 式を用いて、各分割バンドごとに、スケール変換後の周波数成分  $A_k(j)$ ,  $B_k(j)$  に対するエネルギー総和値  $E_{k1}(h)$ ,  $E_{k2}(h)$ ,  $E_{k3}(h)$ ,  $E_{k4}(h)$  を計算する。

$$\begin{aligned} E_{k1}(h) &= \sum_{j=m, m+2p-1} \{A_k(j, h)^2 + B_k(j, h)^2\}. \\ E_{k2}(h) &= \sum_{j=m+2p, m+4p-1} \{A_k(j, h)^2 + B_k(j, h)^2\}. \\ E_{k3}(h) &= \sum_{j=mh, mh+ph-1} \{A_k(j, h)^2 + B_k(j, h)^2\}. \\ E_{k4}(h) &= \sum_{j=mh+ph, mh+2ph-1} \{A_k(j, h)^2 + B_k(j, h)^2\}. \end{aligned} \quad (k=1, 2) \quad (9)$$

#### 4.2.3. 残響補正とビットデータの判定

前節で算出したエネルギー総和値には残響成分として1フレーム前の各エネルギー総和値  $E_{k1}^p(h)$ ,  $E_{k2}^p(h)$ ,  $E_{k3}^p(h)$ ,  $E_{k4}^p(h)$  が所定割合  $q$  ( $0 \leq q \leq 1$ ) だけ含まれていると仮定し、以下のように補正を加えながらエネルギー成分を算出する。尚、 $q$  の目安としては一般的な室内残響で運用する場合を  $q=0.06 \cdot 4096/N$  とし、窓幅に反比例させた値を与える。ただし、 $E_{k1}'(h) \geq 0$ ,  $E_{k2}'(h) \geq 0$ ,  $E_{k3}'(h) \geq 0$ ,  $E_{k4}'(h) \geq 0$  とする。

$$\begin{aligned} E_{11}'(h) &= E_{11}(h) - qE_{11}^p(h). & E_{21}'(h) &= E_{21}(h) - qE_{21}^p(h). \\ E_{12}'(h) &= E_{12}(h) - qE_{22}^p(h). & E_{22}'(h) &= E_{22}(h) - qE_{12}^p(h). \\ E_{13}'(h) &= E_{13}(h) - qE_{23}^p(h). & E_{23}'(h) &= E_{23}(h) - qE_{13}^p(h). \\ E_{14}'(h) &= E_{14}(h) - qE_{24}^p(h). \\ E_{24}'(h) &= E_{24}(h) - qE_{14}^p(h). \end{aligned} \quad (10)$$

本残響補正值を使用するか否かについてはフレームごとに位相適合性を含めて判断するようにし、以下位相適合性の評価式  $D(h)$ ,  $D'(h)$  を算出する。

$$\begin{aligned} D(h) &= [\{E_{11}(h)E_{22}(h)E_{13}(h)E_{24}(h)\}^{1/4} \\ &\quad - \{E_{12}(h)E_{21}(h)E_{14}(h)E_{23}(h)\}^{1/4}] \\ &\quad / [\{E_{11}(h)E_{22}(h)E_{13}(h)E_{24}(h)\}^{1/4} \\ &\quad + \{E_{12}(h)E_{21}(h)E_{14}(h)E_{23}(h)\}^{1/4}]. \\ D'(h) &= [\{E_{11}'(h)E_{22}'(h)E_{13}'(h)E_{24}'(h)\}^{1/4} \\ &\quad - \{E_{12}'(h)E_{21}'(h)E_{14}'(h)E_{23}'(h)\}^{1/4}] \\ &\quad / [\{E_{11}'(h)E_{22}'(h)E_{13}'(h)E_{24}'(h)\}^{1/4} \\ &\quad + \{E_{12}'(h)E_{21}'(h)E_{14}'(h)E_{23}'(h)\}^{1/4}]. \end{aligned} \quad (11)$$

位相  $h$  に対応する符号判定テーブルを  $B(h)$ 、位相判定テーブルを  $T(h)$  (初期値を全て0に設定) とすると、以下の優先順位で符号の仮判定を行なう。

- 1)  $|D(h)| > |D'(h)|$  かつ  $E_{11}(h)E_{13}(h) > E_{21}(h)E_{23}(h)$  かつ  $E_{22}(h)E_{24}(h) > E_{12}(h)E_{14}(h)$  の場合：  
 $B(h)=0, T(h) \leftarrow T(h)+D(h)$ .
- 2)  $|D(h)| > |D'(h)|$  かつ  $E_{11}(h)E_{13}(h) < E_{21}(h)E_{23}(h)$  かつ  $E_{22}(h)E_{24}(h) < E_{12}(h)E_{14}(h)$  の場合：  
 $B(h)=1, T(h) \leftarrow T(h)-D(h)$ .
- 3)  $|D'(h)| > |D(h)|$  かつ  $E_{11}'(h)E_{13}'(h) > E_{21}'(h)E_{23}'(h)$  かつ  $E_{22}'(h)E_{24}'(h) > E_{12}'(h)E_{14}'(h)$  の場合：  
 $B(h)=0, T(h) \leftarrow T(h)+D'(h)$ .
- 4)  $|D'(h)| > |D(h)|$  かつ  $E_{11}'(h)E_{13}'(h) < E_{21}'(h)E_{23}'(h)$  かつ  $E_{22}'(h)E_{24}'(h) < E_{12}'(h)E_{14}'(h)$  の場合：  
 $B(h)=1, T(h) \leftarrow T(h)-D'(h)$ .
- 5)  $|D(h)| > |D'(h)|$  かつ  $E_{11}(h)E_{13}(h)E_{22}(h)E_{24}(h) > E_{21}(h)E_{23}(h)E_{12}(h)E_{14}(h)$  の場合：  
 $B(h)=0, T(h) \leftarrow T(h)+D(h)$ .
- 6)  $|D(h)| > |D'(h)|$  かつ  $E_{11}(h)E_{13}(h)E_{22}(h)E_{24}(h) \leq E_{21}(h)E_{23}(h)E_{12}(h)E_{14}(h)$  の場合：  
 $B(h)=1, T(h) \leftarrow T(h)-D(h)$ .
- 7)  $|D'(h)| > |D(h)|$  かつ  $E_{11}'(h)E_{13}'(h)E_{22}'(h)E_{24}'(h) > E_{21}'(h)E_{23}'(h)E_{12}'(h)E_{14}'(h)$  の場合：  
 $B(h)=0, T(h) \leftarrow T(h)+D'(h)$ .
- 8)  $|D'(h)| > |D(h)|$  かつ  $E_{11}'(h)E_{13}'(h)E_{22}'(h)E_{24}'(h) \leq E_{21}'(h)E_{23}'(h)E_{12}'(h)E_{14}'(h)$  の場合：  
 $B(h)=1, T(h) \leftarrow T(h)-D'(h)$ .

全ての位相  $h$  において上記仮符号を判定したら、初期フレームから積算される位相判定テーブル  $T(h)$  の値が最大となる  $h_{max}$  を探索し、これを最適位相と判断し、そのときの候補符号  $B(h_{max})$  を本フレームにおける符号と確定する。

#### 4.3. バイトデータの埋め込みシーケンス

前節までは、フレーム単位にビットデータを埋め込み、抽出するアルゴリズムを説明してきたが、本節では実用頻度の高い7ビットのASCII文字に誤り検出符号を付加した一連のバイトデータを信号に埋め込む手法について述べる。

通信の分野で一般に使われる誤り検出符号は、予測不能な種々の外乱によるビット誤りを対象としたもので、本稿のような電子透かし抽出においてはそれ以外の要因でもビット誤りが発生する。例えば、原信号の状態により4.1.2節で述べた最適埋め込みを実現できない場合があり、例えば4.1.1節で述べた白色雑音成分の

表1 [16,7]ビット非組織的ハミング符号表

ASCII	16-bits Unstructured Hamming Codes							
00-07	1	62	454	505	1610	1653	1933	1970
08-0F	2708	2731	2899	2924	3295	3296	3352	3367
10-17	12364	12403	12683	12724	13831	13880	14272	14335
18-1F	15065	15078	15134	15137	15506	15533	15701	15722
20-27	20631	20648	20816	20847	22236	22243	22299	22308
28-2F	23042	23101	23493	23546	23625	23670	23950	23985
30-37	24794	24805	24861	24866	26257	26286	26454	26473
38-3F	27215	27248	27528	27575	27652	27707	28099	28156
40-47	37462	37481	37777	37806	37917	37922	38362	38373
48-4F	39107	39164	39172	39227	40584	40631	40783	40816
50-57	41499	41508	41948	41955	42064	42095	42391	42408
58-5F	43150	43185	43337	43382	44741	44794	44802	44861
60-67	49856	49919	49927	49976	50315	50356	50508	50547
68-6F	51285	51306	51602	51629	52766	52769	53209	53222
70-77	62093	62130	62282	62325	62662	62713	62721	62782
78-7F	63512	63527	63967	63968	65107	65132	65428	65451

[Caution] Three codes of 75(62713),7D(65132),7E(65428) are duplicated for cyclic bit patterns.

表2 [16,7]ビット非組織的ハミング符号の巡回ビットパターン表

ASCII	Hamming Code	Cyclic Bit Pattern (16 patterns for each code)							
		2	4	8	16	32	64	128	256
0	1	512	1024	2048	4096	8192	16384	32768	
		124	248	496	992	1984	3968	7936	15872
1	62	31744	63488	61441	57347	49159	32783	31	
		908	1816	3632	7264	14528	29056	58112	50689
2	454	35843	6151	12302	24604	49208	32881	227	
		1010	2020	4040	8080	16160	32320	64640	63745
3	505	61955	58375	51215	36895	8255	16510	33020	
		3220	6440	12880	25760	51520	37505	9475	18950
4	1610	37900	10265	20530	41060	16585	33170	805	
		...	...	...	...	...	...	...	...
127	65451	65367	65199	64863	64191	62847	60159	54783	44031
		22527	45054	24573	49146	32757	65514	65493	

加算が品質上行なえない場合、原信号レベルが0に近い区間ではデータを適切に埋め込むことができない。そのため、その箇所では抽出時にビットエラーになる可能性が高い。

また、非同期シリアル通信においてはバイトデータの先頭ビットを示す同期符号を挿入する方法がとられるが、文献[5][6]では、同期符号を挿入せず、誤り検出符号で代用する方法をとっていた。

このように、外乱によるビット誤り以外に、原信号レベルに起因する埋め込み処理の失敗、バイト先頭ビットの識別という多機能をもたせた誤り検出符号を検討した結果、文献[6]では最小ハミング距離4を確保した非組織的ハミング符号を採用した。その後、それでも誤り検出能力が十分でないことが判明し、本稿では、最小ハミング距離6を確保した16ビットの非組織的ハミング符号を採用することにした。

表1に示される16ビットの各符号間には6以上のハ

ミング距離を確保してあり、7ビットのASCII符号を表1に従って16ビットに変換して埋め込むと、5ビットまでのエラー検出と2ビットのエラー自動訂正が行なえる。従って、一度16ビットのデータがエラー無しに抽出できれば、バイトデータの同期はとれていると判断できるため、それ以降は2ビットの自動訂正機能を働かせることができる。ただし、本符号だけでは、同期不良を完璧に判定することはできないことが判明した。

実際の運用では、数バイトのID情報を時系列に繰り返し埋め込む手法がとられる。そこで、本稿では、16ビットの単一ハミング符号が繰り返し埋め込まれている場合に限定して、埋め込み時のバイトシーケンスに同期させずに、16ビットの符号を抽出する手法を提案する。16フレーム分の抽出が完了し、16ビットのデータが揃ったら、表2で示すようなハミング符号の巡回ビットパターンのいずれかに合致するかの判定を行う。本表は各16ビットのハミング符号に対して1ビットづつ巡回シフトさせた16通りの巡回ビットパターンをあらかじめ作成したものである。本手法を用いる場合は、2ビットのエラー自動訂正(1ビットは可)は使用できず、後半の75(62713),7D(65132),7E(65428)の3種の符号は前半の巡回パターンと重複するため使用することができない。

### 5. 提案手法の評価実験

文献[3]と同様に2分のCD音楽素材(44.1kHz/16bits/2-ch)の左右チャンネルを同一信号に変換したステレオ信号に対して、本稿で提案したステレオ埋め込み手法を適用してデータを埋め込んだ。左側にはASCII"0"、右側には"G"を表1の16ビットの非組織的ハミング符号に変換して全信号区間にフレームサイズ2048サンプルで繰り返し埋め込み、中央部には"1"が抽出されるように埋め込みレベルを4.1.2節で述べたように2段階に変位させた。この電子透かしが埋め込まれたステレオ信号に対して、左右チャンネルの音量を次の15通りに変化させながら、15通りの混合モノラル信号を作成し、各々作成されたWAVファイルより電子透かし抽出を試みた。

{(L:100%, R:0%), (L:79.34%, R:20.42%), (L:72.53%, R:26.79%), (L:66.22%, R:32.10%), (L:60.46%, R:38.46%), (L:55.27%, R:42.12%), (L:50.47%, R:46.13%), (L:46.13%, R:46.13%), (L:46.13%, R:50.47%), (L:42.12%, R:55.27%), (L:38.46%, R:60.46%), (L:32.10%, R:66.22%), (L:26.79%, R:72.53%), (L:20.42%, R:79.34%), (L:0%, R:100%)}

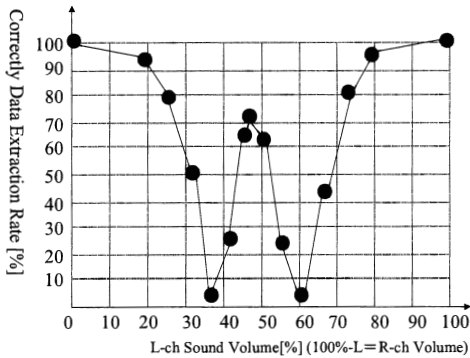


図8 左右チャンネル信号の混合音量バランスとデータ抽出率の測定結果

その結果を図8に示す。中央部では左右近辺での抽出に比べ若干感度が悪く、抽出可能な位置範囲が狭いが目論見通り第3の符号が安定して抽出できていることがわかった。図8は抽出時に表2の巡回ビットパターンによる判定を用いた場合であるが、追加実験で、巡回ビットパターンを使用せずハミング符号によるエラー判定を用いた場合についても試した結果、同様な特性でカーブが更に急峻になり、下方向の2つの抽出不能ピークの区間が延びるという結果になった。

次に、定量的な解析は行っていないが、2台のスピーカーで再生し、マイクロフォンを左右にゆっくり動かすと、抽出される3種の符号が切り替わることがアナログ空間でも確認できた。この時、同様に、巡回ビットパターンによる判定を用いた方がスムーズに符号の切り替えが行われることが確認できた。

## 6. あとがき

本稿で提案した改良型電子透かし埋め込み方式により、ステレオ信号の左右チャンネルに異なる符号を埋め込み、左右のスピーカー近辺では各々のチャンネルに埋め込まれた符号が抽出されるとともに、左右チャンネルが混合する中央部では第3の符号を安定して抽出できることが確認できた。ただし、中央部では感度が若干悪く、抽出可能な空間が狭いという問題も明確になり、今後の改善課題としたい。本技術を例えば文献[7]に記載のミュージアムでの展示解説に応用すれば、1台のステレオ音楽プレーヤーで3種のIDを同時に送出することができ、設置する音楽プレーヤーの台数を削減することができる。

本稿では左右チャンネルの中央部で第3の符号が抽出できるようにする手法について述べてきたが、左右チャンネル間で音量バランスを変化させたり、埋め込

みレベルの変位パラメータを変化させれば、第3の符号が抽出される位置を中央部から左右いずれかの方向に偏移させることもできる。これらの偏移操作を時系列に行えば、より木目細かい位置検出が実現できるといふ期待がもてるので、今後の検討課題としたい。

更に、スピーカを2次的に配置し、その際の検出分解能等を調査し、本稿で提案した手法が本格的な2次元位置検出システムへ展開できるか、GPS等の既存の位置検出システムを補完できるか等についても今後検討してゆきたい。

また、ユビキタス音響空間を実現するため、抽出ソフトウェアを携帯電話等に実装し、携帯端末を音源に近づけなくても、できれば衣服のポケットやカバンに入れた状態で、周囲に流れているBGMやアナウンスに埋め込まれた位置情報をリアルタイムに受信し、行動履歴が記録できるようにすることを目標に開発を進めて行く予定である。

## 文 献

- [1] 小野東：“第4章. 音に入る電子透かし”，「電子透かしとコンテンツ保護」，オーム社，第1版，pp.139-168，2001.
- [2] 九州大学・中島祥好「聴覚心理学」研究室 <http://www.design.kyushu-u.ac.jp/~ynhome/JPN/index.html>
- [3] 茂出木敏雄：“携帯電話で非接触抽出可能な音楽への電子透かし埋め込み技術の開発，”電気学会・電子情報システム部門誌，Vol.126-C，No.7，pp.825-831，July，2006.
- [4] 茂出木敏雄：“音響空間のユビキタス化に向けた電子透かし埋め込み容量の拡大技術，”電気学会・電子情報システム部門誌，Vol.127-C，No.7，pp.1013-1021，July，2007.
- [5] Toshio Modegi：“Construction of Ubiquitous Acoustic Spaces Using Audio Watermark Technology and Mobile Terminals，”IEEJ Transactions on Electrical and Electronic Engineering，Vol.2，No.6，pp.608-619，November，2007.
- [6] 茂出木敏雄：“音脈分擬を活用した携帯電話で抽出可能な音楽電子透かし技術—ゲンコード Marker—，”電気学会・電子情報システム部門誌，Vol.128-C，No.7，pp.1087-1095，July，2008.
- [7] 茂出木敏雄：“音楽電子透かしを用いたプッシュ型情報配信システムの試作，”電子情報通信学会2008年総合大会，DS-4-5，pp.S27-28，March，2008.