

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 9 日現在

機関番号：13601

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500148

研究課題名(和文) 補助楽曲データを用いたペンコンピュータ上での楽譜作成支援システムの構築

研究課題名(英文) Development of a music score production system on a pen-based computer by using auxiliary music data

研究代表者

宮尾 秀俊 (MIYAO, Hidetoshi)

信州大学・工学部・准教授

研究者番号：10239353

交付決定額(研究期間全体)：(直接経費) 4,000,000円、(間接経費) 1,200,000円

研究成果の概要(和文)：作曲過程で事前に演奏入力された楽曲中から、現在ペン筆記した楽譜に類似した箇所を検出するために、双方をピアノロールで表現し、それらの図形的な重なりによって検索する手法を提案した。実験結果として、3音以上の楽譜記号を筆記した場合に3位以内に7割以上の確率で目的の演奏箇所を出力できることを示した。一方、手書きで書かれた検索クエリに基づき手書きメモを検索する手法を考案した。具体的には、筆記ストロークを円で表現し、それらの重なり面積に基づく類似度を定義することによって検索を実現した。1425本のストロークを対象に30クエリの検索を行った結果、10位以内に97%の確率で目的の要素を検出できることを示した。

研究成果の概要(英文)：In a common music composition process, a user composes a music phrase by a trial and error using a musical instrument. In order to retrieve similar music phrases in the musical performance based on a handwritten music query, we have proposed that the both music data are represented by piano rolls and then the similarity of them are calculated based on the overlap area of them. In our experiments, top-3 retrieval accuracy of 70% or more was obtained by using a query consisting of three or more handwritten musical symbols. Next, to retrieve an arbitrary part of handwritten objects from a database by a handwritten query, we have proposed a new similarity measure based on an overlap area of circles whose positions and sizes are determined by positions and category types of handwritten strokes. We examined the effectiveness of the method using a database of 1425 strokes and 30 handwritten queries. As a result, top-10 retrieval accuracy of 97% was obtained.

研究分野：総合領域

科研費の分科・細目：情報学・メディア情報学・データベース

キーワード：楽譜作成支援 ペンコンピュータ 手書き楽譜認識 楽曲検索

1. 研究開始当初の背景

音楽分野において作曲・編曲活動を行う場合、一般的に作曲者は次のような作業を行っていると考えられる。(1). 曲の一部(フレーズ)をイメージし、演奏してみる。演奏したフレーズが意図したものと異なる場合は、試行錯誤しながら演奏を繰り返す。(2). 自分の意図したフレーズが得られた場合、その部分を手書き楽譜として五線譜に記述する。現在創作しているフレーズの前に、すでに作曲した部分がある場合は、その部分の楽譜を修正する場合もある。(3). (1)と(2)を繰り返すことにより、曲全体を完成させる。このような手書きによる楽譜作成作業は、思考したフレーズを即座に楽譜として記録できるという利点を備える反面、楽譜の後利用・編集が難しいといった問題を抱えている。

一方、楽譜作成ソフトウェアの利用や MIDI 楽器入力による楽譜作成は、後に楽譜の自動演奏・編集操作などが可能になる反面、ソフトウェアを用いる場合は、作曲者はその使用方法に習熟する必要があり、創作したフレーズも即座に記録できない、楽器入力の場合は、入力した楽譜の編集作業に時間がかかるなどの問題を抱えている。このような理由から、いまだに手書き楽譜によって作曲を行なっている作曲者が多いのが現状である。このような問題点を解決するために、研究代表者らは、パソコン上で書かれた手書き楽譜を逐次に認識してコンピュータが理解できる表現形式に変換する手法を開発してきた。しかし、人によって大きく形状が変化する手書き楽譜記号を精度良く認識することは難しい。そこで、本研究では、作曲過程で演奏された情報を有効に利用したパソコン上で手書き楽譜認識の精度向上と楽譜の編集性向上を目指した楽譜作成システムの開発を考えた。

2. 研究の目的

本研究では、使い勝手の良いパソコン上で楽譜作成システムの構築を目的としている。そのために下記手法を考案する。

- (1) 作曲過程の楽曲情報を利用したオンライン手書き楽譜認識の精度向上手法の考案
- (2) ペン筆記を用いた楽譜編集のための優れた GUI の考案
- (3) 既存楽曲からの効率の良い検索手法の考案

3. 研究の方法

本研究では、システム構築のために特に重要と考えられる下記の2つの事項について研究を行った。

- 現在ペン筆記で入力した楽譜データを検索クエリとして、事前の作曲過程で入力された演奏データ中から、類似の演奏データを抽出して出力する手法を考える必要がある。ここでは、抽出データを

楽譜化して、候補提示し、楽譜入力の手間を軽減する手法を提案する。

- 楽譜の編集作業においては、以前に入力した楽譜データを利用したい場合がある。その際、手書きで書かれた楽譜記号を検索クエリとして、以前に書かれた楽譜記号を検索する必要がある。ここでは、楽譜記号にこだわらず、一般的な手書き記号をも対象とした検索手法を提案する。

上記2つの事項に関する手法を下記(1)、(2)において説明する。

- (1) 手書き楽譜認識システムにおける演奏情報に基づく楽譜作成支援

本手法では、研究代表者らが従来に開発したオンライン手書き楽譜認識システムを使用する。このシステムは、ペンタブレット等により手書きで楽譜記号を入力すると、それらを逐次に自動認識し、音楽記号単位で出力してくれる。今回は認識結果に誤りがなく、MusicXML データ形式で出力されると仮定し、このデータから、時系列の音高・音長、和音、タイに関する情報を抽出し、縦軸を音高、横軸を音長で表したピアノロール(図1)に変換して、これを検索クエリとして使用する。検索対象のデータベースとして、楽譜を書く前に MIDI 楽器で演奏された演奏情報を使用する。この MIDI データから音高・音長を抽出し、検索クエリと同様にピアノロール形式で表し、これらを検索対象とする。検索処理では、検索クエリとデータベースのピアノロールを図形的に重ね合わせることで、検索を行う。具体的な手順は次のとおりである。検索クエリの先頭音が単音の場合、その先頭音を基準音とする。データベース中で、この基準音が鳴っているフレーズを見つけ、検索クエリ全体を、検索クエリの基準音の長さが、データベース中の基準音の長さに合うように正規化する。次に、データベース中のフレーズと正規化した検索クエリを図形的に重ね合わせ、一致率を計算する。いま、検索クエリ中で音が鳴っている部分のピアノロールの面積を A、データベースのフレーズ中での該当箇所での鳴っている部分の面積を B、検索クエリに出現する音高において、音が鳴っていない部分の面積を C、データベースのフレーズ中での該当箇所での音が鳴っていない部分の面積を D とすると(図1)、一致率は、 $(A \text{ と } B \text{ の重なり部分の面積} + C \text{ と } D \text{ の重なり部分の面積}) / (A + C)$ で計算する。この操作をデータベース中に出現する基準音すべてについて行う。その結果、計算された一致率が最も大きいものから順に候補として提示する。一方、マッチングで得られた部分の演奏を忠実に楽譜に変換して出力すると、音長は、正確に演奏されていることが少ないため、複雑なリズムの楽譜になる可能性が高い。このため、演奏情報に補正をかけ、意図したフ

レーズに整形して出力する必要がある。そこで、本手法では、候補提示するデータベース中のフレーズの基準音の長さを検索クエリの基準音の長さに合うように正規化し、8分音符を基準に量子化して、候補提示することにした。また、検索クエリの先頭音が和音の場合、この先頭音に出現する和音の各音すべてで、基準音をずらしながら検索を行う。先頭和音の各音で計算した一致率を平均化し、その平均化された一致率を使い、順位付けを行う。候補提示を行う前に、先頭和音中で最も一致率が高かった音を基準に、量子化を行う。このように、量子化し、整形された演奏情報を予測候補として楽譜で提示し、作曲者がこの候補を選択できるようにした。

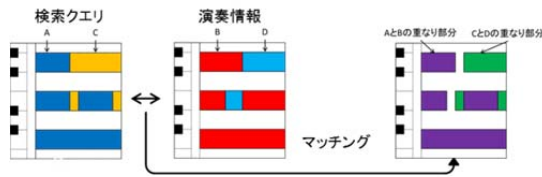


図1 検索クエリと演奏情報のピアノロール表現とそれらのマッチング結果

(2) 円を用いたストローク表現によるオンライン手書き記号の検索

①チェーンコード生成

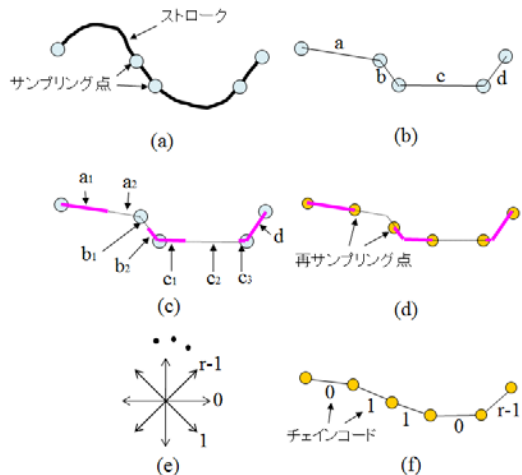


図2 チェーンコードの生成

本研究では、ペンが筆記面についてから離れるまで、つまり一筆で筆記された線を1ストロークとし、全ての筆記要素をストローク単位に分割する。ここで、各ストロークはペン位置の2次元座標列で表現される。そして、図2の例のように、以下の手順でこのストロークをストローク長に依存せずコード長が等しいチェーンコードに変換する。

- サンプル点を直線で結び(図2(b))、各線分の長さの和を求める。そして、この長さの和をあらかじめ指定したコード長 t で割ることで、平均の長さ L が得られる。図2(b)の例では、

$L = (a + b + c + d)/t$ となる。

- 図2(b)の線分上で距離 L ごとに再サンプリングする。図2(c)では、 $L = a_1 = a_2 + b_1 = b_2 + c_1 = c_2 = c_3 + d$ である。そして、この再サンプリング点を直線で結ぶ。
- 円を等角度に r 方向に分割し、各方向に0から $r-1$ までの番号を与える(図2(e))。再サンプリング点を結んだ各線分に最も近い方向の番号を付与し、チェーンコードとする(図2(f))。

以上のように、各ストロークを長さ t で r 方向のチェーンコードで表現する。

②各ストロークのクラスタリング

データベース内の全ストロークに k-means クラスタリング手法を適用する。その結果、 k 個のクラスタ中心が得られ、全ストロークは k グループに分類される。ここで、クラスタ数 k はクラスタリングを行う前に指定しておく。クラスタリングの過程で、2つのストローク間の距離の定義と、複数ストロークの平均を求める計算が必要になる。

2つのストローク \mathbf{x} と \mathbf{y} のチェーンコードを $\mathbf{x} = \{x_1, x_2, \dots, x_i, \dots, x_t | x_i \in 0, 1, \dots, r-1\}$ 、 $\mathbf{y} = \{y_1, y_2, \dots, y_i, \dots, y_t | y_i \in 0, 1, \dots, r-1\}$ と表した場合、ストローク間の距離 d を次のように定義する。

$$d = \sum_{i=1}^t \rho_i$$

$$\rho_i = |x_i - y_i|$$

$$(\rho_i > \lfloor r/2 \rfloor \text{ の場合 } \rho_i \leftarrow r - \rho_i)$$

そして、 n 個のストロークの平均ストローク $\bar{\mathbf{c}}$ を $\bar{\mathbf{c}} = \{\bar{c}_1, \bar{c}_2, \dots, \bar{c}_i, \dots, \bar{c}_t | \bar{c}_i \in 0, 1, \dots, r-1\}$ と表す。ここで、 \bar{c}_i は次のベクトル $\bar{\mathbf{v}}_i$ に最も近い図2(e)の方向コードである。

$$\bar{\mathbf{v}}_i = \sum_{p=0}^{r-1} m_p \mathbf{v}_p$$

ここで、 \mathbf{v}_p は方向コード p の単位ベクトルであり、 m_p はストローク集合中の j 番目のストローク \mathbf{c}_j を $\mathbf{c}_j = \{c_{j1}, c_{j2}, \dots, c_{ji}, \dots, c_{jt} | c_{ji} \in 0, 1, \dots, r-1\}$ としたときのコード集合 $\{c_{1i}, c_{2i}, \dots, c_{ni}\}$ 中の方向 p のコードの数である。

③類似検索アルゴリズム

検索したい手書き要素は複数のストロークから成り、各ストロークは所属するクラスタのラベルとストロークの中心点座標によって表現される。 $D = \{(a_p, \mathbf{x}_p)\}_{p=1}^N$ をデータベースのストローク集合、 $Q = \{(b_q, \mathbf{u}_q)\}_{q=1}^n$ をクエリのストローク集合とする。ここで、 a_p, b_q はラベル、 $\mathbf{x}_p, \mathbf{u}_q$ は中心点座標である。そして、 D の各ストローク (a_p, \mathbf{x}_p) に対して以下の手順を実行する。

- クエリストロークの正規化を行う。 (b_1, \mathbf{u}_1) の中心点座標を $\mathbf{u}'_1 = \mathbf{x}_p$ とし、 (a_p, \mathbf{x}_p) と (b_1, \mathbf{u}_1) のストローク長の比 σ を求める。そして、 Q の各ストローク

の中心点座標を $\mathbf{u}'_q = \mathbf{u}'_1 + \sigma(\mathbf{u}_q - \mathbf{u}_1)$ とする。図3の例では、まずクエリストローク q_1 の中心点座標をデータベースstroーク d_1 と一致させる。この際の q_1 と d_1 のstroーク長の比 σ を求め、残りのクエリストローク q_2 と q_3 の中心点座標を q_1 とのベクトルの差が σ 倍になる位置に設定する。

- 正規化後の各クエリストローク (b_j, \mathbf{u}'_j) に対して類似度 $s(a_i, \mathbf{x}_i; b_j, \mathbf{u}'_j)$ が最も大きくなるようなデータベースstroーク $(a_{\varphi(j)}, \mathbf{x}_{\varphi(j)})$ を D から探し出す。
 - 全クエリストロークの最大類似度の合計 $\sum_{j=1}^n s(a_{\varphi(j)}, \mathbf{x}_{\varphi(j)}; b_j, \mathbf{u}'_j)$ を求める。
- 以上の手順をデータベースの全てのstroークに対して行い、合計最大類似度の大きい箇所から降順に結果として出力する。

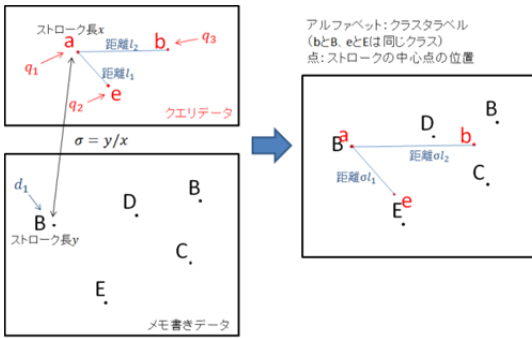


図3 クエリストロークの正規化と重ね合せ

類似度 $s(a, \mathbf{x}; b, \mathbf{u})$ の定義として、ラベル a, b のクラスタ中心間の距離 $\delta(a, b)$ と中心点座標の幾何学的距離 $l = \|\mathbf{x} - \mathbf{u}\|$ をどちらも同時に考慮するために、本研究では円の重なり面積を基にした類似度尺度を提案する。

- クエリストローク (b, \mathbf{u}) に関する円 C_0 を作成する。 C_0 の中心点の位置は \mathbf{u} 、半径はあらかじめ指定する値 r_0 とする。
- データベースstroーク (a, \mathbf{x}) に対しても円 C_1 を作成する。 C_1 の中心点の位置は \mathbf{x} 、半径 r_1 はラベルのクラスタ間距離 $\delta(a, b)$ に応じて次の式で決定される。

$$r_1 = r_0 \left(1 - \frac{\delta(a, b)}{\delta_{max}}\right)^\alpha \quad (\alpha \geq 0)$$

ここで、 δ_{max} は全てのクラスタ中心の組み合わせでの最大のクラスタ間距離である。クラスタ間距離が小さいほど、つまり2つのstroークの形状に近いほど半径 r_1 は大きくなる。

- C_0 と C_1 の重なっている部分の面積 $f(l, r_1)$ を計算する。 $f(l, r_1)$ を類似度尺度とすることでクラスタ間距離と幾何学的距離を同時に考慮することができる。
- $f(l, r_1)$ に基づき、次式で類似度を定める。

$$\hat{f}(l, r_1) = f(0, r_1) \left\{ \frac{f(l, r_1)}{f(0, r_1)} \right\}^\beta \quad (\beta > 0)$$

図4に類似度算出の例を示す。始めにクエリストローク q_1 に対する類似度を求める(図4(a))。 q_1 のクラスタラベルは a なので、データベースの各stroーク円の半径は、そのstroークの所属クラスタとクラスタ a とのクラスタ間距離によって決まる。これにより最大類似度 f_1 と類似stroーク d_1 が求められる。残りのクエリストローク q_2, q_3 に対しても同様に最大類似度 f_2, f_3 と類似stroーク d_2, d_3 を求める(図4(b), (c))。そして、検索クエリとデータベース中のこの箇所の適合度は合計最大類似度 $\hat{f}_1 + \hat{f}_2 + \hat{f}_3$ によって計算される。

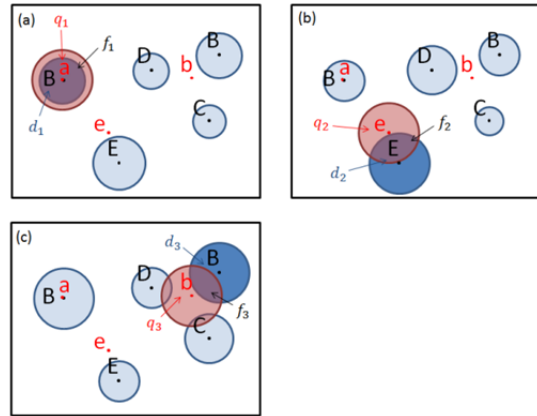


図4 類似度の算出

4. 研究成果

(1) 3節の手法(1)に対する実験結果を示す。実験では、和音を含む曲(203音, 18小節)を演奏し、手書き楽譜情報20箇所でもッチングを行い、1位と3位以内に目的の出力が出現する確率を調査した。手書き楽譜情報の検索クエリの音数を2音~6音まで変化させて検索率を測定した結果を表1に示す。

表1 楽曲検索の結果

| | 2音 | 3音 | 4音 | 5音 | 6音 |
|------|-----|-----|-----|-----|-----|
| 1位 | 35% | 50% | 40% | 75% | 80% |
| 3位以内 | 45% | 80% | 70% | 80% | 90% |

この結果より、検索クエリが3音以上であれば、70%以上の確率で3位以内までに候補を提示することが可能となった。これより、手書き楽譜として記号を書いた場合に、それまでに演奏入力されたデータから該当する演奏部分の候補を抽出することができる。よって、この成果は、研究目的で述べた手書き楽譜認識の誤認識の補正にも十分活用できると考えられる。

(2) 次に3節の手法(2)に対する実験結果を示す。実験では、被験者一人がペンタブレット(Wacom CTE-650)で筆記したメモ書きデータを用いて検索を行った。このデータには数式、グラフ、図、絵、地図記号など様々な

要素を含み、合計 1425 本のストロークで構成されている。チェインコードはコード長 $t=10$ 、方向の数 $r=16$ とした。また、クラス数 $k=30$ としてクラスタリングを行った。検索クエリとしてはストローク数 3 から 36 (平均 11.9) の 30 パターンを用意し、それらの要素はデータベース中にそれぞれ 1 つずつ存在する。なお、30 パターンの検索クエリは普通に筆記したものと、意識的に雑に筆記したものをそれぞれ用意した。上記の環境、データにおいて、実験を行った結果、 α と β に適切な値を選ぶことで、正しい検索結果が候補の 10 位以内に出現する割合が 97%、雑に筆記したクエリデータを用いても 90% という高い検索率が得られた。よって、この手法は有効であると結論づけることができる。さらに、本手法ではデータベース中の要素の任意の部分を検索できるという特長もある。この成果より、手書きの楽譜記号を検索クエリとして入力することにより、すでに筆記した手書き楽譜中から目的の部分容易に検索できるという研究目的が達成できると考えられる。

5. 主な発表論文等

[学会発表] (計 12 件)

- ① 小林 晋也、宮尾 秀俊、丸山 稔、タブレット端末による楽譜作成システムの構築、電子情報通信学会信越支部大会、2013 年 10 月 5 日、長岡技術科学大学
- ② 林 直哉、宮尾 秀俊、丸山 稔、手書き楽譜認識システムにおける演奏情報に基づく楽譜作成支援、電子情報通信学会信越支部大会、2013 年 10 月 5 日、長岡技術科学大学
- ③ Tsubasa Yagi, Hidetoshi Miyao, Minoru Maruyama, Robust On-Line Handwritten Object Retrieval using a Similarity Measure Based on Overlap Area of Circles, International Conference on Pattern Recognition (ICPR2012), 2012 年 11 月 13 日、つくば国際会議場
- ④ Hidetoshi Miyao, Rei Maruyama, On-line handwritten flowchart recognition, beautification, and editing system, International Conference on Frontiers in Handwriting Recognitions (ICFHR2012), 2012 年 9 月 18 日、イタリア
- ⑤ 飯島 愛子、宮尾 秀俊、丸山 稔、オンライン手書き楽譜認識システム用識別器に関する検討、電子情報通信学会信越支部大会、2011 年 10 月 8 日、新潟工科大学
- ⑥ 林 直哉、宮尾 秀俊、ハミングによる楽曲検索システムにおける検索の高速化、電子情報通信学会信越支部大会、2011 年 10 月 8 日、新潟工科大学

6. 研究組織

(1) 研究代表者

宮尾 秀俊 (MIYAO, Hidetoshi)

信州大学・工学部・准教授

研究者番号：10239353