

# 音響空間のユビキタス化に向けた非接触音源情報獲得技術の開発

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

モノとサイバー空間との連携を可能にする技術として、ICタグやQRコードをモノに付加して、モノのID情報を非接触に読み取る手法が既に実用化されている。一方、印刷物や電子ディスプレイなどに表示される画像コンテンツに対しても、QRコードを付加したり、画像に電子透かしを埋め込むことにより、カメラ付き携帯電話でコンテンツのID情報を非接触に読み取る手法が既に確立されている[2]。また、画像の特徴パターンをあらかじめデータベース化して、類似検索を行なうことにより、画像コンテンツ自身に全く手を加えずにID情報を抽出する手法も利用されている。

これに対し、音響空間とサイバー空間との連携を可能にする技術として、音声認識技術が代表的であるが、対象が音声に限定される。音楽については、前述の類似検索技法を適用し、流れているメロディを携帯電話で録音し、音響特徴パターンをデータベースと照合して、対象楽曲IDを抽出するフィンガープリント技術は実用化されている。ただし、対象音響コンテンツがデータベース登録された楽曲に限定され、リアルタイムな抽出が行えないという問題があった。

そこで、本発表では、筆者らが開発を進めてきた非接触抽出可能な電子透かし手法[3]と新規な音響特徴抽出手法を組み合わせた統合的な音源情報獲得技術を提案する。本デモでは、電子透かしが埋め込まれている音楽素材と未加工の一般自然音をスピーカで再生しながら、各音源のID情報をマイクロフォンや携帯電話で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、心理聴覚分析法、MDCT係数操作法、スペクトル拡散法などを併用する方式であり[1]、品質の劣化が比較的少なく、データ圧縮やアナログ変換などへの耐性がそこそ

こあるが、品質と耐性はトレードオフの関係である。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、従来手法で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[3]。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。即ち、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正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。

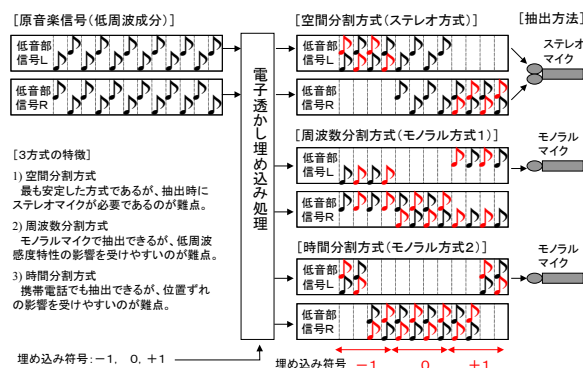


図1 提案した音楽電子透かし埋め込み方式

### 3. 提案する音響素材の特徴抽出技術

図2は本稿で提案する音響特徴データベース検索方式による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。

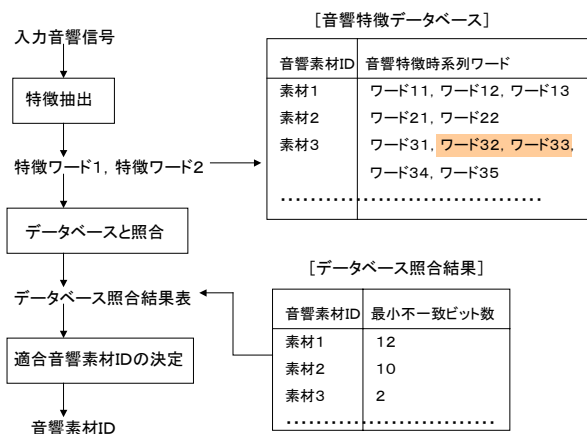


図2 音響特徴データベース検索方式の概要

特徴抽出は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. 提案する音源情報獲得技術

前節の音響特徴データベース検索法は、第2節の方式と比較して、原素材に対して全く改変を加えないこと、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ビット特徴ワード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節の電子透かし埋込み方式では、全く同じ素材でも埋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、図3に示すように第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ければ、音響特徴データベース検索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。

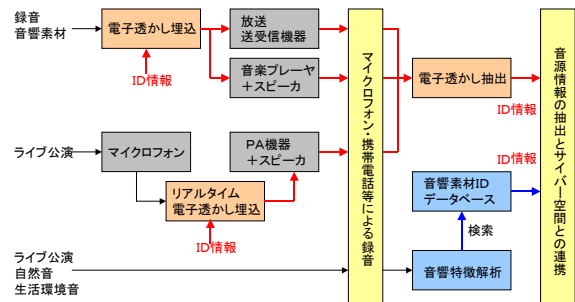


図3 提案する音源情報獲得技術

### 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。音響特徴データベース検索法については、携帯電話での音響解析を前提として各種DBを試作し、素材識別能を評価する予定である。

#### [参考文献]

[1] 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。  
 [2] 中村高雄、片山淳、山室雅司、曾根原登：「カメラ付携帯電話機を用いたアナログ画像からの高速電子透かし検出方式」、電子情報通信学会誌, Vol.J87-D2, No.12, pp.2145-2155, (2004-12)。  
 [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」、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)。