

携帯電話を用いた音響情報の非接触 I D 検出技術の開発

茂出木 敏雄 (e-mail: Modegi-T@mail.dnp.co.jp)[†]

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[†]

1. はじめに

昨年大会において、ステレオ音楽に対して低音部の定位を移動させることにより、殆どロスレスに電子透かしを埋め込むことができる手法を提案した[1]。その後、ステレオの片側チャンネルだけで電子透かしの抽出を可能にし、マイクロフォンや携帯電話を用いて非接触に埋め込まれたデータを抽出できる手法を確立した[2]。ただし、先提案手法は、再生環境がステレオである必要があり、スピーカやマイクロフォンに低音部の再生/検出能力が要求される。また、ライブ演奏音や生活環境音など事前に電子透かしの埋め込みが不可能な音源には適用できないという制約があった。

そこで本発表では、フィンガープリント法による音響情報の I D 抽出技術を先提案に付加した統合方式を提案する。この追加手法は、指紋照合と同様に、与えられた音響信号の特徴パターンをあらかじめ構築しておいたデータベースと照合することにより、音響信号の I D を検索する方法である。ソース音響信号はモノラルでも良く、信号自体に手を加える必要が無く、スピーカやマイクロフォンの周波数特性を問わず運用可能である。これにより、先の電子透かし埋め込み手法が適用できない用途を補完することができる。本デモでは、電子透かしが埋め込まれている音楽と一般の自然音をスピーカで再生しながら、各音響情報の I D を非接触に抽出するデモを行なう。

2. 既提案の音楽への電子透かし埋め込み技術

音楽電子透かしの主流は、心理聴覚分析、MDCT 係数操作、スペクトル拡散法などを併用する方式であるが[3]、原品質の保持とデータ圧縮やアナログ変換など信号加工に対する耐性力とは、トレードオフの関係にあった。

したがって、携帯電話で読み取れるようにするためには、相当強力な耐性のある透かしを埋め込む必要があり、実用上の品質を維持することは難しかった。

そこで筆者らは、図 1 に示されるように、ステレオの左右チャンネル間で低音部の成分を移動させることにより、ほぼロスレスにデータを埋め込める方式を提案した[2]。ヒトの聴覚は低音部の定位感覚が鈍いという特徴があるため、約 200Hz 以下にこのような操作を施せば、定位の移動は認識されない。従って、顕著に定位を動かすことができ、埋め込まれたデータはアナログ音響空間に放射されても消失することは無く、強力なアナログ耐性をもたせることが可能になる。

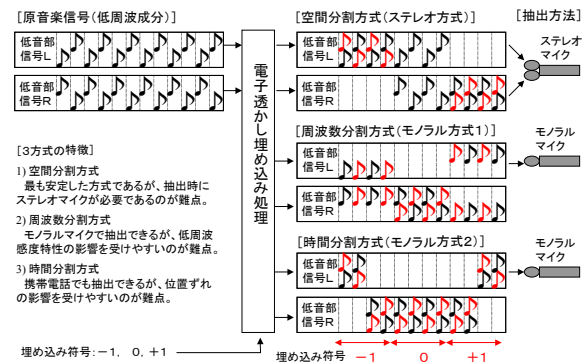


図 1 既提案の電子透かし埋め込み方式

図 1 に示されるように、これまで 3 方式を提案してきたが、下 2 つの周波数分割方式および時間分割方式は、左チャンネルの信号だけをモノラルマイクで拾えば透かしを抽出できるようにした。この場合、右チャンネルは左チャンネル側で失われている信号成分を補完する役目をするため、原素材がモノラルの場合でも、品質維持のため右チャンネルを必要とする。3 番目の時間分割方式であれば、大抵の携帯電話で抽出できることを確認している。

Development of Contactless Acoustic Information Identification Technique Using Cell Phone

[†]Toshio Modegi, Media Technology Research Center, Dai Nippon Printing Co., Ltd. (Modegi-T@mail.dnp.co.jp)

3. 提案する音響フィンガープリント方式

図2は本稿で提案する音響フィンガープリント方式によるID抽出処理の概要である。入力音響信号は特徴抽出により4～8万分の1に圧縮され、32ビットワードの時系列パターンに変換される(図では2ワード)。音響特徴データベースに登録されている各音響特徴時系列ワードと照合し、音響素材IDごとに最小となる不一致ビット数を算出する。この値が最小となる音響素材IDを求める。

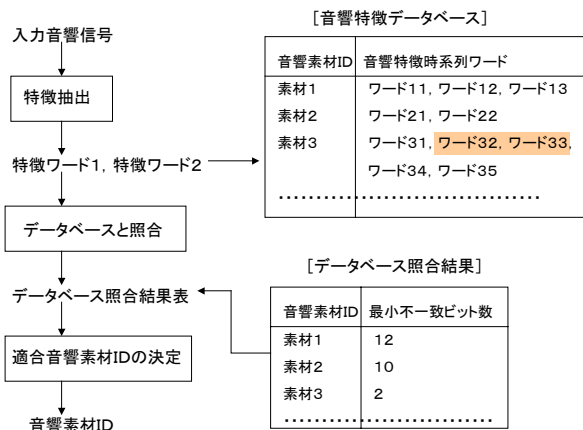


図2 フィンガープリント方式の概要

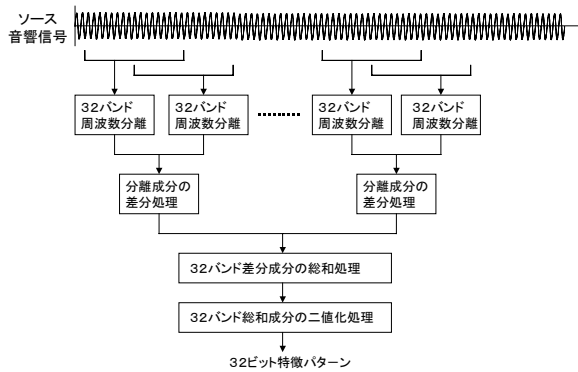


図3 音響信号の特徴抽出概要

図3は特徴抽出処理の具体例で、ソース音響信号に対して、32バンドの周波数成分を算出し、32ビットの特徴ワードに変換する。ソースが44.1kHzの場合、フレーム幅を4096に設定し、ハニング窓をかけて半フレームのピッチでフーリエ変換を行う。携帯電話での入力音声に対応するため4kHz以下の周波数成分のみを使用し、32バンドに分離する。周波数成分比率を強調するため、直前フレームとの差分をとり、更に位置ずれの影響を抑えるため、10～20フレーム分の総和をとり、2値化して32ビット

ワードに変換する。2値化にあたっては、レベル変動の影響を抑えるため、0と1のビット数が同一になるようにしきい値を設定する。

4. 先提案方式との統合化

前節で述べたフィンガープリント法は、第2節で述べた方式と比較して、原素材に対して全く変更を加えないこと、対象音響素材は広くモノラルでも適用できる点で優位である。しかし、32ビットワード列で識別できる音楽素材には限界があり、例えば、同一楽曲に対して、異なるアーティストが演奏した音楽素材を識別することは難しい。

一方、第2節で述べた電子透かし埋め込み方式では、全く同じ素材でも埋め込むデータを変えれば識別でき、楽曲の進行に合わせて時系列に埋め込みデータを変化させることができるという利点がある。しかしライブ演奏の場合、PAを伴えばリアルタイム電子透かし埋め込みにより対応できるが、自然音など生音楽素材に対しては基本的に無力である。

そこで、第2・第3節で述べた両方式を併用する手法を提案する。ID情報を抽出する場合は、電子透かしの抽出を先に行ない、所定のフォーマットでデータが記録されていない場合は、フィンガープリント法を適用すれば良い。これにより、電子透かしの埋め込みが行なえないライブ演奏音や生活環境音などあらゆる音に対して、ID検出を実現することができる。

5. おわりに

今後は抽出ソフトウェアを携帯電話に実装し、新規な広告媒体として展開を図る予定である。フィンガープリント法については、携帯電話の帯域制約での音響解析を前提として素材識別能を評価するため、比較的大規模なデータベースを構築して実験を行なう予定である。

[参考文献]

- [1] 茂出木 敏雄：「オーディオのアナログ耐性電子透かし埋め込み技術の開発」、情報処理学会第67回全国大会論文集、Vol. 3, pp. 271-272, (2005-3).
- [2] 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」、情報処理学会・コンピュータセキュリティシンポジウム 2005, pp. 547-552, (2005-10).
- [3] 小松尚久、田中賢一監修、画像電子学会編、「電子透かし技術 デジタルコンテンツのセキュリティ」、東京電機大学出版局、2004.