

## ハフ変換のパラメータ空間を用いた 手書き音符の認識に関する研究

### 1 はじめに

音楽情報を計算機に入力する方法の1つとして、印刷楽譜を光学的に読み取り、パターン認識を行う楽譜認識と採譜<sup>1</sup>がある。採譜では、楽譜上に表記されていた強弱記号や発想記号などの情報を復元することは非常に困難であるが、楽譜認識は楽譜上の記号を直接読み取ることができるので、特に楽譜情報の計算機入力方法としての期待は大きいとされている [1]。

しかしながら、現在市販されている楽譜認識ソフトウェアは、手書きされた楽譜の読取りに対応していない [2]。手書きの音符の形状は、書き手によって様々な違いがあるため、通常のパターン認識では正しい結果を期待できない。そこで手書きされた音符・記号の特徴を抽出し、それによって判断する方法が考えられる。

本論文では、音符・記号の特徴抽出の方法としてハフ変換を用いた手書きの楽譜の認識手法を提案する。

### 2 楽譜認識における問題点

現在までに行われてきた研究はいずれも印刷楽譜を対象としたものであるが、それらの成果によって、実用的な楽譜認識システムを構築するためには、次のような問題点を解決する必要があることが明らかになっている。

1. 高分解能での画像入力  
楽譜には微少な構造の記号が多く含まれているので、高解像度の入力画像と処理能力の高い計算機が必要となる。
2. 2 値化しきい値の決定  
2 値化の際のしきい値を尾適切に決定できるかどうかは認識結果に大きく影響する。
3. 入力楽譜自体の画質や誤値  
インクの汚れや擦れなどの楽譜の劣化、作曲時の音符の間違いなどを単に削除するのではなく適切に扱う必要がある。
4. 認識対象外記号の扱い  
楽譜には楽譜固有の記号以外に、日本語、外国語、数字などが、表記されている場合があり、しかも出現位置の特定も困難なことが多い。
5. 楽譜記号の多様性  
楽譜には活字のように変形しない記号と、変形するが意味的に同じ記号が存在する。変形しない記号でも、作者によりその形状が異なるため、パターンマッチングのみによる認識には限界がある。

### 6. 楽譜の難易度差

音符密度が低く、限られた数の基本的な記号しか表記されない楽譜から、音譜が密集し、様々な記号が書き込まれている楽譜までその差は様々である。

7. 認識結果の評価認識の評価は、音高・音価の情報を重視するものから、種類・大きさ・配置などの図形的な情報が重要となる楽譜出版まで、楽譜から得られる情報のうち必要とする情報には大きな幅があるためいろいろな面からの評価が必要である。

### 3 手書き音符の認識

ハフ変換とは式 (1) で示されるもので、画像中のデータ点をパラメータ空間への写像を行うことで、あらかじめ決められた線を検出することができる手法である。図 1 に  $x$ - $y$  空間と  $\theta$ - $\rho$  空間相互に対応する写像である。

$$\rho = x \cos \theta + y \sin \theta \quad (1)$$

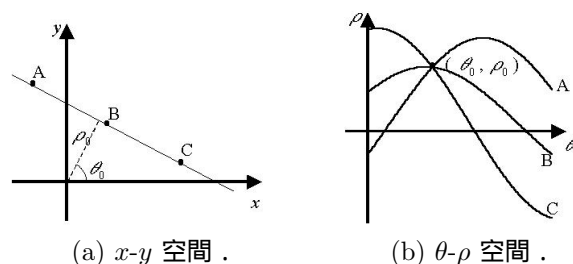


図 1:  $x$ - $y$  空間と  $\theta$ - $\rho$  空間相互に対応する写像

#### 3.1 音符の領域の取得

まず、認識対象である音符の検出を考える。楽譜の認識において、五線を検出することは音譜の位置を示すためにとても重要である。五線を検出することで音符のおおよその位置や領域の限定を行うことが可能となる。

##### 3.1.1 五線領域の推定

五線の検出は画像の水平方向のヒストグラムに基づいて行う。楽譜は 2 値画像であることから、楽譜の中では、五線が水平方向に一番長い直線であるので、その部分の画像数が極めて大きい値をとる。従来の方では、この五線の示すピーク値からその領域を確定させていたが、長いスラーがある場合だと五線の示すピーク値に近い値が出る場合がある。そこで、提案手法では五線と五線の間の空白を利用して領域の推定を行う。

<sup>1</sup>音響情報からの認識

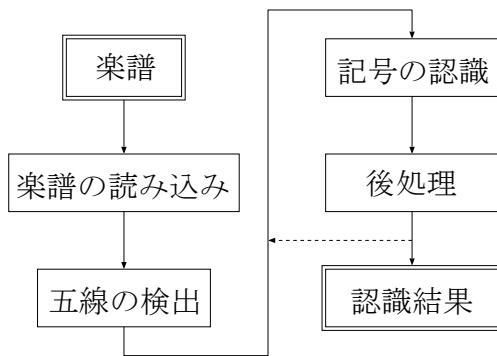


図 2: 楽譜認識の手順の例

### 3.1.2 五線の消去

先に用いた水平方向のヒストグラムに基づき五線と考えられる部分を消去する。これは後の認識の作業の妨げになるからである。

また、消去の際に五線と判断した直線領域には、ラベルをつけておく。これは後で音符の音程を推定するために使用する。第一線の考えられる直線領域には”1”，第二線には”2”，というように第五線までラベルをつける。また、各線同士の間隔の平均分だけ第一線から下に平行移動した直線領域を下第一線として”0”のラベルをつける。同様に上第一線として”6”のラベルをつける。

### 3.1.3 音符の取得

各五線領域における垂直方向のヒストグラムに基づいて行う。

符幹のある音符はその部分の分布が大きな値をとるので、それを基準に音符の領域を定め手でもいいが、音符は符幹のないものもあり、また休符などの記号場合は符幹のような大きな値をとる部分はないため符幹を基準にするだけではすべての音符を取り出すことはできない。よって、分布の集中している部分に音符があると考えて、領域を求める。

### 3.1.4 音程の推定

切り出した音符の符頭から下に向かって探索し、最初に見つかるラベルを  $L_1$  とする。また、最上部からも下に向かって探索し、同じように見つけたラベルを  $L_2$  とする。 $L_1$  と  $L_2$  の関係により、符頭の位置を決定し、音程を取得する。

1.  $L_1 = L_2$  ならば符頭は第  $L_1$  間に存在することになる。
2.  $L_1 \neq L_2$  ならば符頭は第  $L_2$  線上に存在することになる。

### 3.1.5 重心の統一

切り出した音符の重心を統一することで画像の音符の位置を統一する。切り出した文字の重心  $(J_X, J_Y) = (\frac{X}{SUM}, \frac{Y}{SUM})$  が画像の中心となるように平行移動する。

## 3.2 評価方法

### 3.2.1 ピーク取得

ハフ変換によって生成されたパラメータ空間上で濃度値の高い点を取り出す。この点が画像の特徴となる [3]。

### 3.2.2 ピーク位置の照合

データベース画像のピーク位置を中心に  $\theta$  方向に  $\pm A$  画素、 $\rho$  方向に  $\pm B$  の範囲内に入っている切り出し画像のピークをペアとする。また、その範囲内に複数個のピークが入っている場合は最も距離の近いものをペアとする。全く同じ画像ならばピーク同士の距離はゼロであるので、ペアとしたピーク同士の距離の合計が一番小さいものが最も類似している音符、記号として判断する。

## 4 検証実験

### 4.1 評価方法

認識した結果を楽譜に再現することで認識結果の確認を目視にて行う。

### 4.2 印刷楽譜についての検証

提案手法は、手書きの楽譜の認識を目的としたものであるが、形状特徴の抽出に基づく手法であることから、印刷楽譜の認識も当然可能であり、また印刷楽譜の方がより形状特徴を抽出しやすいと考えられる。そこでまず、印刷楽譜に対して提案手法と、KAWAI スコアメーカーとで比較を行った。

### 4.3 手書き楽譜についての検証

同様に、手書きの楽譜に対して提案手法と、KAWAI スコアメーカーとで比較を行った。

結果より、全体的な認識率は提案手法が低いものの、形状が複雑なためにスコアメーカーでは認識率の低い休符について高い認識率が得られた。一方、八分音符を四分音符に認識してしまう場合が多いが、これは連桁と五線が重なるからだと考えられる。

## 5 おわりに

ハフ変換を利用した手書き楽譜の認識手法を提案した。検証実験より、市販されているソフトウェアでは認識が難しい部分で認識率の向上が確認できた。今後の課題としては、五線領域の推定および音符の切り出し精度の向上が挙げられる。

## 参考文献

- [1] 長島洋一, 橋本周司, 平賀譲, 平田圭二” コンピュータと音楽の世界,” 共立出版, 1998.
- [2] (株)河合楽器製作所, ”スコアメーカー製品紹介,”
- [3] 永井嘉高, 寺田賢治, 石橋裕美, 大橋剛介, ”Hough 変換を用いた筆跡による個人認証,” SSI'pp.205-210, 2003.