

2 値多重音響特徴ベクトルを用いた類似音楽探索法の頑健性評価

永野 秀尚[†] 柏野 邦夫[†] 藤原 融[‡]

[†] 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所

[‡] 大阪大学 大学院情報科学研究科

{nagano, kunio}@eye.brl.ntt.co.jp, fujiwawra@ist.osaka-u.ac.jp

あらまし 我々は、多重奏音響信号をクエリーとして多重奏音響信号から類似する音楽を探索する類似音楽探索法を提案している。本提案手法は、テンポの異なる演奏や移調した演奏などのような変動があっても、曲自体が同一であれば探索できることを目的としている。今回、本手法の頑健性を、原曲または再演奏のみの変動、テンポ変換、移調、楽器変換の各変動ごとに実験評価した。216曲の実験用音楽データベースを用いた探索実験では、クエリーの長さが平均19秒のとき、テンポ変換、移調の変動に対しては90%以上の検出率であった。また、楽器変換についても77%と高い検出率が得られていることがわかった。

Robustness Evaluation of a Similar Music Retrieval Method Using Polyphonic Binary Feature Vectors

Hidehisa Nagano[†] Kunio Kashino[†] Toru Fujiwara[‡]

[†] NTT Communication Science Laboratories, NTT Corporation

[‡] Graduate School of Information Science and Technology, Osaka University

{nagano, kunio}@eye.brl.ntt.co.jp, fujiwawra@ist.osaka-u.ac.jp

ABSTRACT We propose a method for retrieving similar music from a polyphonic-music audio database using a polyphonic audio signal as a query. Our method is also aiming at retrieving *similar music* such as that played at different tempo, played with different instruments, and so on. In this paper, we evaluate the robustness of our method against these changes. Experiments using a test database containing 216 music pieces show that the search accuracy is quite good. For example, when the average length of the queries is 19 s, the detection rate is more than 90% even when the tempo or pitch differs. It is also shown that the detection rate is about 77% even when the instruments were changed.

1 まえがき

音や映像のメディア情報の増加、多様化により様々なメディア情報の探索技術が必要とされている。我々は特に、メディア探索、すなわち、長時間の音や映像の信号(蓄積信号)またはそのデータベース(DB)と、探したい音や映像の信号(参照信号)がクエリーとして与えられたとき、蓄積信号中の参照信号に類似する区間を探し出す探索技術について研究を進めてきた [1, 2, 3, 4, 5]。特に文献 [4] においては、音楽演奏の音響信号の探索において、多重奏音響信号を参照信号とし、同じく多重奏音響信号の中から、参照信号に類似する楽曲を見つけ出す手法とその高速化手法を提案している。このような多重奏音響信号による類似音楽探索は、

例えば、通常多重奏音響信号である音楽CDの断片などを参照信号とし、それに類似する音楽を音楽音響信号のDBから探索するために必要な技術である。さらに、この手法においては、参照信号と同一の音楽音響信号だけではなく、類似する音楽、例えば、再演奏されたもの、テンポの異なる演奏、移調されたもの、もしくは他の楽器を用いて演奏されたものなど、同一の音響信号ではないが同じ音楽である類似音楽も探索することを目的としている。

メディア探索の重要な応用例である音楽探索についてはこれまで多くの研究がなされてきたが、それらは、その目的により二つに大別できる。一つは音楽CDなどと信号レベルでほぼ同一の音楽の探索をねらう一致探索であり、この場合、音響信号から得

られるスペクトル特徴およびそのヒストグラムを用いる手法 [1, 3, 5] や音響特徴間の類似度をベクトル量子化により高速に計算する手法 [6] などが提案されている。一方、先に述べたように、信号レベルでは必ずしも一致していないが、演奏者が異なる同一の曲や、主旋律が同じ曲や、アレンジの異なる曲などのように、何らかの意味で類似しているものの探索をねらう類似探索がある [7, 8, 9, 10, 11, 12, 4]。このような類似音楽探索においては、音響特徴が蓄積信号と参照信号で大きく異なるためや、音響信号以外に楽譜などで音楽が表現されるため、音響特徴以外の特徴を用いた探索が必要となる。そこで、Ghiasらは単旋律の音高系列を“S(同じ)”, “U(上がる)”, “D(下がる)”の相対音高の文字列で表す melodic contour を用いた単旋律の照合手法を提案し、ハミングをクエリーとし、類似する単旋律を探索する類似音楽探索においてその有効性を示した [8]。melodic contour は旋律の表現としては楽譜などに比べ単純であるが、単旋律の識別、照合には有効であり、その他の研究においても単旋律の探索における特徴として用いられている。一方、我々はクエリーとDBも多重奏音響信号とし、さらに、再演奏されたもの、テンポの異なる演奏、移調されたもの、もしくは他の楽器を用いて演奏されたものなども類似音楽として探索することを目的とした類似音楽探索法を提案している [4]。

文献 [4] の手法が対象とする類似音楽探索においては、同一の曲であっても、演奏の変化、演奏速度の違い、楽器の変更などにより音響信号が激しく変化するため、スペクトル特徴のような音響信号を用いた探索では探索精度の低下が予測される。また、従来の単旋律の照合を基本とした手法では、音響信号からの旋律の抽出や多重奏と多重奏の間での主旋律の照合およびそのための探索時間の増加などの困難な問題点があった。そこで、文献 [4] ではこのような多重奏音楽の音響信号の照合のための2値多重音響特徴ベクトルと、この特徴を用いた類似音楽探索法とその高速化手法が提案されている。この手法は、Ghiasらの melodic contour と同様に、音楽演奏の時系列を単純な符号の一次元系列で表現し、この符号系列に対し、文字列照合を行うことでロバストかつ高速な類似音楽の探索を狙うものである。2値多重音響特徴ベクトルは、各成分を2値とする特徴ベクトルであり、この各成分は各々対応する音高の音の有無に大まかに対応する。これを、多重奏音楽の音響信号から順次抽出し2値多重音響特徴ベクトルの時系列を得る。そして、蓄積信号と参照信号から得られた2値多重音響特徴ベクトルの系列について、2値多重音響特徴ベクトルの類似度に基づいた文字列照合を

行うことにより、類似音楽探索を行う。文献 [4] では、2値多重音響特徴ベクトルにより、従来のスペクトル特徴を用いた手法に比べ、目的の類似音楽探索を高精度に行うことが可能であることが示されている。そして、本稿では、文献 [4] の実験を詳細に調べ、提案手法が移調や楽器変換などの各変動に対し、どれだけの探索精度が得られているかを評価する。

以下、2では文献 [4] の2値多重音響特徴ベクトルを用いた類似音楽探索手法とその高速化手法の概要について述べる。次に、3では、文献 [4] の実験における探索精度を変動ごとに評価した結果を示す。そして、4でまとめる。

2 2値多重音響特徴ベクトルを用いた類似音楽探索法 [4]

2.1 探索手続きの概要

図1に探索手続きの概要を示す。ここで、蓄積信号 (stored signal) と参照信号 (reference signal) は多重奏音楽の音響信号であり、探索においては、蓄積信号中の参照信号に類似するすべての区間を検出することを目的とする。まず、探索の準備として、各音響信号について周波数分析を行い特徴ベクトルを時系列順に抽出する。そして、抽出された特徴ベクトルを順次符号化し符号系列を得る。ここで蓄積信号と参照信号から得られる符号系列を、各々、蓄積符号系列 (stored feature vector string)、参照符号系列 (reference feature vector string) と呼ぶ。そして、探索においては、蓄積符号系列中の参照符号系列に類似する長さ w のすべての部分系列を探索し、検出する。ここで w は探索窓 (window) の幅として与えられるものとする。そして、参照符号系列と長さ w の部分系列の類似度は符号間の類似度に基づいた DP 照合により求める。これは、演奏速度の違いなどによる時間軸方向の信号伸縮に対応するためである。なお、この DP 照合により得られる類似度を系列間類似度と呼び、照合を行う符号系列の長さ、すなわち参照符号系列の長さ w により正規化されているものとする。以上、

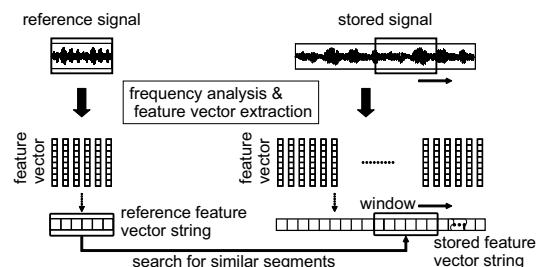


図1: 探索手続きの概要

本探索で行う処理についてその概要を述べたが、本探索は、形式的には、参照符号系列との系列間類似度が探索閾値より大きい蓄積符号系列中の長さ w の部分系列をすべて検出することとする。

2.2 2値多重音響特徴ベクトル

2値多重音響特徴ベクトルは波形信号からパワースペクトルを抽出後、SNAP (Simultaneous-Note-set Alteration Point ; 同時単音集合変化時刻) を区切りとする snapshot 単位で抽出する。SNAP とは多重奏音楽において、単音の立ち上がりにより、同時に鳴っている単音の組合せの変化が起きた時刻である。そして、隣り合う SNAP と SNAP の間の一定の組合せの音が鳴っている区間を snapshot と呼ぶ。そして、この snapshot ごとに一つの2値多重音響特徴ベクトルを抽出し、一つの2値多重音響特徴ベクトルを一つの符号に符号化する。この snapshot 単位の符号化には、分析窓に重なりを持たせながら抽出した多数の短時間スペクトル特徴の系列を順に一つ一つ符号化する場合に比べ、一般に、符号系列長を短くできるというメリットがある。また、snapshot 単位の符号化では、演奏速度の違いなどによる音の長さの伸縮を吸収することも期待できる。

SNAP の抽出後、各 snapshot で、各周波数における snapshot 内でのパワーの最大値を、その snapshot におけるその周波数でのパワーとする。2値多重音響特徴ベクトルの抽出においては、上記で得られた各 snapshot において、倍音除去を行った後、各音 (半音単位の音高で表現) の有無を調べる。そして、得られた音の有無に着目し、各 snapshot について 12 ビットの 2 値多重音響特徴ベクトルを構成する。特徴ベクトルの各ビットは下位から順に A, A#, B, ..., G# の各音に対応しており、この snapshot 内で同じ音名となる音が存在したとき、対応するビットを 1 に、そうでないとき 0 にする。例えば、A (220Hz), E (330Hz), A (440Hz), C (523Hz), A (880Hz) が snapshot 内に存在した場合、この snapshot の特徴ベクトルは 2 進表現で "000010001001" とする。2値多重音響特徴ベクトルは、多重音響信号において同時に発生する音の有無に大まかに対応していることになる。

そして、2値多重音響特徴ベクトル $x = (x_1, x_2, \dots, x_{12})$ と $y = (y_1, y_2, \dots, y_{12})$ の間の類似度は

$$\frac{\sum_{k=1}^{12} u_k}{\sum_{j=1}^{12} (x_j + y_j)} \quad (1)$$

と定義する。ここで u_k は、 $x_k = y_k = 1$ のとき $u_k = 2$ で、そうでないとき $u_k = 0$ と定義される。

なお、 $x = y = 0$ の場合、 x と y の類似度は 1 とする。これは無音の snapshot どうしは同じ和音として取り扱うことを意味している。

2値多重音響特徴ベクトルの符号化は、2値多重音響特徴ベクトルを単に 12 ビットの 2 進数とみなすことで行う。この符号化により、2値多重音響特徴ベクトルを簡潔に符号化できると同時に、探索時に参照符号系列の符号をシフトしてから探索することにより、移調された音楽の探索も可能となる。なお、符号間の類似度 (符号間類似度) も対応する 2 値多重音響特徴ベクトル間の類似度 (式 (1)) により定義するものとする。

2.3 探索の高速化の概要

図 2 に文献 [4] で提案されている探索の高速化の概要を示す。本高速化手法においては、参照符号系列と類似する符号を持つ部分系列とのみ、照合する区間を短くしながら照合することで探索を高速化している。なお、ここで、類似度行列は、行列の (u, v) 成分が符号 u と v の符号間類似度を表す行列であり、探索においては類似度行列を用いて、符号間類似度を参照する。なお、二つの符号が類似するとは、これらの符号間類似度が 0 でないことをいう。そして、提案手法では、符号間類似度をスパースにする (0 となる成分を増やす) ことで、探索をさらに高速化できる。なお、同じ類似度行列を用いた場合、提案手法と全探索の探索結果は同じである。

3 実験

実験では、異なる種類の変動がある類似楽曲を探し出す際に、2値多重音響特徴ベクトルを用いた場合、2値多重音響特徴ベクトルを用い類似度行列をスパース化した場合について、変動の種類に応じて検出性能がどのように変化するかを調べ

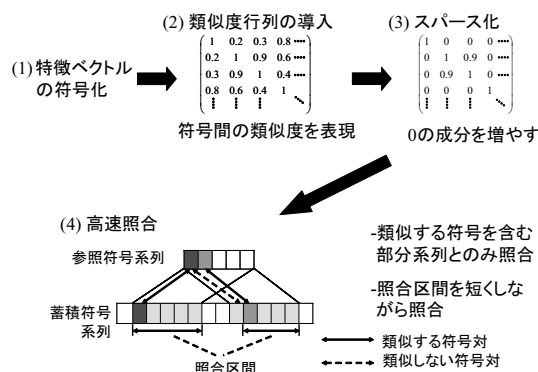


図 2: 探索の高速化の概要

た。また、比較のため、従来のスペクトル特徴を用いた場合についても、同様の実験を行った。

3.1 実験条件

実験は実験用の音楽 DB を用い、この中から実際に類似音楽を探索することで行った。音楽 DB には、ポップス、クラシック、インストルメンタルなどの様々なジャンルの多重奏の楽曲 216 曲の音響信号を用いた。なお、ここでは、これらの 216 曲の録音の一つ一つを楽曲と呼ぶ。この DB の音響信号の長さは合計で約 480 分であった。なお、音響信号は 48kHz サンプリング、16bit 量子化、ステレオ録音のものをモノラル化して用いた。実験用の参照信号については、参照信号の長さにより探索精度が変化するため、各実験で示す長さごとに 15 個ずつ用意し探索実験を行った。各参照信号は複数の三重奏アンサンブル音楽の音響信号から切り出した楽曲の一部である。なお、これらのアンサンブル音楽については、それぞれ、参照信号を切り出した楽曲（原曲）を含め、計 8 曲の類似楽曲が DB に含まれている。これらは、原曲と、原曲にテンポ変換、移調、楽器変換のアレンジを行った各編曲を実際に演奏家に二回ずつ演奏してもらい録音したものである。実験ではこれらの 8 つの類似楽曲（目的楽曲）の検出を行うことを探索の目的とした。なお、これらのアンサンブル音楽の演奏にはピアノ、フルート、オーボエ、バスーン、バイオリン、チェロを用いた。そして、楽器変換の際には一つまたは二つのパートを異なる楽器で演奏した。

次に、比較する 3 つの探索手法 (a) スペクトル特徴を用いた場合 (b) 2 値多重音響特徴ベクトルを用いた場合 (c) 2 値多重音響特徴ベクトルを用いスパースな類似度行列を用いた場合の実験条件について述べる。まず (a) のスペクトル特徴および (b) と (c) の 2 値多重音響特徴ベクトルの抽出に用いるパワースペクトルの抽出には、帯域通過フィルタを用いることとし、75Hz から 9600Hz まで 1 オクターブごとに 48 ずつ 7 オクターブ分、計 336 個のフィルタを周波数の対数軸上で等間隔に配置したフィルタバンクを用いて、各周波数におけるパワーを抽出した。なお (a) のスペクトル特徴の抽出には、探索時の極端な速度低下を避けるため、分析区間を 400ms として 100ms ごとに抽出することとした。また (b) と (c) の場合は、各フィルタでは 44ms の時間区間の分析を、11ms ごとに行った。そして (a) の符号化は LBG アルゴリズムを用いて学習した符号帳に基づくベクトル量子化 (VQ) とし、符号間類似度は VQ コードが同じ場合に 1、それ以外は 0 とした。なお、ベクトル量子化の際の符号帳の大きさは、2 値多重音響特徴ベ

クトルとの比較のため 4096 ($=2^{12}$) とした (b) の場合は式 (1) にしたがった符号間類似度により定義される類似度行列を用い (c) の場合においては (b) の類似度行列の 0.85 より小さい成分の値を 0 として得られるスパースな類似度行列を用いた。なお、2 値多重音響特徴ベクトルを用いた場合の符号長は DB 全体で 94744 であった。

上記の条件の下、実験では三通りいずれの場合も、各特徴および類似度に基づき 2.1 の探索手続きにより探索した。ただし、参照符号系列の符号長 m に対し、探索窓の符号長 w については、 $w = m$ とし、DP 照合における整合窓については、 $|x - y| < \max(2, \lfloor 0.06m \rfloor)$ となる格子点 (x, y) は整合窓内の点であるとして定義した。 w は m に対し可変であるが、本探索実験においては参照信号と同じ演奏区間をもつ楽曲を検出することを目的としているため $w = m$ とした。

なお (b) と (c) の場合は参照符号系列の 2 値多重音響特徴ベクトルを 1 ビットずつ 12 通りにシフトしたものについても探索を行った。

3.2 全体の探索精度

まず、8 つの目的楽曲を探索した場合の全体の探索精度を表 1 に示す。参照信号の長さは 2 値多重音響特徴ベクトルで抽出した際の符号長、または、実際の演奏における音符による長さを表している。なお、実際の参照信号の平均の長さは符号長 10, 30, 50, 70 の場合で、各々、約 3, 8, 14, 19 秒であった。このとき、標準偏差は、各々、約 0.2, 0.5, 0.9, 0.9 であり、いずれの場合も、平均値からの偏差は平均値の 20% 以下であり、平均値からの変動はほとんどなかった。四分音符 24 音符分の場合は、参照信号の長さは平均約 21 秒であり、符号長の平均は約 74 であったが、符号長の最小値は 21、最大値は 114、標準偏差は 34.7 とばらつきがあった。

各参照信号に対する一度の探索においては、楽曲中に類似区間が一つでも検出された場合、その楽曲は検出されたとし、探索結果に含まれるとした。そして、各探索において、探索結果として得られた楽曲のうちの目的楽曲の割合 (P) と、目的楽曲のうちの探索結果に含まれた割合 (R) が等しくなる探索閾値でのこれらの割合の値をこの参照信号を用いた探索における探索精度とした。なお、両割合が等しくなることが複数回ある場合は、両割合が等しくなったときの値のうちの最大値とした¹。

¹実際には P と R は探索閾値に対し不連続であるため、探索閾値を 0.02 単位で変化させその間は直線で結んだ P と R のグラフの交点を、両割合が等しくなるときとしている。また、3.3 の検出率は探索精度として選ばれた交点に最も近い探索閾値での検出率である。

そして、表1の値は、各長さごとに用意した15個の参照信号での探索精度の平均値である。

表1からは、実用上重要と思われる符号長が30から70、または、四分音符24音符分の場合において2値多重音響特徴ベクトルの探索精度はスペクトル特徴を大きく上回っていることがわかる。また、2値多重音響特徴ベクトルで類似度行列をスパース化しても探索精度はあまり低下しないことがわかる。

3.3 変動ごとの検出率

次に、表1の実験において、原曲、テンポ変換、移調、楽器変換の各変動ごとに、検出率を調べた結果を表2, 3, 4, 5に示す。ここで、検出率とは表1での各探索において、各変動をもつ目的楽曲が探索結果に含まれた割合であり、表2, 3, 4, 5の値は15個の参照信号での各変動における検出率の平均値である。

表2, 3を見ると、(a), (b), (c)のいずれの場合にも、高い検出率が得られていることがわかる。これは、いずれの場合にも、照合においてはDP照合を用いており、信号の時間伸縮が吸収されているためと思われる。ただし、(a)の場合はスペクトル特徴を細かい分析間隔でとる必要があり、探索時間の観点からは(b), (c)のようにSNAP単位での特徴抽出が望ましい。

また、表4, 5からは、(b)と(c)の検出率が(a)の検出率を大きく上回っていることがわかる。

4 むすび

本稿では、文献[4]で提案されている2値多重音響特徴ベクトルを用いた類似音楽探索手法について、その頑健性を、原曲、テンポ変換、移調、楽器変換の各変動ごとに調べた。実験結果では、2値多重音響特徴ベクトルを用いた手法は移調や楽器変換の変動に対して、スペクトル特徴に比べより高い検出率を示すことがわかった。しかし、2値多重音響特徴ベクトルを用いた手法においても原曲(録りなおし)や、テンポ変換、移調等の変動に比べると、楽器変換による検出率の低下が目立つ。そこで、今後は、このような変動に対する性能の改善のため、楽器変換などにより頑健な2値多重音響特徴ベクトルの抽出法を検討する予定である。

参考文献

- [1] 柏野邦夫, ガピンスミス, 村瀬洋, “ヒストグラム特徴を用いた音響信号の高速探索法—時系列アクティブ探索法—,” 信学論(D-II), vol. J82-D-II, no. 9, pp. 1365–1373, Sep. 1999.
- [2] 黒住隆行, 柏野邦夫, 村瀬洋, “時系列アクティブ探索法のための特徴ひずみに頑健な確率ディザボータイング,” 信学論(D-II), vol. J84-D-II, no. 8, pp. 1817–1825, Aug. 2001.
- [3] 木村昭悟, 柏野邦夫, 黒住隆行, 村瀬洋, “グローバルな枝刈りを導入した音や映像の高速探索,” 信学論(D-II), vol. J85-D-II, no. 10, pp. 1552–1562, Oct. 2002.
- [4] 永野秀尚, 柏野邦夫, 村瀬洋, “2値多重音響特徴ベクトルを用いた類似音楽探索とその高速化,” 信学論(D-II), vol. J86-D-II, no. 11, pp. 1657–1667, Nov. 2003.
- [5] 永野秀尚, 柏野邦夫, 村瀬洋, “多数の小領域スペクトログラムの探索に基づく背景音楽の高速探索法,” 信学論(D-II), vol. J87-D-II, no. 5, pp. 1179–1188, May 2004.
- [6] 成田智也, 杉山雅英, “楽曲の高速検索手法の検討,” 信学技報, SP2000-16, pp. 1–8, June 2000.
- [7] 蔭山哲也, 高島洋典, “ハミング歌唱を手掛かりとするメロディ検索,” 信学論(D-II), vol. J77-D-II, no. 8, pp. 1543–1551, Aug. 1994.
- [8] A. Ghias, J. Logan, D. Chamberlin, and B. C. Smith, “Query By Humming: Musical Information Retrieval in An Audio Database,” Proc. ACM Multimedia '95, pp. 231–236, Nov. 1995.
- [9] L. A. Smith, R. J. McNab, and I. H. Witten, “Sequence-Based Melodic Comparison: A Dynamic Programming Approach,” in Melodic Similarity: Concepts, Procedures, and Applications, vol. 11 of Computing in Musicology, eds. W. B. Hewlett and E. Selfridge-Field, pp. 101–117, The MIT Press, England, 1998.
- [10] 園田智也, 後藤真孝, 村岡洋一, “WWW上での歌声による曲検索システム,” 信学論(D-II), vol. J82-D-II, no. 4, pp. 721–731, Apr. 1999.
- [11] N. Kosugi, Y. Nishihara, S. Kon'ya, M. Yamamuro, and K. Kushima, “Music Retrieval by Humming—Using Similarity Retrieval over High Dimensional Feature Vector Space—,” Proc. IEEE PACRIM '99, pp. 404–407, Aug. 1999.
- [12] 橋口博樹, 西村拓一, 張建新, 滝田順子, 岡隆一, “モデル依存傾斜制限型の連続DPを用いた鼻歌入力による楽曲信号のスポッティング検索,” 信学論(D-II), vol. J84-D-II, no. 12, pp. 2479–2488, Dec. 2001.
- [13] A. L. Uitdenbogerd and J. Zobel, “Manipulation of Music For Melody Matching,” Proc. ACM Multimedia '98, pp. 235–240, Sept. 1998.
- [14] T. Crawford, C. S. Iliopoulos, and R. Raman, “String-Matching Techniques for Musical Similarity and Melodic Recognition,” in Melodic Similarity: Concepts, Procedures, and Applications, vol. 11 of Computing in Musicology, eds. W. B. Hewlett and E. Selfridge-Field, pp. 73–100, The MIT Press, England, 1998.

表 1: 探索精度の比較 (全体)

参照信号の長さ	(a) スペクトル特徴	(b) 2値多重音響特徴ベクトル	(c) 2値多重音響特徴ベクトル (スパース化)
符号長 10	45 %	58 %	55 %
符号長 30	56 %	82 %	77 %
符号長 50	60 %	84 %	81 %
符号長 70	63 %	89 %	86 %
四分音符 24 音符分	65 %	81 %	82 %

表 2: 検出率 (原曲)

参照信号の長さ	(a) スペクトル特徴	(b) 2値多重音響特徴ベクトル	(c) 2値多重音響特徴ベクトル (スパース化)
符号長 10	97 %	90 %	93 %
符号長 30	97 %	100 %	100 %
符号長 50	100 %	100 %	100 %
符号長 70	100 %	100 %	100 %
四分音符 24 音符分	100 %	97 %	100 %

表 3: 検出率 (テンポ変換)

参照信号の長さ	(a) スペクトル特徴	(b) 2値多重音響特徴ベクトル	(c) 2値多重音響特徴ベクトル (スパース化)
符号長 10	50 %	63 %	70 %
符号長 30	90 %	83 %	80 %
符号長 50	100 %	80 %	77 %
符号長 70	100 %	90 %	90 %
四分音符 24 音符分	100 %	87 %	90 %

表 4: 検出率 (移調)

参照信号の長さ	(a) スペクトル特徴	(b) 2値多重音響特徴ベクトル	(c) 2値多重音響特徴ベクトル (スパース化)
符号長 10	27 %	50 %	40 %
符号長 30	27 %	87 %	83 %
符号長 50	30 %	93 %	97 %
符号長 70	40 %	97 %	100 %
四分音符 24 音符分	50 %	90 %	93 %

表 5: 検出率 (楽器変換)

参照信号の長さ	(a) スペクトル特徴	(b) 2値多重音響特徴ベクトル	(c) 2値多重音響特徴ベクトル (スパース化)
符号長 10	7 %	27 %	17 %
符号長 30	10 %	53 %	47 %
符号長 50	10 %	70 %	53 %
符号長 70	13 %	77 %	53 %
四分音符 24 音符分	10 %	50 %	47 %