

## 作曲や即興演奏における 旋律創作ルール抽出支援手法の検討

西本 一志      間瀬 健二  
(株) ATR 知能映像通信研究所

音楽の演奏表現に関するルールを抽出する手法の研究が従来から多く行われている。しかしこれらの手法は、演奏する楽曲があらかじめ与えられていることを前提としており、特に音高の構成に関するルール抽出が不十分であるため、楽曲自体を新たに創作する場合にどのような旋律作りをすれば良いのかという音楽表現ルールを獲得する目的には使えない。そこで本稿では、新たな旋律の創造を主とする音楽の創作活動の支援を目的として、特に音が人に与える情動的影響に注目した分析と、楽曲を構成するフレーズおよび各音楽プリミティブの関連を時系列的でなく空間的な構造として表現することにより、音楽表現ルールの抽出を支援する手法を提案する。

### A Supporting Method to Extract Rules for Melody Creation in Music Composition and Improvisation

Kazushi Nishimoto and Kenji Mase  
ATR Media Integration & Communications research Laboratories

A lot of research effort has been made on extracting rules for music expression. However, the resulting methods have dealt only with music pieces previously composed, and hence their functions to extract rules on pitch structure have been insufficient. Therefore, it is impossible to apply them to obtain rules in terms of how a music piece should be structured. In this paper, aiming at support in the creation of novel melodies, the authors propose a method of extracting rules by analyzing melodies based on the functions of notes that affect human emotions.

#### 1 はじめに

人間の演奏家による演奏データから、演奏の表情づけなどに関する演奏ルールを抽出する手法の研究が従来から多数なされており [11, 12, 14, 3], さらに何らかの与えられた楽曲をコンピュータによって演奏する際に、抽出された演奏ルールを利用して人間的な表情づけをすることにより、演奏における創造性の発露を実現しようとする試みもなされている [4, 8, 10]. これらの研究では、演奏すべき楽曲が与えられていることを前提としている。したがって、演奏ルール抽出の対象となる音楽的特徴ないし構造は、拍節構造、音の強弱、音程や音長のゆらぎなどに限られる。しかしながら、作曲という行為を取り扱おうとした場合、上記のようなルールの抽出範囲では不十分である。なぜならば、作曲においては当然旋律が事前に与えられてはいないので、楽曲自体を構成する音の音高 (pitch) とその組み合わせを取り扱う必要があるからである。

しかしながら、音高をそのまま用いてなんらかのルールを抽出することは一般に困難をはらむ。たとえば、440Hzの音高の音 (Aの音) について考えてみると、様々な楽曲において使用されるAの音、あるいはある単一の楽曲中であっても違う箇所で使用されるAの音が共通に示すルールというものが存在するとは考えにくい。これは、純粋な物理量である音高を、音楽「表現」におけるルールという感性に基づく対象へ直接に対応づけようとすることに起因する。この点について、従来有効な解は示されていなかった。

また従来の手法では、ほとんどの場合時間軸に沿った方向で演奏ルール抽出を行っている。音楽は時間軸上の芸術であるため、時間軸に沿った解析を行うことは自然であるし、また必要である。しかしながら、音楽作品を創作するためにはそれだけでは不十分であり、時間的には隣接していない離れた部分同士の関係なども考慮することにより、作品全体の構造を一貫性あるものにするのが不可欠であ

る。たとえば交響曲の創作などはその典型であり、長大な作品全体に対する構造を、複数の楽章にまたがって構築しなければならぬ。そのためには、時間軸から離れた視点が必要となる。

そこで、本稿では音楽の創作支援を意図し、音高に関連する音楽表現ルールの抽出、および時間軸から離れた楽曲構造の分析実現のための基礎的な手法を提案する。さらに提案した手法に基づき構築したプロトタイプシステムと、その適用例を示す。以下、2章では、音高に基づく表現ルール抽出を実現するための基礎となる、音の「機能」分析について概説する。3章では、統計手法を応用し、フレーズの相互関係を空間構造化することによる楽曲の全体構造分析の手法について説明する。4章では、前2章で説明した手法に基づき構築したプロトタイプシステムについて説明する。5章では、プロトタイプシステムを用いての適用例と分析例を示し、本手法の有効性について考察する。6章はまとめである。

## 2 音の機能に基づく分析

筆者らは文献 [6] において、音の機能という考え方を提案している。詳細は文献 [6] にゆずり、本章では音の機能とその求め方について簡単に述べる。

音の機能とは、なんらかの音楽的環境  $E_i$  のもとで、ある音  $s_j$  が人に特定の情動的影響を与える作用のことである。なお、 $s_1$  と  $s_2$ 、あるいは  $E_1$  と  $E_2$  とを異なった音ないし音楽的環境とする要因としては、1章で述べた理由により、本稿では音高のみを考慮する。したがって、言い替えば、音の機能とは、なんらかの音高の音（単音ないし和音）が鳴っている環境  $E_i$  のもとで、ある音高の音  $s_j$  が人に特定の情動的影響を与える作用のことである。たとえば、Cの音が鳴っている環境では、Cに対して完全5度の音程となるGの音は「完全協和」という安定感を感じさせる機能を持つ。

ただし音楽的環境や、ある音楽的環境のもとのある音の持つ機能は、時代や音楽の種類によって異なり、必ずしも一定ではない。そこで本研究では、演奏対象を1940年代中期に始まったビ・バップと呼ばれる形式のジャズ（以下、単に「ジャズ」と呼ぶ）を具体的事例としてとりあげ、楽曲各部におけるコードと楽曲の進行に伴うコード進行のみを、その楽曲の音楽的環境とする。以下、ごく簡単な場合について、ジャズ理論 [13] による一般的な分析手法を極めて簡略に説明する。

まず、コード進行をもとに、楽曲各部の調性（主音および長調／短調）を決定する。具体的には楽曲

中のドミナント7thのコードに注目し、そのコードの根音の完全5度下の音をその箇所周辺の主音であると仮定する。一般にドミナント7thの後にはトニックが来るので、ドミナント7th後のコードの根音が先に仮定した主音と一致するかどうか確認する。一致すればそのコードはトニックであると確定されるため、そのトニックの調性が、楽曲のその部分周辺の調性であるということになる。調性が決定されたら、次はその部分の各コードを、その部分の主音に対する相対音高表記で書き直す。たとえば、調性がへ長調の部分に含まれる  $G_{m7}$  のコードは、根音Gが主音Fに対して長2度にあるので  $II_{m7}$  となる。ついで、ジャズの理論で経験的に与えられている、相対音高表記でのコードに対する使用可能な音列（これをアヴェイラブル・ノート・スケールと呼ぶ）データを参照し、各コードに対するアヴェイラブル・ノート・スケールを決定する。たとえば、根音がGの  $II_{m7}$  のコードには、G dorian と呼ばれるスケールがアヴェイラブル・ノート・スケールとなる。このスケールが得られれば、各スケールの根音から数えていくつめの音にあたるかによって、各音のその部分における機能を決定することができる。各音の機能的名称は、根音を1度音とし、以下根音に近いものから順に上へ2度音、3度音…とする。

本来、オクターブ中にある12の音は全て個々に異なる機能を持つが、本研究では、作品の特徴や表現ルールの把握を容易にするために、12の機能を以下の5つの機能グループにまとめて扱う。

無彩色コード音 1度音と5度音。強い安定感を表現するが、調性感（長調、短調の別などを感じさせる機能）に乏しい。

有彩色コード音 3度音と7度音。コードの機能（調性感や終止感など）を強く表現する。

テンション音 アヴェイラブル・ノート・スケール中の音で、無彩色コード音、有彩色コード音、アヴォイド音のいずれでもない音。浮遊感や緊張感を表現する。

アウト音 アヴェイラブル・ノート・スケールに含まれない音。本来のコードから外れた、強い浮遊感を表現する。

アヴォイド音 アヴェイラブル・ノート・スケール中で Avoid Note に指定される音。本来のコードが別のコードに変化したような印象を与える。

このように、環境を考慮して音高を音の機能に表現しなおすことにより、音高と関連する音楽表現のルール（この機能の組み合わせは心地好い、など）の

抽出が可能となり、異なる環境における異なる音高の音で構成された異なる旋律が与える情動的作用を、相互に比較することができるようになる。

### 3 フレーズの空間構造化による楽曲構造分析

本研究では、楽曲はフレーズの時系列連鎖で構成され、各フレーズはその中に含まれる音楽プリミティブとその重みによって特徴づけられるものとして楽曲をモデル化する。ここに、フレーズとは「文章の句読点に相当するような段落感で区切られるひとまとまりの楽句」[1]のことである。本来その長さは不定で一概に規定できないが、本研究では、便宜的に4小節を1つのフレーズとして扱う。現代の一般的な西洋音楽では、これはさほど無理な設定ではない。また、音楽プリミティブとは、演奏を構成するすべての基本的な音楽要素のことである（たとえば、コード、ハーモニー、ピッチ、音量、テンションなど）[5]。実装にあたってどのような音楽プリミティブを用いたかについては、4.2節で示す。

上記のモデルに基づき、各フレーズ相互の関係を空間構造として可視化する。この可視化には、双対尺度法[7]を応用した概念可視化手法[9]を適用した。双対尺度法とは、複数の数量化属性で構成されたオブジェクト集合が与えられたときに、オブジェクト集合と属性集合にそれぞれ得点数量を与えることによって、オブジェクト同士の属性共有性と属性同士の共起性を空間における相対的な位置関係として表現する手法である。本研究では、個々のフレーズをオブジェクト、個々の音楽プリミティブをオブジェクト属性、音楽プリミティブの重みを属性値として双対尺度法を適用した。これによって、各フレーズと各音楽プリミティブとの関連性が1つの空間上（本研究では2次元空間）に表現される。こうして得られる空間によって、各フレーズの時系列的ではない関連が提示される。この空間を分析することにより、楽曲の全体的な構造を調べることができる。実際の空間表現とその分析例については、5章で示す。

## 4 実験システムの構成

### 4.1 システムの概要

図1に実験システムの構成の概略を示す。まず、処理対象とする楽曲のコード進行を2章に示した手法に基づき分析し、各コードにおけるアヴェイラブル・ノート・スケールを求めておく（自動分析ツ

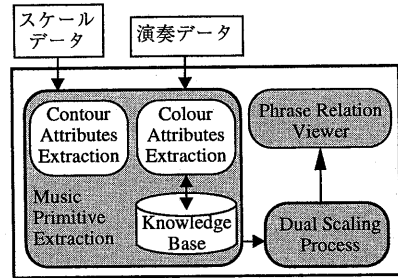


図1: 実験システムの構成図

ルも準備している[6]。作成されたスケールデータを、Music Primitive Extraction モジュールに与える。ついで、あらかじめ打ち込んでおいた演奏データを、同じく Music Primitive Extraction モジュールに与える。Music Primitive Extraction モジュールは、演奏データから重み付き音楽プリミティブを抽出する。得られた重み付き音楽プリミティブに基づき、Dual Scaling Process モジュールが個々のフレーズおよび音楽プリミティブの関連性を3章で示した方法によって2次元空間構造として表現する。この結果がPhrase Relation Viewer モジュールによって空間構造として表示される。

### 4.2 音楽プリミティブ抽出

Music Primitive Extraction モジュールによって抽出される音楽プリミティブは、輪郭属性と色彩属性とに分類される。

#### 4.2.1 輪郭属性

輪郭属性とは、演奏されるフレーズにおける音高の時系列的動きから得られる属性である。なお、音高はすべてMIDIノートナンバーで扱っており、実際に発音された音の周波数を処理しているのではない。本研究では、以下の9種類のプリミティブを定義する。

1. 音数：1フレーズ中で演奏された音の数。個々の音の長さは不問。
2. 平均音高：1フレーズ中の全ての音の音高の、各音の長さを考慮した加重平均値。
3. 音高分散：1フレーズ中の全ての音の音高の、各音の長さを考慮した分散。
4. 代表繰り返し長：正規化した音高の推移の自己相関を計算した結果、最も強い自己相関係数が得られた周期。

5. 極長周期脈動強さ：音高の推移を3次スプライン補間し，FFTによって得たパワースペクトルの，2分音符長以上のパワーの積分。
6. 長周期脈動強さ：同じく4分音符以上2分音符未満のパワーの積分。
7. 中周期脈動強さ：同じく8分音符以上4分音符未満のパワーの積分。
8. 短周期脈動強さ：同じく16分音符以上8分音符未満のパワーの積分。
9. 極短周期脈動強さ：同じく16分音符未満のパワーの積分。

#### 4.2.2 色彩属性

色彩属性とは，演奏されるフレーズ中の音を，各演奏時点におけるアヴェイラブル・ノート・スケールに照合して得られる音機能に基づく属性である。本研究では，2章で示した，音の5つの機能グループを，そのまま5つのプリミティブとして扱う。各プリミティブの重みは，フレーズ中でのそれらの音の長さ強さの積の総和とする。

### 5 実験と考察

プロトタイプシステムを用いていくつかの適用実験を行った。ここでは，プロによる既存の演奏データに対する分析例を示し，本システムの有効性を論じる。

#### 5.1 プロの演奏データの分析例

本実験で使用した演奏事例は，アルバム *Somethin' Else* に収録されている，ジャズのスタンダードナンバー「枯葉」の演奏における，Miles Davisによるトランペットのソロ即興演奏と，Julian “Cannonball” Adderleyによるアルト・サクソのソロ即興演奏である。この枯葉の演奏は非常に有名であり，その演奏の質に対する評価は高い。そこで，このような「良い」事例について分析することにより，どのようなソロを構築することが良い演奏（創作）につながるのかを検討する。

図2にMiles Davisによる64小節（16フレーズ）の演奏を2次元空間構造化したものを，図3にCannonballによる64小節（16フレーズ）の演奏を2次元空間構造化したものを，また図4には，この両者によるフレーズ全部によって1つの空間を構築したものを示す。

まずこの図の見方の概略を説明する。この空間において，水平軸は第1主成分に対応し，垂直軸は第2主成分に対応している。原点は空間の中心にお

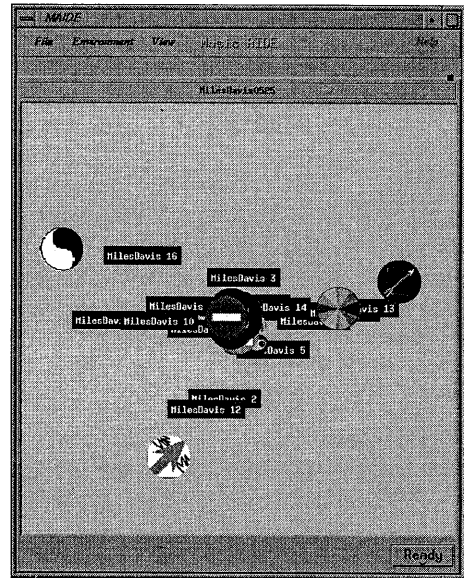


図 2: Miles Davis による即興演奏の構造

かれる。空間上の長方形のアイコンは各フレーズに対応し，アイコン中には演奏者名と，その演奏者によるソロの何フレーズ目であるかが示されている。また，様々な図柄のアイコンは，それぞれいずれかの音楽プリミティブに対応している。たとえば，図2の最も左にあるアイコンは無彩色コード音に，最も下にあるアイコン（矢印が何かを打ち破っている図柄）はアウト音に，最も右のアイコン（弓矢の図柄）はテンション音に，画面中央にあるアイコン（交通標識の進入禁止）はアヴォイド音に，図3の最も右にあるアイコン（多数の色の濃さの違う扇形の集合）は有彩色コード音に，それぞれ対応する。また，重なっていて判別しにくい，いずれの図においても画面中央付近に密集している音符の図柄のアイコンは，5種類の脈動強さプリミティブに，目玉の図柄のアイコンは音高分散のアイコンに，それぞれ対応している。この他のプリミティブのアイコンは重なっているため，この3つの図では見えない。

これらの図から，図2では第1主成分は主に無彩色コード音，テンション音，有彩色コード音の3つのプリミティブで特徴づけられ，第2主成分はアウト音とわずかに無彩色コード音で特徴づけられていることが即座に読み取れる。同時に，空間中心（原点）付近にあるその他の属性は，各フレーズの差別化に寄与していないこともわかる。したがって，

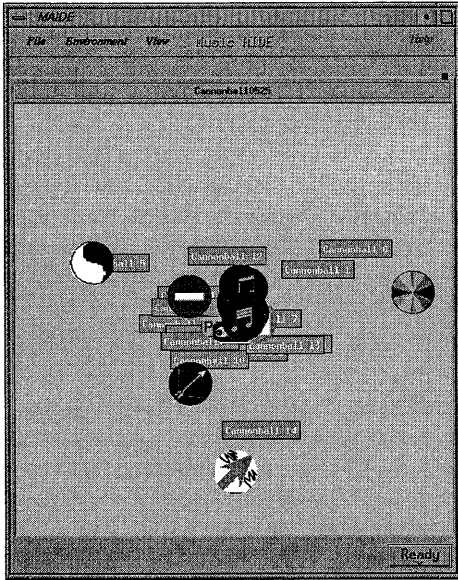


図 3: Cannonball による即興演奏の構造

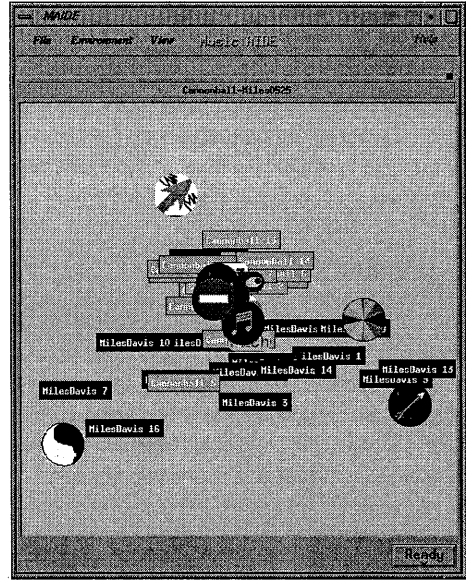


図 4: 両者による即興演奏の構造

Miles Davis のフレーズ作りの特徴は、無彩色コード音、テンション音、有彩色コード音という、ジャズにおいては基本的で理論に沿った音の使い方にあることがわかる。特に興味深いのは、テンション音と有彩色コード音には強い正の相関がある（両者のアイコンが近くに配置されている）こと、そしてこれら2つのプリミティブと無彩色コード音とは強い負の相関を持つこと（水平方向の両端にアイコンが配置されている）である。この結果をジャズギターを真剣に勉強している者に見せた所、特に有彩色コード音と無彩色コード音の負の相関に興味を示した。従来そのような音使い、すなわち有彩色コード音と無彩色コード音とを排他的に使用するという方法は聞いたことがなく、考えたこともなかったというのである。そして、即座にギターを持っていくつかのフレーズを試しに即興演奏してみた結果、特に無彩色コード音の5度音を有彩色コード音と同時に使用しない方が「イキな」フレーズとなることを発見した。これはこのような分析が直接に有益なルールの抽出に貢献した好例であると言えるだろう。

一方、図3では、無彩色コード音と有彩色コード音が第1主成分に強く寄与しているのは Miles Davis の場合と同じであるが、さらにアヴォイド音も比較的大きな貢献をしていることが読み取れる。また、第2主成分はアウト音とテンション音によって主に

特徴づけられている。Cannonball のこの枯葉における即興演奏は、非常に不安定で、時にはまったくコード進行から逸脱したような印象を与える演奏を行っている。これは、Cannonball が、枯葉本来のコード進行をさらに細分化し、より複雑なコード進行に再構成して演奏を行っている一方、伴奏者側は基本的に本来のコード進行に準じて演奏を進めているためである。そしてこの事実は、この図上でアヴォイド音の使用、およびテンション音とアウト音という、共に不安定感を産み出すプリミティブが正の相関を持つことに現れており、Cannonball の演奏のルール的一端が把握できる。また、Miles の場合に比べて、わずかではあるが輪郭属性に属するプリミティブのアイコンが原点から分散していることは、音の動きに変化が激しいことを示している。

さらに、両者の演奏を1つの空間上に表現することにより、同じ軸と尺度のもとで両者の差を浮彫りにする。図4において、第1主成分は無彩色コード音、テンション音、有彩色コード音に特徴づけられており、第2主成分はアウト音とアヴォイド音、および一部の輪郭属性プリミティブに特徴づけられている。そして、明確に Miles の演奏は水平方向に分布し垂直方向の分布は少なく、逆に Cannonball の演奏は垂直方向、特に中央より上方に分布し、水平方向の分布は少ない。この結果から、Miles は理論

に沿った素直な音使いによる演奏の特徴づけをしているのに対し、Cannonballは緊張感の高い音によるフレーズ作りをしていることがより明確になる。

以上のような分析から、Miles風のフレーズ作り、あるいはCannonball風のフレーズ作りにはどのような音使いをすればよいかという創作におけるルール的一端を抽出することが可能となる。

## 5.2 問題点

現状のシステムの第1の問題は、輪郭属性の特徴抽出が現状ではうまくいっているとは言えないことである。図2~4などを見ればわかる通り、輪郭属性は主成分に対してほとんど寄与していない。これについては、1章で示した、従来の演奏ルール抽出手法から適当なものを採用することが必要であろう。

第2の問題は、旋律の創作に重要なもう一つの要素であるリズムをどのように取り扱えば良いのか(音高に対する音の機能のように、リズムに対するリズムの機能のような概念はあるのか、あるならばどう抽出すればいいのか)についての解が得られていないことである。これは、より具体的な形で音楽創作に関するルールを獲得するために非常に重要な課題であるので、今後研究を進めたい。

## 6 おわりに

本稿では、作曲や即興演奏などの、新たな旋律の創造を主とする音楽の創作活動の支援を目的とした、音楽表現ルールの抽出手法について検討した。すなわち、音が人に与える情動的影響としての音の機能を分析し、また楽曲を構成するフレーズおよび各音楽プリミティブの関連を時系列的でなく空間的な構造として表現することにより、音楽表現ルールの抽出を支援する。実際の演奏データに対して適用した結果から、本手法の有効性が示唆された。今後はさらにリズムなどに対する処理手法を考案し組み込むことにより、より実用的なシステムを構築したいと考えている。

## 謝辞

本研究の機会を与えて下さった(株)ATR知能映像通信研究所の酒井保良会長ならびに中津良平社長に感謝致します。また、双対尺度法処理ライブラリを御提供下さった学術情報センターの杉本雅則氏、空間構造の分析について有益な御示唆を下さった角康之氏、実験に御協力下さった馬田一郎氏に感謝致します。

## 参考文献

- [1] 浅香淳(編):音楽中辞典,音楽之友社,1994.
- [2] 藤井英一監修:完全コピー ジャズスコア, Vol.2,株式会社ヤマハミュージックメディア,1995.
- [3] Hiraga, R.: Objective Evaluation for Computer Generated Musical Performance, IJCAI-97 Workshop Program Working Notes "Issues in AI and Music —Evaluation and Assessment—," pp.23-26, 1997.
- [4] 五十嵐滋, 弥富あかね, 小川大典, 清野桂子: PSYCHE から: フレーズ表情の多角形, 音楽構造に基づく表情づけ, 自動連弾システム, 情処研報, 音楽情報科学 9-1, pp.1-6, 1995.
- [5] 金森 務, 片寄 晴弘, 新美 康永, 平井 宏, 井口 征士: ジャズセッションシステムのための音楽認識処理の一実現法, 情処論, Vol.36, No.1, pp.139-152, 1995.
- [6] 西本一志, 渡邊 洋, 馬田一郎, 間瀬健二, 中津良平: 創造的音楽表現を可能とする音楽演奏支援手法の検討 —音機能固定マッピング楽器の提案, 情処論, Vol.39, No.5, pp.1556-1567, 1998.
- [7] 西里 静彦: 質量データの数量化 — 双対尺度法とその応用 —, 朝倉出版, 1982.
- [8] 白川健一, 小田彦彦, 熊谷俊行, 梶川嘉延, 野村康雄: 演奏者情報を加味したピアノの自動演奏システムの構築, 情処研報, 音楽情報科学 10-1, pp.1-6, 1995.
- [9] 角康之, 小川 竜太, 堀 浩一, 大須賀 節雄, 間瀬 健二: 思考空間の可視化によるコミュニケーション支援手法, 信学論, Vol.J79-A, No.2, pp. 251-260, 1996.
- [10] 竹内好宏, 片寄晴弘: Two Finger Piano による曲想の表現, 情処研報, 音楽情報科学 11-6, pp.37-44, 1995.
- [11] 竹内好宏, 保科 洋: 演奏表情を知覚させる演奏変数の研究, 情処研報, 音楽情報科学 15-2, pp.7-12, 1996.
- [12] 上符裕一, 青野裕司, 片寄晴弘, 井口征士: 演奏ルールの抽出について, 情処研報, 音楽情報科学 15-14, pp.79-84, 1996.
- [13] 渡辺貞夫: ジャズ・スタディ, 株式会社エー・ティ・エヌ, 1996.
- [14] 矢向正人, 土屋景一, 荒木敏規: 旋律パターンの分類 - 類似性判断と分析例 -, 情処研報, 音楽情報科学 16-5, pp.27-32, 1996.