

音楽に著作権保護機能を埋め込む音響信号に対する ハイディング技術に関する研究

茂出木 敏雄

Research on Hiding Techniques for Acoustic Signals by Embedding Copyright Protection Functions into Music Signals

Toshio Modegi

Abstract

Watermarking techniques, with which we can embed copyright information without degrading content quality, have been developed in order to track illegal copies of image or audio contents. By development of robust embedding techniques against data compressions and analogue conversions, we can extract embedded information just by photographing images or recording sounds from a loudspeaker with our mobile phones. These advanced technologies are expected as a novel advertisement media. However, as illegal copy actions cannot be suppressed by a watermarking technique itself, researches on embedding copy-disturbance noises have also begun. The embedded copy-disturbance noises can be visible or audible by giving the content some changes such as analogue conversion with copy actions, and we can suppress distribution of illegally copied products. In this paper, we will overview fundamental information hiding techniques for audio signals. Then we present the other types of hiding researches on embedding secondary audio signals and copy-disturbance noises in order to protect audio contents from illegal copy actions.

Key Word

digital watermark, information hiding, G-encoder Mark, copyright protection,
copy-disturbance noise, auditory masking

[抄録]

コンテンツの違法コピーを追跡する方法として電子透かし手法が開発され、画像や音楽に対して原コンテンツの品質を劣化させずに著作権情報等を埋め込むことができる。この技術を発展させて、各種データ圧縮やアナログ変換に対する耐性をもたせたロバストな埋め込み技術の開発が進み、現状では、携帯電話やスマートフォンを用いて画像をカメラで撮影したり、スピーカ再生音を録音するだけで、画像や音楽に埋め込まれている情報を抽出することができ、新規な広告媒体として期待されている。しかし、当初の著作権保護という用途において、電子透かしだけでは違法コピー自体を阻止することはできないため、妨害雑音を埋め

込む手法も研究されている。これはコピー操作に伴いアナログ変換などコンテンツに何らかの改変を施すと、埋め込まれている妨害雑音が可視化または可聴化されるようにしたもので、違法コピー品の流通を阻止できる。本稿では、音響信号に情報を埋め込む情報ハイディング手法を概説し、情報以外の第2の音響信号や妨害雑音を埋め込み、音楽著作権を保護する研究事例について紹介する。

[キーワード]

電子透かし、情報ハイディング、ゲンコーダ Mark,
著作権保護、妨害雑音、聴覚マスキング

1. はじめに

先報告¹⁾において、「音楽作品に耳に聞こえないサインを埋め込む音楽電子透かし技術”ゲンコーダ®Mark”」というツールについて言及したが、本稿ではこの音楽電子透かし技術にフォーカスし、音楽に対して、耳に聞こえない情報あるいは種々の著作権保護機能を埋め込む研究事例について紹介する。

音楽というのは単に鑑賞して楽しむというエンターテインメント用途や、BGM²⁾のように環境を構築するための空間演出用途といった受動的な活用以外に、多くのヒトの精神に同時に訴える洗脳の効果を活用する宗教音楽や、精神または肉体的に病んでいるヒトに療養的効果をもたらす音楽療法のように、音楽そのものに備わる精神作用や療養機能を能動的に活用した試みも古くからなされている。また、BGMには空間演出効果だけでなくマスキング効果があり、飲食店やショッピングセンターなど人が集まるところにBGMを流すのは、人々の会話音が気にならないようにするといった騒音快音化やスピーチプライバシー保護効果¹²⁾¹³⁾を活用しているという能動的な面もある。

90年代にマルチメディアの情報処理技術が発展し、録音された音響信号に対してデジタルサンプリングを施しコンピュータ内に取り込み、ソフトウェアにより音楽の各種加工がデジタル的に行えるようになった。これは音楽制作にDTM (Desk Top Music) という革命をもたらただけでなく、MPEGなどのデータ圧縮技術の開発を推進させ、ネットを介して音楽作品が簡単に交換できるようになってきた。このデジタル加工技術の進展はユーザ（制作者および視聴者）に利便性を与える反面、悪意のある人々にも利便性を与え、違法コピーを作成することが、より簡便に、より高品質に行えるようになってきた。そのような背景から、著作権保護のニーズが高まり、画像や音楽の信号に著作権情報を埋め込み、違法コピーを追跡したり牽制したりする機能を画像や音楽コンテンツそのものにもたせるといった電子透かし技術の開発が始まった。

電子透かし技術には後述するように、透かしを可視または可聴な形態で埋め込む手法もあるが、知覚されないように情報を埋め込む技術（情報ハイディング、ステガノグラフィーとも言われる）の方が、著作権保護以外の広範な用途に活用できる可能性を秘めているため進展した。特に、各種データ圧縮やアナログ変換に対する耐性をもたせたロバストな埋め込み

技術の開発が進み、現状では、印刷画像やディスプレイ画面を携帯電話やスマートフォンのカメラで撮影したり、再生されているスピーカに携帯電話やスマートフォンをかざすだけで、画像や音楽に埋め込まれている情報を抽出することができ、新規な広告媒体として期待されている³⁾⁻¹¹⁾。

このように電子透かし技術の活用範囲は拡大したが、著作権保護という用途においては、違法コピー行為を監視するといった間接的な抑止効果しかなく、当初の活用範囲からは逸脱していない。逆に、電子透かし技術の進展により、透かしが埋め込まれた画像や音楽の品質は原素材と遜色無いレベルを維持していることが裏面に出て、透かしが埋め込まれた状態で高品質な違法コピー製品が容易に作成されてしまう。そこで、情報を埋め込む代わりに、妨害雑音を埋め込む手法も研究されている¹⁴⁾⁻¹⁷⁾。電子透かしと同様に、画像や音楽に埋め込まれた妨害雑音は通常再生状態では知覚されないが、アナログ変換などコンテンツに何らかの改変を伴うコピー操作を施すと、埋め込まれている妨害雑音が可視化・可聴化されるようにしたもので、違法コピー品の流通を阻止できる期待がもてる。

埋め込み対象のコンテンツを音楽に限定すると、情報以外の耳に聞こえない種々の付加価値機能を埋め込む手法の研究事例として、別の音楽や副音声の埋め込んだり、マスキング機能を強調させる効果を埋め込む試みも行われている¹²⁾。後者はスピーチプライバシー保護に活用されており、特殊加工された BGM を流すことにより周囲から漏れ聞こえてくる会話音を不明瞭にすることができる¹²⁾。

以下、音楽に対する情報ハイディング技術に焦点を当て、その基本手法やロバストな埋め込み手法について述べ、音楽著作権を保護するための新規なハイディング技術開発事例について紹介する。

2. 音楽電子透かし技術の基本的手法

電子透かし技術には紙幣の透かしのよう可視（音楽では可聴）のタイプと不可視（音楽では不可聴）のタイプに大別され、紙幣のように複写をとると透かしが消える弱い（fragile）タイプと複写をとっても透かしが消えない頑強な（robust）タイプにも大別できる。また、紙幣の透かしとは逆に、不可視（不可聴）の透かしが埋め込まれている媒体に対して複写をとると可視（可聴）のタイプになるように開発された手法もある¹⁴⁾⁻¹⁷⁾。以下、音楽に可聴または不可聴な透かしを埋め込む基本的な手法について述べる。

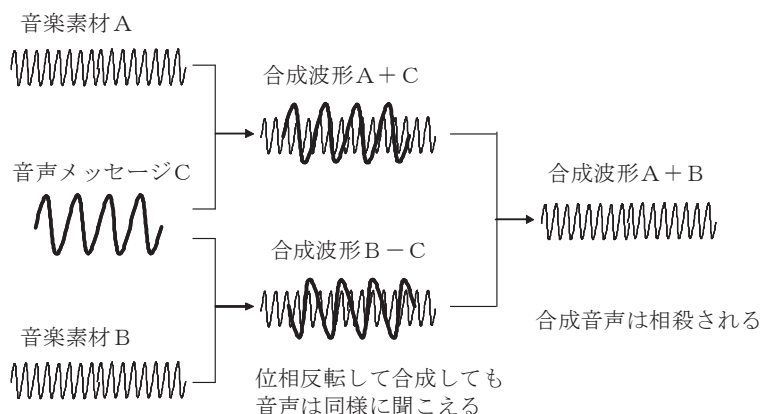


図1 波形合成音の原素材に可聴の電子透かしを埋め込む方法

2. 1. 可聴な電子透かしの埋め込み手法と波形合成による不可聴化

可聴な電子透かしとは、図1の上段に示されるように、著作権保護対象の音楽素材Aに著作権主張を音声で表現した音声メッセージCを波形合成させることにより実現できる。音楽素材のサンプル提供に効果的で、正規ライセンスが締結された段階で電子透かしが付加されていない原素材を提供するか、電子透かしを除去する方法をとる。また、文献2)ではマトリックス音楽と称するミックスダウン前の原素材をセットにして提供し、視聴者は5種の素材を選択の上、波形合成をしながら視聴させる方法を提案している。図1は文献2)の方法を2種の波形合成に簡略化したもので、視聴者には音楽素材Aと音楽素材Bを渡し、両者を合成しないと意味のある音楽にならないように制作されている。音楽素材Aと音楽素材Bを視聴者に渡す際に、あらかじめ双方に音声メッセージCを合成しているため、各々を単独に持ち出そうとすると、音声メッセージCが同時に聴取され他用途に使用することはできない。しかし、音声素材B側に合成されている音声メッセージCの位相は反転されているため、両者を合成して再生すると音声メッセージCがキャンセルされて音楽素材Aと音楽素材Bとの合成音のみが再生される。

以上のように可聴な電子透かしは、正規視聴者に対しては電子透かしが削除された状態になるため、この段階で違法コピーが自由に行われてしまうという問題点が残る。そこで、次節以降に説明する不可聴な電子透かしを埋め込む技術が重要になる。

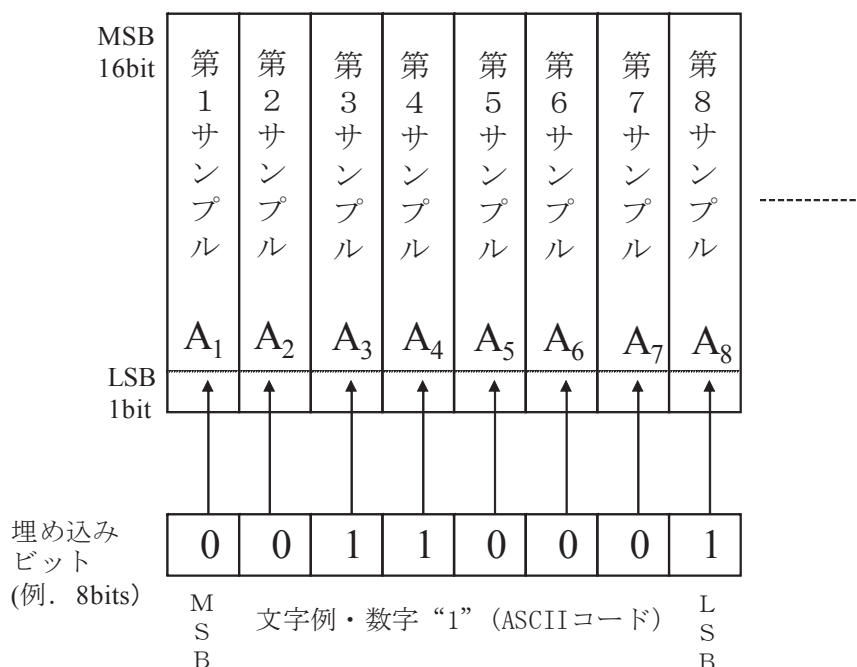


図2 下位ビット置換法による不可聴な電子透かしの埋め込み手法

2. 2. 下位ビット置換法による不可聴な電子透かしの埋め込み手法

コンピュータに取り込まれる波形形式の音楽信号は、所定のサンプリング周波数でサンプ

リングされた正負符号付きの多値配列になっている。例えば、音楽CDの場合、44.1kHz のサンプリング周波数でステレオLRの2チャンネルの音楽信号が各々 16bits でサンプリングされ、Windows ではWAV 形式ファイル、Mac ではAIFF 形式ファイルで保存される。図2の例では、16 ビットの A_1 から A_8 の最初の8個のサンプルが図示されているが、音楽CDの場合、奇数番目のサンプルがLチャンネル、偶数番目のサンプルがRチャンネルとなる。このように各サンプルが 16 ビット以上あると（業務用では 24 ビットが多用される）、各サンプルの最下位ビット LSB の値を反転させても、原音との相違を再生音で識別することは Gold Ear をもつ専門家でも難しい。

そこで、図2のように、埋め込み対象の電子透かしとして8ビットのバイト配列（文字列など）を準備し、各バイトの8ビットのデータを8サンプルに対応させ、各サンプルのLSBの値と置換する。このような方法で埋め込まれた音楽は、原音との相違を再生音で識別することは難しく、後述する種々の電子透かし手法の中で最もアルゴリズムが単純である割に、最も品質劣化が少ないという特徴をもつ。埋め込まれた電子透かしの抽出するには、各サンプルのLSBを順次抽出し、8サンプルごとに1バイトのデータを生成すればよい。

しかし、アナログ変換（マイクやオーディオライン経由で再録音する）やMP3などのデータ圧縮を施したり、波形データの一部を切り出すという改変を行うと、埋め込まれた電子透かしが簡単に破壊されてしまうという問題がある。そこで、次節以降に説明するロバストな電子透かしの埋め込み手法が提案されている。

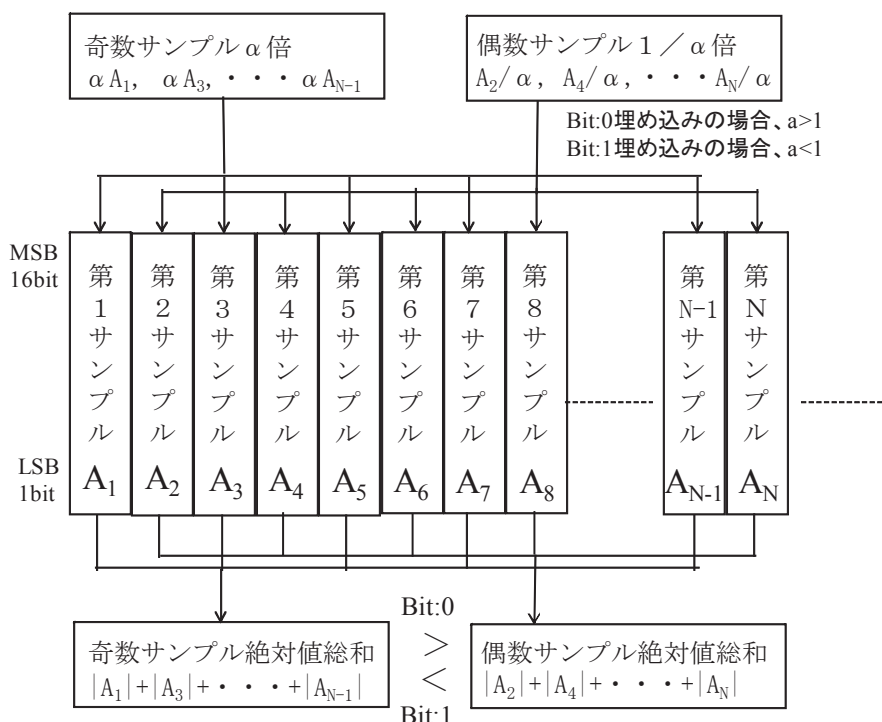


図3 統計的偏りを利用したロバストな電子透かしの埋め込み手法

2. 3. データ圧縮に対する耐性をもたせた不可聴な電子透かしの埋め込み手法

前節で述べた手法に対して、データ圧縮等の改変に対するロバスト性をもたせるには、最下位ビットだけでなく、もう少し上位のビットにも埋め込みを行えば良い。しかし、それに伴って音質劣化が顕著になるという問題が発生する。そこで、個々のサンプルに対してビットを埋め込むのではなく、サンプル群にビットを埋め込むという統計的アプローチをとる。即ち、平均的には均等な特性をもつ2つのサンプル群に対して、ビットの0または1に応じて統計的な偏りを施すと、個々のサンプルに対する改変量は少なくすむ上、データ圧縮等の改変に対するロバスト性が増大する。

図3はその一例で、奇数番目のサンプルと偶数番目のサンプルを各々 $N/2$ 個（例、 $N=4096$ ）を組みにして2つのサンプル群を構成する。各サンプル群のサンプル絶対値の総和は通常同じような値になる。そこで、ビット0を埋め込む時は、奇数サンプル群のサンプル絶対値の総和が偶数サンプル群のサンプル絶対値の総和より顕著に大きくなるように、奇数サンプル群の各サンプル値に所定の1以上の実数値 α を乗算し、偶数サンプル群の各サンプル値に所定の1以上の実数値 α を除算する。ビット1を埋め込む時は逆の操作を施し、奇数サンプル群のサンプル絶対値の総和が偶数サンプル群のサンプル絶対値の総和より顕著に小さくなるようにする。実数値 α は、ビットを埋め込むサンプル群ごとに変更する必要がある、ビットを埋め込む対象の奇数および偶数のサンプル群のサンプル絶対値の総和を各々算出し、双方のサンプル群が所定の大小関係になる範囲で、できるだけ1に近い値になるように調整する。

このような方法でビットを埋め込むと、埋め込み可能な情報量は N サンプルごとに1ビットとなり、前節で述べた方法（1サンプルごとに1ビット）に比べると極めて少なくなるが、データ圧縮等の改変を施しても電子透かしが維持される特徴をもつ。しかし、アナログ変換（マイクやオーディオライン経由で再録音する）や波形データの一部を切り出すという位相を変える改変を行うと、埋め込まれた電子透かしが破壊されてしまうという問題は相変わらず残り、次節で述べるロバストな電子透かし埋め込み手法が提案されている。

3. 携帯端末で非接触抽出可能な電子透かしの埋め込み手法「ゲンコード®Mark」

本節では、前節までに指摘した電子透かし埋め込み手法における問題点を解消し、文献1)でも概要紹介したロバストな電子透かし埋め込み手法「ゲンコード®Mark」について述べる。この手法は、音楽電子透かし技術を音楽の著作権管理だけでなく、ユビキタス音響空間⁵⁾⁶⁾への応用を目指して開発されたものである。ユビキタス音響空間⁵⁾⁶⁾とは、空間に分布する音源の情報を携帯端末で非接触に抽出し、音源に関連するサイバー空間と連動できるようにする構想で、これを実現する具体的手法として、携帯電話やスマートフォン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た。本手法では、前節までに指摘した、アナログ変換（マイクやオーディオライン経由で再録音する）やMP3などの各種データ圧縮を施したり、波形データの一部を切り出すといった種々の改変に対する強固なロバスト性をもたせる必要がある。

音楽著作権保護を目的とした電子透かし手法では、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に各種圧縮・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、音源に埋め込まれた電子透かしを携帯電話で非接触に抽出する機能を実現するためには、アナログ空間を経由することに対する耐性と、携帯電話の音声入力信号処理系に対する耐性の双方を備えている必要がある。本技術の開発に着手した当時（2005 年頃）はスマートフォンが普及しておらず 2G から 3G 携帯電話への移行時期であったため、電子透かしを抽出する携帯端末として 3G 携帯電話にフォーカスを当てた。また、この当時スマートフォンと同等な処理能力をもつ携帯端末として、Windows Mobile を搭載した PDA(Personal Digital Assistant)、任天堂 DS があり、これらも電子透かし抽出の実験に使用した。

3. 1. 3G 携帯電話で非接触抽出可能な音楽電子透かし技術の開発要件

3G 携帯電話では電話回線帯域と同様に 4kHz を越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。更に、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の要件として、4kHz 未満の聴感特性が最も鋭敏な音域に、振幅歪みが加わ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。手始めに、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⁾。空間分割法、周波数分割法、時分割法の 3 つの基本方式を順次提案し、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)・6)}。しかし、これらのステレオ方式の提案手法はモノラル再生やヘッドフォ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 側信号による透かし抑圧効果を受けやすく、抽出精度が低下するという問題があり実用化が困難であった。

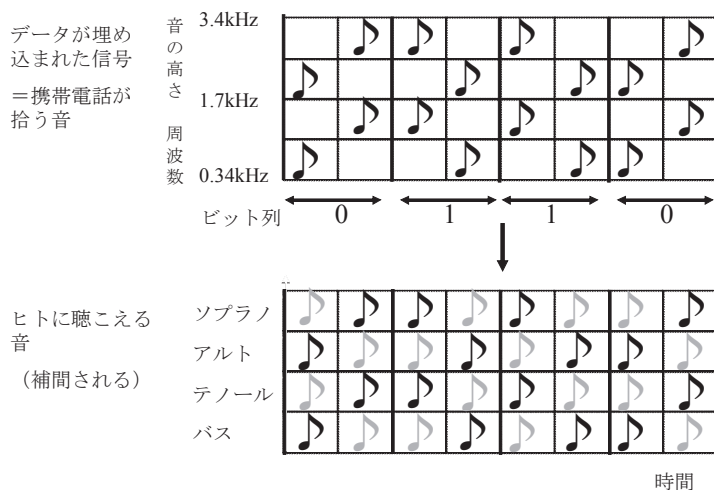


図4 音脈分疑を活用した口バストな音楽電子透かし技術の開発思想

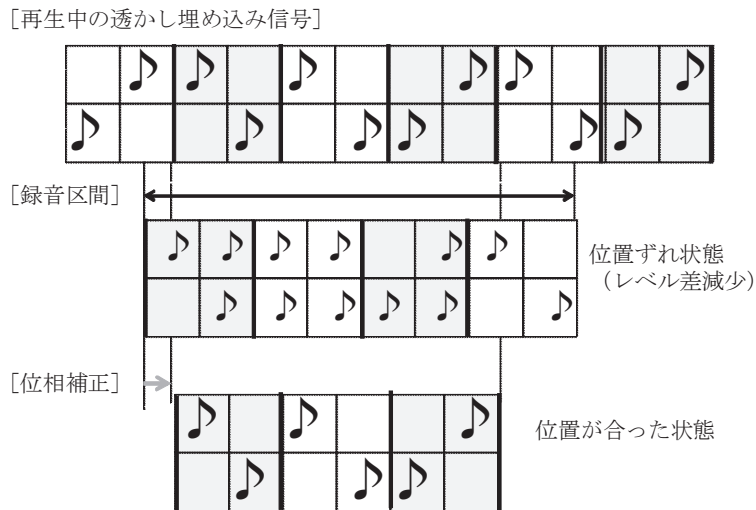


図5 録音信号中の先頭フレームの探索処理

3. 2. 音脈分凝を活用した口バストな電子透かし埋め込み技術の開発思想

前節で述べた問題点を克服するアイデアとして、聴覚心理学の分野で知られているヒト聴覚系の錯覚現象である音脈分凝（おんみゃくぶんぎょう）⁷⁾⁹⁾に着目した。これは音声の分野ではカクテルパーティ効果として知られている現象と同様な原理に基づくもので、ヒトの会話音声や音楽のメロディー・和声進行などの音素の流れ（音脈）が、途中で雑音等で遮られ分断されても、ヒトの脳聴覚野で補間し音脈を自然と再構築する現象である。図4はこれを電子透かし埋め込みに応用する手法を示したものである。横軸は時間で埋め込み対象の音楽信号を 50[msec] 程度のフレーム間隔で分割し、各フレームに1ビットのデータを埋め込む様子を示している。図の縦軸は周波数で埋め込み対象の周波数を携帯電話で受信可能な電話回線帯域（0.34 ～ 3.4kHz）とし、この帯域を4分割する。この4つの分割バンドは音楽でいうソプラノ・アルト・テノール・バスにほぼ対応する。各フレームに対して時間軸方向に更に2分割し、周波数方向を含め8分割の信号成分に対し、図4上に示すように市松模様のように信号成分を抜いてゆく。このとき、埋め込むビット値に基づいて、2通りの方向で信号成分を抜いてゆく。図4上に示されるパターンをヒトが聴取すると、図4下のように抜けた信号成分が補間されて聞こえるため、原音とあまり差異無く聞こえてしまう。しかし、マイクロフォンや携帯電話は図4上に示される市松模様のパターンをそのまま読み取るため、埋め込まれたビットパターンを抽出することが可能になる。このような方法で、視聴者には気づかれずに 20 [bps] のレートで音楽信号に情報を埋め込むことができることを確認した⁷⁾⁹⁾。

3. 3. 音楽の局所録音区間からの電子透かし抽出における技術課題と対策

図2・図3で示した方法では、電子透かしが埋め込まれた音楽ファイルの先頭から透かしデータを順次抽出する方法しか想定していなかったが、本節のように携帯端末で透かしを抽

出する運用では、再生中の音楽の任意の区間から透かしの抽出を行えるようにすることが必須となる。前節で述べたように、電子透かしは与えられた音響信号に対して、50[msec] 程度のフレーム間隔で1ビットずつ埋め込んでいる。埋め込まれているビットデータは文字列などバイト配列の各バイトを8ビットにシリアル展開したビットデータとなる。このような形態で埋め込まれた音響信号に対して、アナログ経由で任意の区間を録音した信号から透かしバイト配列を抽出するにあたり、以下2つの不確定要素がある。

(1) 録音信号中のビットが埋め込まれているフレーム範囲を特定できない。(先頭フレームの開始時間位置が不明)

(2) フレームに埋め込まれているビットが特定できても、バイト配列中のどのビットか特定できない。(バイトデータの先頭ビット位置が不明)

上記(1)の問題については、図5に示されるように、録音信号中に設定する解析フレームをフレーム幅の範囲で数段階ずらしながら抽出処理を行い、2つの対角要素(左上および右下成分と左下および右上成分)の差が最も大きくなる位置が埋め込み時に設定したフレーム位置であると推定する。文献3)ではフレーム間隔50[msec]の範囲で6段階に位相をずらして6通りの信号解析を行いながら、最良の位置での抽出結果を採用する方法を提案している。

表1 16ビット・ハミング巡回符号表の例

基準符号 (7 ビット)	ハミング符号 (16 ビット)				ハ ミ ン グ 距 離 3
0	1	0000	0000	0000 0001	
1	15	0000	0000	0000 1111	
2	51	0000	0000	0011 0011	
3	85	0000	0000	0101 0101	
4	91	0000	0000	0101 1011	
5	157	0000	0000	1001 1101	
...					
129	22363	0101	0111	0101 1011	
130	23295	0101	1010	1111 1111	
131	28655	0110	1111	1110 1111	
132	65535	1111	1111	1111 1111	

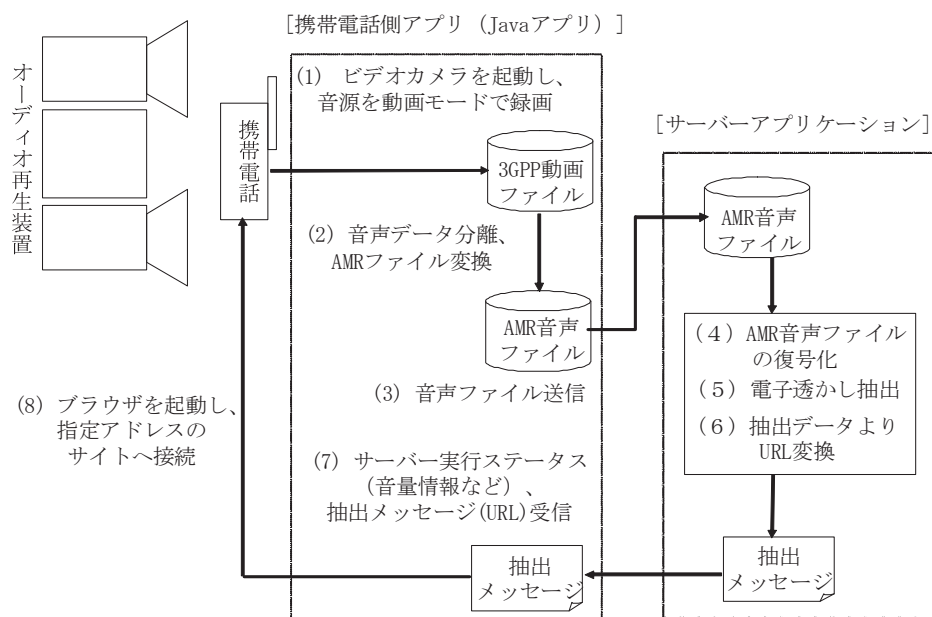


図7 3G携帯電話を用いた電子透かし抽出システムの構築例

スでは録音制御が困難で、処理能力的にも透かし抽出ソフトウェアを端末内に実装することは困難である。そこで、携帯電話とサーバーを連携させて透かし抽出を実現する手法を考案した。図7に示される系で、音楽に埋め込まれた透かし情報をもとに携帯電話を所定のサイトへ誘導させることができることを確認した。まず、ユーザは携帯電話内のJavaアプリを起動し、同アプリにより動画カメラが起動され、ユーザのクリック指示により3秒程度の撮影（実際は音声録音）を実行し、カメラ撮影が終了すると3GPP形式で動画ファイルが保存され所定のサーバーへ送信される。サーバー側では受信した動画ファイルより音声データのみを復号化し、透かし抽出を行い、抽出した情報をもとにURLに変換し、待機中の携帯電話側のJavaアプリへ返送する。これにより、携帯電話側は取得されたURLのサイトへ接続できる。本システムを試作した結果、携帯電話からサーバーへ動画ファイルを伝送する時間がネックになることが判明し、ダミーの映像フレームを伝送する時間が予想外に冗長であった。そこで、Javaアプリにより携帯電話側で、動画ファイルよりAMR形式の音声データ部のみ抽出して送信するように改善し、応答性が倍増した。

3. 5. ステレオ音響信号に対する1次元空間位置検出機能の埋め込み

本提案の電子透かし埋め込み方式では、音響信号がステレオの場合、左右のチャンネルに同一の符号を埋め込めば、左右スピーカ近傍や中央のいずれの場所に携帯端末が位置しても同一の符号が抽出でき、電子透かし抽出の感度範囲が広がる。電子透かしを埋め込む際、3.2節で述べたように対角2対の信号成分の一对の強度を下げて強度差を形成するが、他方の対の強度を上げることもでき、強度差を調整することができる。そこで、対角2対の信号成

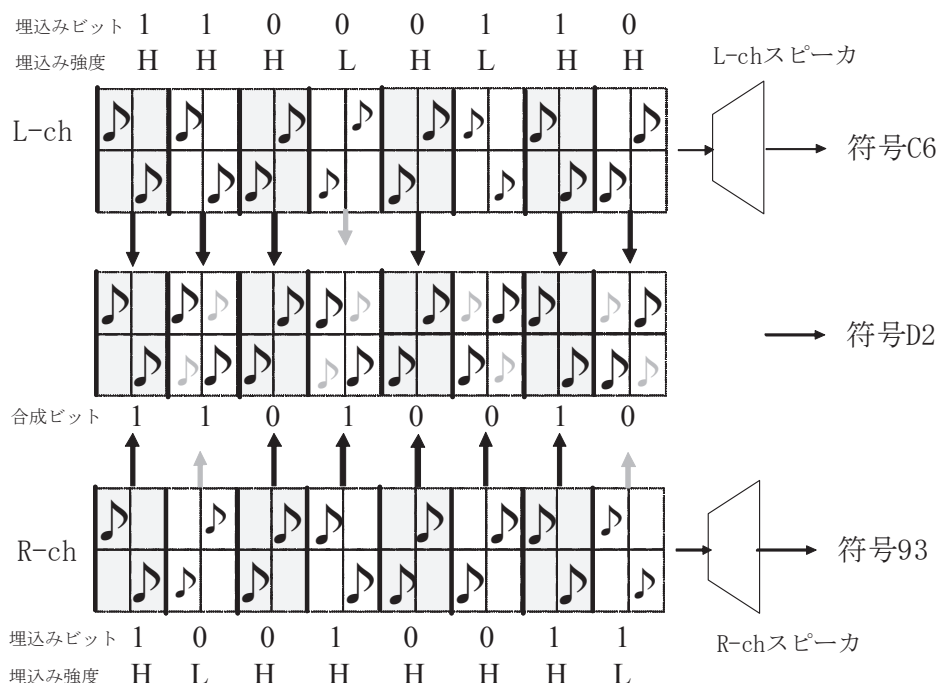


図8 ステレオ音響信号に対する1次空間位置検出機能の埋め込み事例

分の強度差を大きめにする強い埋め込み H と、強度差を小さめにする弱い埋め込み L の2種の埋め込みパターンを用いて、音響信号の左右のチャンネルに異なる符号を埋め込むようにした事例が図8である。図8で、Lチャンネル側には、8ビット符号C6 "11000110" が埋め込み強度パターン "HHHLHLHH" というモードで埋め込まれており、Rチャンネル側には、8ビット符号93 "10010011" が埋め込み強度パターン "HLHHHHHL" というモードで埋め込まれている。このように埋め込まれたステレオ音響信号を2つのスピーカで再生すると、携帯端末がLチャンネルスピーカに近いと符号C6が抽出され、Rチャンネルスピーカに近いと符号93が抽出される。ところが、双方のスピーカの中央では、C6でも93でもない符号D2が抽出される。これは一方のチャンネルに弱いモードで埋め込まれ、かつ他方のチャンネルに強いモードで埋め込まれているビットは、強いモードで埋め込まれているビットの値が優先され、中央部で新規な符号D2 "11010010" が動的に形成されるためである。

文献¹⁰⁾¹¹⁾では、埋め込み強度パターンをL,M,Hの3段階に拡張してステレオ信号の左右のチャンネルに異なる符号を埋め込み、左右スピーカの中央部ではL側によっているか、R側によっているかでも中央部と異なる符号が抽出できる、1次元方向に5種類の符号の抽出を可能にする方法も提案している。また、1次元方向に3種類の符号が抽出できる機能を2次元的に拡張させ、5. 1チャンネルサラウンド音響のFL、FR、BL、BRの4つのチャンネルに4種類の符号を埋め込むと、3×3の9種類の符号が抽出できるようにする方法も提案している。

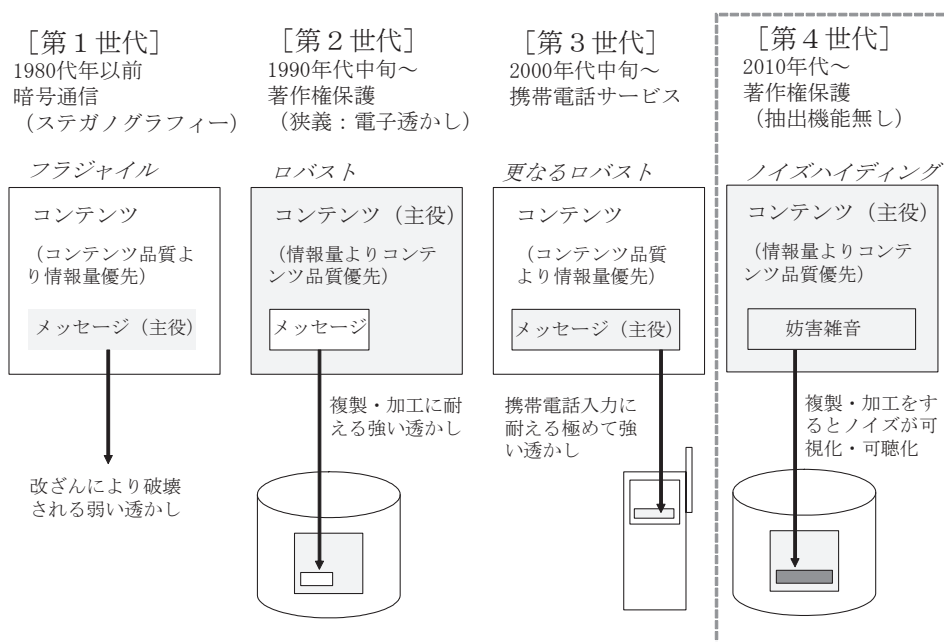


図9 電子透かし技術開発の変遷と第4世代の登場

4. 副音響信号や妨害雑音の埋め込みにより違法コピーを妨害する手法

前節までに述べてきた種々の電子透かし手法、具体的には画像や音楽に情報を埋め込む情報ハイディング手法をまとめると、図9のようになる。はじめに同図に基づいて、これまでの電子透かし技術開発の変遷を概観する⁴⁾。

第1世代は、ステガノグラフィーといわれ主に軍事目的の暗号通信に使用されたもので、コンテンツを隠れ蓑にして秘密のメッセージを伝送する方法である。一般的な暗号通信では意味不明なデータが流れるため、何らかの暗号化が施されていることが類推され、暗号破りなどの行為を助長してしまう。一方、ステガノグラフィーは見かけ上、意味のある画像や音楽が伝送されるため、暗号化が施されていること自体を秘密にすることができる。この世代では、電子透かしの耐性やコンテンツの品質はあまり要求されず、できるだけ多くの情報を埋め込むことが重要視された。

第2世代は、著作権保護が主目的になり、コンテンツのデジタル化およびネットによるコンテンツ配信が推進されるに伴い、種々のデータ圧縮技術の開発がさかんになった。そこで、暗号化を主体とした著作権保護技術(DRM Digital Right Management)の開発も進められた。しかし、暗号化技術には暗号解除されたコンテンツからアナログ経由で無尽蔵にコピーされてしまうという弱点がある(アナログホールと呼ばれる)。例えば、著作権保護されている画像が表示されているディスプレイ画面をカメラで撮影したり、音楽再生されているスピーカをマイクで拾ったりして違法コピー品を容易に作成できる。これらの違法コピー品を追跡できる手段として電子透かし技術が注目された。そのため、電子透かし技術においては、種々のデータ圧縮に対する耐性とともアナログ変換への耐性が求められた。一方、埋め込み情

報量はそれほど必要とされないが、コンテンツの品質はできる限り劣化のないことが求められた。

第2世代に電子透かし技術は高度に進展したが、著作権保護技術としては脇役でしかなく、これを事業の基幹技術として展開するには弱い面があった。そこで、携帯電話の急激な普及に伴って新たな電子透かしの活用が試行されたのが第3世代である。きっかけは、カメラ付携帯電話とQRコードの発明である。携帯電話は電話機の延長でテンキー入力が主体であるため、URLなど長い文字列を入力するのが不便である。そこで、印刷物や画面の片隅にリンク先のURLを記載したQRコードを印字し、携帯電話のカメラで撮影することにより、携帯電話をリンク先にネット接続できるようにする技術が開発された。しかし、QRコードは目立つ割にデザイン的に美しいものではなく、音楽に付加することが難しいという問題があった。そこで、電子透かし技術により画像や音楽にQRコードに相当するURL情報を埋め込み、携帯電話のカメラで撮影したり、音楽を部分的に録音することにより、携帯電話をリンク先にネット接続できるようにする技術開発が行われた³⁾⁻⁹⁾。携帯電話で取得される画像や音響信号は高度に圧縮されている上に歪みが大きく、第2世代より高度なロバスト性が求められた。

4. 1. 日本国内におけるデジタルコンテンツの違法コピーの現状

文献 16) によると、DVD/BD など映像パッケージメディアについては、リッピングのハードルが高く、海賊版の被害は殆ど無いが、映画スクリーン盗撮による違法コピー動画の問題は深刻である。日本国際映画著作権協会によると、日本国内では、商用の海賊版は殆ど作成されないが、違法録画ファイルがボランティア的に作成されネットに流されることが問題になっている。2007 年より施行された「映画の盗撮の防止に関する法律」により、スクリーン盗撮行為に対する罰則が強化されたため、日本国内では映像の盗撮は行われなくなった。しかし、海外では相変わらずであり、盗撮洋画がネットを通じて日本に上陸する。そこで、海外から移入された盗撮洋画の日本語吹き替え版が上映されている国内の映画館で音声だけ隠し録りされ、海外盗撮映像のサウンドトラックを盗音と差し替えてネットにアップされる事例が目立ってきた。商用でないため、ボイスレコーダやスマホでモノラル収録された音声品質で運用されている点が犯行の阻止をより困難にしている¹⁶⁾。

一方、音楽 CD メディアについては現状では著作権保護対策が解除されているため、(一社)日本レコード協会によると、特にレンタル CD からの違法音源がサイトに挙がり深刻な被害を受けている。また、パチンコ台などアミューズメント機器で流れる新曲がサイトにアップされる被害も深刻である。以上、日本国内では、映像に対する違法録画は法整備により阻止できたが、音響に対する違法録音に対しては刃が立たない。また、ボランティア的に行われ商用レベルの品質が問われないという特性から、違法コピー対策をより困難なものにしている¹⁶⁾。

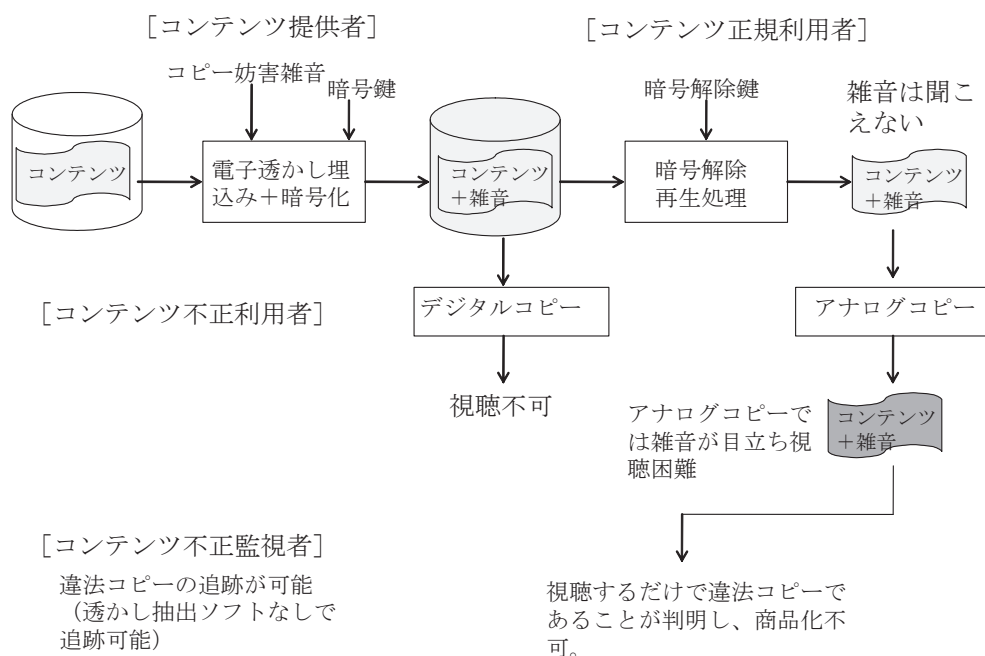


図10 ノイズハイディング手法を導入した新規著作権保護システムの構成

4. 2. 著作権保護技術としての情報ハイディング手法の限界と新規手法の提案

第3世代までに提案されてきた各種コンテンツへの電子透かし埋め込み方式は、いずれも付加情報をコンテンツに不知覚な形態で埋め込み、アナログ系を介したり非可逆データ圧縮を施すような改変を加えたコンテンツからも付加情報を抽出できるようにしている。これらの方式では、違法コピーを行ったという証拠を追跡することは可能であるが、違法コピー行為自体を阻止できないという難点があった。また、前節の日本国内のように超低ビットレートで違法コピーが行われると、ロバスト性が高い電子透かしが求められ、保護対象のコンテンツ品質をある程度犠牲にしないと、電子透かし抽出そのものが困難になっている。更に、電子透かしの抽出においては、埋め込み側のソフトウェアと対になった専用の抽出ソフトウェアを必要とし、現状では電子透かし技術は MPEG などのデータ圧縮技術のように標準化されていないため、多様な抽出ソフトウェアを入手しなくてはならず、万能な電子透かし追跡・抽出システムを構築することが困難である。

一方、金券・証書印刷分野では写真複写機に対する各種偽造防止技術が提案され実用化されている。例えば、「禁複写」などの牽制メッセージを浮かび上らせる技術が実用化されている¹⁷⁾。これと同様な手法を動画に対して適用した事例として文献¹⁸⁾があり、縮小操作を施すとエイリアシングにより警告メッセージが表示される。しかし、解像度変換を伴わない非可逆圧縮やアナログ的なコピーを施せば容易に回避できる。また、音楽に対しても、情報の代わりに、牽制メッセージに相当する副音声データやコピー妨害雑音をハイディングさせ、非可逆圧縮やアナログ的なコピーなどコンテンツになんらかの改変を加えると、埋め込まれ

ている副音声データやコピー妨害雑音が可聴化される手法が提案されている¹⁴⁾⁻¹⁶⁾。

図10は妨害雑音を埋め込むノイズハイディング手法を導入した著作権保護システムの構成である。上左端の与えられたコンテンツに妨害雑音を埋め込み、更に暗号化を施した上でユーザに配信したものが上中央である。これに対して、ユーザがデジタル的にコピーをとっても暗号鍵を入手しない限り視聴できない。そして、ユーザが正規ライセンスを購入し暗号鍵を入手すると上右端のようになりコンテンツは視聴可能になるが不可視・不可聴の妨害雑音が埋め込まれている。これに対してデータ圧縮を施したりアナログ的にコピーをとるなど改変を加えると、妨害雑音が可視化・可聴化される。専用の電子透かし抽出ソフトウェアがなくても、違法コピー品であることを特定でき、違法コピーの商品化や流通を阻止することができる。

4. 3. 音響信号に副音響信号を埋め込む方法

文献 18) では動画像にコピー牽制を促す別の副画像を埋め込み、ユーザが画像に対してサムネールを作成するなど縮小操作を施すエイリアシングが発生し、埋め込まれている副画像が可視化される手法を提案している。本節ではこれと同様な手法を音響信号に対して適用した事例¹⁵⁾を紹介する。エイリアシングとは空間周波数や時間軸の周波数が無限大の成分をもつアナログ映像や音響信号

に対して、限定された周波数 F_s でサンプリングを行いデジタル変換を行おうとすると発生する現象で、デジタル化された信号ではナイキスト周波数 $F_s/2$ 以上の周波数成分は収納できないため、 $F_s/2$ 以下の低域側に折り返され、 $F_s/2$ 以下の周波数成分に重畳される。

図11は、音響信号（音でない波形信号も含む）に対してエイリアシング現象をスペクトルで説明したもので、横軸は周波数で縦軸は左側のアナログ系は信号強度で右側のデジタル系は複素強度（スカラー化した値）である。サンプリング周波数 F_s でサンプリング

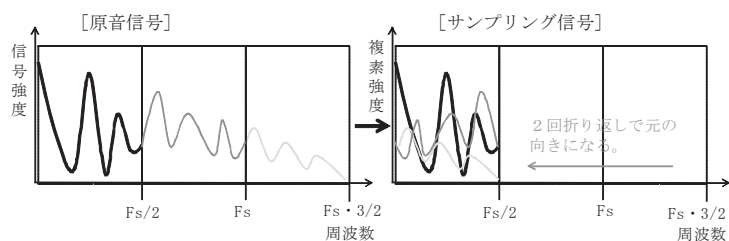


図11 波形信号のデジタルサンプリングによるエイリアシング現象

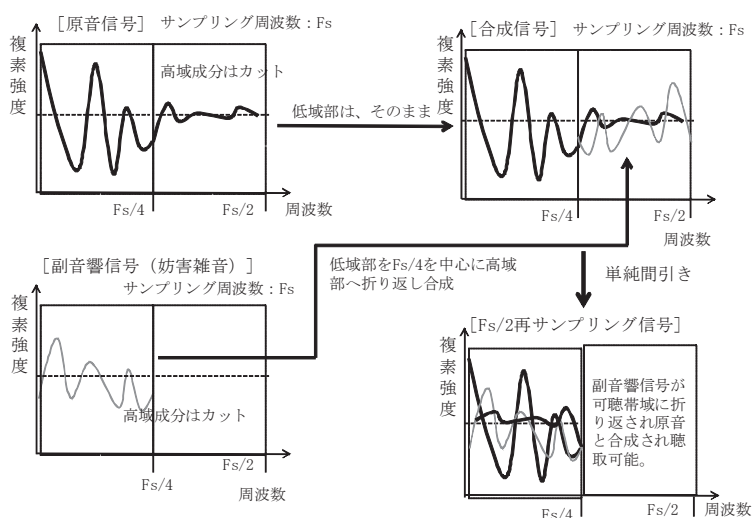


図12 エイリアシングを用いた副音響信号の埋め込み方法

を行うと、 $F_s/2$ を超える成分が $F_s/2$ を中心に折り返され、 $F_s/2$ 以下の信号成分に重畳される。即ち、ナイキスト周波数 $F_s/2$ を超える高周波の信号成分が、 $F_s/2$ 以下の低域信号成分として誤解析され重畳されてしまう。（対策として、サンプリングを行う前にハイカットフィルタ処理を加え、 $F_s/2$ 以上の高域成分を削除するフィルタ処理が一般に行われる。）尚、同図右では、3種のスペクトル成分を重ねて表示しているが、実際には複素ベクトルで加算される。

この現象を応用して音響信号に副音響信号を不可聴な高域部に埋め込み、再サンプリングを行うとエイリアシングが発生して高域部に埋め込まれている信号成分が可聴な低域部に降りてきて聴取できるようにすることができる。この具体的構成例を図12に示す。同図はすべてスペクトルで、左側の2つは埋め込み対象の原音響信号と副音響信号で、各々事前にサンプリング周波数 F_s でサンプリングされ、 $F_s/2$ 以上の信号成分はない。ここで、副音響信号に対して、 $F_s/4 \sim F_s/2$ の信号成分を削除し、 $F_s/4$ 以下の低域の信号成分を $F_s/4$ を対称軸に高域へ折り返し、原音信号の $F_s/4 \sim F_s/2$ の信号成分に重畳したものが同図右上である。原音信号には $F_s/4 \sim F_s/2$ の信号成分が残っているため、右半分では2種のスペクトル成分を重ねて表示しているが、実際には複素ベクトルで加算される。この合成信号を再生すると、 $F_s/4 \sim F_s/2$ の高域にノイズが重畳された状態になるが、副音響信号の振幅を事前に弱くしておけば聴取困難になる。この右上の合成信号に対して、サンプリング周波数 $F_s/2$ で再サンプリングを行うと、同図右下のように、 $F_s/4 \sim F_s/2$ の信号成分が $F_s/4$ を対称軸に低域へ折り返され、合成信号の $F_s/4$ 以下の低域信号成分に重畳される。この再サンプリングされた合成信号を再生すると、原音響信号と副音響信号との合成音が聴取される。

本提案方式では、図12右上のサンプリング周波数 F_s の合成信号に対して、サンプリング周波数 $F_s/2$ で再サンプリングする場合、デジタル状態で1サンプル置きに間引くなど、デジタル信号処理を加えない限り、図12右下のような折り返し効果を得ることは実際には難しい。それはサンプラーの前処理部において、再サンプリング時のナイキスト周波数 $F_s/4$ より高域成分をカットする低域通過フィルタが入っており、図12右上の $F_s/4$ 以上の成分が減衰し、エイリアシング自体は図12右下のように起きても、低域への折り返し成分は原音成分によりマスキングされ聴取不能になってしまうからである。また、デジタル的な再サンプリング処理においても線形補間が組み込まれていれば、低域通過フィルタと同等な効果が生じ、低域への折り返し成分は原音成分によるマスキングレベルの範囲内まで減衰してしまう。

4. 4. 音響信号にコピー妨害雑音を埋め込む方法

前節で提案した方式で妨害雑音を埋め込むことができるが、違法コピー操作により所望の再サンプリング処理を実行させ、所望の折り返し効果を起こさせることは難しい。本節では図13に示されるような信号処理を加えることにより、違法コピーにより帯域幅が狭くなることにより不快なコピー妨害雑音を再生させる妨害雑音の埋め込む手法を提案する。

図13(2)は埋め込み対象の音響コンテンツの帯域幅を示す。第3章で述べた電子透かし埋め込み手法では、携帯電話による抽出に対応できるように、図13(2)の中央部の「既提案の電子透かし埋め込み領域M」という聴覚感度が比較的高い領域に電子透かしの埋め込み

でいた。ここより低域または高域のマイクロフォンで集音可能または音声・音楽非可逆圧縮で符号化対象としている下限周波数または上限周波数近辺に、コピー妨害雑音を重畳する。この際、通常再生時には重畳されたコピー妨害雑音が聴取されないようにすることが技術的なポイントになる。

本節で提案するコピー妨害雑音の埋め込む方法は次の通りである。所定の帯域内で逆回し再生あるいは時間軸方向にスクランブルを加えることにより、コピー妨害雑音を埋め込む。スクランブルを加えた帯域幅が上下の帯域に比べ狭く、上下に十分な信号成分が存在すれば、聴覚マスキングにより逆回し成分が順方向に揃いコピー妨害雑音は聴取されない。しかし、違法コピーによりマイクロフォンによる再録音または非可逆データ圧縮等が施され、上下のいずれかの帯域が欠損すると、聴覚マスキングが働かなくなり、逆回し成分がそのまま不快な成分として聴取されるようになる。

図13 (2) 左側の低音帯域および右側の高音帯域に、埋め込み妨害信号Lおよび妨害信号Hとして、図14に示されるような狭帯域のスクランブル処理を加える。このとき、原音信号L、原音信号Mおよび原音信号Hのレベルが、埋め込み妨害信号Lおよび妨害信号Hと同等以上に大きければ、原音信号L、原音信号Mおよび原音信号Hをマスキャーとしてマスキャーである埋め込み妨害信号Lおよび妨害信号Hに対して聴覚マスキングが働く。即ち、図13上段に示されるように、ヒト聴覚系には埋め込み妨害信号Lおよび妨害信号Hは知覚されない。しかし、図13 (3) に示されるように、アナログ的なコピーの過程で原音信号Lの帯域または原音信号Hの帯域のいずれか、または双方が減衰してマスキャー信号が欠落すると、先の聴覚マスキングは働かなくなり、埋め込み妨害信号Lまたは妨害信号Hのいずれか、または双方が知覚されるようになる。

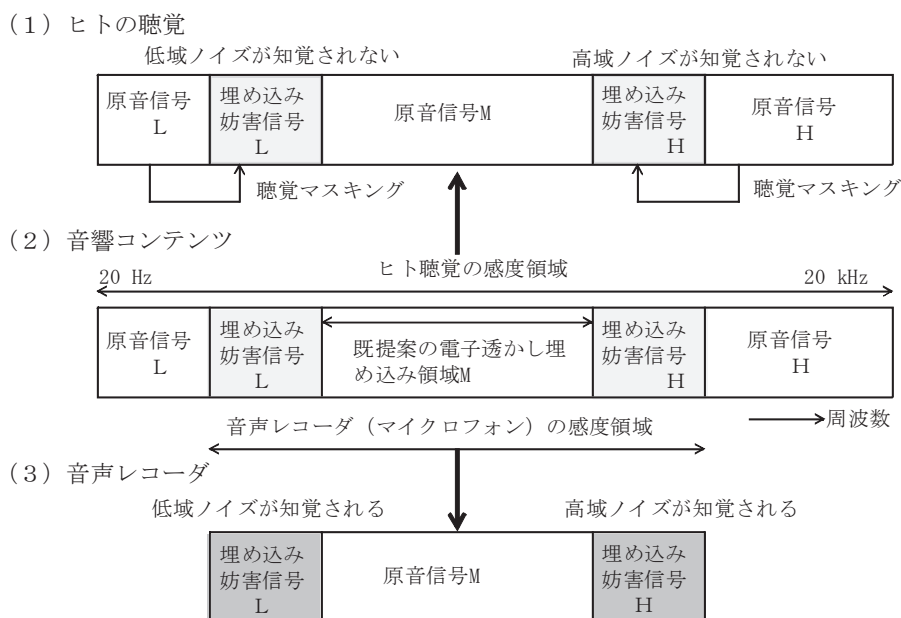
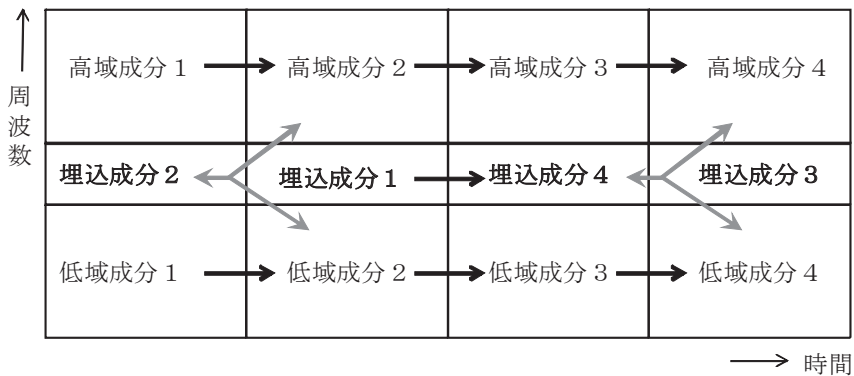


図13 帯域幅が狭くなることにより再生されるコピー妨害雑音の埋め込み方法

(1) 狭帯域信号成分のスクランブル処理 (妨害雑音埋め込み)



(2) 録音により高域側の信号成分が減衰した場合

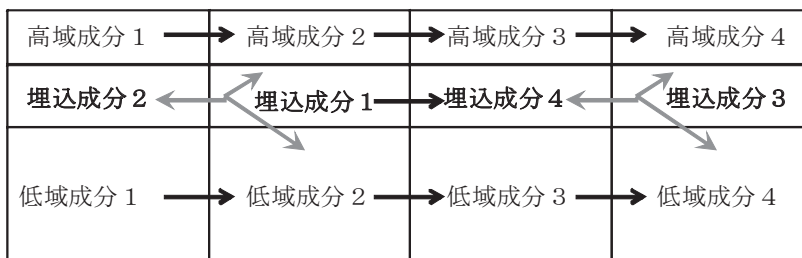


図14 狭帯域スクランブル処理によるコピー妨害雑音の作用

図14は狭帯域スクランブル処理によるコピー妨害雑音の具体的作用例を示す。図14において横軸は時間で所定の間隔(例、50msec)でフレーム分割されており、縦軸は周波数で、妨害雑音を埋め込む帯域をはさんで、上下に原音信号の帯域が(高域成分および低域成分)存在する。図14の埋込成分は図13で妨害信号Lまたは妨害信号Hのいずれかの埋め込み帯域の信号成分である。図14の埋込成分が妨害信号Lの場合は、図14の高域成分は図13の原音信号Mに対応し、図14の低域成分は図13の原音信号Lに対応し、図14の埋込成分が妨害信号Hの場合は、図14の高域成分は図13の原音信号Hに対応し、図14の低域成分は図13の原音信号Mに対応する。

図14(1)は、妨害雑音が埋め込まれた状態を示し、図中の埋込成分1と埋込成分2および埋込成分3と埋込成分4の時間軸順番を逆転させる処理を行っている。一方、高域成分と低域成分の時間軸順番は原音通りであるため、聴覚マスキング作用により、逆転された埋込成分2と埋込成分1および埋込成分4と埋込成分3は、図中の矢印で示すように原音の順番に揃わせるような作用が聴覚的に働き、不快な雑音は聴取されない。図14(2)は図14(1)の信号に対して、高域成分の帯域が狭くなり高域の信号成分が弱くなった状態を示し、図中の矢印で示すように原音の順番に揃わせるような作用の高域側が働かなくなり、不快な雑音が聴取される。図示していないが、図14(1)の信号に対して、低域成分の帯域が狭くなり低域の信号成分が弱くなった状態でも、同様に図中の矢印で示すように原音の順番に揃わせるような作用の低域側が働かなくなり、不快な雑音が聴取される。

5. おわりに

本稿では、音楽に対する情報ハイディング技術に焦点を当て、その基本的な手法からロバストな埋め込み手法までの一連について述べ、加えて、音楽著作権を保護するために副音響信号や妨害雑音を埋め込むなどの新規なハイディング技術開発事例について紹介した。

音楽作品に耳に聞こえない情報を埋め込む音楽電子透かし技術「ゲンコード Mark」については、技術的には完成に近づいており、埋め込み容量の拡大、室内残響への対策と受信可能な距離拡大などの課題は残されているものの、これらはアプリケーションが特定されないと技術目標も定まらない。最大の問題は、著作権管理などの用途以外に産業上有望なキラーアプリケーションが開拓されていないことにあると考えられる。本開発に着手した頃は、RFIDなどワイヤレス技術が発展途上で、電波より音波を媒介とした情報伝達の方が低コストに行えるという優位性があったが、近年はそれが揺らいできており、電波の使用制限がある場面など適用分野の探索が益々困難になってきている。

一方、著作権管理という面でも、情報ハイディング手法だけでは不十分で、本稿で紹介した副音響信号や妨害雑音を埋め込むなどの新規なハイディング技術に関する研究も今後更に推進する必要がある。映画業界もこのノイズハイディング技術に注目しているが、違法コピー時に再生される妨害雑音の業界からの要件として、音声が不明瞭になるレベルに妨害雑音の音量が大きく、長時間の視聴に耐えない不快感を与えることが要望されている。残念ながら、現状の研究レベルでは業界が要望するレベルには程遠い。また、使用が想定されるレコーダの周波数特性のバリエーションは非常に大きいため、録音に使用されるマイクの帯域が広いとマイク録音後の音を再生しても妨害雑音が聴取されないとともに、録音に使用されるマイクの帯域が極端に狭くてもマイク録音後の音を再生しても妨害雑音が聴取されなくなる。そこで、埋め込み周波数帯を低域側と高域側の2箇所ではなく、各々2箇所以上設定することなどを含めて今後更なる研究を重ねたい。

以上、本稿で紹介した種々のハイディング手法については、全て Windows で稼働するC言語ソフトウェアで実装し、筆者が担当している本学・情報表現学科の「マルチフィールド体験演習」等の授業で既に活用しており、随時拡充や改良開発も進めている。本学の学生や教職員の皆様とのインタラクションが本研究やソフトウェア開発の推進に多大な影響を与えたものと考えており、改めて本学関係者の皆様に謝意を示す。本稿で紹介した手法を実装した Windows 版ソフトウェアにつき、教育・研究・その他にご活用されたい場合はソースコードを含め提供可能ですので、ご連絡ください。

引用文献

- 1) 茂出木敏雄, “聴覚芸術への情報学的アプローチと音楽情報処理ツールの開発事例,” 尚美学園大学芸術情報研究, Vol.18, pp.15-35, November 2010.
- 2) 茂出木敏雄, 「映像コンテンツ解析によるBGMサウンドトラックの自動生成」, 『電気学会 電子情報システム部門誌』, Vol.125-C, No.7, オーム社, pp.1004-1010, 2005-7.
- 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, オーム社, pp.825-831, 2006-7.

- 4) 茂出木敏雄、「電子透かし技術の概要と今後の展望 -- 携帯電話社会に到来した第3の波」、『マテリアルステージ』, No.7(2)、技術情報協会、pp.70-76、2007-5.
- 5) 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、オーム社、pp.1013-1021、2007-7.
- 6) Modegi, Toshio, '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, Wiley, 2007-11.
- 7) 茂出木敏雄、「音脈分凝を活用した携帯電話で抽出可能な音楽電子透かし技術ーゲンコード Markー」、『電気学会 電子情報システム部門誌』, Vol.128-C, No.7、オーム社、pp.1087-1095、2008-7.
- 8) Modegi, Toshio, 'Increasing the Audio Watermark Data Rate in the Construction of Ubiquitous Acoustic Spaces' , Electrical Engineering in Japan, Vol.165, No.1, pp.42-51, Wiley InterScience , 2008-10.
- 9) Modegi, Toshio, 'Audio Watermark Embedding Technique Applying Auditory Stream Segregation: G-encoder Mark, Extractable by Mobile Phone' , Electronics and Communications in Japan, Vol.92, No.12, pp.1-12, Wiley InterScience , 2009-11.
- 10) Modegi, Toshio, 'Spatial and Temporal Position Information Delivery to Mobile Terminals Using Audio Watermark Techniques' , Proceedings of Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IHHMSP2009, pp.499-502, 2009-9.
- 11) Modegi, Toshio, 'Chapter 9: Spatial and Temporal Position Information Delivery to Mobile Terminals Using Audio Watermark Techniques' , "Multimedia Information Hiding Technologies and Methodologies for Controlling Data (edited by Kazuhiro Kondo)" , pp.182-207, IGI Global, 2012-10.
- 12) 茂出木敏雄, 橋本聡, 片山貴信, 庄司藤男、「スピーチプライバシー保護機能をもたせた電子ポスターーポスラサウンドパネル for スピーチプライバシーー」、情報処理学会インタラクシオン 2011 予稿集、ISCL-22, 2011-3.
- 13) 茂出木敏雄, 「第15節：電子ポスターのサウンドデザイン」、『製品音の快音技術～感性にアピールする製品の音作り (岩宮眞一郎・高田正幸監修)』, S & T 出版、pp.423-443, 2012-7.
- 14) 茂出木敏雄:「コピー妨害雑音のハイディング機能をもたせた音楽電子透かし技術 “ゲンコード Mark for COPY PROTECT”」, 電子情報通信学会・第4回マルチメディア情報ハイディング・エンリッチメント研究会資料, IEICE-EMM2011-42, pp.13-18, 2011-11.
- 15) 茂出木敏雄:「副音響信号のハイディング機能をもたせた音楽電子透かし技術 “ゲンコード Mark for COPY PROTECT”」, 電子情報通信学会・第6回マルチメディア情報ハイディング・エンリッチメント研究会資料, Vol.111, No.496, pp.7-12, 2012-3.
- 16) 茂出木敏雄, 宮崎慎一, 関口剛, 小須田祐二:「狭帯域成分の時間軸スクランブルによるコピー妨害雑音のハイディング技術 “ゲンコード Mark for COPY PROTECT”」, 電子情

報通信学会・第1回マルチメディア情報ハイディング・エンリッチメント研究会資料,
Vol.112, No.129, pp.43-49, 2012-7.

17) カシオ計算機ホームページ「コピー牽制地紋」: <http://casio.jp/ppr/software/copyguard/>.

18) 日立製作所ニュースリリース「不正コピーをけん制する動画電子透かし技術を開発」:
<http://www.hitachi.co.jp/New/cnews/month/2011/10/1020a.html>.