

簡易な手書き譜面を利用した演奏手法の検討と実装

馬場 哲晃^{1,a)} 菊川 裕也¹ 串山 久美子¹ 青木 允²

概要: 本稿は、以前の研究会にて報告した「ラインイメージセンサを利用した簡易な手書き五線譜読取りアルゴリズム」に関する経過報告となる。本稿では主に五線と符頭からなる手書き譜面を利用した演奏インタフェース及びインタラクションに関して検討を行った。結果として、音量、発音時間、ピッチバンド、楽器選択操作等を独自の操作方法によって実現した。本システムは、演奏パフォーマンスや幼少時における音楽学習教材等への展開を目的としている。

キーワード: インタラクションデザイン, 演奏インタフェース, 光学楽譜認識 (OMR), 手書き譜面

Prototyping of Musical Interface based on Handwritten Musical Notation

TETSUAKI BABA^{1,a)} YUYA KIKUKAWA¹ KUMIKO KUSHIYAMA¹ MAKOTO AOKI²

Abstract: In this paper, we shall describe a progress report of last year's paper titled "Algorithm that Enables Users to Read Hand-Written Staff Notations with a Line Image Sensor". We mainly describe musical interface / interaction that uses handwritten musical notating. As a result, we developed original musical interaction for articulation such as velocity, beat, and pitch, and so on. Our goal is to develop the system that enables users to read easy notations with a image sensor and perform with our system.

Keywords: Interaction Design, Musical Interface, Optical Music Recognition(OMR), Handwritten notation

1. はじめに

著者らはこれまで簡易な手書き譜面を利用した演奏インタラクションを検討してきた [1]。図 1 に示すように、五線と符頭からなる簡易な手書き譜面を、ユーザが小型デバイスでスキャン操作をすることで、即座に任意の音程を再生するシステムである。実験機を制作し、各音程位置情報を検出できることを示した。しかし研究を継続した結果、ラインイメージセンサだけではユーザの手動スキャン操作時において誤検出が多くみられた。そこで著者らはラインイメージセンサを CMOS イメージセンサに変更し、プロト

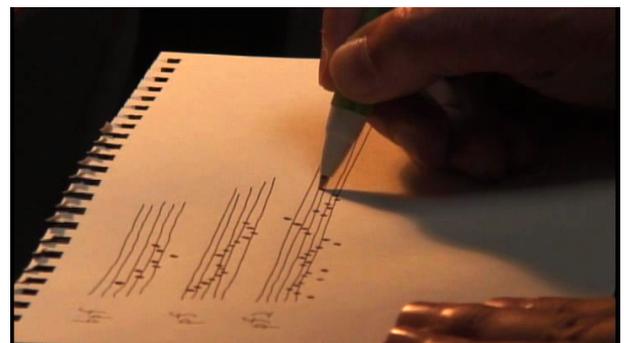


図 1 記譜の様子

Fig. 1 A snapshot of handwritten notation.

タイプ制作を行った。

本研究のシステム概観を図 2 に示す。まずユーザは五線及び符頭を記述する。符頭とは図 3 に示すような、音符を構成する一つの記号部位を指す。五線及び符頭を記述後、

¹ 首都大学東京大学院
Tokyo Metropolitan Univeristy, 6-6, Asahigaoka, Hino,
Tokyo 191-0065, Japan

² 株式会社ウィズコーポレーション
With Corporation Co.,Ltd.

a) baba@sd.tmu.ac.jp

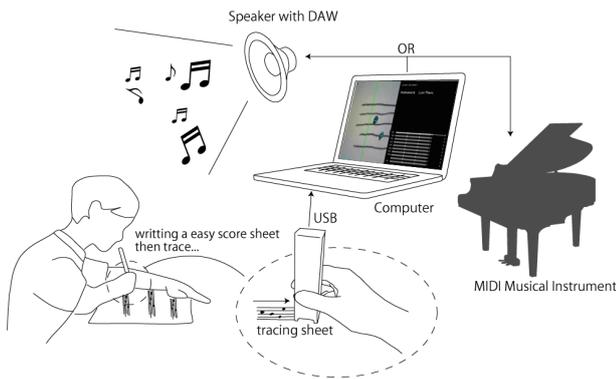


図 2 システム概観

Fig. 2 A system abstract of our system

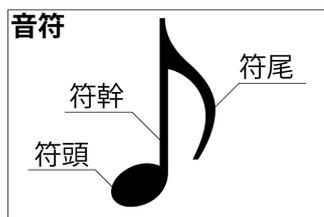


図 3 音符を構成する符頭, 符尾, 符幹の関係

Fig. 3 Compositions of musical note

ユーザはスキャンデバイスを記述箇所に合わせてスキャン操作することで、任意の音程を再生できる。ユーザが手動でスキャン操作を行うことで、拍情報が意味をなさなくなるため、処理の簡易化も兼ねて符頭のみを音符記号として利用する。これら一連の操作に加え、本研究では演奏時における表情付けや楽器選択、シーケンサ機能を実装することで、本格的な音楽演奏を楽しめるシステムを提案する。

記譜とは楽曲の保存手法の一つである。口承などとは異なり、曖昧性を排除し、正確に楽曲を保存することができる。一方で、私たちは音楽記録媒体の発達によって、記譜に関する特別な訓練を必要とせず、楽曲を保存・再生することが可能になった。しかし実際に楽器を演奏する場合、正確な音程やリズム等の情報を、音楽記録媒体を利用した再生（いわゆる耳コピなど）だけで再現することは一般のユーザには困難である。このように、音楽や楽器演奏を深く楽しむ場合、楽譜を読む力は必要不可欠である。楽曲を保存する「記譜」と、楽曲を再生する「演奏」を同時に楽しめることで、楽譜に新たなアーカイブ手法を付加する。

1.1 関連研究

Optical Music Recognition(OMR) によるバッチ処理型楽譜認識は著者がこれまで述べたとおり [1] である。このようなバッチ処理型楽譜認識に対して、本研究が従来研究と大きく異なる点は、簡易な手書き譜面であること、任意の音程を鳴らすだけでなく対話的な演奏表情付が可能である 2 点である。

手書きとは異なるが、五線譜上に音符記号オブジェクト

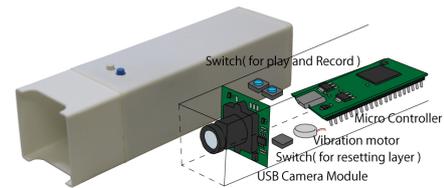


図 4 デバイス構成

Fig. 4 Compositions of our device

を置くことで楽曲演奏を行うシステムとして、Bottello による Tangible user interface for music learning[2] がある。予めマーカをつけた実物体オブジェクトを印刷された五線紙上に配置することで、楽曲演奏が可能になる。実物体オブジェクトを置くだけでなく、引き伸ばしや回転動作に対応しており、それぞれの動作によって拍の長さや臨時記号に対応できる。このような実物体をタイムライン上に配置し、それに応じて楽曲を再生する手法は Levin[3] や岩井による「音楽のチェス」[4] 等多くの先行事例がある。前述の Bottello らはこれらタンジブルな電子楽器を応用して、音楽教育への可能性を踏まえている点で、本研究に類似している。

本研究に最も類似した事例として Hoerter ら [5] による「The Music Wand」があげられる。これは本研究と同様に手持ち型のデバイスのスキャン操作により、印刷楽譜をリアルタイムに再生するものである。しかし音程認識が 70% の認識率であったり、各種臨時記号の認識に問題があるなどの問題点が報告されている。この他 2 値化画像取得の為にトレースボックスを必要とする等、ユーザビリティに問題がある。本研究では手書き譜面を対象とする他、操作方法によって各種アーティキュレーションや楽器選択ができるなど、Hoerter らの「印刷楽譜を手動スキャナ操作で再生する」といった設計と異なる。

2. システム概要

システムは図 2 に示したとおり、主にデバイス、コンピュータ、音源の 3 つからなる。デバイスは主にカメラモジュール、スイッチ群、マイクロコントローラ、振動モータからなる。内部の構成を図 4 に示す。

ユーザのスキャン操作によって読み込まれた画像をコンピュータで処理を行い、その結果を適合する音程や楽器音で出力する。ユーザが記述した手書き譜面上にデバイスを置くことで、図 5 に示される画像がコンピュータ画面に表示される。符頭位置が十字にマークされ、ユーザがデバイスを操作し、画面中央の緑色の演奏ラインに重畳させることで、符頭の音程を出力できる。なお十字の縦はその符頭の音量情報を、横は符頭認識領域幅を示している。ユーザは図 5 に示すカメラ取得画像だけでなく、図 6 に示す画面を利用して演奏を行う。図 6 にはカメラ画像、選択楽器

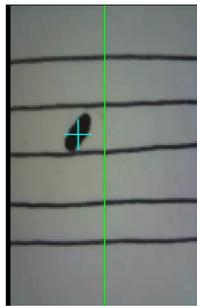


図 5 デバイスが取得した画像を元に、符頭検出を行った画像
Fig. 5 A Processed image that is captured by camera device.



図 7 OCR 機能により、楽器文字を認識した時の様子
Fig. 7 A screenshot of recognition of instrument name character by ocr.

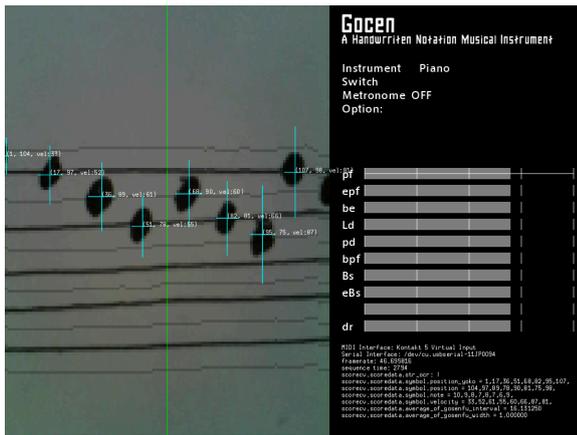


図 6 ユーザインタフェース画面
Fig. 6 Screenshot of User Interface

名、オプション選択、タイムライン確認、デバッグのそれぞれの項目が確認できるようになっている。演奏ボタンを押しながらスキャン行為を行うことで、ユーザは楽器音の発音が可能になる。

2.1 楽器選択

ここまで、ユーザは任意の楽器選択が可能であると述べてきた。この機能を実装するために本研究では OCR ライブラリによる処理を利用した。一般的に記譜上の楽器を示す場合、ピアノであれば pf, ギターであれば gt といった省略文字列が利用される。そこで本システムにおいても同様に、手書きで記述した文字列をカメラで取得、処理することで任意の楽器選択機能を実現した。現在利用できる楽器記号は pf(ピアノ), epf(エレКТРИック・ピアノ), be(ベル), Ld(シンセリード), pd(シンセパッド), bs(ベース), dr(ドラムセット) 等計 9 種類の利用が可能である。図 7 は、楽器 Bell(be) を認識した際の UI 画面におけるスクリーンショットである。本機能は、取得画像上に五線譜エリアが 1 箇所も認識されない場合、自動的に OCR モードに切り替わるため、ユーザは任意のボタンを押すなどの操作を必要としない。

Ocrad:The Gnu OCR ライブラリを使用した。
<http://www.gnu.org/software/ocrad/>

2.2 シーケンス演奏機能

本システムには対話的な演奏方法に加え、ユーザの演奏をタイムライン上に録音し、それらを重ね合わせることで複雑な演奏を可能にするシーケンス演奏機能を実装した。シーケンス演奏を行う為には、デバイス上の録音ボタンを押しながら、演奏を行えばよい。各楽器に 1 レイヤを設定し、ユーザは楽器を切り替えると同時にシーケンス演奏用のレイヤを切り替えることになる。これによりユーザはレイヤ概念を意識せずに演奏を楽しめる。ユーザは対話的に切り替え可能なメトロノーム音を基準に、録音時間に応じて伸縮されるタイムライン上に発音ポイントを作成できる。録音長は 4 小節 $\times n$ 長を基準に伸縮できる。なお録音機能にはクオンタイズ機能が実装されているため、ユーザが乱雑に発音を行った場合における僅かな発音ポイントの揺れを修正できる。一度録音したタイムラインはデバイス上のリセットボタンを押すことで消去することができる。一方で、一度録音したタイムラインに更に同じ楽器音で録音を行う場合、再度録音ボタンを押しながら演奏することで録音ポイントが追加できる。

3. 演奏手法

音楽を演奏する際、音の強弱、音高のゆらぎ(ピッチベンド、モジュレーション、ビブラート)、和音、発音長さ等のアーティキュレーションによって、演奏に表情付けを行う。本研究においても、このような表情付けを実現するための実装を行った。

3.1 発音長

図 5 に示される緑色の発音ラインと符頭を重畳することで、任意の音程を発音できる。図 8 に発音前から発音、発音終了の様子を示す。発音ラインが符頭上にある間、ユーザはその符頭音を note on 状態にし、発音ラインを符頭上から外すことで note off 状態にできる。これを利用することで、スラーやスタッカートが表現可能になる。

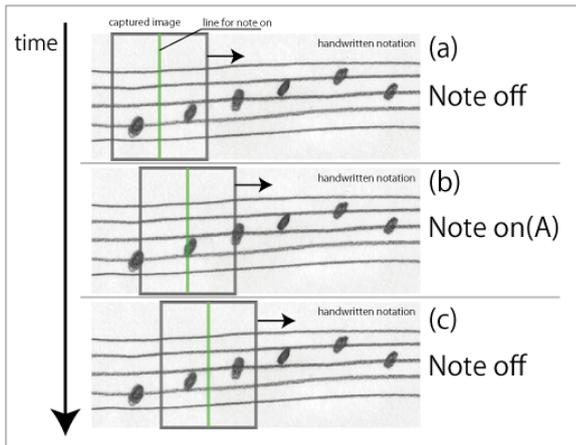


図 8 発音の流れ
Fig. 8 The flow of making sounds

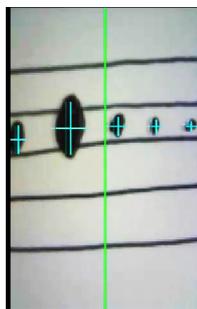


図 9 符頭の大小とペロシティの関係
Fig. 9 Relation of velocity and size of a notation

3.2 音の強弱

譜面上で音量を指示する場合、フォルテやピアノ、アクセント記号などを用いるのが一般的である。本研究ではこれに対し、より直感的な操作インタフェースを提供するために、符頭の大小を音量として扱う。図9に示す通り、符頭の大小に合わせてペロシティ(音量)が異なっているのが分かる。

3.3 音高のゆらぎ

主に弦楽器を利用する場合、発音後に音程を一定に保つだけでなく、音程にゆらぎをつけることで、ビブラートや(逆)ターンといったアーティキュレーションを表現する。本研究において、ユーザは発音後、デバイスを上下に動かすことで音程にピッチバンド効果を与えることができる。これによって前述のアーティキュレーションを表現可能である。発音時の符頭位置を基にカメラの上下動作量を検出している。

4. 音程位置認識結果

制作したシステムにて、音程の認識に関する簡単な検証を行った。五線と符頭からなる単純な楽譜ではあるが、印刷楽譜と異なり、記述にゆらぎが含まれる為、認識結果に個人差が生じる場合も考えられる。そこで実際に2人のユー

ザに記譜をしてもらい、認識精度を確かめた。被験者には日常的に記譜経験があったユーザ1名と、そうでないユーザ2名を選んだ。被験者に対し、まず本システムの簡単な説明及び記述に際する注意事項(1. カメラ画像で読み取る範囲より大きく記述しないこと。2. 符頭検出領域におけるx軸が重畳しないこと。3. 符頭は塗りつぶすこと)を告げた。符頭の大きさや五線の間隔等については説明をしなかった。次に被験者に五線とその中に符頭を自由に記述してもらい、記述後、各符頭位置を被験者に確認し、被験者が意図した正しい符頭位置を記録する。その後本システムを利用し、記述された譜面を読み取ることで認識率を確認した。結果として各ユーザに対していずれも100%の認識精度を実現できた。結果を表1に示す。なお、筆記具には0.7mmボールペンを、紙は一般的なコピー用紙(白)を利用した。

表 1 実験結果

Table 1 Experimental results

被験者	符頭数	誤認識	認識率 [%]
1	69	0	100
2	87	0	100
3	67	0	100

5. 今後の展望

本稿では演奏手法を中心に論じたが、符頭以外の記号対応や五線位置検出における詳細なアルゴリズム検討等、画像処理手法に関する詳細な検討が未熟である。今後はこれら機能追加・改善に取り組む。

本システムでは録音機能によって、シーケンス演奏が可能であるため、一人でもある程度の合奏を楽しむことができる。一方で複数のユーザ同士にて本システムを利用する際、より複雑な演奏実演が可能である。よって演奏パフォーマンス向上の為に複数のデバイスを制作する。

参考文献

- [1] 馬場哲晃, 青木允. ラインイメージセンサを利用した簡易な手書き五線譜読取りアルゴリズム. 情報処理学会研究報告. EC, エンタテインメントコンピューティング, Vol. 2011, No. 9, pp. 1-4, 2011-08-23.
- [2] Loris Bottello. Tangible user interface for music learning, 2010. <http://cloverleafdesign.weebly.com/>.
- [3] Golan Levin. The table is the score: An augmented-reality interface for real-time, tangible, spectrographic performance. In *Proceedings of the International Conference on Computer Music 2006 (ICMC'06)*, 2006.
- [4] 岩井俊雄. 音楽のチェス, 1997.
- [5] Nick Hoerter and Tom Chatt. The music wand: Real-time optical scanning of sheet music, 2008. http://people.ece.cornell.edu/land/courses/ece4760/FinalProjects/s2008/tjc42_nah35/tjc42_nah35/index.html.