

# 研究速報

## 複数情報が重畳した尺八譜認識に関する一検討

澤木美奈子<sup>†</sup> (正員)      村瀬 洋<sup>†</sup> (正員)  
 萩田 紀博<sup>††</sup> (正員)      石井健一郎<sup>†</sup> (正員)

A Study on Syakuhachi Score Recognition with Embedded Symbols

Minako SAWAKI<sup>†</sup>, Hiroshi MURASE<sup>†</sup>, Norihiro HAGITA<sup>††</sup>, and Kenichiro ISHII<sup>†</sup>, Members

<sup>†</sup> NTT 基礎研究所, 厚木市

NTT Basic Research Laboratories, Atsugi-shi, 243-0198 Japan

<sup>††</sup> NTT コミュニケーション科学研究所, 京都府

NTT Communication Science Laboratories, Kyoto-fu, 619-0237 Japan

あらまし 文字や記号などの複数の情報が重畳して記述されている尺八譜を認識する手法について述べる。補完類似度を用いることにより、複数の情報を独立に扱うことが可能となり、高い認識率が得られることを示す。

キーワード 楽譜認識, 類似度

### 1. まえがき

楽譜認識の研究は、楽譜情報からの自動演奏、転調した楽譜の作成、楽譜データベースの検索、などを目的に、多く取り組まれてきた[1], [2]。これらの多くは五線譜を対象としており、邦楽を対象としたものは、志村らによる尺八譜認識編集演奏システム[3]や、祖慶らによる琉球サンシン楽譜(工工四)の認識[4]、などがあるものの、研究例は非常に少ない。しかし、邦楽譜は、楽器や流派により記述法が異なるために理解することが困難である。そのため、コンピュータによって五線譜への変換や自動演奏を行うことにより容易に理解できる形式に変換することは非常に重要である。

五線譜または邦楽譜によらず、一般に楽譜では、音の高さや長さ、歌詞などの多くの情報を効率良く伝達するために、文字や記号などが重畳して記述されている場合が多い。しかし、従来のパターン認識手法は、単独の情報のみからなるパターンを対象としており、複数の情報が重畳したパターンを認識することは困難であった。そのため、楽譜を認識するためには、個々の情報(五線譜の例では、五線、音符の頭、音符の棒、など)を、黒画素の射影などに基づき順々に切り出し、切り出された個別の情報ごとに認識する手法がとられていた。

しかし、ピッチや大きさなど重なり合っている情報の性質が異なるために、単純な黒画素の射影で個別の情報を取り出すことは困難であった。個々の情報の特

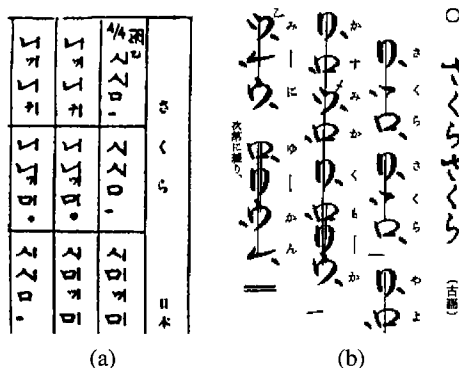


図1 尺八譜の例 (a)都山流, (b)琴古流  
 Fig. 1 Examples of Syakuhachi scores. (a)Tozan style, (b)Kinko style

徴や複数の情報が重畳する条件などのさまざまな制約を記述して切出しに利用することも考えられるが、実際の楽譜では、印刷上の都合や作曲家の多様な意図を表現するために、記述された制約に反する状況が多々発生する。その結果、個別の情報をきれいな状態で抽出することができず、他の情報の一部が残ったり、着目している情報の一部が欠損したりするため、認識が困難になっていた。

尺八譜に対しても、同様の問題がある。尺八には都山流と琴古流の二つの主流な流派がある。これらの流派の譜面例を図1に挙げる。志村らの研究では、都山流の譜面を取り上げ、黒画素数の射影による個別の情報の切出しを行った後にパターン照合により認識する[3]。都山流の譜面では、基本的に複数の情報を重畳せずに記述するため、黒画素数の射影を再帰的に行うことにより、個別の情報の切出しが可能となるのである。しかし、もう一つの主流な流派である琴古流などのように、複数の情報を重畳して記述する記法には対応できなかった。

本手法では、補完類似度[5]とずらし照合を組み合わせることにより、個別の情報を切り出すことをせずに、複数情報が重畳したままのパターンを用いて認識する。補完類似度は、汚れやかすれの雑音にロバストであるという性質があるため、複数の情報が重畳していても、着目した情報以外を自動的に汚れ雑音とみなすことにより、着目している情報を正しく認識することができる。そのため、個別の情報の切出しの際に発生する画質の劣化を受けることがない。更に複数の情報を独立に扱えるため、情報が重畳する条件というような複数の情報間での影響に関する知識をもつ必要が

ない。また、譜面の老化やコピーなどにより画像が汚れたりかすれたりした場合にも、雑音の影響をあまり受けずに認識できる。更に、ずらし照合を用いることにより、ピッチの異なったカテゴリーに対応することができ、個別文字の際に発生する切出し誤りの影響を受けない。これらの結果、さまざまな記述に柔軟に対応した認識が可能となる。

## 2. 尺八譜の記述

尺八譜で用いられる情報は、主に、以下の3種類に分けられる(図2参照)。

- (a) 音の高さを表す「文字」
- (b) 音の長さを表す「直線」および「点」
- (c) 「歌詞」など

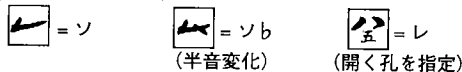
(a)「文字」は、音の高さを示す。基本となる音程を示す文字(ㇿなど)、その文字に点を重畳したもの(ㇿなど)(以下、この点を付点と呼ぶ)、基本となる文字に小さい漢数字を付与したもの(ㇿなど)、などからなる。付点は約半音の音程の変化を指示し、位置により音程の変化量が異なる。また、文字に付与された小さい漢数字は、指使いにおける開く孔の位置を指示している。(b)「直線」と「点」は、音の長さを示す。記号のない「文字」を基準の長さとし、線や点を組み合わせることにより、異なった長さを示す。「線」では、重畳した縦線の数が増えるごとに長さが半分になり、「点」では、文字の左右に配した点の数が増えるごとに長さが2倍になる。(c)その他、譜面には、歌詞や特殊な指示などが記述されることが多い。

本論文では、この中で、(a)、(b)を認識対象とする。これらが認識できれば、音の高さと長さの情報についての情報が得られ、尺八譜による自動演奏が可能となる。

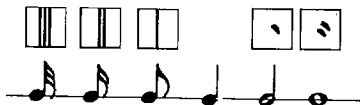
## 3. 記号が重畳した文字列認識方法

認識は、「文字列切出し」と「文字・直線・点認識」の2段階からなる。

### (a) 文字 --- 音の高さ



### (b) 直線や点 --- 音の長さ



### (c) 歌詞、その他(ビブラートなど)

図2 尺八譜の記述法

Fig. 2 Notation of Syakuhachi score.

識」の2段階からなる。

### 3.1 文字列切出し

文字列切出しでは、黒画素の縦方向の射影値に基づき文字列の位置および幅の推定を行う。具体的には、切出し観測窓(図3(a))内の領域に対し黒画素数の垂直方向の射影分布を求め、射影値があらかじめ定められたしきい値を超えた場合文字列候補とみなし、その位置および幅を求める。切出し観測窓を縦方向にずらしながら、各位置で得られた文字列候補の位置と幅を平均化することにより、安定した文字列の位置と幅が確定できる。

切出し

### 3.2 文字・直線・点認識

次に、補完類似度を用いたずらし照合により、文字・直線・点の認識をそれぞれ独立に行う。譜面上の文字のピッチは異なっており、更にさまざまな情報が記述されているため、従来法である黒画素数の射影による個々の情報の切出しは困難である。そのため、個々の情報の切出しを行う必要のないずらし照合を採用する。また、同時に、複数の情報が重畳したパター



図3 認識実験結果(各行の右端の「平仮名」は今回認識対象外)(a)入力画像と観測窓(大きい方形:切出し観測窓,小さい方形:認識観測窓)(b)認識結果(c)認識結果を五線上に示したもの

Fig. 3 Recognition results (small Hiragana characters on the right side of each line were not recognized in experiments) (a) input image and observation windows (large rectangle: observation window for text-line extraction, small rectangle: observation window for recognition) (b) recognition results (c) recognition results on staff lines.

ンを認識するために、ノイズにロバストな認識尺度である補完類似度 $S_c$ [5]を採用する。これにより、複数の情報が重畳していても、一つの情報のみに着目して認識を行うことができる。なお、 $n$ 画素からなる入力パターンベクトル $F$ の参照パターンベクトル $T$ に対する補完類似度 $S_c$ の定義は以下のとおりである。

$$S_c(F, T) = \frac{n \cdot a - F \cdot T}{\sqrt{T \cdot (n - T)}} \quad (1)$$

ここで、

$$a = \langle F, T \rangle, F = \|F\|, T = \|T\|. \quad (2)$$

切出された文字列単位に認識観測窓(図3(a))を垂直方向に走査し、各位置で認識観測窓と文字辞書または直線辞書、点辞書との照合を行う。点の認識では、認識観測窓を窓幅の1/2だけ左右に移動して照合を行う。各認識観測窓に対する最大類似度が、あらかじめ設定されたしきい値以上の値となった場合、そのカテゴリーを認識結果とする。

#### 4. 実験結果

尺八譜2曲分(「春の小川」「夕焼け小焼け」)を実験に用いた。雑音耐性を調べるために、譜面から直接スキャンした良質な画像と、裏面が使用済の再生紙にコピーした後にスキャンした劣化画像(図4)の2種類を用いた。認識に用いた各辞書のカテゴリー数は、文字辞書24カテゴリー、直線辞書1カテゴリー、点辞書2カテゴリーとした。辞書内の参照パターンは、カテゴリー当り1個、大きさは32×32画素とした。文字部分の認識結果を表1に、良質な画像に対する本手法による認識結果例を図3(b), (c)に示す。

本手法により、良質な画像の場合、文字部分の84%、直線部分の99%、点部分の70%が正しく認識できた。文字部分の誤読の主な原因は、認識誤り(付点の有無の判別誤り等)であった。

比較のために、個別に文字部分を切り出し認識する実験を行った。上記と同じ方法で文字列切出しを行った後に、黒画素数の水平方向の射影分布にしきい値処理を施すことにより文字部分の抽出を行い認識した。このような個々の情報の切出しを行った後に認識を行う手法では、良質な画像に対しても42%の認識率が得られなかった。誤読の原因は、切出し誤り29%、認識誤り29%であった。本手法ではずらし照合を用いることにより、切出し誤り、認識誤りともに削減することができた。

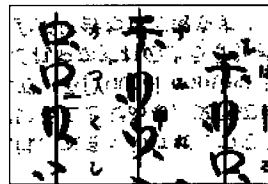


図4 劣化画像の一部

Fig. 4 Example of degraded image.

表1 文字部分の認識率(%) (a)良質な画像, (b)劣化画像  
Table 1 Recognition results for character parts (%). (a)good quality image (b)degraded image

|         | (a)   |     | (b)   |     |
|---------|-------|-----|-------|-----|
|         | $S_c$ | $S$ | $S_c$ | $S$ |
| ずらし照合   | 84    | 85  | 84    | 64  |
| 個別文字切出し | 42    | 42  | 23    | 25  |

更に、補完類似度を用いることによる効果を調べるために、類似尺度として単純類似度 $S(S=a/\sqrt{F \times T})$ を用いた認識実験を行った。表1より、良質な画像に対しては、補完類似度を用いた場合と単純類似度を用いた場合とで認識率の差が見られないが、劣化画像に対しては補完類似度を用いることにより、高い認識率を保つことが示された。

これらの結果より、提案手法の有効性が示された。

#### 5. むすび

複数情報が重畳している尺八譜を認識する手法を提案した。補完類似度を用いたずらし照合により、複数情報を個別の情報に分ける必要がないため、高い認識率が得られることを示した。今後の課題としては、手書き譜のような字形変形が大きい場合の辞書作成法、五線譜への応用、などが挙げられる。

#### 文 献

- [1] B.Blostain and H.S.Baird, "A critical survey of music image analysis," Proc. of SSPR '90, 1990.
- [2] G.E.Kopec, P.A.Chou, and D.A.Maltz, "Markov source model for printed music decoding," Proc. of SPIE '95, vol. 2422, pp.115-125, San Jose, U.S.A, 1995.
- [3] 志村 哲, 坪井邦明, 松島俊明, "日本音楽の情報処理 - 尺八の場合," 情処学音情研報, 93-MUS-2, pp.17-24, June 1996.
- [4] 祖慶実利, 山城 毅, 伊波善清, 渡久地 實, "琉球サンシン楽譜(工工四)の認識に関する研究," 信学ソ大, D-12-18, p.210, Sept. 1997.
- [5] M.Sawaki and N.Hagita, "Text-line extraction and character recognition of Japanese newspaper headlines with graphical designs," Proc. of 13th ICPR, Vienna, Austria, Track C, pp.73-78, 1996.

(平成10年2月4日受付, 5月12日再受付)